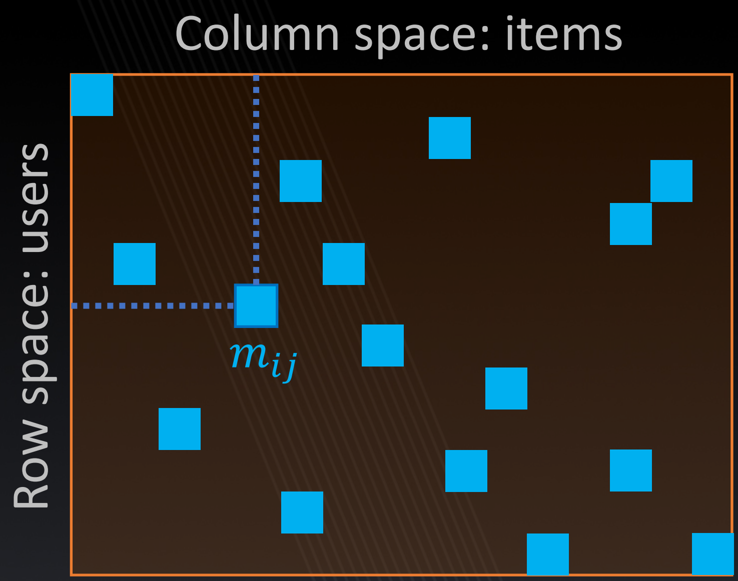
Spectral Geometric Matrix Completion

Amit Boyarski, CS@Technion

Joint work with Sanketh Vedula & Alex Bronstein

MSML 2021

Matrix completion



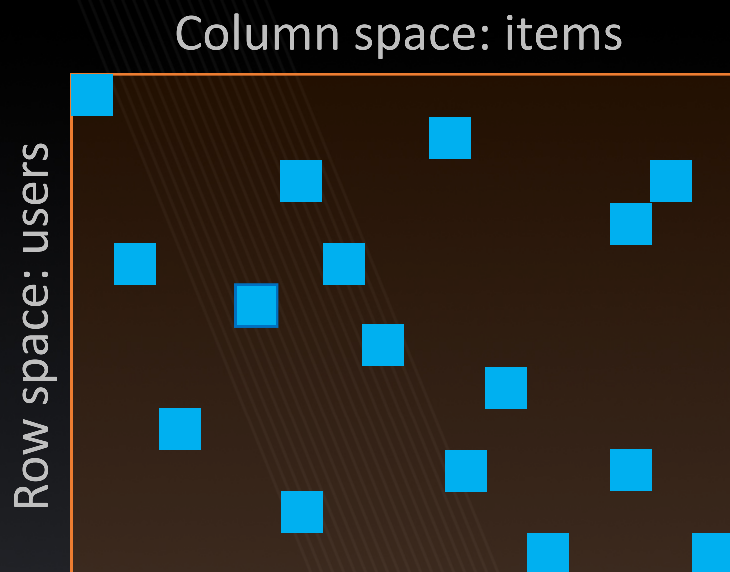
Incomplete measurements

$$m_{ij} : (i,j) \in \mathcal{S}$$

Goal: estimate full matrix $\mathbf{X}$ such that

$$\mathbf{X} \odot \mathbf{S} = \mathbf{M} \odot \mathbf{S}$$

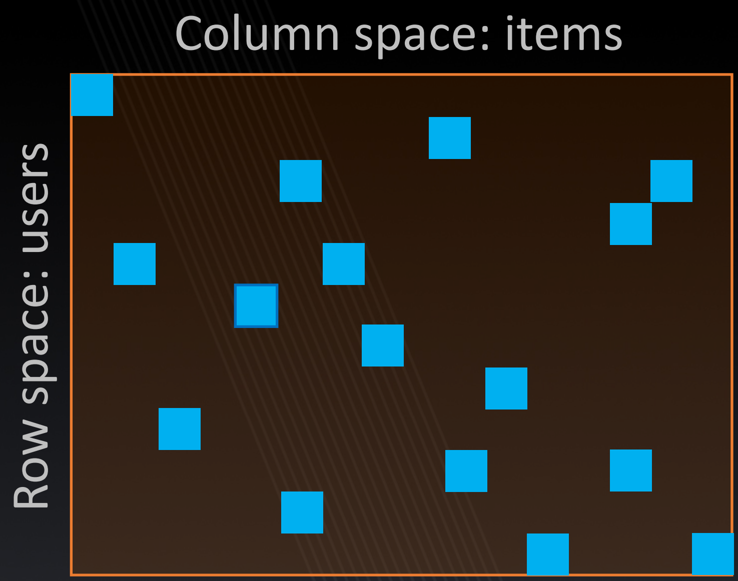
Matrix completion



$$\min_{\mathbf{X}} \|(\mathbf{X}-\mathbf{M})\odot\mathbf{S}\|_F^2 + \mathrm{prior}(\mathbf{X})$$

Problem is under-determined!

Rank-regularized matrix completion



$$\min_{\mathbf{X}} \|(\mathbf{X} - \mathbf{M}) \odot \mathbf{S}\|_F^2 + \mathrm{rank}(\mathbf{X})$$

Problem: rank is intractable!

Rank-regularized matrix completion

$$\min_{\mathbf{X}} \|(\mathbf{X}-\mathbf{M})\odot\mathbf{S}\|_F^2 + \mu\,\mathrm{rank}(\mathbf{X})$$

Convex relaxation

$$\min_{\mathbf{X}} \|(\mathbf{X}-\mathbf{M})\odot\mathbf{S}\|_F^2 + \mu\,\|\mathbf{X}\|_*$$

$$\|\mathbf{X}\|_* = \sigma_1 + \cdots + \sigma_n$$

Factorized form

$$\min_{\mathbf{Y},\mathbf{Z}^{\mathrm{T}}} \left\|\left(\mathbf{Y}\mathbf{Z}^{\mathrm{T}}-\mathbf{M}\right)\odot\mathbf{S}\right\|_F^2$$

Deep matrix factorization (DMF)

$$\min_{\mathbf{Y}_1,\ldots,\mathbf{Y}_N} \|(\mathbf{Y}_1\cdot\cdots\cdot\mathbf{Y}_N-\mathbf{M})\odot\mathbf{S}\|_F^2$$

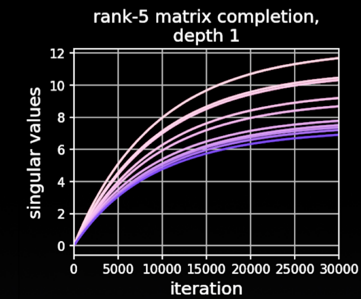
Candes & Recht, 2009 · Gunasekar *et al.*, 2017 · Arora *et al.*, 2019

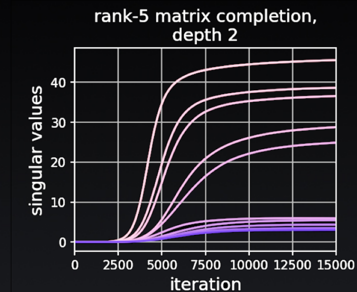
Deep matrix factorization (DMF)

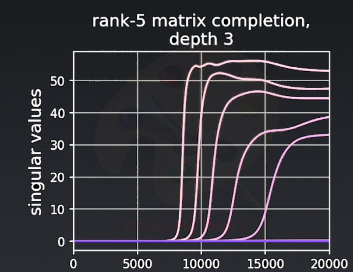
$$\min_{\mathbf{Y}_1,\ldots,\mathbf{Y}_N} \left\|\left(\mathbf{Y}_1 \cdot \cdots \cdot \mathbf{Y}_N - \mathbf{M}\right) \odot \mathbf{S}\right\|_{\mathrm{F}}^2$$

- Overparametrized representation
- Gradient descent dynamics promote implicit rank regularization
- The effect is stronger with depth N

$$\dot{\sigma}_i(t) = -N \cdot \left(\sigma_i^2(t)\right)^{1-\frac{1}{N}} \langle \nabla l(\mathbf{X}(t)), u_i(t) v_i^T(t) \rangle$$

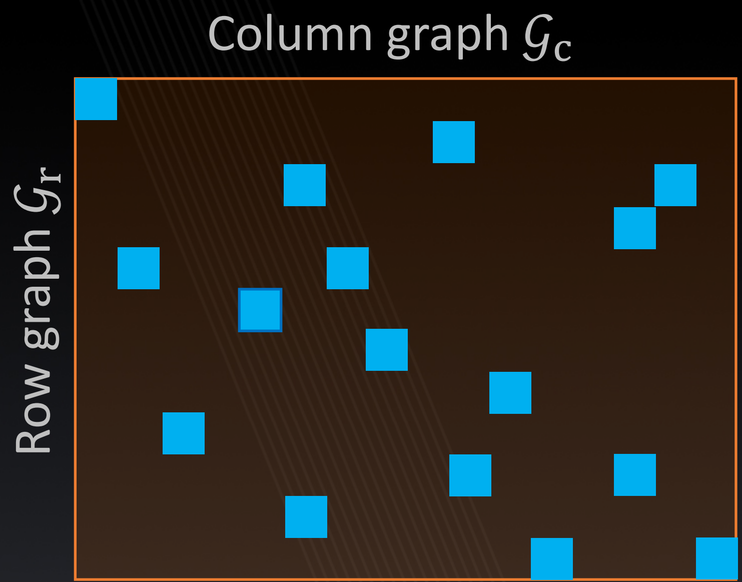






Gunasekar *et al.*, 2017 · Arora *et al.*, 2019

Geometric matrix completion



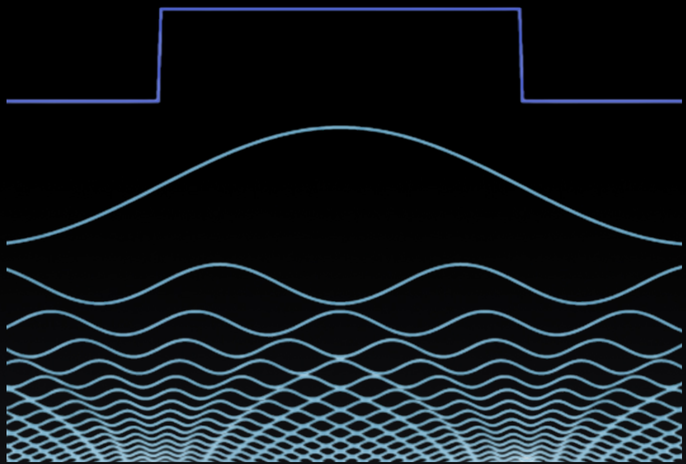
Matrix $\mathbf{X}$ lives on the product graph $\mathcal{G} = \mathcal{G}_{\mathrm{c}} \boxtimes \mathcal{G}_{\mathrm{r}}$

Prior: $\mathbf{X}$ is band-limited on $\mathcal{G}$

Assume $\mathcal{G}_{\mathrm{r}}$ and $\mathcal{G}_{\mathrm{c}}$ are known

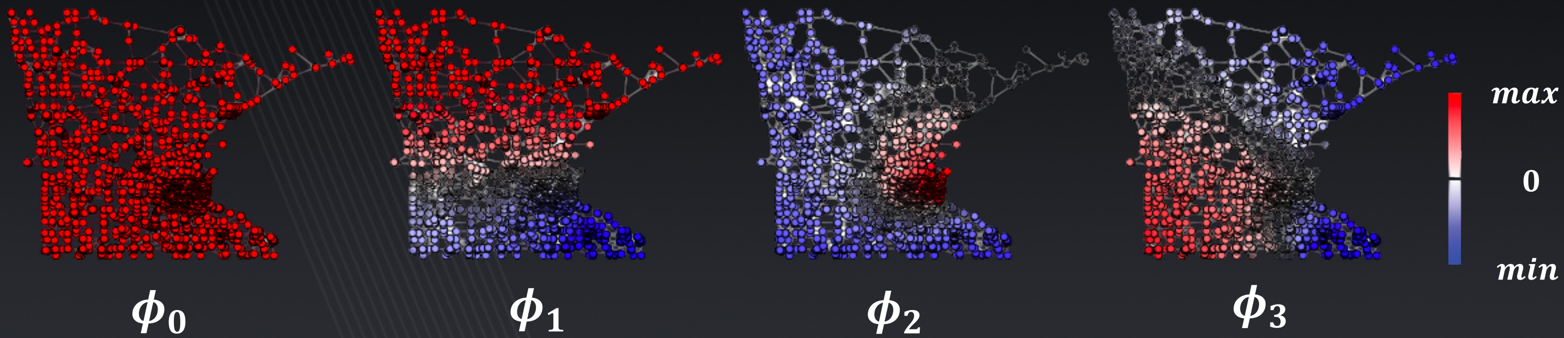
Kalofolias *et al.*, 2014 · Monti *et al.*, 2017

Spectral graph theory



Undirected graph $\mathcal{G}$

$\mathbf{W}$	adjacency matrix
$\mathbf{D}$	degree matrix
$\mathbf{L} = \mathbf{D} - \mathbf{W}$	graph Laplacian
$\mathbf{L} = \mathbf{D}^{-1}\mathbf{W}$	random walk Laplacian



Product graph Laplacian

Row graph $\mathcal{G}_{\mathrm{r}}$

$\mathbf{A}_{\mathrm{r}}$	adjacency matrix
$\mathbf{D}_{\mathrm{r}}$	degree matrix
$\mathbf{L}_{\mathrm{r}} = \mathbf{D}_{\mathrm{r}} - \mathbf{A}_{\mathrm{r}}$	graph Laplacian

Column graph $\mathcal{G}_{\mathrm{c}}$

$\mathbf{A}_{\mathrm{c}}$	adjacency matrix
$\mathbf{D}_{\mathrm{c}}$	degree matrix
$\mathbf{L}_{\mathrm{c}} = \mathbf{D}_{\mathrm{c}} - \mathbf{A}_{\mathrm{c}}$	graph Laplacian

$\mathbf{L} = \mathbf{L}_{\mathrm{r}} \oplus \mathbf{L}_{\mathrm{c}}$ Laplacian of product graph $\mathcal{G} = \mathcal{G}_{\mathrm{c}} \boxtimes \mathcal{G}_{\mathrm{r}}$

Dirichlet energy

Prior: matrix $\mathbf{X}$ is band-limited on $\mathcal{G} = \mathcal{G}_{\mathrm{c}} \boxtimes \mathcal{G}_{\mathrm{r}}$

Dirichlet energy $E(\mathbf{X}) = \langle \mathbf{X}, \mathbf{L}_{\mathrm{r}} \oplus \mathbf{L}_{\mathrm{c}} \mathbf{X} \rangle_{\mathcal{G}}$

$$= \mathrm{tr}\left(\mathbf{X}^{\mathrm{T}} \mathbf{L}_{\mathrm{r}} \mathbf{X}\right) + \mathrm{tr}\left(\mathbf{X} \mathbf{L}_{\mathrm{c}} \mathbf{X}^{\mathrm{T}}\right)$$

$$\min_{\mathbf{X}} \ \|(\mathbf{X} - \mathbf{M}) \odot \mathbf{S}\|_{\mathrm{F}}^2 + \mu_{\mathrm{r}} \, \mathrm{tr}\left(\mathbf{X}^{\mathrm{T}} \mathbf{L}_{\mathrm{r}} \mathbf{X}\right) + \mu_{\mathrm{c}} \, \mathrm{tr}\left(\mathbf{X} \mathbf{L}_{\mathrm{c}} \mathbf{X}^{\mathrm{T}}\right)$$

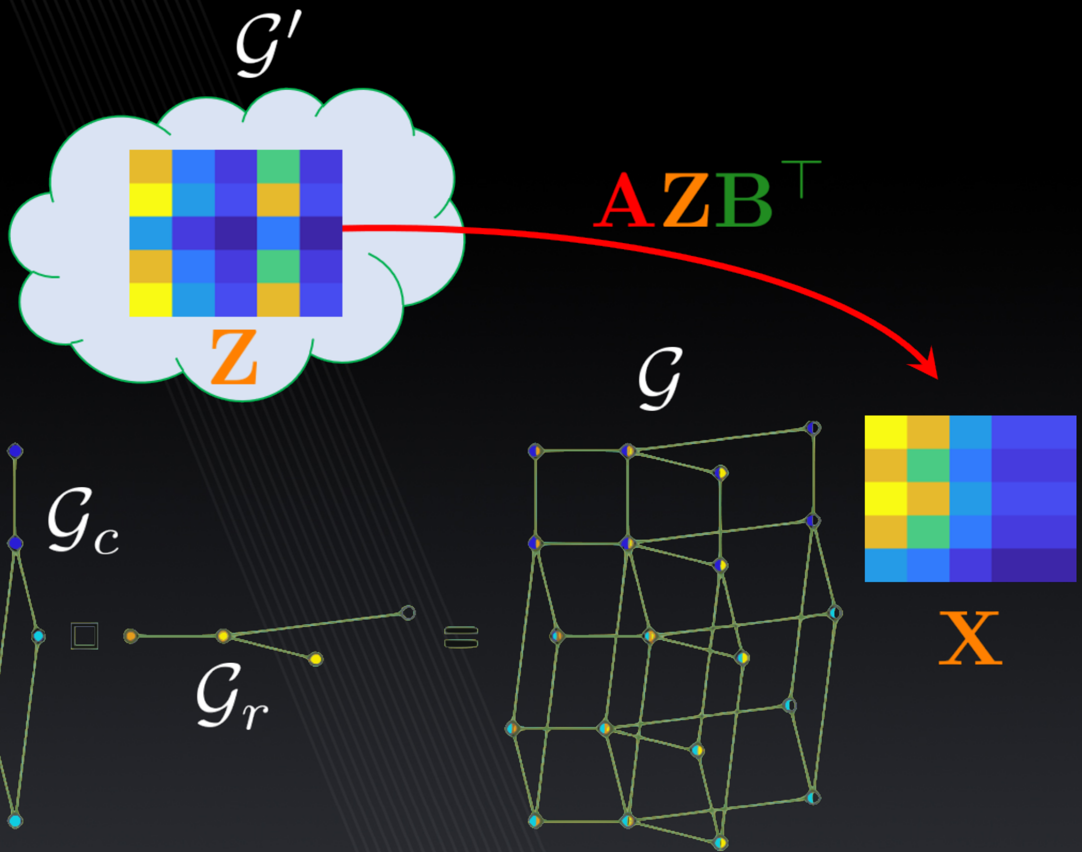
s.t. $\mathrm{rank}(\mathbf{X}) \leq \mathrm{k}$

DMF + spectral regularization

$$\min_{\mathbf{X}} \|(\mathbf{X}-\mathbf{M})\odot\mathbf{S}\|_F^2 + \mu_r \operatorname{tr}\left(\mathbf{X}^T\mathbf{L}_r\mathbf{X}\right) + \mu_c \operatorname{tr}\left(\mathbf{X}\mathbf{L}_c\mathbf{X}^T\right)$$
$$\text{s.t. } \operatorname{rank}(\mathbf{X}) \le k$$

(over)-parametrize with $\mathbf{X} = \mathbf{A}\mathbf{Z}\mathbf{B}^T$

$$\min_{\mathbf{A},\mathbf{Z},\mathbf{B}} \left\|\left(\mathbf{A}\mathbf{Z}\mathbf{B}^T-\mathbf{M}\right)\odot\mathbf{S}\right\|_F^2 + \mu_r \operatorname{tr}\left(\mathbf{B}\mathbf{Z}^T\mathbf{A}^T\mathbf{L}_r\mathbf{A}\mathbf{Z}\mathbf{B}^T\right) + \mu_c \operatorname{tr}\left(\mathbf{A}\mathbf{Z}\mathbf{B}^T\mathbf{L}_c\mathbf{B}\mathbf{Z}^T\mathbf{A}^T\right)$$
$$\text{s.t. } \cancel{\operatorname{rank}(\mathbf{A}\mathbf{Z}\mathbf{B}^T)} \le k$$



Example:

If $\mathcal{G}' \cong \mathcal{G}$

$$\mathbf{X} = \mathbf{\Pi_1 Z \Pi_2^T}$$

Spectral geometric matrix completion (SGMC)

$$\min_{\mathbf{C},\mathbf{P},\mathbf{Q}} \left\| \left(\mathbf{\Phi} \boxed{\mathbf{P}\mathbf{C}\mathbf{Q}^{\mathrm{T}}} \mathbf{\Psi}^{\mathrm{T}} - \mathbf{M} \right) \odot \mathbf{S} \right\|_{\mathrm{F}}^{2}$$ data term

$$+ \mu_{\mathrm{r}} \operatorname{tr}\left(\mathbf{Q}\mathbf{C}^{\mathrm{T}}\mathbf{P}^{\mathrm{T}} \mathbf{\Lambda}_{\mathrm{r}} \mathbf{P}\mathbf{C}\mathbf{Q}^{\mathrm{T}} \right)$$

smoothness

$$+ \mu_{\mathrm{c}} \operatorname{tr}\left(\mathbf{P}\mathbf{C}\mathbf{Q}^{\mathrm{T}} \mathbf{\Lambda}_{\mathrm{c}} \mathbf{Q}\mathbf{C}^{\mathrm{T}}\mathbf{P}^{\mathrm{T}} \right)$$

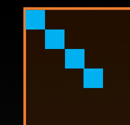
SGMC = a form of **DMF** + geometric regularization

Spectral geometric matrix completion (SGMC-Z)

$$\min_{\mathbf{C},\mathbf{P},\mathbf{Q}} \sum_{p,q} \left\| \left(\mathbf{\Phi} \mathbf{P} \mathbf{F}_{\mathrm{p}} \mathbf{C} \mathbf{G}_{\mathrm{q}}^{\mathrm{T}} \mathbf{Q}^{\mathrm{T}} \mathbf{\Psi}^{\mathrm{T}} - \mathbf{M} \right) \odot \mathbf{S} \right\|_{\mathrm{F}}^{2}$$

$\mathbf{F}_{\mathrm{p}}, \mathbf{G}_{\mathrm{q}} \equiv$ spectral filters

data term



$$+ \mu_{\mathrm{r}} \operatorname{tr}\left(\mathbf{Q} \mathbf{C}^{\mathrm{T}} \mathbf{P}^{\mathrm{T}} \mathbf{\Lambda}_{\mathrm{r}} \mathbf{P} \mathbf{C} \mathbf{Q}^{\mathrm{T}} \right)$$

smoothness

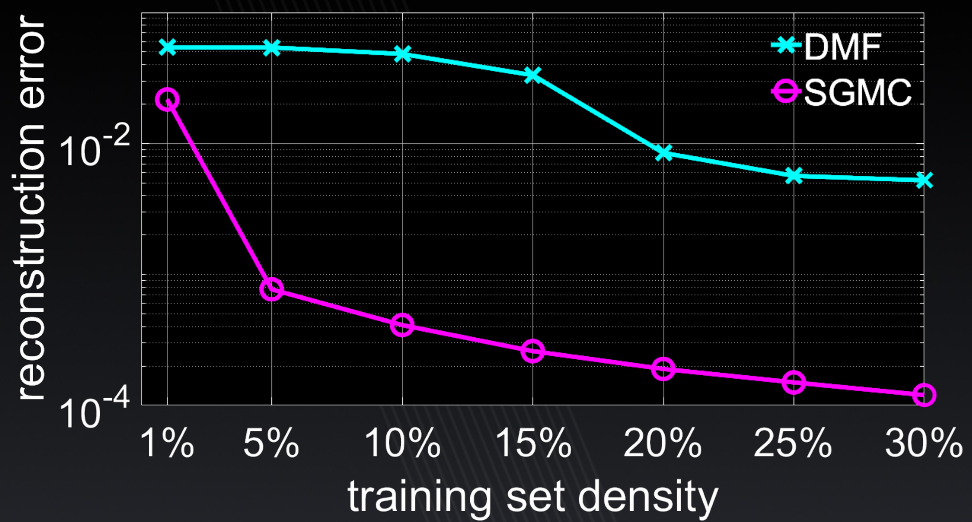
$$+ \mu_{\mathrm{c}} \operatorname{tr}\left(\mathbf{P} \mathbf{C} \mathbf{Q}^{\mathrm{T}} \mathbf{\Lambda}_{\mathrm{c}} \mathbf{Q} \mathbf{C}^{\mathrm{T}} \mathbf{P}^{\mathrm{T}} \right)$$

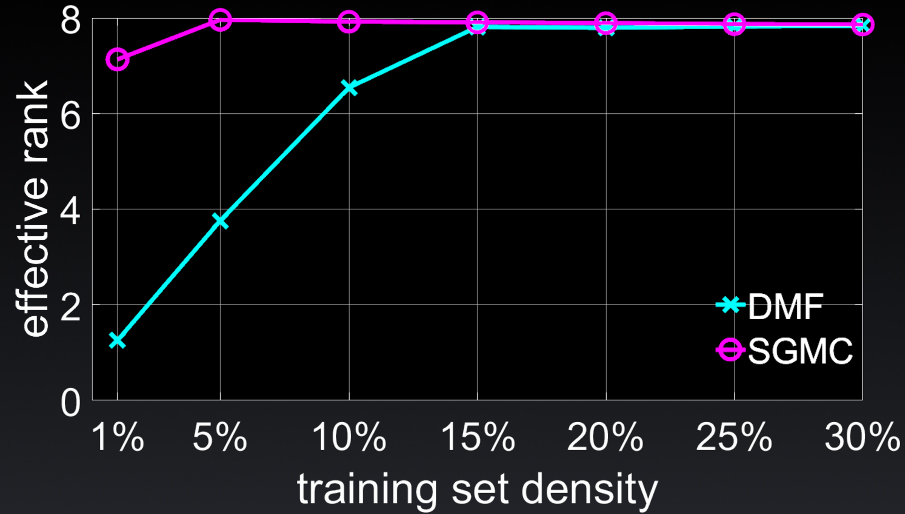
$$+ \rho_{\mathrm{r}} \left\| \operatorname{off}\left(\mathbf{P}^{\mathrm{T}} \mathbf{\Lambda}_{\mathrm{r}} \mathbf{P} \right) \right\|_{\mathrm{F}}^{2} + \rho_{\mathrm{c}} \left\| \operatorname{off}\left(\mathbf{Q}^{\mathrm{T}} \mathbf{\Lambda}_{\mathrm{c}} \mathbf{Q} \right) \right\|_{\mathrm{F}}^{2}$$

approximate diagonalization

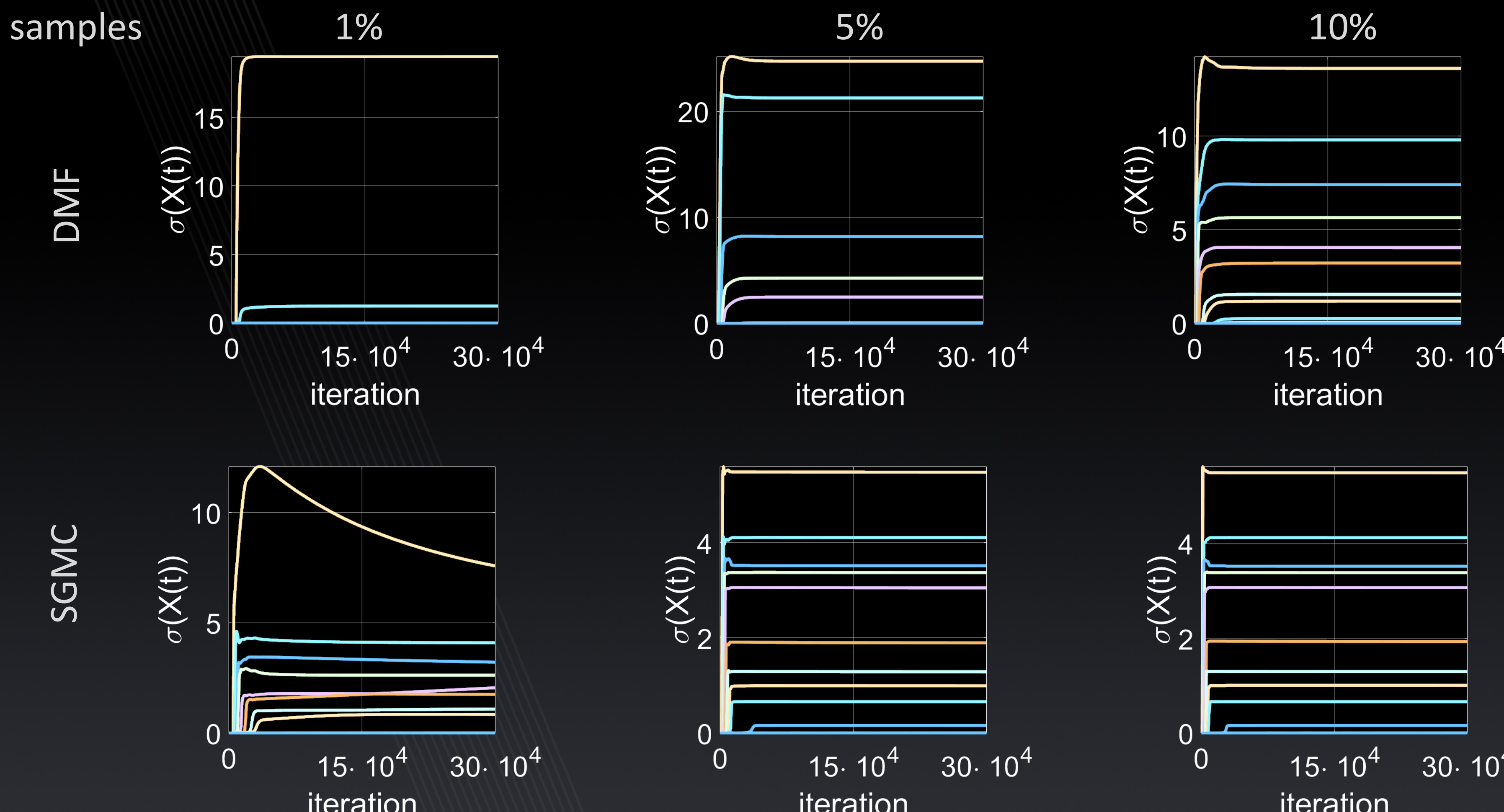
SGMC = a form of **DMF** + geometric regularization

Comparison to DMF – sample density

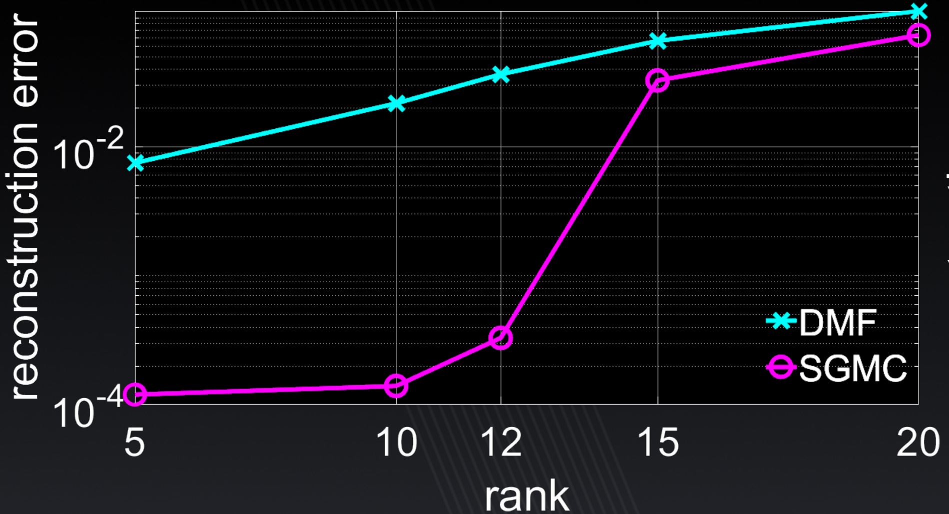


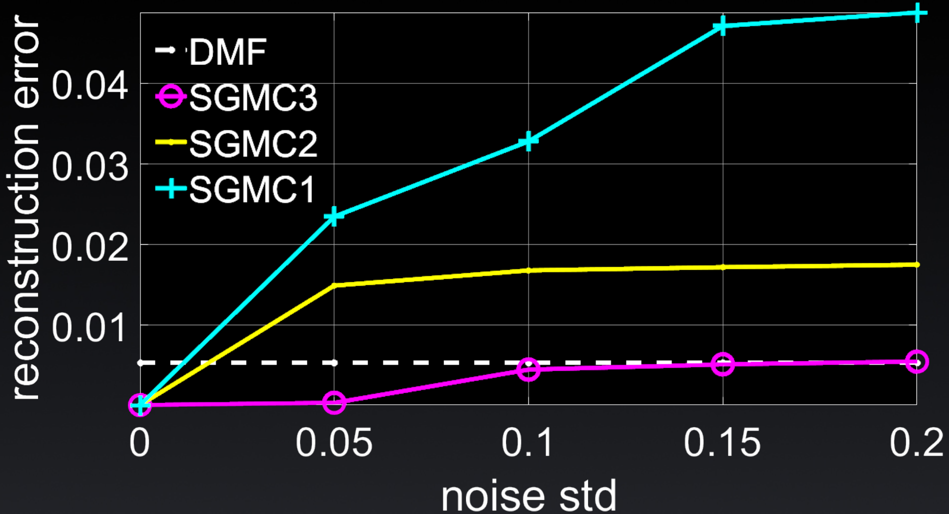


Data: Synthetic Netflix (variant)


samples
1%
5%
10%
DMF
SGMC
σ(X(t))
iteration
0
15· 10^4
30· 10^4

Comparison to DMF – rank and noise in graphs





Data: Synthetic Netflix (variant)

Results

Model	Synthetic Neflix	Flixster	Douban	ML-100K
MC (Candès & Recht, 2009)	–	1.533	0.845	0.973
GMC (Kalofolias et al., 2014)	0.3693	–	–	0.996
GRALS (Rao et al., 2015)	0.0114	1.313/1.245	0.833	0.945
RGCNN (Monti et al., 2017)	0.0053[a]	1.179/0.926	0.801	0.929
GC-MC (Berg et al., 2017)	–	**0.941**/0.917	0.734	0.910[b]
FM	0.0064	3.32	3.15	1.10
DMF (Arora et al., 2019)	0.0468[d]	1.06	0.732	0.918[c]/ 0.922
SGMC (ours)	**0.0021**	0.971 / 0.900	**0.731**	0.912
SGMC-Z (ours)	0.0036	0.957 / **0.888**	0.733	**0.907**[c]/ 0.913

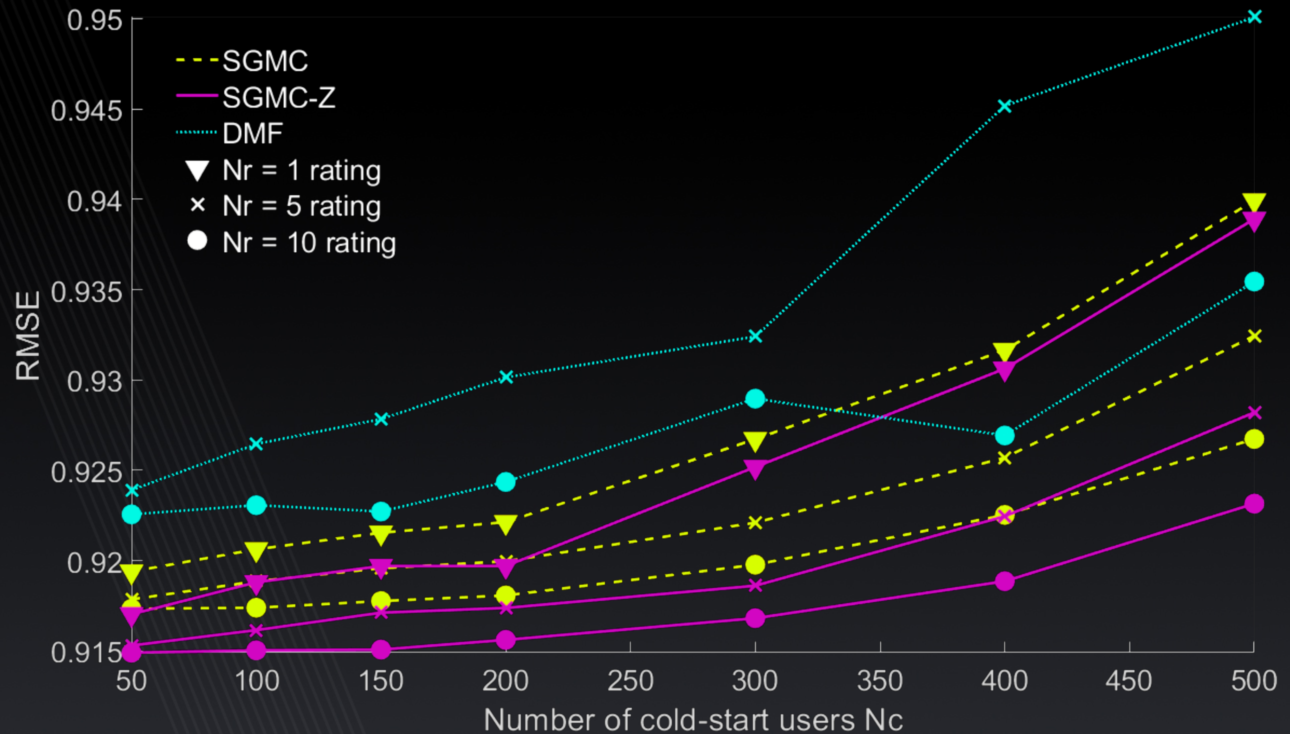
[a] This number corresponds to the inseparable version of MGCNN.
[b] This number corresponds to GC-MC.
[c] Early stopping.
[d] Initialization with $0.01\mathbf{I}$.

Visual Sensing Theory & Applications Laboratory

Computer Science Department

Technion Israel Institute of Technology

Cold Start Analysis



Data: Movielens 100K

Conclusion

- Simple spectral technique for matrix completion
- Performs better than many more complex methods
- Includes implicit graph learning
- Variant of DMF with different parametrization + geometric priors
- Useful interpretations related to shape analysis

Paper:

https://arxiv.org/pdf/1911.07255

Code:

https://colab.research.google.com/drive/1OkNEiTHok14gcVf3NxFIbAFutDN6-Tx6

https://github.com/amitboy/SGMC