Particle Filters for Markov-Switching Stochastic-Correlation Models

Giannni Amisano[1]
Dep. of Economics, University of Brescia
European Central Bank
E-mail: amisano@eco.unibs.it

Roberto Casarin [2]
Dep. of Economics, University of Brescia,
Dep. of Mathematics, University Paris Dauphine
E-mail: casarin@eco.unibs.it

Abstract: This work deals with multivariate stochastic volatility models that account for time-varying stochastic correlation between the observable variables. We focus on the bivariate models. A contribution of the work is to introduce Beta and Gamma autoregressive processes for modelling the correlation dynamics. Another contribution of our work is to allow the parameter of the correlation process to be governed by a Markov-switching process. Finally we propose a simulation-based Bayesian approach, called regularised sequential Monte Carlo. This framework is suitable for on-line estimation and the model selection.

Keywords: Markov Switching; Stochastic Correlation; Multivariate Stochastic Volatility; Particle Filters; Sequential Monte Carlo; Beta Autoregressive Process

1. Introduction

After the introduction of the *Dynamic Conditional Correlation* (DCC) model due to Engle (2002), the literature mainly focuses on the application of a time-varying correlation structure to the class of multivariate GARCH models. For an updated review on multivariate GARCH models we refer to Bauwens *et al.* (2006).
In stochastic volatility modeling, the results obtained in the univariate case (see for example Taylor (1986, 1994) and Jacquier *et al.* (1994)) have been extended to the multivariate case (see Harvey *et al.* (1994), Aguilar and West (2000) and Chib *et al.* (2006)) by considering a constant correlation structure. See also Asai and McAleer (2006a) and Asai *et al.* (2006) for an updated review on *Multivariate Stochastic Volatility* MSV.
Recent studies start dealing with time-varying correlation in MSV models. The models due to Philipov and Glickman (2006) and Gourieroux *et al.* (2004) allow for time-varying variances and covariances. However the stochastic correlation structure is implicitly determined by the dynamics of variances and covariances, thus a direct modeling of the correlation between asset returns is not possible. In this work we follow instead an alternative route. We treat variances and correlations separately. This allows us to better

---

[2]Address of correspondence: Dep. of Economics, University of Brescia, San Faustino, 74/b, 25122 Brescia, Italy

describe the dynamics of the two set of variables and the causality relations between them. For an updated discussion on the different ways of introducing time-varying correlation into MSV models see Asai and McAleer (2006b, 2005).

In our stochastic-correlation MSV models both the variance vector and the correlation matrix are stochastic unobservable (or latent) variables. The use of latent variables allows a more flexible modeling of the volatility and correlation dynamics than the DCC-GARCH approach. Unobservable components provide a natural setting to introduce exogenous explanatory variables in the correlation and variance dynamics. Nevertheless stochastic correlation MSV models are difficult to estimate and a suitable inference approach is the object of many latter works.

The first aim of our work is to propose new classes of stochastic models for representing the correlation dynamics. We propose the use of *Gamma Autoregressive Processes* due to Gourieroux and Jasiak (2006) and of the *Beta Autoregressive Processes* (BAR), which have the advantage to be naturally defined on a bounded interval. BAR are weak autoregressive processes and, similarly to the Gamma autoregressive, can accomodate various nonlinearities in the correlation dynamics. Our approach borrows from the earlier Bayesian literature on hazard rate modeling (see Nieto-Barajas and Walker (2002)), and on the construction of general autoregressive processes (see Mena and Walker (2004) and Pitt *et al.* (2002)), but represents an original application to the dynamic correlation modeling.

Another contribution of the paper is to introduce a new class of models which account for sudden changes of regimes in the correlation dynamics. The proposed stochastic correlation models can be extended by considering heavy-tailed observation noise. We propose skewed Student-t distributions, which allow for different magnitude of kurtosis and skewness in each direction of the observation space.

The last aim of our work is to apply a full Bayesian approach to the sequential estimation of the unknown parameter and to the nonlinear filtering problem which arises when estimating the hidden states. The Bayesian inference approach is powerful in dealing with the estimation of nonlinear models (Jacquier *et al.* (1994)). The existing literature on Bayesian inference for MSV and for dynamic correlation models mainly focused on traditional MCMC methods (see Asai and McAleer (2006a)). In the Bayesian framework, alternative estimation methods have been recently proposed. These methods rely upon a convenient nonlinear state-space representation of the econometric model of interest and upon *Sequential Monte Carlo* (SMC) techniques also called *Particle Filter* (PF). SMC allows to deal with nonlinear state-space models and is particularly suitable for on-line applications (see Doucet *et al.* (2001), Pitt and Shephard (1999), Polson *et al.* (2002, 2003)).

The work is organised as follows. Section 2 first defines bivariate stochastic volatility models. Then stochastic correlation models, such as Gamma and Beta autoregressive processes are introduced. Finally the Markov-switching correlation is considered. Section 3 presents the particle filter approach to the on-line estimation of the latent variables and the unknown parameters and for the selection of the model. Section 4 concludes.

2. Markov-Switching Stochastic-Correlation

Let $\mathbf{y}_t = (y_{1t}, y_{2t})' \in \mathbb{R}^2$ be a 2-dimensional vector of observable variables and $\mathbf{h}_t = (h_{1t}, h_{2t})' \in \mathbb{R}^2$ a 2-dimensional vector of stochastic log-volatilities. Denote the stochastic

correlation matrix by Ω_t and let $\Lambda_t = \text{diag}\{\exp(h_{1t}/2), \exp(h_{2t}/2)\}$ be a diagonal matrix with standard deviations on the main diagonal. The time-varying variance-covariance matrix of the observable factorises as follow: $\Sigma_t = \Lambda_t \Omega_t \Lambda_t$. Note that this decomposition allows us to model volatilities and correlations separately.
The proposed multivariate stochastic volatility model is

$$\begin{aligned} \mathbf{y}_t &= A\mathbf{y}_{t-1} + \Sigma_t^{1/2}\varepsilon_t, & \varepsilon_t \sim \mathcal{N}_2\left(\mathbf{0}, Id_2\right) \quad (1)\\ \mathbf{h}_{t+1} &= \mathbf{c} + B\mathbf{h}_t + \Xi^{1/2}\eta_t, & \eta_t \sim \mathcal{N}_2\left(\mathbf{0}, Id_2\right) \quad (2) \end{aligned}$$

where $\varepsilon_t \perp \eta_s$, $\forall\, s,t$, which denotes the absence of leverage. Id_n represents the n-dimensional identity matrix and $\mathcal{N}_n\left(\mathbf{0}, Id_n\right)$ the n-variate standard normal. $\Xi^{1/2}$ denotes the Cholesky decomposition of Ξ, which represents the variance-covariance matrix of the log-volatility process and captures instantaneous spill-over effects across asset volatilities. The autoregressive coefficient B is constant and determines the causality structure (spill-over effects) in the log-volatility. In the following we focus on the dynamic of the stochastic correlation.
Let $(s_t)_{t\geq 0}$ be a univariate Markov-switching process with values in $E = \{1, \ldots, L\} \subset \mathbb{N}$ and transition density

$$\mathbb{P}(s_{t+1} = i | s_t = j) = p_{ij}, \quad \text{with } i,j \in E. \quad (3)$$

Let $(\psi_t)_{t\geq 0}$ be a sequence of i.i.d. white noises. In this work we assume: $\psi_t \perp \varepsilon_s$ and $\psi_t \perp \eta_s$, $\forall\, s,t$, but the proposed model and inference framework can be extended to include dependence between those processes.
The (i,j)-th element of the stochastic correlation matrix Ω_t is defined as follows

$$\begin{aligned} \rho_{ij,t} &= (1 - \delta_i(j))\phi(\omega_t) + \delta_i(j) & (4)\\ \omega_{t+1} &= g(\omega_t, s_{t+1}, \psi_t) & (5) \end{aligned}$$

with $i,j \in \{1,2\}$ and $\phi : \mathbb{R} \to [-1,+1]$ a smooth function. In the previous equation ω_t is a latent process (customarily called *transformed correlation process*) driving the correlation between observable processes. The function $g : (\mathbb{R} \times E \times \mathbb{R}) \to \mathbb{R}$ defines the transition dynamics of ω_t.
In the following examples we discus some stochastic correlation models. In particular we focus on some alternative ways to specify the stochastic correlation process. We show also how to introduce a switching-regimes process in the dynamics of the correlation process.

Example 1 - (Switching Gaussian Autoregressive)
We assume the transformed correlation process ω_t follows a Markov-switching first order Gaussian autoregressive process

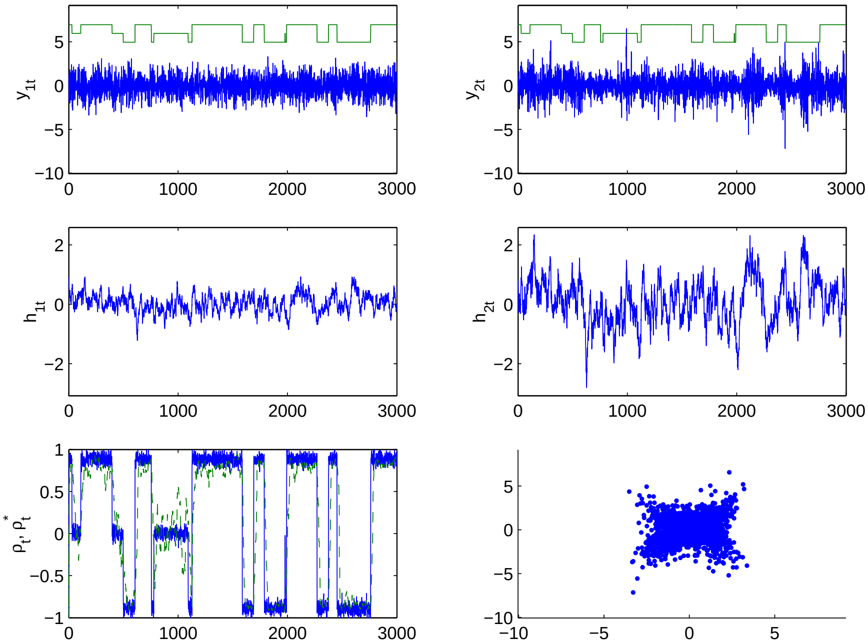
$$\omega_{t+1} = \bar{\omega}_{s_{t+1}} + \lambda_{s_{t+1}}\omega_t + \gamma_{s_{t+1}}^{1/2}\psi_t, \quad \psi_t \sim \mathcal{N}(0,1), \quad (6)$$

with $\gamma_k > 0$, $\forall k$. Note that our model has the pure Markov-switching correlation process as a special case, when $\lambda_k = 0$ and $\gamma_k = 0\ \forall k$.
In order to obtain a process which lives in the interval $[-1,+1]$, we could apply to ω_t the Fisher's transform

$$\phi(\omega_t) = \frac{\exp(\omega_t) - 1}{\exp(\omega_t) + 1},$$

. A simulated example of Markov-switching stochastic-correlation is given in Fig. 1.

Figure 1: *Simulated bivariate stochastic volatility model with Markov-switching stochastic correlation. Top left and right simulated observable process assuming* $A = diag\{0.05, 0.05\}$*. In the second line simulated stochastic log-volatility process assuming* $\mathbf{c} = (0.001, 0.001)$, $B = diag\{0.95, 0.97\}$ *and* $\Xi = diag\{0.1, 0.2\}$*. The bottom-left chart shows the stochastic correlation process (continuous line) with parameters:* $\lambda = 0.1$, $\gamma = 0.04$, $\bar{\omega}_1 = -0.8$, $\bar{\omega}_2 = 0$, $\bar{\omega}_3 = 0.8$ *and the recursive estimation of the correlation (dashed line) with a moving window of 30 observations. The bottom-right chart shows the scatter plot of the observations.*



In these models, even if the dynamic of the latent process ω_t is a linear one, conditionally on s_t, the proposed transforms produce nonlinearities in the dynamic of the correlation process making the inference and forecast more difficult. For example, when the Fisher's transform is used, the first conditional moment is

$$\mathbb{E}(\rho_t|s_t, \rho_{t-1}) = \sum_{k=0}^{\infty} \sum_{j=0}^{k} \left(\begin{array}{c} k \\ j \end{array} \right) \frac{(-1)^{2k+1-j}}{2^{k+1}} \exp\{\bar{\omega}_{s_t} j + j^2 \gamma/2\} \left(\frac{1+\rho_{t-1}}{1-\rho_{t-1}} \right)^{j\lambda}, \quad (7)$$

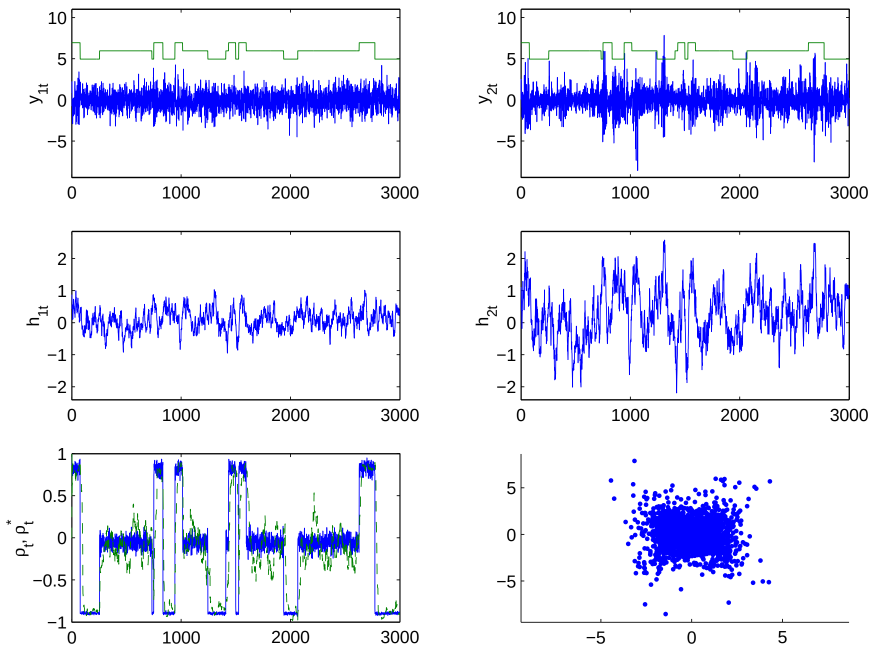
which is nonlinear in the lagged value of the correlation process. A proof is in Amisano and Casarin (2007).

□

Example 2 - (Switching Gamma Autoregressive) A class of model which could replace the Gaussian autoregressive is the autoregressive Gamma process due to Gourieroux and Jasiak (2006). We propose a Markov-switching Autoregressive Gamma of the first order (MS-ARG(1)), with transition dynamics

$$\omega_{t+1} \sim \mathcal{G}a\left(\delta, \alpha_{s_{t+1}} \omega_t, \beta_{s_{t+1}}\right) \quad (8)$$

Figure 2: *Simulated bivariate stochastic volatility model with Markov-switching stochastic correlation. Top left and right simulated observable process assuming* $A = diag\{0.05, 0.05\}$*. In the second line simulated stochastic log-volatility process assuming* $\mathbf{c} = (0.001, 0.001)$, $B = diag\{0.95, 0.97\}$ *and* $\Xi = diag\{0.1, 0.2\}$*. The bottom-left chart shows the stochastic correlation process (continuous line) with parameters:* $\alpha_1 = 1.8$, $\alpha_2 = 0.111$, $\alpha_3 = 0.009$, $\beta = 100$, $\gamma_1 = 100$, $\gamma_2 = 11.11$, $\gamma_1 = 5.747$ *and the recursive estimation of the correlation (dashed line) with a moving window of 30 observations. The bottom-right chart shows the scatter plot of the observations.*



with $\delta, \alpha_k, \beta_k \in \mathbb{R}\ \forall k$ and where $\mathcal{G}a(\delta, \alpha, \beta)$ denotes the noncentered gamma distribution. In order to obtain a stochastic correlation process we can transform ω_t as follow: $\rho_t = 1 - \exp(-\omega_t)$. The first conditional of the correlation process is

$$\mathbb{E}(\rho_{t+1}|\rho_t, s_{t+1}) = 1 - \left((1 + \beta_{s_{t+1}})^{-\delta}\left(\rho_t^{\alpha_{s_{t+1}}\beta_{s_{t+1}}/(1+\beta_{s_{t+1}})} - 1\right)\right) \tag{9}$$

that follows by conditioning on s_{t+1} and applying Proposition 4..2 in Appendix B.

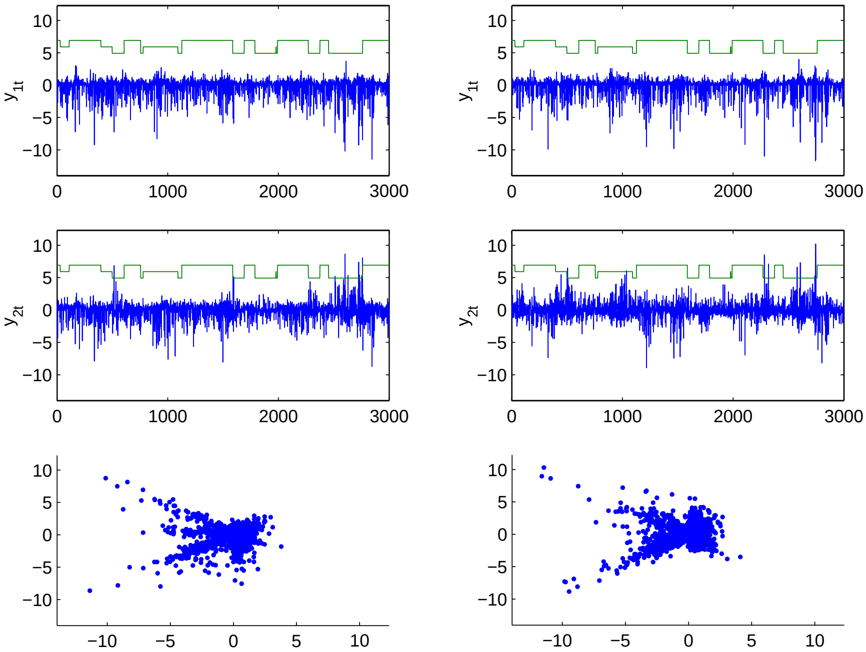
□

Example 3 - (Switching Beta Autoregressive)
As alternative to the Gaussian and Gamma autoregressive processes we could use a stochastic process naturally defined on $[0, 1]$. We propose a Markov-switching Beta Autoregressive Process of the first order (MS-BAR(1)), with transition dynamics

$$\omega_{t+1} \sim \mathcal{B}e\left(\alpha_{s_{t+1}} + \gamma_{s_{t+1}}\omega_t, \beta_{s_{t+1}} + \gamma_{s_{t+1}}(1 - \omega_t)\right), \tag{10}$$

with $\alpha_k, \beta_k, \gamma_k \in \mathbb{R}\ \forall k$ and where $\mathcal{B}e(a, b)$ denotes a beta distribution of the Type I with $a, b > 0$. For $K = 1$ we obtain the time-homogeneous BAR(1) process given in Definition 4..2, Appendix C. See also Amisano and Casarin (2007) for further details.

Figure 3: *Simulated bivariate stochastic volatility model with Markov-switching stochastic correlation and skewed Studentt observations. Left column shows the observations and the scatter plot for a skewed Student-t with $\xi = (0.5, 0.5)'$ and $\nu = (30, 80)'$, the right column for a skewed Student-t with $\xi = (0.5, 1.4)'$ and $\nu = (23, 23)'$.*



Following our definition of BAR(1) process and applying Proposition 4..4 in Appendix C, the first two conditional moments are easily found

$$\mathbb{E}(\omega_{t+1}|\omega_t, s_{t+1}) = \frac{\alpha_{s_{t+1}} + \gamma_{s_{t+1}}\omega_t}{\alpha_{s_{t+1}} + \beta_{s_{t+1}} + \gamma_{s_{t+1}}} \tag{11}$$

$$\mathbb{V}ar(\omega_{t+1}|\omega_t, s_{t+1}) = \frac{\alpha_{s_{t+1}} + \gamma_{s_{t+1}}\omega_t}{\alpha_{s_{t+1}} + \beta_{s_{t+1}} + \gamma_{s_{t+1}} + 1} \frac{\beta_{s_{t+1}} + \gamma_{s_{t+1}}(1-\omega_t)}{(\alpha_{s_{t+1}} + \beta_{s_{t+1}} + \gamma_{s_{t+1}})^2}. \tag{12}$$

The BAR(1) process has a nonlinear dynamics, but the conditional mean of the process depends linearly on the lagged value of the process. This makes easier the interpretation of the parameter of process. Moreover the BAR(1) shares some properties with the Gamma processes and can capture complex nonlinear dynamics. Its conditional heteroscedasticity is a quadratic function of the lagged value and its over- and underdispersion is time-varying (see Appendix C). Finally note that if the stationarity conditions given in Appendix C are satisfied, the three parameter of the process determine the serial dependence: corr$(\omega_t, \omega_{t-r}) = (\gamma(\gamma+\beta+\alpha)^{-1})^r$.
We apply the linear transform $\phi(x) = 1 - 2x$ in order to obtain a process $\rho_t = \phi(\omega_t)$ with values in $[-1, +1]$. Note however that the class of the BAR(1) processes is closed under linear transform. Thus all the properties discussed above apply to the ρ_t. A simulated example of MS-BAR(1) is in Fig. 2.

□

Example 4 - (Heavy-Tails Stochastic-Correlation Models)

As alternative to the Gaussian VAR process we could use a process which account also for skewness and kurtosis. We propose a skewed Student-t model

$$\mathbf{y}_t = A_t\mathbf{y}_{t-1} + \Sigma_t^{1/2}\varepsilon_t, \qquad \varepsilon_t \sim \mathcal{S}k\mathcal{T}_2\left(\mathbf{0}, Id_2, \nu, \xi\right) \tag{13}$$

$$\mathbf{h}_{t+1} = \mathbf{c} + B\mathbf{h}_t + \Xi^{1/2}\eta_t, \qquad \eta_t \sim \mathcal{N}_2\left(\mathbf{0}, Id_2\right) \tag{14}$$

where $\mathcal{S}k\mathcal{T}_n\left(\mathbf{0}, Id_n, \nu, \xi\right)$ is the n-variate skewed Student-t with $\nu = (\nu_1, \ldots, \nu_n)'$ degrees of freedom vector and skewness parameter $\xi = (\xi_1, \ldots, \xi_n)'$. For a definition of multivariate skewed Student-t distribution see Azzalini and Dalla Valle (1996), Branco and Dey (2001) and Sahu *et al.* (2003). In this work we adopt the definition due to Ferreira and Steel (2003) (but see also Bauwens and Laurent (2002) for a quite similar definition). Their constructive method for skewed Student-t makes the simulation from this distribution simple. The existence of the moments is guaranteed by the existence of the moments of the underlying univariate distributions. Finally the resulting multivariate distribution accounts for heterogenous components, i.e. it allows for different magnitudes and directions of kurtosis and skewness. Two simulated examples of skewed Student-t observations are given in Fig. 3.

□

3. Bayesian Inference

3.1 Estimation of Latent Variables and Parameters

In the following we deal with the inference problems in the nonlinear dynamic models presented in previous sections. We follow the nonlinear filtering approach. As suggested by Berzuini *et al.* (1997) we include the parameters into the state vector and then, following Liu and West (2001), we apply a *Regularised Auxiliary Particle Filter* (R-APF) for filtering the hidden states and estimating the unknown parameters of the model.
We assume that the Bayesian nonlinear model is represented in a distributional state-space form, that is defined by the following measurement, transition and initial densities

$$\mathbf{y}_t \sim p(\mathbf{y}_t|\mathbf{x}_t, \boldsymbol{\theta}_t) \tag{15}$$

$$(\mathbf{x}_t, \boldsymbol{\theta}_t) \sim p(\mathbf{x}_t, \boldsymbol{\theta}_t|\mathbf{x}_{t-1}, \boldsymbol{\theta}_{t-1}) \tag{16}$$

$$(\mathbf{x}_0, \boldsymbol{\theta}_0) \sim p(\mathbf{x}_0|\boldsymbol{\theta}_0)p(\boldsymbol{\theta}_0) \tag{17}$$

with $t = 1, 2, \ldots, T$. In this general and possibly nonlinear model, $\mathbf{y}_t \in \mathcal{Y} \subset \mathbb{R}^{n_y}$ represents the observable variable, $\mathcal{Y}$ the observations space, $\mathbf{x}_t \in \mathcal{X} \subset \mathbb{R}^{n_x}$ the hidden state (i.e. the latent variable) and $\mathcal{X}$ the state space.
We assume the transition density of the parameter vector is trivially $\delta_{\boldsymbol{\theta}_{t-1}}(\boldsymbol{\theta}_t)$, where $\delta_x(y)$ denotes the Dirac's mass centered in x. The last line shows the prior distribution on the parameter vector $\boldsymbol{\theta}_0 = \boldsymbol{\theta}$, with $\boldsymbol{\theta} \in \Theta \subset \mathbb{R}^{n_{\boldsymbol{\theta}}}$. Note that the prior distribution on the parameter represents the Bayesian part of the model.
In the case of the three-regime models given in the previous examples, $\mathbf{y}_t \in \mathbb{R}^2$ and $\mathbf{x}_t = (\mathbf{h}_t', \omega_t, s_t)' \in (\mathbb{R}^2 \times \mathbb{R} \times \{1, 2, 3\})$. The parameter vectors in examples 1 and 3 are $\boldsymbol{\theta} = ((a_{11}, a_{22})'$, $\mathbf{c}'$, $(b_{11}, b_{22})'$, vec$\{\Xi\}'$, λ, γ, $\bar{\omega}_1$, $\bar{\omega}_2$, $\bar{\omega}_3)'$ and $\boldsymbol{\theta} = ((a_{11}, a_{22})'$, $\mathbf{c}'$, $(b_{11}, b_{22})'$, vec$\{\Xi\}'$, λ, α_1, α_2, α_3, β, γ_1, γ_2, $\gamma_3)'$ respectively. We denote with *vec* the matrix operator, which stacks into a vector the columns of a given matrix.

Let us define $\mathbf{z}_t = (\mathbf{x}'_t, \boldsymbol{\theta}'_t)'$, $\mathcal{Z} = \mathcal{X} \times \Theta$ and $\mathbf{z}_{s:t} = (\mathbf{z}_s, \mathbf{z}_{s+1}, \dots, \mathbf{z}_t)$. The optimal prediction, filtering and smoothing densities, for the model in Eq. (15), (16) and (17), are

$$p(\mathbf{z}_{t+1}|\mathbf{y}_{1:t}) = \int_{\mathcal{Z}} p(\mathbf{x}_{t+1}|\mathbf{x}_t, \boldsymbol{\theta}_{t+1})\delta_{\boldsymbol{\theta}_t}(\boldsymbol{\theta}_{t+1})p(\mathbf{z}_t|\mathbf{y}_{1:t})d\mathbf{z}_t \tag{18}$$

$$p(\mathbf{y}_{t+1}|\mathbf{y}_{1:t}) = \int_{\mathcal{Z}} p(\mathbf{y}_{t+1}|\mathbf{z}_{t+1})p(\mathbf{z}_t|\mathbf{y}_{1:t})d\mathbf{z}_{t+1} \tag{19}$$

$$p(\mathbf{z}_{t+1}|\mathbf{y}_{1:t+1}) = \frac{p(\mathbf{y}_{t+1}|\mathbf{x}_{t+1}, \boldsymbol{\theta}_{t+1})p(\mathbf{x}_{t+1}|\mathbf{x}_t, \boldsymbol{\theta}_{t+1})\delta_{\boldsymbol{\theta}_t}(\boldsymbol{\theta}_{t+1})}{p(\mathbf{y}_{t+1}|\mathbf{y}_{1:t})} \tag{20}$$

$$p(\mathbf{z}_s|\mathbf{y}_{1:t}) = p(\mathbf{z}_s|\mathbf{y}_{1:s}) \int_{\mathcal{Z}} \frac{(\mathbf{z}_{s+1}|\mathbf{z}_s)p(\mathbf{z}_{s+1}|\mathbf{y}_{1:t})}{p(\mathbf{z}_{s+1}|\mathbf{y}_{1:t})}d\mathbf{z}_{s+1} \tag{21}$$

with $s < t$.
The analytical solution of the general stochastic filtering problem described in Eq. (18)-(21) is known in only few cases. Inference problems in nonlinear and/or non-Gaussian state-space models are usually solved by introducing some approximations. In this work we bring into action *Particle Filters* (PF) (Doucet *et al.* (2001)). In particular we apply the regularised particle filters due to Liu and West (2001) and Musso *et al.* (2001).
Let $(\mathbf{z}_0^i, w_0^i)$ be a weighted sample from the prior distribution given in Eq. (17). The sample is also called particle and the collection of N samples, $\{\mathbf{z}_0^i, w_0^i\}_{i=1}^N$ is called particle set. Assume that at time t a particle set $\{\mathbf{z}_t^i, w_t^i\}_{i=1}^N$ is approximating the density in Eq. (20), then the density in Eq. (18) is approximated

$$p_N(\mathbf{z}_{t+1}|\mathbf{y}_{1:t}) \propto \sum_{i=1}^N \frac{1}{N|H_t^i|^d} w_t^i p(\mathbf{z}_{t+1}|\mathbf{z}_t^i) K_{H_t^i}(\mathbf{z}_{t+1} - \mathbf{z}_t^i) \tag{22}$$

$$p_N(\mathbf{y}_{t+1}|\mathbf{y}_{1:t}) \propto \sum_{i=1}^N \frac{1}{N|H_t^i|^d} w_t^i p(\mathbf{y}_{t+1}|\mathbf{x}_{t+1}, \boldsymbol{\theta}_{t+1}) K_{H_{it}}(\mathbf{z}_{t+1} - \mathbf{z}_t^i) \tag{23}$$

$$p_N(\mathbf{z}_{t+1}|\mathbf{y}_{1:t+1}) \propto \sum_{i=1}^N \frac{1}{N|H_t^i|^d} w_t^i p(\mathbf{y}_{t+1}|\mathbf{z}_{t+1}) \tag{24}$$

where $K_H(\mathbf{x}) = K(H^{-1}\mathbf{x})$ is a multivariate Gaussian kernel and H a p.d. scale matrix (*bandwidth*). $|H|$ denotes the determinant of H. Note that with respect to Liu and West (2001), we consider a matrix-variate bandwidths to allow for different jittering dimension in each direction of the augmented state space. See Amisano and Casarin (2007) for details.

3.2 Number of Regimes

Let L_t represent the current number of regimes. Following Chopin Chopin (2001) and Chopin and Pelgrin Chopin and Pelgrin (2004), the state vector is augmented with the auxiliary variable, m_{t+1} defined as follows

$$m_{t+1} = \max\{L_{t+1}, m_t\} \tag{25}$$

This augmented state allow us to estimate sequentially the number of regimes. See Amisano and Casarin (2007) for further details.

4. Conclusion

We introduce some new Markov-switching stochastic-correlation models. For the inference, a simulation-based Bayesian approach is considered, which relies upon sequential Monte Carlo algorithms. SMC is particularly suitable for on-line model selection and joint estimation of the latent variables and the parameters. The on-line context allows us to evaluate sequentially also risk and the performance of the optimal portfolio when the stochastic correlation is governed by a Markov-switching process.

References

Aguilar O. and West M. (2000) Bayesian Dynamic Factor Models and Portfolio Allocation, *Journal of Business and Economic Statistics*, 18, 338–357.

Amisano G. and Casarin R. (2007) Particle filters for markov-switching stochastic-correlation models.

Arulampalam S., Maskell S., Gordon N. and Clapp T. (2001) A Tutorial on Particle Filters for On-line Nonlinear/Non-Gaussian Bayesian Tracking.

Asai M. and McAleer M. (2005) Dynamic Correlations in Symmetric Multivariate SV Models, in: *MODSIM 2005 International Congress on Modelling and Simulation. Modelling and Simulation Society of Australia and New Zealand*, Zerger A. and Argent R., eds.

Asai M. and McAleer M. (2006a) Multivariate Stochastic Volatility Models: Bayesian Estimation and Model Comparison, *Econometric Reviews*, 25, 361–384.

Asai M. and McAleer M. (2006b) The Structure of Dynamic Correlations in Multivariate Stochastic Volatility Models, Econometric Reviews, forthcoming.

Asai M., McAleer M. and Yu J. (2006) Multivariate Stochastic Volatility: A Review, *Econometric Reviews*, 25, 145–175.

Azzalini A. and Dalla Valle A. (1996) The multivariate skew-normal distribution, *Biometrika*, , 83, 715–726.

Bauwens L. and Laurent S. (2002) A new class of multivariate skew densities, with application to GARCH models, technical report, CORE.

Bauwens L., Laurent S. and Rombouts J. (2006) Multivariate GARCH models: a survey, *Journal of Applied Econometrics*, 21, 1, 79–109.

Berzuini C., Best N.G., Gilks W.R. and Larizza C. (1997) Dynamic conditional independence models and Markov chain Monte Carlo methods, *Journal of the American Statistical Association*, 92, 440, 1403–1441.

Branco M. and Dey D.K. (2001) A general class of multivariate skew elliptical distributions, *J. Multivariate Anal.*, , 79, 99–113.

Chib S., Nardari F. and Shephard N. (2006) Analysis of High-Dimensional Multivariate Stochastic Volatility Models, *Journal of Econometrics*, 134, 341–371.

Chopin N. (2001) Sequential inference and state number determination for discrete state-space models through particle filtering, technical Report N. 34, CREST.

Chopin N. and Pelgrin F. (2004) Bayesian inference and state number determination for hidden Markov models: an application to the information content of the yield curve about inflation, *Journal of Econometrics*, 123, 2, 327–344.

Doucet A., Freitas J.G. and Gordon J. (2001) *Sequential Monte Carlo Methods in Practice*, Springer Verlag, New York.

Engle R. (2002) Dynamic Conditional Correlation: A Simple Class of Multivariate Generalized Autoregressive Conditional Heteroskedasticity Models, *Journal of Business and Economic Statistics*, 20, 3, 339–350.

Ferreira J.T.A.S. and Steel M.F.J. (2003) Bayesian Multivariate Regression Analysis with a New Class of Skewed Distributions, statistics Research Report 419, University of Warwick.

Gourieroux C. and Jasiak J. (2006) Autoregressive Gamma Processes, *Journal of Forecasting*, , 25, 129–152.

Gourieroux C., Jasiak J. and Sufana R. (2004) The Wishart Autoregressive Process of Multivariate Stochastic Volatility, working Paper, CREST.

Harvey A., Ruiz E. and Shephard N. (1994) Multivariate Stochastic Volariance Models, *Review of Economic Studies*, 61, 247–264.

Jacquier E., Polson N.G. and Rossi P. (1994) Bayesian Analysis of Stochastic Volatility Models, *Journal of Business and Economic Statistics*, 12, 281–300.

Julier S.J. and Uhlmann J.K. (1996) A general method for approximating nonlinear transformations of probability distributions, technical Report RRG, Department of Engineering Science, University of Oxford.

Liu J. and West M. (2001) Combined Parameter and State Estimation in Simulation Based Filtering, in: *Sequential Monte Carlo Methods in Practice*, Doucet A. F.J. and J. G., eds., Springer-Verlag, New York.

Mena R.H. and Walker S.G. (2004) Stationary Autoregressive Models via a Bayesian Nonparametric Approach, *Journal of Time Series Analysis*, 26, 789–805.

Musso C., Oudjane N. and LeGland F. (2001) Improving Regularised Particle Filters, in: *Sequential Monte Carlo in Practice*, Doucet A., Freitas J. and Gordon J., eds., Springer Verlag, New York.

Nieto-Barajas L.E. and Walker S.G. (2002) Markov beta and gamma processes for modelling hazard rates, *Scandinavian Journal of Statistics*, 29, 413–424.

Philipov A. and Glickman M.E. (2006) Multivariate Stochatic Volatility via Wishart Processes, *Journal of Economics and Business Statistics*, 24, 313–328.

Pitt M. and Shephard N. (1999) Filtering via Simulation: Auxiliary Particle Filters, *Journal of the American Statistical Association*, 94, 446, 590–599.

Pitt M.K., Chatfield C. and Walker S.G. (2002) Constructing First Order Stationary Autoregressive Models via Latent Processes, *Scandinavian Journal of Statistics*, 29, 657–663.

Polson N.G., Stroud J.R. and Müller P. (2002) Practical Filtering with Sequential Parameter Learning, tech. report, Graduate School of Business, University of Chicago.

Polson N.G., Stroud J.R. and Müller P. (2003) Practical Filtering for Stochastic Volatility Models, tech. report, Graduate School of Business, University of Chicago.

Sahu S., Dey D.K. and Branco D. (2003) A new class of multivariate skew distributions with applications to Bayesian regression models, *Canadian Journal of Statistics*, 31, 129–150.

Taylor S.J. (1986) *Modelling Financial Time Series*, Chichester, Wiley.

Taylor S.J. (1994) Modelling Stochastic Volatility, *Mathematical Finance*, 4, 183–204.

Appendix A - Gaussian AR(1)

We find the conditional and unconditional first order moments for a transformed AR(1) process, when applying the Fisher's transform.

Proposition 4..1. *Let ω_t follow the AR(1) process*

$$\omega_t = \bar{\omega} + \lambda\omega_{t-1} + \gamma^{1/2}\psi_t, \quad \psi_t \overset{i.i.d}{\sim} \mathcal{N}(0,1) \tag{26}$$

and $\mathcal{F}_t = \sigma(\{\omega_u\}_{u\leq t})$ be the sigma algebra generated by ω_t. The process $\rho_t = (\exp\{\omega_t\} - 1)(\exp\{\omega_t\}+1)^{-1}$ has the following first order conditional and unconditional moments

1. $\mathbb{E}_{t-1}(\rho_t) = \sum_{k=0}^{\infty}\sum_{j=0}^{k}\begin{pmatrix} k \\ j \end{pmatrix}\frac{(-1)^{2k+1-j}}{2^{k+1}}\exp\{j\bar{\omega} + j^2\gamma/2\}\left(\frac{1+\rho_{t-1}}{1-\rho_{t-1}}\right)^{j\lambda}$;
2. $\mathbb{E}(\rho_t) = \sum_{k=0}^{\infty}\sum_{j=0}^{k}\begin{pmatrix} k \\ j \end{pmatrix}\frac{(-1)^{2k+1-j}}{2^{k+1}}\exp\left\{\frac{j^2\gamma}{2(1-\lambda^2)}\right\}$.

Proof. See Amisano and Casarin (2007). □

Appendix B - ARG(1) Process

Definition 4..1. *The stochastic process ω_t is a time-homogeneous Autoregressive Gamma Process (ARG(1)) with invariant parameters $\alpha, \beta, \delta \in \mathbb{R}$ if and only if*

$$\omega_{t+1} \sim \mathcal{G}a\left(\delta, \alpha\omega_t, \beta\right) \tag{27}$$

where $\mathcal{G}a(\delta, \alpha, \beta)$ is a noncentered gamma distribution.

The transition density of a ARG(1) process is

$$f(y|x) \sim \exp\left(\frac{y}{\beta}\right)\sum_{k=0}^{\infty}\left(\frac{y^{\delta+k-1}}{\beta^{\delta+k}\Gamma(\delta+k)}\frac{\exp(-\alpha^k)}{k!}\right)\mathbb{I}(y)_{(0,\infty)} \tag{28}$$

with $\Gamma(c)$ denoting the gamma function.

Proposition 4..2. *Let ω_t follow a ARG(1) process. The conditional Laplace transform is*

$$\mathbb{E}_{t-1}\left(\exp(-u\omega_t)\right) = (1+\beta u)^{\delta}\exp\left(-\omega_{t-1}\frac{\alpha\beta u}{1+\beta u}\right) \tag{29}$$

Proof. See Gourieroux and Jasiak Gourieroux and Jasiak (2006). □

Appendix C - BAR(1) Process

B.1 - BAR(1) Stationarity conditions

Definition 4..2. *The stochastic process ω_t is a time-homogeneous Beta Autoregressive Process (BAR(1)) with invariant parameters $\alpha, \beta, \gamma \in \mathbb{R}$ if and only if*

$$\omega_{t+1} \sim \mathcal{B}e\left(\alpha + \gamma\omega_t, \beta + \gamma(1-\omega_t)\right) \tag{30}$$

where $\mathcal{B}e(c, d)$ is a central beta distribution of the first Type with parameters $c, d > 0$.

The transition density of a BAR(1) process is the central Beta density

$$f(y|x) \sim \frac{1}{B(\alpha+\gamma x, \beta+\gamma(1-x))} y^{\alpha+\gamma x-1}(1-y)^{\beta+\gamma(1-x)-1}\mathbb{I}_{[0,1]}(y), \quad (31)$$

with $B(c, d)$ denoting the Beta function defined as $B(c, d) = \Gamma(c)\Gamma(d)/\Gamma(c + d)$.
Let $\mathcal{F}_t = \sigma(\{\omega_s\}_{s\leq t})$ be the σ-algebra generated by ω_t and denote with $\mathbb{E}_t(\cdot) = \mathbb{E}(\cdot|\mathcal{F}_t)$ the conditional expectation operator. The conditional and unconditional moments and the serial dependence of the BAR(1) are given in the following.

Proposition 4..3. *Let ω_t follow a BAR(1) process. The first two conditional and unconditional noncentral moments of a BAR(1) are*

1. $\mathbb{E}_{t-1}(\omega_t) = (\alpha + \gamma\omega_{t-1})(\alpha + \beta + \gamma)^{-1}$;
2. $\mathbb{E}_{t-1}(\omega_t^2) = \frac{(\alpha+\gamma\omega_{t-1})(\alpha+\gamma\omega_{t-1}+1)}{(\alpha+\beta+\gamma)(\alpha+\beta+\gamma+1)}$;
3. $\mathbb{E}(\omega_t) = \alpha(\alpha + \beta)^{-1}$;
4. $\mathbb{E}(\omega_t^2) = \frac{(\alpha^2+\alpha)(\alpha+\beta)+\alpha\gamma(2\alpha+1)}{(\alpha+\beta)((\alpha+\beta+\gamma)(\alpha+\beta+\gamma+1)-\gamma^2)}$.

Proof. See Amisano and Casarin (2007). □

Proposition 4..4. *Let ω_t follow a BAR(1) process. The autocorrelation of order r, $\rho(r)$ of a BAR(1) is $\rho(r) = (\gamma/(\alpha + \beta + \gamma))^r$.*

Proof. See Amisano and Casarin (2007). □

B.2 - Overdispersion properties of the BAR(1) process

Another feature of the proposed BAR(1) process is that it can be both over- or underdispersed as showed in the following propositions.

Proposition 4..5. *The first two conditional moments of a stationary BAR(1) process are*

$$\mathbb{E}_{t-1}(\omega_t) = \frac{(\alpha+\gamma\omega_{t-1})}{(\alpha+\beta+\gamma)} \quad (32)$$

$$\mathbb{V}_{t-1}(\omega_t) = \frac{(\alpha+\gamma\omega_{t-1})}{(\alpha+\beta+\gamma+1)}\frac{(\beta+\gamma(1-\omega_{t-1}))}{(\alpha+\beta+\gamma)^2}. \quad (33)$$

Proposition 4..6. *For a stationary homogeneous BAR(1) with parameters $\alpha, \beta > 0$, $\gamma < 0$ there exists conditional overdispersion, $\mathbb{V}_{t-1}(\omega_t) > (\mathbb{E}_{t-1}(\omega_t))^2$, if and only if $\omega_{t-1} \in (\omega^{(1)}, \omega^{(2)})$ for $(\gamma > -\beta-\alpha)$, where $\omega^{(1)} = (\gamma(2+\alpha+\beta+\gamma))^{-1}(\beta + \gamma - \alpha(1 + \beta + \gamma + \alpha))$ and $\omega^{(2)} = -\alpha/\gamma$.*

Proof. See Amisano and Casarin (2007). □