A Rank-Corrected Procedure for Matrix Completion with Fixed Basis Coefficients

Defeng Sun

Department of Mathematics and Risk Management Institute
National University of Singapore

Based on a joint work with Weimin Miao and Shaohua Pan

Correlation matrix completion

A correlation matrix $\overline{X}$:

symmetric positive semidefinite, diag$(\overline{X}) = e$.

<u>Question</u>: How to recover $\overline{X}$ from observations of entries

$$\underbrace{\overline{X}_{(a_1,b_1)}, \ \overline{X}_{(a_2,b_2)}, \ \ldots, \ \overline{X}_{(a_m,b_m)}}_{m \text{ upper off-diagonal entries}}, \qquad a_k < b_k.$$

(Multiple observations of one entry are allowed.)

Density matrix completion

A density matrix $\overline{X}$ of a quantum system:

Hermitian positive semidefinite matrix, Tr$(\overline{X}) = 1$.

Observations: Pauli measurements, i.e., Re$($Tr$(\Theta_i \overline{X}))$,

$$\Theta_i \in \text{ Pauli basis } : \left\{\sigma_{s_1} \otimes \cdots \otimes \sigma_{s_l} \mid (s_1, \cdots, s_l) \in \{0, 1, 2, 3\}^k\right\},$$

where

$$\sigma_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \ \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \ \sigma_2 = \begin{pmatrix} 0 & -\sqrt{-1} \\ \sqrt{-1} & 0 \end{pmatrix}, \ \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

How to recover the unknown matrix $\overline{X}$ in quantum state tomograph?[1]

[1] D. Gross, Y.K. Liu, S.T. Flammia, S. Becker, and J. Eisert. Quantum state tomography via compressed sensing. *Physical review letters*, 105 (2010).

Matrix completion with fixed basis coefficients

Correlation matrix completion (an example):

- ◇ all diagonal entries are fixed, i.e.,$\langle e_i e_i^T, \overline{X}\rangle = 1, i = 1, \ldots, n;$
- ◇ some off-diagonal entries may be fixed as well, [e.g., the correlations among pegged currencies.]

Consider the orthonormal basis $(d = d_1 + d_2 = n(n+1)/2)$:

$$\{\Theta_i\}_{i=1}^d := \underbrace{\{e_i e_i^T \mid 1 \le i \le n\}}_{\Theta_\alpha \text{(fixed)}} \bigcup \underbrace{\left\{\frac{1}{\sqrt{2}}(e_i e_j^T + e_j e_i^T) \,\Big|\, 1 \le i < j \le n\right\}}_{\Theta_\beta}.$$

$$\alpha = \{1, \ldots, d_1\} \qquad \beta = \{d_1 + 1, \ldots, d_1 + d_2\}$$

The observation model

- Fixed basis coefficients: $\langle \Theta_k, \overline{X} \rangle$, $k \in \alpha$.
- The observation model:

$$y_i = \langle \Theta_{\omega_i}, \overline{X} \rangle + \nu \xi_i, \quad \omega_i \in \beta, i = 1, \cdots, m,$$

where ξ_i are i.i.d noises with mean 0 and variance 1.

  - The sampling probability:

$$p = (\underbrace{0, \cdots, 0}_{\alpha}, \underbrace{p_{d_1+1}, \cdots p_{d_1+d_2}}_{\beta})^T, \qquad p_k > 0 \; \forall \, k \in \beta.$$

  Uniform sampling: $p_k = 1/d_2 \; \forall \, k \in \beta$.

A problem formulation

- $\mathcal{R}_\alpha(X) = \big(\langle \Theta_k, X\rangle\big)^T_{k\in\alpha} \in \mathbb{R}^{d_1}$.
- $\mathcal{R}_\beta(X) = \big(\langle \Theta_k, X\rangle\big)^T_{k\in\beta} \in \mathbb{R}^{d_2}$.
- $\mathcal{R}_\Omega(X) := (\langle \Theta_{\omega_1}, X\rangle, \cdots, \langle \Theta_{\omega_m}, X\rangle)^T \in \mathbb{R}^m$.

Suppose that $\overline{X}$ has a low-rank structure. One may recover $\overline{X}$ by solving the following problem

$$\min_{X\in\mathcal{S}^n} \ \frac{1}{2m}\|y - \mathcal{R}_\Omega(X)\|_2^2 + \rho_m \text{rank}(X)$$
$$\text{s.t.} \ \ \mathcal{R}_\alpha(X) = \mathcal{R}_\alpha(\overline{X}), \quad X \in \mathcal{S}^n_+.$$

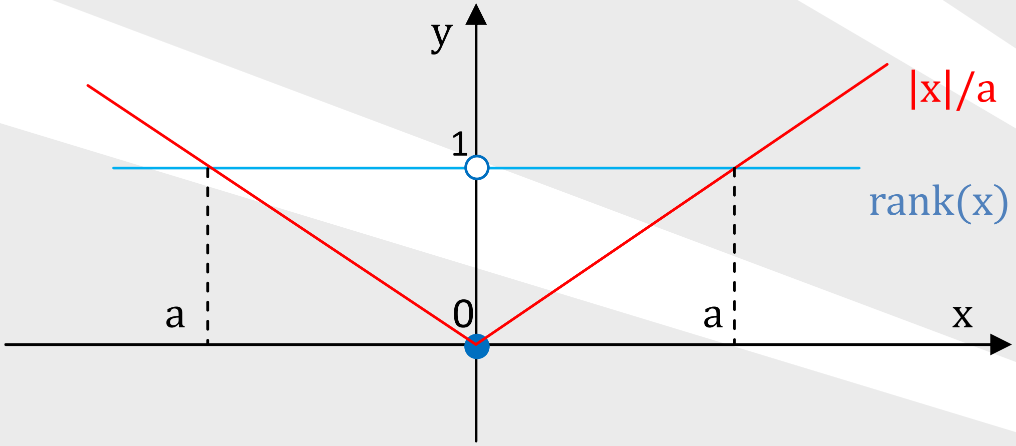
- NP-hard.
- This model is also applicable to covariance matrix completion.

Nuclear norm always fails.

A popular convex relaxation to encourage low-rank solutions:

$$\text{rank}(X) \quad \Longrightarrow \quad \|X\|_* := \sum_{i=1}^{n} \sigma_i(X).$$

◇ Nuclear norm — convex envelope of the rank function over the unit ball of the spectral norm.



◇ Correlation matrices: $\|X\|_* = $ constant $\Longrightarrow$ Nuclear norm fails!

The rank constrained problem

A majorized penalty approach proposed by Gao and Sun[2]:

$$\min_{X\in\mathcal{C}} \left\{ h(X) : \ \text{rank}(X) \le r \right\}$$

$$\Downarrow \qquad \left[\text{rank}(X) \le r \iff \sigma_{r+1}(X) + \cdots + \sigma_n(X) = 0\right]$$

$$\min_{X\in\mathcal{C}} \ h(X) + \rho\big(\|X\|_* - s_r(X)\big)$$

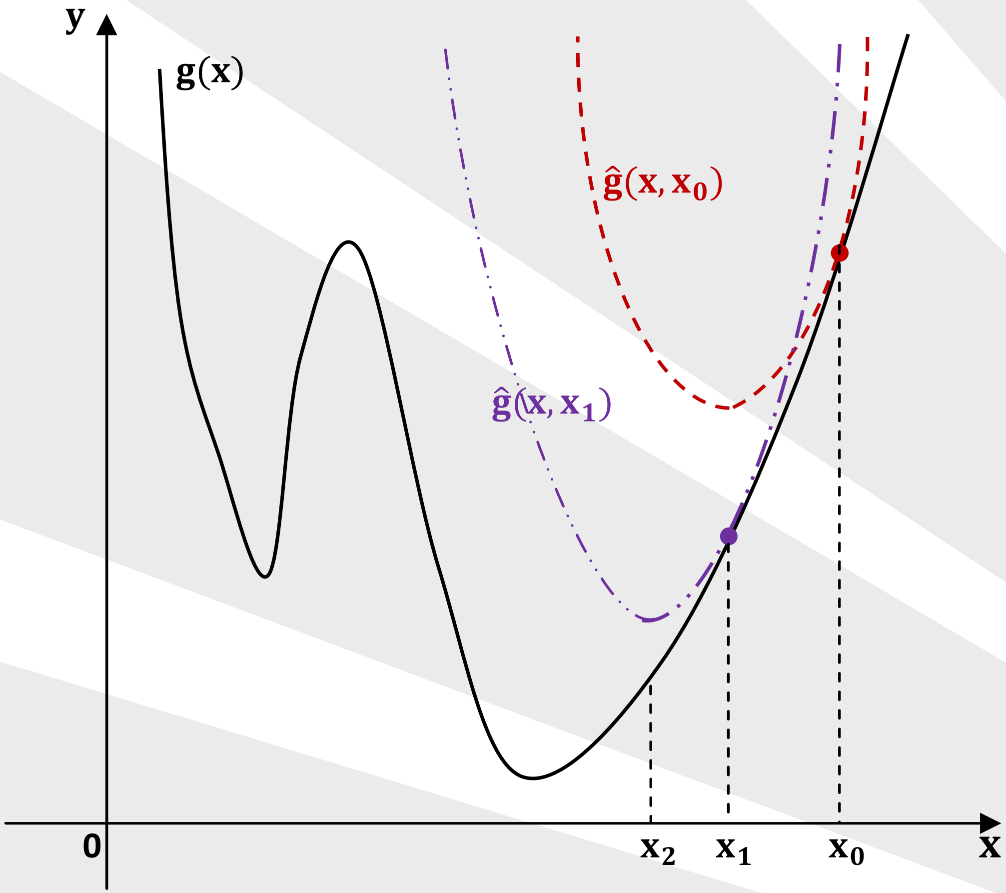
$$\Downarrow \qquad \left[s_r(X) \ge s_r(Y) + \langle G^Y, X - Y\rangle, G^Y \in \partial s_r(Y)\right]$$

$$X^{k+1} = \arg\min_{X\in\mathcal{C}} \ \widehat{h}^k(X) + \rho\left(\|X\|_* - \langle G^k, X\rangle + \frac{\gamma_k}{2}\|X - X^k\|_F^2 \right)$$

where $s_r(X) := \sum_{i=1}^r \sigma_i(X)$, $G^k \in \partial s_r(X^k)$ and $\widehat{h}^k$ is a majorized convex function to h at X^k.

[2] Gao, Y. and Sun, D., A majorized penalty approach for calibrating rank constrained correlation matrix problems, 2010.

The majorization method



A majorization function $\widehat{g}(x, x^k)$ of g at x^k satisfies

$$\widehat{g}(x^k, x^k) = g(x^k) \quad \text{and} \quad \widehat{g}(x, x^k) \geq g(x) \quad \forall\, x.$$

Our proposed rank-correction step

Our proposed rank-correction step:

$$\min_{X\in\mathcal{S}^n} \ \frac{1}{2m}\|y-\mathcal{R}_\Omega(X)\|_2^2+\rho_m\Big(\langle I_n - F(\widetilde{X}_m), X\rangle+\frac{\gamma_m}{2}\|X-\widetilde{X}_m\|_F^2\Big)$$

$$\text{s.t.}\quad \mathcal{R}_\alpha(X)=\mathcal{R}_\alpha(\overline{X}),\qquad X\in\mathcal{S}_+^n,$$

- ◇ F: a spectral operator.
- ◇ $\widetilde{X}_m$: a reasonable initial estimator, say the (nuclear norm regularized) least squares estimator.
- ◇ $\gamma_m \geq 0$: ensure the boundness of the optimal solution.

The spectral operator

A function $f:\mathbb{R}^n\to\mathbb{R}^n$ is said to be symmetric if

$$f(x)=Q^{\mathbb{T}}f(Qx)\quad \forall \text{ signed permutation matrix } Q \text{ and } x\in\mathbb{R}^n.$$

- ◇ $f_i(x)=0$ if $x_i=0$.

The spectral operator[3] $F:\mathbb{R}^{n_1\times n_2}\to\mathbb{R}^{n_1\times n_2}$ associated with the symmetric function $f:\mathbb{R}^{\min(n_1,n_2)}\to\mathbb{R}^{\min(n_1,n_2)}$ is given by

$$F(X):=U\mathrm{Diag}\big(f(\sigma(X))\big)V^T,\quad (U,V)\in\mathcal{O}^{n_1,n_2}(X)^4, X\in\mathbb{R}^{n_1\times n_2}.$$

- ◇ For any $X\in\mathbb{S}^n_+$, $F(X)=P\mathrm{Diag}\big(f(\lambda(X))\big)P^T,\ P\in\mathcal{O}^n(X)^5$.

[3] Ding, C, Ph.D. thesis, National University of Singapore, (2012).
[4] For any $X\in\mathbb{R}^{n_1\times n_2}$, $\mathcal{O}^{n_1,n_2}(X):=\big\{(U,V)\in\mathcal{O}^{n_1}\times\mathcal{O}^{n_2}\mid X=U\mathrm{Diag}\big(\sigma(X)\big)V^T\big\}$.
[5] For any $X\in\mathcal{S}^n$, $\mathcal{O}^n(X):=\big\{P\in\mathcal{O}^n\mid X=P\mathrm{Diag}(\lambda(X))P^T\big\}$.

Orthogonal decomposition

Suppose that $\mathrm{rank}(\overline{X}) = r$. Decompose $\mathcal{S}^n = T \oplus T^\perp$ with

$$T := \left\{ X \in \mathcal{S}^n \mid X = Y + Y^T \text{ with } \mathrm{row}(Y) \subseteq \mathrm{row}(\overline{X}) \right\},$$
$$T^\perp := \left\{ X \in \mathcal{S}^n \mid \mathrm{row}(X) \perp \mathrm{row}(\overline{X}) \right\},$$

where $\mathrm{row}(X)$ denotes the row space of X.

Then, we have that for any $X \in \mathcal{S}^n$,

$$\mathcal{P}_T(X) = \overline{P}_1\overline{P}_1^T X + X\overline{P}_1\overline{P}_1^T - \overline{P}_1\overline{P}_1^T X\overline{P}_1\overline{P}_1^T,$$
$$\mathcal{P}_{T^\perp}(X) = \overline{P}_2\overline{P}_2^T X\overline{P}_2\overline{P}_2^T,$$

where $[\overline{P}_1 \ \ \overline{P}_2] \in \mathcal{O}^n(\overline{X})$ with $\overline{P}_1 \in \mathbb{R}^{n\times r}$ and $\overline{P}_2 \in \mathbb{R}^{n\times(n-r)}$.

Error bounds

Let $\widehat{X}_m$ be the estimator generated from the rank-correction step.

For simplicity, we set $\gamma_m := 0$. Let

$$a_m := \|\overline{P}_1\overline{P}_1^T - \mathcal{P}_T(F(\widetilde{X}_m))\| \quad \text{and} \quad b_m := 1 - \|\mathcal{P}_{T^\perp}(F(\widetilde{X}_m))\|.$$

Assume $b_m > 0$. For any constant $\kappa > 1$, if

$$\rho_m \geq \frac{\kappa\nu}{b_m}\left\|\frac{1}{m}\mathcal{R}_\Omega^*(\xi)\right\|$$

then, from the optimality of $\widehat{X}_m$, we have

$$\frac{1}{2m}\|\mathcal{R}_\Omega(\widehat{X}_m - \overline{X})\|_2^2 \leq \sqrt{2r}\left(a_m + \frac{b_m}{\kappa}\right)\rho_m\|\widehat{X}_m - \overline{X}\|_F.$$

Error bounds (Cont.)

- The sampling operator $\mathcal{R}_\Omega$ does not satisfy the restricted isometric property (RIP).
- However, $\mathcal{R}_\Omega$ has a similar property with high probability under certain conditions, such that

$$\frac{1}{2m}\|\mathcal{R}_\Omega(\widehat{X}_m - \overline{X})\|_2^2 \geq C\|\widehat{X}_m - \overline{X}\|_F^2 - \text{a small term}$$

for some constant C.

Error bounds (Cont.)

We adopt the setting of Klopp (2012)[6] and correspondingly modify it. Assume that

- $\overline{X}$ is bounded in terms of $\|\mathcal{R}_\beta(\overline{X})\|_\infty \leq c$ for some constant c.
- ξ_i are subexponential [7] with mean 0 and variance 1.

Let $\widehat{X}^c_m$ be generated from the rank-correction step with an additional constraint $\|\mathcal{R}_\beta(X)\|_\infty \leq c$ to the optimization problem.

- $\widehat{X}^c_m = \widehat{X}_m$ if the bound c is not tight.

---

[6] Klopp, O., Noisy low-rank matrix completion with general sampling distribution, (2012).

[7] ξ_i is said to be subexponential, i.e., there exists some $C, c, a > 0$ such that for all $t > 0$, $\mathbb{P}(|\xi_i| \geq t) \leq C\exp(-ct^\alpha)$.

Error bounds (Cont.)

Theorem 1. *For any given* $\kappa > 1$, *choose* ρ_m *by*

$$\rho_m = \frac{\kappa\nu}{\beta_m} C^* \sqrt{\frac{\mu_2 \log(n)}{mn}}.$$

Then, $\exists$ *a constant* C *s.t. with probability at least* $1 - 1.5/n$,

$$\frac{\|\widehat{X}_m^c - \overline{X}\|_F^2}{d_2} \leq C \max\left\{ \eta_m \mu_1^2 \mu_2 \frac{d_2 r \log(n)}{mn}, c^2 \mu_1 \sqrt{\frac{\log(n)}{m}} \right\},$$

where $$\eta_m := \left(\left(1 + \kappa \frac{a_m}{b_m}\right)^2 \nu^2 + \left(\frac{\kappa}{\kappa - 1}\right)^2 \left(1 + \frac{a_m}{b_m}\right)^2 c^2 \right).$$

- μ_1 and μ_2 are constants, irrelevant to n and d_2.
- The sample size to control the error bound is $O(nr\log(n)) \approx$ the order of degree of freedom, since $d_2 \leq n^2$.

The power of the correction term

- The (nuclear norm penalized) least squares estimator:

$$F \equiv 0 \quad \Longrightarrow \quad a_m = b_m = \frac{a_m}{b_m} = 1.$$

- The rank-correction step:

$$\frac{a_m}{b_m} \le \frac{\varepsilon_1}{1-\varepsilon_2} \quad \text{if} \quad \begin{cases} \left\| \mathcal{P}_T\big(F(\widetilde{X}_m)\big) - \overline{P}_1 \overline{P}_1^T \right\| \le \varepsilon_1, \\ \left\| \mathcal{P}_{T^\perp}\big(F(\widetilde{X}_m)\big) \right\| \le \varepsilon_2 < 1. \end{cases}$$

- If we have a reasonable $\widetilde{X}_m$, why not use it as a correction?
- We should construct a spectral operator F such that

$$F(\widetilde{X}_m) \quad \text{is close to} \quad \overline{P}_1 \overline{P}_1^T.$$

Rank consistency

Not only potentially reduce the recovery error, but also the rank!

Definition 1 (Bach, 2008)**.** *An estimator X_m of the true matrix $\overline{X}$ is said to be rank consistent if*

$$\lim_{m\to\infty} \mathbb{P}\big(\mathrm{rank}(X_m) = \mathrm{rank}(\overline{X})\big) = 1.$$

Assumption 1.

- *The spectral operator F is continuous at $\overline{X}$.*
- *The initial estimator $\widetilde{X}_m$ satisfies $\widetilde{X}_m \xrightarrow{p} \overline{X}$ as $m \to \infty$.*[8]

[8] "The notation $\xrightarrow{p}$" means convergence in probability.

A local necessary condition for rank consistency

Let $\mathcal{Q}_\beta(X) := \sum_{k\in\beta} p_k\langle\Theta_k, X\rangle\Theta_k$ and $\mathcal{Q}_\beta^\dagger(X) := \sum_{k\in\beta} \frac{1}{p_k}\langle\Theta_k, X\rangle\Theta_k$.

Proposition 1. *If* $\rho_m \to 0, \sqrt{m}\rho_m \to \infty$ *and* $\gamma_m = O_p(1)$, *then* $\rho_m^{-1}(\widehat{X}_m - \overline{X}) \xrightarrow{p} \widehat{\Delta}$, *where* $\widehat{\Delta}$ *is the unique optimal solution to*

$$\begin{array}{cl} \min\limits_{\Delta\in\mathcal{S}^n} & \frac{1}{2}\langle\mathcal{Q}_\beta(\Delta), \Delta\rangle + \langle I_n - F(\overline{X}), \Delta\rangle \\ \text{s.t.} & \mathcal{R}_\alpha(\Delta) = 0, \quad \overline{P}_2^{\mathbb{T}}\Delta\overline{P}_2 \in \mathcal{S}_+^{n-r}. \end{array}$$

A local necessary condition for rank consistency:

$$\overline{P}_2^T\widehat{\Delta}\overline{P}_2 = 0.$$

A sufficient condition for rank consistency

Assume that the Slater condition holds. Consider the linear system:

$$\overline{P}_2^T \mathcal{Q}_\beta^\dagger(\overline{P}_2 \Lambda \overline{P}_2^T)\overline{P}_2 = \overline{P}_2^T \mathcal{Q}_\beta^\dagger\big(I_n - F(\overline{X})\big)\overline{P}_2. \qquad (1)$$

Theorem 2. *If $\rho_m \to 0$, $m^{1/2}\rho_m \to \infty$ and $\gamma_m = O_p(1)$, then for the rank consistency of $\widehat{X}_m$,*

- *a necessary condition: (1) has a solution $\widehat{\Lambda} \in \mathcal{S}_+^{n-r}$.*
- *a sufficient condition: (1) has a unique solution $\widehat{\Lambda} \in \mathcal{S}_{++}^{n-r}$.*

The linear system can be written concisely as

$$\mathcal{B}_2(\Lambda) = \mathcal{B}_2(I_{n-r}) + \mathcal{B}_1\big(\mathrm{Diag}(\widehat{g}(\overline{X}))\big), \quad \Lambda \in \mathcal{S}^{n-r},$$

where $\mathcal{B}_1(Y) := \overline{P}_2^T \mathcal{Q}_\beta^\dagger(\overline{P}_1 Y \overline{P}_1^T)\overline{P}_2$, $\mathcal{B}_2(Z) := \overline{P}_2^T \mathcal{Q}_\beta^\dagger(\overline{P}_2 Z \overline{P}_2^T)\overline{V}_2$ and $\widehat{g}(\overline{X}) := \big(1 - f_1(\sigma(\overline{X})), \ldots, 1 - f_r(\sigma(\overline{X}))\big)^T$.

Constraint nondegeneracy

We say that the constraint nondegeneracy at $\overline{X}$ if

$$\mathcal{R}_\alpha\big(\text{lin}(\mathcal{T}_{\mathcal{S}^n_+}(\overline{X}))\big) = \mathbb{R}^{d_1},$$

where

$$\text{lin}(\mathcal{T}_{\mathcal{S}^n_+}(\overline{X})) = \big\{X \in \mathcal{S}^n \mid \overline{P}_2^T X \overline{P}_2 = 0\big\}.$$

If constraint nondegeneracy holds at $\overline{X}$, then

- ◇ the linear operators $\mathcal{B}_2$ is self-adjoint and positive definite;
- ◇ the sufficient condition for rank consistency reduces to

$$I_{n-r} + \mathcal{B}_2^{-1}\mathcal{B}_1\big(\text{Diag}(g_r(\overline{X}))\big) \in \mathcal{S}^{n-r}_{++}.$$

In general, $F(\overline{X})$ is desired to be close to $\overline{P}_1^T\overline{P}_1$ for rank consistency.

Rank consistency for correlation matrix compl.

Theorem 3. *For the correlation matrix completion problems with all diagonal entries being fixed as ones under uniform sampling, if $\rho_m \to 0$, $\sqrt{m}\rho_m \to \infty$, $\gamma_m = O_p(1)$ and F is a spectral operator associated with a symmetric function $f : \mathbb{R}^n \to \mathbb{R}^n$ such that for $i = 1, \ldots, n$,*

$$f_i(x) \geq 0 \ \forall\, x \in \mathbb{R}^n_+ \quad \text{and} \quad f_i(x) = 0 \ \text{if and only if} \ x_i = 0, \qquad (2)$$

then the estimator $\widehat{X}_m$ generated from the rank-correction step is rank consistent.

- ◇ Constraint nondegeneracy holds at $\overline{X}$ automatically.
- ◇ This theorem is also applicable to covariance matrix completion with partial positive diagonal entries being fixed.

The construction of F

The results of recovery error and rank consistency suggest a consistent criterion for the construction of the rank-correction function F, if possible, such that

$$F(X) \to \overline{P}_1 \overline{P}_1^T \quad \text{as} \quad X \to \overline{X}.$$

<u>When the true rank is known:</u>

$$F(X) := P_1 P_1^T, \tag{3}$$

where $([P_1, P_2]) \in \mathcal{O}^n(X)$, $X \in \mathcal{S}^n$ with $P_1 \in \mathbb{R}^{n\times r}$, $P_2 \in \mathbb{R}^{n\times(n-r)}$.

◇ The rank-correction step reduces to one step of the majorized penalty approach proposed by Gao and Sun (2010).

The construction of F (Cont.)

When the true rank is unknown:

$$F(X) := P\mathrm{Diag}\big(f(\sigma(X))\big)P^T,$$

associated with the symmetric function $f : \mathbb{R}^n \to \mathbb{R}^n$ defined as

$$f_i(x) = \begin{cases} \phi\left(\dfrac{x_i}{\|x\|_\infty}\right) & \text{if } x \in \mathbb{R}^n \backslash \{0\}, \\ 0 & \text{if } x = 0, \end{cases}$$

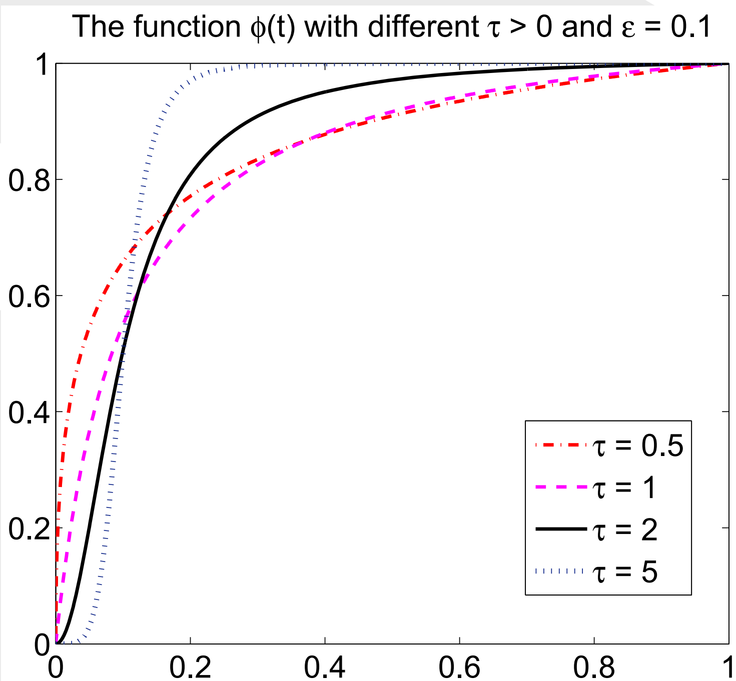
where $P \in \mathcal{O}^n(X)$ and $\phi : \mathbb{R} \to \mathbb{R}$ takes the form

$$\phi(t) := \mathrm{sgn}(t)(1+\varepsilon^\tau)\frac{|t|^\tau}{|t|^\tau + \varepsilon^\tau}, \quad \forall\, t \in \mathbb{R},$$

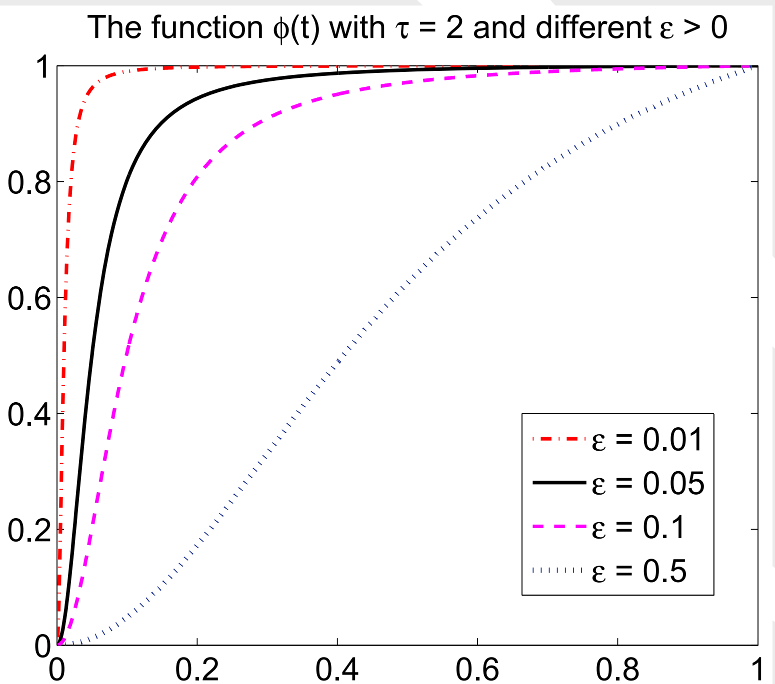
for some $\tau > 0$ and $\varepsilon > 0$.

The construction of F (Cont.)

- Rank consistency: $\varepsilon \downarrow 0$.
- $\phi(t)$ over $[0, 1]$ is concave if $0 < \tau \leq 1$ and S-shaped if $\tau > 1$.



(a) $\varepsilon = 0.1$

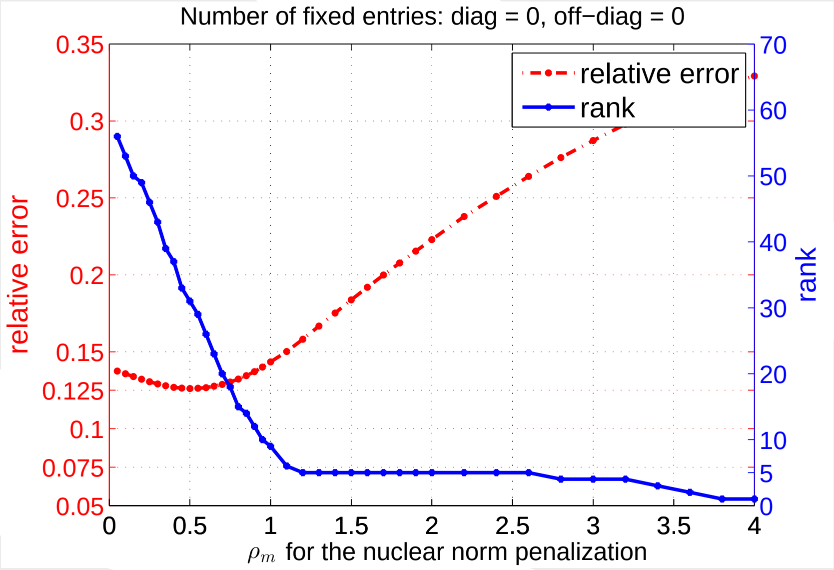


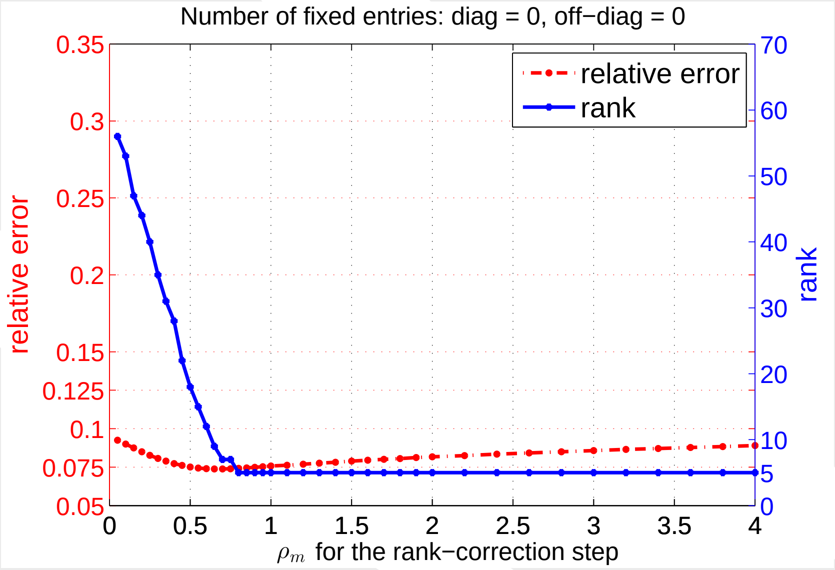
(b) $\tau = 2$

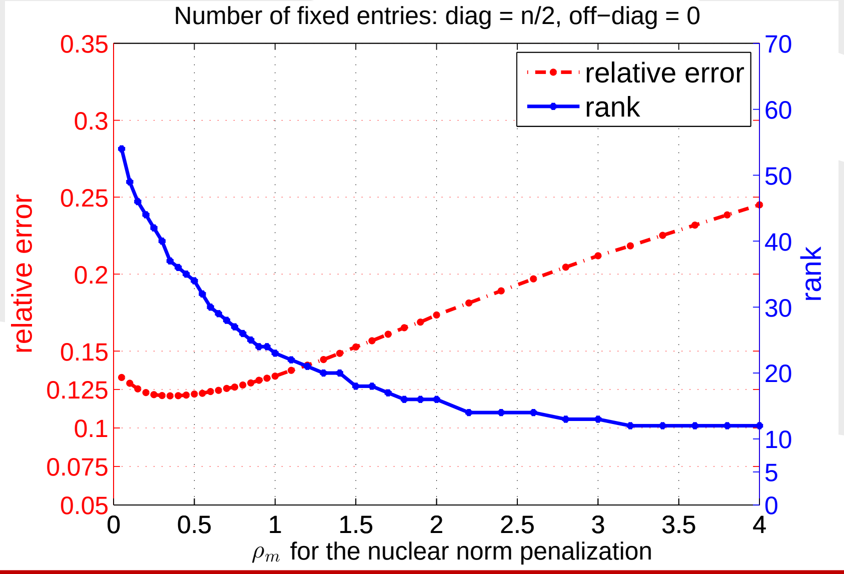
We recommend the choices $\tau = 1, 2$ and $\varepsilon = 0.01 \sim 0.1$ by considering the optimality and robustness of recovery.

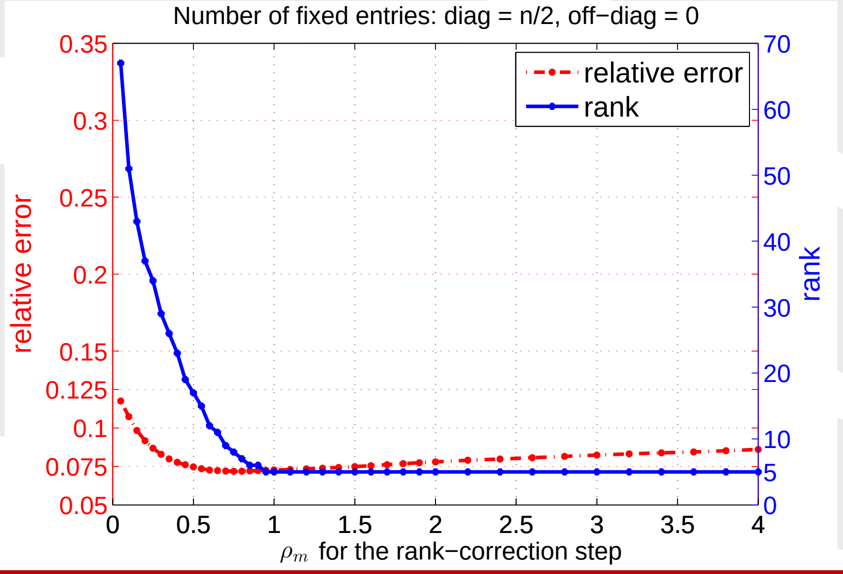
Influence of fixed basis coefficients

$n = 1000$, $\mathrm{rank} = 5$, noise level $= 10\%$, sample ratio $= 6.38\%$, $\tau = 2$, $\varepsilon = 0.02$.

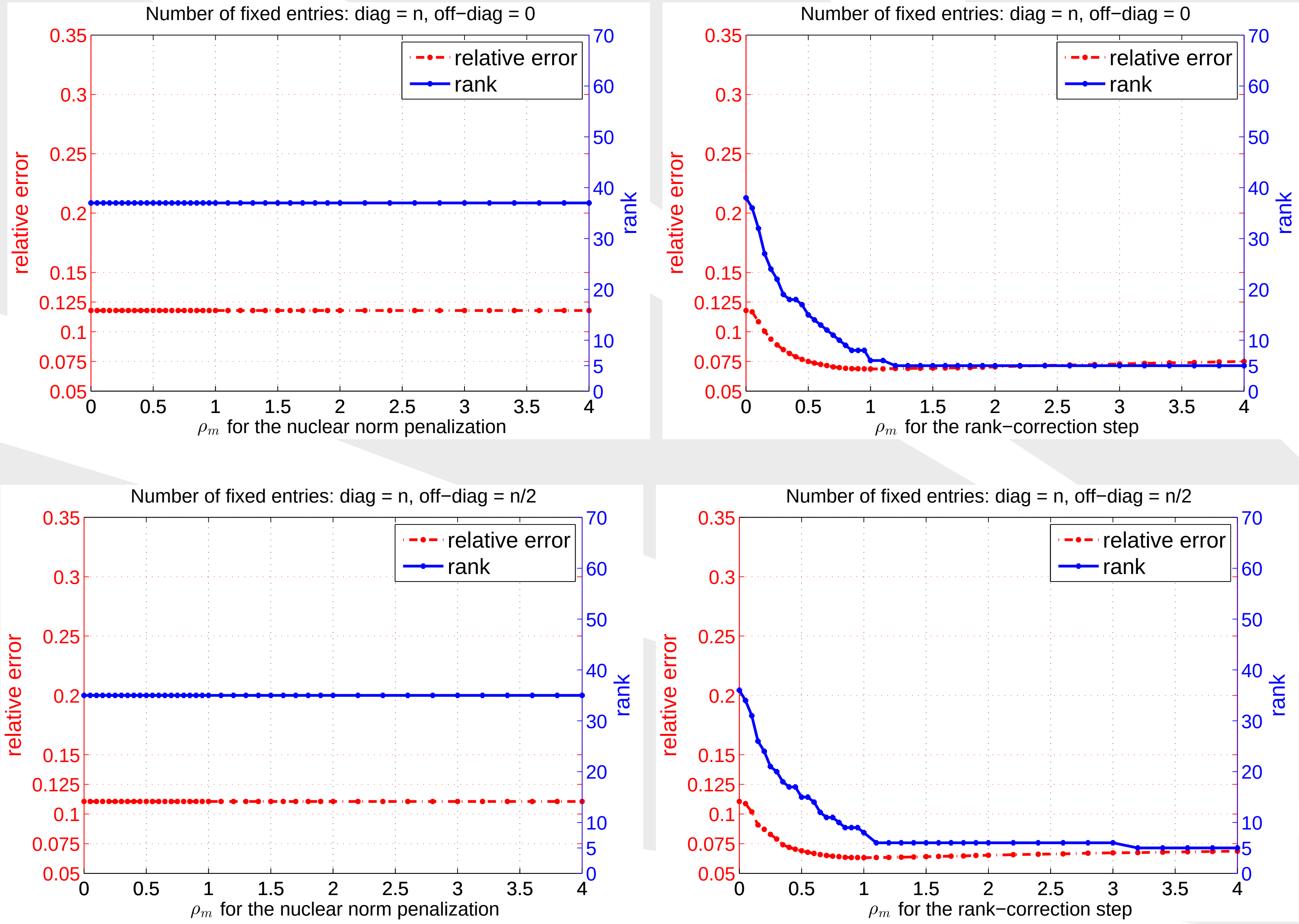








Influence of fixed basis coefficients (Cont.)



Performance of different F

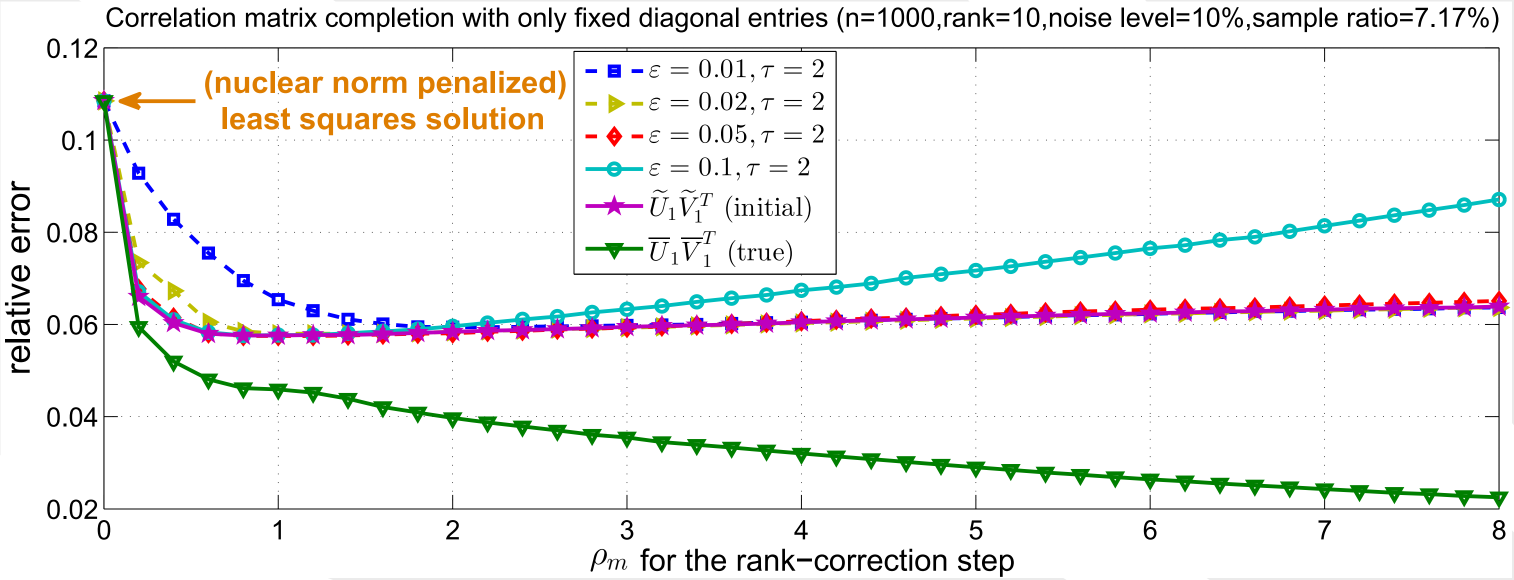
Correlation matrix completion with only diagonal entries being fixed:

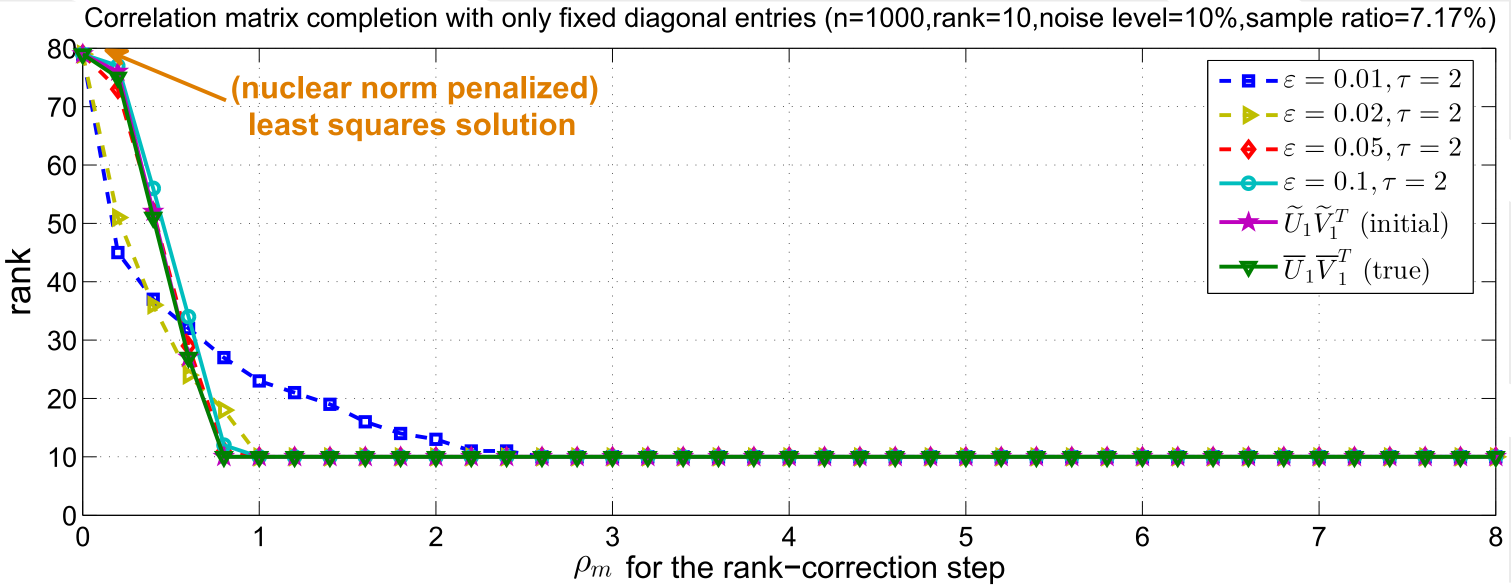
$n = 1000$, $\text{rank} = 10$, noise level $= 10\%$, sample ratio $= 7.17\%$.

Initial estimator: (nuclear norm penalized) least squares estimator.

rank-correction function	a_m	b_m	a_m/b_m	optimal relerr
zero function	1	1	1	10.85%
$\varepsilon = 0.01, \tau = 2$	0.1420	0.2351	0.6038	5.96%
$\varepsilon = 0.02, \tau = 2$	0.1459	0.5514	0.2646	5.80%
$\varepsilon = 0.05, \tau = 2$	0.1648	0.8846	0.1863	5.75%
$\varepsilon = 0.1, \tau = 2$	0.2399	0.9681	0.2478	5.77%
$\widetilde{U}_1\widetilde{V}_1^{\mathbb{T}}$ (initial)	0.1445	0.9815	0.1472	5.75%
$\overline{U}_1\overline{V}_1^{\mathbb{T}}$ (true)	0	1	0	2.25%

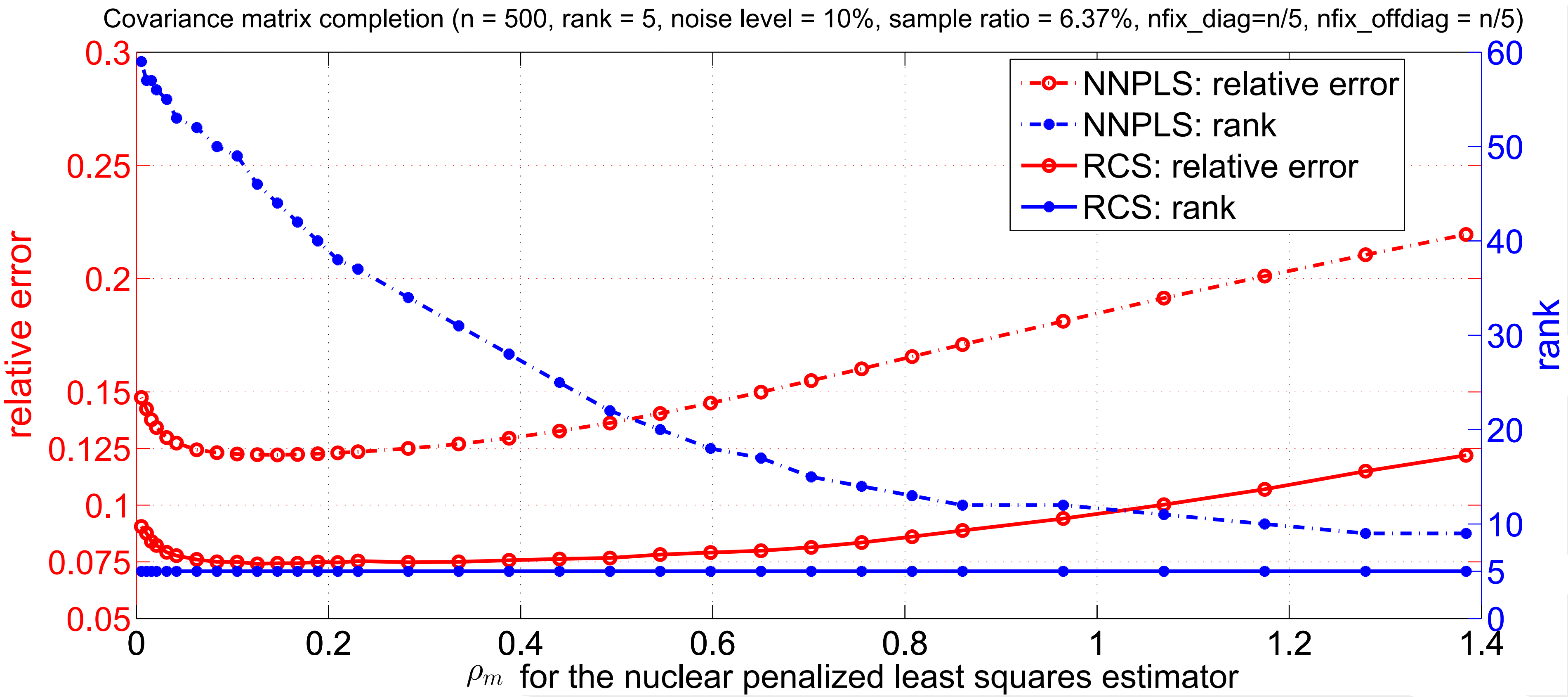
Performance of different F (Cont.)





Covariance matrix completion

$n = 500$, $\mathrm{rank} = 5$, noise level $= 10\%$, sample ratio $= 6.37\%$, $\tau = 2$, $\varepsilon = 0.02$.
number of fixed diagonal entries $= n/5$, number of fixed off-diagonal entries $= n/5$,



Correlation / covariance matrix completion

r	diag/ off-diag	sample ratio	NNPLS	1st RCS	2st RCS	3rd RCS
			relerr (rank)	relerr(rank)	relerr (rank)	relerr (rank)
5	1000/0	2.40%	1.95e-1 (47)	1.27e-1 (5)	1.18e-1 (5)	1.12e-1 (5)
	1000/0	7.99%	6.10e-2 (51)	3.41e-2 (5)	3.37e-2 (5)	3.36e-2 (5)
	500/50	2.39%	2.01e-1 (45)	1.10e-1 (5)	9.47e-2 (5)	8.97e-2 (5)
	500/50	7.98%	7.19e-2 (32)	3.77e-2 (5)	3.59e-2 (5)	3.58e-2 (5)
10	1000/0	5.38%	1.32e-1 (74)	7.68e-2 (10)	7.39e-2 (10)	7.36e-2 (10)
	1000/0	8.96%	9.18e-2 (78)	5.15e-2 (10)	5.08e-2 (10)	5.08e-2 (10)
	500/100	5.37%	1.58e-1 (57)	8.66e-2 (10)	7.74e-2 (10)	7.60e-2 (10)
	500/100	8.96%	1.02e-1 (49)	5.36e-2 (10)	5.24e-2 (10)	5.25e-2 (10)

- $n = 1000$.
- The algorithm is based on an inexact APG method by Jiang, Sun and Toh (2012).[9]

[9] Jiang, K., Sun, D., and Toh, K.C., An inexact accelerated proximal gradient method for large scale linearly constrained convex SDP, *SIAM Journal on Optimization 22*, 22 (2012).

Conclusions

- Our propose rank-correction procedure is also applicable to the general low-rank matrix completion problems.
- For additional linear constraints, all the theoretical results hold with slight modifications.
- This approach can substantially overcome the limitation of the nuclear norm penalization for recovering a low-rank matrix.
- This approach can significantly improve the recovery performance in the sense of both the recovery error and the rank.
- It would be of great interest to extend the asymptotic rank consistency results to the case that the matrix size is allowed to grow.