The Canadian Journal of Statistics | La revue canadienne de statistique
Vol. 000, No. 00, 2025, e70030

Hidden Markov graphical models with state-dependent generalized hyperbolic distributions

Beatrice FORONI @
MEMOTEF Department, Sapienza University of Rome, Rome, Italy

Luca MERLO
Department of Human Sciences, Link Campus University, Rome, Italy

Lea PETRELLA
MEMOTEF Department, Sapienza University of Rome, Rome, Italy

Abstract In this article, we develop a novel hidden Markov graphical model to investigate time-varying interconnectedness between different financial markets. To identify conditional correlation structures under varying market conditions and accommodate shape features embedded in financial time series, we rely upon the generalized hyperbolic family of distributions with time-dependent parameters evolving according to a latent Markov chain. We exploit its location-scale mixture representation to build a penalized EM algorithm for estimating the state-specific sparse precision matrices by means of an L_1 penalty. The proposed approach leads to regime-specific conditional correlation graphs that allow us to identify different degrees of network connectivity of returns over time. The methodology's effectiveness is validated through simulation exercises under different scenarios. In the empirical analysis, we apply our model to daily returns of a large set of market indices, cryptocurrencies and commodity futures over the period 2017–2023.

Résumé Dans cet article, nous développons un nouveau modèle graphique de Markov caché pour étudier l'interconnexion variable dans le temps entre différents marchés financiers. Afin d'identifier les structures de corrélation conditionnelle sous des conditions de marché variables et d'intégrer les caractéristiques de forme présents dans les séries temporelles financières, nous nous appuyons sur la famille des distributions hyperboliques généralisées avec des paramètres dépendants du temps évoluant selon une chaîne de Markov latente. Nous exploitons sa représentation en mélange de localisation-échelle pour construire un algorithme EM pénalisé permettant d'estimer les matrices de précision clairsemées spécifiques à chaque état à l'aide d'une pénalité L_1. L'approche proposée conduit à des graphes de corrélation conditionnelle spécifiques aux régimes, nous permet-tant d'identifier différents degrés de connectivité du réseau des rendements au fil du temps. L'efficacité à travers des exercices de simulation sous différents scénarios. Dans l'analyse empirique, nous appliquons notre modèle aux rendements quotidiens d'un large ensemble d'indices de marché, de cryptomonnaies et de contrats à terme sur matières premières sur la période 2017–2023.

Keywords Cryptocurrencies; EM algorithm; financial networks; generalized hyperbolic family; hidden Markov models.

MSC2020 Primary: 62M05; 62A09; Secondary: 62P05.

@ Corresponding author beatrice.foroni@uniroma1.it

DOI: 10.1002/cjs.70030

1 Introduction

The financial system is a complex, dynamic, and interconnected world. Observing the extreme financial integration in the recent global crisis, researchers soon noted the crucial importance of identifying how the impact of stress events can spread across the whole global financial system. In the last few years, the development of statistical techniques to accurately quantify and investigate the interrelations among financial institutions has been at the centre of attention not only of investors and fund managers, but also of regulators for early identification of systemic risk and proactively engaging preventive measures to control financial stability (Silva et al., 2017, 2018; Giudici and Parisi, 2018; Brunetti et al., 2019). For this reason, network science has emerged as a useful tool for describing the propagation of systemic risk, where the interconnectedness between markets is represented by a graph whose nodes stand for companies, commodities, and institutions, and edges represent their interactions. In this context, Gaussian Graphical Models (GGMs) have received considerable attention because they provide a simple method to model the pairwise conditional dependencies of a collection of stochastic variables. As is well known, for normally distributed data, the underlying conditional dependence structure is completely characterized by the inverse of the covariance matrix, also known as the precision or concentration matrix, of the corresponding GGM (see Lauritzen (1996) for a general background). Thus, parameter estimation in GGMs is a fundamental issue for the identification of the zero entries in the concentration matrix. When dealing with large dimensional problems, we are interested in identifying only the variables that exhibit the most relevant and strongest connections. Among the several graphical methods proposed in the literature (Dempster, 1972; Drton and Perlman, 2004; Banerjee et al., 2008; Drton and Perlman, 2008; Cai et al., 2011; Liu et al., 2012; Liu and Wang, 2017), there exists the popular and computationally efficient *Graphical Lasso* (*glasso*) algorithm of Friedman et al. (2008), which maximizes the likelihood of the model penalized by the L_1-norm of the elements of the precision matrix.

Moreover, financial time series are characterized by well-known shape features like fat tails, leptokurtosis, and deviations from normality, that the analysis needs to properly take into account. Especially in the highly turbulent cryptocurrency market, risk managers and regulators are increasingly interested in determining whether, and how, the temporal evolution and volatility clustering of returns can be influenced by hidden variables, such as the state of the market, during tranquil and crisis periods. In this context, hidden Markov models (HMMs, see MacDonald and Zucchini, 1997; Zucchini et al., 2016) have been successfully employed in the analysis of financial time series data, with applications to asset allocation and stock returns as discussed in Mergner and Bulla (2008), De Angelis and Paas (2013), Nystrup et al. (2017), Maruotti et al. (2019, 2021), Pennoni et al. (2022), and Foroni et al. (2024a, 2024b). In the context of undirected graphs, estimation of graphical models in HMMs has been addressed by Städler and Mukherjee (2013) using multivariate Gaussian emission distributions with sparse precision matrices which can be interpreted as state-specific conditional independence graphs. More recently, Bianchi et al. (2019) introduced a Markov switching graphical seemingly unrelated regression model to investigate time-varying systemic risk based on a range of multifactor asset pricing models. Nevertheless, "nonstandard" features of returns cannot be accommodated by standard models such as those based on normality assumptions. Unfortunately, the literature regarding non-GGMs is fairly limited. Potential modelling strategies could rely on semiparametric Gaussian copula models (Liu et al., 2012; Xue and Zou, 2012), such as the nonparanormal model of Liu et al. (2009), or power transformations of the data. Alternatively, Finegold and Drton (2011) have introduced a robust graphical model based on the multivariate t distribution called *tlasso*.

In this article, we contribute to the existing literature by introducing a sparse hidden Markov graphical model to investigate time-varying conditional correlation structures in multivariate time series data, without assuming normally distributed returns. To build our network model, we consider multivariate symmetric generalized hyperbolic (GH, McNeil et al., 2015) distributions with time-dependent parameters evolving according to a discrete, homogeneous latent Markov chain. Within the financial literature, this family of densities has garnered significant attention for describing pertinent features of the distribution of returns (Chen et al., 2008; Necula, 2009; Ignatieva and Landsman, 2015) but also for its considerable flexibility in modelling financial data (Konlack Socgnia and Wilcox, 2014; Zhang et al., 2019) which includes the multivariate normal, t, Laplace, and several others as particular cases (McNeil et al., 2015; Browne and McNicholas, 2015). Following Finegold and Drton (2011), we demonstrate that, conditionally on each latent state, the inverse of the state-specific scale matrix of the multivariate GH completely characterizes the conditional correlation structure among the random variables. To induce sparsity in the inverse of the scale matrices, that is, the precision matrices, and identify whether two nodes are connected by an edge, we exploit the Gaussian location-scale mixture representation of

the GH family to build a suitable penalized expectation-conditional maximization either (ECME) algorithm. This enables us to include in our estimation procedure the *glasso* approach of Friedman et al. (2008) accounting for an L_1 penalty on the off-diagonal elements of the precision matrices in the M-step of the algorithm. We call this method the *hidden Markov generalized hyperbolic graphical model* (HMGHGM). Within this scheme, our modelling framework allows us to construct a set of regime-specific graphs whose set of edges is determined by the nonzero elements of the estimated state-specific precision matrices. As opposed to GGMs, the proposed methodology has several advantages. Firstly, it enables us to estimate levels of network connectivity among asset returns in different market phases corresponding to different states of the latent process. Secondly, we do not rely on the restrictive assumption of normally distributed data, but instead each state-dependent GH distribution has shape parameters that are free to vary within the GH family to provide the best fit to the data. Using simulation exercises, we validate the ability of our method to correctly (i) recover the true values of the parameters under different states of the Markov chain, (ii) identify the true HMM clustering partition, and (iii) retrieve the graphical model by recovering the true edges of the graphs.

Empirically, we analyze daily returns of a large set of financial assets, including the world's most important stock market indices, commodity futures, and the largest cryptocurrencies by market capitalization. In the last 10 years, the emergence of the cryptocurrency market has increasingly attracted the attention of market participants. The shortage of safe assets after the global financial crisis of 2008 has raised several concerns among investors and researchers about whether digital currencies can offer hedging and safe-haven abilities for equity investments (Bouri et al., 2020a). In the current literature, there are very few studies examining how traditional asset classes, (i.e., stocks, bonds, and foreign currencies), commodities and different cryptocurrencies interact with each other. The existing approaches have mainly focused on the relationships between a limited collection of cryptocurrencies and assets (Baur et al., 2018; Corbet et al., 2018; Ji et al., 2018; Bouri et al., 2020b; Chen et al., 2020; Giudici and Polinesi, 2021) or relied on conditional means to identify correlation structures (Bouri et al., 2017), which cannot provide a complete picture of the dependencies and the transmission path of risks between markets. Here, we implement the proposed HMGHGM to analyze the conditional correlation structure among the cryptocurrency, commodity, and stock market sectors from 2017 to 2023 and evaluate how it may vary when considering different volatility clusters. During the period considered, there were significant episodes of price fluctuations and instability, such as the explosion in cryptocurrencies at the start of 2018, the COVID-19 pandemic erupting in 2020, and the EU sanctions to Russia dictated by the Ukrainian war, which have triggered unexpected levels of uncertainty and high volatility. In this context, the latent process captures structural changes or regimes that influence the observed data's correlation patterns. These latent states represent "market conditions", such as periods of high and low volatility, that emerge in response to shocks and structural shifts in the markets.

To the best of our knowledge, this is the first attempt to build a non-Gaussian hidden Markov graphical model for estimating time-varying cross-market conditional association structures of financial returns. The rest of the article is organized as follows. In Section 2, we briefly review the GH distribution and formally introduce the HMGHGM. Section 3 proposes the ECME-based maximum likelihood approach and the related penalized algorithm for sparse estimation of the state-specific precision matrices. In the Supplementary Material we provide simulation results, while the empirical application is presented in Section 4. Section 5 summarizes our conclusions.

2 Methodology

In this section we introduce the hidden Markov generalized hyperbolic graphical model (HMGHGM). Before describing the model, we briefly revise the symmetric multivariate GH distribution, its location-scale mixture representation and its limiting cases. Subsequently, we show how it is possible to build sparse state-specific graphical models for characterizing time-varying conditional correlation relations among variables.

2.1 The GH distribution and its special cases

Formally, let $\mathbf{Y}_t = [Y_t^{(1)}, \dots, Y_t^{(d)}]'$ denote a continuous d-dimensional random vector for $t \in \{1, \dots, T\}$. The joint probability density function of $\boldsymbol{Y}_t$ following the (symmetric) GH distribution can be written as

$$f_{\boldsymbol{Y}_t}(\boldsymbol{y}_t; \boldsymbol{\mu}, \boldsymbol{\Sigma}, \lambda, \chi, \psi) = \frac{1}{(2\pi)^{d/2}|\boldsymbol{\Sigma}|^{1/2} K_\lambda(\sqrt{\psi\chi})} \left[\frac{\chi + \delta(\boldsymbol{y}_t; \boldsymbol{\mu}, \boldsymbol{\Sigma})}{\psi}\right]^{\frac{\lambda - d/2}{2}} K_{\lambda - d/2}\left(\sqrt{[\chi + \delta(\boldsymbol{y}_t; \boldsymbol{\mu}, \boldsymbol{\Sigma})]\psi}\right) \quad (1)$$

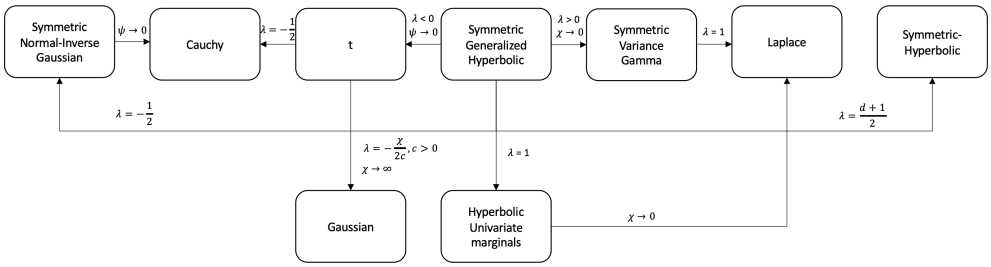


Figure 1: Special and limiting cases of the symmetric GH distribution in terms of λ, χ, and ψ.

where $\boldsymbol{\mu} \in \mathbb{R}^d$ is the location parameter, $\boldsymbol{\Sigma}$ is a $d \times d$ positive definite and symmetric scale matrix, such that $|\boldsymbol{\Sigma}| = 1$ for identifiability purposes (see McNeil et al. (2015) for details), $\lambda \in \mathbb{R}$ is the index parameter, $\chi > 0$ and $\psi > 0$ are concentration parameters, $\delta(\boldsymbol{y}_t; \boldsymbol{\mu}, \boldsymbol{\Sigma}) = (\boldsymbol{y}_t - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\boldsymbol{y}_t - \boldsymbol{\mu})$ is the squared Mahalanobis distance between $\boldsymbol{y}_t$ and $\boldsymbol{\mu}$ with scale matrix $\boldsymbol{\Sigma}$ and finally $K_{\lambda - d/2}(\cdot)$ denotes the modified Bessel function of the third kind of order $\lambda - d/2$. We adopt the compact notation $\boldsymbol{Y}_t \sim \mathcal{GH}_d(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \lambda, \chi, \psi)$. One of the key benefits of the GH distribution is that, using equation (1), $\boldsymbol{Y}_t$ admits the following location-scale mixture representation:

$$\mathbf{Y}_t = \boldsymbol{\mu} + \sqrt{W_t} \boldsymbol{\Sigma}^{1/2} \boldsymbol{Z}_t \tag{2}$$

where $\boldsymbol{Z}_t \sim \mathcal{N}_d(\mathbf{0}_d, \boldsymbol{I}_d)$ denotes a d-variate standard normal distribution whose covariance matrix is the identity matrix and W_t has a generalized inverse Gaussian (GIG) distribution, $W_t \sim \mathcal{GIG}(\lambda, \chi, \psi)$, with $\boldsymbol{Z}_t$ being independent of W_t. From equation (2), we can refer to the following hierarchical representation of $\boldsymbol{Y}_t \sim \mathcal{GH}_d(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \lambda, \chi, \psi)$:

$$\begin{aligned} W_t &\sim \mathcal{GIG}(\lambda, \chi, \psi), \\ \boldsymbol{Y}_t | W_t = w_t &\sim \mathcal{N}_d(\boldsymbol{\mu}, w_t \boldsymbol{\Sigma}) \end{aligned} \tag{3}$$

which is useful for random data generation and for the implementation of our ECME algorithm.

The GH family encompasses several well-known applied models for financial data by varying appropriately the values of the parameters of the GIG distribution, (λ, χ, ψ), including the multivariate Laplace, t, and normal distribution as presented in Figure 1.

A complete taxonomy of all the models belonging to the class of GH densities can be found, for instance, in McNeil et al. (2015), Browne and McNicholas (2015), and Bagnato et al. (2024).

2.2 Hidden Markov graphical models with state-specific GH densities

In this section we describe the proposed hidden Markov graphical model with state-specific multivariate GH emission densities. Formally, let $\{S_t\}_{t=1}^T$ be a latent, homogeneous, first-order Markov chain defined on the discrete state space $\{1, \dots, K\}$. Let $\pi_k = \Pr(S_1 = k)$ be the initial probability of state k, $k \in \{1, \dots, K\}$, and $\pi_{k|j} = \Pr(S_{t+1} = k | S_t = j)$, with $\sum_{k=1}^K \pi_{k|j} = 1$ and $\pi_{k|j} \geq 0$, denote the transition probability between states j and k, that is, the probability of visiting state k at time $t+1$ from state j at time t, $j, k \in \{1, \dots, K\}$, and $t \in \{1, \dots, T\}$. We collect the initial and transition probabilities in the K-dimensional vector $\boldsymbol{\pi}$ and in the $K \times K$ matrix $\boldsymbol{\Pi}$, respectively.

In our model we assume that the conditional distribution of $\boldsymbol{Y}_t$ given the state occupied by the latent process at time t corresponds to a GH distribution in equation (1) whose parameters depend on the value of the Markov chain S_t, namely, $\boldsymbol{Y}_t | S_t = k \sim \mathcal{GH}_d(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k, \lambda_k, \chi_k, \psi_k)$. We define with $\boldsymbol{\Theta}_k = \boldsymbol{\Sigma}_k^{-1}$ for $k \in \{1, \dots, K\}$, the precision matrices used to build state-specific undirected graphs that convey the conditional intercorrelation structure among the elements of $\boldsymbol{Y}_t$ given the latent process S_t. More precisely, supposing $S_t = k$, let $G_k = (V, E_k)$ be an undirected state-dependent graph where $V \in \{1, \dots, d\}$ denotes the set of state-invariant nodes, such that each component of the random variable $\boldsymbol{Y}_t$ corresponds to a node in V, and $E_k \subseteq V \times V$ represents the set of undirected edges in the kth state. In order to study the conditional correlation structure of $\boldsymbol{Y}_t$ within the kth state through the graph G_k, we establish a useful result that allows us to make inference on the edge set E_k based on the precision matrix $\boldsymbol{\Theta}_k$. Following Finegold and Drton (2011) and exploiting the mixture representation of the symmetric GH in equation (2), we can characterize the conditional correlation structure as shown in the following proposition.

Proposition 1. *Consider $S_t = k$ and let $\boldsymbol{Y}_t|S_t = k \sim \mathcal{GH}_d(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k, \lambda_k, \chi_k, \psi_k)$. If two nodes j and h, with $j, h \in V$ and $j \neq h$, of the graph are non adjacent in the sense that all paths between them are blocked by a set of nodes $C \in V$, then $Y_t^{(j)}$ and $Y_t^{(h)}$ are conditionally uncorrelated given $\boldsymbol{Y}_t^{(C)}$ and $S_t = k$.*

Proof. Without loss of generality, throughout the proof we omit the subscript k indicating the latent state yet all the following equalities are conditional on $S_t = k$. To prove the result it is sufficient to show that $Y_t^{(j)}$ and $Y_t^{(h)}$ are conditionally uncorrelated given $\boldsymbol{Y}_t^{(V\setminus\{j,h\})}$. Partition V into $a = \{j, h\}$ and $b = V/\{j, h\}$. For a given value of W_t and given $\boldsymbol{\Theta} = \boldsymbol{\Sigma}^{-1}$:

$$\left(\boldsymbol{Y}_t^{(a)}|\boldsymbol{Y}_t^{(b)}, W_t\right) \sim \mathcal{N}_2\left(\boldsymbol{\mu}^{(a)} - \boldsymbol{\Theta}_{a,a}^{-1}\boldsymbol{\Theta}_{a,b}\left(\boldsymbol{Y}_t^{(b)} - \boldsymbol{\mu}^{(b)}\right), W_t\boldsymbol{\Theta}_{a,a}^{-1}\right) \tag{4}$$

and

$$\left(Y_t^{(j)}|\boldsymbol{Y}_t^{(h\cup b)}, W_t\right) \sim \mathcal{N}_1\left(\mu^{(j)} - \Theta_{j,j}^{-1}\boldsymbol{\Theta}_{j,h\cup b}\left(\boldsymbol{Y}_t^{(h\cup b)} - \boldsymbol{\mu}^{(h\cup b)}\right), W_t\Theta_{j,j}^{-1}\right), \tag{5}$$

where $\boldsymbol{\Theta}_{a,b}$ is the submatrix of $\boldsymbol{\Theta}$ with rows and columns indexed by the sets a and b. Since $\Theta_{j,h} = 0$,

$$\begin{aligned} E\left[Y_t^{(j)}|\boldsymbol{Y}_t^{(h\cup b)}, W_t\right] &= \mu^{(j)} - \Theta_{j,j}^{-1}\boldsymbol{\Theta}_{j,b}\left(\boldsymbol{Y}_t^{(b)} - \boldsymbol{\mu}^{(b)}\right) \\ &= E\left[Y_t^{(j)}|\boldsymbol{Y}_t^{(b)}, W_t\right] \end{aligned} \tag{6}$$

for any value of W_t. Therefore,

$$\begin{aligned} E\left[Y_t^{(j)}|\boldsymbol{Y}_t^{(h\cup b)}\right] &= E\left[E\left[Y_t^{(j)}|\boldsymbol{Y}_t^{(h\cup b)}, W_t\right]|\boldsymbol{Y}_t^{(h\cup b)}\right] = E\left[E\left[Y_t^{(j)}|\boldsymbol{Y}_t^{(b)}, W_t\right]|\boldsymbol{Y}_t^{(b)}\right] \\ &= E\left[Y_t^{(j)}|\boldsymbol{Y}_t^{(b)}\right] \end{aligned} \tag{7}$$

which implies that $Y_t^{(j)}$ and $Y_t^{(h)}$ are conditionally uncorrelated given $\boldsymbol{Y}_t^{(b)}$. □

Based on equation (4), we can exploit the properties of GGMs to characterize conditional correlation relationships between the elements of $\boldsymbol{Y}_t$ in each state $k \in \{1, \ldots, K\}$, estimating the precision matrices $\boldsymbol{\Theta}_k$'s. More formally, for each state and for any pair of nodes j, h, with $j \neq h$, if the (j, h)th element of the matrix $\boldsymbol{\Theta}_k$, $\Theta_{k,j,h}$, is equal to 0, then $Y_t^{(j)}$ and $Y_t^{(h)}$ are conditionally uncorrelated given $\boldsymbol{Y}_t^{(V\setminus\{j,h\})}$ and $S_t = k$. Hence, for each $k \in \{1, \ldots, K\}$, the edge set E_k of the graph G_k describing the distribution of $\boldsymbol{Y}_t|S_t = k$ is completely encoded by the matrix $\boldsymbol{\Theta}_k$ of the GH distribution, that is, $(j, h) \in E_k$ if and only if $\Theta_{k,j,h} \neq 0$. The proposed methodology allows us to capture regime-specific conditional correlation structures for probability distributions within the GH family. With respect to alternative strategies, our framework is a generalization of the graphical model of Finegold and Drton (2011) when $K = 1$ where the GH reduces to the multivariate t distribution. In addition, we encompass the hidden Markov Gaussian graphical model (HMGlasso) of Städler and Mukherjee (2013) when the state-specific GH emission densities reduce to the multivariate normal distributions.

In this setting, it is therefore crucial to accurately determine the matrices $\boldsymbol{\Theta}_1, \ldots, \boldsymbol{\Theta}_K$ for a correct interpretation of the graphs and to visualize the true interactions among the variables. In Section 3, we introduce a maximum likelihood approach to estimate the model parameters and make inference on the sparsity pattern in the $\boldsymbol{\Theta}_k$'s.

3 Estimation

As shown in the previous sections, the location-scale mixture representation of the GH is a convenient tool to build a sparse graphical model. To estimate the model parameters of the method proposed based on the observed data, one could use direct numerical maximization of the likelihood (Ötting et al., 2021; Maruotti and Punzo, 2021) or exploit an EM algorithm (Baum et al., 1970). Here we focus on the latter. Since we are interested in detecting only the most important connections, we also propose a penalized version of the EM by considering the Lasso-type regularization of Tibshirani (1996) for sparse matrix estimation in high-dimensional settings where a large number of variables is available.

3.1 The EM algorithm

For a given number of hidden states K, the EM algorithm runs on the complete log-likelihood function of the model introduced, which is defined as

$$\ell_c(\boldsymbol{\theta}) = \sum_{k=1}^{K} \gamma_1(k) \log \pi_k + \sum_{t=2}^{T} \sum_{k=1}^{K} \sum_{j=1}^{K} \xi_t(j,k) \log \pi_{k|j} + \sum_{t=1}^{T} \sum_{k=1}^{K} \gamma_t(k) \log f_Y(\boldsymbol{y}_t | S_t = k), \tag{8}$$

where $\boldsymbol{\theta} = (\boldsymbol{\mu}_1, \dots, \boldsymbol{\mu}_K, \boldsymbol{\Sigma}_1, \dots, \boldsymbol{\Sigma}_K, \lambda_1, \dots, \lambda_K, \psi_1, \dots, \psi_K, \chi_1, \dots, \chi_K, \boldsymbol{\pi}, \boldsymbol{\Pi})$ represents the vector of all model parameters, $\gamma_t(k)$ denotes a dummy variable equal to 1 if the latent process is in state k at occasion t and 0 otherwise, and $\xi_t(j,k)$ is a dummy variable equal to 1 if the process is in state j in $t-1$ and in state k at time t and 0 otherwise. Unfortunately, the log-likelihood of the GH distribution does not yield any closed-form expression in the optimization process. Nevertheless, following Chatzis (2010), the issue can be resolved by exploiting the data augmentation scheme of equation (3). We exploit the conditional structure by writing

$$\log f_Y(\boldsymbol{y}_t | S_t = k) = l_1(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k | S_t = k) + l_2(\lambda_k, \chi_k, \psi_k | S_t = k) \tag{9}$$

where

$$l_1(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k | S_t = k) = \sum_{t=1}^{T} \left[-\frac{d}{2} \log(2\pi) - \frac{d}{2} \log(w_{tk}) - \frac{1}{2} \log |\boldsymbol{\Sigma}_k| - \frac{\delta(\boldsymbol{y}_t; \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)}{2 w_{tk}} \right] \tag{10}$$

and

$$l_2(\lambda_k, \chi_k, \psi_k | S_t = k) = \sum_{t=1}^{T} \left\{ (\lambda_k - 1) \log(w_{tk}) - \frac{1}{2} \frac{\chi_k}{w_{tk}} - \frac{1}{2} \psi_k w_{tk} - \frac{1}{2} \lambda_k \log(\chi_k) + \frac{1}{2} \lambda_k \log(\psi_k) - \log \left[2K_{\lambda_k} \left(\sqrt{(\psi_k \chi_k)} \right) \right] \right\}. \tag{11}$$

Working on $\ell_c(\boldsymbol{\theta})$, we adopt the ECME algorithm of Liu and Rubin (1994). The ECME algorithm is an extension of the expectation-conditional maximum (ECM) algorithm which replaces the M-step of the EM algorithm by a number of computationally simpler conditional maximization (CM) steps. The ECME algorithm generalizes the ECM algorithm by conditionally maximizing on some or all of the CM-steps the incomplete-data log-likelihood. In our case, the ECME algorithm iterates between four steps, one E-step and three CM-steps, until convergence. The three CM-steps arise from the update of the partition of $\boldsymbol{\theta}$ as $\{\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \boldsymbol{\theta}_3\}$, where $\boldsymbol{\theta}_1 = \{\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1, \dots, \boldsymbol{\mu}_K, \boldsymbol{\Sigma}_K\}$, $\boldsymbol{\theta}_2 = \{\lambda_1, \chi_1, \psi_1, \dots, \lambda_K, \chi_K, \psi_K\}$ and $\boldsymbol{\theta}_3 = \{\boldsymbol{\pi}, \boldsymbol{\Pi}\}$.

3.1.1 E-step

In the E-step, at the generic $(h+1)$ th iteration, the unobservable indicator variables $\gamma_t(k)$ and $\xi_t(j,k)$ in equation (8) are replaced by their conditional expectations given the observed data and the current parameter estimates $\boldsymbol{\theta}^{(h)}$. To compute such quantities we require the calculation of the probability of being in state k at time t given the observed sequence

$$\gamma_t^{(h)}(k) = P_{\boldsymbol{\theta}^{(h)}}(S_t = k | \boldsymbol{y}_1, \dots, \boldsymbol{y}_T) \tag{12}$$

and the probability that at time $t-1$ the process is in state j and then in state k at time t, given the observed sequence

$$\xi_t^{(h)}(j,k) = P_{\boldsymbol{\theta}^{(h)}}(S_{t-1} = j, S_t = k | \boldsymbol{y}_1, \dots, \boldsymbol{y}_T). \tag{13}$$

The quantities in equations (12) and (13) can be obtained using the forward-backward algorithm of Welch (2003).

Then, we use these to calculate the conditional expectation of the complete log-likelihood function (8), given the observed data and the current estimates:

$$Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(h)}) = \sum_{k=1}^{K} \gamma_1^{(h)}(k) \log \pi_k + \sum_{t=1}^{T}\sum_{k=1}^{K}\sum_{j=1}^{K} \xi_t^{(h)}(j,k) \log \pi_{k|j}$$
$$+ \sum_{t=1}^{T}\sum_{k=1}^{K} \gamma_t^{(h)}(k)\left[Q_1(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k|\boldsymbol{\theta}^{(h)}, S_t = k) + Q_2(\lambda_k, \chi_k, \psi_k|\boldsymbol{\theta}^{(h)}, S_t = k)\right]. \tag{14}$$

In equation (14), $Q_1(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k|\boldsymbol{\theta}^{(h)}, S_t = k)$ and $Q_2(\lambda_k, \chi_k, \psi_k|\boldsymbol{\theta}^{(h)}, S_t = k)$ are respectively the conditional expectations of $l_1(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k|S_t = k)$ and $l_2(\lambda_k, \chi_k, \psi_k|S_t = k)$ given the observed data, using the current $\boldsymbol{\theta}^{(h)}$ for $\boldsymbol{\theta}$. To compute Q_1 and Q_2, we have to consider the expected value of any function of the latent variable W_{tk} in equation (10). In particular the functions are w_{tk}, $1/w_{tk}$, and $\log(w_{tk})$ and their expected values must be calculated with respect to the current $\boldsymbol{\theta}^{(h)}$. Since the conditional distribution of W_{tk} given $\boldsymbol{Y}_t$ corresponds to a GIG distribution with parameters $\lambda_k - d/2, \delta(\boldsymbol{y}_t; \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k) + \chi_k, \psi_k$, that is,

$$f_{W_{tk}|\boldsymbol{Y}_t}(w_{tk}|\boldsymbol{Y}_t = \boldsymbol{y}_t, S_t = k) \sim GIG(\lambda_k - d/2, \delta(\boldsymbol{y}_t; \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k) + \chi_k, \psi_k), \tag{15}$$

it follows that (see Chatzis, 2010)

$$\upsilon_{tk}^{(h)} = E^{(h)}[W_{tk}|\boldsymbol{Y}_t = \boldsymbol{y}_t, S_t = k] =$$
$$= \left(\frac{\delta(\boldsymbol{y}_t; \boldsymbol{\mu}_k^{(h)}, \boldsymbol{\Sigma}_k^{(h)}) + \chi_k^{(h)}}{\psi_k^{(h)}}\right)^{\frac{1}{2}} \frac{K_{\lambda_k^{(h)} - d/2 + 1}\left(\sqrt{\psi_k^{(h)}\left[\delta(\boldsymbol{y}_t; \boldsymbol{\mu}_k^{(h)}, \boldsymbol{\Sigma}_k^{(h)}) + \chi_k^{(h)}\right]}\right)}{K_{\lambda_k^{(h)} - d/2}\left(\sqrt{\psi_k^{(h)}\left[\delta(\boldsymbol{y}_t; \boldsymbol{\mu}_k^{(h)}, \boldsymbol{\Sigma}_k^{(h)}) + \chi_k^{(h)}\right]}\right)}, \tag{16}$$

$$u_{tk}^{(h)} = E^{(h)}[W_{tk}^{-1}|\boldsymbol{Y}_t = \boldsymbol{y}_t, S_t = k] =$$
$$= \left(\frac{\delta(\boldsymbol{y}_t; \boldsymbol{\mu}_k^{(h)}, \boldsymbol{\Sigma}_k^{(h)}) + \chi_k^{(h)}}{\psi_k^{(h)}}\right)^{-\frac{1}{2}} \frac{K_{\lambda_k^{(h)} - d/2 + 1}\left(\sqrt{\psi_k^{(h)}\left[\delta(\boldsymbol{y}_t; \boldsymbol{\mu}_k^{(h)}, \boldsymbol{\Sigma}_k^{(h)}) + \chi_k^{(h)}\right]}\right)}{K_{\lambda_k^{(h)} - d/2}\left(\sqrt{\psi_k^{(h)}\left[\delta(\boldsymbol{y}_t; \boldsymbol{\mu}_k^{(h)}, \boldsymbol{\Sigma}_k^{(h)}) + \chi_k^{(h)}\right]}\right)}$$
$$- \frac{2\left(\lambda_k^{(h)} - d/2\right)}{\delta(\boldsymbol{y}_t; \boldsymbol{\mu}_k^{(h)}, \boldsymbol{\Sigma}_k^{(h)}) + \chi_k^{(h)}} \tag{17}$$

and

$$z_{tk}^{(h)} = E^{(h)}[\log(W_{tk})|\boldsymbol{Y}_t = \boldsymbol{y}_t, S_t = k] = \log\left(\sqrt{\frac{\delta(\boldsymbol{y}_t; \boldsymbol{\mu}_k^{(h)}, \boldsymbol{\Sigma}_k^{(h)}) + \chi_k^{(h)}}{\psi_k^{(h)}}}\right)$$
$$+ \frac{\partial}{\partial \lambda_k^{(h)}} \log K_{\lambda_k^{(h)} - d/2}\left(\sqrt{\psi_k^{(h)}\left[\delta(\boldsymbol{y}_t; \boldsymbol{\mu}_k^{(h)}, \boldsymbol{\Sigma}_k^{(h)}) + \chi_k^{(h)}\right]}\right). \tag{18}$$

As a consequence, substituting $w_{tk}, 1/w_{tk}$, and $\log(w_{tk})$ with $\upsilon_{tk}^{(h)}, u_{tk}^{(h)}$, and $z_{tk}^{(h)}$ respectively in $l_1(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k|S_t = k)$ and $l_2(\lambda_k, \chi_k, \psi_k|S_t = k)$, we obtain

$$Q_1(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k|\boldsymbol{\theta}^{(h)}, S_t = k) = \sum_{t=1}^{T}\left[-\frac{1}{2}\log|\boldsymbol{\Sigma}_k| - \frac{u_{tk}^{(h)}\delta(\boldsymbol{y}_t; \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)}{2}\right] \tag{19}$$

$$Q_2(\lambda_k, \chi_k, \psi_k|\boldsymbol{\theta}^{(h)}, S_t = k)$$
$$= \sum_{t=1}^{T}\left\{(\lambda_k - 1)z_{tk}^{(h)} - u_{tk}^{(h)}\frac{\chi_k}{2} - \frac{\psi_k}{2}\upsilon_{tk}^{(h)} - \frac{\lambda_k}{2}[\log(\chi_k) + \log(\psi_k)] - \log\left[2K_{\lambda_k}\left(\sqrt{(\psi_k\chi_k)}\right)\right]\right\} \tag{20}$$

where in equation (19), we dropped the terms which are constant with respect to $\boldsymbol{\mu}_k$ and $\boldsymbol{\Sigma}_k$.

3.1.2 CM-step 1

The initial probabilities π_k and transition probabilities $\pi_{k|j}$ of the partition θ_3 are updated using

$$\pi_k^{(h+1)} = \gamma_1^{(h)}(k), \quad k \in \{1, \dots, K\} \tag{21}$$

and

$$\pi_{k|j}^{(h+1)} = \frac{\sum_{t=2}^{T} \xi_t^{(h)}(j,k)}{\sum_{t=2}^{T}\sum_{k=1}^{K} \xi_t^{(h)}(j,k)}, \quad j,k \in \{1, \dots, K\}. \tag{22}$$

3.1.3 CM-step 2

The second maximization step requires the calculation of $\theta_1^{(h+1)}$ as the value of θ_1 that maximizes $Q_1(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k | \theta^{(h)}, S_t = k)$ in equation (19), with θ_2 fixed at $\theta_2^{(h+1)}$. The first-order conditions with respect to $\boldsymbol{\mu}_k$ and $\boldsymbol{\Theta}_k = \boldsymbol{\Sigma}_k^{-1}$ yield

$$\frac{\partial Q_1(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k | \theta_2^{(h)})}{\partial \boldsymbol{\mu}_k} = \sum_{t=1}^{T}\sum_{k=1}^{K} u_{tk}^{(h)} (\boldsymbol{y}_t - \boldsymbol{\mu}_k) \gamma_t^{(h)}(k) \tag{23}$$

$$\frac{\partial Q_1(\boldsymbol{\mu}_k, \boldsymbol{\Theta}_k | \theta_2^{(h)})}{\partial \boldsymbol{\Theta}_k} = \sum_{t=1}^{T}\sum_{k=1}^{K} \gamma_t^{(h)}(k) \left[\frac{1}{2} \frac{\partial}{\partial \boldsymbol{\Theta}_k} \log |\boldsymbol{\Theta}_k| - \frac{1}{2} u_{tk}^{(h)} \frac{\partial}{\partial \boldsymbol{\Theta}_k} (\boldsymbol{y}_t - \boldsymbol{\mu}_k)' \boldsymbol{\Theta}_k (\boldsymbol{y}_t - \boldsymbol{\mu}_k) \right] \tag{24}$$

so the CM-step update expressions for $\boldsymbol{\mu}_k$ and $\boldsymbol{\Sigma}_k$ are

$$\boldsymbol{\mu}_k^{(h+1)} = \frac{\sum_{t=1}^{T}\sum_{k=1}^{K} \gamma_t^{(h)}(k) u_{tk}^{(h)} \boldsymbol{y}_t}{\sum_{t=1}^{T}\sum_{k=1}^{K} \gamma_t^{(h)}(k) u_{tk}^{(h)}} \tag{25}$$

and

$$\boldsymbol{\Sigma}_k^{(h+1)} = |\boldsymbol{\Sigma}_k^{*(h+1)}|^{-1/d} \boldsymbol{\Sigma}_k^{*(h+1)} \tag{26}$$

where

$$\boldsymbol{\Sigma}_k^{*(h+1)} = \frac{\sum_{t=1}^{T}\sum_{k=1}^{K} \gamma_t^{(h)}(k) u_{tk}^{(h)} (\boldsymbol{y}_t - \boldsymbol{\mu}_k^{(h+1)})' (\boldsymbol{y}_t - \boldsymbol{\mu}_k^{(h+1)})}{\sum_{t=1}^{T}\sum_{k=1}^{K} \gamma_t^{(h)}(k)}. \tag{27}$$

In equation (26), the scalar $|\boldsymbol{\Sigma}_k^{*(h+1)}|^{-1/d}$ is needed to ensure the identifiability constraint $|\boldsymbol{\Sigma}_k^{(h+1)}| = 1$.

3.1.4 CM-step 3

In the third CM-step, we choose the value of θ_2 that maximizes $\ell_c(\theta)$ in equation (8), with θ_1 fixed at $\theta_1^{(h+1)}$. As a closed-form solution for $\theta_2^{(h+1)}$ is not analytically available, numerical optimization can be used with this aim. As in Bagnato et al. (2024), we perform an unconstrained maximization on $\mathbb{R}^3$, based on a (log/exp) transformation/back-transformation approach for χ_k and ψ_k.

The E- and CM-steps are alternated until convergence, that is, when the observed likelihood between two consecutive iterations is smaller than a predetermined threshold. In this article, we set the threshold criterion equal to 10^{-8}. For fixed K, we initialize the ECME algorithm by providing the initial states partition, $\{S_t^{(0)}\}_{t=1}^{T}$, according to the K-means algorithm. From the generated partition, the elements of $\boldsymbol{\Pi}^{(0)}$ are computed as proportions of transition. The location parameters $\boldsymbol{\mu}_1, \dots, \boldsymbol{\mu}_K$ are obtained from the centroid of the derived clusters while we set the initial values of the $\boldsymbol{\Sigma}_k$'s as the empirical covariance matrices of the obtained clusters with unit determinant using equation (26). Finally, the $(\lambda_k,\ \chi_k,\ \psi_k)$'s are initialized from uniform distributions. To deal with the possibility of multiple roots of the likelihood equation and to better explore the parameter space, we fit the proposed HMGHGM using a multiple random starts strategy with different starting partitions and retain the solution corresponding to the maximum likelihood value.

All the computations have been conducted using the R software, version 4.4.0 (R Core Team, 2024).

3.2 Penalized inference with the GH distribution

In this section, we extend the procedure described above for estimating high-dimensional graphs where the number of model parameters grows with the dimension of the problem. Indeed, the correct identification of the sparsity patterns is a fundamental issue for capturing the most relevant interconnections, which motivates us to use sparse estimators that automatically shrink the elements of the inverse of the scale matrix. In the literature, within the GGM context, several works have been put forward to obtain sparse estimates of the concentration matrix. Friedman et al. (2008) introduced the efficient coordinate descent algorithm, called *glasso*, to estimate a sparse graph using the Lasso L_1 penalty of Tibshirani (1996). Gao and Massam (2015) proposed to estimate the concentration matrix using further penalties, that is, the smoothly clipped absolute deviation of Fan and Li (2001) and the minimax concave penalty of Zhang (2010). Alternatively Meinshausen and Bühlmann (2006) presented a neighbourhood selection scheme that estimates the conditional independence relations separately for each node in the graph by using L_1-penalized regressions. In this work, we exploit the efficient *glasso* approach of Friedman et al. (2008) to induce sparsity in the precision matrices of the GH for each latent state, without relying on the limitation of normally distributed returns. Specifically, starting from the EM algorithm of Section 3 and following Green (1990), we construct a penalized EM (PEM) algorithm by adding to the complete likelihood in equation (8) an L_1-norm penalty that shrinks to zero the off-diagonal elements of the $\boldsymbol{\Theta}_k^{-1}$ matrices. The penalized version of the conditional normal log-likelihood in equation (10) is reported in the following definition.

Definition 1. The conditional normal log-likelihood function is proportional to:

$$\ell_{1,pen}(\boldsymbol{\mu}_k, \boldsymbol{\Theta}_k | \boldsymbol{Y}_t, W_{tk}, S_t = k) \propto \frac{1}{2} \sum_{t=1}^{T} \gamma_t(k) \log |\boldsymbol{\Theta}_k| - \frac{1}{2} \text{tr}\{\tilde{\boldsymbol{S}}_k \boldsymbol{\Theta}_k\} - \rho \sqrt{\nu_k} ||\boldsymbol{\Theta}_k||_1 \tag{28}$$

where $||\boldsymbol{\Theta}_k||_1$ is the sum of the absolute values of the off-diagonal entries of the matrix $\boldsymbol{\Theta}_k$ and $\tilde{\boldsymbol{S}}_k$ is the weighted empirical covariance matrix defined as

$$\tilde{\boldsymbol{S}}_k = \sum_{t=1}^{T} \frac{\gamma_t(k)}{W_{tk}} (\boldsymbol{Y}_t - \boldsymbol{\mu}_k)(\boldsymbol{Y}_t - \boldsymbol{\mu}_k)'. \tag{29}$$

Additionally, $\nu_k > 0$ is a state-dependent weight to allow for penalty adaptation to state-sizes such that $\sum_{k=1}^{K} \nu_k = 1$ and $\rho \geq 0$ is the tuning parameter.

As one can see, the penalized complete likelihood in equation (28) is exactly the objective function maximized in a GGM where the empirical covariance matrix of the data is substituted by the matrix $\tilde{\boldsymbol{S}}_k$ defined in equation (29), for $k \in \{1, \ldots, K\}$. This suggests that we can exploit the simple and fast *glasso* algorithm of Friedman et al. (2008) to update the estimate of $\boldsymbol{\Theta}_k$. More specifically, for a given value of ρ, the proposed PEM alternates between the E-step described in Section 3.1 and modifies the CM-step 2 by maximizing the quantity in equation (28). Furthermore, when dealing with HMMs a scaling issue may emerge due to the presence of latent states. Different states may exhibit variations in scale, but due to the unknown state assignments a priori, standardization cannot be implemented as a preprocessing step (Städler and Mukherjee, 2013). For this reason, in the Supplementary Material we consider a state-specific penalty that not only adjusts for state sizes but also addresses scaling issues.

4 Empirical application

4.1 Data description

In recent decades, hedge fund managers and regulators have highlighted the need to accurately assess contagion and systemic risk, depending on the situation of the economy. In this article, we are interested in investigating how the degree of network connectivity across different asset classes changes in severe market turbulence, which may threaten the integrity of the financial system and represent a potential source of financial instability. In our application, the underlying market dynamics and economic conditions are captured via latent states (or "regimes") that shape how the observable returns of different assets co-move over time.

For this purpose, we apply the proposed methodology to daily returns of $d = 29$ financial assets comprising stock market indices, commodity futures, and digital currencies. The set of assets includes the 10 largest cryp-

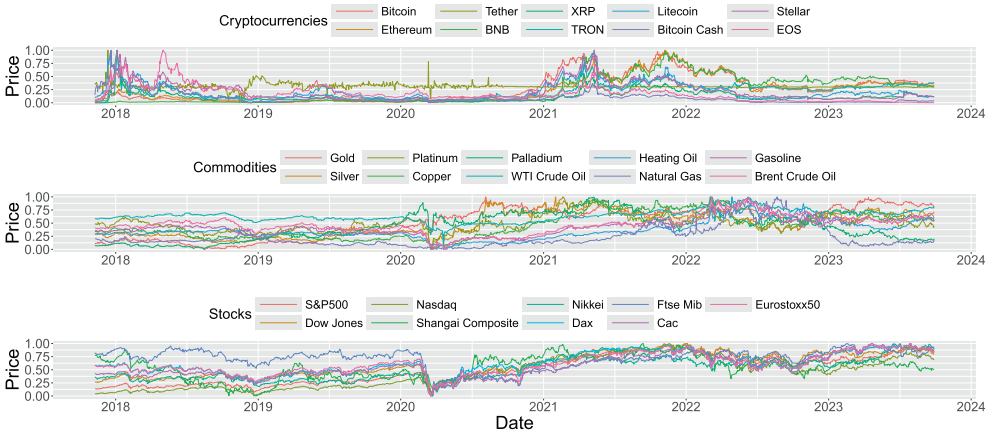


Figure 2: From top to bottom daily time-series of normalized prices of the assets considered. Observations span from November 10, 2017 to September 28, 2023.

tocurrencies in terms of market capitalization, including Bitcoin, Ethereum, Ripple and Stellar, the 10 most exchanged metal and energy commodities, such as Gold, Silver, Crude Oil, and Natural Gas, and the 9 major world stock indices, including the S&P 500, FTSE MIB, Nikkei 225, and the Shanghai Composite Index.

The sample dataset was collected from Yahoo! Finance and the study period starts on November 10, 2017 and ends on September 28, 2023, for a total of $T = 1250$ observations after removing missing data. This timespan is marked by numerous crises that may have impacted cross-market correlation patterns, spanning from the cryptocurrency bubble crisis in 2017-2018 to the global crisis sparked by the COVID-19 pandemic in 2020-2021, which caused unprecedented levels of uncertainty and risk. Daily returns with continuous compounding are calculated by taking the logarithm of the difference between closing prices on consecutive trading days. In Table 1 we report the list of examined variables and the summary statistics for the whole sample. As one can see, each asset displays the typical shape features that characterize the financial markets, and cryptocurrencies are generally much more volatile than commodity futures and stock market indices, having the highest standard deviations. Such findings reflect the large price variations during the period under consideration, which were triggered by the burst of the cryptocurrency bubble in 2018 and by the COVID-19 outbreak at the dawn of 2020. In conclusion, the Augmented Dickey-Fuller (ADF) test (Dickey and Fuller, 1979) shows that all daily returns are stationary at the 1% level of significance. Figure 2 shows the daily prices of the 29 assets over the entire period of observation. At first glance we immediately recognize the high volatility that characterizes the market of the cryptocurrencies and the waves of exponential price increases. The first wave started at the end of 2018, the first big "bubble" that brought BTC above $20,000 USD, another one in mid-2019, and the most recent and bigger one started in late 2020, which developed over the entire course of 2021. With regard to commodities and stock indices, we observe quite a different path with respect to digital currencies. The collapse of the financial markets at the beginning of 2020 caused by the COVID-19 pandemic represents a watershed for volatility regimes. After March 2020 we see an increase in price volatility for a large proportion of commodities and stock indices. We also observe price peaks of gasoline, WTI crude and Brent crude during the course of 2022 caused by the sanctions imposed against Russia in response to the Ukrainian war.

4.2 Results

Following these considerations, the proposed HMGHGM can provide insights into the temporal evolution of the conditional correlations of asset returns, where the assets' multivariate distribution is not constrained to be normal but can be assumed to be any fat-tailed distribution within the GH family. As a first step of the empirical analysis, in order to select the optimal ρ in equation (28) and number of latent states K, we fit the proposed HMGHGM for a sequence of 300 values of ρ for $K = \{1, \dots, 4\}$. We also include a state-specific penalty term by setting the weight ν_k in equation (28) as the (scaled) effective sample size of state k, that is, $\nu_k = \sum_{t=1}^{T} \gamma_t(k)/T$, where $\gamma_t(k)$ is defined in equation (12). To select the best pair (K, ρ), we consider two model selection criteria, namely, a Bayesian Information Criterion (BIC) type index and the mixture minimum description length

DOI: 10.1002/cjs.70030

Table 1: Summary statistics of daily log-returns over the entire period. The Augmented Dickey-Fuller (ADF) test statistic is displayed in boldface when the null hypothesis is rejected at the 1% significance level.

Sector	Variable	Minimum	Mean	Maximum	Stdev	Skewness	Kurtosis	ADF test
Cryptocurrency	1: Bitcoin	−46.47	0.11	22.51	5.03	−0.76	9.01	**−10.02**
	2: Ethereum	−55.07	0.13	32.5	6.47	−0.66	7.73	**−10.09**
	3: Tether	−5.26	0	5.66	0.48	0.77	49.86	**−14.98**
	4: BNB	−54.31	0.38	84.62	7.46	1.19	22.42	**−9.19**
	5: XRP USD	−69.26	0.07	74.65	8.21	1.13	20.88	**−9.9**
	6: TRON	−69.21	0.29	174.43	9.7	5.14	94.96	**−8.37**
	7: Litecoin	−44.91	0	53.98	6.88	0.18	8.55	**−10.22**
	8: Bitcoin Cash	−56.13	−0.08	50.38	8.01	0.21	9.13	**−10.66**
	9: Stellar	−42.96	0.08	80.74	8.06	1.9	20.88	**−9.94**
	10: EOS	−50.42	−0.06	65.12	8.18	0.33	8.78	**−10.37**
Commodities	11: Gold	−5.11	0.03	5.81	1.01	−0.22	4.68	**−10.5**
	12: Silver	−16.08	0.03	14.71	2.02	−0.38	9.08	**−10.25**
	13: Platinum	−12.35	0	11.18	1.99	−0.27	4.07	**−11.32**
	14: Copper	−8.14	0.02	6.91	1.55	−0.3	2.26	**−10.93**
	15: Palladium	−23.4	0.03	22.6	2.73	−0.43	10.5	**−12.38**
	16: Crude Oil	−28.22	0.09	42.31	3.57	0.99	33.4	**−8.87**
	17: Heating Oil	−26.14	0.06	15.42	2.91	−1.21	13.78	**−11.36**
	18: Natural Gas	−30.05	−0.01	38.17	4.34	0.17	7.92	**−9.74**
	19: RBOB Gasoline	−50.89	0.05	28.88	3.46	−2.51	48.5	**−9.23**
	20: Brent Crude Oil	−27.58	0.06	27.42	2.88	0.09	18.14	**−9.81**
Stock Indices	21: S&P500	−12.77	0.04	8.97	1.38	−0.94	13.81	**−10.1**
	22: Dow Jones	−13.84	0.03	10.76	1.38	−1.12	18.85	**−10.4**
	23: Nasdaq	−13.15	0.06	8.93	1.62	−0.62	7.07	**−10.1**
	24: Shanghai Composite Index	−8.04	−0.01	7.55	1.17	−0.39	4.98	**−11.33**
	25: Nikkei	−6.27	0.03	7.73	1.27	−0.01	3.48	**−11.02**
	26: Dax	−13.05	0.01	10.41	1.37	−0.56	12.89	**−10.25**
	27: Ftse Mib	−11.51	0	8.67	1.15	−1.05	14.21	**−10.59**
	28: Cac	−13.1	0.02	8.06	1.35	−0.92	12.45	**−10.56**
	29: Eurostoxx50	−13.24	0.01	8.83	1.35	−0.88	12.75	**−10.45**

Table 2: BIC and MMDL values corresponding to the optimal ρ with varying number of states for HMGHGM and HMGlasso. Bold font highlights the best values for the considered criteria, where lower is better.

	$K=1$	$K=2$	$K=3$	$K=4$
Panel A: HMGHGM				
BIC	148666.4	**141381.2**	141727.2	143860
MMDL	148711.1	140743.5	**140252.9**	141313
Panel B: HMGlasso				
BIC	151328.2	**142899.7**	143277.9	143261.8
MMDL	151328.2	142076.0	141627.1	**140479.1**

(MMDL, Figueiredo et al., 1999) defined, respectively, as

$$\text{BIC}_{K,\rho} = -\ell(\boldsymbol{\theta}|\boldsymbol{y}_1,\dots,\boldsymbol{y}_T) + \frac{1}{2}\log(T)K(K-1) + \frac{1}{2}\log(T)\sum_{k=1}^{K} Df(k,\rho) \tag{30}$$

$$\text{MMDL}_{K,\rho} = -\ell(\boldsymbol{\theta}|\boldsymbol{y}_1,\dots,\boldsymbol{y}_T) + \frac{1}{2}\log(T)K(K-1) + \sum_{k=1}^{K}\frac{1}{2}\log(T\nu_k)Df(k,\rho), \tag{31}$$

where $\ell(\boldsymbol{\theta}|\boldsymbol{y}_1,\dots,\boldsymbol{y}_T)$ denotes the observed log-likelihood and we set the degrees of freedom as $Df(k,\rho) = d + 3 + \sum_{l\geq l'} \mathbf{1}_{\boldsymbol{\theta}^{ll'}_{k,\rho}\neq 0}$. The two criteria are identical, with the only difference being that the MMDL adjusts the penalty term for the effective sample size $T\nu_k$ of each state k. The optimal K and ρ are obtained by selecting the model with the lowest criterion value. To compare models with differing number of states, Table 2 (Panel A) reports the BIC and MMDL values for each K at the optimal ρ. As can be seen, the BIC selects 2 states, while MMDL chooses $K = 3$ at the optimal ρ. For comparison, we also fitted the HMGlasso of Städler and Mukherjee (2013) and reported the corresponding BIC and MMDL values in Panel B of the table. By comparing these goodness-of-fit criteria, we observe that the proposed model based on GH distributions provides a better fit to the data than the normal distribution, as evidenced by both criteria. Following the works of Figueiredo et al. (1999) and Städler and Mukherjee (2013), demonstrating that, in general, the MMDL outperforms the BIC, we thus select the proposed HMGHGM with $K = 3$ states for our analyses. To obtain standard errors of the parameter estimates we adopt a parametric bootstrap approach. That is, we refitted the model to 500 bootstrap resamples from the estimated model, and 95% confidence intervals are obtained via the percentile method.

In light of these points, we begin commenting on the results by the hidden process. The top of Figure 3 represents the estimated posterior probabilities of being in latent state j, at time t, where $j \in \{1,\dots,K\}$ and $t \in \{1,\dots,T\}$, conditional on the observed time-series and with an over-imposed trend according to a smoothed local regression. The shaded areas represent 95% bootstrap confidence intervals. The bottom of Figure 3 shows the decoded states obtained through local decoding by considering the maximum of the posterior probabilities and the Viterbi algorithm. The two methods convey almost identical results. The predicted trajectories indicate that the three states are visited 38%, 39%, and 23% of the entire period. Overall, the following conclusions can be drawn. The distribution of states closely mirrors the dynamics of the considered portfolio during the last five years. In particular, State 3 manages to capture the speculative waves of the cryptocurrency market at the end of 2017, in mid-2019, and at the most recent surge during 2021. As can be observed in Figure 2, these periods are characterized by extreme price movements with exponential price change behaviours and can represent subsequent dramatic losses for investors. States 1 and 2 instead succeed in mimicking the behaviour of commodities and stock markets during the considered time frame. State 1, which identifies the period that goes from the end of 2017 to the collapse of financial markets of March 2020 induced by the COVID-19 pandemic, is related to a phase of low prices. State 2 instead indicates the phases of the markets that, after the second half of 2020, are characterized by a rise in prices. Overall, Figure 3 suggests to us that the proposed model efficiently manages to identify high and low prices and volatility trends that alternate over the years. These patterns are also demonstrated by the estimated transition matrix reported in Table 3, alongside 95% bootstrap confidence intervals (in parentheses). We observe that State 2 is highly persistent, suggesting that market conditions characterized by rising prices endure for long periods of time, whereas States 1 and 3 are less persistent. Moreover, markets are more likely to transition from State 3 to State 1, whereas the opposite switch is less likely to occur.

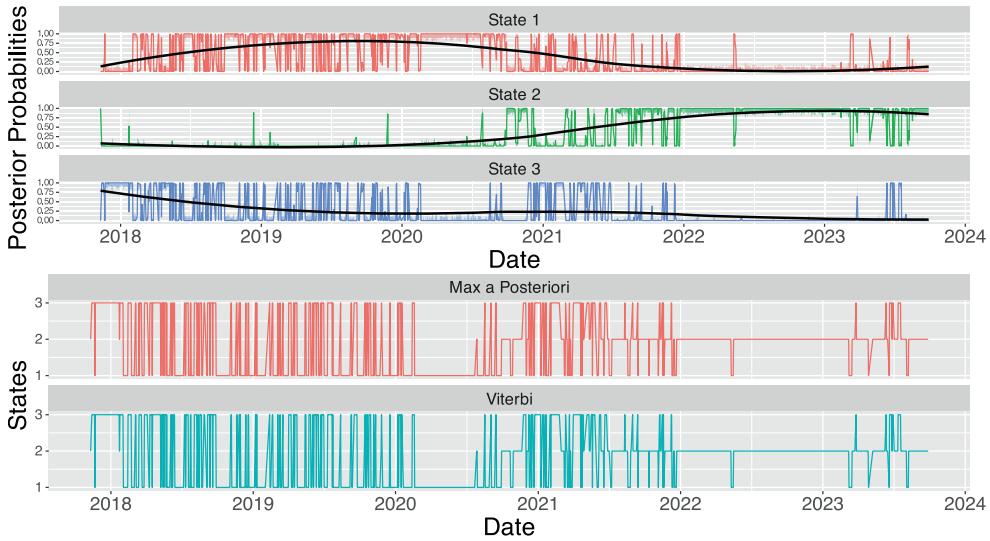


Figure 3: From top to bottom, predicted posterior probabilities with 95% bootstrap confidence intervals (shaded areas) and predicted sequence of hidden states over time with $K = 3$ states. The over-imposed black lines denote the trend estimated by a local linear regression.

Table 3: Estimated probability transition matrix $\mathbf{\Pi}$ with 95% bootstrap confidence intervals in parentheses obtained over 500 resamples.

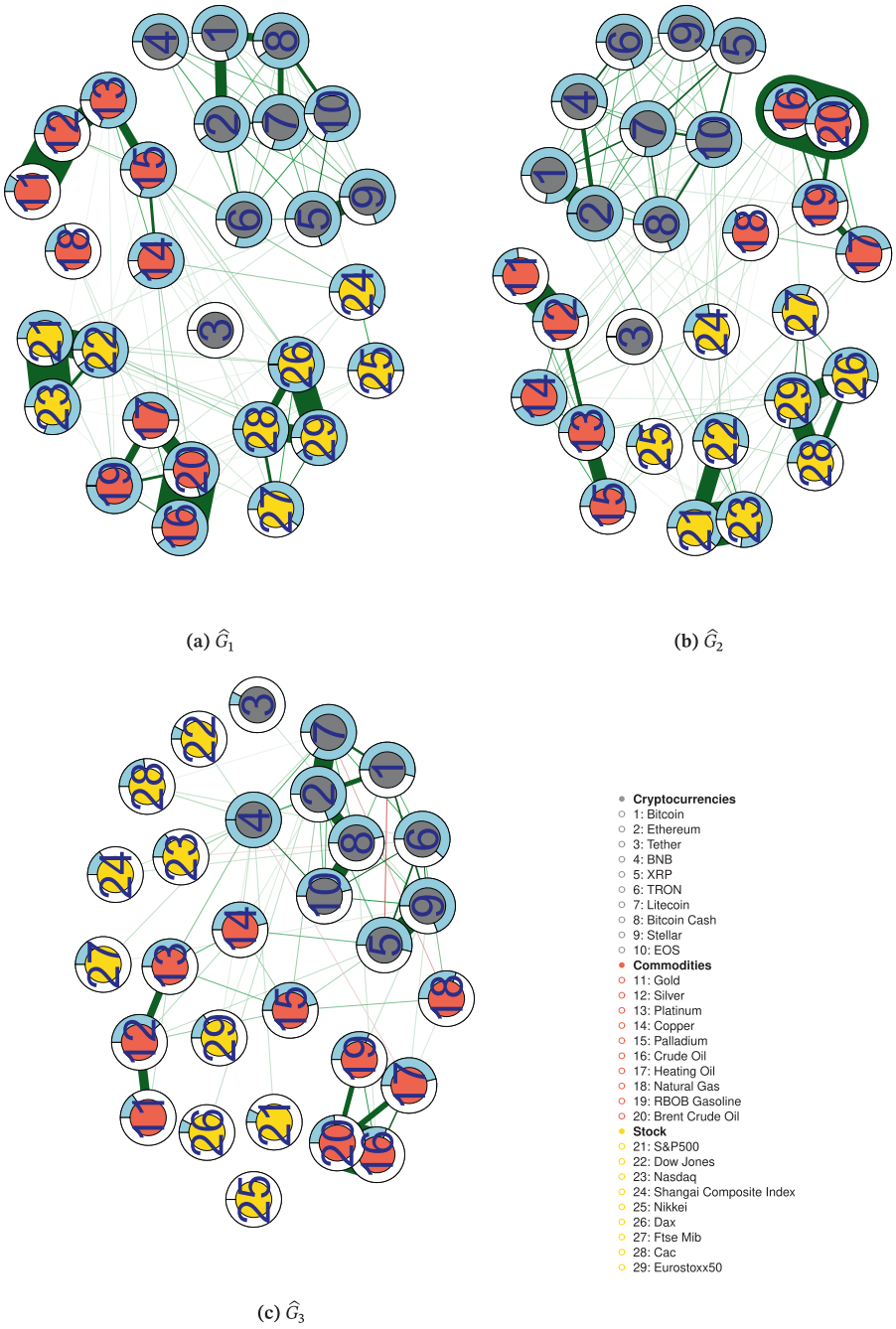
	State 1	State 2	State 3
State 1	0.776 (0.736, 0.812)	0.045 (0.028, 0.068)	0.179 (0.145, 0.213)
State 2	0.040 (0.023, 0.065)	0.924 (0.890, 0.944)	0.037 (0.019, 0.057)
State 3	0.313 (0.258, 0.370)	0.056 (0.031, 0.087)	0.630 (0.571, 0.688)

Table 4: State-specific parameters λ_k, χ_k, and ψ_k, $k = 1, \dots, K$, with 95% bootstrap confidence intervals in parentheses obtained over 500 resamples.

Parameter	State 1	State 2	State 3
λ_k	−1.819 (−1.876, −1.054)	−3.158 (−3.408, −1.975)	−1.856 (−2.179, −1.556)
χ_k	5.254 (4.372, 6.085)	8.454 (6.847, 9.759)	7.325 (5.399, 8.258)
ψ_k	0.000 (0.000, 0.043)	0.000 (0.000, 0.142)	0.027 (0.008, 0.511)

Moving on to the observable process, we begin by analyzing the shape of the state-dependent distributions. Table 4 presents the estimates and 95% bootstrap confidence intervals of the state-specific parameters λ_k, χ_k, and ψ_k, $k = 1, \dots, K$. Clearly, the estimated GH distributions provide a better fit to the empirical kurtosis in the data than the Gaussian and indicate severe departures from normality. In all three states, the estimated densities show heavy tails. Particularly, State 2 exhibits the most pronounced extreme tail behaviour and a leptokurtic-shaped distribution.

In order to summarize the interconnectedness among the considered assets during different market conditions, we build the regime-specific graphs $\widehat{G}_k = (V, \widehat{E}_k)$, $k \in \{1, \dots, K\}$, where the variables in Table 1 represent the vertices in V, identified by the corresponding estimated precision matrix $\widehat{\mathbf{\Theta}}_k$, $k \in \{1, \dots, K\}$. Specifically, an edge between two nodes is created in the kth graph if the estimated $\widehat{\Theta}_{k,i,l} \neq 0$, that is, $(i, l) \in \widehat{E}_k$ if and only if $\widehat{\Theta}_{k,i,l} \neq 0$, for $i, l \in V, i \neq l$. Figure 4a–c reports the estimated graphs for the selected three-states HMGHGM. To highlight the most important variables in the network, pies in light blue indicate the degree centrality measure of each node, while the edge colours specify the sign of the corresponding interaction (green = positive, red = negative), and the edge thickness the magnitude of association. In each graph, stock market indices are coloured in yellow, cryptocurrencies in grey while commodities are shown in red, and the vertex labels are



(a) $\widehat{G}_1$ (b) $\widehat{G}_2$

(c) $\widehat{G}_3$

Figure 4: Estimated graphs for $k = 1$ (a), $k = 2$ (b), and $k = 3$ (c). Yellow, grey, and red nodes represent respectively indices, cryptocurrencies, and commodities while the vertex labels are illustrated in Table 1. Green edges in the networks depict positive associations, while red edges represent negative associations. Pies in light blue indicate degree centrality. The edge thickness represents the magnitude of association. (a) $\widehat{G}_1$, (b) $\widehat{G}_2$, (c) $\widehat{G}_3$.

reported in Table 1. In Figure 4, each graph identifies a different regime of network connectivity consistently with the remarks given above. We identify a high asymmetry in network connectivity: State 3 represents the speculative waves of the cryptocurrency market, and the third graph shows, as expected, a strong centrality and clustering pattern of the crypto assets. The most central nodes (BNB, TRON, Litecoin, Stellar) represent some of the assets with the most "explosive" trends during the crypto bubble. The strong clustering pattern is present across the three estimated graphs. Specifically, digital currencies are strongly tied to each other and are rather

isolated from other assets (Dyhrberg, 2016; Bouri et al., 2017; Corbet et al., 2018; Giudici and Polinesi, 2021). It emerges how they might act as safe-haven assets offering protection to investors against losses to offset market risk, being uncorrelated with stocks during market up and downturns. In all three networks, Bitcoin is one of the least central nodes among the cryptos. This is in line with the work of Bouri et al. (2019), which explains that the behaviour of crypto-traders is not to anchor to a price level: after an exponential price appreciation in one cryptocurrency (e.g., Bitcoin), they start to look for more attractive cryptocurrencies that have a better risk-reward profile. One example is represented by Stellar (9) and Litecoin (7), which appear always very central in each state of the chain. On the other hand, volatility of commodities and stock indices is closely related to the uncertainty of the economic outlook. As one can see, in the first two states of the chain the most central nodes are represented by stocks (in particular Dow Jones, Nasdaq, Dax, and Eurostoxx50) and commodities (i.e., Platinum, Copper, Crude Oil, and Gasoline), which have been particularly affected by the interruptions in global production chains over the course of the pandemic. We detect strong clustering patterns for all three sectors considered. In particular, we highlight the strongest connections within metals, energy, European and US indices, representing possible co-movements during both upward and downward trends, probably due to the fact that assets of the same sectors or indices of the same regions have responded similarly to the economic crisis caused by the pandemic. Finally, we emphasize that gold and natural gas are two of the most disconnected nodes in the graph for every volatility level. While the expectation for gold's disconnection aligns with its role as a safe haven asset, the disconnection observed for natural gas is likely attributable to the anomalous and sudden price surge of this commodity in 2022 due to the Russo-Ukrainian conflict. The results of the analysis with an alternative penalty able to handle state-specific scaling are shown in the Supplementary Material.

5 Conclusions

This article proposes a new sparse dynamic graphical model to study time-varying conditional correlations between important commodities, cryptocurrencies, and stock indices in a period of high volatility and market turbulence, without relying on the assumption of normally distributed data. Starting from Finegold and Drton (2011) and Städler and Mukherjee (2013), we introduce a hidden Markov graphical model where, conditional on the latent state, emission densities follow a (symmetric) GH distribution with state-dependent parameters. Since the index and concentration parameters are completely free to vary within the GH family they allow us to accommodate well a wide range of empirical characteristics of the data in real applications. To induce sparsity and recover only the most prominent relations, we develop a penalized ECME algorithm exploiting the Gaussian location-scale mixture representation of the GH distribution with a Lasso L_1 penalty on the elements of the state-specific precision matrices. Our procedure is computationally efficient and can be implemented by making use of the *glasso* of Friedman et al. (2008) within the conditional M-step of the proposed algorithm. The performance of our model is evaluated using numerical simulations under different scenarios.

In the real data analysis, we considered daily returns of 29 cryptocurrencies, commodities, and market indices from November 2017 to September 2023 to investigate the dynamics of a conditional correlation structure across different financial markets during low and high volatility periods. According to the MMDL criteria, we identified three latent regimes corresponding to different degrees of network connectivity. The first two regimes show negative and positive phases of the market, respectively, while the third manages to capture the peculiar explosive trends of the cryptocurrencies. The estimated graphs are consistent with the hypothesis that cryptocurrencies are highly connected to each other and disconnected from traditional asset types (Bouri et al., 2017; Corbet et al., 2018; Giudici and Polinesi, 2021), resulting in possible safe-haven assets during tumultuous times. Commodities that stand out in terms of degree centrality are the ones that have been the most affected by the interruptions in global production chains over the course of the pandemic, namely, palladium, copper, Crude Oil and gasoline.

Outside the financial world, a similar methodology could be employed, for example, in the field of dynamic mixed graphical models, which could convey critical information when employing not just continuous but also discrete datasets. Future research may also extend the proposed graphical model by using multivariate asymmetric distributions within the GH family.

Data sharing

The data used for the analysis were downloaded from Yahoo! Finance and are made available in `.RData` format at the following link: https://doi.org/10.17605/OSF.IO/7ETYC. These data were derived from the following resources available in the public domain: https://finance.yahoo.com/.

Acknowledgement

Open access publishing facilitated by Universita degli Studi di Roma La Sapienza, as part of the Wiley - CRUI-CARE agreement.

ORCID

Beatrice Foroni: https://orcid.org/0000-0002-4683-2879
Luca Merlo: https://orcid.org/0000-0001-5267-5390
Lea Petrella: https://orcid.org/0000-0002-1653-5443

Supplementary Material

The Supplementary Material consists of a PDF file that includes the simulation studies and the application of HMGHGM with a different penalty function.

References

Bagnato, L., Farcomeni, A., and Punzo, A. (2024). The generalized hyperbolic family and automatic model selection through the multiple-choice LASSO. *Statistical Analysis and Data Mining: The ASA Data Science Journal*, 17(1):e11652.

Banerjee, O., Ghaoui, L. E., and d'Aspremont, A. (2008). Model selection through sparse maximum likelihood estimation for multivariate Gaussian or binary data. *Journal of Machine Learning Research*, 9(Mar):485–516.

Baum, L. E., Petrie, T., Soules, G., and Weiss, N. (1970). A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains. *The Annals of Mathematical Statistics*, 41(1):164–171.

Baur, D. G., Hong, K., and Lee, A. D. (2018). Bitcoin: Medium of exchange or speculative assets? *Journal of International Financial Markets, Institutions and Money*, 54:177–189.

Bianchi, D., Billio, M., Casarin, R., and Guidolin, M. (2019). Modeling systemic risk with Markov switching graphical SUR models. *Journal of Econometrics*, 210(1):58–74.

Bouri, E., Molnár, P., Azzi, G., Roubaud, D., and Hagfors, L. I. (2017). On the hedge and safe haven properties of Bitcoin: Is it really more than a diversifier? *Finance Research Letters*, 20:192–198.

Bouri, E., Shahzad, S. J. H., and Roubaud, D. (2019). Co-explosivity in the cryptocurrency market. *Finance Research Letters*, 29:178–183.

Bouri, E., Shahzad, S. J. H., and Roubaud, D. (2020a). Cryptocurrencies as hedges and safe-havens for US equity sectors. *The Quarterly Review of Economics and Finance*, 75:294–307.

Bouri, E., Shahzad, S. J. H., Roubaud, D., Kristoufek, L., and Lucey, B. (2020b). Bitcoin, gold, and commodities as safe havens for stocks: New insight through wavelet analysis. *The Quarterly Review of Economics and Finance*, 77:156–164.

Browne, R. P. and McNicholas, P. D. (2015). A mixture of generalized hyperbolic distributions. *The Canadian Journal of Statistics*, 43(2):176–198.

Brunetti, C., Harris, J. H., Mankad, S., and Michailidis, G. (2019). Interconnectedness in the interbank market. *Journal of Financial Economics*, 133(2):520–538.

Cai, T., Liu, W., and Luo, X. (2011). A constrained ℓ_1 minimization approach to sparse precision matrix estimation. *Journal of the American Statistical Association*, 106(494):594–607.

Chatzis, S. P. (2010). Hidden Markov models with non-elliptically contoured state densities. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(12):2297–2304.

Chen, Y., Giudici, P., Hadji Misheva, B., and Trimborn, S. (2020). Lead behaviour in Bitcoin markets. *Risks*, 8(1):4.

Chen, Y., Härdle, W., and Jeong, S.-O. (2008). Nonparametric risk management with generalized hyperbolic distributions. *Journal of the American Statistical Association*, 103(483):910–923.

Corbet, S., Meegan, A., Larkin, C., Lucey, B., and Yarovaya, L. (2018). Exploring the dynamic relationships between cryptocurrencies and other financial assets. *Economics Letters*, 165:28–34.

De Angelis, L. and Paas, L. J. (2013). A dynamic analysis of stock markets using a hidden Markov model. *Journal of Applied Statistics*, 40(8):1682–1700.

Dempster, A. P. (1972). Covariance selection. *Biometrics*, 28:157–175.

Dickey, D. A. and Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, 74(366a):427–431.

Drton, M. and Perlman, M. D. (2004). Model selection for Gaussian concentration graphs. *Biometrika*, 91(3):591–602.

Drton, M. and Perlman, M. D. (2008). A SINful approach to Gaussian graphical model selection. *Journal of Statistical Planning and Inference*, 138(4):1179–1200.

Dyhrberg, A. H. (2016). Bitcoin, gold and the dollar–a GARCH volatility analysis. *Finance Research Letters*, 16:85–92.

Fan, J. and Li, R. (2001). Variable selection via non-concave penalized likelihood and its oracle properties. *Journal of the American statistical Association*, 96(456):1348–1360.

Figueiredo, M. A., Leitao, J. M., and Jain, A. K. (1999). On fitting mixture models. In *International Workshop on Energy Minimization Methods in Computer Vision and Pattern Recognition*, Springer, Berlin Heidelberg. 54–69.

Finegold, M. and Drton, M. (2011). Robust graphical modeling of gene networks using classical and alternative t-distributions. *The Annals of Applied Statistics* 5(2A):1057–1080.

Foroni, B., Merlo, L., and Petrella, L. (2024a). Expectile hidden Markov regression models for analyzing cryptocurrency returns. *Statistics and Computing*, 34(2):66.

Foroni, B., Merlo, L., and Petrella, L. (2024b). Quantile and expectile copula-based hidden Markov regression models for the analysis of the cryptocurrency market. *Statistical Modelling*, 25:454–472. doi: 10.1177/1471082X241279513.

Friedman, J., Hastie, T., and Tibshirani, R. (2008). Sparse inverse covariance estimation with the graphical Lasso. *Biostatistics*, 9(3):432–441.

Gao, X. and Massam, H. (2015). Estimation of symmetry-constrained Gaussian graphical models: Application to clustered dense networks. *Journal of Computational and Graphical Statistics*, 24(4):909–929.

Giudici, P. and Parisi, L. (2018). Corisk: Credit risk contagion with correlation network models. *Risks*, 6(3):95.

Giudici, P. and Polinesi, G. (2021). Crypto price discovery through correlation networks. *Annals of Operations Research*, 299(1):443–457.

Green, P. J. (1990). On use of the EM algorithm for penalized likelihood estimation. *Journal of the Royal Statistical Society: Series B (Methodological)*, 52(3):443–452.

Ignatieva, K. and Landsman, Z. (2015). Estimating the tails of loss severity via conditional risk measures for the family of symmetric generalised hyperbolic distributions. *Insurance: Mathematics and Economics*, 65:172–186.

Ji, Q., Bouri, E., Gupta, R., and Roubaud, D. (2018). Network causality structures among Bitcoin and other financial assets: A directed acyclic graph approach. *The Quarterly Review of Economics and Finance*, 70:203–213.

Konlack Socgnia, V. and Wilcox, D. (2014). A comparison of generalized hyperbolic distribution models for equity returns. *Journal of Applied Mathematics*, 2014:1–15.

Lauritzen, S. L. (1996). *Graphical Models*, Vol. 17, Clarendon Press, Oxford, UK.

Liu, C. and Rubin, D. B. (1994). The ECME algorithm: A simple extension of EM and ECM with faster monotone convergence. *Biometrika*, 81(4):633–648.

Liu, H., Han, F., Yuan, M., Lafferty, J., and Wasserman, L. (2012). High-dimensional semiparametric Gaussian copula graphical models. *The Annals of Statistics*, 40(4):2293–2326.

Liu, H., Lafferty, J., and Wasserman, L. (2009). The nonparanormal: Semiparametric estimation of high dimensional undirected graphs. *Journal of Machine Learning Research*, 10(10):2295–2328.

Liu, H. and Wang, L. (2017). TIGER: A tuning-insensitive approach for optimally estimating Gaussian graphical models. *Electronic Journal of Statistics*, 11(1):241–294.

MacDonald, I. L. and Zucchini, W. (1997). *Hidden Markov and Other Models for Discrete-Valued Time Series*, Vol. 110, CRC Press, Boca Raton, Florida.

Maruotti, A., Petrella, L., and Sposito, L. (2021). Hidden semi-Markov-switching quantile regression for time series. *Computational Statistics & Data Analysis*, 159:107208.

Maruotti, A. and Punzo, A. (2021). Initialization of hidden Markov and semi-Markov models: A critical evaluation of several strategies. *International Statistical Review*, 89(3):447–480.

Maruotti, A., Punzo, A., and Bagnato, L. (2019). Hidden Markov and semi-Markov models with multivariate leptokurtic-normal components for robust modeling of daily returns series. *Journal of Financial Econometrics*, 17(1):91–117.

McNeil, A. J., Frey, R., and Embrechts, P. (2015). *Quantitative Risk Management: Concepts, Techniques and Tools - Revised Edition*, Princeton University Press, New Jersey, USA.

Meinshausen, N. and Bühlmann, P. (2006). High-dimensional graphs and variable selection with the Lasso. *The Annals of Statistics*, 34(3):1436–1462.

Mergner, S. and Bulla, J. (2008). Time-varying beta risk of Pan-European industry portfolios: A comparison of alternative modeling techniques. *The European Journal of Finance*, 14(8):771–802.

Necula, C. (2009). Modeling heavy-tailed stock index returns using the generalized hyperbolic distribution. *Romanian Journal of Economic Forecasting*, 10(2):118–131.

Nystrup, P., Madsen, H., and Lindström, E. (2017). Long memory of financial time series and hidden Markov models with time-varying parameters. *Journal of Forecasting*, 36(8):989–1002.

Ötting, M., Langrock, R., and Maruotti, A. (2021). A copula-based multivariate hidden Markov model for modelling momentum in football. *AStA Advances in Statistical Analysis*, 107:1–19.

Pennoni, F., Bartolucci, F., Forte, G., and Ametrano, F. (2022). Exploring the dependencies among main cryptocurrency log-returns: A hidden Markov model. *Economic Notes*, 51(1):e12193.

R Core Team (2024). *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/.

Silva, T. C., da Silva Alexandre, M., and Tabak, B. M. (2018). Bank lending and systemic risk: A financial-real sector network approach with feedback. *Journal of financial Stability*, 38:98–118.

Silva, W., Kimura, H., and Sobreiro, V. A. (2017). An analysis of the literature on systemic financial risk: A survey. *Journal of Financial Stability*, 28:91–114.

Städler, N. and Mukherjee, S. (2013). Penalized estimation in high-dimensional hidden Markov models with state-specific graphical models. *The Annals of Applied Statistics*, 7(4):2157–2179.

Tibshirani, R. (1996). Regression shrinkage and selection via the Lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, 58(1):267–288.

Welch, L. R. (2003). Hidden Markov models and the Baum-Welch algorithm. *IEEE Information Theory Society Newsletter*, 53(4):10–13.

Xue, L. and Zou, H. (2012). Regularized rank-based estimation of high-dimensional nonparanormal graphical models. *The Annals of Statistics*, 40(5):2541–2571.

Zhang, C.-H. (2010). Nearly unbiased variable selection under minimax concave penalty. *The Annals of Statistics*, 38(2):894–942.

Zhang, Y., Chu, J., Chan, S., and Chan, B. (2019). The generalised hyperbolic distribution and its subclass in the analysis of a new era of cryptocurrencies: Ethereum and its financial risk. *Physica A: Statistical Mechanics and its Applications*, 526:120900.

Zucchini, W., MacDonald, I. L., and Langrock, R. (2016). *Hidden Markov Models for Time Series: An Introduction Using R*, Chapman and Hall/CRC, Boca Raton, Florida.

Received 10 September 2024 — Accepted 30 August 2025