ICML | 2019
Thirty-sixth International Conference
on Machine Learning

Neurally-Guided Structure Inference

http://ngsi.csail.mit.edu

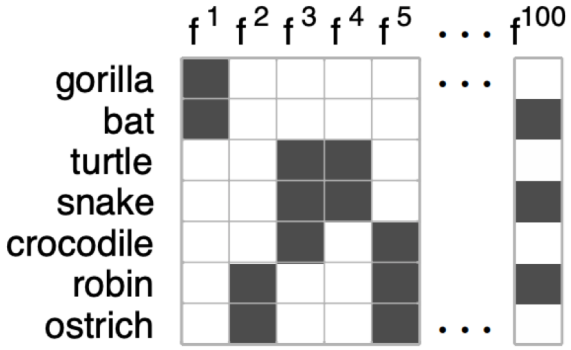
Sidi Lu*, Jiayuan Mao*, Josh Tenenbaum, and Jiajun Wu

(* indicates equal contributions)

Structure Inference

Structure Inference

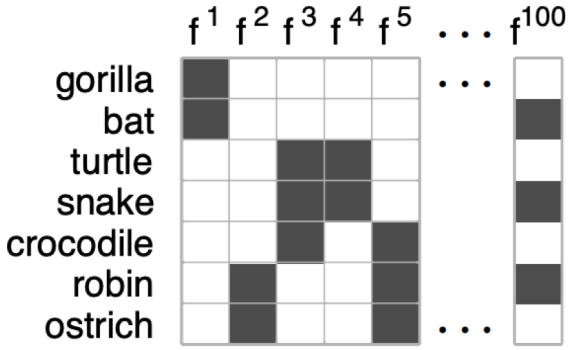
Data



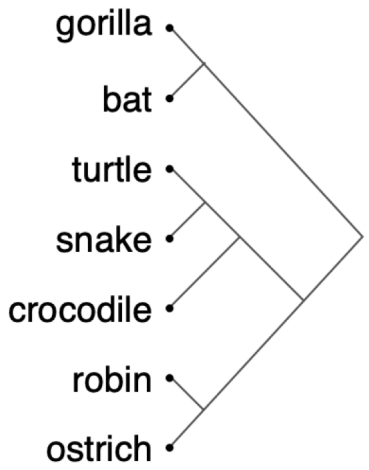
[Kemp et al. 2008]

Structure Inference

Data



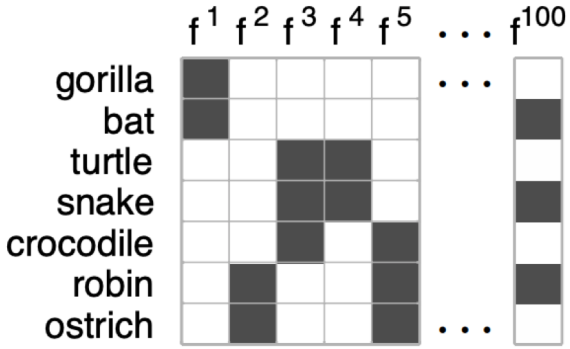
Structure



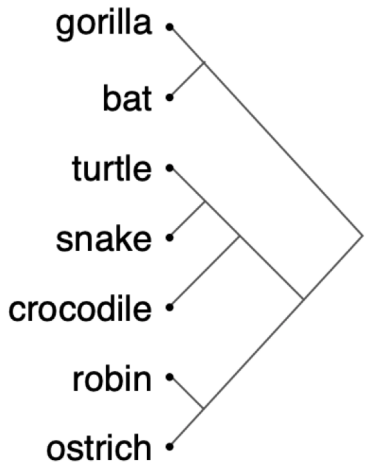
[Kemp et al. 2008]

Structure Inference

Data



Structure



[Kemp et al. 2008]

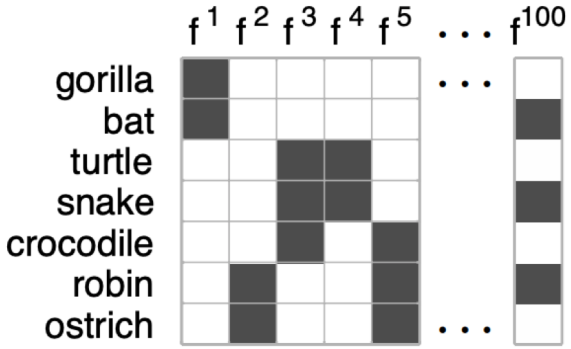
[LeCun et al. 1998]

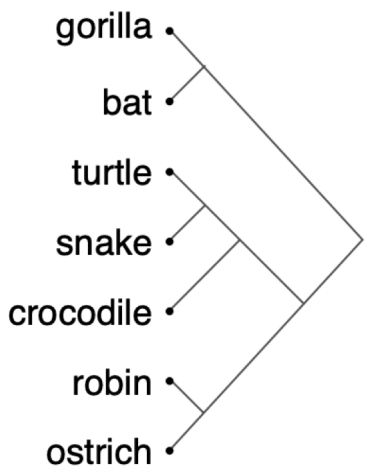
Structure Inference

Data

Structure Inference

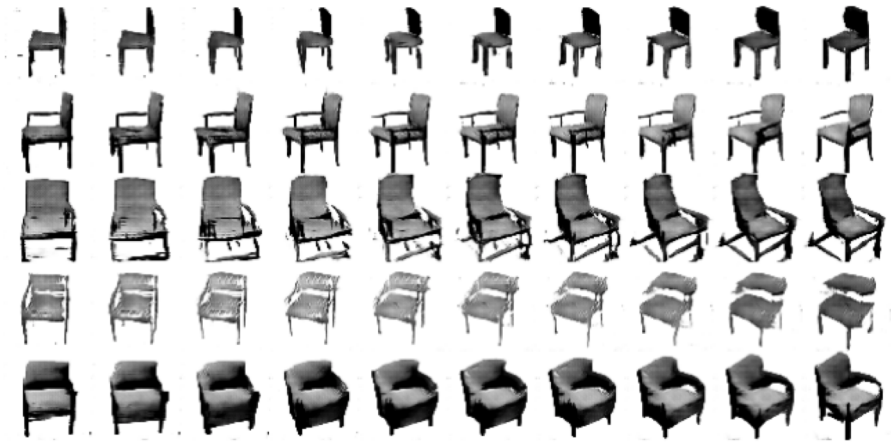
Structure





[Kemp et al. 2008]

[LeCun et al. 1998]



[Chen et al. 2016]

Matrix Decomposition Models

- Clustering $MG+G$
- Low-Rank Approximation $GG+G$
- Binary Features $BG+G$
- Random Walk $CG+G$
- Co-Clustering $M(GM^T+G)+G$
- Clustered Matrix Decomp. $(MG+G)(GM^T+G)+G$
- Binary Matrix Factorization $(BG+G)(GB^T+G)+G$
- Dependent GSM $(\exp(GG+G)\circ G)+G$
- ……

Structure Inference with Matrix Decomposition Models

[Grosse et al. 2012]

Structure Inference with Matrix Decomposition Models

$$\begin{bmatrix} -1.01 & 1.36 & -1.64 \\ -0.76 & 0.48 & -0.50 \\ \vdots & \vdots & \vdots \\ -0.24 & 0.89 & -1.12 \end{bmatrix}$$

Input Matrix

20×3

[Grosse et al. 2012]

Structure Inference with Matrix Decomposition Models

$$\begin{bmatrix} -1.01 & 1.36 & -1.64 \\ -0.76 & 0.48 & -0.50 \\ \vdots & \vdots & \vdots \\ -0.24 & 0.89 & -1.12 \end{bmatrix}$$

Input Matrix

20×3

Structure G

[Grosse et al. 2012]

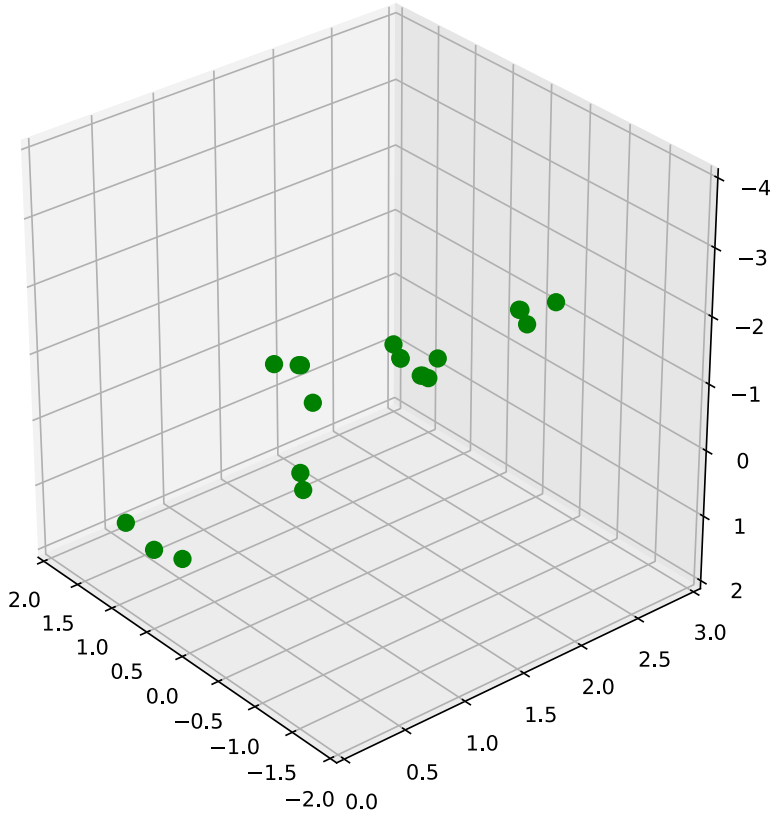
Structure Inference with Matrix Decomposition Models

$$\begin{bmatrix} -1.01 & 1.36 & -1.64 \\ -0.76 & 0.48 & -0.50 \\ \vdots & \vdots & \vdots \\ -0.24 & 0.89 & -1.12 \end{bmatrix}$$

Input Matrix

20×3

Structure G



[Grosse et al. 2012]

Structure Inference with Matrix Decomposition Models

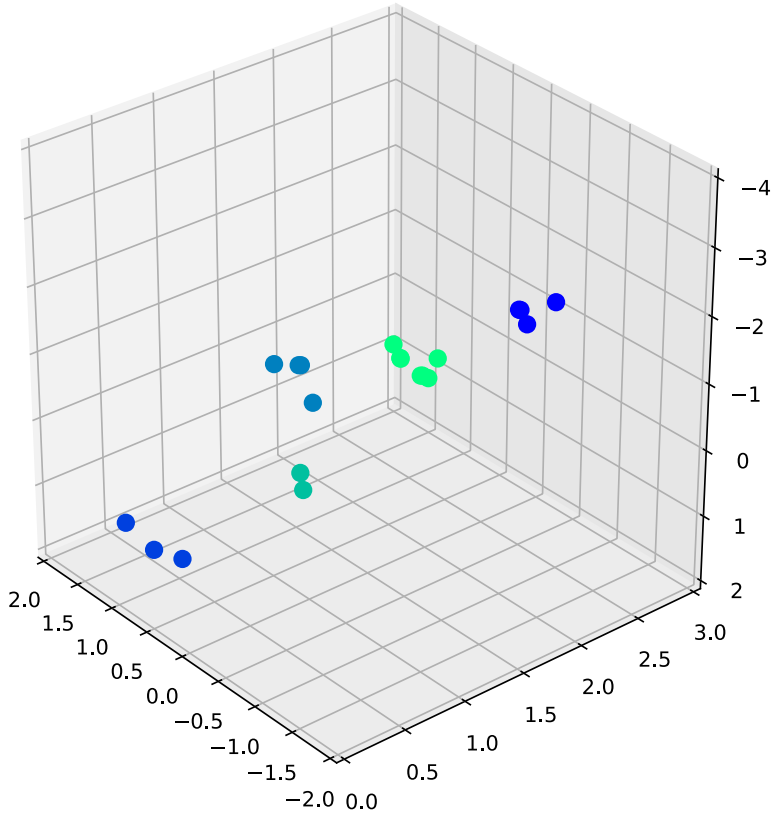
$$\begin{bmatrix} -1.01 & 1.36 & -1.64 \\ -0.76 & 0.48 & -0.50 \\ \vdots & \vdots & \vdots \\ -0.24 & 0.89 & -1.12 \end{bmatrix}$$

Cluster →

Input Matrix

20×3

Structure G



[Grosse et al. 2012]

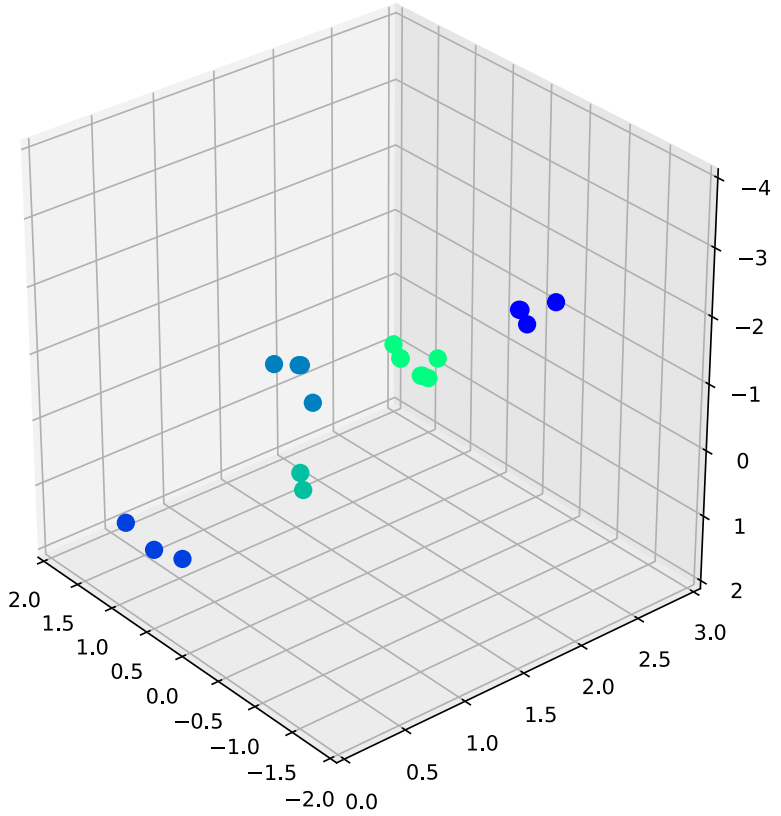
Structure Inference with Matrix Decomposition Models

$$\begin{bmatrix} -1.01 & 1.36 & -1.64 \\ -0.76 & 0.48 & -0.50 \\ \vdots & \vdots & \vdots \\ -0.24 & 0.89 & -1.12 \end{bmatrix} \xrightarrow{\text{Cluster}} \begin{bmatrix} 0\,0\,0\,0\,1 \\ 0\,0\,0\,1\,0 \\ \vdots \\ 0\,0\,1\,0\,0 \end{bmatrix}$$

Input Matrix
20×3

Cluster Label
20×5

Structure G



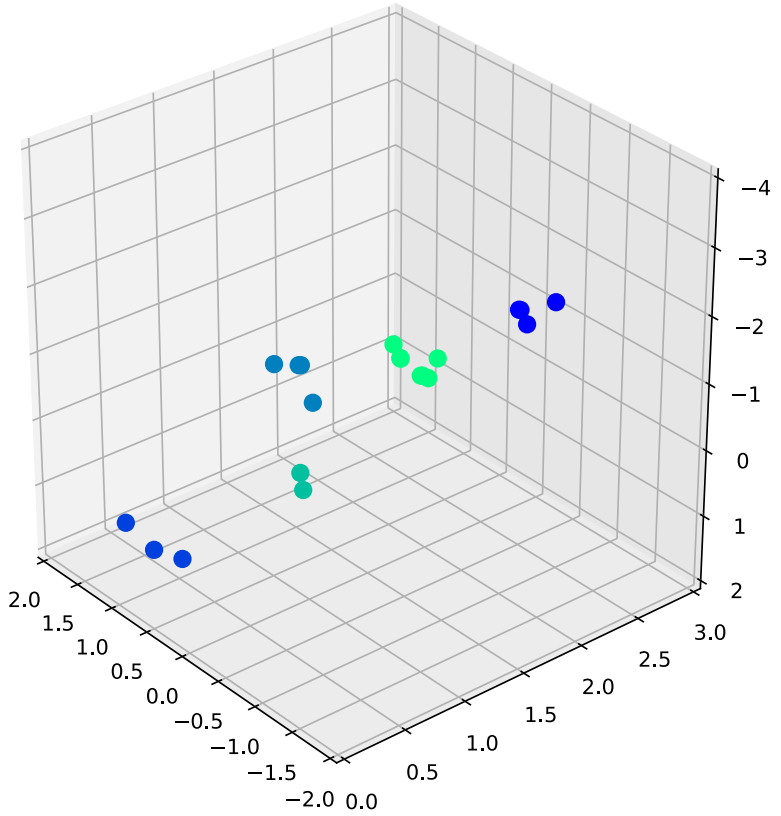
[Grosse et al. 2012]

Structure Inference with Matrix Decomposition Models

$$\begin{bmatrix} -1.01 & 1.36 & -1.64 \\ -0.76 & 0.48 & -0.50 \\ \vdots & \vdots & \vdots \\ -0.24 & 0.89 & -1.12 \end{bmatrix} \xrightarrow{\text{Cluster}} \begin{bmatrix} 0\,0\,0\,0\,1 \\ 0\,0\,0\,1\,0 \\ \vdots \\ 0\,0\,1\,0\,0 \end{bmatrix} \begin{bmatrix} -1.32 & +2.01 & -2.33 \\ +0.04 & -0.23 & -0.30 \\ +0.08 & +0.95 & -1.27 \\ -0.68 & +0.52 & -0.54 \\ -0.77 & +1.40 & -1.66 \end{bmatrix}$$

Input Matrix 20×3 — Cluster Label 20×5 — Cluster Center 5×3

Structure G



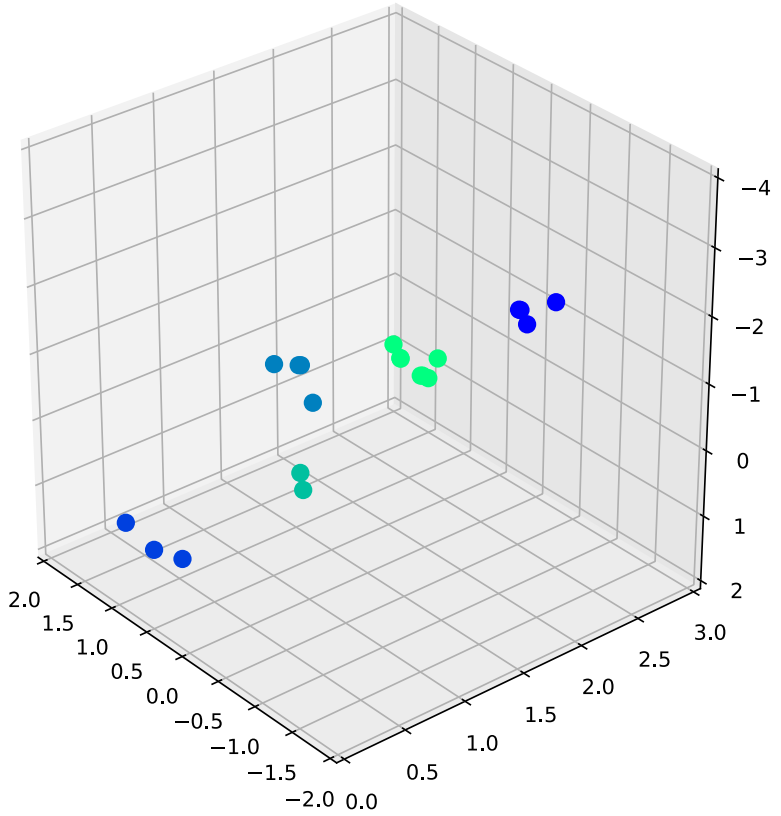
[Grosse et al. 2012]

Structure Inference with Matrix Decomposition Models

$$\underbrace{\begin{bmatrix} -1.01 & 1.36 & -1.64 \\ -0.76 & 0.48 & -0.50 \\ \vdots & \vdots & \vdots \\ -0.24 & 0.89 & -1.12 \end{bmatrix}}_{\text{Input Matrix } 20\times 3} \xrightarrow{\text{Cluster}} \underbrace{\begin{bmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ & & \vdots & & \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix}}_{\text{Cluster Label } 20\times 5} \underbrace{\begin{bmatrix} -1.32 & +2.01 & -2.33 \\ +0.04 & -0.23 & -0.30 \\ +0.08 & +0.95 & -1.27 \\ -0.68 & +0.52 & -0.54 \\ -0.77 & +1.40 & -1.66 \end{bmatrix}}_{\text{Cluster Center } 5\times 3} + \underbrace{\begin{bmatrix} -0.24 & -0.04 & 0.01 \\ -0.61 & 0.76 & -04 \\ \vdots & \vdots & \vdots \\ -0.09 & 0.90 & -1.84 \end{bmatrix}}_{\text{Cluster Noise } 20\times 3}$$

Structure G

Structure $MG + G$



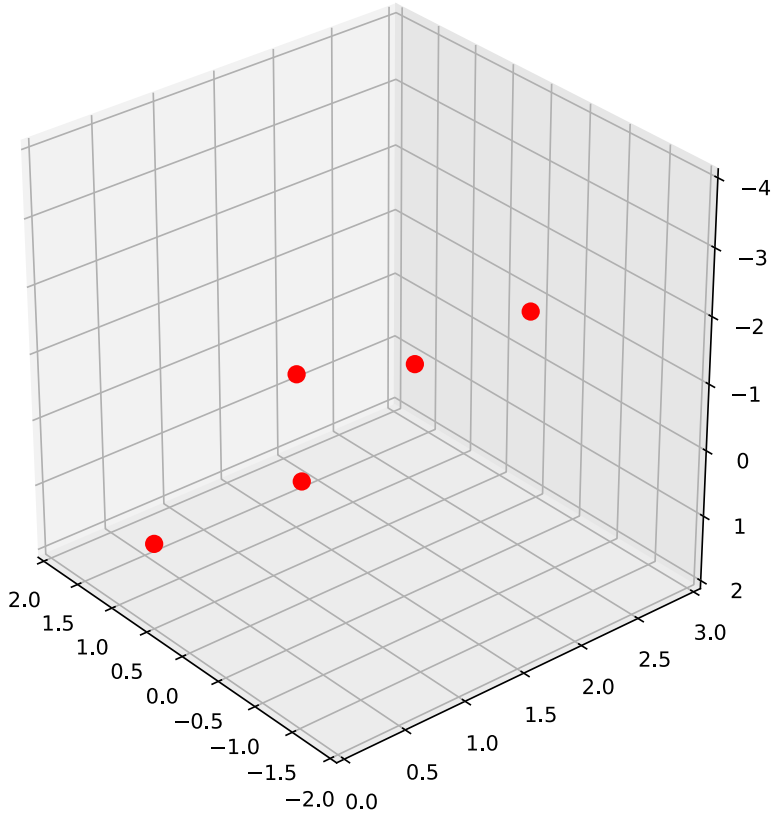
[Grosse et al. 2012]

Structure Inference with Matrix Decomposition Models

$$\underbrace{\begin{bmatrix} -1.01 & 1.36 & -1.64 \\ -0.76 & 0.48 & -0.50 \\ \vdots & \vdots & \vdots \\ -0.24 & 0.89 & -1.12 \end{bmatrix}}_{\text{Input Matrix } 20\times 3} \xrightarrow{\text{Cluster}} \underbrace{\begin{bmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ & & \vdots & & \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix}}_{\text{Cluster Label } 20\times 5} \underbrace{\begin{bmatrix} -1.32 & +2.01 & -2.33 \\ +0.04 & -0.23 & -0.30 \\ +0.08 & +0.95 & -1.27 \\ -0.68 & +0.52 & -0.54 \\ -0.77 & +1.40 & -1.66 \end{bmatrix}}_{\text{Cluster Center } 5\times 3} + \underbrace{\begin{bmatrix} -0.24 & -0.04 & 0.01 \\ -0.61 & 0.76 & -04 \\ \vdots & \vdots & \vdots \\ -0.09 & 0.90 & -1.84 \end{bmatrix}}_{\text{Cluster Noise } 20\times 3}$$

Structure G

Structure $MG + G$

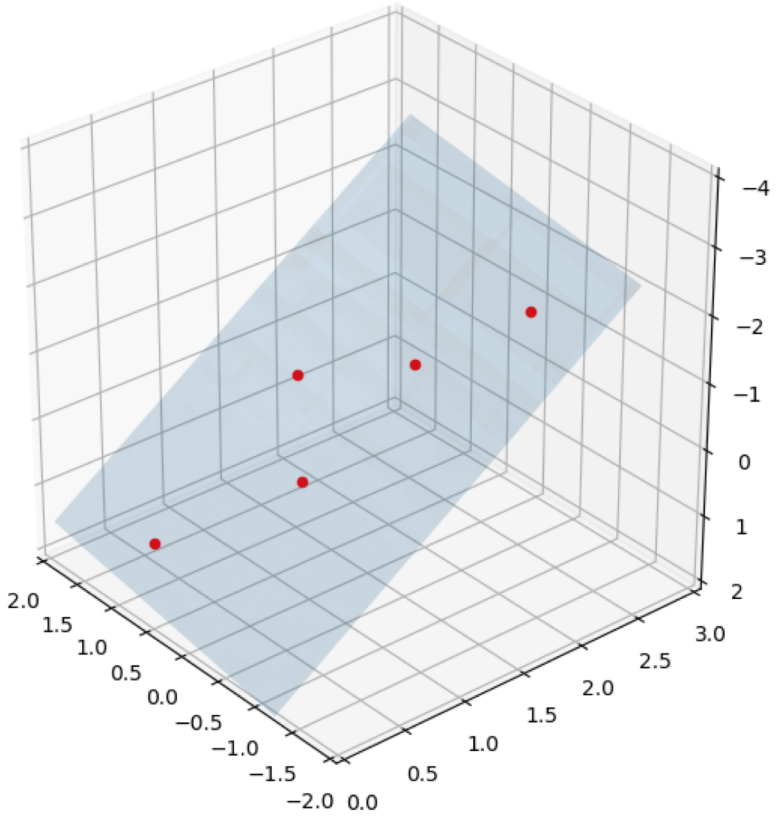


[Grosse et al. 2012]

Structure Inference with Matrix Decomposition Models

$$\underbrace{\begin{bmatrix} -1.01 & 1.36 & -1.64 \\ -0.76 & 0.48 & -0.50 \\ \vdots & \vdots & \vdots \\ -0.24 & 0.89 & -1.12 \end{bmatrix}}_{\text{Input Matrix } 20\times 3} \xrightarrow{\text{Cluster}} \underbrace{\begin{bmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ & & \vdots & & \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix}}_{\text{Cluster Label } 20\times 5} \underbrace{\begin{bmatrix} -1.32 & +2.01 & -2.33 \\ +0.04 & -0.23 & -0.30 \\ +0.08 & +0.95 & -1.27 \\ -0.68 & +0.52 & -0.54 \\ -0.77 & +1.40 & -1.66 \end{bmatrix}}_{\text{Cluster Center } 5\times 3} + \underbrace{\begin{bmatrix} -0.24 & -0.04 & 0.01 \\ -0.61 & 0.76 & -04 \\ \vdots & \vdots & \vdots \\ -0.09 & 0.90 & -1.84 \end{bmatrix}}_{\text{Cluster Noise } 20\times 3}$$

Structure G — Structure $MG + G$

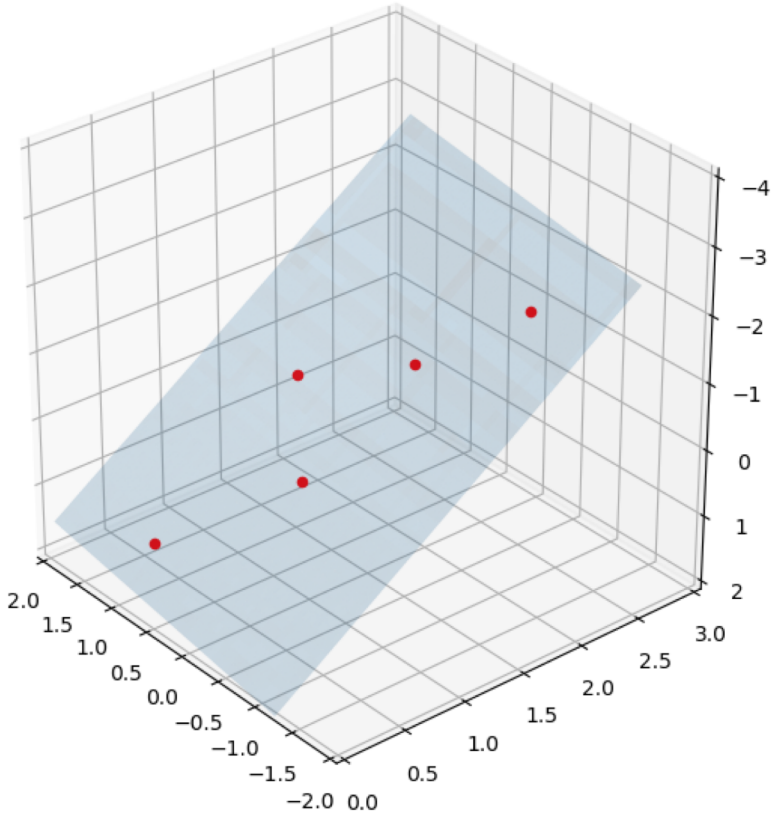


[Grosse et al. 2012]

Structure Inference with Matrix Decomposition Models

$$\underset{\text{Input Matrix } 20\times3}{\begin{bmatrix} -1.01 & 1.36 & -1.64 \\ -0.76 & 0.48 & -0.50 \\ \vdots & \vdots & \vdots \\ -0.24 & 0.89 & -1.12 \end{bmatrix}} \xrightarrow{\text{Cluster}} \underset{\text{Cluster Label } 20\times5}{\begin{bmatrix} 0&0&0&0&1 \\ 0&0&0&1&0 \\ & \vdots & & & \\ 0&0&1&0&0 \end{bmatrix}} \underset{\text{Cluster Center } 5\times3}{\begin{bmatrix} -1.32 & +2.01 & -2.33 \\ +0.04 & -0.23 & -0.30 \\ +0.08 & +0.95 & -1.27 \\ -0.68 & +0.52 & -0.54 \\ -0.77 & +1.40 & -1.66 \end{bmatrix}} + \underset{\text{Cluster Noise } 20\times3}{\begin{bmatrix} -0.24 & -0.04 & 0.01 \\ -0.61 & 0.76 & -04 \\ \vdots & \vdots & \vdots \\ -0.09 & 0.90 & -1.84 \end{bmatrix}}$$

Structure G — Structure $MG + G$



LowRank →

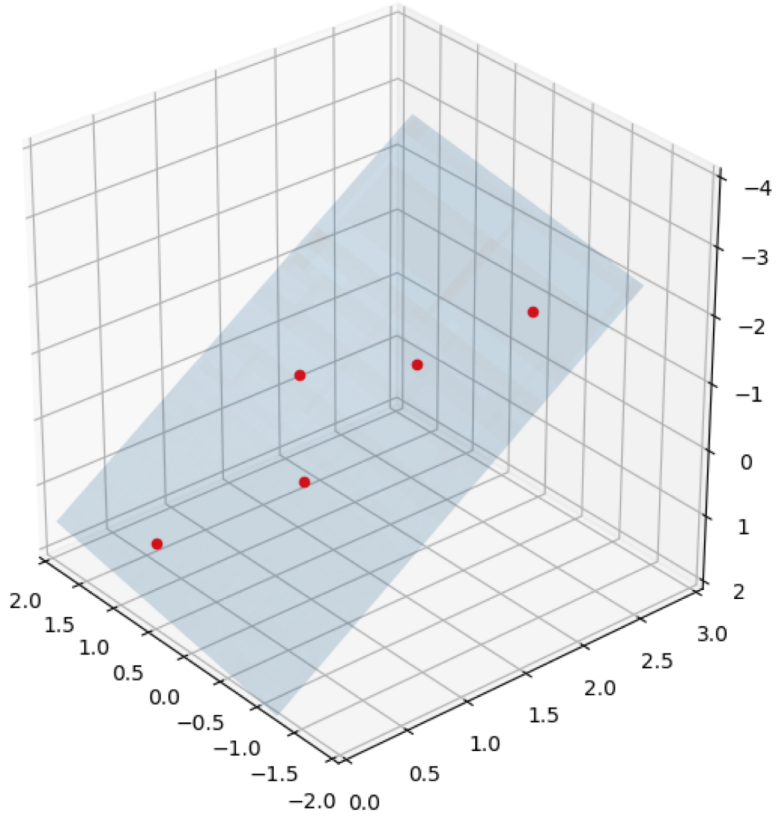
[Grosse et al. 2012]

Structure Inference with Matrix Decomposition Models

$$\underset{\text{Input Matrix } 20\times3}{\begin{bmatrix} -1.01 & 1.36 & -1.64 \\ -0.76 & 0.48 & -0.50 \\ \vdots & \vdots & \vdots \\ -0.24 & 0.89 & -1.12 \end{bmatrix}} \xrightarrow{\text{Cluster}} \underset{\text{Cluster Label } 20\times5}{\begin{bmatrix} 0&0&0&0&1 \\ 0&0&0&1&0 \\ & \vdots & & & \\ 0&0&1&0&0 \end{bmatrix}} \underset{\text{Cluster Center } 5\times3}{\begin{bmatrix} -1.32 & +2.01 & -2.33 \\ +0.04 & -0.23 & -0.30 \\ +0.08 & +0.95 & -1.27 \\ -0.68 & +0.52 & -0.54 \\ -0.77 & +1.40 & -1.66 \end{bmatrix}} + \underset{\text{Cluster Noise } 20\times3}{\begin{bmatrix} -0.24 & -0.04 & 0.01 \\ -0.61 & 0.76 & -04 \\ \vdots & \vdots & \vdots \\ -0.09 & 0.90 & -1.84 \end{bmatrix}}$$

Structure G — Structure $MG + G$

$$\xrightarrow{\text{LowRank}} \underset{\text{Cluster Label } 20\times5}{\begin{bmatrix} 0&0&0&0&1 \\ 0&0&0&1&0 \\ & \vdots & & & \\ 0&0&1&0&0 \end{bmatrix}} \left(\quad \underset{\text{Cluster Center}}{} \quad \right) + \underset{\text{Cluster Noise } 20\times3}{\begin{bmatrix} -0.24 & -0.04 & 0.01 \\ -0.61 & 0.76 & -04 \\ \vdots & \vdots & \vdots \\ -0.09 & 0.90 & -1.84 \end{bmatrix}}$$



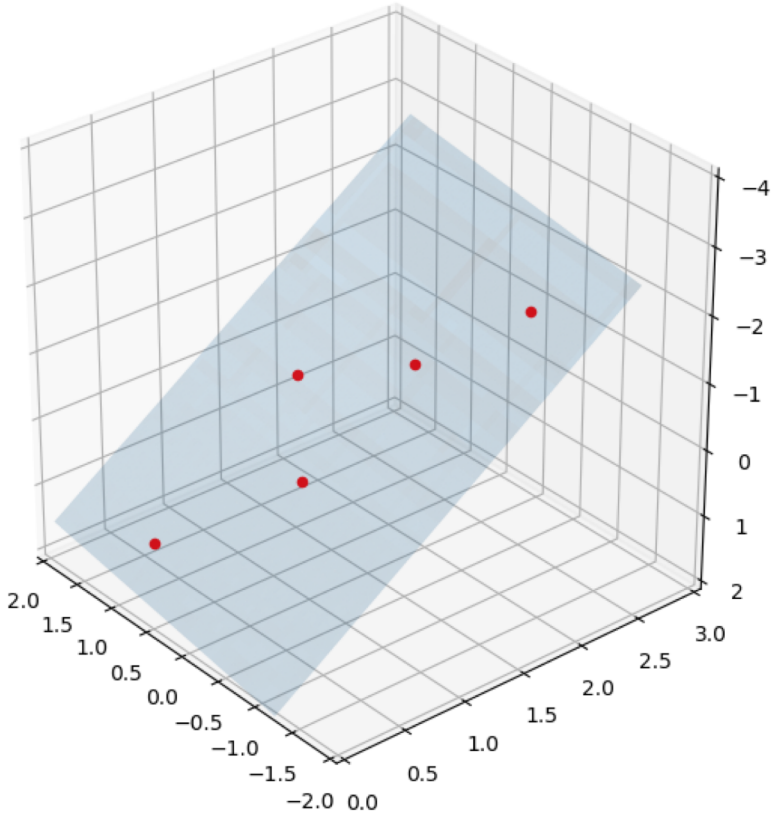
[Grosse et al. 2012]

Structure Inference with Matrix Decomposition Models

$$\underbrace{\begin{bmatrix} -1.01 & 1.36 & -1.64 \\ -0.76 & 0.48 & -0.50 \\ \vdots & \vdots & \vdots \\ -0.24 & 0.89 & -1.12 \end{bmatrix}}_{\text{Input Matrix } 20\times3} \xrightarrow{\text{Cluster}} \underbrace{\begin{bmatrix} 0&0&0&0&1 \\ 0&0&0&1&0 \\ & \vdots & \\ 0&0&1&0&0 \end{bmatrix}}_{\text{Cluster Label } 20\times5} \underbrace{\begin{bmatrix} -1.32 & +2.01 & -2.33 \\ +0.04 & -0.23 & -0.30 \\ +0.08 & +0.95 & -1.27 \\ -0.68 & +0.52 & -0.54 \\ -0.77 & +1.40 & -1.66 \end{bmatrix}}_{\text{Cluster Center } 5\times3} + \underbrace{\begin{bmatrix} -0.24 & -0.04 & 0.01 \\ -0.61 & 0.76 & -04 \\ \vdots & \vdots & \vdots \\ -0.09 & 0.90 & -1.84 \end{bmatrix}}_{\text{Cluster Noise } 20\times3}$$

Structure G

Structure $MG + G$



$$\xrightarrow{\text{LowRank}} \underbrace{\begin{bmatrix} 0&0&0&0&1 \\ 0&0&0&1&0 \\ & \vdots & \\ 0&0&1&0&0 \end{bmatrix}}_{\text{Cluster Label } 20\times5} \left(\underbrace{\begin{bmatrix} -1.32 & +2.01 \\ +0.04 & -0.23 \\ +0.08 & +0.95 \\ -0.68 & +0.52 \\ -0.77 & +1.40 \end{bmatrix}}_{\text{Cluster Center } (5\times2)} \underbrace{\begin{bmatrix} 1 & 0 & -0.22 \\ 0 & 1 & -1.30 \end{bmatrix}}_{(2\times3)} \right) + \underbrace{\begin{bmatrix} -0.24 & -0.04 & 0.01 \\ -0.61 & 0.76 & -04 \\ \vdots & \vdots & \vdots \\ -0.09 & 0.90 & -1.84 \end{bmatrix}}_{\text{Cluster Noise } 20\times3}$$

Structure $M(GG + G) + G$

[Grosse et al. 2012]

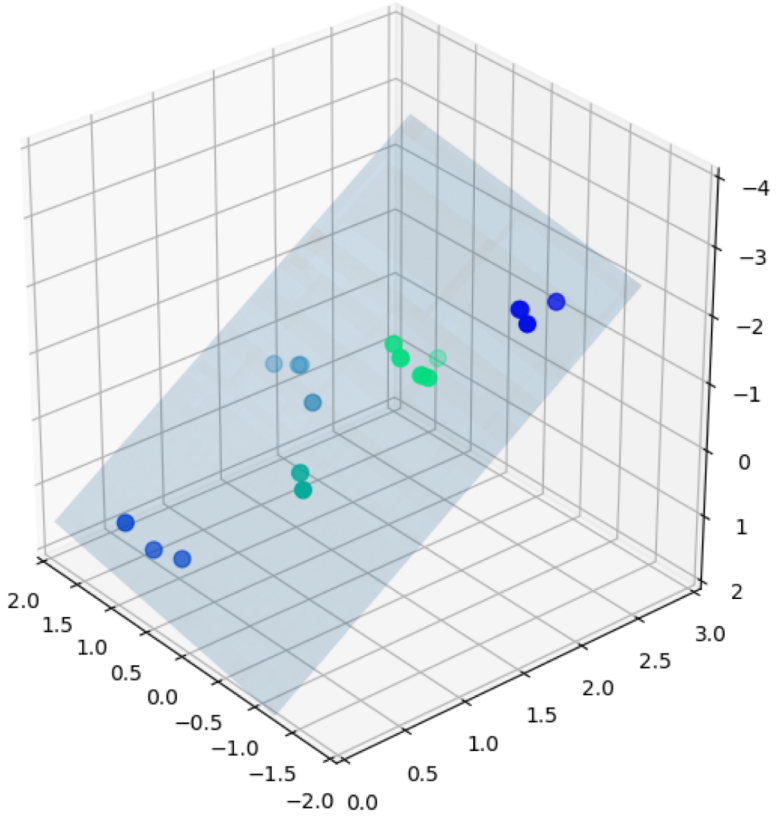
Structure Inference with Matrix Decomposition Models

$$\underbrace{\begin{bmatrix} -1.01 & 1.36 & -1.64 \\ -0.76 & 0.48 & -0.50 \\ \vdots & \vdots & \vdots \\ -0.24 & 0.89 & -1.12 \end{bmatrix}}_{\text{Input Matrix } 20\times3} \xrightarrow{\text{Cluster}} \underbrace{\begin{bmatrix} 0&0&0&0&1 \\ 0&0&0&1&0 \\ & \vdots & \\ 0&0&1&0&0 \end{bmatrix}}_{\text{Cluster Label } 20\times5} \underbrace{\begin{bmatrix} -1.32 & +2.01 & -2.33 \\ +0.04 & -0.23 & -0.30 \\ +0.08 & +0.95 & -1.27 \\ -0.68 & +0.52 & -0.54 \\ -0.77 & +1.40 & -1.66 \end{bmatrix}}_{\text{Cluster Center } 5\times3} + \underbrace{\begin{bmatrix} -0.24 & -0.04 & 0.01 \\ -0.61 & 0.76 & -04 \\ \vdots & \vdots & \vdots \\ -0.09 & 0.90 & -1.84 \end{bmatrix}}_{\text{Cluster Noise } 20\times3}$$

Structure G → Structure $MG + G$

$$\xrightarrow{\text{LowRank}} \underbrace{\begin{bmatrix} 0&0&0&0&1 \\ 0&0&0&1&0 \\ & \vdots & \\ 0&0&1&0&0 \end{bmatrix}}_{\text{Cluster Label } 20\times5} \left(\underbrace{\begin{bmatrix} -1.32 & +2.01 \\ +0.04 & -0.23 \\ +0.08 & +0.95 \\ -0.68 & +0.52 \\ -0.77 & +1.40 \end{bmatrix}}_{\text{Cluster Center } (5\times2)@(2\times3)} \begin{bmatrix} 1 & 0 & -0.22 \\ 0 & 1 & -1.30 \end{bmatrix} \right) + \underbrace{\begin{bmatrix} -0.24 & -0.04 & 0.01 \\ -0.61 & 0.76 & -04 \\ \vdots & \vdots & \vdots \\ -0.09 & 0.90 & -1.84 \end{bmatrix}}_{\text{Cluster Noise } 20\times3}$$

Structure $M(GG + G) + G$



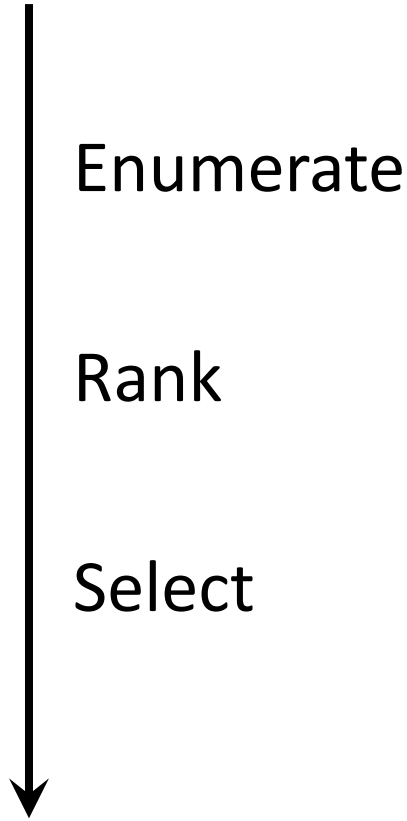
[Grosse et al. 2012]

Naïve Exhaustive Search

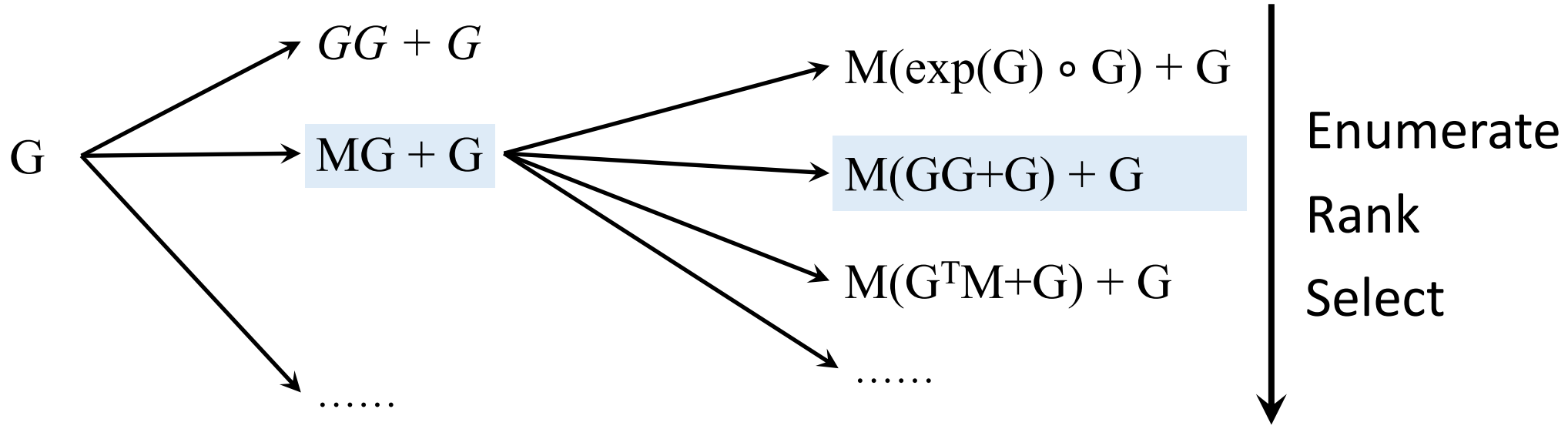
- Clustering $MG+G$
- Low-Rank Approximation $GG+G$
- Binary Features $BG+G$
- Random Walk $CG+G$
- Co-Clustering $M(GM^T+G)+G$
- Clustered Matrix Decomp. $(MG+G)(GM^T+G)+G$
- Binary Matrix Factorization $(BG+G)(GB^T+G)+G$
- Dependent GSM $(\exp(GG+G)\circ G)+G$
- ……

Naïve Exhaustive Search

• Clustering	$MG+G$
• Low-Rank Approximation	$GG+G$
• Binary Features	$BG+G$
• Random Walk	$CG+G$
• Co-Clustering	$M(GM^T+G)+G$
• Clustered Matrix Decomp.	$(MG+G)(GM^T+G)+G$
• Binary Matrix Factorization	$(BG+G)(GB^T+G)+G$
• Dependent GSM	$(\exp(GG+G)\circ G)+G$
• ……	



Layer-wise Exhaustive Search



$$\begin{bmatrix} -1.01 & 1.36 & -1.64 \\ -0.76 & 0.48 & -0.50 \\ \vdots & \vdots & \vdots \\ -0.24 & 0.89 & -1.12 \end{bmatrix}$$

Input Matrix
20×3

Structure G

$$\begin{bmatrix} 0\,0\,0\,0\,1 \\ 0\,0\,0\,1\,0 \\ \vdots \\ 0\,0\,1\,0\,0 \end{bmatrix} \begin{bmatrix} -1.32 & +2.01 & -2.33 \\ +0.04 & -0.23 & -0.30 \\ +0.08 & +0.95 & -1.27 \\ -0.68 & +0.52 & -0.54 \\ -0.77 & +1.40 & -1.66 \end{bmatrix} + Noise$$

Cluster Label
20×5

Cluster Center
5×3

Structure $MG + G$

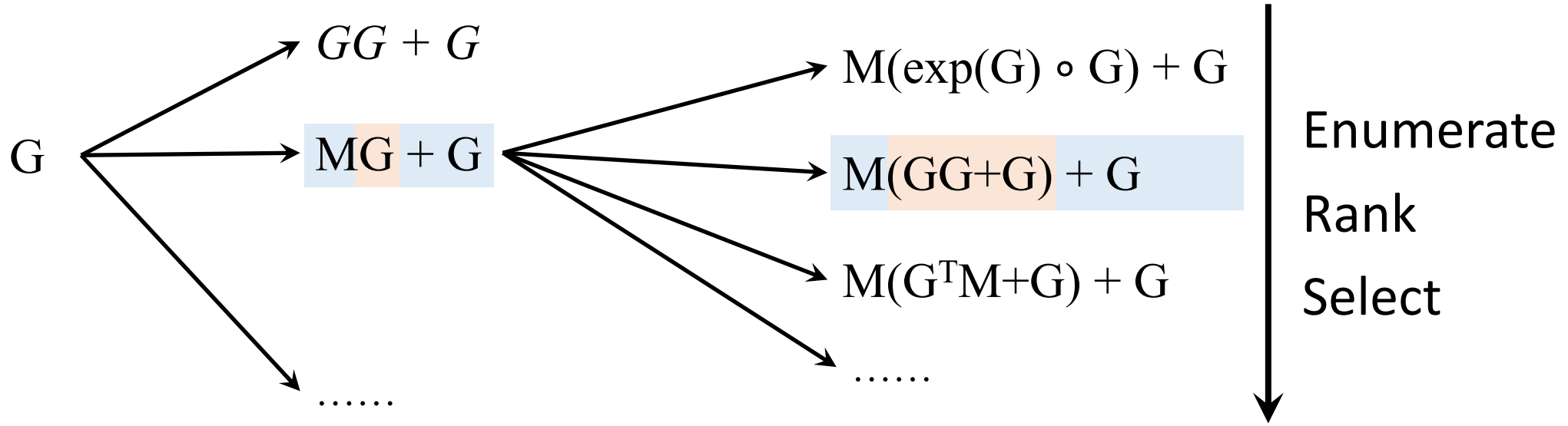
$$\begin{bmatrix} 0\,0\,0\,0\,1 \\ 0\,0\,0\,1\,0 \\ \vdots \\ 0\,0\,1\,0\,0 \end{bmatrix} \left(\begin{bmatrix} -1.32 & +2.01 \\ +0.04 & -0.23 \\ +0.08 & +0.95 \\ -0.68 & +0.52 \\ -0.77 & +1.40 \end{bmatrix} \begin{bmatrix} 1 & 0 & -0.22 \\ 0 & 1 & -1.30 \end{bmatrix} \right) + Noise$$

Cluster Center
(5×2)@(2×3)

Structure $M(GG + G) + G$

Layer-wise Exhaustive Search

Key Observation: Each step involves the **same** sub-problem.



$$\begin{bmatrix} -1.01 & 1.36 & -1.64 \\ -0.76 & 0.48 & -0.50 \\ \vdots & \vdots & \vdots \\ -0.24 & 0.89 & -1.12 \end{bmatrix}$$

Input Matrix
20×3

Structure G

$$\begin{bmatrix} 0\,0\,0\,0\,1 \\ 0\,0\,0\,1\,0 \\ \vdots \\ 0\,0\,1\,0\,0 \end{bmatrix} \begin{bmatrix} -1.32 & +2.01 & -2.33 \\ +0.04 & -0.23 & -0.30 \\ +0.08 & +0.95 & -1.27 \\ -0.68 & +0.52 & -0.54 \\ -0.77 & +1.40 & -1.66 \end{bmatrix} + Noise$$

Cluster Label 20×5 | Cluster Center 5×3

Structure $MG + G$

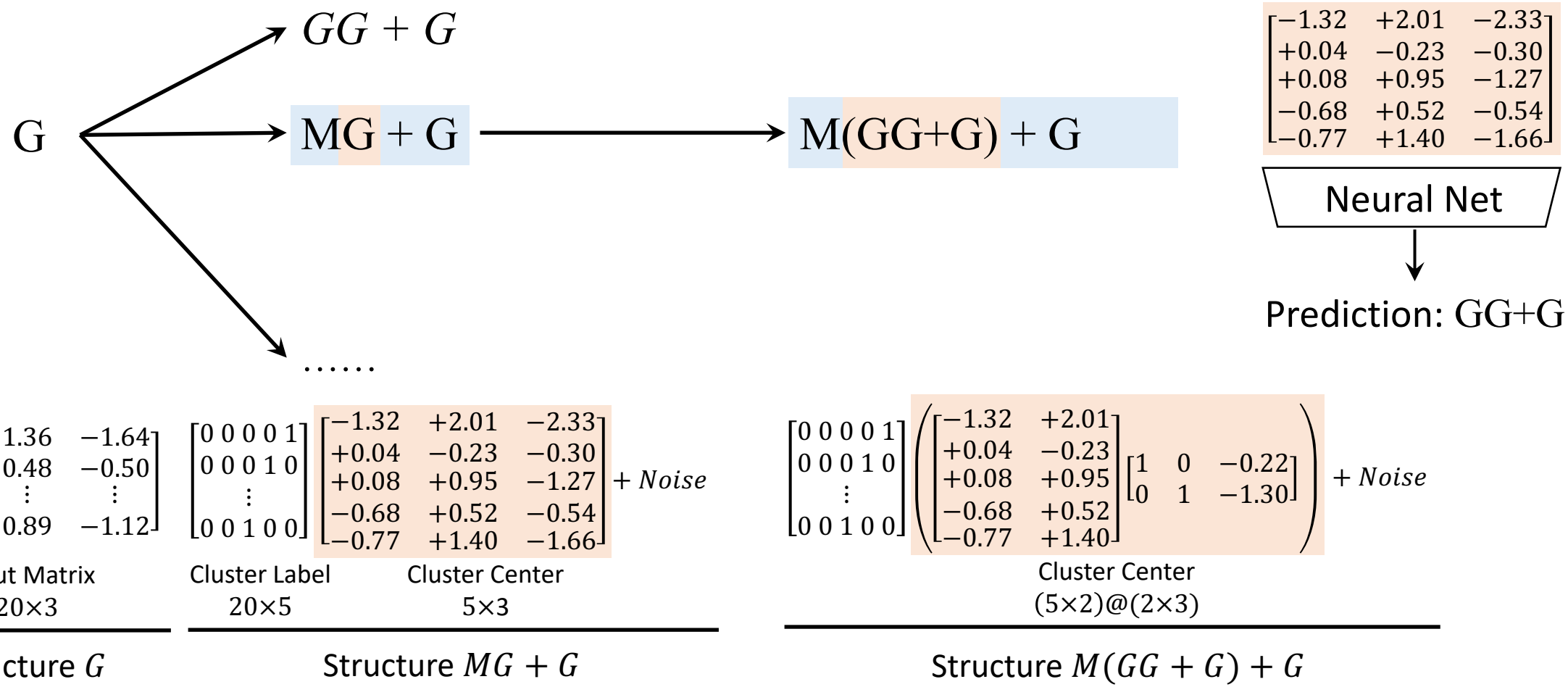
$$\begin{bmatrix} 0\,0\,0\,0\,1 \\ 0\,0\,0\,1\,0 \\ \vdots \\ 0\,0\,1\,0\,0 \end{bmatrix} \left(\begin{bmatrix} -1.32 & +2.01 \\ +0.04 & -0.23 \\ +0.08 & +0.95 \\ -0.68 & +0.52 \\ -0.77 & +1.40 \end{bmatrix} \begin{bmatrix} 1 & 0 & -0.22 \\ 0 & 1 & -1.30 \end{bmatrix} \right) + Noise$$

Cluster Center
(5×2)@(2×3)

Structure $M(GG + G) + G$

Neurally Guided Search

Use Neural Network for Layer-wise Amortized Inference.



Experiment – Accuracy

Data	Ground Truth	Grosse et al. (2012)	Mat2Seq	NG-SI (ours)
Low-rank	$GG+G$	Correct	Correct	Correct
Clustering	$MG+G$	Correct	Correct	Correct
Binary latent features	$BG+G$	Correct	Correct	Correct
Random walk	$CG+G$	Correct	Correct	Correct
Co-clustering	$M(GM^T+G)+G$	Correct	$MG+G$	Correct
Binary matrix factorization	$B(GB^T+G)+G$	Correct	$BG+G$	Correct
BCTF	$(MG+G)(GM^T+G)+G$	Correct	$BG+G$	Correct
Sparse coding	$s(G)G+G$	Correct	$s(G)$	Correct
Dependent GSM	$s(GG+G)G+G$	$s(G)G+G$	$s(G)$	$s(G)G+G$
Linear dynamical system	$(CG+G)G+G$	Correct	$CG+G$	$B(s(G)G+G)+G$
Motion Capture - Level 1	-	$CG+G$	$CG+G$	$CG+G$
Motion Capture - Level 2	-	$C(GG+G)+G$	$CG+G$	$C(GG+G)+G/CG+G$
Image Patch - Level 1	-	$GG+G$	$GG+G$	$GG+G$
Image Patch - Level 2	-	$s(G)G+G$	$GG+G$	$s(G)G+G$
Image Patch - Level 3	-	$s(GG+G)G+G/s(G)G+G$	$GG+G$	$s(G)G+G$

Experiment – Accuracy

Data	Ground Truth	Grosse et al. (2012)	Mat2Seq	NG-SI (ours)
Low-rank	$GG+G$	Correct	Correct	Correct
Clustering	$MG+G$	Correct	Correct	Correct
Binary latent features	$BG+G$	Correct	Correct	Correct
Random walk	$CG+G$	Correct	Correct	Correct
Co-clustering	$M(GM^T+G)+G$	Correct	$MG+G$	Correct
Binary matrix factorization	$B(GB^T+G)+G$	Correct	$BG+G$	Correct
BCTF	$(MG+G)(GM^T+G)+G$	Correct	$BG+G$	Correct
Sparse coding	$s(G)G+G$	Correct	$s(G)$	Correct
Dependent GSM	$s(GG+G)G+G$	$s(G)G+G$	$s(G)$	$s(G)G+G$
Linear dynamical system	$(CG+G)G+G$	Correct	$CG+G$	$B(s(G)G+G)+G$
Motion Capture - Level 1	-	$CG+G$	$CG+G$	$CG+G$
Motion Capture - Level 2	-	$C(GG+G)+G$	$CG+G$	$C(GG+G)+G/CG+G$
Image Patch - Level 1	-	$GG+G$	$GG+G$	$GG+G$
Image Patch - Level 2	-	$s(G)G+G$	$GG+G$	$s(G)G+G$
Image Patch - Level 3	-	$s(GG+G)G+G/s(G)G+G$	$GG+G$	$s(G)G+G$

Train

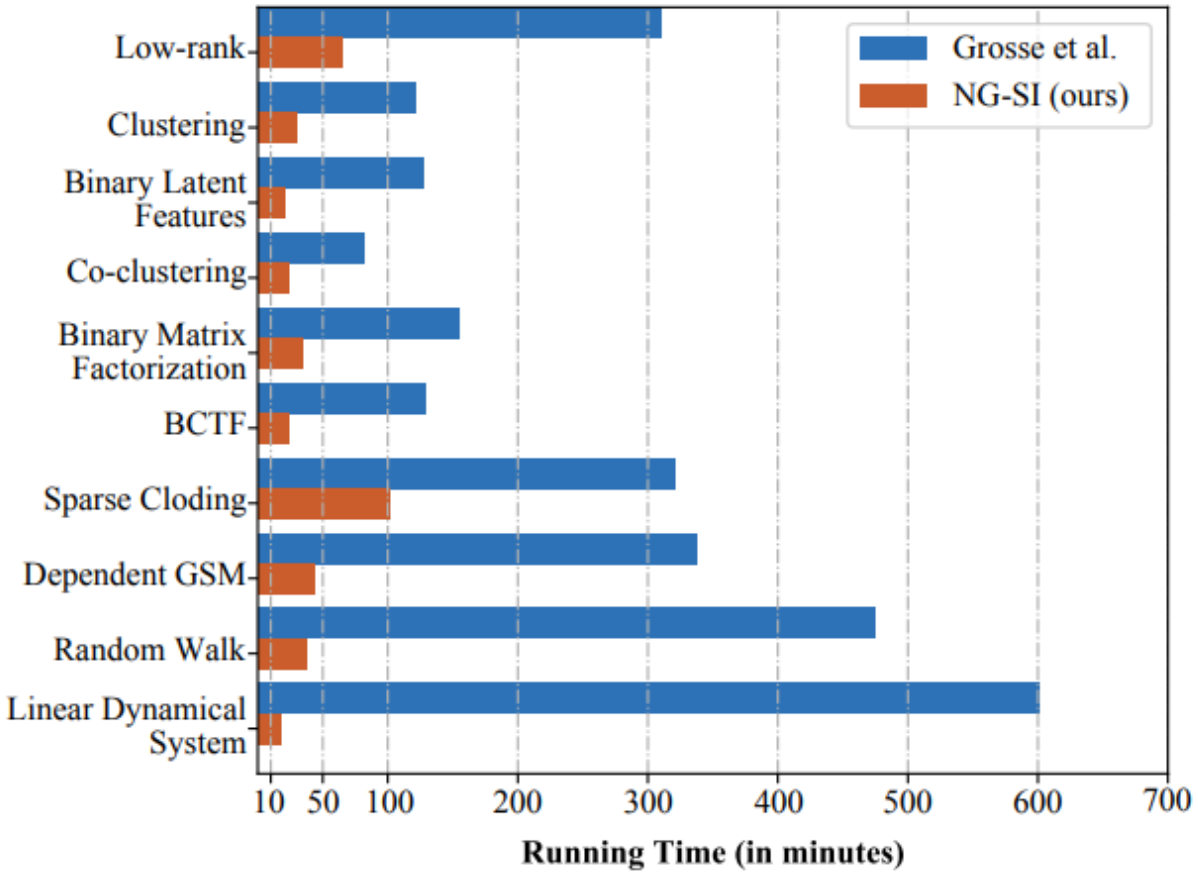
Experiment – Accuracy

Data	Ground Truth	Grosse et al. (2012)	Mat2Seq	NG-SI (ours)
Low-rank	$GG+G$	Correct	Correct	Correct
Clustering	$MG+G$	Correct	Correct	Correct
Binary latent features	$BG+G$	Correct	Correct	Correct
Random walk	$CG+G$	Correct	Correct	Correct
Co-clustering	$M(GM^T+G)+G$	Correct	$MG+G$	Correct
Binary matrix factorization	$B(GB^T+G)+G$	Correct	$BG+G$	Correct
BCTF	$(MG+G)(GM^T+G)+G$	Correct	$BG+G$	Correct
Sparse coding	$s(G)G+G$	Correct	$s(G)$	Correct
Dependent GSM	$s(GG+G)G+G$	$s(G)G+G$	$s(G)$	$s(G)G+G$
Linear dynamical system	$(CG+G)G+G$	Correct	$CG+G$	$B(s(G)G+G)+G$
Motion Capture - Level 1	-	$CG+G$	$CG+G$	$CG+G$
Motion Capture - Level 2	-	$C(GG+G)+G$	$CG+G$	$C(GG+G)+G/CG+G$
Image Patch - Level 1	-	$GG+G$	$GG+G$	$GG+G$
Image Patch - Level 2	-	$s(G)G+G$	$GG+G$	$s(G)G+G$
Image Patch - Level 3	-	$s(GG+G)G+G/s(G)G+G$	$GG+G$	$s(G)G+G$

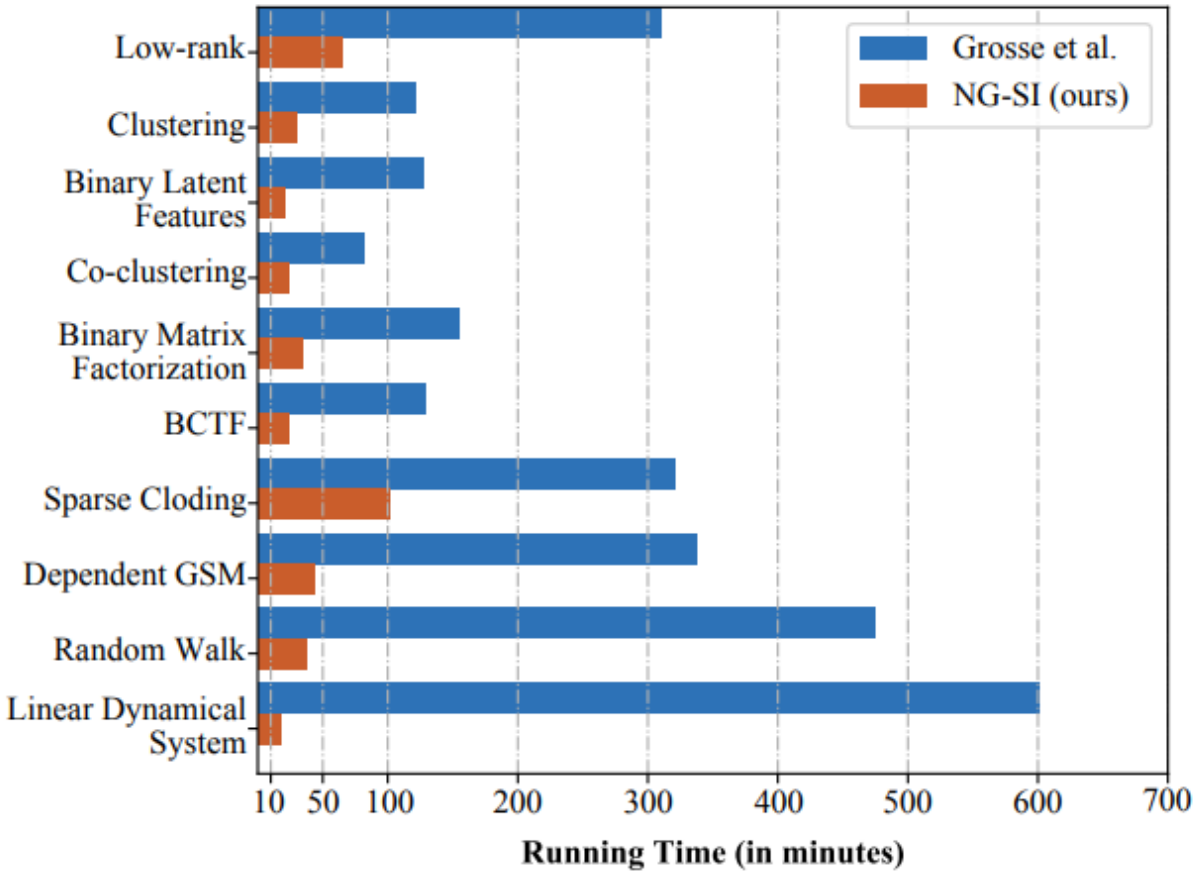
Train

Test

Experiment – Running Time Comparison



Experiment – Running Time Comparison



Speed Up: 3×∼39×.

Conclusion

Conclusion

- Combine search-based algorithms and data-driven models.

Conclusion

- Combine search-based algorithms and data-driven models.
- Use neural networks to guide layer-wise search.

Conclusion

- Combine search-based algorithms and data-driven models.
- Use neural networks to guide layer-wise search.

- Advantage:
  - Accurate: c.f. search-based algorithms.

Conclusion

- Combine search-based algorithms and data-driven models.
- Use neural networks to guide layer-wise search.

- Advantage:
  - Accurate: c.f. search-based algorithms.
  - Efficient: c.f. data-driven models.

Conclusion

- Combine search-based algorithms and data-driven models.
- Use neural networks to guide layer-wise search.

- Advantage:
  - Accurate: c.f. search-based algorithms.
  - Efficient: c.f. data-driven models.
  - Combinatorial generalization.

Conclusion

Poster #233

- Combine search-based algorithms and data-driven models.
- Use neural networks to guide layer-wise search.

- Advantage:
  - Accurate: c.f. search-based algorithms.
  - Efficient: c.f. data-driven models.
  - Combinatorial generalization.

Project Page