MODELLING MULTIVARIATE VOLATILITY PROCESSES USING TEMPORAL INDEPENDENT COMPONENT ANALYSIS

Paola Palmitesta

Department of Quantitative Methods, University of Siena (Italy)

Corrado Provasi

Department of Statistical Sciences, University of Padua (Italy)

Abstract

Forecasting temporal dependence in second order moments of returns is a relevant problem in many contexts of financial econometrics. It is commonly accepted that financial volatilities move together over time across assets and markets. For this reason in this paper we propose an approach based on the analysis of independent temporal components to model the multivariate volatility. We have assumed that the underlying factors or sources of the model are AR-APARCH processes with errors interpreted by the Meixner distribution. An application with two sets of real data shows the use of the model in the analysis of parallel financial series.

Keywords: independent component analysis, APARCH process, maximum likelihood estimation, Meixner distribution.

1. INTRODUCTION

Forecasting temporal dependence in second order moments of returns is a relevant problem in many contexts of financial econometrics. It is commonly accepted that financial volatilities move together over time across assets and markets. For this reason, from the seminal papers of Engle (1982) and Bollerslev (1986), the generalized autoregressive conditional models (GARCH) have been extended to the multivariate case (MGARCH) (see Bollerslev, Engle, and Wooldridge, 1988). Nevertheless, the implementation of multivariate GARCH models in more than a few dimensions is extremely difficult: since the model has many parameters, the likelihood function becomes very flat, and consequently the optimization of the likelihood becomes practically impossible. There is simply no way that full multivariate GARCH models can be used to estimate directly the very large covariance matrices that are required to net all the risks in a large trading book.

To compensate for this problem, different parametrizations of the original model have been proposed (for a review see Bawens, Laurent, and Rombouts, 2006); in particular, Alexander and Chibumba (1996) (see also Alexander, 2001) proposed the principal component GARCH or "orthogonal GARCH" (O-GARCH) model for generating large GARCH covariance matrices (for a generalization of the model see van der Weide, 2002). The O-GARCH model is an accurate and efficient method for generating large covariance matrices that only requires the estimation of univariate GARCH models. Hence, it has many practical advantages, for example in value-at-risk models. It works best in highly correlated systems, such as term structures. A limit of this approach in the analysis of multivariate structures of financial data is the fact that it accounts only for the first and second order moments of the multivariate return distribution, whereas it is known that this distribution is characterized to have marginals with heavy tails which, consequently, have a high kurtosis. In this direction Chen, Härdle, and Spokoiny (2007) have proposed and analyzed a technology that is based on performing an independent component search instead of an analysis of the principal components and adaptively fitting the resulting independent marginals by generalized hyperbolic distributions.

In this paper we expand the methodology proposed by Chen, Härdle, and Spokoiny (2007) using a model that integrates the three properties commonly used for the analysis of independent components: nongaussianity, distinct autocorrelation and a smoothly changing stationary variance (Hyvärinen, 2005). The model can adopt different distributions as independent components, as the distributions considered, for example, by Palmitesta and Provasi (2007). Nevertheless, Chen, Härdle, and Spokoiny (2007) have shown that a subset of the generalized hyperbolic distributions, the Normal Inverse Gaussian distribution (NIG), has particular characteristics to be useful to interpret financial returns. Since the derivatives of the NIG log-likelihood involve the Bessel function, direct likelihood maximization for this distribution is difficult. As a further complication, the NIG kurtosis parameter is constrained to be larger than the absolute value of the skewness parameter. As a consequence, maximization algorithms suffer from problems like non-convergence and the need for good initial values. This is especially true when the NIG is used within GARCH-type models, and many parameters need to be estimated simultaneously. As an alternative approach, we have considered the adoption of the Meixner distribution, which has properties similar to those of the NIG. In fact, it has semi-heavy tails and is infinitely divisible. However, the derivatives of the Meixner log-likelihood are more analytically tractable, and the Meixner joint parameter space is the Cartesian product of each parameter space; likelihood maximization is therefore much easier (Grigoletto and Provasi, 2007).

The paper is organized as follows. In the next Section we discuss the MGARCH models, while in the following Section we introduce the independent component analysis, assuming Meixner sources and temporal dynamics rendered by GARCH-type models; therefore, in Section 4 we cope with the problem maximum likelihood estimate of the temporal model with independent components. Then, in Section 5 we show the results of two empirical experiments. Concluding remarks are discussed in Section 6.

2. MGARCH MODELS

Let $\{x_t\}_{t=1}^{T}$ be a vector of k return series with

$$\mathrm{E}\left(\boldsymbol{x}_t \left|\boldsymbol{F}_{t-1}\right.\right) \equiv \mu_t, \quad \mathrm{E}\left(\boldsymbol{x}_t \boldsymbol{x}_t^{'} \left|\boldsymbol{F}_{t-1}\right.\right) \equiv \boldsymbol{\Sigma}_t \equiv \left(\sigma_{t,ij}\right),$$

where F_{t-1} is the σ-algebra generated by $\left\{\boldsymbol{x}_t, \boldsymbol{x}_{t-1}, \ldots\right\}$ and $\boldsymbol{\Sigma}_t$ is an F_{t-1}-measurable $k \times k$ semi-positive definite matrix . Without loss of generality, we assume that the mean μ_t is zero. Many specifications have been proposed for $\boldsymbol{\Sigma}_t$.

The BEKK representation of the multivariate GARCH(p,q) model by Engle and Kroner (1995) is

$$\boldsymbol{\Sigma}_t = \boldsymbol{C} + \sum_{i=1}^{p}\sum_{j=1}^{M} \boldsymbol{A}_{ij}\left(\boldsymbol{x}_{t-i}\boldsymbol{x}_{t-1}^{'}\right)\boldsymbol{A}_{ij}^{'} + \sum_{i=1}^{q}\sum_{j=1}^{M} \boldsymbol{B}_{ij}\boldsymbol{\Sigma}_{t-1}\boldsymbol{B}_{ij}^{'},$$

where $\boldsymbol{C}$, $\boldsymbol{A}_{ij}$, $\boldsymbol{B}_{ij}$, are matrices $k \times k$ and $\boldsymbol{C}$ is positive definite. The factor GARCH model can be viewed as a special case of this model (e.g. Lanne and Saikkonen, 2007). Also if the shape of the BEKK model is very general, above all when M is reasonably big, it suffers problems related to overparametrization.

In order to go over estimation problems due to the overparametrization, a dynamic correlation model has been proposed (Engle, 2002; Engle and Sheppard, 2001).It is based on the decomposition

$$\boldsymbol{\Sigma}_t = \boldsymbol{D}_t^{1/2}\boldsymbol{R}_t\boldsymbol{D}_t^{1/2},$$

where $\boldsymbol{D}_t = \mathrm{diag}\left(\sigma_{t,11}^2, \ldots, \sigma_{t,kk}^2\right)$, $\sigma_{t,ii}^2$ is the conditional variance of the i-th component of $\boldsymbol{x}_t$ and $\boldsymbol{R}_t = \left(\rho_{t,ij}\right)$ is the correlation matrix. Bollerslev (1990) has proposed a constant correlation model assuming $\boldsymbol{R}_t = \boldsymbol{R}$.

The O-GARCH(p,q) model assumes that the $\boldsymbol{\Sigma}_t$ matrix is decomposed as

$$\boldsymbol{\Sigma}_t = \boldsymbol{B}'\boldsymbol{D}_t\boldsymbol{B}, \tag{1}$$

where $\boldsymbol{B}$ is an orthogonal positive definite matrix. $\boldsymbol{B}$, in general, is constituted by the eigenvectors of the variance-covariance sampling matrix $\boldsymbol{S}_X$ and it satisfies the condition

$$\boldsymbol{B}'\boldsymbol{S}_X\boldsymbol{B} = \boldsymbol{\Lambda},$$

with Λ being the matrix of eigenvalues. Note that in the O-GARCH setup the factor loadings $\boldsymbol{B}$ are assumed constant so its application is limited to short time periods. Other specifications of $\boldsymbol{\Sigma}_t$ can be found in Bawens, Laurent, and Rombouts (2006).

As an alternative to the (1), in the next Section we show a possible specification of $\boldsymbol{\Sigma}_t$ based on the independent component analysis.

3. INDEPENDENT COMPONENT ANALYSIS

In independent component analysis (ICA), the observed variable $\boldsymbol{x}_t \in \mathbb{R}^k$ is modelled as

$$\boldsymbol{x}_t = \boldsymbol{V}\boldsymbol{y}_t \tag{2}$$

where the underlying factors or sources, $\boldsymbol{y}_t \in \mathbb{R}^k$, are statistically independent and $\boldsymbol{V}$ is the invertible $k\times k$ *mixing* matrix. ICA can be seen also as a linear transformation to a new set of variables

$$\boldsymbol{y}_t = \boldsymbol{W}\boldsymbol{x}_t,$$

whose components, y_{ti}, are statistically independent. The *demixing* matrix $\boldsymbol{W}$ is the inverse of the $\boldsymbol{V}$ matrix.

ICA can be placed in a maximum likelihood framework. In fact, the likelihood functions of the observations can be obtained by the likelihood of the source densities because multivariate probability density function (pdf) are transformed as follows:

$$\boldsymbol{f}_x\left(\boldsymbol{x}_t\right) = \frac{\boldsymbol{f}_y\left(\boldsymbol{y}_t\right)}{\left|\boldsymbol{J}\right|},$$

where $\boldsymbol{f}$ indicates the pdf and $\boldsymbol{J}$ is the Jacobian of the transformation. Since in ICA the transformation is linear and we have assumed independent sources, we have

$$|\boldsymbol{J}| = \frac{1}{|\det(\boldsymbol{J})|}$$

and

$$\boldsymbol{f}_{\boldsymbol{y}}(\boldsymbol{y}_{\boldsymbol{t}}) = \prod_{\boldsymbol{i}=1}^{\boldsymbol{k}} \boldsymbol{f}_{y_{\boldsymbol{ti}}}(\boldsymbol{y}_{\boldsymbol{ti}}),$$

where $\boldsymbol{f}_{y_{ti}}$, is the pdf of the *i*th source component. These results motivate us to write the log-likelihood function as a function of $\boldsymbol{W} = (\boldsymbol{w}_1, \ldots, \boldsymbol{w}_{\boldsymbol{k}})'$ and $\boldsymbol{x}_{\boldsymbol{t}}$, giving

$$\log \boldsymbol{f}_{\boldsymbol{x}}(\boldsymbol{x}_{\boldsymbol{t}}) = \log|\det(\boldsymbol{W})| + \sum_{\boldsymbol{i}=1}^{\boldsymbol{k}} \log \boldsymbol{f}_{y_{\boldsymbol{ti}}}(\boldsymbol{w}_{\boldsymbol{i}}'\boldsymbol{x}_{\boldsymbol{t}}). \tag{3}$$

Based $\sigma^2_{\boldsymbol{t},\boldsymbol{ii}}$, on this result, it is immediate to obtain the covariance of $\boldsymbol{x}_{\boldsymbol{t}}$:

$$\boldsymbol{\Sigma}_{\boldsymbol{t},\boldsymbol{x}} = \boldsymbol{V}\boldsymbol{D}_{\boldsymbol{t}}\boldsymbol{V}',$$

where

$$\boldsymbol{D}_{\boldsymbol{t}} = \mathrm{diag}(\sigma^2_{\boldsymbol{t},11}, \ldots, \sigma^2_{\boldsymbol{t},\boldsymbol{kk}}),$$

being $\sigma^2_{\boldsymbol{t},\boldsymbol{ii}}$, the variances of the independent components $\boldsymbol{y}_{ti}$, $\boldsymbol{i}=1, \ldots, K$. We now describe a source density model.

3.1 MEIXNER SOURCES

In order for the ICA model to be identified, we need that at most one of the sources has Gaussian distribution and the moments of the sources exist up to the necessary order (e.g. Hyvärinen, Karhunen, and Oja, 2001). Therefore, the context in which we consider the ICA model needs the specification of the distributions of the independent components in the (2) *super-Gaussian*, that is with kurtosis greater than 3. In this sense, a very well-versed distribution which can be assumed as source for an ICA model is the Meixner distribution, recently studied also in the framework of GARCH-type models by Grigoletto and Provasi (2007) (for an in-depth examination see the bibliography of that paper).

A random variable ***Y*** follows a Meixner distribution with parameter vector (*m*,*a*,*b*,*d*), $a > 0$, $-\pi < b < \pi$, $d > 0$, and $m \in \mathbb{R}$, in symbolic notation Y~MXN (*m*,*a*,*b*,*d*), if its pdf for $y \in \mathbb{R}$ is

$$f_{\mathrm{MXN}}\left(y;m,a,b,d\right) = a^{-1} f_{\mathrm{MXN}}\left(\frac{y-m}{a};0,1,b,d\right),$$

where

$$f_{\mathrm{MXN}}(z;0,1,b,d) = \frac{\left(2\cos\left(\frac{b}{2}\right)\right)^{2d}}{2\pi\Gamma\left(2d\right)} e^{bz} \left|\Gamma\left(d+\iota z\right)\right|^2,$$

where $\iota = \sqrt{-1}$ and Γ indicates the gamma function. Note that *m* and *a* are position and scale parameters, while *b* and *d* control the shape of the density; i.e. if $Y \sim \mathrm{MXN}(m,a,b,d)$, then $Z = (Y - m)/a$ is the standardized version of Y~MXN (0,1,*b*,*d*) .

The characteristic function of *Y* is given by

$$\phi_{\mathrm{MXN}}\left(u\right) = e^{\iota mu}\left(\frac{2\cos\left(\frac{b}{2}\right)}{\cosh\left(\frac{au-ib}{2}\right)}\right)^{2d},$$

from which we deduce that the Meixner distribution is infinitely divisible. By means of the usual relationships which permit to derive the moments from the origin of the characteristic function, it is immediate to obtain the mean, the variance and skewness and kurtosis indices of MXN(*m*,*a*,*b*,*d*), which are given by, respectively,

$$\begin{aligned}
\mathrm{E}\left[Y\right] &= {}_{\mathrm{MXN}}\mu = m + ad\tan\left(\tfrac{b}{2}\right),\\
\mathrm{Var}\left[Y\right] &= {}_{\mathrm{MXN}}\sigma^2 = \tfrac{a^2 d}{\cos b + 1},\\
\mathrm{Skew}\left[Y\right] &= {}_{\mathrm{MXN}}\kappa_1 = \sin b\sqrt{\tfrac{1}{d\left(\cos b + 1\right)}},\\
\mathrm{Kurt}\left[Y\right] &= {}_{\mathrm{MXN}}\kappa_2 = 3 - \tfrac{\cos b - 2}{d}.
\end{aligned} \tag{4}$$

Here, Skew[Y] = E[Y–E[Y])^3]/Var(Y)$^{3/2}$ and Kurt[Y] = E[Y–E[Y])^4]/Var(Y)^2. Note that Y is symmetric when b=0 and the kurtosis is always greater than the kurtosis of the Normal distribution.

An important property of the Meixner distribution, that makes it potentially useful in financial applications, is that it has semi-heavy tails (Grigelionis, 2001). Formally, this implies that, for a MXN(m,a,b,d) distribution, we have

$$f_{\mathrm{MXN}}(y;m,a,b,d) \sim C_- |x|^{\rho} e^{-\sigma_-|x|} \qquad x \to -\infty,$$

$$f_{\mathrm{MXN}}(y;m,a,b,d) \sim C_+ |x|^{\rho} e^{-\sigma_+|x|} \qquad x \to +\infty,$$

for some $\rho \in \mathbb{R}$ and $C_-, C_+, \sigma_-, \sigma_+ \geq 0$, with

$$\rho = 2d - 1, \quad \sigma_- = \frac{\pi - b}{a} \quad \text{and} \quad \sigma_+ = \frac{\pi + b}{a}.$$

4. TICA-GARCH MODEL

The underlying temporal structure present in the volatility of financial series induces to hypothesize a temporal model for the ICA (TICA) components, as in Cheung and Xu (2003). In this sense, note that MGARCH models discussed in Section 2 can be also included in the temporal ICA assuming that $\boldsymbol{\Sigma_t} = \boldsymbol{V D_t V}'$ and assuming that the independent components follow a GARCH-type model with source *super-Gaussian*. Then, according to what we stated in the previous Section, we propose to interpret a vector of k return series $\{\boldsymbol{x_t}\}_{\boldsymbol{t}=1}^{\boldsymbol{T}}$ using a TICA model written as

$$\boldsymbol{x_t} = \boldsymbol{V y_t},$$

where the ith component of $\boldsymbol{y_t}$ follows the model

$$\begin{aligned} y_{ti} &= \mu_{ti}(\vartheta_i) + \varepsilon_{ti}, \\ \varepsilon_{ti} &= \sigma_{t,ii}(\vartheta_i) z_{ti}, \end{aligned} \tag{5}$$

where the z_{ti} are iid random variables with zero mean, unit variance, known pdf g_{η_i} with derivative g'_{η_i} and known cdf G_{η_i}, which can also depend from unknown shape parameters η_i; $\mu_{ti} \equiv \mu_{ti}(\vartheta)$; and $\sigma^2_{t,ii} \equiv \sigma^2_{t,ii}(\vartheta_i)$ are, respectively, the conditional mean and variance of y_{ti} based on the set of information F_{t-1} at time t–1; ϑ_i is the vector M×1 formed with all the parameters involved by the two equations, $\vartheta_i \in \mathbb{R}^M$. Note that this context folds many GARCH-type models with conditional heteroskedasticity normally used in literature.

Note that in the TICA-GARCH, as in the O-GARCH, the mixing matrix is assumed constant, so its application is limited to short time periods.

4.1 LOG-LIKELIHOOD FUNCTION

From the (3) it is immediate to infer the average log-likelihood function of the TICA-GARCH model, which is given by

$$\ell(W,\vartheta,\eta) = \log\left|\det(W)\right| - \frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{k}\log\sigma_{t,ii} + \frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{k}\log g_{\eta_i}(z_{ti};\eta_i), \quad (6)$$

where $\vartheta = (\vartheta_1,\ldots,\vartheta_k)'$, $\eta = (\eta_1,\ldots,\eta_k)'$ and $z_{ti} = (w_i' x_t - \mu_{ti})/\sigma_{ti}$ denotes the calculated residuals of the ith component. Therefore, the estimates of the interest parameters can be obtained using an optimization algorithm which maximizes the $\ell(W,\vartheta,\eta)$.

At this point note that the estimates of g_{η_i} in (6) must be very accurate so that $\boldsymbol{W}$ doesn't go to zero or infinity (Hyvärinen, 2005). A manner to avoid this drawback is to prewhiten x_t and standardize y_t. This aim can be obtained using the Mahalanobis transformation $S_X^{-1/2}(x_t - \bar{x})$, where $\bar{x}$ and S_X are, respectively, the sampling mean vector and the variance-covariance matrix. Denote by $\tilde{x}_t$ the prewhitened x_t and $y_t = \tilde{W}\tilde{x}_t$ the corresponding independent components. Then, $W = \tilde{W}S_X^{-1/2}$ is the ICA transformation for the original observations.

4.2 GARCH-TYPE SPECIFICATION

The most elaborated GARCH-type model suggested so far seems to be the model presented in Hentschel (1995) or Duan (1997). However, a GARCH-type model very used in financial applications is the AR(r)-APARCH(p,q) model (Ding, Granger and Engle, 1993; see also He and Teräsvirta, 1999), which is specified in (5) assuming that z_{ti} follows a Meixner distribution conveniently transformed to have zero mean and unit variance, that is with pdf

$$g_{\eta_i}(z_{ti};\eta_i) = \frac{{}_{\mathrm{MXN}}\sigma_i\left(2\cos\left(\frac{b_i}{2}\right)\right)^{2d_i}}{2\pi\Gamma\left(2d_i\right)}\exp\left(b_i\left({}_{\mathrm{MXN}}\mu_i + {}_{\mathrm{MXN}}\sigma_i z_{ti}\right)\right)$$
$$\cdot\left|\Gamma\left(d_i + \iota\left({}_{\mathrm{MXN}}\mu_i + {}_{\mathrm{MXN}}\sigma_i z_{ti}\right)\right)\right|^2,$$

with $\eta_i = \left(b_i, d_i\right)$ and the meaning of symbols already seen previously.

The AR(r)-APARCH(p,q) model specifies for t = 1, …, T the conditional means as

$$\mu_{ti} = \nu_i + \sum_{j=1}^{r}\phi_{ij}y_{t-j,i}$$

and the conditional variances as

$$\sigma_{t,ii}^{\delta_i} = \omega_i + \sum_{j=1}^{p}\alpha_{ij}\left(\left|\varepsilon_{t-j,i}\right| - \gamma_{ij}\varepsilon_{t-j,i}\right)^{\delta_i} + \sum_{j=1}^{q}\beta_{ij}\sigma_{t-j,ii}^{\delta_i},$$

where $\nu_i \in \mathbb{R}$, $-1<\phi_{ij}<1$ (j =1, …, r), $\omega_i > 0$, $\alpha_{ij} \geq 0$ (j =1, …, p), $\beta_{ij} \geq 0$ (j =1, …, q), $\delta_i \geq 0$ and $-1 < \gamma_{ij} < 1$. The most relevant characteristics of the model are the presence of a Box-Cox power transformation of the conditional variances and the asymmetric absolute errors. The Bollerslev's GARCH(p,q) model can be obtained with $\gamma_{ij} = 0$ and $\delta_i = 2$.

We have obtained the first derivatives of the log-likelihood function of the TICA model with temporal components AR(r)-APARCH(1,1) and Meixner sources using *MathStatica* (Rose and Smith, 2002), a package of the computer algebra system *Mathematica* (Wolfram, 1999); then, we have implemented the model in a Fortran 90 program. In order to maximize the log-likelihood function we have used a Newton-type numerical algorithm, appropriate for solving huge optimization problems (Nash, 1984)[1]. In the next Section we show the results obtained with this program in two experiments with real data.

---

[1] *The explicit derivatives of the AR(r)-APARCH(1,1) model with errors distributed as a Meixner can be seen in Grigoletto and Provasi (2007).*

5. EMPIRICAL STUDIES

In this Section we show the application of the model proposed to two sets of real data.

Tab. 1: Summary Statistics of the SCI Data Set.

	S&P 500	CISCO	INTEL
T	1258	1258	1258
Mean	0.0163	0.0279	-0.0351
Stdev	1.0144	2.4224	2.4536
Skewness	0.2394	0.4161	-0.7588
Kurtosis	6.0950	11.3736	11.6593
Min	-4.2423	-11.9982	-20.4794
1st Qu.	-0.5341	-1.1542	-1.2683
Median	0.0558	0.0538	-0.0826
3st Qu.	0.5298	1.2146	1.2383
Max	5.5744	21.8239	10.2419
$Q(10)$	15.1221	13.0035	20.9924
$Q(15)$	27.0528	21.7680	39.1356
$Q(20)$	47.1867	35.0565	64.9559

Note: Q(k) is referred to the Ljung-Box statistics.

The first data set, denoted as SCI, consists of the 1258 daily log returns (in percentage) of the S&P 500 Index, the stock price of Cisco Systems and the stock price of Intel Corp. in the period January, 2, 2002 - December, 29, 2006. The three assets have been considered also by Fan, Wang, and Yao (2005). Figure 1 depicts the time series plot of the three series. Descriptive statistics are listed in Table 1. The correlation matrix among the returns is

$$\hat{R} = \begin{bmatrix} 1.000 & 0.6352 & 0.6714 \\ 0.6352 & 1.000 & 0.6248 \\ 0.6714 & 0.6248 & 1.000 \end{bmatrix}$$

Obviously, the unconditional distribution of the three series exhibits a kurtosis excess, indicating significative departure from normal distributions.

The high values of the Ljung-Box statistics Q suggest some plausible autocorrelation in the series, but this can be due to heteroskedasticity. Then, we have considered for the SCI data set a TICA model with GARCH(1,1) components, as in Fan, Wang, and Yao (2005):

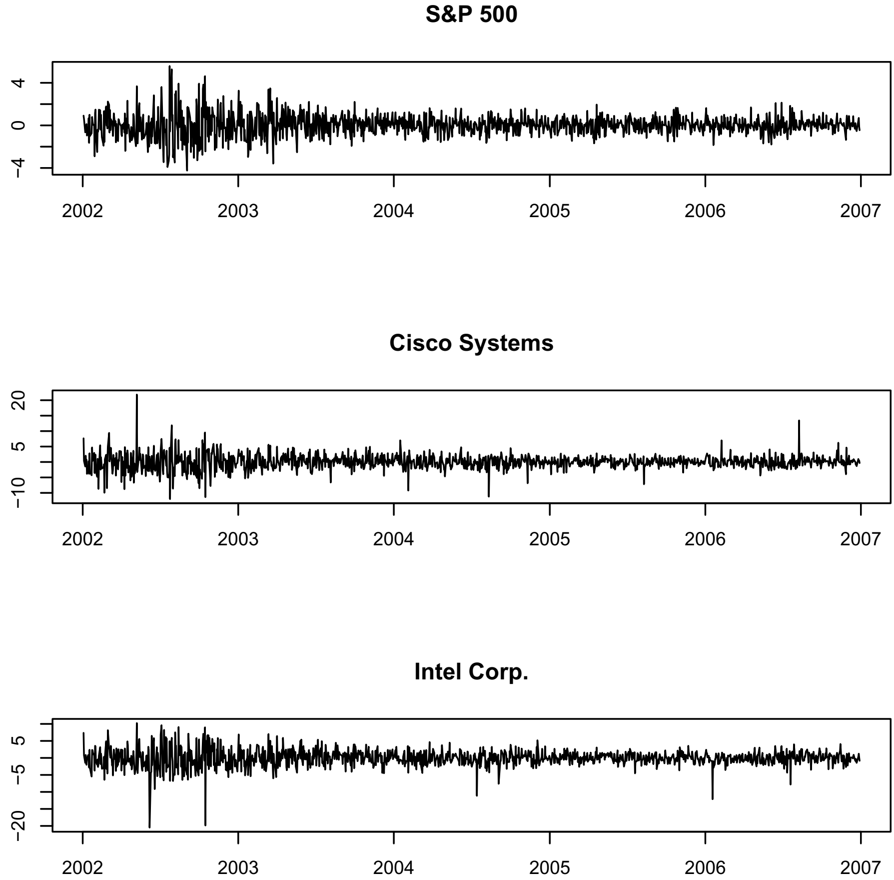

Fig. 1: Plots of daily log return of S&P 500 index, Cisco systems stock and Intel Corp. stock. Time span is from January 2, 2002 to December 29, 2006 with 1258 observations.

$$
\begin{aligned}
y_{ti} &= \nu_i + \varepsilon_{ti},\\
\varepsilon_{ti} &= \sigma_{t,ii} z_{ti},\\
\sigma^2_{t,ii} &= \omega_i + \alpha_{i1}\varepsilon^2_{ti} + \beta_i \sigma^2_{t-1,i}.
\end{aligned}
$$

We have assumed that z_{ti} follows the Meixner distribution with zero mean and unit variance, and shape parameters b_i and d_i.

In Table 2 we quote the maximum likelihood estimates of the parameters with the estimate of the asymptotic standard errors of the estimators. We quote the estimates both when the model components are independent among them and when the mixing matrix is estimated in the optimization process. The matrix is given by

$$\hat{V} = \begin{bmatrix} 0.0395 & 0.6520 & 0.3822 \\ -0.0752 & -0.3112 & 0.6798 \\ -0.8090 & 0.0485 & 0.0832 \end{bmatrix}$$

Note that both the models bring to a remarkable reduction of the Ljung-Box statistics computed on the standardized residuals. As was fairly expected, we have different results in both cases for the S&P 500 index and the two assets.

Tab. 2: Fitted GARCH(1,1) and TICA-GARCH(1,1) models for SCI.

	GARCH(1,1)			TICA-GARCH(1,1)		
	S&P 500	CISCO	INTEL	S&P 500	CISCO	INTEL
$\hat{v}$	0.0495	0.0794	0.0176	0.0387	0.0045	0.0015
	(0.1279)	(0.3367)	(0.3596)	(0.1593)	(0.2021)	(0.1760)
$\hat{\omega}$	0.0221	0.2287	0.2005	0.0338	0.1645	0.0743
	(0.0243)	(0.1936)	(0.2124)	(0.0476)	(0.1667)	(0.0890)
$\hat{\alpha}_1$	0.0947	0.1103	0.0860	0.0918	0.0977	0.0643
	(0.1178)	(0.1540)	(0.1216)	(0.1090)	(0.1573)	(0.1189)
$\hat{\beta}_1$	0.8810	0.8557	0.8805	0.8862	0.8180	0.8764
	(0.1177)	(0.1387)	(0.1332)	(0.1091)	(0.1889)	(0.1493)
$\hat{b}$	0.0025	0.1241	0.1084	-0.1906	0.0107	-0.0912
	(1.0474)	(0.7164)	(0.8101)	(1.1847)	(0.6901)	(0.7689)
$\hat{d}$	1.5956	0.4000	0.5133	1.7702	0.2448	0.2735
	(3.7025)	(0.3351)	(0.4665)	(4.7838)	(0.1660)	(0.1881)
$Q(10)$	13.1571	9.9101	3.0338	11.6492	4.7240	9.0358
	0.2150	0.4484	0.9806	0.3092	0.9088	0.5287
$Q(15)$	18.7388	14.4680	8.7323	17.5754	9.8102	18.2614
	0.2258	0.4904	0.8911	0.2856	0.8315	0.2492
$Q(20)$	23.3706	19.2122	16.8890	20.8993	16.5167	25.0822
	0.2710	0.5081	0.6602	0.4031	0.6841	0.1983

Note: Q(k) is referred to the Ljung-Box portmanteau test statistics computed on the standardized residuals with the corresponding p values. Figures in parentheses are the estimate of the asymptotic standard errors of the estimators.

The second data set, denoted as HDMA, consists of the 2515 daily log returns (in percentage) of stock prices of Hewlett-Packard Co. (HP), Dell Inc. (DELL), Microsoft Corp. (MSFT) and Advanced Micro Devices Inc. (AMS) in the period January, 2, 1997 - December, 29, 2006. These four assets have been analyzed with the ICA methodology by Wu and Yu (2003). Figure 2 depicts the time series plot of the four series. Descriptive statistics are listed in Table 3. The correlation matrix among the returns of the four assets is given by

$$\hat{R} = \begin{bmatrix} 1.0000 & 0.4664 & 0.4075 & 0.3773 \\ 0.4664 & 1.0000 & 0.5544 & 0.3993 \\ 0.4075 & 0.5544 & 1.0000 & 0.3840 \\ 0.3773 & 0.3993 & 0.3840 & 1.0000 \end{bmatrix}$$

Also with this data set the unconditional distribution of the four series exhibits a high kurtosis, indicating significative departure from normal distributions.

Tab. 3: Summary Statistics of the HDMA Data Set.

	HP	DELL	MSFT	AMD
T	2515	2515	2515	2515
Mean	0.0342	0.0819	0.0485	0.0161
Stdev	2.7925	2.9945	2.2510	4.4284
Skewness	-0.1165	-0.1096	-0.1565	-0.4915
Kurtosis	7.9075	6.5600	9.0794	10.1619
Min	-19.9994	-20.9838	-16.9627	-39.1595
1st Qu.	-1.3382	-1.4432	-1.0225	-2.2606
Median	0.0000	0.0000	0.0000	0.0000
3st Qu.	1.4183	1.5378	1.1420	2.2446
Max	15.8989	16.3532	17.8798	23.2193
Q(10)	12.6034	15.3156	14.6426	8.2881
Q(15)	23.8512	30.8006	24.4327	16.1990
Q(20)	31.8651	39.1942	28.6215	17.3810

Note: Q(k) is referred to the Ljung-Box statistics.

The model considered for the HDMA data set is a TICA model with temporal components AR(1)-APARCH(1,1):

$$\begin{aligned} y_{ti} &= \nu_i + \phi_{i1} y_{t-1,i} + \varepsilon_{ti}, \\ \varepsilon_{ti} &= \sigma_{t,ii} z_{ti}, \\ \sigma_{t,ii}^{\delta_i} &= \omega_i + \alpha_{i1}\left(\left|\varepsilon_{t-1,i}\right| - \gamma_{i1}\varepsilon_{t-1,i}\right)^{\delta_i} + \beta_{i1}\sigma_{t-1,ii}^{\delta_i}. \end{aligned}$$

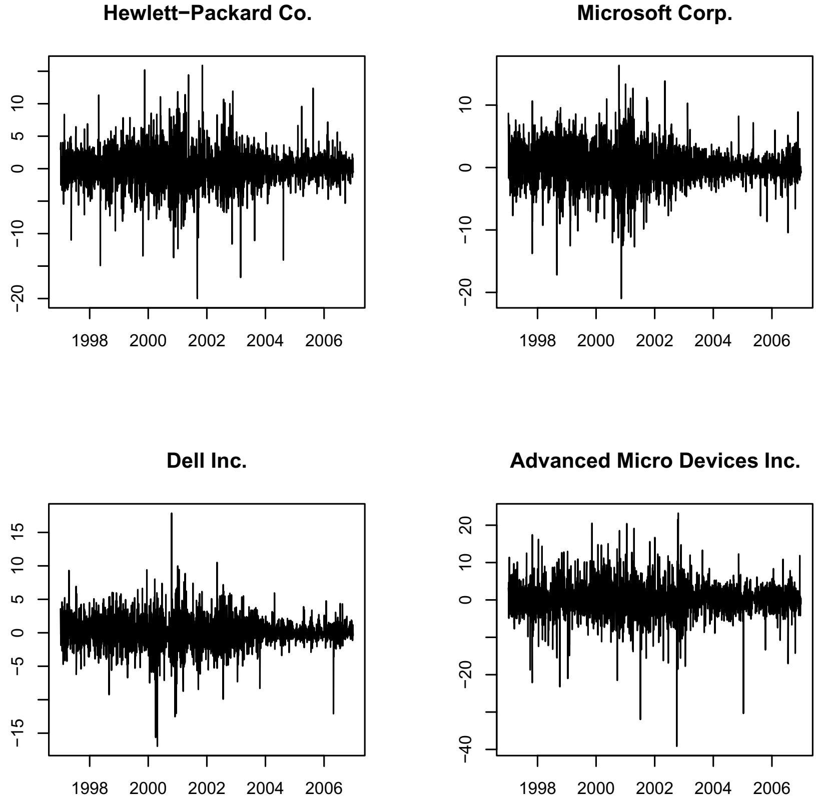

Fig. 2: Plots of daily log return of Hewlett-Packard Co. stock, Dell Inc. stock, Microsoft Corp. stock, and Advanced Micro Devices stock. Time span is from January 2, 1997 to December 29, 2006 with 2515 observations.

Also here we have assumed that z_{ti} follows the Meixner distribution with zero mean and unit variance, and shape parameters b_i and d_i.

In Table 4 we quote the estimates of the parameters obtained maximizing the likelihood function with the estimate of the asymptotic standard errors of the estimators. The estimated mixing matrix is given by

$$\hat{V} = \begin{bmatrix} 0.4909 & 0.4263 & -0.2503 & 0.4264 \\ -0.0987 & -0.5598 & -0.4966 & 0.3869 \\ -0.4422 & 0.1981 & 0.2797 & 0.5564 \\ -0.4632 & 0.3422 & -0.5372 & -0.2632 \end{bmatrix}$$

Tab. 4: TICA-APARCH(1,1) models for HDMA.

	HP	DELL	MSFT	AMD
$\hat{\nu}$	-0.0057	0.0150	-0.0196	-0.0028
	(0.1180)	(0.1460)	(0.1477)	(0.1396)
$\hat{\phi}_1$	-0.0528	0.0066	-0.0552	0.0168
	(0.1401)	(0.1281)	(0.1289)	(0.1399)
$\hat{\omega}$	0.0192	0.0437	0.0256	0.0379
	(0.0341)	(0.0775)	(0.0461)	(0.0661)
$\hat{\alpha}_1$	0.1219	0.0743	0.0632	0.1536
	(0.1322)	(0.1069)	(0.0930)	(0.1761)
$\hat{\gamma}_1$	0.0852	-0.1381	0.0061	0.1030
	(0.4227)	(0.5941)	(0.6435)	(0.4081)
$\hat{\beta}$	0.8816	0.9057	0.9318	0.8446
	(0.0993)	(0.1168)	(0.0855)	(0.1407)
$\hat{\delta}$	1.8160	1.5206	1.5397	1.8745
	(0.9037)	(1.4225)	(0.9655)	(1.0818)
$\hat{b}$	0.2911	0.0451	-0.1163	0.2550
	(0.5728)	(0.6030)	(0.5599)	(0.6330)
$\hat{d}$	0.3987	0.3401	0.3287	0.3591
	(0.2826)	(0.2249)	(0.2006)	(0.2727)
$Q(10)$	8.3591	10.6351	13.8424	7.2462
	0.5938	0.3866	0.1803	0.7020
$Q(15)$	15.5739	20.0346	16.5139	12.9584
	0.4109	0.1706	0.3487	0.6055
$Q(20)$	18.4507	22.3400	23.6091	18.0190
	0.5577	0.3223	0.2599	0.5862

Note: Q(k) is referred to the Ljung-Box portmanteau test statistics computed on standardized residuals with the corresponding p values. Figures in parentheses are the estimate of the asymptotic standard errors of the estimators.

The estimated volatility processes are drawn in Figure 3. The TICA-APARCH model brings also for this data set to a remarkable reduction of the Ljung-Box statistic computed on standardized residuals.

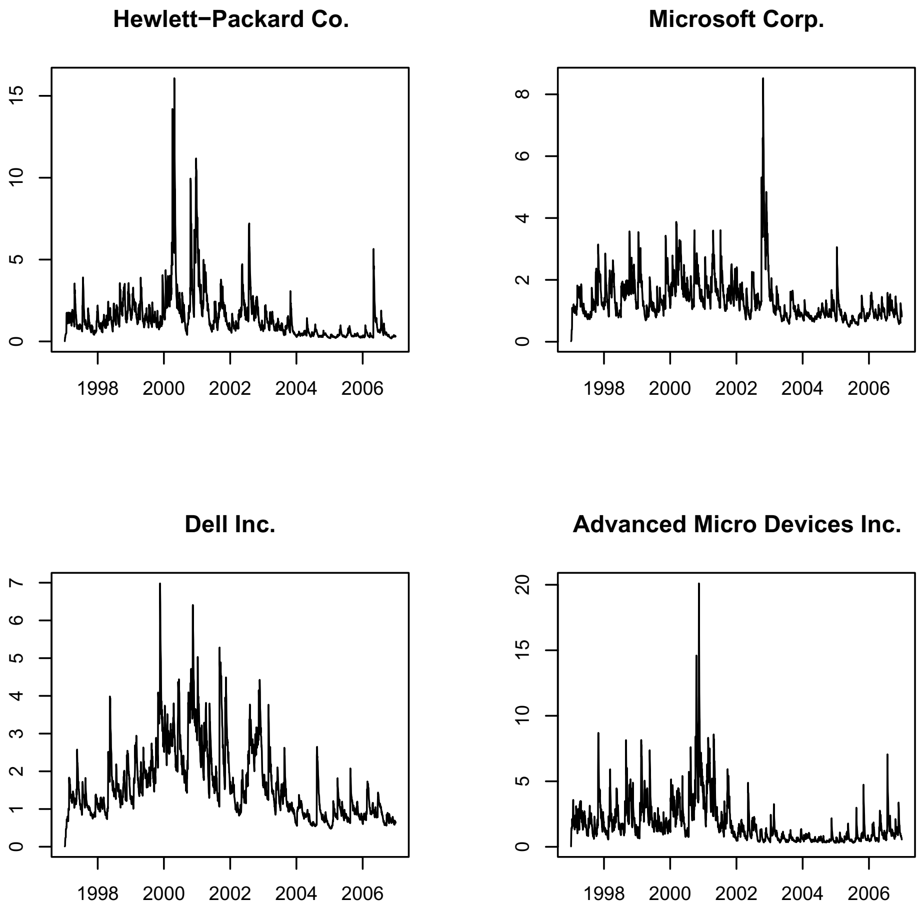

Fig. 3: ***Fitted volatility processes TICA-APARCH(1,1) model for daily log return of Hewlett-Packard Co. stock, Dell Inc. stock, Microsoft Corp. stock, and Advanced Micro Devices stock.***

6. FINAL REMARKS

On the basis of the results obtained on the two data sets considered, we can remark that

i) the ICA methodology can be a valid instrument to interpret the multivariate series of financial returns for temporal periods not very long;

ii) the estimates of the parameters of the Meixner distribution show that it is necessary to account for the moments higher than the second in interpreting temporal series; therefore, the TICA-GARCH model is better than the O-GARCH model because it separates the components of the model on the basis of the statistical independence instead of the incorrelation; similar remarks can be done for factorial GARCH models;

iii) increasing the number of assets in the portfolio, computational troubles increase; the construction of new algorithms can simplify the optimization of the likelihood function (Cheung and Xu, 2003; Hyvärinen, 2005).

Finally, we have verified that an ICA approach to the analysis of parallel financial series can lead to identify the underlying temporal structure of sources which can be used, for example, in the optimization of an asset portfolio.

REFERENCES

ALEXANDER C. (2001): *Market Models*, Wiley, New York.

ALEXANDER, C., and A. CHIBUMBA (1997): Multivariate orthogonal factor GARCH, University of Sussex, Mimeo.

BAWENS, L., S. LAURENT, and J. ROMBOUTS (2006): Multivariate GARCH models: a survey, *Journal of Applied Econometrics*, 21, 79-109.

BOLLERSLEV T. (1986): Generalized autoregressive conditional heteroskedasticity, *Journal of Econometrics*, 31, 307-327.

BOLLERSLEV T. (1990): Modelling the coherence in short-run nominal exchange rates: A multivariate generalized ARCH model, *Review of Economics and Statistics*, 72, 498-505.

BOLLERSLEV, T., R.F. ENGLE, and J. WOOLDRIDGE (1988): A capital asset pricing model with time varying covariances, *Journal of Political Economy*, 96, 116-131.

CHEN, Y., W. HÄRDLE, and S.C. SPOKOINY (2007): Portfolio value at risk based on independent component analysis, *Journal of Computational and Applied Mathematics*, 205, 594-607.

CHEUNG, YIU-MING, and LEI, XU (2003): Dual multivariatre auto-regressive modeling in state-space for temporal signal separation, *IEEE Transactions on Systems, Man, and Cybernetics - Part B: Cybernetics*, 33, 386-398.

DING, Z., C.W.J. GRANGER, and R.F. ENGLE (1993): A long memory property of stock market returns and a new model, *Journal of Empirical Finance*, 1, 83-106.

DUAN, J. (1997): Augmented GARCH(p,q) process and its diffusion limit, *Journal of Econometrics*, 79, 97-127.

ENGLE, R.F. (1982): Autoregressive, heteroskedasticity with estimates of the variance of United Kindgom inflation, *Econometrica*, 50, 987-1007.

ENGLE, R.F. (2002): Dynamic conditional correlation - a simple class of multivariate GARCH models, *Journal of Business and Economic Statistics*, 20, 339-350.

ENGLE, R.F. and KRONER, K.F. (1995): Multivariate simultaneous generalized ARCH, *Econometric Theory*, 11, 122-150.

ENGLE, R.F. and SHEPPARD, K.F. (2001): Theoretical and empirical properties of dynamic conditional correlation multivariate GARCH, *Mimeo, UCSD*.

FAN, JIANQING, MINGJIN WANG and QIWEI YAO (2005): Modelling multivariate volatilies via conditionally uncorrelated components. *arXiv:math/0506027v1 [math.ST]*.

GRIGELIONIS, N. (2001): Generalized z-distributions and related stochastic processes, *Lithuanian Mathematical Journal*, 39(1), 33-41.

GRIGOLETTO, M. and C. PROVASI (2007): Simulation and estimation of the Meixner distribution, *A preprint*.

HE, C. and T. TERÄSVIRTA (1999): Statistical properties of the Asymmetric Power ARCH process, in: R. Engle and H. White, eds., *Cointegration, causality and forecasting. Festschrift in honour of Clive W.J. Granger*, Oxford University Press, Oxford, 462-474.

HENTSCHEL, L. (1995): All in the family: nesting symmetric and asymmetric GARCH models, *Journal of Financial Economics*, 39, 71-104.

HYVÄRINEN, A. (2005): A unifying model for blind separation of independent sources, *Signal Processing*, 85, 1419-1427.

HYVÄRINEN, A., J. KARHUNEN, and E. OJA (2001): *Independent Component Analysis*, Wiley, New York.

LANNNE, N. and P. SAIKKONEN (2007): A multivariate generalized orthogonal factor GARCH model, *Jornal of Business & Economic Statistics*, 25, 61-75.

NASH, S.G. (1984): Newton-type minimization via the the Lanczos method, *SIAM Journal of Numerical Analysis*, 21, 770-778.

PALMITESTA, P. and C. PROVASI (2007): Maximum likelihood estimation of the APARCH model with skew distributions for the innovation process, *To appear*.

ROSE, C. and M.D. SMITH (2002): *Mathematical Statistics with Mathematica*, Springer, Berlin.

VAN DER WEIDE, R. (2002): GO-GARCH: a multivariate generalized orthogonal GARCH model, *Journal of Applied Econometrics*, 17, 549-564.

WOLFRAM, S. (1999): *The Mathematica Book, 4th edition*, Wolfram Media/Cambridge University Press, Cambridge.

WU, E.H.C. and YU P.L.H. (2005): Volatility modelling of multivariate financial time series by using ICA-GARCH models, in: M. Gallagher, J. Hoga, and F. Maire, eds., *IDEAL 2005, LNCS 3578*, Springer, Berlin, 571-579.

UN MODELLO PER L'ANALISI TEMPORALE DELLE COMPONENTI INDIPENDENTI DI PROCESSI DI VOLATILITÀ MULTIVARIATA

Riassunto

Un aspetto statistico dei rendimenti comunemente accettato in letteratura è che le volatilità di serie storiche parallele osservate nei mercati finanziari siano fra loro dipendenti; di conseguenza, un importante problema che sorge in molti contesti dell'econometria finanziaria è la previsione della dipendenza temporale di momenti del secondo ordine. In questo lavoro proponiamo un modello basato sull'analisi temporale delle componenti indipendenti per interpretare la volatilità multivariata di serie storiche parallele, assumendo che i fattori non osservabili siano processi AR - APARCH con errori che si distribuiscono come una Meixner. Una applicazione con due insiemi di rendimenti osservati mostra la possibilità di utilizzo del modello per l'analisi dei rendimenti finanziari.