Available online at www.sciencedirect.com

SCIENCE DIRECT®

ELSEVIER

Statistics & Probability Letters ▮ (▮▮▮▮) ▮▮▮–▮▮▮

STATISTICS & PROBABILITY LETTERS

www.elsevier.com/locate/stapro

Latent variable analysis and partial correlation graphs for multivariate time series

Roland Fried[a,*,1], Vanessa Didelez[b]

[a]*Departamento de Estadística, Universidad Carlos III de Madrid, 28903 Getafe, Spain*
[b]*University College London, London WC1E 6BT, UK*

Received 14 January 2003

Abstract

We investigate the possibility of exploiting partial correlation graphs for identifying interpretable latent variables underlying a multivariate time series. It is shown how the collapsibility and separation properties of partial correlation graphs can be used to understand the relation between a factor model and the structure among the observable variables.

Keywords: Time series analysis; Dimension reduction; Factor analysis; Partial correlations

---

1. Introduction

Statistical modelling should appropriately reflect the correlations among the components of a multivariate time series. This claim usually leads to complex models involving numerous parameters and requiring a high amount of data to enable reliable inference. Thus, suitable strategies for dimension reduction are called for when analyzing high-dimensional processes as the available data does not often suffice to consider the full set of variables. This problem is known as the curse of dimensionality.

---

[*] Corresponding author. Tel.: 34 91 624 5854; fax: 34 91 624 9849.
E-mail address: rfried@est-econ.uc3m.es (R. Fried).

[1] Financial Support of the Deutsche Forschungsgemeinschaft (SFB 475) is gratefully acknowledged.

0167-7152/$ - see front matter © 2005 Published by Elsevier B.V.
doi:10.1016/j.spl.2005.04.002

Factor analysis is a well-known approach to reduce the observed variables to a few underlying latent variables. Peña and Box (1987) suggest the following simple generalization to model a d-variate stationary time series $\{Y_V(t) = (Y_1(t), \ldots, Y_d(t))', t \in \mathbb{Z}\}$, $V = \{1, \ldots, d\}$. They assume that there is an l-variate latent factor process $\{X_F(t), t \in \mathbb{Z}\}$ following a VARMA(p, q) model and driving the observable variables, i.e. for each time point t

$$Y_V(t) = \Lambda X_F(t) + \varepsilon(t) \tag{1}$$

is assumed, where Λ is a $d \times l$-matrix of loadings and $\{\varepsilon(t), t \in \mathbb{Z}\}$, $\varepsilon(t) \sim N(0, \Sigma_\varepsilon)$, is a d-variate white noise process, which is independent of $\{X_F(t), t \in \mathbb{Z}\}$. If model (1) holds with independent factors, i.e. if all matrices in the VARMA(p, q) model are diagonal, the autocovariance matrices $\Gamma_Y(h)$ of $\{Y_V(t), t \in \mathbb{Z}\}$ are symmetrical for $h \geqslant 1$ and the columns of Λ are the common eigenvectors of $\Gamma_Y(h)$ while the corresponding eigenvalues $\gamma_i(h)$, $i = 1, \ldots, l$, are the diagonal elements of the autocovariance matrices $\Gamma_X(h)$ of $\{X_F(t), t \in \mathbb{Z}\}$. These relations can be exploited to identify factor models.

For illustration, we analyze an 11-variate time series of vital signs (different types of blood pressures, heart rate, pulse, and blood temperature) of a critically ill patient. In a first rough analysis using model (1), we compute the eigenvalues and eigenvectors of the autocorrelation matrices $\Gamma_Y(h), h = 1, 2, 3$, i.e. the autocovariance matrices of the standardized time series (Table 1). Based on these values it seems reasonable to assume that there are four or five underlying factors. Gather et al. (2001) use four factors for a similar data situation, but without pulsoximetry, and so we decide to use five factors, here. In the present example, it is important that the factors can be interpreted by the physician who has to make decisions regarding changes of treatments. Therefore we rotate the factors in the l-dimensional space using the automatic 'varimax' procedure. The resulting loadings, shown in Table 2, allow to relate each of the factors with a physiological meaningful subset of the variables, e.g. the second factor consists mainly of the arterial pressures. In order to further improve the interpretation of the factors, Gather et al. (2001) suggest to impose restrictions on the loading matrix using physiological knowledge and the results obtained from an analysis of the partial correlations among the component processes. This seems even more important given the problems that may occur with automatic rotations w.r.t. the identification of underlying dependence structures even for i.i.d. data (Jolliffe, 1989). In the following we put the suggestions of Gather et al. (2001) on a sound basis by exploiting the factorization properties of partial correlation graphs and relating them to dynamic factor models.

Table 1
Eigenvalues of the autocorrelation matrices at the first three time lags

Lag	EV1	EV2	EV3	EV4	EV5	EV6	EV7	EV8	EV9	EV10	EV11
1	3.772	2.163	1.279	0.962	0.650	0.390	0.310	0.193	0.013	0.007	0.005
2	3.623	2.036	1.196	0.895	0.590	0.357	0.265	0.156	0.010	0.004	0.003
3	3.520	1.968	1.167	0.853	0.533	0.342	0.234	0.133	0.010	0.005	0.001

Table 2
Left: Factor loadings for physiological time series after varimax rotation. The first rotated factor can be identified mainly with the intrathoracic pressures (PAPx and CVP), the second with HR and Puls, the third with the arterial pressures, the fourth with the temperature and the fifth with SPO2. Right: Factor loadings for 'partitioned factor model'

Var.	fac. 1	fac. 2	fac. 3	fac. 4	fac. 5	fac. 1	fac. 2	fac. 3	fac. 4	fac. 5
PAPS	0.380	−0.142	−0.016	−0.031	−0.346	0.484	0	0	0	0
PAPM	0.558	−0.001	0.041	−0.067	−0.049	0.565	0	0	0	0
PAPD	0.582	0.041	−0.001	0.003	0.145	0.510	0	0	0	0
CVP	0.424	0.048	−0.047	0.413	−0.004	0.432	0	0	0	0
APS	0.002	−0.102	0.592	−0.046	0.273	0	0.530	0	0	0
APM	0.039	0.018	0.604	−0.025	−0.015	0	0.622	0	0	0
APD	−0.037	0.092	0.535	0.085	−0.276	0	0.577	0	0	0
HR	0.008	0.690	0.011	0.003	0.003	0	0	0.702	0	0
Puls	0.001	0.698	−0.003	−0.010	0.015	0	0	0.712	0	0
Temp	−0.138	−0.036	0.023	0.909	0.014	0	0	0	1	0
SPO2	0.084	−0.015	−0.010	0.011	0.844	0	0	0	0	1

2. Graph notations

Graphical models aim at analyzing the associations among a vector of variables such that they can uniquely be represented by a graph (Lauritzen, 1996). A *graph* $G = (V, E)$ consists of a finite set of *vertices* V and a set of *edges* $E \subseteq V \times V$. If only (a, b) is in E, we draw a *directed* edge (arrow) from a to b, $a \to b$, and call a a *parent* of b, and b a *child* of a. If both $(a, b) \in E$ and $(b, a) \in E$, we use an *undirected* edge (line) a–b and call a and b *neighbors*. Directed and undirected edges typically encode different dependence structures subject to the kind of graphical model. The sets of parents, children and neighbors of $a \in V$ are denoted by $\mathrm{pa}(a)$, $\mathrm{ch}(a)$ and $\mathrm{ne}(a)$, respectively. Similarly, we define the parents, children and neighbors of a subset $A \subseteq V$ by $\mathrm{pa}(A) = \bigcup_{a \in A} \mathrm{pa}(a) \backslash A$, $\mathrm{ch}(A) = \bigcup_{a \in A} \mathrm{ch}(a) \backslash A$ and $\mathrm{ne}(A) = \bigcup_{a \in A} \mathrm{ne}(a) \backslash A$. The *boundary* of A is $\mathrm{bd}(A) = \mathrm{pa}(A) \cup \mathrm{ne}(A)$. If $\mathrm{bd}(A) = \emptyset$ we call A an *ancestral* set. The *closure* $\mathrm{cl}(A)$ of A is $A \cup \mathrm{bd}(A)$. The *subgraph* G_A of G induced by A is obtained by eliminating all vertices except those in A and all edges (a, b) not contained in $A \times A$. A *path* from $a \in V$ to $b \in V$ is a sequence of vertices $a = a_0, \ldots, a_m = b$, $m \geqslant 1$, such that $(a_{i-1}, a_i) \in E$, $i = 1, \ldots, m$, and is denoted by $a \mapsto b$. If both $a \mapsto b$ and $b \mapsto a$ we say that a and b are *connected*. Connectivity defines an equivalence relation and the equivalence classes are called *connectivity components*.

In order to address factor models we will make use of *chain graphs*. The vertex set V of such a chain graph can be partitioned into disjoint subsets $B(j)$, $V = B(1) \cup \cdots \cup B(k)$, such that all edges between vertices in the same subset are undirected and all edges between different subsets are directed, pointing from the subset with the lower number to the subset with the higher number. We assume w.l.o.g. that $B(1), \ldots, B(k)$ are connectivity components and call them *chain components*, while $C(j) = B(1) \cup \cdots \cup B(j)$ is called set of *concurrent variables*, $j = 1, \ldots, k$. For a chain graph G we define its *moral graph* G^m as the undirected graph with the same vertex set but with a–b in G^m iff, in G, we have a–b, $a \to b$, $b \to a$ or if there are c_a, c_b in the same chain component such that $a \to c_a$ and $b \to c_b$.

Undirected graphs (no directed edges) are special cases of chain graphs, where $V = B(1)$ in case of a single connectivity component. In such graphs, subsets $A, B \subset V$ are *separated* by $S \subset V$ if any path from every $a \in A$ to $b \in B$ intersects S. An undirected graph that contains all possible edges is called *complete*. It typically represents the saturated model.

3. Partial correlation graphs

Brillinger (1996) and Dahlhaus (2000) introduce partial correlation graphs for multivariate time series to represent the essential linear, possibly time-lagged relations among the components which remain after eliminating the linear effects of the other variables. We consider throughout the paper a vector-valued weakly stationary time series $\{Y_V(t), t \in \mathbb{Z}\}$, $V = \{1, \ldots, d\}$, and denote it briefly by Y_V. Similarly, for $A \subseteq V$ we denote the subprocess of all variables $a \in A$ by Y_A. We further assume that the covariance function $\gamma_{ab}(h) = \mathrm{Cov}(Y_a(t+h), Y_b(t))$ is absolutely summable with respect to all time lags $h \in \mathbb{Z}$ for all pairs $a, b \in V$. Then the *cross spectrum* between the time series Y_a and Y_b is defined as the Fourier transform of their covariance function:

$$f_{Y_aY_b}(\lambda) = \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} \gamma_{ab}(h) \exp(-\mathrm{i}\lambda h).$$

The variables Y_a and Y_b are uncorrelated at all time lags h iff $f_{ab}(\lambda)$ equals zero for all frequencies.

We are interested in the *partial* correlations, adjusting for the linear effects of the remaining variables on Y_a and Y_b. This is done by considering the residual series $\varepsilon_a(t)$ and $\varepsilon_b(t)$ obtained by subtracting all linear influences of $Y_{V\setminus\{a,b\}}$ from $Y_a(t)$ and $Y_b(t)$, respectively (Brillinger, 1981). The cross spectrum between the series ε_a and ε_b then yields the *partial cross spectrum* of Y_a and Y_b, $f_{Y_aY_b\cdot V\setminus\{a,b\}}(\lambda) = f_{\varepsilon_a\varepsilon_b}(\lambda)$. The (partial) cross spectrum between two vector time series Y_A and Y_B, $A, B \subset V$, can be defined in a similar way. The *partial spectral coherency* is a standardization of the partial cross spectrum

$$R_{Y_aY_b\cdot Y_{V\setminus\{a,b\}}}(\lambda) = \frac{f_{Y_aY_b\cdot Y_{V\setminus\{a,b\}}}(\lambda)}{[f_{Y_aY_a\cdot Y_{V\setminus\{a,b\}}}(\lambda) f_{Y_bY_b\cdot Y_{V\setminus\{a,b\}}}(\lambda)]^{1/2}}. \tag{2}$$

With these definitions, the *partial correlation graph* of a multivariate time series is given as the undirected graph $G = (V, E)$, where two vertices a and b are connected by an undirected edge whenever the partial spectral coherency $R_{Y_aY_b\cdot Y_{V\setminus\{a,b\}}}(\cdot)$ is not identical to zero. A missing edge between a and b is denoted by $a \perp b | V\setminus\{a,b\}$ and indicates that the linear relation between these two variables given all the others is zero at all time lags. This relation between a graph and the partial correlation structure is known as *undirected pairwise Markov property* (PU). Under the assumption that the spectral density matrix is regular for all frequencies, the PU implies the undirected *global Markov property*, a stronger property in general. The latter states that $A \perp B | S$ for all subsets $A, B, S \subset V$, whenever S separates A and B in G. It is plausible to consider undirected graphs because the residual series are adjusted not only for the past but also for the future effects so that the graph cannot carry any information on the dynamics of the dependencies.

4. Chain graphs and dynamic factor models

In the following, we derive what a partial correlation graph of an observed time series should look like given an underlying factor model. This allows to derive suitable restrictions for a factor model from a preliminary data analysis using partial correlation graphs. Particularly, the resulting graph provides an assistance in identifying the number and types of factors. Throughout this section, we assume that the spectral density matrix of the multivariate stationary time series Y_V is regular at all frequencies.

The first proposition needed gives a condition which ensures that missing edges in a subgraph can still be regarded as zero partial correlations within the corresponding subprocess after marginalizing over the remaining components (see Fried and Didelez, 2003).

Proposition 1. *Let $G = (V, E)$ be the partial correlation graph of a multivariate time series. If the boundary of each connectivity component of $B \subset V$ is complete then $G_{V \backslash B}$ is not smaller than the partial correlation graph of the subprocess $X_{V \backslash B}$, i.e. $G_{V \backslash B}$ has the same or more edges than the latter. We say that G is collapsible on to $V \backslash B$ (or over B).*

In order to derive partial correlation graphs for time series models with latent variables, we next define *partial correlation chain graphs*. The idea is that factor models consist of two building blocks: The first one reflects the assumptions about the interdependence among the underlying factors; this constitutes the first chain component $B(1)$. Then we model the distribution of the observable variables given the factors; this constitutes $B(2)$, and the conditional distribution of $B(2)$ given $B(1)$ is specified by some suitable model.

The implementation of this idea requires the generalization of the notion of a chain graph to time series. While time series models are often thought to be causal in time, some time series methods like dynamic principal component analysis (Brillinger, 1981) apply noncausal filters with nonzero weights for past and future observations. The following definition is designed for the latter case due to our interest for such latent variable techniques. We define a partial correlation chain graph $G = (V, E)$ by the *pairwise block-recursive Markov property* (PB) relatively to a dependence chain $B(1), \ldots, B(k)$. It states that for any pair a, b of nonadjacent vertices we have

$$a \perp b | C(j^\star) \backslash \{a, b\},$$

where $j^\star$ is the smallest $j \in \{1, \ldots, k\}$ with $a, b \in C(j)$. We consider two further Markov properties that are commonly used for i.i.d. data. The *global chain graph Markov property* (GC) states that $A \perp B \mid S$ for all subsets A, B, S of V such that S separates A and B in $(G_{\mathrm{An}(A \cup B \cup S)})^m$, which is the moral graph of the smallest ancestral subgraph containing $A \cup B \cup S$. The *pairwise chain Markov property* (PC) states $a \perp b \mid \mathrm{nd}(a) \backslash \{b\}$, whenever a, b are nonadjacent and $b \in \mathrm{nd}(a)$. Obviously, we have (GC) $\Rightarrow$ (PC) $\Rightarrow$ (PB). In order to prove that these properties are even equivalent, provided that the spectral density matrix is regular everywhere, we first state another result, which is also interesting by itself.

Proposition 2. *If the PC is satisfied with respect to a partial correlation chain graph G, then the PU is satisfied w.r.t. G^m, too.*

Proof. Proposition 2 can be proven along the same lines as Lemma 3.33 in Lauritzen (1996, p. 56f) using Lemma 3.1(ii) in Dahlhaus (2000). It only requires the property $X \perp Y \mid Z \Rightarrow h(X) \perp Y \mid Z$ for any component selection function h, which is satisfied for zero partial correlation. □

Proposition 3. *If the spectral density matrix is regular everywhere then the PB, the PC and the pairwise GC for partial correlation chain graphs are equivalent.*

Proof. Proposition 3 can be proven in the same way as Theorem 3.34 in Lauritzen (1996, p. 57f) using Lemma 3.1(ii) in Dahlhaus (2000) and Proposition 2. □

Partial correlation chain graphs are most useful for hierarchical time series models of which factor models are a special case. Assume that

$$Y_{B(j)}(t) = \sum_{i=1}^{j-1} \sum_{h=-\infty}^{\infty} \Lambda^{j,i}(h) Y_{B(i)}(t-h) + \varepsilon_{B(j)}(t),$$

$$\varepsilon_{B(j)}(t) = \sum_{h=1}^{p} \Theta_j(h) \varepsilon_{B(j)}(t-h), \quad j = 1, \ldots, k,$$

i.e. $\varepsilon_{B(j)}$ follows a VAR-model, where the elements $\Theta_j(h)_{b,a}$ of $\Theta_j(h)$ denote the influence of variable a in the regression of b on the other variables. The partial correlation chain graph of the whole multivariate time series obeying the above model is given by the following algorithm, where we make use of the results of Dahlhaus (2000) for VAR-processes.

Construction of partial correlation chain graph:

1. Starting with $B(1)$. Connect each pair $(a, b) \in B(1) \times B(1)$ whenever $\Theta_j(h)_{a,b} \neq 0$ or $\Theta_j(h)_{b,a} \neq 0$ for any $h \in \{1, \ldots, p\}$, or if $c \in B(1)$ and $h_a, h_b \in \{1, \ldots, p\}$ exist such that $\Theta_j(h_a)_{c,a} \neq 0$ and $\Theta_j(h_b)_{c,b} \neq 0$.
2. Draw vertices for the variables in $B(2)$, connect the pairs of variables in $B(2)$ by a line using an analogous rule as in step 1, and draw an arrow from $a \in B(1)$ to $b \in B(2)$ if (with obvious notation) $\Lambda^{2,1}(u)_{b,a} \neq 0$ for any $u \in \mathbb{Z}$.
3. Repeat step 2 for $B(3), \ldots, B(k)$ drawing an arrow from a variable $a \in B(i)$ to a variable $b \in B(j), j > i$, if $\Lambda^{j,i}(u)_{b,a} \neq 0$ for any $u \in \mathbb{Z}$, and using the rule stated above for connecting pairs of variables in $B(j)$ by lines.

To show that this construction is valid, we only need to prove that steps 1–3 are correct for the construction of the partial correlation chain graph, i.e. we have to prove (PB) for the resulting graph. This can be done by induction on j. The correctness for $j = 1$ is verified by Dahlhaus (2000) as $\varepsilon_{B(1)}$ is a VAR(p)-process. Now assume that the statement is true for all $j \leqslant n$. In order to prove correctness for $j = n + 1$ let w.l.o.g. $b \in B(n+1)$, and assume $a \in C(n+1)$ is nonadjacent to b. As $\varepsilon_b \perp Y_{C(n)}$ the regression coefficients for variables $a \in C(n)$ when regressing Y_b on $Y_{C(n+1)\backslash\{b\}}$ are given by the elements $(\Lambda^{j,i}(u))_{b,a}, u \in \mathbb{Z}$. Hence, $a \perp b \mid C(n+1)\backslash\{a, b\}$ for any nonadjacent $a \in C(n)$ follows from Proposition 3 in Fried and Didelez (2003). The coefficients for $a \in B(n+1)$ are the same as the coefficients for a in the regression of ε_b on $Y_{B(n+1)\backslash\{b\}}$. As $\varepsilon_b \perp \{\sum_{i=1}^{n} \sum_{u=-\infty}^{\infty} \Lambda^{j,i}(u) Y_{B(i)}(t-u)\}$ these coefficients are zero if the coefficients of ε_a in the regression of ε_b on $\varepsilon_{B(n+1)\setminus\{b\}}$ are zero, and this in turn is equivalent to $\varepsilon_b \perp \varepsilon_a | \varepsilon_{B(n+1)\setminus\{a,b\}}$, which proves the result. □

Now we have all necessary tools available for constructing the partial correlation graph for the observed variables generated by a dynamic factor model, where $k = 2$. A general stationary dynamic factor model is given by

$$Y_V(t) = \sum_{h=-\infty}^{\infty} \Lambda(h) X_F(t-h) + \varepsilon_V(t),$$

with an unobserved factor series and an error series following VAR(p)-processes:

$$X_F(t) = \sum_{h=1}^{p} \Phi(h) X_F(t-h) + \eta(t), \quad \varepsilon_V(t) = \sum_{h=1}^{p} \Theta(h) \varepsilon_V(t-h) + \delta(t).$$

We note that the model in this very general form is not identifiable but it can serve to investigate which information on the model structure can be gained from partial correlation graphs without imposing any further restrictions.

First, we have to construct the partial correlation chain graph, according to the above algorithm, with $Y_{B(1)} = X_F$, $Y_{B(2)} = Y_V$. Then we moralize this chain graph, according to Proposition 2, obtaining the partial correlation graph for (Y_V, X_F). Finally, we marginalize this moral graph w.r.t. X_F by applying Proposition 1 for all collapsible connectivity components of $Y_B(1) = X_F$ and completing the boundaries of noncollapsible components in $Y_B(2) = Y_V$. This yields the partial correlation graph G_Y of Y_V. It is easy to see that all subgraphs of G_Y on variables that are affected by the same underlying factor will be complete. Therefore, it is straightforward to detect possible factors from the partial correlation graph of the observable time series by identifying such complete subsets. However, the identification of common factors can be obscured since dependencies within the error process $\varepsilon_V(t)$ can cause additional edges in G_Y. Nevertheless, it seems reasonable to attribute strong relations to the factors and weaker ones to the errors.

5. Application to physiological time series

The ideas of the previous section are now applied to detect the partial linear relations and underlying factors in the physiological time series mentioned in the Introduction. To begin with, the cross spectra are estimated from the data, and then the partial spectral coherencies are computed using Eq. (2). For our calculations we use the program Spectrum (Dahlhaus and Eichler, 2000) which is based on a nonparametric kernel estimator. In this first step, the partial spectral coherencies are estimated in the saturated model.

As relations among (physiological) variables may have different strengths we classify the empirical partial relations into strong (S), moderate (M), weak (W) and negligible (N) partial correlation on the basis of the area under the estimated partial spectral coherence. This area can be measured by the partial mutual information between Y_a and Y_b:

$$-\frac{1}{2\pi} \int \log\{1 - |R_{Y_a Y_b \cdot Y_{V\setminus\{a,b\}}}(\lambda)|^2\} \, \mathrm{d}\lambda.$$

The resulting partial correlation graph is shown in Fig. 1 with distinct edges for different classifications and negligible edges omitted.

In a second step, we verify the obtained graph by exploiting its collapsibility properties such as described in Proposition 1 (cf. also Fried and Didelez, 2003). Consider a missing edge (a, b): If G is the partial correlation graph for Y_V then Y_V also satisfies the pairwise Markov property w.r.t. the graph G' with $\mathrm{cl}(a)$ as well as $\mathrm{bd}(a) \cup \{b\}$ made complete. Then Proposition 1 applies to G' with $B = V \backslash (\mathrm{cl}(a) \cup \{b\})$ and we find that an edge between a and b is missing in G if it is missing in G'_A, where $A = (\mathrm{cl}(a) \cup \{b\})$. Therefore, we can restrict testing the existence of an edge (a, b) to the subprocess Y_A. This allows to double check the previous classifications in a stepwise procedure. However, we do not change the initial classification by more than one class.

Since false omission of an edge is more serious than false inclusion because it induces more restrictions than supported by the data, we start by verifying the edges classified as (N). We can, e.g. check the missing edges (HR,APD), (HR,APM) and (HR,APS) applying Proposition 1 to $\{\mathrm{APM}, \mathrm{APD}, \mathrm{APS}, \mathrm{HR}, \mathrm{SPO2}\}$. We find that only the partial mutual information for (HR,APM) is increased while the others remain about the same. Therefore, we reclassify this edge as (W). A similar argument leads to the reclassification of the edges (APM,PULS), (APM,PAPM) and (APM,CVP) as (W).

Next we look at the edges in (W). We find the partial mutual information for (CVP,HR) to be very small when considering the subgraph on $\{\mathrm{CVP}, \mathrm{HR}, \mathrm{PULS}, \mathrm{SPO2}, \mathrm{APM}\}$. Hence, we reclassify this edge as (N). Similarly, we find (SPO2,PULS) and (APS,SPO2) to be negligible based on $\{\mathrm{SPO2}, \mathrm{PULS}, \mathrm{HR}, \mathrm{CVP}, \mathrm{TEMP}\}$ and $\{\mathrm{APS}, \mathrm{SPO2}, \mathrm{Temp}, \mathrm{Puls}, \mathrm{HR}\}$.

Since we could eliminate some edges in the previous step we obtain more graph separations, that can be used for further double checking. In particular, we reinvestigate the relations between CVP and the pulmonary pressures based on $\{\mathrm{CVP}, \mathrm{SPO2}, \mathrm{Temp}, \mathrm{PAPx}\}$ with $\mathrm{PAPx} \in \{\mathrm{PAPD}, \mathrm{PAPM}, \mathrm{PAPS}\}$, where APM has to be included when $\mathrm{PAPx} = \mathrm{PAPM}$. We find all these edges to be significant and the partial mutual information to be much higher for (CVP,PAPx) than, e.g. for (CVP,SPO2). This suggests that conditioning on the other pulmonary

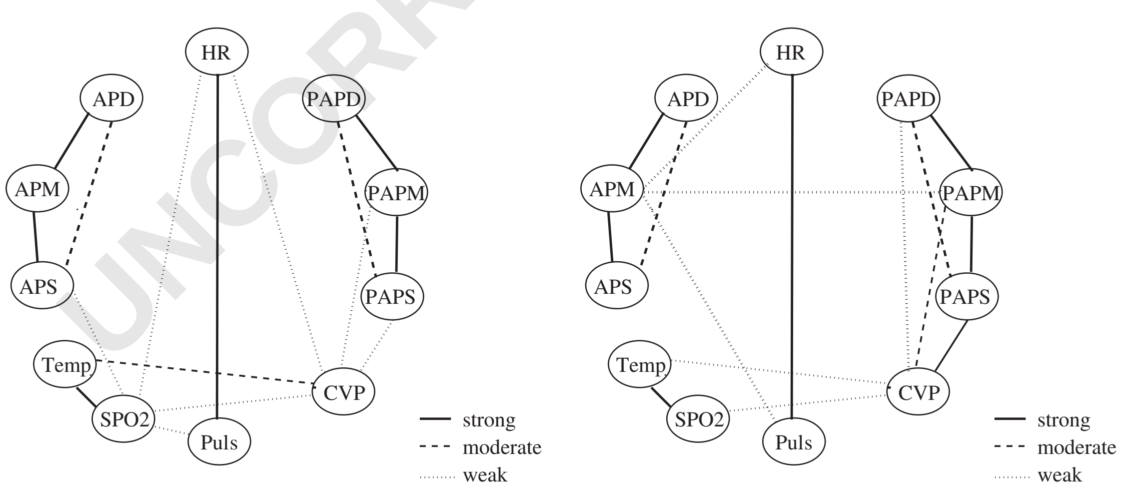

Fig. 1. Partial correlation graph for hemodynamic system, one-step selection (left) and final selection (right).

Table 3
Percentage of nonexplained variation: Factor model (top), partitioned factor model (bottom).

PAPS	PAPM	PAPD	CVP	APS	APM	APD	HR	Puls	Temp	SPO2
0.268	0.041	0.181	0.199	0.146	0.024	0.168	0.022	0.019	0.053	0.128
0.309	0.070	0.234	0.412	0.271	0.027	0.227	0.013	0.012	0.000	0.000

pressures hides some of the relations, in particular those to CVP. Indeed, the pulmonary arterial pressures and CVP are jointly denoted as intrathoracic pressures because of their well-known physiological association.

Further double checking of the remaining edges does not lead to any more alterations of the graph. The final model found by our stepwise search is also depicted in Fig. 1. It shows strong relations among the arterial pressures, among the heart rate and the pulse, as well as among the intrathoracic pressures. In addition, there are some weak relations. The strong relation between SPO2 and Temp is caused by a systematic error of the measurement instruments, of which the physicians were unaware before. The other results agree with medical knowledge.

Disregarding the edges classified as (W), the final partial correlation graph consists of four complete subgraphs, just like the partial correlation graph for a dynamic factor model with four independent factors. This seems to justify the assumption of a separate factor for each of these groups of variables, respectively. As we believe the relation between Temp and SPO2 to be a measurement artifact, we also treat them separately.

When applying the Peña–Box dynamic factor model to the clusters of variables identified above, we find one factor to be sufficient for each group. The resulting factor loadings are provided in Table 2, and a comparison of the residual variances for the factor model for all variables and the 'partitioned' factor model is given in Table 3. Most of the variables are explained almost equally well by both models. The residual variance in the simpler partitioned model is substantially larger for CVP, only. If we assume two factors for the group of intrathoracic pressures, we find the second factor to be essentially the difference between PAPS and CVP.

6. Conclusion

Statistical methods for dimension reduction aim at condensing the information provided by a high-dimensional time series into a few essential variables. In this regard, partial correlation graphs are a suitable tool: On the one hand, they help to explore the relations among the observable variables. On the other hand, they can be used to identify suitable rotations of the loading matrices in dynamic factor analysis, or even to partition the variables according to clusters of closely related variables. With this kind of information we can identify meaningful and interpretable factor models as we have demonstrated in the present paper. This is particularly important as automatic rotations are difficult to apply when a more complicated dynamic factor model with nonzero loadings at various time lags is used. However, very strong relations among some of the variables may hide other, weaker relations or even cause spurious relations, thus misleading the initial analysis of the partial correlation structure. Using the stepwise selection procedure suggested by Fried and Didelez (2003), and further refined here, seems a promising alternative.

Acknowledgements

We thank Rainer Dahlhaus and Michael Eichler for providing the program ‘Spectrum’.

References

Brillinger, D.R., 1981. Time Series. Data Analysis and Theory. Holden Day, San Francisco.

Brillinger, D.R., 1996. Remarks concerning graphical models for time series and point processes. Rev. Econom. 16, 1–23.

Dahlhaus, R., 2000. Graphical interaction models for multivariate time series. Metrika 51, 157–172.

Dahlhaus, R., Eichler, M., 2000. SPECTRUM, a C program to calculate and test partial spectral coherences. Available via http://www.statlab.uni-heidelberg.de/projects/.

Fried, R., Didelez, V., 2003. Decomposability and selection of graphical models for multivariate time series. Biometrika 90, 251–267.

Gather, U., Fried, R., Lanius, V., Imhoff, M., 2001. Online monitoring of high-dimensional physiological time series—a case study. Estadística 53, 259–298.

Jolliffe, I.T., 1989. Rotation of ill-defined principal components. Appl. Statist. 38, 139–147.

Lauritzen, S.L., 1996. Graphical Models. Clarendon Press, Oxford.

Peña, D., Box, G.E.P., 1987. Identifying a simplifying structure in time series. J. Amer. Statist. Assoc. 82, 836–843.