Composed and Factor Composed Multivariate GARCH Models

Sebastian Kring
Department of Econometrics, Statistics and Mathematical Finance
School of Economics and Business Engineering
University of Karlsruhe
Postfach 6980, 76128 Karlsruhe, Germany
E-mail: sebastian.kring@statistik.uni-karlsruhe.de

Svetlozar T. Rachev
Chair-Professor, Chair of Econometrics, Statistics and Mathematical Finance
School of Economics and Business Engineering
University of Karlsruhe
Postfach 6980, 76128 Karlsruhe, Germany
and
Department of Statistics and Applied Probability
University of California, Santa Barbara
CA 93106-3110, USA
E-mail: rachev@statistik.uni-karlsruhe.de

Markus Höchstötter
Department of Econometrics, Statistics and Mathematical Finance
School of Economics and Business Engineering
University of Karlsruhe
Postfach 6980, 76128 Karlsruhe, Germany

Frank J. Fabozzi
Professor in the Practice of Finance
School of Management
Yale University
New Haven, CT USA

Abstract

In this paper we present a new type of multivariate GARCH model which we call the composed MGARCH and factor composed MGARCH models. We show sufficient conditions for the covariance stationarity of these processes and proof of the invariance of the models under linear combinations, an important property for factor modeling. Furthermore, we introduce an α-stable version of these models and fit a four dimensional α-stable composed MGARCH process to the returns on four German stocks included in the DAX index. We show in an in-sample analysis as well as in an out-of-sample analysis that the model outperforms the classical exponentially weighted moving average (EMWA) model introduced by RiskMetrics.

Keywords and Phrases: Volatility, Multivariate GARCH models, α-stable distributions, RiskMetrics, Risk Management, DAX index.

JEL Classification: C32, G0, G1

Acknowledgment: The authors would like to thank Stoyan Stoyanov and Borjana Racheva-Iotova from FinAnalytica Inc for providing ML-estimators encoded in MATLAB. For further information, see Stoyanov and Racheva-Iotova (2004).

1 Introduction

In modern risk management and factor modeling it is important to understand and to predict the temporal dependence structure of assets and risk factor returns in a multivariate time series framework. It is now widely accepted, and some researchers call it even a stylized fact (see McNeil, Frey, and Embrechts (2005)), that the conditional volatilities and the conditional correlation of multivariate financial time series vary over time and occur in clusters.

These style facts are very well understood and modeled by univariate GARCH models for one-dimensional time series. It is straightforward to generalize univariate GARCH models to multivariate GARCH models. Although they are the natural candidate to capture these stylized facts about multivariate financial time series, multivariate GARCH modeling has not been applied very often in the financial industry. The reason is that the implementation of these models is extremely difficult even in low dimensions, their major problem being that the number of parameters tends to explode with the dimension of the model. Because of this, the maximum likelihood function becomes very flat and optimization of the likelihood is practicably impossible in higher dimensions, as stressed by Alexander (2002). But from the an asset manager's perspective, a multivariate modeling framework is desirable since it opens the door to better decision tools in various areas, such as asset pricing, portfolio selection, factor modeling, and risk management.

Multivariate GARCH models were introduced by Bollerslev, Engle, and Wooldridge (1988). At the beginning of the 1990s new models were developed such as the constant conditional correlation (CCC) GARCH model by Bollerslev (1990), the principal component GARCH model by Ding (1994), the BEKK model of Baba, Engle, Kroner and Kraft (1995), and others. At the beginning of the 2000s Christodoulakis and Satchell (2002), Engle (2002), and Tse and Tsui (2002) developed the dynamic conditional correlation (DCC)-GARCH model that can be considered to be an extension of the CCC-GARCH model. Furthermore, Patton (2000) and Jondeau and Rockinger (2001) introduced copula-GARCH models.

The most common applications of multivariate GARCH models are for the study of the conditional covariance and correlation between several markets. Multivariate GARCH models can help asset managers understand if the volatility of one market (e.g., the Dow Jones 30), leads the volatilities of several other markets (such as Euro Stoxx 50, DAX 30 or Nikkei).

In asset pricing theory, the asset excess returns are modeled as linear combinations of factors (e.g., market return). In arbitrary approaches, the coefficients of the factors are assumed to be constant and estimated by an ordinary least squares (OLS-regression. Since these coefficients are the covariance between the asset excess return and the factor returns divided by the variance of the factor returns, these coefficients can be modeled as time varying by a multivariate GARCH model.

In asset management it is not recommended modeling directly all assets in a large portfolio by a multivariate GARCH model since the parameters of the model explode as noted above. Instead, an asset manager should use factor-model strategies in order to reduce the overall dimension of the time series modeling problem. After that the factors obtained can be modeled thoroughly by a multivariate GARCH or, even better, VARMA-MGARCH model.

In this paper we introduce two new multivariate GARCH models, which we refer to as the composed and factor composed MGARCH models. The idea behind these models comes from a common technique in portfolio risk management: Risk managers of large portfolios have to forecast risk functionals such as value-at-risk (VaR) or expected shortfall of the underlying portfolio. A common approach is to generate a univariate return series from the current asset shares and the multivariate return series of the assets in the portfolio. A univariate model such as a GARCH or ARMA-GARCH model is fitted to this time series allowing the calculation of these risk functionals. However, the univariate model is only valid for the current weights. Since the weights change daily, we have to repeat this procedure every day. Furthermore, the univariate model does not provide any information about the dependence structure of the assets, which is important for the portfolio risk manager. The basic idea behind the composed and factor composed MGARCH models is to use many linear combinations of the multivariate asset return series in the portfolio in order to reconstruct the conditional covariance matrix Σ_t. The matrix Σ_t can be reconstructed by solving an optimization problem ensuring the positivity of Σ_t.

We extend the composed and factor composed MGARCH model to an α-stable version with multivariate α-stable sub-Gaussian innovations. According to Rachev and Mittnik (2000), there are empirical as well as theoretical evidences that α-stable laws are the fundamental ”building blocks” (i.e., innovations) that drive asset return series in many sectors of the financial market.

This paper is organized as follows. In Section 2 we provide a short review of ARMA-GARCH models since they are the key device for the composed MGARCH models. In Section 3 we present and discuss the advantages and disadvantages of the most common multivariate GARCH models. In Section 4 we introduce the composed and factor composed GARCH models, put them in the context of the former models, and propose methods to fit composed and factor composed MGARCH models to data. We introduce an α-stable version of the composed and factor composed MGARCH model, that is based on the α-stable power GARCH processes introduced by Mittnik, Paolella, and Rachev (2002) in Section 5. In Section 6, the α-stable composed MGARCH model is applied to the returns on four German stocks included in the DAX index. We compare the performance of the proposed model with the exponentially weighted moving average (EWMA) model of RiskMetrics. Section 7 concludes the paper.

2 Univariate GARCH Models

The univariate generalized autoregressive conditional heteroscedasticity (GARCH) models have been successfully applied in financial econometrics since their introduction by Engle (1982) and Bollerslev (1986). They have been used with great success in volatility forecasting in several financial markets.

The voluminous literature related to GARCH models spans modeling exchange rates, equity returns, convergent term structure volatility forecast, and stochastic volatility models for option pricing and hedging. For a survey of ARCH-type models, see Bollerslev et al. (1992), Bera and Higgins (1993), Shephard (1996), Alexander (2001), among others.

In this section we review the basic definitions and properties in the field of univariate GARCH models. We do so because they are the fundamental device for multivariate GARCH modeling. We denote with $\mu_t = E(X_t|\mathcal{F}_{t-1})$, $t \in \mathbf{Z}$ the conditional mean of the time series $(X_t)_{t\in\mathbf{Z}}$, where $\mathcal{F}_t = \sigma(\{X_s : s \leq t\})$, $t \in \mathbf{Z}$, is the sigma field generated by the past and present values of $(X_t)_{t\in\mathbf{Z}}$.

Definition 1. *$(Z_t)_{t\in\mathbf{Z}}$ is a strict white noise (SWN) process if it is a series of identically distributed, finite-variance random variables.*

An important property of financial return series is whether they are *strictly* stationary or *covariance* stationary (see, e.g., McNeil, Frey, and Embrechts (2005) for a definition of this properties). Both of these definitions attempt to formalize the notion that the behavior of a time series is similar in any epoch in which we might observe it. Systematic changes in mean, variance, or the covariances between equally spaced observations are inconsistent with stationarity. We require these notions for the next definition.

Definition 2. *Let $(Z_t)_{t\in\mathbf{Z}}$ be $SWN(0,1)$. The process $(X_t)_{t\in\mathbf{Z}}$ is a GARCH(p,q) process if it is strictly stationary and if it satisfies, for all $t \in \mathbf{Z}$ and some strictly positive-valued process $(\sigma_t)_{t\in\mathbf{Z}}$, the equations*

$$X_t = \sigma_t Z_t,\ \sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i X_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2$$

where $\alpha_0 > 0$, $\alpha_i \geq 0$, $i = 1, ..., p$, and $\beta_j \geq 0$, $j = 1, ..., q$.

We demand strictly stationarity of univariate GARCH processes, since most financial return series seem to have this property.

It is straightforward to generalize GARCH processes to so-called ARMA-GARCH processes $(X_t)_{t\in\mathbf{Z}}$ satisfying the equation

$$X_t = \mu_t + \sigma_t Z_t,$$

where $(X_t - \mu_t)_{t\in\mathbf{Z}}$ follows a GARCH(p,q) process and $(\mu_t)_{t\in\mathbf{Z}}$ an ARMA process. (For an introduction to ARMA processes, see Hamilton (1994).) In daily return series volatility effects captured by the GARCH part are much more important than the mean effects modeled by the ARMA part of the model.[1] Because of this fact and for notational ease we do not consider ARMA processes in this paper.

In the next theorem we give sufficient and necessary conditions for covariance stationarity of GARCH processes.

Theorem 1. *A GARCH(p,q) process is a covariance-stationary white noise process if and only if $\sum_{i=1}^{p} \alpha_i + \sum_{j=1}^{q} \beta_j < 1$. The variance of the covariance-stationary process is given by $\alpha_0/(1 - \sum_{i=1}^{p} \alpha_i + \sum_{j=1}^{q} \beta_j)$.*

Proof. See McNeil, Frey, and Embrechts (2005). □

We will see in Section 4 that Theorem 1 is very useful to ensure covariance stationarity of composed and factor composed MGARCH processes.

[1] For a more detailed discussion of this issue, see RiskMetrics (1996) and McNeil, Frey, and Embrechts (2005).

3 Multivariate GARCH Models

In this section we give an historical overview[2] of the more important multivariate GARCH models. As stressed in Section 1 we always have to consider a trade-off between the complexity of the model (i.e., amount of parameters) and its applicability for financial modeling. A sophisticated multivariate GARCH specification might have the capability to capture all the phenomenons in the underlying multivariate time series, but if there may not exist an estimation procedure to fit the model to data, the model is not applicable. On the other hand, if the model has a too parsimonious parametrization, we can fit it easily to data but it might be worthless since it does not model the data appropriately.

Additional important properties of multivariate GARCH models are if the definition of the model ensures the positive definiteness of the conditional covariance matrix, the covariance stationarity of the process and the invariance of the model under linear transformation. A positive definite conditional covariance matrix can be achieved in most models, whereas the covariance stationarity is difficult to derive. For practical purposes, the former property is more important because we require a positive definite covariance matrix for the *Cholesky decomposition* in the definition of MGARCH processes.[3]

At the end of this section we describe how to integrate MGARCH models into factor models. This is an important issue in risk management since it is still not possible to model all risk factors of a large portfolio in one MGARCH model. Instead, we have to identify the common underlying risk factors of the portfolio and thoroughly model them by an MGARCH process.

But before beginning with an historical overview of the most common MGARCH models, we present the basic definitions and properties of this model class.

3.1 Basic Definitions and Properties

Consider a d-dimensional multivariate time series $(X_t)_{t\in\mathbf{Z}}$ defined on some probability space $(\Omega, \mathcal{F}, P)$. We assume for the rest that $(X_t)_{t\in\mathbf{Z}}$ is always a d-dimensional time series. We denote with $\mathcal{F}_t = \sigma_t(\{X_s : s \leq t\})$ the sigma field generated by the past and present values of the time series $(X_t)_{t\in\mathbf{Z}}$. Based on the efficient market hypothesis (see Fama (1991)) the sigma field $\mathcal{F}_t$ can be interpreted as representing the publicly available information at time t. Furthermore, we refer to

$$\mu_t = E_{t-1}(X_t) = E(X_t|\mathcal{F}_{t-1})$$

as the conditional mean and to

$$\mathrm{Var}_{t-1}(X_t) = E((X_t - \mu_t)(X_t - \mu_t)'|\mathcal{F}_{t-1})$$

as the conditional covariance matrix. The conditional covariance matrix P_t is defined by

$$P_t = \mathcal{P}(\Sigma_t) = \Delta(\Sigma_t)^{-1}\Sigma_t\Delta(\Sigma_t)^{-1},$$

[2]Detailed surveys about multivariate GARCH models can be found in Bauwens, Laurent, and Rombouts (2006) and McNeil, Frey, and Embrechts (2005).

[3]See Hamilton (1994) for a discussion of the Cholesky decomposition.

where the operator $\mathcal{P}(.)$ extracts the correlation matrix from the covariance matrix and the operator Δ satisfies

$$\Delta(\Sigma) = \mathrm{diag}(\sqrt{\sigma_{11}}, ..., \sqrt{\sigma_{dd}}).$$

A d-dimensional time series $(Z_t)_{t\in\mathbf{Z}}$ is called *multivariate strict white noise*, denoted by SNW(μ,Σ), if it is a series of independent elliptically distributed random vectors with mean μ and covariance matrix Σ.

Definition 3. *A process $(X_t)_{t\in\mathbf{Z}}$ has a multivariate martingale difference property with respect to the filtration $(\mathcal{F}_t)_{t\in\mathbf{Z}}$ if it satisfies*

$$E|X_t| < \infty \textit{ and } E(X_t|\mathcal{F}_{t-1}) = 0,$$

for all $t \in \mathbf{Z}$.

The martingale difference property corresponds to the stylized fact about daily financial return series that conditional expected returns are close to zero.[4] We will see below that a multivariate GARCH process fulfills this property.

Further important properties of multivariate time series can be captured by the following definitions.

Definition 4. *The multivariate time series $(X_t)_{t\in\mathbf{Z}}$ is strictly stationary if*

$$(X'_{t_1}, ..., X'_{t_n}) \stackrel{d}{=} (X'_{t_1+k}, ..., X'_{t_n+k}),$$

for all $t_1, ..., t_n, k \in \mathbf{Z}$ and for all $n \in \mathbf{N}$.

Definition 5. *The multivariate time series $(X_t)_{t\in\mathbf{Z}}$ is covariance stationary if the first two moments exist and satisfy*

$$\begin{aligned} E(X_t) &= \mu(t) = \mu,\ t \in \mathbf{Z} \\ \mathrm{Cov}(X_t, X_s) &= \mathrm{Cov}(X_{t+k}, X_{s+k})\ t, s, k \in \mathbf{Z}. \end{aligned}$$

Definitions 4 and 5 formalize the notion that the behavior of a time series is similar in any epoch in which we might observe it. Systematic changes in mean, variance or covariances between equally spaces observation are inconsistent with stationarity.

We turn now our attention to the definition of a multivariate GARCH process.

Definition 6. *Let $(Z_t)_{t\in\mathbf{Z}}$ be SWN($0, Id$). The process $(X_t)_{t\in\mathbf{Z}}$ is said to be a multivariate GARCH process if it is strictly stationary and satisfies equations of the form*

$$X_t = \Sigma_t^{1/2} Z_t,\ t \in \mathbf{Z},$$

where $\Sigma^{1/2}$ is the Cholesky factor of a positive-definite matrix Σ_t which isH measurable with respect to $\mathcal{F}_{t-1}$.

[4]For a more detail treatment of this topic see McNeil, Frey, and Embrechts (2005).

The most important property about multivariate GARCH models is that the conditional covariance Σ_t is measurable with respect to $\mathcal{F}_{t-1}$. This means that the covariance matrix of tomorrow's asset returns is known today. Note that in contrast to the definition of univariate GARCH processes, there is no functional specification of Σ_t in Definition 6. The functional form will depend on the specific model we define. Because of missing specification we cannot derive the general conditions that are necessary or sufficient for covariance stationarity of a multivariate GARCH process.

It is an immediate conclusion that any pure MGARCH process $(X_t)_{t\in\mathbf{Z}}$ has the martingale difference property, since we have

$$E(X_t|F_t) = E(\Sigma_t^{1/2} Z_t|\mathcal{F}_{t-1}) = \Sigma_t^{1/2} E(Z_t) = 0.$$

Furthermore, Σ_t is the conditional covariance matrix of any MGARCH process, since we have

$$\mathrm{Var}_{t-1}(X_t) = E(X_t X_t'|\mathcal{F}_{t-1}) = \Sigma_t^{1/2} E(Z_t Z_t')(\Sigma_t^{1/2})' = \Sigma_t^{1/2}(\Sigma_t^{1/2})' = \Sigma_t.$$

In particular, in the context of MGARCH models we use Σ_t interchangeably with conditional covariance $\mathrm{Var}(X_t|\mathcal{F}_{t-1})$.

Proposition 1. Let $(X_t)_{t\in\mathbf{Z}}$ be a multivariate GARCH process with conditional covariance matrix process $(\Sigma_t)_{t\in\mathbf{Z}}$. Then the univariate process $(a'X_t)_{t\in\mathbf{Z}}$ has a conditional variance process $(a'\Sigma_t a)_{t\in\mathbf{Z}}$ that is conditioned on the filtration $(\mathcal{F}_t)_{t\in\mathbf{Z}}$ for all $a \in \mathbf{R}^d$.

Proof.

$$\mathrm{Var}_{t-1}(a'X_t) = \mathrm{Var}(a'X_t|\mathcal{F}_{t-1}) = a'\,\mathrm{Var}(X_t|\mathcal{F}_{t-1})a = a'\Sigma_t a,$$

for all $a \in \mathbf{R}^d$. □

As in the univariate case, one can extend the definition of an multivariate GARCH process to a VARMA-GARCH process satisfying equations of the form

$$X_t = \mu_t + \Sigma_t^{1/2} Z_t,$$

where $(X_t - \mu_t)_{t\in\mathbf{Z}}$ follows a MGARCH process and $(\mu_t)_{t\in\mathbf{Z}}$ a VARMA process (see Hamilton (1994)). But such as in the univariate case volatility effects are much more important than mean effects with respect to daily return series. Because of this and the notational ease, we do not model the conditional mean process $(\mu_t)_{t\in\mathbf{Z}}$ in this paper.

3.2 MGARCH Models - an Historical Overview

Multivariate GARCH models were introduced by Bollerslev, Engle, and Wooldridge (1988) in the familiar half-vec *(vech)* form, providing a general framework for multivariate volatility models. In their paper they suggest the vector GARCH or VEC model.

Definition 7. *The process $(X_t)_{t\in\mathbf{Z}}$ is a VEC process if it has the general structure given in Definition 6, and the dynamics of the conditional covariance matrix Σ_t are given by the equations*

$$\operatorname{vech}\Sigma_t = a_0 + \sum_{i=1}^{p} \tilde{A}_i \operatorname{vech}(X_{t-i}X_{t-i}') + \sum_{j=1}^{q} \tilde{B}_j \operatorname{vech}(\Sigma_{t-j}),$$

for $a_0 \in \mathbf{R}^{d(d+1)/2}$ and matrices $\tilde{A}_i$ and $\tilde{B}_j$ in $\mathbf{R}^{(d(d+1)/2)\times(d(d+1)/2)}$.

The operator vech in Definition 7 stacks the columns of the lower triangle of a symmetric matrix in a single column vector of the length $d(d+1)/2$. In this general definition each element of Σ_t is a linear function of the lagged squared errors and cross-products of errors and the values of the lagged conditional covariance matrices. The fully unrestricted VEC model requires $O(d^4)$ parameters to be estimated by maximum likelihood, where d denotes the dimension of the underlying multivariate time series. The VEC model is certainly the most general MGARCH model, but it has too many parameters for practical purposes and is only of theoretical interest. It is also difficult to ensure the positive definiteness of the conditional covariance matrix. In order to overcome the drawbacks of the VEC model, Bollerslev, Engle, and Wooldbridge proposed the diagonal VEC or DVEC model in the same paper. The DVEC model is essential in the VEC model, but with the additional restriction that the matrices $\tilde{A}_i$ and $\tilde{B}_j$ in Definition 7 have to be diagonal. The DVEC model can be formulated elegantly in terms of the *Hadamard product*, denoted $\circ$, which signifies element-by-element multiplication of two matrices of the same size.

Definition 8. *The process $(X_t)_{t\in\mathbf{Z}}$ is called DVEC process if it has the general structure given in Definition 6 and satisfies equations of the form*

$$\Sigma_t = A_0 + \sum_{i=1}^{p} A_i \circ (X_{t-i}X_{t-i}') + \sum_{j=1}^{q} B_j \circ \Sigma_{t-j}$$

where A_0, A_i and B_j are symmetric matrices in $\mathbf{R}^{d\times d}$ such that A_0 has positive diagonal elements and all others matrices have non-negative diagonal elements.

The conditional covariance matrix Σ_t is a linear combination of own lagged squared errors and cross-products of errors. The advantage of the model compared to former ones is that only $O(d^2)$ parameters needed to be estimated by maximum likelihood. Furthermore, because of the Hadamard representation of the model it is easy to guarantee that Σ_t is positive definite for all t: Provided that A_0, A_i, B_j and the initial covariance matrix Σ_0 are positive definite for all t Attansio (1991) showed that Σ_t is positive definite for all t. Certainly, a disadvantage of the DVEC specification is that, in contrast to the VEC model, the volatility of a single component series cannot be affected directly by large lagged values in other time series. It should be mentioned that the DVEC model is still highly parameterized and large-scale systems are difficult to estimate in practice.

Bollerslev (1990) proposed the constant conditional correlation (CCC) multivariate GARCH specification. The CCC-GARCH model is the simplest representative of the class of MGARCH processes, where the marginals and the dependence structure of the multivariate time series are modeled separately. In this class, the marginals are modeled by univariate GARCH processes, whereas the dependence structure is defined model specific. In the case of the CCC-GARCH model, the dependence structure is captured by a constant correlation matrix leading to the following definition.

Definition 9. *The process $(X_t)_{t\in\mathbf{Z}}$ is called a CCC-GARCH process if it is a process with the general structure given in Definition 6. The conditional covariance matrix is of the form $\Delta_t P_c \Delta_t$, where*

(i) P_c is a constant, positive definite correlation matrix; and

(ii) Δ_t is a diagonal volatility matrix with elements $\sigma_{t,k}$ satisfying

$$\sigma_{t,k}^2 = \alpha_{k0} + \sum_{i=1}^{p_k} \alpha_{ki} X_{t-i,k}^2 + \sum_{j=1}^{q_k} \beta_{kj} \sigma_{t-j,k}^2, \ k = 1, ..., d, \tag{1}$$

where $\alpha_{k0} > 0$, $\alpha_{ki} \geq 0$, $i = 1, ..., p_k$, $\beta_{kj} \geq 0$, $j = 1, ..., q_k$.

It is easy to show that the design of the models guarantees a positive definite conditional covariance matrix. Because of the separation of the marginals and the dependence structure, an efficient two-step estimation procedure is available. In the first step we fit univariate GARCH models to the marginals and in the second step we use the devolatized residuals $Y_t = \Delta^{-1} X_t$ to estimate the constant correlation matrix P_c. This approach has the advantage that it opens the door to modeling large-scale systems. On the other hand, the model has a very parsimonious specification and the assumption of constant correlation may seem to be questionable in empirical work. In particular, Tsui and Yu (1999) have found that constant correlation can be rejected for certain multivariate time series. However, the CCC-GARCH model is more popular in the financial industry than the models described before and because of its simplicity it is a good starting point for MGARCH modeling.

The BEKK model of Baba, Engle, Kroner and Kraft was published in Engle and Kroner (1995). The model was also named after the two other authors who co-authored an earlier unpublished manuscript.

Definition 10. *The process $(X_t)_{t\in\mathbf{Z}}$ is a BEKK process if it has the general structure given in Definition 6 and if the conditional covariance matrix Σ_t follows the specification*

$$\Sigma_t = A_0 + \sum_{k=1}^{K} \sum_{i=1}^{p} A'_{k,i} X_{t-i} X'_{t-i} A_{k,i} + \sum_{k=1}^{K} \sum_{i=1}^{q} B_{k,j} \Sigma_{t-j} B_{k,j},$$

where $t \in \mathbf{Z}$, all matrices $A_{k,i}$ and $B_{k,j}$ are in $\mathbf{R}^{d\times d}$ and A_0 is symmetric and positive definite.

The advantage of the model is that it guarantees the positivity of the conditional covariance matrix Σ_t without imposing further restrictions. This is because of the general quadratic structure of the model. One can show that the BEKK model is a special case of the VEC model. The parameter K determines the generality of the process and one can show that the BEKK model covers all DVEC models. In practical applications the parameter K equals 1; even in this case the model is difficult to fit to data and it is rarely used in dimension larger than 3 or 4. In the most common version of the BEKK model $O(d^2)$ parameters have to be estimated. Certainly, a further disadvantage of the model is that the exact interpretation of the individual parameters is not obvious.

Ding (1994) described the principal component GARCH (PC-GARCH) model for the first time. This model was extensively investigated by Alexander (2002) under the name orthogonal GARCH.

Definition 11. *The process $(X_t)_{t\in\mathbf{Z}}$ follows a PC-GARCH model if it has the general structure of the process described in Definition 6 and if there exists some orthogonal matrix $\Gamma \in \mathbf{R}^{d\times d}$ with $\Gamma\Gamma' = \Gamma'\Gamma = Id$ such that $(\Gamma' X_t)_{t\in\mathbf{Z}}$ follows a pure diagonal GARCH model. The conditional covariance matrix Σ_t satisfies for all t*

$$\Sigma_t = \Gamma\Delta_t\Gamma',$$

where Δ_t is defined as in Definition 9.

It can be seen that the model ensures a positive definite covariance matrix Σ_t for all t without imposing further constrains. The strength of this approach is its simplicity and the possibility for dimensionality reduction. The model allows the estimation of large conditional covariance matrices since we have a straightforward estimation technique: In the first step we estimate the sample covariance matrix and by using the Spectral Decomposition Theorem, we calculate the sample principal components. In a second step we fit univariate GARCH models to the principal components. Furthermore, if certain components do not contribute much to the variability of the whole system, they can be neglected, leading to a dimensionality reduction. As Alexander (2002) stresses, the strength of the approach relies on modeling highly correlated systems such as the term structure of commodities futures or interest rates, where only a few principal components capture the behavior of the underlying multivariate time series. On the other hand, the simplicity of the model permits only a very limited evolution of the time series $(\Sigma_t)_{t\in\mathbf{Z}}$. If we have in mind that there is a one-to-one correspondence between a covariance matrix and an ellipsoid, we can visualize the evolution of $(\Sigma_t)_{t\in\mathbf{Z}}$: The corresponding ellipsoid can only be diluted and edged along its principal components, a rotation of the ellipsoid is not possible. As a result, the model only works well in those time series where the directions of the components do not vary over time since the principal components vary their directions over time. This is why the model reveals its weakness in modeling conditional correlation of asset returns.

In 1996 RiskMetrics suggested the exponentially weighted moving average scheme for modeling the conditional covariance matrix in RiskMetrics (1996).

Definition 12. *The process $(X_t)_{t\in\mathbf{Z}}$ is an exponentially weighted moving average (EWMA) process if it is a VEC process satisfying the updating scheme*

$$\Sigma_t = (1-\lambda)X_{t-1}X'_{t-1} + \lambda\Sigma_{t-1},$$

or equivalent,

$$\Sigma_t = (1-\lambda)\sum_{i=-\infty}^{t-1} X_i X'_i,$$

for all t.

The RiskMetrics model is widely used in industry, especially for portfolio VaR and is now considered to be an industry standard for market risk. The primary advantage of the RiskMetrics model is that it is extremely easy to estimate, since it has no parameters to be estimated. RiskMetrics suggested the smoothing factor λ to be 0.94 for daily log-returns and $\lambda = 0.97$ for monthly log-returns based on extensive data analysis in various markets and countries. Since in practice we use only the last M observations, we have to rescale the updating scheme in Definition 12, leading to

$$\Sigma_t = \frac{1-\lambda}{1-\lambda^{M+1}} \sum_{i=1}^{M} \lambda^i X_{t-i} X'_{t-i}.$$

The obvious drawback of the model is that it has no estimated parameters, and that it forces all assets to have the same decay coefficient irrespective of the asset type. It is necessary to assume the same decay coefficient for all assets to guarantee a positive definite conditional covariance matrix. The EWMA model of RiskMetrics can be regarded as the benchmark model that all other MGARCH models have to outperform.

As mentioned earlier, the assumption of constant correlation in the CCC-GARCH model seems unrealistic in empirical application. Christodoulakis and Satchell (2002), Engle and Sheppard (2001), and Tse and Tsui (2002) suggest a generalization of the CCC-GARCH model the so-called dynamic conditional correlation (DCC) model. There are different versions of the DCC model, the two most common being those of Tse and Tsui (2002), denoted DCC_T, and the one of Engle (2002), denoted DCC_E.

Definition 13. *The process* $(X_t)_{t\in\mathbf{Z}}$ *is a* DCC_T*-GARCH process if it is a process with the general structure given in Definition 6. The conditional covariance matrix is of the form* $\Sigma_t = \Delta_t P_t \Delta_t$*. The volatility matrix* Δ_t *is defined as in Definition 9 and* P_t *satisfies*

$$P_t = (1-\alpha-\beta)P_c + \alpha\Psi_{t-1} + \beta P_{t-1} \tag{2}$$

where $\alpha \geq 0$*,* $\beta \geq 0$ *and* $\alpha + \beta < 1$*,* $P_c \in \mathbf{R}^{d\times d}$ *is a positive definite matrix and* $\Psi_{t-1} \in \mathbf{R}^{d\times d}$ *is the correlation matrix of* $(Y_{t-1}, ..., Y_{t-M})$*, where* $(Y_t)_{t\in\mathbf{Z}} = (\Delta_t^{-1} X_t)_{t\in\mathbf{Z}}$ *is the devolatized process.*

Definition 14. *The process* $(X_t)_{t\in\mathbf{Z}}$ *is a* DCC_E*-GARCH process if it has the structure of the process given in Definition 13, but* P_t *satisfies*

$$P_t = \mathcal{P}(Q_t) \text{ and } Q_t = (1 - \sum_{i=1}^{p}\alpha_i - \sum_{j=1}^{q}\beta_j)Q_c + \sum_{i=1}^{p}\alpha_i Y_{t-i}Y'_{t-i} + \sum_{j=1}^{q}\beta_j Q_{t-j} \tag{3}$$

for all t*, where* Q_c *is the unconditional covariance matrix of the time series* $(Y_t)_{t\in\mathbf{Z}} = (\Delta_t^{-1}X_t)_{t\in\mathbf{Z}}$ *and the coefficients satisfy* $\alpha_i \geq 0$*,* $\beta_j \geq 0$ *and* $\sum_{i=1}^{p}\alpha_i + \sum_{j=1}^{q}\beta_j < 1$*.*

The two versions of the DCC-GARCH model permit estimating large conditional covariance matrices since we have the same two-step estimation procedure as in the CCC-GARCH model. The chief difference is that the dependence structure is modeled by a time dependent correlation matrix which is defined by equation (2) and equation

(3), respectively. In particular, we can divide the second step into sub-steps. In the first sub-step, we estimate the matrices P_c and Q_c and in the second sub-step we estimate the scalars α and β and α_i and β_j, respectively. The DCC-GARCH model guarantees the positive definiteness of the sample covariance matrix without imposing further constraints. Since in the $\mathrm{DCC_T}$ model P_c, Ψ_{t-i}, and P_{t-1} are positive definite, so is P_t and since in the $\mathrm{DCC_E}$ model $Q_c, Y_{t-i}Y'_{t-i}$ and Q_{t-i}, so is Q_t and P_t. If we set $\alpha = \beta = \alpha_i = \beta_i = 0$ we observe that the $\mathrm{DCC_T}$ and the $\mathrm{DCC_E}$ reduce to the CCC-GARCH model. It can be tested if $\alpha = \beta = 0$ and $\alpha_i = \beta_j = 0$ in order to check whether imposing constant correlation is empirically relevant. Certainly, a drawback of the DCC model is that α and β in the $\mathrm{DCC_T}$ model and α_i and β_j in the $\mathrm{DCC_E}$ model are scalars instead of matrices. Hence, all entries of the conditional correlation matrix are influenced by the same coefficients which might not be realistic in empirical work. However, these conditions are necessary in order to maintain the positivity of the conditional correlation matrix. In the literature there are extensions of the DCC-GARCH model to overcome the scalar problem. For a further discussion see Billio et al. (2003), Engle (2002) and Pelletier (2003).

Patton (2000) and Jondeau and Rockinger (2001) were the first to propose a copula-factor model. These models are specified in such a way that the marginals follow GARCH processes and their time varying dependence structure is modeled by a copula. The following definition formalizes this class of processes.

Definition 15. *The process $(X_t)_{t\in\mathbf{Z}}$ is a copula-GARCH model iff it is a process satisfying*

(i) the marginals $(X_{t,k})_{t\in\mathbf{Z}}$, $k = 1, ..., d$, follow a GARCH(p_k, q_k) process;

(ii) the dependence structure of the marginals is modeled by a copula

$$C(u_1, ..., u_d|R_t),$$

where R is the parameter set defining the copula C and R_t follows an updating scheme $R_t = f(X_{t-1}, X_{t-2}, ...)$;

(iii) the conditional distribution is given by

$$X_t|\mathcal{F}_{t-1} = C(F^{-1}_{X_{t,1}|\mathcal{F}_{t-1}}(X_{t,k}), ..., F^{-1}_{X_{t,d}|\mathcal{F}_{t-1}}(X_{t,d})|R_t).$$

Note that R_t is measurable with respect to $\mathcal{F}_{t-1}$ and time varying. Patton and Jondeau and Rockinger highlighted in both papers the need to allow for a time-variation in the conditional copula function. The copula function is rendered time varying through its parameters, which can be functions of past data. The copula-MGARCH model can be viewed as an extension of the CCC and DCC-GARCH model.

In Table 1[5] we show an overview of the number of parameters used in the models presented in this section. The VEC, BEKK and DVEC models are only applied in low dimensions ($d \leq 10$) and the VEC is purely of theoretical interest. The CCC-, DCC- and PC-GARCH models are implemented in dimensions larger than 10 in the financial industry.

[5]Table 1 is adopted from McNeil, Frey, and Embrechts (2005).

Model	Parameter count	≥ 10	2	5	10
VEC	$d(d+1)(1+(p+q)d(d+1)/2)/2$	No	21	465	6105
BEKK	$d(d+1)/2 + Kd^2(p+q)$	No	11	65	255
DVEC	$d(d+1)/2(1+p+q)$	No	9	45	165
CCC	$d(d+1)/2 + d(p+q)$	Yes	7	25	77
DCC_T	$d(d+1)/2 + 2 + d(p+q)$	Yes	9	27	77
DCC_E	$d(d+1)/2 + (p+q)(d+1)$	Yes	9	27	77
PC	$d(d+1)/2 + (p+q)d$	Yes	7	25	75
EWMA	$d(d+1)/2$	Yes	3	15	55

Table 1: Summary of numbers of parameters in various multivariate GARCH models: in the CCC, DCC_T and DCC_E it is assumed that the numbers of GARCH terms are p and q; in the DCC_T we assume that the conditional correlation matrix has 2 parameters and in the DCC_E we suppose that the conditional correlation matrix has $p+q$ parameters. The second column gives the general formula. The final columns give the numbers for models of dimensions $2, 5$, and 10 when $p = q = 1$.

3.3 Factor Modeling with MGARCH Models

The material presented in this section follows McNeil, Frey, and Embrechts (2005). It is still not recommended to model all financial risk factors with general multivariate GARCH models. Rather, these models have to be combined with factor-model strategies to reduce the overall dimension of the time series modeling problem.

A fundamental consideration is whether factors are identified *a priori* and treated as exogenous variables, or whether they are treated as endogenous variables and statistical factors manufactured from the observed data.

Suppose we adopt the former approach and identify a small number of common factors F_t to explain the variation in many equity returns X_t. These common factors can be modeled by multivariate GARCH models. The dependence of the individual returns on the factor returns can then be modeled by calibrating a factor model of the type

$$X_t = a + BF_t + \epsilon,\ t = 1, ..., n.$$

We assume then that, conditional on the factors F_t, the errors form a multivariate white noise process with GARCH volatility structure.

The latter approach is based on a linear transformation of the equity returns X_t to define factors

$$F_t \quad = \quad (F_{t,1}, ..., F_{t,k})' = \Gamma_1' X_t,$$

where $\Gamma_1 \in \mathbf{R}^{d\times k}$ and $k << d$. The factors F_t can be modeled by a transformation invariant multivariate GARCH model and should explain most of the variability of the equity returns X_t. This approach leads to a factor model of the form

$$X_t = \Gamma_1 F_t + \epsilon_t,\ t = 1, ..., n,$$

where the error term is usually ignored in practice.

4 Composed and Factor Composed MGARCH Models

For high dimensional multivariate GARCH modeling it is indispensable that the model definition permits an efficient estimation procedure. In the previous section we have seen that only those models that allow for a multi-step estimation procedure can be applied in higher dimensions ($d \geq 10$). For example, the specification of the CCC-, DCC-, and copula-GARCH model admits a two-step estimation procedure to estimate the dynamics of the marginals and the temporal dependence structure separately. Similar, in the PC-GARCH model the temporal dependence structure is captured by modeling the principal components of the unconditional covariance matrix through univariate GARCH processes. In addition, since we are interested in statistical factor modeling it is essential that the presented models are invariant under linear transformation. In this section we show that composed and factor composed MGARCH models exhibit the invariance property and allow for two-step estimation procedures.

4.1 Definitions and Properties

The key idea behind the specification of the composed and factor composed MGARCH model introduced in this section is to identify the temporal dependence structure of the multivariate time series $(X_t)_{t\in\mathbf{Z}}$ through linear combinations of this time series. These linear combinations are modeled by univariate GARCH processes. In a second step the dependence structure is reconstructed by solving an optimization problem. But before defining these two models, we have to introduce additional notions.

We assume in the following that all processes $(X_t)_{t\in\mathbf{Z}}$ exhibit unconditional and conditional second moments. Let $(X_t)_{t\in\mathbf{Z}}$ be a d- dimensional process and we denote by $\mathcal{F}_t(a) = \sigma(\{a'X_s : s \leq t\})$ the sigma field generated by the past and present values of the univariate time series $(a'X_t)_{t\in\mathbf{Z}}$. If $(a'X_t)_{t\in\mathbf{Z}}$ follows a GARCH(p,q) process, we write $\sigma_t^2(a)$ for the conditional variance $\mathrm{Var}(a'X_t|\mathcal{F}_{t-1}(a))$.

The sigma field $\mathcal{F}_t$ (defined in Section 3) includes more information than $\mathcal{F}_t(a)$. It is important to note that mathematically

$$\mathrm{Var}(a'X_t|\mathcal{F}_{t-1}) = \mathrm{Var}(a'X_t|\mathcal{F}_{t-1}(a)) \tag{4}$$

is *not* true since we are dealing with different sigma fields.

But we reasonable assume that equation (4) holds for many multivariate financial return series at least approximately. In the univariate case, we know that GARCH models based on the filtration $\mathcal{F}(e_i)$ have been successfully applied in volatility forecasting, implying immediately

$$\begin{aligned}\mathrm{Var}(e_i'X_t|\mathcal{F}_{t-1}) &= e_i'\Sigma_t e_i = \sigma_{t,ii} \approx \mathrm{Var}(e_i'X_t|\mathcal{F}_{t-1}(e_i)) \\ &= \alpha_0 + \sum_{i=1}^{p} X_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j,i}^2. \end{aligned} \tag{5}$$

Hence, equation (5) justifies equation (4) for the marginals, i.e. $a = e_i$, $i = 1, ..., d$, from an empirical point of view. Another argument as to why equation (4) holds is the efficient market hypothesis (see Fama (1991)) which asserts that all relevant information of an asset is represented by the past and present values of the time series

$(X_{t,i})_{t\in\mathbf{Z}}$. Hence, we obtain

$$\mathrm{Var}(e_i' X_t | \mathcal{F}_{t-1}) \quad = \quad \mathrm{Var}(e_i' X_t | \mathcal{F}_{t-1}(e_i)).$$

Concerning non-trivial linear combinations of multivariate financial time series, we observe that a very widespread and successfully applied technique in the risk management of large portfolios is to take the current weights $w_t \in \mathbf{R}^d$, where d denotes the number of assets in the portfolio, and to generate univariate time series $(w_t' X_t)_{t\in\mathbf{Z}}$ from the portfolio's log-returns $(X_t)_{t\in\mathbf{Z}}$. Then a univariate GARCH process is fitted to the time series $(w_t' X_t)_{t\in\mathbf{Z}}$. The success of this technique is based on the empirical fact that we have at least approximately

$$\begin{aligned}
\mathrm{Var}(w_t' X_t | \mathcal{F}_{t-1}) &= w_t' \Sigma_t w_t \approx \sigma_t^2(w_t) = \mathrm{Var}(w_t' X_t | \mathcal{F}_{t-1}(w_t)) \\
&= \alpha_0 + \sum_{i=1}^{p} \alpha_i (w_t' X_{t-i})^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2(w_t),
\end{aligned}$$

where $w_t \in \mathbf{R}^d$. This consideration is evidence for equation (4).

In contrast to the marginals, it is at least not immediately clear how to justify

$$\mathrm{Var}(a' X_t | \mathcal{F}_{t-1}) = \mathrm{Var}(a' X_t | \mathcal{F}_{t-1}(a))$$

with the efficient market hypothesis, since $(a' X_t)_{t\in\mathbf{Z}}$ is an artificial time series that is not observable and traded in financial markets.

These semi-theoretical considerations are summed up in the following definitions.

Definition 16. *A multivariate time series $(X_t)_{t\in\mathbf{Z}}$ is projection-efficient if it satisfies*

$$a' X_t | \mathcal{F}_{t-1} \quad \overset{d}{=} \quad a' X_t | \mathcal{F}_{t-1}(a), \tag{6}$$

for all $t \in \mathbf{Z}$ and $a \in \mathbf{R}^d$.

Note that the last definition implies

$$\mathrm{Var}(a' X_t | \mathcal{F}_{t-1}(a)) = \mathrm{Var}(a' X_t | \mathcal{F}_{t-1}) = a' \Sigma_t a.$$

The next definition tells us how to model $\mathrm{Var}(a' X_t | \mathcal{F}_{t-1})$.

Definition 17. *A multivariate time series $(X_t)_{t\in\mathbf{Z}}$ is called GARCH-projection-efficient if it is projection-efficient and satisfies*

$$\mathrm{Var}(a' X_t | \mathcal{F}_{t-1}) = \alpha_0 + \sum_{i=1}^{p} \alpha_i (a' X_{t-i})^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2(a),$$

for all $t \in \mathbf{Z}$ and $a \in \mathbf{R}^d$.

The notion of projection-efficient is derived from the efficient market hypothesis in the sense that all information about the projective time series $(a' X_t)_{t\in\mathbf{Z}}$ included in the sigma algebra $\mathcal{F}_{t-1}$ equals the information in the sigma algebra $\mathcal{F}_{t-1}(a)$. Furthermore, the term GARCH-projection-efficient stresses that all information about the volatility $a' X_t$ is captured by $\mathcal{F}_{t-1}(a)$ and can be modeled by a univariate GARCH(p, q) process.

In particular, a GARCH-projection-efficient time series $(X_t)_{t\in\mathbf{Z}}$ has the property that the variance of $a'X_t|\mathcal{F}_{t-1}(a)$ can be modeled by a GARCH process. Due to the consideration above there should be many multivariate financial time series which are at least approximately projection-efficient and GARCH-projection efficient. It is now straightforward and consistent to define the following multivariate GARCH process.

Definition 18. *The process $(X_t)_{t\in Z}$ follows a composed MGARCH (CMGARCH) process if it is a process with the general structure given in Definition 6 and the conditional matrix $\Sigma_t = (\sigma_{t,ij})$satisfies*

(i) $(\sigma^2_{t,k})_{t\in\mathbf{Z}}$ *follows a univariate GARCH(p_k,q_k) process for* $k = 1, ..., d$.

(ii) For all $i, j = 1, ..., d$, $i \neq j$ *we have*

$$\sigma_{t,ij} = \frac{1}{4}(\sigma^2_t(e_i + e_j) - \sigma^2_t(e_i - e_j)),$$

where $(\sigma^2_t(e_i + e_j))_{t\in\mathbf{Z}}$ *and* $\sigma^2_t(e_i - e_j))_{t\in\mathbf{Z}}$ *follow univariate GARCH(p^+_{ij}, q^+_{ij}) and GARCH(p^-_{ij}, q^-_{ij}) processes, respectively.*

The composed MGARCH model does not impose any explicit functional form of the conditional covariance matrix Σ_t such as the other models in Section 3. We only have to assume that $(X_t)_{t\in\mathbf{Z}}$ follows a multivariate GARCH process which is GARCH-projection-efficient in order to be consistent. This idea is formalized in the next theorem.

Theorem 2. *Let $(X_t)_{t\in\mathbf{Z}}$ be a GARCH-projection-efficient MGARCH process with conditional covariance time series $(\Sigma_t)_{t\in\mathbf{Z}}$, then $(\Sigma_t)_{t\in\mathbf{Z}}$ can be modeled by a composed MGARCH process.*

Proof. Let $i, j = 1, ..., d$, then we have

$$\begin{aligned}
\sigma_{t,ij} &= \mathrm{Cov}(X_{t,i}, X_{t,j}|\mathcal{F}_{t-1}) \\
&= E((X_{t,i} - \mu_{t,i})(X_{t,j} - \mu_{t,j})|\mathcal{F}_{t-1}) \\
&= E\left(\frac{1}{4}(X_{t,i} + X_{t,j} - (\mu_{t,i} + \mu_{t,j})^2))\right. \\
&\quad \left. -\frac{1}{4}(X_{t,i} - X_{t,j} - (\mu_{t,i} - \mu_{t,j}))^2)|\mathcal{F}_{t-1}\right) \\
&= \frac{1}{4}E((X_{t,i} + X_{t,j} - (\mu_{t,i} + \mu_{t,j}))^2|\mathcal{F}_{t-1}) \\
&\quad -\frac{1}{4}E((X_{t,i} - X_{t,j} - (\mu_{t,i} - \mu_{t,j}))^2)|\mathcal{F}_{t-1}) \\
&= \frac{1}{4}(\mathrm{Var}((e_i + e_j)'X_t|\mathcal{F}_{t-1}) - \mathrm{Var}((e_i - e_j)'X_t|\mathcal{F}_{t-1})) \\
&\overset{(*)}{=} \frac{1}{4}(\mathrm{Var}((e_i + e_j)'X_t|\mathcal{F}_{t-1}(e_i + e_j)) - \mathrm{Var}((e_i - e_j)'X_t|\mathcal{F}_{t-1}(e_i - e_j))) \\
&= \frac{1}{4}(\sigma^2_t(e_i + e_j) - \sigma^2_t(e_i - e_j))
\end{aligned}$$

(*) holds in the last equation, since the process is projection-efficient. In particular, $(\sigma^2_t(e_i + e_j))_{t\in\mathbf{Z}}$ and $(\sigma^2_t(e_i - e_j))_{t\in\mathbf{Z}}$ follow GARCH processes, since $(X_t)_{t\in\mathbf{Z}}$ is GARCH-projection-efficient. The same arguments hold also for diagonal entries $\sigma_{t,ii}$ of Σ_t. □

The property GARCH-projection-efficient is essential for CMGARCH processes since it determines the class of processes that can be modeled by them. In contrast to many applied MGARCH models, we have a motivation for the CMGARCH models.[6] The CMGARCH model resembles the PC-GARCH (see Definition 11) in the sense that we use linear combinations of the multivariate time series $(X_t)_{t\in\mathbf{Z}}$ to model the conditional covariance matrix Σ_t and hence, the temporal dependence structure. But in the CMGARCH approach we extend this idea to a new level, since the conditional covariance matrix is determined solely by univariate processes, which is not the case in the PC-GARCH model. In contrast to most of the other multivariate GARCH models we reviewed in Section 3 we can easily derive sufficient conditions for the covariance-stationarity of a CMGARCH model.

Theorem 3. *Let the time series $(X_t)_{t\in\mathbf{Z}}$ follow a composed MGARCH process. The process is covariance stationary if all GARCH processes $(\sigma^2_{t,i})_{t\in\mathbf{Z}}$, $(\sigma^2_t(e_i+e_j))_{t\in\mathbf{Z}}$ and $(\sigma^2_t(e_i-e_j))_{t\in\mathbf{Z}}$ are covariance stationary, or equivalently, if the coefficients of the GARCH processes $(X_{t,i})_{t\in\mathbf{Z}}$, $((e_i+e_j)'X_t))_{t\in\mathbf{Z}}$ and $((e_i-e_j)'X_t)_{t\in\mathbf{Z}}$ satisfy*

$$\begin{aligned}\sum_{k=1}^{p_i}\alpha_k^{(i)}+\sum_{k=1}^{q_i}\beta_k^{(i)} &< 1,\\ \sum_{k=1}^{p_{ij}^+}\alpha_k^{(ij+)}+\sum_{k=1}^{q_{ij}^+}\beta_k^{(ij+)} &< 1, \textit{ and}\\ \sum_{k=1}^{p_{ij}^+}\alpha_k^{(ij-)}+\sum_{k=1}^{q_{ij}^-}\beta_k^{(ij-)} &< 1,\end{aligned}$$

for all $i,j=1,...,d$.

Proof. Since $(\sigma_{t,i})_{t\in\mathbf{Z}}$ is covariance stationary, we obtain

$$E(X^2_{t,i})=E(\sigma^2_{t,i}Z^2_t)=E(\sigma^2_{t,i})=\sigma^2_i,$$

for $1=1,...,d$. Furthermore, for the unconditional covariances we have

$$\begin{aligned}E(X_{t,i}X_{t,j}) &= E(\mathrm{Cov}(X_{t,i},X_{t,j}|\mathcal{F}_{t-1}))\\ &= E(\sigma_{t,ij})\\ &= E(\frac{1}{4}(\sigma^2_t(e_i+e_j))-\sigma^2_t(e_i-e_j))\\ &= \frac{1}{4}(E(\sigma^2_t(e_i+e_j))-E(\sigma^2_t(e_i-e_j)))\\ &= \frac{1}{4}(\sigma^2(e_i+e_j)-\sigma^2(e_i-e_j))\end{aligned}$$

for all $i,j=1,...,d$. In particular, we have $\sigma_{ij}=E(X_{t,i}X_{t,j})$ for all $t\in\mathbf{Z}$. Since $(X_t)_{t\in\mathbf{Z}}$ is a multivariate martingale difference with finite, non-time-dependent second moments σ_{ij}, $i,j=1,...,d$, it is covariance-stationary white noise. □

[6] See the arguments leading to Definitions 16 and 17.

One further important property of a multivariate GARCH model is the invariance of the model with respect to linear combinations, that is, the times series $(Y_t)_{t\in\mathbf{Z}} = (FX_t)_{t\in\mathbf{Z}}$ belongs to the same model class, where $F \in \mathbf{R}^{k\times d}$. If $(X_t)_{t\in\mathbf{Z}}$ is a time series of asset returns, a linear transformation $(FX_t)_{t\in\mathbf{Z}}$ corresponds to new assets (portfolios combining the original assets). It seems sensible that a model should be invariant, otherwise the question arises as to which basic assets should be modeled. This aspect becomes very important when we are interested in statistical factor modeling in order to reduce the dimensionality of the portfolio. Statistical risk factors are linear combinations of the underlying assets. Modeling the factors and the assets, respectively, should lead to the same results.

Theorem 4. *Let $(X_t)_{t\in\mathbf{Z}}$ follow a GARCH-projection-efficient CMGARCH process. Then the CMGARCH process is invariant under linear transformation, i.e., the process $(Y_t)_{t\in\mathbf{Z}} = (FX_t)_{t\in\mathbf{Z}}$ follows a CMGARCH process in terms of $(\mathcal{G}_t)_{t\in\mathbf{Z}}$ and*

$$\mathrm{Var}(Y_t|\mathcal{F}_{t-1}) = F\Sigma_t F' = \mathrm{Var}(Y_t|\mathcal{G}_{t-1}), \tag{7}$$

where $F \in \mathbf{R}^{k\times d}$, $k \in \mathbf{N}$, and $\mathcal{G}_t$ is the sigma field generated by $\sigma(\{Y_s : s \le t\})$.

Proof. Note, that we have $\mathcal{G}_t \subset \mathcal{F}_t$ for all $t \in \mathbf{Z}$. First we show equation (7).

$$\begin{aligned}
\mathrm{Cov}(Y_{t,i}, Y_{t,j}|\mathcal{G}_{t-1}) &= \frac{1}{4}(\mathrm{Var}((e_i+e_j)'Y_t|\mathcal{G}_{t-1}) - \mathrm{Var}((e_i-e_j')Y_t|\mathcal{G}_{t-1}))\\
&= \frac{1}{4}(\underbrace{E(((e_i+e_j)'Y_t)^2|\mathcal{G}_{t-1})}_{(1)} - (\underbrace{E((e_i+e_j)'Y_t|\mathcal{G}_{t-1})}_{(2)})^2\\
&\quad - \underbrace{E(((e_i-e_j)'Y_t)^2|\mathcal{G}_{t-1})}_{(3)} - (\underbrace{E((e_i-e_j)'Y_t|\mathcal{G}_{t-1})}_{(4)})^2)
\end{aligned} \tag{8}$$

Since we have $\mathcal{F}_t((e_i+e_j)'F) \subset \mathcal{G}_t \subset F_t$ we can derive from term (1) in equation (8) that we have

$$\begin{aligned}
E(((e_i+e_j)'Y_t)^2|\mathcal{G}_{t-1}) &= E(E(((e_i+e_j)'Y_t)^2|\mathcal{F}_{t-1})|\mathcal{G}_{t-1}))\\
&= E(E(((e_i+e_j)'Y_t)^2|\mathcal{F}_{t-1}((e_i+e_j)'F))|\mathcal{G}_{t-1}))\\
&= E(((e_i+e_j)'Y_t)^2|\mathcal{F}_{t-1}((e_i+e_j)'F))\\
&= E(((e_i+e_j)'Y_t)^2|\mathcal{F}_{t-1}).
\end{aligned} \tag{9}$$

Analogously, equation (9) holds also for terms (2), (3), and (4) in equation (8). Hence, we obtain

$$\begin{aligned}
\mathrm{Cov}(Y_{t,i}, Y_{t,j}|\mathcal{G}_{t-1}) &= \frac{1}{4}(E(((e_i+e_j)'Y_t)^2|\mathcal{F}_{t-1}) - E((e_i+e_j)'Y_t|\mathcal{F}_{t-1})^2\\
&\quad - E(((e_i-e_j)'Y_t)^2|\mathcal{F}_{t-1}) - E((e_i-e_j)'Y_t|\mathcal{F}_{t-1})^2)\\
&= \frac{1}{4}(\mathrm{Var}((e_i+e_j)'Y_t|\mathcal{F}_{t-1}) - \mathrm{Var}((e_i-e_j')Y_t|\mathcal{F}_{t-1}))\\
&= \mathrm{Cov}(Y_{t,i}, Y_{t,j}|\mathcal{F}_{t-1}).
\end{aligned}$$

Furthermore, we observe that we have

$$\mathrm{Cov}(Y_{t,i}, Y_{t,j}|\mathcal{F}_{t-1}) = e_i' F\Sigma_t F' e_j$$

and we have proved equation (7). We can derive from equations (8) and (9) that we have

$$\begin{aligned}\mathrm{Cov}(Y_{t,i}, Y_{t,j}|\mathcal{G}_{t-1}) &= \frac{1}{4}(\mathrm{Var}((e_i+e_j)'Y_t|\mathcal{G}_{t-1}(e_i+e_j)) \\ &\quad - \mathrm{Var}((e_i+e_j)'Y_t|\mathcal{G}_{t-1}(e_i-e_j))),\end{aligned}$$

where $\mathcal{G}_t(e_i+e_j) = \mathcal{F}_t((e_i+e_j)'F)$ and $\mathcal{G}_t(e_i-e_j) = \mathcal{F}_t((e_i-e_j)'F)$. Since the process $(X_t)_{t\in\mathbf{Z}}$ is GARCH-projection-efficient, $\mathrm{Var}((e_i+e_j)'Y_t|\mathcal{G}_{t-1}(e_i+e_j))$ and $\mathrm{Var}((e_i-e_j)'Y_t|\mathcal{G}_{t-1}(e_i-e_j))$ follow a GARCH process. Hence, the time series $(Y_t)_{t\in\mathbf{Z}}$ is a CMGARCH process. □

Because of Theorem 6 we can consistently define an extension of the composed multivariate GARCH model what we call the factor composed multivariate GARCH (FCMGARCH) model.

Definition 19. *The process $(X_t)_{t\in\mathbf{Z}}$ follows a factor composed MGARCH (FCMGARCH) process, if there exists some orthogonal matrix $\Gamma \in \mathbf{R}^{d\times d}$ satisfying $\Gamma\Gamma' = \mathrm{Id}$ such that $(\Gamma' X_t)_{t\in\mathbf{Z}}$ follows a composed MGARCH process.*

The definition of the factor composed MGARCH model resembles the definition of the PC-GARCH model. As with the PC-GARCH model we are interested in modeling the principal components $(Y_t) = (\Gamma' X_t)_{t\in\mathbf{Z}}$ of the unconditional covariance matrix $\mathrm{Cov}(X_t)$. If the multivariate time series $(X_t)_{t\in\mathbf{Z}}$ is highly correlated, this approach has the advantage that we can model the system through $d_1 << d$ principal components appropriately. But in contrast to the PC-GARCH model, the factor composed GARCH model is more flexible because not only are the principal components modeled by univariate GARCH processes but also the conditional covariance between these factors. Furthermore, the FCMGARCH model offers the opportunity of reducing the dimensionality of the estimation problem since the number of parameters needed to be estimated is proportional to d^2. This follows from the next proposition.

Proposition 2. Let $(X_t)_{t\in\mathbf{Z}}$ be a d-dimensional CMGARCH process and the time series $(X_{t,i})_{t\in\mathbf{Z}}$, $((e_i+e_j)'X_t)_{t\in\mathbf{Z}}$ and $((e_i-e_j)'X_t)_{t\in\mathbf{Z}}$ follow GARCH(p,q) processes, then we have to estimate $(p+q+1)d^2$ parameters.

Proof. Since the conditional volatility process $(\sigma_{t,ii})_{t\in\mathbf{Z}}$ of $(X_{t,i})_{t\in\mathbf{Z}}$ follows a GARCH(p,q) process, we have to estimate $1+p+q$ coefficients for one GARCH process. Since there are d different marginal processes we have to estimate $d(1+p+q)$ parameters. For each conditional covariance process $(\sigma_{t,ij})_{t\in\mathbf{Z}}$, $i \neq j$, we estimate $2(1+p+q)$ parameters due to the formula

$$\sigma_{t,ij} = \frac{1}{4}(\sigma_t^2(e_i+e_j) - \sigma_t^2(e_i-e_j)),$$

for $i,j = 1,...,d$. Since we have

$$\sigma_{t,ij} = \sigma_{t,ji}$$

we have $d(d-1)/2$ different conditional covariance processes. Hence we have to estimate $d(d-1)(1+p+q)$ parameters for the conditional covariance processes $(\sigma_{t,ij})_{t\in\mathbf{Z}}$. In conclusion, we estimate $(1+p+q)d^2$ parameters. □

In many financial applications, it is sufficient to use a GARCH$(1,1)$ to model the linear combinations. In the case of a d-dimensional CMGARCH process $(X_t)_{t\in\mathbf{Z}}$, we have to estimate $3d^2$ coefficients. In the CMGARCH model, we have to estimate more parameters than in the CCC-GARCH, DCC-GARCH or the PC-GARCH model as can be seen from Table 1. But we stress, that the CMGARCH model is more flexible than these models since we do not have a restrictive functional form which allows only a very constrained evolution of the conditional covariance process $(\Sigma_t)_{t\in\mathbf{Z}}$. A further advantage of this model is that we estimate Σ_t only through univariate GARCH processes. This approach allows us to circumvent the problem of applying maximum likelihood estimation in high dimension. As stressed in Alexander (2001), this is the fundamental problem of multivariate GARCH modeling.

Nevertheless, in multivariate GARCH modeling for large portfolios several researchers such as Alexander (2001) and McNeil, Frey, and Embrechts (2005) recommend a factor model approach in order to reduce the dimensionality of the portfolio. According to Theorem 6, the FCMGARCH model is consistent with the CMGARCH and, as mentioned before, is predestinated to model the principal components of the unconditional covariance matrix $\mathrm{Cov}(X_t)$.[7] In addition, in many financial time series we observe the so-called "80/20 rule" or "Pareto principle" which says that 20% of the largest eigenvalues account for 80% of the overall variability. Hence, if we model 20% of the "largest" principal components we can decrease the parameters needed to be estimated by 96%.

Certainly, a drawback of the CMGARCH model is that its definition does not ensure a positive definite conditional covariance matrix Σ_t, meaning that if the estimation error of $\hat{\Sigma}_t$ becomes too large the matrix is not necessarily positive definite. In the next section we present a method to deal with this problem.

4.2 Estimation of the Models

In this section, we introduce two approaches to estimate the CMGARCH and FCMGARCH model. The two approaches have in common that in the first step the problem of estimating the conditional covariance matrix Σ_t is decomposed into $n \in \mathbf{N}$ simpler estimation problems. For these estimation problems, efficient solving algorithms are available. In the following steps we apply these solutions of the n estimation problems to reconstruct the conditional covariance matrix Σ_t. A similar approach has been successfully applied to estimate multivariate α-stable sub-Gaussian distributions (see Nolan (2005) and Kring et al. (2007)).

Since the CMGARCH and FCMGARCH model specification does not guarantee the positivity of the conditional covariance matrix Σ_t, fortunately, the second presented estimation procedure ensures a positive definite estimate $\hat{\Sigma}_t$ of the conditional covariance matrix by applying the Cholesky Decomposition Theorem.

The first estimation approach is immediately derived from the definition of the CMGARCH process.

(1) Let $X_1, ..., X_t$ be a sequence of return data. Fit univariate GARCH processes to the projective data sets $X_{1,i}, ..., X_{t,i}$, where $(e_i + e_j)'X_1, ..., (e_i + e_j)'X_t$

[7] Due to Theorem 3, we can assume $(X_t)_{t\in\mathbf{Z}}$ to be covariance stationary.

and $(e_i - e_j)'X_1, ..., (e_i - e_j)'X_t$, $1 \leq i < j \leq d$. Denote the corresponding volatility estimates with $\hat{\sigma}_t^2(e_i)$, $\hat{\sigma}_t^2(e_i + e_j)$ and $\hat{\sigma}_t^2(e_i - e_j)$.

(2) Reconstruct the conditional covariance matrix $\hat{\Sigma}_t$ by

$$\begin{aligned}\hat{\Sigma}_{t,ii} &= \hat{\sigma}_t^2(e_i),\\ \hat{\Sigma}_{t,ij} &= \frac{1}{4}(\hat{\sigma}_t^2(e_i + e_j) + \hat{\sigma}_t^2(e_i - e_j)),\end{aligned}$$

or alternatively,

(2') Reconstruct the conditional covariance matrix $\hat{\Sigma}_t$ by

$$\begin{aligned}\hat{\Sigma}_{t,ii} &= \hat{\sigma}_t^2(e_i),\\ \hat{\Sigma}_{t,ij} &= \frac{1}{2}(\hat{\sigma}_t^2(e_i + e_j) - \hat{\sigma}_t^2(e_i) - \hat{\sigma}_t^2(e_j)),\end{aligned}$$

where $1 \leq i < j \leq d$.

This approach has the advantage that it is computationally straightforward. In step (1) we have to fit d^2 GARCH processes in order to calculate $\hat{\Sigma}$ in step (2) and for the alternative approach we have to fit only $d(d+1)/2$ GARCH processes in step (1) since we do not have to estimate $\hat{\sigma}_t^2(e_i - e_j)$ in step (2'). This method has the drawback that we cannot ensure the positivity of the conditional covariance matrix $\hat{\Sigma}$. Hence, we always have to check whether $\hat{\Sigma}_t$ is positive definite. One way of doing this is to apply the Spectral Decomposition Theorem.

If all eigenvalues are positive, then $\hat{\Sigma}_t$ is positive definite.

The second estimation approach is called the *regression approach* since we reconstruct the conditional covariance matrix Σ_t using a regression.

(1) Let $X_1, ..., X_t$ be a sequence of return data. Fit univariate GARCH processes to the projective time series $u_i'X_1, ..., u_i'X_t$, where $u_i \in \mathbf{R}^d$ and $i = 1, ..., n$. Denote the volatility estimates with $\hat{\sigma}_t^2(u_i)$, $i = 1, ..., n$.

(2) Reconstruct the conditional covariance matrix by

$$\hat{\Sigma}_t^{(2)} = \text{argmin}_{\Sigma \in S^{d\times d}} \sum_{i=1}^{n} (u_i'\Sigma u_i - \sigma_t^2(u_i))^2,$$

where $S^{d\times d} = \{\Sigma | \Sigma \in \mathbf{R}^{d\times d}, \Sigma' = \Sigma\}$.

The regression approach may be more accurate than the former approach because it uses multiple directions, whereas the first method only uses the directions e_i, e_j, $e_i + e_j$ and $e_i - e_j$. In addition, this approach allows for more flexibility since the directions u_i are not predefined and their number n is also variable. For example, it might be better to model the principal components Y_i, $i = 1, ..., d$ of the sample covariance matrix Σ and their linear combinations than the marginals in order to estimate $\hat{\Sigma}_t$. Or, one might increase the accuracy of the estimate for Σ_t by increasing the number of directions. But still, so far we cannot ensure $\hat{\Sigma}_t^{(2)}$ to be positive definite. We overcome this drawback through an additional step (3).

(3) In case $\det(\hat{\Sigma}_t^{(2)}) > 0$, set

$$\hat{\Sigma}_t = \hat{\Sigma}_t^{(2)}.$$

Otherwise reconstruct the conditional covariance matrix $\hat{\Sigma}_t^{(3)}$ by

$$\begin{aligned} \hat{\Delta}_t &= \operatorname{argmin}_{\Delta \in D_d} \sum_{i=1}^{n} (u_i' \Delta \Delta' u_i - \hat{\sigma}_t^2(u_i))^2, \text{ and} \\ \hat{\Sigma}_t^{(3)} &= \hat{\Delta}_t \hat{\Delta}_t', \end{aligned}$$

where $D_d = \{\Delta | \Delta \in \mathbf{R}^{d\times d}, \Delta$ regular upper triangular matrix$\}$. Finally set

$$\hat{\Sigma}_t = \hat{\Sigma}_t^{(3)}.$$

Due to the Cholesky Decomposition Theorem, the optimization problem in step (3) is equivalent to

$$\hat{\Sigma} = \operatorname{argmin}_{\Sigma \in D_d^2} \sum_{i=1}^{n} (u_i' \Sigma u_i - \hat{\sigma}_t^2(u_i))^2$$

where D_d^2 is the set of all positive definite $d \times d$ matrices. Hence, the conditional step (3) guarantees $\hat{\Sigma}_t$ to be positive definite. While the optimization problem in step (3) is computationally much more involved than in the one in step (2)[8], it is important to note that if $\hat{\Sigma}_t^{(2)}$ is positive definite then $\hat{\Sigma}_t^{(2)}$ equals $\hat{\Sigma}_t^{(3)}$.

It is straightforward to show that $\hat{\Sigma}_t^{(3)} = \hat{\Delta}_t \hat{\Delta}_t'$ is positive definite, since we have

$$\begin{aligned} u' \hat{\Sigma}_t u &= u' \hat{\Delta}_t \hat{\Delta}_t' u \\ &= ||\hat{\Delta}_t' u||^2 \\ &> 0, \end{aligned}$$

for all $u \in \mathbf{R}^d \setminus \{0\}$ and $\hat{\Delta}_t \in D_d$.

5 α-Stable Composed and Factor Composed MGARCH Models

5.1 α-Stable Power-GARCH Processes

It is often observed when fitting GARCH models to financial time series that univariate GARCH residuals tend to be heavy tailed. To accommodate this, GARCH models with heavier conditional innovation distributions than those of the normal have been proposed, among them the Student's t and the Generalized Hyperbolic Distribution. To allow for heavy-tailed, conditional distributions, GARCH processes with α-stable error distributions have been considered by McCulloch (1985), Panorska, Mittnik, and Rachev (1995), Mittnik, Paolella, and Rachev (1998), Rachev and Mittnik (2000), among others.

[8] The optimization problem in step (2) can be solved by OLS regression.

An objection against the use of the α-stable distribution is that it has no second moments. This seems to contradict empirical studies suggesting the existence of third or fourth moments for various financial return data. But as Mittnik, Paolella, and Rachev (2002) stressed, these findings had been almost exclusively obtained by the use of the Hill (1975) or related tail estimators, which are known to be highly unreliable.

In this section, we present α-stable power-GARCH processes which were originally introduced by Rachev and Mittnik (2000) and Mittnik, Paolella, and Rachev (2002).

Definition 20. *An univariate process $(X_t)_{t\in\mathbf{Z}}$ is called an α-stable Paretian power-GARCH process, in short, an $S_{\alpha,\beta,\delta}$GARCH(r,s) process, if it is described by*

$$X_t = \mu_t + \sigma_t Z_t, \text{ where } Z_t \sim S_\alpha(1,\beta,0), \tag{10}$$

and

$$\sigma_t^\delta = \alpha_0 + \sum_{i=1}^{p} \alpha_i |X_{t-i} - \mu_{t-i}|^\delta + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^\delta,$$

where $\alpha_0 > 0$, $\alpha_i \geq 0$, $i = 1,...,r$, $\beta_j \geq 0$, $j = 1,...,q$, $0 < \delta < \alpha$ and $S_\alpha(1,\beta,0)$ denotes the α-stable distribution with tail index $\alpha \in (1,2]$, skewness parameter $\beta \in [-1,1]$, zero location parameter and unit scale parameter. The location parameter process $(\mu_t)_{t\in\mathbf{Z}}$ in (10) follows an ARMA process.

Since for $\alpha < 2$ Z_t in Definition 20 does not possess moments of order α or higher, we restrict α to be in the set $(1,2]$ in order to possess first moments.[9] This restriction is consistent with financial return data (see among others Höchstötter, Rachev, and Fabozzi (2005) and Rachev and Mittnik (2000)), where we observe α to be in the same range. Furthermore, it is important to note, that σ_t is just a time-varying scaling parameter, implying $\sigma_t Z_t|\mathcal{F}_{t-1} \sim S_\alpha(\sigma_t, \sigma_t\beta, 0)$. Hence, in an α-stable power-GARCH process we forecast the scale parameter of the α-stable innovation distribution.

Mittnik, Paolella, and Rachev (2002) show the following proposition.

Proposition 3. The $S_{\alpha,\beta,\delta}$GARCH process has unique, strictly stationary solution if

$$\lambda_{\alpha,\beta,\delta} \sum_{i=1}^{p} \alpha_i + \sum_{j=1}^{q} \beta_j \leq 1, \tag{11}$$

where $\lambda_{\alpha,\beta,\delta} = E(|\mathbf{Z}_t|^\delta)$ and $Z_t \sim S_\alpha(1,\beta,0)$.

Proposition 3 allows us to guarantee a unique, strictly stationary solution of $S_{\alpha,\beta,\delta}$GARCH process by imposing equation (11) during estimation.

5.2 A Multivariate α-stable GARCH Model

In this section, we propose an α-stable version of the composed and factor composed MGARCH models, allowing similar estimation procedures as in the ordinary versions

[9] See Samorodnitsky and Taqqu (1994) for the existence of moments of an α-stable random variable.

of these processes. In particular, we are dealing with processes with α-stable innovation we believe, according to Rachev and Mittnik (2000), that these are the fundamental ”building blocks” that drive asset return processes. But the main problem we face in defining an α-stable MGARCH model with multivariate α-stable innovations is that we do not possess second moments and the conditional covariance matrix or any covariance matrix are not defined. [10] We overcome this problem by choosing the α-stable sub-Gaussian distribution for the innovations. In this particular case, we obtain a substitute for the covariance matrix, the so-called dispersion matrix. The dispersion matrix has the same interpretation in terms of the scaling properties of the distribution (see Samorodnitsky and Taqqu (1994) and Kring et al. (2007) for a discussion of this issue). But before defining these processes, we have to introduce additional notions.

Definition 21. *$(Z_t)_{t\in\mathbf{Z}}$ is multivariate α-stable strict white noise if it is a series of independent and identically distributed α-stable sub-Gaussian random vectors with dispersion matrix Σ.*

An α-stable strict white noise process with mean μ and covariance matrix Σ will be denoted by α-SWN(μ,Σ). It can be shown easily that a dispersion matrix of an α-stable sub-Gaussian random vector has to be non-negative definite.

Definition 22. *Let $(Z_t)_{t\in\mathbf{Z}}$ be α-stable strict white noise α-SWN($0, \mathrm{Id}$). The process $(X_t)_{t\in\mathbf{Z}}$ is said to be an α-stable multivariate GARCH process if it is strictly stationary and satisfies equations of the form*

$$X_t = \Sigma_t^{1/2} Z_t, t \in \mathbf{Z},$$

where $\Sigma^{1/2}$ is the Cholesky factor of a positive-definite matrix Σ_t which is measurable with respect to $\mathcal{F}_{t-1}$.

As in Section 4, we take no account of the conditional mean vector for notational ease. It is usually specified as function of the past, through a vectorial autoregressive moving average (VARMA) representation.

Due to Kring et al. (2007, pp. 12-13), it follows immediately that we have

$$\Sigma_t^{1/2} Z_t|\mathcal{F}_{t-1} \sim E_d(0, \Sigma_t, \psi_{sub}(., \alpha)) \tag{12}$$

In order to shorten the notation we introduce the dispersion operator

$$\mathrm{Disp}(X) = \Sigma,$$

where X is an α-stable sub-Gaussian random vector with dispersion matrix Σ. In particular, we define by

$$\mathrm{Disp}(X_t|\mathcal{F}_{t-1}) = \Sigma_t$$

the conditional dispersion matrix of X_t given $\mathcal{F}_{t-1}$. This notion is well defined because of equation (12). Furthermore, we have

$$\mathrm{Disp}(a'X_t|\mathcal{F}_{t-1}) = a'\Sigma_t a,$$

[10] See also Doganoglu, Hartz, and Mittnik (2006) for a multivariate model with conditionally varying and heavy-tailed risk factors.

since we have $a'X_t|\mathcal{F}_{t-1} \sim E_1(0, a'\Sigma_t a, \psi_{sub}(., \alpha))$ (see Kring et al. (2007, p. 13) or Samorodnitsky and Taqqu (1994, p. 77 et seq.)).

In Section 4 we argued that equation (4) holds at least approximately for many financial return time series possessing second moments. We can now repeat those arguments for processes with α-stable sub-Gaussian innovations. Hence, we obtain

$$\mathrm{Disp}(a'X_t|\mathcal{F}_{t-1}) = \mathrm{Disp}(a'X_t|\mathcal{F}_{t-1}(a))$$

holds at least approximately. We can now rephrase Definition 16.

Definition 23. *An α-stable multivariate GARCH process $(X_t)_{t\in\mathbf{Z}}$ is α-stable projection-efficient if it satisfies for all $t \in \mathbf{Z}$ and $a \in \mathbf{R}^d$*

$$a'X_t|\mathcal{F}_{t-1} \stackrel{d}{=} a'X_t|\mathcal{F}_{t-1}(a) \sim E_1(0, a'\Sigma_t a, \psi_{sub}(., \alpha)),$$

where Σ_t is the conditional dispersion matrix.

According to Kring et al. (2007) Proposition 2 and Samorodnitsky and Taqqu (1994, p. 77 et seq.) the following equation holds for the scaling and dispersion parameter of $a'X$

$$\sigma(a) = \left(\frac{1}{2}a'\Sigma a\right)^{1/2} = (\mathrm{Disp}(a'X))^{1/2}, \tag{13}$$

where $X \sim E_d(0, \Sigma, \psi_{sub}(., \alpha))$. Note that in the classical case where second moments exist, we have

$$\sigma(a) = (a\Sigma a)^{1/2}, \tag{14}$$

where $\sigma(a)$ and $(a\Sigma a)^{1/2}$ can be considered as the standard deviation and variance of $a'X$, respectively. In the α-stable case, we have to take the factor $1/2$ in the relation between scaling parameter and dispersion parameter of dX due to equation (13). In particular, we can write

$$E_1(0, a'\Sigma a, \psi_{sub}(., \alpha)) \text{ or } S_\alpha\left(\left(\frac{1}{2}a'\Sigma a\right)^{1/2}, 0, 0\right)$$

for the distribution of $a'X$. In addition, if $a'X_t|\mathcal{F}_{t-1}(a)$ is α-stable distributed, we denote the scaling parameter with $\sigma_t(a)$ and we can write

$$\sigma_t(a) = \left(\frac{1}{2}\,\mathrm{Disp}(a'X_t|\mathcal{F}_{t-1}(a))\right)^{1/2}. \tag{15}$$

We restate Definition 17 in terms of α-stable power GARCH processes.

Definition 24. *An α-stable multivariate GARCH process $(X_t)_{t\in\mathbf{Z}}$ is called power-GARCH-projection-efficient if it is α-stable projection efficient and satisfies*

$$(\sigma_t(a))^\delta = \alpha_0 + \sum_{i=1}^{p} \alpha_i |a'X_{t-i}|^\delta + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^\delta(a), \tag{16}$$

for all $t \in \mathbf{Z}$, $\alpha_0 > 0$, $\alpha_i \geq 0$, $i = 1, ..., r$, $\beta_j \geq 0$, $j = 1, ..., q$, $0 < \delta < \alpha$.

We have now all notions to define an α-stable version of the CMGARCH model.

Definition 25. *The process $(X_t)_{t\in\mathbf{Z}}$ follows an α-stable composed MGARCH process if it is a process with the general structure given in Definition 22 and the conditional dispersion matrix $\Sigma_t = (\sigma_{t,ij})$ satisfies*

(i) $(\sigma_t(e_i))_{t\in\mathbf{Z}} = ((\frac{1}{2}\sigma_{t,ii})^{1/2})_{t\in\mathbf{Z}}$ *follows a $S_{\alpha,0,\delta}$GARCH(p_i, q_i) process for $i = 1, ..., d$.*

(ii) For all $i, j = 1, ..., d$ $i \neq j$ we have

$$\sigma_{t,ij} = \frac{1}{2}(\sigma_t^2(e_i + e_j) - \sigma_t^2(e_i - e_j)),$$

where $(\sigma_t(e_i + e_j))_{t\in\mathbf{Z}}$ and $(\sigma_t(e_i - e_j))_{t\in\mathbf{Z}}$ follow a $S_{\alpha,0,\delta}$GARCH(p_{ij}^+, q_{ij}^+) and $S_{\alpha,0,\delta}$GARCH(p_{ij}^-, q_{ij}^-) processes, respectively.

An α-stable CMGARCH model does not impose any explicit functional form of the conditional dispersion matrix Σ_t. The next theorem shows which α-stable MGARCH processes can be modeled by α-stable CMGARCH processes.

Theorem 5. *Let $(X_t)_{t\in\mathbf{Z}}$ be an α-stable multivariate GARCH process which is power-GARCH-projection-efficient with conditional dispersion time series $(\Sigma_t)_{t\in\mathbf{Z}}$, then this time series can be modeled by an α-stable composed MGARCH process.*

Proof.

$$\begin{aligned}
\sigma_{t,ij} &= \frac{1}{4}((e_i + e_j)\Sigma_t(e_i - e_j) - (e_i - e_j)\Sigma_t(e_i - e_j)) \\
&= \frac{1}{4}(\mathrm{Disp}((e_i + e_j)'X_t|\mathcal{F}_{t-1}) - \mathrm{Disp}((e_i - e_j)'X_t|\mathcal{F}_{t-1})) \\
&\overset{(*)}{=} \frac{1}{4}(\mathrm{Disp}((e_i + e_j)'X_t|\mathcal{F}_{t-1}(e_i + e_j)) \\
&\quad - \mathrm{Disp}((e_i - e_j)'X_t|\mathcal{F}_{t-1}(e_i - e_j))) \\
&\overset{\text{equation (15)}}{=} \frac{1}{4}(2\sigma_t^2(e_i + e_j) - 2\sigma_t^2(e_i - e_j)) \\
&= \frac{1}{2}(\sigma_t^2(e_i + e_j) - \sigma_t^2(e_i - e_j))
\end{aligned}$$

equation $(*)$ holds because the process is α-stable-projection-efficient. $\sigma_t^2(e_i+e_j)$ and $\sigma_t^2(e_i - e_j)$ are modeled by power-GARCH-processes. ☐

Unfortunately, we cannot rephrase Theorem 3 for an α-stable CMGARCH process. This is because of the fact that we do not know the unconditional distribution of X_t, $t \in \mathbf{Z}$, so we cannot ensure if the dispersion operator is well defined. But it seems reasonable to impose that the univariate $S_{\alpha,\beta,\delta}$GARCH processes of an α-stable CM-GARCH model should be strictly stationary, i.e., $\lambda_{\alpha,\beta,\delta}\sum_{i=1}^p \alpha_i + \sum_{j=1}^q \beta_j \leq 1$.

We show that an α-stable CMGARCH model is invariant under linear transformation. This is essential, since this result enables us to define consistently an α-stable factor composed MGARCH model.

Theorem 6. *Let $(X_t)_{t\in\mathbf{Z}}$ follow an α-stable GARCH-projection-efficient CMGARCH process. Then the CMGARCH process is invariant under linear transformation, i.e. the process $(Y_t)_{t\in\mathbf{Z}} = (FX_t)_{t\in\mathbf{Z}}$ follows an α-stable CMGARCH process in terms of the filtration $(\mathcal{G})_{t\in\mathbf{Z}}$, and we have*

$$\mathrm{Disp}(Y_t|\mathcal{F}_{t-1}) = F\Sigma_t F' = \mathrm{Disp}(Y_t|\mathcal{G}_{t-1}) \tag{17}$$

where $F \in \mathbf{R}^{k\times d}$, $k \in \mathbf{N}$, and $\mathcal{G}_t$ is the sigma field generated by $\sigma(\{Y_s : s \le t\})$.

Proof. We show $\mathrm{Disp}(FX_t|\mathcal{F}_{t-1}) = F\Sigma_t F' = \mathrm{Disp}(FX_t|\mathcal{G}_{t-1})$. We know

$$\mathrm{Disp}(a'X_t|\mathcal{F}_{t-1}) \;=\; \mathrm{Disp}(a'FX_t|\mathcal{F}_{t-1}(a'F)) = a'F\Sigma_t F'a. \tag{18}$$

for all $a \in R^d$. Hence, by using the characteristic function, we conclude

$$\begin{aligned}\psi_{sub}(x^2(a'F\Sigma_t F'a),\alpha) &= E(e^{ix(a'FX_t)|\mathcal{F}_{t-1}}) \\ &= E(e^{ix(a'FX_t)}|\mathcal{F}_{t-1}(a'F))\end{aligned}$$

for all $a \in \mathbf{R}^d$ and $x \in \mathbf{R}$. Since we have $\mathcal{F}_t(a'F) \subset \mathcal{G}_t \subset \mathcal{F}_t$ for all $a \in \mathbf{R}^d$ and $t \in \mathbf{Z}$, we obtain

$$\psi(x^2(a'F\Sigma_t F'a),\alpha) \;=\; E(e^{ix(a'FX_t)}|\mathcal{G}_{t-1}).$$

Since $a \in \mathbf{R}^d$ and $x \in \mathbf{R}^d$ are arbitrary, we follow

$$\psi_{sub}(s'F\Sigma_t F's,\alpha) \;=\; E(e^{is'(FX_t)}|\mathcal{G}_{t-1})$$

for all $s \in \mathbf{R}^d$. Hence, we can conclude

$$FX_t|\mathcal{G}_{t-1} \sim E_k(0, F\Sigma_t F', \psi_{sub}(.,\alpha))$$

and we have $\mathrm{Disp}(FX_t|\mathcal{F}_{t-1}) = F\Sigma_t F' = \mathrm{Disp}(FX_t|\mathcal{G}_{t-1})$. We write for $F\Sigma_t F'$ shortly Σ_t^Y. Since $FX_t|\mathcal{G}_{t-1}$ is sub-Gaussian, we can write

$$\begin{aligned}\sigma^Y_{t,ij} &= \frac{1}{4}(\mathrm{Disp}((e_i+e_j)'FX_t|\mathcal{G}_{t-1}) - \mathrm{Disp}((e_i-e_j)'FX_t|\mathcal{G}_{t-1}) \\ &= \frac{1}{2}(\sigma_t^2((e_i+e_j)'F) - \sigma_t^2((e_i-e_j)'F),\end{aligned}$$

where $(\sigma_t((e_i+e_j)'F))_{t\in\mathbf{Z}}$ and $(\sigma_t((e_i-e_j)'F))_{t\in\mathbf{Z}}$ follow $S_{\alpha,0,\delta}GARCH(p^+_{ij}, q^+_{ij})$ and $S_{\alpha,0,\delta}GARCH(p^-_{ij}, q^-_{ij})$. Hence, we have demonstrated that $(Y_t)_{t\in\mathbf{Z}}$ follows an α-stable CMGARCH process. □

We can consistently define an extension of the α-stable composed multivariate GARCH model, what we label the α-stable factor composed multivariate GARCH model.

Definition 26. *The process $(X_t)_{t\in\mathbf{Z}}$ follows an α-stable factor composed MGARCH (α-FCMGARCH) process, if there exists some orthogonal matrix $\Gamma \in \mathbf{R}^{d\times d}$ satisfying $\Gamma\Gamma' = \mathrm{Id}$ such that $(\Gamma' X_t)_{t\in\mathbf{Z}}$ follows an α-stable composed MGARCH process.*

As with the FCMGARCH model, the α-stable version allows for statistical factor modeling and dimensionality reduction. Again, we estimate the sample dispersion matrix of the process $(X_t)_{t\in\mathbf{Z}}$ and model its principal components $(Y_t)_{t\in\mathbf{Z}}$ by an α-stable FCMGARCH model. In contrast to a α-stable version of the PC-GARCH model, we have the advantage that we can capture the conditional dependence of the components.

As with Proposition 2 in Section 3, we have to estimate $(p+q+1)d^2$ parameters for a d-dimensional α-stable CMGARCH process whose projective time series follow α-stable power-GARCH(p,q) processes.

5.3 Estimation of the Models

In principle, we can employ the same two estimation procedures as presented in Section 4.2. In the first step, we have to use algorithms fitting α-stable power-GARCH processes to data in both algorithms.

In the first algorithm we reconstruct the conditional dispersion matrix by

$$\begin{aligned}\hat{\Sigma}_{t,ii} &= 2\hat{\sigma}_t^2(e_i),\\ \hat{\Sigma}_{t,ij} &= \frac{1}{2}(\hat{\sigma}_t^2(e_i+e_j)+\hat{\sigma}_t^2(e_i-e_j)).\end{aligned}$$

In the second approach we apply the univariate α-stable process $(\sigma_t(u_i))_{t\in\mathbf{Z}}$, $u_i \in \mathcal{S}^{d-1}$, $i=1,...,n$ to reconstruct the conditional dispersion. This leads to the following optimization problem.

$$\hat{\Sigma}_t = \text{argmin}_{\Sigma\in S^{d\times d}} \sum_{i=1}^{d} (u_i'\Sigma u_i - 2\hat{\sigma}_t^2(u_i))^2$$

in step (2) and

$$\hat{\Sigma}_t = \text{argmin}_{\Delta\in D_d} \sum_{i=1}^{n} (u_i'\Delta\Delta' u_i - 2\hat{\sigma}_t^2(u_i))^2$$

in step (3).

When fitting α-stable power-GARCH processes to the projective time series $(u_i'X_t)_{t\in\mathbf{Z}}$, $i=1,...,n$ via the standard maximum likelihood method, we need a global tail parameter α in order to be consistent with the model specification. The simplest way to obtain such an α is to estimate the unconditional tail parameter $\alpha(u_i)$ of the projective time series $(u_i'X_t)_{t\in\mathbf{Z}}$, $i=1,...,n$. Then, the global tail parameter α is defined by

$$\alpha = \frac{1}{n}\sum_{i=1}^{n}\alpha(u_i).$$

This is certainly a very heuristic method, since we do not estimate the global tail parameter of the innovations Z_t but of the returns X_t. But since we estimate the parameter $\alpha_0(u_i),...,\alpha_r(u_i)$ and $\beta_1(u_i),...,\beta_s(u_i)$ of the power-GARCH(r,s)-process $(u_i'X_t)_{t\in\mathbf{Z}}$ via the classical ML-method, these estimates are robust under misspecification of the tail parameter α. For larger sample sizes, $u_i'X_1, ..., u_i'X_{t_0}$ (t_0 large) the estimates for the scale parameters and the tail parameter are nearly independent (see DuMouchel (1973) for further information). Thus the estimates of the power-GARCH parameters are nearly independent of the tail parameter α and hence the time series of scale parameters $(\sigma_t(u_i))_{t\in\mathbf{Z}}$.

In addition, after having estimated the time series of conditional dispersion matrices $(\hat{\Sigma}_t)_{t\in\mathbf{Z}}$ we can use the residuals

$$\hat{Z}_t = \hat{\Sigma}_t^{-1/2} X_t$$

to estimate the tail parameter of the innovations where $\hat{\Sigma}_t^{1/2}$ is the inverse of the Cholesky factor of $\hat{\Sigma}_t$.

6 Applications

For the empirical analysis of the α-stable CMGARCH model, we investigated the daily logarithmic return series for four German stocks included in the DAX index: Adidas, Allianz, Altana, and BASF. The period analyzed is January 2, 2001 through March 31, 2006 (1,338 daily observations for each stock). For the estimation of our model, we selected the first $1,000$ returns i.e. the period form January 2, 2001 until December 7, 2004. The balance of the observed returns are held out for an out-of-sample analysis of the model.

The plots of the individual returns series in the estimation period for the four stocks are shown in Figure 1. One can easily detect times of intense and less pronounced volatility which is to be attributed to the well-known effect of volatility clustering. In Table 2 the maximum likelihood estimates of the four return time series are listed. The tail parameters α appear to be in a tight range around 1.69. Hence, assuming the same parameter α is justifiable. The scale parameter is within in a range of roughly 0.1 and 0.15. The location parameters μ are close to zero.

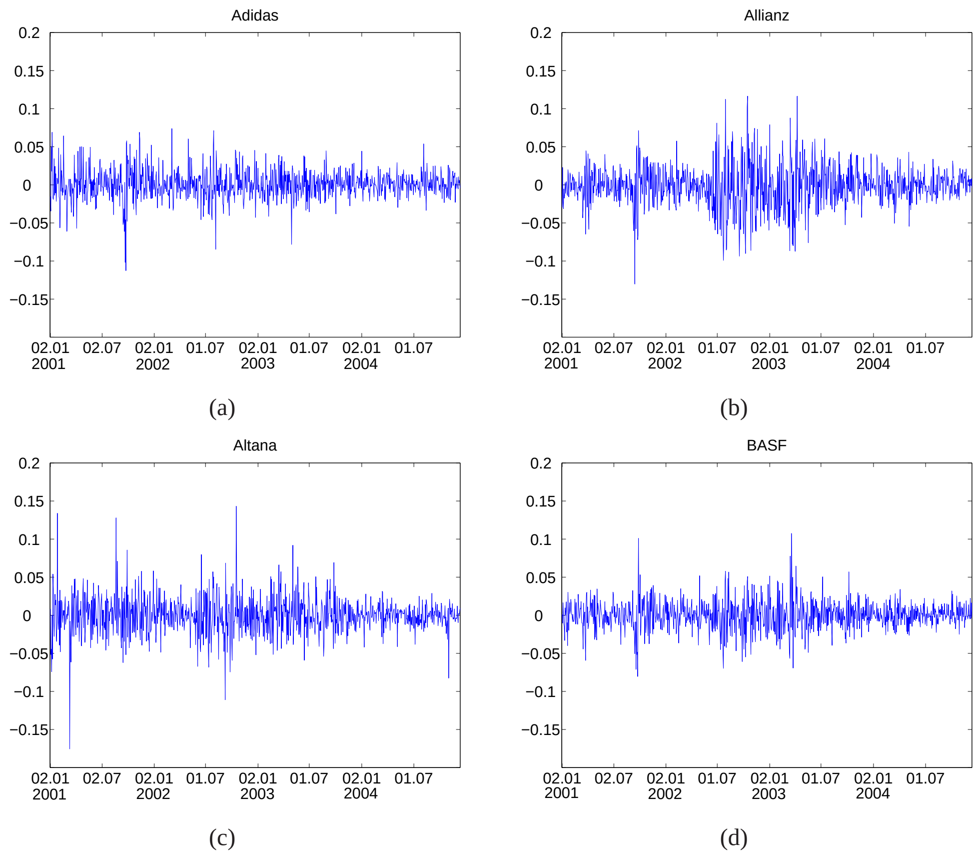


Figure 1: 1,000 daily returns during the period from January 2, 2001 until December 7, 2004.

We assume that the four univariate return series follow $S_{\alpha,0,1}GARCH(1,1)$ processes. More precisely,

$$\sigma_t(e_i) = \alpha_0(e_i) + \alpha_1(e_i)|X_{t-1} - \mu(e_i)| + \beta_1(e_i)\sigma_{t-1}(e_i)$$

and

$$X_{t,i} = \mu(e_i) + \sigma_t(e_i)Z_t,$$

where $Z_t \sim S_\alpha(1,0,0)$ and $\mu(e_i)$ is the unconditional mean of the ith time series, $i = 1, ..., 4$ and $\alpha = 1.69$.[11] The estimated parameters based on the period from January 2, 2007 until December 7, 2007 are reported in the right half of Table 2. As discussed in Section 5.3, a misspecification of the tail parameter α has only a minor influence on the estimated power-GARCH parameters.

[11] For daily log-returns it is not necessary to model the daily mean process $(\mu_t)_{t\in\mathbf{Z}}$ (see RiskMetrics (1996)).

Stock	α	σ	μ	α_0	α_1	β_1	LL
Adidas	1.667	0.0109	0.0001	0.0002	0.0413	0.9316	2617.30
Allianz	1.6672	0.0155	-0.0011	0.0002	0.0841	0.8854	2356.2
Altana	1.7337	0.0138	-0.0003	0.0004	0.0706	0.8837	2423.4
BASF	1.6992	0.011	0.0001	0.0003	0.0661	0.8966	2650.6

Table 2: The left half of the table depicts the unconditional stable estimates of the returns time series. The right half shows the estimated parameters of the univariate stable GARCH(1,1) processes for the returns. The time period is January 2, 2001 to December 7, 2004.

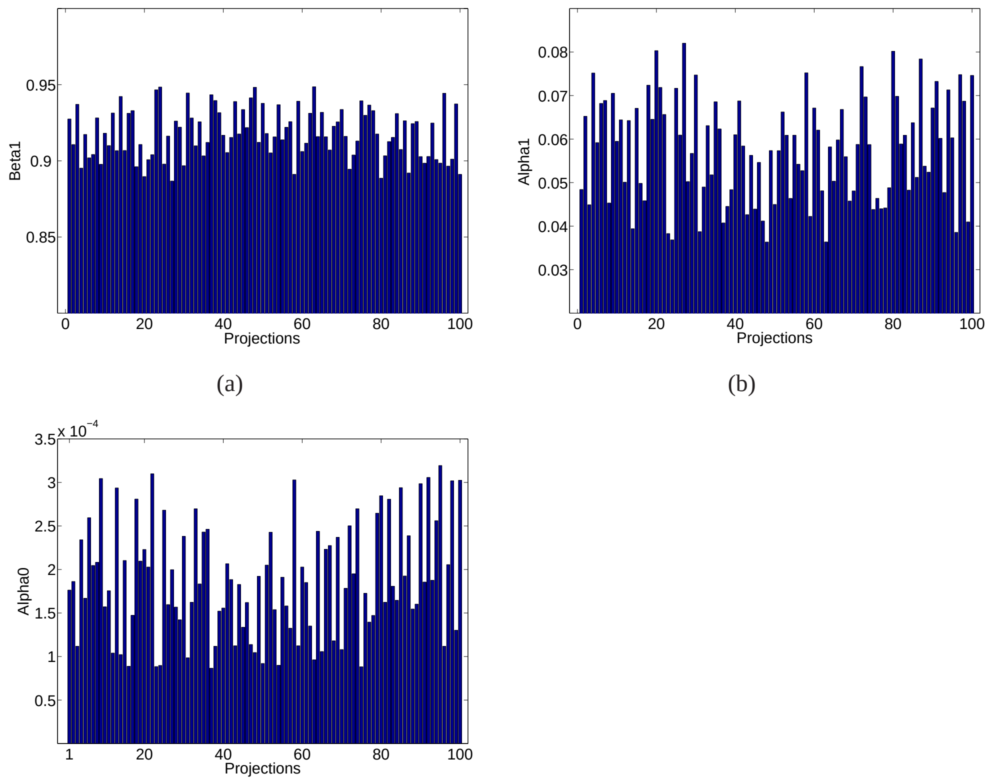
The absolute mean $\lambda_{\alpha,0}$ of a centered unit scale variable $Z_t \sim S_\alpha(1,0,0)$ is given by

$$\lambda_{\alpha,0} = \frac{2}{\pi}\Gamma\left(1 - \frac{1}{\alpha}\right) = 1.38.$$

Due to Proposition 3, all processes are strict stationary since we have $\alpha_0(e_i)$, $\alpha_1(e_i)$, $\beta_1(e_i) > 0$, and $\lambda_{\alpha,0}\alpha_1(e_i) + \beta_1(e_i) < 1$, $i = 1, ..., 4$. This purely univariate analysis of the return data does not reveal any contradiction to an α-stable CMGARCH modeling.

6.1 In-Sample Analysis of the α-stable CMGARCH Model

In the following we assume that the return data $X_1, ..., X_{1000}$ of the four stocks in our study are power-GARCH-projection-efficient with $\delta = 1$ and follow an α-stable CMGARCH model. In order to estimate the α-stable CMGARCH process we generate random vectors $u_i \in \mathcal{S}^3$, $i = 1, ..., 100$, that are uniformly distributed on $\mathcal{S}^3$. The projective time series $(u_i' X_t)_{t\in\mathbf{Z}}$ follow again an α-stable power-GARCH process. According to the estimation procedure described in Sections 4.2 and 5.3, we have to estimate the parameters $\alpha_0(u_i), \alpha_1(u_i)$, and $\beta_1(u_i)$ of the corresponding power-GARCH$(1,1)$-processes. Figure 2 (a), (b), and (c) illustrate the estimates. The parameter estimates $\hat{\beta}_1(u_i)$ are tightly scattered around 0.91; $\hat{\alpha}_1(u_i)$ ranges from 0.038 to 0.082 and for $\hat{\alpha}_0(u_i)$ from $0.7 \cdot 10^{-4}$ to $3 \cdot 10^{-4}$. All projective time series $(u_i' X_t)_{t\in\mathbf{Z}}$ have a low market reaction ($\hat{\alpha}_1(u_i)$ small) but a high persistence ($\hat{\beta}_1(u_i)$ high).



(a) (b)

(c)

Figure 2: (a),(b),(c) show the stable GARCH(1,1) estimates for the 100 projective time series $(u_i' X_t)_{t \in I}$.

In particular, we see that each projective time series is strict stationary. Figure 4 depicts the 100 corresponding time series $(\sigma_t(u_i))$ of the conditional, time-varying scale parameters. The effects of volatility clustering can be seen.

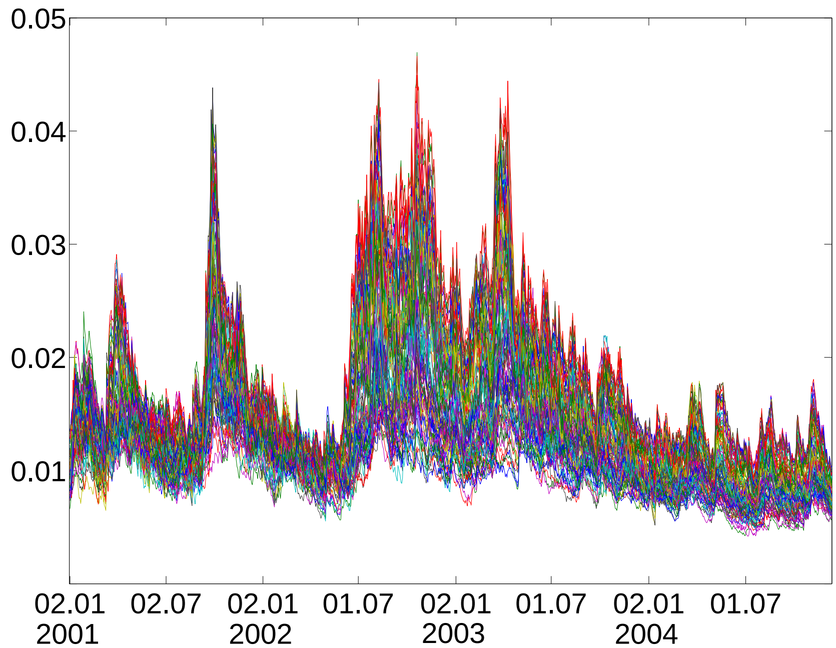


Figure 3: The figure shows the 100 different time series $(\sigma_t(u_i))_{t\in I}$, $i = 1, ..., 100$, of the stable GARCH$(1, 1)$-Processes.

In order to obtain the time series of the conditional dispersion matrices $\Sigma_t \in \mathbf{R}^{4\times4}$, we have to apply steps (2) and (3) of the estimation algorithm given in Sections 4.2 and 5.3, respectively. Figure 2 (d) shows the number of not positive definite matrices Σ_t, $t = 1, ..., 1000$, obtained after applying step (2) subject to the number of projective time series $(\sigma_t(u_i))_{t\in\mathbf{Z}}$ used in the regression.

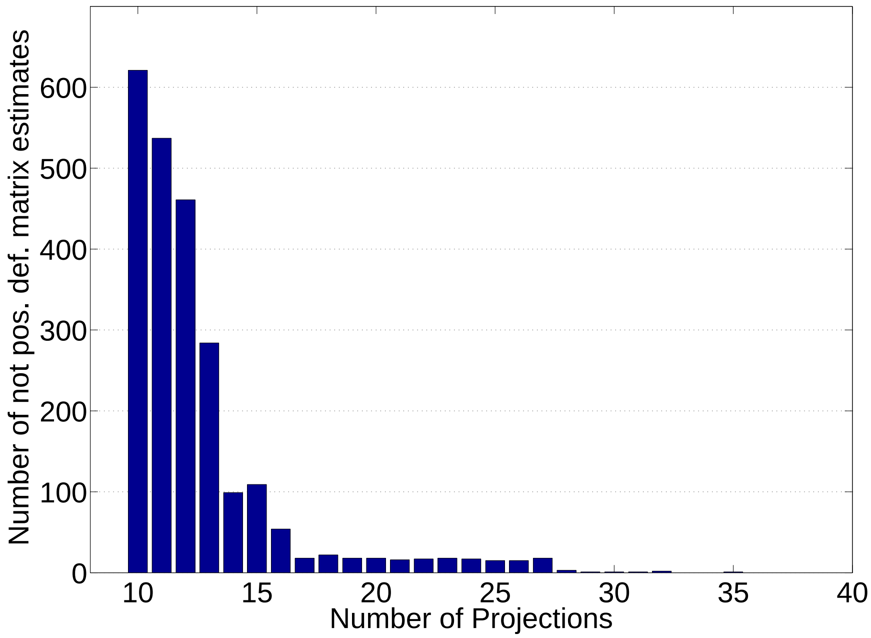


Figure 4: The figure illustrates the number of not positive definite matrices subject to the number of projection used in the regression.

By increasing the number of projections in step (2), the number of these matrices that exhibit this characteristic decreases fast. We notice the last not positive definite matrix when applying 35 projections. In the range from 36 until 100 projections, all matrices are positive definite. Furthermore, we observed a fast stabilization of the entries of the time series (Σ_t) subject to the number of projections used in step (2). This observation supports the assumption that the considered multivariate time series is GARCH-projection efficient.

In order to obtain a very high accuracy of our estimates, we use 100 projections for the reconstruction of the time series (Σ_t). In particular, we do not need to apply the optional step (3) since all conditional dispersion matrices are positive definite.

Figure 5 shows the 2-dimensional scatterplots between the different returns pairs. We find that Adidas and Altana as well as Altana and BASF exhibit a very low cross dispersion due to the scatterplots, while the ones of BASF-Allianz and Adidas-BASF illustrate stronger cross dispersion. Figure 6 depicts the time series of the conditional dispersion matrices $(\Sigma_t)_{t=1,\ldots,1000}$ for the period January 2, 2001 to December 7, 2004.

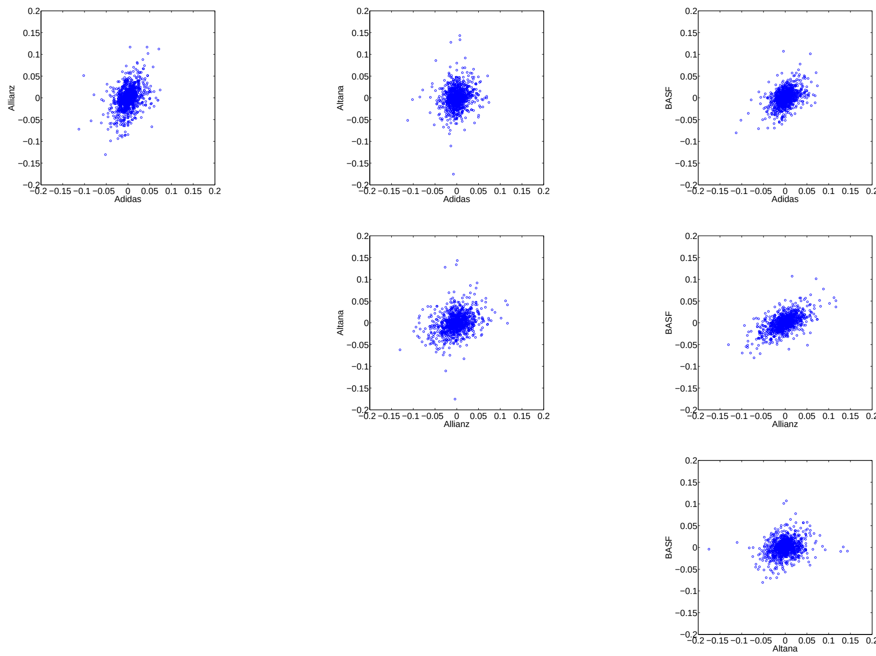


Figure 5: Two dimensional scatterplots of the returns in the period January 2, 2001 until December 7, 2004.

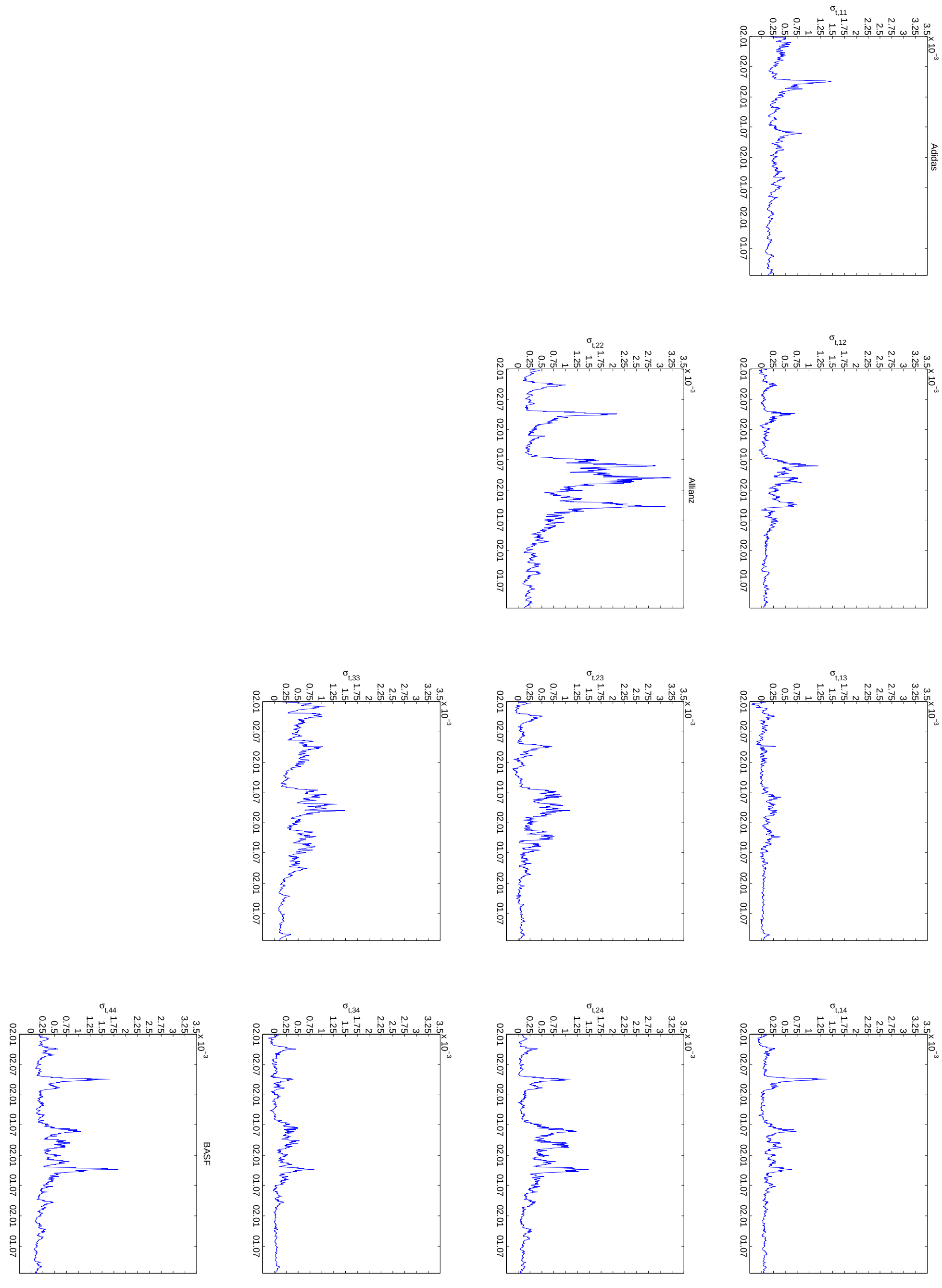


Figure 6: The figure depicts the estimated conditional dispersion matrices Σ_t in the period January 2, 2001 until December 7, 2004.

In particular, the conditional dispersions $(\sigma_{t,13})$ and $(\sigma_{t,34})$ corresponding to Adidas-Altana and Altana-BASF are low, consistent with the observation made about Figure 5. Definitely the highest conditional dispersion can be observed between Allianz and BASF, $(\sigma_{t,34})$, which is also the case in the unconditional graphical consideration of Figure 5. For the time period investigated, the returns of Allianz are the most volatile especially in the period July 2002 until July 2003. Besides Altana, we observe an increase in the conditional dispersion and cross dispersion in all stocks because of September 11, 2001.

For a quantitative analysis of the α-stable CMGARCH model we examine its residuals given by

$$\hat{Z}_t = \hat{\Sigma}_t^{-1/2}(X_t - \hat{\mu}),$$

where $t = 1, ..., 1000$, $\hat{\mu} \in \mathbf{R}^4$ the unconditional mean of the four individual return time series, and $\hat{\Sigma}_t^{-1/2}$ the inverse of the Cholesky factor of $\hat{\Sigma}_t$. To test whether the generated innovations are strict white noise (SNW(0,Id)) (see Section 3 and Definition 21), we estimate their unconditional dispersion matrix $\hat{\Sigma}_Z$ by using the spectral estimator.[12] The spectral estimator is a robust estimator of the dispersion and covariance matrix up to a scaling constant. To have a unique dispersion matrix we demand $\hat{\sigma}_{11}$ to be 1.

For comparison, we first list the normalized dispersion matrix of the original returns given by

$$\hat{\Sigma}_0(X_1, ..., X_{1000}) = \begin{pmatrix} 1.0000 & 0.5294 & 0.1929 & 0.3724 \\ 0.5294 & 1.8865 & 0.5637 & 0.8195 \\ 0.1929 & 0.5637 & 1.4942 & 0.3121 \\ 0.3724 & 0.8195 & 0.3121 & 0.9499 \end{pmatrix}.$$

One can clearly detect the non-zero cross dispersion between the stock returns, because $\hat{\sigma}_{ij}(X_1, ..., X_{1000})$, $i, j = 1, ..., 4, i \neq j$, deviate significantly from zero. Moreover, the diagonal entries reveal non-standardized quantities, since $\hat{\sigma}_{ii}(X_1, ..., X_{1000})$, $i = 1, ..., 4$, differ significantly from 1. However, the residuals $\hat{Z}_1, ..., \hat{Z}_{1000}$ of the α-stable CMGARCH model yield the unconditional normalized dispersion matrix

$$\hat{\Sigma}_0(\hat{Z}_1, ..., \hat{Z}_{1000}) = \begin{pmatrix} 1.0000 & 0.0381 & -0.0223 & 0.0046 \\ 0.0381 & 1.0392 & 0.0305 & 0.0440 \\ -0.0223 & 0.0305 & 1.1323 & 0.0146 \\ 0.0046 & 0.0440 & 0.0146 & 1.0991 \end{pmatrix},$$

which indicates that the scales of the individual stocks are close to one and zero, respectively. In addition, the cross dispersion between the returns are definitely minimized if not removed.

We compare these residuals with those of the EWMA model introduced by RiskMetrics.[13] The EWMA model is still the industry standard for multivariate conditional risk modeling (see the discussion in Section 3 and RiskMetrics (1996)). In the EWMA updating scheme

$$\Sigma_t = \frac{1 - \lambda}{1 - \lambda^{M+1}} \sum_{i=1}^{M} \lambda^i X_{t-i} X'_{t-i}$$

[12] See Tyler (1987a) for further information about the spectral estimator.
[13] See Definition 12.

we choose M to be $112, 50$, and 20. For daily returns, RiskMetrics (1996) recommends an optimal decay factor $\lambda = 0.94$. By using $\lambda = 0.94$, 99.9% of the information is contained in the last 112 days and the classical RiskMetrics updating scheme

$$\Sigma_t = (1-\lambda)X_{t-1}X'_{t-1} + \lambda\Sigma_{t-1},$$

is captured very well. In this case, the residuals are denoted by $\hat{Z}_1^{E(112)}, ..., \hat{Z}_{1000}^{E(112)}$ and their unconditional normalized dispersion matrix satisfies

$$\hat{\Sigma}_0(\hat{Z}_1^{E(112)}, ..., \hat{Z}_{1000}^{E(112)}) = \begin{pmatrix} 1.0000 & 0.1562 & -0.2340 & -0.2519 \\ 0.1562 & 1.1071 & -0.1848 & -0.2198 \\ -0.2340 & -0.1848 & 1.5471 & 0.6278 \\ -0.2519 & -0.2198 & 0.6278 & 1.8321 \end{pmatrix}.$$

We see that the diagonal elements of this matrix deviate significantly from one and cross-dispersion is definitely not zero. It is obvious that the results of the α-stable CMGARCH are superior. Incorporating the last 112 might be too much.

In our statistical analysis we test different values of M (i.e., $M = 10, 20, ..., 100$). We obtain the best results for $M = 20$. In this case, the dispersion matrix of the residuals satisfies

$$\hat{\Sigma}_0(\hat{Z}_1^{E(20)}, ..., \hat{Z}_{1000}^{E(20)}) = \begin{pmatrix} 1.0000 & 0.0408 & -0.0162 & 0.0335 \\ 0.0408 & 1.2124 & -0.0036 & -0.0287 \\ -0.0162 & -0.0036 & 1.2769 & 0.0647 \\ 0.0335 & -0.0287 & 0.0647 & 1.5789 \end{pmatrix}.$$

The cross dispersions are similar to those in the α-stable CMGARCH model but the diagonal elements of the normalized dispersion matrix differ significantly from one. This behavior is expected since we know from univariate GARCH- and EWMA-modeling that volatility processes of univariate return series are captured much better by a GARCH process. Moreover, we have one decay factor $\lambda = 0.94$ that should be valid for all stocks simultaneously, which is not realistic. In the case of $M = 50$, we obtain

$$\hat{\Sigma}_0(\hat{Z}_1^{E(50)}, ..., \hat{Z}_{1000}^{E(50)}) = \begin{pmatrix} 1.0000 & 0.0680 & -0.0406 & -0.0480 \\ 0.0680 & 1.1271 & 0.0404 & -0.0197 \\ -0.0406 & 0.0404 & 1.1219 & 0.0771 \\ -0.0480 & -0.0197 & 0.0771 & 1.7043 \end{pmatrix},$$

which does not significantly differ from the case $M = 20$.

The maximum likelihood estimates of the residuals $\hat{Z}_1, ..., \hat{Z}_{1000}$ are depicted in Table 3. We see that the innovations have a larger tail parameter than the unconditional returns depicted in Table 2. This phenomenon is to be expected because the leptokurtosis in the unconditional distribution of the process $(X_t)_{t\in\mathbf{Z}}$ is attributed to the GARCH structure of the process (see RiskMetrics (1996) and McNeil, Frey, and Embrechts (2005) for further information). By removing the MGARCH effects from the data, we decrease the leptokurtosis and thereby increase the tail parameter.

Stock	α	Disp	σ
Adidas	1.76	0.94	0.68
Allianz	1.92	1.12	0.75
Altana	1.8	1.11	0.74
BASF	1.9	1.18	0.77

Table 3: The table depicts the stable estimates of the α-stable CMGARCH residuals.

In particular, $\hat{Z}_{t,1}, ..., \hat{Z}_{t,1000}$ has a scale parameter of 0.9440, thus the residuals have an estimated dispersion matrix satisfying

$$\hat{\Sigma}(\hat{Z}_1, ..., \hat{Z}_{1000}) = \begin{pmatrix} 0.9440 & 0.0359 & -0.0211 & 0.0044 \\ 0.0359 & 0.9810 & 0.0288 & 0.0416 \\ -0.0211 & 0.0288 & 1.0689 & 0.0138 \\ 0.0044 & 0.0416 & 0.0138 & 1.0375 \end{pmatrix},$$

that is close the identity. However, there is a difference between the estimates $\hat{\sigma}_{ii}(Z)$ of the spectral estimator and those of the univariate estimates depicted in Table 3. We have greater confidence in the spectral estimates for two reasons. First, the spectral estimator uses a four dimensional sample for its estimates whereas in the other case we use only the univariate time series. Second, the spectral estimator estimates $\hat{\sigma}_{ii}$, $i = 1, ..., 4$, immediately, whereas in the other method we estimate the scale parameter $\hat{\sigma}(e_i)$ and then we calculate the dispersion by the formula $\hat{\sigma}_{ii} = 2\hat{\sigma}(e_i)^2$ increasing the estimation error.

In order to complete our sample analysis, we analyze the autocorrelation of the squared returns and squared residuals in the different models. While one might correctly object that due to the model specification second moments of the residuals and returns do not exist, nevertheless the estimators of the autocorrelation function (ACF) have distributions that have lower and upper confidence bounds under the independent and identically distributed (i.i.d.) assumption. The range of these bounds is larger than the ones in the case where second moments exist. Hence, if the residuals are within the normal bounds, they are also in model specific confidence bounds. The results are reported in Table 4. The lower and upper 95% confidence bounds are -0.0632 and 0.0632, respectively. The squared returns significantly violate these bounds (29 violations) and the hypothesis of independence can be rejected. In the case of the α-stable CMGARCH residuals we find three violations of these bounds. In the classical EWMA model ($M = 112$) there are eight violations. The EWMA model using the 20 last observations only violates these bounds twice and for $M = 50$ we observe five values out of these bounds. All the models work well and remove a lot of autocorrelation in the squared return data. The α-stable CMGARCH model seems to be superior to the classical EWMA model since the estimates are closer to zero in most cases and there are less violations of the confidence bounds.

Stock	1	2	3	4	5	6	7
$ACF((X_{t,1}^2))$	0.2	0.305	0.119	0.188	0.159	0.116	0.083
$ACF((X_{t,2}^2))$	0.236	0.232	0.23	0.23	0.353	0.19	0.233
$ACF((X_{t,3}^2))$	0.046	0.016	0.082	0.012	0.028	0.004	0.021
$ACF((X_{t,4}^2))$	0.176	0.246	0.263	0.355	0.133	0.2	0.13
$ACF((X_{t,1}^2))$	0.2	0.305	0.119	0.188	0.159	0.116	0.083
$ACF((\hat{Z}_{t,1}^2))$	0.028	0.054	0.004	0.046	0.041	0.04	-0.03
$ACF((\hat{Z}_{t,2}^2))$	0.101	0.045	0.073	0.059	0.081	-0.006	-0.015
$ACF((\hat{Z}_{t,3}^2))$	0.002	-0.009	0.033	-0.015	-0.009	-0.018	-0.021
$ACF((\hat{Z}_{t,4}^2))$	-0.009	0.01	0.042	-0.003	0.083	0.001	-0.018
$ACF((\hat{Z}_{t,1}^{2,E(112)}))$	0.043	0.058	0.012	0.066	0.059	0.098	-0.004
$ACF((\hat{Z}_{t,2}^{2,E(112)}))$	0.155	0.067	0.137	0.12	0.14	0.03	0.023
$ACF((\hat{Z}_{t,3}^{2,E(112)}))$	0.009	-0.016	0.062	0.01	0	0.004	-0.007
$ACF((\hat{Z}_{t,4}^{2,E(112)}))$	0.049	0.024	0.095	0.032	0.072	0.037	-0.008
$ACF((\hat{Z}_{t,1}^{2,E(20)}))$	0.011	0.003	-0.018	0.013	0.018	0.044	-0.037
$ACF((\hat{Z}_{t,2}^{2,E(20)}))$	0.097	0.007	0.026	0.028	0.051	-0.014	0.01
$ACF((\hat{Z}_{t,3}^{2,E(20)}))$	-0.01	-0.012	0.005	-0.008	-0.01	-0.014	-0.015
$ACF((\hat{Z}_{t,4}^{2,E(20)}))$	0.038	-0.018	0.041	0.021	0.033	-0.016	-0.006
$ACF((\hat{Z}_{t,1}^{2,E(50)}))$	0.023	0.034	-0.001	0.05	0.038	0.082	-0.022
$ACF((\hat{Z}_{t,2}^{2,E(50)}))$	0.128	0.048	0.086	0.079	0.1	0.018	0.045
$ACF((\hat{Z}_{t,3}^{2,E(50)}))$	0.003	-0.024	0.054	0.006	-0.008	-0.003	-0.014
$ACF((\hat{Z}_{t,4}^{2,E(50)}))$	0.036	0.01	0.077	0.018	0.059	0.027	-0.014

Table 4: The table depicts the autocorrelation of the squared returns and the squared residuals in the different models. The lower and upper 95% confidence bounds are -0.0632 and 0.0632, respectively. The period covered is January 2, 2001 to December 7, 2004.

6.2 Out-Of-Sample Analysis of the α-stable CMGARCH Model

For the out-of-sample analysis we use the period from December 8, 2004 to March 31, 2006 (observations $1,001$ to $1,338$). The normalized dispersion matrix of the observed returns satisfies

$$\hat{\Sigma}_0(X_{1001}, ..., X_{1338}) = \begin{pmatrix} 1.0000 & 0.3756 & 0.0896 & 0.3487 \\ 0.3756 & 1.1663 & 0.2344 & 0.5109 \\ 0.0896 & 0.2344 & 0.7438 & 0.1429 \\ 0.3487 & 0.5109 & 0.1429 & 0.9180 \end{pmatrix}$$

and the α-stable CMGARCH residuals is given by

$$\hat{\Sigma}_0(\hat{Z}_{1001}, ..., \hat{Z}_{1338}) = \begin{pmatrix} 1.0000 & -0.0419 & 0.0954 & -0.0118 \\ -0.0419 & 1.0558 & -0.0354 & 0.0240 \\ 0.0954 & -0.0354 & 0.9994 & 0.0025 \\ -0.0118 & 0.0240 & 0.0025 & 1.1954 \end{pmatrix}.$$

The normalized dispersion matrix of the returns once again exhibit a significant cross dispersion. Furthermore, the diagonal entries of the normalized dispersion matrix are not close to one, suggesting that the univariate return series exhibit different scale properties. In contrast, the normalized dispersion matrix of the CMGARCH residuals are much closer to the identity matrix. This means that the forecasted conditional dispersion matrix explains fairly well the common scaling properties of the returns.

Again, comparing the model with the classical EWMA model ($M = 112$) we see that the cross dispersion is explained well by the model but exhibits weakness on the diagonal entries

$$\hat{\Sigma}_0(\hat{Z}_{1001}^{E(112)}, ..., \hat{Z}_{1338}^{E(112)}) = \begin{pmatrix} 1.0000 & 0.0335 & 0.0576 & 0.0555 \\ 0.0335 & 1.3439 & -0.0199 & -0.0647 \\ 0.0576 & -0.0199 & 0.8381 & 0.0503 \\ 0.0555 & -0.0647 & 0.0503 & 1.4686 \end{pmatrix}.$$

The two alternative EWMA-models ($M = 20, 50$) show similar behavior: The off-diagonal entries are close to zero while the entries on the diagonal exhibit poor behavior. In particular, the standardized dispersion matrices of these residuals satisfy

$$\hat{\Sigma}_0(\hat{Z}_{1001}^{E(20)}, ..., \hat{Z}_{1338}^{E(20)}) = \begin{pmatrix} 1.0000 & -0.0106 & 0.0980 & 0.0122 \\ -0.0106 & 1.3794 & -0.0501 & -0.0509 \\ 0.0980 & -0.0501 & 1.2486 & 0.0194 \\ 0.0122 & -0.0509 & 0.0194 & 1.8280 \end{pmatrix}$$

and

$$\hat{\Sigma}_0(\hat{Z}_{1001}^{E(50)}, ..., \hat{Z}_{1338}^{E(50)}) = \begin{pmatrix} 1.0000 & 0.0260 & 0.0671 & 0.0495 \\ 0.0260 & 1.3679 & -0.0200 & -0.0717 \\ 0.0671 & -0.0200 & 0.9335 & 0.0309 \\ 0.0495 & -0.0717 & 0.0309 & 1.5179 \end{pmatrix}.$$

As already mentioned, this is due to the well-known univariate phenomenon that the volatility structure of univariate return series is captured better by GARCH processes than EWMA processes.

Again, we consider the autocorrelation of the squared returns and squared residuals in the period of December 8, 2004 until March 31, 2006. The results are depicted in Table 5. The lower and upper 95% confidence bounds are -0.109 and 0.109. The observed squared returns do not exhibit significant autocorrelation. This might be explained by the low volatility in this period. We observe only three violations of the 95% confidence bounds. The out-of-sample residuals of the α-stable CMGARCH model violate the these bounds only once. In the classical EWMA model ($M = 112$), we have three autcorrelation estimates out of this range. In the EWMA model using only the last 20 observations, there are no the violation of these bounds and in the last model we observe one violation.

6.3 Summary of the Results

Summing up, the α-stable CMGARCH model outperforms the EWMA models because the normalized dispersion matrix of its residuals is closer to strict white noise than the ones in the other EWMA models. The EWMA models, in particular, reveal its weakness in estimating the diagonal entries. Furthermore, the autocorrelation in the squared return data is captured better by the α-stable CMGARCH model than by the classical EWMA model ($M = 112$) because we observe in the former one less violation of the confidence bounds and, in general, the estimates are closer to zero.

These observations hold for the in-sample as well as the out-of-sample analysis. The good empirical performance of the α-stable CMGARCH model is clear evidence for the GARCH-projection-efficiency of the return series $(X_t)_{t\in\mathbf{Z}}$ investigated.

Stock	1	2	3	4	5	6	7
$ACF((X_{t,1}^2))$	-0.006	-0.018	-0.012	-0.025	-0.03	-0.048	-0.043
$ACF((X_{t,2}^2))$	0.112	0.023	0.123	0.106	0.121	0.062	0.041
$ACF((X_{t,3}^2))$	-0.007	0.02	-0.003	-0.015	-0.012	-0.01	-0.01
$ACF((X_{t,4}^2))$	0.028	0.043	0.098	-0.083	0.015	-0.092	0.011
$ACF((\hat{Z}_{t,1}^2))$	-0.025	-0.009	-0.025	-0.026	-0.021	-0.04	-0.04
$ACF((\hat{Z}_{t,2}^2))$	-0.013	-0.041	0.007	-0.015	0.012	0.055	-0.015
$ACF((\hat{Z}_{t,3}^2))$	-0.009	-0.005	-0.006	-0.011	-0.007	-0.01	-0.011
$ACF((\hat{Z}_{t,4}^2))$	0.132	-0.077	-0.095	-0.061	-0.015	-0.043	-0.029
$ACF((\hat{Z}_{t,1}^{2,E(112)}))$	0.043	0.058	0.012	0.066	0.059	0.098	-0.004
$ACF((\hat{Z}_{t,2}^{2,E(112)}))$	0.155	0.067	0.137	0.12	0.14	0.03	0.023
$ACF((\hat{Z}_{t,3}^{2,E(112)}))$	0.009	-0.016	0.062	0.01	0	0.004	-0.007
$ACF((\hat{Z}_{t,4}^{2,E(112)}))$	0.049	0.024	0.095	0.032	0.072	0.037	-0.008
$ACF((\hat{Z}_{t,1}^{2,E(20)}))$	0.011	0.003	-0.018	0.013	0.018	0.044	-0.037
$ACF((\hat{Z}_{t,2}^{2,E(20)}))$	0.097	0.007	0.026	0.028	0.051	-0.014	0.01
$ACF((\hat{Z}_{t,3}^{2,E(20)}))$	-0.01	-0.012	0.005	-0.008	-0.01	-0.014	-0.015
$ACF((\hat{Z}_{t,4}^{2,E(20)}))$	0.038	-0.018	0.041	0.021	0.033	-0.016	-0.006
$ACF((\hat{Z}_{t,1}^{2,E(50)}))$	0.023	0.034	-0.001	0.05	0.038	0.082	-0.022
$ACF((\hat{Z}_{t,2}^{2,E(50)}))$	0.128	0.048	0.086	0.079	0.1	0.018	0.045
$ACF((\hat{Z}_{t,3}^{2,E(50)}))$	0.003	-0.024	0.054	0.006	-0.008	-0.003	-0.014
$ACF((\hat{Z}_{t,4}^{2,E(50)}))$	0.036	0.01	0.077	0.018	0.059	0.027	-0.014

Table 5: The table depicts the autocorrelation of the squared returns and the squared residuals for the different models. The lower and upper 95% confidence bounds are -0.109 and 0.109, respectively. The period covered is December 8, 2004 to March 31, 2006.

7 Conclusion

In this paper we introduce a new class of multivariate GARCH models that is flexible enough to model multivariate time series appropriately and allow for estimation procedures that are applicable even in higher dimensions. We motivate these models by introducing the notions of projection-efficient and GARCH-projection-efficient that are fundamental for the working of these models.

Moreover, in this paper we demonstrate that α-stable multivariate GARCH modeling is feasible. To do so, we develop α-stable versions of the CMGARCH and FCMGARCH model. We demonstrate the applicability of the model and report empirical evidence that indicates that it outperforms the classical EWMA model introduced by RiskMetrics.

References

Alexander, C.O. 2001. *Market Models: A Practitioner's Guide to Financial Data Analysis.* Wiley.

Alexander, C.O. 2002. *Principal Component Models for Generating Large GARCH Covariance Matrices.* Economic Notes, 31(2), pp. 337-359.

Attansio, O. 1991. *Risk, time-varying second moments and market efficiency.* Review of Economic Studies 58:479-494.

Bauwens, L., S. Laurent and J.V.K. Rombouts. 2006.*Multivariate GARCH models: A Survey*. J. Appl. Econ. 21: 79-109 (2006)

Bera, A.K.,M.L. Higgins 1993. *ARCH models:properties, estimation and testing*. Journal of Economic Surveys 7:305-362.

Billio M., M. Caporin and M. Gobbo 2003. *Block dynamic conditional correlation multivariate GARCH models* Greta Working Paper 03.03.

Bollerslev, T. 1986. *Generalized autoregressive conditional heteroscedasticity*. Journal of Econometrics 31: 307-327.

Bollerslev, T., F. Engle and J. M. Wooldrige *A capital asset pricing model with time-varying covariances.* Journal of Political Economy, 96 (1988), pp. 116-131.

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

Bollerslev, T., R.Y. Chou, K.F. Kroner. *ARCH modeling in finance: a review of the theory and empirical evidence*. Journal of Econometrics 52:5-59.

Christodoulakis G.A. and S.E. Satchell 2002. *Correlated ARCH: modeling the time-varying correlation between financial asset returns.* European Journal of Operations 139:351-370.

DuMouchel, W.H. 1973. *On the asymptotic normality of the Maximum-Likelihood Estimate when Sampling from a Stable Distribution.* Annals of Statistics, 1, 948-57.

Doganoglu, T., C. Hartz, S. Mittnik 2006. *Portfolio Optimization when Risk Factors are Conditionally Varying and Heavy Tailed.* Center for Financial Studies, University of Frankfurt.

Ding, Z. 1994. *Time Series Analysis of Speculative Returns.* PhD thesis, University of California, San Diego.

Engle, R. F. 1982. *Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation*. Econometrica 50:987-1007.

Engle, R. F. and K. F. Kroner 1995. *Multivariate simultaneous generalized arch.* Econometric Theory, 11 (1995), pp. 122-150.

Engle, R. F. and K. Sheppard 2001 *Theoretical and empirical properties of dynamic conditional correlation multivariate GARCH*. Mimeo, UCSD.

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

Fama, E. 1991. *Efficient Capital Markets:II.* Journal of Finance, Vol. 46, No. 5., pp. 1575-1617.

Hamilton, J.D. 1994 *Time Series Analysis.* Princetion University Press 1994.

Hill, B.M. 1975. *A simple general approach to inference about the tail of a distribution.* Annals of Statistics 3, 1163-1174.

Höchstötter, M., Fabozzi, F.J., and Rachev, S.T. 2005. *Distributional Analysis of the Stocks Comprising the DAX 30*. Probability and Mathematical Statistics, Vol. 25, 363-383.

Jondeau, E., M. Rockinger 2001. *The copula-GARCH model of conditional dependencies: an international stock-market application.* Journal of International Money and Finance.

Kring, S., S.T. Rachev, M. Höchstötter, F.J. Fabozzi 2007. *Estimation of α-Stable Sub-Gaussian Distributions for Asset Returns.* Risk Assessment: Decisions in Banking and Finance (eds. G. Bol, S.T. Rachev, and R. Würth), Physica Verlag Heidelberg 2007. (to appear)

McCulloch, J.H., 1985. *Interest-risk sensitive deposit insurance premia: stable ACH estimates.* Journal of Banking and Finance 9, 137-156.

McNeil, A. J., R. Frey and P. Embrechts. 2005. *Quantitative Risk Management*, Princeton University Press.

Mittnik, S., M.S. Paolella, and S.T. Rachev 1998. *A tail estimator for the index of the stable Paretian distribution.* Communications in Statistics–Theory and Methods 27, 1239-1262.

Mittnik, S., M.S. Paolella and S.T. Rachev. 2002 *Stationarity of stable power-GARCH processes* Journal of Econometrics 106 (2002) 97-107.

Nolan, J. P., A.K. Panorska, J.H. McCulloch 2001. *Estimation of stable spectral measure.* Mathematical and Computer Modelling 34, 1113-1122.

Nolan, J. P. 2005. *Multivariate stable densities and distribution functions: general and elliptical case*. Deutsche Bundesbank's 2005 Annual Fall Conference.

Rachev, S. T. and S. Mittnik 2000. *Stable Paretian Models in Finance*. Wiley.

Patton, A. 2000. *Modelling time-varying exchange rate dependence using the conditional copula.* University of California, San Diego, Discussion Paper 01-09.

Panorska, A., S. Mittnik, S.T. Rachev 1995. *Stable GARCH Models for Financial Time Series.* Applied Mathematics Letters 8, 33-37.

Pelletier, D. 2003. *Regime switching for dynamic correlations.* Journal of Econometrics.

RiskMetrics 1996.*Riskmetrics Technical Document*, 4th edn. J.P. Morgan: New York.

Shephard, N. 1996. *Statistical aspects of ARCH and stochastic volatility.* In Time Series Models in Econometrics, Finance and Other Fields, Hinkley D.V., Cox D.R., Barndorff-Nielsen O.E. (eds). Chapman & Hall: London.

Samorodnitsky, G. and M. Taqqu. 1994. *Stable Non-Gaussian Random Processes*, Chapmann & Hall

Stoyanov, S., V., B. Racheva-Iotova 2004. *Univariate stable laws in the fields of finance-approximations of density and distribution functions*. Journal of Concrete and Applicable Mathematics, 2/1, 38-57.

Tse, Y. K. and A.K.C. Tsui 2002. *A multivariate GARCH model with time-varying correlations.* Journal of Business and Economic Statistics 20:351-362.

Tsui, A.K. and Q. Yu 1999. *Constant conditional correlation in a bivariate garch model: Evidence from the stock market in china.* Mathematics and Computers in Simulation, 48 (1999), pp. 503-509.

Tyler, D.E. 1987a. *A distribution-free M-estimator of multivariate scatter*. The Annals of Statistics 15: pp. 234-251.