Supplementary Material

EC.1. Proof of Proposition 1

Proof of Proposition 1 We decompose each matrix $\boldsymbol{Y}^t$ into $\boldsymbol{Y}^t = \boldsymbol{u}_t\boldsymbol{u}_t^\top$ with $\|\boldsymbol{u}_t\|^2 = \mathrm{tr}(\boldsymbol{Y}^t) = 1$. Hence, for any pair (t,t'), $\langle \boldsymbol{Y}^t, \boldsymbol{Y}^{t'}\rangle = \boldsymbol{u}_t^\top \boldsymbol{Y}^{t'}\boldsymbol{u}_t = (\boldsymbol{u}_t^\top \boldsymbol{u}_{t'})^2 \geq 0$.

($\Rightarrow$) If $\boldsymbol{Y} := \sum_{t'\in[r]} \boldsymbol{Y}^{t'} \preceq \mathbb{I}$, then, for any $t \in [r]$, $\boldsymbol{u}_t^\top \boldsymbol{Y}\boldsymbol{u}_t \leq \|\boldsymbol{u}_t\|^2 = 1$. However, $\boldsymbol{u}_t^\top \boldsymbol{Y}\boldsymbol{u}_t = 1 + \sum_{t'\neq t}\langle \boldsymbol{Y}^t, \boldsymbol{Y}^{t'}\rangle$. Hence, for all $t' \neq t$, we must have $\langle \boldsymbol{Y}^t, \boldsymbol{Y}^{t'}\rangle = 0$.

($\Leftarrow$) If $\langle \boldsymbol{Y}^t, \boldsymbol{Y}^{t'}\rangle = 0$ for all $t' \neq t$, then $\{\boldsymbol{u}_t\}_{t\in[r]}$ is an orthonormal family that can be completed to form an orthonormal basis $\{\boldsymbol{u}_t\}_{t\in[p]}$. For any $t \in [r], t' \in [r]$, $\boldsymbol{u}_t^\top \boldsymbol{Y}^{t'}\boldsymbol{u}_t = 1$ if $t = t'$, 0 otherwise so for any $t \in [r]$, $\boldsymbol{u}_t^\top \boldsymbol{Y}\boldsymbol{u}_t \leq \|\boldsymbol{u}_t\|^2$ and $\boldsymbol{Y} = \sum_{t'\in[r]} \boldsymbol{Y}^{t'} \preceq \mathbb{I}$. ☐

EC.2. Proof of Theorem 2

In this section, we derive the valid inequalities introduced in Theorem 2 from first principles. We proceed in two ways. First, we derive inequalities which hold for each $\boldsymbol{Y}^t$ separately. Second, we observe that these inequalities can be generalized to also apply for $\boldsymbol{Y} = \sum_{t\in[r]} \boldsymbol{Y}^t$, and also derive new inequalities which reflect the interaction of the sparsity and rank constraints.

We repeatedly reference two results on the convex hulls of convex quadratic functions under logical constraints that were established by Wei et al. (2022), building upon the work of Günlük and Linderoth (2010), Atamtürk and Gomez (2025):

Lemma EC.1. *(Theorem 3 of Wei et al. 2022) The convex closure of the set*

$$\mathcal{S} = \left\{(\boldsymbol{x},\boldsymbol{z},t) \in \mathbb{R}^n \times \{0,1\}^n \times \mathbb{R} : \quad t \geq \sum_{i=1}^n x_i^2,\ \boldsymbol{e}^\top \boldsymbol{z} \leq k,\ x_i = 0 \text{ if } z_i = 0,\ \forall i \in [n]\right\}$$

is given by

$$\mathcal{S}^c = \left\{(\boldsymbol{x},\boldsymbol{z},\boldsymbol{\theta},t) \in \mathbb{R}^n \times [0,1]^n \times \mathbb{R}^n \times \mathbb{R} : \quad t \geq \sum_{i=1}^n \theta_i,\ \boldsymbol{e}^\top \boldsymbol{z} \leq k,\ \theta_i z_i \geq x_i^2\ \forall i \in [n]\right\}.$$

The above result is sometimes known as a perspective reformulation, since we strengthen the quadratic constraint $t \geq \sum_i x_i^2$ by replacing x_i^2 with its perspective $z_i(x_i/z_i)^2$.

Lemma EC.2. *(Proposition 4 of Wei et al. 2022) Let $k \geq 2$. Then, the convex closure of the set*

$$\mathcal{T} = \left\{(\boldsymbol{x},\boldsymbol{z},t) \in \mathbb{R}^n \times \{0,1\}^n \times \mathbb{R} : \quad t \geq (\boldsymbol{e}^\top \boldsymbol{x})^2,\ \boldsymbol{e}^\top \boldsymbol{z} \leq k,\ x_i = 0 \text{ if } z_i = 0,\ \forall i \in [n]\right\},$$

is given by

$$\mathcal{T}^c = \left\{(\boldsymbol{x},\boldsymbol{z},t) \in \mathbb{R}^n \times [0,1]^n \times \mathbb{R} : \quad t \cdot \min(1, \boldsymbol{e}^\top \boldsymbol{z}) \geq (\boldsymbol{e}^\top \boldsymbol{x})^2,\ \boldsymbol{e}^\top \boldsymbol{z} \leq k\right\}.$$

Lemma EC.2 is extremely useful when a single continuous variable depends upon multiple indicator variables, as occurs in certain substructures of our reformulations of Problem (3).

Rank-One Valid Inequalities First, inspired by Bertsimas and Cory-Wright (2020), we observe that in a feasible solution to Problem (8) the 2×2 minors of $\boldsymbol{Y}^t$ are certainly non-negative, i.e.,

$$(Y^t_{i,j})^2 \leq Y^t_{i,i} Y^t_{j,j}, \ \forall i, j \in [p].$$

These constraints are implied by $\boldsymbol{Y}^t \succeq \boldsymbol{0}$ and hence redundant in and of themselves. However, we can sum over all such constraints $i \in [p]$ and use $\mathrm{tr}(\boldsymbol{Y}^t) = 1$, to obtain the constraint

$$\sum_{i \in [p]} \left(Y^t_{i,j}\right)^2 \leq Y^t_{j,j}, \ \forall j \in [p].$$

This constraint is a sum of redundant constraints and hence redundant. However, we can strengthen it, by noting that it is a separable convex quadratic inequality under logical constraints. Indeed, by Lemma EC.1, its convex closure under the logical constraints $Y^t_{i,j} = 0$ if $Z_{i,t} = 0$ is given by:

$$\sum_{j=1}^{p} \left(Y^t_{i,j}\right)^2 \leq Y^t_{i,i} Z_{i,t}, \ \forall i \in [p], t \in [r]. \tag{EC.1}$$

Rank-r Valid Inequalities In the same spirit as in the rank one case, we can obtain strong valid inequalities by summing the 2×2 minors of $Y_{i,i} = \sum_{t=1}^{r} Y^t_{i,i}$. Indeed, since $\boldsymbol{Y}$ is positive semidefinite, summing its 2×2 minors implies that:

$$\sum_{j=1}^{p} \left(Y_{i,j}\right)^2 \leq r Y_{i,i}.$$

Moreover, since $Y_{i,j} = \sum_{t=1}^{r} \boldsymbol{Y}^t_{i,j}$ is a rank-one quadratic under logical constraints $Y^t_{i,j} = 0$ if $Z_{i,t} = 0$, invoking Lemma EC.2 reveals that the convex closure of this quadratic constraint under these logical constraints is given by the strengthened inequality:

$$\sum_{j=1}^{p} \left(Y_{i,j}\right)^2 \leq r Y_{i,i} \min\left(1, \sum_{t=1}^{r} Z_{i,t}\right), \ \forall i \in [p]. \tag{EC.2}$$

Second, in any feasible solution we have:

$$|Y_{i,j}| \leq \sum_{t=1}^{r} |Y^t_{i,j}| = \sum_{t=1}^{r} |U_{i,t}||U_{j,t}|.$$

Let us denote by k_t the sparsity of the tth column of $\boldsymbol{U}$. Then, it is well known that $\|\boldsymbol{U}_t\|_1 \leq \sqrt{k_t}$. Therefore:

$$\sum_{j=1}^{p} |Y_{i,j}| \leq \sum_{j=1}^{p} \left(\sum_{t=1}^{r} |U_{i,t}||U_{j,t}|\right) \leq \sum_{t=1}^{r} \sqrt{k_t} |U_{i,t}|.$$

Next, squaring both sides and invoking the Cauchy-Schwarz inequality reveals that

$$\left(\sum_{j=1}^{p} |Y_{i,j}|\right)^2 \leq \left(\sum_{t=1}^{r} U_{i,t}^2\right) \left(\sum_{t=1}^{r} k_t\right) = k Y_{i,i}.$$

Finally, noting that the expression $\left(\sum_{j=1}^{p} |Y_{i,j}|\right)^2 \leq k Y_{i,i}$ is a convex quadratic under logical constraints $Y^t_{i,j} = 0$ if $Z_{i,t} = 0$ and invoking Lemma EC.2 to obtain its convex closure yields the strengthened second-order cone inequality

$$\left(\sum_{j=1}^{p}|Y_{i,j}|\right)^2 \leq kY_{i,i}\min\left(1, \sum_{t\in[r]} Z_{i,t}\right),\ \forall i\in[p]. \tag{EC.3}$$

Third, in the same spirit, the 2×2 minors of $\boldsymbol{Y} \preceq \mathrm{Diag}\left(\min\left(\boldsymbol{e}, \sum_{t\in[r]}\boldsymbol{Z}_t\right)\right)$ are

$$\left(\min\left(1,\sum_{t\in[r]}Z_{i,t}\right)-Y_{i,i}\right)\left(\min\left(1,\sum_{t\in[r]}Z_{j,t}\right)-Y_{j,j}\right) \geq \left(\min\left(1,\sum_{t\in[r]}Z_{i,t}\right)\delta_{i,j}-Y_{i,j}\right)^2,$$

where $\delta_{i,j}=\mathbf{1}\{i=j\}$ is an indicator denoting whether $i=j$. Summing these constraints over all indices $i\neq j$ and using $k-r+1$ as an upper bound on $\sum_{i\in[p]:i\neq j}\sum_{t\in[r]}Z_{i,t}-Y_{i,i}$ then yields

$$(k-r+1)\left(\min\left(1,\sum_{t\in[r]}Z_{j,t}\right)-Y_{j,j}\right)\geq \sum_{i\in[p]:i\neq j}Y_{i,j}^2,\ \forall j\in[p].$$

Finally, we recognize the right hand side as a sum of rank-one quadratic terms $(\sum_{t=1}^{r}Y_{i,j}^t)^2$ under logical constraints $Y_{i,j}^t=0$ if $Z_{j,t}=0$ and invoke Lemma EC.2 to obtain the convex closure, giving:

$$(k-r+1)\min\left(1,\sum_{t\in[r]}Z_{j,t}\right)(\min\left(1,\sum_{t\in[r]}Z_{j,t}\right)-Y_{j,j})\geq \sum_{i\in[p]:i\neq j}Y_{i,j}^2\ \forall j\in[p]. \tag{EC.4}$$

The result then follows by introducing a vector $\boldsymbol{w}$ such that w_i models $\min(1,\sum_{t\in[r]}Z_{i,t})$ via $\boldsymbol{w}\in[0,1]^p$, $\boldsymbol{w}\leq \boldsymbol{Z}\boldsymbol{e}$, and noting that we can replace $\boldsymbol{Y}\preceq\mathrm{Diag}(\min(\boldsymbol{e},\boldsymbol{Z}\boldsymbol{e})$ with $\boldsymbol{Y}\preceq\mathrm{Diag}(\boldsymbol{w})$.

EC.3. Proof of Proposition 2

Proof of Proposition 2 First, let us observe that if $\sum_{i\in[p]}Z_{i,t}\leq k$ and $\boldsymbol{Y}^t=\boldsymbol{U}_t\boldsymbol{U}_t^\top$ is a rank-one matrix such that $\|\boldsymbol{U}\|_2=1$ then we have $\|\boldsymbol{U}\|_1\leq\sqrt{k_t}$ by norm equivalence. Therefore

$$\sum_{j=1}^{p}|Y_{i,j}^t|\leq\sum_{j=1}^{p}|U_{i,t}||U_{j,t}|\leq\sqrt{k_t}|U_{i,t}|.$$

Squaring both sides of this inequality then yields

$$\left(\sum_{j=1}^{p}|Y_{i,j}^t|\right)^2\leq k_tY_{i,i}^t,$$

and combining Lemma EC.1 with this inequality yields (10).

Second, in the same spirit, since $\boldsymbol{U}_t\boldsymbol{U}_t^\top$ is only supported on indices where $\boldsymbol{Z}_t$ is non-zero, we have that $\boldsymbol{Y}^t\preceq\mathrm{Diag}(\boldsymbol{Z}_t)$. This constraint implies the following 2×2 minors are non-negative

$$(Z_{i,t}-Y_{i,i}^t)(Z_{j,t}-Y_{j,j}^t)\geq(\delta_{i,j}-Y_{i,j}^t)^2,\ \forall i,j\in[p],$$

where $\delta_{i,j}=1$ if $i=j$ and 0 otherwise. Summing these inequalities over indices $i\neq j$ and setting k_t-1 as a valid upper bound on $\sum_{i\in[p]:i\neq j}Z_{i,t}-Y_{i,i}^t$ whenever $Z_{j,t}=1$ (as $Y_{i,j}^t=0$ if $Z_{j,t}=0$) gives

$$(k_t-1)(Z_{j,t}-Y_{j,j}^t)\geq\sum_{i\in[p]:i\neq j}{Y_{i,j}^t}^2,\ \forall j\in[p].$$

Finally, using Lemma EC.1 to take the convex closure of this inequality under the logical constraints $Y^t_{i,j} = 0$ if $Z_{j,t} = 0$ gives Equation (11). $\square$

EC.4. Complete Formulations for the Semidefinite and Second-Order Cone Relaxations

In this section, we provide the complete formulation for our SDP relaxation (12) as well as its second-order cone approximation.

EC.4.1. Semidefinite Relaxation

In Section 2.3, we proposed a semidefinite relaxation, (12), in the case where a sparsity budget for each PC, k_t, is provided. In particular, this relaxation involves valid inequalities that Kim et al. (2022) have derived in the single-PC case. To the best of our knowledge, their formulation leads to the strongest known relaxation for sparse PCA with $r = 1$ which can be solved in polynomial time. We note however that invoking a fixed but sufficiently large level of the sum-of-squares hierarchy may give tighter relaxations, although we do not write these relaxations down as they involve very large semidefinite constraints and are therefore intractable in practice (see also Dey et al. (2022a) for an NP-hard relaxation that uses the ℓ_1 norm).

In (12), we concisely denoted $(\boldsymbol{Y}^t, \boldsymbol{Z}_t) \in \mathcal{T}(k_t)$ the set of valid inequalities involved in Kim et al. (2022)'s "T-relaxation". We now elicit the constraints involved in the set $\mathcal{T}(k_t)$ and provide the complete formulation of the SDP relaxation (12). For each $t \in [r]$, we introduce an additional variable $\boldsymbol{F}^t$ to capture the entry-wise absolute value of $\boldsymbol{Y}^t$, and an additional matrix $\boldsymbol{G}^t$ which contains a sorted version of $\boldsymbol{F}^t$. We obtain:

$$
\begin{aligned}
\max_{\substack{\boldsymbol{Z}\in[0,1]^{p\times r}:\\ \langle \boldsymbol{E},\boldsymbol{Z}\rangle\leq k,\\ \boldsymbol{w}\in[0,1]^p}} \quad \max_{\substack{\boldsymbol{Y}\in\mathcal{S}^p_+,\boldsymbol{Y}^t,\boldsymbol{F}^t,\boldsymbol{G}^t\in\mathcal{S}^p_+,\\ \boldsymbol{T}^t\in\mathbb{R}^{p\times p}_+,\\ \boldsymbol{r}^{t,D}\in\mathbb{R}^{p-1},\boldsymbol{t}^{t,D}\in\mathbb{R}^{p\times p-1}_+}} \quad & \langle \boldsymbol{Y}, \boldsymbol{\Sigma} \rangle && \text{(EC.5)}\\
\text{s.t.} \quad & \boldsymbol{Y} \preceq \mathrm{Diag}(\boldsymbol{w}),\ \boldsymbol{Y} = \sum_{t=1}^{r} \boldsymbol{Y}^t,\ \boldsymbol{w} \leq \boldsymbol{Z}\boldsymbol{e}, \\
& \sum_{j=1}^{p} Y_{i,j}^2 \leq r Y_{i,i} w_i && \forall i \in [p], \\
& \left(\sum_{j=1}^{p} |Y_{i,j}| \right)^2 \leq k Y_{i,i} w_i && \forall i \in [p], \\
& \sum_{i\in[p]: i\neq j} Y_{i,j}^2 \leq (k - r + 1) w_j (w_j - Y_{j,j}) && \forall j \in [p], \\
& \pm \boldsymbol{Y}^t \leq \boldsymbol{F}^t && \forall t \in [r], \\
& G^t_{i,1} \geq G^t_{i,2} \geq \ldots \geq G^t_{i,k_t} && \forall i \in [k_t],\ \forall t \in [r], \\
& G^t_{i,j} = 0 && \forall i > k_t \text{ or } j > k_t,\ \forall t \in [r],
\end{aligned}
$$

$$
\begin{aligned}
&\mathrm{tr}(\boldsymbol{Y}^t) = \mathrm{tr}(\boldsymbol{G}^t) = \mathrm{tr}(\boldsymbol{F}^t) = 1 && \forall t \in [r],\\
&\langle \boldsymbol{E}, \boldsymbol{G}^t \rangle = \langle \boldsymbol{E}, \boldsymbol{F}^t \rangle && \forall t \in [r],\\
&\sum_{i=1}^{j} G^t_{i,i} \geq j r^{D,t}_j + \sum_{j=1}^{n} t^{t,D}_{i,j} && \forall j \in [p-1],\ t \in [r],\\
&Y^t_{i,i} \leq r^{t,D}_j + t^{t,D}_{i,j} && \forall i \in [p], j \in [p-1], t \in [r],\\
&\left(F^t_{i,j}\right)^2 \leq T^t_{i,j} T^t_{j,i},\ T^t_{i,i} = F^t_{i,i} && \forall i \in [p], j \in [i-1],\ t \in [r],\\
&\sum_{j \in [p]} T^t_{i,j} = Z_{i,t},\ \sum_{i \in [p]} T^t_{i,j} = k_t F^t_{j,j} && \forall i \in [p], \forall j \in [p], t \in [r],\\
&0 \leq T^t_{i,j} \leq F^t_{i,j} && \forall i, j \in [p], t \in [r].
\end{aligned}
$$

The additional variables $\boldsymbol{r}^{t,D}$ and $\boldsymbol{t}^{t,D}$ are introduced to enforce coupling constraints between the diagonal entries of $\boldsymbol{F}^t$ and $\boldsymbol{G}^t$ (Kim et al. 2022, eq. 44), while $\boldsymbol{T}^t$ allows to couple $\boldsymbol{F}^t$ with the binary variables $\boldsymbol{Z}$ (Kim et al. 2022, eq. 50). In contrast to Kim et al. (2022), we explicitly require that each $\boldsymbol{F}^t$ is positive semidefinite (rather than that its 2×2 minors are), in order to obtain a stronger relaxation; we consider the 2×2 minors when developing a more tractable relaxation in the next section.

EC.4.2. A Second-Order Cone Relaxation for Large-Scale Instances

Unfortunately, (12) cannot scale beyond $p = 100$, at least with current technology, due to the presence of multiple semidefinite matrices and constraints.

We now develop a more tractable, albeit less tight, version of the relaxation of (12) which scales to $p > 100$ features. Namely, we replace all semidefinite constraints of the form $\boldsymbol{X} \in \mathcal{S}^p_+$ with the non-negativity of their 2×2 minors, $X_{i,i} X_{j,j} \geq X^2_{i,j}\ \forall i, j \in [p]$, as presented by Bertsimas and Cory-Wright (2020) and references therein. This gives the following second-order cone relaxation of (12):

$$
\max_{\substack{\boldsymbol{Z} \in [0,1]^{p \times r}: \\ \langle \boldsymbol{E}, \boldsymbol{Z} \rangle \leq k, \boldsymbol{w} \in [0,1]^p}} \quad \max_{\substack{\boldsymbol{Y} \in \mathcal{S}^p, \boldsymbol{Y}^t, \boldsymbol{F}_t, \boldsymbol{G}_t \in \mathcal{S}^p, \\ \boldsymbol{T}_t \in \mathbb{R}^{p \times p}_+\ \forall t \in [r], \\ \boldsymbol{r}^{t,D} \in \mathbb{R}^{p-1}, \boldsymbol{t}^{t,D} \in \mathbb{R}^{p \times p-1}_+}} \quad \langle \boldsymbol{Y}, \boldsymbol{\Sigma} \rangle \tag{EC.6}
$$

$$
\begin{aligned}
\text{s.t.}\quad &\boldsymbol{Y} = \sum_{t=1}^{r} \boldsymbol{Y}^t,\ \mathrm{tr}(\boldsymbol{Y}^t) = 1, \boldsymbol{w} \leq \boldsymbol{Z}\boldsymbol{e} && \forall t \in [r],\\
&{Y^t_{i,j}}^2 \leq Y^t_{i,i} Y^t_{j,j} && \forall i, j \in [p], \forall t \in [r],\\
&(\delta_{i,j} - Y_{i,j})^2 \leq (w_i - Y_{i,i})(w_j - Y_{j,j}) && \forall i, j \in [p],\\
&\sum_{j=1}^{p} {Y_{i,j}}^2 \leq r Y_{i,i} w_i && \forall i \in [p]\\
&\left(\sum_{j=1}^{p} |Y_{i,j}|\right)^2 \leq k Y_{i,i} w_i,\ \pm \boldsymbol{Y}^t \leq \boldsymbol{F}_t && \forall i \in [p],\ \forall t \in [r],
\end{aligned}
$$

$$
\begin{aligned}
&\sum_{i\in[p]:i\neq j} Y_{i,j}^2 \le (k-r+1)w_j(w_j - Y_{j,j}) && \forall j\in[p],\\
&Y^t_{i,i} \le t^{t,D}_{i,j} + r^{t,D}_j && \forall i\in[p], j\in[p-1], t\in[r],\\
&\mathrm{tr}(\boldsymbol{Y}^t) = \mathrm{tr}(\boldsymbol{G}^t) = \mathrm{tr}(\boldsymbol{F}_t) = 1 && \forall t\in[r],\\
&\langle \boldsymbol{E}, \boldsymbol{G}_t - \boldsymbol{F}_t\rangle = 0 && \forall t\in[r],\\
&{G^t_{i,j}}^2 \le G^t_{i,i}G^t_{j,j} && \forall i,j\in[p], \forall t\in[r],\\
&G^t_{i,1} \ge G^t_{i,2} \ge \ldots \ge G^t_{i,k_t} && \forall i\in[k_t],\ \forall t\in[r],\\
&G^t_{i,j} = 0 && \forall i > k_t \text{ or } j > k_t\ \forall t\in[r],\\
&\sum_{i=1}^{j} G^t_{i,i} \ge j r^{D,t}_j + \sum_{j=1}^{n} t^{t,D}_{i,j} && \forall j\in[p-1],\ t\in[r],\\
&{F^t_{i,j}}^2 \le T^t_{i,j}T^t_{j,i},\ T^t_{i,i} = F^t_{i,i} && \forall i\in[p], j\in[i-1],\ t\in[r],\\
&\sum_{j\in[p]} T^t_{i,j} = Z_{i,t},\ \sum_{i\in[p]} T^t_{i,j} = k_t F^t_{j,j} && \forall i\in[p], \forall j\in[p], t\in[r],\\
&0 \le T^t_{i,j} \le F^t_{i,j}, {F^t_{i,j}}^2 \le F^t_{i,i}F^t_{j,j} && \forall i,j\in[p], \forall t\in[r].
\end{aligned}
$$

EC.5. Proof of Theorem 3

We prove that Problem (13) is NP-hard via a reduction from a variant of the exact 3-set cover problem.

Restricted Exact Cover by 3-Sets Problem (X3C-R) Consider a set X with $|X| = 3q$ and a family $\mathcal{C} = \{S_1, \ldots, S_m\}$ of 3-element subsets of X of size $m := |\mathcal{C}|$ with $m \ge q$. Further, assume that for any pair of sets $S, S' \in \mathcal{C}$, $|S \cap S'| \le 1$. Then, X3C-R corresponds to asking whether there exists an exact cover of X, namely a subfamily $\mathcal{C}' \subseteq \mathcal{C}$ of size q, whose sets are pairwise disjoint and whose union is X.

Problem X3C-R is NP-complete, as we prove in Proposition EC.2. Leveraging that result, we now prove the following proposition, which implies Theorem 3.

Proposition EC.1. *For any instance of X3C-R, we can construct integers* (p, r)*, a binary mask* $\hat{\boldsymbol{Z}} \in \{0,1\}^{p\times r}$*, a symmetric matrix* $\boldsymbol{\Sigma} \in \mathcal{S}^p$*, and a scalar* τ*, such that X3C-R is equivalent to finding a feasible solution to* (13) *with objective value at least* τ*. Moreover,* $p = |X| + |\mathcal{C}| = 3q + m$*,* $r = |\mathcal{C}|$*,* $\tau = 2q$ *are all polynomial in the size of X3C-R.*

Remark EC.1. Given the orthogonality constraints in (13), we can replace the matrix $\boldsymbol{\Sigma}$ in Problem (13) by $\boldsymbol{\Sigma} + \beta \boldsymbol{I}_p$ for any $\beta > 0$ and obtain the same optimal solution (it only shifts the objective value by βr). In particular, if $\boldsymbol{\Sigma} \in \mathcal{S}^p$ in Problem (13), we can assume $\boldsymbol{\Sigma} \succeq \boldsymbol{0}$ without loss of generality.

Proof of Proposition EC.1 Consider an instance of X3C-R and let us construct an instance of (13).

Let us first define the coordinate space: We create one coordinate for each element $x \in X$, and one "dummy" coordinate d_t for each 3-set $S_t \in \mathcal{C}$. Hence, $p = |X| + |\mathcal{C}| = 3q + m$.

We consider one PC per 3-set. Hence, $r = |\mathcal{C}| = m$. Each column $t \in [r]$ is allowed to be nonzero only on the coordinates corresponding to $x \in S_t$ and on the dummy d_t. In other words,

$$\begin{aligned} \hat{Z}_{i,t} &= 1, \quad \text{if } i \in S_t \cup \{d_t\}, \\ \hat{Z}_{i,t} &= 0, \quad \text{otherwise.} \end{aligned}$$

Note that $r \leq m$ and Problem (13) is always feasible (a feasible solution is obtained by setting $\boldsymbol{u}_t$ equal to 1 on its dummy coordinate d_t and 0 elsewhere).

Finally, the matrix $\boldsymbol{\Sigma}$ is non-zero on the block indexed by $X \times X$ only, i.e.,

$$\boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{A} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix}, \quad \text{with} \quad A_{xy} = \begin{cases} 1, & \text{if there exists } S \in \mathcal{C} \text{ with } x, y \in S, \\ 0, & \text{otherwise.} \end{cases}$$

Consider one column $t \in [r]$. It is associated with one 3-set S_t. On this set, the submatrix of $\boldsymbol{A}$ is equal to

$$\boldsymbol{A}_{S_t,S_t} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} = \boldsymbol{e}\boldsymbol{e}^\top - \boldsymbol{I}_3.$$

So, for any non-zero vector $\boldsymbol{z} \in \mathbb{R}^3$, we have by Cauchy-Schwarz

$$\boldsymbol{z}^\top \boldsymbol{A}_{S_t,S_t} \boldsymbol{z} = (\boldsymbol{e}^\top \boldsymbol{z})^2 - \|\boldsymbol{z}\|^2 \leq (\|\boldsymbol{z}\|_0 - 1)\|\boldsymbol{z}\|^2, \tag{EC.7}$$

with $\|z\|_0 = |\{j : z_j \neq 0\}|$.

Set $\tau = 2q$. Let us now show that there exists a feasible solution to (13) with objective value at least τ iff there is an affirmative answer to X3C-R.

Consider a solution to X3C-R, $\mathcal{C}'$. For each set $S_t \in \mathcal{C}'$, define $\boldsymbol{u}_t$ as equal to $\frac{1}{\sqrt{3}}(1,1,1)$ on the coordinates indexed by S_t and 0 on the dummy coordinate. Alternatively, for each set $S_t \notin \mathcal{C}'$, set $\boldsymbol{u}_t$ equal to $\boldsymbol{0}$ on the coordinates indexed by S_t and 1 on the dummy coordinate. By construction, because each element $x \in X$ is covered by exactly one 3-set of $\mathcal{C}'$, for any $t \neq t'$, the columns $\boldsymbol{u}_t$ and $\boldsymbol{u}_{t'}$ have disjoint supports and $\boldsymbol{u}_t^\top \boldsymbol{u}_{t'} = \delta_{t,t'}$. Second,

$$\sum_{t \in [r]} \boldsymbol{u}_t^\top \boldsymbol{\Sigma} \boldsymbol{u}_t = \frac{1}{3} \sum_{S_t \in \mathcal{C}'} \boldsymbol{e}^\top \boldsymbol{A}_{S_t,S_t} \boldsymbol{e} = 2q = \tau.$$

Conversely, assume there exists a solution to (13) with objective value at least τ. For any $x \in X$, there exists at most one column $\boldsymbol{u}_t$ such that $u_{t,x} \neq 0$. Indeed, by contradiction, assume there exist two distinct columns t, t' such that $u_{t,x} \neq 0$ and $u_{t',x} \neq 0$. Then, $x \in S_t \cap S_{t'}$ so we must have $\{x\} = S_t \cap S_{t'}$ by assumption. In this case, the constraint $\boldsymbol{u}_t^\top \boldsymbol{u}_{t'} = 0$ reads $u_{t,x} u_{t',x} = 0$ and forces one of the coefficients $u_{t,x}$ or $u_{t',x}$ to be equal to 0, which is a contradiction. We say that a column of $\boldsymbol{U}$, $\boldsymbol{u}_t$, is s-heavy ($s \in \{0,1,2,3\}$) if s of its coordinates indexed by X are non-zeros. Let n_s denote the number of s-heavy columns. From our earlier reasoning, each $x \in X$ is covered by at most one column of $\boldsymbol{U}$ so we must have $3n_3 + 2n_2 \leq 3n_3 + 2n_2 + n_1 \leq 3q$. The solution $\boldsymbol{U}$ achieves an objective value at least equal to $\tau = 2q$ so (EC.7) implies $2q = \tau \leq 2n_3 + n_2$. However,

$$2n_3 + n_2 = \frac{2}{3}(3n_3 + \frac{3}{2}n_2) \leq \frac{2}{3}(3n_3 + 2n_2) \leq 2q,$$

so $2q = 2n_3 + n_2$ and the inequalities above are tight, i.e., $n_2 = 0$ and $n_3 = q$. Consequently, $n_1 = 0$ as well. In conclusion, define $\mathcal{C}' = \{S_t : \boldsymbol{u}_t \text{ is 3-heavy}\}$. For any $x \in X$, there exists at most one column $\boldsymbol{u}_t$ such that $u_{t,x} \neq 0$, so the sets in $\mathcal{C}'$ are pairwise disjoint. Furthermore, $|\mathcal{C}'| = n_3 = q = |X|/3$ so $\mathcal{C}'$ must cover the entire set X. □

Remark EC.2. Observe that the proof of Proposition EC.1 shows that there exists a solution to X3C-R iff there exists a feasible solution to Problem (13) with objective value $\tau - \eta$ with $\tau = 2q$ being the optimal value and $\eta \in [0,1)$. Hence, it shows that it is NP-hard to approximate Problem (13) within a factor $1-\epsilon$ with $\epsilon < 1/\tau = 1/(2q)$. Since $p = 3q + m$ and $r = m$, it is therefore NP-hard to approximate Problem (13) within a factor $1-\epsilon$ with $\epsilon < \dfrac{3}{2(p-r)}$.

EC.5.1. X3C-R is NP-Complete

To the best of our knowledge, there is no existing proof of the complexity of X3C-R. Problem X3C-R is a restriction of the classical exact cover by 3-sets problem (X3C), which is known to be NP-complete (Garey and Johnson 1979, chapter A3.1).

Exact Cover by 3-Sets Problem (X3C) Consider a set X with $|X| = 3q$ and a family $\mathcal{C} = \{S_1, \ldots, S_m\}$ of 3-element subsets of X, of size $m := |\mathcal{C}|$. Does there exist a subfamily $\mathcal{C}' \subseteq \mathcal{C}$ of size q whose sets are pairwise disjoint and whose union is X?

Formally, we have the following result:

Proposition EC.2. *Any instance of X3C can be reduced to an instance of X3C-R in polynomial time, and hence Problem X3C-R is NP-hard.*

Any instance of X3C-R is definitionally an instance of X3C, so Proposition EC.2 implies that X3C-R is NP-complete.

Proof of Proposition EC.2 We consider an instance $(X, \mathcal{C})$ of X3C. Define the set

$$Y = X \cup \{z_{S,t} : S \in \mathcal{C},\ t \in \{1, \ldots, 6\}\}.$$

So, $|Y| = |X| + 6|\mathcal{C}| = 3q + 6m$.

For any 3-set $C = \{x_1, x_2, x_3\} \in \mathcal{C}$ (the indexing is arbitrary and fixed), define the following sets

$$\begin{array}{lllllllllll}
D_{S,1} = \{ & x_1 & & & z_{S,1} & & & z_{S,4} & & & \}, \\
D_{S,2} = \{ & & x_2 & & & z_{S,2} & & & z_{S,5} & & \}, \\
D_{S,3} = \{ & & & x_3 & & & z_{S,3} & & & z_{S,6} & \}, \\
D_{S,4} = \{ & & & & z_{S,1} & z_{S,2} & z_{S,3} & & & & \}, \\
D_{S,5} = \{ & & & & & & & z_{S,4} & z_{S,5} & z_{S,6} & \},
\end{array}$$

and define $\mathcal{D} = \{D_{S,t} : S \in \mathcal{C},\ t \in \{1, \ldots, 5\}\}$. Hence, $|\mathcal{D}| = 5|\mathcal{C}| = 5m$. By construction, for any pair of sets $D, D' \in \mathcal{D}$, we must have $|D \cap D'| \leq 1$. In other words, $(Y, \mathcal{D})$ is a valid X3C-R instance.

Let us now show that there exists an exact cover of $(X, \mathcal{C})$ iff there exists an exact cover of $(Y, \mathcal{D})$.

Let $\mathcal{C}'$ be an exact cover of $(X, \mathcal{C})$. Define $\mathcal{D}' = \{D_{S,1}, D_{S,2}, D_{S,3} : S \in \mathcal{C}'\} \cup \{D_{S,4}, D_{S,5} : S \in \mathcal{C} \setminus \mathcal{C}'\}$. We have $|\mathcal{D}'| = 3|\mathcal{C}'| + 2|\mathcal{C} \setminus \mathcal{C}'| = |\mathcal{C}'| + 2|\mathcal{C}| = q + 2m = |Y|/3$. Because $\mathcal{C}'$ is a cover, for any $x \in X$, there exists $S \in \mathcal{C}'$ such that $x \in S$. So there exists $t \in \{1, 2, 3\}$ such that $x \in D_{S,t}$ and $D_{S,t} \in \mathcal{D}'$. For any $z_{S,t}, S \in \mathcal{C}, t \in \{1, \ldots, 6\}$, either $S \in \mathcal{C}'$ and there must exists $t' \in \{1, 2, 3\}$ such that $z_{S,t} \in D_{S,t'}$ and $D_{S,t'} \in \mathcal{D}'$, or $S \notin \mathcal{C}'$ and there must exists $t' \in \{4, 5\}$ such that $z_{S,t} \in D_{S,t'}$ and $D_{S,t'} \in \mathcal{D}'$. Altogether, we showed that any element of Y belongs to an element of $\mathcal{D}'$ (i.e., $\mathcal{D}'$ is a cover). Because elements of $\mathcal{D}'$ are 3-sets and $|\mathcal{D}'| = |Y|/3$, $\mathcal{D}'$ is an exact cover.

Let $\mathcal{D}'$ be an exact cover of $(Y, \mathcal{D})$ (in particular, $|\mathcal{D}'| = q + 2m$). Define $\mathcal{C}' = \{S \in \mathcal{C} : \exists t \in \{1, 2, 3\}, D_{S,t} \in \mathcal{D}'\}$. For any $x \in X$, $x \in Y$ and $\mathcal{D}'$ is a cover, so there exist $S \in \mathcal{C}, t \in \{1, 2, 3\}$ such that $x \in D_{S,t}$ and $D_{S,t} \in \mathcal{C}'$. Hence, $\mathcal{C}'$ covers X. Let us consider a set $S \in \mathcal{C}$ such that $D_{S,1} \in \mathcal{D}'$, we must have $D_{S,2}, D_{S,3} \in \mathcal{D}'$. Indeed, we must have $D_{S,4}, D_{S,5} \notin \mathcal{D}'$, because they both overlap with $D_{S,1}$ and elements of $\mathcal{D}'$ are pairwise disjoint. The variable $z_{S,2}$ is covered by $\mathcal{D}'$ so necessarily, we must have $D_{S,2} \in \mathcal{D}'$. Similarly for $z_{S,3}$. Hence, $3|\mathcal{C}'| + 2|\mathcal{C} \setminus \mathcal{C}'| = |\mathcal{C}'| + 2m$. Therefore, $|\mathcal{C}'| = q = |X|/3$ and $\mathcal{C}'$ is an exact cover of X. ☐

EC.6. Asymptotic Feasibility of the Solutions Returned by Algorithm 2

In this section, we analyze the asymptotic feasibility of solutions generated by Algorithm 2 and show how they depend on the method used to solve each subproblem. First, we prove that Algorithm 2 generates a subsequence of solutions that are asymptotically feasible, provided that each subproblem in Algorithm 2 is solved to near-optimality (Section EC.6.1). We acknowledge that this assumption may not hold for the truncated power method of Yuan and Zhang (2013) that we use in our implementation (Section EC.6.2). Secondly, under less restrictive assumptions that are satisfied by the truncated power method, we show convergence to stationary points of the function $\boldsymbol{U} \mapsto \sum_{t'>t} (\boldsymbol{U}_t^\top \boldsymbol{U}_{t'})^2$ over the feasible region. Unfortunately, this result cannot rule out convergence to infeasible solutions, which is a known limitation of the quadratic penalty method, even when subproblems are solved with increasing accuracy (see Nocedal and Wright 2006, Theorem 17.2).

We introduce some compact notation from Nocedal and Wright (2006, chapter 17). Letting $\mathbb{S}_k^p$ denote the set of k-sparse vectors from the p-dimensional unit sphere, we rewrite Problem (3) as a generic-equality constrained maximization problem of the form

$$\max_{\boldsymbol{U}\in\mathcal{U}} f(\boldsymbol{U}) \quad \text{s.t.} \quad \boldsymbol{c}(\boldsymbol{U})=\boldsymbol{0},$$

where the feasible set $\mathcal{U}:=\{\boldsymbol{U}\in\mathbb{R}^{p\times r} : \forall t\in[r],\ \boldsymbol{U}_t\in\mathbb{S}_{k_t}^p\}$ is non-convex, the objective function f is defined as $f(\boldsymbol{U})=\sum_{t\in[r]}\boldsymbol{U}_t^\top\boldsymbol{\Sigma}\boldsymbol{U}_t$, and the constraint function $\boldsymbol{c}:\mathbb{R}^{p\times r}\to\mathbb{R}^{r(r-1)/2}$ is defined as $c_{t,t'}(\boldsymbol{U})=\boldsymbol{U}_t^\top\boldsymbol{U}_{t'}$ for $t>t'$.

EC.6.1. Asymptotic Convergence of Algorithm 2 With Branch-and-Bound for Subproblems

Suppose that, at each iteration ℓ of Algorithm 2, the vector $\boldsymbol{U}_t^{(\ell)}$ is computed by solving

$$\max_{\boldsymbol{u}\in\mathbb{S}_{k_t}^p} \quad \left\langle \boldsymbol{u}\boldsymbol{u}^\top, \boldsymbol{\Sigma}-\sum_{t'\in[r]:t'\neq t}\lambda_{t,t'}^{(\ell)}\boldsymbol{U}_{t'}^{(\ell-1)}\boldsymbol{U}_{t'}^{(\ell-1)\top}\right\rangle. \tag{EC.8}$$

Under this assumption, we now establish the asymptotic feasibility of a subsequence of the solutions generated by Algorithm 2, provided the subproblems (EC.8) are solved to finite (but not necessarily improving) accuracy, and $k_t\geq r$ for each $t\in[r]$.

Remark EC.3. The subproblem (EC.8) for updating column t at iteration ℓ corresponds to a block-Jacobi update rule where the penalty on orthogonality violation is expressed using the value of $\boldsymbol{U}$ at the previous iteration $\ell-1$. In particular, it requires storing both $\boldsymbol{U}^{(\ell)}$ and $\boldsymbol{U}^{(\ell-1)}$ and does not use the updated values of $\boldsymbol{U}_{t'}^{(\ell)}$ for $t'<t$. Instead, we can implement our algorithm (and adapt our analysis) using Gauss-Seidel updates instead, which would involve subproblems of the form

$$\max_{\boldsymbol{u}\in\mathbb{S}_{k_t}^p} \quad \left\langle \boldsymbol{u}\boldsymbol{u}^\top, \boldsymbol{\Sigma}-\sum_{t'<t}\lambda_{t,t'}^{(\ell)}\boldsymbol{U}_{t'}^{(\ell)}\boldsymbol{U}_{t'}^{(\ell)\top}-\sum_{t'>t}\lambda_{t,t'}^{(\ell)}\boldsymbol{U}_{t'}^{(\ell-1)}\boldsymbol{U}_{t'}^{(\ell-1)\top}\right\rangle.$$

Proposition EC.3. *Assume that $k_t\geq r$, for all $t\in[r]$. Consider a sequence of penalty parameters $\{\lambda_{t,t'}^{(\ell)}\}$ with $\mu_\ell:=\min_{t,t'\in[r]}\lambda_{t,t'}^{(\ell)}\to+\infty$ and a sequence of iterates $\{\boldsymbol{U}^{(\ell)}\}$ such that, for any $\ell\geq 1$, any $t\in[r]$, and any sparse vector $\boldsymbol{u}\in\mathbb{S}_{k_t}^p$,*

$$\boldsymbol{U}_t^{(\ell)\top}\boldsymbol{\Sigma}\boldsymbol{U}_t^{(\ell)}-\sum_{t'\neq t}\lambda_{t,t'}^{(\ell)}(\boldsymbol{U}_t^{(\ell)\top}\boldsymbol{U}_{t'}^{(\ell-1)})^2\geq\boldsymbol{u}^\top\boldsymbol{\Sigma}\boldsymbol{u}-\sum_{t'\neq t}\lambda_{t,t'}^{(\ell)}(\boldsymbol{u}^\top\boldsymbol{U}_{t'}^{(\ell-1)})^2-\epsilon_\ell, \tag{EC.9}$$

with $\epsilon_\ell/\mu_\ell\to 0$. Then, for any convergent subsequence of $\{\boldsymbol{U}^\ell\}$, its limit point $\bar{\boldsymbol{U}}$ is feasible.

Remark EC.4. Because $\boldsymbol{U}^{(\ell)}\in\mathcal{U}$, which is compact, a converging subsequence exists by the Bolzano-Weierstrass theorem.

When using a provably optimal method, like the branch-and-bound algorithm of Berk and Bertsimas (2019), for solving the subproblems (EC.8), we can set the suboptimality gap ϵ_ℓ explicitly as the termination criterion and ensure the condition of Proposition EC.3 holds (e.g., by keeping

ϵ_ℓ constant). However, empirically, solutions from the truncated power method (TPM) are very close to the optimal solution, but we cannot guarantee that TPM solves (EC.8) to ϵ_ℓ-optimality for some controllable $\epsilon_\ell > 0$.

Proof of Proposition EC.3 Consider the ℓth iteration and fix $t \in [r]$. Since $k_t \geq r$, we can construct $\boldsymbol{u} \in \mathbb{S}^p_{k_t}$ such that $\boldsymbol{u}^\top \boldsymbol{U}_{t'}^{(\ell-1)} = 0, \forall t' \neq t$ (e.g., select r coordinates, the restriction of the $r-1$ vectors $\boldsymbol{U}_{t'}^{(\ell-1)}, t' \neq t$ to these r coordinates cannot span the entire space $\mathbb{R}^r$ so we can find a unit vector orthogonal to all of them, which we can view as an r-sparse unit vector in $\mathbb{R}^p$ by padding with zeros). For this vector, we have, by assumption

$$\boldsymbol{U}_t^{(\ell)\top} \boldsymbol{\Sigma} \boldsymbol{U}_t^{(\ell)} - \sum_{t' \neq t} \lambda_{t,t'}^{(\ell)} (\boldsymbol{U}_t^{(\ell)\top} \boldsymbol{U}_{t'}^{(\ell-1)})^2 \geq \boldsymbol{u}^\top \boldsymbol{\Sigma} \boldsymbol{u} - \epsilon_\ell,$$

leading to

$$\sum_{t' \neq t} (\boldsymbol{U}_t^{(\ell)\top} \boldsymbol{U}_{t'}^{(\ell-1)})^2 \leq \frac{1}{\mu_\ell} \left(\boldsymbol{U}_t^{(\ell)\top} \boldsymbol{\Sigma} \boldsymbol{U}_t^{(\ell)} - \boldsymbol{u}^\top \boldsymbol{\Sigma} \boldsymbol{u} \right) + \frac{\epsilon_\ell}{\mu_\ell} \leq \frac{\lambda_{\max}(\boldsymbol{\Sigma}) + \epsilon_\ell}{\mu_\ell}.$$

Taking limits when $\ell \to \infty$, we get $\sum_{t' \neq t} (\bar{\boldsymbol{U}}_t^\top \bar{\boldsymbol{U}}_{t'})^2 = 0$, hence $\bar{\boldsymbol{U}}_t^\top \bar{\boldsymbol{U}}_{t'} = 0, \ \forall t \neq t'$. □

EC.6.2. Asymptotic Stationarity of Algorithm 2 With Truncated Power Method for Subproblems

In this section, we analyze the convergence of Algorithm 2 when the subproblems (EC.8) are solved using the truncated power method of Yuan and Zhang (2013). In particular, we leverage the following property, which follows immediately from the description of the method in Yuan and Zhang (2013):

Lemma EC.3. *Consider a symmetric matrix $\boldsymbol{A} \in \mathbb{R}^{p \times p}$, $k \leq p$, and the sparse quadratic problem*

$$\max_{\boldsymbol{x} \in \mathbb{R}^p} \quad \boldsymbol{x}^\top \boldsymbol{A} \boldsymbol{x} \ \ s.t. \ \boldsymbol{x} \in \mathbb{S}^p_k, \tag{EC.10}$$

The truncated power method of Yuan and Zhang (2013) applied to (EC.10) *converges to a sparse eigenvector of $\boldsymbol{A}$, i.e., a vector $\bar{\boldsymbol{x}} \in \mathbb{R}^p : \|\bar{\boldsymbol{x}}\|_2 = 1$, such that the support of $\bar{\boldsymbol{x}}$, $S := \{j \, : \, \bar{x}_j \neq 0\}$ is of size at most k and such that there exists $\lambda \in \mathbb{R}$ such that $\boldsymbol{A}_{S,S} \bar{\boldsymbol{x}}_S = \lambda \bar{\boldsymbol{x}}_S$.*

For any $\boldsymbol{x} \in \mathbb{S}^p_k$, we denote $\mathcal{T}_{\boldsymbol{x}}(\mathbb{S}^p_k)$ the tangent cone to $\mathbb{S}^p_k$ at $\boldsymbol{x}$, and $\mathrm{Proj}_{\boldsymbol{x}}$ the projection onto $\mathcal{T}_{\boldsymbol{x}}(\mathbb{S}^p_k)$. In particular, we have

$$\mathcal{T}_{\boldsymbol{x}}(\mathbb{S}^p_k) = \{ \boldsymbol{y} \in \mathbb{R}^p \, : \, \mathrm{supp}(\boldsymbol{y}) \subseteq \mathrm{supp}(\boldsymbol{x}), \ \boldsymbol{y}^\top \boldsymbol{x} = 0 \}.$$

With these notations, the equation $\boldsymbol{A}_{S,S} \bar{\boldsymbol{x}}_S = \lambda \bar{\boldsymbol{x}}_S$ in Lemma EC.3 implies

$$\boldsymbol{y}^\top \boldsymbol{A} \bar{\boldsymbol{x}} = 0, \quad \forall \boldsymbol{y} \in \mathcal{T}_{\boldsymbol{x}}(\mathbb{S}^p_k), \qquad \text{and} \qquad \|\mathrm{Proj}_{\boldsymbol{x}}(\boldsymbol{A} \bar{\boldsymbol{x}})\| = \boldsymbol{0}.$$

At iteration ℓ of Algorithm 2, the vector $\boldsymbol{U}_t^\ell$ is computed by solving a problem of the form (EC.10) with $\boldsymbol{A} = \boldsymbol{\Sigma} - \sum_{t' \neq t} \lambda_{t,t'}^{(\ell)} \boldsymbol{U}_{t'}^{(\ell-1)} \boldsymbol{U}_{t'}^{(\ell-1)\top}$ using the truncated power method. Hence, Lemma EC.3 yields

$$\left\| \mathrm{Proj}_{\boldsymbol{U}_t^{(\ell)}} \left(\boldsymbol{\Sigma} \boldsymbol{U}_t^{(\ell)} - \sum_{t' \neq t} \lambda_{t,t'}^{(\ell)} (\boldsymbol{U}_t^{(\ell)\top} \boldsymbol{U}_{t'}^{(\ell-1)})\, \boldsymbol{U}_{t'}^{(\ell-1)} \right) \right\| = 0.$$

Thanks to this condition, we can show that the sequence of iterates generated by Algorithm 2 converges to a stationary point of the quadratic penalty function. Unfortunately, stationarity does not imply feasibility and Algorithm 2 can converge to an infeasible, yet stationary, point. This behavior is a known limitation of the quadratic penalty method, which occurs even when subproblems are solved to increasing accuracy (see Nocedal and Wright 2006, Theorem 17.2).

Formally, we have the following result:

Proposition EC.4. *Consider a sequence of homogeneous penalty parameters* $\{\lambda_{t,t'}^{(\ell)}\} = \{\mu_\ell\}$ *with* $\mu_\ell \to +\infty$ *and a sequence of iterates* $\{\boldsymbol{U}^{(\ell)}\}$ *such that*

$$\max_{t \in [r]} \left\| \mathrm{Proj}_{\boldsymbol{U}_t^{(\ell)}} \left(\boldsymbol{\Sigma} \boldsymbol{U}_t^{(\ell)} - \mu_\ell \sum_{t' \neq t} (\boldsymbol{U}_t^{(\ell)\top} \boldsymbol{U}_{t'}^{(\ell-1)})\, \boldsymbol{U}_{t'}^{(\ell-1)} \right) \right\| \leq \tau_\ell,$$

with $\tau_\ell / \mu_\ell \to 0$. *Then, for any converging subsequence of* $\{\boldsymbol{U}^\ell\}$ *such that the sequence* $\{\mathrm{supp}(\boldsymbol{U}_t^{(\ell)})\}, t \in [r]$ *converges, its limit point* $\bar{\boldsymbol{U}}$ *is a stationary point of the function* $\boldsymbol{U} \mapsto \|\boldsymbol{c}(\boldsymbol{U})\|^2$ *on* $\mathcal{U}$*, namely*

$$\mathrm{Proj}_{\bar{\boldsymbol{u}}_t} \left(\sum_{t' \neq t} (\bar{\boldsymbol{u}}_t^\top \bar{\boldsymbol{u}}_{t'})\, \bar{\boldsymbol{u}}_{t'} \right) = \boldsymbol{0}.$$

Proposition EC.4 asserts that limit points are stationary points of $\boldsymbol{U} \mapsto \|\boldsymbol{c}(\bar{\boldsymbol{U}})\|^2$ over the feasible set, meaning that for each column t of a limit point $\bar{\boldsymbol{U}}$, the directional derivative of $\|\boldsymbol{c}(\bar{\boldsymbol{U}})\|^2$ vanishes along every direction in $\mathcal{T}_{\bar{U}_t}(\mathbb{S}_{k_t}^p)$. To see that stationarity does not imply feasibility, observe, for example, that with $r = 2$, Proposition EC.4 does not rule out convergence to a matrix $\boldsymbol{U} = [\boldsymbol{u}_1, \boldsymbol{u}_1] \in \mathbb{R}^{p \times 2}$ with $\boldsymbol{u}_1 \in \mathbb{S}_k^p$. Provided the truncated power method is run until convergence at each iteration, Algorithm 2 satisfies the assumptions of Proposition EC.4 with $\tau_\ell = 0$. Nonetheless, Proposition EC.4 shows convergence to a stationary point even when this condition is approximately satisfied.

Proof of Proposition EC.4 We follow the proof steps of Nocedal and Wright (2006, Theorem 17.2). Denoting $\boldsymbol{w}_t^{(\ell)} := \boldsymbol{\Sigma} \boldsymbol{U}_t^{(\ell)} - \mu_\ell \sum_{t' \neq t} (\boldsymbol{U}_t^{(\ell)\top} \boldsymbol{U}_{t'}^{(\ell-1)})\, \boldsymbol{U}_{t'}^{(\ell-1)}$, we have

$$\mathrm{Proj}_{\boldsymbol{U}_t^{(\ell)}} \left(\sum_{t' \neq t} (\boldsymbol{U}_t^{(\ell)\top} \boldsymbol{U}_{t'}^{(\ell-1)}) \boldsymbol{U}_{t'}^{(\ell-1)} \right) = \frac{1}{\mu_\ell} \mathrm{Proj}_{\boldsymbol{U}_t^{(\ell)}} \left(\boldsymbol{\Sigma} \boldsymbol{U}_t^{(\ell)} - \boldsymbol{w}_t^{(\ell)} \right)$$

and our assumption leads to

$$\left\| \mathrm{Proj}_{\boldsymbol{U}_t^{(\ell)}}\left(\sum_{t'\neq t}(\boldsymbol{U}_t^{(\ell)\top}\boldsymbol{U}_{t'}^{(\ell-1)})\boldsymbol{U}_{t'}^{(\ell-1)}\right)\right\| \leq \frac{1}{\mu_\ell}\left(\| \mathrm{Proj}_{\boldsymbol{U}_t^{(\ell)}}\left(\boldsymbol{\Sigma}\boldsymbol{U}_t^{(\ell)}\right)\| + \| \mathrm{Proj}_{\boldsymbol{U}_t^{(\ell)}}\left(\boldsymbol{w}_t^{(\ell)}\right)\|\right)$$
$$\leq \frac{1}{\mu_\ell}\left(\lambda_{\max}(\boldsymbol{\Sigma})+\tau_\ell\right)$$

Let us consider a converging subsequence of $\{\boldsymbol{U}^\ell\}$ such that the sequences $\{\mathrm{supp}(\boldsymbol{U}_t^{(\ell)})\}$ for $t\in[r]$ converge. Since the sequences $\{\mathrm{supp}(\boldsymbol{U}_t^{(\ell)})\}$ take discrete values and converge, there exists ℓ_0 such that, for any $\ell\geq\ell_0$, $t\in[r]$, $\mathrm{supp}(\boldsymbol{U}_t^{(\ell)})=\mathrm{supp}(\bar{\boldsymbol{U}}_t)$. Hence, for any $\ell\geq\ell_0$, $\mathrm{Proj}_{\boldsymbol{U}_t^{(\ell)}}(\cdot)\to\mathrm{Proj}_{\bar{\boldsymbol{U}}_t}(\cdot)$ (for $\ell\geq\ell_0$, the projection operators only differ in the linear constraint defining the tangent cone) and taking the limit in the inequality above, for each $t\in[r]$,

$$\mathrm{Proj}_{\bar{\boldsymbol{U}}_t}\left(\sum_{t'\neq t}(\bar{\boldsymbol{U}}_t^\top\bar{\boldsymbol{U}}_{t'})\,\bar{\boldsymbol{U}}_{t'}\right)=\boldsymbol{0}. \quad \square$$

REMARK EC.5. At the expense of more complicated notation, Proposition EC.4 can be extended to non-homogeneous penalty sequences $\{\lambda_{t,t'}^{(\ell)}\}$ such that $\mu_\ell:=\min_{t,t'}\lambda_{t,t'}^{(\ell)}\to+\infty$ and such that, for each pair (t,t'), there exists $\lambda_{t,t'}>0$ s.t. $\lambda_{t,t'}^{(\ell)}/\mu_\ell\to\lambda_{t,t'}$ (i.e., all $\lambda_{t,t'}^{(\ell)}$s grow at the same rate), in which case we have convergence to a stationary point of the function $\boldsymbol{U}\mapsto\sum_{t>t'}\lambda_{t,t'}(\boldsymbol{U}_t^\top\boldsymbol{U}_{t'})^2$.

EC.7. Alternative Proof of Theorem 4

For a fixed $\boldsymbol{Z}$, the objective value in

$$\max_{\boldsymbol{U}\in\mathbb{R}^{p\times r}} \quad \langle\boldsymbol{U}\boldsymbol{U}^\top,\boldsymbol{\Sigma}\rangle \quad \text{s.t.} \quad \boldsymbol{U}^\top\boldsymbol{U}=\mathbb{I},\, U_{i,t}=0 \text{ if } Z_{i,t}=0,\ \forall i\in[p],\ \forall t\in[r],$$

is equal to

$$\langle\boldsymbol{U}\boldsymbol{U}^\top,\boldsymbol{\Sigma}\rangle=\sum_{t\in[r]}\langle\boldsymbol{U}_t\boldsymbol{U}_t^\top,\boldsymbol{\Sigma}\rangle=\sum_{t\in[r]}\langle\boldsymbol{U}_t\boldsymbol{U}_t^\top,\mathrm{Diag}(\boldsymbol{Z}_t)\boldsymbol{\Sigma}\mathrm{Diag}(\boldsymbol{Z}_t)\rangle=\mathrm{vec}(\boldsymbol{U})^\top\boldsymbol{S}\,\mathrm{vec}(\boldsymbol{U}),$$

with

$$\boldsymbol{S}:=\begin{pmatrix}\mathrm{Diag}(\boldsymbol{Z}_1)\boldsymbol{\Sigma}\mathrm{Diag}(\boldsymbol{Z}_1) & & \\ & \ddots & \\ & & \mathrm{Diag}(\boldsymbol{Z}_r)\boldsymbol{\Sigma}\mathrm{Diag}(\boldsymbol{Z}_r)\end{pmatrix},$$

and $\mathrm{vec}(\boldsymbol{U})\in\mathbb{R}^{pr}$ the vector obtained by concatenating the columns of $\boldsymbol{U}$, $\boldsymbol{u}_t$, vertically. As in the proof of (18), we use the fact that $\boldsymbol{S}\preceq\mathrm{Diag}(\boldsymbol{s})$ with $\boldsymbol{s}\in\mathbb{R}^{pr}$ the vector of absolute row-sums of $\boldsymbol{S}$, i.e., $s_{(t-1)\times p+i}=\sum_{j\in[p]}Z_{i,t}Z_{j,t}|\Sigma_{i,j}|$ for $i\in[p]$, $t\in[r]$. This leads to

$$\langle\boldsymbol{U}\boldsymbol{U}^\top,\boldsymbol{\Sigma}\rangle\leq\mathrm{vec}(\boldsymbol{U})^\top\,\mathrm{Diag}(\boldsymbol{s})\,\mathrm{vec}(\boldsymbol{U})=\sum_{t\in[r]}\sum_{i\in[p]}U_{i,t}^2\left(\sum_{j\in[p]}Z_{i,t}Z_{j,t}|\Sigma_{i,j}|\right).$$

Hence, we can bound the objective value by

$$\max_{\boldsymbol{U}\in\mathbb{R}^{p\times r}} \quad \sum_{t\in[r]}\sum_{i\in[p]} U_{i,t}^2 \left(\sum_{j\in[p]} Z_{i,t}Z_{j,t}|\Sigma_{i,j}|\right) \quad \text{s.t.} \quad \boldsymbol{U}^\top\boldsymbol{U}=\mathbb{I}.$$

On one side, the orthogonality constraint implies $\|\boldsymbol{U}_t\|_2^2=\sum_{i\in[p]}U_{i,t}^2=1$, for any $t\in[r]$. On the other side, $\boldsymbol{U}^\top\boldsymbol{U}=\mathbb{I} \implies \boldsymbol{U}\boldsymbol{U}^\top \preceq \mathbb{I} \implies \sum_{t\in[r]}U_{i,t}^2\le 1, \forall i\in[p]$. Hence, optimizing for $\mu_{i,t}:=U_{i,t}^2$, we get a bound of the form

$$\max_{\boldsymbol{\mu}\ge 0} \quad \sum_{t\in[r]}\sum_{i\in[p]} \mu_{i,t} \left(\sum_{j\in[p]} Z_{i,t}Z_{j,t}|\Sigma_{i,j}|\right) \quad \text{s.t.} \sum_{i\in[p]}\mu_{i,t}=1, \forall t\in[r]$$
$$\sum_{t\in[r]}\mu_{i,t}\le 1, \forall i\in[p].$$

Finally, we recognize that some optimal solution $\boldsymbol{\mu}$ to the above problem is an extreme point (i.e., binary) by the linearity of the objective, giving the overall result.

EC.8. Algorithmic Benchmark: Non-Convex QCQP Solvers

The sparse PCA problem with multiple PCs, either in its original formulation (3) or its equivalent reformulation (7), can be seen as a non-convex mixed-integer quadratically constrained problem—for (7), the rank constraints can be encoded as non-convex quadratic constraints $(\boldsymbol{Y}^t)^2=\boldsymbol{Y}^t$ or $\boldsymbol{Y}^t=\boldsymbol{U}_t\boldsymbol{U}_t^\top$. However, current non-convex MIQCP solvers cannot handle SDP variables and constraints. Accordingly, we now discuss how to solve our sparse PCA problem exactly using commercial global optimization solvers via formulation (3), and evaluate this option numerically.

EC.8.1. Solving (3) With an Off-the-Shelf MIQCQP Solver

Since feasible solutions are extremely challenging for spatial branch-and-bound solvers to recover when quadratic equality constraints are imposed exactly, especially when these solutions are irrational or of exponential size (cf. Ramana 1997, Bienstock et al. 2023), we relax the constraint $\boldsymbol{U}^\top\boldsymbol{U}=\mathbb{I}$ to require that it is satisfied to within an elementwise tolerance of ϵ. This gives:

$$\begin{aligned}\max_{\substack{\boldsymbol{Z}\in\{0,1\}^{p\times r}:\\ \langle\boldsymbol{E},\boldsymbol{Z}\rangle\le k}} \max_{\boldsymbol{U}\in\mathbb{R}^{p\times r}} \quad & \langle\boldsymbol{U}\boldsymbol{U}^\top,\boldsymbol{\Sigma}\rangle \\ \text{s.t.} \quad & \|\boldsymbol{U}^\top\boldsymbol{U}-\mathbb{I}\|_\infty\le\epsilon, \\ & U_{i,t}=0 \text{ if } Z_{i,t}=0, \qquad \forall i\in[p],\ t\in[r].\end{aligned} \tag{EC.11}$$

We set $\epsilon=10^{-4}/r^2$ so that the total constraint violation does not exceed 10^{-4}. Problem (EC.11) is a non-convex quadratically constrained mixed-integer problem with pr continuous variables, pr binaries, and r^2 quadratic constraints.

In addition, we strengthen Problem (EC.11) with valid inequalities derived from the ℓ_1 relaxation of sparse PCA, as explored by Dey et al. (2022b,a). Indeed, if the sparsity of each PC, k_t, is specified a priori, we have the valid inequalities

$$\|\boldsymbol{U}_t\|_1 \leq \sqrt{k_t},\ \forall t \in [r]. \tag{EC.12}$$

Moreover, if k is specified but k_t is not, we instead impose the second-order cone inequalities

$$\|\boldsymbol{U}_t\|_1^2 \leq \sum_{i \in [p]} Z_{i,t},\ \forall t \in [r], \tag{EC.13}$$

which allows us to model $k_t = \sum_{j=1}^{p} Z_{j,t}$ in a tractable fashion.

In practice, this approach allows MIQCP solvers to solve Problem (7) to optimality for $pr < 100$ (Section 5.2) and obtain high-quality solutions at larger problem sizes. Note that we avoid mixing both sets of inequalities, as we observed in some preliminary numerical experiments that this sometimes induces numerical instability.

To further improve branch-and-bound, we consider two acceleration strategies:

- Using the solution generated by Algorithm 1 as a warm-start;
- Inform branching decisions by the combinatorial upper bound derived in Section 4.

As spatial branch-and-bound technology improves over time, we believe that it should be possible to solve Problem (EC.11) exactly at larger problem sizes. Indeed, recent works, e.g. Gupta et al. (2023), solve some quadratically constrained problems with up to 50 variables to optimality using custom branch-and-bound solvers, and Gupta et al. (2023) reports that `Gurobi`'s off-the-shelf QCQP solver has achieved a machine-independent speedup factor of 67.5 in less than two years, which suggests that larger instances of (13)–(EC.11) may soon be in reach.

EC.8.2. Numerical Performance on Pitprops

In this section, we numerically evaluate the quality of our approaches on the `pitprops` dataset (Table EC.1) and an approach based on the currently available non-convex MIQCQP technology.

Tables EC.2-EC.3 compare commercial spatial branch-and-bound with/without the combinatorial upper bound (19), and with/without a warmstart from Algorithm 1.

In comparison with the performance of Algorithms 2–3, we observe that (i) Gurobi's upper bounds are uniformly worse than the SDP relaxation for $r \geq 4$; (ii) on instances where $\sum_t k_t \leq p$ it explains a comparable amount of variance to Algorithms 2–3, although it explains significantly less variance than Algorithms 2 on instances where $\sum_t k_t > p$; (iii) It requires 3–4 orders of magnitude more time than any of our algorithms.

r	k_t	Alg. 1				Alg. 2			Alg. 3		
		UB	Obj.	Viol.	T(s)	Obj.	Viol.	T(s)	Obj.	Viol.	T(s)
2	2	0.295	**0.295**	0	20.64	**0.295**	0	7.54	**0.295**	0	0.02
2	4	0.408	0.378	0	20.80	0.400	0	4.94	**0.404**	0	0.52
2	6	0.477	0.437	0	20.62	0.452	0	6.77	0.443	0	0.62
2	8	0.501	0.375	0	20.93	**0.476**	0	7.83	0.446	0	0.64
2	10	0.507	0.463	0	21.01	**0.500**	0	6.36	0.464	0	0.72
3	2	0.435	**0.435**	0	20.81	0.424	0	9.98	**0.435**	0	0.03
3	4	0.572	0.525	0	21.03	0.551	0	7.24	**0.555**	0	1.01
3	6	0.641	0.463	0	23.03	**0.608**	0	10.21	0.569	0	0.92
3	8	0.652	0.580	0	20.94	**0.638**	0	8.88	0.569	0	1.06
3	10	0.652	0.392	0	20.81	**0.650**	0	11.43	0.569	0	1.39
4	2	0.554	**0.554**	0	21.22	**0.554**	0	11.05	**0.554**	0	0.97
4	4	0.704	0.470	0	21.07	0.657	0.003	13.81	**0.657**	0	3.25
4	6	0.737	0.537	0	22.22	**0.697**	0.002	12.14	0.644	0	2.07
4	8	0.737	0.553	0	22.71	**0.720**	0	11.84	0.644	0	3.48
4	10	0.737	0.508	0	21.03	**0.736**	0	11.22	0.644	0	3.16
5	2	0.657	0.455	0	20.87	0.647	0	12.23	**0.648**	0	1.57
5	4	0.795	0.586	0	21.27	**0.743**	0	15.60	0.709	0	8.12
5	6	0.807	0.538	0	23.47	**0.779**	0.016	13.95	0.713	0	16.54
5	8	0.807	0.563	0	21.02	**0.800**	0.004	17.94	0.713	0	24.78
5	10	0.807	0.525	0	21.97	**0.807**	0.001	15.01	0.713	0	7.39
6	2	0.749	0.581	0	21.45	0.746	0	12.61	**0.749**	0	4.87
6	4	0.866	0.576	0	21.23	0.807	0.035	23.08	**0.780**	0	54.2
6	6	0.870	0.617	0	23.45	**0.839**	0.044	15.63	0.780	0	12.54
6	8	0.870	0.628	0	21.69	**0.849**	0.011	17.20	0.780	0	8.45
6	10	0.870	0.664	0	21.12	**0.866**	0.006	25.91	0.780	0	86.88
Avg		0.668	0.508	0	21.46	**0.649**	0.005	12.42	0.610	0	9.85

Table EC.1 Performance of Algorithms 1–3 on the pitprops dataset ($p = 13$) using the experimental setup laid out in Section 5.2. We denote the best-performing solution (in terms of the proportion of variance explained minus the total orthogonality constraint violation) in bold. We report the semidefinite upper bound obtained from solving Problem (12) as part of our analysis of Algorithm 1 since the semidefinite upper bound is the tightest bound proposed in the paper, but do not report the upper bound from the other methods to avoid redundancy. Note that k_t denotes the sparsity of each individual component, meaning a set of r PCs have a collective sparsity budget of $k_t r$, and that all objective values are reported in terms of the proportion of variance explained by dividing by p, the number of features. Note that the relative optimality gap from Gurobi was less than 10^{-4} at termination for all results in this table.

r	k_t	Branch-and-Bound						Branch-and-Bound with (19)					
		UB	Obj.	Viol.	Nodes	Gap (%)	T(s)	UB	Obj.	Viol.	Nodes	Gap (%)	T(s)
2	2	0.295	**0.295**	0	5100	0.00	10.9	0.295	**0.295**	0	3557	0.00	19.42
2	4	0.404	**0.404**	0	99800	0.01	59.49	0.404	**0.404**	0	109485	0.01	238.48
2	6	0.514	**0.456**	0	1405600	12.72	>600	0.521	0.452	0	1038059	15.20	>600
2	8	0.595	0.467	0	1266600	27.28	>600	0.604	0.465	0	1413722	29.70	>600
2	10	0.633	0.486	0	640900	30.13	>600	0.635	0.488	0	848164	30.03	>600
3	2	0.435	**0.435**	0	22400	0.01	17.92	0.435	**0.435**	0	17924	0.00	176.79
3	4	0.717	0.530	0	542700	35.43	>600	0.753	0.524	0	222894	43.68	>600
3	6	0.846	0.551	0	518100	53.62	>600	0.879	0.560	0	387403	56.96	>600
3	8	0.933	0.585	0	564200	59.45	>600	0.935	0.570	0	612382	64.06	>600
3	10	0.962	0.627	0	383100	53.51	>600	0.963	0.595	0	455231	61.95	>600
4	2	0.566	**0.554**	0	783600	2.02	>600	0.774	**0.554**	0	36000	39.64	>600
4	4	1.089	0.636	0	269900	71.24	>600	1.114	0.610	0	203231	82.70	>600
4	6	1.213	0.642	0	268300	88.92	>600	1.226	0.617	0	201271	98.73	>600
4	8	1.267	0.648	0	251800	95.45	>600	1.266	0.699	0	308274	81.18	>600
4	10	1.289	0.713	0	163200	80.67	>600	1.289	0.714	0	266105	80.62	>600
5	2	0.946	0.641	0	702700	47.60	>600	1.073	0.603	0	44795	77.98	>600
5	4	1.430	0.697	0	215200	105.20	>600	1.468	0.652	0	113951	125.31	>600
5	6	1.555	0.692	0	199400	124.65	>600	1.555	0.686	0	152744	126.63	>600
5	8	1.592	0.761	0	191100	109.14	>600	1.610	0.714	0	156767	125.47	>600
5	10	1.616	0.801	0	147500	101.68	>600	1.620	0.804	0	108311	101.51	>600
6	2	1.323	0.702	0	323700	88.34	>600	1.430	0.698	0	43880	104.84	>600
6	4	1.822	0.761	0	139900	139.43	>600	1.807	0.711	0	93572	154.35	>600
6	6	1.903	0.771	0	114000	146.82	>600	1.907	0.776	0	43754	145.87	>600
6	8	1.932	**0.846**	0	141700	128.42	>600	1.946	0.833	0	62657	133.63	>600
6	10	1.947	**0.868**	0	31200	124.29	>600	1.947	0.864	0	68429	125.36	>600
Avg		1.113	0.623	0	375700	69.04	533.28	1.138	0.613	0	280500	76.22	547.62

Table EC.2 Performance of branch-and-bound without warmstart on the pitprops dataset ($p = 13$) using the experimental setup laid out in Section 5.2, except we use a time limit of 600s for branch-and-bound. We report the performance of branch-and-bound with and without the upper bound developed in Section 5.1 separately. The column "UB" reports the upper bound obtained by the branch-and-bound scheme at the time limit. We use > 600 to denote an instance where branch-and-bound terminates at the 600s time limit. The column gap denotes the relative optimality gap reported by Gurobi at termination (in %). We denote the best-performing solution (in terms of the proportion of variance explained minus the total orthogonality constraint violation) in bold (cont.).

r	k_t	Branch-and-Bound (warm-start)						Branch-and-Bound with (19) (warm-start)					
		UB	Obj.	Viol.	Nodes	Gap (%)	T(s)	UB	Obj.	Viol.	Nodes	Gap (%)	T(s)
2	2	0.295	**0.295**	0	6400	0.00	10.90	0.295	**0.295**	0	143	0.00	19.42
	4	0.404	**0.404**	0	75200	0.01	59.49	0.404	**0.404**	0	88735	0.01	238.5
	6	0.521	0.453	0	1051700	15.10	> 600	0.524	0.445	0	744000	17.89	> 600
	8	0.598	0.463	0	826700	29.31	> 600	0.604	0.462	0	1447055	30.73	> 600
	10	0.633	0.489	0	686100	29.55	> 600	0.636	0.487	0	687944	30.63	> 600
3	2	0.435	**0.435**	0	48600	0.00	17.92	0.435	**0.435**	0	2490	0.00	176.8
	4	0.716	0.536	0	491800	33.59	> 600	0.752	0.524	0	240207	43.42	> 600
	6	0.863	0.560	0	470900	54.16	> 600	0.855	0.567	0	598767	50.86	> 600
	8	0.935	0.578	0	383000	61.89	> 600	0.937	0.566	0	239381	65.68	> 600
	10	0.963	0.603	0	297600	59.74	> 600	0.963	0.595	0	355350	61.91	> 600
4	2	0.554	**0.554**	0	604900	0.01	> 600	0.554	**0.544**	0	29467	38.09	> 600
	4	1.107	0.627	0	264300	76.52	> 600	1.126	0.646	0	136767	74.26	> 600
	6	1.208	0.633	0	250300	90.89	> 600	1.220	0.625	0	247504	95.30	> 600
	8	1.273	0.676	0	225600	88.26	> 600	1.271	0.651	0	216736	95.07	> 600
	10	1.289	0.702	0	210000	83.73	> 600	1.291	0.712	0	165362	81.32	> 600
5	2	0.907	**0.656**	0	279800	38.27	> 600	1.132	0.616	0	28514	83.71	> 600
	4	1.419	0.718	0	251000	97.50	> 600	1.475	0.680	0	141932	116.99	> 600
	6	1.549	0.699	0	225900	121.73	> 600	1.555	0.677	0	229664	129.75	> 600
	8	1.609	0.743	0	223700	116.69	> 600	1.603	0.743	0	200306	115.90	> 600
	10	1.619	0.800	0	126300	102.49	> 600	1.621	0.794	0	135142	103.99	> 600
6	2	1.285	**0.749**	0	253200	71.53	> 600	1.647	**0.749**	0	11818	119.79	> 600
	4	1.821	0.778	0	89900	133.89	> 600	1.859	0.737	0	141476	152.33	> 600
	6	1.891	0.780	0	185600	142.52	> 600	1.906	0.768	0	126387	148.36	> 600
	8	1.927	0.833	0	124300	131.24	> 600	1.942	0.856	0	91093	126.86	> 600
	10	1.948	0.867	0	43200	124.74	> 600	1.948	0.866	0	82759	124.83	> 600
Avg		1.111	0.625	0	307800	68.13	> 600	1.150	0.618	0	255600	76.30	> 600

Table EC.3 Performance of branch-and-bound with warmstart on the pitprops dataset ($p = 13$) using the experimental setup laid out in Section 5.2, except we use a time limit of 600s for branch-and-bound. We report the performance of branch-and-bound with and without the upper bound developed in Section 5.1 separately. We use > 600 to denote an instance where branch-and-bound terminates at the 600s time limit. The column gap denotes the relative optimality gap reported by Gurobi at termination (in %). We denote the best-performing solution (in terms of the proportion of variance explained minus the orthogonality violation) in bold (cont.).

We observe that including the combinatorial upper bound developed in Section 4 within the branch-and-bound scheme does more harm than good, and that using Algorithm 1 as a warmstart marginally improves the performance.

Furthermore, we observe that the upper bound returned by branch-and-bound outperforms the semidefinite upper bound from Problem (12) for the smallest combinations of r and k, but rapidly becomes worse as k and r increases, to the extent that it is unable to provide an upper bound better than the trivial bound of 1 for the largest combinations of r and k. This suggests that the upper bound from branch-and-bound is not practically useful for larger problem instances.

EC.9. Supplementary Numerical Results on UCI Datasets

This section provides supplementary results supporting the numerical experiments performed in Section 5 on UCI datasets.

EC.9.1. Description of the Experimental Setup

All experiments were performed on MIT's supercloud cluster (Reuther et al. 2018), which hosts Intel Xeon Platinum 8260 processors and Intel Xeon Gold 6248 processors. For experiments where $p < 100$, we use Platinum processors with 32 GB RAM, for experiments where $p \in [100, 250]$, Platinum processors with 100 GB RAM; and for $p > 250$, Gold processors with 370 GB RAM.

We also implement some existing algorithmic strategies from the literature, to provide a baseline for the performance of our methods. To abide by software licensing restrictions, all existing strategies from the literature were benchmarked using a MacBook Pro laptop with a 2.9GHz 6-Core Intel i9 CPU, using 16 GB DDR4 RAM. Therefore, runtimes are not directly comparable across strategies.

EC.9.2. Description of the Data Sources

We perform experiments on eleven datasets from the frequently used UCI database in Sections 5.1-5.2 and 5.4. Of the eleven datasets, six datasets are overdetermined (meaning $n > p$), while five datasets are underdetermined (meaning $p > n$). Moreover, many existing works on sparse PCA report results on similar datasets. For instance, the pitprops dataset was also considered by Jolliffe et al. (2003), Zou et al. (2006), Journée et al. (2010) among others, and three of the datasets studied by Berk and Bertsimas (2019) are included within our suite of datasets. Thus, our experimental setup is broadly representative of both the underdetermined and the overdetermined regimes, as well as of the literature. For completeness, we summarize the datasets we benchmark on and their dimensionality in Table EC.4.

Dataset	p	n
Pitprops	13	180
Wine	13	178
Ionosphere	34	351
Lung (Lung cancer)	54	32
Geographical (Geographical Origin of Music)	68	1059
Communities (Communities and Crime)	101	1994
Arrhythmia	274	452
Voice (LSVT Voice Rehabilitation)	310	126
Gait (Gait Classification)	320	48
Gastro (Gastrointestinal Lesions in Regular Colonoscopy)	466	152
Micromass	1300	931

Table EC.4 Summary of the 11 datasets in our library, where n denotes the number of observations and p the number of features. For conciseness, the names of certain datasets are abbreviated throughout. For these datasets, we first state the abbreviation used, followed by their full names in brackets. Further, we report the dimensionality of each dataset after preprocessing, removing all features with missing values. All datasets can be found in the UCI database, except the pitprops dataset, which is due to Jeffers (1967) and distributed via the R package ElasticNet.

EC.9.3. Preliminary Experiments With Pitprops Dataset

We now provide instance-wise results for different variants of our methods on the `pitprops` dataset. In particular, we consider invoking the valid inequalities (21) derived in Section 4 to improve branch-and-bound further. When we do so, we also invoke a branching callback each time we expand a node to determine whether the subtree rooted at this node can improve upon the incumbent solution. This is justified by the fact that at each node, some variables $Z_{i,t}$ are fixed to 0, some to 1, and some are not fixed. Accordingly, we can compute an upper bound on any solution with the same fixed variables by relaxing the orthogonality constraint and applying the Gershgorin circle theorem to each component separately; see Bertsimas et al. (2022b, Section 2.4) for a discussion of this callback in the rank-one case. In particular, if the Gershgorin bound for a given subtree is weaker than an incumbent solution, then this subtree does not contain any optimal solutions, and we can prune it from our search tree.

EC.9.4. Instance-Wise Results on Larger UCI Datasets

Next, we provide an instance-by-instance account of the results summarized in Table 4–Table 5, in Tables EC.5-EC.10.

Dataset	p	r	k_t	Alg. 1				Alg. 2				Branch-and-bound		
				UB	Obj.	Viol.	T(s)	UB	Obj.	Viol.	T(s)	Obj.	Viol.	T(s)
Pitprops	13	2	5	0.449	0.429	0	0.99	0.524	0.433	0	5.01	**0.439**	0	>7200
		2	10	0.507	0.380	0	0.34	0.642	**0.500**	0	4.10	0.498	0	>7200
		3	5	0.616	0.541	0	0.44	0.786	**0.582**	0	6.48	0.555	0	>7200
		3	10	0.652	0.511	0	0.46	0.963	**0.650**	0	6.41	0.618	0	>7200
Wine	13	2	5	0.458	0.401	0	0.35	0.529	0.446	0	2.19	**0.448**	0	1073
		2	10	0.554	0.508	0	0.31	0.707	**0.544**	0	4.03	0.537	0	>7200
		3	5	0.632	0.446	0	0.48	0.794	**0.613**	0	3.56	0.576	0	>7200
		3	10	0.665	0.528	0	0.46	1.060	**0.660**	0	6.47	0.640	0	>7200
Ionosphere	34	2	5	0.209	0.203	0	5.97	0.221	**0.204**	0	3.18	0.202	0	>7200
		2	10	0.305	0.265	0	8.72	0.361	**0.285**	0	3.95	**0.285**	0	>7200
		2	20	0.378	0.286	0	23.31	0.500	**0.360**	0	7.64	0.344	0	>7200
		3	5	0.297	0.287	0	31.04	0.331	0.279	0	6.72	**0.289**	0	>7200
		3	10	0.411	0.305	0	39.00	0.542	0.397	0	11.89	0.375	0	>7200
		3	20	0.464	0.390	0	10.26	0.749	**0.458**	0	12.17	0.383	0	>7200
Lung	54	2	5	0.119	**0.119**	0	17.84	0.124	0.110	0	3.81	0.113	0	>7200
		2	10	0.176	**0.175**	0	30.81	0.178	0.171	0	2.06	0.168	0	>7200
		2	20	0.234	0.170	0	30.82	0.262	0.217	0	4.12	0.185	0	>7200
		3	5	0.173	0.165	0	40.83	0.185	0.160	0	4.45	0.169	0	>7200
		3	10	0.249	0.194	0	90.20	0.267	0.240	0	4.31	0.188	0	>7200
		3	20	0.324	0.308	0	34.50	0.393	0.303	0	9.51	0.213	0	>7200
Geography	68	2	5	0.147	0.145	0	99.45	0.147	**0.147**	0	2.99	0.145	0	>7200
		2	10	0.294	0.290	0	107.27	0.294	**0.294**	0	1.81	0.292	0	>7200
		2	20	0.433	0.393	0	1213.3	0.564	0.376	0	6.32	0.327	0	>7200
		3	5	0.221	0.213	0	119.3	0.221	**0.221**	0	2.58	0.215	0	>7200
		3	10	0.410	0.342	0	1453.19	0.441	0.348	0	5.39	0.355	0	>7200
		3	20	0.529	0.457	0	1571.86	0.846	0.345	0	11.51	0.352	0	>7200
Communities	101	2	5	0.095	**0.095**	0	484.0	0.096	0.078	0	3.73	**0.095**	0	>7200
		2	10	0.169	**0.169**	0	1327	0.175	0.159	0	4.81	0.160	0	>7200
		2	20	0.268	0.219	0	2438	0.284	0.244	0	6.26	0.198	0	>7200
		3	5	0.141	**0.141**	0	979.7	0.144	0.119	0	7.63	**0.141**	0	>7200
		3	10	0.246	0.242	0	3553	0.262	0.243	0	8.92	0.205	0	>7200
		3	20	0.385	0.267	0	3231	0.425	**0.370**	0	8.28	0.300	0	>7200
Arrhythmia	274	2	5	0.031	0.021	0	583.7	0.031	0.027	0	16.57	0.027	0	>7200
		2	10	0.055	0.035	0	555.4	0.055	0.047	0	24.72	0.044	0	>7200
		2	20	0.086	0.067	0	622.8	0.084	0.071	0.002	49.41	0.059	0	>7200
		3	5	0.047	0.031	0	1423.0	0.046	0.039	0	27.99	0.044	0	>7200
		3	10	0.083	0.044	0	1085.8	0.083	0.067	0	38.16	0.065	0	>7200
		3	20	0.129	0.083	0	1059.7	0.126	0.105	0	28.96	0.068	0	>7200

Table EC.5 Performance of Algorithms 1–2 and branch-and-bound on UCI datasets. k_t denotes the sparsity of each individual component, meaning a set of r PCs have a collective sparsity budget of $k_t r$. Note that all objective values are reported in terms of the proportion of correlation explained by dividing by p, the number of features.

Dataset	p	r	k_t	Alg. 1				Alg. 2				Branch-and-bound		
				UB	Obj.	Viol.	T(s)	UB	Obj.	Viol.	T(s)	Obj.	Viol.	T(s)
Voice	310	2	5	0.032	0.024	0	375.0	0.032	**0.032**	0	21.14	**0.032**	0	>7200
		2	10	0.064	**0.064**	0	741.2	0.064	0.063	0	21.99	**0.064**	0	>7200
		2	20	0.127	**0.127**	0	630.6	0.127	0.124	0	20.97	0.109	0	>7200
		3	5	0.048	**0.048**	0	758.3	0.048	0.047	0	29.94	**0.048**	0	>7200
		3	10	0.096	0.079	0	797.1	0.096	0.093	0	34.52	**0.096**	0	>7200
		3	20	0.191	0.125	0	721.5	0.191	0.183	0	31.06	0.155	0	>7200
Gait	320	2	5	0.031	0.018	0	399.2	0.031	0.028	0	19.40	0.028	0	>7200
		2	10	0.057	0.036	0	450.7	0.057	0.050	0	24.64	0.047	0	>7200
		2	20	0.103	0.062	0	392.3	0.103	0.081	0	24.51	0.067	0	>7200
		3	5	0.046	0.036	0	933.7	0.046	0.041	0	27.80	**0.045**	0	>7200
		3	10	0.085	0.049	0	829.3	0.085	0.077	0	34.99	0.060	0	>7200
		3	20	0.154	0.060	0	792.1	0.154	0.121	0	34.11	0.111	0	>7200
Gastro	466	2	5	0.021	**0.021**	0	1633	0.021	**0.021**	0	452.9	**0.021**	0	>7200
		2	10	0.043	**0.043**	0	1753	0.043	**0.043**	0	38.95	**0.043**	0	>7200
		2	20	0.086	0.085	0	2166	0.086	0.085	0	535.4	**0.086**	0	>7200
		3	5	0.032	**0.032**	0	2682	0.032	**0.032**	0	59.97	**0.032**	0	>7200
		3	10	0.064	**0.064**	0	4307	0.064	**0.064**	0	73.39	**0.064**	0	>7200
		3	20	0.129	0.126	0	5544	0.128	0.121	0	529.8	**0.128**	0	>7200
Micromass	1300	2	5	0.008	0.005	0	1089	0.008	0.006	0	170.5	0.004	0	>7200
		2	10	0.015	0.008	0	13620	0.014	0.011	0	158.6	0.011	0	6100
		2	20	0.027	0.018	0	9213	0.023	0.019	0	440.3	0.018	0	6826
		3	5	0.012	0.008	0	7953	0.011	0.010	0	238.0	0.006	0	>7200
		3	10	0.023	0.009	0	19640	0.021	0.018	0	208.5	0.009	0	>7200
		3	20	0.043	0.029	0	18630	0.034	0.030	0	205.9	0.010	0	>7200
Avg				0.213	0.176	0.000	1899	0.257	0.199	0.000	61.38	0.187	0	>7200

Table EC.6 Performance of Algorithms 1–2 and branch-and-bound on UCI datasets (cont). k_t denotes the sparsity of each individual component, meaning a set of r PCs has a collective sparsity budget of $k_t r$. Note that all objective values are reported in terms of the proportion of correlation explained by dividing by p, the number of features.

Dataset	p	r	k_t	Berk and Bertsimas (2019)			Hein and Bühler (2010)			Zou et al. (2006)		
				Obj.	Viol.	T(s)	Obj.	Viol.	T(s)	Obj.	Viol.	T(s)
Pitprops	13	2	5	0.421	0.168	1.67	0.418	0	0.11	0.177	1.341	0.12
		2	10	0.502	0.008	0.14	0.502	0.008	0.01	0.139	1.827	0.22
		3	5	0.592	0.675	0.08	0.575	0.166	0.02	0.169	3.462	0.04
		3	10	0.648	0.073	0.07	0.647	0.084	0	0.181	3.771	0.36
Wine	13	2	5	**0.448**	0	0.04	0.422	0.004	0.01	0.127	0.315	0.04
		2	10	0.545	0.020	0.04	0.545	0.02	0	0.068	0.731	0.06
		3	5	0.610	0.019	0.06	0.559	0.092	0	0.225	2.830	0.05
		3	10	0.654	0.059	0.06	0.655	0.093	0	0.232	2.771	0.32
Ionosphere	34	2	5	**0.205**	0	0.08	0.153	0	0.08	0.078	0	0.02
		2	10	0.289	0	0.30	0.288	0	0.01	0.106	0	0.04
		2	20	0.369	0.058	4.45	**0.370**	0.010	0.17	0.147	0.305	0.12
		3	5	0.291	0	0.14	0.227	0	0.02	0.097	1.666	0.07
		3	10	0.392	0.109	0.38	0.365	0.255	0.01	0.100	1.909	0.12
		3	20	0.449	0.183	0.27	0.451	0.037	0.03	0.111	2.111	4.51
Lung	54	2	5	**0.119**	0	0.43	0.107	0	0.34	0.040	0.587	0.04
		2	10	**0.176**	0	0.10	0.170	0	0.03	0.044	0.639	0.12
		2	20	0.220	0.008	0.44	0.184	0	0.05	0.044	0.908	0.63
		3	5	**0.172**	0	0.10	0.149	0	0.03	0.061	2.755	0.15
		3	10	**0.243**	0	0.11	0.234	0.113	0.05	0.054	1.593	0.20
		3	20	0.300	0.219	0.16	0.261	0.081	0.04	0.044	1.703	1.20
Geography	68	2	5	**0.147**	0	0.09	0.097	0	0.01	0.034	1.793	0.41
		2	10	**0.294**	0	0.08	0.164	0	0	0.068	1.939	0.66
		2	20	0.395	0	5.95	0.316	0.135	0.04	0.062	1.754	0.53
		3	5	**0.221**	0	0.13	0.122	0	0.01	0.061	2.720	0.57
		3	10	**0.389**	0	0.18	0.192	0	0.01	0.054	4.021	0.90
		3	20	0.484	0.273	23.66	0.387	0.261	0.06	0.090	5.009	1.40
Communities	101	2	5	**0.095**	0	0.73	0.093	0	0	0.032	0.576	0.05
		2	10	**0.169**	0	1.68	0.154	0	0	0.029	0.605	0.18
		2	20	**0.258**	0	120	**0.258**	0	0.07	0.027	0.090	1.49
		3	5	**0.141**	0	1.18	0.129	0	0.01	0.050	1.854	0.29
		3	10	**0.245**	0	2.85	0.181	0	0.02	0.044	1.504	1.76
		3	20	0.361	0.058	180.1	0.350	0.064	0.02	0.043	1.869	5.27
Arrhythmia	274	2	5	**0.031**	0	2.81	0.012	0	0.02	0.007	1.799	0.71
		2	10	**0.052**	0	61.25	0.011	0	0.03	0.007	1.143	1.08
		2	20	**0.077**	0	120.0	0.043	0.005	0.06	0.006	1.140	4.62
		3	5	**0.046**	0	5.35	0.016	0	0.02	0.012	1.076	0.53
		3	10	**0.074**	0	121.6	0.018	0	0.05	0.012	0.876	3.82
		3	20	**0.109**	0	180.0	0.074	0.005	0.07	0.012	0.694	10.65

Table EC.7 Performance of the methods of Berk and Bertsimas (2019), Hein and Bühler (2010), and Zou et al. (2006) on UCI datasets.

Dataset	p	r	k_t	Berk and Bertsimas (2019)			Hein and Bühler (2010)			Zou et al. (2006)		
				Obj.	Viol.	T(s)	Obj.	Viol.	T(s)	Obj.	Viol.	T(s)
Voice	310	2	5	**0.032**	0	1.04	**0.032**	0	0.05	0.006	0.874	0.71
		2	10	**0.064**	0	1.09	**0.064**	0	0.04	0.006	0.907	2.68
		2	20	**0.127**	0	0.66	**0.127**	0	0.02	0.006	1.017	16.02
		3	5	**0.048**	0	1.72	0.039	0	0.16	0.009	1.834	0.94
		3	10	**0.096**	0	1.56	0.069	0	0.08	0.012	0.242	12.8
		3	20	**0.190**	0	1.22	0.187	0	0.12	0.021	2.121	26.18
Gait	320	2	5	**0.030**	0	0.93	0.027	0	0.02	0.006	1.071	1.75
		2	10	**0.055**	0	0.61	0.051	0	0.05	0.004	1.054	1.27
		2	20	0.094	0	1.36	0.080	0	0.080	0.005	0.852	3.88
		3	5	**0.045**	0	1.02	0.041	0	0.06	0.01	1.16	4.81
		3	10	**0.082**	0	1.64	0.070	0	0.06	0.009	1.821	5
		3	20	**0.135**	0	1.94	0.095	0	0.15	0.008	1.477	9.49
Gastro	466	2	5	**0.021**	0	2.71	0.020	0	0.05	0.007	1.154	1.14
		2	10	**0.043**	0	1.42	0.039	0	0.08	0.007	0.178	1.62
		2	20	**0.086**	0	1.95	0.076	0	0.07	0.005	0.456	3.61
		3	5	**0.032**	0	2.79	0.029	0	0.05	0.006	2.303	1.59
		3	10	**0.064**	0	3.51	0.053	0	0.12	0.007	1.826	3.33
		3	20	**0.128**	0	2.81	0.085	0	0.24	0.008	1.139	46.05
Micromass	1300	2	5	**0.008**	0	45.4	0.004	0	0.77	0.002	0.014	18.05
		2	10	**0.014**	0	120.2	0.007	0	1.09	0.002	0.323	41.97
		2	20	**0.023**	0	120.2	0.012	0	1.34	0.002	0.361	2.64
		3	5	**0.011**	0	71.97	0.005	0	1.54	0.002	3.004	24.30
		3	10	**0.020**	0	180.3	0.008	0	1.80	0.002	2.301	31.89
		3	20	**0.034**	0	180.3	0.013	0	2.64	0.002	1.252	2.02
Avg				0.205	0.031	25.57	0.180	0.023	0.19	0.049	1.458	4.95

Table EC.8 Performance of the methods of Berk and Bertsimas (2019), Hein and Bühler (2010), and Zou et al. (2006) on UCI datasets (cont.).

Dataset	p	r	k_t	Deshpande and Montanari (2014b)			Algorithm 3			
				Obj.	Viol.	T(s)	UB	Obj.	Viol.	T(s)
Pitprops	13	2	5	0.422	0.226	1.00	0.559	0.422	0	0.53
		2	10	0.501	0.104	0.00	0.803	0.456	0	0.05
		3	5	0.592	0.661	0.00	0.827	0.568	0	0.13
		3	10	0.644	0.214	0.00	1.198	0.569	0	0.13
Wine	13	2	5	0.434	0.211	0.02	0.579	0.447	0	0.02
		2	10	0.545	0.021	0	0.876	0.508	0	0.06
		3	5	0.546	0.510	0	0.853	0.577	0	0.21
		3	10	0.656	0.228	0	1.296	0.580	0	0.23
Ionosphere	34	2	5	0.203	0	0.12	0.228	0.202	0	0.08
		2	10	**0.287**	0	0.02	0.401	**0.290**	0	0.14
		2	20	0.368	0.011	0.02	0.618	0.357	0	0.2
		3	5	0.276	0	0.02	0.340	**0.292**	0	0.28
		3	10	0.357	0.143	0.02	0.597	**0.398**	0	1.51
		3	20	0.447	0.166	0.02	0.920	0.408	0	2.32
Lung	54	2	5	0.117	0	0.45	0.139	0.118	0	0.71
		2	10	0.154	0.291	0.04	0.218	0.171	0	0.1
		2	20	0.217	0.495	0.03	0.345	0.216	0	0.12
		3	5	0.164	0	0.03	0.204	0.169	0	0.23
		3	10	0.200	0.558	0.03	0.326	0.225	0	0.56
		3	20	0.290	0.591	0.03	0.514	0.271	0	3.54
Geography	68	2	5	0.130	0	0.09	0.147	**0.147**	0	0.64
		2	10	0.294	0	0.07	0.294	**0.294**	0	0.81
		2	20	0.395	0	0.08	0.568	**0.396**	0	0.61
		3	5	0.184	0	0.08	0.221	**0.221**	0	2.79
		3	10	0.378	0	0.07	0.441	**0.389**	0	2.04
		3	20	0.462	0.073	0.08	0.852	0.479	0	2.56
Communities	101	2	5	0.089	0	0.19	0.097	**0.095**	0	2.3
		2	10	0.158	0	0.23	0.180	0.167	0	1.37
		2	20	0.249	0.145	0.19	0.320	**0.259**	0	1.94
		3	5	0.140	0	0.19	0.146	**0.141**	0	1.14
		3	10	0.208	0.138	0.20	0.270	0.242	0	1.61
		3	20	0.317	0.669	0.19	0.476	0.369	0	3.62
Arrhythmia	274	2	5	0.030	0	1.79	0.032	**0.031**	0	13.71
		2	10	0.051	0.113	1.84	0.060	0.051	0	11.26
		2	20	0.074	0.019	1.88	0.105	0.075	0	8.04
		3	5	0.039	0	1.97	0.049	**0.046**	0	699.5
		3	10	0.072	0	2.35	0.089	0.072	0	223.1
		3	20	0.103	0.243	2.23	0.155	0.107	0	105.5

Table EC.9 Performance of the method of Deshpande and Montanari (2014b) and Algorithm 3 on UCI datasets.

Dataset	p	r	k_t	Deshpande and Montanari (2014b)			Algorithm 3			
				Obj.	Viol.	T(s)	UB	Obj.	Viol.	T(s)
Voice	310	2	5	**0.032**	0	60.46	0.032	**0.032**	0	16.14
		2	10	0.063	0	54.59	0.064	**0.064**	0	14.22
		2	20	0.124	0	39.19	0.128	**0.127**	0	28.01
		3	5	0.047	0	1.42	0.048	**0.048**	0	64.25
		3	10	0.093	0	1.75	0.097	**0.096**	0	66.68
		3	20	0.184	0	1.42	0.192	**0.190**	0	93.89
Gait	320	2	5	**0.030**	0	1.45	0.031	**0.030**	0	12.51
		2	10	0.054	0	1.45	0.058	**0.055**	0	10.94
		2	20	0.089	0	1.42	0.110	**0.096**	0	12.27
		3	5	0.042	0	1.85	0.046	**0.045**	0	47.73
		3	10	0.078	0	1.65	0.087	**0.082**	0	29.65
		3	20	0.128	0	3.04	0.164	**0.135**	0	46.5
Gastro	466	2	5	**0.021**	0	6.59	0.021	**0.021**	0	35.41
		2	10	0.042	0	4.33	0.043	**0.043**	0	30.51
		2	20	0.083	0	3.67	0.086	**0.086**	0	54.39
		3	5	0.031	0	3.26	0.032	**0.032**	0	97.05
		3	10	0.061	0	3.12	0.064	**0.064**	0	115.2
		3	20	0.117	0	3.11	0.129	**0.128**	0	153.1
Micromass	1300	2	5	0.007	0	75.38	0.008	**0.008**	0	982.8
		2	10	0.012	0	147.9	0.014	0.013	0	615.4
		2	20	0.022	0	98.34	0.025	**0.022**	0	481.8
		3	5	0.010	0	71.05	0.012	0.011	0	1023.8
		3	10	0.018	0	71.88	0.021	0.020	0	1015.1
		3	20	0.033	0	73.00	0.038	**0.033**	0	593.0
Avg				0.197	0.094	12.04	0.289	0.199	0	108.6

Table EC.10 Performance of the method of Deshpande and Montanari (2014b) and Algorithm 3 on UCI datasets (cont.)

EC.9.5. Instance-Wise Plots of Symmetry vs. Proportion of Correlation Explained

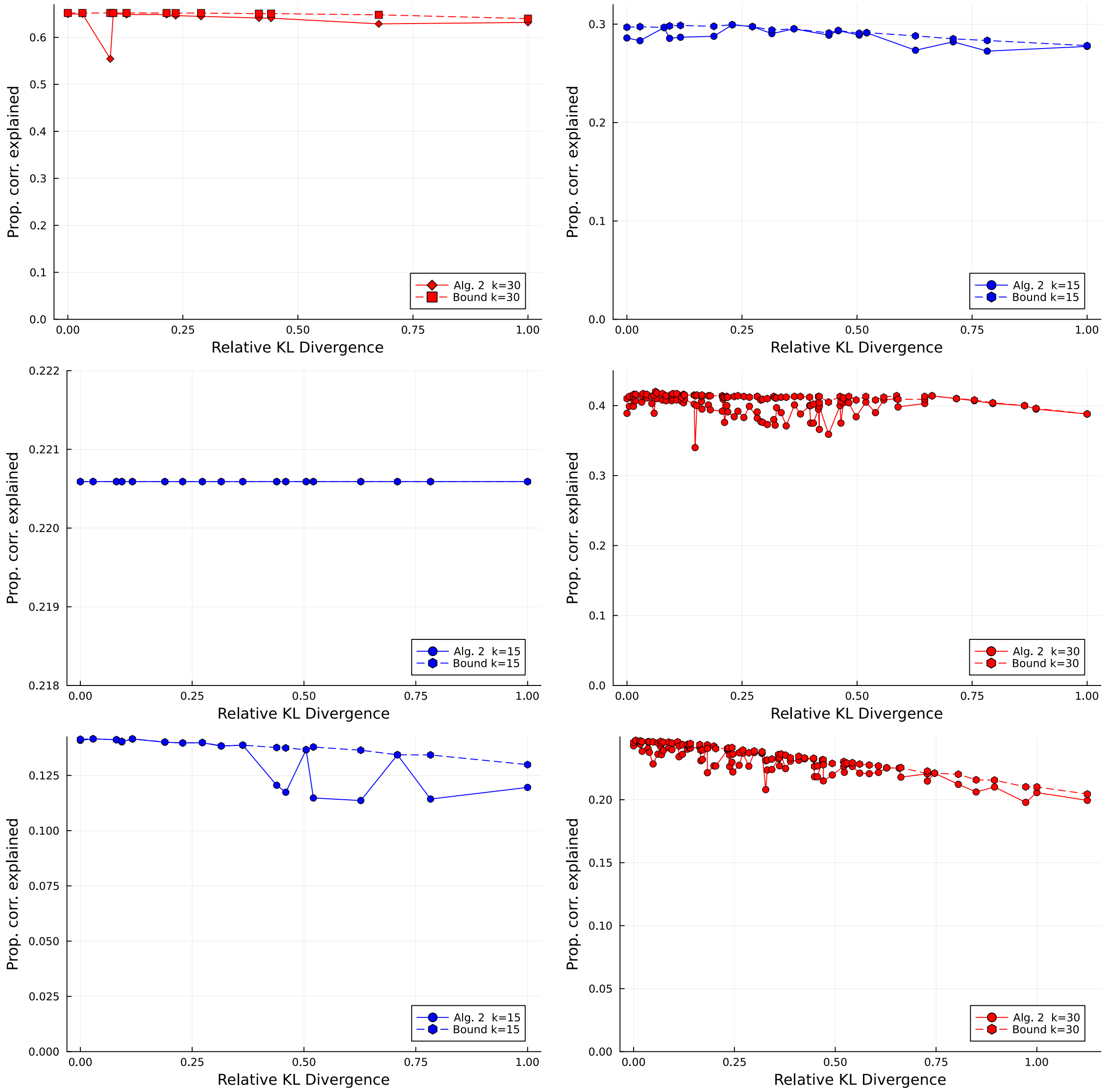

Figure EC.1 Symmetry of sparsity budget allocation vs. proportion of correlation in the dataset explained for pitprops $k = 30$ (top left), ionosphere $k = 15$ (top right), geographical $k = 15$ (middle left), geographical $k = 30$ (middle right), communities $k = 15$ (bottom left), and communities $k = 30$ (bottom right). Note that we normalize the KL divergence for $k = 15$ and $k = 30$ separately.

EC.10. Non-Redundancy of Rank Constraints in Problem (7)

We claimed in Remark 2 that the rank-one constraints in Problem (7) are not redundant. We now demonstrate this by example, by providing an example where, after constraining the support pattern, Problem (7) attains a different optimal value than the following optimization problem:

$$\max_{\substack{\boldsymbol{Z}\in\{0,1\}^{p\times r}:\\ \langle \boldsymbol{E},\boldsymbol{Z}\rangle\le k}} \max_{\boldsymbol{Y}\in\mathcal{S}^p,\boldsymbol{Y}^t\in\mathcal{S}^p_+} \langle \boldsymbol{Y},\boldsymbol{\Sigma}\rangle \text{ s.t. } \quad \boldsymbol{Y}\preceq \mathrm{Diag}\left(\min\left(\boldsymbol{e},\sum_t \boldsymbol{Z}_t\right)\right), \boldsymbol{Y}=\sum_{t=1}^r \boldsymbol{Y}^t, \qquad \text{(EC.14)}$$

$$\mathrm{tr}(\boldsymbol{Y}^t)=1,\ \forall t\in[r],\ Y^t_{i,j}=0 \text{ if } Z_{i,t}=0,\ \forall t\in[r], i,j\in[p],$$

where we take $r=3, p=4$, and fix the support in both problems by setting

$$\boldsymbol{Z}=\begin{pmatrix}1&1&0\\1&0&1\\0&1&1\\0&1&1\end{pmatrix},\ \boldsymbol{\Sigma}=\begin{pmatrix}1&1&0&0\\1&8&2&1\\0&2&2&1\\0&1&1&2\end{pmatrix}.$$

Solving Problem (7) by letting $\boldsymbol{Y}_t=\boldsymbol{u}_t\boldsymbol{u}_t^\top$ via Gurobi with `NonConvex`$=2$ gives an optimal objective value of 12.14006. On the other hand, solving Problem (EC.14) via Mosek gives an optimal objective value of 12.25765. Thus, Problems (7)–(EC.14) cannot be equivalent, as they give a different optimal objective value for a fixed binary support.