Sankhyā : The Indian Journal of Statistics
Special Issue on Bayesian Analysis
1998, Volume 60, Series A, Pt. 3, pp. 459-475

BAYESIAN INFERENCE FOR VECTOR ARMA MODELS WITH STABLE INNOVATIONS*

By ZUQIANG QIOU
Palisades Research Inc.
and
NALINI RAVISHANKER
University of Connecticut, Storrs

SUMMARY. This article describes Bayesian inference for a multivariate time series model with infinite variance stable innovations. Specifically, the innovations are generated from a class of multivariate symmetric stable distributions while a vector autoregressive moving average (VARMA) process characterizes the stochastic evolution of the multiple time series. The Bayesian approach facilitates simultaneous estimation of the parameters characterizing the stable law, together with the parameters of the VARMA model. Our approach uses a Metropolis-Hastings algorithm to generate samples from the joint posterior distribution of all the parameters and is an extension to multivariate time series processes of Qiou and Ravishanker's (1997) approach for the univariate case. We illustrate our approach using simulated data.

1. **Introduction**

Vector autoregressive moving average (VARMA) models with Gaussian innovations are widely used to model multiple time series in order to characterize the stochastic dependence within series and between series. However, not much work has been done on modeling multiple time series with heavy-tailed innovations. In univariate time series, infinite variance may be modeled by assuming non-Gaussian stable innovations instead of the usual assumption of Gaussian innovations. Since the work of Mandelbrot (1960, 1963, 1969), stable distributions have been used to model outlier-prone time series data (i.e. data suspected to have heavy-tailed distributions) including stock price changes (see DuMouchel,

AMS (1991) *subject classifications.* 62A15, 62E25, 62F15, 62M10.
Key words and phrases. Infinite variance, metropolis- hastings algorithm, posterior distribution, VARMA models.
* Work partly supported by an NSF Grant DMS-9510348.

1981 and Granger and Orr, 1972) and teletraffic data (Resnick, 1997). Inference for such processes is discussed by Bhansali (1993), Cline and Brockwell (1985), Davis and Resnick (1985), Knight (1987) and Qiou and Ravishanker (1997) among others. Unlike the other approaches, the sampling based Bayesian approach of Qiou and Ravishanker (1997) enables the simultaneous estimation of the stable law parameters and the parameters of the linear ARMA model. Their approach uses a modified Gibbs sampling algorithm to generate samples from the joint posterior distribution of the parameters and is an extension to time series of Buckle's (1995) approach for independent stable variables. A summary of the literature is given in Ravishanker and Qiou (1998). In this paper, we describe the details of the sampling based Bayesian approach for multiple time series with infinite variance stable innovations.

Let $\tilde{Z}_t$ denote a k-variate time series generated by a VARMA(p,q) process (we use ˜ to denote a vector):

$$\Phi(B)\left(\tilde{Z}_t-\tilde{\mu}\right)=\Theta(B)\tilde{\epsilon}_t, \qquad \ldots(1)$$

where $\Phi(B)=\mathrm{I}-\Phi_1B-\cdots-\Phi_pB^p$ is a matrix polynomial of degree p, $\Theta(B)=\mathrm{I}-\Theta_1B-\cdots\Theta_qB^q$ is a matrix polynomial of degree q, $\tilde{\mu}=(\mu_1,\ldots,\mu_k)^T$ is the location vector and T denotes transpose. We assume that Φ and Θ obey the usual stationarity and invertibility conditions. In order to ensure model identifiability, we further assume that Φ and Θ are left coprime and rank $(\Phi_p,\Theta_q)=k$ (see, *e.g.*, Reinsel, 1993, p.36). Let $\tilde{\Phi}=Vec(\Phi_1,\ldots,\Phi_p)$ and $\tilde{\Theta}=Vec(\Theta_1,\ldots,\Theta_q)$. In (1), $\tilde{\epsilon}_t$ are k-variate *iid* symmetric stable random variables, a definition of which follows.

The multivariate stable distribution (Press, 1972a, 1972b, Samorodnitsky and Taqqu, 1994) is a direct generalization of the definition of a univariate stable distribution so that all linear combinations of elements of the multivariate stable random variable are univariate stable and all linear combinations of multivariate stable variables are also multivariate stable. A k-dimensional random vector $\tilde{X}=(X_1,X_2,\ldots,X_k)^T$ with characteristic function $\phi_{\tilde{X}}(\tilde{t})$ where $\tilde{t}=(t_1,t_2,\ldots,t_k)^T$ follows a multivariate stable distribution if, and only if

$$\phi_{\tilde{X}}\left(\tilde{t}\right)=\exp\left(i\mu\left(\tilde{t}\right)-\sigma\left(\tilde{t}\right)\left[1+i\beta\left(\tilde{t}\right)\omega\left(1,\alpha\right)\right]\right) \qquad \ldots(2)$$

where $i=\sqrt{-1}$ and functions $\mu(\tilde{t})$, $\sigma(\tilde{t})$ and $\beta(\tilde{t})$ are such that for every scalar s, they must satisfy

$$\begin{array}{lcll}
\sigma\left(\tilde{t}s\right) & = & |s|^{\alpha}\sigma\left(\tilde{t}\right) & \sigma\left(\tilde{t}\right)\in(0,\infty)\\
\beta\left(\tilde{t}s\right) & = & \frac{s}{|s|}\beta\left(\tilde{t}\right) & \beta\left(\tilde{t}\right)\in[-1,1]\\
\mu\left(\tilde{t}s\right) & = & \mu\left(\tilde{t}\right)s-\sigma\left(\tilde{t}\right)\beta\left(\tilde{t}\right)|s|^{\alpha}\frac{s}{|s|}\left[\omega\left(s,\alpha\right)-\omega\left(1,\alpha\right)\right] & \mu\left(\tilde{t}\right)\in(-\infty,\infty)
\end{array}$$

where $\omega=\tan\frac{\alpha\pi}{2}$ if $\alpha\neq1$ and $\omega=\frac{2}{\pi}\log|s|$ if $\alpha=1$. In what follows, we assume that the VARMA process has sub-Gaussian symmetric stable innovations (defined in Section 2) and describe Bayesian inference using Markov chain Monte

Carlo methods. The outline of the paper is as follows. Section 2 presents a form of the density of a sub-Gaussian symmetric stable vector. Section 3 presents the Bayesian framework for inference using the Gibbs sampler. Complete conditional densities associated with the parameters are presented and the sampling algorithm is described. The Appendix contains a proof that the joint posterior density is proper. Section 4 presents two illustrations of this approach using simulated data.

2. Form of the Multivariate Stable Density Function

Let A denote a univariate stable random variable with stability parameter $\alpha/2$, skewness parameter $\beta = 1$, scale parameter $\sigma_\alpha = (\cos\frac{\pi\alpha}{4})^{2/\alpha}$ and location parameter $\delta = 0$ (see Samorodnitsky and Taqqu, 1994). Let $\tilde{G} = (G_1, G_2, ..., G_k)^T$ be a zero mean Gaussian vector independent of A and let $\tilde{\mu} = (\mu_1, \mu_2, ..., \mu_k)^T$ be a fixed vector. Then the random vector $\tilde{X} = \tilde{\mu} + (A^{1/2}G_1, A^{1/2}G_2, ..., A^{1/2}G_k)^T$ has a k-dimensional symmetric stable distribution with characteristic function

$$E\exp\left\{i\sum_{j=1}^{k}\theta_j X_j\right\} = \exp\left\{i\sum_{l=1}^{k}\mu_l\theta_l - |\frac{1}{2}\sum_{j=1}^{k}\sum_{l=1}^{k}\theta_j\theta_l\omega_{jl}|^{\alpha/2}\right\} \qquad \ldots(3)$$

where $\omega_{jl} = EG_jG_l,\ j,l = 1,\cdots,k$ are the covariances of the underlying Gaussian random vector $(G_1, G_2, ..., G_k)^T$. Let $\Omega = \{\omega_{jl}\}$. Now if we make a transformation by letting $\tilde{G} = \frac{\tilde{X}-\tilde{\mu}}{A^{1/2}}$ and $A = S$, then the Jacobian of the transformation is $S^{-k/2}$ and the joint density of $(\tilde{X}, S)$ is:

$$f(\tilde{x}, s) = f(\tilde{g}, a)|J| = f_g\left(\frac{\tilde{x}-\tilde{\mu}}{s^{1/2}}\right) f_s(s) s^{-k/2} \qquad \ldots(4)$$

where

$$f_g\left(\frac{\tilde{x}-\tilde{\mu}}{s^{1/2}}\right) = \frac{1}{(2\pi)^{k/2}|\Omega|^{1/2}} e^{-\frac{1}{2}(\tilde{x}-\tilde{\mu})^t\frac{\Omega^{-1}}{s}(\tilde{x}-\tilde{\mu})} \qquad \ldots(5)$$

and

$$f_s(s) = \frac{\alpha}{|\alpha-2|}\int e^{-|\frac{s}{\sigma_\alpha t_\alpha(y)}|^{\frac{\alpha}{\alpha-2}}} |\frac{s}{\sigma_\alpha t_\alpha(y)}|^{\frac{\alpha}{\alpha-2}} \frac{1}{s} dy. \qquad \ldots(6)$$

We define $t_\alpha(y) = t_{\frac{\alpha}{2},\beta}(y)$ with $\beta = 1$ where $t_{\alpha,\beta}(y)$ is

$$t_{\alpha,\beta}(y) = \left(\frac{\sin[\pi\alpha y + \frac{\pi\beta}{2}\min(\alpha, 2-\alpha)]}{\cos(\pi y)}\right)\left(\frac{\cos(\pi y)}{\cos[\pi(\alpha-1)y + \frac{\pi\beta}{2}\min(\alpha, 2-\alpha)]}\right)^{\frac{(\alpha-1)}{\alpha}},$$

$\sigma_\alpha = (\cos\frac{\pi\alpha}{4})^{\frac{2}{\alpha}}$ and $0 < \alpha < 2$. If we further consider an "augmented variable" Y (see Buckle, 1995), we can write the joint density of $(\tilde{X}, S, Y)$ as:

$$f(\tilde{x}, s, y) = \frac{1}{(2\pi)^{k/2}|\Omega|^{1/2}} \frac{\alpha}{|\alpha-2|} \exp\{-|\frac{s}{\sigma_\alpha t_\alpha(y)}|^{\frac{\alpha}{\alpha-2}} - \frac{1}{2}(\tilde{x}-\tilde{\mu})^T \frac{\Omega^{-1}}{s}(\tilde{x}-\tilde{\mu})\} |\frac{s}{\sigma_\alpha t_\alpha(y)}|^{\frac{\alpha}{\alpha-2}} \frac{1}{s^{\frac{k}{2}+1}}. \quad \ldots (7)$$

The next section describes the Bayesian framework for inference; the complete conditional distributions for the parameters that facilitate the Gibbs sampling algorithm are obtained through (7).

3. Bayesian Inference through Gibbs Sampling

Let $\tilde{Z}_n = (\tilde{z}_1, ..., \tilde{z}_n)$ denote n observations from a k-variate VARMA process. In general, given data $\tilde{Z}_n$, along with model parameters, the Bayesian specification requires a likelihood and a prior, from which, using Bayes theorem, we obtain the posterior density as a normalized product of the likelihood and the prior. Suppose we denote the prior of all the parameters by $\pi(\alpha, \tilde{\mu}, \Omega, \tilde{\Phi}, \tilde{\Theta})$ and denote $\Theta^{-1}(B)\Phi(B)$ by $\Theta^{-1}\Phi(B)$; then from (7) and (1), the likelihood function $L(\tilde{Z}_n|\alpha, \tilde{\mu}, \Omega, \tilde{\Phi}, \tilde{\Theta})$ is proportional to

$$\left(\frac{\alpha}{|\Omega|^{1/2}|\alpha-2|}\right)^n \prod_{j=1}^{n} \int\int \exp\{-(|\frac{s_j}{\sigma_\alpha t_\alpha(y_j)}|^{\frac{\alpha}{\alpha-2}} + \frac{1}{2}[\Theta^{-1}\Phi(B)(\tilde{z}_j-\tilde{\mu})]^T \frac{\Omega^{-1}}{s_j} [\Theta^{-1}\Phi(B)(\tilde{z}_j-\tilde{\mu})])\} |\frac{s_j}{\sigma_\alpha t_\alpha(y_j)}|^{\frac{\alpha}{\alpha-2}} \frac{1}{s_j^{\frac{k}{2}+1}} dy_j ds_j.$$

For the prior specifications, we assume that $\pi(\alpha) = 1/2, 0 < \alpha \le 2$, $\pi(\mu_i) = 1, -\infty < \mu_i < \infty, i = 1, 2, ..., k$, $\pi(\tilde{\mu}) = \prod_{i=1}^k \pi(\mu_i) = 1$ and $\pi(\Omega) = \frac{1}{|\Omega|^{\frac{k+1}{2}}}$. We adopt a uniform prior for $(\tilde{\Phi}, \tilde{\Theta})$ over the stationarity and invertibility region. By an assumption of independence, $\pi(\alpha, \tilde{\mu}, \Omega, \tilde{\Phi}, \tilde{\Theta}) = \pi(\alpha)\pi(\tilde{\mu})\pi(\Omega)\pi(\tilde{\Phi}, \tilde{\Theta}) = \frac{1}{2|\Omega|^{\frac{k+1}{2}}}$. Let $\tilde{y} = (y_1, y_2, ..., y_n)$ and $\tilde{s} = (s_1, s_2, ..., s_n)$. The joint posterior density of the parameters has the form

$$\pi(\alpha, \tilde{\mu}, \Omega, \tilde{\Phi}, \tilde{\Theta}|\tilde{Z}_n) \propto \frac{1}{|\Omega|^{(n+k+1)/2}}\left(\frac{\alpha}{|\alpha-2|}\right)^n \times \prod_{j=1}^{n} \int\int \exp\{-(|\frac{s_j}{\sigma_\alpha t_\alpha(y_j)}|^{\frac{\alpha}{\alpha-2}} + \frac{1}{2}[\Theta^{-1}\Phi(B)(\tilde{z}_j-\tilde{\mu})]^T \frac{\Omega^{-1}}{s_j}[\Theta^{-1}\Phi(B)(\tilde{z}_j-\tilde{\mu})])\} |\frac{s_j}{\sigma_\alpha t_\alpha(y_j)}|^{\frac{\alpha}{\alpha-2}} \frac{1}{s_j^{\frac{k}{2}+1}} dy_j ds_j. \quad \ldots (8)$$

In the Appendix, we prove that for a stationary and invertible VARMA process, (8) is proper. Although (8) involves a double integration, we shall see that simulation through Markov chain Monte Carlo methods, together with the form of the joint density representation (7) enables us to generate samples from the joint posterior density of the parameters. The idea is to run the Gibbs sampler on an augmented vector of unknowns,viz. $\tilde{y}$ and $\tilde{s}$ and generate samples from $\pi(\alpha, \tilde{\mu}, \Omega, \tilde{\Phi}, \tilde{\Theta}|\tilde{Z}_n)$ as described below. This approach is an extension to multivariate time series of Buckle's (1995) idea for independent observations and Qiou and Ravishanker's (1997) application to univariate time series.

Given initial values α_0, $\tilde{\mu}_0$, Ω, $\tilde{\Phi}$, $\tilde{\Theta}$ and $\tilde{s}_0 = (s_{1,0}, s_{2,0}, ..., s_{n,0})$, we repeatedly generate samples of parameters $\alpha, \tilde{\mu}$, Ω and $(\tilde{\Phi}, \tilde{\Theta})$ as well as the augmented vectors $\tilde{y}$ and $\tilde{s}$ from their respective complete conditional posterior densities, viz., $\pi(\tilde{y}|\alpha, \tilde{\mu}, \Omega, \tilde{\Phi}, \tilde{\Theta}, \tilde{Z}_n, \tilde{s})$, $\pi(\tilde{s}|\alpha, \tilde{\mu}, \Omega, \tilde{\Phi}, \tilde{\Theta}, \tilde{Z}_n, \tilde{y})$, $\pi(\alpha|\tilde{\mu}, \Omega, \tilde{\Phi}, \tilde{\Theta}, \tilde{Z}_n, \tilde{y}, \tilde{s})$, $\pi(\Omega|\alpha,\ \tilde{\mu}, \tilde{\Phi}, \tilde{\Theta}, \tilde{Z}_n, \tilde{y}, \tilde{s})$, $\pi(\tilde{\mu}|\alpha, \Omega, \tilde{\Phi}, \tilde{\Theta}, \tilde{Z}_n, \tilde{y}, \tilde{s})$ and $\pi((\tilde{\Phi}, \tilde{\Theta})|\alpha, \Omega, \tilde{\mu}, \tilde{Z}_n, \tilde{y}, \tilde{s})$. Although the use of a stochastic EM algorithm to obtain initial values was suggested in Qiou (1996), the procedure is time consuming. Moreover, our experience shows that convergence of the sampler is achieved even with arbitrary initial starting values that are relatively distant from the true values, subject to the usual model restrictions (such as stationarity) and positive elements of $\tilde{s}_0$. Given these initial values, we run a single Markov chain for a fairly large number of iterations and monitor convergence. After the chain has converged, say at the j^{th} iteration, a set of samples (of size M) is obtained by choosing every h^{th} sample, where the sample autocorrelations of the chain at lags $l \geq h$ are very small. Various summary features of interest are obtained from these "independent samples" from the posterior. The next six subsections present the complete conditional distributions of the parameters which facilitate the sampling.

3.1. *Complete conditional distribution for $\tilde{y}$.* From (7), the complete conditional distribution for y_j is:

$$\begin{aligned}\pi(y_j|\alpha, \tilde{\mu}, \Omega, \tilde{\Phi}, \tilde{\Theta}, \tilde{Z}_n, \tilde{s}) &= \pi(y_j|\alpha, s_j) \\ &= Ce^{-\left(\frac{s_j}{\sigma_\alpha t_\alpha(y_j)}\right)^{\frac{\alpha}{\alpha-2}}} \left(\frac{s_j}{\sigma_\alpha t_\alpha(y_j)}\right)^{\frac{\alpha}{\alpha-2}}\end{aligned}$$

where $s_j > 0$, and $y_j \in (l_\alpha, \frac{1}{2})$ with $l_\alpha = -\frac{\min(\alpha/2, 2-\alpha/2)}{\alpha}$ and $\sigma_\alpha = (\cos\frac{\pi\alpha}{4})^{\frac{2}{\alpha}}$. For the univariate stable distribution, Buckle (1995) showed that $t_\alpha(y_j)$ is continuous and monotonic on (-1/2,1/2), $t_\alpha(1/2) = \infty$, $t_\alpha(-1/2) = -\infty$ and $t_\alpha(l_\alpha) = 0$. These desirable properties of t_α together with the unimodality of the function ue^{-u} for any nonnegative u guarantees that the density function attains a maximum of e^{-1} at $t_\alpha(y_j) = s_j$. A slight modification yields

$$\pi(y_j|\alpha, s_j) = Ce^{1-\left(\frac{s_j}{\sigma_\alpha t_\alpha(y_j)}\right)^{\frac{\alpha}{\alpha-2}}} \left(\frac{s_j}{\sigma_\alpha t_\alpha(y_j)}\right)^{\frac{\alpha}{\alpha-2}} = Cg(y_j)h(y_j) \qquad \ldots(9)$$

where $g(y_j) = 1$ is Uniform on $(l_\alpha, \frac{1}{2})$ and $h(y_j) = e^{1-(\frac{s_j}{\sigma_\alpha t_\alpha(y_j)})^{\frac{\alpha}{\alpha-2}}} (\frac{s_j}{\sigma_\alpha t_\alpha(y_j)})^{\frac{\alpha}{\alpha-2}}$ with a range of $(0, 1]$. Hence, a rejection algorithm (Devroye, 1986) can be reasonably applied here and the generation proceeds as follows:

Step (i): Generate y_j from $Y \sim$ Uniform on $[l_\alpha, \frac{1}{2}]$.

Step (ii): Generate u from $U \sim$ Uniform on [0,1].

Step (iii): Reject the generated y_j until $u \leq h(y_j)$.

3.2. *Complete conditional distribution for* $\tilde{s}$. From (7), the complete conditional distribution of each s_j is

$$\pi(s_j|\alpha, \tilde{\mu}, \Omega, \tilde{\Phi}, \tilde{\Theta}, \tilde{Z}_n, \tilde{y}) = \pi(s_j|\alpha, \tilde{\mu}, \Omega, \tilde{\Phi}, \tilde{\Theta}, \tilde{Z}_j, y_j)$$

$$\propto s_j^{-\frac{k}{2}-1+\frac{\alpha}{\alpha-2}} e^{-\{(\frac{s_j}{\sigma_\alpha t_\alpha(y_j)})^{\frac{\alpha}{\alpha-2}} + \frac{1}{2}[\Theta^{-1}\Phi(B)(\tilde{z}_j-\tilde{\mu})]^T \frac{\Omega^{-1}}{s_j}[\Theta^{-1}\Phi(B)(\tilde{z}_j-\tilde{\mu})]\}}. \quad \ldots (10)$$

There is no obvious standard form of distribution that matches the posterior of s_j. Since a function of s of the form $f(s) = \frac{1}{s^{a+c}} \exp\{-\frac{b}{s^c\}-\frac{d}{s}}$ for all possible positive values of a, b, c, d is easily proved to be unimodal on the interval $(0, \infty)$ and to be bounded from above, the ratio of Uniforms method of generating random variates (Kinderman and Monahan, 1977, Wakefield *et al*, 1991), which requires the density in question to be known only up to proportionality and can besides easily deal with distributions with unbounded sample spaces, is appropriate here. Using this approach, we generate s_j's as follows:

Step (i): Compute $a_j = \sqrt{\sup[h(s_j)]}$, $\quad b_j = \sqrt{\sup[s_j^2 h(s_j)]}$, where

$$h(s_j) = s_j^{-\frac{k}{2}-1+\frac{\alpha}{\alpha-2}} e^{-\{(\frac{s_j}{\sigma_\alpha t_\alpha(y_j)})^{\frac{\alpha}{\alpha-2}} + \frac{1}{2}[\Theta^{-1}\Phi(B)(\tilde{z}_j-\tilde{\mu})]^T \frac{\Omega^{-1}}{s_j}[\Theta^{-1}\Phi(B)(\tilde{z}_j-\tilde{\mu})]\}}.$$

Although this cannot be obtained analytically, it can be easily obtained through a numerical search algorithm; we use a very simple bisection method.

Step (ii): Generate u from Uniform $(0, a_j)$, and v from Uniform $(0, b_j)$.

Step (iii): If $u > \sqrt{h(\frac{v}{u})}$ go to *Step (ii)*, otherwise, $s_j = \frac{v}{u}$.

3.3. *Complete conditional distribution for* α. From (7), we have

$$\begin{aligned}\pi(\alpha|\tilde{\mu}, \Omega, \tilde{\Phi}, \tilde{\Theta}, \tilde{Z}_n, \tilde{y}, \tilde{s}) &= \pi(\alpha|\tilde{y}, \tilde{s}) \\ &\propto \frac{\alpha^n}{|\alpha-2|^n} e^{-\sum_{j=1}^{n}\left(\frac{s_j}{\sigma_\alpha t_\alpha(y_j)}\right)^{\frac{\alpha}{\alpha-2}}} \prod_{j=1}^{n}\left(\frac{s_j}{\sigma_\alpha t_\alpha(y_j)}\right)^{\frac{\alpha}{\alpha-2}}.\end{aligned}$$

Transforming $v_j = t_\alpha(y_j)$ gives

$$\pi(\alpha|\Omega,\tilde{v},\tilde{s}) \propto \frac{\alpha^n}{|\alpha-2|^n} e^{-\sum_{j=1}^{n}\left(\frac{s_j}{\sigma_\alpha v_j}\right)^{\frac{\alpha}{\alpha-2}}} \prod_{j=1}^{n}\left(\frac{s_j}{\sigma_\alpha v_j}\right)^{\frac{\alpha}{\alpha-2}} |\frac{dt_\alpha(y)}{dy}|^{-1}_{t_\alpha(y)=v_j}.$$

In the univariate case, the advantage of this transformation, as pointed out by Buckle(1995), is that it supports the parameter range more evenly and makes the density under the transformation unimodal. However, in the multivariate case, the posterior of α is still quite undulating and at high values of α, it could become very spiked under the above transformation. This happens as a consequence of the big difference between the following two situations. First, the scale parameter σ of the univariate stable variate s is a function of α viz., $\sigma = (\cos\frac{\pi\alpha}{4})^{\frac{2}{\alpha}}$, when α approaches 2, σ approaches zero. Secondly, unlike the univariate case where the vector $\tilde{s}$ is fixed, here, it is an augmented parameter vector and gets updated in every iteration.

Since the complete conditional distribution for α does not have a standard form, we use the Metropolis-Hastings algorithm (Chib and Greenberg, 1995, Hastings, 1970) with a linearly transformed Beta distribution as the proposal density since α is bounded both above and below and the density appears to be unimodal (see Buckle, 1995). Therefore, at the k-th iteration, a new α is generated as follows:

Step (i): Generate a sample value α^* from $g(\theta|a,b)$ where

$$g(\theta|a,b) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}(\frac{\theta}{2})^{a-1}(\frac{2-\theta}{2})^{b-1}$$

is the linearly transformed Beta distribution on $(0,2)$, a and b are determined such that the current value α^* is the mode of the distribution and $a+b = 5\log n$, a constant suggested by Buckle (1995), n being the sample size.

Step (ii): Compute $u = \frac{\pi(\alpha^*|\tilde{v},\tilde{s})g(\alpha^k|\alpha^*)}{\pi(\alpha^k|\tilde{v},\tilde{s})g(\alpha^*|\alpha^k)}$.

Step (iii): Set $\alpha^{k+1} = \alpha^*$ with probability $\min(1,u)$, while $\alpha^{k+1} = \alpha^k$ with probability 1-$\min(1,u)$.

3.4. *Complete conditional distribution for* Ω. From (7), we see that

$$\pi(\Omega|\alpha,\tilde{\mu},\tilde{\Phi},\tilde{\Theta},\tilde{Z}_n,\tilde{y},\tilde{s}) = \pi(\Omega|\tilde{\mu},\tilde{\Phi},\tilde{\Theta},\tilde{Z}_n,\tilde{s})$$

$$\propto \frac{1}{|\Omega|^{(n+k+1)/2}} e^{-\sum_{j=1}^{n}\frac{1}{2}[\Theta^{-1}\Phi(B)(\tilde{z}_j-\tilde{\mu})]^T\frac{\Omega^{-1}}{s_j}[\Theta^{-1}\Phi(B)(\tilde{z}_j-\tilde{\mu})]}$$

$$\propto \frac{1}{|\Omega|^{(n+k+1)/2}} e^{-\frac{1}{2}\mathrm{tr}\Omega^{-1}\sum_{j=1}^{n}\frac{[\Theta^{-1}\Phi(B)(\tilde{z}_j-\tilde{\mu})][\Theta^{-1}\Phi(B)(\tilde{z}_j-\tilde{\mu})]^T}{s_j}} . \quad \ldots(11)$$

It is obvious from the above equation that the conditional posterior density of the inverted Ω is a Wishart distribution $W_k(B^{-1}, n+k+3)$ where n is the sample size, k is the dimension and $B = \sum_{j=1}^{n} \frac{[\Theta^{-1}\Phi(B)(\tilde{z}_j - \tilde{\mu})][\Theta^{-1}\Phi(B)(\tilde{z}_j - \tilde{\mu})]^T}{s_j}$. The generation of samples of Ω is straightforward; we generate a sample matrix from the Wishart distribution mentioned above and compute its inverse.

3.5. *Complete conditional distribution for* $\tilde{\mu}$. From (7),

$$\pi\left(\tilde{\mu}|\alpha, \Omega, \tilde{\Phi}, \tilde{\Theta}, \tilde{Z}_n, \tilde{y}, \tilde{s}\right) = \pi\left(\tilde{\mu}|\Omega, \tilde{\Phi}, \tilde{\Theta}, \tilde{Z}_n, \tilde{s}\right)$$

$$\propto \exp\{-\frac{1}{2}\left(\tilde{\mu} - \Phi^{-1}(1)\Theta(1)\Theta^{-1}(B)\Phi(B)\sum_{j=1}^{n}\frac{\tilde{z}_i}{s_i}/\sum_{j=1}^{n}\frac{1}{s_j}\right)^T$$

$$\Sigma\left(\tilde{\mu} - \Phi^{-1}(1)\Theta(1)\Theta^{-1}(B)\Phi(B)\sum_{j=1}^{n}\frac{\tilde{z}_i}{s_i}/\sum_{j=1}^{n}\frac{1}{s_j}\right)\} \qquad \ldots (12)$$

where $\Sigma = \Phi^T(1)\Theta^{-T}(1)\Omega^{-1}\Theta^{-1}(1)\Phi(1)\sum_{j=1}^{n}\frac{1}{s_j}$. It is obvious that the posterior of $\tilde{\mu}$ is Normal with mean

$$\left(\tilde{\mu} - \Phi^{-1}(1)\Theta(1)\Theta^{-1}(B)\Phi(B)\sum_{j=1}^{n}\frac{\tilde{z}_i}{s_i}/\sum_{j=1}^{n}\frac{1}{s_j}\right)$$

and variance

$$\Phi^{-1}(1)\Theta(1)\Omega\Theta^T(1)\Phi^{-T}(1)/\sum_{j=1}^{n}\frac{1}{s_j}$$

from which the generation of $\tilde{\mu}$ is straightforward.

3.6. *Complete conditional distribution for* $\tilde{\Phi}, \tilde{\Theta}$. From (7), we have

$$\pi(\tilde{\Phi}, \tilde{\Theta}|\alpha, \tilde{\mu}, \Omega, \tilde{Z}_n, \tilde{y}, \tilde{s}) = \pi(\tilde{\Phi}, \tilde{\Theta}|\tilde{\mu}, \Omega, \tilde{Z}_n, \tilde{s})$$

$$\propto \exp\{-\frac{1}{2}\sum_{j=1}^{n}[\Theta^{-1}\Phi(B)(\tilde{z}_j - \tilde{\mu})]^T\frac{\Omega^{-1}}{s_j}[\Theta^{-1}\Phi(B)(\tilde{z}_j - \tilde{\mu})]\}. \qquad \ldots (13)$$

In general, the conditional distribution of $(\tilde{\Phi}, \tilde{\Theta})$ does not have any standard form; we use the Metropolis-Hastings algorithm with an appropriate driving function. Note that the mode and the Hessian at the mode of the likelihood function are excellent choices for the mean and the covariance matrix of this Normal proposal. Restriction of $(\tilde{\Phi}, \tilde{\Theta})$ to the stationarity and invertibility region is handled through rejection. The procedure simplifies considerably for the pure VAR(p) case with $q = 0$. Then, it can be easily derived that the conditional density of $\phi_{j,l,m}$, the $\{l,m\}^{th}$ element of the j^{th} coefficient matrix Φ_j, given all other parameters (denoted by *rest* below) is $\pi(\phi_{j,l,m}|rest) \sim N(\mu_{j,l,m}, \sigma_{j,l,m})$ where

$$\begin{aligned}
\mu_{j,l,m} &= \sum_{i=1}^{n} \frac{(z_{i-j,m} - \mu_m)}{s_i}(a_i + b_i)/\sigma_{l,l} \sum_{i=1}^{n} \frac{(z_{i-j,m} - \mu_m)^2}{s_i} \\
a_i &= [(z_{i+1,l} - \mu_l - \sum_{u=1,u\neq m}^{k} \phi_{j,l,u}(z_{i-j,u} - \mu_u) \\
&- \sum_{u=1,u\neq j}^{p} \sum_{v=1}^{k} \phi_{u,l,v}(z_{i-u,v} - \mu_v)]\sigma_{l,l} \\
b_i &= \sum_{u=1,u\neq l}^{k} [(z_{i+1,u} - \mu_u) - \sum_{v=1}^{p} \sum_{w=1}^{k} \phi_{v,u,w}(z_{i-v,w} - \mu_w)]\sigma_{l,u} \\
\sigma_{j,l,m} &= 1/\sigma_{l,l} \sum_{i=1}^{n} \frac{(z_{i-j,m} - \mu_m)^2}{s_i} \\
& (j = 1, 2, ..., p), (l = 1, 2, ..., k), (m = 1, 2, ..., k)
\end{aligned}$$

4. **Illustrative Examples**

We present two examples to illustrate our approach. In both cases, we use data simulated from a VAR(1) model with sub-Gaussian symmetric stable innovations. In the generation of a bivariate VAR(1) time series, the generation of the innovations is the crucial step. We generate a univariate stable random variable with $0 < \alpha < 1$, $\beta = 1$, $\delta = 0$ and $\sigma = (\cos \frac{\alpha\pi}{4})^{\frac{2}{\alpha}}$ using the algorithm of Chambers *et al* (1976). If $\sigma = 1$, then the standard stable random variable can be written as $(a(U)/W)^{\frac{1-\alpha}{\alpha}}$ where $a(u) = \sin((1-\alpha)u)(\sin \alpha u)^{\frac{\alpha}{1-\alpha}}/(\sin(u))^{\frac{1}{1-\alpha}}, 0 < u < \pi$ where U is Uniform on $(0, \pi)$, W has a standard exponential density and U and W are mutually independent. The stable variable with scale σ is then obtained by multiplying the standard stable variable by σ. Given the univariate stable innovations, we obtain sub-Gaussian symmetric stable innovations following the definition in Section 2 and then generate the observations from a VAR(1) process using the VARMA(p, q) model function (1) with $p = 1, q = 0$. In each case, we use the Gibbs sampling procedure described in the previous sections to generate samples from the joint posterior distribution of all the parameters, thus facilitating simultaneous inference in the Bayesian framework.

Table 1. Posterior means and posterior standard deviations for VAR(1) models

Model	VAR(1) with α=1.5		VAR(1) with α=1.0	
True Param.	Post. mean	Post. stdev	Post. mean	Post. stdev
α(1.5)(1.0)	1.5283	0.2139	1.0305	0.4125
μ_1(-2.0)	-1.9902	0.0481	-1.9656	0.0942
μ_2(2.0)	1.9983	0.1324	1.9941	0.0179
σ_{11}(2.0)	2.3042	0.9115	2.4722	1.3203
$\sigma_{12}=\sigma_{21}$(0.0)	0.0068	0.0586	-0.0094	0.0658
σ_{22}(1.0)	1.1316	0.4288	1.1304	0.5536
ϕ_{11}(0.6)	0.5942	0.0080	0.5832	0.0008
ϕ_{12}(-0.8)	-0.7452	0.0129	-0.7648	0.0012
ϕ_{21}(0.3)	0.3087	0.0063	0.3029	0.0011
ϕ_{22}(-0.4)	-0.4186	0.0160	-0.4070	0.0025

Example 1. We generate $n = 1,000$ observations of a bivariate ($k = 2$) time series $\tilde{Z}_t$ as follows. We first generate the multivariate symmetric stable innovations $\tilde{\epsilon}_t$, with parameters $\alpha = 1.5$, $\mu_1 = -2.0$, $\mu_2 = 2.0$, $\omega_{11} = 2.0$, $\omega_{12} = \omega_{21} = 0.0$, $\omega_{22} = 1.0$ where $\omega_{11}, \omega_{12}, \omega_{21}$ and ω_{22} are the elements of the covariance matrix Ω and $\beta = 0$. Since $\alpha \in [1.0, 2.0)$, $\tilde{\mu}$ is the mean of the time series $\tilde{Z}_t$.

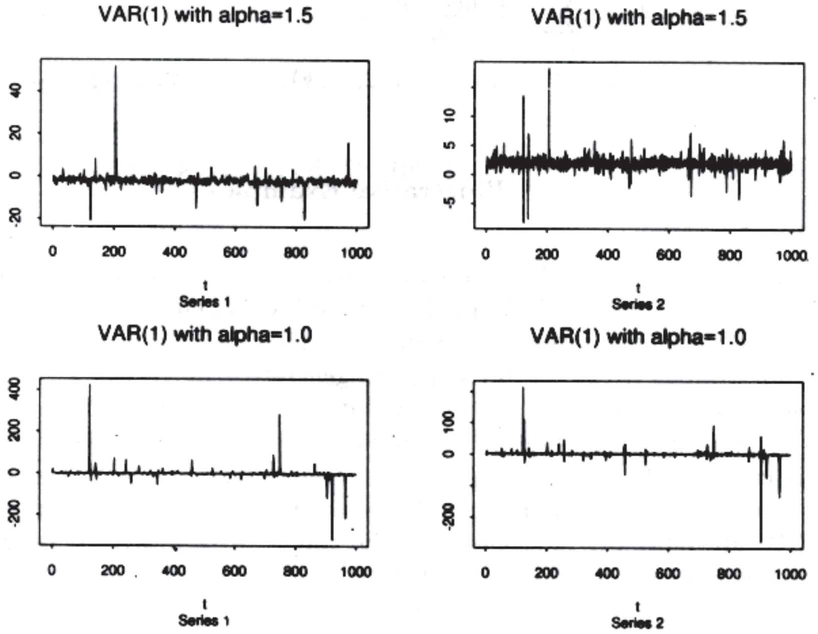

Figure 1. Simulated data from VAR(1) model with bivariate stable innovations with $\alpha = 1.5$ in (b). In (b), Series 1 si at t=120, 13416, t=121, 10675, t=122, 2133, t=903, -854, t=904, -562, t=964, -540 and Series 2 is at t=120, -3282, t=121, 5341 and t=122, 1070.

Assuming that $\tilde{Z}_1 = \tilde{\epsilon}_1$, we generate observations from a VAR(1) process with elements of Φ_1 given by $\phi_{11} = 0.6$, $\phi_{12} = -0.8$, $\phi_{21} = 0.3$, $\phi_{22} = -0.4$, based on the model function (1). The data is shown in Figure 1 (a). The sampling algorithm is implemented with 10,000 iterations of a single chain. We select every 5^{th} generation from the last 5,000 iterations to obtain a sample of 1,000 observations from the joint posterior distribution of the parameters.

Table 1 presents the posterior means and standard deviations of the parameters. The posterior means are all quite close to the true parameter values used to simulate the data, which confirms the accuracy of our approach. Estimated kernel density plots show that the estimated posterior densities of all the parameters except σ_{11} and σ_{22} are fairly symmetric. Pairwise scatter plots of the estimated parameters indicate that α is positively correlated with σ_{11} and σ_{22} (as expected), whereas there appears to be no correlation between α and the elements of Φ_1. Apart from a positive correlation between ϕ_{11} and ϕ_{12}, the scatterplot between the elements of Φ_1 show no useful information.

EXAMPLE 2. We generate $n = 1,000$ observations from a bivariate ($k = 2$) time series $\tilde{Z}_t$, with $\alpha = 1.0$, while the values of all other parameters are the same as in Example 1. Since $\alpha \in (0, 2)$, $\tilde{\mu}$ cannot be regarded as the mean of the process. Assuming that $\tilde{Z}_1 = \tilde{\epsilon}_1$, we again generate observations from a VAR(1) process; the data is shown in Figure 1 (b). The generated data has some very extreme observations. For clarity of the plot, we omitted these extreme values and constructed the plot to the scale of the remaining observations. The positions and magnitudes of these extreme observations are as follows: For Series 1, $Z_{1,120} = 13416.6, Z_{1,121} = 10675.3, Z_{1,122} = 2132.7, Z_{1,903} = -853.9, Z_{1,904} = -561.7, Z_{1,964} = -540.4$. For Series 2, $Z_{2,120} = -3281.6, Z_{2,121} = 5341.2, Z_{2,122} = 1069.6$. Table 1 presents the posterior means and standard deviations of the parameters obtained from an implementation of the sampling algorithm. Once again, the posterior modes are close to the true simulation values of the parameters and the plots offer no other interesting difference from the $\alpha = 1.5$ case, except that the densities for μ_1 and μ_2 appear to be slightly more skewed to the right. The scatter plots of α versus μ_1 and μ_2 are interesting, presenting a vertical band of α values between 0 and 2 at $\mu_1 = -2.0$ and $\mu_2 = 2.0$. A similar, though less pronounced effect is seen in the plots of α versus σ_{11}, σ_{12} and σ_{22}. There is a positive correlation between ϕ_{11} and ϕ_{12} and a negative correlation between ϕ_{21} and ϕ_{22}.

Appendix

Propriety of the posterior distribution. Let $\tilde{u}_j = \Theta^{-1}\Phi(B)\tilde{z}_j$ and let $\tilde{U}_n = (\tilde{u}_1, \tilde{u}_2, ..., \tilde{u}_n)$. Under the prior specification in Section 3, the posterior density may be written as

$$\pi(\alpha,\tilde{\xi},\Omega|\tilde{U}_n) \quad \propto \quad \frac{1}{|\Omega|^{\frac{n+k+1}{2}}}\int_0^{\infty}\exp\{-\frac{1}{2}\sum_{j=1}^{n}(\tilde{u}_j-\tilde{\xi})^T\frac{\Omega^{-1}}{s_j}(\tilde{u}_j-\tilde{\xi})\prod_{j=1}^{n}f_s(s_j)s_j^{-k/2}d\tilde{s}$$

where $\tilde{\xi}=\Theta^{-1}\Phi(B)\tilde{\mu}=\Theta^{-1}\Phi(1)\tilde{\mu}$ and $f_s(s_j)$ is defined in (6). We denote the integrand in the above expression by I, which we rewrite as

$$I=\frac{1}{|\Omega|^{\frac{n+k+1}{2}}}\prod_{j=1}^{n}f_s(s_j)s_j^{-k/2}\exp\{-\frac{1}{2}\left(\tilde{\xi}-\sum_{j=1}^{n}\frac{\tilde{u}_j}{s_j}/\sum_{j=1}^{n}\frac{1}{s_j}\right)^T$$
$$\left(\Omega/\sum_{j=1}^{n}\frac{1}{s_j}\right)^{-1}\left(\tilde{\xi}-\sum_{j=1}^{n}\frac{\tilde{u}_j}{s_j}/\sum_{j=1}^{n}\frac{1}{s_j}\right)-\frac{1}{2}\mathrm{tr}\Omega^{-1}G$$

where G is a function of $\tilde{s}$, defined as

$$G(\tilde{s})=\sum_{j=1}^{n}\frac{\tilde{u}_j\tilde{u}_j^T}{s_j}-\left(\sum_{j=1}^{n}\frac{\tilde{u}_j}{s_j}\right)\left(\sum_{j=1}^{n}\frac{\tilde{u}_j}{s_j}\right)^T/\sum_{j=1}^{n}\frac{1}{s_j}.$$

We first prove some properties of the matrix $G(\tilde{s})$.

Property 1. The matrix $G(\tilde{s})$ is positive definite for any $\tilde{U}_n$ and for any positive vector $\tilde{s}$.

PROOF. Suppose $\tilde{y}=(y_1,...,y_k)^T$ is any nonzero vector. Then

$$\begin{aligned}\tilde{y}^TG(\tilde{s})\tilde{y} &= \sum_{j=1}^{n}\frac{(\tilde{y}^T\tilde{u}_j)^2}{s_j}-\frac{\left(\sum_{j=1}^{n}\frac{\tilde{y}^T\tilde{u}_j}{s_j}\right)^2}{\sum_{j=1}^{n}\frac{1}{s_j}}=\frac{\sum_{l=1}^{n-1}\sum_{m>l}^{n}\frac{[\tilde{y}^T(\tilde{u}_l-\tilde{u}_m)]^2}{s_ls_m}}{\sum_{h=1}^{n}\frac{1}{s_h}}\\ &= \tfrac{1}{2}\sum_{l=1}^{n}\frac{1}{s_l}\left[\sum_{m=1,m\neq l}^{n}\frac{(\tilde{y}^T(\tilde{u}_l-\tilde{u}_m)^2}{s_m}/\sum_{h=1}^{n}\frac{1}{s_h}\right]\\ &= \tfrac{1}{2}\sum_{l=1}^{n}\frac{1}{s_l}\sum_{m=1,m\neq l}^{n}\left[\frac{1}{s_m}/\sum_{h=1}^{n}\frac{1}{s_h}\right]\left[\tilde{y}^T(\tilde{u}_l-\tilde{u}_m)\right]^2.\end{aligned}$$

Clearly, since $\tilde{s}$ is positive, $\tilde{y}^TG(\tilde{s})\tilde{y}\geq 0$ and it equals zero only if $\tilde{y}^T\tilde{u}_j$ is a constant for all j; but this means that either $\tilde{y}=0$ or there must exist linear relations among $\tilde{u}_j's$, which is obviously impossible. Hence, $\tilde{y}^TG(\tilde{s})\tilde{y}=0$ only if $\tilde{y}=0$, so $G(\tilde{s})$ is positive definite.

Property 2. The integral $\int \frac{1}{|\Omega|^{\frac{n+k}{2}}} e^{-\frac{1}{2}\mathrm{tr}\Omega^{-1}G(\tilde{s})} d\Omega$ is proportional to $|G(\tilde{s})|^{\frac{-n+1}{2}}$.

PROOF. Since Ω is assumed to be positive definite , the matrix $G(\tilde{s})$ is positive definite (by Property 1), and moreover since $|\frac{\partial\Omega}{\partial\Omega^{-1}}| = |\Omega|^{k+1}$, it follows that

$$\int \frac{1}{|\Omega|^{\frac{n+k}{2}}} \exp\{-\tfrac{1}{2}\mathrm{tr}\Omega^{-1}G(\tilde{s})\}d\Omega = \int \frac{1}{|\Omega|^{\frac{n+k}{2}}} \exp\{-\tfrac{1}{2}\mathrm{tr}\Omega^{-1}G(\tilde{s})\}|\tfrac{\partial\Omega}{\partial\Omega^{-1}}|d\Omega^{-1}$$
$$= \int |\Omega^{-1}|^{\frac{n-k}{2}-1} \exp\{-\tfrac{1}{2}\mathrm{tr}\Omega^{-1}G(\tilde{s})\}d\Omega^{-1} = |G(\tilde{s})|^{-\frac{1}{2}(n-1)} 2^{\frac{1}{2}k(n-1)} \Gamma_k(\tfrac{n-1}{2})$$

where $\Gamma_a(b)$ is a generalized gamma function.

Property 3. For some positive M and for any positive vector $\tilde{s}$, we have $|G(\tilde{s})|^{1/2}/(\sum_{j=1}^{n} \frac{1}{s_j})^{(k)/2} \leq M$.

PROOF. Define $\omega_j = \frac{1}{s_j}/\sum_{h=1}^{n} \frac{1}{s_h}, j = 1, \cdots, n$ and $\tilde{u} = \sum_{j=1}^{n} \omega_j \tilde{u}_j$. Then

$$G(\tilde{s}) = \sum_{j=1}^{n} \frac{\tilde{u}_j \tilde{u}_j^T}{s_j} - \left(\sum_{j=1}^{n} \frac{\tilde{u}_j}{s_j}\right)\left(\sum_{j=1}^{n} \frac{\tilde{u}_j}{s_j}\right)^T / \sum_{j=1}^{n} \frac{1}{s_j}$$
$$= \sum_{j=1}^{n} \frac{\tilde{u}_j \tilde{u}_j^T}{s_j} - \tilde{u} \sum_{j=1}^{n} \frac{\tilde{u}_j^T}{s_j} = \sum_{j=1}^{n} \frac{(\tilde{u}_j - \tilde{u})\,\tilde{u}_j^T}{s_j}$$

and hence

$$|G(\tilde{s})|^{1/2}/\left(\sum_{j=1}^{n} \frac{1}{s_j}\right)^{(k)/2} = |\sum_{j=1}^{n} \frac{(\tilde{u}_j - \tilde{u})\tilde{u}_j^T}{s_j} / \sum_{h=1}^{n} \frac{1}{s_h}|^{1/2}$$
$$= |\sum_{j=1}^{n} \omega_j(\tilde{u}_j - \tilde{u})\tilde{u}_j^T|^{1/2} = |\sum_{j=1}^{n} \omega_j \tilde{u}_j \tilde{u}_j^T - \tilde{u}\tilde{u}^T|^{1/2}.$$

Since for any positive vector $\tilde{s}$, $0 \leq \omega_j \leq 1$ for all j, $\sum_{j=1}^{n} \omega_j = 1$ and $\tilde{U}_n$ is uniformly bounded, there must exist an $M > 0$ such that $|\sum_{j=1}^{n} \omega_j \tilde{u}_j \tilde{u}_j^T - \tilde{u}\tilde{u}^T|^{1/2} \leq M$ for all $\tilde{s}$.

Property 4. Let $f(\tilde{s}) = 1/|G(\tilde{s})|^{n/2} \prod_{j=1}^{n} s_j^{k/2}$, then there exists an $N > 0$ such that $f(\tilde{s}) \leq N$ uniformly on the whole space of vector $\tilde{s}$.

PROOF. Since $f(\tilde{s})$ is continuous if $0 < s_j < \infty, j = 1, \cdots, n$, we only need to check the behavior of $f(\tilde{s})$ when at least some of the components of $\tilde{s}$ approach either 0 or ∞. From Property 2,

$$G(\tilde{s}) = \sum_{j=1}^{n} \frac{(\tilde{u}_j - \tilde{u})\tilde{u}_j^T}{s_j},$$

so that

$$\begin{aligned}|G(\tilde{s})|^{n/2}\prod_{j=1}^{n} s_j^{k/2} &= \prod_{j=1}^{n}|s_jG(\tilde{s})|^{1/2}\\ &= \prod_{j=1}^{n}|(\tilde{u}_j-\tilde{u})\tilde{u}_j^T+\sum_{m=1,m\neq j}^{n}\frac{s_j}{s_m}(\tilde{u}_m-\tilde{u})\tilde{u}_m^T|^{1/2}.\end{aligned}$$

Since each $\tilde{u}_j$ is uniformly bounded, $\tilde{u}$ depends on $\tilde{s}$ only through $(\omega_1,\cdots\omega_n)$, it is a bounded vector; it is easier and more direct to verify the following:

(1)$f(\tilde{s})\to c_1>0$ ($constant$) if $s_j\to 0$ for all j at the same rate ;
(2)$f(\tilde{s})\to 0$ if $s_j\to 0$ for some but not all j ;
(3)$f(\tilde{s})\to 0$ if $s_j\to 0$ for all j but not all at the same rate ;
(4)$f(\tilde{s})\to c_2>0$ ($constant$) if $s_j\to\infty$ for all j at the same rate;
(5)$f(\tilde{s})\to 0$ if $s_j\to\infty$ for some but not all j ;
(6)$f(\tilde{s})\to 0$ if $s_j\to\infty$ for all j but not at the same rate.

We only verify (1) here; (2) - (6) may be verified similarly. Suppose $\tilde{k}=(k_1,k_2,...,k_n)$ is an n- vector of positive constants and $\frac{s_j}{t}\to k_j$ for all j as $t\to\infty$. Then, since $G(\tilde{s})$ is positive definite,

$$|G(\tilde{s})|^{n/2}\prod_{j=1}^{n}s_j^{k/2}=|\sum_{j=1}^{n}\frac{\tilde{u}_j\tilde{u}_j^T}{s_j}-\frac{\left(\sum_{j=1}^{n}\frac{\tilde{u}_j}{s_j}\right)\left(\sum_{j=1}^{n}\frac{\tilde{u}_j}{s_j}\right)^T}{\sum_{j=1}^{n}\frac{1}{s_j}}|^{n/2}\prod_{j=1}^{n}s_j^{k/2}$$

$$\longrightarrow|\sum_{j=1}^{n}\frac{\tilde{u}_j\tilde{u}_j^T}{tk_j}-\left(\sum_{j=1}^{n}\frac{\tilde{u}_j}{tk_j}\right)\left(\sum_{j=1}^{n}\frac{\tilde{u}_j}{tk_j}\right)^T/\sum_{j=1}^{n}\frac{1}{tk_j}|^{n/2}\prod_{j=1}^{n}(tk_j)^{k/2}$$

$$=\frac{1}{t^{\frac{nk}{2}}}|\sum_{j=1}^{n}\frac{\tilde{u}_j\tilde{u}_j^T}{k_j}-\left(\sum_{j=1}^{n}\frac{\tilde{u}_j}{k_j}\right)\left(\sum_{j=1}^{n}\frac{\tilde{u}_j}{k_j}\right)^T/\sum_{j=1}^{n}\frac{1}{k_j}|^{n/2}t^{\frac{nk}{2}}\prod_{j=1}^{n}k_j^{k/2}$$

$$=|G(\tilde{k})|^{\frac{n}{2}}\prod_{j=1}^{n}k_j^{k/2}=c_1>0\ (constant),$$

thus verifying (1). When (1) –(6) hold, $f(\tilde{s})$ never becomes infinite in the limit as some or all of the components of $\tilde{s}$ tend either to 0 or to ∞ at the same or at different rates. Thus there must be some $N>0$ such that $f(\tilde{s})\leq N$ uniformly on the whole space of $\tilde{s}$, thus proving Property 4. Now it is easy to see that

$$\int\int\int\int\int\int Id\tilde{\mu}d\Omega d\tilde{s}d\alpha d\tilde{\Phi}d\tilde{\Theta}=\int\int\int\int\int\int\frac{1}{|\Theta^{-1}\Phi(1)|}Id\tilde{\xi}d\Omega d\tilde{s}d\alpha d\tilde{\Phi}d\tilde{\Theta}$$

$$= \int\int\int\int\int \frac{1}{|\Theta^{-1}\Phi(1)|} \frac{1}{|\Omega|^{\frac{n+k+1}{2}}} \prod_{j=1}^{n} f_s(s_j) s_j^{-k/2} \exp\{-\frac{1}{2}\mathrm{tr}\Omega^{-1}G\}$$

$$\left[\int \exp\{-\frac{1}{2}\left(\tilde{\xi} - \sum_{j=1}^{n}\frac{\tilde{u}_j}{s_j}/\sum_{j=1}^{n}\frac{1}{s_j}\right)^T\right.$$

$$\left.\left(\Omega/\sum_{j=1}^{n}\frac{1}{s_j}\right)^{-1}\left(\tilde{\xi} - \sum_{j=1}^{n}\frac{\tilde{u}_j}{s_j}/\sum_{j=1}^{n}\frac{1}{s_j}\right)\}d\tilde{\xi}\right] d\Omega d\tilde{s} d\alpha d\tilde{\Phi} d\tilde{\Theta}$$

$$\propto \int\int 1/|\Theta^{-1}\Phi(1)| \left\{\int\int \prod_{j=1}^{n} f_s(s_j) s_j^{-k/2} / \left(\sum_{j=1}^{n}\frac{1}{s_j}\right)^{\frac{k}{2}}\right.$$

$$\left.\left[\int 1/|\Omega|^{\frac{n+k}{2}} e^{-\frac{1}{2}\mathrm{tr}\Omega^{-1}G} d\Omega\right] d\tilde{s} d\alpha\right\} d\tilde{\Phi} d\tilde{\Theta}$$

$$\propto \left[\int\int 1/|\Theta^{-1}\Phi(1)| d\tilde{\Phi} d\tilde{\Theta}\right] \left\{\int\left[\int\left(|G|^{1/2}/\left(\sum_{j=1}^{n}\frac{1}{s_j}\right)^{k/2}\right)\right.\right.$$

$$\left.\left.\left(1/|G|^{n/2}\prod_{j=1}^{n} s_j^{k/2}\right)\prod_{j=1}^{n} f_s(s_j) d\tilde{s}\right] d\alpha\right\}$$

since

$$\int\left[\int\left(|G|^{1/2}/\left(\sum_{j=1}^{n}\frac{1}{s_j}\right)^{k/2}\right)\left(1/|G|^{n/2}\prod_{j=1}^{n} s_j^{k/2}\right)\prod_{j=1}^{n} f_s(s_j) d\tilde{s}\right] d\alpha$$

$$\leq MN\int\left[\int\prod_{j=1}^{n} f_s(s_j) d\tilde{s}\right] d\alpha = MN\int\left[\prod_{j=1}^{n}\int f_s(s_j) ds_j\right] d\alpha = 2MN < \infty.$$

Therefore, the propriety of the posterior density depends on the integral $\int\int \frac{1}{|\Theta^{-1}(1)\Phi(1)|}$ being finite. For this, we require that $|\Psi(1)| = |\Phi^{-1}(1)\Theta(1)|$ be finite, which is true for a stationary and invertible VARMA process (see Reinsel, 1993, page 8).

Acknowledgement. The authors are grateful to a referee and the editors for their helpful suggestions.

References

BHANSALI, R.J, (1993). Estimation of the impulse response coefficients of a linear process with infinite variance. *Jour. Multivariate Analysis* **45**, 274-290.

BUCKLE, D.J, (1995). Bayesian inference for stable distributions. *Jour. Amer. Statist. Assoc.* **90**, 605- 613.

CHAMBERS, J.M., MALLOWS, C.L. AND STUCK, B.W, (1976). A method for simulating stable random variables. *Jour. Amer. Statist. Assoc.* **71**, 340-344.

CHIB, S. AND GREENBERG, E, (1995). Understanding the metropolis-hastings algorithm. *American Statistician* **49**, 327-334.

CLINE, D.B.H. AND BROCKWELL, P.J, (1985). Linear prediction of ARMA processes with infinite variance. *Stochastic Processes and their Applications* **19**, 281-296.

DAVIS, R. AND RESNICK, S, (1985). Limit theory for moving averages of random variables with regularly varying tail probabilities. *Ann. Probab.* **13**, 179-195.

DEVROYE, L, (1986). *Nonuniform Random Variate Generation* Springer-Verlag:New York.

DUMOUCHEL, W.H, (1981). Estimating the stable index α in order to measure tail thickness: a critique. *Ann. Statist.* **11**, 1019-1031.

GELFAND, A.E. AND SMITH, A.F.M, (1990). Sampling based approaches to calculating marginal densities, *Jour. Amer. Statist. Assoc.* **85**, 398–409.

GRANGER, C.W.J. AND ORR, D, (1972). "Infinite Variance" and research strategy in time series analysis. *Jour. Amer. Statist. Assoc.* **67**, 275-285.

HASTINGS, W.K, (1970). Monte Carlo sampling methods using Markov chains and their applications, *Biometrika* **57**, 97-109.

KINDERMAN, A.J. AND MONAHAN, J.F, (1977). Computer generation of random variables using the ratio of uniform deviates, *ACM Transactions of Mathematical Software* **3**, 257-260.

KNIGHT, K, (1987). Rate of convergence of centered estimates of autoregressive parameters for infinite variance autoregressions. *Jour. Time Series Analysis* **8**, 51-60.

MANDELBROT, B, (1960). The Pareto-Levy law and the distribution of income. *International Economic Review* **1**, 79-106.

– – –– (1963). The variation of certain speculative prices. *Jour. Business University of Chicago* **26**, 394-419.

– – –– (1969). Long-run linearity, locally Gaussian processes, H-spectra and infinite variances. *International Economic Review* **10**, 82-111.

PRESS, S.J, (1972a). Estimation in univariate and multivariate stable distributions. *Jour. Amer. Statist. Assoc.* **67**, 842-846.

– – –– (1972b). Multivariate Stable Distributions, *Jour. Multivariate Analysis* **2**, 444-462.

QIOU, Z, (1996). Bayesian inference for Stable processes. *Doctoral dissertation*, University of Connecticut.

QIOU, Z. AND RAVISHANKER, N, (1997). Bayesian inference for time series with stable innovations. *Jour. Time Series Analysis* **19**, 235-249.

RAVISHANKER, N. AND QIOU, Z, (1998). Bayesian inference for time series with infinite variance stable innovations. In *A User's Guide to Heavy Tails: Statistical techniques for analyzing heavy tailed distributions and processes*, eds. R.J. Adler, R. Feldman, M.S. Taqqu. Birkhauser:Boston, 259-280.

REINSEL, G.C, (1993). *Elements of Multivariate Time Series Analysis* Springer Series in Statistics, Springer Verlag: New York.

RESNICK, S.I, (1997). Heavy tail modeling and teletraffic data (with discussion), *Ann. Statist.,* **25**, 1805–1869.

SAMORODNITSKY, G. AND TAQQU, M. S, (1994) *Stable Non-Gaussian Random Processes: Stochastic Models with Infinite Variance* Chapman & Hall: New York.

TANNER, M, (1993). *Tools for Statistical Inference: Observed Data and Data Augmentation methods* Lecture Notes in Statistics, Springer Verlag: New York.

WAKEFIELD, J.C., GELFAND, A.E. AND SMITH, A.F.M., (1991). Efficient generation of random variates via the ratio-of-uniforms method. *Statistics and Computing* **1**, 129–133.

NALINI RAVISHANKER
DEPARTMENT OF STATISTICS
UNIVERSITY OF CONNECTICUT
U-120, 196 AUDITORIUM ROAD,
STORRS, CT 06269
e-mail : ravishan@uconnvm.uconn.edu