Appendix to

# Learning Block Structured Graphs in Gaussian Graphical Models

published in the Journal of Computational and Graphical Statistics

Alessandro Colombi[1], Raffaele Argiento[2], Lucia Paci[3] and Alessia Pini[3]

[1]Department of Economics, Management and Statistics, Università degli studi di Milano-Bicocca

[2]Department of Economics, Università degli studi di Bergamo

[3]Department of Statistical Sciences, Università Cattolica del Sacro Cuore.

## A Deriving the acceptance-rejection probability

The BDRJ algorithm considers switching between $(\boldsymbol{K}^{[s]}, G^{[s]}, \widetilde{\boldsymbol{W}}, G')$ to the alternative $(\boldsymbol{K}', G', \boldsymbol{W}^0, G^{[s]})$, where $\boldsymbol{W}^0 \in \mathbb{P}_{G^{[s]}}$, by performing two Reversible Jump moves: (i) a dimension increasing step from $(\boldsymbol{K}^{[s]}, G^{[s]})$ to $(\boldsymbol{K}', G')$ according to posterior parameters $b+n$ and $\boldsymbol{D}+\boldsymbol{U}$ and (ii) a dimension decreasing step from $(\widetilde{\boldsymbol{W}}, G')$ to $(\boldsymbol{W}^0, G^{[s]})$ according to prior parameters $b$ and $\boldsymbol{D}$.

As mentioned in Section 3, the proposed graph $G'$ is obtained from Equation (4) by adding the edge $(l, m)$ to the multigraph representation of $G^{[s]}$. Regarding the precision matrix, the double reversible jump is performed by leveraging on the change of variables $\boldsymbol{K} \mapsto \boldsymbol{\Phi}$, where $\boldsymbol{\Phi}$ is an upper triangular matrix such that $\boldsymbol{K} = \boldsymbol{\Phi}^T\boldsymbol{\Phi}$, see Section 3. We set $\boldsymbol{\Phi}'_{ij} = \boldsymbol{\Phi}_{ij}$ for all $(i, j) \in \nu\left(G^{[s]}\right)$. The free elements are the ones in the set $L$ that are proposed by perturbing the old values independently and with the same variance $\sigma_g^2$, which is a tuning parameter. Namely, we draw $\eta_h \overset{\text{ind}}{\sim} \mathrm{N}\left(\boldsymbol{\Phi}_h^{[s]}, \sigma_g^2\right)$ and set $\boldsymbol{\Phi}'_h = \eta_h$ for each $h \in L$. This defines all the free elements of $\boldsymbol{\Phi}'$, while the non-free elements are determined through the completion operation (Atay-Kayis and Massam 2005). Hence, $\boldsymbol{\Phi}'$ is well defined as well as $\boldsymbol{K}' = (\boldsymbol{\Phi}')^T \boldsymbol{\Phi}'$.

The probability to accept the proposed values of $(\boldsymbol{K}', G')$ is equal to $\min(1, R^{+})$, where

$$R^{+} = \frac{p\left(\boldsymbol{K}', G', \boldsymbol{W}^0, G^{[s]} \mid \mathbf{y}\right)}{p\left(\boldsymbol{K}^{[s]}, G^{[s]}, \widetilde{\boldsymbol{W}}, G' \mid \mathbf{y}\right)} \frac{J\left(\boldsymbol{K}' \to \boldsymbol{\Phi}'\right) J\left(\boldsymbol{W}^0 \to \boldsymbol{\Phi}^0\right)}{J\left(\boldsymbol{K}^{[s]} \to \boldsymbol{\Phi}^{[s]}\right) J\left(\widetilde{\boldsymbol{W}} \to \widetilde{\boldsymbol{\Phi}}\right)} \frac{q\left(\boldsymbol{K}' \mid \boldsymbol{K}^{[s]}\right)}{q\left(\boldsymbol{W}^0 \mid \widetilde{\boldsymbol{W}}\right)}, \tag{11}$$

where $J(A \to B)$ denotes the Jacobian of the transformation from A to B. As usual with discrete spaces, in Equation (11), the Jacobian needed for matching the dimensions of the compared states has been omitted since it reduces to the determinant of the identity matrix.

First, we recall that the Cholesky decomposition of the precision matrix discussed in Section 3 allows us to easily compute the determinant of $\boldsymbol{K}$, see Roverato (2002), that is

$$\det(\boldsymbol{K}) = \prod_{i=1}^{p} \boldsymbol{\Phi}_{ii}^2. \tag{12}$$

Note that this formulation involves only diagonal values of $\boldsymbol{\Phi}$, which are free elements by definition. Hence, Equation (12) implies that $\det(\boldsymbol{K}') = \det\left(\boldsymbol{K}^{[s]}\right)$.

The first ratio in Equation (11) can be factorized as follows:

$$
\begin{aligned}
\frac{p\left(\boldsymbol{K}', G', \boldsymbol{W}^0, G^{[s]} \mid \mathbf{y}\right)}{p\left(\boldsymbol{K}^{[s]}, G^{[s]}, \widetilde{\boldsymbol{W}}, G' \mid \mathbf{y}\right)} &= \frac{p\left(\mathbf{y} \mid \boldsymbol{K}', G'\right)}{p\left(\mathbf{y} \mid \boldsymbol{K}^{[s]}, G^{[s]}\right)} \frac{p(\boldsymbol{K}' \mid G')}{p\left(\boldsymbol{K} \mid G^{[s]}\right)} \frac{q\left(G^{[s]} \mid G'\right)}{q\left(G' \mid G^{[s]}\right)} \\
&\times \frac{p\left(\boldsymbol{W}^0 \mid G^{[s]}\right)}{p\left(\widetilde{\boldsymbol{W}} \mid G'\right)} \frac{\pi(G')}{\pi\left(G^{[s]}\right)} \\
&= \frac{\sqrt{|\boldsymbol{K}'|}}{\sqrt{|\boldsymbol{K}|}} \exp\left\{-\frac{1}{2}\left\langle \boldsymbol{K}' - \boldsymbol{K}^{[s]},\ U\right\rangle\right\} \\
&\times \frac{I_{G^{[s]}}(b, D)}{I_{G'}(b, D)} \exp\left\{-\frac{1}{2}\left\langle \boldsymbol{K}' - \boldsymbol{K}^{[s]},\ D\right\rangle\right\} \frac{\left|nbd_M^{\mathscr{B},+}\left(G_B^{[s]}\right)\right|}{\left|nbd_M^{\mathscr{B},-}\left(G'_B\right)\right|} \\
&\times \frac{I_{G'}(b, D)}{I_{G^{[s]}}(b, D)} \frac{1}{\exp\left\{-\frac{1}{2}\left\langle \widetilde{\boldsymbol{W}} - \boldsymbol{W}^0,\ D\right\rangle\right\}} \frac{\pi(G')}{\pi(G^{[s]})} \\
&= \frac{\exp\left\{-\frac{1}{2}\left\langle \boldsymbol{K}' - \boldsymbol{K}^{[s]},\ D + U\right\rangle\right\}}{\exp\left\{-\frac{1}{2}\left\langle \widetilde{\boldsymbol{W}} - \boldsymbol{W}^0,\ D\right\rangle\right\}} \frac{\left|nbd_M^{\mathscr{B},+}\left(G_B^{[s]}\right)\right|}{\left|nbd_M^{\mathscr{B},-}\left(G'_B\right)\right|} \frac{\pi(G')}{\pi\left(G^{[s]}\right)},
\end{aligned} \tag{13}
$$

where $\langle A, B\rangle$ denotes the trace of the product between $A$ and $B$. Note that the two ratios of G-Wishart densities allow us to eliminate the presence of their normalizing constants. Also, note that, thanks to Equation (12), all the determinants of the matrices in Equation (13) canceled out.

For what concerns the change of variable from a precision matrix to its Cholesky decomposition, the Jacobian of such a transformation is

$$
J(\boldsymbol{K} \mapsto \boldsymbol{\Phi}) = 2^p \prod_{i=1}^{p} \boldsymbol{\Phi}_{ii}^{\nu_i^G}, \tag{14}
$$

where $\nu_i^G = |\{j : j > i \text{ and } (i, j) \in E\}|$ is the sum of elements in $i$-th row of the adjacency matrix, from position $i + 1$ up to the end. Then, the ratio of the Jacobians appearing in Equation (11) is readily computed using Equation (14). That is,

$$
\frac{J(\boldsymbol{K}' \to \boldsymbol{\Phi}')}{J(\boldsymbol{K}^{[s]} \to \boldsymbol{\Phi}^{[s]})} = \frac{2^p}{2^p} \frac{\prod_{i=1}^{p} \left(\boldsymbol{\Phi}'_{ii}\right)^{\nu_i^{G'}+1}}{\prod_{i=1}^{p} \left(\boldsymbol{\Phi}_{ii}^{[s]}\right)^{\nu_i^G+1}} = \prod_{i \in V(L)} \left(\boldsymbol{\Phi}_{ii}^{[s]}\right)^{\nu_i^{G'} - \nu_i^{G^{[s]}}}. \tag{15}
$$

The last equality follows by noticing that $\nu_i(G) = \nu_i(G')$ for all $i \neq V(L)$ and those diagonal elements are not modified by construction. Analogously, one can show that

$$\frac{J(\boldsymbol{W}^0 \to \boldsymbol{\Phi}^0)}{J(\widetilde{\boldsymbol{W}} \to \widetilde{\boldsymbol{\Phi}})} = \prod_{i \in V(L)} \left(\boldsymbol{\Phi}_{ii}^0\right)^{\nu_i^{G'} - \nu_i^{G^{[s]}}} .$$

Under the assumption that $G'$ is obtained by adding edge $(l, m)$ to the multigraph representation of $G^{[s]}$, the exponent in (15) reduces to

$$\nu_i^{G'} - \nu_i^{G^{[s]}} = |\{j \in B_m : j > i\}|,$$

which is equal to the number of nodes in group $m$ whose index is greater than $i$.

Finally, the last ratio in Equation (11) is due to the randomness in the construction of the proposed and the auxiliary matrices. By definition, each term is just the ratio of independent multivariate Gaussian densities, i.e.,

$$q\left(\boldsymbol{K}' \mid \boldsymbol{K}^{[s]}\right) = \left(\frac{1}{\sqrt{2\pi\sigma_g^2}}\right)^{|L|} \exp\left\{-\frac{1}{2\sigma_g^2} \sum_{h \in L} \left(\boldsymbol{\Phi}'_h - \boldsymbol{\Phi}_h\right)^2\right\},$$

where, for sake of clarity, we explicitly wrote the ratio in terms of $\boldsymbol{\Phi}'$. Similarly, we obtain the quantity $q\left(\boldsymbol{W}^0 \mid \widetilde{\boldsymbol{W}}\right) \propto \exp\left\{-\frac{1}{2\sigma_g^2} \sum_{h \in L} \left(\boldsymbol{\Phi}_h^0 - \widetilde{\boldsymbol{\Phi}}_h\right)^2\right\}$.

Wrapping everything together, we end up with

$$R^+ = \frac{\exp\left\{-\frac{1}{2}\left\langle \boldsymbol{K}' - \boldsymbol{K}^{[s]},\ D + U\right\rangle\right\}}{\exp\left\{-\frac{1}{2}\langle \widetilde{\boldsymbol{W}} - \boldsymbol{W}^0,\ D\rangle\right\}} \prod_{i \in V(L)} \left(\frac{\boldsymbol{\Phi}_{ii}^{[s]}}{\boldsymbol{\Phi}_{ii}^0}\right)^{\nu_i^{G'} - \nu_i^{G^{[s]}}}$$
$$\times \exp\left\{\frac{1}{2\sigma_g^2} \sum_{h \in L} \left[\left(\boldsymbol{\Phi}'_h - \boldsymbol{\Phi}_h^{[s]}\right)^2 - \left(\boldsymbol{\Phi}_h^0 - \widetilde{\boldsymbol{\Phi}}_h\right)^2\right]\right\} \frac{\pi(G')}{\pi\left(G^{[s]}\right)}.$$

**Algorithm 1:** Block Double Reversible Jump

Suppose the chain to be in state $\left(\boldsymbol{K}^{[s]}, G^{[s]}\right)$, with $\boldsymbol{K}^{[s]} = \left(\boldsymbol{\Phi}^{[s]}\right)^{\top}\left(\boldsymbol{\Phi}^{[s]}\right) \in P_{G^{[s]}}$ and $G^{[s]} \in \mathscr{B}$.

For each iteration:

**Step 1.** Updating the graph $G$

1.1. Sample $G'_B$ from $q\left(G'_B \mid G^{[s]}\right)$ given by (4). Set $G' = \rho^{-1}\left(G^{[s]}\right)$. Suppose an addition move is selected. Call $L$ the set of new edges.

1.2. Draw $\widetilde{\boldsymbol{W}} \mid G' \sim$ G-Wishart$(b, D)$ from an exact sampler (Lenkoski 2013).

1.3. For each $h \in L$, draw $\eta_h \sim N\left(\boldsymbol{\Phi}_h^{[s]}, \sigma_g^2\right)$

1.4. Set $(\boldsymbol{\Phi}')^{\nu\left(G^{[s]}\right)} = \left(\boldsymbol{\Phi}^{[s]}\right)^{\nu\left(G^{[s]}\right)}$ and $\boldsymbol{\Phi}'_h = \eta_h \quad \forall h \in L$.

Derive the remaining elements by completion operation and define $\boldsymbol{K}' = (\boldsymbol{\Phi}')^{\top}\boldsymbol{\Phi}'$.

1.5. Set $(\boldsymbol{\Phi}^0)^{\nu\left(G^{[s]}\right)} = (\widetilde{\boldsymbol{\Phi}})^{\nu\left(G^{[s]}\right)}$. Derive the remaining elements by completion operation and define $\boldsymbol{W}^0 = \left(\boldsymbol{\Phi}^0\right)^{\top}\boldsymbol{\Phi}^0$.

1.6. Compute $\gamma\left(\left(\boldsymbol{K}^{[s]}, G^{[s]}\right) \to (\boldsymbol{K}', G')\right) = \min\{1, R^+\}$ where

$$R^+ = \frac{\exp\left\{-\frac{1}{2}\left\langle \boldsymbol{K}' - \boldsymbol{K}^{[s]},\ D + U\right\rangle\right\}}{\exp\left\{-\frac{1}{2}\left\langle \widetilde{\boldsymbol{W}} - \boldsymbol{W}^0,\ D\right\rangle\right\}} \prod_{i \in V(L)} \left(\frac{\boldsymbol{\Phi}_{ii}^{[s]}}{\boldsymbol{\Phi}_{ii}^{0}}\right)^{\nu_i^{G'} - \nu_i^{G^{[s]}}}$$
$$\times \exp\left\{\frac{1}{2\sigma_g^2}\sum_{h \in L}\left[\left(\boldsymbol{\Phi}'_h - \boldsymbol{\Phi}_h^{[s]}\right)^2 - \left(\boldsymbol{\Phi}_h^0 - \widetilde{\boldsymbol{\Phi}}_h\right)^2\right]\right\}\frac{\pi(G')}{\pi\left(G^{[s]}\right)}.$$

1.7. Draw $c \sim$ Unif$[0, 1]$. if $c < \gamma$ then set $G^{[s+1]} = G'$.

**Step 2.** Updating the precision matrix $\mathbf{K}$

Draw $\boldsymbol{K}^{[s+1]} \mid G^{[s+1]}, \mathbf{y} \sim$ G-Wishart$(b + n, D + U)$.