6 Supplementary

Proof of Lemma 1:

Proof. In the 2-step procedure of Algorithm 1, $S_0$ is obvious the optimal solution to the sub-problem with parameter $|S_0|$, that is, $S_0 = S(|S_0|)$. Then for the second model selection step, $S_0 = S^*$ due to global optimality of $S_0$. \hfill \Box 

To prove Thm.3, we need the following Finsler’s lemma.

**Lemma 8.** (Finsler) Let $x \in \mathbb{R}^n$, $B \in \mathbb{R}^{m \times n}$ and $Q \in \mathbb{R}^{n \times n}$ such that rank($B$) < $n$, $Q$ symmetric and positive semi-definite. Then the following two statements are equivalent:

\[ x'Qx > 0, \forall x \neq 0, Bx = 0 \iff \exists \gamma > 0: Q + \gamma B'B > 0. \] (19)

**Proof of Thm.3:**

Proof. Assume w.l.o.g. that $S = \{1, 2, ..., k\}$ consists of the first $k$ nodes. Then $A \circ M$ exactly captures the adjacency matrix of the induced sub-graph:

\[ A \circ M = \begin{pmatrix} A_S & 0 \\ 0 & 0 \end{pmatrix} \] (20)

In the fashion, $\text{diag} ((A \circ M)1_n) - (A \circ M)$ captures the Laplacian matrix of $S$:

\[ \text{diag} ((A \circ M)1_n) - (A \circ M) = \begin{pmatrix} L_S & 0 \\ 0 & 0 \end{pmatrix}. \] (21)

By Lemma 2 and Rayleigh-Ritz theorem, we want the following to hold on $L_S$:

\[ x' L_S x > 0, \forall 0 \neq x \in \mathbb{R}^k, x'1_k = 0. \] (22)

By Lemma 8, the above condition can be converted into:

\[ L_S + \gamma 1_k1_k^T \succeq \epsilon I_k, \] (23)

where $\gamma k \geq \epsilon$. Now we place this LMI back to the large matrix and notice the fact that:

\[ \text{diag}(M1_n) = \begin{pmatrix} kI_k & 0 \\ 0 & 0 \end{pmatrix}, \] (24)

the equivalent LMI for the large matrix is:

\[ \text{diag} ((A \circ M)1_n) - (A \circ M) + \gamma M \succeq \frac{\epsilon}{k} \text{diag}(M1_n), \] (25)

where $\gamma k \geq \epsilon$ should be satisfied. Let $\epsilon = \gamma k$, and the proof is done. \hfill \Box 

**Proof of Corollary 4:**

Proof. Let $\gamma = \lambda_2(\Lambda_k)/k$. Then every $S$ satisfying $Q(M; \gamma) \succeq 0$ and $\text{diag}(M)1_n = k$ is connected by Thm.3 and of size $k$. So $S \in \Lambda_k$.

On the other hand, for any $S \in \Lambda_k$, $\lambda_2(S) \geq \lambda_2(\Lambda_k) \geq \gamma k$. From the proof of Thm.3, the indicator matrix $M$ corresponding to $S$ satisfies $Q(M; \gamma) \succeq 0$ and $\text{diag}(M)1_n = k$. Proof is done. \hfill \Box 

**Proof of Lemma 5:**
Proof. If $M \in \{0,1\}^{n \times n}$, by constraints of Eq.(10), $M_{ij} = 1$ if and only if $M_{ii} = 1$ and $M_{jj} = 1$. Thus $M = \text{diag}(M)\text{diag}(M)'$ is rank-1.

On the other hand, if $M$ is rank-1, or $M = ff'$, Consider any two non-zero entries of $f$: $f_i = a \neq 0$, $f_j = b \neq 0$. Then by $M_{ij} \leq \min\{M_{ii}, M_{jj}\}$, we have $a = b$. So every non-zero entry of $f$ is equal. The node with $M_{ii} = 1$ ensures that all non-zero entries of $f$ is 1. Proof is done.

Proof of Theorem 6:

Proof. For part (a), assume on the contrary that the support of $\text{diag}(M)$ is disconnected: $S = C \cup \bar{C}$, where $\bar{C} = S - C$. Let $|S| = k, |C| = k_1, \bar{C} = k_2$. W.l.o.g. assume $M_{11} = 1$, and $C$ consists of nodes $\{1,2,\ldots,k_1\}$.

Consider the $k \times k$ sub-matrix $Q_S$ of $Q$ corresponding to $S$, since the rest part are all 0. Now we use the vector $g = [1_{k_1}; -1_{k_2}]$ to hit $Q_S$:

$$g'Q_Sg = g'\text{diag}((A_S \circ M_S)1_n) - (A_S \circ M_S)) g - \gamma g'\text{diag}(M_S1_n) - M_S) g \geq 0.$$  
(26)

Note that $A_S$ has the form:

$$A_S = \begin{pmatrix} AC & 0 \\ 0 & A_{\bar{C}} \end{pmatrix},$$  
(27)

where the off-diagonal block is zero because by assumption $C$ and $\bar{C}$ is disconnected. Then:

$$\text{diag}((A_S \circ M_S)1_n) - (A_S \circ M_S) = \begin{pmatrix} \tilde{L}_C & 0 \\ 0 & \tilde{L}_C \end{pmatrix},$$  
(28)

where $\tilde{L}_C$ is the Laplacian matrix of $C$ weighted by $M_C$. Notice it still holds that $\tilde{L}_C1_{k_1} = 0$. This means $g'\text{diag}((A_S \circ M_S)1_n) - (A_S \circ M_S)) g = 0$.

On the other hand, let $\text{diag}(M_S1_n) - M_S$ be:

$$\text{diag}(M_S1_n) - M_S = \begin{pmatrix} L_1 & L_3 \\ L_3 & L_2 \end{pmatrix}.$$  
(29)

Using $g_1 = [1_{k_1}; 0]$ and $g_2 = [0; 1_{k_2}]$ to hit $Q_S$ will yield: $1_{k_1}L_11_{k_1} = 0$ and $1_{k_2}L_21_{k_2} = 0$. Apparently $g'\text{diag}(M_S1_n) - M_S) g \geq 0$ due to positive semi-definiteness of Laplacian matrix. If it’s strictly positive, proof is done. Otherwise this means $1_{k_1}L_31_{k_2} = 0$. Note that all entries of $L_3$ are either 0 or negative due to non-negativity of $M_S$. This means $L_i = 0$, or equivalently $M_{ij} = 0$ for any $i \in C, j \in \bar{C}$. But this can not happen, because $M_{11} = 1$ and $M_{1j} \geq 1 + M_{jj} - 1 = M_{jj} > 0$ for any $j \in \bar{C}$. Contradiction!

Part (b) is straightforward by using $g = 1_C - 1_{\bar{C}}$ to hit $Q_S$. Proof is done.

Proof of Proposition 7:

Proof. The proof is similar to the proof of Thm.6, by using $g = 1_{C_1} - 1_{C_2}$ to hit $Q$. □