Noname manuscript No.
(will be inserted by the editor)

Fourier Inference for Stochastic Volatility Models with Heavy–Tailed Innovations

Bruno Ebner · Bernhard Klar · Simos G. Meintanis

Received: date / Accepted: date

Abstract We consider estimation of stochastic volatility models which are driven by a heavy–tailed innovation distribution. Exploiting the simple structure of the characteristic function of suitably transformed observations we propose an estimator which minimizes a weighted L_2–type distance between the theoretical characteristic function of these observations and an empirical counterpart. A related goodness–of–fit test is also proposed. Monte–Carlo results are presented. The procedures are also applied to real data from the financial markets.

Keywords Stochastic volatility model · Minimum distance estimation · Heavy–tailed distribution · Characteristic function

Bruno Ebner
Institute of Stochastics, Karlsruhe Institute of Technology, Karlsruhe, Germany
Tel.: +49-721-60847636
Fax: +49-721-60846066
E-mail: bruno.ebner@kit.edu

Bernhard Klar
Institute of Stochastics, Karlsruhe Institute of Technology, Karlsruhe, Germany
E-mail: bernhard.klar@kit.edu

Simos G. Meintanis
Department of Economics, National and Kapodistrian University of Athens, Athens, Greece, Unit for Business Mathematics and Informatics, North–West University, South Africa E-mail: simosmei@econ.uoa.gr
On sabbatical leave from the University of Athens

1 Introduction

Consider observations $\{y_t\}$ driven by the stochastic volatility model (SVM)

$$\begin{cases} y_t = \delta + c_t \varepsilon_t \\ c_t = e^{\frac{1}{2}h_t}, \; h_t = \mu + \gamma h_{t-1} + \sigma_v v_t, \quad t = 1, 2, ..., \end{cases} \tag{1}$$

where $\{\varepsilon_t\}$ and $\{v_t\}$ are sequences of $i.i.d.$ random variables each with mean zero, and $(\delta, \mu, \gamma, \sigma_v)$, denote unknown parameters. In the classical Gaussian SVM (ε_t, v_t) are assumed to be jointly normally distributed with zero covariance. Even in this benchmark case however, standard methods such as maximum likelihood are difficult to apply requiring high–dimensional integration over the unobserved volatilities. This drawback inspired two separate approaches. On the one hand, there are computationally intensive methods which approximate the efficiency of maximum likelihood, whereas a second approach employs relatively simple estimators based on moments. The former approach is made more feasible with present time computers. In this connection, [20] suggest an approximation of the SVM likelihood via hidden Markov models which can be made arbitrary accurate and that is competitive in terms of computing time. Also the second approach has attained recent results with good efficiency properties; see for instance [24]. Nevertheless efficient and easily implemented estimation of SVM still remains a challenging problem. Good review articles on methods of estimation for SVM are [28] and [16].

In this paper we adopt a generalized version of model (1), in which the innovation distributions $\{\varepsilon_t\}$ and $\{v_t\}$ may not be Gaussian, and could involve an extra vector of parameters. In this connection we note that new results with data from the financial markets suggest that a SVM involving heavy–tailed and possibly asymmetric distributions might yield a better fit (see for instance [11,21]). For this non–Gaussian model we propose an alternative method of estimation, first put forward by [17] in the context of the classical Gaussian SVM, and was recently applied by [31] to a variation of the classical Gaussian SVM which involves a threshold. This method is based on the fact that, unlike the density, the joint characteristic function (CF) of $\log(y_t - \delta)^2$ in (1) may be theoretically computed. As particular instances of non–Gaussian innovation distributions we consider the stable Pareto, the variance gamma, and the normal inverse Gaussian distributions, all of which have the normal distribution as a special (or limiting) case, thus generalizing the approach of [17]. Note that these distributions are amongst the most popular alternatives to the Gaussian law, and have often been employed in modeling financial data; see for instance [1,2,8–10,19,23,32,26], among others. Moreover they enjoy the feature of being more conveniently parameterized by the CF rather than by the corresponding density.

The remainder of the paper is as follows. In Section 2 we introduce the new estimation procedure, while in Sections 3, 4 and 5 the new method is particularized to certain heavy–tailed SVM and specific computational aspects of the procedure are considered. A Monte Carlo study for the finite–sample properties of the methods and empirical applications are presented in Sections 6 and 7, respectively. Section 8 presents a goodness–of–fit test for SVM with arbitrary innovation distribution. In Section 9 we consider an extension of the estimation procedure to multivariate SVM. The paper concludes in Section 10 with discussion and outlook.

2 Fourier estimation

Let $\mathbf{Z}_t = (z_t, ..., z_{t-k})'$ denote a vector of size $k+1$ with elements $z_t := \log(y_t - \delta)^2$, $t = 1, ...$ and denote the corresponding CF by $\varphi(\mathbf{u})$, $\mathbf{u} \in \mathbb{R}^{k+1}, k \geq 0$. We shall write $\varphi(\mathbf{u}) := \varphi(u; \boldsymbol{\vartheta})$, where $\boldsymbol{\vartheta} = (\boldsymbol{\theta}', \boldsymbol{\lambda}')'$, $\boldsymbol{\vartheta} \in \varXi \subseteq \mathbb{R}^d$, includes the entire set of parameters involved in the SVM of eqn. (1), with $\boldsymbol{\theta}' = (\delta, \mu, \gamma, \sigma_v) \in \varTheta \subseteq \mathbb{R}^4$ the vector of SVM parameters, and $\boldsymbol{\lambda} \in \varLambda \subseteq \mathbb{R}^{d-4}$ the set of extra distributional parameters pertaining to the joint law of (ε_t, v_t).

Then in view of observations $\{y_t,\ t = 1, ..., T\}$, the proposed estimator is defined as

$$\widehat{\boldsymbol{\vartheta}}_T = \arg\min_{\boldsymbol{\vartheta} \in \varXi} \varDelta_W(\boldsymbol{\vartheta}), \tag{2}$$

where

$$\varDelta(\boldsymbol{\vartheta}) := \varDelta_W(\boldsymbol{\vartheta}) = \int_{\mathbb{R}^{k+1}} |\varphi_T(\mathbf{u}) - \varphi(\mathbf{u}; \boldsymbol{\vartheta})|^2 W(\mathbf{u}) d\mathbf{u}. \tag{3}$$

In equation (3)

$$\varphi_T(\mathbf{u}) := \frac{1}{T}\sum_{t=1}^{T} e^{i\mathbf{u}'\mathbf{Z}_t},$$

is the empirical CF computed on the basis of the transformed variables $\mathbf{Z}_t$ and evaluated at $\mathbf{u} = (u_0, u_1, ..., u_k)'$, and $W(\cdot)$ is a weight function which is introduced in order to smooth out the persistent periodic behavior of $\varphi_T(\cdot)$.

The empirical CF has been successfully employed in time series under varying frameworks of dependence. A good review article is [35]. Other sources containing results within fairly general contexts include [18, 19, 25], and [15]. The asymptotic properties of the estimator based on the joint CF are established in [18] showing that the resulting ECF estimator is strongly consistent and asymptotically normally distributed with convergence rate $\sqrt{n}$ even for processes with stable noise. Although estimation based on (2) may also be carried out with arbitrary modes of dependence, in this work we shall operate within the specific set–up of (1) and under certain parametric distributional assumptions.

3 Characteristic function of heavy–tailed SVM

Consider the SVM in (1) and assume that ε_t and v_t are independent, following arbitrary distributions. Write

$$z_t = \log(y_t - \delta)^2 = h_t + \log\varepsilon_t^2 := h_t + e_t.$$

Clearly, for $k = 0$, the CF of $\mathbf{Z}_t = z_t$ is

$$\varphi_{z_t}(u) = \mathbb{E}\left(e^{iu(h_t+e_t)}\right) = \varphi_e(u)\varphi_h(u),$$

where $\varphi_X(\cdot)$ denotes the CF of the random variable X. To obtain the joint CF of $\mathbf{Z}_t,\ k = 1, 2, ...$, we start with the bivariate case ($k = 1$ and $\mathbf{u} = (u_0, u_1)'$). Then we have

$$\begin{aligned}\varphi_{\mathbf{Z}_t}(\mathbf{u}) &= \mathbb{E}\left(e^{iu_0z_t+iu_1z_{t-1}}\right)\\ &= \mathbb{E}_{h_{t-1}}\mathbb{E}\left(e^{iu_0z_t+iu_1z_{t-1}}\,\middle|\, h_{t-1}\right)\\ &= \mathbb{E}_{h_{t-1}}\mathbb{E}\left(e^{iu_0(h_t+e_t)+iu_1(h_{t-1}+e_{t-1})}\,\middle|\, h_{t-1}\right)\end{aligned}$$

$$= \mathbb{E}_{h_{t-1}} \mathbb{E}\left(e^{iu_0(\mu+\gamma h_{t-1}+\sigma_v v_t+e_t)+\mathrm{i}u_1(h_{t-1}+e_{t-1})} \middle| h_{t-1}\right)$$
$$= e^{i\mu u_0}\mathbb{E}(e^{iu_0\sigma_v v_t})\mathbb{E}(e^{iu_0 e_t})\mathbb{E}(e^{iu_1 e_{t-1}})\mathbb{E}_{h_{t-1}}\left(e^{i(u_0\gamma+u_1)h_{t-1}}\right)$$
$$= e^{i\mu u_0}\varphi_v(\sigma_v u_0)\varphi_e(u_0)\varphi_e(u_1)\varphi_h(u_0\gamma+u_1).$$

The argument readily generalizes from $\mathbf{u}\in\mathbb{R}^2$ to $\mathbf{u}\in\mathbb{R}^{k+1},\ k>1$. Specifically, by conditioning on h_{t-k}, we compute the following formula for the joint CF of $\mathbf{Z}_t$ as,

$$\varphi_{\mathbf{Z}_t}(\mathbf{u}) = \exp\left\{ i\mu \sum_{j=0}^{k-1} u_j \frac{1-\gamma^{k-j}}{1-\gamma}\right\} \Phi_{\mathbf{e}}(\mathbf{u})\ \Phi_{\mathbf{v}}(\mathbf{u}_\gamma)\ \varphi_h\left(u_\gamma\right), \tag{4}$$

where $u_\gamma = \sum_{j=0}^{k}\gamma^{k-j}u_j,\ \ u_{j,\gamma}=\sum_{m=0}^{k-j}\gamma^m u_{k-j-m},\ k\geq j$, and

$$\Phi_{\mathbf{e}}(\mathbf{u}) = \prod_{j=0}^{k}\varphi_e(u_j), \quad \Phi_{\mathbf{v}}(\mathbf{u}_\gamma) = \Phi_{\mathbf{v}}(u_{1,\gamma},...,u_{k,\gamma}) = \prod_{j=1}^{k}\varphi_v\left(\sigma_v u_{j,\gamma}\right).$$

To complete the derivation note that from (1), $h_t = \mu/(1-\gamma)+\sum_{k=0}^{\infty}\gamma^k\sigma_v v_{t-k}$, so that the CF of h_t is given by

$$\varphi_h(u) = e^{i\frac{\mu}{1-\gamma}u}\prod_{k=0}^{\infty}\varphi_v(\sigma_v u\gamma^k). \tag{5}$$

Equation (4) is fairly general in that it applies to SVM with arbitrary laws of (ε_t, v_t) as long as ε_t and v_t are assumed independent (although a more general equation may also be derived under dependence). In what follows however we present special instances of (4) in which v_t follows specific parametric distributions of particular interest:

(i) Stable Pareto distribution

The CF of a Stable Pareto (SP) distribution is given by

$$\varphi_v(u) = \begin{cases} e^{-|u|^{\theta_2}\{1-i\theta_1\mathrm{sgn}(u)\tan(\pi\theta_2/2)\}}, & \theta_2\neq 1,\\ e^{-|u|\{1+i\theta_1\frac{2}{\pi}\mathrm{sgn}(u)\log|u|\}}, & \theta_2 = 1,\end{cases} \tag{6}$$

where $-1\leq\theta_1\leq 1$, $0<\theta_2\leq 2$, and $\mathrm{sgn}(u)=1, u>0$, $\mathrm{sgn}(0)=0$, and $\mathrm{sgn}(u)=-1,\ u<0$. Note that θ_1 (resp. θ_2) is a skewness (resp. shape) parameter, and that for $(\theta_1,\theta_2)=(0,2)$ the zero–mean normal distribution results, while $(\theta_1,\theta_2)=(0,1)$ corresponds to the Cauchy distribution. Symmetric distributions result for $\theta_1=0$. Otherwise the distribution is asymmetric. In terms of tail–heaviness, the family includes at the one end the (medium–tailed) normal distribution for $\theta_2 = 2$, while at the other extreme, that is as $\theta_2 \to 0^+$, the corresponding distribution does not have moments of any order. Nevertheless, and since the mean of the SP distribution is finite only if $\theta_2 > 1$, we will exclude here those extreme cases and work with SP having finite mean.

(ii) Variance gamma and normal inverse Gaussian distributions
The parametrizations used below refer to zero–mean/unit–variance variance gamma (VG) and normal inverse Gaussian (NIG) distributions. These parametrizations although nonstandard they are particularly convenient for the definition of these distributions in terms of their CF. (The relation with other more common forms is deferred to Section 5). Specifically, the CF of a VG distribution is given by

$$\varphi_v(u) = e^{i\ c\,\theta_1\theta_2 u}(1 + i\ \theta_1 cu + c^2u^2)^{-\theta_2},\ \theta_1 \in \mathbb{R},\ \theta_2 > 0, \tag{7}$$

$c := c(\theta_1, \theta_2) = 1/\sqrt{(2 + \theta_1^2)\theta_2}$.
Likewise, the CF of a NIG distribution is given by

$$\varphi_v(u) = e^{i\ c\,\theta_1\theta_2 u}e^{\theta_2\left(1-\sqrt{1+2i\theta_1 cu+c^2u^2}\right)},\ \theta_1 \in \mathbb{R},\ \theta_2 > 0, \tag{8}$$

where $c := c(\theta_1, \theta_2) = 1/\sqrt{(1 + \theta_1^2)\theta_2}$.

The VG and the NIG distributions are also very flexible. They can be symmetric ($\theta_1 = 0$) or asymmetric ($\theta_1 \neq 0$), with arbitrary high value of skewness if the value of θ_1 is large and the value of θ_2 is small. On the other hand, the same distributions can assume a normal–like value of kurtosis if θ_2 is large, but for values of θ_2 near zero, both distributions become highly leptokurtic. Finally a common theme which encompasses the SP, as well as the VG and the NIG is that these distributions admit stochastic representations as mixture of normal distributions with stochastic variance. In the case of the VG (resp. the NIG) distribution the variance follows a gamma (resp. an inverse Gaussian) distribution. The SP case is more complicated to describe, but for the stochastic representation in the special case of *symmetric* SP distributions, the law of the variance is again SP with $\theta_1 = 1/2$ and $\theta_2 \leq 1$.

4 Estimation of heavy–tailed SVM

We come now to the estimation of the parameters of model (1) by means of the estimator defined by (2) under specific parametric assumptions. In particular assume that $\varepsilon_t \sim N(0,1)$ and that v_t follows a SP distribution with $\theta_2 > 1$. Then the CF of $e_t = \log \varepsilon_t^2$ (see [17]) and h_t (insert (6) in (5)) are given by (see [17]),

$$\varphi_e(u) = \frac{2^{iu}}{\sqrt{\pi}} \Gamma\left(iu + \frac{1}{2}\right), \tag{9}$$

and

$$\varphi_h(u) = e^{i\frac{\mu}{1-\gamma}u} \prod_{k=0}^{\infty} e^{-|\sigma_v u \gamma^k|^{\theta_2}\{1 - i\theta_1 \mathrm{sgn}(u\gamma^k)\tan(\pi\theta_2/2)\}} \tag{10}$$

$$= e^{i\frac{\mu}{1-\gamma}u} e^{-\frac{1}{1-\gamma^{\theta_2}}|\sigma_v u|^{\theta_2}\{1 - i\theta_1 \mathrm{sgn}(u)\tan(\pi\theta_2/2)\}}, \quad \text{if } \gamma > 0, \tag{11}$$

respectively. Hence for the SP SVM, the estimator (2) may be found by minimizing (3) over $\boldsymbol{\vartheta} = (\delta, \mu, \gamma, \sigma_v, \theta_1, \theta_2)'$, with $\varphi(\cdot)$ given in (4), and $\varphi_v(\cdot)$, $\varphi_e(\cdot)$ and $\varphi_h(\cdot)$ as in (6), (9) and (10), respectively.

Likewise if v_t is distributed as VG then by making use of (7) and (5) we have

$$\varphi_h(u) = \exp\left(iu\,\frac{\mu + c\theta_1\theta_2\sigma_v}{1-\gamma}\right)\left(\prod_{k=0}^{\infty}\left((1 + i\theta_1 c\gamma^k \sigma_v u + \left(c\gamma^k\sigma_v u\right)^2\right)\right)^{-\theta_2}, \tag{12}$$

whereas if v_t is distributed as NIG then (8) and (5) yield

$$\varphi_h(u) = \exp\left(iu\,\frac{\mu + c\theta_1\theta_2\sigma_v}{1-\gamma}\right)\left(\prod_{k=0}^{\infty} e^{1-\sqrt{1+2i\theta_1 c\gamma^k\sigma_v u + (c\gamma^k\sigma_v u)^2}}\right)^{\theta_2}. \tag{13}$$

As with the SP case, the corresponding SVM may be estimated by means of the estimator defined in (2), with $\varphi(\cdot)$ and $\varphi_e(\cdot)$ given by (4) and (9), respectively, and $\varphi_v(\cdot)$ and $\varphi_h(\cdot)$ as in (7) and (12) for the VG SVM, or as in (8) and (13) for the NIG SVM.

5 Parameterizations

We clarify the relationship between the parameterization of the VG and NIG distributions in (7) and (8) and more common parameterizations used in the literature and statistical software.

In [30] the CF φ of a random variable X following a VG distribution with parameters $(\tilde{c}, \sigma, \theta, \tilde{v})$ is given by

$$\varphi_X(u) = e^{i\tilde{c}u}(1 - i\theta\tilde{v}u + (\sigma^2\tilde{v}u^2)/2)^{-1/\tilde{v}},$$

where $\mathbb{E}(X) = \tilde{c} + \theta$ and $\mathrm{Var}(X) = \sigma^2 + \theta^2\tilde{v}$. Further, skewness and excess kurtosis are

$$S(X) = \frac{2\theta^3\tilde{v}^2 + 3\sigma^2\tilde{v}\theta}{(\sigma^2 + \theta^2\tilde{v})^{3/2}} \quad \text{and} \quad K(X) = \frac{3\sigma^4\tilde{v} + 12\sigma^2\theta^2\tilde{v}^2 + 6\theta^4\tilde{v}^3}{(\sigma^2 + \theta^2\tilde{v})^2}.$$

This parameterization is also used in the `R` library `VarianceGamma` ([29]) which we used for random number generation. Using new parameters $\theta_2 = 1/\tilde{v} > 0$ and $\theta_1 = -\theta/(c\theta_2) \in \mathbb{R}$, where $c = ((2 + \theta_1^2)\theta_2)^{-1/2}$, together with the restrictions $\mathbb{E}(X) = 0$ and $\mathrm{Var}(X) = 1$ leads to the CF in the form given in (7). Skewness and excess kurtosis are then given by

$$S(X) = -c\theta_1(3 - c^2\theta_1^2\theta_2) \quad \text{and} \quad K(X) = 3/\theta_2 + 3c^2\theta_1^2(2 - c^2\theta_1^2\theta_2).$$

In [4,5] the CF of a random variable X following a NIG distribution with parameters $(\mu, \alpha, \beta, \delta)$ (where μ denotes the location parameter, α is the tail heaviness, β is the asymmetry parameter and δ is the scale parameter) is given by

$$\varphi_X(u) = e^{i\ \mu u + \delta\left(\sqrt{\alpha^2 - \beta^2} - \sqrt{\alpha^2 - (\beta + iu)^2}\right)}.$$

This parameterization is also used in the `R` library `fBasics` ([34]) which we used for random number generation. In this notation we have with $\gamma = \sqrt{\alpha^2 - \beta^2}$

$$\mathbb{E}(X) = \mu + \delta\beta/\gamma \quad \text{and} \quad \mathrm{Var}(X) = \delta\alpha^2/\gamma^3,$$

skewness and excess kurtosis are given by

$$S(X) = \frac{3\beta}{\alpha\sqrt{\gamma\delta}} \quad \text{and} \quad K(X) = \frac{3(1 + 4\beta^2/\alpha^2)}{\gamma\delta}.$$

Using new parameters $\theta_1 = -\beta/\gamma$ and $\theta_2 = \gamma\delta$ together with the restrictions $\mathbb{E}(X) = 0$ and $\mathrm{Var}(X) = 1$ we obtain

$$\alpha = \left(c^2\sqrt{\theta_2}\right)^{-1}, \quad \beta = -\theta_1/c, \quad \delta = c\theta_2, \quad \mu = c\theta_1\theta_2, \quad \gamma = 1/c,$$

where $c = ((1 + \theta_1^2)\theta_2)^{-1/2}$ as above. Skewness and excess kurtosis are then given by

$$S(X) = -3c\theta_1 \quad \text{and} \quad K(X) = 3(1/\theta_2 + 4c^2\theta_1^2),$$

and the CF has the form given in (8).

6 Simulations

We present the results of a Monte Carlo study with $m = 500$ replications for the univariate ($k = 0$) and bivariate ($k = 1$) case. All simulations are performed using the statistical computing environment `R` (R Core Team 2014). In every simulation run we generated a time series of length 1000 after a burn-in period of 200.

In all simulations, we assume $\delta = 0$, and choose $W(u) = \exp(-\|u\|^2)$ as weight function. Since even the five parameter model allows for a vast variety of parameter configurations, and our focus is on estimation of heavy tailed models, we vary the distributional parameters θ_1 and θ_2, and fix the model parameters to $(\mu, \gamma, \sigma_v) = (-1, 0.5, 1.5)$. Further simulations showed no considerable effect in varying μ and σ_v, while augmenting γ to the upper bound results in unstable simulations due to the optimization routine `BBoptim` ([33]) where either boundary solutions or convergence failures occurred rather often. The starting values for the optimization routine were chosen by uniformly distributed shifts on the interval $[-0.2, 0.2]$ of the true parameter values. In Tables 1 and 2 we summarize the results for $k = 0$ and in Tables 3 and 4 for $k = 1$ respectively. The only variation between the simulation runs consists in the family of distributions of v_t and their parameters, which is reflected in the nomenclature of the rows, where e.g. $NIG(0, 3)$ stands for the NIG distribution with parameters $\theta_1 = 0$ and $\theta_2 = 3$ and the other parameters are fixed due to the standardization as described in the beginning of this section. Standard deviation and mean absolute deviation (MAD) are defined by

$$\text{St.Dev.} = \sqrt{\frac{1}{m}\sum_{j=1}^{m}(\widehat{\boldsymbol{\vartheta}}_{T,j} - \bar{\boldsymbol{\vartheta}})^2}, \quad \text{MAD} = \frac{1}{m}\sum_{j=1}^{m}|\widehat{\boldsymbol{\vartheta}}_{T,j} - \boldsymbol{\vartheta}|,$$

where $\widehat{\boldsymbol{\vartheta}}_{T,j}$ denotes the estimated parameter value as in (2), $\bar{\boldsymbol{\vartheta}}$ the arithmetic mean of the estimated parameters and $\boldsymbol{\vartheta}$ the true underlying parameter. We fixed the parameters of the NIG and VG distribution to fit some previously chosen skewness and kurtosis values, given in the second and third column of Tables 1 and 3 (clearly, the SP family does not allow the calculation of these moment–based measures).

Table 1 $k = 0$, Normal Inverse Gaussian and Variance Gamma

Dist.	Skew.	Kurt.		μ	γ	σ_v	θ_1	θ_2
$NIG(0, 3)$	0	1	Bias.	-.0884	-.0448	.0277	.0016	-.0048
			St.Dev.	.3246	.1654	.1132	.3658	.1984
			MAD	.1634	.0849	.0881	.2836	.1093
$NIG(-0.5, 1.8)$	1	3	Bias.	-.0876	-.0428	.0467	-.0211	-.0574
			St.Dev.	.2539	.1298	.1639	.3744	.3213
			MAD	.1704	.0882	.1174	.2867	.1795
$NIG(-1, 0.5)$	3	18	Bias.	-.0476	-.0216	.0466	.0032	-.0018
			St.Dev.	.1925	.0983	.1796	.4453	.2321
			MAD	.1265	.0622	.1432	.3542	.1910
$NIG(1, 0.5)$	-3	18	Bias.	-.0145	-.0082	.0338	.0188	.0418
			St.Dev.	.2030	.1063	.2173	.3765	.2749
			MAD	.1350	.0694	.1731	.2805	.2213
$VG(0, 3)$	0	1	Bias.	-.0664	-.0345	.0085	-.0573	-.0242
			St.Dev.	.3088	.1581	.1452	.4597	.2783
			MAD	.1874	.0966	.1018	.3513	.1440
$VG(-0.6, 1.2)$	1	3.2	Bias.	-.0980	-.0529	.0493	-.0912	-.0342
			St.Dev.	.2810	.1505	.1916	.4597	.2948
			MAD	.2180	.1168	.1463	.3593	.2237
$VG(-2, 0.4)$	3	14	Bias.	-.0572	-.0310	.0496	.0172	.0289
			St.Dev.	.1883	.1007	.1690	.3927	.1796
			MAD	.1266	.0674	.1398	.2225	.1386
$VG(2, 0.4)$	-3	14	Bias.	-.0182	-.0105	.0348	-.0115	.0580
			St.Dev.	.1715	.0938	.2092	.2513	.2475
			MAD	.1222	.0649	.1678	.1579	.1910

Table 2 $k = 0$, Stable Pareto Distribution

Dist.		μ	γ	σ_v	θ_1	θ_2
$SP(0, 1.5)$	Bias.	-.0280	-.0328	.0100	.0213	-.0015
	St.Dev.	.1336	.1233	.1140	.1250	.0950
	MAD	.0999	.0708	.0877	.0992	.0731
$SP(0.5, 1.5)$	Bias.	.0080	-.6094	.3058	-.1318	.0309
	St.Dev.	.3290	.2693	.1363	.1769	.0804
	MAD	.2538	.6097	.3065	.1726	.0654
$SP(-0.5, 1.5)$	Bias.	-.2510	.1140	.1704	.0730	.0585
	St.Dev.	.2559	.0796	.1467	.1370	.1093
	MAD	.2520	.1248	.1860	.1223	.1000
$SP(0, 1)$	Bias.	-.1053	-.0679	.1235	-.0147	.0065
	St.Dev.	.7275	.1814	.3782	.0933	.1011
	MAD	.3350	.1047	.2046	.0376	.0793
$SP(0.5, 1)$	Bias.	-.6797	-.2625	.2702	-.4886	.0201
	St.Dev.	1.381	.2125	.4081	.1183	.1335
	MAD	.7978	.2645	.3674	.4886	.0951
$SP(-0.5, 1)$	Bias.	-.0212	-.2836	.2341	.4975	.0089
	St.Dev.	.9063	.2288	.3107	.1073	.1197
	MAD	.4478	.2854	.3151	.4976	.0869

We can see from the simulations that the presented method works very well in the vast majority of cases. The calculations although time consuming (depending on the parameter configuration and the underlying distribution family, the computation time may vary considerably), they are numerically stable. For all distributions considered, there is no general improvement by increasing the window-size parameter k; mean absolute deviation, for example, is sometimes much smaller and sometimes much larger for $k = 1$ than for $k = 0$. In the NIG family there is no great difference in the estimation results between a positive and a negative skewness parameter. Likewise, for the VG distribution, we're not able to characterize a special behaviour in this respect. For the SP family, we see a remarkable difference between a symmetric and an asymmetric underlying distribution, see Figure 1. Namely, for the distributions

Table 3 $k = 1$, Normal Inverse Gaussian and Variance Gamma

Dist.	Skew.	Kurt.		μ	γ	σ_v	θ_1	θ_2
$NIG(0,3)$	0	1	Bias.	-.0067	-.0002	.0071	.0060	.0082
			St.Dev.	.1946	.0892	.1401	.3707	.1434
			MAD	.1569	.0715	.1093	.2820	.1037
$NIG(-0.5,1.8)$	1	3	Bias.	-.0210	-.0076	.0204	.0789	-.0308
			St.Dev.	.1995	.0905	.1622	2.888	.6133
			MAD	.1573	.0716	.1265	.4162	.3401
$NIG(-1,0.5)$	3	18	Bias.	-.0316	-.0186	.0154	-.0530	.1297
			St.Dev.	.1933	.0886	.2034	.3849	.3987
			MAD	.1530	.0708	.1588	.2599	.2714
$NIG(1,0.5)$	-3	18	Bias.	-.0281	-.0156	-.0126	-.0570	.1493
			St.Dev.	.1917	.0887	.2049	.3971	.4295
			MAD	.1517	.0702	.1608	.2698	.2887
$VG(0,3)$	0	1	Bias.	.1853	.0940	.0142	-.0068	-.0553
			St.Dev.	.1295	.0604	.1322	.1077	.1253
			MAD	.1923	.0968	.1073	.0874	.1138
$VG(-0.6,1.2)$	1	3.2	Bias.	-.0148	-.0008	.0078	.9420	-.1309
			St.Dev.	.1821	.0845	.1301	.2405	.1299
			MAD	.1424	.0677	.1032	.9420	.1488
$VG(-2,0.4)$	3	14	Bias.	.0768	.1244	-.4524	.0545	.2633
			St.Dev.	.3161	.1226	.2211	.3464	.7747
			MAD	.2492	.1367	.4558	.1985	.3714
$VG(2,0.4)$	-3	14	Bias.	.4028	.1111	-.4647	-.0306	.3278
			St.Dev.	.2276	.1575	.2343	.2890	.4506
			MAD	.4096	.1607	.4660	.1654	.3502

Table 4 $k = 1$, Stable Pareto Distribution

Dist.		μ	γ	σ_v	θ_1	θ_2
$SP(0, 1.5)$	Bias.	-.0117	-.0099	-.0049	.0140	.0026
	St.Dev.	.1137	.0478	.1169	.1055	.0902
	MAD	.0908	.0386	.0910	.0846	.0707
$SP(0.5, 1.5)$	Bias.	.5299	-.0253	.0429	-.1768	.0352
	St.Dev.	.2122	.0520	.1391	.1770	.0971
	MAD	.5299	.0461	.1170	.2110	.0817
$SP(-0.5, 1.5)$	Bias.	-.5091	.0236	-.0233	.1543	.0332
	St.Dev.	.1989	.0529	.1273	.1435	.0981
	MAD	.5093	.0461	.1020	.1751	.0830
$SP(0, 1)$	Bias.	-.0199	-.0227	.0904	-.0089	.0392
	St.Dev.	.8521	.0877	.2501	.1782	.1809
	MAD	.3238	.0527	.1820	.0662	.1069
$SP(0.5, 1)$	Bias.	-.1646	-.0618	.0692	-.4400	.1056
	St.Dev.	2.225	.1337	1.252	.1496	.1814
	MAD	.5200	.0755	.2382	.4485	.1474
$SP(-0.5, 1)$	Bias.	.0575	-.0677	.0521	.4090	.1625
	St.Dev.	2.007	.1866	.2556	.1927	.2808
	MAD	.5871	.1006	.2045	.4168	.2383

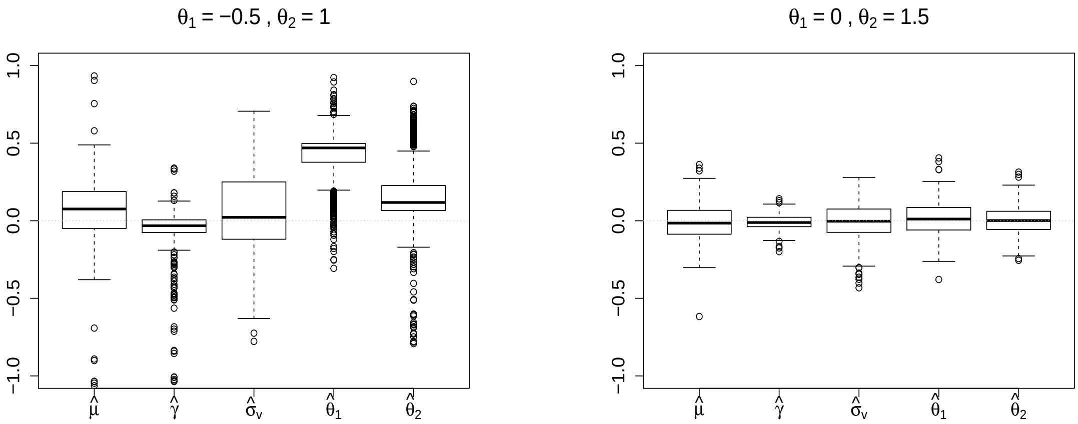

Fig. 1 Boxplot of the bias of the estimates for the Stable Pareto family ($k = 1$).

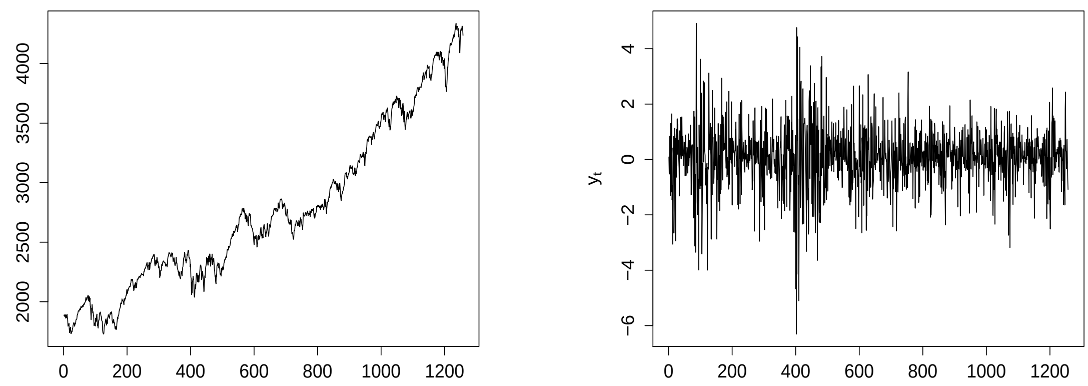

Fig. 2 NASDAQ Stock Index observed daily from 1st January 2010 to 1st January 2015 and log returns y_t.

$SP(\pm 0.5, 1)$, the bias of the estimators for θ_1 is close to the negative of the true parameter value, indicating that, in the present case, the deviation from symmetry can not be detected. As expected, increasing the sample size to $T = 10000$ results in a more accurate estimation (to save space, results are not reported here).

7 Real Data Example

In this section the estimation procedure is applied to daily observations of the NASDAQ stock price index to illustrate with real data the differences between the three distribution families. The series is observed daily from 1st of January 2010 to 1st of January 2015 with $T = 1258$. The prices p_t are transformed into log returns by $y_t = 100 \log \left(\frac{p_t}{p_{t-1}} \right)$ and centered around the sample mean $\bar{y}$. Estimates of the parameters of the stochastic volatility model by the proposed methods for this data set are given in Table 5. One clearly sees that the estimates of the model parameters (μ, γ, σ_v) are quite similar for the three different families with the exception of $\widehat{\sigma_v}$ in the SP case. Also, the estimated NIG and VG distributions have similar shapes as the values -0.70 and 3.6 (NIG) and -0.49 and 3.1 (VG) for skewness and excess kurtosis

Table 5 Results of the parameter estimation for the different distribution families.

Family	$\widehat{\mu}$	$\widehat{\gamma}$	$\widehat{\sigma}_v$	$\widehat{\theta}_1$	$\widehat{\theta}_2$
NIG	-.1733	.3925	1.0936	.2419	1.0169
VG	-.1670	.3954	1.0826	.2385	1.0211
SP	-.1442	.4089	.6972	-.0150	1.9344

indicate. Clearly, the SP distribution is not comparable with the other distributions. Simulated time series for the estimated parameters and the three different distributions can be found in Figure 3. We see that at least the first models mimic the original data set quite well, whereas the spikes for the fitted Pareto model seem to be too pronounced. This visual impression should be verified by a formal goodness of fit test as described in the next section.

8 Goodness–of–fit tests

In this section we shall be concerned with Fourier methods for testing the goodness–of–fit of a given data–set $\{y_t,\ t = 1, ..., T\}$, to the SVM model (1). Specifically we consider the distribution of ε_t as fixed, and we wish to test that the law of v_t belongs to a specific family of distributions indexed by the parameter $\boldsymbol{\lambda}$. To this end write $\varphi_v(u; \boldsymbol{\lambda})$ for the CF of v_t, and recall the notation $\widehat{\boldsymbol{\vartheta}}_T$ for the estimator of the parameter of the assumed SVM satisfying (2). We suggest the test statistic

$$\widehat{\Delta} := \Delta_W(\widehat{\boldsymbol{\vartheta}}_T) = \int_{\mathbb{R}^{k+1}} \left|\varphi_T(\mathbf{u}) - \varphi(\mathbf{u}; \widehat{\boldsymbol{\vartheta}}_T)\right|^2 W(\mathbf{u}) d\mathbf{u}, \tag{14}$$

where $\varphi(\cdot; \widehat{\boldsymbol{\vartheta}}_T)$ is given by (4) with $\boldsymbol{\vartheta}$ replaced by $\widehat{\boldsymbol{\vartheta}}_T$. Rejection then is for large values of $\widehat{\Delta}$.

Since the asymptotic distribution of the test statistic under the null hypothesis is complicated, we suggest to use the bootstrap in order to approximate this distribution and actually carry out the test. Specifically, a particular version of the bootstrap termed *parametric bootstrap* (PB) is particularly suited in the current setting. In this connection, [22] applied the PB for a continuous time SVM and showed that it consistently estimates the limit null distribution

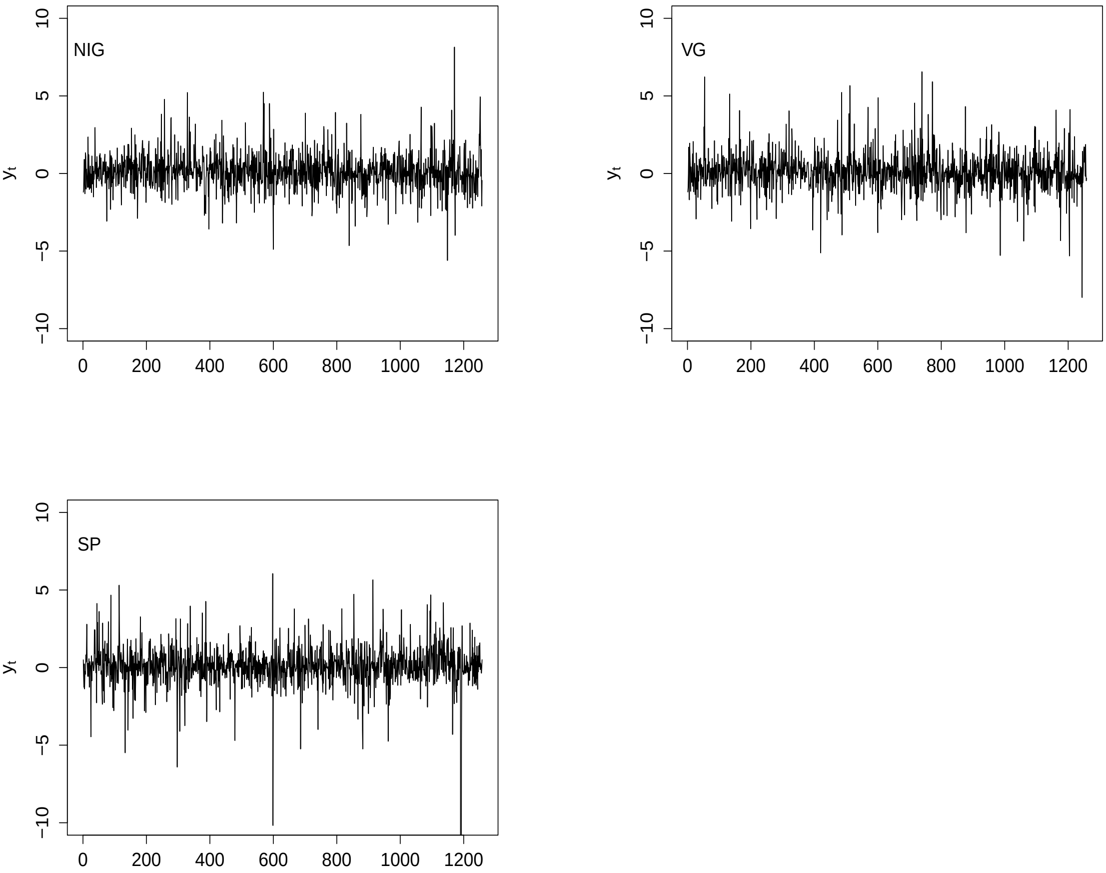

Fig. 3 Simulated time series of log returns for the NIG, VG and SP models with the estimated parameters.

of their test statistic. We note that the test statistic of [22] is also based on the empirical CF and it is used in order to test goodness–of–fit of the volatility process based on the CF of the corresponding discretized process but their method only applies to goodness–of–fit testing for the marginal law of the underlying volatility process.

In our current context of SVM in (1) with $\varepsilon_t \sim N(0,1)$, the PB takes the following form: Conditionally on $\{y_t, 1 \leq t \leq T\}$, obtain the aforementioned estimator $\widehat{\boldsymbol{\vartheta}}_T$ of the parameter, and write $\widehat{\Delta} := \widehat{\Delta}(z_1, ..., z_T; \widehat{\boldsymbol{\vartheta}}_T)$ for the test statistic in (14) corresponding to the transformed data $z_t = \log(y_t - \widehat{\delta})^2,\ t = 1, ..., T$. Then,

(i) Generate i.i.d. observations, $\{\varepsilon_t^*, 1 \leq t \leq T\}$ from $N(0,1)$ and, independently, $\{v_t^*, 1 \leq t \leq T\}$ from the assumed distribution with CF $\varphi_v(u; \widehat{\boldsymbol{\lambda}})$.

(ii) With initial value h_0^*, compute the pseudo–observations $z_t^* := \log(y_t^* - \widehat{\delta})^2,\ t = 1, ..., T$, where

$$\begin{cases} y_t^* = \widehat{\delta} + c_t^* \varepsilon_t^* \\ c_t^* = e^{\frac{1}{2} h_t^*},\ h_t^* = \widehat{\mu} + \widehat{\gamma}\, h_{t-1}^* + \widehat{\sigma} v_t^*. \end{cases}$$

(iii) On the basis of the pseudo–observations $\{z_t^*, 1 \leq t \leq T\}$, obtain a bootstrap estimate $\widehat{\boldsymbol{\vartheta}}_T^*$ of the parameter $\boldsymbol{\vartheta}$ from (2).

(iv) Calculate the value of the test statistic, $\widehat{\Delta}^* = \widehat{\Delta}(z_1^*, ..., z_T^*; \widehat{\boldsymbol{\vartheta}}_T^*)$.

(v) Repeat steps (i) and (iv) a number of times, say B, and obtain $\{\widehat{\Delta}_b^*\}_{b=1}^{\mathrm{B}}$.

(vi) Calculate the p–value as $\varpi = k/(\mathrm{B}+1)$ where $k := \sum_{b=1}^{\mathrm{B}} \mathbb{I}(\widehat{\Delta}_b^* \geq \widehat{\Delta})$ denotes the number of times that $\widehat{\Delta}_b^*$, $(b = 1, ..., \mathrm{B})$, was greater than or equal to $\widehat{\Delta}$.

(vii) Reject the null hypothesis if $\varpi \leq \alpha$, where α denotes the designated level of significance.

Given that bootstrap methods are computationally demanding, we can employ the warp–speed method of [12] in order to approximate rejection regions of the goodness–of–fit statistic $\widehat{\Delta}$ of (14). This method circumvents the need for costly bootstrap replications for each Monte Carlo sample. Specifically with the warp–speed method, rather than generating B bootstrap resamples for each Monte Carlo sample, we generate just one bootstrap resample for each Monte Carlo replication and compute the bootstrap test statistic $\widehat{\Delta}^*$ for that resample. By performing a number m of Monte Carlo replications and collecting the corresponding m bootstrap statistics $\widehat{\Delta}_b^*,\ b = 1, ..., m$, then the p–value is computed as in step (vi) above by replacing m for B.

We conducted a simulation study with 500 Monte Carlo replications and using the warp-speed method just described. In every run we simulated a SVM as in (1), where $\{\varepsilon_t\}$ were i.i.d standard normal and independently $\{v_t\}$ followed a χ^2 distribution with different degrees of freedom to show the behaviour for different degrees of skewness. As parameter vector for the SVM we used $(\delta, \mu, \gamma, \sigma_v) = (0, 0, 0.5, 1)$. Table 6 summarizes rejection rates, where the null hypothesis was either a symmetric NIG, symmetric VG or the Gaussian case in the SP model.

Table 6 Rejection rates for the goodness–of–fit tests for the different models (symmetric)

	sample size	χ^2_1	χ^2_5	χ^2_{50}
NIG	1000	0.418	0.128	0.066
SP: $N(0, 2)$	1000	0.276	0.094	0.056
VG	1000	0.220	0.254	0.084
VG	5000	0.984	0.516	0.178

Furthermore we tested the stochastic volatility model in (1) with Gaussian innovations (ε_t, v_t) against the model

$$\begin{cases} y_t = \delta + c_t \varepsilon_t \\ c_t = e^{\frac{1}{2}h_t},\ h_t = \mu + \gamma h_{t-1} + \sigma_v v_t + f(\varepsilon_t), \quad t = 1, 2, ..., \end{cases},$$

where $f(z) = \kappa_1 z + \kappa_2(z^2 - 1)$, with $(\kappa_1, \kappa_2) = (-0.05, 0.1)$, again with Gaussian innovations (ε_t, v_t). This model is motivated by the realized GARCH model in [13] and presents a slight deviation from the classical SVM in (1). The resulting rejection rate was 0.209 ($T = 1000$). Additional innovation–specifications as well as alternative model configurations were simulated but are not reported here in order to save space. These Monte Carlo results, which are available from the authors upon request, further corroborate the conclusions which can be drawn, namely that the test statistic (14) can detect not only deviations from the assumed distribution of the innovations $\{v_t\}$ but also model deviations from the assumed SVM (1). However it takes large sample sizes for such deviations to be detected by the suggested criterion.

Revisiting the data example of Section 7, we now test the goodness–of–fit of the stochastic volatility model for the NASDAQ data. For the three distribution families considered, we used the parameter estimates $\widehat{\boldsymbol{\vartheta}}_T$ given in Table 5 (and put $\widehat{\delta} = 0$). The number of bootstrap replications was B=500

Table 7 Results of the goodness–of–fit test for the different distribution families.

Family	$\widehat{\Delta}$	ϖ
NIG	.001186032	.3892216
VG	.001234078	.4111776
SP	.001225735	.2574257

for NIG and VG and B=100 for the SP family. The initial value in step (ii) was fixed to $h_0^* = \widehat{\mu}$. The restriction to less bootstrap repetitions for the SP family was due to a much longer simulation time required in this case. One can clearly see a difference between the three families to a smaller p–value for the SP family although none of the tests would reject the hypothesis on a significance level of 5%.

9 A note on multivariate extension

There exist many versions of the multivariate SVM (MSVM) model with varying complexities. Good review articles on the multitude of MSVM considered in the literature, including corresponding methods of estimation, are [3] and [7]. Here however we shall adopt the most basic model. This model was introduced by [14], and was subsequently analysed and generalized in several works. For dimension $p \geq 2$, the model admits the following formalization:

$$\begin{cases} \mathbf{y}_t = \mathbf{C}_t^{1/2} \boldsymbol{\varepsilon}_t \\ \mathbf{h}_t = \boldsymbol{\Gamma} \mathbf{h}_{t-1} + \mathbf{v}_t, \quad t = 1, 2, ..., \end{cases} \tag{15}$$

where $\mathbf{y}_t = (y_{1t}, ..., y_{pt})'$, $\mathbf{h}_t = (h_{1t}, ..., h_{pt})'$, and

$$\mathbf{C}_t = \text{diag}\left(e^{h_{1t}}, ..., e^{h_{pt}}\right), \quad \boldsymbol{\Gamma} = \text{diag}\left(\gamma_1, ..., \gamma_p\right).$$

Again we shall assume that $\boldsymbol{\varepsilon}_t$ and $\mathbf{v}_t$ are i.i.d. and mutually independent with mean zero. In the classical formulation $(\boldsymbol{\varepsilon}_t, \mathbf{v}_t)$ are assumed to be jointly normally distributed, but several authors considered alternative innovation distributions, always with an eye on heavier–tailed models; see for instance [36] and [14] for a Student t MSVM, and [7] for innovations following a generalized hyperbolic distribution.

We will present the CF $\varphi(\mathbf{U};\boldsymbol{\vartheta}) := \mathbb{E}(\mathbf{e}^{\mathbf{i}\mathbf{U}'\boldsymbol{\Psi}_\mathbf{t}})$ of $\boldsymbol{\Psi}_t = (\mathbf{Z}'_t, ..., \mathbf{Z}'_{t-(p-1)})'$ for the MSVM in eqn. (15), where $\mathbf{Z}_\tau = (\log y_{1\tau}^2, ..., \log y_{p\tau}^2)'$ and $\mathbf{U} = (\mathbf{u}'_0, ..., \mathbf{u}'_{p-1})'$, $\mathbf{u}_j \in \mathbb{R}^p,\ j = 0, ..., p-1$. Specifically, by conditioning on $\mathbf{h}_{t-p}$ and arguing along similar lines as in the derivation of eqn. (4) we obtain

$$\varphi(\mathbf{U};\boldsymbol{\vartheta}) = \varphi_\mathbf{h}(\mathbf{u}_{\boldsymbol{\Gamma}}) \prod_{j=0}^{p-1} \varphi_\mathbf{e}(\mathbf{u}_j) \prod_{j=1}^{p-1} \varphi_\mathbf{v}(\mathbf{u}_{j,\boldsymbol{\Gamma}}), \tag{16}$$

where

$$\mathbf{u}_{j,\boldsymbol{\Gamma}} = \sum_{m=0}^{p-1-j} \boldsymbol{\Gamma}^m \mathbf{u}_{p-1-j-m}, \quad \mathbf{u}_{\boldsymbol{\Gamma}} = \sum_{j=0}^{p-1} \boldsymbol{\Gamma}^{p-1-j} \mathbf{u}_j.$$

In eqn. (8.2) we write $\varphi_\mathbf{v}(\cdot)$ for the CF of $\mathbf{v}_t$ and denote by $\varphi_\mathbf{e}(\cdot)$ the CF of $\mathbf{e}_t = (e_{1t}, ..., e_{pt})'$ where $e_{jt} = \log \varepsilon_{jt}^2,\ j = 1, ..., p$. Also in eqn. (16), the CF $\varphi_\mathbf{h}(\cdot)$ of $\mathbf{h}_t$ may be further reduced and expressed in terms of $\varphi_\mathbf{v}(\cdot)$ following backward iteration as in eqn. (5).

The proposed estimator is then defined by eqn. (2) with

$$\Delta(\boldsymbol{\vartheta}) = \int_{\mathbb{R}^{p\times p}} |\varphi_T(\mathbf{U}) - \varphi(\mathbf{U};\boldsymbol{\vartheta})|^2 W(\mathbf{U}) d\mathbf{U},$$

where

$$\varphi_T(\mathbf{U}) = \frac{1}{T} \sum_{t=1}^{T} e^{i\mathbf{U}'\boldsymbol{\Psi}_t}.$$

10 Conclusion

We propose a method of estimation of SVM which is convenient to apply when the innovations driving the volatility parameter have a simple characteristic function. This case includes certain heavy–tailed SVM which are popular with applied researchers. Monte Carlo results show that overall the method works quite well provided that the sample size is adequately large. In this connection we note that while these results pertain to the estimators of the underlying SVM, we expect that resulting volatility forecasts would also be fairly accurate, and possibly would outperform corresponding forecasts based on a GARCH model; see for instance [6]. We also suggest a goodness–of–fit test which can be readily applied in order to decide which parametric distribution in the model is more compatible with a given set of data. In principle the method is shown to extend to multivariate SVM in a straightforward manner. However given the high–dimensionality of the parameter space with such models, more work is needed in this direction in order to verify the actual performance of the method both in terms of computational cost as well as in terms of estimation efficiency.

Acknowledgements The research of Simos Meintanis was partially supported by grant Nr. 11699 of the Special Account for Research Grants (EΛKE) of the National and Kapodistrian University of Athens.

References

1. Abanto–Valle, C.A., Bandyopadhyay, D., Lachos, V.H., and Enriquez, I., Robust Bayesian analysis of heavy–tailed stochastic volatility models using scale mixtures of normal distributions, Comput. Statist. Dat. Anal., 54, 2883–2898 (2010).
2. Andersson, J., The normal inverse Gaussian stochastic volatility model, J. Bus. Econom. Statist., 19, 44–54 (2001).
3. Asai, M., McAleer, M., and Yu, J., Multivariate stochastic volatility models: A review, Econometr. Rev., 25, 145–175 (2006).
4. Barndorff-Nielsen, O.E., Exponentially decreasing distributions for the logarithm of particle size, Proc. Roy. Soc. Lond., Series A, 353, 401-409 (1977).
5. Barndorff-Nielsen, O.E., Normal inverse Gaussian distributions and stochastic volatility modelling, Scand. J. Statist., 24, 1–13 (1997).
6. Boscher, H., Fronk, E.V., and Pigeot, I., Forecasting interest rates volatilities by GARCH(1,1) and stochastic volatility models, Stat. Papers, 41, 409–422 (2000).
7. Chib, S., Omori, Y., and Asai, M., Multivariate stochastic volatility. In T.G. Andersen et al. (Eds.): Handbook of Financial Time Series, 365–400. Springer, Berlin (2009).
8. Curto, J.D., Pinto, J.C., and Tavares, G.N., Modeling stock markets' volatility using GARCH models with Normal, Student's t and Stable Paretian distributions, Stat. Papers, 50, 311–321 (2009).
9. Figueroa–López, J.E., Lancette, S.R., Lee, K., and Mi, Y., Estimation of NIG and VG models for high frequency financial data. In Viens et al. (Eds.): Handbook of Modeling High–Frequency Data in Finance, 3–26. Wiley, New Jersey (2012).

10. Forsberg, L., On the Normal Inverse Gaussian Distribution in Modeling Volatility in the Financial Markets, PhD Dissertation, Uppsala University (2002).
11. Geske, R., and Torous, W., Skewness, kurtosis, and black–scholes option mispricing, Stat. Papers, 32, 299–309 (1991).
12. Giacomini, R., Politis, D. N. and White, H., A warp-speed method for conducting monte carlo experiments involving bootstrap estimators, Econometr. Theor., 29, 567–589 (2013).
13. Hansen, P.R., Huang, Z., and Shek, H.H., Realized GARCH: A joint model for returns and realized measures of volatility, J. Appl. Econometr., 27, 877–906 (2012).
14. Harvey, A.C., Ruiz, E., and Shephard, N., Multivariate stochastic variance models, Rev. Econom. Stud., 61, 247–264 (1994).
15. Jiménez–Gamero, M.D., On the empirical characteristic function process of the residuals in GARCH models and applications, TEST, 23, 409–432 (2014).
16. Jungbacker, B., and Koopman, S.J., Parameter estimation and practical aspects of modeling stochastic volatility, In Andersen et al. (Eds.): Handbook of Financial Time Series, 313–344, Springer, Berlin (2009).
17. Knight, J.L., Satchell, S.E., and Yu, J., Estimation of the stochastic volatility model by the empirical characteristic function method, Austral. New Zeal. J. Statist., 44, 319–335 (2002).
18. Knight, J.L., and Yu, J., Empirical characteristic function in time series estimation, Econometr. Theor., 18, 691–721 (2002).
19. Kotchoni, R., Applications of the characteristic function–based continuum GMM in finance, Comput. Statist. Dat. Anal., 56, 3599–3622 (2012).
20. Langrock, R., MacDonald, I.L., and Zucchini, W., Some nonstandard stochastic volatility models and their estimation using structured hidden Markov models, J. Empir. Finan., 19, 147–161 (2012).
21. Liesenfeld, R., and Jung, R.C., Stochastic volatility models: Conditional normality versus heavy–tailed distributions, J. Appl. Econometr., 15, 137–160 (2000).
22. Lin, L.C, Lee, S., and Guo, M., Goodness–of–fit tests for stochastic volatility models, J. Multivar. Anal., 116, 473–498 (2013).
23. Meintanis, S.G., and Taufer, E., Inference procedures for stable–Paretian stochastic volatility models, Math. Comput. Model., 55, 1199–1212 (2012).
24. Nagel, H., and Schöbel, R., Volatility and GMM–Monte Carlo studies and empirical estimations, Stat. Papers, 40, 297–321 (1999).
25. Ngatchou–Wandji, J., Harel, M., A Cramér–von Mises test for symmetry of ther error distribution in asymptotically stationary stochastic models, Statist. Infer. Stoch. Process., 16, 207–236 (2013).
26. Qin, R., Tian, Z., and Jin, H., Truncating estimation for the change in stochastic trend with heavy–tailed innovations, Stat. Papers, 52, 203–217 (2011).
27. R Core Team, R: A language and environment for statistical computing, Statistical Computing, Vienna, Austria (2014).
28. Renault, E., Moment–based estimation of stochastic volatility models, In Andersen et al. (Eds.): Handbook of Financial Time Series, 269–311. Springer, Berlin (2009).

29. Scott, D., and Yang Dong, C., VarianceGamma: The Variance Gamma Distribution, R package version 0.3-1 (2012).
30. Seneta, E., Fitting the Variance-Gamma Model to Financial Data, Stochastic Methods and Their Applications, 41, 177–187 (2004).
31. Stojanović, V.S, Popović, B.C., and Milovanović, G.V., The split–SV model, Comput. Statist. Dat. Anal., 100, 560–581 (2016).
32. Taufer, E., Leonenko, N., and Bee, M., Characteristic function estimation of Orstein–Uhlenbeck–based stochastic volatility models, Comput. Statist. Dat. Anal., 55, 2525–2539 (2011).
33. Varadhan, R., and Gilbert, P.D., BB: An R Package for Solving a Large System of Nonlinear Equations and for Optimizing a High-Dimensional Nonlinear Objective Function, J. Statistical Software, Vol. 32, Iss. 4., 1–26 (2009).
34. Wuertz, D., and Rmetrics core team members ,fBasics: Rmetrics - Markets and Basic Statistics, R package version 3010.86 (2013).
35. Yu, J., Empirical characteristic function estimation and its applications, Econometr. Theor., 18, 691–721 (2004).
36. Yu, J., and Meyer, R., Multivariate stochastic volatility models: Bayesian estimation and model comparison, Econometr. Rev., 25, 361–384 (2006).