Journal of Econometrics 241 (2024) 105747

Contents lists available at ScienceDirect

Journal of Econometrics

journal homepage: www.elsevier.com/locate/jeconom

Dynamic partial correlation models

Enzo D'Innocenzo [a,*], Andre Lucas [b]

[a] Università di Bologna, Italy
[b] Vrije Universiteit Amsterdam and Tinbergen Institute, The Netherlands

ARTICLE INFO

Keywords:
Dynamic correlations
Score-driven models
Stationarity
Filter invertibility

ABSTRACT

We introduce a new scalable model for dynamic conditional correlation matrices based on a recursion of dynamic bivariate partial correlation models. By exploiting the model's recursive structure and the theory of perturbed stochastic recurrence equations, we establish stationarity, ergodicity, and filter invertibility in the multivariate setting using conditions for bivariate slices of the data only. From this, we establish consistency and asymptotic normality of the maximum likelihood estimator for the model's static parameters. The new model outperforms benchmarks like the t-cDCC and the multivariate t-GAS, both in simulations and in an in-sample and out-of-sample asset pricing application to US stock returns.

1. Introduction

Modeling multivariate covariance and correlation structures is a well-established research topic in the econometric literature; see for instance the overviews of Bauwens et al. (2012) and Francq and Zakoian (2019). Since Engle (2002) and Tse and Tsui (2002), the typical way to empirically model the time-variation in conditional covariance matrices is via its decomposition into a variance and a correlation related component; see, e.g., the standard benchmark DCC model of Engle (2002). The DCC first accounts for time-variation in the variances of the individual time series, and then investigates whether there is any time-variation left in the correlations. The second step of this approach requires a model for the dynamics of the conditional correlation matrix.

Modeling the dynamics of conditional correlation matrices is challenging given the joint restrictions that should hold for such matrices: they (i) need to be positive (semi)-definite and (ii) have ones on the diagonal. The DCC of Engle (2002) ensures this by an algorithmically simple, but theoretically hard, non-linear matrix-transformation. The drawback of this approach is that it becomes hard to formulate conditions for stationarity, ergodicity, and filter invertibility. Establishing such stochastic properties is crucial for dynamic models: it opens the door to a rigorous econometric analysis of the asymptotic statistical properties of such models. For example, due to the complexity of the non-linear matrix transformation in the DCC, the asymptotic properties of the quasi maximum likelihood estimator for the DCC are as yet still unknown; see also the related (heuristic) discussion in Aielli (2013).

Alternative parameterizations of correlation matrices have been proposed in the literature. These include the hypersphere parameterization of the Cholesky decomposition of a correlation matrix as in Rapisarda et al. (2007), Creal et al. (2011), and Buccheri et al. (2021), or the log correlation matrix transformation of Archakov and Hansen (2021) as also used by Hafner and Wang (2021). These parameterizations result automatically in positive (semi) definite matrices with ones on the diagonal. For all these parameterizations, however, formulating conditions for filter invertibility remains hard. In addition, all these models are cast in

* Correspondence to: Department of Economics, Università di Bologna, Piazza Antonio Scaravilli, 2, 40126 Bologna, Italy.
E-mail address: enzo.dinnocenzo2@unibo.it (E. D'Innocenzo).

https://doi.org/10.1016/j.jeconom.2024.105747
Received 27 September 2022; Received in revised form 23 November 2023; Accepted 12 April 2024

matrix format, which means that contraction conditions like that of Bougerol (1993) can become increasingly strict in higher dimensions due to the use of matrix norms.[1]

In this paper, we contribute to the literature by introducing a novel class of non-linear, heavy-tailed time-series models for dynamic conditional correlation matrices. The new model avoids most of the drawbacks mentioned earlier. In particular, instead of considering a full multivariate model for the entire dynamic conditional correlation matrix at once, we define univariate nonlinear filters for conditional *partial* correlation coefficients based on bivariate slices of the data only. This also allows us to easily impose zero restrictions on particular partial correlations in case this is theoretically or empirically desirable. By stacking the different bivariate models and relying on Anderson (1958) and Joe (2006), we can recursively reconstruct the full multivariate dynamic correlation matrix. The matrix constructed in this way automatically has ones on the diagonal and satisfies the restrictions of positive-definiteness. The idea has also been used in for instance (Barthel et al., 2020) to model realized covariance matrices and forecast the individual Fisher transforms of the realized partial correlations using HAR and ARFIMA models. In our setting, we do not observe realized correlations or variances, but only the daily return data. Our model therefore uses a non-linear time series set-up based on a similar vine structure to model the correlation dynamics of the observed returns directly. In addition, we are also able to provide a full asymptotic analysis of the model and its maximum likelihood estimator.

We endow the models for the bivariate data slices with score-driven dynamics for the univariate partial correlation parameter. We assume the multivariate data have a Student's t distribution, such that the bivariate data slices have a conditional Student's t distribution with parameters known from the multivariate distribution. These pairwise distributions give rise directly to the score-driven dynamics. In this way we obtain a robust filter for the entire correlation matrix; see Creal et al. (2013) and Harvey (2013) for an introduction to score-driven dynamics. As the full model is a cascade of bivariate models, each of these bivariate models can in principle use its own parameters to govern the dynamics of that pair-specific filtering equation. This flexibility can be empirically relevant as shown in our application to US stock returns in Section 4. We stress that the flexibility of all these pairwise model does not jeopardize the positive definiteness of the implied multivariate conditional Pearson correlation matrix for the entire system.

Splitting the modeling approach from a multivariate problem into a recursion of conditional models for bivariate slices of the data not only provides benefits from a computational or model design perspective. We show that the approach also leads to advantages for a rigorous theoretical analysis of the model's asymptotic properties; compare (Blasques et al., 2022). We consider an asymptotic setting where the sample size T goes to infinity for a fixed dimension N of the time series, and leave a setting with both N and T going to infinity to a future paper. By using the theory on perturbed stochastic recurrence equations of Straumann and Mikosch (2006), we are able to provide clear conditions for stationarity, ergodicity, and filter invertibility, as well as conditions for consistency and asymptotic normality of the maximum likelihood estimator. All these conditions only make use of univariate contraction requirements based on bivariate data slices, even if the dimension of the entire data vector is substantially larger than two. An important advantage of this approach is that the restrictions can be more relaxed than dealing with the entire multivariate system at once; see also Footnote 1. In essence, we prove that the conditions for bivariate models like (Blasques et al., 2018b) continue to hold in slightly modified form for the fully multivariate setting. Similar rigorous results for these non-linear correlation models were not available before. We also mention that due to the use of a robust filtering method, we only require a limited ($2+\delta$ for some small $\delta > 0$) number of moments for the observations in order for the model and filter to behave well. This stands in sharp contrast with the asymptotic theory developed for MGARCH models, like in the BEKK-GARCH models where at least 6-order moments of the observations may be required; see Comte and Lieberman (2003), Hafner and Preminger (2009), and Pedersen and Rahbek (2014).

The new model performs well in a controlled simulation setting, where it outperforms typical strong benchmarks like the cDCC of Engle (2002) and Aielli (2013) based on the Student's t distribution, the t-GAS model with hypersphere parameterization of Creal et al. (2011) and Buccheri et al. (2021), and the recursive SCC model of Palandri (2009). We also apply the model both in-sample and out-of-sample to study its asset pricing implications for time-series of US stock returns over the period 1980–2021 as in Engle (2016), Boudt et al. (2017) and Darolles et al. (2018). The empirical application considers time-varying betas in a risk attribution model with a market (MKT - RF), size (SMB), and value (HML) risk factor and assesses performance in terms of tracking errors rather than statistical measures of fit only. We consider the risk factors jointly with an industry return series in a four-dimensional system (for 12 industries), as well as a 23-dimensional setting with 20 stocks and the three risk factors. The results reveal that the new model outperforms its benchmarks both in-sample and out of sample. The dynamic partial correlation model enters the model confidence set (MCS) of Hansen et al. (2011) most often compared to the other benchmarks.

The rest of this paper is organized as follows. Section 2 introduces the model. The asymptotic properties of the model are derived in Section 3. Section 4 provides the empirical application. Section 5 concludes. All proofs and other supplementary materials are provided in the online Appendix.

2. The model

2.1. Approaches to modeling correlation matrices

Consider a real-valued N-dimensional time series $\{\boldsymbol{y}_t\}_{t\in\mathbb{Z}}$ and a sequence of corresponding information sets $\mathcal{F}_{t-1} = \{\boldsymbol{y}_{t-1}, \boldsymbol{y}_{t-2}, \dots\}$. We focus on modeling the dynamics of the conditional Pearson correlation matrix $\boldsymbol{R}_t$ of $\boldsymbol{y}_t$ given $\mathcal{F}_{t-1}$. More specifically, we consider

[1] For instance, for the Frobenius norm of a matrix A we have $\|A\| = \sqrt{\sum_{i,j} \left|A_{i,j}\right|^2}$. If this norm has to be bounded in a (Bougerol, 1993) type condition, this becomes increasingly restrictive if the dimension of A grows.

the case

$$\boldsymbol{y}_t|\mathcal{F}_{t-1} \sim t\left(\boldsymbol{0}_N, \left(1-2\nu^{-1}\right)\cdot \boldsymbol{R}_t, \nu\right), \qquad \nu > 2, \tag{1}$$

where $t(\mu, \Omega, \nu)$ denotes an N-dimensional Student's t distribution with location μ, scale matrix Ω, and $\nu > 2$ degrees of freedom.[2] We assume $\boldsymbol{R}_t$ is a measurable function of $\mathcal{F}_{t-1}$, such that the model is observation-driven. The model can easily be extended to allow for a non-zero location and for non-unit variances as well as for other distributions. In addition, with a slight extension, our model can be extended into a dynamic Student's t copula framework. For expositional purposes, however, we focus on the current more constrained set-up in (1) to better highlight what is new in our approach.

As mentioned in the introduction, one of the challenges in models such as (1) is the parameterization of the dynamic conditional correlation matrix $\boldsymbol{R}_t$. The matrix $\boldsymbol{R}_t$ not only has to be positive definite, but also needs to have unit entries on the diagonal. So far, three main approaches to tackle this issue have been put forward in the literature. The first approach is that of Engle (2002). It models the covariance matrix directly and standardizes it by pre- and post-multiplying by the square root inverse of its diagonal to ensure the correlation matrix structure with unit entries on the diagonal. A second approach casts the correlation matrix entries into hypersphere coordinates and models the dynamic behavior of these spherical coordinates rather than of the original correlations themselves; see Rapisarda et al. (2007), Creal et al. (2011), and Buccheri et al. (2021). Finally, Archakov and Hansen (2021) introduce the possibility of modeling the strictly lower-half of the log-correlation matrix. Separate models can be used for each of these unconstrained entries. Putting the individual entries back into a matrix and taking the matrix exponential, one automatically recovers a proper correlation matrix. This approach is extended to a dynamic setting by Hafner and Wang (2021) using score-driven dynamics. These non-linear re-parameterizations used to obtain proper correlation matrices complicate a rigorous analysis of the asymptotic properties of these models.

In this paper, we do not consider the Pearson correlations themselves, but instead consider the dynamics of pairwise *partial* correlations using the work of Anderson (1958) and Joe (2006). As a result, we need not worry about positive definiteness of the implied full correlation matrix: pairwise partial correlation coefficients can be modeled freely and independently with the only restriction being that the partial correlations lie in the interval $(-1, 1)$. As long as all pairwise partial correlations (as defined further below) lie in this interval, the implied full Pearson correlation matrix is always a proper correlation matrix. As we see later, this has important advantages, both in terms of the flexibility of the model construction, the model's computational and stability aspects, its theoretical statistical properties, and its empirical performance.

2.2. From partial correlations to correlation matrices

A conditional partial correlation $\rho_{i,j|L_{ij};t}$ for a set of indices L_{ij} with $i, j \notin L_{ij}$ is defined as the correlation between $\boldsymbol{y}_{i,t}$ and $\boldsymbol{y}_{j,t}$, conditional on $\mathcal{F}_{t-1}$ and on $\boldsymbol{y}_{L_{ij},t}$, where $\boldsymbol{y}_{L_{ij},t}$ is a vector containing the values of $\boldsymbol{y}_{k,t}$ for $k \in L_{ij}$. In this paper we follow Joe (2006) and use $i < k < j$, with $L_{ij} = \emptyset$ for $j = i+1$ and the conditional partial correlation collapsing to the standard conditional Pearson correlation coefficient.[3] Joe (2006) notes that every $N \times N$ correlation matrix can be parameterized in terms of $N(N-1)/2$ partial correlation parameters. The first $N-1$ parameters are standard pairwise Pearson conditional correlations $\rho_{i,i+1;t}$ for $i = 1, \dots, N-1$ and $L_{ij} = \emptyset$. The remaining $(N-2)(N-1)/2$ parameters are the conditional partial correlations $\rho_{i,j|L_{ij};t}$ for $L_{ij} = \{i+1, \dots, j-1\}$ for $i = 1, \dots, N-1$ and $j = i+1, \dots, N$, i.e., the conditional partial correlations between $\boldsymbol{y}_{i,t}$ and $\boldsymbol{y}_{j,t}$ conditioning on all intermediate coordinates k between i and j for $i < k < j$.

Define $\boldsymbol{V}_{i,j;t} = \rho_{i,j;t} - \boldsymbol{R}_{i,L_{ij};t}\boldsymbol{R}^{-1}_{L_{ij},L_{ij};t}\boldsymbol{R}_{L_{ij},j;t}$. Then the link between pairwise and partial correlations is obtained from Anderson (1958) and Joe (2006) via the recursive formula

$$\rho_{i,j|L_{ij};t} = \frac{\rho_{i,j;t} - \boldsymbol{R}_{i,L_{ij};t}\boldsymbol{R}^{-1}_{L_{ij},L_{ij};t}\boldsymbol{R}_{L_{ij},j;t}}{\sqrt{\left(1-\boldsymbol{R}_{i,L_{ij};t}\boldsymbol{R}^{-1}_{L_{ij},L_{ij};t}\boldsymbol{R}_{L_{ij},i;t}\right)\cdot\left(1-\boldsymbol{R}_{j,L_{ij};t}\boldsymbol{R}^{-1}_{L_{ij},L_{ij};t}\boldsymbol{R}_{L_{ij},j;t}\right)}} = \frac{\boldsymbol{V}_{i,j|L_{ij};t}}{\sqrt{\boldsymbol{V}_{i,i|L_{ij};t}\cdot\boldsymbol{V}_{j,j|L_{ij};t}}}, \tag{2}$$

for $i = 1, \dots, N-1$, $j = i+1, \dots, N$, and $L_{ij} = \{i+1, \dots, j-1\}$, where

$$\operatorname{corr}(\boldsymbol{y}_{i:j;t}) = \begin{bmatrix} 1 & \boldsymbol{R}_{i,L_{ij};t} & \rho_{i,j;t} \\ \boldsymbol{R}_{L_{ij},i;t} & \boldsymbol{R}_{L_{ij},L_{ij};t} & \boldsymbol{R}_{L_{ij},j;t} \\ \rho_{i,j;t} & \boldsymbol{R}_{j,L_{ij};t} & 1 \end{bmatrix}, \tag{3}$$

and $\boldsymbol{y}_{i:j;t} = (\boldsymbol{y}_{i,t}, \dots, \boldsymbol{y}_{j,t})^\top$. Inverting (2), we easily obtain the Pearson correlation $\rho_{i,j;t}$ as a function of the partial correlation $\rho_{i,j|L_{ij};t}$, and the Pearson correlations in $\boldsymbol{R}_{i,L_{ij};t}$, $\boldsymbol{R}_{j,L_{ij};t}$, and $\boldsymbol{R}_{L_{ij},L_{ij};t}$:

$$\rho_{i,j;t} = \boldsymbol{R}_{i,L_{ij};t}\boldsymbol{R}^{-1}_{L_{ij},L_{ij};t}\boldsymbol{R}_{L_{ij},j;t} + \rho_{i,j|L_{ij};t}\sqrt{\boldsymbol{V}_{i,i|L_{ij};t}\cdot\boldsymbol{V}_{j,j|L_{ij};t}}. \tag{4}$$

[2] Alternatively, we could use $\boldsymbol{R}_t$ as a scaling matrix and relax subsequent moment conditions even further. The current parameterization with $\nu > 2$, however, allows us to interpret $\boldsymbol{R}_t$ directly as a real Pearson correlation matrix.

[3] There are many different ways to construct the full correlation matrix from a sequence of pairs. In the main text, we adhere to the original proposal of Joe (2006) to simplify the notation. All results remain valid, however, for alternative vine structures to decompose the multivariate process.

Interestingly, as Joe (2006) points out, the $N-1$ pairwise correlations and the $(N-2)(N-1)/2$ partial correlations can vary independently in the interval $(-1,1)$. The implied Pearson correlation matrix will always be positive definite by construction. Thus, by modeling the dynamics of the partial correlations, we can use (4) to obtain a dynamic positive definite conditional correlation matrix $\boldsymbol{R}_t$ for all t.

A major advantage of parameterizing a correlation matrix in terms of its partial correlations is that we only have to consider bivariate relationships. The full multivariate nature of the problem can be deferred until we have to evaluate the full likelihood function, if so desired. In addition, parameter restrictions on the dynamic partial correlations take a much simpler form than when dealing with the entire matrix $\boldsymbol{R}_t$ in one step. Finally, estimating a sequence of bivariate models can lead to computational gains compared to a fully-fledged likelihood optimization of the multivariate model, if only to obtain good starting values for the latter. We return to this recursive way of estimating the multivariate model by a cascade of bivariate estimations in Section 2.5.

2.3. Dynamic specification of the partial correlations

As a tool to describe the dynamics of the correlation matrix $\boldsymbol{R}_t$ via its partial correlations, we use score-driven dynamics as introduced by Creal et al. (2013) and Harvey (2013). For a hypersphere and a log correlation matrix parameterization this was done (Creal et al., 2011) and (Hafner and Wang, 2021), respectively. In our setting, however, we do not require the matrix-valued full $\boldsymbol{R}_t$, but only work with bivariate partial correlations instead.

The key step in making our approach feasible and scalable is obtained by observing that for $j>i$ the conditional distribution of $(\boldsymbol{y}_{i,t},\boldsymbol{y}_{j,t})^\top$ in (1) conditional on $\mathcal{F}_{t-1}$ *and* $\boldsymbol{y}_{L_{ij},t}=\{\boldsymbol{y}_{k,t}\}_{k\in L_{ij}}$, i.e., on all intermediate coordinates, is Student's t

$$\begin{pmatrix}\boldsymbol{y}_{i,t}\\ \boldsymbol{y}_{j,t}\end{pmatrix} \mid \mathcal{F}_{t-1},\boldsymbol{y}_{L_{ij},t} \quad\sim\quad t\Big(\ \boldsymbol{\mu}_{i,j|L_{ij};t}\ ,\ \boldsymbol{D}^{1/2}_{i,j|L_{ij};t}\,\boldsymbol{R}_{i,j|L_{ij};t}\,\boldsymbol{D}^{1/2}_{i,j|L_{ij};t}\ ,\ \nu_{i,j|L_{ij}}\Big), \tag{5}$$

where $\boldsymbol{R}_{i,j|L_{ij};t}$ is the conditional partial (bivariate) correlation matrix

$$\boldsymbol{R}_{i,j|L_{ij};t}=\begin{bmatrix}1 & \rho_{i,j|L_{ij};t}\\ \rho_{i,j|L_{ij};t} & 1\end{bmatrix},$$

$\nu_{i,j|L_{ij}}=\nu+j-i-1$ is the degrees of freedom parameter,

$$\boldsymbol{\mu}_{i,j|L_{ij};t}=\begin{pmatrix}\boldsymbol{R}_{i,L_{ij};t}\\ \boldsymbol{R}_{j,L_{ij};t}\end{pmatrix}\boldsymbol{R}^{-1}_{L_{ij},L_{ij};t}\ \boldsymbol{y}_{L_{ij},t}, \tag{6}$$

is the location parameter, and

$$\boldsymbol{D}_{i,j|L_{ij};t}=\frac{\nu-2+\boldsymbol{y}^\top_{L_{ij},t}\boldsymbol{R}^{-1}_{L_{ij},L_{ij};t}\boldsymbol{y}_{L_{ij},t}}{\nu_{i,j|L_{ij}}}\begin{pmatrix}\boldsymbol{V}_{i,i|L_{ij};t} & 0\\ 0 & \boldsymbol{V}_{j,j|L_{ij};t}\end{pmatrix} \tag{7}$$

a diagonal matrix holding the coordinate wise scale parameters; see Roth (2013) or Ding (2016). Note that for $j=i+1$, L_{ij} is the empty set as there are no intermediate coordinates between i and $i+1$, such that for $j=i+1$ the location $\boldsymbol{\mu}_{i,j|L_{ij};t}$ parameter collapses to zero, $\nu_{i,j|L_{ij}}=\nu$, and $\boldsymbol{D}^{1/2}_{i,j|L_{ij};t}\boldsymbol{R}_{i,j|L_{ij};t}\boldsymbol{D}^{1/2}_{i,j|L_{ij};t}$ collapses to $(1-2\nu^{-1})$ times the pairwise Pearson correlation matrix for $(\boldsymbol{y}_{i,t},\boldsymbol{y}_{j,t})^\top$; compare Eq. (1).

We can use (5) to *recursively* build the dynamic correlation matrix via univariate transition equations for the partial correlations $\rho_{i,j|L_{ij};t}$ using the bivariate data slice for pair (i,j). To see this, consider a trivariate example. In a first step, we use (5) to model $(\boldsymbol{y}_{1,t},\boldsymbol{y}_{2,t})^\top$. This gives a filter for the dynamics of $\rho_{1,2,t}$. We choose scaled Fisher transform of a correlation parameter $\rho_{1,2,t}=\epsilon\cdot\tanh(f_{1,2,t})$ for $0<\epsilon<1$ to ensure that the (partial) correlation always lies in the interval $(-1,1)$ for any $f_{1,2,t}\in\mathbb{R}$.[4] Next, we repeat this procedure for $(\boldsymbol{y}_{2,t},\boldsymbol{y}_{3,t})^\top$, obtaining a model for the dynamics of $\rho_{2,3,t}$. Finally, we consider (5) for $(\boldsymbol{y}_{1,t},\boldsymbol{y}_{3,t})^\top$ conditional on $\boldsymbol{y}_{2,t}$, obtaining the dynamics for (a possibly re-parameterized version of) $\rho_{1,3|2;t}$. To recover the Pearson correlation $\rho_{1,3,t}$ and thus the entire correlation matrix, we use $\rho_{1,3|2;t}$ and the correlations $\rho_{1,2,t}$ and $\rho_{2,3,t}$ obtained in the previous steps together with the inverse mapping from $\rho_{1,3|2;t}$ to $\rho_{1,3,t}$ in Eq. (4).

Because we only have to work with the bivariate conditional distributions in (5), all transition equations for the (partial) correlation parameters $\rho_{i,j|L_{ij};t}$ take a similar form. In particular, we have the following result.

Proposition 1 (*Score Recursions*). *Define* $\boldsymbol{y}_{i,j;t}=(\boldsymbol{y}_{i,t},\boldsymbol{y}_{j,t})^\top$ *for* $j>i$ *and moreover, let* $p(\boldsymbol{y}_{i,j;t}\mid\boldsymbol{y}_{L_{ij},t},\mathcal{F}_{t-1})$ *be the Student's* t *pdf corresponding to* (5). *Using the scaled Fisher transform* $\rho_{i,j|L_{ij};t}=\epsilon\cdot\tanh(f_{i,j|L_{ij};t})$ *for* $f_{i,j|L_{ij};t}\in\mathbb{R}$ *and* $0<\epsilon<1$, *the score expression for the score-driven dynamics of Creal et al. (2013) is given by*

$$s_{i,j|L_{ij};t}=\frac{\partial\log p(\boldsymbol{y}_{i,j;t}\mid\boldsymbol{y}_{L_{ij},t},\mathcal{F}_{t-1})}{\partial f_{i,j|L_{ij};t}}$$

[4] Alternative choices for the mapping from $f_{i,j|L_{ij};t}$ to $\rho_{i,j|L_{ij};t}$ result in a different model with possibly different dynamic properties, i.e., stationarity and ergodicity and invertibility properties, similar to a GARCH context where conditions for stationarity and invertibility are quite different between for instance GARCH, log-GARCH, and EGARCH models. The current Fisher transform is not only intuitive and well-known, but also allows us to formulate sharp and easy to verify contraction condition which ensure stationarity and ergodicity properties; see Section 3.1.

$$= \frac{1}{2} \boldsymbol{G}_{i,j|L_{ij};t}^{\top} \left(\boldsymbol{R}_{i,j|L_{ij};t}^{-1} \otimes \boldsymbol{R}_{i,j|L_{ij};t}^{-1} \right) \times \operatorname{vec}\left(w_{i,j|L_{ij};t} \cdot \boldsymbol{D}_{i,j|L_{ij};t}^{-1/2} \left(\boldsymbol{y}_{i,j;t} - \boldsymbol{\mu}_{i,j|L_{ij};t} \right) \left(\boldsymbol{y}_{i,j;t} - \boldsymbol{\mu}_{i,j|L_{ij};t} \right)^{\top} \boldsymbol{D}_{i,j|L_{ij};t}^{-1/2} - \boldsymbol{R}_{i,j|L_{ij};t} \right), \tag{8}$$

for $i = 1, \dots, N-1$, $j = i+1, \dots, N$, *and* $L_{ij} = \{i+1, \dots, j-1\}$, *with*

$$w_{i,j|L_{ij};t} = \frac{\nu_{i,j|L_{ij}} + 2}{\nu_{i,j|L_{ij}} + (\boldsymbol{y}_{i,j;t} - \boldsymbol{\mu}_{i,j|L_{ij};t})^{\top} \boldsymbol{D}_{i,j|L_{ij};t}^{-1/2} \boldsymbol{R}_{i,j|L_{ij};t}^{-1} \boldsymbol{D}_{i,j|L_{ij};t}^{-1/2} (\boldsymbol{y}_{i,j;t} - \boldsymbol{\mu}_{i,j|L_{ij};t})},$$

$$\boldsymbol{G}_{i,j|L_{ij};t} = \partial \operatorname{vec}(\boldsymbol{R}_{i,j|L_{ij};t}) / \partial f_{i,j|L_{ij};t} = \epsilon \cdot \left(1 - \tanh(f_{i,j|L_{ij};t})^2\right) \cdot \begin{pmatrix} 0 & 1 & 1 & 0 \end{pmatrix}^{\top}.$$

This leads to the score transition equation

$$f_{i,j|L_{ij};t+1} = \omega_{i,j|L_{ij}} + \beta_{i,j|L_{ij}} \, f_{i,j|L_{ij};t} + \alpha_{i,j|L_{ij}} \, s_{i,j|L_{ij};t}, \tag{9}$$

where we use unit score scaling in the sense of Creal et al. (2013).[5]

The score equation in (8) makes intuitive sense: it reacts to the outer product of the bivariate observations to drive the correlation parameter, which includes the cross-product of $\boldsymbol{y}_{i,t}$ and $\boldsymbol{y}_{j,t}$. Unlike the score equation of Creal et al. (2011) for the multivariate volatility model, however, the score in (8) reacts to the observations minus their *conditional* mean, and divided by their *conditional* standard deviations, where the conditioning is on all observations $\boldsymbol{y}_{k,t}$ for intermediate coordinates $i < k < j$, i.e., $k \in L_{ij}$. The partial correlation dynamics are thus driven by the squares and cross products of $\boldsymbol{y}_{i,t}$ and $\boldsymbol{y}_{j,t}$ *after correcting for* the values of the intermediate variables. This makes perfect intuitive sense for the dynamics of a *partial* rather than a standard correlation coefficient.

The result in Proposition 1 has a number of key differences with earlier score-driven dynamic correlation models. We mention five of them. First, unlike the matrix equations in for instance (Creal et al., 2011; Opschoor et al., 2018, 2021), and Hafner and Wang (2021), the recursions in (9) are all univariate for $i = 1, \dots, N-1$, $j = i+1, \dots, N$, and $L_{ij} = \{i+1, \dots, j-1\}$. Second, as a result of this, the parameters in (9) can be estimated recursively for a given value of ν, starting with the pairs $(i, i+1)$ for $i = 1, \dots, N-1$, followed by the pairs $(i, i+2)$ for $i = 1, \dots, N-2$, and so on, up to the last pair $(1, N)$. We discuss this approach further in Section 2.5. Third, the dynamic parameters $\boldsymbol{\mu}_{i,j|L_{ij};t}$, $\boldsymbol{D}_{i,j|L_{ij};t}$, and $\boldsymbol{R}_{i,j|L_{ij};t}$ for $i < j$ all depend on the data and on dynamic correlations $\rho_{k_1,k_2,;t}$ for $i \le k_1 < k_2 \le j$ and $(k_1, k_2) \neq (i, j)$ obtained in an earlier step of the vine cascade. The system of equations is therefore recursive rather than simultaneous, which has the potential to substantially simplify the estimation; see Section 2.5. Still, the dynamics for the full correlation matrix are multivariate rather than bivariate, as the bivariate filtering equations are locked together (non-linearly) via the expressions for $\boldsymbol{\mu}_{i,j|L_{ij};t}$, $\boldsymbol{D}_{i,j|L_{ij};t}$, and $\boldsymbol{R}_{i,j|L_{ij};t}$. Fourth, because of its bivariate nature, the current set-up of designing a dynamic correlation matrix is perfectly scalable to higher dimensions: there is no worry about the final correlation matrix $\boldsymbol{R}_t$ not being positive definite, as we have modeled the partial correlations rather than their Pearson correlation counterparts. Fifth, the approach based on partial correlations allows us to easily impose zero (or sign) restrictions on some of the partial correlations if this is desirable from a theoretical perspective. Imposing such restrictions on partial correlations when modeling a dynamic Pearson correlation matrix directly, by contrast, is much harder.

2.4. Maximum likelihood estimation

As our model is observation-driven, the likelihood is known in closed form as

$$\hat{L}_T(\boldsymbol{\theta}) = \sum_{t=1}^{T} \hat{\ell}_t(\boldsymbol{\theta}), \tag{10}$$

$$\hat{\ell}_t(\boldsymbol{\theta}) = \left\{ \log \Gamma\left(\frac{\nu+N}{2}\right) - \log \Gamma\left(\frac{\nu}{2}\right) - \frac{N}{2} \log((\nu-2)\pi) - \frac{1}{2} \log |\hat{\boldsymbol{R}}_t(\boldsymbol{\theta})| - \frac{\nu+N}{2} \log\left(1 + \frac{\boldsymbol{y}_t^{\top} \hat{\boldsymbol{R}}_t(\boldsymbol{\theta})^{-1} \boldsymbol{y}_t}{\nu-2}\right) \right\},$$

where $\boldsymbol{\theta}$ contains ν, $\omega_{i,j|L_{ij}}$, $\alpha_{i,j|L_{ij}}$, $\beta_{i,j|L_{ij}}$, for $i = 1, \dots, N-1$ and $j = i+1, \dots, N$, and $\{\hat{\boldsymbol{R}}_t(\boldsymbol{\theta})\}_{t=1}^{T}$ contains the filtered correlation matrices using the score driven recursions from Proposition 1, initialized at some $\hat{\boldsymbol{R}}_1$. In our empirical application, we set $\hat{\boldsymbol{R}}_1$ to the correlation matrix of the first 100 observations, but other choices are possible as well.

The likelihood in (10) can be optimized numerically using standard software to yield the maximum likelihood estimator (MLE)

$$\hat{\boldsymbol{\theta}}_T = \arg\max_{\boldsymbol{\theta} \in \boldsymbol{\Theta}} \hat{L}_T(\boldsymbol{\theta}). \tag{11}$$

Note that the optimization could make use of the sequential estimation approach of Section 2.5 to obtain good starting values of $\boldsymbol{\theta}$.

[5] We have also experimented with alternative forms of scaling, such as scaling the score by an additional factor $(1 - \rho_{i,j|L_{ij}}^2)$ to mitigate score step sizes near the edges of the domain. This results in somewhat smoother paths of the empirical correlations in Section 4 and modest changes in the stationarity and invertibility conditions.

2.5. Sequential estimation strategy

The full maximum likelihood approach of Section 2.4 does not exploit the recursive structure of the model set-up in any way. As the cross-sectional dimension N of the data increases, this standard approach may become more and more computer intensive. In this section we propose an alternative estimation strategy that uses the recursive set-up of the model and that can be solved by a sequence of small-scale estimation problems rather than one full multivariate estimation procedure. In essence, we follow an equation-by-equation ML (EbE-ML) approach. A similar procedure was also considered by Francq and Zakoïan (2016), where the authors also developed the corresponding asymptotic theory with an application to multivariate GARCH models such as the BEKK and DCC model. Also Darolles et al. (2018) use the EbE-ML for the Cholesky GARCH model. Within the context of our partial dynamic correlation model, this alternative strategy can be used on its own, or as a tool to obtain good starting values for the full ML approach.

The core element of the EbE-ML estimator (EbE-MLE) in the context of the dynamic partial correlation model is the bivariate conditional distribution in (5) for each data slice $(\boldsymbol{y}_{i,t}, \boldsymbol{y}_{j,t})^\top$ for $i < j$. To start off the EbE-ML estimator, we first consider all pairs (i, j) for $j = i + 1$ and $i = 1, \dots, N - 1$. For these pairs, we have $L_{ij} = \emptyset$, such that the conditional bivariate Student t density in (5) only conditions on $\mathcal{F}_{t-1}$, and not on $\boldsymbol{y}_{L_{ij}}$. For each pair $(i, i+1)$, we optimize the bivariate log-likelihood for that particular data slice with respect to the static parameters $\nu_{i,j|L_{ij}}$, $\omega_{i,j|L_{ij}}$, $\alpha_{i,j|L_{ij}}$, and $\beta_{i,j|L_{ij}}$, for $j = i + 1$. This is feasible, because for each of these data slices $(i, i + 1)$ the likelihood only depends on the single dynamic partial conditional correlation coefficient $\rho_{i,j|L_{ij};t} = g\left(f_{i,j|L_{ij};t}\right)$, and on none of the others. The optimizations can therefore even be parallelized for $i = 1, \dots, N - 1$ to speed up computations.

Given estimates of the parameters $\nu_{i,j|L_{ij}}$, $\omega_{i,j|L_{ij}}$, $\alpha_{i,j|L_{ij}}$, and $\beta_{i,j|L_{ij}}$ for $j = i + 1$, and the resulting filtered dynamic partial correlations $\rho_{i,j|L_{ij};t}$ for this first set of pairs $(i, i + 1)$, we can use Eqs. (5)–(7) to compute the conditional means for the next set of pairs (i, j) for $j = i + 2$, conditional on the intermediate data series $L_{ij} = \{i + 1\}$. This gives rise to a new set of $N - 2$ small-scale bivariate conditional log-likelihood optimizations with respect to the new parameters $\nu_{i,j|L_{ij}}$, $\omega_{i,j|L_{ij}}$, $\alpha_{i,j|L_{ij}}$, and $\beta_{i,j|L_{ij}}$ for $j = i + 2$ and $L_{ij} = \{i + 1\}$. This set of optimizations can again be parallelized. The procedure is repeated for all pairs $(i, j = i + k)$ for $k = 1, \dots, N - 1$ and $L_{ij} = \{i + 1, \dots, i + k - 1\}$. The algorithm is summarized in Fig. 1.

Finally, the different estimates $\hat{\nu}_{i,j|L_{ij}}$ of $\nu_{i,j|L_{ij}} = \nu + j - i - 1$ can be combined into a single estimate of ν by a simple moments estimator, such as for instance $\hat{\nu} = 2(N(N-1))^{-1} \sum_{i<j} \nu_{i,j|L_{ij}} + i + 1 - j$. As mentioned earlier, the estimates resulting from this sequential algorithm can be considered on their own, or used as starting values for the full ML approach from Section 2.4.

3. Asymptotic properties

In this section, we study the asymptotic properties of the model. We first study the stationarity properties of the model as a DGP in Section 3.1, followed by filter invertibility in Section 3.2. Finally, we study the consistency and asymptotic normality of the maximum likelihood estimator for the static parameters of the model in Section 3.3. The recursive structure of the partial correlation model will prove extremely useful here: the exponentially fast almost sure convergence of the filtered time-varying parameters $\rho_{i,j|L_{ij};t}$ for $j = i + k$ allows us to use them as a plug-in estimators in subsequent recursions for $j = i + k + 1$ without loosing filter invertibility. As a result, we can obtain consistency and asymptotic normality of the maximum likelihood estimator of the static parameters $\boldsymbol{\theta}$.

3.1. Stationarity and ergodicity of the model

To establish stationarity and ergodicity of $\boldsymbol{y}_t$, we first consider the model as a DGP. Using (1), we can rewrite (9) in the stochastic recurrence equation (SRE) representation defined by Bougerol (1993) and Straumann and Mikosch (2006). In this subsection and the next, we are somewhat more meticulous regarding notation. We write $\hat{\boldsymbol{R}}_t$ as the true sequence of bivariate correlation matrices, initialized at a fixed $\hat{\boldsymbol{R}}_1$. We write $\boldsymbol{R}_t$ without a hat to indicate its uninitialized stationary and ergodic limit sequence, if it exists. Similar notation is used for the partial correlations $\rho_{i,j|L_{ij};t}$ and their transformations $f_{i,j|L_{ij};t}$. Based on the SRE representation, we formulate conditions for the convergence of the random sequences $\{\hat{f}_{i,j|L_{ij};t}\}_{t\in\mathbb{N}}$ initialized at fixed values $\hat{f}_{i,j|L_{ij};1}$ to unique strictly stationary and ergodic sequences $\{f_{i,j|L_{ij};t}\}_{t\in\mathbb{Z}}$. We make the following three assumptions.

Assumption 1. The partial correlation coefficients are defined via the scaled Fisher transformation $\rho_{i,j|L_{ij};t} = \epsilon \cdot \tanh(f_{i,j|L_{ij};t})$ for $i = 1, \dots, N - 1$, and $j = i + 1, \dots, N$ and some constant $0 < \epsilon < 1$.

Assumption 2. The degrees of freedom parameter ν of the Student's t density satisfies $2 + \delta < \nu < \infty$ for some $\delta > 0$.

Assumption 3. For $i = 1, \dots, N - 1$ and $j = i + 1, \dots, N$, let

$$\mathbb{E}\left[\log \sup_{|\rho|<\epsilon} \left|\beta_{i,j|L_{ij}} + \alpha_{i,j|L_{ij}} \cdot \left(\tilde{h}(\rho) \cdot \left(e_1 - \rho\, e_2\right) - \tilde{h}(\rho) \cdot \epsilon \cdot \left(1 - \rho^2/\epsilon^2\right) \cdot e_2\right)\right|\right] < 0, \tag{13}$$

for $\tilde{h}(\rho) = \epsilon \cdot (1 - \rho^2/\epsilon^2)/(1 - \rho^2)$, and $\tilde{h}(\rho) = 2(\epsilon^2 - 1)\rho(1 - \rho^2/\epsilon^2)/(1 - \rho^2)^2$, where $e_1 = (\nu_{i,j|L_{ij}} + 2)\boldsymbol{\eta}_{1,t}\boldsymbol{\eta}_{2,t}/(\nu_{i,j|L_{ij}} + \boldsymbol{\eta}_t^\top \boldsymbol{\eta}_t)$ and $e_2 = \frac{1}{2}(\nu_{i,j|L_{ij}} + 2)\boldsymbol{\eta}_t^\top \boldsymbol{\eta}_t/(\nu_{i,j|L_{ij}} + \boldsymbol{\eta}_t^\top \boldsymbol{\eta}_t) - 1$ for a standard Student $t(\nu_{i,j|L_{ij}})$ distributed bivariate random variable $\boldsymbol{\eta}_t$.

Step 1: initialize $\hat{\boldsymbol{\theta}}^{(k=0)} = \emptyset$, for $k = 1, \dots, N-1$

Step 2: for each pair (i, j) with $j = i + k$ and $i = 1, \dots, N - k$

Step 3a: define the parameter vector $\boldsymbol{\theta}^{(i,j|L_{ij})} = (\omega_{i,j|L_{ij}}, \alpha_{i,j|L_{ij}}, \beta_{i,j|L_{ij}}, \nu_{i,j|L_{ij}})^\top$ for this pair, with $L_{ij} = \{i+1, \dots, i+k-1\}$ the set of coordinates between i and $i+k$ (which is empty for $k = 1$)

Step 3b: if $k = 1$, set $\hat{\boldsymbol{y}}^{\star}_{i,j|L_{ij};t} = \sqrt{\nu/(\nu-2)}\, \boldsymbol{y}_{i,j|L_{ij};t}$, otherwise set

$$\hat{\boldsymbol{y}}^{\star}_{i,j|L_{ij};t} = \boldsymbol{y}^{\star}_{i,j|L_{ij};t}(\hat{\boldsymbol{\theta}}^{(k-1)}) = \boldsymbol{D}_{i,j|L_{ij};t}(\hat{\boldsymbol{\theta}}^{(k-1)})^{-1/2} \left(\boldsymbol{y}_{i,j;t} - \boldsymbol{\mu}_{i,j|L_{ij};t}(\hat{\boldsymbol{\theta}}^{(k-1)}) \right),$$

for $\boldsymbol{\mu}_{i,j|L_{ij};t}(\boldsymbol{\theta}^{(k-1)})$ and $\boldsymbol{D}_{i,j|L_{ij};t}(\boldsymbol{\theta}^{(k-1)})$ as defined in equations (6) and (7)

Step 3c: using $\hat{\boldsymbol{y}}^{\star}_{i,j|L_{ij};t}$, maximize

$$\hat{L}_{i,j|L_{ij};T}(\boldsymbol{\theta}^{(i,j|L_{ij})}) = \sum_{t=1}^{T} \hat{\ell}_{i,j|L_{ij};t} \left(\boldsymbol{\theta}^{(i,j|L_{ij})} \right), \tag{12}$$

$$\hat{\ell}_{i,j|L_{ij};t}(\boldsymbol{\theta}^{(i,j|L_{ij})}) = -\log(2\pi) - \frac{1}{2}\log(1-\rho^2_{i,j|L_{ij};t}) - \frac{1}{2}\log|\hat{\boldsymbol{D}}_{i,j|L_{ij};t}(\hat{\boldsymbol{\theta}}^{(k-1)})|$$
$$- \frac{\nu_{i,j|L_{ij}}+2}{2}\log\left(1 + \frac{\hat{\boldsymbol{y}}^{\star\top}_{i,j|L_{ij};t} \begin{pmatrix} 1 & \rho_{i,j|L_{ij};t} \\ \rho_{i,j|L_{ij};t} & 1 \end{pmatrix}^{-1} \hat{\boldsymbol{y}}^{\star}_{i,j|L_{ij};t}}{\nu_{i,j|L_{ij}}}\right),$$

with respect to $\boldsymbol{\theta}^{(i,j|L_{ij})}$; for fixed k, this step can be parallelized over i

Step 3d: with $\hat{\boldsymbol{\theta}}^{(i,j|L_{ij})}$, compute the filtered parameters $\hat{\rho}_{i,j|L_{ij};t}$ and $\hat{\rho}_{i,j;t}$ and store them in $\hat{\boldsymbol{R}}_t$ for the computations in Step 3b

Step 3e: add all parameter estimates $\hat{\boldsymbol{\theta}}^{(i,j|L_{ij})}$ for $i = 1, \dots, N-k$ to $\hat{\boldsymbol{\theta}}^{(k-1)}$ to form the new $\hat{\boldsymbol{\theta}}^{(k)}$

Step 4: increase k, and return to Step 2

Step 5: compute the moments estimator $\hat{\nu} = 2(N(N-1))^{-1}\sum_{i<j}\nu_{i,j|L_{ij}} + i + 1 - j$

Fig. 1. Sequential estimation algorithm (Equation-by-Equation).

Assumption 1 is common in the literature on dynamic conditional correlation models. It can be found in for instance Harvey (2013), or Blasques et al. (2018b). In our case, it ensures that the partial correlations are never equal to ±1, such that the correlation matrix $\boldsymbol{R}_t$ implied by the partial correlations is always (strictly) positive definite. Assumption 2 is a standard moment condition that is needed for second moments (and thus the correlation matrix) to exist. If we choose to model a scaling matrix instead, this assumption can be further relaxed to the existence of an arbitrarily small moment. Assumption 3 formulates a sufficient condition for ensuring that the recursions for $\hat{f}_{i,j|L_{ij};t}$ are contracting on average. This in turn allows us to apply Theorem 3.1 of Bougerol (1993) and conclude stationarity and ergodicity properties of the model as a DGP. The restrictions on the parameter space imposed by Eq. (13) can easily be checked numerically for specific values of $\alpha_{i,j|L_{ij}}$, $\beta_{i,j|L_{ij}}$, and $\nu_{i,j|L_{ij}}$.

To visualize the restrictions imposed on the parameter space by the contraction condition in Eq. (13) of Assumption 3, Fig. 2 plots the boundary for the stationarity and ergodicity region for different values of $\nu_{i,j|L_{ij}}$. Combinations of $\alpha_{i,j|L_{ij}} > 0$ and $\beta_{i,j|L_{ij}} > 0$ below the boundary curve satisfy Assumption 3. The region has an irregular space due to the heavy non-linearity of the recursive filtering equations.[6]

Using the above assumptions, we can now prove the following proposition.

Proposition 2 (*Strict Stationarity and Ergodicity*). *Let Assumptions 1–3 hold true. Let $\hat{\boldsymbol{R}}_1$ denote a fixed initial correlation matrix with implied partial correlations $\hat{\rho}_{i,j|L_{ij};1}$ and their transforms $\hat{f}_{i,j|L_{ij};1}$. Then, the solutions $\hat{f}_{i,j|L_{ij};t}$ of model (5)–(9) for $t \in \mathbb{N}$, initialized at $\hat{f}_{i,j|L_{ij};1}$ for $i = 1, \dots, N-1$, $j = i+1, \dots, N$, converge e.a.s. to unique strictly stationary and ergodic solutions $\{f_{i,j|L_{ij};t}\}_{t\in\mathbb{Z}}$. In addition,*

[6] The region can even be further enlarged if we consider higher order iterates of the transition equation for $f_{i,j|L_{ij};t}$; see for instance Bougerol (1993) or D'Innocenzo et al. (2023). Computational details for Fig. 2 can be found at the end of the proof of Proposition 2.

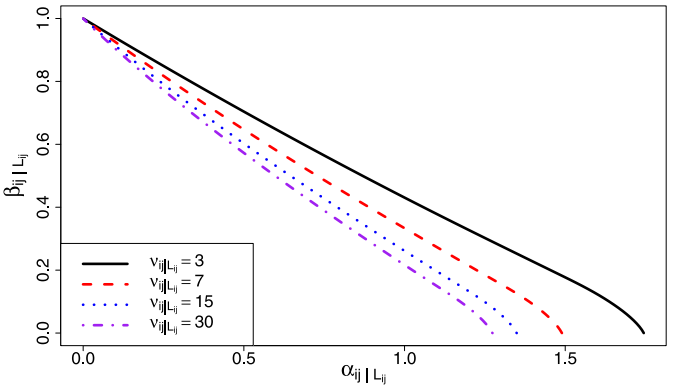

Fig. 2. Stationarity and ergodicity boundary.

the (initialized) partial correlations $\hat{\rho}_{i,j|L_{ij};t} = \epsilon \cdot \tanh(\hat{f}_{i,j|L_{ij};t})$ *and the Pearson correlations* $\hat{\rho}_{i,j;t}$ *converge e.a.s. to their unique stationary and ergodic limits* $\{\rho_{i,j|L_{ij};t}\}_{t\in\mathbb{Z}} = \{\epsilon \cdot \tanh(f_{i,j|L_{ij};t})\}_{t\in\mathbb{Z}}$ *and* $\{\rho_{i,j;t}\}_{t\in\mathbb{Z}}$.

Proposition 2 establishes that the model has a stationary solution. Irrespective of any fixed initialization $\hat{f}_{i,j|L_{ij};1}$, the model's solution converges to this same stationary limit. We can even allow the initial condition $\hat{f}_{i,j|L_{ij};1}$ to be random, as long as a logarithmic moment exists of $|\omega_{i,j|L_{ij}} + \beta \hat{f}_{i,j|L_{ij};1} + \alpha s_{i,j|L_{ij};1}|$. In the next section, we extend the result of Proposition 2 to the model used as a filter rather than a DGP, and formulate conditions under which the asymptotic solution of the filter does not depend on its initialization.

3.2. Filter invertibility

Naturally, the true time-varying partial correlation processes $\{\rho_{i,j|L_{ij};t}\}_{t\in\mathbb{Z}}$ are unobserved. However, due to the observation-driven nature of the model, we can easily replace every unobserved $\rho_{i,j|L_{ij};t}$ by its initialized (at $f_{i,j|L_{ij};1}$) filtered counterpart $\hat{\rho}_{i,j|L_{ij};t}(\boldsymbol{\theta}) = \epsilon \cdot \tanh(\hat{f}_{i,j|L_{ij};t}(\boldsymbol{\theta}))$ for $t = 1, \dots, T$, where we add the argument $\boldsymbol{\theta}$ to the notation to indicate that the filter is evaluated at an arbitrary $\boldsymbol{\theta} \in \boldsymbol{\Theta}$. These filtered partial correlations map into the filtered Pearson correlation matrices via Eq. (4).

To study the asymptotic properties of the MLE $\hat{\boldsymbol{\theta}}_T$, we first need to study the stochastic limit properties of the filtered processes $\{\hat{f}_{i,j|L_{ij};t}(\boldsymbol{\theta})\}_{t=1}^{T}$, since the likelihood function depends on both the data and these filtered processes. The appropriate convergence result for the filter is known in the literature as filter invertibility; see Straumann and Mikosch (2006), Wintenberger (2013), and Blasques et al. (2018a). A complication in our setting is that *all* partial correlations are needed to construct the full correlation matrix. This is important, as unlike Wintenberger (2013) or Blasques et al. (2018a, 2022) which also deal with non-linear filtering methods for time-varying parameters, we cannot rely on standard contraction theorems such as the Bougerol's Theorem 3.1. The novelty in the result below lies in the fact that we show that the multivariate convergence follows easily from the individual univariate convergence results for the pairwise partial correlation filters based on bivariate data slices. This provides a substantial simplification of the proof. To accomplish this, we lean on the theory for perturbed stochastic recurrence equations (SREs) of Straumann and Mikosch (2006, Theorem 2.10) using a sequence of cascading SREs.

To formulate the result, we recall that the demeaned and standardized bivariate observation vectors $\boldsymbol{y}^{\star}_{i,j|L_{ij};t}(\boldsymbol{\theta})$ for a general $\boldsymbol{\theta} \in \boldsymbol{\Theta}$ are defined as

$$\boldsymbol{y}^{\star}_{i,j|L_{ij};t}(\boldsymbol{\theta}) = \boldsymbol{D}_{i,j|L_{ij};t}(\boldsymbol{\theta})^{-1/2} \left(\boldsymbol{y}_{i,j;t} - \boldsymbol{\mu}_{i,j|L_{ij};t}(\boldsymbol{\theta}) \right), \tag{14}$$

with $\boldsymbol{\mu}_{i,j|L_{ij};t}(\boldsymbol{\theta})$ and $\boldsymbol{D}_{i,j|L_{ij};t}(\boldsymbol{\theta})$ as defined in Eqs. (6) and (7), respectively. These standardized observations make up the main input of the bivariate conditional Student's t distributions in (5). Note that $\boldsymbol{y}^{\star}_{i,j|L_{ij};t}(\boldsymbol{\theta})$ not only depends on $\boldsymbol{y}_{i,t}$ and $\boldsymbol{y}_{j,t}$, but also on the pairwise correlations as gathered in $\boldsymbol{R}_{i,L_{ij};t}(\boldsymbol{\theta})$ and $\boldsymbol{R}_{L_{ij},L_{ij};t}(\boldsymbol{\theta})$, i.e., the pairwise correlations for all coordinates strictly between i and j. These correlations have been estimated in a previous step of the cascade; see also the algorithm in Section 2.5. We therefore also introduce the perturbed counterpart $\hat{\boldsymbol{y}}^{\star}_{i,j|L_{ij};t}(\boldsymbol{\theta})$ of $\boldsymbol{y}^{\star}_{i,j|L_{ij};t}(\boldsymbol{\theta})$, where we replace the elements of the uninitialized stationary and ergodic $\boldsymbol{R}_{i,L_{ij};t}(\boldsymbol{\theta})$ and $\boldsymbol{R}_{L_{ij},L_{ij};t}(\boldsymbol{\theta})$ in (14) by those of the initialized (non-stationary) $\hat{\boldsymbol{R}}_{i,L_{ij};t}(\boldsymbol{\theta})$ and $\hat{\boldsymbol{R}}_{L_{ij},L_{ij};t}(\boldsymbol{\theta})$, respectively. We also distinguish three different filtered sequences: (i) the filter sequence $\hat{f}_{i,j|L_{ij};t}(\boldsymbol{\theta})$, initialized at $\hat{f}_{i,j|L_{ij};1}$ an taking $\hat{\boldsymbol{y}}^{\star}_{i,j|L_{ij};t}(\boldsymbol{\theta})$ as inputs; (ii) the filter sequence $\hat{\tilde{f}}_{i,j|L_{ij};t}(\boldsymbol{\theta})$, initialized at the same $\hat{f}_{i,j|L_{ij};1}$ but taking the stationary and ergodic $\boldsymbol{y}^{\star}_{i,j|L_{ij};t}(\boldsymbol{\theta})$ as inputs; and (iii) the sequence $f_{i,j|L_{ij};t}(\boldsymbol{\theta})$, denoting the uninitialized stationary and ergodic limiting filter that takes $\boldsymbol{y}^{\star}_{i,j|L_{ij};t}(\boldsymbol{\theta})$ as inputs. The first of these three is the one that is actually computed in empirical applications via the MLE procedure and is available to the user.

To formulate our proposition, we make the following assumption.[7]

Assumption 4. The set $\boldsymbol{\Theta} \subset \mathbb{R}^d$ is a compact parameter space satisfying $\nu \geq 2+\delta$ for some $\delta > 0$ and $\alpha_{i,j|L_{ij}} \neq 0$ for $i = 1, \dots, N-1$ and $j = i+1, \dots, N$, with

$$\mathbb{E}\left[\log \sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} \sup_{f} \left| \beta_{i,j|L_{ij}} + \alpha_{i,j|L_{ij}} \cdot \frac{\partial s_{i,j|L_{ij};t}\left(f, \boldsymbol{y}^{\star}_{i,j|L_{ij};t}(\boldsymbol{\theta}); \boldsymbol{\theta}\right)}{\partial f} \right| \right] < 0. \tag{15}$$

As in GARCH models, the requirement that $\alpha_{i,j|L_{ij}} \neq 0$ in Assumption 4 is necessary in order to avoid the potential lack of identification. If $\alpha_{i,j|L_{ij}} = 0$ and $0 \leq \beta_{i,j|L_{ij}} < 1$, then $f_{i,j|L_{ij};t} \to \omega_{i,j|L_{ij}}/(1-\beta_{i,j|L_{ij}})$, such that these parameters cannot be identified separately; see also Darolles et al. (2018) for a similar discussion. We also point out that, thought the condition in (15) looks bivariate in nature, the full multivariate structure of the model is retained as $\boldsymbol{y}^{\star}_{i,j|L_{ij};t}(\boldsymbol{\theta})$ depends on all data and parameters for the intermediate coordinates between $i < j$.

Assumption 4 ensures that the initialized filter is contracting on average when taking the unperturbed $\boldsymbol{y}^{\star}_{i,j|L_{ij};t}(\boldsymbol{\theta})$ as inputs, i.e., $\hat{\hat{f}}_{i,j|L_{ij};t}(\boldsymbol{\theta}) \xrightarrow{e.a.s.} f_{i,j|L_{ij};t}(\boldsymbol{\theta})$. An approach based on $\hat{\hat{f}}_{i,j|L_{ij};t}(\boldsymbol{\theta})$ is, however, infeasible: the MLE procedure can only use the perturbed $\hat{\boldsymbol{y}}^{\star}(\boldsymbol{\theta})$ based on all previously filtered pairs of (initialized) correlation estimates. Therefore, the empirical procedure produces $\hat{f}_{i,j|L_{ij};t}(\boldsymbol{\theta})$ rather than $\hat{\hat{f}}_{i,j|L_{ij};t}(\boldsymbol{\theta})$. Only for $j-i=1$ we observe $\boldsymbol{y}^{\star}(\boldsymbol{\theta})$ directly because $\boldsymbol{\mu}_{i,i+1|L_{ij};t}(\boldsymbol{\theta}) = 0$ and $\boldsymbol{D}_{i,i+1|L_{ij};t}(\boldsymbol{\theta}) = (1-2\nu^{-1})\mathbf{I}_2$. For $j-i = k > 1$, however, the score recursions for the filter also use the initialized sequence $\hat{f}_{i,j|L_{ij};t}(\boldsymbol{\theta})$ for $j-i = 1, \dots, k-1$. The latter are not stationary and ergodic, which prevents us from applying (Bougerol, 1993) as this theorem requires stationary and ergodic inputs.

The way out of this challenge is as follows. If the filters $\hat{f}_{i,j|L_{ij};t}(\boldsymbol{\theta})$ for $j-i<k$ converge exponentially fast and almost surely to their stationary and ergodic limits $f_{i,j|L_{ij};t}(\boldsymbol{\theta})$, then we can use the results on perturbed SREs from Straumann and Mikosch (2006). In particular, under condition (15) the desired filter invertibility for $\hat{f}_{i,j|L_{ij};t}(\boldsymbol{\theta})$ can then still be established for $j-i=1,\dots,N-1$. The composite procedure boils down to the following. Starting from $j-i=1$, we recursively obtain invertibility for $\hat{f}_{i,j|L_{ij};t}(\boldsymbol{\theta})$ for all $j-i=k=1,\dots,N-1$. Finally, by standard continuity arguments, we conclude that filter invertibility holds for the pairwise conditional correlation coefficients $\rho_{i,j|L_{ij};t}(\boldsymbol{\theta})$ and the Pearson correlation matrices $\hat{\boldsymbol{R}}_t(\boldsymbol{\theta})$. We summarize this in the following proposition.

Proposition 3 (*Filter Invertibility*). *Let Assumptions 1–4 hold true. Then, the filter processes* $\{\hat{f}_{i,j|L_{ij};t}(\boldsymbol{\theta})\}_{t\in\mathbb{N}}$ *initialized at fixed values* $\hat{f}_{i,j|L_{ij};1}$ *converge exponentially fast almost surely to the unique stationary and ergodic sequences* $\{f_{i,j|L_{ij};t}(\boldsymbol{\theta})\}_{t\in\mathbb{Z}}$ *uniformly over the parameter space* $\boldsymbol{\Theta}$*, that is*

$$\sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} \left|\hat{f}_{i,j|L_{ij};t}(\boldsymbol{\theta}) - f_{i,j|L_{ij};t}(\boldsymbol{\theta})\right| \xrightarrow{e.a.s.} 0,$$
$$\sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} \left|\hat{\rho}_{i,j|L_{ij};t}(\boldsymbol{\theta}) - \rho_{i,j|L_{ij};t}(\boldsymbol{\theta})\right| \xrightarrow{e.a.s.} 0,$$
$$\sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} \left|\hat{\rho}_{i,j;t}(\boldsymbol{\theta}) - \rho_{i,j;t}(\boldsymbol{\theta})\right| \xrightarrow{e.a.s.} 0,$$

as $t \to \infty$.

As a result of Proposition 3, the impact of starting values for the filters becomes negligible asymptotically. In Appendix B we show that this result extends to the derivative processes of $\hat{f}_{i,j|L_{ij};t}(\boldsymbol{\theta})$ with respect to $\boldsymbol{\theta}$. These derivative processes play a crucial role for proving the asymptotic normality of the MLE. Filter invertibility in the multivariate model thus simplifies for the setting at hand to a sequence of univariate invertibility conditions, which are much easier to handle.

3.3. Consistency and asymptotic normality of the MLE

Our approach to establish strong consistency and asymptotic normality of the MLE for our dynamic partial correlation model relies on similar arguments as discussed in Straumann and Mikosch (2006) and Blasques et al. (2022). The idea consists in first showing that the nonstationary average log-likelihood function $T^{-1}\hat{L}_T(\boldsymbol{\theta})$ in (10) converges to its stationary counterpart $T^{-1}L_T(\boldsymbol{\theta})$, which uses $f_{i,j|L_{ij};t}(\boldsymbol{\theta})$ rather than $\hat{f}_{i,j|L_{ij};t}(\boldsymbol{\theta})$. We can then apply the uniform strong law of large numbers for stationary and ergodic processes of Rao (1962) to show that $T^{-1}L_T(\boldsymbol{\theta}) \to \mathbb{E}[\ell_t(\boldsymbol{\theta})]$ almost surely and uniformly over $\boldsymbol{\theta}\in\boldsymbol{\Theta}$. The strong consistency of $\hat{\boldsymbol{\theta}}_T$ then follows by checking standard identifiability arguments. The result is stated in the following theorem.

Theorem 1. *Under Assumptions 1–4,* $\hat{\boldsymbol{\theta}}_T \xrightarrow{a.s.} \boldsymbol{\theta}_0$ *for every fixed set of starting values* $\hat{f}_{i,j|L_{ij};1} \in \mathbb{R}$ *for the filter for* $i = 1, \dots, N-1$ *and* $j = i+1 \dots, N$.

[7] In fact, Assumption 4 may be further relaxed by replacing the supremum over $\boldsymbol{\theta}$ in (15) by a supremum over $(\omega_{i,j|L_{ij}}, \alpha_{i,j|L_{ij}}, \beta_{i,j|L_{ij}}, \nu)$ only. In order not to overburden the (already heavy) notation further, we opt for the current simpler, though somewhat more restrictive formulation.

To establish the asymptotic normality of the MLE, the following two additional assumptions are needed, which are rather standard in the literature.

Assumption 5. $\theta_0 \in \text{interior}(\boldsymbol{\Theta})$, i.e., the true parameter vector θ_0 lies in the interior of the (compact) parameter space $\boldsymbol{\Theta}$.

Assumption 6. For some $\delta > 0$, it holds that

$$\mathbb{E}\left[\sup_{\theta\in\boldsymbol{\Theta}}\sup_{f}\left|\beta_{i,j|L_{ij}}+\alpha_{i,j|L_{ij}}\cdot\frac{\partial s_{i,j|L_{ij};t}\left(f,\boldsymbol{y}^{\star}_{i,j|L_{ij};t}(\theta);\theta\right)}{\partial f}\right|^{2+\delta}\right]<1. \tag{16}$$

Assumption 5 excludes situations where the true parameter lies on the boundary of the parameter space. Assumption 6 in addition requires that the score-driven filters and their derivative processes have second moments. This allows us to appeal to an appropriate central limiting result. Combining these assumptions, we obtain the following theorem, which is proved in Appendix A.

Theorem 2. *Under Assumptions 1–6, and for every fixed set of starting values for the filter, $\hat{f}_{i,j|L_{ij};1} \in \mathbb{R}$ for $i = 1,\dots,N-1$ and $j = i+1\dots,N$, we have $\sqrt{T}(\hat{\theta}_T - \theta_0) \Rightarrow \mathbb{N}(\mathbf{0}, \boldsymbol{\mathcal{I}}^{-1}(\theta_0))$, where $\boldsymbol{\mathcal{I}}(\theta_0)$ is the Fisher information matrix evaluated at the true parameter vector θ_0.*

The results obtained in Theorems 1 and 2 are for the full ML estimator. In Fig. 1 in Section 2.5 we also introduced a sequential estimation approach. The approach used for proving Theorems 1 and 2 carries over directly to this sequential estimation strategy, resulting in a consistency and asymptotic normality result for the sequential estimator. For the first step in the vine structure, i.e., the pairs $(i, i+1)$, this is self-evident as no auxiliary (partial) correlations have to be estimated that relate to other coordinate pairs. For the remaining vine layers, consistency and asymptotic normality follows entirely along the same lines as in the proofs of the theorems above, building on the exponentially fast almost sure convergence of the (conditionally) demeaned and standardized data $\hat{\boldsymbol{y}}^{\star}_{i,j|L_{ij};t}$. Efficiency of the sequential estimator, however, cannot be guaranteed and is likely to be lost. In the empirical application in Section 4 we investigate the performance of both the full ML and the sequential estimator and show that the full ML estimator results in the best performance, with the sequential estimator being a good runner-up.

4. Empirical application

4.1. Data and benchmark models

In this section we provide two examples. In Section 4.3 we consider 4-dimensional applications to US industry returns with three Fama–French risk factors. In Section 4.4, we consider a 23-dimensional application of 20 individual blue-chip stocks with three Fama–French risk factors. We take daily return data from January 3, 1980 to December 31, 2021. The data are obtained from Ken French's website and CRSP.[8] We assess model adequacy in economic terms from an asset pricing perspective as in Boudt et al. (2017). The three risk factors are the excess market factor Mkt-RF, the size factor SMB (Small Minus Big), and the value factor HML (High Minus Low). Plots and descriptive statistics of some of the series can be found in Appendix C. The series display the familiar stylized facts of high kurtosis and clear volatility clustering. The fat-tailedness motivates the use of the Student's t distribution for the analysis.

We label the new dynamic partial correlation model as *PCorr* in the subsequent discussion. The model's performance is benchmarked against three proven models from the literature: (i) the multivariate GAS(1, 1) model of Creal et al. (2011) (labeled t-GAS), where the correlation matrix is modeled using the hypersphere parameterization, (ii) the cDCC model of Engle (2002) and Aielli (2013) endowed with a Student's t distribution and labeled t-cDCC, and (iii) the SCC model of Palandri (2009), which also uses a recursive set-up of the model. For the t-cDCC model, we use the standard targeting approach to estimate the (matrix-valued) intercept parameter of the correlation transition equation. For the partial correlation model and the matrix t-GAS model such a targeting approach is not available, and we estimate the intercept terms as part of the static parameter vector using standard numerical optimization.

We also note that the new PCorr model has pair-specific parameters $\alpha_{i,j|L_{ij}}$ and $\beta_{i,j|L_{ij}}$, unlike the standard versions of the t-cDCC and matrix t-GAS. The latter typically only use a common scalar α and β. To put the different models on a more equal footing, we introduce the same number of pair-specific $\alpha_{i,j}$ and $\beta_{i,j}$ into the t-GAS model. This can be done without further complications due to the hypersphere parameterization in the t-GAS. For the t-cDCC, we impose a BEKK type specification with diagonal A and B matrices holding N parameters $\alpha_{i,i}$ and $\beta_{i,i}$, respectively. This ensures positive definiteness of the correlation matrix at all times.[9]

To fully concentrate on the differences in modeling correlations, we first de-volatilize all return series in the same way for every correlation model using the score-driven volatility models of Creal et al. (2011) and Creal et al. (2013) based on the Student's t distribution, also known as the Beta-t-GARCH(1,1) model of Harvey (2013). The de-volatilized series are then used as inputs for the correlation-based models. As all correlation models now work with the same input series, any differences can no longer be attributed to differences in univariate volatility filters. Of course, alternative ways to filter the volatility can also be implemented, including the use of realized volatilities or other volatility models than the univariate t-GAS.

[8] http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

[9] In particular, we use $Q^{\star}_t = \Omega + B^{1/2}Q^{\star}_{t-1}B^{1/2} + A^{1/2}\boldsymbol{y}_{t-1}\boldsymbol{y}^{\top}_{t-1}A^{1/2}$, where A and B are diagonal matrices with parameters $\alpha_{i,i}$ and $\beta_{i,i}$, respectively.

Table 1
MSE and *MAE* simulation results.

	PCorr	t-GAS	t-DCC	SCC	PCorr	t-GAS	t-DCC	SCC
	Gaussian				Student t_7			
	Single rolling window empirical correlation path							
$\overline{MSE}$	**0.0174**	0.0248	0.0270	0.0318	**0.0192**	0.0247	0.0349	0.0292
$\overline{MAE}$	**0.1106**	0.1140	0.1223	0.1338	**0.1106**	0.1123	0.1370	0.1246
	Block-bootstrapped rolling window empirical correlation paths							
$\overline{MSE}$	**0.0427**	0.0448	0.0463	0.0480	**0.0420**	0.0459	0.0463	0.0472
$\overline{MAE}$	**0.1700**	0.1769	0.1754	0.1772	**0.1680**	0.1734	0.1756	0.1792

Note: the labels PCorr, t-GAS and t-DCC indicate the new score-driven partial correlation model discussed in Section 2, the Student's t GAS model of Creal et al. (2011) with hypersphere parameterization, and the t-cDCC model of Engle (2002) with a multivariate Student's t log-likelihood, respectively. Results are based on 300 Monte Carlo (top panel) or bootstrap (bottom panel) experiments with sample size $T = 1000$ and $N = 4$. The true correlation path used in the data generating process for the top panel is given by the 100-day rolling window estimates of empirical correlation matrices of the series (HML, SMB, Mkt-RF, BusEq). For the bottom panel, the 100-day rolling window estimate is block-bootstrapped in every replication (block length 50) to get a new correlation path in every simulation. Results are averaged over time, replications, and across pairs (i, j), $1 \leq i < j \leq 4$. Further results for the individual pairs are found in Tables C.2 and C.3 in Appendix C and confirm the results in the table.

4.2. Simulation results

Before turning to the empirical applications, we first investigate the performance of the new model in a controlled simulation setting. To generate time series with empirically relevant correlation dynamics, we use a 100-day rolling window to estimate time-varying empirical correlation matrices for the four series (HML, SMB, Mkt-RF, BusEq), where BusEq (business equipment) holds the returns on one of the sector portfolios from Section 4.3. These rolling window estimates produce paths for $4(4-1)/2 = 6$ different pairwise correlations. We fix these paths and then generate 300 realizations of the returns $\boldsymbol{y}_t$ for $T = 1000$, using either a Gaussian or a Student's t ($\nu = 7$) distribution, where the latter is close to the empirical estimate. For each of the simulated return series, we estimate the new time-varying partial correlation model as well as the benchmark models.

To compare the different models, we consider the mean squared error (MSE) and the mean absolute error (MAE), where

$$MSE_{i,j} = \frac{1}{T}\sum_{t=1}^{T}\left(\hat{\rho}_{i,j;t} - \rho_{i,j;t}\right)^2, \qquad MAE_{i,j} = \frac{1}{T}\sum_{t=1}^{T}\left|\hat{\rho}_{i,j;t} - \rho_{i,j;t}\right|,$$

$\overline{MSE}$ and $\overline{MAE}$ are the averages of $MSE_{i,j}$ and $MAE_{i,j}$ across all pairs $i < j$, respectively, and $\hat{\rho}_{i,j;t}$ denotes the filtered path of the conditional (Pearson) correlation coefficient. As all measures consider the performance of the correlation models in terms of pairwise Pearson correlations rather than partial correlations, the benchmark models (which all operate on the Pearson correlations directly) are put at an advantage compared to our new model (which operates on the partial correlations).

The results in Table 1 present a clear outcome: in both the Gaussian and Student's t case and both performance measures, the new partial correlation model outperforms the three benchmarks. The extended results per pair (i, j) in Table C.2 in Appendix C show that the conclusion is uniform across all pairwise correlation paths. For the empirically more relevant Student's t case, the average MAE and MSE are in the same ballpark for the different models. This seems realistic: we expect all three models to do reasonably well for typical stock return series. This is confirmed by the filtered correlation paths in Fig. 3. The black pattern gives the true path of the (six) correlations used in the simulations. Most of the time, the different correlation models follow each other quite closely. However, there are also marked differences such as for $\rho_{2,3,t}$. The differences in filtered correlations between models also shows up in terms of the improvement in MAE in Table 1. The MAE improves by around 2/11/20 percent for the new partial correlation model vis-à-vis the t-GAS/SCC/t-cDCC model. For the MSE, differences are even larger with 22/35/55 percent improvements compared to these three models, respectively. Differences are much smaller if instead we compare the filtered correlations between models with different distributions, such as the Gaussian versus Student's t distribution for the PCorr model; see Figure C.2 in Appendix C.

The results do not depend on the particular correlation path. In the bottom panel of Table 1 we provide averages of MSE and MAE for a second simulation experiment. Here we take the correlation pattern from the first experiment, but block-bootstrap it to obtain new, empirically relevant correlation patterns. For each bootstrapped path, we generate return data, estimate the 8 different models, and compute the MAE and MSE. A summary of the results across different correlation paths is given in the bottom panel in Table 1. It confirms our earlier findings: across all performance measures, the PCorr model outperforms the other models. Table C.3 and Figure C.4 in Appendix C for each individual pair (i, j) again show that the conclusion is robust for all individual elements of the correlation matrix. The partial correlation model thus appears to do a better job in cases where it produces different correlations than its competitors. We continue to investigate this for the empirical data in the remainder of this section.

4.3. Application to US industry returns

4.3.1. In-sample analysis

In our first application, we consider a four-dimensional system of US stock returns based on the three Fama–French risk factor portfolios and one industry portfolio. We first compare the different models in-sample over the period 1980–2009. Next, we consider an out-of-sample analysis over the period 2010–2021.

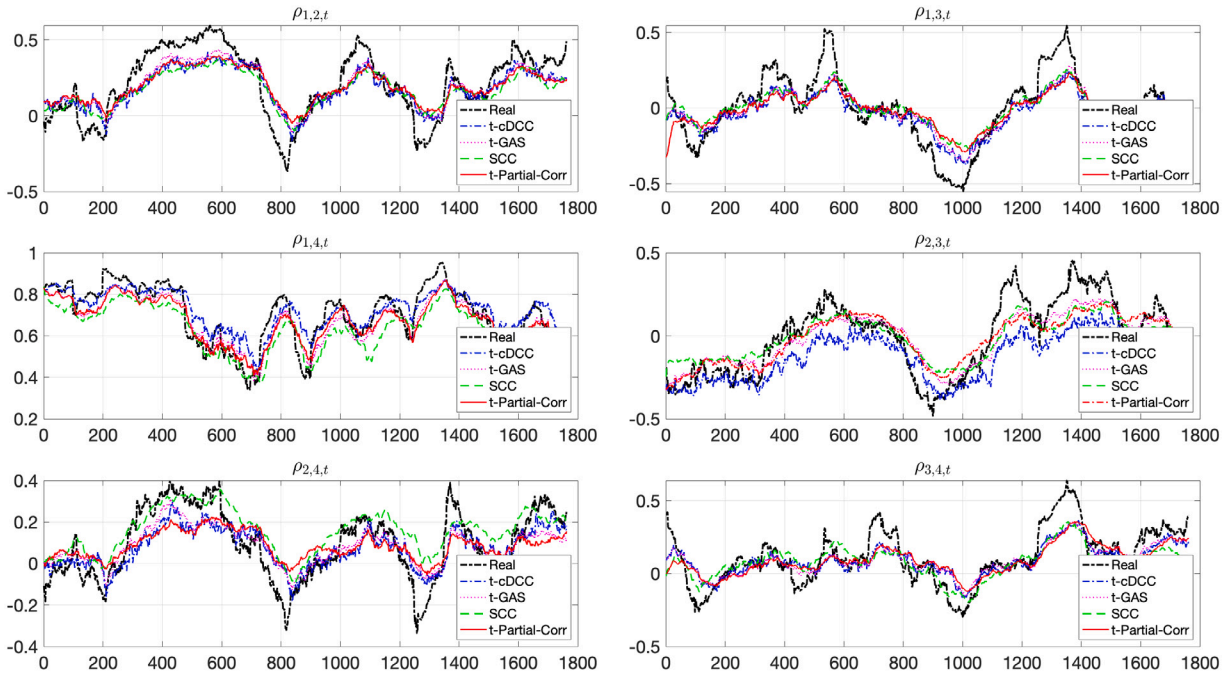

Fig. 3. Comparison of the mean of the Monte Carlo simulation of the filtered conditional correlation coefficients with Student's t DGP with $\nu = 7$.

Table 2 holds the differences in log-likelihood values between the different models for each of the industries. In all cases we take the PCorr model as the benchmark, such that positive values in the log-likelihood column signal that the new model outperforms the benchmark. The results clearly show that the PCorr model always outperforms the t-cDCC for each of the 12 industries. In comparison with the t-GAS, the new model also fits better in 8 out of the 12 industries, performs less well in only 2 cases, and at par in 2 others. Most gains are in the range of 20–50 likelihood points for PCorr versus t-GAS. Improvements are even higher at 60–110 likelihood points when comparing the PCorr versus the t-cDCC model. Note that these results hold despite the fact that we made the t-GAS and t-cDCC models more flexible by endowing them with different $\alpha_{i,j}$ and $\beta_{i,j}$ parameters for different return pairs as compared to their standard scalar form from the literature.

Besides evaluating the models in terms of likelihood fit, we also compare the models in terms of their asset pricing implications as in Hansen et al. (2014), Boudt et al. (2017), and Darolles et al. (2018).[10] For this, we consider the tracking errors

$$e_t = r_{i,t} - \gamma_{Mkt,t}(r_t^{Mkt} - r_t^F) - \gamma_{SMB,t}SMB_t - \gamma_{HML,t}HML_t, \tag{17}$$

where $r_{i,t}$ denotes the industry portfolio return for $i = 1, \ldots, 12$, and

$$\begin{pmatrix} \gamma_{Mkt,t} \\ \gamma_{SMB,t} \\ \gamma_{HML,t} \end{pmatrix} = \begin{pmatrix} \rho_{Mkt,Mkt,t} & \rho_{Mkt,SMB,t} & \rho_{Mkt,HML,t} \\ \rho_{SMB,Mkt,t} & \rho_{SMB,SMB,t} & \rho_{SMB,HML,t} \\ \rho_{HML,Mkt,t} & \rho_{HML,SMB,t} & \rho_{HML,HML,t} \end{pmatrix}^{-1} \begin{pmatrix} \rho_{Mkt,i,t} \\ \rho_{SMB,i,t} \\ \rho_{HML,i,t} \end{pmatrix}. \tag{18}$$

As all models considered are observation-driven, the $\rho_{i,j;t}$ and thus also the $\gamma_{i,j;t}$ are known at time $t-1$. Remember that all returns enter the tracking error equation after being de-volatilized in order to fully concentrate on the differences due to correlation modeling.

Table 2 displays the tracking-error-based MSE and MAE comparisons. Significance is assessed using the Diebold–Mariano t-test. Negative values in the MSE or MAE column indicate that the PCorr model outperforms the benchmark. The results confirm the earlier log-likelihood analysis. Also in terms of its asset pricing implications, the PCorr model outperforms the other models. Interestingly, for a few industries the statistical log-likelihood criterion and the economic performance criteria based on tracking error MSE and MAE point to a different model ranking. Based on the economic criteria, the overall conclusion seems even more robust: the PCorr model outperforms the benchmarks in-sample. In terms of MAE, the outperformance is unanimous and strongly statistically significant across all industries. For MSE, the results are similar, with a few exceptions.

[10] As remarked by one of the referees, the comparisons could also be carried out by comparing the filtered correlations with realized correlations based on intraday data. Though potentially interesting, we have not pursued this in this paper and leave this for future research. Instead, we concentrate on the comparisons in terms of model fit (likelihood) and in terms of economic implications (asset pricing errors).

Table 2
In sample performance of the four correlation models.

	PCorr versus t-GAS			PCorr versus t-cDCC			PCorr versus SCC		MCS			
	log-Lik	DM_{MSE}	DM_{MAE}	log-Lik	DM_{MSE}	DM_{MAE}	DM_{MSE}	DM_{MAE}	PCorr	t-GAS	t-cDCC	SCC
NoDur	21.8*	−6.88***	−7.69***	91.7***	−6.05***	−7.78***	−5.60***	−6.62***	✓			
Durl	31.9**	−9.21***	−10.20***	84.6***	−7.85*	−8.71***	−24.23***	−31.11***	✓			
Manuf	21.9*	−9.70***	−11.44***	64.2***	−7.85***	−10.26***	−5.20***	−8.32***	✓			
Enrgy	−1.0*	−4.47***	−4.84***	88.4***	−4.45***	−4.08***	−7.33***	−8.02***	✓			
Chems	40.6*	−6.68*	−7.01***	94.1***	−6.15***	−7.15***	−6.40***	−7.79***	✓			
BusEq	−46.7***	−3.62***	−5.84***	64.1***	−2.47***	−5.91***	−7.68***	−7.79***	✓	✓	✓	
Telcm	32.5***	−7.28***	−7.94***	110.9***	−7.81***	−7.10***	7.31***	−12.32***	✓			
Utils	0.1	−1.41	−2.87***	87.0***	−1.08	−2.92***	−3.94***	−8.39***	✓	✓	✓	
Shops	28.8**	−7.21***	−9.00***	83.6***	−7.03***	−8.37***	−4.07***	−5.21***	✓			
Health	23.7*	−5.90***	−6.23***	91.7***	−5.31***	−6.68***	−5.77***	−8.43***	✓			
Money	−30.4*	−5.66***	−8.51***	76.5***	−5.73***	−9.99***	−10.41***	−15.14***	✓			
Other	50.3***	−8.61***	−10.01***	89.6***	−7.36***	−8.02***	−3.54***	−4.69***	✓			

Note: the PCorr, t-GAS, and t-cDCC models are estimated over the sample 03 January 1980 to 31 December 2009. The log-Lik indicates the differences in log-likelihood value at the optimum. Diebold–Mariano t statistics are reported based on the MSE and MAE criterion, related to the differences in mean squared and mean absolute pricing errors, $e_t = r_{i,t} - \hat{\gamma}_{Mkt,t}(r_t^{Mkt} - r_t^F) - \hat{\gamma}_{SMB,t}SMB_t - \hat{\gamma}_{HML,t}HML_t$, where all return series are volatility filtered, and $\hat{\gamma}_{Mkt/SMB/HML,t}$ is obtained as in (18). The symbols ∗, ∗∗, and ∗∗∗ denote statistical significance at the 10, 5, and 1 percent level, respectively. The MCS columns indicate whether the model is selected for the 95% model confidence set based on MSE. For the MAE criterion, the GAS and DCC model are only selected for Utils. The PCorr, t-GAS, t-cDCC, and SCC have 19, 19, 19, and 40 parameters, respectively, and take on average 218, 250, 100, and 46 s to estimate for the current four-dimensional setting.

To account for the cross-sectional correlation of the different tests and the possibly inflated type I error due to multiple pairwise tests, we also compute the model confidence set (MCS) of Hansen et al. (2011) based on the MSE criterion. We see that the PCorr model is always in the 95% MCS. The t-GAS and t-cDCC models, by contrast, only enter the MCS for two industries, and even then at lower p-values. The SCC model never enters the MCS for any of the industries. For the MCS based on the MAE (not shown) the pattern is similar: whereas the PCorr model is in the model confidence set for each of the 12 industries, the t-GAS and t-cDCC model enter only for one industry (Utils), and the SCC does not enter at all.

As we have 4 models with at least 19 parameters each, estimated across 12 industries, we have estimated almost 700 parameters in total and their standard errors. In Figure C.3, we visually present part of these estimates. Three main findings emerge. First, the estimates of ν are relatively stable between 6 and 7.5 across industries and models, indicating a realistic, moderate degree of fat-tailedness. Second, the (partial) correlations in all models have a high degree of persistence: all $\beta_{i,j|L_{ij}}^{PCorr}$, $\beta_{i,i}^{DCC}$, and $\beta_{i,j}^{GAS}$ parameters are close to one across all industries and models. Third, the adjustment speeds ($\alpha_{i,j}$) for the t-GAS and t-cDCC appear much more homogeneous than those of the PCorr model. In particular the adjustment speed of the partial correlation of SMB with the market return (MKT) is much higher in the PCorr model. The partial correlation between the industry returns and SMB given the market return ($IND, SMB \mid MKT$), as well as that between the industry return and HML given SMB and the market return ($IND, HML \mid MKT, SMB$) are both substantially lower. Such heterogeneity can easily be allowed for in the PCorr model. This is more complicated in t-GAS model, which scrambles this linkage between $\alpha_{i,j}$ and $\rho_{i,j}$ via the hypersphere re-parameterization. As a result, the heterogeneity in $\alpha_{i,j}$ is more blurred for the t-GAS model.

We now proceed to investigate whether the in-sample outperformance of the benchmark models by the PCorr model in an asset pricing context also persists in an out-of-sample forecasting setting.

4.3.2. Out-of-sample analysis

In our out-of-sample analysis, we fully focus on the tracking errors defined in (17), similar to Hansen et al. (2014), Boudt et al. (2017), and Darolles et al. (2018). We perform a recursive out-of-sample analysis. First, we estimate all models on the in-sample period 1980–2009. We then fix the static parameter estimates and run the filter up to 2010 to obtain the one year out-of-sample model-implied correlation matrices $\boldsymbol{R}_t$ as well as the implied coefficients $\gamma_{Mkt,t}$, $\gamma_{SMB,t}$, and $\gamma_{HML,t}$ from Eqs. (17)–(18). These result in predicted returns (conditional on the risk factors) and the corresponding tracking errors. After obtaining the tracking errors for 2010, we then add 2010 to the sample, and re-estimate the model over 1980–2010 to obtain tracking errors for 2011. We repeat this process up till the last year in the sample, giving us 2978 tracking errors.

Table 3 presents the results. The first three columns provide the pairwise out-of-sample predictive log likelihood comparisons using the Diebold–Mariano test. It is clear that the PCorr model also significantly outperforms the benchmarks out-of-sample across all industries. This is confirmed by the model confidence sets (MCS) based on the predictive log-likelihood (PLL) as well as on the tracking error MSE criterion. The PCorr model is always in the MCS, whereas other models enter only sporadically. Also the Mincer–Zarnowitz tests confirm this result. For the PCorr model, the test only rejects in three industries, whereas the t-GAS rejects in five, followed by the t-cDCC and SCC models, which reject in 10 of the 12 industries.

Summarizing, also in the out-of-sample analysis the results clearly point towards the PCorr model. We attribute this to the flexibility of the PCorr model to adapt itself to each (partial) correlation separately, with a robust propagation system due to the use of the Student's t distribution and the score-driven dynamics. From both our in-sample and out-of-sample analysis, it appears that both properties are useful for typical empirical data. The t-GAS shares the robust score-driven propagation of the PCorr model, but lacks the direct link to each (partial) correlation due to the use of the complex hypersphere transformation, which may explain

Table 3
Out-of-sample results.

	PCorr vs.	PCorr vs.	PCorr vs.	MCS-PLL				MCS-MSE				Mincer–Zarnowitz			
	t-GAS	*t*-cDCC	SCC	PCorr	*t*-GAS	*t*-cDCC	SCC	PCorr	*t*-GAS	*t*-cDCC	SCC	PCorr	*t*-GAS	*t*-cDCC	SCC
NoDur	49.98***	24.52***	27.72***	✓				✓						✗✗✗	✗✗
Durbl	37.98***	26.11***	24.54***	✓				✓					✗✗		✗✗✗
Manuf	25.29***	28.69***	25.38***	✓				✓						✗✗✗	
Enrgy	0.59	30.52***	25.49***	✓	✓			✓				✗✗		✗✗✗	✗✗✗
Chems	32.29***	28.45***	23.85***	✓				✓						✗✗✗	✗✗✗
BusEq	50.93***	32.90***	28.38***	✓				✓				✗✗✗	✗✗✗	✗✗✗	✗✗✗
Telcm	30.85***	26.75***	28.86***	✓				✓						✗✗✗	✗✗✗
Utils	51.03***	25.81***	28.47***	✓				✓		✓					✗✗
Shops	37.75***	25.47***	25.83***	✓				✓						✗✗✗	✗
Health	43.75***	25.44***	29.14***	✓				✓					✗	✗✗✗	✗
Money	38.58***	32.74***	26.92***	✓				✓	✓			✗	✗✗	✗✗✗	✗✗✗
Other	41.00***	30.79***	23.17***	✓				✓					✗	✗✗✗	

This table contains Diebold–Mariano t statistics are reported based on the predictive log-likelihood criterion, related to the differences in predicted log-likelihood. The symbols $*$, $**$, and $***$ denote statistical significance at the 10, 5, and 1 percent level, respectively. The MCS columns indicate whether the model is selected for the 95% model confidence of Hansen et al. (2011) based on predictive log-likelihood (PLL) or the tracking error MSE. Results are similar for the 99% MCS. The final columns give the results for the Mincer–Zarnowitz regressions $r_{i,t} = a_0 + a_1 \hat{r}_{i,t} + u_{i,t}$, where $\hat{r}_{i,t} = r_{i,t} - \hat{e}_{i,t}$ is the (recursive) return forecast using one of the models using Eqs. (17)–(18), and $\hat{e}_{i,t}$ evaluated at the least-squares estimates. The symbols ✗, ✗✗, and ✗✗✗ indicate rejection of $H_0 : a_0 = 0, a_1 = 1$, at the 10%, 5%, and 1% significance level, respectively, using the suitable heteroskedasticity and autocorrelation consistent (HAC) estimator for the covariance matrix of the regression parameters as suggested by for instance White (1980) and MacKinnon and White (1985).

Table 4
In-sample performance in the 23-dimensional application.

	PCorr-Full		PCorr-EbE		*t*-cDCC		SCC	
	log-Lik	AIC	Δlog-Lik	ΔAIC	Δlog-Lik	ΔAIC	Δlog-Lik	ΔAIC
ALL	**−130 750.99**	**263 021.99**	−200.45	400.91	−358.70	708.30	−363.39	1740.13

In-sample log-likelihood and AIC for the PCorr model as values and increments for the dynamic partial correlations model (PCorr) estimated by full maximum likelihood (PCorr-Full) or by the sequential equation-by-equation approach of Section 2.5 (PCorr-EbE), the t-cDCC model of Creal et al. (2011), or the SCC model of Palandri (2009). Values of log-likelihood and AIC are presented for the benchmark (PCorr-Full), and changes compared to the benchmark for the other methods.

why it performs less well. The t-cDCC, on the other hand, retains the direct link to the pairwise correlations, but lacks the robust propagation mechanism. Finally, the SCC model also uses a recursive set-up like the PCorr model does, but does so in a very different way and also lacks the robust propagation mechanism of the PCorr model.

Fig. 4 shows the different $\gamma_{j,t}$ parameters for $j = MKT, SMB, HML$. The figure shows that though the secular movements of the models align, there can also be episodes where the time-varying parameters differ substantially between models. The γs in this industry for the PCorr model also appear somewhat smoother than for the benchmark models. As argued in Francq and Zakoian (2019), this may increase the applicability of the model if one believes the γs to be more stable over time and not change erratically from one day to another.

4.4. Application to 20 US stocks and 3 risk factors

In our second empirical application, we consider a larger number of assets. We use 20 individual US stocks form the Dow Jones index for which we have complete time series available over the period 1992–2022. The stocks are given in Table 5. Next to the 20 stocks, we also again consider the three Fama–French risk factors, making up a 23-dimensional setting in total. We compare the PCorr, the t-DCC and the SCC model, but leave out the t-GAS model in the current comparison. The t-GAS uses a matrix to scale the score, which has $0.5 \cdot 23(23-1)$ rows and columns, so over 64k elements at each time point. Handling the t-GAS in this setting therefore becomes computationally cumbersome, particularly when searching over the current high-dimensional parameter space.

To lever on the recursive structure of the PCorr model, we estimate its parameters in this second application using both the sequential estimation approach (PCorr-EbE) explained in Section 2.5 and a full ML optimization (PCorr-Full) starting from the sequential estimates. The sequential estimation gives a similar computation time as for the SCC model; see the note to Table 5. Computing the full MLE, obviously, is computationally much more cumbersome as there are $\frac{1}{2} \cdot 23 \cdot 22 = 253$ dynamic correlations, each with their own $\omega_{i,j|L_{ij}}$, $\alpha_{i,j|L_{ij}}$, and $\beta_{i,j|L_{ij}}$ parameters. Together with the degrees of freedom parameter, this gives 760 parameters to be estimated via a non-linear numerical optimization. Still, the estimator converges when started from the sequential estimates.

We perform an in-sample analysis over the entire data range 1992–2022, as well as an out-of-sample analysis. The initial sample for the latter consists of the years 1992–2009. Next, we recursively use the model to forecast one-period-ahead for a one year period, after which we update the sample by that year and re-estimate all models. The in-sample results can found in Table 4 and in Table C.5 in the online appendix. The results for the out-of-sample analysis are presented in Table 5.

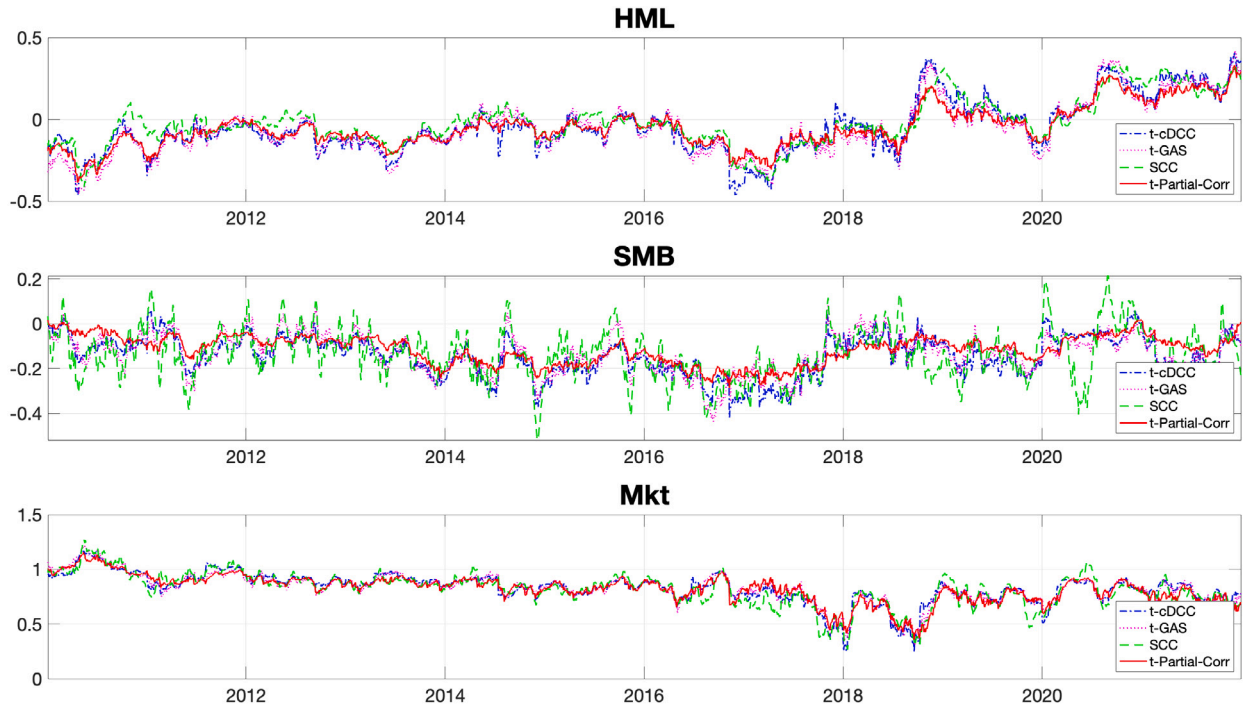

Fig. 4. Recursive one-step-ahead forecasts of the conditional risk loadings $\gamma_{Mkt,t}$, $\gamma_{SMB,t}$, $\gamma_{HML,t}$ from (17)–(18), for the *NoDur* industry (model re-estimated annually).

In-sample, Table 4 shows that the PCorr model estimated by full-maximum likelihood has the best log-likelihood value. The changes with respect to this benchmark are all negative for the log-likelihood, and positive for the AIC. As shown in Table C.5, these results are statistically significant using both a Diebold–Mariano test or a Model Confidence Set.

In Table 5 we present the out-of-sample results for the full 23-dimensional system. The top part of the table shows that when comparing the predictive log-likelihood values, the PCorr model again outperforms its competitors also in this higher dimensional out-of-sample setting. Both the Diebold–Mariano tests and the MCS point in the same direction: the PCorr estimated by full maximum likelihood performs best, followed by the PCorr estimated by the sequential estimation strategy. Both are in the model confidence set. Also when looking at the asset pricing implications of the models for the individual assets, we see that the PCorr model estimated by full maximum likelihood mostly outperforms the other models and is most often in the MCS. For 19 out of the 20 stocks, the PCorr estimated using the full ML estimator enters the MCS, accompanied by the PCorr estimated with the sequential approach for 6 stocks. Only for 1 out of the 20 stocks, the PCorr with the full ML is not in the MCS, whereas the PCorr with the sequential estimation approach is together with the other two models. In the end, there are only 2 stocks for which the t-cDCC and the SCC are in the MCS.

The results for the in-sample analysis in Table C.5 in the appendix are even more supportive of the PCorr model: the model is always in the in-sample MCS for the full ML estimator (20 out of 20 stocks), and for 15 out of 20 stocks also the PCorr with the sequential estimation approach is in the MCS. Only for 4 stocks, the SCC is also in the MCS, whereas the t-cDCC only enters the MCS once. We conclude that the current higher dimensional application confirms our earlier empirical results: also in these higher dimensional settings, and using the sequential estimation strategy from Section 2.5, the PCorr model provides good correlation dynamics in a flexible and robust way. The model can therefore be a good benchmark in other empirical analyses.

5. Conclusions

In this paper we introduced a recursive model for correlation matrix dynamics based on partial correlations and score-driven dynamics. The model's structure provided flexibility and interpretability, without loosing computational tractability. We provided two estimation strategies for the model's static parameters: a full multivariate maximum likelihood optimization, or a cascade of small-scale maximum likelihood estimations based on bivariate data slices.

The recursive structure of our model ensured stationarity and ergodicity as well as filter invertibility for any fixed dimension. The required conditions remained of similar complexity as in the univariate time-varying correlation setting and are based on bivariate data slices only. The proof relied on the theory of perturbed stochastic recurrence equations and could be applied to a cascade

Table 5
Out-of-sample performance in 23-dimensional application.

	PCorr-Full vs. PCorr-EbE		PCorr-Full vs. *t*-cDCC		PCorr-Full vs. SCC		MCS			
							PCorr-Full	PCorr-EbE	*t*-cDCC	SCC
	DM_{PLL}		DM_{PLL}		DM_{PLL}					
ALL	107.71***		197.42***		149.97***		✓	✓		
	DM_{MSE}	DM_{MAE}	DM_{MSE}	DM_{MAE}	DM_{MSE}	DM_{MAE}				
APPL	1.58	0.45	−3.55***	−6.75***	−7.75***	−7.73***	✓	✓		
AXP	−1.55	−2.99***	−4.67***	−8.31***	−5.17***	−7.15***	✓			
BA	−2.68***	−2.72***	−10.07***	−11.89***	−4.80***	−6.83***	✓			
CAT	−2.26***	−0.75	−6.08***	−9.35***	−6.14***	−8.05***	✓			
CSCO	−0.36	−1.41	−5.30***	−8.38***	−7.59***	−10.76***	✓	✓		
DOW	4.00***	3.79***	1.95**	0.21	1.10	−0.00		✓	✓	✓
HD	−2.57***	−2.52***	−5.45***	−8.67***	−5.30***	−8.11***	✓			
IBM	0.12	−0.72	−8.10***	−9.92***	−7.83***	−10.37***	✓	✓		
INTC	0.74	0.17	−7.21***	−11.63***	−6.70***	−8.71***	✓	✓		
JNJ	−2.49***	−2.28***	−7.44***	−10.51***	−7.09***	−8.25***	✓			
JPM	−8.81***	−8.74***	−5.62***	−10.34***	−4.63***	−6.26***	✓			
KO	−6.83***	−7.56***	−3.86***	−8.56***	−5.90***	−9.10***	✓			
MCD	−2.14***	−3.50***	−4.56***	−7.78***	−3.25***	−5.31***	✓			
MMM	−4.32***	−5.50***	−5.33***	−8.17***	−5.28***	−7.47***	✓			
MRK	−2.43***	−2.43***	−5.36***	−7.91***	−7.28***	−8.35***	✓			
PFE	−4.10***	−4.55***	−4.45***	−8.79***	−3.62***	−7.30***	✓			
PG	−4.14***	−3.37***	−4.72***	−6.89***	−5.66***	−6.68***	✓			
UTX	−0.65	−2.09***	−6.33***	−10.15***	−6.15***	−8.40***	✓			
V	−6.80***	−7.96***	−2.09***	−2.47***	−5.27***	−6.39***	✓			
WMT	4.59***	3.55***	1.55	−1.21	0.20	−2.33***	✓	✓	✓	✓

In the top-part of the table, we report the Diebold–Mariano statistics and MCS based on the predictive log-likelihood for the full 23-variate system. In the bottom part of the table, we report the asset pricing implications for each of the individual stocks based on the tracking errors. Here, the Diebold–Mariano t statistics are reported based on the tracking error MSE and MAE from Eqs. (17)–(18). The MCS columns indicate whether the model is selected for the 95% model confidence set based on MSE. The PCorr, t-cDCC, and SCC model contain 760/47/1265 parameters, respectively, and take 2116/1019/2309 s to estimate. The initial in-sample estimation period is from 1992–2009. The models are used for a one-year period to generate one-step-ahead forecasts. Then the sample is updated and all models are re-estimated. The entire out-of-sample period is from 2010–2022.

of bivariate (conditional) models. Using the stationarity and invertibility properties of the model and its filter, we were also able to prove consistency and asymptotic normality of the maximum likelihood estimator. Both in simulations and in in-sample and out-of-sample applications to US stock returns, the new model outperformed benchmarks such as the Student's t based cDCC and multivariate volatility GAS models, as well as the SCC model.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.jeconom.2024.105747.

References

Aielli, G.P., 2013. Dynamic conditional correlation: On properties and estimation. J. Bus. Econom. Statist. 31 (3), 282–299.

Anderson, T.W., 1958. An Introduction to Multivariate Statistical Analysis. Wiley New York.

Archakov, I., Hansen, P.R., 2021. A new parametrization of correlation matrices. Econometrica 89 (4), 1699–1715.

Barthel, N., Czado, C., Okhrin, Y., 2020. A partial correlation vine based approach for modeling and forecasting multivariate volatility time-series. Comput. Statist. Data Anal. 142, 106810.

Bauwens, L., Hafner, C., Laurent, S., 2012. Handbook of Volatility Models and their Applications. In: Wiley Handbooks in Financial Engineering and Econometrics, John Wiley & Sons..

Blasques, F., Gorgi, P., Koopman, S.J., Wintenberger, O., 2018a. Feasible invertibility conditions and maximum likelihood estimation for observation-driven models. Electron. J. Stat. 12, 1019–1052.

Blasques, F., Lucas, A., Silde, E., 2018b. A stochastic recurrence equations approach for score driven correlation models. Econometric Rev. 37 (2), 166–181.

Blasques, F., van Brummelen, J., Koopman, S.J., Lucas, A., 2022. Maximum likelihood estimation for score-driven models. J. Econometrics 227 (2), 325–346.

Boudt, K., Laurent, S.b., Lunde, A., Quaedvlieg, R., Sauri, O., 2017. Positive semidefinite integrated covariance estimation, factorizations and asynchronicity. J. Econometrics 196 (2), 347–367.

Bougerol, P., 1993. Kalman filtering with random coefficients and contractions. SIAM J. Control Optim. 31 (4), 942–959.

Buccheri, G., Bormetti, G., Corsi, F., Lillo, F., 2021. A score-driven conditional correlation model for noisy and asynchronous data: An application to high-frequency covariance dynamics. J. Bus. Econom. Statist. 39 (4), 920–936.

Comte, F., Lieberman, O., 2003. Asymptotic theory for multivariate GARCH processes. J. Multivariate Anal. 84 (1), 61–84.

Creal, D., Koopman, S.J., Lucas, A., 2011. A dynamic multivariate heavy-tailed model for time-varying volatilities and correlations. J. Bus. Econom. Statist. 29 (4), 552–563.

Creal, D., Koopman, S.J., Lucas, A., 2013. Generalized autoregressive score models with applications. J. Appl. Econometrics 28 (5), 777–795.

Darolles, S., Francq, C., Laurent, S.b., 2018. Asymptotics of Cholesky GARCH models and time-varying conditional betas. J. Econometrics 204 (2), 223–247.

Ding, P., 2016. On the conditional distribution of the multivariate t distribution. Amer. Statist. 70 (3), 293–295.

D'Innocenzo, E., Lucas, A., Schwaab, B., Zhang, X., 2023. Modeling Extreme Events: Time-Varying Extreme Tail Shape. Working Paper, Vrije Universiteit.

Engle, R., 2002. Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. J. Bus. Econom. Statist. 20 (3), 339–350.
Engle, R.F., 2016. Dynamic conditional beta. J. Financ. Econom. 14 (4), 643–667.
Francq, C., Zakoïan, J.-M., 2016. Estimating multivariate volatility models equation by equation. J. R. Stat. Soc. Ser. B Stat. Methodol. 78 (3), 613–635.
Francq, C., Zakoian, 2019. GARCH Models, second ed. John Wiley & Sons..
Hafner, C.M., Preminger, A., 2009. Asymptotic theory for a factor GARCH model. Econometric Theory 25 (2), 336–363.
Hafner, C.M., Wang, L., 2021. A dynamic conditional score model for the log correlation matrix. J. Econometrics.
Hansen, P.R., Lunde, A., Nason, J.M., 2011. The model confidence set. Econometrica 79 (2), 453–497.
Hansen, P.R., Lunde, A., Voev, V., 2014. Realized beta garch: a multivariate garch model with realized measures of volatility. J. Appl. Econometrics 29 (5), 774–799.
Harvey, A.C., 2013. Dynamic Models for Volatility and Heavy Tails. Cambridge University Press.
Joe, H., 2006. Generating random correlation matrices based on partial correlations. J. Multivariate Anal. 97 (10), 2177–2189.
MacKinnon, J.G., White, H., 1985. Some heteroskedasticity-consistent covariance matrix estimators with improved finite sample properties. J. Econometrics 29 (3), 305–325.
Opschoor, A., Janus, P., Lucas, A., Van Dijk, D., 2018. New HEAVY models for fat-tailed realized covariances and returns. J. Bus. Econom. Statist. 36 (4), 643–657.
Opschoor, A., Lucas, A., Barra, I., Van Dijk, D., 2021. Closed-form multi-factor copula models with observation-driven dynamic factor loadings. J. Bus. Econom. Statist. 39 (4), 1066–1079.
Palandri, A., 2009. Sequential conditional correlations: Inference and evaluation. J. Econometrics 153 (2), 122–132.
Pedersen, R.S., Rahbek, A., 2014. Multivariate variance targeting in the BEKK-GARCH model. Econom. J. 17 (1), 24–55.
Rao, R.R., 1962. Relations between weak and uniform convergence of measures with applications. Ann. Math. Stat. 33 (2), 659–680.
Rapisarda, F., Brigo, D., Mercurio, F., 2007. Parameterizing correlations: a geometric interpretation. IMA J. Manag. Math. 18 (1), 55–73.
Roth, M., 2013. On the Multivariate t Distribution. Technical Report from Automatic Control at Linköpings Universitet LiTH-ISY-R-3059, Department of Electrical Engineering, Linköpings Universitet.
Straumann, D., Mikosch, T., 2006. Quasi-maximum-likelihood estimation in conditionally heteroscedastic time series: A stochastic recurrence equations approach. Ann. Statist. 34 (5), 2449–2495.
Tse, Y.K., Tsui, A.K.C., 2002. A multivariate generalized autoregressive conditional heteroscedasticity model with time-varying correlations. J. Bus. Econom. Statist. 20 (3), 351–362.
White, H., 1980. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. Econometrica 48 (4), 817–838.
Wintenberger, O., 2013. Continuous invertibility and stable QML estimation of the EGARCH(1,1) model. Scand. J. Stat. 40 (4), 846–867.