SUPPLEMENT TO "BIDIMENSIONAL LINKED MATRIX FACTORIZATION FOR PAN-OMICS PAN-CANCER ANALYSIS"

By Eric F. Lock[1], Jun Young Park[2] and Katherine A. Hoadley[3]

[1]Division of Biostatistics, School of Public Health, University of Minnesota, elock@umn.edu

[2]Department of Statistical Sciences, Faculty of Arts & Science, University of Toronto, junjy.park@utoronto.ca

[3]Department of Genetics, Computational Medicine Program, University of North Carolina, hoadley@med.unc.edu

APPENDIX A: MODULE ENUMERATION

As the default representation of model (6) in the main article, set $K = (2^I - 1)(2^J - 1)$ and let $\mathbf{R}$ and $\mathbf{C}$ enumerate all possible modules as follows. For $k = 1, \ldots, K$, let $\mathbf{R}[\bullet, k]$ be the I-digit binary representation for $k \bmod (2^I - 1) + 1$, where mod gives the modulo (remainder) operator. For $k = 1, \ldots, K$, let $\mathbf{C}[\bullet, k]$ give the J-digit binary representation for $\lceil k/(2^I - 1) \rceil$, where $\lceil \cdot \rceil$ gives the ceiling operator.

APPENDIX B: PROOFS

B.1. Proof of Proposition 1.

Proof. Let $\{\hat{\mathbf{S}}_{\bullet\bullet}^{(k)}\}_{k=1}^K \in \mathbb{S}_{\hat{\mathbf{X}}}$ be a minimizer of the objective function $f(\bullet)$. Assume a violation of condition 1., wherein $\lambda_{k'} \geq \lambda_k$. Consider another minimizer $\{\tilde{\mathbf{S}}_{\bullet\bullet}^{(k)}\}_{k=1}^K$, where $\widetilde{\mathbf{S}}_{\bullet\bullet}^{(k)} = \mathbf{0}$ and $\widetilde{\mathbf{S}}_{\bullet\bullet}^{(k')} = \widehat{\mathbf{S}}_{\bullet\bullet}^{(k)} + \widehat{\mathbf{S}}_{\bullet\bullet}^{(k')}$, and all other modules are equal. Then, using the triangle inequality,

$$
\begin{aligned}
f(\{\hat{\mathbf{S}}_{\bullet\bullet}^{(k)}\}_{k=1}^K) - f(\{\tilde{\mathbf{S}}_{\bullet\bullet}^{(k)}\}_{k=1}^K) &= \lambda_k ||\widehat{\mathbf{S}}_{\bullet\bullet}^{(k)}||_* + \lambda_{k'} ||\widehat{\mathbf{S}}_{\bullet\bullet}^{(k')}||_* - \lambda_k ||\widehat{\mathbf{S}}_{\bullet\bullet}^{(k)} + \widehat{\mathbf{S}}_{\bullet\bullet}^{(k')}||_* \\
&\geq \lambda_k ||\widehat{\mathbf{S}}_{\bullet\bullet}^{(k)}||_* + \lambda_{k'} ||\widehat{\mathbf{S}}_{\bullet\bullet}^{(k')}||_* - \lambda_k (||\widehat{\mathbf{S}}_{\bullet\bullet}^{(k)}||_* + ||\widehat{\mathbf{S}}_{\bullet\bullet}^{(k')}||_*) \\
&\geq \lambda_k ||\widehat{\mathbf{S}}_{\bullet\bullet}^{(k)}||_* + \lambda_{k'} ||\widehat{\mathbf{S}}_{\bullet\bullet}^{(k')}||_* - \lambda_k (||\widehat{\mathbf{S}}_{\bullet\bullet}^{(k)}||_*) - \lambda_{k'} ||\widehat{\mathbf{S}}_{\bullet\bullet}^{(k')}||_*) \\
&= 0,
\end{aligned}
$$

and thus there is a solution in which module k is $\mathbf{0}$, regardless of the data $\mathbf{X}_{\bullet\bullet}$.

Now assume a violation of condition 2., wherein $\lambda_k \geq \sum_{j \in \mathcal{I}_k} \lambda_j$. Let $\widehat{\mathbf{S}}^{(k)} = \sum_{j \in \mathcal{I}_k} \widehat{\mathbf{S}}_{\bullet\bullet}^{j\prime}$, where $\widehat{\mathbf{S}}_{\bullet\bullet}^{j\prime}$ contains the submatrix of $\widehat{\mathbf{S}}_{\bullet\bullet}^{(k)}$ corresponding to $\mathbf{R}[\bullet, j]$ and $\mathbf{C}[\bullet, j]$ and $\mathbf{0}$ otherwise. Consider another decomposition $\{\tilde{\mathbf{S}}_{\bullet\bullet}^{(k)}\}_{k=1}^K$, where $\widetilde{\mathbf{S}}_{\bullet\bullet}^{(k)} = \mathbf{0}$ and $\widetilde{\mathbf{S}}_{\bullet\bullet}^{(j)} = \widehat{\mathbf{S}}_{\bullet\bullet}^{(j)} + \widehat{\mathbf{S}}_{\bullet\bullet}^{(j)\prime}$ for all $j \in \mathcal{I}_k$. Then,

$$
\begin{aligned}
f(\{\hat{\mathbf{S}}_{\bullet\bullet}^{(k)}\}_{k=1}^K) - f(\{\tilde{\mathbf{S}}_{\bullet\bullet}^{(k)}\}_{k=1}^K) &= \lambda_k ||\widehat{\mathbf{S}}_{\bullet\bullet}^{(k)}||_* + \sum_{j \in \mathcal{I}_k} \lambda_j ||\widehat{\mathbf{S}}_{\bullet\bullet}^{(j)}||_* - \sum_{j \in \mathcal{I}_k} \lambda_j ||\widehat{\mathbf{S}}_{\bullet\bullet}^{(j)} + \widehat{\mathbf{S}}_{\bullet\bullet}^{(j)\prime}||_* \\
&\geq \lambda_k ||\widehat{\mathbf{S}}_{\bullet\bullet}^{(k)}||_* + \sum_{j \in \mathcal{I}_k} \lambda_j ||\widehat{\mathbf{S}}_{\bullet\bullet}^{(j)}||_* - \sum_{j \in \mathcal{I}_k} \lambda_j ||\widehat{\mathbf{S}}_{\bullet\bullet}^{(j)}||_* - \sum_{j \in \mathcal{I}_k} \lambda_j ||\widehat{\mathbf{S}}_{\bullet\bullet}^{(j)\prime}||_* \\
&= \lambda_k ||\widehat{\mathbf{S}}_{\bullet\bullet}^{(k)}||_* - \sum_{j \in \mathcal{I}_k} \lambda_j ||\widehat{\mathbf{S}}_{\bullet\bullet}^{(j)\prime}||_* \\
&\geq \lambda_k ||\widehat{\mathbf{S}}_{\bullet\bullet}^{(k)}||_* - \sum_{j \in \mathcal{I}_k} \lambda_j ||\widehat{\mathbf{S}}_{\bullet\bullet}^{(k)}||_* \\
&\geq 0,
\end{aligned}
$$

and thus there is a solution in which module k is 0, regardless of the data $\mathbf{X}_{\cdot\cdot}$. □

B.2. Proof of Proposition 4.

PROOF. We show that $\lambda_k = \sqrt{\mathbf{R}[\bullet,k]\cdot\mathbf{M}} + \sqrt{\mathbf{C}[\bullet,k]\cdot\mathbf{N}}$ satisfies the necessary conditions of Proposition **??**. For condition 1., note that $\sqrt{\mathbf{R}[\bullet,k]\cdot\mathbf{M}} + \sqrt{\mathbf{C}[\bullet,k]\cdot\mathbf{N}} > \sqrt{\mathbf{R}[\bullet,j]\cdot\mathbf{M}} + \sqrt{\mathbf{C}[\bullet,j]\cdot\mathbf{N}}$.

For condition 2., note that

$$\begin{aligned}
\sum_{j\in\mathcal{I}_k}\sqrt{\mathbf{R}[\bullet,j]\cdot\mathbf{M}} + \sqrt{\mathbf{C}[\bullet,j]\cdot\mathbf{N}} &\geq \sqrt{\sum_{j\in\mathcal{I}_k}\mathbf{R}[\bullet,j]\cdot\mathbf{M}} + \sqrt{\sum_{j\in\mathcal{I}_k}\mathbf{C}[\bullet,j]\cdot\mathbf{N}} \\
&= \sqrt{r\cdot\mathbf{R}[\bullet,k]\cdot\mathbf{M}} + \sqrt{c\cdot\mathbf{C}[\bullet,k]\cdot\mathbf{N}} \\
&> \sqrt{\mathbf{R}[\bullet,k]\cdot\mathbf{M}} + \sqrt{\mathbf{C}[\bullet,k]\cdot\mathbf{N}}
\end{aligned}$$

□

B.3. Proof of Proposition 5. Lemmas 1 and 2 below are used to establish Proposition 5.

LEMMA 1. *Take two decompositions* $\{\hat{\mathbf{S}}_{\cdot\cdot}^{(k)}\}_{k=1}^K \in \mathbb{S}_{\hat{\mathbf{X}}}$ *and* $\{\tilde{\mathbf{S}}_{\cdot\cdot}^{(k)}\}_{k=1}^K \in \mathbb{S}_{\hat{\mathbf{X}}}$, *and assume that both minimize the structured nuclear norm penalty:*

$$f_{pen}(\{\hat{\mathbf{S}}_{\cdot\cdot}^{(k)}\}_{k=1}^K) = f_{pen}\left(\{\tilde{\mathbf{S}}_{\cdot\cdot}^{(k)}\}_{k=1}^K\right) = \min_{\mathbb{S}_{\hat{\mathbf{X}}}} f_{pen}(\{\mathbf{S}_{\cdot\cdot}^{(k)}\}_{k=1}^K).$$

Then, for any $\alpha\in[0,1]$,

$$||\alpha\hat{\mathbf{S}}_{\cdot\cdot}^{(k)} + (1-\alpha)\tilde{\mathbf{S}}_{\cdot\cdot}^{(k)}||_* = \alpha||\hat{\mathbf{S}}_{\cdot\cdot}^{(k)}||_* + (1-\alpha)||\tilde{\mathbf{S}}_{\cdot\cdot}^{(k)}||_*$$

for $k = 1,\dots,K$.

PROOF. Because $\mathbb{S}_{\hat{\mathbf{X}}}$ is a convex space and f_{pen} is a convex function, the set of minimizers of f_{pen} over $\mathbb{S}_{\hat{\mathbf{X}}}$ is also convex. Thus,

$$f_{\text{pen}}\left(\{\alpha\hat{\mathbf{S}}_{\cdot\cdot}^{(k)} + (1-\alpha)\tilde{\mathbf{S}}_{\cdot\cdot}^{(k)}\}_{k=1}^K\right) = \min_{\mathbb{S}_{\hat{\mathbf{X}}}} f_{\text{pen}}(\{\mathbf{S}_{\cdot\cdot}^{(k)}\}_{k=1}^K).$$

The result follows from the convex property of the nuclear norm operator, which implies that for any two matrices of equal size $\hat{\mathbf{A}}$ and $\tilde{\mathbf{A}}$,

$$||\alpha\hat{\mathbf{A}} + (1-\alpha)\tilde{\mathbf{A}}||_* \leq \alpha||\hat{\mathbf{A}}||_* + (1-\alpha)||\tilde{\mathbf{A}}||_*. \tag{1}$$

Applying (1) to each additive term in f_{pen} gives

$$\begin{aligned}
f_{\text{pen}}\left(\{\alpha\hat{\mathbf{S}}_{\cdot\cdot}^{(k)} + (1-\alpha)\tilde{\mathbf{S}}_{\cdot\cdot}^{(k)}\}_{k=1}^K\right) &\leq \alpha f_{\text{pen}}(\{\hat{\mathbf{S}}_{\cdot\cdot}^{(k)}\}_{k=1}^K) + (1-\alpha)f_{\text{pen}}(\{\tilde{\mathbf{S}}_{\cdot\cdot}^{(k)}\}_{k=1}^K) \\
&= \min_{\mathbb{S}_{\hat{\mathbf{X}}}} f_{\text{pen}}(\{\mathbf{S}_{\cdot\cdot}^{(k)}\}_{k=1}^K).
\end{aligned} \tag{2}$$

Because $\{\alpha\hat{\mathbf{S}}_{\cdot\cdot}^{(k)} + (1-\alpha)\tilde{\mathbf{S}}_{\cdot\cdot}^{(k)}\}_{k=1}^K \in \mathbb{S}_{\hat{\mathbf{X}}}$, the inequality in (2) must be an equality, and it follows that the inequality (1) must be an equality for each penalized term in the decomposition. □

LEMMA 2. *Take two matrices* $\hat{\mathbf{A}}$ *and* $\tilde{\mathbf{A}}$. *If* $||\hat{\mathbf{A}}+\tilde{\mathbf{A}}||_* = ||\hat{\mathbf{A}}||_* + ||\tilde{\mathbf{A}}||_*$, *and* $\mathbf{U}\mathbf{D}_+\mathbf{V}^T$ *is the SVD of* $\hat{\mathbf{A}}+\tilde{\mathbf{A}}$, *then* $\hat{\mathbf{A}} = \hat{\mathbf{U}}\hat{\mathbf{D}}\hat{\mathbf{V}}^T$ *where* $\hat{\mathbf{D}}$ *is diagonal and* $||\hat{\mathbf{A}}||_* = ||\hat{\mathbf{D}}||_*$, *and* $\tilde{\mathbf{A}} = \mathbf{U}\tilde{\mathbf{D}}\mathbf{V}^T$ *where* $\tilde{\mathbf{D}}$ *is diagonal and* $||\tilde{\mathbf{A}}||_* = ||\tilde{\mathbf{D}}||_*$.

PROOF. Here we use the fact that the spectral norm is dual to the nuclear norm (Fazel et al., 2001). That is, if $\sigma_1(\mathbf{Z})$ is the maximum singular value of $\mathbf{Z}$ (i.e., the spectral norm), then

$$||\mathbf{A}||_* = \sup_{\sigma_1(\mathbf{Z})=1} \langle \mathbf{Z}, \mathbf{A} \rangle.$$

Thus,

$$\sup_{\sigma_1(\mathbf{Z})=1} \langle \mathbf{Z}, \hat{\mathbf{A}} + \tilde{\mathbf{A}} \rangle = \sup_{\sigma_1(\mathbf{Z})=1} \langle \mathbf{Z}, \hat{\mathbf{A}} \rangle + \sup_{\sigma_1(\mathbf{Z})=1} \langle \mathbf{Z}, \tilde{\mathbf{A}} \rangle. \tag{3}$$

By the properties of the SVD,

$$\langle \mathbf{U}\mathbf{V}^T, \tilde{\mathbf{A}} \rangle + \langle \mathbf{U}\mathbf{V}^T, \hat{\mathbf{A}} \rangle = \langle \mathbf{U}\mathbf{V}^T, \hat{\mathbf{A}} + \tilde{\mathbf{A}} \rangle = \sup_{\sigma_1(\mathbf{Z})=1} \langle \mathbf{Z}, \hat{\mathbf{A}} + \tilde{\mathbf{A}} \rangle. \tag{4}$$

By (3) and (4),

$$\langle \mathbf{U}\mathbf{V}^T, \tilde{\mathbf{A}} \rangle + \langle \mathbf{U}\mathbf{V}^T, \hat{\mathbf{A}} \rangle = \sup_{\sigma_1(\mathbf{Z})=1} \langle \mathbf{Z}, \hat{\mathbf{A}} \rangle + \sup_{\sigma_1(\mathbf{Z})=1} \langle \mathbf{Z}, \tilde{\mathbf{A}} \rangle. \tag{5}$$

Because $\mathbf{U}\mathbf{V}^T \in \{\mathbf{Z} : \sigma_1(\mathbf{Z}) = 1\}$ it follows that

$$\langle \mathbf{U}\mathbf{V}^T, \hat{\mathbf{A}} \rangle \leq \sup_{\sigma_1(\mathbf{Z})=1} \langle \mathbf{Z}, \hat{\mathbf{A}} \rangle \text{ and } \langle \mathbf{U}\mathbf{V}^T, \tilde{\mathbf{A}} \rangle \leq \sup_{\sigma_1(\mathbf{Z})=1} \langle \mathbf{Z}, \tilde{\mathbf{A}} \rangle,$$

and so (5) implies

$$\langle \mathbf{U}\mathbf{V}^T, \hat{\mathbf{A}} \rangle = \sup_{\sigma_1(\mathbf{Z})=1} \langle \mathbf{Z}, \hat{\mathbf{A}} \rangle = ||\hat{\mathbf{A}}||_*,$$

and similarly $\langle \mathbf{U}\mathbf{V}^T, \tilde{\mathbf{A}} \rangle = ||\tilde{\mathbf{A}}||_*$.

Let $\tilde{\mathbf{U}}\tilde{\mathbf{D}}\tilde{\mathbf{V}}^T$ give the SVD of $\tilde{\mathbf{A}}$. Note that

$$\langle \mathbf{U}\mathbf{V}^T, \tilde{\mathbf{U}}\tilde{\mathbf{D}}\tilde{\mathbf{V}}^T \rangle = \text{Tr}(\mathbf{V}\mathbf{U}^T\tilde{\mathbf{U}}\tilde{\mathbf{D}}\tilde{\mathbf{V}}^T) = \text{Tr}(\mathbf{V}^T\tilde{\mathbf{V}}\mathbf{U}^T\tilde{\mathbf{U}}\tilde{\mathbf{D}}),$$

and

$$\text{Tr}(\mathbf{V}^T\tilde{\mathbf{V}}\mathbf{U}^T\tilde{\mathbf{U}}\tilde{\mathbf{D}}) = ||\tilde{\mathbf{A}}||_* = \sum_i \tilde{\mathbf{D}}[i,i]$$

if and only if $\mathbf{V}^T\tilde{\mathbf{V}}\mathbf{U}^T\tilde{\mathbf{U}}[i,i] = 1$ where $\tilde{\mathbf{D}}[i,i] > 0$. It follows that the left and right singular vectors of $\tilde{\mathbf{A}}$ that correspond to non-zero singular values must also be singular vectors of $\hat{\mathbf{A}} + \tilde{\mathbf{A}}$. By an analogous argument, the left and right singular vectors that correspond to non-zero singular values in $\hat{\mathbf{A}}$ must also be singular vectors of $\hat{\mathbf{A}} + \tilde{\mathbf{A}}$. □

Proposition 5 is a direct corollary of Lemmas 1 and 2, as Lemma 1 implies $||\hat{\mathbf{S}}_{\cdot\cdot}^{(k)} + \tilde{\mathbf{S}}_{\cdot\cdot}^{(k)}||_* = ||\hat{\mathbf{S}}_{\cdot\cdot}^{(k)}||_* + ||\tilde{\mathbf{S}}_{\cdot\cdot}^{(k)}||_*$ for each k, and then Lemma 2 implies the result.

B.4. Proof of Theorem 1.

PROOF. Take two decomposition $\{\hat{\mathbf{S}}_{\cdot\cdot}^{(k)}\}_{k=1}^K$ and $\{\tilde{\mathbf{S}}_{\cdot\cdot}^{(k)}\}_{k=1}^K$ that satisfy properties 1., 2., and 3. of Theorem 1; we will show that $\{\hat{\mathbf{S}}_{\cdot\cdot}^{(k)}\}_{k=1}^K = \{\tilde{\mathbf{S}}_{\cdot\cdot}^{(k)}\}_{k=1}^K$. For each $k = 1, \ldots, K$, write $\hat{\mathbf{S}}_{\cdot\cdot}^{(k)} = \mathbf{U}_{\cdot}^{(k)}\hat{\mathbf{D}}\mathbf{V}_{\cdot}^{(k)T}$ and $\hat{\mathbf{S}}_{\cdot\cdot}^{(k)} = \mathbf{U}_{\cdot}^{(k)}\tilde{\mathbf{D}}^{(k)}\mathbf{V}_{\cdot}^{(k)T}$ as in Proposition **??**. Then, it suffices to show that $\hat{\mathbf{D}}^{(k)}[r,r] = \tilde{\mathbf{D}}^{(k)}[r,r]$ for all k, r.

First, consider module $k = 1$ with $\mathbf{R}[\bullet, 1] = [1\ 0\ \cdots\ 0]^T$ and $\mathbf{C}[\bullet, 1] = [1\ 0\ \cdots\ 0]^T$. By way of contradiction, assume $\hat{\mathbf{D}}^{(1)}[1,1] > 0$ and $\tilde{\mathbf{D}}^{(1)}[1,1] = 0$. The linear independence of $\{\mathbf{V}_j^{(k)}[\bullet, r] : \hat{\mathbf{D}}^{(k)}[r,r] > 0\}$ and $\{\mathbf{V}_j^{(k)}[\bullet, r] : \tilde{\mathbf{D}}^{(k)}[r,r] > 0\}$ implies that

$$\text{row}(\mathbf{X}_{\cdot\cdot}) = \text{span}\{\mathbf{U}_{\cdot}^{(k)}[\bullet, r] : \hat{\mathbf{D}}^{(k)}[r,r] > 0\} = \text{span}\{\{\mathbf{U}_{\cdot}^{(k)}[\bullet, r] : \tilde{\mathbf{D}}^{(k)}[r,r] > 0\}.$$

Thus, $\mathbf{U}^{(1)}[\bullet,1]] \in \text{span}\{\{\mathbf{U}_{\bullet}^{(k)}[\bullet,r] : \tilde{\mathbf{D}}^{(k)}[r,r] > 0\}$, and it follows from the orthogonality of $\mathbf{U}^{(1)}[\bullet,1]$ and $\{\mathbf{U}^{(1)}[\bullet,r], r > 1\}$ that

$$\mathbf{U}_{\bullet}^{(1)}[\bullet,1] \in \text{span}\{\{\mathbf{U}_{\bullet}^{(k)}[\bullet,r] : \tilde{\mathbf{D}}^{(k)}[r,r] > 0 \text{ and } k > 1\}.$$

Moreover, because $\mathbf{U}_i^{(1)} = \mathbf{0}$ for any $i > 1$ and $\{\mathbf{U}_i^{(k)}[\bullet,r] : \tilde{\mathbf{D}}^{(k)}[r,r] > 0\}$ are linearly independent it follows that

$$\mathbf{U}_{\bullet}^{(1)}[\bullet,1] \in \text{span}\{\mathbf{U}_{\bullet}^{(k)}[\bullet,r] : \tilde{\mathbf{D}}^{(k)}[r,r] > 0,\ k > 1,\ \text{ and } \mathbf{R}[i,k] = 0 \text{ for any } i > 1\}. \tag{6}$$

Note that (6) implies $\mathbf{U}_1^{(1)}[\bullet,1] \in \text{row}(\mathbf{X}_{12} + \cdots + \text{row}(\mathbf{X}_{1J})$, however, this is contradicted by the linear independence of $\mathbf{U}_1^{(1)}[\bullet,1]$ and $\{\mathbf{U}_i^{(k)}[\bullet,r] : \hat{\mathbf{D}}^{(k)}[r,r] > 0, k > 1\}$. Thus, we conclude that $\tilde{\mathbf{D}}^{(1)}[1,1] > 0$ implies $\tilde{\mathbf{D}}^{(1)}[1,1] > 0$. Analogous arguments show that $\tilde{\mathbf{D}}^{(k)}[r,r] > 0$ if and only if $\tilde{\mathbf{D}}^{(k)}[r,r] > 0$ for any pair (r,k). It follows that $\{\mathbf{U}_i^{(k)}[\bullet,r] : \hat{\mathbf{D}}^{(k)}[r,r] > 0 \text{ or } \tilde{\mathbf{D}}^{(k)}[r,r] > 0\}$ are linearly independent for $i = 1,\ldots I$, and $\{\mathbf{V}_j^{(k)}[\bullet,r] : \hat{\mathbf{D}}^{(k)}[r,r] > 0 \text{ or } \tilde{\mathbf{D}}^{(k)}[r,r] > 0\}$ are linearly independent for $j = 1,\ldots,J$. Thus,

$$\sum_{k=1}^{K} \mathbf{U}_{\bullet}^{(k)}(\hat{\mathbf{D}}^{(k)} - \tilde{\mathbf{D}}^{(k)})\mathbf{V}_{\bullet}^{(k)T} = \sum_{k=1}^{K} \hat{\mathbf{S}}_{\bullet\bullet}^{(k)} - \tilde{\mathbf{S}}_{\bullet\bullet}^{(k)} = \mathbf{X}_{\bullet\bullet} - \mathbf{X}_{\bullet\bullet} = \mathbf{0}$$

implies that $\hat{\mathbf{D}}^{(k)}[r,r] = \tilde{\mathbf{D}}^{(k)}[r,r]$ for all k,r.

□

REFERENCES

Fazel, M., Hindi, H., Boyd, S. P., et al. (2001). A rank minimization heuristic with application to minimum order system approximation. In *Proceedings of the American control conference*, volume 6, pages 4734–4739. Citeseer.