Change point detection in dynamic Gaussian graphical models: the impact of COVID-19 pandemic on the US stock market

Beatrice Franzolini[1], Alexandros Beskos[2], Maria De Iorio[3], Warrick Poklewski Koziell[4], and Karolina Grzeszkiewicz[5]

[1,3]Singapore Institute for Clinical Sciences, Agency for Science, Technology and Research
[2,3,4]Department of Statistical Science, University College London, UK
[3]Yong Loo Lin School of Medicine, National University of Singapore
[5]Yale-NUS College, Singapore

Abstract

Reliable estimates of volatility and correlation are fundamental in economics and finance for understanding the impact of macroeconomics events on the market and guiding future investments and policies. Dependence across financial returns is likely to be subject to sudden structural changes, especially in correspondence with major global events, such as the COVID-19 pandemic. In this work, we are interested in capturing abrupt changes over time in the dependence across US industry stock portfolios, over a time horizon that covers the COVID-19 pandemic. The selected stocks give a comprehensive picture of the US stock market. To this end, we develop a Bayesian multivariate stochastic volatility model based on a time-varying sequence of graphs capturing the evolution of the dependence structure. The model builds on the Gaussian graphical models and the random change points literature. In particular, we treat the number, the position of change points, and the graphs as object of posterior inference, allowing for sparsity in graph recovery and change point detection. The high dimension of the parameter space poses complex computational challenges. However, the model admits a hidden Markov model formulation. This leads to the development of an efficient computational strategy, based on a combination of sequential Monte-Carlo and Markov chain Monte-Carlo techniques. Model and computational development are widely applicable, beyond the scope of the application of interest in this work.

1 Introduction

Understanding the temporal evolution of the dependence structure among time series is a fundamental topic in many fields, such as psychology, speech recognition, genomics, and, in particular, finance. In this latter context, estimates of volatility and correlation of different financial instruments are largely used for portfolio allocation, option-pricing, and to draw conclusions about the impact of macroeconomics events on the markets with the goal of guiding future investments and policies. In particular, estimates of correlation are key to minimise risk of investment portfolios and define hedging strategies (see, among others, Lien et al., 2002; Lee, 2010; Thampanya et al., 2020; Dutta et al., 2021). Changes in correlation modify the return/risk profile of the investments and are of interest for both investors and policy markers. To understand how to better prepare for and deal with future major global events, it is important to estimate the impact of the COVID-19 pandemic

on volatility and dependence structure of financial instruments (Just and Echaust, 2020; Sakurai and Kurosaki, 2020; Alqaralleh and Canepa, 2021; Guidolin et al., 2021; Yousfi et al., 2021; Derbali et al., 2022; Dey et al., 2022). In correspondence with global "catastrophic" events, such as COVID-19 or financial crises, dependence across time series is likely to be subjected to sudden structural changes. However, standard statistical approaches assume time-varying dependence to change smoothly over time. In this manuscript, we develop a statistical machinery to detect abrupt changes in the correlation structure among time series. Such machinery is employed to detect the impact of the COVID-19 pandemic on the US stock market and, in particular, on the cross-industry relationships.

There exists a vast literature on models for time-varying second moments. More specifically, there are two main approaches: *conditional volatility models*, as the well-known ARCH and GARCH (Engle, 1982; Bollerslev et al., 1994; Bollerslev, 1986; Engle and Bollerslev, 1986; Bauwens et al., 2006; Silvennoinen and Teräsvirta, 2009; Boudt et al., 2019), and *stochastic volatility models* (e.g., Taylor, 1982; Wiggins, 1987; Hull and White, 1987; Asai et al., 2006). The former class specifies second moments at a certain time t as deterministic function of past values of observations, volatility, and possibly covariance, given model parameters. The latter assumes second moments to follow a latent stochastic process, typically of Markovian structure, so that, even conditionally on all past information, volatility and correlations are unobservable random variables evolving over time. While stochastic volatility models are often more flexible and may achieve better inferential performances when compared to conditional volatility approaches (Chan, 2013; Clark and Ravazzolo, 2015), they are more difficult to estimate since the likelihood is typically intractable, see, e.g., Nilsson (2016).

Within both classes, a further distinction may be made between models that explicitly target the covariance matrix Σ_t and those focusing on the precision matrix $\Omega_t = \Sigma_t^{-1}$, specifically allowing for zero entries in Ω_t to favour parsimony. In this work, we develop a Bayesian stochastic volatility model for the precision matrix. More in details, the precision matrix at time t is modelled conditionally on a graph at time t, which describes the dependence structure among time series. As such, our work lies within the literature on Gaussian graphical models (GGMs) (see, e.g., Carvalho and West, 2007; Wang and West, 2009; Prado and West, 2010; Wang, 2010; Chandra et al., 2021). This approach presents an important advantage: GGMs target conditional independence instead of marginal, leading to possible identification of macro-components (represented, for instance, by hubs and cliques in the graph) and safeguarding against spurious relationships. The identification of graph substructures is of particular interest in finance, where hubs may be interpreted as risk factors driving the market, while cliques represent financial instruments exposed to the same unobserved risk factor. See Figure 1 for a toy example clarifying the role of graph substructures in financial markets. Moreover, marginal dependence and Pearson correlation simply measure the pairwise co-movement between two investments, but do not provide any indication on whether a risk factor generating the co-movement is specific to the two investments or it is common also to other financial instruments of interest. On the contrary, entries of the precision matrix represent co-movements conditionally on the effect of all the other instruments considered in the model (Michis, 2022), so that an entry is non-zero if and only if the two returns are dependent conditionally on all other investments.

Changes over time of second moments can be smooth or abrupt. The focus of this work is on changes of the second type. Standard versions of the models cited so far assume variances and covariances changing smoothly over time and, in particular, between any two consecutive time points. For instance, Carvalho and West (2007) propose a Bayesian dynamic stochastic volatility model based on GGMs and conditional independence. In

| Market n.1 | Market n.2 | Market n.3 | Market n.4 |

$X = Z + \epsilon_x$ $X = W + \epsilon_x$ $X = W + \epsilon_x$ $X = \epsilon_x$

$Y = Z + \epsilon_y$ $Y = W + \epsilon_y$ $Y = W + \epsilon_y$ $Y = \epsilon_y$

$Z = \epsilon_z$ $Z = W + \epsilon_z$ $Z = \epsilon_z$ $Z = \epsilon_z$
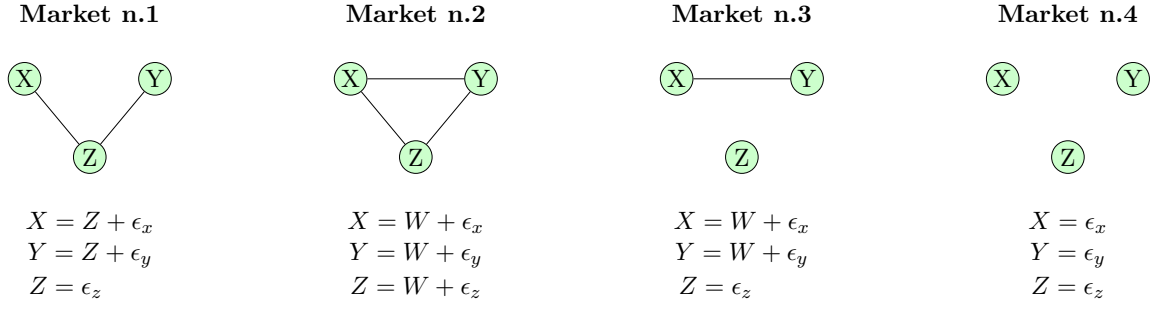
Figure 1: Toy example of graphs substructures in a financial market with three assets. Market n.1: the asset Z acts as common risk factor in the market driving the dependence and is an hub. Market n.2: the return of all three assets X, Y, and Z are driven by an unobservable risk factor W and the tree assets form a clique. Market n.3: X and Y are driven by a common risk factor, not affecting the asset Z. Market n.4: the three assets are independent and the graph does not include any edge. Here $\epsilon_z, \epsilon_x, \epsilon_y$ are white noises.

their construction, the graph structure is kept constant over time, while the covariance matrix changes smoothly between any two consecutive time points. However, this feature is often in contrast with what is observed in financial markets, where volatility clusters (i.e., periods with a persistent value of volatility, that are interrupted by sudden changes) and correlation breakdowns (i.e., substantial changes in correlations during stressed times and financial crises) are well documented (see, for example, Von Furstenberg et al., 1989; Contessi et al., 2014).

To detect the possible effect of COVID-19 pandemic on the US stock market, we analyse the correlation structure between nine industry portfolios, considering logarithmic weekly returns in the years 2019, 2020, and 2021. Weekly returns are computed starting from the daily returns available at Kennth R. French's Data Library. Figure 2 shows empirical estimates of a graph describing the conditional dependence structure over time. The estimates are obtained using a moving window of 80 weeks, shifting in steps of 20 weeks from one graph to the next. Graph estimates are obtained using an adaptive lasso approach as implemented in the R package GGMselect. From the preliminary analysis, changes in correlation are already evident, as well as the role of hub of the manufacturing and consumer non-durables industries. However, it is difficult to determine whether the pandemic had an effect on the overall structure and when. Moreover, it is well known that edge estimation in GGMs is sensitive to sample size and this empirical approach is highly dependent on the arbitrary choice of the window size (and corresponding sample size). To achieve our inferential goals, a sound modelling strategy is needed to be able to effectively infer the existence and location of change points due to sudden changes, still borrowing information across the entire time horizon,

Sudden changes in volatility and correlation have been modelled generalizing either conditional or stochastic volatility models with the introduction of Markov switching regimes (see, among others, So et al., 1998; Haas et al., 2004; Caporale and Zekokh, 2019). However, in a frequentist framework, Markov switching approaches require an arbitrary choice of the number of different regimes, requiring ad hoc criteria for model choice. Alternatively, penalised likelihood techniques have been successfully employed for estimating dynamic GGMs (Zhou et al., 2010; Danaher et al., 2014; Yang et al., 2015; Gibberd and Nelson, 2017), however, such approaches do not allow for uncertainty quantification on the number and temporal location of the abrupt changes.

In this work, we introduce a Bayesian dynamic GGM to detect abrupt changes in the
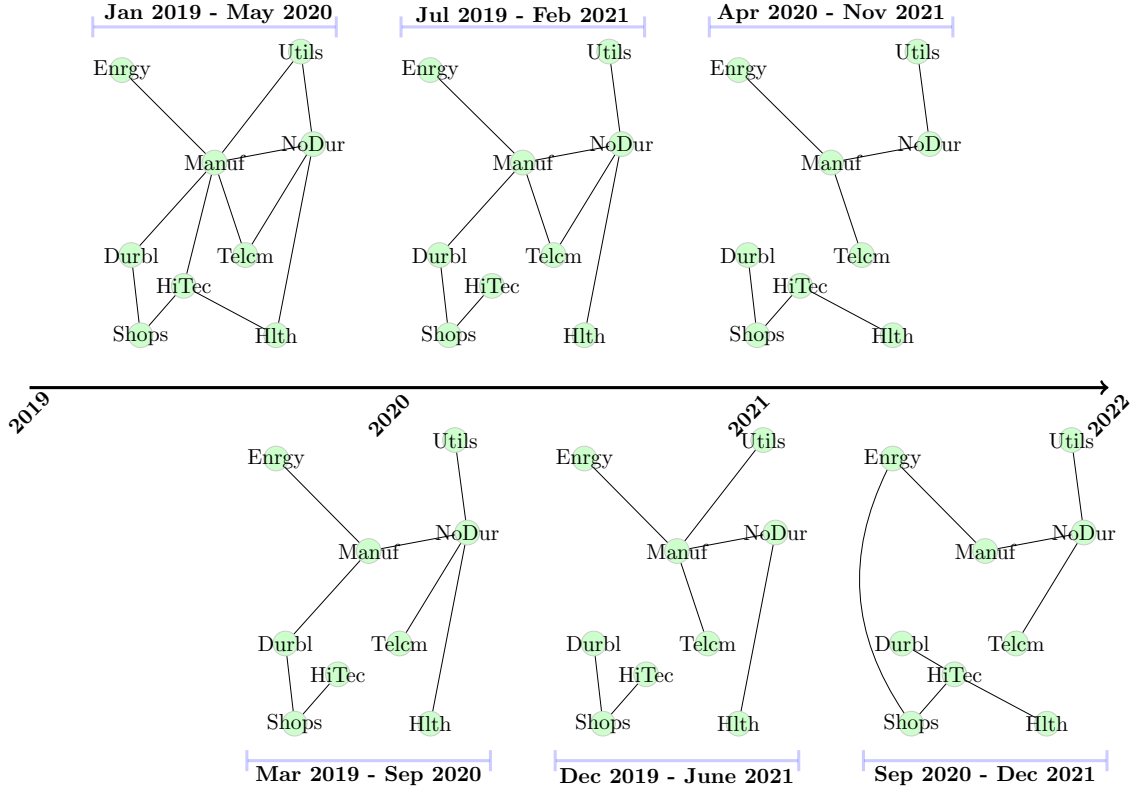
Figure 2: Moving window: graph estimated with the R package GGMselect. Each graph is computed on 80 time points. The window is moved by 20 time points at the time.

conditional dependence structure between time series. In brief, this is a piece-wise constant stochastic volatility model for graphs, which favours sparsity at three levels, by explicitly penalizing: (i) the number of change points; (ii) the number of edges within each graph; (iii) the number of edges which are activated (appear) and deactivated (disappear) at each change point. In a Bayesian framework, it is straightforward, at least in principle, to perform posterior inference also on the number of change points and on their location. Finally, we notice that our model allows for deviations from the assumption of global Gaussianity, which implies that the overall distribution of the data is a Gaussian distribution and poses challenges in the analysis of financial returns, often characterised by heavy tails. Indeed, we assume a Gaussian distribution locally, between two consecutive change points, possibly accommodating for the excess of kurtosis typically observed in financial returns' distributions. More precisely, the introduction of change points allows the Multivariate Gaussian distribution to change along the overall time horizon, so that the observed empirical distribution can be though as been drawn from a mixture of Gaussians, which can accommodate heavy tails (Cui, 2012).

The paper is structured as follows. In Section 2, the dynamic GGM is presented. Section 3 contains a discussion of the computational challenges, the proposed algorithm, and a simulation study to assess the performance of the sequential Monte-Carlo procedure. Results on simulated data and on the US stock market data can be found in Section 4 and Section 5, respectively. Section 6 concludes the paper with a discussion about future directions and extensions. In Supplementary Material we provide the dataset, R codes to replicate all the results in this work, and additional results on the simulation studies and the application.

2 The dynamic Gaussian graphical model

We first introduce some definitions and notation. Let $G = (V, E)$ represent an undirected acyclic graph, where $V = \{1, \ldots, p\}$, $p \geq 1$, corresponds to the set of labelled nodes and $E \subseteq \{(h, k) \in V \times V : h < k\}$ the set of edges linking pairs of nodes. There is a one-to-one correspondence between G and its $p \times p$ adjacency binary matrix A, which is defined as follows. The element $A[h, k]$ on the h−th row and k−th column is equal to 1 when an edge exists between nodes h and k, and to 0 otherwise. Note that A is symmetric with zeros on the main diagonal, since G is simple. When each node corresponds to a random variable, the graph structure can be used to encode conditional independence so that an edge is present between vertices h and k if and only if the h−th and k−th random variables are dependent conditionally on all other variables in the graph (Lauritzen, 1996).

A powerful modelling tool is offered by GGMs, which assume that the distribution of the random variables represented by the nodes in V is Multivariate Gaussian. Then, the precision matrix Ω can be modelled conditionally on the graph, so that the presence of an edge between two nodes in G implies a non-zero entry in the precision matrix between the corresponding random variables, while the absence of an edge implies a zero entry. Let the cone M^+ be the space of symmetric positive-definite matrices on $\mathbb{R}^{p \times p}$. For graph G and adjacency matrix A, $M^+(G) \subset M^+$ denotes the set of the matrices, M, with $M[h, k] = 0$ if and only if $A[h, k] = 0$, for any $h \neq k$, so that $\Omega \in M^+(G)$.

In a time series setting, let $G_t = (V, E_t)$ describe the (conditional) dependence structure at time t between p time series, each corresponding to one node in V. We propose a prior distribution for the process $\{G_t, t \geq 1\}$, obtained by letting t vary, which lies in the class of stochastic volatility models.

Data are collected at common discrete time points $t = 1, 2, \ldots, T$. We denote with Y_t the vector of observations at time t on the p variables (i.e., returns at week t for the considered industry portfolios) and with $Y_{1:T} = [Y_t]_{t=1}^T$ the $T \times p$ data matrix. We assume that, conditionally on a time-indexed collection of precision matrices $\{\Omega_t, t = 1, \ldots, T\}$, the vectors of observations are normally distributed and independent over time, i.e.,

$$Y_t \mid \Omega_t \overset{ind}{\sim} N_p(0, \Omega_t^{-1}) \qquad \text{for } t = 1, \ldots, T \tag{1}$$

where $N_p(\mu, \Sigma)$ denotes a p-variate Gaussian distribution with mean μ and covariance matrix Σ.

We model Ω_t conditionally on a graph at time t, G_t. Then, to allow for time-varying dependence structure among the p variables, we introduce a sequence of random change points. A time point t is said to be a *change point* if the dependence structure among the p observable variables changes between $t - 1$ and t, i.e., if $G_t \neq G_{t-1}$ and, consequently, $\Omega_t \neq \Omega_{t-1}$. Let $c_{1:\kappa} = (c_1, c_2, \ldots, c_\kappa)$ be the vector of ordered change points, which, similarly to the precision matrices and the graphs, are unobserved. Here, κ denotes the (random) number of change points. In the following we use the conventions $c_0 = 1$, $c_{\kappa+1} = T + 1$. Note that between consecutive change points the graph and the corresponding precision matrix are kept constant. Given the sequence of graphs, $G_{1:T} = \{G_t, t = 1, \ldots, T\}$, and change points, we assume that

$$\Omega_1 \mid G_1 \sim W_{G_1}(\delta, D) \tag{2}$$

and, for $t \geq 2$,

$$\Omega_t \mid \Omega_{t-1}, G_t, c_{1:\kappa} \sim \begin{cases} W_{G_t}(\delta, D), & \text{if } t \in \{c_1, \ldots, c_k\} \\ \delta_{\Omega_{t-1}}, & \text{otherwise} \end{cases} \tag{3}$$

where, δ_x denotes the Dirac delta distribution at x and $W_G(\delta, D)$ the G-Wishart distribution (see Roverato, 2002; Dobra et al., 2011), with shape parameter $\delta > 2$ and inverse scale matrix parameter $D \in M^+$. Its density w.r.t. the Lebesgue measure of dimension equal to the free elements of a matrix in $M^+(G)$ is

$$P(\Omega|G) = \frac{1}{I_G(\delta, D)} \, |\Omega|^{(\delta-2)/2} \exp\big\{ -\tfrac{1}{2}\mathrm{tr}(D\,\Omega)\big\}, \quad \Omega \in M^+(G)$$

The constraints for hyper-parameters δ and D suffice to ensure integrability of the above density (Diaconis and Ylvisaker, 1979). The normalizing constant is equal to

$$I_G(\delta, D) = \int_{M^+(G)} |\Omega|^{(\delta-2)/2} \exp\big\{ -\tfrac{1}{2}\mathrm{tr}(D\,\Omega)\big\}\,\mathrm{d}\Omega$$

and will be used later to compute the marginal likelihood of the data conditionally only on the graph structure.

To complete the model, we next describe the graph dynamics. Denote with A_t the adjacency matrix corresponding to G_t, then, for all $h, k \in \{1, \ldots, p\}$ with $h < k$, we specify the prior distributions

$$A_1[h, k] \mid \omega \overset{iid}{\sim} \mathrm{Bernoulli}\left(\frac{2\omega}{p-1}\right) \tag{4}$$

and, for $t \geq 2$,

$$A_t[h, k] \mid A_{t-1}[h, k], c_{1:\kappa}, z \overset{ind}{\sim} \begin{cases} |\, A_{t-1}[h, k] - \mathrm{Bernoulli}\left(\frac{2z}{p-1}\right) \,|, & \text{if } t \in \{c_1, \ldots, c_k\} \\ \delta_{A_{t-1}[h,k]}, & \text{otherwise} \end{cases} \tag{5}$$

Notice that the hyper-parameter $\omega \in [0, (p-1)/2]$ controls the graph sparsity, so that the expected number of edges for the initial graph a priori equals $p\,\omega$, while the hyper-parameter $z \in [0, (p-1)/2]$ controls the impact of an event on graph structure when a change point is reached, in particular the (a priori) expected number of edges that will change is equal to $p\,z$. Our prior choice is reminiscent of the one proposed in Jones et al. (2005), who recommend setting a prior edge inclusion probability equal to $2/(p-1)$ so that the expected number of edges is p. The extra parameters ω and z allow for more control on graph sparsity and temporal dependence.

Equations (1) and (2)-(5) can be viewed as observation and state dynamics, respectively, of a hidden Markov model with the unobserved signal corresponding to the pair $\{(G_{c_i}, \Omega_{c_i}), i = 0, 1, \ldots, \kappa\}$ (see Figure 3 for a graphical representation). For more details see, for example, West and Harrison (2006).

Finally, the prior distribution for $c_{1:\kappa}$ is chosen as

$$\begin{aligned} c_{1:\kappa}|\kappa &\sim \mathrm{Uniform}(\mathcal{T}_{\kappa\ell}) \\ \kappa \mid p_0 &\sim \mathrm{Truncated-Geometric}(p_0) \qquad \text{for } \kappa = 0, 1, \ldots, K_{T\ell} \end{aligned} \tag{6}$$

for hyper-parameter $p_0 \in (0, 1)$, so that the (a priori) number of expected change points is

$$\mathbb{E}[\kappa] = \frac{1 - p_0}{p_0} \frac{1 - (1 - p_0)^{K_{T\ell}} (p_0 K_{T\ell} + 1)}{1 - (1 - p_0)^{K_{T\ell}+1}}$$

Through appropriate choice of p_0 in (6), we are able to enforce the desired level of sparsity on the number of change points. Here, $\mathcal{T}_{\kappa\ell}$ is the space of ordered κ-tuples $c_1 < \cdots < c_\kappa$, with c_j in $\{2, \ldots, T\}$, under the minimum-span constraint that $c_{j+1} - c_j \geq \ell$, for any $j = 0, 1, \ldots, \kappa$,
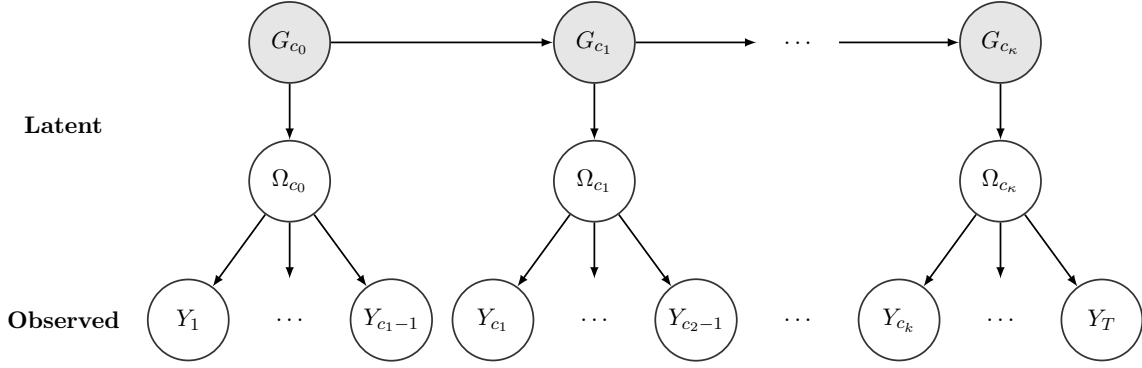
Figure 3: Graphical representation of the model conditionally on change points. Both graph and precision matrix are constant between change points. Observation Y_t are independent conditionally on the model parameters over time, while are iid between change points. Moreover, the precision matrices $\{\Omega_{c_i}, i = 1, \ldots, \kappa\}$ are conditionally independent given the sequence $\{G_{c_i}, i = 1, \ldots, \kappa\}$.

with the convention $c_0 = 1$ and $c_{\kappa+1} = T + 1$. Moreover, $K_{T\ell}$ is the maximum number of change points compatible with the minimum-span constraint. The imposition of such constraint ensures likelihood identifiability between change points, leading to more stable computations and robust inference. In the simulation study and in the application, we set ℓ equal to $p + 2$.

3 Bayesian inference via sequential Monte Carlo

The dynamic GGM proposed herein is a hierarchical model, with the first layer represented by the change points, $c_{1:\kappa}$, the second by graphs and precision matrices (G_{c_j}, Ω_{c_j}), and the third by the observations. Markov chain Monte-Carlo (MCMC) methods developed directly on such space of unobserved variables would face major challenges. We ignore the precision matrices in this discussion as they are later integrated out. Gibbs-type approaches would involve reversible-jump MCMC (Green, 1995), thus requiring the design of a joint update on the "model" space (as determined by the change-points) and model "parameters" of varying dimension (corresponding to the graphs). This joint space is entangled, with very limited space for maneuvering, as updates on the graph space would be heavily constrained by the strong prior Markovian dependencies amongst graphs.

Instead, we perform computationally effective posterior inference for the dynamic GGM through a tailored sequential Monte-Carlo (SMC) algorithm. The proposed Particle MCMC (PMCMC) method is quite appropriate for exploiting the hidden Markov model structure conditionally on the change points, and naturally disentangles the updates on the change points and the latent Markovian signal. See, e.g., Karagiannis and Andrieu (2013) and Persing et al. (2015) for related ideas. The proposed PMCMC algorithm is better understood as comprised of an 'outer' cycle and an 'inner' cycle. In the former, the change points are updated via a reversible jump Metropolis-Hasting (M-H) algorithm. In the latter, a particle filter, of enhanced performance due to a combination of adaptive tempering, dynamic resampling, and mutation steps, is employed to sample the sequence of graphs and compute the acceptance probability of the outer algorithm. In particular, for each M-H step, the inner component provides an unbiased estimate of the conditional likelihood given the proposed change point sequence and the corresponding proposed particles of graphs sequences. Adaptive tempering and resampling steps are used to improve the robustness of

such estimate, while the mutation step is used both to bring particles closer to the modal region of the likelihood and to avoid depletion of the number of unique particles which can otherwise be a consequence of successive resampling and tempering.

The overall PMCMC algorithm is well-understood as an "exact-approximate" one, in the sense that it targets the correct posterior on the space of graphs and change points, thanks to the unbiasedness (and positivity) of the estimator provided by the inner particle filter.

3.1   Outer component

The key component in the development of the M-H step is the choice of proposal distribution, $q(c'_{1:\kappa'} \mid c_{1:\kappa})$, where $c_{1:\kappa}$ and $c'_{1:\kappa'}$ are the current and proposed collections of ordered change points, respectively. Starting from $c_{1:\kappa}$, one of four alternative events (namely a birth, a death, a global move, or a local move) generates the proposed new value. With probabilities equal to $P(B \mid c_{1:\kappa})$, $P(D \mid c_{1:\kappa})$, $P(M_{glob} \mid c_{1:\kappa})$ and $P(M_{loc} \mid c_{1:\kappa})$, one of the following four events takes place, respectively: a new change point is added to the current set (birth); a change point is removed from the current set (death); one of the existing change points is moved to another position (global move); one of the existing change points is moved to another position in-between its neighbours (local move). When a new change point, c^*, is created, c^* is chosen uniformly over the set $\mathsf{B}(c_{1:\kappa}, \ell) \subset \{2, \ldots, T\}$ of allowed positions (i.e., satisfying the minimum-span constraint), of size $|\mathsf{B}(c_{1:\kappa}, \ell)| =: n(c_{1:\kappa}, \ell) \geq 0$. Thus, in the birth scenario

$$q(c'_{1:\kappa'} \mid c_{1:\kappa}) = \frac{P(B \mid c_{1:\kappa})}{n(c_{1:\kappa}, \ell)}, \qquad \kappa' = \kappa + 1, \quad c'_{1:\kappa'} = c_{1:\kappa} \cup c^*, \quad c^* \in \mathsf{B}(c_{1:\kappa}, \ell) \quad (7)$$

When a change point, c', is removed, the change point is chosen uniformly among the current change points, i.e., in the death scenario

$$q(c'_{1:\kappa'} \mid c_{1:\kappa}) = \frac{P(D \mid c_{1:\kappa})}{\kappa}, \qquad \kappa' = \kappa - 1, \quad c'_{1:\kappa'} = c_{1:\kappa} \setminus c', \quad c' \in c_{1:\kappa} \quad (8)$$

To improve mixing and posterior exploration, we introduce also two move steps. When a change point is moved, firstly, a change point $c' \in c_{1:\kappa}$ is selected uniformly among the current change points and removed. Then, if the step is a global move a new change point is selected uniformly in $\mathsf{B}(c_{1:\kappa} \setminus c', \ell)$. If instead the step is a local move a new change point c^* is selected with probability proportional to

$$\exp\{-\lambda \, |c^* - c'|\} \mathbb{1}_{\{c^* \in [\bar{c}_l, \bar{c}_r]\}} \tag{9}$$

for algorithmic parameter $\lambda > 0$, with

$$\bar{c}_l = c_l + \ell, \quad \bar{c}_r = c_r - \ell$$

where c_l and c_r denote the left-side and right-side neighbours of c' in $c_{0:\kappa+1}$. Thus, the proposal kernel for a move step is

$$q(c'_{1:\kappa'} \mid c_{1:\kappa}) = \frac{P(M_{glob} \mid c_{1:\kappa})}{\kappa \, n(c_{1:\kappa} \setminus c', \ell)} + \frac{P(M_{loc} \mid c_{1:\kappa}) \, e^{-\lambda \, |c^* - c'|}}{\kappa \sum_{\chi = \bar{c}_l}^{\bar{c}_r} e^{-\lambda \, |c^* - \chi|}} \mathbb{1}_{\{c^* \in [\bar{c}_l, \bar{c}_r]\}}$$

$$\kappa' = \kappa, \quad c'_{1:\kappa'} = c_{1:\kappa} \setminus c' \cup c^*, \quad c' \in c_{1:\kappa}, \quad c^* \in \mathsf{B}(c_{1:\kappa} \setminus c', \ell) \tag{10}$$

Algorithm 1 Outer algorithm - Reversible jump M-H

    Input: change point sequence $c_{1:\kappa}$; $\mathsf{B}(c_{1:\kappa}, \ell)$; likelihood value $P(Y_{1:T}|c_{1:\kappa})$; prior value $P(c_{1:\kappa})$.

    Output: new change point sequence $\tilde{c}_{1:\tilde{\kappa}}$; $\mathsf{B}(\tilde{c}_{1:\tilde{\kappa}}, \ell)$; likelihood value $P(Y_{1:T}|\tilde{c}_{1:\tilde{\kappa}})$; prior value $P(\tilde{c}_{1:\tilde{\kappa}})$.

Sample event E from $\{B, D, M_{glob}, M_{loc}\}$ according to (11);

if $E = B$ **then**

    Sample a new change point uniformly from $\mathsf{B}(c_{1:\kappa}, \ell)$ and propose $c'_{1:\kappa'} = c_{1:\kappa} \cup c'$;

    Compute $q(c'_{1:\kappa'} \mid c_{1:\kappa})$ according to (7);

else

    Sample uniformly and remove a change point c' from $c_{1:\kappa}$;

    if $E = M_{glob}$ or M_{loc} **then**

        if $E = M_{glob}$ **then**

            Sample a new change point c^* uniformly from $\mathsf{B}(c_{1:\kappa} \setminus c', \ell)$;

        else

            Sample a new change point c^* from the interval $[\bar{c}_l, \bar{c}_r]$ according to (9);

        Propose $c'_{1:\kappa'} = c_{1:\kappa} \setminus c' \cup c^*$ and compute $q(c'_{1:\kappa'} \mid c_{1:\kappa})$ according to (10);

    else

        Propose $c'_{1:\kappa'} = c_{1:\kappa} \setminus c'$ and compute $q(c'_{1:\kappa'} \mid c_{1:\kappa})$ according to (8);

    Determine $\mathsf{B}(c'_{1:\kappa'}, \ell)$ and compute $q(c_{1:\kappa} \mid c'_{1:\kappa'})$ according to (7), (8), or (10), respectively;

Compute prior $P(c'_{1:\kappa'})$ and likelihood $P(Y_{1:T} \mid c'_{1:\kappa'})$ for proposed configuration;

Sample u from a Uniform$(0, 1)$;

if $u \leq \frac{P(Y_{1:T}|c'_{1:\kappa'})P(c'_{1:\kappa'})\,q(c_{1:\kappa}|c'_{1:\kappa'})}{P(Y_{1:T}|c_{1:\kappa})P(c_{1:\kappa})\,q(c_{1:\kappa}|c'_{1:\kappa'})}$ **then**

    Return $c'_{1:\kappa'}$, $\mathsf{B}(c'_{1:\kappa'}, \ell)$, $P(Y_{1:T}|c'_{1:\kappa'})$, $P(c'_{1:\kappa'})$.

else

    Return $c_{1:\kappa}$, $\mathsf{B}(c_{1:\kappa}, \ell)$, $P(Y_{1:T}|c_{1:\kappa})$, $P(c_{1:\kappa})$.

Note that we prefer to write the joint kernel for both the global and local move, as the same configuration $c'_{1:\kappa}$ might be reached by both type of moves. This needs to be accounted for when computing the M-H acceptance probability to ensure detailed balance.

Finally, we choose the event probabilities as

$$P(B \mid c_{1:\kappa}) = \begin{cases} 1, & \text{if } \kappa = 0 \\ 0, & \text{if } n(c_{1:\kappa}, \ell) = 0 \\ q_B, & \text{otherwise} \end{cases} \qquad P(D \mid c_{1:\kappa}) = \begin{cases} 0, & \text{if } \kappa = 0 \\ q'_D, & \text{if } n(c_{1:\kappa}, \ell) = 0 \\ q_D, & \text{otherwise} \end{cases} \tag{11}$$

$$\text{and} \quad P(M_{glob} \mid c_{1:\kappa}) = P(M_{loc} \mid c_{1:\kappa}) = [1 - P(B \mid c_{1:\kappa}) - P(D \mid c_{1:\kappa})]/2$$

The Metropolis-Hastings acceptance probability is equal to

$$1 \wedge \frac{P(Y_{1:T}|c'_{1:\kappa'})P(c'_{1:\kappa'})\,q(c_{1:\kappa}|c'_{1:\kappa'})}{P(Y_{1:T}|c_{1:\kappa})P(c_{1:\kappa})\,q(c'_{1:\kappa'}|c_{1:\kappa})}$$

where $P(Y_{1:T} \mid c_{1:\kappa})$ is the marginal likelihood of the data given the change points, i.e.,

$$P(Y_{1:T} \mid c_{1:\kappa}) = \int P(Y_{1:T} \mid G_{1:T})P(G_{1:T}|c_{1:\kappa})dG_{1:T}$$

with

$$P(Y_{1:T} \mid G_{1:T}) = \int P(Y_{1:T} \mid G_{1:T}, \Omega_{1:T})P(\Omega_{1:T} \mid G_{1:T})d\Omega_{1:T}$$

Since $P(Y_{1:T} \mid c_{1:\kappa})$ is not available in closed form, it needs to be estimated. In the next section we describe the SMC algorithm to approximate the marginal likelihood.

3.2   Inner Component: Particle Filter

As already mentioned, the inner component of the algorithm is used to compute $P(Y \mid c_{1:\kappa})$, i.e., the likelihood values required in the acceptance probability of the outer M-H, and to provide proposed samples of the graphs to be accepted or rejected by the outer algorithm.

To compute the marginal likelihood given the change point sequence, a standard bootstrap particle filter with multinomial resampling carried out at each change point, samples $N \geq 1$ particles $\{G_{c_j}^{(n)}\}_{n=1}^N$, for $0 \leq j \leq \kappa$, from the joint distribution

$$\prod_{n=1}^N P(G_{c_0}^{(n)}) \times \prod_{j=1}^\kappa \left\{ \prod_{n=1}^N \left(\sum_{l=1}^N \frac{w_{j-1}^{(l)}}{\sum_{m=1}^N w_{j-1}^{(m)}} P(G_{c_j}^{(n)} \mid G_{c_{j-1}}^{(l)}) \right) \right\}$$

where the unnormalised weights are defined as

$$w_j^{(n)} = P(Y_{c_j:c_{j+1}-1} \mid G_{c_j}^{(n)}), \qquad 1 \leq n \leq N, \quad 0 \leq j \leq \kappa$$

The unbiased estimate $\hat{P}(Y \mid c_{1:\kappa})$ of $P(Y \mid c_{1:\kappa})$ could then be obtained as

$$\hat{P}(Y_{1:T} \mid c_{1:\kappa}) = \prod_{j=0}^\kappa \left(\frac{1}{N} \sum_{n=1}^N w_j^{(n)} \right) \tag{12}$$

However, it is often the case that further algorithmic advances must complement the standard particle filter to control the variance of the estimate (12). It is well-understood that such variability is critically linked to the performance of the overall PMCMC algorithm. See e.g. Pitt et al. (2012); Doucet et al. (2015); Sherlock et al. (2015) where, in various model settings, standard deviations centred around 1 are proposed for the estimate of the logarithm of the normalising constant, with exponential decay in performance for PMCMC reported when the standard deviation exceeds a (not too high) threshold.

A standard approach to reduce standard deviation for given number of particles, is via the application of tempering, i.e. introduction of a sequence of temperatures together with corresponding mutation steps. Such approach has been shown, in cases, to reduce the required number of particles for a target error from exponential to quadratic in the number T of log-likelihood terms, see e.g. Beskos et al. (2014); Ruzayqat et al. (2021). Application of tempering will indeed be critical for the class of models we are considering in this work, as shown in Section 3.2.4. The temperatures are determined on-the-fly, thus avoiding the introduction of additional tuning parameters in the algorithm. The complete algorithm can be understood as a particle filter applied on a Feynman-Kac model (Del Moral, 2004) that we extend to include additional Markov iterations and potentials. The overall approach is summarised in Algorithm 2.

Algorithm 2 Inner Algorithm - Particle Filter with Tempering

Input: Data $Y_{1:T}$; change points $c_{1:\kappa}$; number of particles N; ESS threshold ϵ_0; number S_j of temperature for graph j, $\phi_{0,j} \equiv 0$, $\phi_{0,S_j+1} \equiv 1$, $0 \leq j \leq \kappa$; number of mutation steps $M \geq 1$; temperatures $\{\phi_{1,j}, \ldots, \phi_{S_j,j}\}$; M-H kernel $\bar{P}_{j,s}(G_{c_j}|G_{c_j}, G_{c_j-1})$, $0 \leq j \leq \kappa$, $1 \leq s \leq S_j + 1$; $\bar{P}_{j,s}^M$ denotes M iterations of such a kernel; hyper-parameters ω, z, δ, D.

Output: Unbiased estimate $\hat{P}(Y_{1:T}|G_{1:T})$.

     (Actions over n are understood to be repeated for $1 \leq n \leq N$.)

Set $\hat{P} = 1$;

for j in $0 : \kappa$ **do**

    if $j = 0$ **then**

         Sample $G_{c_0}^{0,n} \overset{iid}{\sim} P(G_{c_0})$ and set $w_0^{0,n} = 1$;

    else

         Initialise particles $G_{c_j}^{0,n} \overset{ind.}{\sim} P(G_{c_j}|G_{c_{j-1}}^{S_{j-1},n})$;

         Initialise weights $w_j^{0,n} = w_{j-1}^{S_{j-1}+1,n}$;

    for s in $1 : S_j + 1$ **do**

         Calculate weights $w_j^{s,n} = w_j^{s-1,n} \cdot P^{\phi_{j,s}-\phi_{j,s-1}}(Y_{c_j:c_{j+1}-1}|G_{c_j}^{s-1,n})$;

        if $ESS(w_j^{s,1:N}) < \epsilon$ **then**

             $\hat{P} \leftarrow \hat{P} \cdot \frac{1}{N} \sum_{n=1}^N w_j^{s,n}$;

             Resample $\{G_{c_j}^{s-1,n}, G_{c_j-1}^{S_{j-1},n}\}$ according to the weights $\{w_j^{s,n}\}$;

             Mutate particles, i.e., sample $G_{c_j}^{s,n} \overset{ind.}{\sim} \bar{P}_{j,s}^M(G_{c_j}|G_{c_j}^{s-1,n}, G_{c_{j-1}}^{S_{j-1},n})$;

             Set $w_j^{s,n} = 1$;

        else

             set $G_{c_j}^{s,n} = G_{c_j}^{s-1,n}$;

Return $\hat{P}(Y_{1:T|G_{1:T}}) = \hat{P}$.

3.2.1   Preliminary run − Determination of temperatures

Within Algorithm 2, the sequence of temperatures is treated as given. In practice, at each iteration of the outer algorithm, the temperatures are determined by a separate preliminary and independent execution of the particle filter, that identifies and stores the temperatures, that are later used within Algorithm 2. Similar ideas have been used in the SMC literature, see e.g. Jasra et al. (2011). That is, we first determine the temperatures according to a target effective sample size (ESS), and, then, apply the particle filter in Algorithm 2, with the obtained temperatures, to produce a robust unbiased estimator of the likelihood needed to compute the acceptance probability in Algorithm 1.

     We describe here how to compute the temperatures $\{\phi_{1,j}, \ldots, \phi_{S_j,j}\}$ used in Algorithm 2, for $S_j \geq 0$, where $\phi_{0,j} \equiv 0$, $\phi_{0,S_j+1} \equiv 1$, and $0 \leq j \leq \kappa$. Note that S_j, the number of temperatures, can vary across graphs at different change points, i.e., it depends on j.

Algorithm 3 Inner Algorithm - Temperature Tuning (Preliminary particle filter)

Input: Data $Y_{1:T}$; change points $c_{1:k}$; number of particles N;
hyper-parameters w, z, δ, D; ESS threshold ϵ_0; mutation steps $M \geq 1$;
M-H kernel $\bar{P}_{j,s}(G_{c_j}|G_{c_j}, G_{c_j-1})$ that preserves the law
$P^{\phi_{j,s}}(Y_{c_j:c_{j+1}-1}|G_{c_j}) \cdot P(G_{c_j}|G_{c_j-1})$,
$\bar{P}_{j,s}^M$ denotes M iterations of such a kernel.

Output: Temperatures $\{\phi_{1,j}, \ldots, \phi_{S_j,j}\}$, $S_j \geq 0$, $\phi_{0,j} \equiv 0$, $\phi_{0,S_j+1} \equiv 1$, $0 \leq j \leq \kappa$.

(Actions over n are understood to be repeated for $1 \leq n \leq N$.)

for j in $0 : \kappa$ **do**

    if $j = 0$ **then**

        Sample $G_{c_0}^{0,n} \overset{iid}{\sim} P(G_{c_0})$ and set $w_0^{0,n} = 1$;

    else

        Initialise particles $G_{c_j}^{0,n} \overset{ind.}{\sim} P(G_{c_j} \mid G_{c_{j-1}}^{S_{j-1},n})$;

        Initialise weights $w_j^{0,n} = w_{j-1}^{S_{j-1}+1,n}$;

    $s \leftarrow 1$; $\phi_{0,j} \leftarrow 0$; $\phi_{s,j} \leftarrow 1$;

    while $\phi_{s,j} \neq \emptyset$ **do**

        Find $\phi_{s,j} \in (\phi_{s-1,j}, 1]$ so that $ESS(\phi_{s,j}) \geq \epsilon_0 N$;

        if $\phi_{s,j} = \emptyset$ **then**

            $S_j = s - 1$;

        else

            Set $w_j^{s,n} = P^{\phi_{s,j}-\phi_{s-1,j}}(Y_{c_j:c_{j+1}-1} \mid G_{c_j}^{s-1,n})$;

            Resample $\{G_{c_j:c_{j+1}-1}^{s-1,n}, G_{c_j-1}^{S_{j-1},n}\}$ according to the weights $\{w_j^{s,n}\}$;

            Set $w_j^{s,n} = 1$;

            Mutate particles, i.e., sample $G_{c_j}^{s,n} \overset{ind.}{\sim} \bar{P}_{j,s}^M(G_{c_j} \mid G_{c_j}^{s-1,n}, G_{c_j-1}^{S_{j-1},n})$;
            $s \leftarrow s+1$

    Set $w_j^{S_j+1,n} = P^{1-\phi_{s-1,j}}(Y_{c_j:c_{j+1}-1} \mid G_{c_j}^{s-1,n})$;

  Return $\{\phi_{1,j}, \ldots, \phi_{S_j,j}\}$, $S_j \geq 0$, $0 \leq j \leq \kappa$.

Within this subsection, particles and weights $G_{c_j}^{s,n}$, $w_j^{s,n}$ refer to such a preliminary execution of the particle filter. The temperatures are selected on-the-fly, based on the target ESS, denoted by ESS_0, with $ESS_0 = \epsilon_0 N$, $\epsilon_0 \in (0,1)$. Consider the current collection of particles and weights, $G_{c_j}^{s-1,n}$ and $w_j^{s-1,n}$, generated while filtering data points $Y_{c_j:c_{j+1}-1}$, and the corresponding likelihood factor $P^{\phi_{s-1,j}}(Y_{c_j:c_{j+1}-1} \mid G_{c_j})$ up to the present step. Then, the next temperature ϕ_s is determined so that the ESS equals the target, i.e., $ESS(\phi_s) = ESS_0$. More precisely, define the next set of weights as function of the next temperature

$$w_j^{s,n}(\phi) = P^{\phi-\phi_{s-1,j}}(Y_{c_j:c_{j+1}-1} \mid G_{c_j}^{s-1,n})$$

and consider

$$ESS(\phi) := \frac{\left(\sum_{n=1}^{N} w_{s,j}^n(\phi)\right)^2}{\sum_{n=1}^{N} \left(w_{s,j}^n(\phi)\right)^2} = \epsilon_0 N$$

whose solution – assuming it exists within $(\phi_{s-1,j}, 1]$ – provides the next temperature $\phi_{s,j}$. The solution is obtained with a simple fast bisection method. If $ESS_0(1) \geq \epsilon_0 N$, we simply select $\phi = 1$. With this procedure we obtain all temperatures related to data $Y_{c_j:c_{j+1}-1}$ and we can then proceed to the next filtering step. We set ϵ_0 to $1/2$ to obtain a minimum ESS of $N/2$, which is a common choice (see, e.g., Chopin et al., 2020, p.133). See Algorithm 3 for a detailed description.

3.2.2 Determination of mutation kernel $\bar{P}$

The mutation kernel $\bar{P}$ is used within the algorithm to jitter particles, and move them towards the centre of the support of each filtering distribution under consideration during a full application of the particle filter. The addition of mutation steps has been shown to be, in many cases, critical, both in theoretical and experimental works, see e.g. Beskos et al. (2014); Ruzayqat et al. (2021) and Llopis et al. (2018); van Leeuwen et al. (2021), respectively. In Section 3.2.4, we illustrate such impact for the specific model at hand through a simulation study.

For the overall algorithm to ensure a correct particle filter on an extended space, the user-specified mutation kernel $\bar{P}_{j,s}(G_{c_j} \mid G_{c_j}, G_{c_j})$ must have invariant distribution

$$G_{c_j} \mapsto P^{\phi_{j,s}}(Y_{c_j:c_{j+1}-1} \mid G_{c_j}) \times P(G_{c_j} \mid G_{c_j-1})$$

where we use the convention that $G_{c_0-1} = \emptyset$ in which case the rightmost term becomes the prior defined by (4). This is readily achieved via a M-H step. That is, for each current segment $c_j : c_{j+1} - 1$, temperature $\phi_{j,s}$, graph G_{c_j}, with adjacency matrix $A_{c_j} = A_{c_j}[h,k]$, and given G_{c_j-1}, we define a proposed graph G'_{c_j}, with adjacency matrix $A'_{c_j} = A'_{c_j}[h,k]$, using the symmetric transition

$$A_{c_j}[h,k]' \mid A_{c_j}[h,k] \overset{ind}{\sim} \mid A_{c_j}[h,k] - \text{Bernoulli}\left(\frac{2s_0}{p-1}\right) \mid \tag{13}$$

for algorithmic tuning parameter $s_0 \in (0, (p-1)/2)$. Thus, under the proposal in (13), the expected number of flips in the edges in $s_0 \cdot p$. The acceptance probability for the mutation step is

$$1 \wedge \frac{P^{\phi_{j,s}}(Y_{c_j:c_{j+1}-1} \mid G'_{c_j}) \times P(G'_{c_j} \mid G_{c_j-1})}{P^{\phi_{j,s}}(Y_{c_j:c_{j+1}-1} \mid G_{c_j}) \times P(G_{c_j} \mid G_{c_j-1})}$$

3.2.3 Likelihood given the graph structure

An important quantity required within the particle filter is the marginal likelihood

$$P(Y_{c_j:c_{j+1}-1} \mid G_{c_j}) = \int_{M^+(G_{c_j})} P(Y_{c_j:c_{j+1}-1} \mid \Omega_{c_j}) P(\Omega_{c_j} \mid G_{c_j}) \, d\Omega_{c_j}$$

Since the G-Wishart law is conjugate, we can integrate out the precision matrices $\Omega_{1:T}$. That is, we have (Atay-Kayis and Massam, 2005)

$$P(Y_{c_j:c_{j+1}-1} \mid G_{c_j}) = \frac{1}{(2\pi)^{(c_{j+1}-c_j)p/2}} \frac{I_{G_{c_j}}(\delta + (c_{j+1} - c_j), D + H_j)}{I_{G_{c_j}}(\delta, D)}$$

where, for $j = 0, \ldots, \kappa$,

$$H_j = \sum_{i=c_j}^{c_{j+1}-1} Y_i Y_i^\top$$

The normalising constant of the G-Wishart prior can be factorised (Roverato, 2002; Uhler et al., 2018), i.e., for a given graph G,

$$I_G(\delta, D) = \frac{\prod_{m=1}^r I_{G_{P_m}}(\delta, D_{P_m})}{\prod_{m=2}^r I_{G_{S_m}}(\delta, D_{L_m})} \tag{14}$$

where $P_1; L_2, P_2; \ldots; P_r, L_r$, is a perfect sequence of prime components and corresponding minimal separators of G (see, e.g., Chapter 2 of Lauritzen, 1996, for details) and D_{P_m} is the submatrix of D corresponding to the rows and columns in P_m. In the case of a *decomposable* graph G, all prime components are complete graphs. For complete graphs the G-Wishart distribution coincides with the Hyper-Wishart distribution (Dawid and Lauritzen, 1993), for which an analytical expression for the normalising constant is available:

$$I_{G_{P_m}}(\delta, D_{P_m}) = \frac{2^{(\delta+p_m-1)p_m/2}\Gamma_{p_m}(\frac{\delta+p_m-1}{2})}{|D_{P_m}|^{(\delta+p_m-1)}}$$

Here $\Gamma_d(\cdot)$ is the multivariate Gamma function of dimension d and p_m is the dimension of D_{P_m}. Note that, by construction, the minimal separators are complete sub-graphs of G, thus the terms in the denominator in (14) are analytically computable. For a general, *non-decomposable* graph G, Roverato (2002); Dellaportas et al. (2003); Atay-Kayis and Massam (2005); Carvalho et al. (2007) propose Monte-Carlo methods for the approximation of $I_G(\delta, D)$. Herein, to compute the normalizing constant, we employ the method of Atay-Kayis and Massam (2005) implemented in the function gnorm of the R package BDgraph. When dealing with large number of nodes, the implementation of more sophisticated algorithms, such as the exchange algorithm by Murray et al. (2006), is advisable (see, e.g., Cheng and Lenkoski, 2012; van den Boom et al., 2022, for an application to GGMs).

3.2.4 Evaluation of SMC approximation

The inner SMC algorithm provides unbiased estimates of the marginal likelihood conditionally on the change points. The variability of such estimates depends on the number of particles N, and the effect of the tempering and mutation steps, with the number of the latter, M, specified by the user. Thus, a trade-off is posed between accuracy of estimates and computational time.

To assess the effect of the number of particles and the mutation step, and, in general, obtain insights into the performance of the SMC component, we perform a series of simulation studies. We simulate data for $p = 10$ nodes and $T = 200$ observation instances. We then fix the change point sequence to its known true value, and carry out 30 independent executions of the SMC algorithm, for each different combination of $N \in \{200, 500, 750\}$ and $M \in \{0, 5, 10, 20\}$. Recall that the mutation steps are performed only when the ESS falls below the threshold $\epsilon_0 N$, where ϵ_0 is here fixed to be $N/2$. We consider two data generating mechanisms. The first (Scenario A) has no change points and the p variates are mutually independent (see Figure B.1 (a) in the Supplement for the corresponding graph). In Scenario B we set a change point at $t = 70$, and the two graphs (before and after the change point) encode some non-trivial dependence. The full graph structure of Scenario B is described later in Section 4.2 and displayed in Figure B.2 of the Supplement. Figure A.1
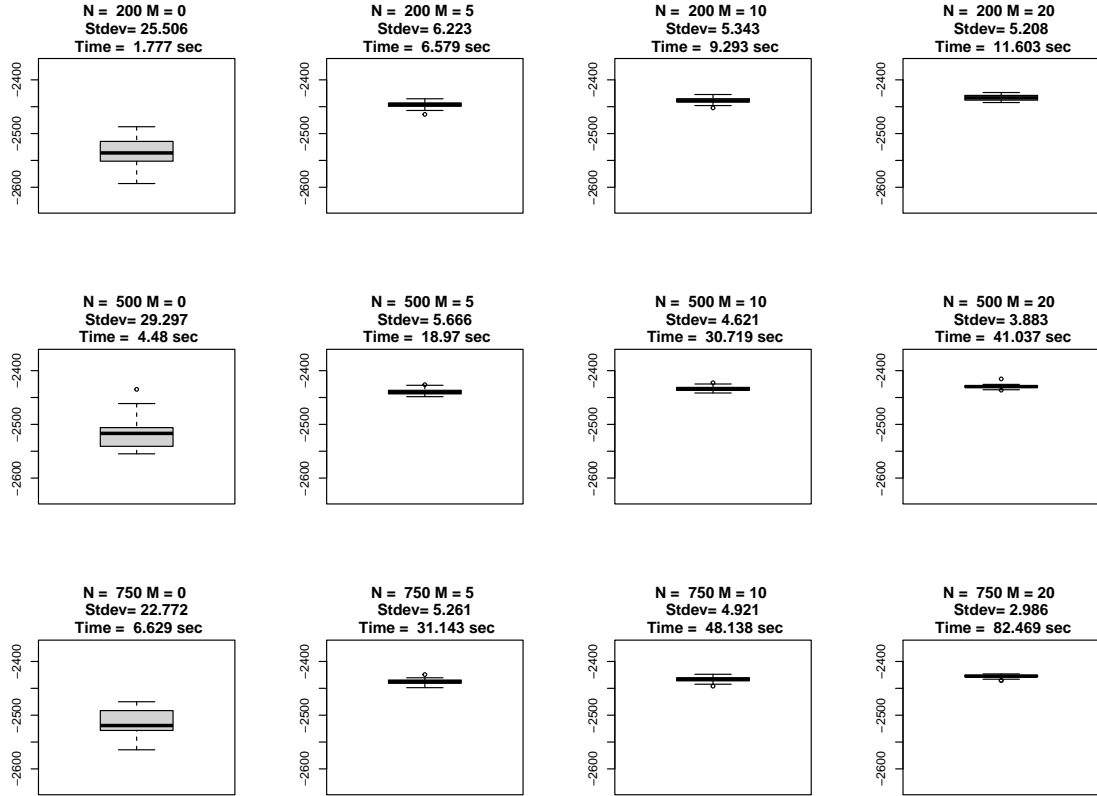
Figure 4: Log-likelihood estimates for Scenario B obtained with the particle filter by fixing the change points to the truth. Distinct box-plots correspond to different numbers of particles N and/or mutations steps M. For each pair (N, M), we run the algorithm 30 times and obtain the log-likelihood estimates. Each box-plot shows the distribution of such estimates. The variability of the estimates is rather limited for all pairs (N, M), provided that $M \geq 5$.

of the Supplement and Figure 4 show the box-plots of the estimates of the log-likelihood, the standard deviation, and the running time of the inner SMC algorithm coded in R (and run with an Intel Xeon W-1250 processor), under scenarios A and B, respectively.

Under Scenario A, the variability of the estimates is limited, as expected, for all pairs (N, M) since data are simulated under the assumption of independence with no change points. However, in real scenarios, as the analysis of financial markets, this is highly unlikely to be the case and the computational machinery here developed is essential. Figure 4 shows similar box-plots obtained under the more realistic and challenging Scenario B. Here, we obtain higher variability with values of standard deviation ranging from 2.986 to 29.297. These results highlight the importance of the tempering and mutation steps for the overall algorithmic performance (even accounting for the increased computational time), and in particular for recovering the complex dependence structures. Their role is essential in reducing variability of the estimates of the normalising constant. Lastly, notice that computational time is increased compared to Scenario A as a consequence of: (i) the presence of a change point; (ii) the computational complexity of the Monte Carlo iterations used to compute $I_G(\delta, D)$, which increases when particles concentrate on less sparse graphs; (iii) and the increased number of times the ESS threshold is reached.

15

(a) False positive rate
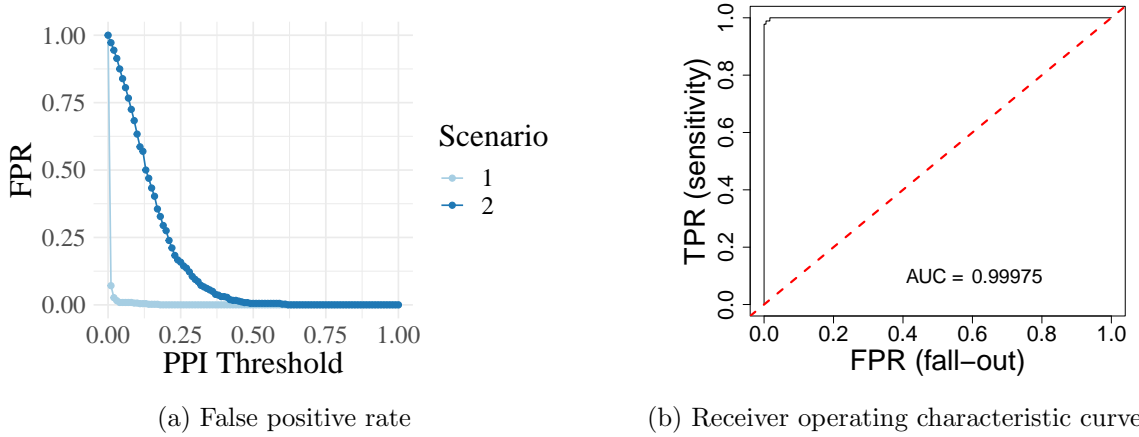
(b) Receiver operating characteristic curve

Figure 5: Panel (a): FPR versus PPI threshold, for simulation Scenarios 1 and 2, computed across 10 replicates. Panel (b): ROC curve in Scenario 2, computed across 10 replicates.

4   Model performance on simulated data

We perform a series of simulation studies and check the model performance under different scenarios for the data generating mechanism. For each scenario, we consider 10 simulation replicates, $p = 10$ variables and $T = 200$ observation times. To carry out posterior inference, we, firstly, run the algorithm to estimate the change point sequence, using $N = 200$ particles, $M = 10$ mutation steps, and performing 5,000 iterations of the outer component, of which the first 2,000 are discarded as burn-in. When the true sequence of change points is the null set (see Section 4.1), we initialise the MCMC chain at $(c_1 = 51, c_2 = 101, c_3 = 151)$, whereas when change points exist (Section 4.2), we initialise the chain to the state of zero change points. The displayed MCMC results are obtained after applying thinning every 10 samples. Secondly, we re-run only the inner SMC with $N = 1,000$ particles and $M = 20$ mutation steps to obtain the graph estimates conditionally on the maximum-a-posteriori (MAP) estimate of the change point sequence obtained in the first step.

4.1   Simulation results with no change points

The first two scenarios do not include change points among the $T = 200$ instances. In Scenario 1 all variables are independent, while in Scenario 2 we assume a non-trivial conditional independence structure represented by a graph with nine edges, shown in Figure B.1 (b) of the Supplement.

   The inference results show extremely good performance of the model in terms of both identification of change points and recovery of the dependence structure. In both scenarios the posterior on the change points concentrates on the true state of no change points, with posterior MAP probability not falling below 0.98 in all simulation replicates. More detailed results about the posterior on change points can be found in Table B.1 of the Supplement. Figure 5a shows the false positive rate (FPR) of edge detection as function of the threshold used for the marginal posterior probability of edge inclusion (PPI, i.e., the posterior probability associated to each edge in the graph, $P(A[h, k] = 1 \mid \text{data})$). FPRs are very low for almost any PPI threshold (indicatively, for a threshold of 0.5, FPR is 0 and 0.006 for Scenarios 1 and 2, respectively). Figure 5b reports the receiver operating characteristic (ROC) curve for Scenario 2, with the area under the curve (AUC) approximately 1.

16

	True change point position		MAP estimate	MAP prob.	75% C.I.
Scenario 3	(70)	Replica 1	(69)	0.317	[69, 73]
		Replica 2	(70)	0.313	[69, 73]
		Replica 3	(70)	0.420	[69, 73]
		Replica 4	(72)	0.183	[68, 74]
		Replica 5	(70)	0.413	[70, 76]
		Replica 6	(73)	0.233	[70, 79]
		Replica 7	(70)	0.230	[67, 72]
		Replica 8	(69)	0.137	[64, 77]
		Replica 9	(72)	0.117	[69, 80]
		Replica 10	(70)	0.300	[70, 74]

Table 1: Scenario 3: Posterior summaries for change points. MAP estimates, MAP probabilities (for the posterior over all configurations of change points), and 75% credible intervals of the position of the change point (conditionally on having one change point).

4.2 Simulation results with changes points

The third simulation scenario (Scenario 3) is obtained by setting one change point at $t = 70$ and generating two precision matrices as in Peterson et al. (2015) and Molinari et al. (2022b). In particular, we first define the Ω_{c_0} and then derive Ω_{c_1} as perturbation of Ω_{c_0}. Ω_{c_0} is obtained by setting diagonal elements equal to 1, the first off-diagonal elements to 0.5, i.e., $\Omega_{c_0}[h, h+1] = \Omega_{c_0}[h+1, h] = 0.5$, for $h = 1, \ldots, 9$, and the remaining elements to 0. To construct Ω_{c_1}, we randomly remove five edges among the active ones in G_{c_0} and set to 0 the corresponding entries in the precision matrix. Then we add five randomly selected edges drawn from the set of inactive edges in G_{c_0}. Finally, a precision entry equal to 0.2 is associated to the new edges. The obtained matrix is not necessary positive-definite, and, to this end, we compute the nearest positive-definite approximation through the R function nearPD (Higham, 2002), available in the R package Matrix. The resulting graphs are shown in Figure B.2 of the Supplement. The hyperparameter ω in (4) is determined using an approach inspired by empirical Bayes techniques, so that the a priori the expected number of edges for the graphs is equal to the number of edges detected by estimating one unique graph using all the time points. To this end, we estimate the graph using an adaptive lasso approach, which is a modification of the estimation procedure proposed by Meinshausen and Bühlmann (2006) inspired by the adaptive lasso of (Zou, 2006), as implemented in the R package GGMselect. For the hyperparameter z in (5), we opt for $z = 0.1$, so that a priori we expect only one edge to change at each change point, favouring graph similarity. Note that this choice also favours more parsimonious models in terms of change points. It also allows us to better understand model performance, as in our simulations we force 10 edges to change across the change point 10, an event to which our prior associates a probability lower than $4 \cdot 10^{-8}$.

In all replicates the posterior of the number of change points is concentrated around 1, with a posterior probability greater or equal to 0.97. Regarding the ability of the model to correctly recover the position of the change point, the performance of the model is summarised in Table 1. Though the space of sequences that satisfy the minimum duration constraint of $\ell = 12$ for $T = 200$ includes more than $4 \cdot 10^{12}$ sequences, in five of the ten replicates the MAP estimate (which minimises the 0-1 loss function) coincides with the true state $c^\star = 70$, and in all replicates the MAP is contained in the interval [69, 73].

(a) False positive rate
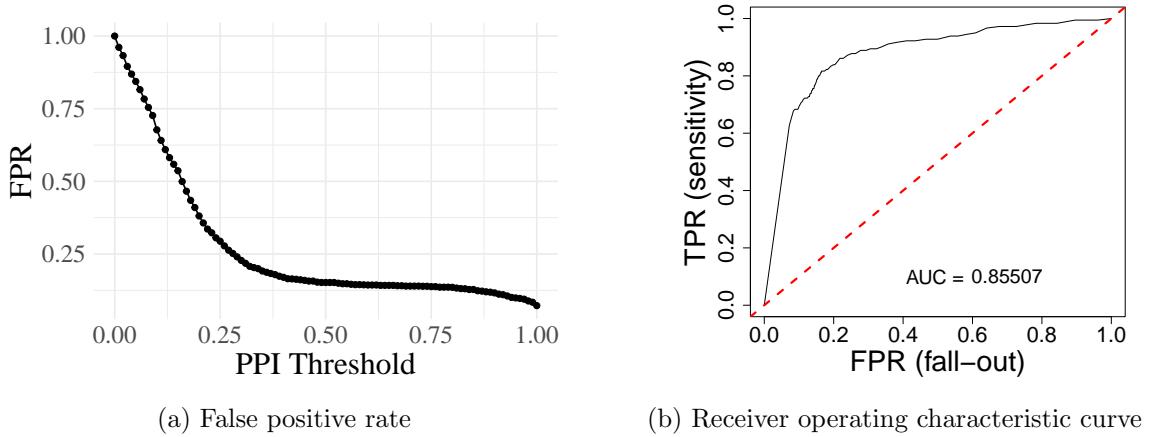

(b) Receiver operating characteristic curve

Figure 6: Panel (a): FPR versus PPI threshold, for Scenario 3, computed via 10 replicates. Panel (b): ROC curve in Scenario 3, computed via 10 replicates.

Moreover, Table 1 reports also the 75% credible intervals for the position of the change point, confirming that the posterior is concentrated around the true state.

Figure 6a shows the combined FPR of edge detection for the two graphs as function of the threshold used for the PPI, for Scenario 3. Again, FPRs are low for almost any PPI threshold, and, for the 0.5 threshold, the FPR is 0.152. Figure 6b displays the combined ROC curve, with an AUC approximately equal to 0.86.

5 Industry returns during COVID-19 pandemic

We apply the model to detect changes in the dependence structure of the nine industry portfolios' weekly returns described in Section 1. We consider weekly data over a time horizon of three years: from January 2019 to December 2021 so that $T = 157$. Logarithmic weekly returns are computed starting from the daily returns available from Kenneth R. French's Data Library at https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html, where the industry classification used to associate each stock to one of the nine portfolios is defined as follows. Stocks listed in the New York Stock Exchange (NYSE), the American Stock Exchange (AMEX), and National Association of Securities Dealers Automatic Quotation System (NASDAQ) are assigned to an industry at the end of June of year t based on their four-digit standard industrial classification (SIC) code at that time. Then, returns are computed from July of year t to June of year $t + 1$. The corresponding standardized time series are represented in Figure C.1 of the Supplement.

To estimate the dependence structure for the weekly returns, we firstly run the algorithm to estimate the change point sequence, using 200 particles, 10 mutation steps, performing 32,000 iterations of the outer component, of which the first 2,000 are discarded as burn-in, and thinning every ten iterations. Secondly, we re-run only the particle filter with 1,000 particles and 20 mutation steps to sample the graphs from their posterior distribution conditionally on the MAP estimate of the change point configuration. The algorithm is initialised assuming no change points and hyperparameters are set as described in Section 4.2.

Figure 7 shows, for each time point, the marginal posterior probability of being a change point and Table 2 reports the joint posterior distribution of the configurations of change points, which have been accepted by the algorithm. The posterior distribution on the number of change points associates a probability of 0.998 to two change points and the
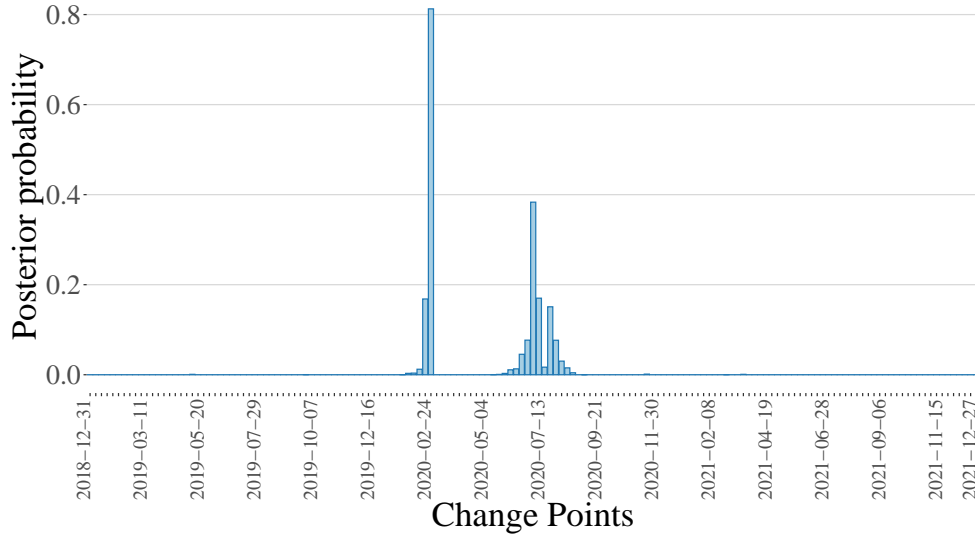
18

Figure 7: Marginal posterior probability of every time point to be a change point.

remaining mass 0.002 to three change points. Our analysis highlights a first structural change at $t = 61$, i.e., during the week starting on February 24, 2020, in correspondence of what appears to be the market's reaction to the first significant world-wide increase in Coronavirus confirmed cases and deaths outside China over the previous weekend. In particular, during the weekend February 21-23, 2020, Italy, the first and hardest-hit country in Europe in 2020, reported the first local cases of COVID-19 (see, e.g., Just and Echaust, 2020). During the week of February 24, 2020, the Dow Jones and S&P 500 fell by 11% and 12%, respectively, marking the biggest weekly declines to occur since the financial crisis of 2008. A second change point is detected at $t = 79$, i.e., during the last week of June 2020, interpretable as a subsequent and partial re-stabilization of the financial markets after the initial and most uncertain period of the pandemic.

Conditionally on the MAP change point configuration, estimates of the three graphs are provided in Figure 8, while the estimated variance and covariance matrices are displayed in Figure C.2.1 of the Supplement. The graphs are obtained based on the marginal PPI of the edges in order to control the corresponding Bayesian false discovery rate (Newton et al.,

change points	post prob	change points	post prob	change points	post prob	change points	post prob
(57 79)	0.0010	(59 82)	0.0073	(60 99)	0.0003	(61 81)	0.0137
(57 82)	0.0010	(59 84)	0.0017	(61 116)	0.0010	(61 82)	**0.1094**
(57 84)	0.0007	(60 78)	0.0020	(61 73)	0.0007	(61 83)	0.0650
(57 85)	0.0003	(60 79)	0.0073	(61 74)	0.0033	(61 84)	0.0150
(58 79)	0.0003	(60 80)	**0.1010**	(61 75)	0.0103	(61 85)	0.0127
(58 82)	0.0033	(60 81)	0.0033	(61 76)	0.0133	(61 86)	0.0043
(59 77)	0.0013	(60 82)	0.0297	(61 77)	0.0447	(61 99)	0.0010
(59 78)	0.0007	(60 83)	0.0100	(61 78)	**0.0737**	(19 61 83)	0.0010
(59 79)	0.0010	(60 84)	0.0127	**(61 79)**	**0.3735**	(39 61 83)	0.0003
(59 81)	0.0003	(60 85)	0.0020	(61 80)	**0.0694**	(61 80 113)	0.0003

Table 2: Posterior distribution of change point configurations. In **bold** we highlight probabilities greater than 0.05 and in **blue** the MAP estimate.
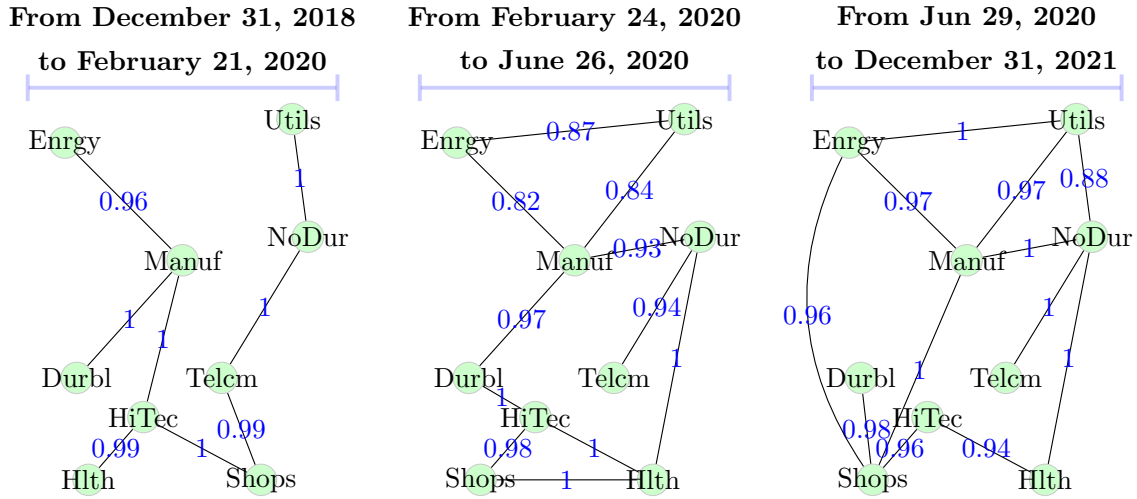
19

Figure 8: Posterior estimates of the graphs and PPI for the selected edges. Threshold of inclusion is set to achieve an expected posterior specificity of at least 95%.

2004). In particular, we set the threshold of inclusion based on the PPI to 0.8 in order to guarantee an expected rate of false detection not higher than 0.05, i.e., a specificity of at least 0.95 (for more details, see, Leday and Richardson, 2019; Williams, 2021). In Section C.1 of the Supplement we report the values of degree centrality, betweenness centrality (Freeman, 1977), local clustering and global clustering coefficients (Watts and Strogatz, 1998) for the estimated graphs, which give insights into the role of each node.

A clearly noticeable feature from Figure 8 is the increase in the number of edges from the first change point (8 edges) to the followings (11 and 12 edges, respectively), reflected also in the global measures of clustering of the graphs which varies from 0 in the pre COVID-19 period, to 0.43 during the first COVID-19 outbreak, to 0.24 to the post COVID-19 outbreak (see Table C.1.3 in the Supplement). Such increase in the connectivity of the graph is coherent with the hypothesis of the COVID-19 outbreak acting as a common non-measurable risk factor driving the returns in the market. As already noticed in the preliminary analysis summarised by Figure 2, the role of hub of the manufacturing portfolio (i.e., machinery, trucks, planes, chemicals, office furniture, paper) over the three years is confirmed. The corresponding degree centrality (i.e., number of vertices in the neighbourhood of the manufacturing portfolio) is the highest in all three graphs (see Table C.1.1 of the Supplement for more details). However, contrary to the conclusions of the initial exploratory analysis, consumer non-durable (i.e., food, tobacco, textiles, apparel, leather, toys) returns appear to play a less central role before the COVID-19 outbreak. Such evidence that was absent in the explanatory analysis is discovered mainly thanks to the automatic detection of the change points, that allow us to determine the most appropriate time window to estimate the graph and capture differences in structure. Moreover, we identify another hub in the Shops portfolio (i.e., wholesale, retail, and some services, as laundries and repair shops) in the time-interval after the last change point.

In terms of volatility, in all three periods the consumer durables (i.e., cars, TVs, furniture, household appliances) and energy (i.e, oil, gas, and coal extraction and products) are characterised by the highest volatility. Moreover, the analysis confirms that the three periods (pre COVID-19 outbreak, during first COVID-19 global outbreak, and after) coincide with small, high, and medium volatility markets, as already evident from the time-series plot (see Figure C.1 in the Supplement). More in details, we note that health (i.e., health-

care, medical equipment, and drugs) is the only industry in the market whose portfolio presents a similar value of volatility before and after the outbreak, while all other portfolios' returns are set to higher levels of variability as consequence of a long-run effect of market uncertainty. Similar conclusions can be drawn also for the pair-wise correlations (that can be easily computed from the values in Figure C.2.1 and are reported in Section C.2 of the Supplement), i.e. correlations are higher during COVID-19 outbreak.

6 Discussion and conclusion

In this work, we study the impact of the COVID-19 pandemic on the US stock market, with a specific focus on changes in dependence structure across stocks related to different industries. To do so, we consider weekly returns recorded for three years starting in January 2019. We identify two structural changes. The first change is in correspondence with the last week of February 2020, a date that for most countries coincided with the beginning of the pandemic. That same week financial markets recorded the weekly biggest losses since the financial crisis of 2008. The second change point is detected after approximately four months, when there is a reduction in market uncertainty, but the dependence structure as well as the volatility are not back to pre COVID-19 levels. Comparing the dependence structure across the three periods (i.e., before February 24th, between February 24th and June 26th, after June 26th) we provide many insights on the impact of the pandemic on the stock market and highlight which effects appears to be persistent up to the end of 2021, the last year considered.

The main methodological contribution of this work is the development of a dynamic GGM, which allows for abrupt changes in the dependence structure of the random variables represented by the nodes of the graph. Our model builds on existing literature on GGMs, as well as random change points. Our model construction allows us to control sparsity in the number of change points and/or in the graph structure. We have designed a tailored SMC algorithm, arguing for its use in such a complex setup over other M-H based alternatives and demonstrating its performance on simulated data and on our motivating application.

Our work opens up several avenues for future research.

(a) Scalability. We have not made use of the full SMC machinery. We briefly discuss two directions for increase in model dimension, along the number of nodes, p, and along the length of time instances for observations T. In terms of the size of the graph, recent works have developed effective proposals, informed by the observations, for MCMC methods on graph posteriors, instead of previously used random-walk-type blind moves. See, e.g., van den Boom et al. (2022), and the references therein, for approaches based on Langevin-type analogues for discrete spaces, with parallelisation employed within the specification of the proposal. Such approaches have been seen to be effective for node sizes of $p = \mathcal{O}(10^2)$. In terms of the length T, recent advances on modelling involving change points and accompanying SMC methodologies, can permit for recasting models so that change points also become part of the hidden Markov process (see, e.g. Yildirim et al., 2013). At the same time, SMC methods based on state-of-art particle Gibbs approaches that incorporate backward steps to improve mixing over the update of the unobserved Markovian states, are shown to provide pseudomarginal methodologies of superior mixing compared to standard PMCMC (Lindsten et al., 2014). Such new algorithms are supported by strong theoretical results. Indicatively, the number of particles can now be allowed to remain constant, $N = \mathcal{O}(1)$, as a function of T, when PMCMC requires $N = \mathcal{O}(T)$. Thus, costs for

the overall SMC algorithm can be brought down to $\mathcal{O}(NT)$, from the previous $\mathcal{O}(T^2)$, for big T – with $\mathcal{O}(NT)$ not taking under consideration the option of parallelisation across particles.

(b) Smooth changes. In this work, we have considered abrupt changes in edge inclusion probabilities. Alternatively, we could model edge inclusion probabilities as a function of time, for example, using autoregressive-type models. In this setup, shrinkage priors could be specified to link the probability of edge inclusion at time t to the same probability at time $t-1$ (see, for instance, Molinari et al., 2022a). This approach is amenable to many generalizations, such as the inclusion of covariates. Moreover, the probability of edge inclusion at time t could be a function of the probabilities of edge inclusions at time $t-1$ of a neighbourhood of each node.

(c) Graph sub-structures. Here, we have presented changes between graphs as captured by edge flips before and after a change point. Edge detection is very sensitive to the number of nodes as well as sample size. It has been argued (van den Boom et al., 2022, and references therein) that in many applications a more robust approach is to shift the focus of inference to graph sub-structures such as hubs and communities, with the goal of capturing the evolution over time of such macro-structures which better describe the underlying phenomenon.

(d) More general response types. The model can be easily extended to accommodate different type of responses, such as binary and count data. An easy solution would be the representation of such data in terms of latent variable (Albert and Chib, 1993; Chib and Greenberg, 1998). Moreover, it is straightforward to include time-homogeneous and time-varying covariates to model the mean of the time series, as well as a trend and seasonal component.

References

Albert, J. H. and S. Chib (1993). Bayesian analysis of binary and polychotomous response data. *Journal of the American Statistical Association 88*, 669–679.

Alqaralleh, H. and A. Canepa (2021). Evidence of stock market contagion during the COVID-19 pandemic: A Wavelet-Copula-GARCH approach. *Journal of Risk and Financial Management 14*, 329.

Asai, M., M. McAleer, and J. Yu (2006). Multivariate stochastic volatility: a review. *Econometric Reviews 25*, 145–175.

Atay-Kayis, A. and H. Massam (2005). A Monte Carlo method for computing the marginal likelihood in nondecomposable Gaussian graphical models. *Biometrika 92*, 317–335.

Bauwens, L., S. Laurent, and J. V. Rombouts (2006). Multivariate GARCH models: a survey. *Journal of Applied Econometrics 21*, 79–109.

Beskos, A., D. Crisan, and A. Jasra (2014). On the stability of sequential Monte Carlo methods in high dimensions. *The Annals of Applied Probability 24*, 1396–1445.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics 31*, 307–327.

Bollerslev, T., R. F. Engle, and D. B. Nelson (1994). ARCH models. *Handbook of Econometrics 4*, 2959–3038.

Boudt, K., A. Galanos, S. Payseur, and E. Zivot (2019). Multivariate GARCH models for large-scale applications: A survey. In *Handbook of Statistics*, Volume 41, pp. 193–242. Elsevier.

Caporale, G. M. and T. Zekokh (2019). Modelling volatility of cryptocurrencies using Markov-switching GARCH models. *Research in International Business and Finance 48*, 143–155.

Carvalho, C. M., H. Massam, and M. West (2007). Simulation of hyper-inverse Wishart distributions in graphical models. *Biometrika 94*, 647–659.

Carvalho, C. M. and M. West (2007). Dynamic matrix-variate graphical models. *Bayesian Analysis 2*, 69–97.

Chan, J. C. (2013). Moving average stochastic volatility models with application to inflation forecast. *Journal of Econometrics 176*, 162–172.

Chandra, N. K., P. Mueller, and A. Sarkar (2021). Bayesian scalable precision factor analysis for massive sparse Gaussian graphical models. *arXiv preprint arXiv:2107.11316*.

Cheng, Y. and A. Lenkoski (2012). Hierarchical Gaussian graphical models: Beyond reversible jump. *Electronic Journal of Statistics 6*, 2309–2331.

Chib, S. and E. Greenberg (1998). Analysis of multivariate probit models. *Biometrika 85*, 347–361.

Chopin, N., O. Papaspiliopoulos, et al. (2020). *An introduction to sequential Monte Carlo*. Springer.

Clark, T. E. and F. Ravazzolo (2015). Macroeconomic forecasting performance under alternative specifications of time-varying volatility. *Journal of Applied Econometrics 30*, 551–575.

Contessi, S., P. De Pace, and M. Guidolin (2014). How did the financial crisis alter the correlations of US yield spreads? *Journal of Empirical Finance 28*, 362–385.

Cui, K. (2012). Semiparametric Gaussian variance-mean mixtures for heavy-tailed and skewed data. *International Scholarly Research Notices 2012*.

Danaher, P., P. Wang, and D. M. Witten (2014). The joint graphical lasso for inverse covariance estimation across multiple classes. *Journal of the Royal Statistical Society: Series B (Statistical Methodology) 76*, 373–397.

Dawid, A. P. and S. L. Lauritzen (1993). Hyper Markov laws in the statistical analysis of decomposable graphical models. *The Annals of Statistics*, 1272–1317.

Del Moral, P. (2004). *Feynman-Kac formulae: genealogical and interacting particle systems with applications*, Volume 88. Springer.

Dellaportas, P., P. Giudici, and G. Roberts (2003). Bayesian inference for nondecomposable graphical Gaussian models. *Sankhyā: The Indian Journal of Statistics*, 43–55.

Derbali, A., K. Naoui, M. B. Sassi, and M. M. Amiri (2022). Do COVID-19 epidemic explains the dynamic conditional correlation between china's stock market index and international stock market indices? *The Chinese Economy 55*, 227–242.

Dey, A. K., G. T. Hoque, K. P. Das, and I. Panovska (2022). Impacts of COVID-19 local spread and google search trend on the us stock market. *Physica A: Statistical Mechanics and its Applications 589*, 126423.

Diaconis, P. and D. Ylvisaker (1979). Conjugate priors for exponential families. *The Annals of Statistics 7*, 269–281.

Dobra, A., A. Lenkoski, and A. Rodriguez (2011). Bayesian inference for general Gaussian graphical models with application to multivariate lattice data. *Journal of the American Statistical Association 106*, 1418–1433.

Doucet, A., M. K. Pitt, G. Deligiannidis, and R. Kohn (2015). Efficient implementation of Markov chain Monte Carlo when using an unbiased likelihood estimator. *Biometrika 102*, 295–313.

Dutta, A., E. Bouri, and M. H. Noor (2021). Climate bond, stock, gold, and oil markets: Dynamic correlations and hedging analyses during the COVID-19 outbreak. *Resources Policy 74*, 102265.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, 987–1007.

Engle, R. F. and T. Bollerslev (1986). Modelling the persistence of conditional variances. *Econometric Reviews 5*, 1–50.

Freeman, L. C. (1977). A set of measures of centrality based on betweenness. *Sociometry*, 35–41.

Gibberd, A. J. and J. D. Nelson (2017). Regularized estimation of piecewise constant Gaussian graphical models: The group-fused graphical lasso. *Journal of Computational and Graphical Statistics 26*, 623–634.

Green, P. J. (1995). Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika 82*(4), 711–732.

Guidolin, M., D. La Cara, and M. G. Marcellino (2021). Boosting the forecasting power of conditional heteroskedasticity models to account for COVID-19 outbreaks. *BAFFI CAREFIN Centre Research Paper* (2021-169).

Haas, M., S. Mittnik, and M. S. Paolella (2004). A new approach to Markov-switching GARCH models. *Journal of Financial Econometrics 2*, 493–530.

Higham, N. J. (2002). Computing the nearest correlation matrix—a problem from finance. *IMA journal of Numerical Analysis 22*, 329–343.

Hull, J. and A. White (1987). The pricing of options on assets with stochastic volatilities. *The journal of Finance 42*, 281–300.

Jasra, A., D. A. Stephens, A. Doucet, and T. Tsagaris (2011). Inference for Lévy-driven stochastic volatility models via adaptive sequential Monte Carlo. *Scandinavian Journal of Statistics 38*(1), 1–22.

Jones, B., C. Carvalho, A. Dobra, C. Hans, C. Carter, and M. West (2005). Experiments in stochastic computation for high-dimensional graphical models. *Statistical Science 20*, 388–400.

Just, M. and K. Echaust (2020). Stock market returns, volatility, correlation and liquidity during the COVID-19 crisis: Evidence from the Markov switching approach. *Finance Research Letters 37*, 101775.

Karagiannis, G. and C. Andrieu (2013). Annealed importance sampling reversible jump MCMC algorithms. *Journal of Computational and Graphical Statistics 22*(3), 623–648.

Lauritzen, S. L. (1996). *Graphical Models*, Volume 17. Clarendon Press.

Leday, G. G. and S. Richardson (2019). Fast Bayesian inference in large Gaussian graphical models. *Biometrics 75*, 1288–1298.

Lee, H.-T. (2010). Regime switching correlation hedging. *Journal of Banking & Finance 34*, 2728–2741.

Lien, D., Y. K. Tse, and A. K. Tsui (2002). Evaluating the hedging performance of the constant-correlation GARCH model. *Applied Financial Economics 12*, 791–798.

Lindsten, F., M. I. Jordan, and T. B. Schon (2014). Particle Gibbs with ancestor sampling. *Journal of Machine Learning Research 15*, 2145–2184.

Llopis, F. P., N. Kantas, A. Beskos, and A. Jasra (2018). Particle filtering for stochastic Navier–Stokes signal observed with linear additive noise. *SIAM Journal on Scientific Computing 40*, A1544–A1565.

Meinshausen, N. and P. Bühlmann (2006). High-dimensional graphs and variable selection with the lasso. *The annals of statistics 34*, 1436–1462.

Michis, A. A. (2022). Multiscale partial correlation clustering of stock market returns. *Journal of Risk and Financial Management 15*, 24.

Molinari, M., A. Cremaschi, M. De Iorio, N. Chaturvedi, A. Hughes, and T. Tillin (2022a). Bayesian dynamic network modelling: an application to metabolic associations in cardiovascular disease. *arXiv preprint arXiv:2207.13418*.

Molinari, M., A. Cremaschi, M. De Iorio, N. Chaturvedi, A. D. Hughes, and T. Tillin (2022b). Bayesian nonparametric modelling of multiple graphs with an application to ethnic metabolic differences. *Journal of the Royal Statistical Society. Series C: Applied Statistics*.

Murray, I., Z. Ghahramani, and D. J. C. MacKay (2006). MCMC for doubly-intractable distributions. *In Proceedings of the Twenty-Second Conference on Uncertainty in Artificial Intelligence, UAI'06.*.

Newton, M. A., A. Noueiry, D. Sarkar, and P. Ahlquist (2004). Detecting differential gene expression with a semiparametric hierarchical mixture method. *Biostatistics 5*, 155–176.

Nilsson, O. (2016). On stochastic volatility models as an alternative to GARCH type models.

Persing, A., A. Jasra, A. Beskos, D. Balding, and M. De Iorio (2015). A simulation approach for change-points on phylogenetic trees. *Journal of Computational Biology 22*(1), 10–24.

Peterson, C., F. C. Stingo, and M. Vannucci (2015). Bayesian inference of multiple Gaussian graphical models. *Journal of the American Statistical Association 110*, 159–174.

Pitt, M. K., R. dos Santos Silva, P. Giordani, and R. Kohn (2012). On some properties of Markov chain Monte Carlo simulation methods based on the particle filter. *Journal of Econometrics 171*, 134–151.

Prado, R. and M. West (2010). *Time series: modeling, computation, and inference.* Chapman and Hall/CRC.

Roverato, A. (2002). Hyper inverse Wishart distribution for non-decomposable graphs and its application to Bayesian inference for Gaussian graphical models. *Scandinavian Journal of Statistics 29*, 391–411.

Ruzayqat, H., A. Er-Raiy, A. Beskos, D. Crisan, A. Jasra, and N. Kantas (2021). A lagged particle filter for stable filtering of certain high-dimensional state-space models. *arXiv preprint arXiv:2110.00884*.

Sakurai, Y. and T. Kurosaki (2020). How has the relationship between oil and the us stock market changed after the covid-19 crisis? *Finance Research Letters 37*, 101773.

Sherlock, C., A. H. Thiery, G. O. Roberts, and J. S. Rosenthal (2015). On the efficiency of pseudo-marginal random walk Metropolis algorithms. *The Annals of Statistics 43*, 238–275.

Silvennoinen, A. and T. Teräsvirta (2009). Multivariate GARCH models. In *Handbook of Financial Time Series*, pp. 201–229. Springer.

So, M. E. P., K. Lam, and W. K. Li (1998). A stochastic volatility model with Markov switching. *Journal of Business & Economic Statistics 16*, 244–253.

Taylor, S. J. (1982). Financial returns modelled by the product of two stochastic processes-a study of the daily sugar prices 1961-75. *Time series Analysis: Theory and Practice 1*, 203–226.

Thampanya, N., M. A. Nasir, and T. L. D. Huynh (2020). Asymmetric correlation and hedging effectiveness of gold & cryptocurrencies: From pre-industrial to the 4th industrial revolution. *Technological Forecasting and Social Change 159*, 120195.

Uhler, C., A. Lenkoski, and D. Richards (2018). Exact formulas for the normalizing constants of Wishart distributions for graphical models. *The Annals of Statistics 46*, 90–118.

van den Boom, W., A. Beskos, and M. De Iorio (2022). The G-wishart weighted proposal algorithm: Efficient posterior computation for Gaussian graphical models. *Journal of Computational and Graphical Statistics*, 1–10.

van den Boom, W., M. De Iorio, and A. Beskos (2022). Bayesian learning of graph substructures. *arXiv preprint arXiv:2203.11664*.

van den Boom, W., A. Jasra, M. De Iorio, A. Beskos, and J. G. Eriksson (2022). Unbiased approximation of posteriors via coupled particle Markov chain Monte Carlo. *Statistics and Computing 32*, 1–19.

van Leeuwen, P. J., D. Crisan, O. Lang, and R. Potthast (2021). Bayesian inference for fluid dynamics: A case study for the stochastic rotating shallow water model. *arXiv preprint arXiv:2112.15216*.

Von Furstenberg, G. M., B. N. Jeon, N. G. Mankiw, and R. J. Shiller (1989). International stock price movements: links and messages. *Brookings Papers on Economic Activity 1989*, 125–179.

Wang, H. (2010). Sparse seemingly unrelated regression modelling: Applications in finance and econometrics. *Computational Statistics & Data Analysis 54*, 2866–2877.

Wang, H. and M. West (2009). Bayesian analysis of matrix normal graphical models. *Biometrika 96*, 821–834.

Watts, D. J. and S. H. Strogatz (1998). Collective dynamics of 'small-world' networks. *nature 393*, 440–442.

West, M. and J. Harrison (2006). *Bayesian forecasting and dynamic models*. Springer Science & Business Media.

Wiggins, J. B. (1987). Option values under stochastic volatility: Theory and empirical estimates. *Journal of Financial Economics 19*, 351–372.

Williams, D. R. (2021). Bayesian estimation for Gaussian graphical models: Structure learning, predictability, and network comparisons. *Multivariate Behavioral Research 56*, 336–352.

Yang, S., Z. Lu, X. Shen, P. Wonka, and J. Ye (2015). Fused multiple graphical lasso. *SIAM Journal on Optimization 25*, 916–943.

Yildirim, S., S. S. Singh, and A. Doucet (2013). An online expectation–maximization algorithm for changepoint models. *Journal of Computational and Graphical Statistics 22*, 906–926.

Yousfi, M., Y. B. Zaied, N. B. Cheikh, B. B. Lahouel, and H. Bouzgarrou (2021). Effects of the COVID-19 pandemic on the us stock market and uncertainty: A comparative assessment between the first and second waves. *Technological Forecasting and Social Change 167*, 120710.

Zhou, S., J. Lafferty, and L. Wasserman (2010). Time varying undirected graphs. *Machine Learning 80*, 295–319.

Zou, H. (2006). The adaptive lasso and its oracle properties. *Journal of the American statistical association 101*, 1418–1429.

Supplement to *"Change point detection in dynamic Gaussian graphical models: the impact of COVID-19 pandemic on US stock market"*

Beatrice Franzolini, Alexandros Beskos, Maria De Iorio, Warrick Poklewski Koziell, and Karolina Grzeszkiewicz

A. Evaluation of SMC approximations under scenario A
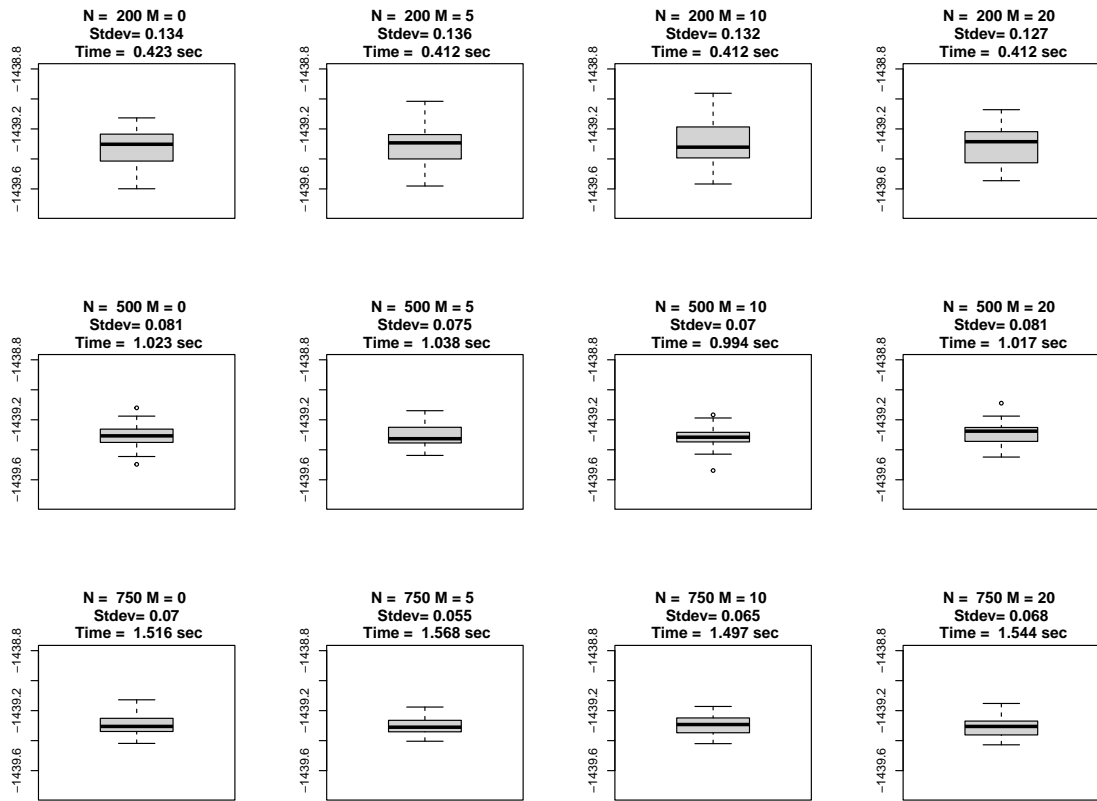
Figure A.1: Log-likelihood estimates for Scenario A obtained with the particle filter by fixing the change points to the truth. Distinct box-plots correspond to different numbers of particles N and/or mutations steps M. For each pair (N, M), we run the algorithm 30 times and obtain the log-likelihood estimates. Each box-plot shows the distribution of such estimates. The variability of the estimates is small for all considered pairs (N, M).

B. Simulation studies: additional figures and tables

(a) No change points and independence

(b) No change points and dependence
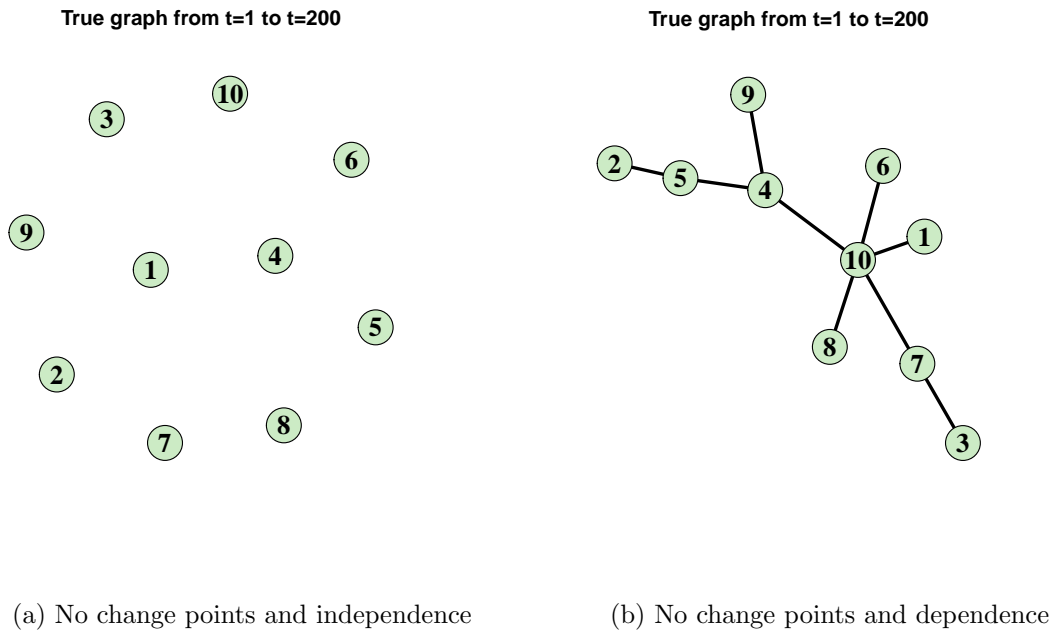
Figure B.1: Graph structure of Scenarios 1 (panel a) and 2 (panel b) used as data generating mechanism for assensing model's performance.
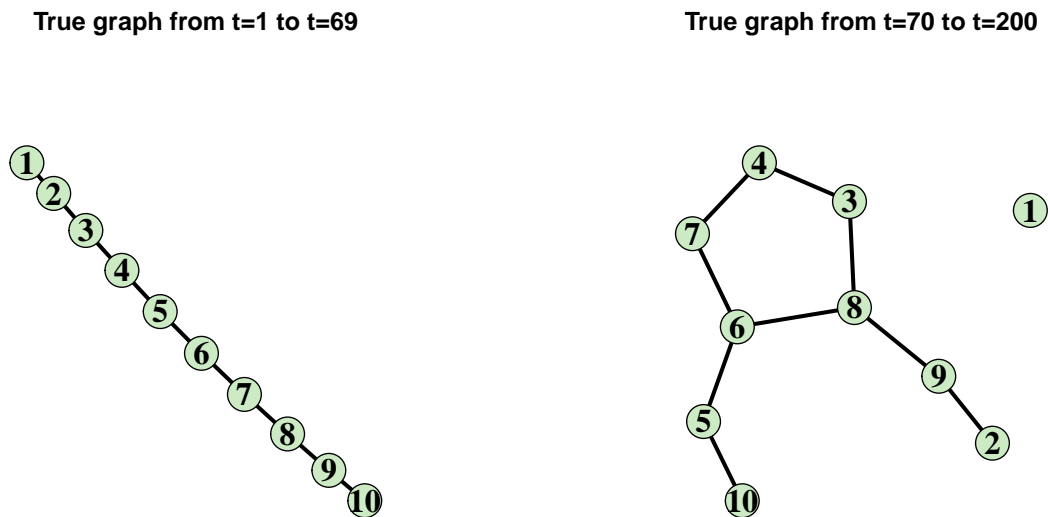
Figure B.2: Scenario 3: graphs before and after the change point at $t = 70$.

	True change point configuration		MAP estimate	MAP prob.
		Replica 1	$\emptyset$	1
		Replica 2	$\emptyset$	1
		Replica 3	$\emptyset$	1
		Replica 4	$\emptyset$	1
Scenario 1	$\emptyset$	Replica 5	$\emptyset$	1
		Replica 6	$\emptyset$	1
		Replica 7	$\emptyset$	1
		Replica 8	$\emptyset$	1
		Replica 9	$\emptyset$	1
		Replica 10	$\emptyset$	1
			MAP estimate	MAP prob.
		Replica 1	$\emptyset$	0.980
		Replica 2	$\emptyset$	1
		Replica 3	$\emptyset$	0.993
		Replica 4	$\emptyset$	1
Scenario 2	$\emptyset$	Replica 5	$\emptyset$	1
		Replica 6	$\emptyset$	1
		Replica 7	$\emptyset$	0.997
		Replica 8	$\emptyset$	1
		Replica 9	$\emptyset$	1
		Replica 10	$\emptyset$	1

Table B.1: Scenario 1 and 2: Posterior summaries for change points. MAP estimates and MAP probabilities (for the posterior over all configurations of change points).

C. US stock market analysis: additional figures and tables

Figure C.1: Weekly Standardised Logarithmic Returns of nine Industry Portfolios from January 2019 to December 2021.

C.1 Descriptive indexes and summaries of the posterior graphs

The following tables contains degree centrality, betweenness centrality, local clustering and global clustering coefficients for the estimated graphs.

Portfolios	pre COVID-19	during COVID-19	post COVID-19
Consumer Nondurables (NoDur)	2	3	**4**
Consumer Durables (Durbl)	1	2	1
Manufacturing (Manuf)	**3**	4	4
Energy (Enrgy)	1	2	3
High Technology (HiTec)	**3**	3	2
Telecommunications (Telcm)	2	1	1
Shops (Shops)	2	2	4
Health (Hlth)	1	3	2
Utilities (Utils)	1	2	3

Table C.1.1: **Degree centrality**. Degree centrality of a certain node is the number of vertices in the neighbourhood of that node. In **bold**, we highlight the highest degree centrality for ech graph.

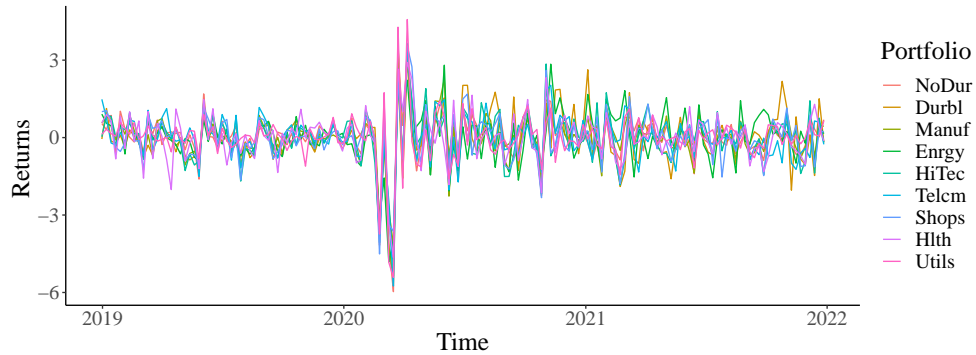| Portfolios	pre COVID-19	during COVID-19	post COVID-19
Consumer Nondurables (NoDur)	7	**11.5**	**10**
Consumer Durables (Durbl)	0	4.5	0
Manufacturing (Manuf)	**13**	14	**6.67**
Energy (Enrgy)	0	0	1.33
High Technology (HiTec)	**19**	3.5	2.33
Telecommunications (Telcm)	12	0	0
Shops (Shops)	**15**	0	**10**
Health (Hlth)	0	**5.5**	2.33
Utilities (Utils)	0	0	1.33

Table C.1.2: **Betweenness centrality**. Between centrality of a certain node v is $c(v) = \sum_{h \neq v \neq k} \sigma_{h,k}(v)\sigma_{h,k}$, where $\sigma_{h,k}(v)$ is the number of geodesics (i.e., shortest paths) between nodes h and k going through node j and $\sigma_{h,k}$ is the number of geodesics between nodes h and k. In **bold**, we highlight the three highest values for each graph.

Portfolios	pre COVID-19	during COVID-19	post COVID-19
Consumer Nondurables (NoDur)	0	0	0.17
Consumer Durables (Durbl)	0	0	0
Manufacturing (Manuf)	0	0.17	0.50
Energy (Enrgy)	0	1	0.67
High Technology (HiTec)	0	0.33	0
Telecommunications (Telcm)	0	0	0
Shops (Shops)	0	1	0.17
Health (Hlth)	0	0.33	0
Utilities (Utils)	0	1	0.67
Global clustering coefficient	0	0.43	0.24

Table C.1.3: **Local clustering coefficients** Local clustering coefficients are the ratio between the number of triangular cliques, of which the node is a part, and $d(d-1)/2$, where d is the degree centrality of that node. Global clustering is the average of local coefficients.

C.2 Posterior estimates of covariance and correlation matrices.

	NoDur	Durbl	Manuf	Enrgy	HiTec	Telcm	Shops	Hlth	Utils
NoDur	1.82	1.04	1.45	1.47	1.09	1.35	0.96	1.4	0.81
Durbl	1.04	7.43	1.89	2.3	2.07	0.77	2.43	1.47	0.48
Manuf	1.45	1.89	3.02	3.07	1.61	1.07	1.75	1.44	0.57
Enrgy	1.47	2.3	3.07	6.93	1.9	1.09	2.13	1.57	0.49
HiTec	1.09	2.07	1.61	1.9	3.32	0.81	1.91	2.11	0.61
Telcm	1.35	0.77	1.07	1.09	0.81	2.69	0.71	1.03	0.6
Shops	0.96	2.43	1.75	2.13	1.91	0.71	2.25	1.37	0.44
Hlth	1.4	1.47	1.44	1.57	2.11	1.03	1.37	3.61	0.96
Utils	0.81	0.48	0.57	0.49	0.61	0.6	0.44	0.96	1.75

	NoDur	Durbl	Manuf	Enrgy	HiTec	Telcm	Shops	Hlth	Utils
NoDur	28.87	38.44	34.51	41.91	22.45	25.55	20.75	21.9	36.3
Durbl	38.44	82.86	48.99	59.91	38.54	34.02	37.92	32.26	50.32
Manuf	34.51	48.99	45.22	56.01	28.18	30.54	26.44	26.57	46.07
Enrgy	41.91	59.91	56.01	88.95	34.47	37.08	32.34	32.5	52.35
HiTec	22.45	38.54	28.18	34.47	24.81	19.87	20.8	19.7	29.02
Telcm	25.55	34.02	30.54	37.08	19.87	25.01	18.36	19.38	32.12
Shops	20.75	37.92	26.44	32.34	20.8	18.36	20.46	17.41	27.16
Hlth	21.9	32.26	26.57	32.5	19.7	19.38	17.41	21.05	27.52
Utils	36.3	50.32	46.07	52.35	29.02	32.12	27.16	27.52	53.58

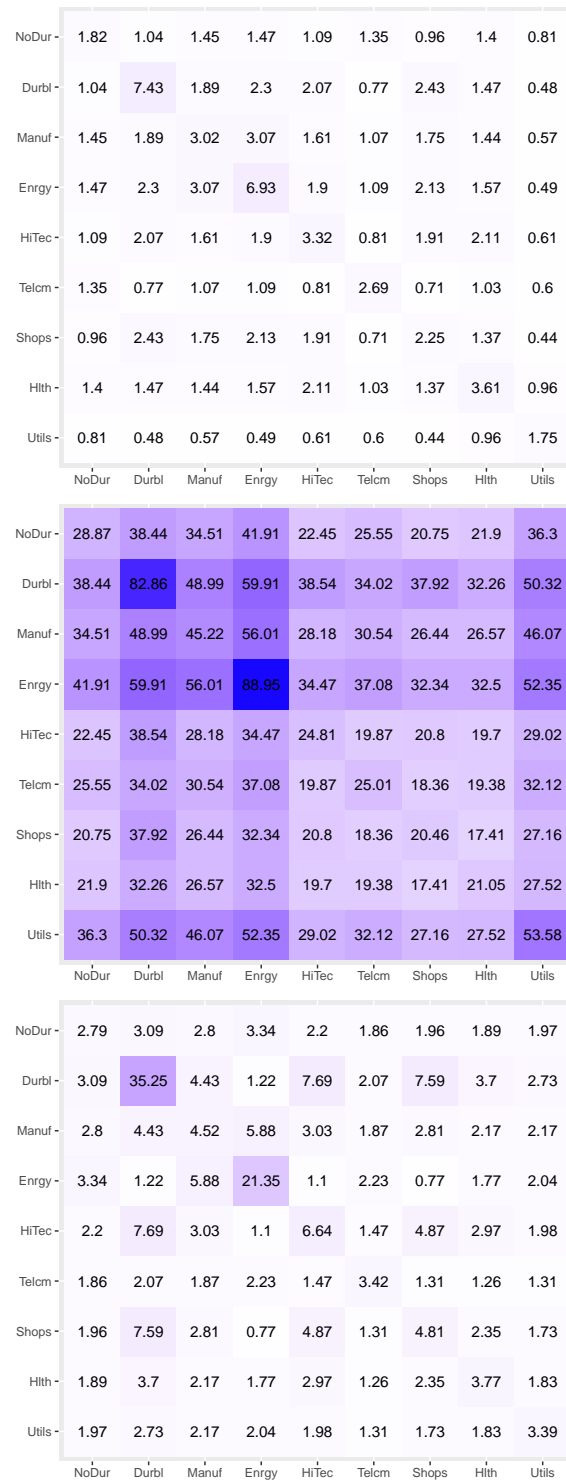|	NoDur	Durbl	Manuf	Enrgy	HiTec	Telcm	Shops	Hlth	Utils
NoDur	2.79	3.09	2.8	3.34	2.2	1.86	1.96	1.89	1.97
Durbl	3.09	35.25	4.43	1.22	7.69	2.07	7.59	3.7	2.73
Manuf	2.8	4.43	4.52	5.88	3.03	1.87	2.81	2.17	2.17
Enrgy	3.34	1.22	5.88	21.35	1.1	2.23	0.77	1.77	2.04
HiTec	2.2	7.69	3.03	1.1	6.64	1.47	4.87	2.97	1.98
Telcm	1.86	2.07	1.87	2.23	1.47	3.42	1.31	1.26	1.31
Shops	1.96	7.59	2.81	0.77	4.87	1.31	4.81	2.35	1.73
Hlth	1.89	3.7	2.17	1.77	2.97	1.26	2.35	3.77	1.83
Utils	1.97	2.73	2.17	2.04	1.98	1.31	1.73	1.83	3.39

Figure C.2.1: Posterior estimates of the variance and covariance matrices for the original non-standardised weekly percentage logarithmic returns.
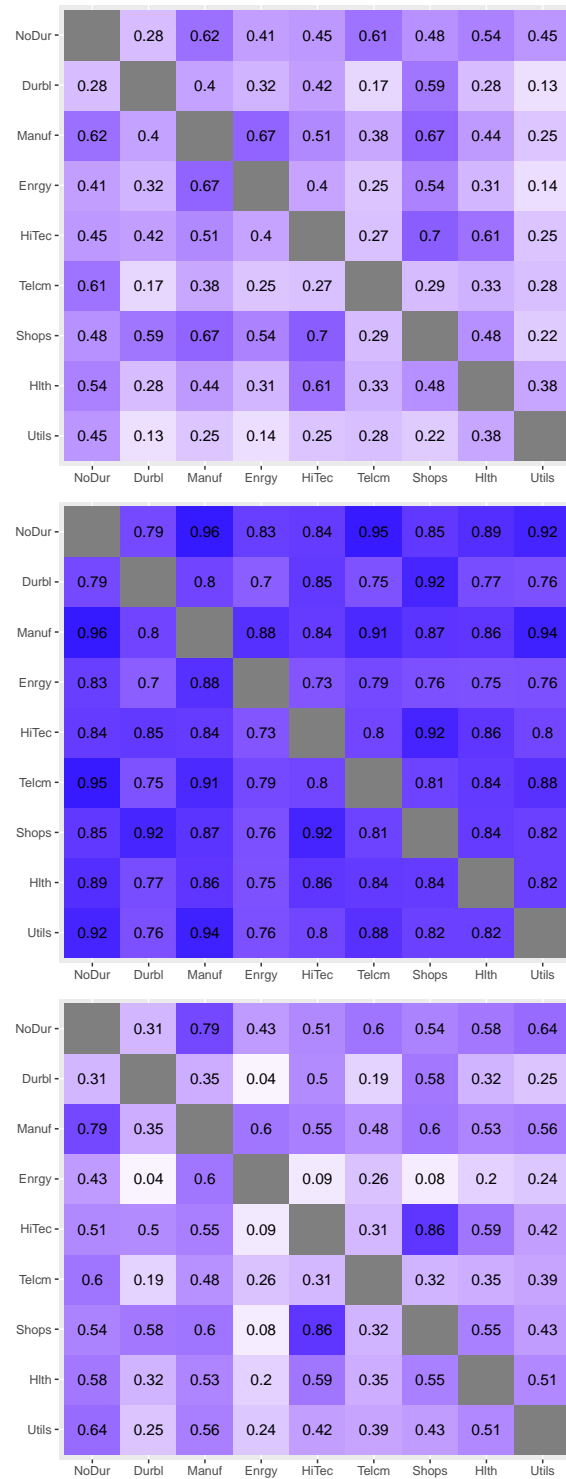
Figure C.2.2: Posterior estimates of the correlation matrices for weekly logarithmic returns.