Computational Economics
https://doi.org/10.1007/s10614-019-09882-2

Markov Regime-Switching in-Mean Model with Tempered Stable Distribution

Yanlin Shi[1] · Lingbing Feng[2] · Tong Fu[2]

Accepted: 17 January 2019

Abstract

Markov Regime-Switching (MRS) model is a widely used approach to model the actuarial and financial data with potential structural breaks. In the original MRS model, the innovation series is assumed to follow a Normal distribution, which cannot accommodate fat-tailed properties commonly present in empirical data. Many existing studies point out that this problem can lead to inconsistent estimates of the MRS model. To overcome it, the Student's t-distribution and General Error Distribution (GED) are two most popular alternatives. However, a recent study argues that those distributions lack in stability under aggregation and suggests using the α-stable distribution instead. The issue of the α-stable distribution is that its second moment does not exist in most cases. To address this issue, the tempered stable distribution, which retains most characteristics of the α-stable distribution and has defined moments, is a natural candidate. In this paper, we conduct systematically designed simulation studies to demonstrate that the MRS model with tempered stable distribution uniformly outperforms that with Student's t-distribution and GED. Our empirical study on the implied volatility of the S&P 500 options (VIX) also leads to the same conclusions. Therefore, we argue that the tempered stable distribution could be widely used for modelling the actuarial and financial data in general contexts with an MRS-type specification. We also expect that this method will be more useful in modelling more volatile financial data from China and other emerging markets.

Keywords Regime-switching · Fat-tailed distribution · Tempered stable distribution

JEL Classification C22 · C51 · G11

✉ Tong Fu
canjianft@hotmail.com

[1] Department of Applied Finance and Actuarial Studies, Macquarie University, Macquarie Park, NSW 2109, Australia

[2] Institute of Industrial Economics/Center for Regulation & Competition, Jiangxi University of Finance and Economics, Nanchang, Jiangxi, China

Published online: 27 February 2019

1 Introduction

Regime-switching models are widely employed to study various types of financial data such as interest rate, equity volatility and inflation rate (see, for example, Zou and Cadenillas 2014; Fan et al. 2015 and Ye et al. 2016). The popularity of those models is due to the fact that they can appropriately control for the effects of structural breaks and business cycle (Hamilton 1988, 1994; Marcucci 2005; Shi and Ho 2015). Hence, more accurate estimates and predictions can be generated, which are essential to financial practice like derivative pricing.

The idea of the regime-switching models can be traced back to the fundamental works of Quandt (1958), Goldfeld and Quandt (1973) and Tong (1983). Based on the earlier literature, the Markov Regime-Switching (MRS) model is proposed in the seminal work of Hamilton (1989). It is originally developed to investigate the impacts of business cycle on financial time series. In this work, Hamilton (1989) argues that the behaviour of financial data tends to be dependent on the states (regimes) of business cycles. To take this impact into consideration, parameters of a general ARIMA model are allowed to switch between states. The MRS model is then constructed by assuming that the latent states follow a Markov chain. Over the past few decades, this model has become one of the standard approaches to model regime switching, structural breaks and business cycle (Diebold and Inoue 2001).

The original MRS model is derived with the assumption that the innovation sequence follows a Normal (Gaussian) distribution. However, significant evidence suggests that the financial time series is rarely Gaussian but typically leptokurtic and exhibits fat-tail behaviour (Bollerslev 1987; Susmel and Engle 1994; Shi and Ho 2015). For the ARIMA model, it is well-known that Maximum Likelihood Estimation (MLE) based on Gaussian distribution will always still lead to consistent estimators.[1] For the MRS model, however, if regimes are not Gaussian but leptokurtic, the use of within-regime normality can seriously affect the identification of the regime process (Klaassen 2002; Ardia 2009; Haas 2009). Moreover, Haas and Paolella (2012) argue that the ML estimators of the original MRS model (with a Normal distribution assumption) for fat-tail-distributed data are not consistent. Shi and Feng (2016) further provide simulation evidence of this inconsistency in the case of a GARCH specification. Consequently, the sought of an appropriate distribution to accommodate the excess kurtosis of the financial time series becomes an essential issue for the MRS-type model.

To solve this problem, a common solution is to employ a fat-tailed distribution such as the Student's t-distribution or General Error Distribution (GED) (Klaassen 2002; Haas et al. 2004; Marcucci 2005). However, a recent study by Calzolari et al. (2014) argues that the widely used Student's t-distribution is problematic. The most serious drawback of this distribution is that it lacks in stability under aggregation, which is of particular importance in portfolio applications and risk management in finance study. The same problem also exists for other non-stable distributions like GED. As a replacement of them, the α-stable distribution is recommended

[1] The estimators are inefficient if the underlying distribution of innovation sequence is not Gaussian.

by Calzolari et al. (2014) to model the fat-tailed behaviour of financial time series. Unfortunately, the second moment of the α-stable distribution does not exist in most cases. Hence, the application of it to the MRS model will lead to questionable interpretation and problematic statistical properties of the ML estimators.

The tempered stable distribution is a natural substitution of the α-stable (Feng and Shi 2016; Shi and Feng 2016). Firstly introduced[2] in Koponen (1995), tempered stable distribution covers a range of well-known subclasses like Variance Gamma, bilateral Gamma and CGMY distributions (Küchler and Tappe 2013). Its most outstanding advantage is that the tempered stable distribution retains most of the attractive properties of the α-stable and has defined moments. Additionally, it has a flexible form which accommodates both the symmetric and asymmetric shapes of density. So unlike the standard Gaussian, Student's t or GED, it can also be used for both symmetric and skewed data.

In this paper, we employ the tempered stable distribution within the MRS framework and argue that it outperforms the commonly used Student's t and GED. To demonstrate that, we conduct a series of simulation studies to compare the performance of the MRS models with the Normal, Student's t, GED and tempered stable distributions. First, we investigate two scenarios where the true distributions are Student's t and GED, respectively. Via twelve combinations of different parameter and sample size settings, the MRS models with four distinct distributions are systematically analysed. It is demonstrated that when the true distribution is Student's t or GED, the MRS model with tempered stable distribution generates very similar results to those of the true model. More importantly, it clearly outperforms the competitors (the MRS models with distributions other than the true density) in terms of consistency, efficiency and overall performance. Second, we set the tempered stable as the true distribution. Twelve sets of simulations are further constructed, including different choices of transition probabilities and tempered stable distribution parameters. In this scenario, none of the Normal, Student's t and GED distributions can perform as well as the tempered stable distribution. Therefore, we argue that the tempered stable distribution could be widely used for modelling the financial data in general contexts with an MRS-type specification.

To empirically analyse the MRS model with different fat-tailed distributions, we work on the daily VIX index, which is the implied volatility of the S&P 500 options. Since the VIX index is skewed to some extent after taking the logarithm transformation, we consider and compare five cases: (skewed) Student's t, (skewed) GED and tempered stable distributions. The results suggest that the MRS model with tempered stable distribution outperforms that with all the other distributions. Besides, the estimated parameters across different distributions are relatively close to each other. The fitted smoothing probability series of the five MRS models also have similar patterns. Hence, our empirical result is robust with the simulation evidence, suggesting that the MRS model with tempered stable distribution is a useful tool in the practical financial study.

[2] In this case, the associated Levy processes are called "truncated Levy flights", the appropriateness of which to be applied in finance studies is discussed in Constantinides and Savel'ev (2013).

The remainder of this paper proceeds as follows. Section 2 describes the specification of the MRS model employed in this study. Section 3 explains how the Student's t, GED and tempered stable distributions can be employed within the MRS framework. We conduct three independent simulation studies in Section 4. The empirical results are discussed in Sect. 5. Section 6 concludes the paper.

2 MRS Model

Financial time series may behave differently depending on the states (regimes) of the business cycle (Hamilton 1989). For instance, the stock return tends to be less volatile in a peak state and be more volatile in a trough state (Ho et al. 2013). To take this impact into consideration, the MRS model is proposed with state-dependent parameters (Hamilton 1988, 1989, 1994)

Let $\{s_t\}$ be a stationary, irreducible Markov process with discrete state space $\{1, 2\}$ and transition matrix $P = [p_{jk}]$, where $p_{jk} = \mathrm{P}(s_{t+1} = k|s_t = j)$ is the one-step transition probability of moving from state j to state k ($j, k \in \{1, 2\}$). Then, we have a standard MRS model[3]:

$$y_t = \mu_{s_t} + \varepsilon_{s_t,t},\ \varepsilon_{s_t,t} = \eta_t \sqrt{\sigma_\varepsilon^2} \quad \text{and} \quad \eta_t \overset{iid}{\sim} N(0, 1) \tag{1}$$

where s_t is the state of y_t at time t. μ_{s_t} is the mean in state s_t. $\varepsilon_{s_t,t}$ is the error at time t in state s_t. η_t is an identical and independent innovation sequence following a Normal distribution, with 0 mean and unit standard deviation. In particular, it is constrained that $\mu_1 < \mu_2$, which indicates that the overall mean in state 2 is greater than that in state 1.

Although the MRS model can have more than two states, we only consider the two-state case in this paper for the following reasons. To the best of our knowledge, there is no effective test developed so far to determine the optimal number of states. Studies like Cho and White (2007) can only test the null hypothesis of one regime against the alternative of regime switching between two states. A recent research of Faias and Nunes (2012) comprehensively reviews the existing regime-switching tests, but none of them test the case of three or more states. Other literature like Carrasco et al. (2014) tests parameters of the MRS model rather than the number of states, which is irrelevant to this purpose. Furthermore, the classic trinity (Wald, likelihood ratio and Lagrange multiplier) tests are not applicable here, due to the fact that the MRS-type model with more states does not nest the one with fewer states (Wilfling 2009). On the other hand, although there are many existing tests examining the number and/or locations of the structural changes (such as the nonparametric test used in Ross et al. 2011), it is not straightforward to extend them to the MRS cases. Another potential problem is that even if the optimal number of states can be tested, their interpretation might not be straightforward. Hamilton's (1989) original work suggests that the two-state case can be linked to the business

[3] In the original MRS model, only the mean is state-dependent, and the variance is not.

cycle. For instance, low- and high-mean states can be referred to as the trough and peak periods, respectively. However, the interpretation for a large number of such states might not be meaningful.

The parameters of the MRS model can be estimated using MLE, and the conditional density of ε_t is given as follows.

$$\begin{aligned} \Omega_{t-1} &= \left\{\varepsilon_{s_{t-1},t-1}, \varepsilon_{s_{t-2},t-2}, \dots, \varepsilon_{s_1,1}\right\} \\ \theta &= (\mu_1, \mu_2, p_{11}, p_{22}, \sigma_\varepsilon^2)' \\ f(\varepsilon_{s_t,t}|s_t = j, \theta, \Omega_{t-1}) &= \frac{1}{\sqrt{2\pi\sigma_\varepsilon^2}} e^{-\frac{\varepsilon_{j,t}^2}{2\sigma_\varepsilon^2}} \end{aligned} \tag{2}$$

where Ω_{t-1} is the information set at time $t-1$. θ is the vector of parameters. $\Gamma(\cdot)$ is the Gamma function and $f(\varepsilon_{s_t,t}|s_t = j, \theta, \Omega_{t-1})$ is the conditional density of $\varepsilon_{s_t,t}$ falling in the state j at time t. This stems from the fact that $\varepsilon_{s_t,t}$ follows a Normal distribution with mean 0 and variance σ_ε^2 given the information at time $t-1$.

Plugging the filtered probability in state j at time $t-1$, $\omega_{j,t-1} = \mathrm{P}(s_{t-1} = j|\theta, \Omega_{t-1})$, into Eq. (2) and integrating out the state variable s_{t-1}, the density function in the Eq. (2) becomes

$$f(\varepsilon_{s_t,t}|\theta, \Omega_{t-1}) = \sum_{j=1}^{2}\sum_{k=1}^{2} p_{jk}\omega_{j,t-1} f(\varepsilon_{s_t,t}|s_t = j, \theta, \Omega_{t-1}). \tag{3}$$

$\omega_{j,t-1}$ can be obtained by an integrative algorithm given in Hamilton (1989). The log-likelihood function corresponding to Eq. (3) is as follows:

$$L(\theta|\varepsilon) = \sum_{t=2}^{T} \ln f(\varepsilon_{s_t,t}|\theta, \Omega_{t-1}) \text{ where } \varepsilon = \left(\varepsilon_{s_t,1}, \varepsilon_{s_t,2}, \dots, \varepsilon_{s_t,T}\right)', \tag{4}$$

and the ML estimator $\hat{\theta}$ is obtained by maximizing Eq. (4).

In order to identify which state y_t lies in at time t, we extract the smoothing probability of state 1 as follows (Hamilton 1988, 1989, 1994):

$$P(s_t = 1|\theta, \Omega_T) = \omega_{1,t}\left[\frac{p_{11}P(s_{t+1} = 1|\theta, \Omega_T)}{P(s_{t+1} = 1|\theta, \Omega_t)} + \frac{p_{12}P(s_{t+1} = 2|\theta, \Omega_T)}{P(s_{t+1} = 2|\theta, \Omega_t)}\right], \tag{5}$$

Using the fact that $P(s_T = 1|\theta, \Omega_T) = \omega_{1,T}$, the smoothing probability series $P(s_t = 1|\theta, \Omega_T)$ can be generated by iterating Eq. (5) backward from T to 1. According to Hamilton (1989), it is expected that y_t lies in state 1 (2) at time t, if $P(s_t = 1|\theta, \Omega_t)$ is greater (smaller) than 0.5.

3 Alternative Distributions for the Innovation Sequence

Although the original MRS model is based on the Gaussian distribution, significant evidence suggests that the financial time series is rarely Gaussian but typically leptokurtic and exhibits fat-tail behaviour (Bollerslev 1987; Susmel and Engle 1994; Stanley et al. 2008; Ho et al. 2013). Further, as noted by Klaassen (2002), Ardia (2009) and Haas (2009), if regimes are not Normal but leptokurtic, the use of within-regime normality can seriously affect the identification of the regime process. The details can be found in Haas and Paolella (2012). They argue that the MLE based on Normal components does not provide consistent estimators of the MRS model, if true distribution of innovation is not Normal. A recent study by Shi and Feng (2016) provides simulation evidence to support this argument in the case of a GARCH specification. Hence, the sought of an alternative distribution to accommodate the excess kurtosis of the financial time series becomes an essential issue for the MRS model.

3.1 Student's t and General Error Distribution

Among the existing literature, Student's t-distribution and GED are two widely employed alternatives in finance studies using MRS-type models (Klaassen 2002; Haas et al. 2004; Marcucci 2005). Both of them can capture leptokurtic and heavy-tail behaviours. When they are applied to the MRS model, the corresponding density functions of Eq. (2) are changed as follows.

$$\begin{aligned} \text{Student's t: } f(\varepsilon_{s_t,t}|s_t = j, \theta, \Omega_{t-1}) &= \frac{\Gamma(\frac{v+1}{2})}{\Gamma(\frac{v}{2})\sqrt{\pi(v-2)\sigma_\varepsilon^2}}\left[1 + \frac{\varepsilon_{j,t}^2}{(v-2)\sigma_\varepsilon^2}\right]^{\frac{v+1}{2}} \\ \text{GED: } f(\varepsilon_{s_t,t}|s_t = j, \theta, \Omega_{t-1}) &= \frac{ve^{-\frac{1}{2}\left|\frac{\varepsilon_{j,t}}{\lambda\sigma_\varepsilon}\right|^v}}{\lambda 2^{(v+1)/v}\Gamma(1/v)} \\ \text{where } \lambda &= \left[\frac{2^{(-2/v)}\Gamma(1/v)}{\Gamma(3/v)}\right]^{\frac{1}{2}} \end{aligned} \tag{6}$$

and v is the degrees of freedom. Then, the ML estimator $\hat{\theta}$ can be obtained in the same way as described in Sect. 2.

3.2 Tempered Stable Distribution

Despite their attractive properties to capture excess kurtosis and fat tails, existing literature argues that the Student's t and GED still have unsolved problems. Calzolari et al. (2014) suggests that the Student's t-distribution lacks the stability-under-addition property. This problem also exists for non-stable distributions like GED. Since stability is desirable in portfolio applications and risk management, a distribution could overcome this issue is of particular importance.

3.2.1 α-Stable distribution

As suggested by Calzolari et al. (2014), the α-stable distribution (also known as stable family of distributions) is a replacement of the traditionally used fat-tail distribution. Its most outstanding characteristic is that α-stable distribution overcomes the stability problem. Additionally, the α-stable distribution constitutes a generalization of the Gaussian distribution by allowing for asymmetry and heavy tails. In general, a random variable x is said to be stably distributed if and only if, for any positive numbers c_1 and c_2, there exists a positive number k and a real number d such that

$$kx + d \overset{d}{=} c_1x_1 + c_2x_2 \tag{7}$$

where x_1 and x_2 are independent variables and have the same distribution as x. The notation $\overset{d}{=}$ indicates equality in distribution. The property described by Equation (7) is also known as the *stability-under-addition* property (Tankov 2003). In particular, if $d = 0$, x is said to be strictly stable. According to Calzolari et al. (2014), theoretical foundations of α-stable distribution lay on the generalized central limit theorem, in which the condition of finite variance is replaced by a much less restricting one concerning a regular behaviour of the tails.

Since α-stable distribution does not have a close form of density function, the best way to describe it is by means of its characteristic function, which has the following form.

$$\phi(t) = \exp\{i\delta t - \sigma^{\alpha}|t|^{\alpha}(1 - i\beta sign(t)\Phi)\} \tag{8}$$

where Φ is $-(2/\pi)\log|t|$ when $\alpha = 1$ and is $\tan(\pi\alpha/2)$ when $\alpha \neq 1$. $\alpha \in [0, 2]$ is the index of stability or characteristic exponent that describes the tail-thickness of the distribution (small values indicating thick tails). $\sigma \in \mathrm{R}^+$ is the scale parameter. $\delta \in \mathrm{R}$ is the location parameter. β describe its skewness. Calzolari et al. (2014) only consider the symmetric α-stable distribution ($\beta = 0$), which is then characterized by (α, σ, δ) and is denoted as $S(\alpha, \sigma, \delta)$. Therefore, the standardized symmetric version is $S(\alpha, 1, 0)$ with the following characteristic function

$$\phi(t) = \exp\{-|t|^{\alpha}\} \tag{9}$$

Despite its attractive properties like stability-under-addition, the second moment of the α-stable distribution does not exist in most cases.[4] Consequently, the application of this distribution to the MRS-type model will cause serious problems. For instance, the interpretation of σ_ε^2 would fail. More importantly, the existence of σ_ε^2 is required for the asymptotic properties of the MLE estimator to hold (Douc et al. 2004). Therefore, the sought of a substitute of the α-stable distribution, which has

[4] The second moment of the α-stable distribution only exists when $\alpha = 2$. In this case, the symmetric α-stable distribution collapses to a Gaussian distribution and cannot describe the fat fails.

similar attractive properties and defined moments, would be of particular interest and importance.

3.2.2 Tempered Stable Distribution

A natural candidate to address this issue is the tempered stable distribution. A general case of it is characterized by six parameters and denoted as $TS(\alpha^+, C^+, \lambda^+; \alpha^-, C^-, \lambda^-)$. The levy measure of such random variable x is

$$\nu(x) = \frac{c^-}{|x|^{1+\alpha^-}} e^{-\lambda^-|x|} 1_{x<0} + \frac{c^+}{|x|^{1+\alpha^+}} e^{-\lambda^+|x|} 1_{x>0} \tag{10}$$

Thus, a tempered stable distribution with zero mean has the following characteristic function (Cont and Tankov 2004).

$$\begin{aligned}\phi(t) = \exp \Bigg\{ &\Gamma(-\alpha^+)(\lambda^+)^{\alpha^+} C^+ \left[(1 - \frac{it}{\lambda^+})^{\alpha^+} - 1 + \frac{it\alpha^+}{\lambda^+} \right] \\ &+ \Gamma(-\alpha^-)(\lambda^-)^{\alpha^-} C^- \left[(1 - \frac{it}{\lambda^-})^{\alpha^-} - 1 + \frac{it\alpha^-}{\lambda^-} \right] \Bigg\}\end{aligned} \tag{11}$$

where $\alpha^+, \alpha^- < 2$ and $C^+, C^-, \lambda^+, \lambda^- > 0$. Its first four cumulants are hereby defined as:

$$\begin{aligned}\kappa_1 &= 0 \\ \kappa_2 &= \Gamma(2-\alpha^+)C^+(\lambda^+)^{\alpha^+-2} + \Gamma(2-\alpha^-)C^-(\lambda^-)^{\alpha^--2} \\ \kappa_3 &= \Gamma(3-\alpha^+)C^+(\lambda^+)^{\alpha^+-3} + \Gamma(3-\alpha^-)C^-(\lambda^-)^{\alpha^--3} \\ \kappa_4 &= \Gamma(4-\alpha^+)C^+(\lambda^+)^{\alpha^+-4} + \Gamma(4-\alpha^-)C^-(\lambda^-)^{\alpha^--4}\end{aligned} \tag{12}$$

Hence, the first four moments of x can be found via the following relations:

$$\begin{aligned}m_1 &= \kappa_1 \\ m_2 &= \kappa_2 + \kappa_1^2 \\ m_3 &= \kappa_3 + 3\kappa_2\kappa_1 + \kappa_1^3 \\ m_4 &= \kappa_4 + 4\kappa_3\kappa_1 + 3\kappa_2^2 + 6\kappa_2\kappa_1^2 + \kappa_1^4\end{aligned} \tag{13}$$

Clearly, the tempered stable distribution has defined moments, which enables it to be further employed to describe the innovation of the MRS-type model.

Küchler and Tappe (2013) suggest that sample paths of x can be classified as some well-known processes, which mainly depend on values of α^+ and α^-. In particular, if $\alpha^+ = \alpha^-$, x follows a classical tempered stable distribution. In addition, if we further require $C^+ = C^-$, then x follows a CGMY distribution (Carr et al. 2002).

- For $\alpha^+, \alpha^- < 0$ we have $\nu(\mathrm{R}) < \infty$, and thus, x is a compound Poisson process.
- For $\alpha^+, \alpha^- \in [0, 1)$ we have $\nu(\mathrm{R}) = \infty$, but $\int_{-1}^{1} |x|\nu(dx) < \infty$. Therefore, x is a finite-variation process making infinitely many jumps in each interval of positive length, which we can express as $x_t = \sum_{s \le t} \Delta x_s$.
- For $\alpha^+, \alpha^- \in (1, 2)$ we have $\int_{-1}^{1} |x|\nu(dx) = \infty$. Thus, x has sample paths of infinite variation.

To apply this distribution into the MRS model, we require x to be standardized. Bianchi et al. (2010) argues that one way to achieve the standardization is letting

$$C^+ = \frac{p(\lambda^+)^{2-\alpha^+}}{\Gamma(2-\alpha^+)} \quad \text{and} \quad C^- = \frac{(1-p)(\lambda^-)^{2-\alpha^-}}{\Gamma(2-\alpha^-)} \tag{14}$$

where $p \in (0, 1)$, then $x \sim TS(\alpha^+, \lambda^+, \alpha^-, \lambda^-, p)$ has zero mean and unit variance.

Combining Eqs. (11) and (14), we now have a standardized tempered stable distribution. This distribution is expected to retain all the attractive properties similar to those of α-stable distribution and has defined moments.

3.2.3 MRS Model with Tempered Stable Distribution

Since the tempered stable distribution has defined moments, the sufficient condition to ensure the asymptotic properties of MLE estimator of the MRS model will not be affected.[5] As the tempered stable distribution can effectively accommodate extreme behaviours like fat tails, we expect that the MLE of MRS model with this density would provide consistent estimators for financial data.

In terms of estimation, the procedures of MLE as discussed in Sect. 2 can still be used. Since the tempered stable distribution also does not have a closed form of density function, a certain numerical algorithm is needed to compute it (Kim et al. 2008). As argued by Mittnik et al. (1999), compared with other approximation methods, discrete Fourier transform is accurate and efficient to estimate parameters of stable family distributions, especially when $N = 2^{13}$ or above. Therefore, we employ the discrete Fourier transform method to obtain the estimated $f(\varepsilon_{s_t,t}|s_t, \theta, \Omega_{t-1})$ via the following iteration steps suggested by Shi and Feng (2016):

1. Acquiring the minimum and maximum of $\eta_t = \varepsilon_{s_t,t}/\sigma_\varepsilon$ as η_1 and η_2, respectively;
2. Calculating the values of $\phi(t)$ for the tempered stable distribution determined by the estimates of $\alpha^+, C^+, \lambda^+, \alpha^-, C^-$ and λ^- via Eq. (11), where t evenly ranges from η_1-0.1 to η_2+0.1 with size $N = 2^{15}$;[6]
3. Using discrete Fourier transform to find the values of the corresponding density function for t; and
4. Employing the linear interpolation to η_t that fall between the prespecified equally-spaced density values of t.

[5] Specific conditions can be found in Douc et al. (2004).

[6] As discussed in Mittnik et al. (1999), discrete Fourier transform works most efficiently for N being expressed in terms of a power of 2.

Hence, the interpolated values will be the estimates of $f(\varepsilon_{s_t,t}|s_t,\theta,\Omega_{t-1})$. By further applying them to Eqs. (3) and (4), the required log-likelihood values can be obtained.

4 Comparisons Between Distributions: Simulation Studies

In this section, we will conduct three simulation studies to compare the performance of the MRS models with Normal, Student's t, GED and tempered stable distributions. The data generation process is the 2-state MRS model with regime-switching in mean only as described equation (1) in all cases.[7] True distributions of the three studies are Student's t, GED and tempered stable, respectively. In terms of the simulation method, we adopt the inverse sampling to simulate the Student's t and GED densities. The approximative acceptance-rejection sampling method proposed by Baeumer and Meerschaert (2010) is employed for the tempered stable distribution.

4.1 Simulation Study: Student's t-Distribution

First, we set the true distribution as Student's t with 3 degrees of freedom.[8] Altogether, twelve sets of simulations of the MRS process with different p_{11}, p_{22}[9] and sample sizes T are generated, where $\mu_1 = -0.5$, $\mu_2 = 0.5$, $\sigma_\varepsilon^2 = 1$ and the number of replicates for each set is 300. To avoid the starting bias, 10000 points are generated for each simulation, and then only the last 3000, 4000 or 5000 are kept. Moreover, to avoid simulation bias, 500 such replicates are produced for each combination, while the first 200 are discarded.

The simulated data are fitted into the MRS models with Normal (MRS-N), Student's t (MRS-t), GED (MRS-G) and tempered stable (MRS-S) distributions, respectively. In Table 1, the mean log-likelihood (LL), bias, standard error (SE) and root-mean-square-error (RMSE) of μ_1, μ_2, p_{11}, p_{22} and σ_ε^2 obtained from the MRS-N, MRS-t and MRS-G models are reported. Bias is the mean difference between the true parameter and its estimates. SE is the standard error of the estimates. RMSE is the square root of the mean of squared difference between the true parameter and its estimates. The results of MRS-S model are presented in panel A of Table 4.

Although tempered stable distribution has four (five) more parameters than Student's t and GED (Normal), the log-likelihood is still a preliminary indicator of the model performance. Not surprisingly, MRS-N has the smallest log-likelihood in all cases, as the true distribution is fat-tailed. Besides, MRS-G leads to smaller log-likelihood than the true model. However, it is worth noticing that MRS-S can

7 The cases of regime-switching in variance only and in both mean and variance are also considered. The results are robust and available upon request.

8 We also consider the cases of 4, 5 and 6 degrees of freedom. The results are robust and available upon request.

9 We also conduct studies with smaller transition probabilities of 0.25, 0.5, 0.75 and 0.9. The results are robust and available upon request.

Table 1 Simulation results: Student's t-distribution

p_{11}	p_{22}	T	$Mean_{ll}$	$Bias_{\mu_1}$	SE_{μ_1}	$RMSE_{\mu_1}$	$Bias_{\mu_2}$	SE_{μ_2}	$RMSE_{\mu_2}$	$Bias_{p_{11}}$	$SE_{p_{11}}$	$RMSE_{p_{11}}$	$Bias_{p_{22}}$	$SE_{p_{22}}$	$RMSE_{p_{22}}$	$Bias_{\sigma^2}$	SE_{σ^2}	$RMSE_{\sigma^2}$
Panel A: Normal Distribution																		
0.999	0.99	3000	-4165	-0.1880	2.1714	2.1795	1.2792	4.4752	4.6544	-0.0088	0.0813	0.0818	-0.2042	0.3838	0.4347	-0.0846	0.1234	0.1496
		4000	-5592	-0.3992	3.9622	3.9822	1.7713	9.8237	9.9821	-0.0122	0.0994	0.1001	-0.1910	0.3731	0.4191	-0.0681	0.1174	0.1357
		5000	-7019	-0.8526	6.4299	6.4862	2.0167	8.6591	8.8908	-0.0214	0.1399	0.1415	-0.1397	0.3302	0.3585	-0.0529	0.0956	0.1092
0.99	0.999	3000	-4198	-1.4489	6.3714	6.5341	0.5608	6.0099	6.0360	-0.1658	0.3533	0.3903	-0.0122	0.0994	0.1002	-0.0643	0.1285	0.1437
		4000	-5603	-1.9501	8.3446	8.5694	0.5096	5.0461	5.0717	-0.1541	0.3452	0.3781	-0.0116	0.0994	0.1001	-0.0574	0.1113	0.1253
		5000	-7007	-1.0897	5.5912	5.6964	0.3613	4.3272	4.3422	-0.1512	0.3417	0.3736	-0.0060	0.0587	0.0590	-0.0584	0.1113	0.1257
0.99	0.99	3000	-4278	-0.1379	2.4391	2.4430	1.3016	7.7240	7.8329	-0.0052	0.0572	0.0574	-0.0318	0.1688	0.1718	-0.0433	0.1367	0.1434
		4000	-5705	-1.0672	7.6045	7.6790	0.3447	4.4594	4.4727	-0.0215	0.1386	0.1403	-0.0082	0.0806	0.0811	-0.0447	0.1111	0.1198
		5000	-7162	-0.7086	6.3366	6.3761	0.7563	6.6835	6.7262	-0.0148	0.1136	0.1146	-0.0147	0.1136	0.1145	-0.0325	0.1104	0.1151
0.999	0.999	3000	-4221	-0.4003	4.0158	4.0357	0.6774	7.7013	7.7310	-0.0258	0.1510	0.1532	-0.0157	0.1148	0.1158	-0.0337	0.1336	0.1377
		4000	-5633	-1.1575	9.6168	9.6862	0.5188	4.9330	4.9602	-0.0278	0.1605	0.1629	-0.0184	0.1280	0.1293	-0.0321	0.1168	0.1211
		5000	-7047	-0.7321	6.4486	6.4900	0.7848	6.1809	6.2305	-0.0146	0.1147	0.1156	-0.0178	0.1280	0.1292	-0.0282	0.1195	0.1228
Panel B: Student's t-Distribution																		
0.999	0.99	3000	-3695	-0.0020	0.0135	0.0136	0.0053	0.0948	0.0950	-0.0002	0.0008	0.0008	-0.0101	0.0313	0.0329	0.0018	0.1023	0.1023
		4000	-4928	0.0010	0.0117	0.0118	-0.0029	0.0541	0.0542	-0.0001	0.0006	0.0006	-0.0046	0.0135	0.0143	0.0096	0.0959	0.0964
		5000	-6161	-0.0005	0.0106	0.0106	-0.0036	0.0650	0.0651	-0.0001	0.0006	0.0006	-0.0032	0.0106	0.0111	0.0070	0.0736	0.0739
0.99	0.999	3000	-3695	0.0139	0.0945	0.0956	0.0001	0.0136	0.0136	-0.0087	0.0263	0.0277	-0.0002	0.0006	0.0006	0.0059	0.1070	0.1071
		4000	-4929	0.0010	0.0506	0.0506	0.0014	0.0118	0.0119	-0.0032	0.0118	0.0122	-0.0000	0.0006	0.0006	0.0052	0.0852	0.0854
		5000	-6164	0.0036	0.0456	0.0458	-0.0004	0.0108	0.0108	-0.0038	0.0149	0.0154	-0.0001	0.0005	0.0005	0.0081	0.0827	0.0831
0.99	0.99	3000	-3779	-0.0003	0.0189	0.0189	-0.0017	0.0188	0.0189	-0.0006	0.0029	0.0030	-0.0004	0.0030	0.0030	0.0051	0.1049	0.1051
		4000	-5033	-0.0001	0.0173	0.0173	0.0028	0.0171	0.0173	-0.0000	0.0027	0.0027	-0.0001	0.0025	0.0026	0.0012	0.0851	0.0851
		5000	-6293	-0.0015	0.0148	0.0149	0.0016	0.0148	0.0149	-0.0003	0.0023	0.0023	-0.0002	0.0023	0.0023	0.0034	0.0759	0.0760
0.999	0.999	3000	-3692	-0.0004	0.0311	0.0311	-0.0011	0.0230	0.0231	-0.0012	0.0061	0.0062	-0.0005	0.0017	0.0018	0.0160	0.1031	0.1043
		4000	-4922	0.0006	0.0402	0.0402	0.0003	0.0175	0.0175	-0.0004	0.0020	0.0020	-0.0005	0.0021	0.0022	0.0161	0.0911	0.0925
		5000	-6157	-0.0008	0.0148	0.0148	0.0018	0.0169	0.0170	-0.0003	0.0011	0.0012	-0.0003	0.0011	0.0011	0.0089	0.0782	0.0788
Panel c: GED Distribution																		
0.999	0.99	3000	-3740	-0.0042	0.0269	0.0272	0.0034	0.1109	0.1109	-0.0044	0.0441	0.0444	-0.0119	0.0547	0.0560	-0.1674	0.0538	0.1758
		4000	-4992	0.0002	0.0127	0.0127	-0.0022	0.0636	0.0637	-0.0001	0.0006	0.0006	-0.0042	0.0126	0.0133	-0.1620	0.0482	0.1690
		5000	-6242	-0.0004	0.0126	0.0126	0.0003	0.0499	0.0499	-0.0000	0.0005	0.0005	-0.0027	0.0098	0.0101	-0.1604	0.0381	0.1649
0.99	0.999	3000	-3743	0.0153	0.0850	0.0864	0.0015	0.0236	0.0236	-0.0112	0.0546	0.0558	-0.0029	0.0349	0.0350	-0.1624	0.0529	0.1708
		4000	-4992	0.0070	0.0692	0.0696	0.0016	0.0136	0.0137	-0.0093	0.0720	0.0726	-0.0000	0.0006	0.0006	-0.1613	0.0437	0.1671
		5000	-6243	0.0065	0.0500	0.0504	-0.0011	0.0126	0.0127	-0.0059	0.0538	0.0541	-0.0000	0.0005	0.0005	-0.1611	0.0436	0.1669
0.99	0.99	3000	-3825	0.0002	0.0217	0.0217	-0.0022	0.0213	0.0214	-0.0003	0.0028	0.0029	-0.0002	0.0029	0.0029	-0.1618	0.0531	0.1703
		4000	-5096	0.0008	0.0196	0.0197	0.0014	0.0199	0.0200	0.0001	0.0026	0.0026	0.0001	0.0025	0.0025	-0.1625	0.0434	0.1682
		5000	-6372	0.0008	0.0172	0.0172	0.0003	0.0176	0.0176	-0.0001	0.0023	0.0023	0.0000	0.0022	0.0022	-0.1613	0.0392	0.1660
0.999	0.999	3000	-3740	0.0000	0.0381	0.0381	-0.0013	0.0264	0.0264	-0.0011	0.0055	0.0056	-0.0005	0.0017	0.0018	-0.1586	0.0529	0.1672
		4000	-4986	0.0022	0.0529	0.0529	0.0006	0.0270	0.0270	-0.0019	0.0272	0.0273	-0.0023	0.0302	0.0303	-0.1585	0.0466	0.1652
		5000	-6238	-0.0007	0.0179	0.0180	0.0016	0.0198	0.0199	-0.0003	0.0011	0.0011	-0.0003	0.0010	0.0011	-0.1587	0.0425	0.1643

yield slightly greater log-likelihood compared with MRS-t. Therefore, it seems that MRS-S model might provide quite satisfied results, even when the true distribution is not tempered stable.

4.1.1 Results of Mean and Variance

As discussed in Sect. 3, ML estimators of the MRS-N model may not be consistent for fat-tailed data. It is evidenced by the large absolute biases observed in all cases of μ_1 and μ_2. some of them even exceed 1. As to the fat-tail models, both MRS-G and MRS-S generate much better estimates. Most of the absolute biases are less than 0.01. For σ_ε^2, however, the story is different. MRS-N and MRS-S yield similar results of bias, which are not too far from those of the MRS-t. The estimates of MRS-G are much smaller than the true values in all cases, with an absolute bias around 0.16 in all cases. Hence, MRS-S can lead to consistent estimators for both the mean and variance. MRS-G (MRS-N) can only provide consistent estimators for the mean (variance), when the true model is MRS-t.

SE is a measurement of the estimation efficiency. It is well known that the 'wrong' model is generally less efficient than the true model. For MRS-N, its SEs are on a much larger scale than the other models for μ_1 and μ_2. All of them are greater than 2, and some are even close to 10. On the other hand, MRS-G and MRS-S still lead to similar results. Although their SEs of the μ_1 and μ_2 are slightly greater than those of the true model, they are basically on the same scale, and none of them is greater than 0.1. For the estimation of variance, MRS-G generates smaller SEs than the true model in all cases. The results of MRS-S are overall better than those of MRS-N and are not far from those of MRS-t. Therefore, MRS-S model can generate similar SEs as the true model to estimate of both the mean and variance. The efficiency of MRS-N is low for the estimation of mean and is acceptable for the estimation of variance. MRS-G has similar efficiency to MRS-S for estimating mean, and it can be more efficient than the true model to estimate variance. However, since the bias of variance produced by MRS-G is quite large, its overall performance of estimating variance may not be optimal.

Finally, RMSE is a combination of bias and SE, which is widely employed as the overall performance indicator in the existing literature. Accordant with the above results, MRS-N is the least preferred model in the estimation of μ_1 and μ_2, while the performance of MRS-G and MRS-S are quite similar and close to that of the true model. As to the variance, MRS-N and MRS-S lead to acceptable results, which are not too far from those of MRS-t. MRS-G, however, tends to generate largest RMSEs in all cases.

4.1.2 Results of Transition Probabilities

MRS-N cannot produce consistent estimators for both transition probabilities in all cases. Its absolute biases are much larger than those of the other models, with the maximum exceeding 0.2. MRS-G and MRS-S can generate consistent estimators in all sets, and their results are very close to those of MRS-t. Their absolute biases are almost all under 0.01. Also, it is worth noticing that MRS-S performs slightly better than MRS-G in most cases.

The efficiency of MRS-N is still the worst among all models. Most of its SEs are over ten times of those of the other models. The three fat-tailed models have similar efficiency, and estimators of MRS-S are still overall more efficient than those of MRS-G. For instance, some SEs of MRS-G are over 0.05, while those of MRS-S are mostly under 0.01. Finally, the RMSE suggests that MRS-N is the least preferred model to estimate transition probabilities in all cases. The performance of the other three are close to each other, while MRS-S slightly outperforms MRS-G.

In conclusion, in the estimation of mean, variance and transition probabilities, MRS-S model can uniformly generate very similar estimates to those of the true model. MRS-N model cannot perform as well as the fat-tailed models in most cases. The estimates of mean and transition probabilities generated by MRS-G are mostly acceptable. But it leads to the worst estimates of variance. Therefore, considering the overall performance of all parameters, MRS-S is preferred to its competitors MRS-G and MRS-N. Additionally, its results are very close to those of the true model in all cases.

4.2 Simulation Study: GED

Next, we set the true distribution as GED with 1.1 degree of freedom.[10] Twelve sets of simulations with the same combinations of parameters as those in Sect. 4.1 are constructed. Replicates and each simulation are also truncated in the same manners to avoid simulation bias.

Simulation results are reported in Table 2 and panel B of Table 4. As preliminary evidence of model performance, log-likelihood results are accordant with those presented in Sect. 4.1. More specifically, MRS-N is the least preferred model, while MRS-t cannot perform as well as the true model in all cases. The results of MRS-S are overall slightly better than those of MRS-G.

4.2.1 Results of Mean and Variance

In the case of bias comparison, MRS-N still leads to much larger absolute values than the others for μ_1 and μ_2. Some of them are greater than 0.1. MRS-t and MRS-S have similar results, which are very close to those of the true model. Most of those absolute biases are less than 0.01. As to the variance, MRS-t cannot provide consistent estimators, since its absolute biases are remarkably large (at around 0.2). MRS-N and MRS-S generate similar biases as the true model (below 0.01). Hence, MRS-S can lead to consistent estimators for both the mean and variance. MRS-t (MRS-N) can only provide consistent estimators for the mean (variance).

Regarding the efficiency, MRS-N cannot perform as well as the others in the estimation of mean. Some of its SEs can be above 0.5. The results of MRS-t and

[10] We also consider the cases of 1.3, 1.5 and 1.8 degrees of freedom. The results are robust and available upon request.

Table 2 Simulation results: GED

p_{11}	p_{22}	T	$Mean_{ll}$	$Bias_{\mu_1}$	SE_{μ_1}	$RMSE_{\mu_1}$	$Bias_{\mu_2}$	SE_{μ_2}	$RMSE_{\mu_2}$	$Bias_{p_{11}}$	$SE_{p_{11}}$	$RMSE_{p_{11}}$	$Bias_{p_{22}}$	$SE_{p_{22}}$	$RMSE_{p_{22}}$	$Bias_{\sigma^2}$	SE_{σ^2}	$RMSE_{\sigma^2}$
Panel A: Normal Distribution																		
0.999	0.99	3000	-4269	-0.0054	0.0204	0.0211	0.1326	0.5293	0.5457	-0.0015	0.0042	0.0045	-0.0659	0.2281	0.2375	-0.0149	0.0490	0.0512
		4000	-5700	-0.0040	0.0173	0.0178	0.0856	0.4424	0.4506	-0.0009	0.0028	0.0030	-0.0430	0.1844	0.1894	-0.0079	0.0401	0.0409
		5000	-7134	-0.0028	0.0162	0.0165	0.0455	0.2722	0.2760	-0.0008	0.0032	0.0033	-0.0281	0.1428	0.1455	-0.0029	0.0398	0.0399
0.99	0.999	3000	-4272	-0.0710	0.3539	0.3609	0.0059	0.0210	0.0218	-0.0384	0.1684	0.1727	-0.0010	0.0033	0.0034	-0.0102	0.0481	0.0492
		4000	-5708	-0.1254	0.4800	0.4961	0.0043	0.0197	0.0202	-0.0737	0.2407	0.2517	-0.0015	0.0044	0.0047	-0.0071	0.0491	0.0496
		5000	-7127	-0.0415	0.2787	0.2818	0.0035	0.0160	0.0164	-0.0241	0.1318	0.1340	-0.0006	0.0026	0.0026	-0.0052	0.0374	0.0377
0.99	0.99	3000	-4346	-0.0059	0.0282	0.0289	0.0036	0.0289	0.0292	-0.0014	0.0035	0.0038	-0.0016	0.0037	0.0040	-0.0025	0.0419	0.0420
		4000	-5796	-0.0032	0.0243	0.0245	0.0069	0.0245	0.0255	-0.0011	0.0031	0.0033	-0.0013	0.0029	0.0031	-0.0018	0.0375	0.0375
		5000	-7242	-0.0017	0.0228	0.0229	0.0015	0.0215	0.0216	-0.0009	0.0028	0.0029	-0.0008	0.0025	0.0027	-0.0017	0.0314	0.0315
0.999	0.999	3000	-4269	-0.0068	0.1039	0.1041	0.0144	0.1590	0.1597	-0.0043	0.0574	0.0575	-0.0072	0.0778	0.0781	-0.0037	0.0439	0.0441
		4000	-5696	0.0033	0.0270	0.0273	-0.0007	0.0265	0.0265	-0.0009	0.0051	0.0052	-0.0007	0.0023	0.0024	-0.0011	0.0306	0.0306
		5000	-7113	-0.0010	0.0257	0.0257	0.0070	0.1164	0.1166	-0.0005	0.0021	0.0021	-0.0035	0.0537	0.0538	-0.0035	0.0339	0.0341
Panel B: Student's t-Distribution																		
0.999	0.99	3000	-4085	-0.0008	0.0156	0.0156	0.0080	0.0950	0.0953	-0.0003	0.0018	0.0018	-0.0121	0.0587	0.0599	0.1855	0.0871	0.2049
		4000	-5452	-0.0006	0.0129	0.0129	-0.0006	0.0591	0.0591	-0.0002	0.0007	0.0007	-0.0060	0.0200	0.0209	0.1862	0.0688	0.1985
		5000	-6826	0.0001	0.0115	0.0115	0.0025	0.0684	0.0684	-0.0002	0.0017	0.0017	-0.0072	0.0537	0.0542	0.1874	0.0685	0.1996
0.99	0.999	3000	-4090	0.0016	0.0906	0.0906	0.0020	0.0151	0.0152	-0.0114	0.0617	0.0627	-0.0002	0.0007	0.0007	0.1896	0.0949	0.2120
		4000	-5461	-0.0005	0.0882	0.0882	-0.0002	0.0139	0.0139	-0.0178	0.0888	0.0906	-0.0004	0.0026	0.0026	0.1950	0.0759	0.2093
		5000	-6818	-0.0008	0.0551	0.0551	0.0001	0.0121	0.0121	-0.0074	0.0582	0.0587	-0.0001	0.0005	0.0005	0.1848	0.0641	0.1956
0.99	0.99	3000	-4168	-0.0038	0.0217	0.0220	0.0009	0.0219	0.0220	-0.0008	0.0032	0.0033	-0.0009	0.0034	0.0035	0.1824	0.0819	0.1999
		4000	-5558	-0.0013	0.0188	0.0188	0.0023	0.0186	0.0187	-0.0005	0.0028	0.0028	-0.0006	0.0027	0.0027	0.1845	0.0751	0.1992
		5000	-6944	-0.0001	0.0179	0.0179	-0.0007	0.0165	0.0165	-0.0003	0.0024	0.0024	-0.0003	0.0024	0.0024	0.1814	0.0669	0.1933
0.999	0.999	3000	-4082	-0.0008	0.0260	0.0260	0.0032	0.0559	0.0560	-0.0012	0.0106	0.0107	-0.0034	0.0409	0.0410	0.1990	0.0912	0.2189
		4000	-5444	0.0028	0.0208	0.0210	-0.0024	0.0207	0.0208	-0.0006	0.0035	0.0035	-0.0006	0.0020	0.0021	0.1905	0.0734	0.2042
		5000	-6802	-0.0015	0.0247	0.0247	0.0003	0.0183	0.0183	-0.0004	0.0016	0.0017	-0.0008	0.0081	0.0081	0.1879	0.0753	0.2025
Panel c: GED Distribution																		
0.999	0.99	3000	-4058	0.0003	0.0136	0.0136	-0.0004	0.0740	0.0740	-0.0002	0.0007	0.0007	-0.0068	0.0162	0.0176	-0.0031	0.0393	0.0394
		4000	-5415	0.0001	0.0115	0.0115	-0.0008	0.0544	0.0544	-0.0001	0.0006	0.0006	-0.0054	0.0184	0.0192	-0.0010	0.0326	0.0326
		5000	-6780	-0.0001	0.0104	0.0104	-0.0020	0.0489	0.0490	-0.0001	0.0005	0.0005	-0.0049	0.0208	0.0214	0.0032	0.0310	0.0312
0.99	0.999	3000	-4061	0.0034	0.0724	0.0725	0.0011	0.0128	0.0128	-0.0051	0.0134	0.0143	-0.0002	0.0006	0.0007	-0.0012	0.0417	0.0417
		4000	-5424	0.0041	0.0647	0.0648	-0.0002	0.0124	0.0124	-0.0078	0.0213	0.0227	-0.0001	0.0006	0.0006	0.0049	0.0356	0.0359
		5000	-6772	0.0020	0.0438	0.0439	-0.0002	0.0108	0.0108	-0.0038	0.0115	0.0121	-0.0001	0.0005	0.0005	-0.0000	0.0314	0.0314
0.99	0.99	3000	-4141	-0.0019	0.0192	0.0193	-0.0004	0.0201	0.0201	-0.0006	0.0031	0.0032	-0.0007	0.0033	0.0034	0.0022	0.0416	0.0417
		4000	-5521	-0.0011	0.0159	0.0159	0.0012	0.0176	0.0177	-0.0004	0.0027	0.0028	-0.0005	0.0026	0.0027	0.0029	0.0361	0.0362
		5000	-6899	0.0005	0.0158	0.0158	-0.0009	0.0148	0.0149	-0.0001	0.0024	0.0024	-0.0001	0.0023	0.0023	0.0020	0.0305	0.0306
0.999	0.999	3000	-4054	0.0002	0.0262	0.0262	-0.0000	0.0274	0.0274	-0.0014	0.0147	0.0148	-0.0010	0.0068	0.0068	-0.0008	0.0409	0.0409
		4000	-5408	0.0026	0.0216	0.0218	-0.0019	0.0186	0.0187	-0.0006	0.0033	0.0034	-0.0006	0.0020	0.0020	-0.0007	0.0299	0.0299
		5000	-6756	-0.0009	0.0221	0.0221	0.0002	0.0171	0.0171	-0.0004	0.0016	0.0016	-0.0008	0.0084	0.0085	-0.0024	0.0329	0.0330

MRS-S are again similar and close to those of the true model. All of them are below 0.1. As to the variance, SEs of MRS-t are overall close to 0.1, while the others are mostly below 0.05. Consequently, RMSE suggests that MRS-N (MRS-t) is the least preferred model in the estimation of the mean (variance). The overall performance of MRS-S is very close to that of the true model in all cases.

4.2.2 Results of Transition Probabilities

Most of the absolute biases obtained from the MRS-N model are below 0.01, but are still the largest among all models. Taking the sets where $p_{11} \neq p_{22}$ as examples, the absolute biases of MRS-N can be larger than 0.06. The results of the other models are less than 0.001 in most cases. Hence, the results of MRS-N are still relatively worse than those of the fat-tailed models.

SEs of the MRS-N are small for the sets where $p_{11} = p_{22}$, but are still considerably large in the other cases. Additionally, all of them are greater than those of the fat-tailed models. MRS-t and MRS-S lead to similar results which are close to those of MRS-G. It is worth noticing that the SEs of MRS-S are relatively smaller than those of MRS-t, especially in the cases where $p_{11} \neq p_{22}$. Therefore, as indicated by RMSE, MRS-S is preferred to the MRS-N and MRS-t in the estimation of transition probabilities. The results of MRS-S are also close to those of the true model in all cases.

In conclusion, in the estimation of mean, variance and transition probabilities, MRS-S model can uniformly generate very similar estimates to those of MRS-G. MRS-N model cannot perform as well as the fat-tailed models in most cases. The estimates of mean and transition probabilities generated by MRS-t are mostly acceptable, but it has the worst estimates of variance. Therefore, considering the overall performance of all parameters, MRS-S outperforms its competitors MRS-t and MRS-N. Apart from that, the results of MRS-S are very close to those of the true model in all cases.

4.3 Simulation Study: Tempered Stable Distribution

Finally, we consider the scenario that the true distribution is tempered stable with three sets of different parameters (where all the ps are set to be 0.5), including one case of the CGMY distribution (when $\alpha^+ = \alpha^- = 0.5$ and $\lambda^+ = \lambda^- = 1.0$) and two general cases. Altogether, twelve sets of simulations are constructed. All the sample sizes are set to 5000. True values of p_{11}, p_{22}, μ_1, μ_2 and σ_ε^2 are not changed. Replicates and each simulation are further truncated in the same manners as in Sects. 4.1 and 4.2 to avoid simulation bias.

Among the existing literature, there are main methods to simulate the tempered stable random variates. They include acceptance-rejection sampling, Gaussian approximation of a small jump component and infinite shot noise series representations (see Kawai and Masuda 2011 for an overview of those approaches). Given a specified computing budget, Kawai and Masuda (2011) suggest that the approximative acceptance-rejection sampling procedure proposed by Baeumer and

Meerschaert (2010) is the most efficient and handiest among the existing methods. Additionally, it is suggested that using this sampling procedure, a high-level accuracy can be reached with only a small amount of extra computing effort. For those reasons, we employ this sampling method to simulate the random disturbance following tempered stable distributions in this paper.

Simulation results are reported in Table 3 and panel C of Table 4. Not surprisingly, log-likelihood values of MRS-N, MRS-t and MRS-G are all smaller than those of the true model. Also, MRS-G outperforms MRS-t only in the CGMY distribution cases, while MRS-t generates lager log-likelihood in the other two general cases.

4.3.1 Results of Mean and Variance

In the case of bias comparison, MRS-N leads to the worse results for μ_1 and μ_2 with the largest values close to 0.8. MRS-t and MRS-G have similar absolute biases which are not far from those of the true model. As to the variance, the absolute biases of MRS-N is smaller than those of the MRS-t and MRS-G. Most absolute biases of MRS-t (MRS-G) can be as large as 0.4 (0.1). Therefore, estimators of MRS-N (MRS-t and MRS-G) are only consistent in the case of variance (mean).

The story of SE is slightly different. MRS-N is the least efficient model in the estimation of mean in most cases. Its largest SE is even above 1. The SEs of MRS-t and MRS-G are close to those of the true model. Most of them are at around 0.02. For the estimation of variance, MRS-N and MRS-G are more efficient than MRS-t. Most of their SEs are below 0.05, while SEs of MRS-t are generally greater than 0.1. In terms of the overall performance, MRS-N is not preferred in the estimation of mean, while it outperforms MRS-t and MRS-G in the estimation of variance.

4.3.2 Results of Transition Probabilities

The absolute biases of MRS-N are below 0.01 when $p_{11} = p_{22}$, but can be above 0.2 when $p_{11} \neq p_{22}$. The results of MRS-t and MRS-G are similar and close to those of the true model (almost all of them are below 0.03). Hence, MRS-N still cannot provide consistent estimators as the fat-tailed models in the estimation of transition probabilities.

The situation of SEs is basically the same as that of the biases. The SEs of MRS-N are below 0.01 when $p_{11} = p_{22}$, but can be as large as 0.4 when $p_{11} \neq p_{22}$. MRS-t and MRS-G lead to similar SEs which are not far from those of MRS-S. Their SEs are generally smaller than those of MRS-N. Therefore, RMSE suggests MRS-N is the least preferred model in the estimation of transition probabilities. MRS-t and MRS-G can provide similar results as the true model.

To summarise, in terms of the RMSE, when the true distribution is Student's t (GED), MRS-S model uniformly outperforms the competing models MRS-N and MRS-G (MRS-N and MRS-t) in the estimation of all parameters. Also, the results of MRS-S and those of the true model are very close in most situations. Besides, MRS-S can even generate larger log-likelihood than the true model. When the true distribution is tempered stable, MRS-N, MRS-t and MRS-G cannot uniformly

Table 3 Simulation results: tempered stable distribution

p_{11}	p_{22}	a_1	a_2	λ_1	λ_2	$Mean_{ll}$	$Bias_{\mu_1}$	SE_{μ_1}	$RMSE_{\mu_1}$	$Bias_{\mu_2}$	SE_{μ_2}	$RMSE_{\mu_2}$	$Bias_{p_{11}}$	$SE_{p_{11}}$	$RMSE_{p_{11}}$	$Bias_{p_{22}}$	$SE_{p_{22}}$	$RMSE_{p_{22}}$	$Bias_{\sigma^2}$	SE_{σ^2}	$RMSE_{\sigma^2}$
Panel A: Normal Distribution																					
0.999	0.99	0.5	0.5	1.0	1.0	-7122	-0.0039	0.0160	0.0165	0.0572	0.3296	0.3345	-0.0007	0.0025	0.0026	-0.0288	0.1462	0.1491	-0.0083	0.0383	0.0392
		0.2	1.2	1.2	0.2	-7035	0.0011	0.0165	0.0165	0.1128	0.4294	0.4440	-0.0011	0.0028	0.0030	-0.0438	0.1802	0.1854	-0.0458	0.0519	0.0692
		1.2	0.2	0.2	1.2	-6990	-0.0174	0.0179	0.0250	0.7974	1.4052	1.6156	-0.0039	0.0048	0.0061	-0.2421	0.4110	0.4770	-0.0834	0.0735	0.1112
0.99	0.999	0.5	0.5	1.0	1.0	-7130	-0.0791	0.4005	0.4082	0.0035	0.0165	0.0169	-0.0367	0.1715	0.1754	-0.0008	0.0026	0.0027	-0.0053	0.0423	0.0426
		0.2	1.2	1.2	0.2	-6995	-0.8170	1.4433	1.6585	0.0199	0.0219	0.0296	-0.2406	0.4092	0.4747	-0.0035	0.0038	0.0051	-0.0809	0.0734	0.1092
		1.2	0.2	0.2	1.2	-7035	-0.2234	0.6229	0.6617	0.0019	0.0166	0.0167	-0.0947	0.2736	0.2896	-0.0019	0.0048	0.0052	-0.0498	0.0586	0.0769
0.99	0.99	0.5	0.5	1.0	1.0	-7242	-0.0052	0.0209	0.0215	0.0033	0.0222	0.0225	-0.0010	0.0025	0.0027	-0.0012	0.0027	0.0029	-0.0032	0.0355	0.0356
		0.2	1.2	1.2	0.2	-7159	-0.0039	0.0200	0.0203	0.0192	0.0215	0.0289	-0.0021	0.0029	0.0036	-0.0023	0.0026	0.0035	-0.0409	0.0460	0.0615
		1.2	0.2	0.2	1.2	-7152	-0.0191	0.0197	0.0274	0.0037	0.0217	0.0220	-0.0023	0.0027	0.0035	-0.0021	0.0028	0.0035	-0.0436	0.0489	0.0655
0.999	0.999	0.5	0.5	1.0	1.0	-7115	0.0008	0.0218	0.0218	-0.0029	0.0277	0.0278	-0.0002	0.0010	0.0011	-0.0006	0.0023	0.0023	-0.0026	0.0321	0.0322
		0.2	1.2	1.2	0.2	-7029	0.0039	0.0234	0.0238	0.0247	0.1689	0.1707	-0.0010	0.0024	0.0026	-0.0076	0.0805	0.0809	-0.0388	0.0462	0.0603
		1.2	0.2	0.2	1.2	-7027	-0.0077	0.0231	0.0243	0.0145	0.2464	0.2468	-0.0006	0.0013	0.0014	-0.0074	0.0814	0.0817	-0.0388	0.0490	0.0625
Panel B: Student's t-Distribution																					
0.999	0.99	0.5	0.5	1.0	1.0	-6773	-0.0006	0.0128	0.0128	-0.0015	0.0524	0.0524	-0.0001	0.0005	0.0005	-0.0036	0.0108	0.0114	0.1599	0.0647	0.1725
		0.2	1.2	1.2	0.2	-6292	-0.0356	0.0120	0.0375	-0.0359	0.0445	0.0571	-0.0006	0.0048	0.0049	-0.0088	0.0535	0.0543	0.4614	0.2035	0.5043
		1.2	0.2	0.2	1.2	-6295	0.0344	0.0106	0.0360	0.0376	0.0477	0.0608	-0.0000	0.0005	0.0005	-0.0031	0.0091	0.0096	0.4610	0.2355	0.5177
0.99	0.999	0.5	0.5	1.0	1.0	-6777	-0.0002	0.0497	0.0497	-0.0000	0.0127	0.0127	-0.0039	0.0122	0.0128	-0.0001	0.0005	0.0005	0.1654	0.0643	0.1775
		0.2	1.2	1.2	0.2	-6296	-0.0323	0.0537	0.0627	-0.0344	0.0114	0.0362	-0.0038	0.0147	0.0152	-0.0000	0.0005	0.0005	0.4754	0.2088	0.5192
		1.2	0.2	0.2	1.2	-6298	0.0298	0.0581	0.0653	0.0373	0.0130	0.0395	-0.0229	0.1175	0.1197	-0.0017	0.0108	0.0109	0.4654	0.1911	0.5031
0.99	0.99	0.5	0.5	1.0	1.0	-6899	-0.0017	0.0168	0.0169	-0.0003	0.0174	0.0174	-0.0003	0.0022	0.0022	-0.0004	0.0024	0.0024	0.1597	0.0690	0.1739
		0.2	1.2	1.2	0.2	-6434	-0.0387	0.0142	0.0412	-0.0367	0.0153	0.0398	-0.0004	0.0022	0.0022	-0.0003	0.0022	0.0022	0.4677	0.1954	0.5069
		1.2	0.2	0.2	1.2	-6428	0.0357	0.0141	0.0384	0.0372	0.0139	0.0397	-0.0004	0.0024	0.0024	-0.0005	0.0021	0.0021	0.4584	0.2072	0.5031
0.999	0.999	0.5	0.5	1.0	1.0	-6756	0.0017	0.0182	0.0183	-0.0030	0.0221	0.0223	-0.0007	0.0100	0.0100	-0.0004	0.0014	0.0015	0.1646	0.0694	0.1786
		0.2	1.2	1.2	0.2	-6281	-0.0345	0.0177	0.0388	-0.0335	0.0187	0.0383	-0.0007	0.0065	0.0066	-0.0035	0.0521	0.0523	0.4507	0.1822	0.4862
		1.2	0.2	0.2	1.2	-6276	0.0356	0.0153	0.0387	0.0347	0.0170	0.0386	-0.0002	0.0008	0.0008	-0.0004	0.0016	0.0016	0.4653	0.2276	0.5180
Panel C: GED Distribution																					
0.999	0.99	0.5	0.5	1.0	1.0	-6758	-0.0009	0.0134	0.0134	-0.0032	0.0524	0.0525	-0.0001	0.0005	0.0005	-0.0034	0.0108	0.0113	-0.0177	0.0318	0.0364
		0.2	1.2	1.2	0.2	-6315	-0.0610	0.0131	0.0624	-0.0648	0.0539	0.0842	-0.0008	0.0093	0.0093	-0.0084	0.0538	0.0544	-0.1010	0.0346	0.1067
		1.2	0.2	0.2	1.2	-6318	0.0606	0.0117	0.0617	0.0638	0.0493	0.0806	-0.0000	0.0005	0.0005	-0.0027	0.0084	0.0088	-0.0975	0.0387	0.1049
0.99	0.999	0.5	0.5	1.0	1.0	-6762	0.0024	0.0515	0.0516	-0.0003	0.0123	0.0123	-0.0034	0.0108	0.0113	-0.0001	0.0005	0.0005	-0.0151	0.0315	0.0349
		0.2	1.2	1.2	0.2	-6319	-0.0536	0.0639	0.0834	-0.0603	0.0131	0.0617	-0.0047	0.0342	0.0346	0.0000	0.0005	0.0005	-0.0950	0.0378	0.1022
		1.2	0.2	0.2	1.2	-6321	0.0717	0.0739	0.1029	0.0621	0.0149	0.0639	-0.0146	0.0739	0.0753	-0.0023	0.0196	0.0197	-0.0990	0.0372	0.1058
0.99	0.99	0.5	0.5	1.0	1.0	-6886	-0.0012	0.0179	0.0179	-0.0014	0.0175	0.0175	-0.0001	0.0021	0.0021	-0.0002	0.0024	0.0024	-0.0154	0.0330	0.0364
		0.2	1.2	1.2	0.2	-6456	-0.0638	0.0166	0.0659	-0.0643	0.0172	0.0666	-0.0002	0.0021	0.0021	-0.0000	0.0021	0.0021	-0.0948	0.0367	0.1016
		1.2	0.2	0.2	1.2	-6451	0.0633	0.0161	0.0653	0.0626	0.0169	0.0649	-0.0002	0.0023	0.0023	-0.0004	0.0020	0.0021	-0.0984	0.0387	0.1057
0.999	0.999	0.5	0.5	1.0	1.0	-6742	0.0023	0.0183	0.0184	-0.0024	0.0222	0.0223	-0.0001	0.0008	0.0008	-0.0004	0.0014	0.0015	-0.0189	0.0300	0.0355
		0.2	1.2	1.2	0.2	-6303	-0.0600	0.0189	0.0630	-0.0598	0.0193	0.0629	-0.0007	0.0058	0.0058	-0.0035	0.0526	0.0527	-0.1001	0.0358	0.1063
		1.2	0.2	0.2	1.2	-6299	0.0612	0.0180	0.0638	0.0598	0.0212	0.0635	-0.0002	0.0008	0.0008	-0.0004	0.0015	0.0016	-0.1007	0.0372	0.1073

Table 4 Summary of the MRS-S models

p_{11}	p_{22}	T	a_1	a_2	λ_1	λ_2	$Mean_{ll}$	$Bias_{\mu_1}$	SE_{μ_1}	$RMSE_{\mu_1}$	$Bias_{\mu_2}$	SE_{μ_2}	$RMSE_{\mu_2}$	$Bias_{p_{11}}$	$SE_{p_{11}}$	$RMSE_{p_{11}}$	$Bias_{p_{22}}$	$SE_{p_{22}}$	$RMSE_{p_{22}}$	$Bias_{\sigma^2}$	SE_{σ^2}	$RMSE_{\sigma^2}$
Panel A: Student's t-Distribution																						
0.999	0.99	3000					-3691	-0.0028	0.0216	0.0218	-0.0086	0.1142	0.1145	-0.0002	0.0009	0.0009	-0.0115	0.0401	0.0417	-0.0325	0.1691	0.1722
		4000					-4924	0.0028	0.0181	0.0184	-0.0028	0.0558	0.0559	-0.0001	0.0006	0.0006	-0.0046	0.0135	0.0143	-0.0411	0.1360	0.1420
		5000					-6157	-0.0004	0.0153	0.0153	-0.0018	0.0412	0.0412	-0.0001	0.0005	0.0005	-0.0032	0.0107	0.0112	-0.0500	0.0968	0.1090
0.99	0.999	3000					-3690	0.0135	0.0783	0.0794	0.0001	0.0222	0.0222	-0.0081	0.0245	0.0258	-0.0002	0.0006	0.0007	-0.0213	0.1531	0.1546
		4000					-4924	0.0032	0.0518	0.0519	0.0019	0.0187	0.0188	-0.0033	0.0121	0.0125	-0.0000	0.0006	0.0006	-0.0465	0.1198	0.1285
		5000					-6160	0.0056	0.0467	0.0470	-0.0002	0.0165	0.0165	-0.0040	0.0162	0.0167	-0.0001	0.0005	0.0005	-0.0489	0.0950	0.1068
0.99	0.99	3000					-3774	0.0000	0.0234	0.0234	-0.0019	0.0237	0.0238	-0.0006	0.0030	0.0030	-0.0005	0.0030	0.0030	-0.0370	0.1370	0.1419
		4000					-5029	-0.0002	0.0213	0.0213	0.0025	0.0228	0.0230	-0.0001	0.0027	0.0027	-0.0001	0.0026	0.0026	-0.0414	0.1156	0.1228
		5000					-6290	-0.0028	0.0195	0.0197	0.0002	0.0200	0.0200	-0.0003	0.0023	0.0023	-0.0002	0.0023	0.0023	-0.0588	0.0960	0.1126
0.999	0.999	3000					-3687	-0.0027	0.0352	0.0353	-0.0046	0.0281	0.0284	-0.0011	0.0058	0.0060	-0.0005	0.0017	0.0018	-0.0154	0.1689	0.1696
		4000					-4917	-0.0005	0.0414	0.0414	-0.0014	0.0235	0.0235	-0.0004	0.0020	0.0020	-0.0005	0.0022	0.0022	-0.0466	0.0983	0.1088
		5000					-6153	-0.0001	0.0198	0.0198	0.0022	0.0204	0.0205	-0.0003	0.0012	0.0012	-0.0003	0.0011	0.0011	-0.0468	0.1098	0.1194
Panel B: GED																						
0.999	0.99	3000					-4055	-0.0010	0.0193	0.0193	-0.0068	0.0947	0.0949	-0.0002	0.0007	0.0007	-0.0063	0.0168	0.0179	-0.0019	0.0393	0.0393
		4000					-5412	-0.0005	0.0160	0.0160	-0.0023	0.0582	0.0583	-0.0001	0.0006	0.0006	-0.0045	0.0157	0.0163	-0.0005	0.0332	0.0332
		5000					-6777	-0.0002	0.0145	0.0145	-0.0025	0.0517	0.0518	-0.0001	0.0005	0.0005	-0.0036	0.0106	0.0112	0.0034	0.0314	0.0315
0.99	0.999	3000					-4058	0.0140	0.0782	0.0794	0.0014	0.0197	0.0198	-0.0052	0.0139	0.0148	-0.0002	0.0007	0.0007	0.0007	0.0422	0.0422
		4000					-5421	0.0130	0.0758	0.0769	-0.0006	0.0177	0.0177	-0.0074	0.0198	0.0211	-0.0001	0.0007	0.0007	0.0053	0.0358	0.0362
		5000					-6769	0.0074	0.0516	0.0521	0.0006	0.0156	0.0156	-0.0041	0.0143	0.0149	-0.0001	0.0006	0.0006	0.0003	0.0312	0.0312
0.99	0.99	3000					-4139	-0.0014	0.0247	0.0248	0.0004	0.0257	0.0257	-0.0006	0.0032	0.0032	-0.0007	0.0033	0.0034	0.0031	0.0418	0.0419
		4000					-5519	0.0005	0.0204	0.0204	0.0021	0.0204	0.0205	-0.0003	0.0027	0.0027	-0.0005	0.0027	0.0027	0.0030	0.0360	0.0362
		5000					-6896	0.0010	0.0188	0.0188	-0.0012	0.0182	0.0182	-0.0001	0.0024	0.0024	-0.0001	0.0023	0.0023	0.0023	0.0308	0.0308
0.999	0.999	3000					-4051	0.0004	0.0289	0.0289	-0.0000	0.0313	0.0313	-0.0010	0.0074	0.0075	-0.0010	0.0072	0.0073	0.0002	0.0411	0.0411
		4000					-5406	0.0030	0.0248	0.0250	-0.0013	0.0228	0.0228	-0.0007	0.0035	0.0036	-0.0006	0.0020	0.0020	-0.0002	0.0305	0.0305
		5000					-6753	-0.0011	0.0236	0.0236	0.0001	0.0207	0.0207	-0.0004	0.0016	0.0016	-0.0008	0.0078	0.0078	-0.0035	0.0331	0.0333
Panel C: Tempered Stable Distribution																						
0.999	0.99	5000	0.5	0.5	1.0	1.0	-6748	-0.0004	0.0152	0.0152	-0.0046	0.0534	0.0536	-0.0001	0.0005	0.0005	-0.0034	0.0105	0.0110	0.0006	0.0344	0.0344
		5000	0.2	1.2	1.2	0.2	-6230	-0.0013	0.0158	0.0159	-0.0042	0.0444	0.0446	-0.0001	0.0005	0.0005	-0.0037	0.0130	0.0135	0.0149	0.0885	0.0897
		5000	1.2	0.2	0.2	1.2	-6232	0.0025	0.0155	0.0157	0.0015	0.0452	0.0452	-0.0001	0.0007	0.0007	-0.0057	0.0325	0.0330	0.0314	0.0931	0.0983
0.99	0.999	5000	0.5	0.5	1.0	1.0	-6752	0.0030	0.0504	0.0505	0.0001	0.0155	0.0155	-0.0035	0.0113	0.0119	-0.0001	0.0005	0.0005	0.0037	0.0342	0.0344
		5000	0.2	1.2	1.2	0.2	-6231	0.0082	0.0881	0.0885	-0.0020	0.0170	0.0171	-0.0060	0.0278	0.0284	-0.0001	0.0006	0.0006	0.0371	0.0903	0.0976
		5000	1.2	0.2	0.2	1.2	-6236	0.0021	0.0459	0.0459	0.0031	0.0146	0.0149	-0.0050	0.0137	0.0146	-0.0001	0.0005	0.0005	0.0259	0.0914	0.0950
0.99	0.99	5000	0.5	0.5	1.0	1.0	-6875	-0.0014	0.0186	0.0187	-0.0011	0.0192	0.0192	-0.0002	0.0021	0.0022	-0.0003	0.0024	0.0024	0.0034	0.0357	0.0359
		5000	0.2	1.2	1.2	0.2	-6372	-0.0025	0.0177	0.0179	-0.0022	0.0192	0.0193	-0.0001	0.0022	0.0022	-0.0002	0.0021	0.0021	0.0340	0.0878	0.0941
		5000	1.2	0.2	0.2	1.2	-6366	0.0022	0.0171	0.0172	0.0017	0.0179	0.0180	-0.0004	0.0023	0.0023	-0.0002	0.0020	0.0020	0.0357	0.1015	0.1076
0.999	0.999	5000	0.5	0.5	1.0	1.0	-6732	0.0026	0.0276	0.0278	-0.0037	0.0239	0.0242	-0.0006	0.0078	0.0078	-0.0004	0.0014	0.0015	-0.0003	0.0325	0.0325
		5000	0.2	1.2	1.2	0.2	-6217	-0.0012	0.0200	0.0200	-0.0020	0.0222	0.0223	-0.0003	0.0010	0.0011	-0.0004	0.0019	0.0020	0.0252	0.0927	0.0961
		5000	1.2	0.2	0.2	1.2	-6214	0.0025	0.0177	0.0178	0.0012	0.0196	0.0197	-0.0002	0.0008	0.0008	-0.0004	0.0016	0.0017	0.0182	0.0785	0.0806

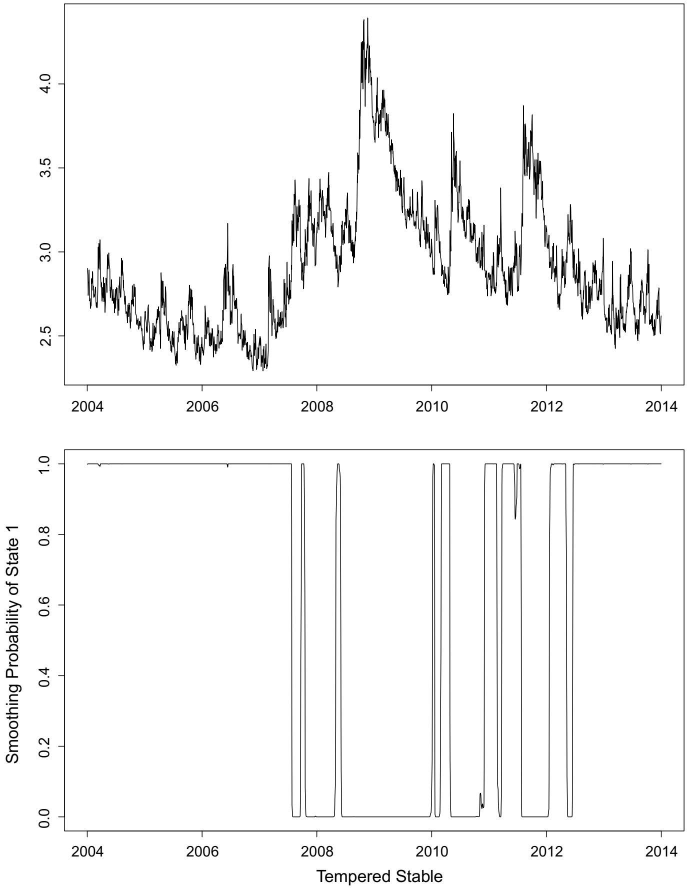

Fig. 1 VIX Index and identified states by the MRS-S model

perform as well as the MRS-S model in the estimation of all parameters. All the above observations are robust across various combinations of different parameter and sample size settings. Hence, we hereby argue that for a given financial dataset from an unknown fat-tailed distribution, the MRS-S model should be employed to accurately study its regime-switching properties.

5 Empirical Results

To empirically compare MRS models with various distributions, we work on the dataset of the daily implied volatility of the S&P 500 options (VIX). The daily closing prices for VIX over the period between 1 January 2004 and 31 December 2013 are obtained from the Thomson Reuters Tick History (TRTH) database, which contains microsecond-time-stamped tick data dating back to January 1996. The database covers 35 million OTC and exchange-traded instruments worldwide, which are provided by the Securities Industry Research Centre of Australasia (SIRCA). Additionally, since volatility is non-negative, we work on the logarithm transformation. Thus, no constraints to ensure non-negativity are needed.

The level of the logarithm of VIX is plotted in Fig. 1. In the this plot, VIX dramatically increases from mid-2007, which is the start of global financial crisis (GFC) period. This suggests that the S&P 500 options return is more volatile during the GFC period. After 2010, the rough end of GFC, it tends to be less volatile with some turbulences around the end of 2010 and the beginning of 2012. Hence, with the presence of GFC and preliminary visual evidence, structural breaks may possibly exist in the VIX. Regarding the descriptive statistics, the mean and standard error of logarithm of VIX are 2.9175 and 0.3909, respectively. The skewness is 0.9741, indicating that the logarithm of VIX is moderately positively skewed. The excess kurtosis is 0.7911, which suggests that a non-Gaussian distribution may be appropriate. Thus, we perform the Kolmogorov–Smirnov and Jarque–Bera normality tests, where the null hypotheses indicating normality are rejected in both cases (p values are 0.0000). Further, we perform the non-parametric change-point test proposed by Ross et al. (2011), which is robust for non-Gaussian data. The corresponding p value is 0.0000, suggesting rejection of the null hypothesis assuming no structural breaks in the data. Consequently, it is interesting to study the regime-switching properties of VIX. As the data follow a non-Gaussian distribution, the MRS-N is not appropriate, and we will employ the MRS-t, MRS-G and MRS-S to investigate the logarithm of VIX. Besides, since its skewness is not negligible, we also consider MRS models with the skewed Student's t-distribution and GED for comparison.

We fit the MRS models with (skewed) Student's t, (skewed) GED and tempered stable distributions for the logarithm of VIX. The results are presented in Table 5. Overall, all estimates are significant at 5% level in all models, and estimates of each parameter from different models are quite close. More specifically, estimates of μ_1 and μ_2 are around 2.7 and 3.3, respectively. Both p_{11} and p_{22} are greater than 0.99 in all models, suggesting that the expectations of staying at both the low- and high-volatility states are quite long (Shi and Ho 2015). The estimated variance is around 0.055 from the MRS models with (skewed) Student's t and (skewed) GED. MRS-S generates an estimate of 0.0762, which is slightly greater. Recall the simulations results of Sect. 4, MRS-S and MRS-G cannot provide consistent estimators of variance if they are not the true distribution. On the other hand, the estimators of MRS-S are consistent for all parameters in all cases. Hence, it is expected that the estimate of 0.0762 is more reliable. To compare the model performance, log-likelihood, Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) are presented. It can be seen that MRS-t and MRS-G cannot perform as well as MRS with skewed Student's t and GED. More importantly, although MRS-S has the largest number of parameters, both AIC and BIC suggest it outperforms all the other four MRS models with fat-tailed distributions.

To identify the underlying states over time, we plot the estimated smoothing probabilities of state 1 (low-volatility state) obtained from all models in Figs. 1 and 2. It is clear that the identified state patterns of VIX are almost the same among all models. It starts from the low-volatility state until the beginning of GFC. Since then, it switches

Table 5 Empirical results: VIX index

	Student's t	*GED*	*Temperd stable*	*Skewed Student's t*	*Skewed GED*
μ_1	2.6777	2.6927	2.6887	2.6993	2.7052
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
μ_2	3.2650	3.3305	3.2779	3.2908	3.3272
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
p_{11}	0.9951	0.9953	0.9950	0.9950	0.9952
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
p_{22}	0.9910	0.9901	0.9907	0.9905	0.9900
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
σ^2	0.0549	0.0535	0.0762	0.0547	0.0539
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
ν	5.8210	1.6777		13.2548	1.8837
	(0.0000)	(0.0000)		(0.0000)	(0.0000)
a_1			− 0.2917		
			(0.0000)		
a_2			− 8.8993		
			(0.0000)		
λ_1			2.8445		
			(0.0000)		
λ_2			1.0848		
			(0.0000)		
p			0.6934		
			(0.0000)		
ξ				1.8878	1.9500
				(0.0000)	(0.0000)
Log Lik.	95.0467	57.1101	250.4691	232.2414	227.3309
AIC	− 178.0933	− 102.2201	− 480.9382	− 450.4828	− 440.6617
BIC	− 143.1084	− 67.2352	− 422.6300	− 409.667	− 399.8460

Log Lik. stands for the log-likelihood. The values in the parentheses are the corresponding p values

to the high-volatility state and stays there for around two years. From 2010 to 2012, the trend is not stable. The periods including end of 2010 and start of 2012 seem to be the high-volatility durations, while the periods between them are mostly low-volatility durations. After the end of 2012, the regime structure is more stable, and it stays in the low-volatility state. These patterns are accordant with those of the real macroeconomic situation: the 2008 GFC causes high volatility, and its effects last for around two years.

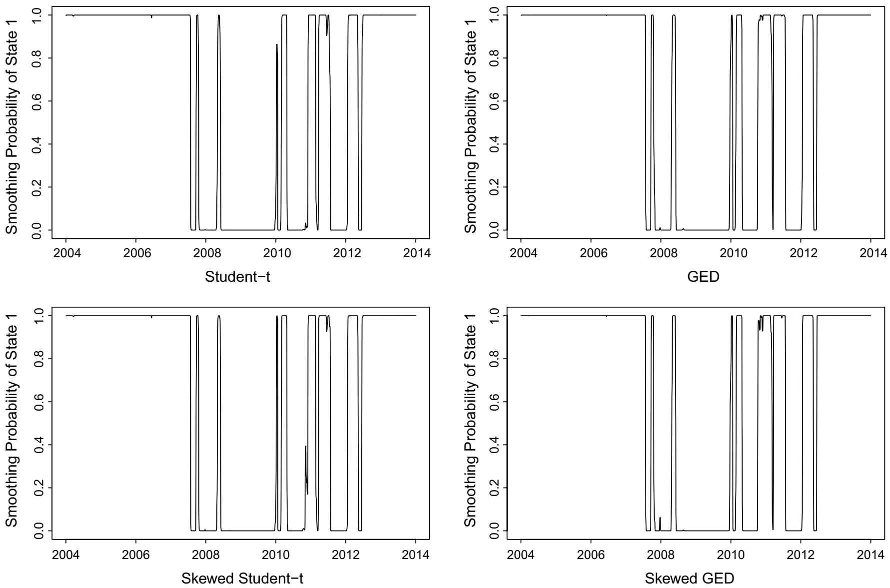

Fig. 2 Smoothing probability of state 1 by the MRS models with (skewed) Student' t-distribution and GED

6 Conclusion

The MRS model has enjoyed particular popularity in the finance research related to structural breaks. However, the original MRS model is based on the Normal distribution, and its estimators will be inconsistent for fat-tail-distributed data. Unfortunately, the financial data are rarely Gaussian in practice. Hence, the sought of an appropriate distribution to accommodate their excess kurtosis becomes an essential issue for the application of the MRS-type model. Student's t-distribution and GED are two widely used fat-tailed densities in the existing literature. However, a recent study by Calzolari et al. (2014) points out that due to the instability under aggregation, those distributions are not optimal choices. To overcome this problem, the α-stable distribution is introduced. Despite its attractive properties, the second moment of the α-stable distribution does not exist. Hence, it can cause more serious issues for the MRS model, which affect the validity of the asymptotic properties of its ML estimators and the interpretation of its estimated parameters.

To address those new problems, this paper suggests that the tempered stable distribution should be used instead of the α-stable. The reason is that tempered stable distribution retains all the attractive properties of the α-stable and has defined moments. Via three different simulation studies of the two-state MRS process, we systematically demonstrate the appropriateness of the tempered stable distribution applied within the MRS framework. The first two studies assume that the true distributions are the Student's t and GED, respectively. In these studies, results of MRS-S are close to those of the true models. Additionally, MRS-S generally outperforms its competitors (models other than the true specification) in terms of consistency, efficiency and overall performance. We consider three different combinations of the underlying tempered stable distribution in the third study. Our results suggest that none of the MRS-N, MRS-t and MRS-G can perform as well as the MRS-S model in the estimation of all interested parameters.

Empirical evidence is further provided to evaluate the performance of MRS-S in practice. We fit the logarithm of daily VIX into MRS models with five fat-tailed distributions. Our results suggest that MRS-S is still preferred to the others, with respect to various model selection criteria. Also, the identified state structures of VIX obtained from the fitted models are accordant with the macroeconomic situation. Therefore, we argue that the tempered stable distribution could be widely used for modelling the financial data in general contexts with an MRS-type specification.

For practitioners and researchers in the area of financial data analysis from China and other emerging markets, we argue that MRS model with α stable distribution will be more promising. The main reason is that financial data from these countries normally have more structural breaks and are more volatile. Models with ordinary innovation distributions are not adequate to fit the excess kurtosis in the data and lead to poor forecasting performance. Therefore, we expect to see more applications of MRS and related models with α stable distributions. In particular, MRS can be useful in the analysis of macroeconomic data like interest rate and CPI, which is also the main direction of our future works.

Acknowledgements We are grateful to Macquarie University and Jiangxi University of Finance and Economics for their support. The authors would also like to thank Dave Allen, Felix Chan, Michael McAleer, Paresh Narayan, Morten Nielson, Albert Tsui, Zhaoyong Zhang, and participants at the 1st Conference on Recent Developments in Financial Econometrics and Applications, the 8th China R Conference (Nanchang), ANU Research School Brown Bag Seminar, Central University of Finance and Economics Seminar, China Meeting of Econometric Society, Chinese Economists Society China Annual Conference, Econometric Society Australasian Meeting, Ewha Woman's University Department of Economics Seminar, International Congress of the Modeling and Simulation Society of Australia and New Zealand and Shandong University Seminar. The usual disclaimer applies. Tong Fu acknowledges financial support by the National Natural Science Foundation of China (Project No. 71763009).

References

Ardia, D. (2009). Bayesian estimation of a markov-switching threshold asymmetric GARCH model with Student-t innovations. *The Econometrics Journal*, *12*, 105–126.

Baeumer, B., & Meerschaert, M. M. (2010). Tempered stable lévy motion and transient super-diffusion. *Journal of Computational and Applied Mathematics*, *233*, 2438–2448.

Bianchi, M. L., Rachev, S. T., Kim, Y. S., & Fabozzi, F. J. (2010). Tempered stable distributions and processes in finance: Numerical analysis. *Mathematical and statistical methods for actuarial sciences and finance* (pp. 33–42). Italy: Springer.

Bollerslev, T. (1987). A conditional heteroskedastic time series model for speculative prices and rates of return. *Review of Economics and Statistics*, *69*, 542–547.

Calzolari, G., Halbleib, R., & Parrini, A. (2014). Estimating GARCH-type models with symmetric stable innovations: Indirect inference versus maximum likelihood. *Computational Statistics & Data Analysis*, *76*, 158–171.

Carr, P., Geman, H., Madan, D. B., & Yor, M. (2002). The fine structure of asset returns: An empirical investigation*. *The Journal of Business*, *75*, 305–333.

Carrasco, M., Hu, L., & Ploberger, W. (2014). Optimal test for markov switching parameters. *Econometrica*, *82*, 765–784.

Cho, J. S., & White, H. (2007). Testing for regime switching. *Econometrica*, *75*, 1671–1720.

Constantinides, A., & Savel'ev, S. (2013). Modelling price dynamics: A hybrid truncated Lévy Flight–GARCH approach. *Physica A*, *392*, 2072–2078.

Cont, R., & Tankov, P. (2004). *Financial modelling with jump processes*. London: CRS Press.

Diebold, F. X., & Inoue, A. (2001). Long memory and regime switching. *Journal of Econometrics*, *105*, 131–159.

Douc, R., Moulines, E., Ryden, T., et al. (2004). Asymptotic properties of the maximum likelihood estimator in autoregressive models with Markov regime. *Annals of Statistics*, *32*, 2254–2304.

Faias, J., & Nunes, L.C. (2012). Testing volatility regime switching. Working paper.

Fan, K., Shen, Y., Siu, T. K., & Wang, R. (2015). Pricing annuity guarantees under a double regime-switching model. *Insurance: Mathematics and Economics*, *62*, 62–78.

Feng, L., & Shi, Y. (2016). Fractionally integrated garch model with tempered stable distribution: A simulation study. *Journal of Applied Statistics*, *44*(16), 1–21.

Goldfeld, S. M., & Quandt, R. E. (1973). A markov model for switching regressions. *Journal of Econometrics*, *1*, 3–15.

Haas, M. (2009). Value-at-risk via mixture distributions reconsidered. *Applied Mathematics and Computation*, *215*, 2103–2119.

Haas, M., Mittnik, S., & Paolella, M. S. (2004). A new approach to Markov-switching GARCH models. *Journal of Financial Econometrics*, *2*, 493–530.

Haas, M., & Paolella, M. S. (2012). Mixture and regime-switching GARCH models. In L. Bauwens, C. Hafner, & S. Laurent (Eds.), *Handbook of volatility models and their applications* (pp. 71–102). London: Wiley.

Hamilton, J. D. (1988). Rational-expectations econometric analysis of changes in regime: An investigation of the term structure of interest rates. *Journal of Economic Dynamics and Control*, *12*, 385–423.

Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica*, *57*, 357–384.

Hamilton, J. D. (1994). *Time series analysis*. Princeton: Princeton University Press.

Ho, K. Y., Shi, Y., & Zhang, Z. (2013). How does news sentiment impact asset volatility? Evidence from long memory and regime-switching approaches. *North American Journal of Economics and Finance*, *26*, 436–456.

Kawai, R., & Masuda, H. (2011). On simulation of tempered stable random variates. *Journal of Computational and Applied Mathematics*, *235*, 2873–2887.

Kim, Y. S., Rachev, S. T., Bianchi, M. L., & Fabozzi, F. J. (2008). Financial market models with lévy processes and time-varying volatility. *Journal of Banking & Finance*, *32*, 1363–1378.

Klaassen, F. (2002). Improving GARCH volatility forecasts with regime-switching GARCH. *Empirical Economics*, *27*, 363–394.

Koponen, I. (1995). Analytic approach to the problem of convergence of truncated lévy flights towards the gaussian stochastic process. *Physical Review E*, *52*, 1197.

Küchler, U., & Tappe, S. (2013). Tempered stable distributions and processes. *Stochastic Processes and their Applications*, *123*, 4256–4293.

Marcucci, J. (2005). Forecasting stock market volatility with regime-switching GARCH models. *Studies in Nonlinear Dynamics & Econometrics*, *9*(4), 1–55.

Mittnik, S., Doganoglu, T., & Chenyao, D. (1999). Computing the probability density function of the stable Paretian distribution. *Mathematical and Computer Modelling*, *29*, 235–240.

Quandt, R. E. (1958). The estimation of the parameters of a linear regression system obeying two separate regimes. *Journal of the American Statistical Association*, *53*, 873–880.

Ross, G. J., Adams, N. M., Tasoulis, D. K., & Hand, D. J. (2011). A nonparametric change point model for streaming data. *Technometrics*, *53*, 379–389.

Shi, Y., & Feng, L. (2016). A discussion on the innovation distribution of the Markov regime-switching GARCH model. *Economic Modelling*, *53*, 278–288.

Shi, Y., & Ho, K. Y. (2015). Long memory and regime switching: A simulation study on the Markov regime-switching ARFIMA model. *Journal of Banking & Finance*, *61*, S189–S204.

Stanley, H. E., Plerou, V., & Gabaix, X. (2008). A statistical physics view of financial fluctuations: Evidence for scaling and universality. *Physica A*, *387*, 3967–3981.

Susmel, R., & Engle, R. F. (1994). Hourly volatility spillovers between international equity markets. *Journal of International Money and Finance*, *13*, 3–25.
Tankov, P. (2003). *Financial modelling with jump processes* (Vol. 2). Boca Raton: CRC Press.
Tong, H. (1983). *Threshold models in non-linear time series analysis*. Berlin: Springer.
Wilfling, B. (2009). Volatility regime-switching in European exchange rates prior to monetary unification. *Journal of International Money and Finance*, *28*, 240–270.
Ye, W., Zhu, Y., Wu, Y., & Miao, B. (2016). Markov regime-switching quantile regression models and financial contagion detection. *Insurance: Mathematics and Economics*, *67*, 21–26.
Zou, B., & Cadenillas, A. (2014). Explicit solutions of optimal consumption, investment and insurance problems with regime switching. *Insurance: Mathematics and Economics*, *58*, 159–167.