Dynamic Principal Component CAW Models for High-Dimensional Realized Covariance Matrices

Bastian Gribisch*
Institute of Econometrics and Statistics, University of Cologne, Germany

Michael Stollenwerk
Alfred-Weber-Institute for Economics, University of Heidelberg, Germany

(December 2, 2019)

Abstract

We propose a new dynamic principal component CAW model (DPC-CAW) for time-series of high-dimensional realized covariance matrices of asset returns (up to 100 assets). The model performs a spectral decomposition of the scale matrix of a central Wishart distribution and assumes independent dynamics for the principal components' variances and the eigenvector processes. A three-step estimation procedure makes the model applicable to high-dimensional covariance matrices. We analyze the finite sample properties of the estimation approach and provide an empirical application to realized covariance matrices for 100 assets. The DPC-CAW model has particularly good forecasting properties and outperforms its competitors for realized covariance matrices.

JEL classification: C32, C58, G17;
Keywords: Realized volatility, Covariance matrix, Spectral Decomposition, Time-Series Models.

1 Introduction

The modeling and forecasting of covariance matrices of asset returns is central to financial decision making since it provides a measurement of the risk involved in different investment allocations. It is

*Corresponding Author. Institute of Econometrics and Statistics, University of Cologne, Universitätsstr. 22a, D-50937 Cologne, Germany. Tel.: +49(0)2214707711; Fax: +49(0)2214705074. *E-mail address:* bastian.gribisch@statistik.uni-koeln.de

specifically used in option pricing, risk management and portfolio allocation.

Traditionally multivariate GARCH (MGARCH) or multivariate stochastic volatility (MSV) models have been applied in order to estimate conditional covariance matrices from daily asset return vectors (see e.g. Bauwens et al., 2006, and Asai et al., 2006, for surveys). Nowadays the increasing availability of intraday asset return information enables the computation of consistent ex-post measures of daily (co)variation of asset prices, so-called realized (co)variances (see e.g. Andersen et al., 2003, and Barndorff-Nielsen and Shephard, 2004). These realized measures can then be modeled directly in order to obtain forecasts of the covariance matrix of asset returns. The literature provides broad evidence that models for realized covariance matrices provide more precise forecasts then MGARCH and MSV models (see e.g. Golosnoy et al., 2012, and the references therein). Pioneering approaches are found in Gourieroux et al. (2009), Chiriac and Voev (2011), Bauer and Vorkink (2011), Noureldin et al. (2012) and Golosnoy et al. (2012).

These models have in common that applications to high-dimensional covariance matrices (say, for more than 10 assets) are complicated if not impossible and empirical applications typically do not exceed the 10-dimensional case.[1] Realistic portfolios however consist of a large number of assets which makes high-dimensional covariance matrix forecasting an important field of research. The development of models for high-dimensional applications is challenging, since the dimension of the object of interest is proportional to the square of the number of assets. This results in a huge number of model parameters and renders one-step maximum likelihood (ML) estimation virtually impossible (the so-called *curse of dimensionality*). An important task is therefore to develop multivariate volatility models which allow for feasible estimation in high-dimensional applications.

One strategy which has been proposed to overcome the curse of dimensionality is the use of sparsity assumptions like e.g. sparse factor structures for the assets' covariance matrix (see e.g. Wang and Zou, 2010, Tao et al., 2011, Shen et al., 2018, Sheppard and Xu, 2019, Asai and McAleer, 2015, Jin et al., 2019). An alternative is to design multivariate volatility models such that their parameters can be iteratively estimated by multistep procedures. In particular, Bauwens et al.

[1] In this paper we follow the convention of labeling covariance matrices of up to ten assets as "small dimensional" and covariance matrices of up to 100 assets as "high-dimensional". We are not concerned with "vast-dimensional" or "large-dimensional" covariance matrices with more than 100 assets (compare e.g. Lunde et al., 2016, Sheppard and Xu, 2019, and Engle et al, 2019, for similar conventions).

(2012) proposed the Realized DCC (Re-DCC) CAW model (see also Bauwens et al., 2014, and Bauwens et al., 2016, for applications and extensions), which resembles the DCC GARCH idea of Engle (2002) under the Conditional Autoregressive Wishart (CAW) setting of Golosnoy et al. (2012) for realized covariance matrices. The model is applicable in high-dimensional settings via three-step estimation with correlation targeting, similar to the corresponding MGARCH model. Bauwens et al. (2012) provide an empirical application for 50 assets. While the DCC idea builds on decomposing the conditional covariance matrix in variances and correlations, which are then estimated independently, an alternative strand of literature constructs orthogonal components via a spectral decomposition (SD) of the covariance matrix. The most prominent model here is the orthogonal GARCH (OGARCH) model of Alexander and Chibumba (1997) and Alexander (2001), where the estimation output can be readily interpreted in terms of (conditional) principal component analysis. Aielli and Caporin (2015) introduce additional flexibility via allowing for dynamic loading matrices. The resulting model is then called Dynamic Principal Component (DPC) GARCH model. Similar to the DCC approach, the framework assumes the presence of an auxiliary process generating orthonormal dynamic eigenvectors and allows for three-step estimation in order to be applicable in high-dimensional settings (the authors provide an application for up to 30 assets).

In this paper we adapt the DPC-GARCH model of Aielli and Caporin (2015) to the modeling of high-dimensional realized covariance matrices. The model structure is based on the CAW framework of Golosnoy et al. (2012) assuming a conditional central Wishart distribution for the realized covariance matrix. This particular distributional assumption allows for a convenient Quasi Maximum Likelihood (QML) interpretation implying consistency of one-step estimation even if the Wishart assumption is violated. We present a scalar version of the resulting DPC-CAW model and its estimation via a three-step approach similar to Aielli and Caporin (2015) in order to enable parameter estimation in high-dimensional settings. The three-step approach suffers from similar inconsistency problems as the DCC GARCH, the Re-DCC CAW and the DPC-GARCH model. We therefore conduct an extensive simulation experiment which shows that biases are present but mainly affect the unconditional variances of lower order principal components which are of minor relevance for covariance forecasting. An out-of-sample forecasting experiment for 100-dimensional realized covari-

ance matrices finally shows that the DPC-CAW model has good forecasting properties in 1-period, 5-period and 10-period ahead forecasting and outperforms its competitors in particular in forecasting the correlation structure and the weights of the Global Minimum Variance Portfolio (GMVP), which are of high relevance in practical portfolio optimization. In particular, the DPC-CAW approach features significantly lower correlation and GMVP losses compared to up-to-date competitors like the flexible Factor HEAVY approach of Sheppard and Xu (2019) and the Factor-CAW of Shen et al. (2018).

The rest of the paper is organized as follows. In section 2 we briefly review the concept of realized covariance measures. Section 3 introduces the scalar DPC-CAW model including one-step and three-step ML estimation. Section 4 presents the results of a simulation experiment analyzing the bias and consistency of estimates obtained via the three-step approach. The empirical application to realized covariance matrices for 100 NYSE traded stocks including in-sample diagnostics and an extensive out-of-sample forecasting experiment is presented in Section 5. Section 6 concludes.

2 Realized Covariance Measures

Consider an n-dimensional vector of log-prices $y(\tau)$, where $\tau \in \mathbb{R}_+$ represents continuous time. Assume that $y(\tau)$ is a Brownian stochastic semimartingale with $(n \times n)$ spot covariance matrix $\Theta(\tau)$. Without loss of generality restricting the trading day to the unit interval we obtain the 'true' integrated covariance matrix at day t as $\Sigma_t = \int_{t-1}^{t} \Theta(\tau)\, d\tau$.

Now assume that we observe $m+1$ uniformly spaced intraday log-prices. Then the j'th intraday return vector on day t $(t = 1, \ldots, T)$ is given by

$$r_{j,t} = y\big((t-1) + j/m\big) - y\big((t-1) + (j-1)/m\big), \quad j = 1, \ldots, m, \quad t = 1, \ldots, T. \tag{1}$$

Let the $(n \times n)$ matrix R_t denote a realized measure, i.e. a nonparametric ex-post estimate of Σ_t exploiting high-frequency asset return information. A well-known example for R_t is the realized

covariance matrix, which is defined as

$$RC_t = \sum_{j=1}^{m} r_{j,t} r'_{j,t}. \tag{2}$$

In the absence of market microstructure noise and discontinuous price jumps it can be shown that RC_t is a consistent estimator of Σ_t as $m \to \infty$ (see Barndorff-Nielsen and Shephard, 2004). If the observed intraday price data contains microstructure noise, jumps or non-synchronous trading one can employ one of several alternatives to the realized covariance matrix, such as the multivariate realized kernel of Barndorff-Nielsen et al. (2011).

3 The DPC-CAW Model

We model the time-evolution of n-dimensional stochastic positive-definite realized covariance measures $\{R_t\}_{t=1}^{T}$. Given the filtration $\mathcal{F}_{t-1} = \{R_{t-1}, R_{t-2}, \dots\}$, R_t is assumed to follow a central Wishart distribution

$$R_t | \mathcal{F}_{t-1} \sim \mathcal{W}_n(\nu, S_t/\nu), \tag{3}$$

where $\nu \geq n$ is the scalar degrees of freedom, and S_t/ν denotes the symmetric, positive definite $n \times n$ scale matrix, such that

$$E[R_t | \mathcal{F}_{t-1}] = S_t. \tag{4}$$

Furthermore let

$$S_t = L_t D_t L'_t \tag{5}$$

denote the SD of the conditional mean of R_t, where the diagonal elements of $D_t = \text{diag}(d_{1,t}, d_{2,t}, \dots, d_{n,t})$ are the eigenvalues of S_t and the columns of L_t are the associated orthonormal eigenvectors (see e.g. Lütkepohl, 1996, p. 69). We are interested in building a forecasting model for R_t where both the eigenvalues and the eigenvectors are allowed to vary persistently over time and which allows for convenient sequential estimation in high-dimensional applications.

3.1 Eigenvector Driving Process

In order to obtain time-varying orthonormal eigenvectors L_t in eq. (5), we introduce a matrix-variate auxiliary process $\{Q_t\}$ from which the eigenvectors L_t are obtained via computing the conditional SD of Q_t. The auxiliary process is defined as a scalar BEKK-type recursion (see Engle and Kroner, 1995) for realized covariance measures:

$$Q_t = (1-a-b)\,S + a\,R_{t-1} + b\,Q_{t-1}, \tag{6}$$

$$Q_t = L_t G_t L_t'. \tag{7}$$

The scalars a and b and the intercept matrix S are parameters to be estimated. Time-varying orthonormal eigenvectors L_t are generated by the conditional SD of Q_t in Eq. (7). The diagonal matrix of eigenvalues G_t obtains as a 'residual', which is of no further interest.

We consider model (6)-(7) as the true data generating process (DGP) for the loading matrices. Note that we may generalize the scalar dynamics of Eq. (6) to full BEKK dynamics (see Aielli and Caporin, 2015, and Noureldin et al., 2014, for details). However, estimation of such a 'complete model' in high-dimensional settings is practically impossible since the number of autoregressive parameters is of order $\mathcal{O}(n^2)$ (the *curse of dimensionality*). We therefore restrict the model to feasible scalar dynamics similar to the popular DCC-GARCH approach.

The SD in Eq. (7) is not uniquely identified. Following Aielli and Caporin (2015) we therefore impose on all SDs within the model except Eq. (5) that the eigenvalues are arranged in strictly decreasing order.

The sign of each eigenvector is still unidentified. However within the model the eigenvector matrix appears only in quadratic form. Hence there is no need for imposing a sign restriction. The implicit assumption that the eigenvalues of Q_t are distinct holds almost surely and is thus mild.

In order to ensure that Q_t is always positive definite we furthermore impose that $0 \leq a,\ 0 \leq b,\ a+b<1$ and S and Q_0 are positive definite.

We require an additional constraint on S in order to ensure a unique sequence of eigenvectors. To see this intuitively, multiply Eq. (6) by some positive constant c. Given the data $\{R_t\}_{t=1}^T$ this

would produce the same eigenvector matrix series $\{L_t\}_{t=1}^T$ since $cQ_t = cL_tG_tL_t' = L_tcG_tL_t' = L_t\tilde{G}_tL_t'$. Identification can be ensured by restricting the magnitude of the intercept matrix S as detailed in Section 3.2 below.

3.2 Eigenvalue Driving Process

The previous Section discussed the eigenvector generating process for the SD in Eq. (5). What remains in order to define the covariance matrix forecast (5) is a model for the dynamics of the eigenvalues in D_t. We employ n independent GARCH-type recursions in order to capture the dynamics of the diagonal elements of D_t. Let

$$d_{i,t} = (1 - \alpha_i - \beta_i)\gamma_i + \alpha_i \, g_{i,t-1} + \beta_i \, d_{i,t-1}, \tag{8}$$

where $g_{i,t} = e_i'L_t'R_tL_te_i$ with e_i being an $n \times 1$ vector of zeros with a 1 at the i'th position. That is, $g_{i,t}$ is the i'th diagonal element of the random matrix $L_t'R_tL_t$. Generalizations of model (8) obtained by increasing the lag order or e.g. including HAR-type dynamics (see Corsi, 2009) are straightforward to implement.

Note that

$$E[L_t'R_tL_t|\mathcal{F}_{t-1}] = L_t'E[R_t|\mathcal{F}_{t-1}]L_t = L_t'L_tD_tL_t'L_t = D_t, \tag{9}$$

such that

$$E[g_{i,t}|\mathcal{F}_{t-1}] = E[e_i'L_t'R_tL_te_i|\mathcal{F}_{t-1}] = e_i' \, E[L_t'R_tL_t|\mathcal{F}_{t-1}] \, e_i = e_i'D_te_i = d_{i,t}. \tag{10}$$

Under the usual stationarity condition we then obtain

$$E[d_{i,t}] = \gamma_i. \tag{11}$$

We now employ the SD of the intercept matrix S of the eigenvector generating auxiliary process Q_t

in Eq. (6), $S = \underline{LDL}'$, where $\underline{D} = \text{diag}(\{\underline{d}_i\}_{i=1}^n)$, and set

$$\gamma_i = \underline{d}_i, \quad i = 1, \ldots, n. \tag{12}$$

That is $\{\underline{d}_i\}_{i=1}^n$ are the eigenvalues of the intercept matrix in the eigenvector driving recursion (see Eq. 6). In summary, we impose that $\gamma_i = \underline{d}_i,\ 0 \leq \alpha_i,\ 0 \leq \beta_i,\ \alpha_i + \beta_i < 1,\ 0 < d_{i,0}\ \ \forall i$. Since all parameters are restricted to be positive this assumption also ensures that $d_{i,t}$ is always positive and consequently S_t is always positive definite.

The targeting-like constraint of setting $\gamma_i = \underline{d}_i$ solves the problem of identifying a unique L_t sequence via the Q_t auxiliary process since it implicitly imposes

$$\text{tr}(E[R_t]) = \text{tr}(S). \tag{13}$$

Hence the magnitude of the intercept matrix S is restricted, which precludes the possibility of scaling the $\{Q_t\}$ sequence by a constant c.

Proof.

$$\begin{aligned} \text{tr}(E[R_t]) = \text{tr}(E[E[R_t|\mathcal{F}_{t-1}]]) = \text{tr}(E[S_t]) = E[\text{tr}(L_t D_t L_t')] = E[\text{tr}(D_t L_t' L_t)] \\ = E[\text{tr}(D_t)] = \text{tr}(\underline{D}) = \text{tr}(\underline{DL}'\underline{L}) = \text{tr}(\underline{LDL}') = \text{tr}(S), \end{aligned} \tag{14}$$

where we used the trace property $\text{tr}(ABC) = \text{tr}(CAB) = \text{tr}(BCA)$ and orthonormality of L_t and $\underline{L}$. □

While this is not the only way to achieve identification of the eigenvector sequence $\{L_t\}$ it does entail an appealing interpretation of the model. Specifically if $a = b = 0$, the eigenvector driving process collapses to the constant matrix $Q_t = S$, such that $L_t = \underline{L}$. The resulting specification resembles the popular orthogonal GARCH model of Alexander and Chibumba (1997) and Alexander (2001), such that the DPC-CAW is regarded as being a dynamic generalization of the OGARCH to the modeling of realized covariance measures.

Recall that we assumed that the diagonal elements of $\underline{D}$ are arranged in decreasing order, in order

to identify the model. This implies that

$$E[d_{1,t}] > E[d_{2,t}] > \ldots > E[d_{n,t}]. \tag{15}$$

This however does not imply that individual elements of d_t themselves are arranged in decreasing order, since the random variables $g_{i,t}$ are not bounded above (recall that we did not need to impose identifying restrictions on the eigenvalue ordering for the SD of the covariance matrix forecast S_t in Eq. 5). A situation where $d_{i,t} < d_{i-1,t}$ happens particularly often in high dimensional applications where the elements of $\underline{d}$ are close to each other. Note that this is not a drawback but merely reflects the fact that the conditional ordering of the eigenvalues may deviate from their unconditional ordering. As an example, unconditionally, the second principal component explains a lower fraction of the total volatility than the first. But conditionally there may exist periods, where the first component dominates the second. Such situations are well known in the context of factor analysis (see also similar argumentations in Aielli and Caporin, 2015). In fact it is possible to restrict the ordering of the conditional eigenvalues by e.g. modeling their positive increments. However, this would impose unnecessary and overidentifying restrictions on the model.

The conditional Wishart assumption for R_t in Eq. (3) implies a conditional Gamma distribution for $g_{i,t}$.

$$g_{i,t}|\mathcal{F}_{t-1} \sim \text{Gamma}(\nu/2, 2d_{i,t}/\nu). \tag{16}$$

Proof. Consider the following theorem of Rao (1965):

Theorem 1. *If an $n \times n$ random matrix Y has a central Wishart distribution with ν degrees of freedom and scale matrix Ω, that is $Y \sim \mathcal{W}_n(\nu, \Omega)$, and X is a $q \times n$ matrix of rank q, then:*

$$XYX' \sim \mathcal{W}_q(\nu, X\Omega X').$$

Set $X = e_i'L_t'$, where e_i is defined as above, $\Omega = S_t/\nu$ and $Y = R_t$ to obtain $XYX' = e_i'L_t'R_tL_te_i = g_{i,t}$

and $X\Omega X' = e_i' L_t' \frac{S_t}{\nu} L_t e_i = \frac{1}{\nu} e_i' L_t' L_t D_t L_t' L_t e_i = \frac{1}{\nu} e_i' D_t e_i = \frac{d_{i,t}}{\nu}$ such that

$$g_{i,t}|\mathcal{F}_{t-1} \sim \mathcal{W}_1\Big(\nu, \frac{d_{i,t}}{\nu}\Big). \tag{17}$$

Since the univariate Wishart resembles the Gamma density, $g_{i,t}$ follows a conditional gamma distribution with shape parameter $\nu/2$ and scale parameter $2d_{i,t}/\nu$:

$$g_{i,t}|\mathcal{F}_{t-1} \sim \text{Gamma}(\nu/2, 2d_{i,t}/\nu).$$

□

Equations (3) - (8) then constitute the scalar DPC-CAW model which is summarized by the distributional assumption $R_t|\mathcal{F}_{t-1} \sim \mathcal{W}_n(\nu, S_t/\nu)$ together with the following set of equations for $t = 1, \ldots, T$:

$$\begin{aligned}
S_t &= L_t D_t L_t' \\
Q_t &= (1 - a - b)\, S + a\, R_{t-1} + b\, Q_{t-1} \\
Q_t &= L_t G_t L_t' \\
d_{i,t} &= (1 - \alpha_i - \beta_i)\gamma_i + \alpha_i\, g_{i,t-1} + \beta_i\, d_{i,t-1}, \quad \gamma_i = \underline{d}_i, \quad i = 1, \ldots, n \\
g_{i,t} &= e_i' L_t' R_t L_t e_i, \quad i = 1, \ldots, n
\end{aligned}$$

where $S = \underline{LDL}'$, $\underline{D} = \text{diag}(\{\underline{d}_i\}_{i=1}^n)$. The parameters of the DPC CAW model are comprised in the parameter vector θ with $\theta = (vech(S)', a, b, \{\alpha_i, \beta_i\}_{i=1}^n, \nu)'$ and $0 \leq a,\ 0 \leq b,\ a + b < 1$ and S and Q_0 are positive definite, $0 \leq \alpha_i,\ 0 \leq \beta_i,\ \alpha_i + \beta_i < 1,\ 0 < d_{i,0}\ \ \forall i$. In our empirical application below we initialize the eigenvector and eigenvalue recursions by $Q_0 = \frac{1}{T}\sum_{t=1}^T R_t$ and $d_{i,0} = \frac{1}{T}\sum_{t=1}^T g_{i,t}$.

3.3 Maximum Likelihood Estimation

3.3.1 One-Step Estimation

Low-dimensional settings (say, up to five assets) allow for one-step estimation of the model parameters $\theta = (vech(S)', a, b, \{\alpha_i, \beta_i\}_{i=1}^n, \nu)'$ of the DPC-CAW model. Estimation can then be carried out by maximizing the log-likelihood function

$$\begin{aligned}\mathcal{L}(\theta) &= \sum_{t=1}^{T}\left[\frac{n\nu}{2}\ln\left(\frac{\nu}{2}\right) - \frac{n(n-1)}{4}\ln(\pi) - \sum_{i=1}^{n}\ln\Gamma\left(\frac{\nu+1-i}{2}\right) + \left(\frac{\nu-n-1}{2}\right)\ln|R_t| \right. \\ &\left. -\frac{\nu}{2}\Big[\ln|S_t(\psi)| + \mathrm{tr}\left(S_t(\psi)^{-1}R_t\right)\Big]\right], \end{aligned} \tag{18}$$

where ψ summarizes the parameters for the Q_t and $d_{i,t}$ recursions, such that $\theta = (\psi, \nu)'$. The parameter ν can be treated as a nuisance parameter due to its irrelevance for the matrix forecast (see Eq. 4). In fact the first order conditions for the maximization of the log-likelihood with respect to ψ are proportional to ν. Then

$$\hat{\psi} = \underset{\psi}{argmax}\ \mathcal{L}^*(\psi), \tag{19}$$

with

$$\mathcal{L}^*(\psi) = \sum_{t=1}^{T} -\frac{1}{2}\left[\ln|S_t(\psi)| + \mathrm{tr}\left(S_t(\psi)^{-1}R_t\right)\right]. \tag{20}$$

The score vector of observation t obtains as

$$s_t(\psi) = \frac{1}{2}\left\{\left[(vec(R_t))' - (vec(S_t))'\right]\left(S_t^{-1}\otimes S_t^{-1}\right)\frac{\partial vec(S_t)}{\partial\psi}\right\}. \tag{21}$$

Assuming a correctly specified mean $E[R_t|\mathcal{F}_{t-1}] = S_t$, $s_t(\psi)$ is a martingale difference sequence since

$$E[s_t(\psi)|\mathcal{F}_{t-1}] = 0. \tag{22}$$

Consequently, as noted by Bauwens et al. (2012) and Noureldin et al. (2012), under usual regularity conditions (see e.g. Wooldridge, 1994) $\hat{\psi}$ is consistent and asymptotically normal even if the Wishart assumption is violated, provided that the conditional mean is correctly specified. Hence Eq. (20)

can be interpreted as a Quasi-Log-likelihood (QL). From the QL function in Eq. (20) we obtain the period-t log-likelihood contribution

$$\begin{aligned}
\ell_t^* &= -\frac{1}{2}\left[\ln|S_t| + \operatorname{tr}\left(S_t^{-1}R_t\right)\right] \\
&= -\frac{1}{2}\left[\ln\left|L_t D_t L_t'\right| + \operatorname{tr}\left((L_t D_t L_t')^{-1}R_t\right)\right] \\
&= -\frac{1}{2}\left[\ln|D_t| + \operatorname{tr}\left(L_t D_t^{-1} L_t' R_t\right)\right] \\
&= -\frac{1}{2}\left[\sum_{i=1}^{n}\ln\left(d_{i,t}\right) + \operatorname{tr}\left(D_t^{-1}L_t'R_tL_t\right)\right] \\
&= -\frac{1}{2}\sum_{i=1}^{n}\left[\ln\left(d_{i,t}\right) + \frac{g_{i,t}}{d_{i,t}}\right]. \qquad (23)
\end{aligned}$$

Standard errors can be obtained by the well known sandwich formula e.g. provided in Bollerslev and Wooldridge (1992). However, initial investigation showed that the QL function is multi-modal, such that standard local gradient based optimization algorithms fail if the realized covariance measure comprises more than a few assets. As an alternative, gradient-free global optimization methods like pattern search (direct search), genetic algorithms and simulated annealing can be employed (see e.g. Kelley, 1999, for details). Moreover and even more importantly, the parameter vector ψ quickly becomes large if the number of assets n increases, in particular since the parameter matrix S of the Q_t process in Eq. (6) comprises $n(n+1)/2$ model parameters, e.g. 465 intercepts for, say, $n = 30$ assets (the so-called *curse of dimensionality*). This causes additional problems in numerically optimizing the likelihood and makes high-dimensional applications (say, for $n > 10$ assets) practically impossible.

The following section therefore proposes a three-step estimation approach which solves the curse of dimensionality via multi-step estimation and *covariance targeting* (see e.g. Bauwens et al., 2006) where the intercept matrix S is replaced by an ex-ante estimate of the unconditional mean of the covariance process, similar to the DCC framework of Engle (2002). The three-step approach can therefore be applied in empirically realistic settings with $n > 10$ assets and also provides a convenient solution to the numerical optimization problems arising for one-step estimation (see above).

3.3.2 Three-Step Estimation

In high-dimensional scenarios the curse of dimensionality precludes one-step estimation of the $n(n+1)/2 + 2(n+1) + 1$ parameters of the DPC-CAW model via the corresponding Wishart likelihood. Aielli and Caporin (2015) propose a three-step estimation technique called the *DPC estimator*, which is easily adapted to the CAW framework. The procedure works as follows:

1. Estimate $S = \underline{LDL}'$ via $\hat{S} = T^{-1}\sum_{t=1}^{T} R_t$ (*covariance targeting*);

2. Conditional on step 1. estimate $(a, b)'$ by fitting a scalar CAW model (see Golosnoy et al., 2012) to the sequence of realized covariance measures, essentially (wrongly) assuming $R_t|\mathcal{F}_{t-1} \sim \mathcal{W}_n(\nu, Q_t/\nu)$, where Q_t is given by Eq. (6) and $S \overset{!}{=} \hat{S}$. Recover $\{\hat{Q}_t\}_{t=1}^{T}$ as the Q_t-sequence computed at the CAW parameter estimates in order to calculate $\{\hat{g}_{i,t}\}_{t=1}^{T}$ for $i = 1, \ldots, n$, where $\hat{g}_{i,t} = e_i'\hat{L}_t'R_t\hat{L}_t e_i$ with $\hat{L}_t$ being the matrix of eigenvectors of $\hat{Q}_t$;

3. Conditional on 1. and 2. estimate $\{\alpha_i, \beta_i\}_{i=1}^{n}$ via univariate QML based on Eqs. (8) and (16) separately $\forall i$ with $g_{i,t}$ replaced by $\hat{g}_{i,t}$ from estimation step 2. The i'th log-likelihood is given by

$$\mathcal{L}_i(\alpha_i, \beta_i) = \sum_{t=1}^{T}\Big[(\nu/2-1)\ln(\hat{g}_{i,t}) - \ln(\Gamma(\nu/2)) - (\nu/2)\ln(2d_{i,t}/\nu) - 0.5\nu\hat{g}_{i,t}/d_{i,t}\Big]. \qquad (24)$$

   Analogous to the Wishart, $\mathcal{L}_i(\alpha_i, \beta_i)$ features a QML interpretation given the previously estimated $\{\hat{g}_{i,t}\}_{t=1}^{T}$.

Steps 1 and 2 estimate the parameters of the eigenvector driving Q_t-process of the DPC-CAW model by fitting a scalar CAW model to the sequence of realized covariance measures and employing covariance targeting in order to alleviate the curse of dimensionality. Under the assumption that the true data-generating process for the realized covariance measures R_t is a DPC-CAW, it is clear from Eqs. (3) and (5) that the conditional mean of R_t is $S_t = L_t D_t L_t'$ and not $Q_t = L_t G_t L_t'$ (see Eq. 7). Hence the first two estimation steps, which essentially assume that the conditional mean of R_t is Q_t rather than S_t, introduce bias and inconsistency in estimation since ML estimators of misspecified mean models are inconsistent (see e.g. Bollerslev and Wooldridge, 1992). However,

since standard CAW models (which in the given context can be interpreted as misspecified conditional volatility models for data generated by the DPC-CAW) typically provide good approximations to the stochastic behavior of realized covariance matrices R_t (compare e.g. the results of Golosnoy et al., 2012), we can argue that the eigenvectors of Q_t-sequences obtained as estimated conditional means from CAW recursions are expected to be rather close to the eigenvectors of S_t, namely L_t. This result is illustrated in Figure 1, which shows consistent estimates of the individual Q_t elements obtained via one-step QML estimation of the DPC-CAW model to a 3-dimensional realized covariance subset of the data discussed in Section 4. The Q_t-dynamics closely follow the pattern of the realized (co)variance data, which is effectively approximated by the CAW in estimation steps 1 and 2.

Recall that steps 1 and 2 result in biased and inconsistent estimates of the parameters a, b and S since the scalar CAW likelihood in step 2 is not correctly specified (the matrix Q_t under the DPC-CAW is not the conditional mean of R_t, as discussed above). Subsequently, conditional on Steps 1 and 2 the parameters of the eigenvalue driving processes are estimated. This last estimation step does not add to the inconsistency due to the QML interpretation of the according likelihoods given in Eq. (24). Notice that the intercept parameters γ_i were fixed in step 1, such that step 3 essentially corresponds to univariate GARCH estimation with variance targeting.

The quasi-likelihood functions in steps 2 and 3 are smooth, hence standard gradient based optimization procedures can be applied. However, estimation of standard errors becomes complicated due to the aforementioned misspecification. In fact we may obtain asymptotic standard errors under the assumption of consistency by applying the GMM-framework of Engle (2009) for the three-step approach (for details see also Aielli and Caporin, 2015). However these standard errors are not valid if the model is inconsistently estimated since the moment conditions (the likelihood scores) are not valid in this case. In fact, our simulation experiments in Section 4 show signs of inconsistency. However note that standard errors are of minor importance for forecasting applications, which are in the focus of the present paper.

The three-step approach is simple and intuitive but comes with the disadvantage of introducing bias and inconsistency in parameter estimation. This inconsistency is unavoidable in high-dimensional applications of the DPC approach and has already been encountered by Aielli and

Caporin (2015) in the corresponding GARCH framework. Section 4 analyzes the properties of obtained estimates in an extensive simulation experiment. The results suggest that bias is present but acceptably small or of reduced impact, especially given the huge dimension of the estimation problem. The forecasting application of Section 5 furthermore shows that these issues do not negatively affect the out-of-sample performance, which is typically in the focus of empirical applications of multivariate volatility models.

The three-step estimator enables quick estimation in high-dimensional settings. The according CAW likelihoods are well-behaved as already found by Golosnoy et al. (2012), resulting in fast convergence of Quasi-Newton based numerical optimizers. Three-step estimation of a 100-dimensional DPC-CAW model with $T = 2500$ takes at most 100 seconds using an Intel Core i7 2.60 GHz processor under Matlab 2018b. However, the estimation of the DPC-CAW model in large/vast-dimensional settings (»100 assets) quickly becomes challenging mainly due to the inversion of the conditional mean S_t in each single likelihood evaluation of estimation step 2. This shortcoming is well known in the literature on multivariate volatility modeling (see e.g. Hafner and Reznikova, 2012). One possibility to overcome this problem is to apply composite likelihood techniques, as e.g. illustrated by Pakel et al (2018) and Engle et al. (2019) under the DCC GARCH framework. Here the composite log-likelihood is computed by summing up the log-likelihoods of pairs of assets. This greatly reduces the dimension of the estimation problem and enables inference for vast-dimensional covariance matrices. We however note that - in contrast to GARCH-type models - the DPC-CAW model is based on observed realized covariance matrices, which have been computed from synchronized intraday asset return observations. The computation of vast-dimensional realized covariance matrices is challenging, involves sparsity assumptions like factor structures, trimming or eigenvalue cleaning procedures (see e.g. Tao et al., 2011, and Lunde et al., 2016), which impose ex-ante structures on the realized measures. As an alternative way of dealing with vast-dimensional applications of the DPC-CAW model this rather suggests to impose a factor structure on the intraday asset returns and apply the DPC-CAW framework to the factor covariance matrix, which is typically of low dimension (e.g. 5 to 10-dimensional). See e.g. Asai and McAleer (2015) and Shen et al. (2018) for similar applications of CAW models to vast-dimensional frameworks. We leave the application of factor structures in the

DPC-CAW framework for future research.

4 Simulation Experiment

We conduct an extensive simulation experiment in order to assess the finite sample properties of the DPC estimator. Since we focus on high-dimensional applications the cross-sectional size is set to $n = 100$.

The following parameter set up is used: The intercept matrix $S = LDL'$ of the Q_t process is set equal to the average realized covariance matrix of the data employed in the empirical application of Section 5.1. We consider 9 distinct eigenvector recursion parameter set-ups, where the ARCH parameter a is set equal to 0.025, 0.035, 0.05 or 0.1 and the GARCH parameter is chosen such that the persistence $(a + b)$ equals 0.9, 0.95, 0.99, or 0.997. Note that the setting $(a, b) = (0.035, 0.997)$ corresponds to our empirical results from Table 6.

In order to achieve some variability in the eigenvalue recursion parameters they are drawn from uniform distributions according to

$$\alpha_i \sim U(0.22, 0.3), \qquad \beta_i|\alpha_i \sim U(0.94 - \alpha_i, 0.99 - \alpha_i). \tag{25}$$

Consequently the persistence parameters $(\alpha_i + \beta_i) \in [0.94, 0.99]$. The degrees of freedom parameter ν is set to $\nu = 100$. This parameter set up is inspired by parameter estimates obtained in the empirical application of Section 5.1. The whole experiment covers 500 independent simulations for each of the four time series lengths $T = 1000$, $T = 2500$, $T = 5000$, $T = 10000$ and each of the $4 \times 4 = 16$ parameter constellations.

All estimation results in this section and the upcoming empirical application in Section 5 are obtained under the 3-step DPC estimator. The CAW- and Gamma likelihoods are maximized via Quasi-Newton methods with BFGS-updating of the Hessian under Matlab 2018b. The eigenvector and eigenvalue recursions are initialized by $Q_0 = \frac{1}{T}\sum_{t=1}^{T} R_t$ and $d_{i,0} = \frac{1}{T}\sum_{t=1}^{T} g_{i,t}$ for $i = 1, \ldots, n$. As starting values for the numerical optimization we choose $(0.05, 0.9)$ for (a, b) and (α_i, β_i). The

numerical optimization however appeared robust to the choice of the starting values.[2]

Estimation Step 1 Note that the symmetric 100×100 parameter matrix S comprises 5050 distinct model parameters. For this reason we focus on the analysis of the 100 eigenvalues d_i, i.e. the diagonal elements of $\underline{D}$, which are of particular importance since they determine the level of the eigenvalue recursions in estimation step 3. Moreover, the d_i estimates can be interpreted as the unconditional variances of the asset returns' principal components, with decreasing fraction of explained asset return variation for decreasing i. Figure 2 shows the average relative (percentage) estimation errors for the d_i estimates, $\frac{1}{500}\sum_{j=1}^{500} 100 \cdot (\hat{d_{i,j}} - d_{i,j})/d_{i,j}$, over the 500 simulated data sets for the 16 parameter setups and four sample sizes outlined above. Note that the setting $(a = 0.035, a + b = 0.997)$ corresponds to our empirical findings of Section 5. The true parameter values for the d_i's are reported in Table 1 and show a sharply decreasing pattern with 50% (90%) of the total asset return variation explained by the first 7 (60) d_i's. For the ten highest, and therefore most important eigenvalues the biases are ranging between -35% and 2% with rather low values between -2% and +2% for settings with low a and low eigenvalue persistence. For the setting which comes closest to our empirical results, $(a = 0.035, a + b = 0.997)$, the biases of the first 10 eigenvalues range from -0.5 to -10%. For persistencies of 0.99 or higher accompanied with comparatively large values for the a parameter we obtain biases of up to 220%. This trend is particularly obvious for the high ARCH, high persistence parameter set $(a = 0.1, (a + b) = 0.997)$ which, however, does not appear to be relevant in practice (see the empirical results of Section 5). Also note that these values correspond to d_i's of very low level (compare Figure 2) whose contribution to the overall asset return volatility is very low. A significant impact of these relative biases on the forecasting performance of the DPC-CAW model is therefore not to be expected (see also our further discussion below on the additional results in Figures 6 and 7).

Estimation Step 2 Figure 3 reports Violin Plots of relative estimation errors for the ARCH parameter a. The biases range from -9.2% to 12.2% and are positive for $a = 0.025$, close to zero for $a = 0.035$ and negative for high ARCH environments with $a > 0.035$. While the biases appear mostly

[2]The Matlab estimation files for the DPC-CAW model are available under `https://github.com/mstollenwerk/dpc_caw`.

stable for increasing sample sizes, we observe a decrease in the biases for the $(a = 0.035, a+b = 0.997)$ setting but also an increasing relative bias pattern for high persistence / high ARCH settings. The smallest biases are obtained for the $(a = 0.035, a + b = 0.997)$ setting, which corresponds to our empirical results of Section 5. The largest biases are obtained for the low persistence / low ARCH setting $(a = 0.025, a + b = 0.9)$ and amount to 12% on average.

Figure 4 depicts the distribution of relative estimation errors for b. The biases appear small, ranging from -0.61% to 5.1%, and decrease with increasing persistence $a + b$. For lower values of the GARCH parameter b the biases tend to increase for high-persistence scenarios. For increasing sample sizes the biases are partly diverging, but come close to zero for our empirical estimates $(b = 0.962, a + b = 0.997)$.

Estimation Step 3 Figure 5 shows relative biases for the eigenvalue persistencies $\alpha_i + \beta_i$ for $i = 1, \ldots, 100$, which appear overall low with values ranging from -3% to 6%. For the individual α_i and β_i we find negative and compensating positive biases between -60% and 10% (α) and -10% and 30% (β) (not reported here).

In order to investigate the effect of the biases on the estimated asset return (co)variances we compute for each of the simulated data-sets the sequence of estimated covariance forecasts $\hat{S}_t(\hat{\theta})$ for $t = 1, \ldots, T$ as a function of the parameter estimates $\hat{\theta}$ discussed above, and compare these estimates to the true simulated forecasts $S_t(\theta)$. Figure 6 reports Violin Plots of the average Euclidean norms $(1/T) \sum_t \text{vech}(\hat{S}_t(\hat{\theta}) - S_t(\theta))' \text{vech}(\hat{S}_t(\hat{\theta}) - S_t(\theta))$ over the 500 simulations, and Figure 7 shows Violin Plots of the corresponding relative (percentage) biases computed as $(1/T) \sum_t (100/5050) \sum_{i>j} (\hat{S}_{ij}(\hat{\theta}) - S_{ij}(\theta))/S_{ij}(\theta)$. The results show a pattern of decreasing Euclidean norms for increasing sample sizes, coming close to zero for $T = 10000$ and all considered parameter settings. The distribution of relative biases shows low dispersion around zero and the variation appears to be decreasing with increasing sample size.

Summarizing the results, the simulation experiment indicates relatively low biases for the eigenvalue persistencies and the ARCH- and GARCH parameters of the eigenvector recursions while the

biases of the d_i-estimates in estimation step 1 amount to up to 220% for eigenvalues of very low level. However, these biases do not significantly affect the covariance forecasts as indicated by the Euclidean norms and biases for S_t reported in Figures 6 and 7. Hence the biases are not expected to significantly affect the forecasting performance of the DPC-CAW model.

5 Empirical Application

5.1 Data

We apply the scalar DPC-CAW model introduced in Section 3 in order to capture the dynamics of 100-dimensional realized covariance measures. The data have been computed from one-minute intraday asset returns by the microstructure-noise and jump robust multivariate realized kernel method of Barndorff-Nielsen at al. (2011). The corresponding ticker symbols are shown in Table 2. Note that the choice of the particular type of realized measure is not an important issue since the model can be fitted to any series of positive-definite realized covariance measures. The sample period starts at January 1, 2002, and ends on December 31, 2014, covering 3271 trading days.

Figure 8 depicts exemplary time-series plots of variance and covariance series and according sample autocorrelation functions for 4 stocks included in the data set. Descriptive statistics are provided in Table 3. The (co)variance processes are highly persistent, skewed to the right, leptokurtic and tend to move parallel to each other.

5.2 In-Sample Estimation Results

We start with analyzing the in-sample fit of the DPC-CAW model for various model-order settings using the BIC information criterion. According to Geweke and Meese (1981) it can be shown that for linear ARMA type models the BIC is consistent in the sense that asymptotically the correct model order is chosen. This however does not hold for nonlinear models such as the CAW of Golosnoy et al (2012) or equivalently the DPC-CAW - even if consistently estimated. Golosnoy et al. (2012) remark that there do not seem to exist published results on the consistency of the BIC for nonlinear time-series models, which would justify its use under the CAW framework. Nevertheless the BIC is

often applied as an 'indicator', although not asymptotically valid in a strict sense. Here we follow Golosnoy et al. (2012) and align the BIC-based model selection by model diagnostics based on Ljung-Box residual autocorrelation tests for the fitted models.

All models are estimated by the 3-step approach. We consider both, order selection for the eigenvector- and for the eigenvalue processes given in Eqs. (6) and (8), jointly. For the eigenvalues we restrict the chosen order to be identical across the 100 assets. Table 4 shows the results on the BIC information criteria. We find a clear indication for the (1,1) specification of the eigenvector recursion, which corresponds to the typically chosen DCC-GARCH specification for correlations. The distribution of BIC values over the various eigenvalue order-constellations is much more even and overall results in the preferred (3,4) model. For comparison we also report the BIC obtained for a standard HAR specification of the eigenvalue dynamics (see Corsi, 2009). The model boils down to a restricted autoregressive specification of order 20. The HAR model, although very popular in empirical applications, is not preferred in any case.

Table 5 shows corresponding model diagnostic results on the in-sample fit to the autocorrelation structure of the underlying realized (co)variance data. The model diagnostics are based on Ljung-Box autocorrelation tests on the standardized Martingale differences ("Pearson residuals") in the $n = 5050$-dimensional vector

$$\begin{aligned} e_t^* &= \Big(\mathrm{Cov}[vech(R_t)|\mathcal{F}_{t-1}]\Big)^{-1/2} \mathrm{vech}\Big(R_t - E[R_t|\mathcal{F}_{t-1}]\Big) \\ &= \Big[\frac{1}{\nu} L_n\big(I_{n^2} + K_{nn}\big)\big(S_t \otimes S_t\big)L_n'\Big]^{-1/2} \mathrm{vech}\Big(R_t - S_t\Big), \end{aligned}$$

where K_{nn} denotes the commutation matrix (see e.g. Lütkepohl, 2005) and L_n denotes the elimination matrix defined by $vech(X) = L_n vec(X)$. Under the null of correct model specification these residuals should be serially uncorrelated. The table reports averaged values of the Ljung-Box test statistics at 200 lags computed over the $100 \times 101/2 = 5050$ time series of Pearson-residuals in the vector e_t^*. The 5% critical value of the Ljung-Box test statistic at 200 lags is $\chi^2_{200} = 233.99$. The results give an impression of the overall fit of the various model constellations to the (co)variance dynamics of the data, since the average Ljung-Box test statistics represent an aggregate of the squared

residual autocorrelations for all variance and covariance series. The minimum average Ljung-Box test statistic is obtained for a (1,2) specification for the eigenvector dynamics combined with a (4,4) lag-order for the eigenvalue processes as the best-fitting model for the in-sample (co)variance dynamics. Compared to the BIC results, the residual diagnostics show a tendency for higher lag orders, which is explained by the absence of a penalty term for the number of model parameters. Similar to the BIC results, the HAR specification is not preferred in any case. Moreover, the minimum average Ljung-Box test statistic of 215.76 is very close to the one obtained for the BIC-preferred (1,1)-(3,4) specification (216.20). Hence the BIC - although not consistent - appears to be a reasonable advice for selecting the lag-order of the DPC-CAW.

Table 6 reports a summary of the obtained estimates for the BIC-preferred (3,4)-(1,1) DPC-CAW specification. The estimated persistence for the eigenvector- and eigenvalue recursions is very high with $(a+b) = 0.997$ and a median of $\sum_{\ell=1}^{p} \alpha_{i,\ell} + \sum_{\ell=1}^{q} \beta_{i,\ell}$ of 0.978. This corresponds to the findings in Aielli and Caporin (2015) and resembles analogous results for scalar DCC-GARCH applications with intercept targeting.

The right panel of Figure 8 shows sample autocorrelation functions of standardized Pearson residuals from the DPC-CAW(3,4)-(1,1) model for exemplary variance and covariance series of four stocks included in the 100-dimensional data set. The results presented in the Figure are representative for the complete set of stocks. The ACFs are depicted together with 95% Bartlett confidence bands for the variance and covariance series separately and illustrate the overall good fit of the DPC-CAW approach. The model successfully reduces the serial dependence to a minimum. We however observe some remaining predictability in the residual series: 441 of the 5050 series do not pass the Ljung-Box test for zero autocorrelation at the 1% level and 100 lags. The literature reports much worse fractions for applications of much lower dimension (see e.g. the model diagnostic results for the flexible CAW specifications in Golosnoy et al., 2012, for a 5-dimensional application). The diagnostics therefore imply a good fit to the complex dynamics of 5050 distinct variance and covariance series. Also note that we may interpret some remaining residual predictability as a result of the sparse scalar model structure of the DPC-CAW which enables applications to high-dimensional covariance matrices, while avoiding overfitting and spuriously uncorrelated residuals. The residual ACFs in Figure 8 show that

remaining predictability is typically found in variance residuals. This may be related to the direct modeling of principal component variances rather than return variances.

5.3 Out-of-Sample Forecasting

We now compare the out-of-sample 1-period, 5-period and 10-period ahead forecasting performance of the DPC-CAW specification to alternative forecasting models proposed in the literature on realized covariance modeling. We consider two out-of-sample windows: The first window starts at January 1, 2009, and ends on December 31, 2011, covering the subprime crisis period. The window exhibits a particularly high volatility level and pronounced volatility peaks. The second window covers a period of low to moderate volatility from January 1, 2012, until December 31, 2014, representing normal stock market fluctuations (see the left panel in Figure 8 for exemplary time series plots). The models are re-estimated daily using a rolling window of the previous 1750 covariance measures, i.e. roughly 7 years of data. New forecasts are generated based on the updated parameter estimates.

5.3.1 Competing Models and Forecast Evaluation

The scalar Re-DCC model of Bauwens et al. (2012) represents the 'natural' competitor for the DPC-CAW approach. The Re-DCC model decouples correlations and variances which facilitates three-step estimation similar to the DPC estimator (see Bauwens et al., 2012, for details). The model assumes a conditional central Wishart distribution for the realized covariance measure and decomposes the scale matrix S_t into

$$S_t = V_t \rho_t V_t, \tag{26}$$

where $V_t = \text{diag}(\sqrt{s_{11,t}}, \sqrt{s_{22,t}}, \ldots, \sqrt{s_{nn,t}})$ and ρ_t is the correlation matrix implied by S_t. We consider GARCH(p, q) recursions for the conditional variances:

$$s_{ii,t} = \gamma_i + \sum_{k=1}^{p} \alpha_{k,i} r_{ii,t-k} + \sum_{l=1}^{q} \beta_{k,i} s_{ii,t-l}. \tag{27}$$

The correlation matrix ρ_t is parameterized as follows:

$$\rho_t = (1 - a - b)\bar{\rho} + aP_{t-1} + b\rho_{t-1}, \tag{28}$$

where P_t is the realized correlation matrix

$$P_t = \{diag(R_t)\}^{-1/2} R_t \{diag(R_t)\}^{-1/2}. \tag{29}$$

$\bar{\rho}$ is estimated by the sample mean of realized correlation measures ("correlation targeting").

We also consider a constant conditional correlation CAW (CCC-CAW) model since it represents a restricted Re-DCC specification where $a = b = 0$. In a similar fashion we restrict the DPC-CAW model to $a = b = 0$ in order to obtain the CAW-analogue to the OGARCH model (O-CAW). Additionally we consider the $DPC - CAW_{0f}$ model which is obtained by restricting the eigenvalue dynamics of the DPC-CAW model to $\alpha_i = \alpha$ and $\beta_i = \beta \ \forall i = 1, \ldots, n$. This particular model restriction turned out to be favorable in forecasting applications.

We furthermore analyze an exponentially weighted moving average (EWMA) specification, called RiskMetrics (see J.P. Morgan, 1996), which boils down to exponential smoothing of realized covariance matrices using a preset smoothing parameter λ. The forecast of the realized covariance matrix is then given by

$$E[R_t|\mathcal{F}_{t-1}] = (1 - \lambda)R_{t-1} + \lambda E[R_{t-1}|\mathcal{F}_{t-2}], \tag{30}$$

where λ is set to its typical value for daily data, i.e. $\lambda = 0.94$.

As further forecasting models we apply the Factor HEAVY approach of Sheppard and Xu (2019), where we use realized variances of the S&P 500 for the market factor (see Section 1.1 and in particular Eq. 12 in Sheppard and Xu, 2019, for details on the model setup), and the PCA based CAW factor approach of Shen et al. (2018) (labeled *Factor-CAW*), with seven factors selected by the eigenvalue criterion as discussed by Shen et al. (2018). While the Factor HEAVY model allows for dynamic loadings and idiosyncratic variances, the Factor-CAW uses the PCA results of Tao et al. (2011), who restrict the loadings and the idiosyncratic covariance matrix to be constant over time. Shen et al.

(2018) first compute the realized factor covariance matrix using PCA techniques and then employ diagonal CAW(p,q) processes in order to forecast the factor covariances.

Multi-step ahead forecasts are obtained by iterating the model recursions and replacing unknown future dependent variables by their forecasts (see the according papers for details). The first two columns of Table 7 provide an overview of all considered model specifications. The (p,q) column describes the number of GARCH lags in the conditional variance specifications (Re-DCC and CCC-CAW) or in the eigenvalue recursions (DPC-CAW and O-CAW), respectively. For the Factor HEAVY model the order refers to the factor-, beta- and idiosyncratic variance dynamics and for the Factor-CAW approach to the CAW(p,q) process for the factor covariance dynamics.

We now turn to the evaluation of the forecasting performance. Let $L(\hat{X}, X)$ denote the Euclidean distance of the half-vectorization of the forecast error matrix given by

$$L(\hat{X}, X) = \text{vech}(\hat{X} - X)' \, \text{vech}(\hat{X} - X), \tag{31}$$

where $\hat{X}$ represents a particular matrix forecast and X the according realization. We apply five different loss functions in order to evaluate the forecasting performance of the considered models:

(i) MSE of predicted covariance matrix: $L(\hat{R}_t, R_t)$;

(ii) MSE of predicted variances: $\big(\text{diag}(\hat{R}_t - R_t)\big)' \big(\text{diag}(\hat{R}_t - R_t)\big)$;

(iii) MSE of predicted correlation matrix: $L(\hat{\rho}_t, \rho_t)$;

(iv) Variance of predicted global minimum variance portfolio (GMVP): $V_{GMPV,t}$;

(v) QLIKE: $QLIKE_t = \ln|\hat{R}_t| + vec\Big(\hat{R}_t^{-1} R_t\Big)' \iota$.

The model-specific forecast of the covariance matrix R_t is given by $\hat{R}_t = E[R_t|\mathcal{F}_{t-1}]$ and accordingly $\hat{\rho}_t = \{diag(\hat{R}_t)\}^{-1/2} \hat{R}_t \{diag(\hat{R}_t)\}^{-1/2}$. We use the realized kernel estimate R_t as unbiased proxy for the true covariance matrix at period t.

Loss function (i) considers whole covariance matrix forecasts, while (ii) and (iii) focus on variances and correlations instead. These quantities are of particular interest since DCC frameworks model

variance and correlation dynamics separately. Loss function (iv) considers economic losses via computing the forecast of the variance of the GMVP given by $V_{GMPV,t} = \hat{w}' R_t \hat{w}$, with $\hat{w} = \hat{R}_t \iota / (\iota' \hat{R}_t^{-1} \iota)$, where ι is an n-dimensional vector of ones. See e.g Patton (2011) for a discussion of the properties of the QLIKE loss function (v), which is known to be robust to noisy volatility proxies.

We compute sample averages of the obtained losses over the respective forecasting windows and asses the significance of differences in losses via the model confidence set (MCS) approach of Hansen et al. (2011). At a given confidence level $(1 - \alpha)$ the MCS contains the single model or the set of models with the best forecasting performance. We select $\alpha = 0.1$ as suggested by Hansen et al. (2011) and compute the confidence sets using the stationary bootstrap method with window lengths determined by the maximum number of significant parameters obtained by fitting an AR(p) process on the loss differences and $5,000$ bootstrap replications.

5.3.2 Forecasting Results

The 1-period ahead forecasting results are summarized in Tables 7 and 8, the 5-period ahead results in Tables 9 and 10 and the 10-period ahead results in Tables 11 and 12.

The DPC-CAW approach (DPC-CAW and DPC-CAW$_{0f}$) provides an overall good forecasting performance over all considered subperiods and forecasting horizons. Moreover, the DPC model provides the overall best forecasts w.r.t the correlation- and the economically important GMVP loss functions (with the exception of 10-period ahead predictions in the calm period). In particular, the DPC-CAW approach features significantly lower correlation and GMVP losses compared to up-to-date competitors like the flexible Factor HEAVY approach of Sheppard and Xu (2019) and the Factor-CAW of Shen et al. (2018) across all subperiods and horizons. The DPC-CAW approach with HAR dynamics for the eigenvalue recursions also appears to be particularly strong in forecasting the whole covariance matrix, as indicated by the respective MSE results. For the variance loss however, the DPC models are significantly outperformed by the EWMA approach in calm periods and forecasting horizons of five and ten trading days ahead. In turbulent periods the MCS for the variance- and covariance matrix losses are very wide due to a huge dispersion of the Euclidean distances in (31), such that differences in MSE losses across models are not significant in most cases. We also note that

EWMA appears to be a serious competitor in multi-period forecasting during calm market phases, while the Factor-HEAVY approach is typically preferred under the QLIKE and turbulent market conditions. The Factor-CAW in contrast does not appear to be a serious competitor in any case, which may be explained by the model's strong restrictions on the covariance dynamics via imposing time-constant idiosyncratic variances and factor loadings. The Re-DCC-CAW, O-CAW and CCC-CAW models do not show an overall remarkable performance and are typically outperformed by their competitors.

Taken all together, the results confirm the presumption that the independent modeling of principal component variances with time-varying eigenvectors offers a precise description of covariance and correlation dynamics. In particular the GMVP forecasting results highlight the importance of capturing correlation dynamics in the portfolio context. We conclude that the DPC-CAW approach has particularly good forecasting properties and notably outperforms its competitors especially in covariance-, correlation- and GMVP forecasting.

6 Conclusion

In this paper we propose a Dynamic Principal Component (DPC) CAW model for time-series of high-dimensional realized covariance measures of asset returns. The model performs a spectral decomposition of the scale matrix of a central Wishart distribution and assumes independent dynamics for the principal components' variances and the eigenvector processes. A three-step estimation procedure similar to the DCC framework for asset returns makes the model applicable to high-dimensional covariance matrices.

We analyze the finite sample properties of the three-step estimation approach in an extensive simulation experiment and provide an empirical application to realized covariance measures for 100 assets traded at the NYSE. The DPC-CAW model has particularly good forecasting properties and outperforms its competitors including DCC-CAW-, Factor HEAVY- and Factor-CAW specifications for realized covariance measures.

7 Acknowledgements

The authors thank two anonymous referees for their helpful and constructive comments. This work was partially performed on the computational resource "BwForCluster MLS&WISO" funded by the Ministry of Science, Research and Arts and the Universities of the State of Baden-Württemberg, Germany, within the framework program bwHPC.

References

[1] Aielli, G.P., Caporin, M., 2015. Dynamic principal components: a new class of multivariate GARCH models. Working Paper.

[2] Alexander, C.O., 2001. Orthogonal GARCH. Mastering Risk 2, 21-38.

[3] Alexander, C.O., Chibumba, A.M., 1997. Multivariate orthogonal factor GARCH. Mimeo: University of Sussex, UK.

[4] Andersen, T.G., Bollerslev, T., Diebold, F.X., Labys, P., 2003. Modeling and forecasting realised volatility. Econometrica 71, 579–625.

[5] Asai, M., McAleer, M., Yu, J., 2006. Multivariate stochastic volatility: a review. Econometric Reviews 25, 145–175.

[6] Asai, M., McAleer, M., 2015. Forecasting co-volatilites via factor models with asymmetry and long memory in realized covariance. Journal of Econometrics 189(2): 251–262.

[7] Barndorff-Nielsen, O.E., Hansen, P.R., Lunde, A., Shephard, N., 2011. Multivariate realised kernels: Consistent positive semi-definite estimators of the covariation of equity prices with noise and non-synchronous trading. Journal of Econometrics 162: 149–169.

[8] Barndorff-Nielsen, O.E., Shephard, N., 2004. Econometric analysis of realised covariation: high frequency covariance, regression and correlation in financial economics. Econometrica 72: 885–925.

[9] Bauer, G.H., Vorkink, K., 2011. Forecasting multivariate realized stock market volatility. Journal of Econometrics 160, 93–101.

[10] Bauwens, L., Braione, M., Storti, G., 2014. Forecasting comparison of long term component dynamic models for realized covariance matrices. Annals of Economics and Statistics 123/124: 103-134.

[11] Bauwens, L., Braione, M., Storti, G., 2016. A dynamic component model for forecasting high-dimensional realized covariance matrices. Econometrics and Statistics 1: 40-61.

[12] Bauwens, L., Laurent, S., Rombouts, J.V.K., 2006. Multivariate GARCH models: a survey. Journal of Applied Econometrics 21, 79–109.

[13] Bauwens, L., Storti, G., and F. Violante. 2012. Dynamic conditional correlation models for realized covariance matrices. CORE Working paper.

[14] Bollerslev, T., Wooldridge, J.M., 1992. Quasi maximum likelihood estimation and inference in dynamic models with time-varying covariances. Econometric Reviews 11:143–172.

[15] Chiriac, R., Voev, V., 2011. Modelling and forecasting multivariate realized volatility. Journal of Applied Econometrics 26, 922–947.

[16] Corsi, F., 2009. A simple approximative long-memory model of realized volatility. Journal of Financial Econometrics.

[17] Engle, R.F., 2002. Dynamic conditional correlation: a simple class of multivariate GARCH models. Journal of Busines and Economic Statistics 20: 339–350.

[18] Engle, R.F., 2009. Anticipating correlations. Princeton University Press.

[19] Engle, R.F., Kroner, K.F. 1995. Multivariate simultaneous generalized ARCH. Econometric Theory 11: 122–150.

[20] Engle, R.F., Ledoit, O., Wolf, M. 2019. Large dynamic covariance matrices. Journal of Business and Economic Statistics 37(2): 363–375.

[21] Geweke, J., Meese, R., 1981. Estimating regression models of finite but unknown order. International Economic Review 22: 55-70.

[22] Golosnoy, V., Gribisch, B., Liesenfeld, R., 2012. The conditional autoregressive Wishart model for multivariate stock market volatility. Journal of Econometrics 167: 211–223.

[23] Gourieroux, C., Jasiak, J., Sufana, R., 2009. The Wishart autoregressive process of multivariate stochastic volatility. Journal of Econometrics 150: 167–181.

[24] Hafner, C.M., Reznikova, O., 2012. On the estimation of dynamic conditional correlation models. Computational Statistics and Data Analysis 56: 3533–3545.

[25] Hansen, P.R., Lunde, A., Nason, J.M. 2011. The model confidence set. Econometrica 79: 453–497.

[26] Jin, X., Maheu, J.M., Yang, Q., 2019. Bayesian parametric and semiparametric factor models for large realized covariance matrices. Journal of Applied Econometrics. Online January 2019.

[27] Kelley, C.T. 1999. Iterative Methods for Optimization. Society for Industrial and Applied Mathematics.

[28] Lütkepohl, H. 1996. Handbook of Matrices. Wiley: Chichester.

[29] Lütkepohl, H., 2005. New Introduction to Multiple Time Series Analysis. Springer, Berlin.

[30] Lunde, A., Shephard, N., Sheppard, K., 2016. Econometric analysis of vast covariance matrices using composite realized kernels and their application to portfolio choice. Journal of Business and Economic Statistics 34(4): 504–518.

[31] Morgan, J.P., 1996. RiskMetrics. Technical Document, Fourth ed., New York.

[32] Noureldin, D., Shephard, N., Sheppard, K., 2012. Multivariate High-Frequency-Based Volatility (HEAVY) Models. Journal of Applied Econometrics 27: 907–933.

[33] Noureldin, D., Shephard, N., Sheppard, K., 2014. Multivariate rotated ARCH models. Journal of Econometrics 179: 16–30.

[34] Pakel, C., Shephard, N., Sheppard, K., Engle, R.F., 2018. Fitting vast dimensional time-varying covariance models. SSRN working paper.

[35] Patton, A.J., 2011. Volatility forecast comparison using imperfect volatility proxies. Journal of Econometrics 16(1): 246–256.

[36] Rao, C.R., 1965. Linear statistical inference and its applications. John Wiley & Sons.

[37] Shen, K., Yao, J., Li, W.K., 2018. Forecasting high-dimensional realized volatility matrices using a factor model. Quantitative Finance. Online June 2018.

[38] Sheppard, K., Xu, W., 2019. Factor high-frequency based volatility (HEAVY) models. Journal of Financial Econometrics 17(1): 33–65.

[39] Tao, M., Wang, Y., Yao, Q., Zou, J., 2011. Large volatility matrix inference via combining low-frequency and high-frequency approaches. Journal of the American Statistical Association 106(495): 1025–1040.

[40] Wang, Y., Zou, J., 2010. Vast volatility matrix estimation for high-frequency financial data. The Annals of Statistics 38(2): 943–978.

[41] Woolridge, J.M., 1994. Estimation and inference for dependent processes. Handbook of econometrics 4, 2639–2738.

125.16	16.39	10.98	9.26	7.99	6.47	6.42	5.57	5.31	5.31
5.22	4.84	4.80	3.97	3.83	3.79	3.69	3.43	3.35	3.22
3.10	3.00	2.95	2.83	2.74	2.71	2.66	2.61	2.55	2.48
2.37	2.28	2.27	2.19	2.17	2.13	2.11	2.07	2.01	1.94
1.93	1.87	1.85	1.82	1.79	1.77	1.72	1.69	1.64	1.62
1.59	1.57	1.54	1.49	1.47	1.47	1.42	1.41	1.37	1.34
1.32	1.29	1.29	1.28	1.25	1.25	1.24	1.23	1.23	1.20
1.19	1.18	1.12	1.12	1.06	1.05	1.05	1.03	1.02	1.00
0.99	0.98	0.97	0.93	0.91	0.88	0.85	0.83	0.83	0.82
0.81	0.81	0.79	0.72	0.68	0.65	0.54	0.51	0.47	0.45

Table 1: Sorted eigenvalues obtained from $\hat{S} = T^{-1} \sum_{t=1}^{T} R_t$ for the data-set described in Section 5.

Symbol	Company	Symbol	Company
a	Agilent Technologies Inc.	gild	Gilead Sciences Inc.
aa	Alcoa Inc.	glw	Corning Incorporated
aapl	Apple Inc.	gps	Gap, Inc.
abt	Abbott Laboratories	gs	Goldman Sachs Group, Inc.
abx	Barrick Gold Corporation	hal	Halliburton Company
adbe	Adobe Systems Incorporated	hd	Home Depot, Inc.
adi	Analog Devices Inc.	hig	Hartford Financial Services Group, Inc.
adp	Automatic Data Processing	hon	Honeywell International Inc.
aig	American International Group Inc.	hpq	Hewlett-Packard Company
all	Allstate Corporation	ibm	International Business Machines Corporation
altr	Altera Corporation	intc	Intel Corporation
amat	Applied Materials Inc.	intu	Intuit Inc.
amd	Advanced Micro Devices Inc.	ip	Internation Paper Company
amgn	Amgen Inc.	jcp	J.C. Penney Company, Inc. Holding Company
amzn	Amazon.com, Inc.	jnj	Johnson & Johnson
apc	Anadarko Petroleum Corporation	jnpr	Juniper Networks, Inc.
axp	American Express Company	jpm	J P Morgan Chase & Co
ba	Boeing Company	klac	KLA-Tencor Corporation
bac	Bank of America Corporation	ko	Cocoa-Cola Company
bax	Baxter International Inc.	kr	Kroger Company
bbby	Bed Bath & Beyond Inc.	kss	Kohl's Corporation
bby	Best Buy Co., Inc.	lb	La Barge Inc.
bhi	Baker Hughes Incorporated	lltc	Linear Technology Corporation
bmy	Bristol-Myers Squibb Company	lly	Eli Lilly and Company
brcm	Broadcom Corporation	lmt	Lockheed Martin Corporation
c	Citigroup Inc.	low	Lowe's Companies, Inc.
cag	ConAgra, Inc.	luv	Southwest Airlines Company
cah	Cardinal Health Inc.	mas	Masco Corporation
cat	Caterpillar, Inc.	mcd	McDonald's Corporation
cbs	CBS Corporation new	mdt	Medtronic Inc.
cien	Ciena Corporation	met	MetLife, Inc.
cl	Colgate-Palmolive Company	mmc	Marsh & McLennan Companies, Inc.
cop	ConocoPhillips	mmm	3M Company
cost	Costco Wholesale Corporation new	mo	Altria Group
csco	Cisco Systems, Inc.	mrk	Merck & Company, Inc. New
ctxs	Citrix Systems, Inc.	ms	Morgan Stanley Dean Witter & Co
cvs	CVS Caremark Corp.	msft	Microsoft Corporation
cvx	Chevron Corporation new	msi	Motorola Solutions, Inc.
dd	E.I. du Pont de Nemours and Company	mu	Micron Technology, Inc.
de	Deere & Company	nem	Newmont Mining Corporation
dis	Walt Disney Company	nke	Nike, Inc.
dow	Dow Chemical Company	ntap	NetApp, Inc.
duk	Duke Energy Corporation new	nvda	NVIDIA Corporation
ea	Electronic Arts Inc.	orcl	Oracle Corporation
ebay	Ebay Inc.	oxy	Occidental Petroleum Corporation
emc	EMC Corporation MA	payx	Paychex, Inc.
emr	Emerson Electric Company	pep	Pepsico, Inc.
f	Ford Motor Company DEL	pfe	Pfizer, Inc.
fitb	Fifth Third Bancorp	pg	Procter & Gamble Company
ge	General Electric Company	qcom	QUALCOMM Incorporated

Table 2: Data set of 100 stocks selected by liquidity from the S&P 500.

	Mean	Min.	Max.	Range	Std. dev.	Skewness	Kurtosis
Realized variances (100 time series)							
Min.	1.01	0.02	48.39	48.35	1.89	3.75	24.30
Median	3.29	0.10	117.00	116.88	5.04	8.90	146.03
Max.	12.51	0.35	7727.54	7727.50	151.31	43.58	2126.80
Realized covariances (4950 time series)							
Min.	0.20	−126.77	14.07	14.97	0.87	−1.09	33.86
Median	1.05	−3.32	63.42	68.12	2.61	10.08	169.49
Max.	3.90	−0.02	1262.30	1282.60	25.51	38.93	1851.31

Table 3: Descriptive statistics for the 5050 realized variance and covariance time series of the 100-dimensional data-set described in Section 5.

Order of eigenvalue process	Order of eigenvector process							
	(1,0)	(1,1)	(2,1)	(1,2)	(2,2)	(3,2)	(2,3)	(3,3)
(1,0)	-2.5515	-2.7671	-2.0818	-2.4645	-2.0505	-2.0298	-2.0460	-2.0998
(1,1)	-2.6289	-2.8342	-2.1779	-2.5363	-2.1531	-2.1402	-2.1502	-2.1892
(2,1)	-2.6287	-2.8340	-2.1778	-2.5361	-2.1531	-2.1404	-2.1503	-2.1890
(1,2)	-2.6304	-2.8361	-2.1793	-2.5377	-2.1546	-2.1418	-2.1517	-2.1905
(2,2)	-2.6303	-2.8360	-2.1793	-2.5376	-2.1546	-2.1420	-2.1518	-2.1904
(3,2)	-2.6301	-2.8358	-2.1791	-2.5374	-2.1544	-2.1418	-2.1516	-2.1902
(2,3)	-2.6311	-2.8371	-2.1801	-2.5384	-2.1553	-2.1428	-2.1525	-2.1911
(3,3)	-2.6309	-2.8370	-2.1799	-2.5382	-2.1552	-2.1426	-2.1524	-2.1909
(4,3)	-2.6308	-2.8368	-2.1798	-2.5380	-2.1550	-2.1425	-2.1522	-2.1908
(3,4)	-2.6314	**-2.8374**	-2.1801	-2.5384	-2.1554	-2.1428	-2.1526	-2.1911
(4,4)	-2.6313	-2.8372	-2.1800	-2.5383	-2.1553	-2.1427	-2.1524	-2.1910
(5,4)	-2.6312	-2.8371	-2.1798	-2.5381	-2.1551	-2.1426	-2.1523	-2.1908
(4,5)	-2.6313	-2.8373	-2.1800	-2.5382	-2.1553	-2.1428	-2.1525	-2.1910
(5,5)	-2.6312	-2.8372	-2.1799	-2.5381	-2.1552	-2.1426	-2.1524	-2.1908
HAR	-2.6280	-2.8341	-2.1773	-2.5357	-2.1525	-2.1397	-2.1497	-2.1883

Table 4: BIC information criteria for various lag-order constellations. BIC values: $\times 10e7$. Models are estimated using the 3-step estimation approach.

Order of eigenvalue process	Order of eigenvector process							
	(1,0)	(1,1)	(2,1)	(1,2)	(2,2)	(3,2)	(2,3)	(3,3)
(1,0)	472.14	229.53	229.64	229.17	230.23	230.02	230.19	229.99
(1,1)	463.15	229.00	230.11	229.31	229.44	228.57	228.70	227.98
(2,1)	410.37	216.57	216.88	216.40	217.04	216.74	216.82	216.59
(1,2)	411.70	216.89	216.90	216.70	216.96	216.89	216.94	216.75
(2,2)	410.12	216.50	216.24	216.24	216.66	216.33	216.57	216.23
(3,2)	413.72	222.43	222.82	222.30	223.76	223.47	223.52	221.87
(2,3)	408.66	216.43	216.28	216.19	216.43	216.68	216.61	216.40
(3,3)	409.83	217.14	216.77	216.74	217.02	216.68	216.71	216.65
(4,3)	417.61	219.02	217.56	218.43	218.08	217.56	218.09	217.61
(3,4)	409.43	216.20	216.27	216.06	216.43	216.30	216.40	216.17
(4,4)	408.87	215.96	216.41	**215.76**	216.32	216.20	216.29	216.10
(5,4)	409.95	216.47	216.66	216.12	216.82	216.28	216.75	216.24
(4,5)	409.37	216.32	216.47	216.20	216.41	216.25	216.36	216.08
(5,5)	408.90	216.35	216.37	216.06	216.29	216.14	216.26	215.99
HAR	408.75	218.19	224.40	218.02	218.04	223.19	217.90	218.98

Table 5: Average Ljung-Box autocorrelation test statistics computed over the $100 \times 101 = 5050$ time series of standardized Martingale differences computed for the realized (co)variance data. The 5% critical value of the Ljung-Box test statistic at 200 lags is $\chi^2_{200} = 233.99$.

	Eigenvalue Process							
	$\alpha_{i,1}$	$\alpha_{i,2}$	$\alpha_{i,3}$	$\beta_{i,1}$	$\beta_{i,2}$	$\beta_{i,3}$	$\beta_{i,4}$	$\sum_{\ell=1}^{p}\alpha_{i,\ell}+\sum_{\ell=1}^{q}\beta_{i,\ell}$
Median	0.311	0.074	0.000	0.132	0.068	0.135	0.135	0.978
Min.	0.025	0.000	0.000	0.000	0.000	0.000	0.000	0.947
Max.	0.492	0.180	0.116	0.519	0.348	0.373	0.378	0.987

Eigenvector Process		
a	b	$a+b$
0.035	0.962	0.997

Table 6: Summary of parameter estimates obtained by the DPC estimator for the 100-dimensional data-set described in Section 5 and the BIC selected model order (3,4)-(1,1).

Volatile Market: 01.01.2009 – 31.12.2011						
Model	(p,q)	Cov matrix	Var	Corr	GMVP $\times 10^2$	QLIKE
DPC-CAW	(1,1)	32289	8322	195.9	37.86	150.8
	(2,2)	32036	8296	195.5	37.74	150.4
	(3,3)	31943	8316	195.1	37.70	149.8
	HAR	**31682**	8275	194.5	37.72	149.5
DPC-CAW$_{0f}$	(1,1)	32596	8511	193.5	37.76	147.0
	(2,2)	32262	8466	193.2	37.67	146.7
	(3,3)	32063	8450	**192.9**	37.65	146.2
	HAR	31981	8415	193.0	**37.64**	146.4
Re-DCC-CAW	(1,1)	32346	8222	229.1	39.72	201.5
	(2,2)	32196	8155	228.8	39.59	200.4
	(3,3)	32335	8239	228.6	39.53	199.3
	HAR	33874	8137	235.9	40.37	186.0
O-CAW	(1,1)	38834	10116	211.6	49.99	148.3
	(2,2)	38626	10112	211.2	49.97	148
	(3,3)	38528	10140	211.3	49.99	147.5
	HAR	38328	10139	210.4	49.91	147.1
CCC-CAW	(1,1)	34359	8222	273.2	41.62	225.2
	(2,2)	34223	8155	273.2	41.54	224.7
	(3,3)	34371	8239	273.2	41.51	223.9
	HAR	35583	8137	273.2	42.40	204.7
EWMA		37178	9823	204.9	38.90	162.9
Factor-HEAVY	(1,1)	32374	8056	208.7	55.11	**118.3**
	(2,2)	32191	**7997**	208.0	54.88	120.3
	(3,3)	31968	8047	207.5	54.76	120.8
	HAR	31935	8053	206.8	55.00	119.9
Factor-CAW	(1,1)	35079	9238	245.4	61.22	216.5
	(2,2)	34658	9157	243.8	60.68	212.3
	(3,3)	34554	9199	243.5	60.71	211.9
	HAR	34237	9118	242.9	60.23	209.6

Table 7: Average daily 1-period ahead forecasting losses for the period 01.01.2009 – 31.12.2011. The loss functions are defined in Section 5.3.1. The smallest value is shown in bold. Grey shaded values indicate that the 90% model confidence set includes the respective model.

Calm Market: 01.01.2012 – 31.12.2014						
Model	(p,q)	Cov matrix	Var	Corr	GMVP $\times 10^2$	QLIKE
DPC-CAW	(1,1)	1586	540.4	226.2	15.48	73.96
	(2,2)	1580	539.7	226.3	15.49	74.09
	(3,3)	1578	540.1	226.4	15.48	74.32
	HAR	**1571**	538.6	227.0	15.49	74.90
DPC-CAW$_{0f}$	(1,1)	1600	542.4	225.4	15.48	**66.76**
	(2,2)	1590	540.7	225.5	15.49	66.99
	(3,3)	1585	540.2	**225.4**	**15.48**	67.05
	HAR	1577	**538.1**	225.7	15.49	67.17
Re-DCC-CAW	(1,1)	1742	597.8	242.8	16.60	94.67
	(2,2)	1732	593.3	242.7	16.59	94.37
	(3,3)	1729	592.8	242.6	16.59	94.10
	HAR	1799	586.4	244.6	16.92	74.21
O-CAW	(1,1)	1756	609.4	243.9	20.08	110.16
	(2,2)	1750	608.2	244.2	20.08	110.43
	(3,3)	1749	607.6	244.5	20.08	110.59
	HAR	1743	605.5	245.2	20.11	111.14
CCC-CAW	(1,1)	1840	597.8	265.3	18.11	96.27
	(2,2)	1828	593.3	265.3	18.12	96.18
	(3,3)	1824	592.8	265.3	18.16	96.23
	HAR	1911	586.4	265.3	18.69	72.33
EWMA		1715	556.3	239.8	15.74	82.57
Factor-HEAVY	(1,1)	1642	595.9	237.5	22.91	68.84
	(2,2)	1639	596.8	237.3	22.93	67.42
	(3,3)	1640	598.4	237.4	22.91	67.66
	HAR	1624	583.5	237.5	23.00	67.00
Factor-CAW	(1,1)	2446	770.9	299.5	41.80	296.47
	(2,2)	2455	771.9	297.4	41.19	282.06
	(3,3)	2455	772.0	296.5	41.09	277.03
	HAR	2435	769.6	299.7	41.52	292.59

Table 8: Average daily 1-period ahead forecasting losses for the period 01.01.2012 – 31.12.2014. The loss functions are defined in Section 5.3.1. The smallest value is shown in bold. Grey shaded values indicate that the 90% model confidence set includes the respective model.

Volatile Market: 01.01.2009 – 31.12.2011						
Model	(p,q)	Cov matrix	Var	Corr	GMVP $\times 10^2$	QLIKE
DPC-CAW	(1,1)	41745	10876	214.3	39.34	162.8
	(2,2)	41309	10828	213.1	39.26	162.9
	(3,3)	40643	10774	211.8	39.21	164.6
	HAR	39061	**10642**	**207.6**	39.06	164.8
DPC-CAW$_{0f}$	(1,1)	40177	10741	210.3	39.21	160.3
	(2,2)	39928	10701	209.8	39.16	160.4
	(3,3)	39719	10671	209.5	39.12	161.0
	HAR	**38883**	10643	207.8	**38.93**	162.4
Re-DCC-CAW	(1,1)	44523	14599	234.6	41.92	209.7
	(2,2)	43588	13916	234.3	41.63	206.2
	(3,3)	42829	13403	234.0	41.30	202.8
	HAR	44485	11301	242.2	44.15	176.9
O-CAW	(1,1)	46821	11917	228.7	50.93	158.3
	(2,2)	46446	11866	227.8	50.86	158.4
	(3,3)	45727	11800	226.6	50.87	159.6
	HAR	44256	11737	221.8	50.36	159.2
CCC-CAW	(1,1)	46200	14599	274.2	43.68	232.4
	(2,2)	45294	13981	274.2	43.52	231.0
	(3,3)	44509	13403	274.2	43.34	227.0
	HAR	45776	11301	274.2	46.38	193.4
EWMA		39319	10726	213.0	40.10	179.1
Factor-HEAVY	(1,1)	46365	15265	240.5	58.82	133.2
	(2,2)	45124	14320	238.5	57.62	143.3
	(3,3)	43859	13690	236.7	57.43	142.0
	HAR	42659	12376	239.5	62.98	**120.4**
Factor-CAW	(1,1)	43528	11866	254.7	62.09	217.7
	(2,2)	42886	11713	252.2	61.10	211.0
	(3,3)	42600	11630	251.2	60.50	207.2
	HAR	41240	11302	248.4	59.02	197.4

Table 9: Average daily 5-period ahead forecasting losses for the period 01.01.2009 – 31.12.2011. The loss functions are defined in Section 5.3.1. The smallest value is shown in bold. Grey shaded values indicate that the 90% model confidence set includes the respective model.

Calm Market: 01.01.2012 – 31.12.2014						
Model	(p,q)	Cov matrix	Var	Corr	GMVP $\times 10^2$	QLIKE
DPC-CAW	(1,1)	1978	595.6	245.0	16.47	69.76
	(2,2)	1933	592.4	243.6	16.39	71.55
	(3,3)	1879	585.4	242.6	16.29	75.90
	HAR	1876	583.2	**241.4**	**16.24**	75.02
DPC-CAW$_{0f}$	(1,1)	1927	592.5	243.4	16.39	68.13
	(2,2)	1909	589.6	243.1	16.37	68.52
	(3,3)	1895	587.2	242.9	16.34	69.12
	HAR	1957	590.6	245.5	16.35	68.18
Re-DCC-CAW	(1,1)	2121	694.8	247.4	17.57	90.19
	(2,2)	2066	670.6	247.2	17.51	88.57
	(3,3)	2029	651.7	247.2	17.43	86.91
	HAR	2347	640.3	249.0	17.79	73.29
O-CAW	(1,1)	2055	651.5	252.3	20.09	102.06
	(2,2)	2024	645.9	252.3	20.18	104.34
	(3,3)	1975	639.7	254.7	20.18	110.69
	HAR	1990	643.7	251.1	20.05	106.14
CCC-CAW	(1,1)	2241	694.8	265.5	19.10	87.23
	(2,2)	2194	675.5	265.5	19.12	86.32
	(3,3)	2140	651.7	265.5	19.12	85.22
	HAR	2516	640.3	265.5	19.71	**67.64**
EWMA		**1798**	**568.3**	248.2	16.31	90.37
Factor-HEAVY	(1,1)	2001	710.2	258.9	24.89	82.52
	(2,2)	1962	683.4	258.2	24.40	85.38
	(3,3)	1921	661.5	257.6	24.31	84.34
	HAR	1849	610.8	256.9	26.15	74.36
Factor-CAW	(1,1)	2659	789.2	298.8	41.41	260.71
	(2,2)	2684	791.2	294.7	40.16	239.35
	(3,3)	2693	791.5	293.2	39.85	230.35
	HAR	2694	789.2	295.3	40.54	234.65

Table 10: Average daily 5-period ahead forecasting losses for the period 01.01.2012 – 31.12.2014. The loss functions are defined in Section 5.3.1. The smallest value is shown in bold. Grey shaded values indicate that the 90% model confidence set includes the respective model.

Volatile Market: 01.01.2009 – 31.12.2011						
Model	(p,q)	Cov matrix	Var	Corr	GMVP $\times 10^2$	QLIKE
DPC-CAW	(1,1)	42407	11254	221.5	40.48	171.3
	(2,2)	42291	11228	220.6	40.36	172.0
	(3,3)	42256	11194	220.0	40.38	175.3
	HAR	**40828**	11211	**214.7**	40.08	173.6
DPC-CAW$_{0f}$	(1,1)	41765	11145	218.7	40.35	169.4
	(2,2)	41672	11126	218.2	40.28	169.6
	(3,3)	41618	**11123**	217.8	40.20	170.2
	HAR	41011	11263	215.6	**39.80**	170.1
Re-DCC-CAW	(1,1)	48946	17292	239.7	45.08	221.7
	(2,2)	47227	15863	239.4	44.57	218.1
	(3,3)	45779	14767	239.1	44.00	214.1
	HAR	48323	12257	247.8	46.78	187.9
O-CAW	(1,1)	46903	12044	233.0	50.63	164.3
	(2,2)	46752	11996	232.3	50.61	164.6
	(3,3)	46622	11937	231.9	50.69	167.5
	HAR	45254	11974	226.5	50.13	165.3
CCC-CAW	(1,1)	50519	17292	274.8	46.29	243.8
	(2,2)	48971	16108	274.8	46.01	243.3
	(3,3)	47281	14767	274.8	45.52	237.4
	HAR	49293	12257	274.8	48.49	201.3
EWMA		41217	11347	223.3	40.92	197.2
Factor-HEAVY	(1,1)	53993	21772	263.0	62.15	149.8
	(2,2)	50542	18477	259.3	60.19	160.6
	(3,3)	48442	16832	256.0	60.35	158.6
	HAR	46241	14408	255.0	64.94	**136.1**
Factor-CAW	(1,1)	44877	12192	257.5	64.27	221.0
	(2,2)	44265	12067	254.5	62.69	212.6
	(3,3)	44012	12008	253.3	61.67	207.1
	HAR	42949	11845	253.9	60.70	203.1

Table 11: Average daily 10-period ahead forecasting losses for the period 01.01.2009 – 31.12.2011. The loss functions are defined in Section 5.3.1. The smallest value is shown in bold. Grey shaded values indicate that the 90% model confidence set includes the respective model.

Calm Market: 01.01.2012 – 31.12.2014						
Model	(p,q)	Cov matrix	Var	Corr	GMVP $\times 10^2$	QLIKE
DPC-CAW	(1,1)	2368	640.9	258.4	17.17	**68.59**
	(2,2)	2250	630.9	254.7	17.04	70.50
	(3,3)	2107	614.2	250.8	16.84	76.54
	HAR	2104	612.0	**249.5**	16.78	76.46
DPC-CAW$_{0f}$	(1,1)	2235	632.1	253.5	16.94	69.15
	(2,2)	2189	625.4	252.8	16.91	69.44
	(3,3)	2143	618.6	251.9	16.87	69.91
	HAR	2220	623.4	254.6	16.87	69.67
Re-DCC-CAW	(1,1)	2528	761.7	250.7	17.98	93.11
	(2,2)	2402	722.2	250.5	17.93	90.65
	(3,3)	2320	694.5	250.5	17.84	88.76
	HAR	2685	687.7	252.1	18.34	79.24
O-CAW	(1,1)	2358	692.5	255.8	20.03	97.69
	(2,2)	2275	678.4	255.5	20.16	100.57
	(3,3)	2142	662.6	257.1	20.14	109.60
	HAR	2180	667.5	254.7	20.11	104.91
CCC-CAW	(1,1)	2662	761.7	265.7	19.37	87.89
	(2,2)	2558	732.8	265.7	19.39	86.99
	(3,3)	2438	694.5	265.7	19.38	84.93
	HAR	2859	687.7	265.7	19.90	73.43
EWMA		**1889**	**577.1**	255.0	**16.77**	98.88
Factor-HEAVY	(1,1)	2299	828.1	263.8	26.35	91.51
	(2,2)	2209	769.0	263.3	25.34	94.09
	(3,3)	2122	727.8	262.7	25.37	91.42
	HAR	2006	654.6	261.0	26.95	81.64
Factor-CAW	(1,1)	2818	804.0	292.7	40.13	224.23
	(2,2)	2867	807.5	287.9	38.04	201.12
	(3,3)	2882	807.9	286.3	37.52	192.07
	HAR	2821	800.4	292.4	39.52	216.50

Table 12: Average daily 10-period ahead forecasting losses for the period 01.01.2012 – 31.12.2014. The loss functions are defined in Section 5.3.1. The smallest value is shown in bold. Grey shaded values indicate that the 90% model confidence set includes the respective model.

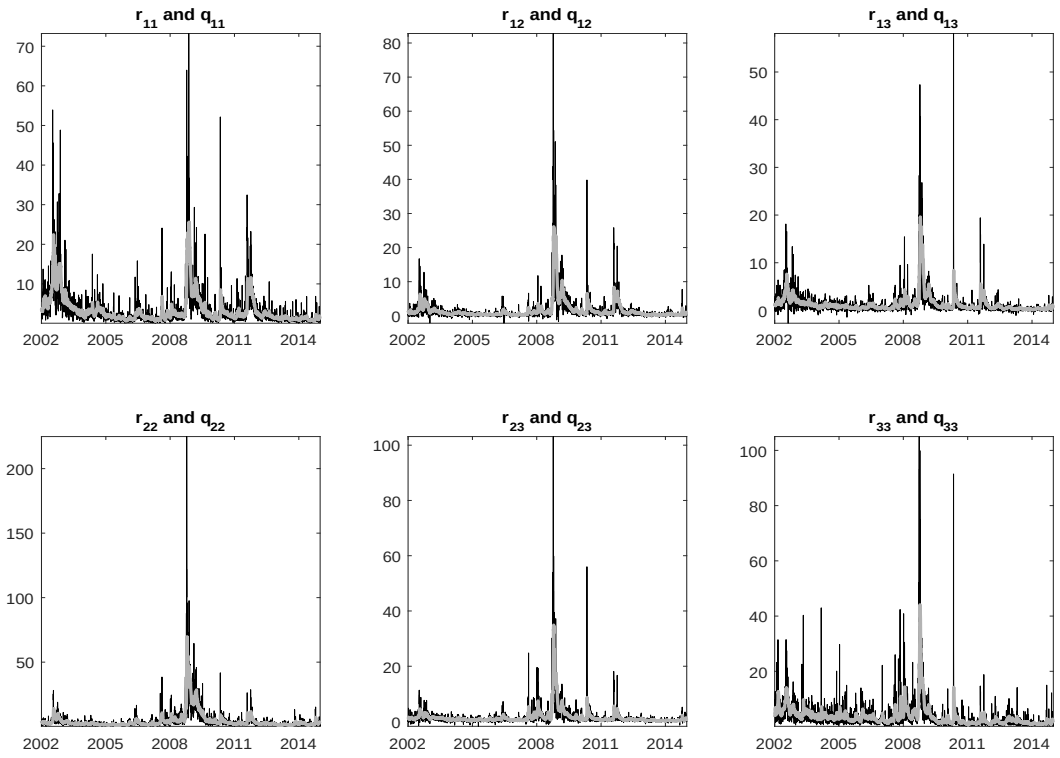

Figure 1: Black line: realized variances and covariances $r_{i,j}$ of A ($i = 1$), AA ($i = 2$) and AAPL ($i = 3$); Gray line: estimates of the individual Q_t elements obtained via one-step QML estimation of the DPC-CAW model as specified in Section 3 to the according set of 3-dimensional realized covariance matrices.

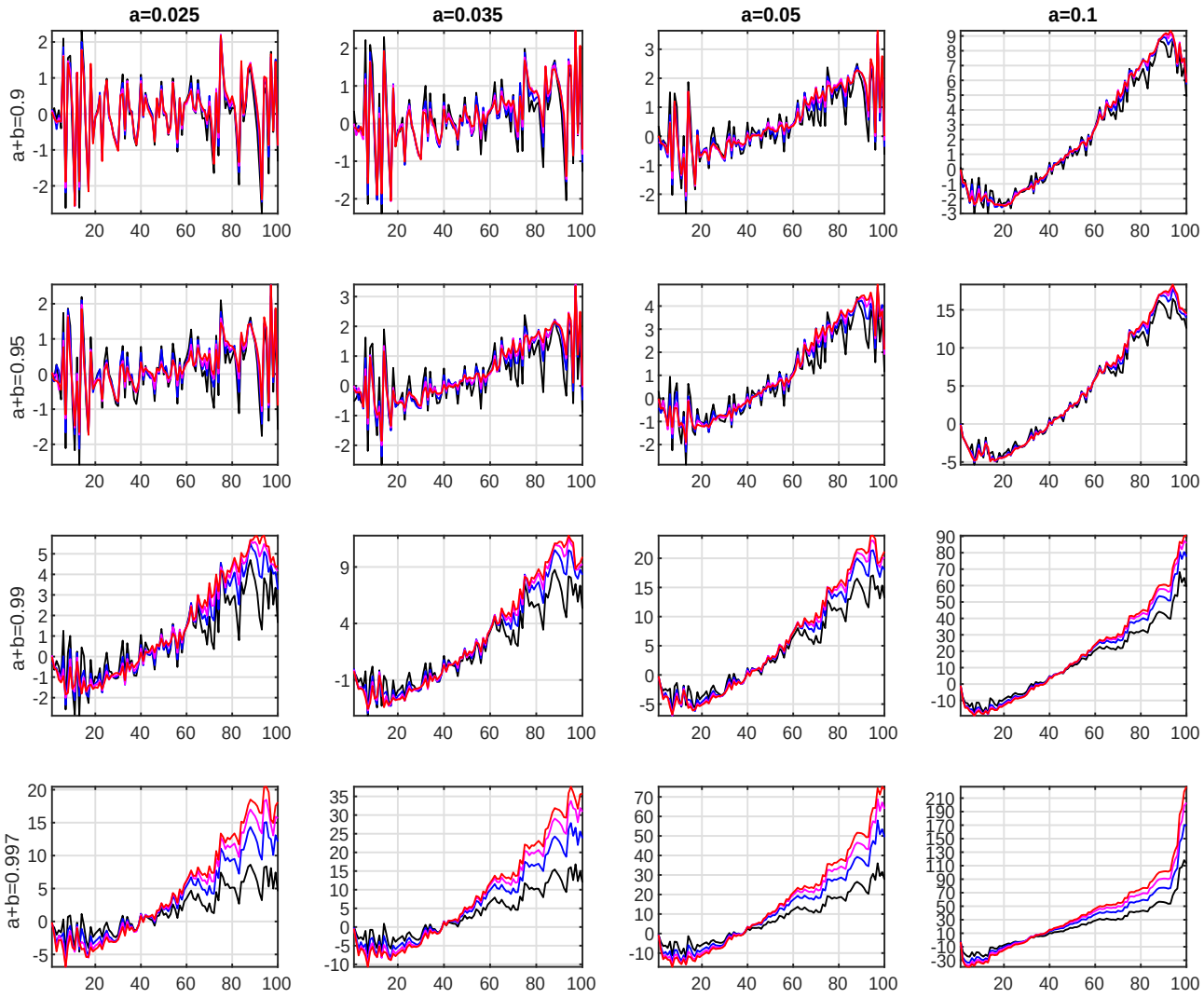

Figure 2: Average relative estimation errors for d_i $(i = 1, \ldots, 100)$ computed as arithmetic mean $\frac{1}{500}\sum_{j=1}^{500} 100 \cdot (\hat{d_{i,j}} - d_{i,j})/d_{i,j}$ over the estimates obtained for the 500 simulations. Black: $T = 1000$; blue: $T = 2500$; magenta: $T = 5000$; red: $T = 10000$. The results are obtained from the DPC estimator for the simulation experiment of Section 4. The DGP parameter values are reported at the top of the panel for a and on the left side of the panel for $(a + b)$. Each line comprises 100 data points, one for each d_i in descending order with d_1 being displayed on the left.

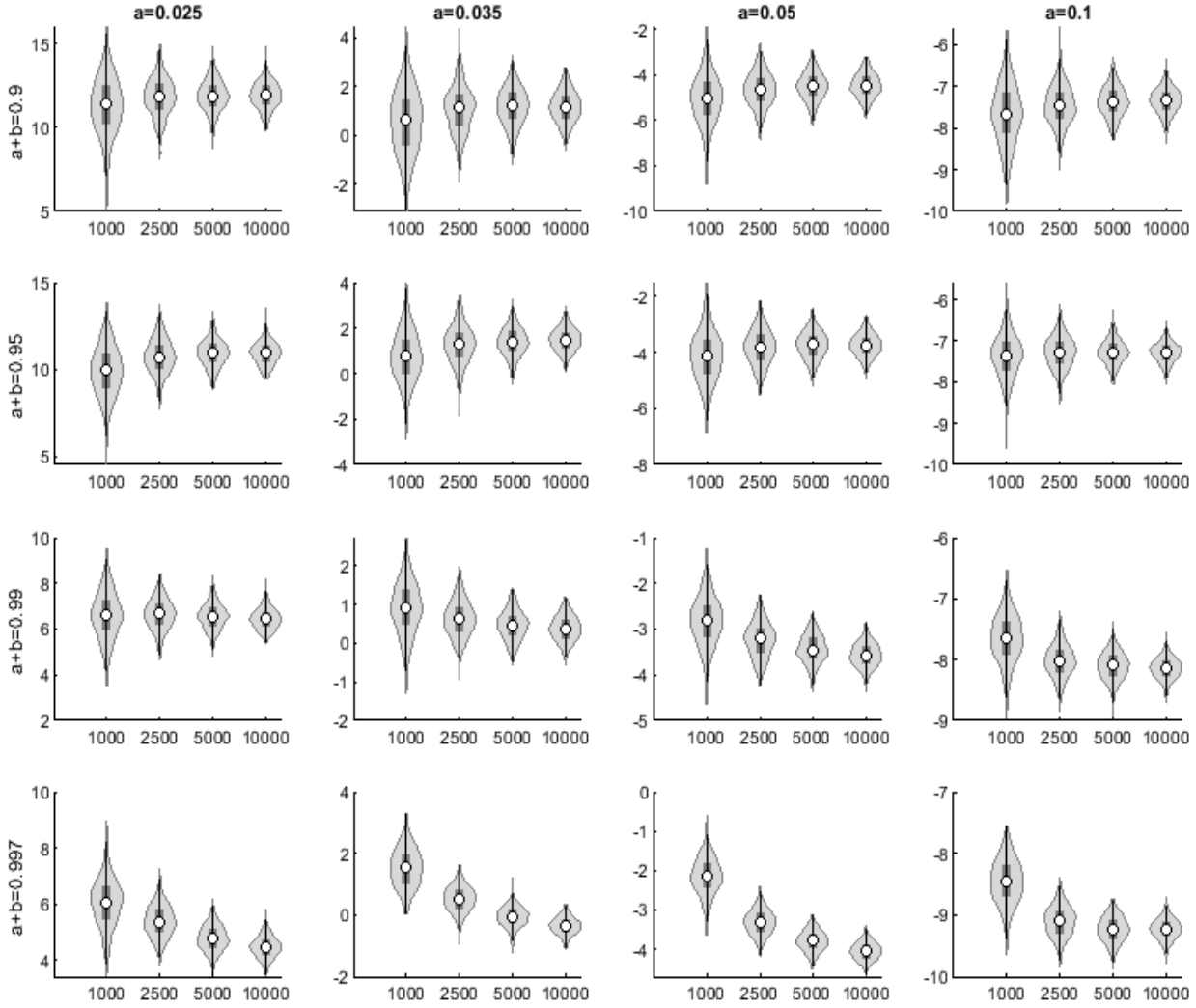

Figure 3: Violin plots of relative estimation errors $100 \cdot (\hat{a} - a)/a$ obtained from the DPC estimator for the simulation experiment of Section 4. The DGP parameter values are reported at the top of the panel for a and on the left side of the panel for $(a + b)$. The first violin plot in each subplot comprises results for $T = 1000$, the second for $T = 2500$, the third for $T = 5000$ and the fourth for $T = 10000$. The white dot within the box indicates the median.

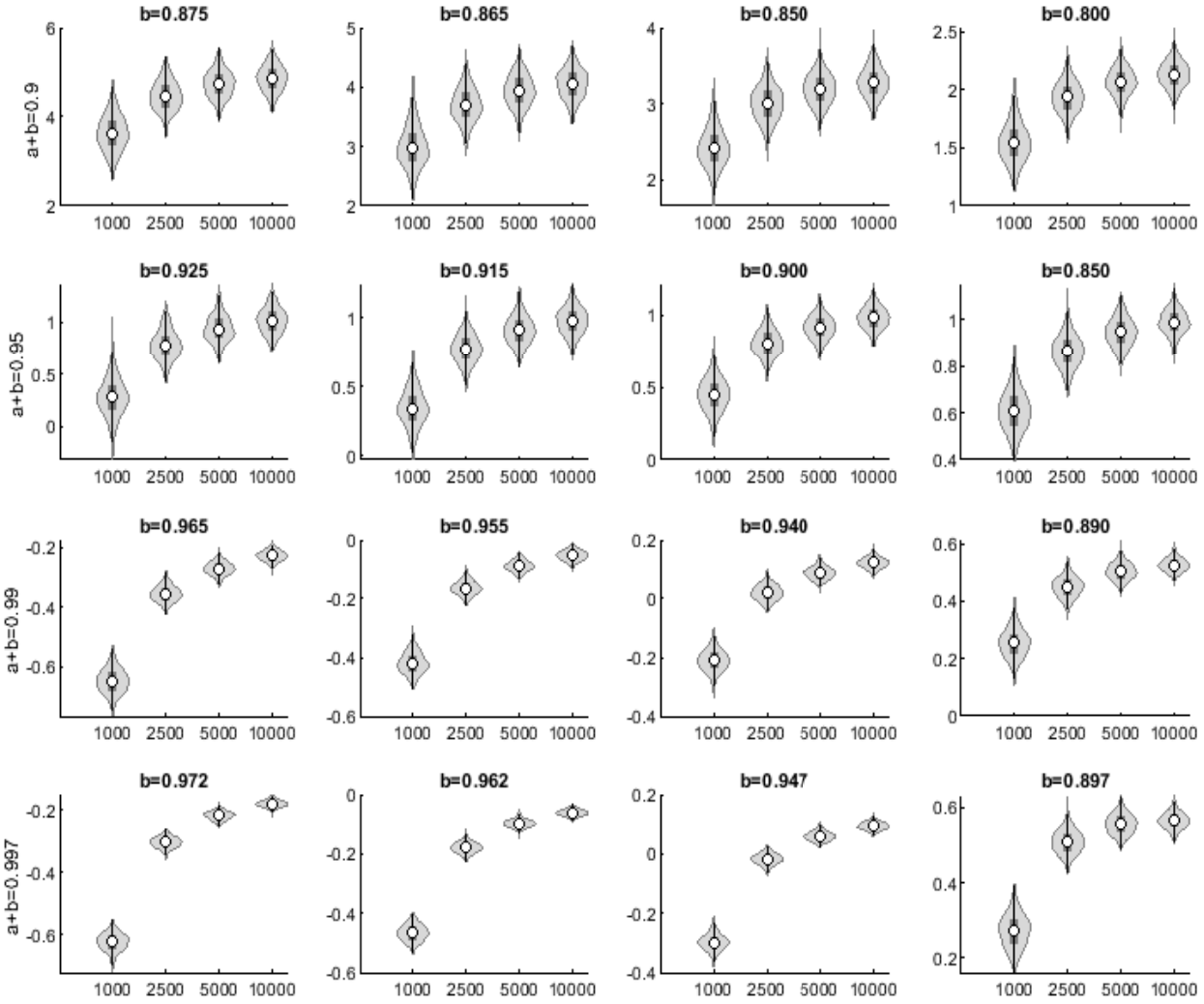

Figure 4: Violin plots of relative estimation errors $100 \cdot (\hat{b} - b)/b$ obtained from the DPC estimator for the simulation experiment of Section 4. The DGP parameter values are reported at the top of the panel for b and on the left side of the panel for $(a + b)$. The first violin plot in each subplot comprises results for $T = 1000$, the second for $T = 2500$, the third for $T = 5000$ and the fourth for $T = 10000$. The white dot within the box indicates the median.

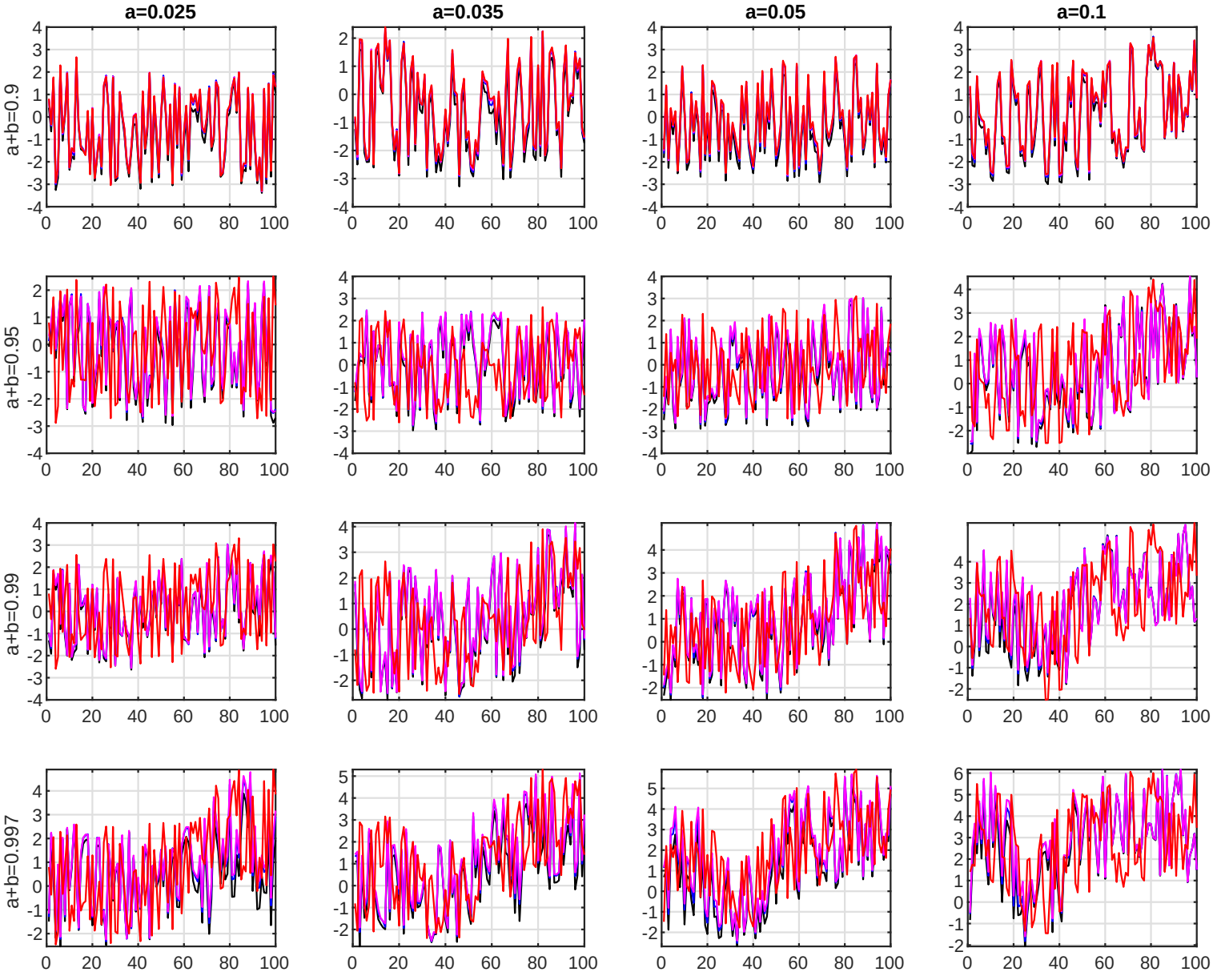

Figure 5: Average relative estimation errors for the eigenvalue persistence $\alpha_i + \beta_i$ computed as arithmetic mean $\frac{1}{500}\sum_{j=1}^{500} 100 \cdot (\widehat{\alpha_{i,j} + \beta_{i,j}} - (\alpha_{i,j} + \beta_{i,j}))/(\alpha_{i,j} + \beta_{i,j})$ over the estimates obtained for the 500 simulations. Black: $T = 1000$; blue: $T = 2500$; magenta: $T = 5000$; red: $T = 10000$. The results are obtained from the DPC estimator for the simulation experiment of Section 4. The DGP parameter values are reported at the top of the panel for a and on the left side of the panel for $(a + b)$. Each line comprises 100 data points, one for each $\alpha_i + \beta_i$ in descending order with $\alpha_1 + \beta_1$ corresponding to the highest eigenvalue being displayed on the left.

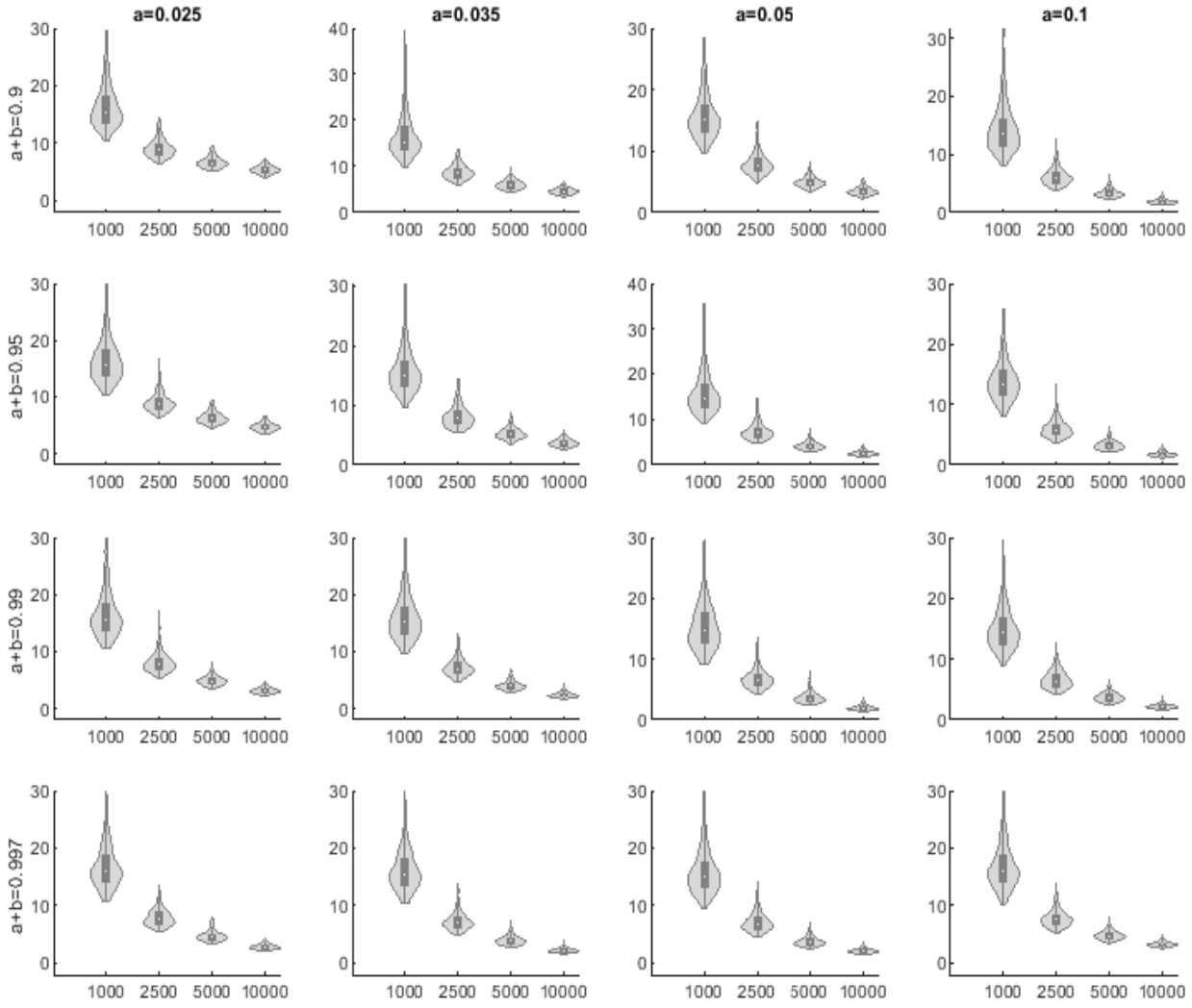

Figure 6: Violin plots of average Euclidean norms $(1/T)\sum_t \text{vech}(\hat{S}_t - S_t)'\text{vech}(\hat{S}_t - S_t)$ over the 500 simulations. The DGP parameter values are reported at the top of the panel for a and on the left side of the panel for $(a+b)$. The first violin plot in each subplot comprises results for $T = 1000$, the second for $T = 2500$, the third for $T = 5000$ and the fourth for $T = 10000$. The white dot within the box indicates the median. The results are obtained from the DPC estimator for the simulation experiment of Section 4.

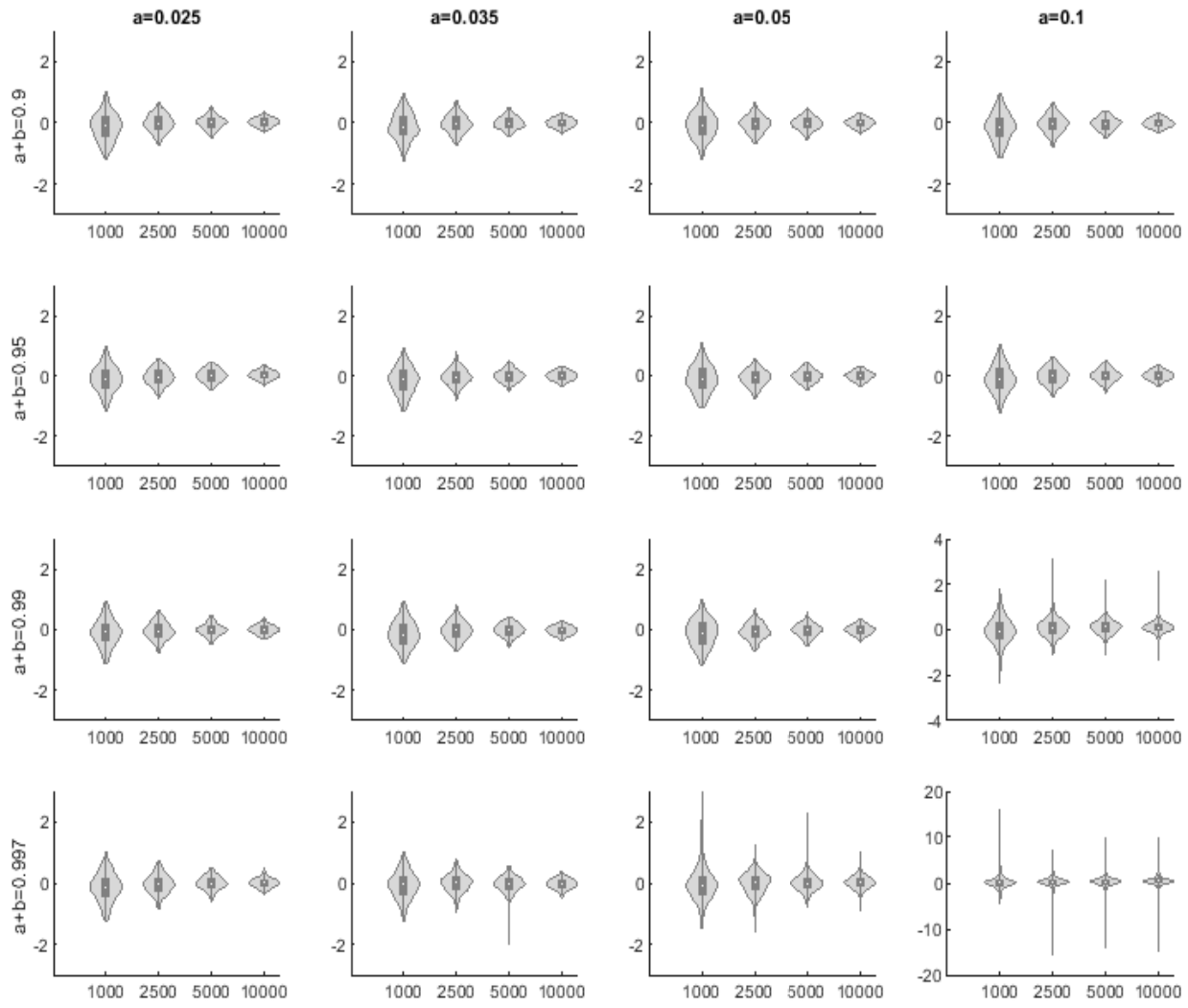

Figure 7: Violin plots of average relative biases $(1/T)\sum_t(1/5050)\sum_{i>j}(\hat{S}_{ij}-S_{ij})/S_{ij}$ over the 500 simulations. The DGP parameter values are reported at the top of the panel for a and on the left side of the panel for $(a+b)$. The first violin plot in each subplot comprises results for $T=1000$, the second for $T=2500$, the third for $T=5000$ and the fourth for $T=10000$. The white dot within the box indicates the median. The results are obtained from the DPC estimator for the simulation experiment of Section 4.

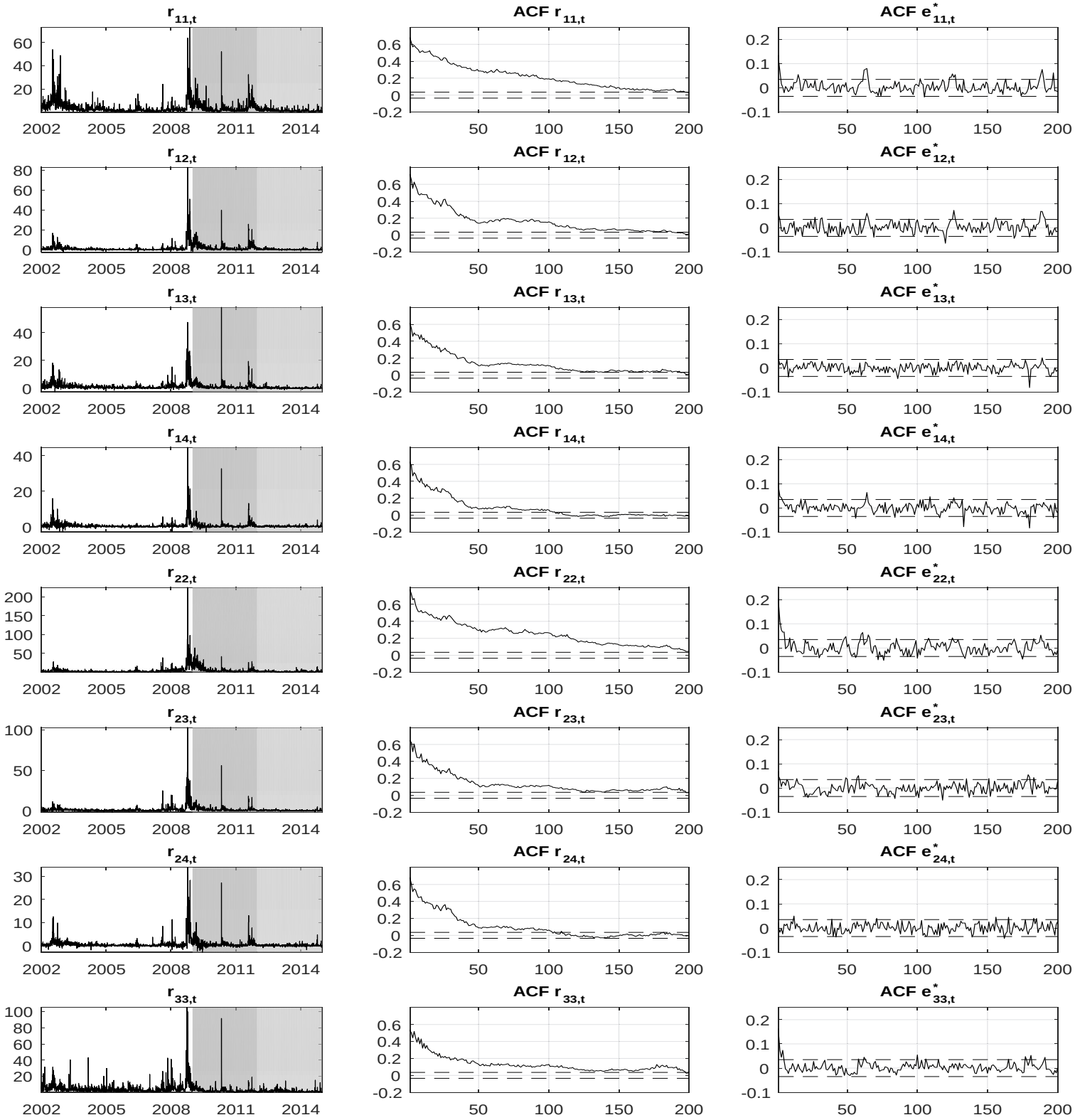

Figure 8: Realized (co)variance plots and sample autocorrelation functions (ACFs). Left panel: Sample of realized variances and covariances $r_{i,j}$ of A ($i = 1$), AA ($i = 2$), AAPL ($i = 3$) and ABT ($i = 4$). Gray shaded areas indicate the periods covered by the forecasting experiment of Section 5.3. Middle panel: Sample ACFs of realized (co)variances together with 95% confidence bounds under the null of zero serial correlation. Right panel: Sample ACFs and according 95% confidence bounds of standardized Pearson residuals obtained from the BIC selected DPC-CAW(3,4)-(1,1) model estimated by the DPC estimator for the 100-dim. data-set illustrated in Section 5.