Metrika
https://doi.org/10.1007/s00184-025-00992-6

Partial correlation graphs for continuous-parameter time series

Vicky Fasen-Hartmann[1] · Lea Schenk[1]

Received: 31 January 2024 / Accepted: 2 February 2025

Abstract

In this paper, we establish the partial correlation graph for multivariate continuous-time stochastic processes, assuming only that the underlying process is stationary and mean-square continuous with expectation zero and spectral density function. In the partial correlation graph, the vertices are the components of the process and the undirected edges represent partial correlations between the vertices. To define this graph, we therefore first introduce the partial correlation relation for continuous-time processes and provide several equivalent characterisations. In particular, we establish that the partial correlation relation defines a graphoid. The partial correlation graph additionally satisfies the usual Markov properties and the edges can be determined very easily via the inverse of the spectral density function. Throughout the paper we compare and relate the partial correlation graph to the mixed (local) orthogonality graph of Fasen-Hartmann and Schenk (Stoch Process Appl 179:104501, 2024. https://doi.org/10.1016/j.spa.2024.104501). Finally, as an example, we explicitly characterise and interpret the edges in the partial correlation graph for the popular multivariate continuous-time AR (MCAR) processes.

Keywords orthogonality graph · Markov property · MCAR process · Partial correlation · Stationary process · Undirected graph

Mathematics Subject Classification Primary 62H22 · 62M20; Secondary 62M10 · 60G25

1 Introduction

Our interest in this paper is in graphical models for wide-sense stationary and mean-square continuous stochastic processes. Graphical models are probabilistic networks,

✉ Vicky Fasen-Hartmann
vicky.fasen@kit.edu

Lea Schenk
lea.schenk@kit.edu

1 Institute of Stochastics, Karlsruhe Institute of Technology, Karlsruhe, Germany

Published online: 11 March 2025

where the vertices represent the components of a random object, e.g., a random vector or a vector-valued stochastic process, and the edges illustrate specific interconnections between them. They are popular because they visualise dependency structures of the random object in a clear and simple way, which can then be analysed, interpreted, and easily communicated. Furthermore, graphical models are an important tool for dimension reduction in high-dimensional models. Due to the growth of complex multivariate data sets and networks, the theory and methodology of graphical models have experienced a surge of research development in probability theory and statistics (Whittaker 2008; Edwards 2000; Lauritzen 2004; Maathuis et al. 2019), and they have been applied in fields as diverse as biology, neuroscience, economics, finance, and psychology, to name just a few.

Although in networks of interconnected processes the data are observed in discrete time, in many situations it is more appropriate to specify the underlying stochastic process in *continuous time*. This is particularly necessary for high-frequency data, irregularly spaced data or data with missing observations, which are common in finance, econometrics, signal processing, and turbulence. In addition, many physical and signal processing models are formulated in continuous time, so such an approach is often more natural.

Overall, however, there is very little theory on graphical models for multivariate stochastic processes in continuous time. The established graphical models are mostly limited to conditional independence and local independence graphs, which have been studied by Mogensen and Hansen (2020, 2022), Didelez (2007, 2008), Aalen (1987), Schweder (1970). They are particularly suitable for semimartingales and point processes, but do not seem to be the right tool for time series. A general approach for graphical continuous-parameter time series models are the (local) orthogonality graphs of Fasen-Hartmann and Schenk (2024a, b), which are mixed graphs representing Granger causalities and contemporaneous correlations. The orthogonality graphs satisfy the usual Markov properties and the theory holds for a very large class of time series models. However, the computation of the edges can be quite challenging in certain examples and the characterisations may not be convenient, as we know for multivariate continuous-time ARMA processes from Fasen-Hartmann and Schenk (2024b), which is problematic in practice. Until now, an undirected graphical model for continuous-time stationary processes has been lacking. Therefore, we aim to fill this gap and provide a user-friendly and powerful undirected graphical model for continuous-time stochastic processes.

In the graphical model in this paper, the vertices $V = \{1, \dots, k\}$ represent the components of a k-dimensional continuous-time process and the edges visualise partial correlations between these components. The concept of *partial correlation* is an important and well-studied measure of dependence in statistics. For an $\mathbb{R}^k$-valued random vector $\mathbf{Y} = (Y_1, \dots, Y_k)^\top$ with positive definite covariance matrix Σ, the partial correlation of Y_a and Y_b given $Y_{V\setminus\{a,b\}} := \overline{\mathbf{Y}} := (Y_l)_{l\in V\setminus\{a,b\}}$ measures the correlation of the real-valued random variables Y_a and Y_b after removing the linear effects of $Y_{V\setminus\{a,b\}}$. The partial correlation is determined as follows: Denote by Σ_{AB} the respective submatrix of Σ for $A, B \subseteq V$ and consider the linear regression problems

$$\beta_l = \text{argmin}_{\beta\in\mathbb{R}^{k-2}}\mathbb{E}(Y_l - \beta^\top\overline{\mathbf{Y}})^2, \quad l \in \{a, b\}. \tag{1.1}$$

These problems have the well-known solution (Fujikoshi et al. 2010; Anderson 1984)

$$\beta_l = \left(\Sigma_{V\setminus\{a,b\}V\setminus\{a,b\}}\right)^{-1} \Sigma_{V\setminus\{a,b\}l}.$$

Furthermore, the residuals $\varepsilon_{a|V\setminus\{a,b\}} := Y_a - \beta_a^\top \overline{\mathbf{Y}}$ and $\varepsilon_{b|V\setminus\{a,b\}} := Y_b - \beta_b^\top \overline{\mathbf{Y}}$ satisfy

$$\mathrm{Cov}(\varepsilon_{a|V\setminus\{a,b\}}, \varepsilon_{b|V\setminus\{a,b\}}) = \Sigma_{ab} - \Sigma_{aV\setminus\{a,b\}} \left(\Sigma_{V\setminus\{a,b\}V\setminus\{a,b\}}\right)^{-1} \Sigma_{V\setminus\{a,b\}b}, \quad (1.2)$$

which is the *partial covariance* of Y_a and Y_b *given* $Y_{V\setminus\{a,b\}}$. Similarly, the correlation of the residuals is called *partial correlation* of Y_a and Y_b *given* $Y_{V\setminus\{a,b\}}$, also known as *coherence*, and is equal to

$$\begin{aligned}\mathrm{Corr}(\varepsilon_{a|V\setminus\{a,b\}}, \varepsilon_{b|V\setminus\{a,b\}}) &= \frac{\mathrm{Cov}(\varepsilon_{a|V\setminus\{a,b\}}, \varepsilon_{b|V\setminus\{a,b\}})}{\sqrt{\mathrm{Cov}(\varepsilon_{a|V\setminus\{a,b\}}, \varepsilon_{a|V\setminus\{a,b\}})\mathrm{Cov}(\varepsilon_{b|V\setminus\{a,b\}}, \varepsilon_{b|V\setminus\{a,b\}})}} \\ &= -\frac{[\Sigma^{-1}]_{ab}}{\sqrt{[\Sigma^{-1}]_{aa}[\Sigma^{-1}]_{bb}}}. \end{aligned} \quad (1.3)$$

From the representation (1.3) we see that the partial correlation is completely determined by the inverse covariance matrix Σ^{-1}, also called *concentration* or *precision matrix*. For a Gaussian random vector, zero partial correlation is even equivalent to Y_a and Y_b being independent given $Y_{V\setminus\{a,b\}}$.

An extension of partial correlation to stationary time series models $(Y_V(n))_{n\in\mathbb{N}}$ in *discrete time* is quite old (Tick 1963) and is ubiquitous in the analysis of multivariate time series (Priestley 1981; Brillinger 2001; Gardner 1988). Recall that wide-sense stationary processes are stochastic processes with a constant expectation at each time point and existing covariance function depending only on the time lags. The covariance function in the time domain is directly related to the spectral density by Fourier transformation which leads to a frequency domain representation of a stationary stochastic process. For these time series models, the partial covariance function of $(Y_a(n))_{n\in\mathbb{N}}$ and $(Y_b(n))_{n\in\mathbb{N}}$ given $(Y_{V\setminus\{a,b\}}(n))_{n\in\mathbb{N}}$ is zero if and only if the partial spectral density function of $(Y_a(n))_{n\in\mathbb{N}}$ and $(Y_b(n))_{n\in\mathbb{N}}$ given $(Y_{V\setminus\{a,b\}}(n))_{n\in\mathbb{N}}$ is zero, such that in the frequency domain, the role of the partial correlation function, the coherence function of the noise processes, is taken over by the *spectral coherence function* of the noise process (cf. Definition 2.3), the normalised cross spectral density. The role of the covariance matrix Σ is taken over by the matrix-valued spectral density function of the process $(Y_V(n))_{n\in\mathbb{N}}$.

The applications of spectral coherence are very broad, especially in signal processing, but the word coherence may have a slightly different meaning in different fields (Gardner 1992). However, to the best of our knowledge, a mathematically rigorous theory for the definition of partial correlation for continuous-parameter time series is missing in the literature, so we include the theory first and relate it to an optimisation problem as in (1.1) in Proposition 3.3 and Remark 3.4. In Lemma 3.5 we present the corresponding result to (1.2) for continuous-time stochastic processes and Proposition 4.4 corresponds to (1.3), respectively. It is important to note that in the above regression problem, $\beta_l^\top \overline{\mathbf{Y}}$ is the linear projection of Y_l on the linear space generated by the components of $Y_{V\setminus\{a,b\}} = \overline{\mathbf{Y}}$ because our approach builds on this idea. In particular, we show that our definition of partial correlation satisfies the important graphoid properties.

The subject of this paper is *partial correlation graphs* for continuous-time wide-sense stationary and mean-square continuous stochastic processes with expectation zero and spectral density. Partial correlation graphs for discrete-time wide-sense stationary stochastic processes with expectation zero and spectral density originated in Brillinger (1996) and Dahlhaus (2000), and are a widely used frequency domain approach for constructing graphs. In our graphical model and in the model of Dahlhaus (2000), the vertices are the components of a multivariate time series and the edges between the vertices are drawn when the spectral coherence function in these components is not the zero function, meaning that the component processes are partially correlated given the remaining process. The method of Dahlhaus (2000) has since been used in a wide variety of applications, including the identification of synaptic connections in air pollution data (Dahlhaus 2000), human tremor data (Dahlhaus and Eichler 2003), vital signs of intensive care patients (Gather et al. 2002), financial data (Abdelwahab et al. 2008), and neuro-physical signals (Dahlhaus et al. 1997; Eichler et al. 2003; Medkour et al. 2009), which demonstrates the popularity of partial correlation graphs in identifying a network structure.

This paper aims to define a probabilistic network of interconnected continuous-time stochastic processes, where the dependence structure in the network is modelled by partial correlation. The proposed partial correlation graph is simple in the sense that there are neither loops from a vertex to itself nor any multiple edges between vertices and it satisfies the required Markov properties that associate the graph factorisation to the partial correlation. Moreover, it is easy to handle in applications because the edges reflect zero entries in the inverse spectral density function. We derive important relations between the undirected partial correlation graph and the recently introduced mixed orthogonality graph of Fasen-Hartmann and Schenk (2024a). In the mixed orthogonality graph, the directed and undirected edges can be defined by conditional orthogonality relations of properly defined linear subspaces generated by the underlying stochastic process, similarly, the edges in the partial correlation graph can be defined by conditional orthogonality. We use this commonality to compare both graphical models and to show the important connection that the edges in the partial correlation graph are also edges in the augmented orthogonality graph. Furthermore, as an example, we apply the partial correlation graph to multivariate continuous-time autoregressive (MCAR) processes and present a perspective on estimation. In the context of MCAR processes, we additionally obtain that the edges of the partial correlation graph are also edges in the corresponding augmented local orthogonality graph of Fasen-Hartmann and Schenk (2024a). Finally, a major conclusion of this paper is that the edges of the continuous-time model are in general not identifiable from equidistantly sampled observations, but this is different for high-frequency data.

Structure of the paper

The paper is structured as follows. In Sect. 2, we lay the groundwork for the paper by introducing relevant properties of multivariate wide-sense stationary and mean-square continuous processes. Then, in Sect. 3, we define the partial correlation relation and establish characterisations and properties. This preliminary work results in the definition of the partial correlation graph $G_{PC} = (V, E_{PC})$ in Sect. 4, where we also discuss edge characterisations, as well as Markov properties, and the relations to the orthogonality graph. As an example, in Sect. 5, we apply the partial correlation graph to MCAR processes and compare it to the local orthogonality graph for MCAR processes. Finally, we complete the paper with a brief conclusion in Sect. 6. The proofs of the paper are given in Sect. 7.

Notation

In the following, $I_k \in \mathbb{R}^{k\times k}$ is the $(k \times k)$-dimensional identity matrix, $0_{k\times d} \in \mathbb{R}^{k\times d}$ is the $(k \times d)$-dimensional zero matrix, and 0_k is either the k-dimensional zero vector or the $(k \times k)$-dimensional zero matrix, which should be clear from the context. The entries and submatrices of a matrix M are denoted by $[M]_{ab}$ for $a, b \in V$ and $[M]_{AB}$ for $A, B \subseteq V$, respectively. The cardinality of a set A is denoted by $|A|$ and the closure of a set A is denoted by $\mathrm{cl}(A)$. For Hermitian matrices $M, N \in \mathbb{C}^{k\times k}$, we write $M \geq N$ if and only if $M - N$ is positive semi-definite. Similarly, we write $M > 0$ if and only if M is positive definite and define $\sigma(M)$ as the set of eigenvalues of M. Finally, for a matrix-valued function $f : \mathbb{R} \to \mathbb{C}^{k\times k}$ with $f(\lambda) > 0$, we define the function $g : \mathbb{R} \to \mathbb{C}^{k\times k}$ by $g(\lambda) = f(\lambda)^{-1}$, $\lambda \in \mathbb{R}$.

2 Preliminaries

In this paper, we consider wide-sense stationary and mean-square continuous stochastic processes $\mathcal{Y}_V = (Y_V(t))_{t\in\mathbb{R}}$ in continuous time with index set $V = \{1, \dots, k\}$, $\mathbb{E}(Y_V(t)) = 0_k$ for $t \in \mathbb{R}$, and existing spectral density function $f_{Y_V Y_V}(\lambda)$ for $\lambda \in \mathbb{R}$. In this section, we present some well-known properties of these processes that are relevant to this paper. These results date back to Khintchine (1934) and Cramér (1940) and were summarised in a comprehensive overview, e.g., by Doob (1953) and Rozanov (1967); see as well the recent monograph by Brockwell and Lindner (2024).

First, note that $\mathcal{Y}_V$ is mean-square continuous if and only if

$$\lim_{t\to 0} c_{Y_V Y_V}(t) = c_{Y_V Y_V}(0), \tag{2.1}$$

where $(c_{Y_V Y_V}(t))_{t\in\mathbb{R}} = (\mathbb{E}[Y_V(t)Y_V(0)^\top])_{t\in\mathbb{R}}$ is the autocovariance function of $\mathcal{Y}_V$. Further, a key property of wide-sense stationary and mean-square continuous stochastic processes with expectation zero and existing spectral density function is their spectral representation

$$Y_V(t) = \int_{-\infty}^{\infty} e^{i\lambda t} Z_V(d\lambda), \quad t \in \mathbb{R}, \tag{2.2}$$

with respect to a random orthogonal measure $Z_V = (Z_1, \ldots, Z_k)^\top$, where

$$\mathbb{E}[Z_V(d\lambda)\overline{Z_V(d\mu)}^\top] = \delta_{\lambda=\mu} f_{Y_V Y_V}(\lambda)d\lambda, \quad \mathbb{E}[Z_V(d\lambda)] = 0_k,$$

and $\delta_{\lambda=\mu}$ is the Kronecker delta. For $A, B \subseteq V$ we refer to the (submatrix) function $f_{Y_A Y_B}(\lambda) = [f_{Y_V Y_Y}(\lambda)]_{AB}$, $\lambda \in \mathbb{R}$, as the *(cross-) spectral density function* of the subprocesses $\mathcal{Y}_A$ and $\mathcal{Y}_B$. Important properties of the spectral density function are the following.

Lemma 2.1 *Let $\lambda, t \in \mathbb{R}$. Then the following statements hold.*

(a) $\int_{-\infty}^{\infty} \| f_{Y_V Y_V}(\lambda)\| d\lambda < \infty$,
(b) $c_{Y_V Y_V}(t) = \int_{-\infty}^{\infty} e^{i\lambda t} f_{Y_V Y_V}(\lambda)d\lambda$,
(c) $f_{Y_V Y_V}(\lambda) \geq 0$ *and* $f_{Y_V Y_V}(\lambda) = \overline{f_{Y_V Y_V}(\lambda)}^\top$.

Remark 2.2 To obtain the one-to-one relationship between $c_{Y_V Y_V}(t)$ and $f_{Y_V Y_V}(\lambda)$ via

$$f_{Y_V Y_V}(\lambda) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\lambda t} c_{Y_V Y_V}(t)dt, \quad \lambda \in \mathbb{R}, \tag{2.3}$$

(complementary to Lemma 2.1(b)) additional integrability assumptions on the covariance function are required. If $\int_{-\infty}^{\infty} \|c_{Y_V Y_V}(t)\|dt < \infty$ the equality (2.3) is the classical Fourier transform and if $\int_{-\infty}^{\infty} \|c_{Y_V Y_V}(t)\|^2 dt < \infty$ we have a relation by the Fourier-Plancharel transform. However, in this paper, we only require the existence of a spectral density function and not the relation (2.3). Therefore, long memory processes such as multivariate fractionally integrated CARMA processes are also covered.

In addition to the spectral density function, we introduce the *(cross-) spectral coherence function* of $\mathcal{Y}_A$ and $\mathcal{Y}_B$, which is obtained by rescaling the cross-spectral density function of $\mathcal{Y}_A$ and $\mathcal{Y}_B$, and which provides a measure of strength.

Definition 2.3 The *(cross-) spectral coherence function* of $\mathcal{Y}_A$ and $\mathcal{Y}_B$ is defined as

$$R_{Y_A Y_B}(\lambda) := \Big(f_{Y_A Y_A}(\lambda)\Big)^{-1/2} f_{Y_A Y_B}(\lambda)\Big(f_{Y_B Y_B}(\lambda)\Big)^{-1/2}, \quad \lambda \in \mathbb{R}. \tag{2.4}$$

If $f_{Y_A Y_A}(\lambda)$ or $f_{Y_B Y_B}(\lambda)$ is singular for some $\lambda \in \mathbb{R}$, we set $R_{Y_A Y_B}(\lambda) := 0_{|A|\times|B|}$.

3 Partial correlation relation

In this section, we introduce the concept of partial correlation for wide-sense stationary, mean-square continuous processes $\mathcal{Y}_V$ that have expectation zero and an existing spectral density function. Therefore, in Sect. 3.1, we define and interpret the so-called partial correlation relation and compute the orthogonal projections therein. Additionally, we discuss properties of $Y_A(t)$ given the linear information of $\mathcal{Y}_C$, i.e., the resulting noise process in the partial correlation relation. Section 3.2 is then devoted to multiple characterisations of the partial correlation relation. We provide characterisations in terms of the spectral density function and the spectral coherence function of the resulting noise processes. Importantly, we present the key characterisation of the partial correlation relation involving the inverse $g_{Y_V Y_V}$ of the spectral density function $f_{Y_V Y_V}$ of the underlying process $\mathcal{Y}_V$. We conclude the section with the main result that partial correlation satisfies the important graphoid properties. Throughout this section, A, B, C are subsets of V.

3.1 Partial correlation relation and orthogonal projections

Let us first introduce the concept of partial correlation and make some comments on this definition. Therefore, we define the linear space of linear transformations of $\mathcal{Y}_C$ as

$$\mathcal{L}_{Y_C} := \left\{ \int_{-\infty}^{\infty} \varphi(\lambda)^\top Z_C(d\lambda) : \varphi : \mathbb{R} \to \mathbb{C}^{|C|} \text{ meas.}, \int_{-\infty}^{\infty} \varphi(\lambda)^\top f_{Y_C Y_C}(\lambda)\overline{\varphi(\lambda)}\, d\lambda < \infty \right\}, \tag{3.1}$$

where Z_C is the random spectral measure from the spectral representation (2.2) of $\mathcal{Y}_C$. Therefore, remark that stochastic integrals of deterministic Lebesgue measurable functions with respect to a random orthogonal measure are defined in the usual L^2-sense. For details on the definition and properties of such integrals, we refer to Doob (1953), Rozanov (1967) and Brockwell and Lindner (2024). In particular, by definition of the linear space $\mathcal{L}_{Y_C}$, we have for any $t \in \mathbb{R}$ and $c \in C$ that $Y_c(t) \in \mathcal{L}_{Y_C}$ and as well any kind of linear combination is in this linear space. Note that we prefer to work with the spectral representation of the linear space $\mathcal{L}_{Y_C}$ as given in (3.1) although there exist a time domain representation because the proofs of the present paper are easier with the spectral domain representation.

Definition 3.1 Two subprocesses $\mathcal{Y}_A$ and $\mathcal{Y}_B$ of $\mathcal{Y}_V$ are *partially uncorrelated* given $\mathcal{Y}_C$ ($\mathcal{Y}_A \perp\!\!\!\perp \mathcal{Y}_B \mid \mathcal{Y}_C$) if and only if

$$\mathbb{E}\left[\left(Y_a(t) - P_{\mathcal{L}_{Y_C}} Y_a(t)\right)\overline{\left(Y_b(t) - P_{\mathcal{L}_{Y_C}} Y_b(t)\right)}\right] = 0 \quad \forall\, a \in A,\ b \in B,\ t \in \mathbb{R},$$

where $P_{\mathcal{L}_{Y_C}}$ denotes the orthogonal projection on $\mathcal{L}_{Y_C}$.

By definition, the linear space $\mathcal{L}_{Y_C}$ describes the linear information provided by $\mathcal{Y}_C$ over all time points. Thus, $\mathcal{Y}_A \perp\!\!\!\perp \mathcal{Y}_B \mid \mathcal{Y}_C$ states, as desired, that, for all $t \in \mathbb{R}$, $Y_A(t)$ and $Y_B(t)$ are uncorrelated given the linear information provided by $\mathcal{Y}_C$. The concept of partial correlation for continuous-time processes can thus be seen as an extension of the definition of partial correlation for random vectors in Sect. 1.

Remark 3.2 (a) Certainly, the partial correlation relation is symmetric and

$$\mathcal{Y}_A \perp\!\!\!\perp \mathcal{Y}_B \mid \mathcal{Y}_C \quad \Leftrightarrow \quad \mathcal{Y}_a \perp\!\!\!\perp \mathcal{Y}_b \mid \mathcal{Y}_C \quad \forall\, a \in A, b \in B,$$

which is useful for verifying zero partial correlation. Furthermore, statements can usually be made without loss of generality for $A = \{a\}$ and $B = \{b\}$. The corresponding results in the multivariate case follow immediately.

(b) In terms of the conditional orthogonality relation $\perp$ (cf. Eichler 2007, Appendix A), the partial correlation relation can be characterized as well. Therefore, we define for $t \in \mathbb{R}$ and $A, B \subseteq V$ the linear spaces generated by the subprocesses $\mathcal{Y}_A$ and $\mathcal{Y}_B$ at time t as

$$L_{Y_A}(t) := \left\{ \sum_{a \in A} \gamma_a Y_a(t) : \gamma_a \in \mathbb{C} \right\} \quad \text{and} \quad L_{Y_B}(t) := \left\{ \sum_{b \in B} \gamma_b Y_b(t) : \gamma_b \in \mathbb{C} \right\}. \tag{3.2}$$

Then

$$\begin{aligned} \mathcal{Y}_A \perp\!\!\!\perp \mathcal{Y}_B \mid \mathcal{Y}_C \quad &\Leftrightarrow \quad \mathbb{E}\left[\left(Y^A - P_{\mathcal{L}_{Y_C}} Y^A\right)\left(Y^B - P_{\mathcal{L}_{Y_C}} Y^B\right)\right] = 0 \\ &\qquad \forall\, Y^A \in L_{Y_A}(t),\ Y^B \in L_{Y_B}(t),\ t \in \mathbb{R}, \\ &\Leftrightarrow: \quad L_{Y_A}(t) \perp L_{Y_B}(t) \mid \mathcal{L}_{Y_C} \quad \forall\, t \in \mathbb{R}, \end{aligned}$$

which is as well the definition of conditional orthogonality of $L_{Y_A}(t)$ and $L_{Y_B}(t)$ given $\mathcal{L}_{Y_C}$ for all $t \in \mathbb{R}$.

Proposition 3.3 *Suppose that $f_{Y_C Y_C}(\lambda) > 0$ for $\lambda \in \mathbb{R}$. Then, for $t \in \mathbb{R}$, the orthogonal projection $P_{\mathcal{L}_{Y_C}} Y_a(t)$ is equal to*

$$P_{\mathcal{L}_{Y_C}} Y_a(t) = \int_{-\infty}^{\infty} e^{i\lambda t} f_{Y_a Y_C}(\lambda) f_{Y_C Y_C}(\lambda)^{-1} Z_C(d\lambda),$$

where Z_C is the random spectral measure from the spectral representation (2.2) of $\mathcal{Y}_C$ and $\mathcal{L}_{Y_C}$ is defined in (3.1). Furthermore, $P_{\mathcal{L}_{Y_C}} Y_a(t)$ is the solution to the optimisation problem

$$\min_{\varphi_{a|C}} \mathbb{E}\left[\left| Y_a(t) - \int_{-\infty}^{\infty} e^{i\lambda t} \varphi_{a|C}(\lambda)^\top Z_C(d\lambda) \right|^2\right], \tag{3.3}$$

where $\varphi_{a|C} : \mathbb{R} \to \mathbb{C}^{|C|}$ is measurable and $\int_{-\infty}^{\infty} \varphi_{a|C}(\lambda)^\top f_{Y_C Y_C}(\lambda) \overline{\varphi_{a|C}(\lambda)} d\lambda < \infty$. Finally, $P_{\mathcal{L}_{Y_C}} Y_A(t) = (P_{\mathcal{L}_{Y_C}} Y_a(t))_{a \in A}$ can be calculated component-wise.

Remark 3.4 The choice of the term partial correlation relation is inspired by the partial correlation relation for discrete-time stationary processes in Brillinger (2001) and Dahlhaus (2000). However, the discrete-time concept is motivated by an optimisation problem similar to (1.1) (Brillinger 2001, Theorem 8.3.1, Dahlhaus 2000, relation (2.1) and Definition 2.1). To see the correspondence to our optimisation problem (3.3), suppose that the function $\varphi_{a|C}$ in (3.3) is the Fourier transform of an integrable function $d_{a|C}$, i.e., $\int_{-\infty}^{\infty} ||d_{a|C}(t)||dt < \infty$, such that

$$\varphi_{a|C}(\lambda) = \int_{-\infty}^{\infty} e^{-i\lambda t} d_{a|C}(t)dt. \tag{3.4}$$

Then, for $t \in \mathbb{R}$, Rozanov (1967), I, Example 8.3, provides

$$\int_{-\infty}^{\infty} e^{i\lambda t} \varphi_{a|C}(\lambda)^\top Z_C(d\lambda) = \int_{-\infty}^{\infty} d_{a|C}(t-s)^\top Y_C(s)ds. \tag{3.5}$$

With this representation, we have the similarity of our optimisation problem (3.3) to the discrete-time optimisation problem

$$\min_{d_{a|C}} \mathbb{E}\left[\left|X_a(n) - \sum_{k=-\infty}^{\infty} d_{a|C}(n-k)^\top X_C(k)\right|^2\right], \tag{3.6}$$

with $d_{a|C} : \mathbb{Z} \to \mathbb{C}^{|C|}$ being integrable and square integrable, and to (1.1). Given this parallelism, similarities with Dahlhaus (2000) are to be expected in various sections of this paper. Note also that Kleiber (2017) considers a time-domain optimisation problem corresponding to (3.6) for bivariate random fields with $|C| = 1$ which is applied in Dey et al. (2021) to construct multivariate graphical Gaussian processes using stitching that crafts cross-covariance functions from graphs.

Finally, we define the multivariate noise process

$$\begin{aligned} \varepsilon_{A|C}(t) &:= Y_A(t) - P_{\mathcal{L}_{Y_C}} Y_A(t) \\ &= Y_A(t) - \int_{-\infty}^{\infty} e^{i\lambda t} f_{Y_A Y_C}(\lambda) f_{Y_C Y_C}(\lambda)^{-1} Z_C(d\lambda), \quad t \in \mathbb{R}, \end{aligned} \tag{3.7}$$

with the following crucial properties.

Lemma 3.5 *Suppose that $f_{Y_C Y_C}(\lambda) > 0$ for $\lambda \in \mathbb{R}$. Then the noise processes $(\varepsilon_{A|C}(t))_{t\in\mathbb{R}}$ and $(\varepsilon_{B|C}(t))_{t\in\mathbb{R}}$ as defined in (3.7) are wide-sense stationary and stationary correlated with (cross-) spectral density function*

$$f_{\varepsilon_{A|C}\varepsilon_{B|C}}(\lambda) = f_{Y_A Y_B}(\lambda) - f_{Y_A Y_C}(\lambda) f_{Y_C Y_C}(\lambda)^{-1} f_{Y_C Y_B}(\lambda) \quad \text{for almost all } \lambda \in \mathbb{R},$$

and (cross-) covariance function

$$c_{\varepsilon_{A|C}\varepsilon_{B|C}}(t) = \int_{-\infty}^{\infty} e^{i\lambda t} \left(f_{Y_A Y_B}(\lambda) - f_{Y_A Y_C}(\lambda) f_{Y_C Y_C}(\lambda)^{-1} f_{Y_C Y_B}(\lambda)\right) d\lambda \quad \text{for all } t \in \mathbb{R}.$$

3.2 Characterisations of the partial correlation relation

In this section, we present several characterisations of the partial correlation relation. We start with simple characterisations in terms of the (cross-) covariance function, the (cross-) spectral density function, and the spectral coherence function of the noise processes $(\varepsilon_{A|C}(t))_{t\in\mathbb{R}}$ and $(\varepsilon_{B|C}(t))_{t\in\mathbb{R}}$, analogous to the discrete-time results in Remark 2.3 of Dahlhaus (2000).

Proposition 3.6 *Suppose that $f_{Y_C Y_C}(\lambda) > 0$ for $\lambda \in \mathbb{R}$. Then with the notation of Lemma 3.5 the following equivalences hold.*

$$\begin{aligned} \mathcal{Y}_A \perp\!\!\!\perp \mathcal{Y}_B \mid \mathcal{Y}_C \quad &\Leftrightarrow \quad c_{\varepsilon_{A|C}\varepsilon_{B|C}}(t) = 0_{|A|\times|B|} \quad \text{for all } t \in \mathbb{R}, \\ &\Leftrightarrow \quad f_{\varepsilon_{A|C}\varepsilon_{B|C}}(\lambda) = 0_{|A|\times|B|} \quad \text{for almost all } \lambda \in \mathbb{R}. \end{aligned}$$

In particular, these conditions imply that the spectral coherence function (2.4) of the noise processes $(\varepsilon_{A|C}(t))_{t\in\mathbb{R}}$ and $(\varepsilon_{B|C}(t))_{t\in\mathbb{R}}$ satisfies $R_{\varepsilon_{A|C}\varepsilon_{B|C}}(\lambda) = 0_{|A|\times|B|}$ for almost all $\lambda \in \mathbb{R}$. If $f_{\varepsilon_{A|C}\varepsilon_{A|C}}(\lambda) > 0$ and $f_{\varepsilon_{B|C}\varepsilon_{B|C}}(\lambda) > 0$ for $\lambda \in \mathbb{R}$, then the converse holds as well.

Remark 3.7 (a) The assumption that $f_{\varepsilon_{A|C}\varepsilon_{A|C}}(\lambda) > 0$ for $\lambda \in \mathbb{R}$ excludes the case $\varepsilon_{A|C}(t) = 0_{|A|}$ $\mathbb{P}$-a.s. for $t \in \mathbb{R}$, e.g., the case where $Y_a(t) \in \mathcal{L}_{Y_C}$ for $a \in A$. This can be explained as follows. If $\varepsilon_{A|C}(t) = 0_{|A|}$ $\mathbb{P}$-a.s. for $t \in \mathbb{R}$, then $f_{\varepsilon_{A|C}\varepsilon_{A|C}}(\lambda) = 0_{|A|} \in \mathbb{R}^{|A|\times|A|}$ is not positive definite for $\lambda \in \mathbb{R}$. We can therefore assume that $A \cap C = \emptyset$.

(b) For $A \cap C = \emptyset$, Bernstein (2009), Proposition 8.2.4 provides that $f_{Y_{A\cup C}Y_{A\cup C}}(\lambda) > 0$ if and only if $f_{Y_C Y_C}(\lambda) > 0$ and $f_{\varepsilon_{A|C}\varepsilon_{A|C}}(\lambda) > 0$.

(c) If $A \cap C = \emptyset$ and $f_{Y_{A\cup C}Y_{A\cup C}}(\lambda) > 0$ for $\lambda \in \mathbb{R}$, then $f_{\varepsilon_{A|C}\varepsilon_{A|C}}(\lambda) > 0$ and Proposition 3.6 results in $\mathcal{Y}_A \not\perp\!\!\!\perp \mathcal{Y}_A \mid \mathcal{Y}_C$. In the following, we always assume a sufficient condition for $f_{Y_{A\cup C}Y_{A\cup C}}(\lambda) > 0$, so we can also exclude the case $A \cap B = \emptyset$ from our analysis and assume throughout the remaining section that $A, B, C \subseteq V$ are disjoint.

Finally, we present a very simple characterisation of the partial correlation relation in terms of the inverse of the spectral density function of the underlying process $\mathcal{Y}_{A\cup B\cup C}$, which we denote, for $\lambda \in \mathbb{R}$, by

$$g_{Y_{A\cup B\cup C}Y_{A\cup B\cup C}}(\lambda) := f_{Y_{A\cup B\cup C}Y_{A\cup B\cup C}}(\lambda)^{-1}. \tag{3.8}$$

The corresponding discrete-time result is given in Theorem 2.4 of Dahlhaus (2000). Due to the importance of the result, we include the proof in the appendix.

Proposition 3.8 *Suppose that $A, B, C \subseteq V$ are disjoint and $f_{Y_{A\cup B\cup C}Y_{A\cup B\cup C}}(\lambda) > 0$ for $\lambda \in \mathbb{R}$ with $g_{Y_{A\cup B\cup C}Y_{A\cup B\cup C}}(\lambda) := f_{Y_{A\cup B\cup C}Y_{A\cup B\cup C}}(\lambda)^{-1}$ as in (3.8). Then the following equivalence holds.*

$$\mathcal{Y}_A \perp\!\!\!\perp \mathcal{Y}_B \mid \mathcal{Y}_C \quad \Leftrightarrow \quad \left[g_{Y_{A\cup B\cup C}Y_{A\cup B\cup C}}(\lambda)\right]_{AB} = 0_{|A|\times|B|} \quad \text{for almost all } \lambda \in \mathbb{R}.$$

The characterisation via the inverse of the spectral density function of $\mathcal{Y}_{A\cup B\cup C}$ can be used to explain the effect of an unobserved multivariate process $\mathcal{Y}_C$, a so-called confounder process. The following lemma introduces a relationship between the inverse of the spectral density function of a full process $\mathcal{Y}_V$ and the inverse of the spectral density function of a process $\mathcal{Y}_V$ reduced by a confounder process $\mathcal{Y}_C$. This result is the continuous-time counterpart to Dahlhaus (2000), Remark 2.5. Since it is a straightforward calculation, we omit the proof.

Lemma 3.9 *Suppose that $A, B, C \subseteq V$ are disjoint and $f_{Y_V Y_V}(\lambda) > 0$ for $\lambda \in \mathbb{R}$. Define $g_{Y_{V\setminus C} Y_{V\setminus C}}(\lambda) := f_{Y_{V\setminus C} Y_{V\setminus C}}(\lambda)^{-1}$ and $g_{Y_V Y_V}(\lambda) := f_{Y_V Y_V}(\lambda)^{-1}$ as in (3.8). Then, for $\lambda \in \mathbb{R}$, the equality*

$$\left[g_{Y_{V\setminus C} Y_{V\setminus C}}(\lambda)\right]_{AB} = \left[g_{Y_V Y_V}(\lambda)\right]_{AB} - \left[g_{Y_V Y_V}(\lambda)\right]_{AC} \left(\left[g_{Y_V Y_V}(\lambda)\right]_{CC}\right)^{-1} \left[g_{Y_V Y_V}(\lambda)\right]_{CB}$$

holds.

Remark 3.10 For an interpretation of this result (cf. Remark 2.5 in Dahlhaus 2000), i.e., the effect of a confounder process, we analyse the univariate case

$$\left[g_{Y_{V\setminus c} Y_{V\setminus c}}(\lambda)\right]_{ab} = \left[g_{Y_V Y_V}(\lambda)\right]_{ab} - \left[g_{Y_V Y_V}(\lambda)\right]_{ac} \left(\left[g_{Y_V Y_V}(\lambda)\right]_{cc}\right)^{-1} \left[g_{Y_V Y_V}(\lambda)\right]_{cb}.$$

This equation explains the relation between the partial correlation structure in the full process $\mathcal{Y}_V$ and the partial correlation structure in the reduced process $\mathcal{Y}_{V\setminus\{c\}}$: If $\mathcal{Y}_a$ and $\mathcal{Y}_b$ are partially uncorrelated given $\mathcal{Y}_{V\setminus\{a,b\}}$ ($\left[g_{Y_V Y_V}(\lambda)\right]_{ab} = 0$ for almost all $\lambda \in \mathbb{R}$), but there is a partial correlation between $\mathcal{Y}_a$ and $\mathcal{Y}_c$ given $\mathcal{Y}_{V\setminus\{a,c\}}$ and between $\mathcal{Y}_c$ and $\mathcal{Y}_b$ given $\mathcal{Y}_{V\setminus\{b,c\}}$ with $\left[g_{Y_V Y_V}(\lambda)\right]_{ac} \neq 0$ and $\left[g_{Y_V Y_V}(\lambda)\right]_{cb} \neq 0$ on some non-zero set, this causes a partial correlation between $\mathcal{Y}_a$ and $\mathcal{Y}_b$ given $\mathcal{Y}_{V\setminus\{a,b,c\}}$ ($\left[g_{Y_{V\setminus c} Y_{V\setminus c}}(\lambda)\right]_{ab} \neq 0$) in the reduced process $\mathcal{Y}_{V\setminus\{c\}}$.

Finally, we establish the main result of this section, namely that the partial correlation relation satisfies the important graphoid properties.

Proposition 3.11 *Suppose that $A, B, C, D \subseteq V$ are disjoint and $f_{Y_V Y_V}(\lambda) > 0$ for $\lambda \in \mathbb{R}$. Then the partial correlation relation defines a graphiod, i.e., it satisfies the following properties:*

(P1) Symmetry: $\mathcal{Y}_A \perp\!\!\!\perp \mathcal{Y}_B \mid \mathcal{Y}_C \Rightarrow \mathcal{Y}_B \perp\!\!\!\perp \mathcal{Y}_A \mid \mathcal{Y}_C$.
(P2) Decomposition: $\mathcal{Y}_A \perp\!\!\!\perp \mathcal{Y}_{B\cup C} \mid \mathcal{Y}_D \Rightarrow \mathcal{Y}_A \perp\!\!\!\perp \mathcal{Y}_B \mid \mathcal{Y}_D$.
(P3) Weak union: $\mathcal{Y}_A \perp\!\!\!\perp \mathcal{Y}_{B\cup C} \mid \mathcal{Y}_D \Rightarrow \mathcal{Y}_A \perp\!\!\!\perp \mathcal{Y}_B \mid \mathcal{Y}_{C\cup D}$.
(P4) Contraction: $\mathcal{Y}_A \perp\!\!\!\perp \mathcal{Y}_B \mid \mathcal{Y}_D$ and $\mathcal{Y}_A \perp\!\!\!\perp \mathcal{Y}_C \mid \mathcal{Y}_{B\cup D} \Rightarrow \mathcal{Y}_A \perp\!\!\!\perp \mathcal{Y}_{B\cup C} \mid \mathcal{Y}_D$.
(P5) Intersection: $\mathcal{Y}_A \perp\!\!\!\perp \mathcal{Y}_B \mid \mathcal{Y}_{C\cup D}$ and $\mathcal{Y}_A \perp\!\!\!\perp \mathcal{Y}_C \mid \mathcal{Y}_{B\cup D} \Rightarrow \mathcal{Y}_A \perp\!\!\!\perp \mathcal{Y}_{B\cup C} \mid \mathcal{Y}_D$.

Let us make some comments on the graphoid properties.

Remark 3.12 (a) The property (P1) is immediately clear. The properties (P2), (P3), and (P5) were already established by Dahlhaus (2000) in his Lemma 3.1 in discrete time, the proof carries over. For property (P4) we apply Proposition 3.8 and Lemma 3.9, which is done in the appendix.

(b) Graphoid properties are important properties of dependence relations such as the conditional orthogonality relation (Fasen-Hartmann and Schenk 2024a, b; Eichler 2007), the conditional independence relation (Eichler 2012), and the partial correlation relation (Dahlhaus 2000). Graphoid properties are not only of interest because of their interpretation (see, e.g., Lauritzen 2004; Pearl 1994), but they also simplify proofs. This is because the graphoid properties can be used directly, one does not have to work with the underlying definition of the relation. Most importantly, in the case of the partial correlation relation, it follows directly from the validity of the graphoid properties that the partial correlation graph satisfies all desired Markov properties of a graphical model (cf. Proposition 4.8 and Lauritzen 2004).

(c) Although partial correlation can be characterised by conditional orthogonality (Remark 3.2), the graphoid properties of Fasen-Hartmann and Schenk (2024a) are not directly applicable. The reason is the following: Due to Remark 3.2, $\mathcal{Y}_A \perp\perp \mathcal{Y}_{B\cup C} \mid \mathcal{Y}_D$ is equivalent to the conditional orthogonality relation $L_{Y_A}(t) \perp L_{Y_{B\cup C}}(t) \mid \mathcal{L}_{Y_D}$ for all $t \in \mathbb{R}$. Thus the weak union property of the conditional orthogonality relation (Fasen-Hartmann and Schenk 2024a, Lemma 2.2) gives $L_{Y_A}(t) \perp L_{Y_B}(t) \mid \text{cl}(\mathcal{L}_{Y_D} + L_{Y_C}(t))$ for all $t \in \mathbb{R}$, where cl denotes the closure of a set. This is not the same as $L_{Y_A}(t) \perp L_{Y_B}(t) \mid \mathcal{L}_{Y_{C\cup D}}$ for all $t \in \mathbb{R}$, i.e., $\mathcal{Y}_A \perp\perp \mathcal{Y}_B \mid \mathcal{Y}_{C\cup D}$. Similar problems occur with (P4) and (P5).

(d) The peculiarity of the continuous-time partial correlation relation is that the graphoid properties hold under the weak assumptions of wide-sense stationarity, zero expectation, mean-square continuity, and a positive definite spectral density function. For many dependence relations which form a graphoid, (P5) is quite difficult to verify and additional, possibly strict, assumptions are required, see, e.g. Eichler (2007), Proposition A.1, Fasen-Hartmann and Schenk (2024a), Assumption 1, Lauritzen (2004), Proposition 3.1, and Eichler (2011), Assumption S.

4 Partial correlation graphs

In Sect. 4.1, we introduce the partial correlation graph $G_{PC} = (V, E_{PC})$, an undirected graph. This graph serves as a simple visual representation of the partial correlation structure within the multivariate stochastic process $\mathcal{Y}_V$. Moreover, for the partial correlation graph, we derive edge characterisations and Markov properties. Finally, in Sect. 4.2, we compare and contrast the partial correlation graph to the orthogonality graph of Fasen-Hartmann and Schenk (2024a).

4.1 Partial correlation graphs and Markov properties

Our approach to visualising the partial correlation structure between the components of the multivariate process $\mathcal{Y}_V$ in the graphical model $G_{PC} = (V, E_{PC})$ is as follows: Each component $\mathcal{Y}_a$, $a \in V$, is represented by a vertex. We then define a missing edge $a \text{ --- } b \notin E_{PC}$ if and only if the components $\mathcal{Y}_a$ and $\mathcal{Y}_b$ are uncorrelated given the linear information provided by $\mathcal{Y}_{V\setminus\{a,b\}}$. As the relation $\mathcal{Y}_a \perp\!\!\!\perp \mathcal{Y}_b \mid \mathcal{Y}_{V\setminus\{a,b\}}$ is symmetric, we use undirected edges in G_{PC}. This leads to the following definition of the partial correlation graph.

Definition 4.1 Suppose that $\mathcal{Y}_V$ is wide-sense stationary with expectation zero, mean-square continuous, and has a spectral density function with $f_{Y_V Y_V}(\lambda) > 0$ for $\lambda \in \mathbb{R}$. Let $V = \{1, \ldots, k\}$ be the vertices and define the edges E_{PC} for $a, b \in V$ with $a \neq b$ as

$$a \text{ --- } b \notin E_{PC} \quad \Leftrightarrow \quad \mathcal{Y}_a \perp\!\!\!\perp \mathcal{Y}_b \mid \mathcal{Y}_{V\setminus\{a,b\}}.$$

Then $G_{PC} = (V, E_{PC})$ is called *partial correlation graph* for $\mathcal{Y}_V$.

Remark 4.2 (a) The name partial correlation graph is based on the partial correlation relation.

(b) For the definition of G_{PC} it is not necessary to require that $f_{Y_V Y_V}(\lambda) > 0$, but it is sufficient that $f_{Y_{V\setminus\{a,b\}} Y_{V\setminus\{a,b\}}}(\lambda) > 0$ for all $a, b \in V$. However, $f_{Y_V Y_V}(\lambda) > 0$ is essential for the graphoid properties and thus for the Markov properties of the partial correlation graph in Proposition 4.8. Note that in general $f_{Y_V Y_V}(\lambda) \geq 0$ holds (cf. Lemma 2.1), so $f_{Y_V Y_V}(\lambda) > 0$ is only a mild assumption.

(c) A direct consequence of Remark 3.7(c) is that for $a \in V$ we would always have $a \text{ --- } a \in E_{PC}$. Since such self-loops do not help to visualise the partial correlation structure and do not change the properties of the graph, we omit them for the sake of simplicity.

Lemma 4.3 *Suppose that $G_{PC} = (V, E_{PC})$ is the partial correlation graph for $\mathcal{Y}_V$. Then, for $a, b \in V$ with $a \neq b$, the following equivalences hold.*

$$\begin{aligned} a \text{ --- } b \notin E_{PC} \quad &\Leftrightarrow \quad c_{\varepsilon_{a|V\setminus\{a,b\}}\varepsilon_{b|V\setminus\{a,b\}}}(t) = 0 \quad \textit{for all } t \in \mathbb{R}, \\ &\Leftrightarrow \quad f_{\varepsilon_{a|V\setminus\{a,b\}}\varepsilon_{b|V\setminus\{a,b\}}}(\lambda) = 0 \quad \textit{for almost all } \lambda \in \mathbb{R}, \\ &\Leftrightarrow \quad R_{\varepsilon_{a|V\setminus\{a,b\}}\varepsilon_{b|V\setminus\{a,b\}}}(\lambda) = 0 \quad \textit{for almost all } \lambda \in \mathbb{R}. \end{aligned}$$

Note that the spectral coherence function of the noise processes $(\varepsilon_{a|V\setminus\{a,b\}}(t))_{t\in\mathbb{R}}$ and $(\varepsilon_{b|V\setminus\{a,b\}}(t))_{t\in\mathbb{R}}$ is well-defined, since $f_{Y_V Y_V}(\lambda) > 0$ by assumption and thus, $f_{\varepsilon_{a|V\setminus\{a,b\}}\varepsilon_{a|V\setminus\{a,b\}}}(\lambda) > 0$ and $f_{\varepsilon_{b|V\setminus\{a,b\}}\varepsilon_{b|V\setminus\{a,b\}}}(\lambda) > 0$, which results in a non-vanishing denominator.

In addition, Proposition 3.8 gives the key representation using the inverse of the spectral density function, the corresponding edge characterisation for time series in discrete time is established in Dahlhaus (2000), Theorem 2.4.

Proposition 4.4 *Suppose that $G_{PC} = (V, E_{PC})$ is the partial correlation graph for $\mathcal{Y}_V$. Then, for $a, b \in V$ with $a \neq b$, the spectral coherence function of the noise processes $(\varepsilon_{a|V\setminus\{a,b\}}(t))_{t\in\mathbb{R}}$ and $(\varepsilon_{b|V\setminus\{a,b\}}(t))_{t\in\mathbb{R}}$ satisfies*

$$R_{\varepsilon_{a|V\setminus\{a,b\}}\varepsilon_{b|V\setminus\{a,b\}}}(\lambda) = -\frac{\left[g_{Y_V Y_V}(\lambda)\right]_{ab}}{\left(\left[g_{Y_V Y_V}(\lambda)\right]_{aa}\left[g_{Y_V Y_V}(\lambda)\right]_{bb}\right)^{1/2}}, \quad \lambda \in \mathbb{R}, \tag{4.1}$$

where $g_{Y_V Y_V}(\lambda) = f_{Y_V Y_V}(\lambda)^{-1}$ as in (3.8). Furthermore,

$$a \text{ --- } b \notin E_{PC} \quad \Leftrightarrow \quad \left[g_{Y_V Y_V}(\lambda)\right]_{ab} = 0 \quad \textit{for almost all } \lambda \in \mathbb{R}. \tag{4.2}$$

Remark 4.5 A significant advantage of (4.2) over other characterisations in Lemma 4.3 is that it is computationally inexpensive. One only needs to know the spectral density function and then perform a singular matrix inversion to obtain all the edges in the graph simultaneously. Furthermore, the relation (4.1) even gives us a simple measure for the strength of the connection between two components a and b given the environment $V \setminus \{a, b\}$.

Remark 4.6 Dahlhaus (2000) remarks that the partial correlation graph can be compared to the so-called concentration graph. The concentration graph for a random vector $X \in \mathbb{R}^k$ with $\mathbb{E}\|X\|^2 < \infty$, $\mathbb{E}(X) = 0_k$, and $\Sigma_X := \mathbb{E}[XX^\top] > 0$ is defined as follows. Let $V = \{1, \ldots, k\}$ be the vertices and define the edges E_{CO} for $a, b \in V$ with $a \neq b$ as

$$\begin{aligned} a \text{ --- } b \notin E_{CO} \Leftrightarrow \quad & \left[\Sigma_X^{-1}\right]_{ab} = 0 \\ \Leftrightarrow \quad & [\Sigma_X]_{ab} - [\Sigma_X]_{aV\setminus\{a,b\}}\left([\Sigma_X]_{V\setminus\{a,b\}V\setminus\{a,b\}}\right)^{-1}[\Sigma_X]_{V\setminus\{a,b\}b} = 0. \end{aligned}$$

Then $G_{CO} = (V, E_{CO})$ is called *concentration graph* of X. The concentration graph G_{CO} describes the sparsity pattern of the inverse covariance matrix Σ_X^{-1} of X, also called *concentration matrix* or *precision matrix*, hence the name concentration graph.

The definition of the concentration graph illustrates why the partial correlation graph for stochastic processes is a generalisation of the concentration graph for random vectors. A missing edge $a \text{ --- } b \notin E_{CO}$ in the concentration graph for X reflects that X_a and X_b are partially uncorrelated given $X_{V\setminus\{a,b\}}$. Similarly, $a \text{ --- } b \notin E_{PC}$ in the partial correlation graph for $\mathcal{Y}_V$ means that the stochastic processes $\mathcal{Y}_a$ and $\mathcal{Y}_b$ are partially uncorrelated given $\mathcal{Y}_{V\setminus\{a,b\}}$. Finally, the edges in the partial correlation graph are characterised by the inverse of the spectral density function, which can be seen as a generalisation of the inverse of a covariance matrix. Indeed, for an independent and identically distributed sequence of random vectors with distribution X, the spectral density is equal to $(2\pi)^{-1}\Sigma_X$. Note that the concentration graph is usually defined only for multivariate Gaussian random vectors (Maathuis et al. 2019, p. 218) and not for general random vectors, but this definition is a natural generalisation. For Gaussian random vectors, however, conditional independence and conditional orthogonality coincide, so missing edges even correspond to conditional independence relations (Maathuis et al. 2019, Corollary 9.1.2).

To conclude this section, we establish the Markov properties of G_{PC}. To do this, we first provide some terminology.

Definition 4.7 For $a \in V$ define $\mathrm{ne}(a) = \{b \in V \mid a \,—\, b \in E_{PC}\}$ as the set of *neighbours* of a. A *path* of length n from a vertex a to a vertex b is a sequence $\alpha_0 = a, \alpha_1, \ldots, \alpha_n = b$ of vertices such that $\alpha_{i-1} \,—\, \alpha_i \in E_{PC}$ for $i = 1, \ldots, n$. For $A, B, C \subseteq V$, we say that C *separates* A and B if every path from an element of A to an element of B contains at least one vertex from the separating set C. We write

$$A \bowtie B \mid C.$$

Now the partial correlation graph satisfies the following Markov properties.

Proposition 4.8 *Suppose that $G_{PC} = (V, E_{PC})$ is the partial correlation graph for $\mathcal{Y}_V$. Then $\mathcal{Y}_V$ satisfies*

(P) the pairwise Markov property with respect to G_{PC}, i.e., for $a, b \in V$ with $a \neq b$,

$$a \,—\, b \notin E_{PC} \quad \Rightarrow \quad \mathcal{Y}_a \perp\!\!\!\perp \mathcal{Y}_b \mid \mathcal{Y}_{V \setminus \{a,b\}},$$

(L) the local Markov property with respect to G_{PC}, i.e., for $a \in V$,

$$\mathcal{Y}_{V \setminus (ne(a) \cup \{a\})} \perp\!\!\!\perp \mathcal{Y}_a \mid \mathcal{Y}_{ne(a)},$$

(G) the global Markov property with respect to G_{PC}, i.e., for disjoint $A, B, C \subseteq V$,

$$A \bowtie B \mid C \quad \Rightarrow \quad \mathcal{Y}_A \perp\!\!\!\perp \mathcal{Y}_B \mid \mathcal{Y}_C.$$

The pairwise Markov property holds by definition. Furthermore, the partial correlation relation defines a graphoid by Proposition 3.11. Thus, Lauritzen (2004) states in Theorem 3.7 that the pairwise, local, and global Markov properties are equivalent, so the local and global Markov properties are also valid. The global Markov property is powerful because it provides a graphical criterion for deciding when two subprocesses $\mathcal{Y}_A$ and $\mathcal{Y}_B$ are partially uncorrelated given a third subprocess $\mathcal{Y}_C$ that does not have to be $\mathcal{Y}_{V \setminus (A \cup B)}$. Most importantly, although the graph itself is defined only by pairwise partial correlation relations, we can obtain partial correlation relations between multivariate subprocesses given any subprocess through path analysis.

4.2 Partial correlation graphs and orthogonality graphs

In this section, we draw parallels between our partial correlation graph and the orthogonality graph of Fasen-Hartmann and Schenk (2024a). First, we introduce the orthogonality graph, using their edge characterisations (cf. Fasen-Hartmann and Schenk 2024a, Lemmas 3.2 and 4.2).

Definition 4.9 Suppose that $\mathcal{Y}_V$ is wide-sense stationary with expectation zero, mean-square continuous, purely non-deterministic, and has a spectral density function with $f_{Y_V Y_V}(\lambda) > 0$ for $\lambda \in \mathbb{R}$ that satisfies Assumption 1 of Fasen-Hartmann and Schenk

(2024a). For $t \in \mathbb{R}$ and $C \subseteq V$ the linear space $L_C(t)$ is defined as in (3.2) and further define the closed linear space

$$\mathcal{L}_{Y_C}(t) := \text{cl}\left\{\sum_{i=1}^{\ell}\sum_{c\in C}\gamma_{c,i}Y_c(t_i) : \gamma_{c,i}\in\mathbb{C},\ -\infty < t_1 \leq \cdots \leq t_\ell \leq t,\ \ell\in\mathbb{N}\right\}.$$

Let $V = \{1, \ldots, k\}$ be the vertices and define the edges E_{OG}, for $a, b \in V$, $a \neq b$, as

(i) $a \longrightarrow b \notin E_{OG} \Leftrightarrow$ $\mathcal{Y}_a$ is *Granger non-causal* for $\mathcal{Y}_b$ with respect to $\mathcal{Y}_V$ $\Leftrightarrow$ $L_{Y_b}(t+h) \perp \mathcal{L}_{Y_a}(t) \mid \mathcal{L}_{Y_{V\setminus\{a\}}}(t)\ \forall\, 0 \leq h \leq 1,\ t \in \mathbb{R}$,

(ii) $a \dashrightarrow b \notin E_{OG} \Leftrightarrow$ $\mathcal{Y}_a$ and $\mathcal{Y}_b$ are *contemporaneously uncorrelated* with respect to $\mathcal{Y}_V$ $\Leftrightarrow$ $L_{Y_a}(t+h) \perp L_{Y_b}(t+h') \mid \mathcal{L}_{Y_V}(t)\ \forall\, 0 \leq h, h' \leq 1,\ t \in \mathbb{R}$.

Then $G_{OG} = (V, E_{OG})$ is called *(mixed) orthogonality graph* for $\mathcal{Y}_V$. The index OG stands for *o*rthogonality *g*raph.

Remark 4.10 To highlight the differences between the undirected edges in the orthogonality graph and in the partial correlation graph, recall from Remark 3.2 that in the partial correlation graph, for $a, b \in V$ with $a \neq b$,

$$a \,—\, b \notin E_{PC} \quad \Leftrightarrow \quad L_{Y_a}(t) \perp L_{Y_b}(t) \mid \mathcal{L}_{Y_{V\setminus\{a,b\}}} \quad \forall\, t \in \mathbb{R}.$$

The concept of contemporaneous uncorrelatedness in Definition 4.9(ii) differs from zero partial correlation in two ways. First, for zero partial correlation, we always project on the linear space of the whole process $\mathcal{Y}_{V\setminus\{a,b\}} = (Y_{V\setminus\{a,b\}}(t))_{t\in\mathbb{R}}$, whereas, for contemporaneous uncorrelatedness, we project on the past $(Y_V(s))_{s\leq t}$. Second, in the case of contemporaneous uncorrelatedness, the correlation has to be considered not only at identical time points but also at mixed time points one time step into the future.

Despite the differences between the two concepts (which is also confirmed by the analysis of MCAR processes in Example 5.10), there are relationships between paths in the mixed orthogonality graph and edges in the partial correlation graph. To show these relations, we first provide the concept of m-separation (cf. Eichler 2007), which is an extension of separation for undirected graphs (cf. Definition 4.7) to mixed graphs.

Definition 4.11 In a mixed graph $G = (V, E)$ an intermediate vertex c on a path π is said to be a *collider*, if the edges preceding and succeeding c on the path both have an arrowhead or a dashed tail at c, i.e., $\longrightarrow c \longleftarrow$, $\longrightarrow c$ - - - , - - - $c \longleftarrow$, or - - - c - - - . A path π between vertices a and b is said to be *m-connecting* given a set C if

(a) every non-collider on π is not in C, and
(b) every collider on π is in C,

otherwise we say π is *m-blocked* given C. If all paths between a and b are m-blocked given C, then a and b are said to be *m-separated* given C, denoted by

$$\{a\} \bowtie_m \{b\} \mid C \quad [G].$$

The first relation between the orthogonality graph and the partial correlation graph follows almost directly from the global AMP Markov property of the orthogonality graph, which Fasen-Hartmann and Schenk (2024a) establish in Theorem 5.15.

Lemma 4.12 *Suppose that $G_{PC} = (V, E_{PC})$ is the partial correlation graph and $G_{OG} = (V, E_{OG})$ is the orthogonality graph for $\mathcal{Y}_V$. Then, for $a, b \in V$ with $a \neq b$, the following implication holds.*

$$\{a\} \bowtie_m \{b\} \mid V \setminus \{a, b\} \quad [G_{OG}] \quad \Rightarrow \quad a \text{ --- } b \notin E_{PC}.$$

The advantage of this result is that the concept of m-separation has several different characterisations in the literature, leading to several sufficient criteria for $a \text{ --- } b \notin E_{PC}$. One approach is to build an undirected graph from the mixed graph G_{OG}, using augmentation. The resulting augmented graph can then be related to the undirected partial correlation graph. The augmented graph is constructed as follows (Richardson 2003, p. 148).

Definition 4.13 Let $G = (V, E)$ be a mixed graph. Two vertices a and b are said to be *collider connected* if they are connected by a pure collider path, which is a path on which every intermediate vertex is a collider. Then the undirected *augmented graph* $G^a = (V, E^a)$ is derived from $G = (V, E)$ via

$$a \text{ --- } b \notin E^a \quad \Leftrightarrow \quad a \text{ and } b \text{ are not collider connected in } G.$$

Note that every single edge is trivially considered to be a collider path. Thus, every directed and undirected edge in the orthogonality graph corresponds to an undirected edge in the augmented orthogonality graph, implicating that the augmented orthogonality graph has more edges than the orthogonality graph.

Lemma 4.14 *Suppose that $G_{PC} = (V, E_{PC})$ is the partial correlation graph, $G_{OG} = (V, E_{OG})$ is the orthogonality graph, and $G^a_{OG} = (V, E^a_{OG})$ is the augmented orthogonality graph for $\mathcal{Y}_V$. Furthermore, we use the notation of Definitions 4.7 and 4.11. For $a, b \in V$ with $a \neq b$, the following equivalences hold.*

$$\begin{aligned} a \text{ --- } b \notin E^a_{OG} \quad &\Leftrightarrow \quad dis\,(a \cup ch(a)) \cap dis\,(b \cup ch(b)) = \emptyset \ \text{ in } G_{OG}, &(4.3)\\ &\Leftrightarrow \quad \{a\} \bowtie \{b\} \mid V \setminus \{a, b\} \quad [G^a_{OG}], &(4.4)\\ &\Leftrightarrow \quad \{a\} \bowtie_m \{b\} \mid V \setminus \{a, b\} \quad [G_{OG}]. \end{aligned}$$

Here, $ch(a) = \{v \in V | a \longrightarrow v \in E_{OG}\}$, $dis(a) = \{v \in V | v \text{ - - - } \cdots \text{ - - - } a \text{ or } v = a\}$, and $dis(A) = \bigcup_{a \in A} dis(a)$. In particular, the statements imply that $a \text{ --- } b \notin E_{PC}$, i.e.,

$$E_{PC} \subseteq E^a_{OG}.$$

Lemma 4.14 gives us several possibilities to make statements about the partial correlation graph from the orthogonality graph. On the one hand, (4.3) is particularly useful, since we can work with the original mixed orthogonality graph G_{OG} and it is easy to implement algorithmically (Eichler 2011), which is not straightforward for the

m-separation criterion from Lemma 4.12. On the other hand, (4.4) is of interest, as it uses the classical separation $\bowtie$ in an undirected graph, which is another common way to define global Markov properties in mixed graphs. Finally, the inclusion property $E_{PC} \subseteq E^a_{OG}$ gives us a simple connection between the edges in both graphs.

Besides the orthogonality graph $G_{OG} = (V, E_{OG})$, Fasen-Hartmann and Schenk (2024a) also introduce the local orthogonality graph $G^0_{OG} = (V, E^0_{OG})$, a mixed graph with $E^0_{OG} \subseteq E_{OG}$. For the augmented local orthogonality graph obviously $E^{a,0}_{OG} \subseteq E^a_{OG}$ holds, but in general the statement $E_{PC} \subseteq E^{a,0}_{OG}$ is probably not possible, since we do not have a global AMP Markov property in the local orthogonality graph. However, if we restrict to MCAR(p) processes, we derive this subset relation in Sect. 5.3.

5 Partial correlation graphs for MCAR processes

In the following, we construct the partial correlation graph for Lévy-driven multivariate continuous-time autoregressive (MCAR) processes to illustrate the partial correlation structure within this important and versatile class of processes. Therefore, in Sect. 5.1, we give a brief introduction to MCAR processes. Subsequently, in Sect. 5.2, we ensure that the partial correlation graph for MCAR processes is well defined and establish the latter. We also provide some edge characterisations by model parameters along with comparisons to the literature. Moving on to Sect. 5.3, we study relations between the partial correlation graph and the (local) orthogonality graph, highlighting both similarities and differences. Finally, in Sect. 5.4, we motivate some methods to estimate the edges in the partial correlation graph for MCAR processes.

5.1 MCAR processes

Early works on univariate and multivariate CAR processes and the more general continuous-time autoregressive moving average (CARMA) processes include those of Doob (1944, 1953), Harvey and Stock (1985a, b), Harvey and Stock (1989); Bergstrom (1997). Since then, these processes have enjoyed great popularity and have stimulated a considerable amount of research in recent years (cf. Brockwell 2014, and Brockwell and Lindner 2024). The following definition of a Lévy-driven MCAR process goes back to Marquardt and Stelzer (2007), Definition 3.20.

Definition 5.1 Let $L = (L(t))_{t\in\mathbb{R}}$ be a Lévy process satisfying $\mathbb{E}[L(1)] = 0_k$ and $\mathbb{E}\|L(1)\|^2 < \infty$ with $\Sigma_L = \mathbb{E}[L(1)L(1)^\top]$. Suppose that $A_1, A_2, \ldots, A_p \in \mathbb{R}^{k\times k}$ and define the matrices

$$\mathbf{A} = \begin{pmatrix} 0_k & I_k & 0_k & \cdots & 0_k \\ 0_k & 0_k & I_k & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0_k \\ 0_k & \cdots & \cdots & 0_k & I_k \\ -A_p & -A_{p-1} & \cdots & \cdots & -A_1 \end{pmatrix} \in \mathbb{R}^{kp\times kp}, \quad \mathbf{B} = \begin{pmatrix} 0_k \\ \vdots \\ 0_k \\ I_k \end{pmatrix} \in \mathbb{R}^{kp\times k},$$
$$\mathbf{C} = \begin{pmatrix} I_k & 0_k & \cdots & 0_k \end{pmatrix} \in \mathbb{R}^{k\times kp}.$$

Finally, suppose $\sigma(\mathbf{A}) \subseteq (-\infty, 0) + i\mathbb{R}$ and that $\mathcal{X} = (X(t))_{t\in\mathbb{R}}$ is the unique kp-dimensional causal strictly stationary solution of the state equation

$$dX(t) = \mathbf{A}X(t)dt + \mathbf{B}dL(t).$$

Then the output process $\mathcal{Y}_V = (Y_V(t))_{t\in\mathbb{R}}$ given by

$$Y_V(t) = \mathbf{C}X(t)$$

is called a *(causal) multivariate continuous-time autoregressive process of order* p, or MCAR(p) process for short.

The MCAR process is the continuous-time counterpart of the well-known discrete-time vector autoregressive (VAR) process. For this correspondence, the idea is that a k-dimensional MCAR(p) $(p \geq 1)$ process $\mathcal{Y}_V$ is the solution to the stochastic differential equation

$$P(D)Y_V(t) = DL(t)$$

where D is the differential operator with respect to t, and

$$P(z) = I_k z^p + A_1 z^{p-1} + \ldots + A_p, \quad z \in \mathbb{C}, \tag{5.1}$$

is the autoregressive (AR) polynomial. However, a Lévy process is not differentiable, so this is not a formal definition of an MCAR process.

The properties of MCAR processes relevant to this paper are summarised below. For additional information we refer to Marquardt and Stelzer (2007) and Schlemm and Stelzer (2012a, b).

Remark 5.2 (a) Since the input process $\mathcal{X}$ is strictly stationary, the MCAR process $\mathcal{Y}_V$ is also strictly stationary. Furthermore, given the finite second moments of the Lévy process, both $\mathcal{X}$ and $\mathcal{Y}_V$ also have finite second moments. Thus, of course, $\mathcal{X}$ and $\mathcal{Y}_V$ are also wide-sense stationary.

(b) The covariance function of the input process $\mathcal{X}$ satisfies

$$c_{XX}(t) = \overline{c_{XX}(-t)}^{\top} = e^{\mathbf{A}t}\Gamma(0), \quad t \geq 0, \quad \text{where} \quad \Gamma(0) = \int_0^\infty e^{\mathbf{A}u}\mathbf{B}\Sigma_L\mathbf{B}^\top e^{\mathbf{A}^\top u}du.$$

The covariance function of the MCAR process $\mathcal{Y}_V$ is then determined as

$$c_{Y_V Y_V}(t) = \mathbf{C} c_{XX}(t) \mathbf{C}^\top, \quad t \in \mathbb{R}.$$

(c) Given that $\sigma(\mathbf{A}) \subseteq (-\infty, 0) + i\mathbb{R}$, it follows that $c_{Y_V Y_V}(t)$ decreases exponentially fast as $t \to \pm\infty$ and $\lim_{t\to 0} c_{Y_V Y_V}(t) = c_{Y_V Y_V}(0)$, so $\mathcal{Y}_V$ is mean-square continuous due to (2.1).
(d) The MCAR(1) process is also known as Ornstein-Uhlenbeck process and in this case, we have $\mathbf{A} = -A_1$ and $\mathbf{B} = \mathbf{C} = I_k$. Furthermore, Gaussian MCAR processes and Gaussian Ornstein-Uhlenbeck processes, where the Brownian motion is the driving Lévy process, are special cases.

5.2 Definition of the partial correlation graph for a MCAR process

We introduce the partial correlation graph for MCAR processes. From Remark 5.2 we already know that the MCAR process is wide-sense stationary with expectation zero and mean-square continuous. Furthermore, the spectral density function is (Marquardt and Stelzer 2007, Eq. (3.43))

$$f_{Y_V Y_V}(\lambda) = \frac{1}{2\pi} P(i\lambda)^{-1} \Sigma_L \left(P(-i\lambda)^{-1} \right)^\top, \quad \lambda \in \mathbb{R},$$

where the AR polynomial P is defined in Eq. (5.1). For the well-definedness of the partial correlation graph we then only need to ensure that $f_{Y_V Y_V}(\lambda) > 0$ for $\lambda \in \mathbb{R}$. But this condition is already met when $\Sigma_L > 0$ and $\sigma(\mathbf{A}) \subseteq (-\infty, 0) + i\mathbb{R}$. Then the inverse spectral density function has the representation

$$g_{Y_V Y_V}(\lambda) = 2\pi P(-i\lambda)^\top \Sigma_L^{-1} P(i\lambda), \quad \lambda \in \mathbb{R}.$$

By Definition 4.1, Propositions 4.4 and 4.8 we then obtain the following result.

Proposition 5.3 *Suppose $\mathcal{Y}_V$ is a causal MCAR(p) process with $\Sigma_L > 0$. Let $V = \{1, \dots, k\}$ be the vertices and define the edges E_{PC}, for $a, b \in V$ with $a \neq b$, via*

$$a \,—\, b \notin E_{PC} \quad \Leftrightarrow \quad \mathcal{Y}_a \perp\!\!\!\perp \mathcal{Y}_b \mid \mathcal{Y}_{V\setminus\{a,b\}} \quad \Leftrightarrow \quad \left[P(-i\lambda)^\top \Sigma_L^{-1} P(i\lambda) \right]_{ab} = 0 \quad \forall \lambda \in \mathbb{R}.$$

Then the partial correlation graph $G_{PC} = (V, E_{PC})$ for the MCAR process $\mathcal{Y}_V$ is well defined and satisfies the pairwise, local, and global Markov properties.

Note that partial correlation graphs can be defined for more general state space models, but we find that MCAR processes are sufficient for our illustrative purposes. Note also that for the MCAR process, $g_{Y_V Y_V}(\lambda) = f_{Y_V Y_V}(\lambda)^{-1}$ has a very simple representation, it is a matrix polynomial. As a result, we can give the following edge characterisation based on the coefficients of the matrices $A_1, A_2, \dots, A_p$ of the AR polynomial, and the covariance matrix Σ_L of the driving Lévy process.

Proposition 5.4 *Suppose that $G_{PC} = (V, E_{PC})$ is the partial correlation graph for the causal MCAR(p) process $\mathcal{Y}_V$ with AR polynomial P given by (5.1), where we define $A_0 := I_k$. For $a, b \in V$ with $a \neq b$, we obtain the edge characterisation*

$$a \text{ --- } b \notin E_{PC} \quad \Leftrightarrow \quad \sum_{m=0\vee n-p}^{n\wedge p} (-1)^m \left[A_{p-m}^\top \Sigma_L^{-1} A_{p-n+m}\right]_{ab} = 0 \;\; \text{for } n = 0, \ldots, 2p.$$

This characterisation is reduced in the following cases.

(i) Suppose $\Sigma_L = \sigma^2 I_k > 0$. Then

$$a \text{ --- } b \notin E_{PC} \quad \Leftrightarrow \quad \sum_{m=0\vee n-p}^{n\wedge p} (-1)^m \left[A_{p-m}^\top A_{p-n+m}\right]_{ab} = 0 \;\; \text{for } n = 0, \ldots, 2p.$$

(ii) Suppose A_j is a diagonal matrix for $j = 1, \ldots, p$. Then

$$a \text{ --- } b \notin E_{PC} \quad \Leftrightarrow \quad \left[\Sigma_L^{-1}\right]_{ab} = 0.$$

Remark 5.5 A consequence of Proposition 5.4(ii) is that for any undirected graph $G = (V, E)$ and any $p \in \mathbb{N}$, there exists an MCAR(p) process with partial correlation graph $G_{PC} = G$. Indeed, we can define

$$\left[\Sigma_L^{-1}\right]_{ab} = \begin{cases} k, & \text{if } a = b, \\ 1, & \text{if } a \neq b \text{ and } a \text{ --- } b \in E, \\ 0, & \text{if } a \neq b \text{ and } a \text{ --- } b \notin E, \end{cases}$$

and $A_m = \binom{p}{m} I_k \in \mathbb{R}^{k\times k}$ for $m = 0, \ldots, p$. Consequently, $\sigma(\mathbf{A}) = \{-1\} \subseteq (-\infty, 0) + i\mathbb{R}$ and Σ_L^{-1} is strictly diagonally dominant, i.e., positive definite. Σ_L is also positive definite and there exists a Lévy process with this covariance matrix. Due to Proposition 5.4(ii), the resulting k-dimensional MCAR(p) process $\mathcal{Y}_V$ generates a partial correlation graph $G_{PC} = (V, E_{PC})$, which is identical to the undirected graph $G = (V, E)$. This is a major advantage over the orthogonality graph in Fasen-Hartmann and Schenk (2024a), where it is not clear if any graph can be constructed by a continuous-time process.

Remark 5.6 The edge characterisations for MCAR(p) processes in Proposition 5.4 are, as might be expected, similar to the edge characterisations for VAR(p) processes in Dahlhaus (2000), Example 2.2. Suppose that the AR coefficient matrices of the VAR(p) process are denoted by $\Phi_m \in \mathbb{R}^{k\times k}$, $m = 1, \ldots, p$, $\Phi_0 = -I_k$, and $0 < \Sigma_\varepsilon \in \mathbb{R}^{k\times k}$ denotes the covariance matrix of the white noise process. Then Dahlhaus (2000) states that in the partial correlation graph $G_{PC}^d = (V, E_{PC}^d)$ for the VAR(p) process we have

$$a \text{ --- } b \notin E_{PC}^d \quad \Leftrightarrow \quad \sum_{m=0\vee n-p}^{p\wedge n} \left[\Phi_m^\top \Sigma_\varepsilon^{-1} \Phi_{m-n+p}\right]_{ab} = 0 \quad \text{for } n = 0, \ldots, 2p.$$

Both characterisations of the continuous-time and the discrete-time multivariate AR processes match exactly if we neglect the factor $(-1)^m$. This small difference is due to the fact that the spectral density of the continuous-time model is defined by the AR polynomial at $\pm i\lambda$ whereas, in the discrete-time model, it is the AR polynomial at $e^{\pm i\lambda}$.

Furthermore, the following sufficient condition for an edge between a and b in the partial correlation graph can be obtained by setting $n = 2p$ in Proposition 5.4.

Lemma 5.7 *Suppose that $G_{PC} = (V, E_{PC})$ is the partial correlation graph for the causal MCAR(p) process $\mathcal{Y}_V$. For $a, b \in V$ with $a \neq b$, the following implication holds.*

$$a \text{ --- } b \notin E_{PC} \quad \Rightarrow \quad \left[\Sigma_L^{-1}\right]_{ab} = 0.$$

Remark 5.8 Note that Σ_L^{-1} is the concentration matrix of the random vector $L(1)$, so it defines the concentration graph $G_{CO} = (V, E_{CO})$ of $L(1)$. Lemma 5.7 therefore gives the subset relation $E_{CO} \subseteq E_{PC}$. In other words, the partial correlation of the random variables $L_a(1)$ and $L_b(1)$ given $L_{V\setminus\{a,b\}}(1)$ imply an edge in the partial correlation graph of the continuous-time process $\mathcal{Y}_V$, i.e., the stochastic processes $\mathcal{Y}_a$ and $\mathcal{Y}_b$ are partially correlated given the process $\mathcal{Y}_{V\setminus\{a,b\}}$. If we additionally assume that A_m, $m = 1, \ldots, p$, are diagonal, then Proposition 5.4(ii) even gives $E_{CO} = E_{PC}$.

Finally, for a visualisation of the previous edge characterisations in Propositions 5.3 and 5.4, we present an example.

Example 5.9 Suppose that $\mathcal{Y}_V$ is a 4-dimensional Ornstein-Uhlenbeck process with $\Sigma_L = I_4$ and

$$\mathbf{A} = \begin{pmatrix} -2 & 0 & 1 & 1 \\ 0 & -2 & -1 & -1 \\ -1 & -1 & -2 & -1 \\ 1 & -1 & -1 & -2 \end{pmatrix}.$$

Then a simple calculation yields $\sigma(\mathbf{A}) = \{-1, -1, -2, -4\} \subseteq (-\infty, 0) + i\mathbb{R}$. For an Ornstein-Uhlenbeck process $\mathcal{Y}_V$ the inverse spectral density function is simplified to $g_{Y_V Y_V}(\lambda) = 2\pi(-i\lambda I_k - \mathbf{A}^\top)\Sigma_L^{-1}(i\lambda I_k - \mathbf{A})$ for $\lambda \in \mathbb{R}$ and we obtain

$$g_{Y_V Y_V}(\lambda) = 2\pi \begin{pmatrix} \lambda^2 + 6 & 0 & 2i\lambda - 1 & -3 \\ 0 & \lambda^2 + 6 & 5 & 5 \\ -2i\lambda - 1 & 5 & \lambda^2 + 7 & 6 \\ -3 & 5 & 6 & \lambda^2 + 7 \end{pmatrix}.$$

The corresponding partial correlation graph $G_{PC} = (V, E_{PC})$ is then given in Fig. 1.

Furthermore, for an Ornstein–Uhlenbeck process, the edge characterisation in Proposition 5.4(i) is simplified to

$$a \text{ --- } b \notin E_{PC} \quad \Leftrightarrow \quad [\mathbf{A}]_{ba} - [\mathbf{A}]_{ab} = 0, \quad \left[\mathbf{A}^\top \mathbf{A}\right]_{ab} = 0. \tag{5.2}$$

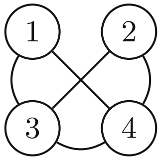

Fig. 1 Partial correlation graph for Example 5.9

Of course, this relation also provides the edges in Fig. 1.

To summarise, Example 5.9 highlights once more the main advantage of the characterisation in Proposition 5.3, which is the ability to obtain all edges simultaneously through the inverse spectral density function.

5.3 Partial correlation graphs and (local) orthogonality graphs

In this section, we relate the partial correlation graph to the orthogonality graph and the local orthogonality graph of Fasen-Hartmann and Schenk (2024a), which can be seen as a continuation of Sect. 4.2. Let us start with the relations between the partial correlation graph and the *orthogonality graph*. In the comparison in Sect. 4.2, we suspected that, in general, there is no direct relationship between the edges in the orthogonality graph and the partial correlation graph, although $E_{PC} \subseteq E^a_{OG}$. We now confirm this conjecture with two counterexamples.

Example 5.10 Recall that for the Ornstein-Uhlenbeck process with $\Sigma_L = I_k$, due to Proposition 5.4 with $p = 1$, the characterisation

$$a \text{ --- } b \notin E_{PC} \quad \Leftrightarrow \quad [\mathbf{A}]_{ba} - [\mathbf{A}]_{ab} = 0, \quad \left[\mathbf{A}^\top \mathbf{A}\right]_{ab} = 0,$$

holds. Additionally, by Corollary 6.21 of Fasen-Hartmann and Schenk (2024a), we have

$$\begin{aligned} a \longrightarrow b \notin E_{OG} \quad &\Leftrightarrow \quad [\mathbf{A}^\alpha]_{ba} = 0, & \alpha = 1, \dots, k-1, \\ a \dashrightarrow b \notin E_{OG} \quad &\Leftrightarrow \quad \left[\mathbf{A}^\alpha \left(\mathbf{A}^\top\right)^\beta\right]_{ab} = 0, & \alpha, \beta = 0, \dots, k-1. \end{aligned} \tag{5.3}$$

(a) Suppose that $\mathcal{Y}_V$ is a 3-dimensional Ornstein-Uhlenbeck process with $\Sigma_L = I_3$ and

$$\mathbf{A} = \begin{pmatrix} -3 & 1 & 1 \\ 1 & -3 & 1 \\ 6 & 1 & -8 \end{pmatrix},$$

where $\sigma(\mathbf{A}) = \{-9, -4, -1\} \subseteq (-\infty, 0) + i\mathbb{R}$. Then

$$[\mathbf{A}]_{21} - [\mathbf{A}]_{12} = 0 \quad \text{and} \quad \left[\mathbf{A}^\top \mathbf{A}\right]_{12} = 0,$$

so $1 \longrightarrow 2 \notin E_{PC}$. However $1 \longrightarrow 2 \in E_{OG}$, $2 \longrightarrow 1 \in E_{OG}$, and $1 \dashrightarrow 2 \in E_{OG}$, since

$$[\mathbf{A}]_{21} \neq 0 \quad \text{and} \quad [\mathbf{A}]_{12} \neq 0.$$

(b) Suppose that $\mathcal{Y}_V$ is a 3-dimensional Ornstein-Uhlenbeck process with $\Sigma_L = I_3$ and

$$\mathbf{A} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 1 & 1 & -2 \end{pmatrix},$$

where $\sigma(\mathbf{A}) = \{-1, -1, -2\} \subseteq (-\infty, 0) + i\mathbb{R}$. Then a simple calculation shows that

$$\left[\mathbf{A}^{\alpha}\right]_{21} = \left[\mathbf{A}^{\alpha}\right]_{12} = 0, \quad \alpha = 1, 2, \quad \text{and} \quad \left[\mathbf{A}^{\alpha}\left(\mathbf{A}^{\top}\right)^{\beta}\right]_{12} = 0, \quad \alpha, \beta = 0, 1, 2.$$

Therefore, $1 \longrightarrow 2 \notin E_{OG}$, $2 \longrightarrow 1 \notin E_{OG}$, and $1 \dashrightarrow 2 \notin E_{OG}$. However, $1 \longrightarrow 2 \in E_{PC}$, since $\left[\mathbf{A}^{\top}\mathbf{A}\right]_{12} = 1$.

In summary, even in the special case $\Sigma_L = I_k$, there are no direct relations between the edges because, in the partial correlation graph the orthogonality of the columns in $\mathbf{A}$ is characteristic, whereas in the orthogonality graph the orthogonality of the rows is relevant for the undirected edges, and the orthogonality of the rows and columns is relevant for the directed edges. Of course, in some special cases, there are simple relations between the edges in the partial correlation graph and the edges in the orthogonality graph. Because of the orthogonality argument, an obvious special case is a symmetric matrix $\mathbf{A}$.

Lemma 5.11 *Suppose that $G_{PC} = (V, E_{PC})$ is the partial correlation graph and $G_{OG} = (V, E_{OG})$ is the orthogonality graph for the causal Ornstein-Uhlenbeck process $\mathcal{Y}_V$, where $\mathbf{A}$ is a symmetric matrix and $\Sigma_L = I_k$. Then, for $a, b \in V$ with $a \neq b$, we receive*

$$a \dashrightarrow b \notin E_{OG} \quad \Rightarrow \quad a \longrightarrow b \notin E_{PC}.$$

Next, we provide a comparison to the *local orthogonality graph* of Fasen-Hartmann and Schenk (2024a). To avoid going too deep into the intricate definition of the local orthogonality graph in its generality here, we present the definition of the local orthogonality graph only for MCAR processes via the characterisations used in Fasen-Hartmann and Schenk (2024a), Propositions 6.12 and 6.13. For a general definition of the local orthogonality graph, we refer to Fasen-Hartmann and Schenk (2024a), Definition 5.9.

Definition 5.12 Suppose $\mathcal{Y}_V$ is a causal MCAR(p) process with $\Sigma_L > 0$. Suppose $V = \{1, \ldots, k\}$ are the vertices and the edges E^0_{OG} for $a, b \in V$ with $a \neq b$ are defined via

(i) $a \longrightarrow b \notin E^0_{OG} \quad \Leftrightarrow \quad [A_j]_{ba} = 0 \quad \text{for } j = 1, \ldots, p,$
(ii) $a \dashrightarrow b \notin E^0_{OG} \quad \Leftrightarrow \quad [\Sigma_L]_{ab} = 0.$

Then $G^0_{OG} = (V, E^0_{OG})$ is called *local orthogonality graph* for $\mathcal{Y}_V$.

Remark 5.13 (a) We emphasise that the undirected edges in the local orthogonality graph are characterised by Σ_L and not Σ_L^{-1} as in the partial correlation graph, and these matrices generally do not match. The local orthogonality graph considers the direct correlation of $L_a(1)$ and $L_b(1)$, while the partial correlation graph considers the correlation of $L_a(1)$ and $L_b(1)$ given the environment $L_{V\setminus\{a,b\}}(1)$.

(b) Due to the different definitions, there are generally no direct relations between the edges in the partial correlation graph and the edges in the local orthogonality graph, not even in the special case $\Sigma_L = I_k$. Note that in this case, $a \dashrightarrow b \notin E^0_{OG}$ is always true. Furthermore, looking at Example 5.10(a), we get $1 \text{ --- } 2 \notin E_{PC}$ but $1 \longrightarrow 2 \in E^0_{OG}$ and $2 \longrightarrow 1 \in E^0_{OG}$. Whereas Example 5.10(b) is an example where $1 \longrightarrow 2 \notin E^0_{OG}$ and $2 \longrightarrow 1 \notin E^0_{OG}$ but $1 \text{ --- } 2 \in E_{PC}$.

(c) In the case of no environment ($k = 2$) we obtain that $[\Sigma_L]_{ab} = 0$ if and only if $[\Sigma_L^{-1}]_{ab} = 0$ and $a \text{ --- } b \notin E_{PC}$ implies $a \dashrightarrow b \notin E^0_{OG}$ and vice versa.

□

However, as for the orthogonality graph, we can establish relations between edges in the partial correlation graph and paths in the local orthogonality graph for MCAR processes via the concept of m-separation and augmentation separation, although no global AMP Markov property could be shown for the local orthogonality graph.

Lemma 5.14 *Suppose that $G_{PC} = (V, E_{PC})$ is the partial correlation graph, $G^0_{OG} = (V, E^0_{OG})$ is the local orthogonality graph, and $G^{0,a}_{OG} = (V, E^{0,a}_{OG})$ is the augmented local orthogonality graph for the causal MCAR(p) process $\mathcal{Y}_V$. Furthermore, we use the notation of Definitions 4.7, 4.11 and Lemma 4.14. Then, for $a, b \in V$ with $a \neq b$, the following equivalences hold.*

$$\begin{aligned} a \text{ --- } b \notin E^{0,a}_{OG} \quad &\Leftrightarrow \quad \{a\} \bowtie \{b\} \mid V \setminus \{a, b\} \quad [G^{0,a}_{OG}], \\ &\Leftrightarrow \quad \{a\} \bowtie_m \{b\} \mid V \setminus \{a, b\} \quad [G^0_{OG}], \\ &\Leftrightarrow \quad dis\,(a \cup ch(a)) \cap dis\,(b \cup ch(b)) \;\; in \;\; G^0_{OG}. \end{aligned}$$

In particular, we then have $a \text{ --- } b \notin E_{PC}$, i.e., $E_{PC} \subseteq E^{0,a}_{OG}$.

Note that the opposite inclusion does in general not hold, there exist examples where $E_{PC} \neq E^{0,a}_{OG}$ as for the orthogonality graph.

As discussed in Lemma 4.14, Lemma 5.14 provides us with several ways to make statements about the partial correlation graph from the local orthogonality graph.

5.4 Estimation

The edges in the partial correlation graph can be found simultaneously and computationally inexpensive using the inverse of the spectral density function (cf. Proposition

4.4). Therefore, in practical applications, we have to estimate the spectral density function from discrete-time observations. Suppose we observe a causal MCAR(p) process $\mathcal{Y}_V$ at equidistant times $0, \Delta, 2\Delta, \ldots$ with $\Delta > 0$ small, as used for modelling high-frequency data. The resulting discrete-time process $\mathcal{Y}_V^{\Delta} = (Y_V(k\Delta))_{k\in\mathbb{N}}$ is also weakly stationary with zero expectation, in fact, it is a vector ARMA process, with spectral density function

$$f^{(\Delta)}_{Y_V Y_V}(\lambda) = \frac{1}{2\pi}\sum_{k=-\infty}^{\infty} c_{Y_V Y_V}(k\Delta)\,\mathrm{e}^{-ik\lambda} = \frac{1}{\Delta}\sum_{k=-\infty}^{\infty} f_{Y_V Y_V}\left(\frac{\lambda+2k\pi}{\Delta}\right), \quad -\pi \le \lambda \le \pi,$$

where the second equality follows from Bloomfield (1976), p. 206.

Low-frequency sampling scheme

But clearly the zero entries of the inverse of $f^{(\Delta)}_{Y_V Y_V}(\lambda)$ for $\lambda \in [-\pi, \pi]$ do not necessarily coincide with the zero entries of the inverse of $f_{Y_V Y_V}(\lambda)$ for $\lambda \in \mathbb{R}$ and hence, there is in general no relationship between the edges in the partial correlation graph for $\mathcal{Y}_V$ and the partial correlation graph for $\mathcal{Y}_V^{(\Delta)}$. This can be seen nicely by looking at an Ornstein-Uhlenbeck process, where, due to Proposition 5.4, we have

$$a \;\text{---}\; b \notin E_{PC} \quad \Leftrightarrow \quad \left[\Sigma_L^{-1}\right]_{ab} = 0, \quad \left[\mathbf{A}^\top \Sigma_L^{-1} - \Sigma_L^{-1}\mathbf{A}\right]_{ab} = 0, \quad \left[\mathbf{A}^\top \Sigma_L^{-1}\mathbf{A}\right]_{ab} = 0.$$

The discrete-time sampled process $\mathcal{Y}_V^{(\Delta)}$ is a VAR(1) process where, due to Remark 5.6, the edges in the partial correlation graph $G_{PC}^{(\Delta)} = (V, E_{PC}^{(\Delta)})$ for $\mathcal{Y}_V^{(\Delta)}$ can be described by the relation

$$\begin{aligned} a \;\text{---}\; b \notin E^{(\Delta)}_{PC} \quad \Leftrightarrow \quad & \left[\left(\Sigma^{(\Delta)}\right)^{-1} + e^{\mathbf{A}^\top\Delta}\left(\Sigma^{(\Delta)}\right)^{-1} e^{\mathbf{A}\Delta}\right]_{ab} = 0, \\ & \left[\left(\Sigma^{(\Delta)}\right)^{-1} e^{\mathbf{A}\Delta}\right]_{ab} = 0, \quad \left[e^{\mathbf{A}^\top\Delta}\left(\Sigma^{(\Delta)}\right)^{-1}\right]_{ab} = 0 \end{aligned}$$

where $\Sigma^{(\Delta)} = \int_0^{\Delta} e^{\mathbf{A}u}\Sigma_L e^{\mathbf{A}^\top u} du > 0$. These characterisations confirm that there do not exist direct relationships between E_{PC} and $E^{(\Delta)}_{PC}$. Therefore, in general, it will be challenging to derive a nonparametric estimator for the edges in the partial correlation graph from a low-frequency sampling scheme.

However, it is possible to derive estimators for a parametric class of continuous-time processes. In the case of MCAR models, the model parameters $A_1, \ldots, A_p, \Sigma_L$ can be estimated from the low-frequency sampled MCAR process $\mathcal{Y}_V^{(\Delta)}$, e.g., by quasi maximum-likelihood estimation as in Schlemm and Stelzer (2012b) or Whittle estimation as in Fasen-Hartmann and Mayer (2022), yielding the parameter estimators $\widehat{A}_1, \ldots, \widehat{A}_p, \widehat{\Sigma}_L$, which are consistent and asymptotically normally distributed. Then an estimator for the inverse of the spectral density is

$$\widehat{g}_{Y_V Y_V}(\lambda) = 2\pi\,\widehat{P}(-i\lambda)^\top \widehat{\Sigma}_L^{-1}\widehat{P}(i\lambda), \quad \lambda \in \mathbb{R} \quad \text{with}$$

$$\widehat{P}(z) = I_k z^p + \widehat{A}_1 z^{p-1} + \ldots + \widehat{A}_p, \quad z \in \mathbb{C},$$

which is also a consistent and asymptotically normally distributed estimator for $g_{Y_V Y_V}(\lambda)$ for fixed $\lambda \in \mathbb{R}$ by an application of the continuous mapping theorem and the delta-method, respectively. By considering the zero entries of this function we receive estimators for the edges in the partial correlation graph for the underlying continuous-time process $\mathcal{Y}_V$.

High-frequency sampling scheme

In the context of high-frequency data where $\Delta \to 0$, we have the relation

$$\lim_{\Delta \to 0} \Delta f^{\Delta}_{Y_V Y_V}(\lambda\Delta) \mathbb{1}_{[-\frac{\pi}{\Delta}, \frac{\pi}{\Delta}]}(\lambda) = f_{Y_V Y_V}(\lambda), \quad \lambda \in \mathbb{R},$$

(Fasen and Fuchs 2013a, Eq. (1.5) for CARMA processes but this is also true for our causal MCAR processes). Roughly speaking, this means that in the limit $\Delta \to 0$, we can identify the edges of the causal MCAR process from the edges of its equidistantly sampled observations. In the special case of univariate CARMA processes, we already know from Fasen and Fuchs (2013a, b) that, under some mild assumptions, the smoothed normalised periodogram is a consistent estimator of the spectral density $f_{Y_V Y_V}(\lambda)$ for the high-frequency sampling scheme, where $\Delta_n \to 0$ and $n\Delta_n \to \infty$ as the number of observations $n \to \infty$. We believe this is still true for multivariate CARMA processes including MCAR processes. An alternative estimator is the lag-window spectral density estimator of Kartsioukas et al. (2023). They develop the statistical inference of this estimator not only for MCAR processes but also for general multivariate stationary processes in Hilbert spaces and, furthermore, they also allow an irregular sampling scheme. For non-Gaussian processes, however, a cumulant condition must be satisfied, which is in the context of MCAR processes a cumulant condition on the driving Lévy process. Due to the generality of this impressive paper, the assumptions for MCAR processes are actually stronger than necessary.

6 Conclusion

The paper establishes and analyses the partial correlation relation for wide-sense stationary and mean-square continuous stochastic processes in continuous time with expectation zero and spectral density function. Based on this, the partial correlation graph for continuous-time stochastic processes is defined, which satisfies the usual Markov properties. Furthermore, we relate the partial correlation graph to the orthogonality and the local orthogonality graph of Fasen-Hartmann and Schenk (2024a) by their augmented graphs and we find some interesting relationships. The derived results for the partial correlation graph in the continuous-time setting correspond to the results for discrete-time processes in Dahlhaus (2000), which we also see by applications to MCAR processes, where we can characterise the edges by the model parameters. In both settings, the low-frequency sampling regime and the high-frequency sampling regime, it is possible to derive some consistent and asymptotically normally distributed estimators for the inverse spectral density of an MCAR process and thus also for the edges in the partial correlation graph. In the high-frequency sampling scheme, the smoothed periodogram and the lag-window spectral density estimator are popular estimators for the spectral density for discrete-time processes (Anderson 1971; Brockwell and Davis 1991; Brillinger 2001; Hannan 1970) and they should also work for a large class of non-parametric continuous-time models. The paper focused on the theoretical properties of the partial correlation graph but statistical methods for estimation and testing for the edges in the continuous-time partial correlation graph are of particular importance and will be the subject of some future work.

7 Proofs

7.1 Proofs of Sect. 3

Proof of Proposition 3.3 Let $t \in \mathbb{R}$ and assume that $\{a\} \cap C = \emptyset$, since the statements apply trivially for $a \in C$. To simplify the notation, we abbreviate

$$\widehat{Y}_{a|C}(t) = \int_{-\infty}^{\infty} e^{i\lambda t} f_{Y_a Y_C}(\lambda) f_{Y_C Y_C}(\lambda)^{-1} Z_C(d\lambda).$$

The proof is divided into three steps. In the first step we derive that $\widehat{Y}_{a|C}(t) \in \mathcal{L}_{Y_C}$ and in the second step we show that $Y_a(t) - \widehat{Y}_{a|C}(t) \in \mathcal{L}_{Y_C}^{\perp}$. Both together then give the assertion $\widehat{Y}_{a|C}(t) = P_{\mathcal{L}_{Y_C}} Y_a(t)$. Then, in a third step, we conclude that the orthogonal projection is the solution to the optimisation problem (3.3).

Step 1: Given $\mathcal{L}_{Y_C}$, we can establish the measurability and integrability of the function $e^{i\lambda t} f_{Y_a Y_C}(\lambda) f_{Y_C Y_C}(\lambda)^{-1}$ for $\lambda, t \in \mathbb{R}$. For the measurability, we first note that $f_{Y_a Y_C}$ and $f_{Y_C Y_C}$ are measurable as derivatives. Furthermore, sums and products of measurable functions are measurable. If we set $\lambda/0 := 0$ for $\lambda \in \mathbb{R}$, then their quotients are also measurable (Klenke (2020), Theorem 1.91). Now we compute $f_{Y_C Y_C}(\lambda)^{-1}$ by Gaussian elimination and find that $f_{Y_C Y_C}(\lambda)^{-1}$ is measurable for $\lambda \in \mathbb{R}$. Thus, $e^{i\lambda t} f_{Y_a Y_C}(\lambda) f_{Y_C Y_C}(\lambda)^{-1}$, $\lambda, t \in \mathbb{R}$, is also measurable.

For the integrability, we first note that $f_{Y_{\{a\}\cup C} Y_{\{a\}\cup C}}(\lambda) \geq 0$ due to Lemma 2.1(c). Furthermore, $f_{Y_C Y_C}(\lambda) > 0$ by assumption, so Proposition 8.2.4 of Bernstein (2009) gives

$$f_{Y_a Y_C}(\lambda) f_{Y_C Y_C}(\lambda)^{-1} f_{Y_C Y_a}(\lambda) \leq f_{Y_a Y_a}(\lambda)$$

for $\lambda \in \mathbb{R}$. Since further $f_{Y_a Y_C}(\lambda) f_{Y_C Y_C}(\lambda)^{-1} f_{Y_C Y_a}(\lambda) \geq 0$ and the integral is monotonous, we obtain the integrability

$$\begin{aligned}
&\int_{-\infty}^{\infty} f_{Y_a Y_C}(\lambda) f_{Y_C Y_C}(\lambda)^{-1} f_{Y_C Y_C}(\lambda) \overline{f_{Y_a Y_C}(\lambda) f_{Y_C Y_C}(\lambda)^{-1}}^{\top} d\lambda \\
&= \int_{-\infty}^{\infty} f_{Y_a Y_C}(\lambda) f_{Y_C Y_C}(\lambda)^{-1} f_{Y_C Y_a}(\lambda) d\lambda \leq \int_{-\infty}^{\infty} f_{Y_a Y_a}(\lambda) d\lambda < \infty,
\end{aligned}$$

where the finiteness follows from Lemma 2.1(a). In summary, $\widehat{Y}_{a|C}(t) \in \mathcal{L}_{Y_C}$ for $t \in \mathbb{R}$.

Step 2: Any element $Y^C \in \mathcal{L}_{Y_C}$ has the spectral representation

$$Y^C = \int_{-\infty}^{\infty} \varphi(\lambda)^\top Z_C(d\lambda) \quad \mathbb{P}\text{-a.s.},$$

where $\int_{-\infty}^{\infty} \varphi(\lambda)^\top f_{Y_C Y_C}(\lambda)\overline{\varphi(\lambda)} < \infty$. Now, writing $Y_a(t)$ in its spectral representation (2.2), it holds that

$$\begin{aligned}
&\mathbb{E}\Big[\Big(Y_a(t) - \widehat{Y}_{a|C}(t)\Big)\overline{Y^C}\Big] \\
&= \mathbb{E}\left[\left(\int_{-\infty}^{\infty} e^{i\lambda t} Z_a(d\lambda) - \int_{-\infty}^{\infty} e^{i\lambda t} f_{Y_a Y_C}(\lambda) f_{Y_C Y_C}(\lambda)^{-1} Z_C(d\lambda)\right) \overline{\int_{-\infty}^{\infty} \varphi(\lambda)^\top Z_C(d\lambda)}\right] \\
&= \int_{-\infty}^{\infty} e^{i\lambda t} f_{Y_a Y_C}(\lambda)\overline{\varphi(\lambda)} d\lambda - \int_{-\infty}^{\infty} e^{i\lambda t} f_{Y_a Y_C}(\lambda) f_{Y_C Y_C}(\lambda)^{-1} f_{Y_C Y_C}(\lambda)\overline{\varphi(\lambda)} d\lambda = 0.
\end{aligned}$$

Thus, $Y_a(t) - \widehat{Y}_{a|C}(t) \in \mathcal{L}_{Y_C}^{\perp}$ for $t \in \mathbb{R}$.

Step 3: From the minimality property of the orthogonal projection, we obtain $Y^C = P_{\mathcal{L}_{Y_C}} Y_a(t)$ is the optimal solution to this optimisation problem. Due to Step 1 and Step 2 the function $\varphi_{a|C}(\lambda) = f_{Y_a Y_C}(\lambda) f_{Y_C Y_C}(\lambda)^{-1}$, $\lambda \in \mathbb{R}$, is then the optimal function in (3.3). □

Proof of Lemma 3.5 We can write

$$\begin{aligned}
\varepsilon_{A|C}(t) &= \int_{-\infty}^{\infty} e^{i\lambda t} Z_A(d\lambda) - \int_{-\infty}^{\infty} e^{i\lambda t} f_{Y_A Y_C}(\lambda) f_{Y_C Y_C}(\lambda)^{-1} Z_C(d\lambda) \\
&= \int_{-\infty}^{\infty} e^{i\lambda t} \left(E_A^\top - f_{Y_A Y_C}(\lambda) f_{Y_C Y_C}(\lambda)^{-1} E_C^\top\right) Z_V(d\lambda),
\end{aligned}$$

where $E_A \in \mathbb{R}^{k\times|A|}$ (and analogously $E_C \in \mathbb{R}^{k\times|C|}$) is the matrix defined by its entries

$$[E_A]_{ij} = \begin{cases} 1, & i = j,\ i, j \in A, \\ 0, & \text{else}. \end{cases}$$

Therefore, the noise process $(\varepsilon_{A|C}(t))_{t\in\mathbb{R}}$ is a linear transformation of the wide-sense stationary process $\mathcal{Y}_V$ with spectral characteristic $E_A^\top - f_{Y_A Y_C}(\lambda) f_{Y_C Y_C}(\lambda)^{-1} E_C^\top$, $\lambda \in \mathbb{R}$. Thus $\mathcal{Y}_V$ is also wide-sense stationary (Rozanov (1967), I, (8.2)). Furthermore, the linear transformation has a spectral density function, which is given by (Rozanov (1967), I, (8.13))

$$\begin{aligned}
&f_{\varepsilon_{A|C}\varepsilon_{A|C}}(\lambda) \\
&= \left(E_A^\top - f_{Y_A Y_C}(\lambda) f_{Y_C Y_C}(\lambda)^{-1} E_C^\top\right) f_{Y_V Y_V}(\lambda) \overline{\left(E_A^\top - f_{Y_A Y_C}(\lambda) f_{Y_C Y_C}(\lambda)^{-1} E_C^\top\right)}^\top \\
&= f_{Y_A Y_A}(\lambda) - f_{Y_A Y_C}(\lambda) f_{Y_C Y_C}(\lambda)^{-1} f_{Y_C Y_A}(\lambda).
\end{aligned}$$

Then Lemma 2.1(b) yields

$$c_{\varepsilon_{A|C}\varepsilon_{A|C}}(t) = \int_{-\infty}^{\infty} e^{i\lambda t}\left(f_{Y_AY_A}(\lambda) - f_{Y_AY_C}(\lambda)f_{Y_CY_C}(\lambda)^{-1}f_{Y_CY_A}(\lambda)\right)d\lambda$$

for $t \in \mathbb{R}$. In particular, the spectral density function of $(\varepsilon_{A\cup B|C}(t))_{t\in\mathbb{R}}$ is given by

$$f_{\varepsilon_{A\cup B|C}\varepsilon_{A\cup B|C}}(\lambda) = f_{Y_{A\cup B}Y_{A\cup B}}(\lambda) - f_{Y_{A\cup B}Y_C}(\lambda)f_{Y_CY_C}(\lambda)^{-1}f_{Y_CY_{A\cup B}}(\lambda).$$

Thus, the cross-spectral density function is, for almost all $\lambda \in \mathbb{R}$,

$$\begin{aligned} f_{\varepsilon_{A|C}\varepsilon_{B|C}}(\lambda) &= E_A^\top\left(f_{Y_{A\cup B}Y_{A\cup B}}(\lambda) - f_{Y_{A\cup B}Y_C}(\lambda)f_{Y_CY_C}(\lambda)^{-1}f_{Y_CY_{A\cup B}}(\lambda)\right)E_B \\ &= f_{Y_AY_B}(\lambda) - f_{Y_AY_C}(\lambda)f_{Y_CY_C}(\lambda)^{-1}f_{Y_CY_B}(\lambda), \end{aligned}$$

and, for all $t \in \mathbb{R}$, it holds that

$$c_{\varepsilon_{A|C}\varepsilon_{B|C}}(t) = \int_{-\infty}^{\infty} e^{i\lambda t}\left(f_{Y_AY_B}(\lambda) - f_{Y_AY_C}(\lambda)f_{Y_CY_C}(\lambda)^{-1}f_{Y_CY_B}(\lambda)\right)d\lambda.$$

Proof of Proposition 3.6 Suppose that $\mathcal{Y}_A \perp\!\!\!\perp \mathcal{Y}_B \mid \mathcal{Y}_C$. By definition of this relation we obtain the first characterisation $c_{\varepsilon_{A|C}\varepsilon_{B|C}}(t) = 0_{|A|\times|B|}$ for $t \in \mathbb{R}$. For the second characterisation suppose that $c_{\varepsilon_{A|C}\varepsilon_{B|C}}(t) = 0_{|A|\times|B|}$ for $t \in \mathbb{R}$. Then the Fourier inversion formula (Pinsky 2009, Proposition 2.2.37) yields $f_{\varepsilon_{A|C}\varepsilon_{B|C}}(\lambda) = 0_{|A|\times|B|}$ for almost all $\lambda \in \mathbb{R}$. If $f_{\varepsilon_{A|C}\varepsilon_{B|C}}(\lambda) = 0_{|A|\times|B|}$ for almost all $\lambda \in \mathbb{R}$, then Lemma 2.1 gives $c_{\varepsilon_{A|C}\varepsilon_{B|C}}(t) = 0_{|A|\times|B|}$ for $t \in \mathbb{R}$. For the third characterisation, suppose that $f_{\varepsilon_{A|C}\varepsilon_{B|C}}(\lambda) = 0_{|A|\times|B|}$ for almost all $\lambda \in \mathbb{R}$. Then $R_{\varepsilon_{A|C}\varepsilon_{B|C}}(\lambda) = 0_{|A|\times|B|}$ holds by Definition 2.3. If we additionally assume that $f_{\varepsilon_{A|C}\varepsilon_{A|C}}(\lambda) > 0$ and $f_{\varepsilon_{B|C}\varepsilon_{B|C}}(\lambda) > 0$, then Definition 2.3 provides the second direction. □

Proof of Proposition 3.8 Let $\lambda \in \mathbb{R}$. For notational convenience, we assume without loss of generality that $A := \{1, 2, \ldots, \alpha\}$, $B := \{\alpha+1, \ldots, \alpha+\beta\}$, and $C := \{\alpha+\beta+1, \ldots, \alpha+\beta+\gamma\}$. Then, we can decompose

$$f_{Y_{A\cup B\cup C}Y_{A\cup B\cup C}}(\lambda) = \begin{pmatrix} f_{Y_{A\cup B}Y_{A\cup B}}(\lambda) & f_{Y_{A\cup B}Y_C}(\lambda) \\ f_{Y_CY_{A\cup B}}(\lambda) & f_{Y_CY_C}(\lambda) \end{pmatrix}.$$

As $f_{Y_{A\cup B\cup C}Y_{A\cup B\cup C}}(\lambda) > 0$ and $f_{Y_CY_C}(\lambda) > 0$ by assumption, e.g., Theorem 2.1 of Lu and Shiou (2002) gives

$$\begin{aligned} &\left[g_{Y_{A\cup B\cup C}Y_{A\cup B\cup C}}(\lambda)\right]_{A\cup B A\cup B} \\ &= \left(f_{Y_{A\cup B}Y_{A\cup B}}(\lambda) - f_{Y_{A\cup B}Y_C}(\lambda)f_{Y_CY_C}(\lambda)^{-1}f_{Y_CY_{A\cup B}}(\lambda)\right)^{-1} \\ &= f_{\varepsilon_{A\cup B|C}\varepsilon_{A\cup B|C}}(\lambda)^{-1}. \end{aligned}$$

The matrix $f_{\varepsilon_{A\cup B|C}\varepsilon_{A\cup B|C}}(\lambda)$ itself has the decomposition

$$f_{\varepsilon_{A\cup B|C}\varepsilon_{A\cup B|C}}(\lambda) = \begin{pmatrix} f_{\varepsilon_{A|C}\varepsilon_{A|C}}(\lambda) & f_{\varepsilon_{A|C}\varepsilon_{B|C}}(\lambda) \\ f_{\varepsilon_{B|C}\varepsilon_{A|C}}(\lambda) & f_{\varepsilon_{B|C}\varepsilon_{B|C}}(\lambda) \end{pmatrix}.$$

Note that $f_{\varepsilon_{A\cup B|C}\varepsilon_{A\cup B|C}}(\lambda) > 0$ and $f_{\varepsilon_{B|C}\varepsilon_{B|C}}(\lambda) > 0$ due to $f_{Y_{A\cup B\cup C}Y_{A\cup B\cup C}}(\lambda) > 0$ and Remark 3.7(b). Thus, we can make use of the matrix inversion formula again and Theorem 2.1 of Lu and Shiou (2002) yields

$$\begin{aligned} &\left[g_{Y_{A\cup B\cup C}Y_{A\cup B\cup C}}(\lambda)\right]_{AB} \\ &= \left[f_{\varepsilon_{A\cup B|C}\varepsilon_{A\cup B|C}}(\lambda)^{-1}\right]_{AB} \\ &= -\left(f_{\varepsilon_{A|C}\varepsilon_{A|C}}(\lambda) - f_{\varepsilon_{A|C}\varepsilon_{B|C}}(\lambda) f_{\varepsilon_{B|C}\varepsilon_{B|C}}(\lambda)^{-1} f_{\varepsilon_{B|C}\varepsilon_{A|C}}(\lambda)\right)^{-1} \\ &\quad \cdot f_{\varepsilon_{A|C}\varepsilon_{B|C}}(\lambda) f_{\varepsilon_{B|C}\varepsilon_{B|C}}(\lambda)^{-1}. \end{aligned}$$

This representation shows that $\left[g_{Y_{A\cup B\cup C}Y_{A\cup B\cup C}}(\lambda)\right]_{AB} = 0_{\alpha\times\beta}$ for almost all $\lambda \in \mathbb{R}$ if and only if $f_{\varepsilon_{A|C}\varepsilon_{B|C}}(\lambda) = 0_{\alpha\times\beta}$ for almost all $\lambda \in \mathbb{R}$. Proposition 3.6 concludes the proof. □

Proof of Proposition 3.11 *(P4)* The relations $\mathcal{Y}_A \perp\!\!\!\perp \mathcal{Y}_B \mid \mathcal{Y}_D$, $\mathcal{Y}_A \perp\!\!\!\perp \mathcal{Y}_C \mid (\mathcal{Y}_B, \mathcal{Y}_D)$, and Proposition 3.8 result in

$$\left[g_{Y_{A\cup B\cup D}Y_{A\cup B\cup D}}(\lambda)\right]_{AB} = 0_{|A|\times|B|} \quad \text{and} \quad \left[g_{Y_{A\cup B\cup C\cup D}Y_{A\cup B\cup C\cup D}}(\lambda)\right]_{AC} = 0_{|A|\times|C|} \tag{A.1}$$

for almost all $\lambda \in \mathbb{R}$. Along with Proposition 3.9, we obtain

$$\begin{aligned} 0_{|A|\times|B|} &= \left[g_{Y_{A\cup B\cup D}Y_{A\cup B\cup D}}(\lambda)\right]_{AB} \\ &= \left[g_{Y_{A\cup B\cup C\cup D}Y_{A\cup B\cup C\cup D}}(\lambda)\right]_{AB} - \left[g_{Y_{A\cup B\cup C\cup D}Y_{A\cup B\cup C\cup D}}(\lambda)\right]_{AC} \\ &\quad \left[\left(g_{Y_{A\cup B\cup C\cup D}Y_{A\cup B\cup C\cup D}}(\lambda)\right]_{CC}\right)^{-1} \left[g_{Y_{A\cup B\cup C\cup D}Y_{A\cup B\cup C\cup D}}(\lambda)\right]_{CB} \\ &= \left[g_{Y_{A\cup B\cup C\cup D}Y_{A\cup B\cup C\cup D}}(\lambda)\right]_{AB} \end{aligned} \tag{A.2}$$

for almost all $\lambda \in \mathbb{R}$. In summary, equations (A.1) and (A.2) give

$$\left[g_{Y_{A\cup B\cup C\cup D}Y_{A\cup B\cup C\cup D}}(\lambda)\right]_{A(B\cup C)} = 0_{|A|\times(|B|+|C|)}$$

for almost all $\lambda \in \mathbb{R}$. Proposition 3.8 implies $\mathcal{Y}_A \perp\!\!\!\perp (\mathcal{Y}_B, \mathcal{Y}_C) \mid \mathcal{Y}_D$. □

7.2 Proofs of Sect. 4

Proof of Lemma 4.12 Theorem 5.15 in Fasen-Hartmann and Schenk (2024a) provides that $\{a\} \bowtie_m \{b\} \mid V \setminus \{a, b\} \quad [G_{OG}]$ implies $\mathcal{L}_{Y_a} \perp \mathcal{L}_{Y_b} \mid \mathcal{L}_{Y_{V\setminus\{a,b\}}}$. This conditional orthogonality relation immediately implies $L_{Y_a}(t) \perp L_{Y_b}(t) \mid \mathcal{L}_{Y_{V\setminus\{a,b\}}}$ for all $t \in \mathbb{R}$ by subset arguments, which in turn yields $a — b \notin E_{PC}$ due to Remark 3.2. □

Proof of Lemma 4.14 By definition and due to Eichler (2011), Theorem 3.1 and Lemma 3.2, we obtain that

$$
\begin{aligned}
a — b \notin E_{OG}^a \quad &\Leftrightarrow \quad a \text{ and } b \text{ are not collider connected in } G_{OG}, \\
&\Leftrightarrow \quad \text{dis}\,(a \cup \text{ch}(a)) \cap \text{dis}\,(b \cup \text{ch}(b)) \quad \text{in } G_{OG}, \\
&\Leftrightarrow \quad \{a\} \bowtie_m \{b\} \mid V \setminus \{a,b\} \quad [G_{OG}], \\
&\Leftrightarrow \quad \{a\} \bowtie \{b\} \mid V \setminus \{a,b\} \quad [G_{OG}^a].
\end{aligned}
$$

These statements are then of course all sufficient for $a — b \notin E_{PC}$ due to the previous Lemma 4.12, and $E_{PC} \subseteq E_{OG}^a$ is valid. □

7.3 Proofs of Sect. 5

Proof of Proposition 5.4 First of all, we insert the AR polynomial P in $g_{Y_V Y_V}(\lambda)$ to get

$$
\begin{aligned}
g_{Y_V Y_V}(\lambda) &= 2\pi \left(\sum_{m=0}^{p} A_{p-m}^\top (-i\lambda)^m \right) \Sigma_L^{-1} \left(\sum_{\ell=0}^{p} A_{p-\ell}(i\lambda)^\ell \right) \\
&= 2\pi \sum_{n=0}^{2p} \sum_{m=0 \vee n-p}^{n \wedge p} (-1)^m A_{p-m}^\top \Sigma_L^{-1} A_{p-n+m} (i\lambda)^n.
\end{aligned}
$$

In the last step, we arrange the addends according to the degree of λ and substitute $n = \ell + m$, where $n = 0, \ldots, 2p$. Since $0 \le \ell = n - m \le p$ and $0 \le m \le p$, we obtain the boundary $0 \vee n - p \le m \le n \wedge p$. Since the components of $g_{Y_V Y_V}$ are polynomials, the components are zero functions if and only if the corresponding coefficients are zero. Then, by Proposition 5.3, we obtain that

$$
a — b \notin E_{PC} \quad \Leftrightarrow \quad \left[\sum_{m=0 \vee n-p}^{n \wedge p} (-1)^m A_{p-m}^\top \Sigma_L^{-1} A_{p-n+m} \right]_{ab} = 0 \quad \text{for } n = 0, \ldots, 2p. \tag{A.3}
$$

(i) Assume that $\Sigma_L = \sigma^2 I_k$. Then $\Sigma_L^{-1} = 1/\sigma^2 I_k$ holds and since $\sigma^2 > 0$, relation (A.3) is equivalent to Proposition 5.4(i).

(ii) Assume that $A_m, m = 1, \ldots, p$, are diagonal matrices. Then the AR polynomial P is a diagonal matrix polynomial and $a — b \notin E_{PC}$ is equivalent to

$$
0 = \left[P(-i\lambda) \Sigma_L^{-1} P(i\lambda) \right]_{ab} = [P(-i\lambda)]_{aa} \left[\Sigma_L^{-1} \right]_{ab} [P(i\lambda)]_{bb}
$$

for all $\lambda \in \mathbb{R}$. Due to the causality assumption $\sigma(\mathbf{A}) \subseteq (-\infty, 0) + i\mathbb{R}$ and the structure of $\mathbf{A}$, the diagonal matrix A_p is not singular and in particular the diagonal elements of

A_p are not zero. Thus the diagonal elements of $P(i\lambda)$ are never zero and a —— $b \notin E_{PC}$ is equivalent to $[\Sigma_L^{-1}]_{ab} = 0$. □

Proof of Lemma 5.11 The assumptions that $\Sigma_L = I_k$, $\mathbf{A}$ is symmetric, and (5.3) imply

$$[\mathbf{A}]_{ba} - [\mathbf{A}]_{ab} = 0, \quad \left[\mathbf{A}^\top \mathbf{A}\right]_{ab} = \left[\mathbf{A}\mathbf{A}^\top\right]_{ab} = 0.$$

Thus, (5.2) yields a —— $b \notin E_{PC}$. □

Proof of Lemma 5.14 In Lemma 4.14, we already established the equivalences

$$\begin{aligned} a \text{ —— } b \notin E_{OG}^{0,a} \quad &\Leftrightarrow \quad a \text{ and } b \text{ are not collider connected in } G_{OG}^0, \\ &\Leftrightarrow \quad \operatorname{dis}(a \cup \operatorname{ch}(a)) \cap \operatorname{dis}(b \cup \operatorname{ch}(b)) = \emptyset \text{ in } G_{OG}^0, \\ &\Leftrightarrow \quad \{a\} \bowtie \{b\} \mid V \setminus \{a, b\} \quad [G_{OG}^{0,a}], \\ &\Leftrightarrow \quad \{a\} \bowtie_m \{b\} \mid V \setminus \{a, b\} \quad [G_{OG}^0], \end{aligned}$$

regardless of the specific definition of the graph. Thus, we only need to prove a —— $b \notin E_{PC}$, which we do by contradiction. So suppose that a —— $b \in E_{PC}$. Then there exists a $\lambda \in \mathbb{R}$, such that

$$0 \neq \left[P(-i\lambda)^\top \Sigma_L^{-1} P(i\lambda)\right]_{ab} = \sum_{c\in V}\sum_{d\in V} [P(-i\lambda)]_{ca} \left[\Sigma_L^{-1}\right]_{cd} [P(i\lambda)]_{db}.$$

Consequently, there exist vertices $c, d \in V$, such that

$$[P(-i\lambda)]_{ca} \neq 0, \quad \left[\Sigma_L^{-1}\right]_{cd} \neq 0, \quad [P(i\lambda)]_{db} \neq 0.$$

The statements about the AR polynomial yield that there exist directed edges $a \longrightarrow c$ and $b \longrightarrow d$ in G_{OG}^0. The statement $[\Sigma_L^{-1}]_{cd} \neq 0$ yields that there exists a path π between c and d of only undirected edges in the local orthogonality graph G_{OG}^0 (Eichler 2007, p. 341).

Indeed, consider an Ornstein-Uhlenbeck process $\widetilde{\mathcal{Y}}_V$ with $\mathbf{A} = -I_k$, that is driven by the same Lévy process as the MCAR process $\mathcal{Y}_V$. Then $\widetilde{\mathcal{Y}}_V$ generates a partial correlation graph $\widetilde{G}_{PC} = (V, \widetilde{E}_{PC})$ and an orthogonality graph $\widetilde{G}_{OG} = (V, \widetilde{E}_{OG})$. For these graphical models, we have c —— $d \notin \widetilde{E}_{PC}$ if and only if $[\Sigma_L^{-1}]_{cd} = 0$ and c - - - $d \notin \widetilde{E}_{OG}$ if and only if $[\Sigma_L]_{cd} = 0$. Additionally, there are no directed edges in the orthogonality graph. Then, a consequence of Lemma 4.14 is that c —— $d \in \widetilde{E}_{PC}$ ($[\Sigma_L^{-1}]_{cd} \neq 0$) implies c —— $d \in \widetilde{E}_{OG}^a$ and $\operatorname{dis}(c) \cap \operatorname{dis}(d) \neq \emptyset$ in $\widetilde{E}_{OG}$. Thus, there exists a path $\widetilde{\pi}$ of only undirected edges between c and d in the orthogonality graph $\widetilde{G}_{OG}$, i.e., for some $c = \alpha_1, \ldots, \alpha_\ell = d \in V$, we have $[\Sigma_L]_{\alpha_i \alpha_{i+1}} \neq 0$ for $i = 1, \ldots, \ell - 1$. This result implies a path $\widetilde{\pi}$ of only undirected edges between c and d in the local orthogonality graph G_{OG}^0.

Finally, if $\widetilde{\pi}$ does not already provide a path between a and b, we complete $\widetilde{\pi}$ with one or both directed edges from above, to get a path π between a and b on which every intermediate vertex is a collider. This path contradicts the premise and the statement $a —— b \notin E_{PC}$ is valid. In particular, $E_{PC} \subseteq E_{OG}^{0,a}$ is satisfied. □

Funding Open Access funding enabled and organized by Projekt DEAL.

Declarations

Conflict of interest On behalf of all authors, the corresponding author states that there is no Conflict of interest.

References

Aalen OO (1987) Dynamic modelling and causality. Scand Actuar J 1987:177–190. https://doi.org/10.1080/03461238.1987.10413826

Abdelwahab A, Amor O, Abdelwahed T (2008) The analysis of the interdependence structure in international financial markets by graphical models. Int Res J Financ Econ 15:291–306

Anderson TW (1971) The statistical analysis of time series. Wiley, New York

Anderson TW (1984) An introduction to multivariate statistical analysis, 2nd edn. Wiley, New York

Bergstrom AR (1997) Gaussian estimation of mixed-order continuous-time dynamic models with unobservable stochastic trends from mixed stock and flow data. Econom Theory 13:467–505. https://doi.org/10.1017/S0266466600005971

Bernstein D (2009) Matrix mathematics: theory, facts, and formulas, 2nd edn. Princeton University Press, Princeton. https://doi.org/10.1515/9781400833344

Bloomfield P (1976) Fourier analysis of time series: an introduction. Wiley, New York

Brillinger DR (1996) Remarks concerning graphical models for time series and point processes. Braz Rev Econom 16:1–23. https://doi.org/10.12660/bre.v16n11996.2878

Brillinger DR (2001) Time series: data analysis and theory. Society for Industrial and Applied Mathematics, Philadelphia. https://doi.org/10.1137/1.9780898719246

Brockwell P (2014) Recent results in the theory and applications of CARMA processes. Ann Inst Stat Math 66:647–685. https://doi.org/10.1007/s10463-014-0468-7

Brockwell P, Davis R (1991) Time series: theory and methods, 2nd edn. Springer, New York. https://doi.org/10.1007/978-1-4419-0320-4

Brockwell PJ, Lindner AM (2024) Continuous-parameter time series. De Gruyter, Berlin

Cramér H (1940) On the theory of stationary random processes. Ann Math 41:215–230. https://doi.org/10.2307/1968827

Dahlhaus R (2000) Graphical interaction models for multivariate time series. Metrika 51:157–172. https://doi.org/10.1007/s001840000055

Dahlhaus R, Eichler M (2003) Causality and graphical models in time series analysis. In highly structured stochastic systems. Oxford University Press, Oxford, pp 115–144. https://doi.org/10.1093/oso/9780198510550.003.0011

Dahlhaus R, Eichler M, Sandkühler J (1997) Identification of synaptic connections in neural ensembles by graphical models. J Neurosci Methods 77:93–107. https://doi.org/10.1016/S0165-0270(97)00100-3

Dey D, Datta A, Banerjee S (2021) Graphical Gaussian process models for highly multivariate spatial data. Biometrika 109:993–1014. https://doi.org/10.1093/biomet/asab061

Didelez V (2007) Graphical models for composable finite Markov processes. Scand J Stat 34:169–185. https://doi.org/10.1111/j.1467-9469.2006.00528.x
Didelez V (2008) Graphical models for marked point processes based on local independence. J R Stat Soc Ser B 70:245–264. https://doi.org/10.1111/j.1467-9868.2007.00634.x
Doob JL (1944) The elementary Gaussian processes. Ann Math Stat 15:229–282. https://doi.org/10.1214/aoms/1177731234
Doob JL (1953) Stochastic processes, 3rd edn. Wiley, New York
Edwards D (2000) Introduction to graphical modelling, 2nd edn. Springer, New York. https://doi.org/10.1007/978-1-4612-0493-0
Eichler M (2007) Granger causality and path diagrams for multivariate time series. J Econom 137:334–353. https://doi.org/10.1016/j.jeconom.2005.06.032
Eichler M (2011) A note on global Markov properties for mixed graphs. https://doi.org/10.48550/arXiv.1107.3036
Eichler M (2012) Graphical modelling of multivariate time series. Probab Theory Relat Fields 153:233–268. https://doi.org/10.1007/s00440-011-0345-8
Eichler M, Dahlhaus R, Sandkühler J (2003) Partial correlation analysis for identification of synaptic connections. Biol Cybern 89:289–302. https://doi.org/10.1007/s00422-003-0400-3
Fasen V, Fuchs F (2013a) Spectral estimates for high-frequency sampled continuous-time autoregressive moving average processes. J Time Ser Anal 34:532–551. https://doi.org/10.1111/jtsa.12029
Fasen V, Fuchs F (2013b) On the limit behavior of the periodogram of high-frequency sampled stable CARMA processes. Stoch. Process. Appl. 123:229–273. https://doi.org/10.1016/j.spa.2012.08.003
Fasen-Hartmann V, Mayer C (2022) Whittle estimation for continuous-time stationary state space models with finite second moments. Ann Inst Stat Math 74:233–270. https://doi.org/10.1007/s10463-021-00802-6
Fasen-Hartmann V, Schenk L (2024a) Mixed orthogonality graphs for continuous-time stationary processes. Stoch Process Appl 179:104501. https://doi.org/10.1016/j.spa.2024.104501
Fasen-Hartmann V, Schenk L (2024b) Mixed orthogonality graphs for continuous-time state space models and orthogonal projections. J Anal Time Ser. https://doi.org/10.1111/jtsa.12787
Fujikoshi Y, Ulyanov VV, Shimizu R (2010) Multivariate statistics. Wiley, New York. https://doi.org/10.1002/9780470539873
Gardner WA (1988) Statistical spectral analysis: a nonprobabilistic theory. Prentice-Hall, New Jersey
Gardner WA (1992) A unifying view of coherence in signal processing. Signal Process 29:113–140. https://doi.org/10.1016/0165-1684(92)90015-O
Gather U, Imhoff M, Fried R (2002) Graphical models for multivariate time series from intensive care monitoring. Stat Med 21:2685–2701. https://doi.org/10.1002/sim.1209
Hannan EJ (1970) Multiple time series. Wiley, New York
Harvey AC, Stock JH (1985a) The estimation of higher-order continuous time autoregressive models. Econom Theory 1:97–112. https://doi.org/10.1017/S0266466600011026
Harvey AC, Stock JH (1985b) Continuous time autoregressive models with common stochastic trends. J Econ Dyn Control 12:365–384. https://doi.org/10.1016/0165-1889(88)90046-2
Harvey AC, Stock JH (1989) Estimating integrated higher-order continuous time autoregressions with an application to money-income causality. J Econom 42:319–336. https://doi.org/10.1016/0304-4076(89)90056-0
Kartsioukas R, Stoev S, Hsing T (2023) Spectral density estimation of function-valued spatial processes. https://doi.org/10.48550/arXiv.2302.02247
Khintchine A (1934) Korrelationstheorie der stationären stochastischen Prozesse. Math Ann 109:604–615. https://doi.org/10.1007/BF01449156
Kleiber W (2017) Coherence for Random Fields. Stat Sin 27:1675–1697. https://doi.org/10.5705/ss.202017.0365
Klenke A (2020) Wahrscheinlichkeitstheorie, 4th edn. Springer, Berlin. https://doi.org/10.1007/978-3-662-62089-2
Lauritzen SL (2004) Graphical models, 2nd edn. Clarendon Press, Oxford
Lu T-T, Shiou S-H (2002) Inverses of 2 × 2 block matrices. Comput Math Appl 43:119–129. https://doi.org/10.1016/S0898-1221(01)00278-4
Maathuis M, Drton M, Lauritzen SL, Wainwright M (2019) Handbook of Graphical Models. CRC Press, Boca Raton. https://doi.org/10.1201/9780429463976

Marquardt T, Stelzer R (2007) Multivariate CARMA processes. Stoch Process Appl 117:96–120. https://doi.org/10.1016/j.spa.2006.05.014
Medkour T, Walden AT, Burgess A (2009) Graphical modelling for brain connectivity via partial coherence. J Neurosci Methods 180:374–383. https://doi.org/10.1016/j.jneumeth.2009.04.003
Mogensen SW, Hansen NR (2020) Markov equivalence of marginalized local independence graphs. Ann Math Stat 48:539–559. https://doi.org/10.1214/19-AOS1821
Mogensen SW, Hansen NR (2022) Graphical modeling of stochastic processes driven by correlated noise. Bernoulli 28:3023–3050. https://doi.org/10.3150/21-bej1446
Pearl J (1994) Probabilistic reasoning in intelligent systems: networks of plausible inference, 3rd edn. Morgan Kaufmann, San Francisco. https://doi.org/10.1016/C2009-0-27609-4
Pinsky MA (2009) Introduction to Fourier analysis and wavelets. American Mathematical Society, Providence. https://doi.org/10.1090/gsm/102
Priestley MB (1981) Spectral analysis and time series, vol II. Academic Press, London
Richardson T (2003) Markov properties for acyclic directed mixed graphs. Scand J Stat 30:145–157. https://doi.org/10.1111/1467-9469.00323
Rozanov JA (1967) Stationary random processes. Holden-Day, San Francisco
Schlemm E, Stelzer R (2012a) Multivariate CARMA processes, continuous-time state space models and complete regularity of the innovations of the sampled processes. Bernoulli 18:46–63. https://doi.org/10.3150/10-BEJ329
Schlemm E, Stelzer R (2012b) Quasi maximum likelihood estimation for strongly mixing state space models and multivariate Lèvy-driven-driven CARMA processes. Electron J Stat 6:2185–2234. https://doi.org/10.1214/12-EJS743
Schweder T (1970) Composable Markov processes. J Appl Probab 7:400–410. https://doi.org/10.2307/3211973
Tick LJ (1963) Conditional spectra, linear systems, and coherency. In: Proceedings of the symposium on time series analysis (Brown University, 1962). Wiley, New York, pp 197–203
Whittaker J (2008) Graphical models in applied multivariate statistics, 2nd edn. Wiley, Chichester