

EXTRACTING ACCURATE SINGULAR VALUES FROM APPROXIMATE SUBSPACES



Mathematical
Institute

LORENZO LAZZARINO, HUSSAM AL DAAS, YUJI NAKATSUKASA

Mathematical Institute - University of Oxford

Computational Mathematics Theme - STFC UKRI

ENUMATH 2025, 1st September 2025



Oxford
Mathematics



EXTRACTING ACCURATE SINGULAR VALUES FROM APPROXIMATE SUBSPACES

- 1 PROBLEM SETTING
- 2 CLASSICAL APPROACHES
- 3 TECHNIQUES FROM (RANDOMIZED) LOW-RANK APPROXIMATIONS
- 4 EXTRACTING SINGULAR VALUES WITH GN
- 5 ANALYSIS AND COMPARISON

PROBLEM SETTING

1

PROBLEM SETTING

$$A = U\Sigma V^*$$

Given \tilde{U} and/or \tilde{V} (orthonormal)
approximations of the leading singular
subspaces of A

$$n \begin{bmatrix} r \\ \tilde{V} \end{bmatrix}, \quad m \begin{bmatrix} r + \ell \\ \tilde{U} \end{bmatrix}$$

AIM: Approximate the leading singular values
 $\{\sigma_i(A)\}_{i=1}^r$

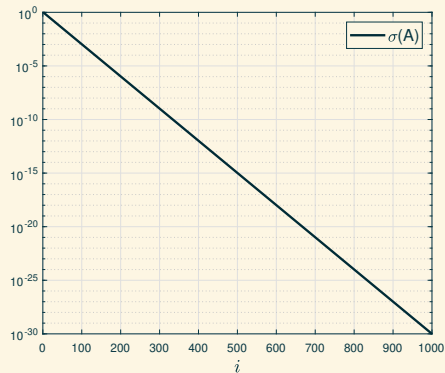
PROBLEM SETTING

$$A = U\Sigma V^*$$

Given \tilde{U} and/or \tilde{V} (orthonormal)
approximations of the leading singular
subspaces of A

$$n \begin{bmatrix} r \\ \tilde{V} \end{bmatrix}, \quad m \begin{bmatrix} r + \ell \\ \tilde{U} \end{bmatrix}$$

AIM: Approximate the leading singular values
 $\{\sigma_i(A)\}_{i=1}^r$



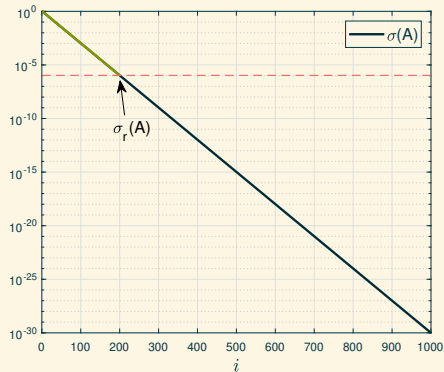
PROBLEM SETTING

$$A = U\Sigma V^*$$

Given \tilde{U} and/or \tilde{V} (orthonormal) approximations of the leading singular subspaces of A

$$n \begin{bmatrix} r \\ \tilde{V} \end{bmatrix}, \quad m \begin{bmatrix} r + \ell \\ \tilde{U} \end{bmatrix}$$

AIM: Approximate the leading singular values $\{\sigma_i(A)\}_{i=1}^r$



CLASSICAL APPROACHES

2

CLASSICAL APPROACHES > *Rayleigh Ritz and (one-sided) SVD approximations*

Rayleigh Ritz (RR)

$$\sigma_i(A) \approx \sigma_i(\tilde{U}^* A \tilde{V}) =: \sigma_i(A_{RR, \tilde{V}, \tilde{U}})$$



(Dax, 2012)
(Saad, 2011)
(Xin-guo, 1992)

CLASSICAL APPROACHES > *Rayleigh Ritz and (one-sided) SVD approximations*

Rayleigh Ritz (RR)

$$\sigma_i(A) \approx \sigma_i(\tilde{U}^* A \tilde{V}) =: \sigma_i(A_{RR, \tilde{V}, \tilde{U}})$$



(Dax, 2012)
(Saad, 2011)
(Xin-guo, 1992)

- ▶ $N_r + \mathcal{O}(mr^2) + \mathcal{O}(r^3)$
- ▶ Single-pass
- ▶ 1 multiplication by A

CLASSICAL APPROACHES › *Rayleigh Ritz and (one-sided) SVD approximations*

Rayleigh Ritz (RR)

$$\sigma_i(A) \approx \sigma_i(\tilde{U}^* A \tilde{V}) =: \sigma_i(A_{RR, \tilde{V}, \tilde{U}})$$



(Dax, 2012)
 (Saad, 2011)
 (Xin-guo, 1992)

- ▶ $N_r + \mathcal{O}(mr^2) + \mathcal{O}(r^3)$
- ▶ Single-pass
- ▶ 1 multiplication by A

$$Q_1 = [\tilde{U} \quad \tilde{U}_\perp], \quad Q_2 = [\tilde{V} \quad \tilde{V}_\perp]$$

CLASSICAL APPROACHES › *Rayleigh Ritz and (one-sided) SVD approximations*

Rayleigh Ritz (RR)

$$\sigma_i(A) \approx \sigma_i(\tilde{U}^* A \tilde{V}) =: \sigma_i(A_{RR, \tilde{V}, \tilde{U}})$$



(Dax, 2012)
 (Saad, 2011)
 (Xin-guo, 1992)

- ▶ $N_r + \mathcal{O}(mr^2) + \mathcal{O}(r^3)$
- ▶ Single-pass
- ▶ 1 multiplication by A

$$Q_1 = [\tilde{U} \quad \tilde{U}_\perp], \quad Q_2 = [\tilde{V} \quad \tilde{V}_\perp]$$

$$\bar{A} = Q_1^* A Q_2$$

$$\begin{aligned} \sigma_i(A_{RR, \tilde{V}, \tilde{U}}) &= \sigma_i(\bar{A}_{RR, \begin{bmatrix} I_r \\ 0 \end{bmatrix}, \begin{bmatrix} I_{r+\ell} \\ 0 \end{bmatrix}}) \\ &= \sigma_i(\bar{A}_{11}) = \sigma_i\left(\begin{bmatrix} \bar{A}_{11} & 0 \\ 0 & 0 \end{bmatrix}\right) \end{aligned}$$

CLASSICAL APPROACHES › *Rayleigh Ritz and (one-sided) SVD approximations*

Rayleigh Ritz (RR)

$$\sigma_i(A) \approx \sigma_i(\tilde{U}^* A \tilde{V}) =: \sigma_i(A_{RR, \tilde{V}, \tilde{U}})$$



(Dax, 2012)
 (Saad, 2011)
 (Xin-guo, 1992)

(one-sided) SVD approximations

$$\sigma_i(A) \approx \sigma_i(A \tilde{V}) =: \sigma_i(A_{SVD, \tilde{V}})$$

- ▶ $N_r + \mathcal{O}(mr^2) + \mathcal{O}(r^3)$
- ▶ Single-pass
- ▶ 1 multiplication by A

$$Q_1 = [\tilde{U} \quad \tilde{U}_\perp], \quad Q_2 = [\tilde{V} \quad \tilde{V}_\perp]$$

$$\bar{A} = Q_1^* A Q_2$$

$$\begin{aligned} \sigma_i(A_{RR, \tilde{V}, \tilde{U}}) &= \sigma_i(\bar{A}_{RR, \begin{bmatrix} I_r \\ 0 \end{bmatrix}, \begin{bmatrix} I_{r+\ell} \\ 0 \end{bmatrix}}) \\ &= \sigma_i(\bar{A}_{11}) = \sigma_i\left(\begin{bmatrix} \bar{A}_{11} & 0 \\ 0 & 0 \end{bmatrix}\right) \end{aligned}$$

CLASSICAL APPROACHES › Rayleigh Ritz and (one-sided) SVD approximations

Rayleigh Ritz (RR)

$$\sigma_i(A) \approx \sigma_i(\tilde{U}^* A \tilde{V}) =: \sigma_i(A_{RR, \tilde{V}, \tilde{U}})$$



(Dax, 2012)
(Saad, 2011)
(Xin-guo, 1992)

(one-sided) SVD approximations

$$\sigma_i(A) \approx \sigma_i(A \tilde{V}) =: \sigma_i(A_{SVD, \tilde{V}})$$

- ▶ $N_r + \mathcal{O}(mr^2) + \mathcal{O}(r^3)$
- ▶ Single-pass
- ▶ 1 multiplication by A

- ▶ $N_r + \mathcal{O}(mr^2)$
- ▶ Single-pass
- ▶ 1 multiplication by A

$$Q_1 = [\tilde{U} \quad \tilde{U}_\perp], \quad Q_2 = [\tilde{V} \quad \tilde{V}_\perp]$$

$$\bar{A} = Q_1^* A Q_2$$

$$\begin{aligned} \sigma_i(A_{RR, \tilde{V}, \tilde{U}}) &= \sigma_i(\bar{A}_{RR, \begin{bmatrix} I_r \\ 0 \end{bmatrix}, \begin{bmatrix} I_{r+\ell} \\ 0 \end{bmatrix}}) \\ &= \sigma_i(\bar{A}_{11}) = \sigma_i\left(\begin{bmatrix} \bar{A}_{11} & 0 \\ 0 & 0 \end{bmatrix}\right) \end{aligned}$$

CLASSICAL APPROACHES > Rayleigh Ritz and (one-sided) SVD approximations

Rayleigh Ritz (RR)

$$\sigma_i(A) \approx \sigma_i(\tilde{U}^* A \tilde{V}) =: \sigma_i(A_{RR, \tilde{V}, \tilde{U}})$$



(Dax, 2012)
(Saad, 2011)
(Xin-guo, 1992)

- ▶ $N_r + \mathcal{O}(mr^2) + \mathcal{O}(r^3)$
- ▶ Single-pass
- ▶ 1 multiplication by A

(one-sided) SVD approximations

$$\sigma_i(A) \approx \sigma_i(A \tilde{V}) =: \sigma_i(A_{SVD, \tilde{V}})$$

- ▶ $N_r + \mathcal{O}(mr^2)$
- ▶ Single-pass
- ▶ 1 multiplication by A

$$Q_1 = [\tilde{U} \quad \tilde{U}_\perp], \quad Q_2 = [\tilde{V} \quad \tilde{V}_\perp]$$

$$\bar{A} = Q_1^* A Q_2$$

$$\begin{aligned} \sigma_i(A_{RR, \tilde{V}, \tilde{U}}) &= \sigma_i(\bar{A}_{RR, \begin{bmatrix} I_r \\ 0 \end{bmatrix}, \begin{bmatrix} I_{r+\ell} \\ 0 \end{bmatrix}}) \\ &= \sigma_i(\bar{A}_{11}) = \sigma_i\left(\begin{bmatrix} \bar{A}_{11} & 0 \\ 0 & 0 \end{bmatrix}\right) \end{aligned}$$

$$\tilde{A} = A Q_2 = [\tilde{A}_1 \quad \tilde{A}_2]$$

$$\begin{aligned} \sigma_i(A_{SVD, \tilde{V}}) &= \sigma_i(\tilde{A}_{SVD, \begin{bmatrix} I_r \\ 0 \end{bmatrix}}) \\ &= \sigma_i([\tilde{A}_1 \quad 0]) \end{aligned}$$

CLASSICAL APPROACHES › Rayleigh Ritz and (one-sided) SVD approximations › Accuracy

Rayleigh Ritz (RR)

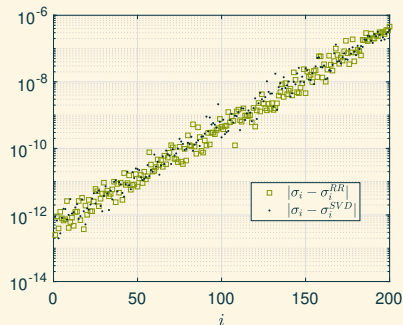
$$\sigma_i(A) \approx \sigma_i(\tilde{U}^* A \tilde{V}) =: \sigma_i(A_{RR, \tilde{V}, \tilde{U}})$$



(Dax, 2012)
(Saad, 2011)
(Xin-guo, 1992)

(one-sided) SVD approximations

$$\sigma_i(A) \approx \sigma_i(A \tilde{V}) =: \sigma_i(A_{SVD, \tilde{V}})$$



CLASSICAL APPROACHES > Rayleigh Ritz and (one-sided) SVD approximations > Accuracy

Rayleigh Ritz (RR)

$$\sigma_i(A) \approx \sigma_i(\tilde{U}^* A \tilde{V}) =: \sigma_i(A_{RR, \tilde{V}, \tilde{U}})$$

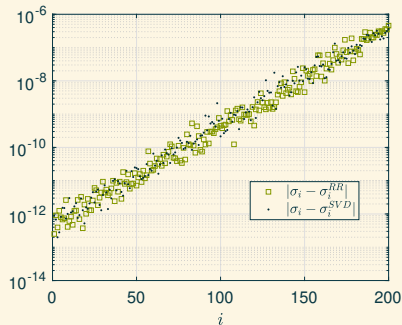


(Dax, 2012)
(Saad, 2011)
(Xin-guo, 1992)

(one-sided) SVD approximations

$$\sigma_i(A) \approx \sigma_i(A \tilde{V}) =: \sigma_i(A_{SVD, \tilde{V}})$$

Not bad...



CLASSICAL APPROACHES > Rayleigh Ritz and (one-sided) SVD approximations > Accuracy

Rayleigh Ritz (RR)

$$\sigma_i(A) \approx \sigma_i(\tilde{U}^* A \tilde{V}) =: \sigma_i(A_{RR, \tilde{V}, \tilde{U}})$$

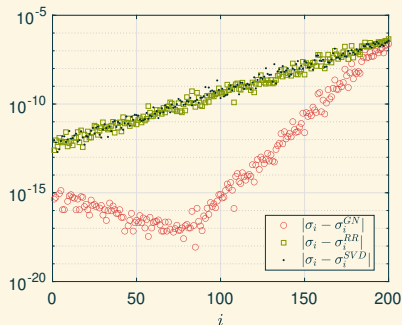


(Dax, 2012)
(Saad, 2011)
(Xin-guo, 1992)

(one-sided) SVD approximations

$$\sigma_i(A) \approx \sigma_i(A \tilde{V}) =: \sigma_i(A_{SVD, \tilde{V}})$$

Not bad...



BUT,
what if we could have this?

TECHNIQUES FROM (RANDOMIZED) LOW-RANK APPROXIMATIONS

3

RANDOMIZED SVD (HMT)

Randomized SVD

$$A \approx (A\Omega)(A\Omega)^\dagger A =: A_{HMT,\Omega}$$



(Clarkson, Woodruff, 2017)
(Halko, Martinsson, Tropp, 2011)
(Rokhlin, Szlam, Tygert, 2009)

1. Choose $\Omega \in \mathbb{R}^{n \times r}$
2. Sketch: $X = A\Omega$
3. $[Q, \sim] = \text{qr}(X, 0)$
4. $A_{HMT,\Omega} = Q(Q^*A)$

RANDOMIZED SVD (HMT)

Randomized SVD

$$A \approx (A\Omega)(A\Omega)^\dagger A =: A_{HMT,\Omega}$$



(Clarkson, Woodruff, 2017)
 (Halko, Martinsson, Tropp, 2011)
 (Rokhlin, Szlam, Tygert, 2009)

1. Choose $\Omega \in \mathbb{R}^{n \times r}$ 2. Sketch: $X = A\Omega$ 3. $[Q, \sim] = \text{qr}(X, 0)$ 4. $A_{HMT,\Omega} = Q(Q^* A)$

- $N_r + \mathcal{O}(mr^2) + \tilde{N}_r$
- Double-pass
- 2 multiplications by A

GENERALIZED NYSTRÖM APPROXIMATION

Generalized Nyström

$$A \approx A\Omega_1(\Omega_2^* A\Omega_1)^\dagger \Omega_2^* A =: A_{GN,\Omega_1,\Omega_2}$$



(Clarkson, Woodruff, 2009)
 (Nakatsukasa, 2020)
 (Woolfe, Liberty, Rokhlin, Tygert, 2008)

1. Choose $\Omega_1 \in \mathbb{R}^{n \times r}, \Omega_2 \in \mathbb{R}^{m \times (r+\ell)}$
2. Two-side Sketch: $X = A\Omega_1$ and $Y = \Omega_2^* A$
3. $[Q,R] = \text{qr}(Y\Omega_1, 0)$
4. $A_{GN,\Omega_1,\Omega_2} = (XR^{-1})(Q^* Y)$

GENERALIZED NYSTRÖM APPROXIMATION

Generalized Nyström

$$A \approx A\Omega_1(\Omega_2^* A\Omega_1)^\dagger \Omega_2^* A =: A_{GN,\Omega_1,\Omega_2}$$



(Clarkson, Woodruff, 2009)
 (Nakatsukasa, 2020)
 (Woolfe, Liberty, Rokhlin, Tygert, 2008)

1. Choose $\Omega_1 \in \mathbb{R}^{n \times r}, \Omega_2 \in \mathbb{R}^{m \times (r+\ell)}$
2. Two-side Sketch: $X = A\Omega_1$ and $Y = \Omega_2^* A$
3. $[Q,R] = \text{qr}(Y\Omega_1, 0)$
4. $A_{GN,\Omega_1,\Omega_2} = (XR^{-1})(Q^* Y)$

$$\triangleright N_{2r+\ell} + \mathcal{O}(r^3 + (m+n)r^2)$$

\triangleright Single-pass

\triangleright 2 multiplications by A

EXTRACTING SINGULAR VALUES WITH GN

4

GN APPROXIMATION AND EXTRACTING SINGULAR VALUES

Generalized Nyström

Given approximations \tilde{U} and \tilde{V} to the leading singular subspaces,

$$\sigma_i(A) \approx \sigma_i \left(A\tilde{V}(\tilde{U}^* A\tilde{V})^\dagger \tilde{U}^* A \right) =: \sigma_i^{GN}$$

$$\sigma_i \left(\begin{array}{c} \boxed{A\tilde{V}} \end{array} \begin{array}{c} \boxed{\tilde{U}^* A\tilde{V}}^\dagger \end{array} \begin{array}{c} \boxed{\tilde{U}^* A} \end{array} \right)$$

$N_{2r+\ell}$

GN APPROXIMATION AND EXTRACTING SINGULAR VALUES

Generalized Nyström

Given approximations \tilde{U} and \tilde{V} to the leading singular subspaces,

$$\sigma_i(A) \approx \sigma_i \left(A \tilde{V} (\tilde{U}^* A \tilde{V})^\dagger \tilde{U}^* A \right) =: \sigma_i^{GN}$$

$$\sigma_i \left(\begin{array}{c} \boxed{Q_L} \end{array} \begin{array}{c} \boxed{R_L} \end{array} \begin{array}{c} \boxed{\tilde{U}^* A \tilde{V}}^\dagger \end{array} \begin{array}{c} \boxed{R_R^*} \end{array} \begin{array}{c} \boxed{Q_R^*} \end{array} \right)$$

$$N_{2r+\ell} + \mathcal{O}((m+n)r^2)$$

GN APPROXIMATION AND EXTRACTING SINGULAR VALUES

Generalized Nyström

Given approximations \tilde{U} and \tilde{V} to the leading singular subspaces,

$$\sigma_i(A) \approx \sigma_i \left(A\tilde{V}(\tilde{U}^* A\tilde{V})^\dagger \tilde{U}^* A \right) =: \sigma_i^{GN}$$

$$\sigma_i \left(\begin{array}{|c|} \hline R_L \\ \hline \end{array} \begin{array}{|c|} \hline \tilde{U}^* A \tilde{V} \\ \hline \end{array}^\dagger \begin{array}{|c|} \hline R_R^* \\ \hline \end{array} \right)$$

$$N_{2r+\ell} + \mathcal{O}((m+n)r^2)$$

GN APPROXIMATION AND EXTRACTING SINGULAR VALUES

Generalized Nyström

Given approximations \tilde{U} and \tilde{V} to the leading singular subspaces,

$$\sigma_i(A) \approx \sigma_i \left(A \tilde{V} (\tilde{U}^* A \tilde{V})^\dagger \tilde{U}^* A \right) =: \sigma_i^{GN}$$

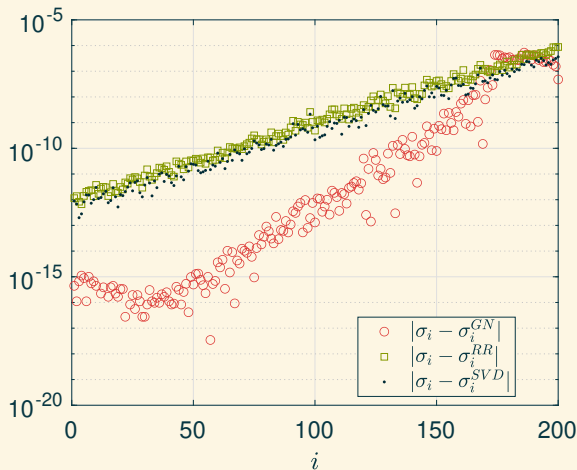
$$\sigma_i \left(\begin{array}{|c|} \hline R_L \\ \hline \end{array} \begin{array}{|c|} \hline R_p^\dagger \\ \hline \end{array} \begin{array}{|c|} \hline Q_p^* \\ \hline \end{array} \begin{array}{|c|} \hline R_R^* \\ \hline \end{array} \right)$$

$$N_{2r+\ell} + \mathcal{O}((m+n)r^2)$$

MOTIVATIONAL COMPARISON

Single-pass methods

- ▶ $\sigma_i^{SVD} = \sigma_i(A\tilde{V})$
- ▶ $\sigma_i^{RR} = \sigma_i(\tilde{U}^* A \tilde{V})$
- ▶ $\sigma_i^{GN} = \sigma_i(A\tilde{V}(\tilde{U}^* A \tilde{V})^\dagger \tilde{U}^* A)$



GN AND MATRIX PERTURBATION THEORY

GN and Orthogonal Transformations

Consider T_1 and T_2 orthogonal matrices, then

$$T_1^*(M_{GN, \tilde{V}, \tilde{U}})T_2 = (T_1^*MT_2)_{GN, T_2^*\tilde{V}, T_1^*\tilde{U}}$$

For any orthonormal \tilde{V} and \tilde{U} , we can:

1. Define $Q_1 = [\tilde{U} \quad \tilde{U}_\perp]$ $Q_2 = [\tilde{V} \quad \tilde{V}_\perp]$;
2. Consider the transformed matrix: $Q_1^*AQ_2$;
3. Consider the transformed GN approximation:

$$Q_1^*A_{GN, \tilde{V}, \tilde{U}}Q_2 = (Q_1^*AQ_2)_{GN, Q_2^*\tilde{V}, Q_1^*\tilde{U}} = (Q_1^*AQ_2)_{GN, \begin{bmatrix} I_r \\ 0 \end{bmatrix}, \begin{bmatrix} I_{r+\ell} \\ 0 \end{bmatrix}}.$$

GN AND MATRIX PERTURBATION THEORY

GN and Orthogonal Transformations

Consider T_1 and T_2 orthogonal matrices, then

$$T_1^*(M_{GN, \tilde{V}, \tilde{U}})T_2 = (T_1^*MT_2)_{GN, T_2^*\tilde{V}, T_1^*\tilde{U}}$$

For any orthonormal \tilde{V} and \tilde{U} , we can:

1. Define $Q_1 = [\tilde{U} \quad \tilde{U}_\perp]$ $Q_2 = [\tilde{V} \quad \tilde{V}_\perp]$;
2. Consider the transformed matrix: $Q_1^*AQ_2$;
3. Consider the transformed GN approximation:

$$Q_1^*A_{GN, \tilde{V}, \tilde{U}}Q_2 = (Q_1^*AQ_2)_{GN, Q_2^*\tilde{V}, Q_1^*\tilde{U}} = (Q_1^*AQ_2)_{GN, \begin{bmatrix} I_r \\ 0 \end{bmatrix}, \begin{bmatrix} I_{r+\ell} \\ 0 \end{bmatrix}}.$$

$$\rightarrow |\sigma_i(A) - \sigma_i(A_{GN, \tilde{V}, \tilde{U}})| = |\sigma_i(Q_1^*AQ_2) - \sigma_i((Q_1^*AQ_2)_{GN, \begin{bmatrix} I_r \\ 0 \end{bmatrix}, \begin{bmatrix} I_{r+\ell} \\ 0 \end{bmatrix}})|$$

GN AND MATRIX PERTURBATION THEORY > Express A_{GN} as a perturbation of the original matrix A

$$\tilde{V} := \begin{matrix} & r \\ r & \begin{bmatrix} I_r \\ - \\ 0 \end{bmatrix} \\ n-r & \end{matrix}, \quad \tilde{U} := \begin{matrix} & r+\ell \\ r+\ell & \begin{bmatrix} I_{r+\ell} \\ - \\ 0 \end{bmatrix} \\ m-(r+\ell) & \end{matrix}, \quad A := \begin{matrix} & r & & n-r \\ r+\ell & \begin{bmatrix} A_{11} \\ - \\ A_{21} \end{bmatrix} & \begin{bmatrix} A_{12} \\ - \\ A_{22} \end{bmatrix} \\ m-(r+\ell) & & & \end{matrix}$$



(Tropp, Webber, 2023)

$$A_{GN, \tilde{V}, \tilde{U}} = A\tilde{V}(\tilde{U}^* A \tilde{V})^\dagger \tilde{U}^* A$$

GN AND MATRIX PERTURBATION THEORY > Express A_{GN} as a perturbation of the original matrix A

$$\tilde{V} := \begin{matrix} & r \\ r & \begin{bmatrix} I_r \\ - \\ 0 \end{bmatrix} \\ n-r & \end{matrix}, \quad \tilde{U} := \begin{matrix} & r+\ell \\ r+\ell & \begin{bmatrix} I_{r+\ell} \\ - \\ 0 \end{bmatrix} \\ m-(r+\ell) & \end{matrix}, \quad A := \begin{matrix} & r & & n-r \\ r+\ell & \begin{bmatrix} A_{11} \\ - \\ A_{21} \end{bmatrix} & \begin{bmatrix} A_{12} \\ - \\ A_{22} \end{bmatrix} \\ m-(r+\ell) & & & \end{matrix}$$

$$A_{GN, \tilde{V}, \tilde{U}} = \begin{bmatrix} A_{11} \\ - \\ A_{21} \end{bmatrix} (\tilde{U}^* A \tilde{V})^\dagger \tilde{U}^* A$$

GN AND MATRIX PERTURBATION THEORY > Express A_{GN} as a perturbation of the original matrix A

$$\tilde{V} := \begin{matrix} & r \\ r & \begin{bmatrix} I_r \\ - \\ 0 \end{bmatrix} \\ n-r & \end{matrix}, \quad \tilde{U} := \begin{matrix} & r+\ell \\ r+\ell & \begin{bmatrix} I_{r+\ell} \\ - \\ 0 \end{bmatrix} \\ m-(r+\ell) & \end{matrix}, \quad A := \begin{matrix} & r & & n-r \\ r+\ell & \begin{bmatrix} A_{11} \\ - \\ A_{21} \end{bmatrix} & | & \begin{bmatrix} A_{12} \\ - \\ A_{22} \end{bmatrix} \\ m-(r+\ell) & & & \end{matrix}$$

$$A_{GN, \tilde{V}, \tilde{U}} = \begin{bmatrix} A_{11} \\ - \\ A_{21} \end{bmatrix} (\tilde{U}^* A \tilde{V})^\dagger \begin{bmatrix} A_{11} & | & A_{12} \end{bmatrix}$$

GN AND MATRIX PERTURBATION THEORY > Express A_{GN} as a perturbation of the original matrix A

$$\tilde{V} := \begin{matrix} r \\ n-r \end{matrix} \begin{bmatrix} I_r \\ 0 \end{bmatrix}, \quad \tilde{U} := \begin{matrix} r+\ell \\ m-(r+\ell) \end{matrix} \begin{bmatrix} I_{r+\ell} \\ 0 \end{bmatrix}, \quad A := \begin{matrix} r+\ell \\ m-(r+\ell) \end{matrix} \left[\begin{array}{c|c} \begin{matrix} r \\ r+\ell \end{matrix} \begin{bmatrix} A_{11} \\ - \end{bmatrix} & \begin{matrix} n-r \\ - \end{matrix} \begin{bmatrix} A_{12} \\ - \end{bmatrix} \\ \hline \begin{matrix} A_{21} \\ A_{22} \end{matrix} \end{array} \right]$$

$$A_{GN, \tilde{V}, \tilde{U}} = \begin{bmatrix} A_{11} \\ - \\ A_{21} \end{bmatrix} (A_{11})^\dagger \left[\begin{array}{c|c} A_{11} & A_{12} \end{array} \right]$$

GN AND MATRIX PERTURBATION THEORY > Express A_{GN} as a perturbation of the original matrix A

$$\tilde{V} := \begin{bmatrix} r & r \\ n-r & 0 \end{bmatrix}, \quad \tilde{U} := \begin{bmatrix} r+\ell & r+\ell \\ m-(r+\ell) & 0 \end{bmatrix}, \quad A := \begin{bmatrix} r & n-r \\ r+\ell & A_{11} & A_{12} \\ m-(r+\ell) & A_{21} & A_{22} \end{bmatrix}$$

$$MM^\dagger M = M$$

$$A_{GN, \tilde{V}, \tilde{U}} = \begin{bmatrix} A_{11} \\ - \\ A_{21} \end{bmatrix} (A_{11})^\dagger \begin{bmatrix} A_{11} & A_{12} \end{bmatrix} = \begin{bmatrix} A_{11} A_{11}^\dagger A_{11} & A_{11} A_{11}^\dagger A_{12} \\ A_{21} A_{11}^\dagger A_{11} & A_{21} A_{11}^\dagger A_{12} \end{bmatrix}$$

GN AND MATRIX PERTURBATION THEORY > Express A_{GN} as a perturbation of the original matrix A

$$\tilde{V} := \begin{matrix} r \\ n-r \end{matrix} \begin{bmatrix} I_r \\ 0 \end{bmatrix}, \quad \tilde{U} := \begin{matrix} r+\ell \\ m-(r+\ell) \end{matrix} \begin{bmatrix} I_{r+\ell} \\ 0 \end{bmatrix}, \quad A := \begin{matrix} r & n-r \\ r+\ell & \end{matrix} \left[\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array} \right]$$

M has linearly independent columns
 $\Rightarrow M^\dagger M = M^{-1}M = M$

$$A_{GN, \tilde{V}, \tilde{U}} = \begin{bmatrix} A_{11} \\ - \\ A_{21} \end{bmatrix} (A_{11})^\dagger \left[\begin{array}{c|c} A_{11} & A_{12} \end{array} \right] = \left[\begin{array}{c|c} \overbrace{A_{11} A_{11}^\dagger A_{11}}^{= A_{11}} & A_{11} A_{11}^\dagger A_{12} \\ \hline A_{21} A_{11}^\dagger A_{11} & A_{21} A_{11}^\dagger A_{12} \end{array} \right]$$

GN AND MATRIX PERTURBATION THEORY > Express A_{GN} as a perturbation of the original matrix A

$$\tilde{V} := \begin{matrix} r \\ n-r \end{matrix} \begin{bmatrix} I_r \\ 0 \end{bmatrix}, \quad \tilde{U} := \begin{matrix} r+\ell \\ m-(r+\ell) \end{matrix} \begin{bmatrix} I_{r+\ell} \\ 0 \end{bmatrix}, \quad A := \begin{matrix} r & n-r \\ r+\ell & \\ m-(r+\ell) & \end{matrix} \left[\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array} \right]$$

$$A_{GN, \tilde{V}, \tilde{U}} = \begin{bmatrix} A_{11} \\ - \\ A_{21} \end{bmatrix} (A_{11})^\dagger \left[A_{11} \mid A_{12} \right] = \left[\begin{array}{c|c} A_{11} & A_{11} A_{11}^\dagger A_{12} \\ \hline \underbrace{A_{21} A_{11}^\dagger A_{11}}_{= A_{21}} & A_{21} A_{11}^\dagger A_{12} \end{array} \right]$$

GN AND MATRIX PERTURBATION THEORY > Express A_{GN} as a perturbation of the original matrix A

$$\tilde{V} := \begin{matrix} r \\ n-r \end{matrix} \begin{bmatrix} I_r \\ 0 \end{bmatrix}, \quad \tilde{U} := \begin{matrix} r+\ell \\ m-(r+\ell) \end{matrix} \begin{bmatrix} I_{r+\ell} \\ 0 \end{bmatrix}, \quad A := \begin{matrix} r & n-r \\ r+\ell & \end{matrix} \left[\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array} \right]$$

$$A_{GN, \tilde{V}, \tilde{U}} = A - \left[\begin{array}{c|c} 0 & A_{12} - A_{11}A_{11}^\dagger A_{12} \\ \hline 0 & A_{22} - A_{21}A_{11}^\dagger A_{12} \end{array} \right] =: A - E_{GN}$$

GN AND MATRIX PERTURBATION THEORY > Express A_{GN} as a perturbation of the original matrix A

$$\tilde{V} := \begin{matrix} r \\ n-r \end{matrix} \begin{bmatrix} I_r \\ 0 \end{bmatrix}, \quad \tilde{U} := \begin{matrix} r \\ m-r \end{matrix} \begin{bmatrix} I_r \\ 0 \end{bmatrix}, \quad A := \begin{matrix} r & n-r \\ m-r \end{matrix} \left[\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array} \right]$$

No-oversample ($\ell = 0$)
 $\rightarrow A_{12} - A_{11}A_{11}^\dagger A_{12} = 0$, but change of
 block sizes!

$$A_{GN, \tilde{V}, \tilde{U}} = A - \left[\begin{array}{c|c} 0 & 0 \\ \hline 0 & A_{22} - A_{21}A_{11}^\dagger A_{12} \end{array} \right] =: A - E_{GN}$$

GN AND MATRIX PERTURBATION THEORY > *Weyl's bound*

Weyl's Theorem

For any matrix M we have that

$$|\sigma_i(M) - \sigma_i(M + E)| \leq \|E\|_2$$



Cor. 7.3.5 (Horn, Johnson, 2012)
Cor. I.4.31 (Stewart, 1998)

Weyl's Theorem

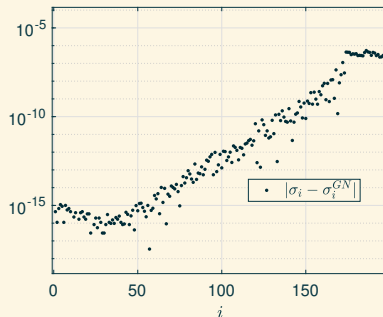
For any matrix M we have that

$$|\sigma_i(M) - \sigma_i(M + E)| \leq \|E\|_2$$



Cor. 7.3.5 (Horn, Johnson, 2012)
Cor. 1.4.31 (Stewart, 1998)

$$|\sigma_i(A) - \sigma_i(A_{GN, \tilde{V}, \tilde{U}})|$$



Weyl's Theorem

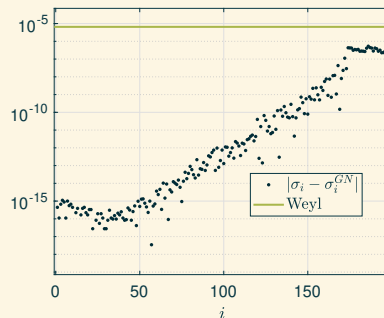
For any matrix M we have that

$$|\sigma_i(M) - \sigma_i(M + E)| \leq \|E\|_2$$



Cor. 7.3.5 (Horn, Johnson, 2012)
Cor. 1.4.31 (Stewart, 1998)

$$|\sigma_i(A) - \sigma_i(A_{GN, \tilde{V}, \tilde{U}})| \leq \|E_{GN}\|_2$$



ANALYSIS AND COMPARISON

5

RESULT ON SYMMETRIC MATRICES

Consider the $n \times n$ symmetric matrices

$$H := \begin{bmatrix} H_{11} & H_{21}^* \\ H_{21} & H_{22} \end{bmatrix}, \quad \hat{H} := H + \begin{bmatrix} E_{11} & E_{21}^* \\ E_{21} & E_{22} \end{bmatrix} =: H + E.$$



Theorem 3.2 (Nakatsukasa, 2012)

RESULT ON SYMMETRIC MATRICES

Consider the $n \times n$ symmetric matrices

$$H := \begin{bmatrix} H_{11} & H_{21}^* \\ H_{21} & H_{22} \end{bmatrix}, \quad \hat{H} := H + \begin{bmatrix} E_{11} & E_{21}^* \\ E_{21} & E_{22} \end{bmatrix} =: H + E.$$



Theorem 3.2 (Nakatsukasa, 2012)

Define

$$\tau_i = \left(\frac{\|H_{21}\|_2 + \|E_{21}\|_2}{\min_j |\lambda_i(H) - \lambda_j(H_{22})| - 2\|E\|_2} \right).$$

Then, for each i , if $\tau_i > 0$, then

$$|\lambda_i(H) - \lambda_i(\hat{H})| \leq \|E_{11}\|_2 + 2\|E_{21}\|_2\tau_i + \|E_{22}\|_2\tau_i^2,$$

RESULT ON SYMMETRIC MATRICES

Consider the $n \times n$ symmetric matrices

$$H := \begin{bmatrix} H_{11} & H_{21}^* \\ H_{21} & H_{22} \end{bmatrix}, \quad \hat{H} := H + \begin{bmatrix} E_{11} & E_{21}^* \\ E_{21} & E_{22} \end{bmatrix} =: H + E.$$



Theorem 3.2 (Nakatsukasa, 2012)

Define

$$\tau_i = \left(\frac{\|H_{21}\|_2 + \|E_{21}\|_2}{\min_j |\lambda_i(H) - \lambda_j(H_{22})| - 2\|E\|_2} \right).$$

Then, for each i , if $\tau_i > 0$, then

$$|\lambda_i(H) - \lambda_i(\hat{H})| \leq \|E_{11}\|_2 + 2\|E_{21}\|_2\tau_i + \|E_{22}\|_2\tau_i^2,$$

- ▶ $\tau_i < 1$ necessary to be better than Weyl
- ▶ If $\|E_{11}\|_2 \ll \|E\|_2$ and λ_i is far from the spectrum of H_{22} then $\tau_i \ll 1$
- ▶ If $E_{11} = E_{21} = 0$ and H_{21} is small, then λ_i is particularly insensitive to the perturbation E_{22}
 \rightarrow bound proportional to $\|E_{22}\|_2\|H_{21}\|_2^2$

FROM THE SYMMETRIC TO THE GENERAL RESULT

General case



Transform to symmetric



Obtain necessary
structure



Apply symmetric Result



Transform back



General Result

Generalize (Nakatsukasa, 2012) to the 2×2 block matrix:

$$G := \begin{bmatrix} G_1 & B \\ C & G_2 \end{bmatrix},$$

and its perturbation:

$$\hat{G} := G + \begin{bmatrix} F_{11} & F_{12} \\ F_{21} & F_{22} \end{bmatrix} =: G + F.$$

Strategy: Use a technique in (Li, Li, 2005)

FROM THE SYMMETRIC TO THE GENERAL RESULT

General case



Transform to symmetric



Obtain necessary
structure



Apply symmetric Result



Transform back



General Result



Thm. 7.3.3 (Horn, Johnson, 2012)

Thm. I.4.2 (Stewart, Sun, 1990)

Jordan-Wielandt (JW) Theorem

Let $\{\sigma_i(M)\}_{i=1}^n$ be the singular values of a matrix $M \in \mathbb{C}^{m \times n}$, with $m \geq n$. Then, the symmetric matrix

$$\begin{bmatrix} 0 & M \\ M^* & 0 \end{bmatrix} \quad (1)$$

has eigenvalues $\pm\sigma_1(M), \dots, \pm\sigma_n(M)$ and $m - n$ zeros eigenvalues.

FROM THE SYMMETRIC TO THE GENERAL RESULT



Thm. 7.3.3 (Horn, Johnson, 2012)
 Thm. I.4.2 (Stewart, Sun, 1990)

General case



Transform to symmetric



Obtain necessary
structure



Apply symmetric Result



Transform back



General Result

Jordan-Wielandt (JW) Theorem

Let $\{\sigma_i(M)\}_{i=1}^n$ be the singular values of a matrix $M \in \mathbb{C}^{m \times n}$, with $m \geq n$. Then, the symmetric matrix

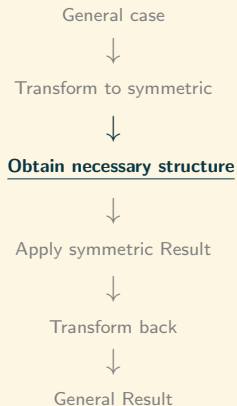
$$\begin{bmatrix} 0 & M \\ M^* & 0 \end{bmatrix} \quad (1)$$

has eigenvalues $\pm\sigma_1(M), \dots, \pm\sigma_n(M)$ and $m - n$ zeros eigenvalues.

$$G \rightarrow G_{JW} := \left[\begin{array}{c|c} 0 & G \\ \hline G^* & 0 \end{array} \right] = \left[\begin{array}{cc|cc} 0 & 0 & G_1 & B \\ 0 & 0 & C & G_2 \\ \hline G_1^* & C^* & 0 & 0 \\ B^* & G_2^* & 0 & 0 \end{array} \right]$$

FROM THE SYMMETRIC TO THE GENERAL RESULT

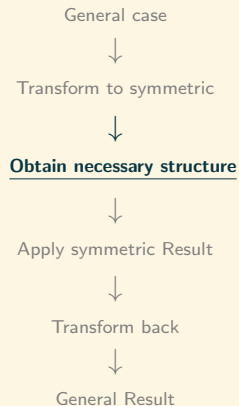
Obtain a matrix similar to G_{JW} suitable for (Nakatsukasa, 2012) and with blocks reasonably related to the blocks of G



$$\left[\begin{array}{cc|cc} 0 & 0 & G_1 & B \\ 0 & 0 & C & G_2 \\ \hline & & & \\ G_1^* & C^* & 0 & 0 \\ B^* & G_2^* & 0 & 0 \end{array} \right]$$

FROM THE SYMMETRIC TO THE GENERAL RESULT

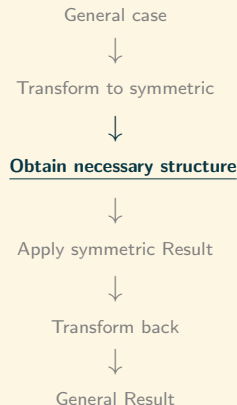
Obtain a matrix similar to G_{JW} suitable for (Nakatsukasa, 2012) and with blocks reasonably related to the blocks of G



$$\left[\begin{array}{cc|cc} 0 & 0 & G_1 & B \\ 0 & 0 & C & G_2 \\ \hline - & - & - & - \\ G_1^* & C^* & 0 & 0 \\ B^* & G_2^* & 0 & 0 \end{array} \right]$$

FROM THE SYMMETRIC TO THE GENERAL RESULT

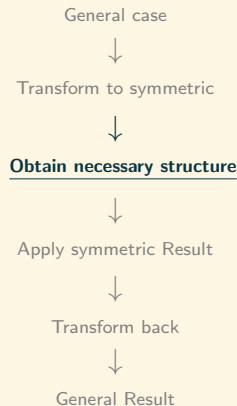
Obtain a matrix similar to G_{JW} suitable for (Nakatsukasa, 2012) and with blocks reasonably related to the blocks of G



$$\left[\begin{array}{cc|cc} 0 & 0 & G_1 & B \\ G_1^* & C^* & 0 & 0 \\ \hline 0 & 0 & C & G_2 \\ B^* & G_2^* & 0 & 0 \end{array} \right]$$

FROM THE SYMMETRIC TO THE GENERAL RESULT

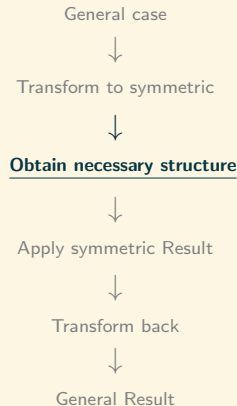
Obtain a matrix similar to G_{JW} suitable for (Nakatsukasa, 2012) and with blocks reasonably related to the blocks of G



$$\left[\begin{array}{cc|cc} 0 & 0 & G_1 & B \\ G_1^* & C^* & 0 & 0 \\ \hline 0 & 0 & C & G_2 \\ B^* & G_2^* & 0 & 0 \end{array} \right]$$

FROM THE SYMMETRIC TO THE GENERAL RESULT

Obtain a matrix similar to G_{JW} suitable for (Nakatsukasa, 2012) and with blocks reasonably related to the blocks of G



$$\left[\begin{array}{cc|cc} 0 & G_1 & 0 & B \\ G_1^* & 0 & C^* & 0 \\ \hline - & - & - & - \\ 0 & C & 0 & G_2 \\ B^* & 0 & G_2^* & 0 \end{array} \right] =: G_p$$

Note: $\lambda_i(G_p) = \lambda_i(G_{JW}) \stackrel{JW}{=} \pm \sigma_i(G)$

FROM THE SYMMETRIC TO THE GENERAL RESULT

General case



Transform to symmetric



Obtain necessary structure



Apply symmetric Result



Transform back



General Result

Obtain a matrix similar to G_{JW} suitable for (Nakatsukasa, 2012) and with blocks reasonably related to the blocks of G

$$G_p = \left[\begin{array}{cc|cc} 0 & G_1 & 0 & B \\ G_1^* & 0 & C^* & 0 \\ \hline 0 & C & 0 & G_2 \\ B^* & 0 & G_2^* & 0 \end{array} \right]$$

$$\hat{G}_p = G_p + \left[\begin{array}{cc|cc} 0 & F_{11} & 0 & F_{12} \\ F_{11}^* & 0 & F_{21}^* & 0 \\ \hline 0 & F_{21} & 0 & F_{22} \\ F_{12}^* & 0 & F_{22}^* & 0 \end{array} \right] =: G_p + F_p.$$

FROM THE SYMMETRIC TO THE GENERAL RESULT

General case



Transform to symmetric



Obtain necessary
structure



Apply symmetric Result



Transform back



General Result

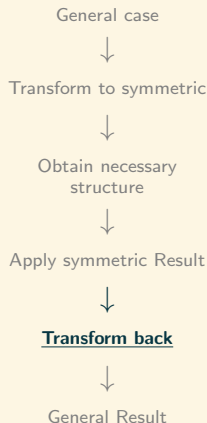
Define

$$\tau_i = \left(\frac{\left\| \begin{bmatrix} 0 & C \\ B^* & 0 \end{bmatrix} \right\|_2 + \left\| \begin{bmatrix} 0 & F_{21} \\ F_{12}^* & 0 \end{bmatrix} \right\|_2}{\min_j |\lambda_i - \lambda_j| \left(\left\| \begin{bmatrix} 0 & G_2 \\ G_2^* & 0 \end{bmatrix} \right\|_2 - 2 \|F_p\|_2 \right)} \right).$$

Then, for each i , if $\tau_i > 0$:

$$|\lambda_i(G_p) - \lambda_i(\hat{G}_p)| \leq \left\| \begin{bmatrix} 0 & F_{11} \\ F_{11}^* & 0 \end{bmatrix} \right\|_2 + 2 \left\| \begin{bmatrix} 0 & F_{21} \\ F_{12}^* & 0 \end{bmatrix} \right\|_2 \tau_i + \left\| \begin{bmatrix} 0 & F_{22} \\ F_{22}^* & 0 \end{bmatrix} \right\|_2 \tau_i^2,$$

FROM THE SYMMETRIC TO THE GENERAL RESULT



$$\left\| \begin{bmatrix} 0 & M_1 \\ M_2 & 0 \end{bmatrix} \right\|_2 = \max\{\|M_1\|_2, \|M_2\|_2\};$$

▶ Jordan-Wielandt theorem

$$\implies |\lambda_i(G_p) - \lambda_i(\hat{G}_p)| = |\sigma_i(G) - \sigma_i(\hat{G})|,$$

for $i = 1, \dots, n$;

▶ By Jordan-Wielandt theorem and by construction of F_p :

$$\|F_p\|_2 = \|F\|_2$$

FROM THE SYMMETRIC TO THE GENERAL RESULT > Generalization of (Nakatsukasa, 2012)

General case



Transform to symmetric



Obtain necessary
structure



Apply symmetric Result



Transform back



General Result



Theorem 4.1 (L., Al Daas, Nakatsukasa, 2024)

Consider the matrices

$$G := \begin{bmatrix} G_1 & B \\ C & G_2 \end{bmatrix}, \quad \hat{G} := G + \begin{bmatrix} F_{11} & F_{12} \\ F_{21} & F_{22} \end{bmatrix} =: G + F,$$

and define

$$\tau_i = \left(\frac{\max\{\|B\|_2, \|C\|_2\} + \max\{\|F_{12}\|_2, \|F_{21}\|_2\}}{\min_j |\sigma_i(G) - \sigma_j(G_2)| - 2\|F\|_2} \right).$$

Then, for each i , if $\tau_i > 0$, then

$$|\sigma_i(G) - \sigma_i(\hat{G})| \leq \|F_{11}\|_2 + 2 \max\{\|F_{12}\|_2, \|F_{21}\|_2\} \tau_i + \|F_{22}\|_2 \tau_i^2,$$

FROM THE SYMMETRIC TO THE GENERAL RESULT > Generalization of (Nakatsukasa, 2012)

General case



Transform to symmetric



Obtain necessary
structure



Apply symmetric Result



Transform back



General Result



Theorem 4.1 (L., Al Daas, Nakatsukasa, 2024)

Consider the matrices

$$G := \begin{bmatrix} G_1 & B \\ C & G_2 \end{bmatrix}, \quad \hat{G} := G + \begin{bmatrix} F_{11} & F_{12} \\ F_{21} & F_{22} \end{bmatrix} =: G + F,$$

and define

$$\tau_i = \left(\frac{\max\{\|B\|_2, \|C\|_2\} + \max\{\|F_{12}\|_2, \|F_{21}\|_2\}}{\min_j |\sigma_i(G) - \sigma_j(G_2)| - 2\|F\|_2} \right).$$

Then, for each i , if $\tau_i > 0$, then

$$|\sigma_i(G) - \sigma_i(\hat{G})| \leq \|F_{11}\|_2 + 2 \max\{\|F_{12}\|_2, \|F_{21}\|_2\} \tau_i + \|F_{22}\|_2 \tau_i^2,$$

- **Generalization to Block Tridiagonal:** A Singular Value is insensitive to blockwise perturbation if it is well-separated from the spectrum of the diagonal blocks near the perturbed blocks.

BOUND ON GN APPROXIMATION ERROR \triangleright Derivation

- $A, \tilde{V}, \tilde{U} \rightarrow A_{GN} = A\tilde{V}(\tilde{U}^* A\tilde{V})^\dagger \tilde{U}^* A$

- Define

$$\bar{A} = [\tilde{U} \ \tilde{U}_\perp]^* A [\tilde{V} \ \tilde{V}_\perp], \quad \bar{A}_{GN} = \left([\tilde{U} \ \tilde{U}_\perp]^* A [\tilde{V} \ \tilde{V}_\perp] \right)_{GN, \begin{bmatrix} I_r \\ 0 \end{bmatrix}, \begin{bmatrix} I_r \\ 0 \end{bmatrix}}$$

$$\Rightarrow \bar{A}_{GN} = \bar{A} - \begin{bmatrix} 0 & 0 \\ 0 & \bar{A}_{22} - \bar{A}_{21} \bar{A}_{11}^\dagger \bar{A}_{12} \end{bmatrix} =: \bar{A} - E_{GN}$$

BOUND ON GN APPROXIMATION ERROR > Derivation

- $A, \tilde{V}, \tilde{U} \rightarrow A_{GN} = A\tilde{V}(\tilde{U}^* A\tilde{V})^\dagger \tilde{U}^* A$

- Define

$$\bar{A} = [\tilde{U} \ \tilde{U}_\perp]^* A [\tilde{V} \ \tilde{V}_\perp], \quad \bar{A}_{GN} = \left([\tilde{U} \ \tilde{U}_\perp]^* A [\tilde{V} \ \tilde{V}_\perp] \right)_{GN, \begin{bmatrix} l_r \\ 0 \end{bmatrix}, \begin{bmatrix} l_r \\ 0 \end{bmatrix}}$$



Corollary 5.1
(L., Al Daas, Nakatsukasa, 2024)

$$\Rightarrow \bar{A}_{GN} = \bar{A} - \begin{bmatrix} 0 & 0 \\ 0 & \bar{A}_{22} - \bar{A}_{21} \bar{A}_{11}^\dagger \bar{A}_{12} \end{bmatrix} =: \bar{A} - E_{GN}$$

Define

$$\tau_i = \frac{\max\{\|\bar{A}_{12}\|_2, \|\bar{A}_{21}\|_2\}}{\min_j |\sigma_i(\bar{A}) - \sigma_j(\bar{A}_{22})| - 2\|E_{GN}\|_2}.$$

Then, for each i , if $\tau_i > 0$

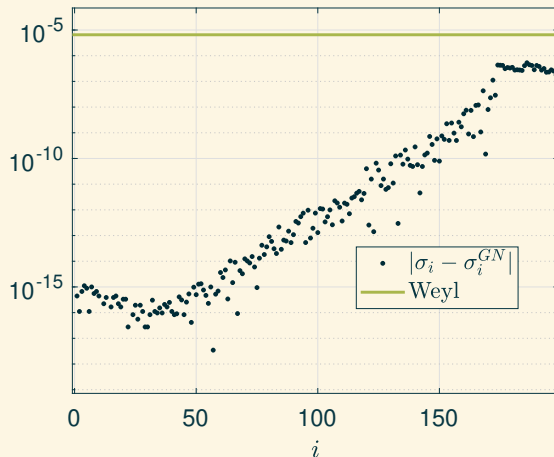
$$|\sigma_i(A) - \sigma_i(A_{GN})| = |\sigma_i(\bar{A}) - \sigma_i(\bar{A}_{GN})| \leq \left\| \bar{A}_{22} - \bar{A}_{21} \bar{A}_{11}^\dagger \bar{A}_{12} \right\|_2 \tau_i^2$$

► $\tau_i < 1$ necessary to be better than Weyl. If $\sigma_i(\bar{A})$ is far from the spectrum of \bar{A}_{22} then $\tau_i \ll 1$

BOUND ON GN APPROXIMATION ERROR > Numerical illustration

- $\ell = 0$
- $A \in \mathbb{R}^{1000 \times 1000}$
- U_{ex}, V_{ex} Haar Matrices
- $\sigma_i(A)$ exponentially decaying
- $[\tilde{V}, \sim] = \text{qr}(A^* \Omega, 0)$
- $[\tilde{U}, \sim] = \text{qr}(A \Omega, 0)$
- $\tilde{V} \in \mathbb{R}^{1000 \times 200}$
- $\tilde{U} \in \mathbb{R}^{1000 \times 200}$
- Compute pseudoinverses by QR factorization

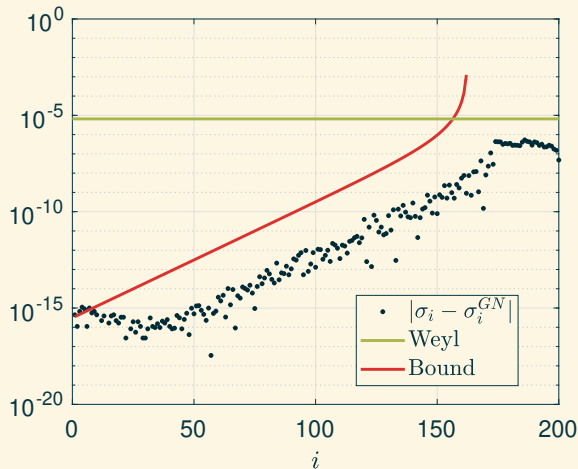
$$\sigma_i(A_{GN, \tilde{V}, \tilde{U}}) = \sigma_i(A \tilde{V} (\tilde{U}^* A \tilde{V})^\dagger \tilde{U}^* A)$$



BOUND ON GN APPROXIMATION ERROR > Numerical illustration

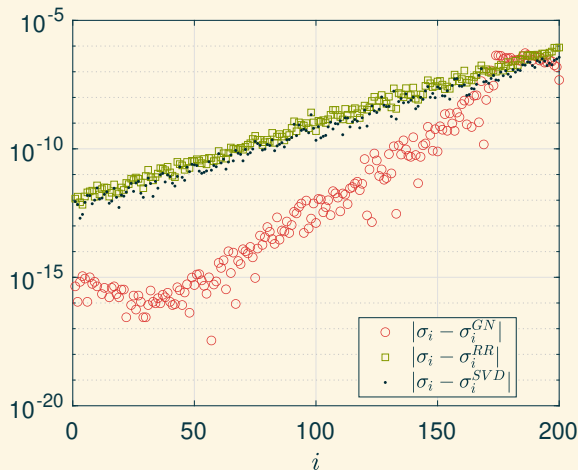
- $\ell = 0$
- $A \in \mathbb{R}^{1000 \times 1000}$
- U_{ex}, V_{ex} Haar Matrices
- $\sigma_i(A)$ exponentially decaying
- $[\tilde{V}, \sim] = \text{qr}(A^* \Omega, 0)$
- $[\tilde{U}, \sim] = \text{qr}(A \Omega, 0)$
- $\tilde{V} \in \mathbb{R}^{1000 \times 200}$
- $\tilde{U} \in \mathbb{R}^{1000 \times 200}$
- Compute pseudoinverses by QR factorization

$$\sigma_i(A_{GN, \tilde{V}, \tilde{U}}) = \sigma_i(A \tilde{V} (\tilde{U}^* A \tilde{V})^\dagger \tilde{U}^* A)$$



COMPARISON OF METHODS > *Idea*Single-pass methods

- ▶ $\sigma_i^{SVD} = \sigma_i(A\tilde{V})$
- ▶ $\sigma_i^{RR} = \sigma_i(\tilde{U}^* A \tilde{V})$
- ▶ $\sigma_i^{GN} = \sigma_i(A\tilde{V}(\tilde{U}^* A \tilde{V})^\dagger \tilde{U}^* A)$



COMPARISON OF METHODS > *Analysis*

$$Q_1 = [\tilde{U} \quad \tilde{U}_\perp], \quad Q_2 = [\tilde{V} \quad \tilde{V}_\perp]$$

$$\triangleright \sigma_i^{RR} = \sigma_i(\tilde{U}^* A \tilde{V})$$

$$\bar{A} = Q_1^* A Q_2$$

$$\sigma_i(A_{RR}, \tilde{V}, \tilde{U}) = \sigma_i(\bar{A}_{RR}, \begin{bmatrix} I_r \\ 0 \end{bmatrix}, \begin{bmatrix} I_{r+\ell} \\ 0 \end{bmatrix}) = \sigma_i(\bar{A}_{11}) = \sigma_i\left(\begin{bmatrix} \bar{A}_{11} & 0 \\ 0 & 0 \end{bmatrix}\right)$$

COMPARISON OF METHODS > Analysis

$$Q_1 = [\tilde{U} \quad \tilde{U}_\perp], \quad Q_2 = [\tilde{V} \quad \tilde{V}_\perp]$$

$$\triangleright \sigma_i^{RR} = \sigma_i(\tilde{U}^* A \tilde{V})$$

$$\bar{A} = Q_1^* A Q_2$$

$$\sigma_i(A_{RR}, \tilde{V}, \tilde{U}) = \sigma_i(\bar{A}_{RR}, \begin{bmatrix} I_r \\ 0 \end{bmatrix}, \begin{bmatrix} I_{r+\ell} \\ 0 \end{bmatrix}) = \sigma_i(\bar{A}_{11}) = \sigma_i\left(\begin{bmatrix} \bar{A}_{11} & 0 \\ 0 & 0 \end{bmatrix}\right)$$

Define

$$\tau_i^{RR} := \frac{2 \max\{\|\bar{A}_{12}\|_2, \|\bar{A}_{21}\|_2\}}{(\min_k |\sigma_i - \sigma_k(\bar{A}_{22})| - 2\|E_{RR}\|_2)} > 0$$

Then, for each i , if $\tau_i > 0$

$$\begin{aligned} |\sigma_i - \sigma_i^{RR}| &\leq 4 \frac{\max\{\|\bar{A}_{12}\|_2, \|\bar{A}_{21}\|_2\}^2}{\min_k |\sigma_i - \sigma_k(\bar{A}_{22})| - 2\|E_{RR}\|_2} \\ &\quad + \|\bar{A}_{22}\|_2 \frac{4 \max\{\|\bar{A}_{12}\|_2, \|\bar{A}_{21}\|_2\}^2}{(\min_k |\sigma_i - \sigma_k(\bar{A}_{22})| - 2\|E_{RR}\|_2)^2} \end{aligned}$$

COMPARISON OF METHODS > Analysis

$$Q_1 = [\tilde{U} \quad \tilde{U}_\perp], \quad Q_2 = [\tilde{V} \quad \tilde{V}_\perp]$$

$$\triangleright \sigma_i^{RR} = \sigma_i(\tilde{U}^* A \tilde{V})$$

$$\triangleright \sigma_i^{SVD} = \sigma_i(A \tilde{V})$$

$$\bar{A} = Q_1^* A Q_2$$

$$\tilde{A} = A Q_2 = [\tilde{A}_1 \quad \tilde{A}_2]$$

$$\sigma_i(A_{RR}, \tilde{V}, \tilde{U}) = \sigma_i(\bar{A}_{RR}, \begin{bmatrix} I_r \\ 0 \end{bmatrix}, \begin{bmatrix} I_{r+\ell} \\ 0 \end{bmatrix}) = \sigma_i(\bar{A}_{11}) = \sigma_i\left(\begin{bmatrix} \bar{A}_{11} & 0 \\ 0 & 0 \end{bmatrix}\right)$$

$$\sigma_i(A_{SVD}, \tilde{V}) = \sigma_i(\tilde{A}_{SVD}, \begin{bmatrix} I_r \\ 0 \end{bmatrix}) = \sigma_i(\begin{bmatrix} \tilde{A}_1 & 0 \end{bmatrix})$$

Define

$$\tau_i^{RR} := \frac{2 \max\{\|\bar{A}_{12}\|_2, \|\bar{A}_{21}\|_2\}}{(\min_k |\sigma_i - \sigma_k(\bar{A}_{22})| - 2\|E_{RR}\|_2)} > 0$$

Then, for each i , if $\tau_i > 0$

$$\begin{aligned} |\sigma_i - \sigma_i^{RR}| &\leq 4 \frac{\max\{\|\bar{A}_{12}\|_2, \|\bar{A}_{21}\|_2\}^2}{\min_k |\sigma_i - \sigma_k(\bar{A}_{22})| - 2\|E_{RR}\|_2} \\ &\quad + \|\bar{A}_{22}\|_2 \frac{4 \max\{\|\bar{A}_{12}\|_2, \|\bar{A}_{21}\|_2\}^2}{(\min_k |\sigma_i - \sigma_k(\bar{A}_{22})| - 2\|E_{RR}\|_2)^2} \end{aligned}$$

COMPARISON OF METHODS > Analysis

$$Q_1 = [\tilde{U} \quad \tilde{U}_\perp], \quad Q_2 = [\tilde{V} \quad \tilde{V}_\perp]$$

$$\triangleright \sigma_i^{RR} = \sigma_i(\tilde{U}^* A \tilde{V})$$

$$\triangleright \sigma_i^{SVD} = \sigma_i(A \tilde{V})$$

$$\bar{A} = Q_1^* A Q_2$$

$$\tilde{A} = A Q_2 = [\tilde{A}_1 \quad \tilde{A}_2]$$

$$\sigma_i(A_{RR, \tilde{V}, \tilde{U}}) = \sigma_i(\bar{A}_{RR, \begin{bmatrix} I_r \\ 0 \end{bmatrix}, \begin{bmatrix} I_{r+\ell} \\ 0 \end{bmatrix}}) = \sigma_i(\bar{A}_{11}) = \sigma_i\left(\begin{bmatrix} \bar{A}_{11} & 0 \\ 0 & 0 \end{bmatrix}\right)$$

$$\sigma_i(A_{SVD, \tilde{V}}) = \sigma_i(\tilde{A}_{SVD, \begin{bmatrix} I_r \\ 0 \end{bmatrix}}) = \sigma_i(\begin{bmatrix} \tilde{A}_1 & 0 \end{bmatrix})$$

Define

$$\tau_i^{RR} := \frac{2 \max\{\|\tilde{A}_{12}\|_2, \|\tilde{A}_{21}\|_2\}}{(\min_k |\sigma_i - \sigma_k(\tilde{A}_{22})| - 2\|E_{RR}\|_2)} > 0$$

Then, for each i , if $\tau_i > 0$

$$|\sigma_i - \sigma_i^{RR}| \leq 4 \frac{\max\{\|\tilde{A}_{12}\|_2, \|\tilde{A}_{21}\|_2\}^2}{\min_k |\sigma_i - \sigma_k(\tilde{A}_{22})| - 2\|E_{RR}\|_2} + \|\tilde{A}_{22}\|_2 \frac{4 \max\{\|\tilde{A}_{12}\|_2, \|\tilde{A}_{21}\|_2\}^2}{(\min_k |\sigma_i - \sigma_k(\tilde{A}_{22})| - 2\|E_{RR}\|_2)^2}$$

Define

$$\tau_i^{SVD} := \frac{2\|\tilde{A}_2\|_2}{\sigma_i - 2\|E_{SVD}\|_2} > 0$$

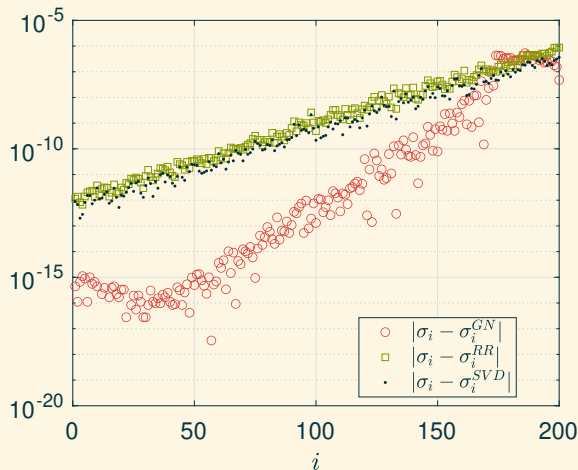
Then, for each i , if $\tau_i > 0$

$$|\sigma_i - \sigma_i^{SVD}| \leq 4 \frac{\|\tilde{A}_2\|_2^2}{\sigma_i - 2\|E_{SVD}\|_2}$$

COMPARISON OF METHODS > Idea

Single-pass methods

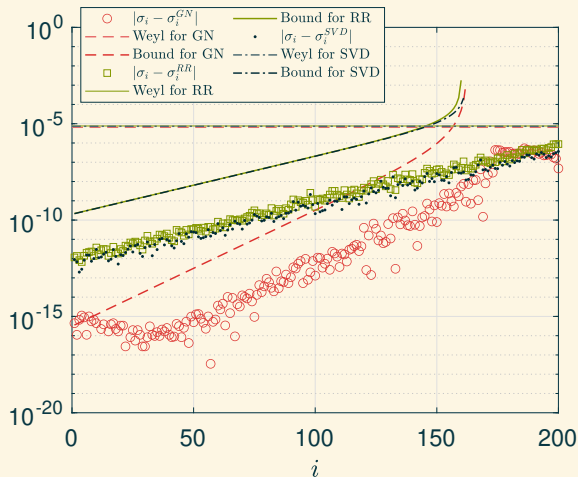
- ▶ $\sigma_i^{SVD} = \sigma_i(A\tilde{V})$
- ▶ $\sigma_i^{RR} = \sigma_i(\tilde{U}^* A \tilde{V})$
- ▶ $\sigma_i^{GN} = \sigma_i(A\tilde{V}(\tilde{U}^* A \tilde{V})^\dagger \tilde{U}^* A)$



COMPARISON OF METHODS > Idea

Single-pass methods

- ▶ $\sigma_i^{SVD} = \sigma_i(A\tilde{V})$
- ▶ $\sigma_i^{RR} = \sigma_i(\tilde{U}^* A \tilde{V})$
- ▶ $\sigma_i^{GN} = \sigma_i\left(A\tilde{V}(\tilde{U}^* A \tilde{V})^\dagger \tilde{U}^* A\right)$



PLUS,

- ▶ Similar results for oversampling case
- ▶ Different approximate singular subspaces
- ▶ Idea on how to modify bound to make it computable

Future work:

- ▶ More on the difference between oversampled and non-oversampled cases
- ▶ Use bounds to formally characterize the differences in behaviors of the different techniques: GN, HMT, Rayleigh-Ritz;

THANK YOU!



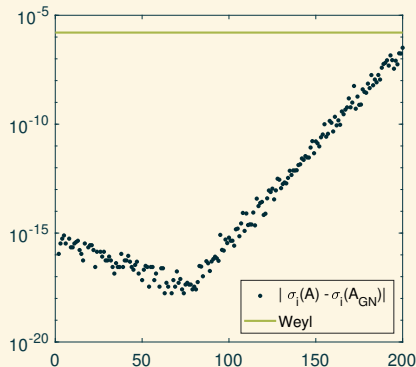
EXTRACTING ACCURATE SINGULAR VALUES FROM APPROXIMATE SUBSPACES

LORENZO LAZZARINO, HUSSAM AL DAAS, YUJI NAKATSUKASA

[1] MATRIX PERTURBATION ANALYSIS OF METHODS FOR EXTRACTING SINGULAR VALUES FROM APPROXIMATE SINGULAR SUBSPACES, L.L., H. AL DAAS, Y. NAKATSUKASA,
2024, TO APPEAR IN SIAM J. MATRIX ANAL. APPL. (SIMAX)

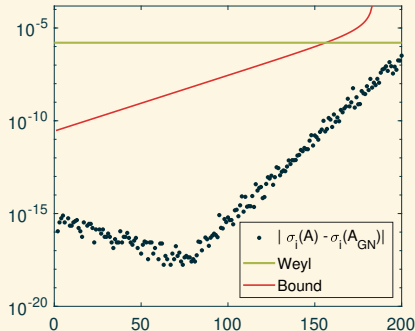
- $r + \ell = 1.5r$
- $A \in \mathbb{R}^{1000 \times 1000}$
- $U_{\text{ex}}, V_{\text{ex}}$ Haar Matrices
- $\sigma_i(A)$ exponentially decaying
- $[\tilde{V}, \sim] = \text{qr}(A^* \Omega, 0)$
- $[\tilde{U}, \sim] = \text{qr}(A \Omega, 0)$
- $\tilde{V} \in \mathbb{R}^{1000 \times 200}$
- $\tilde{U} \in \mathbb{R}^{1000 \times 300}$
- Compute pseudoinverses by QR factorization

$$\sigma_i(A_{GN, \tilde{V}, \tilde{U}}) = \sigma_i(A \tilde{V} (\tilde{U}^* A \tilde{V})^\dagger \tilde{U}^* A)$$



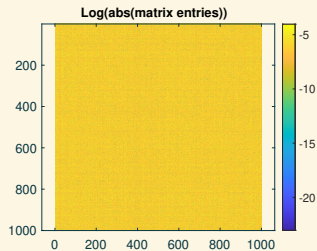
- $r + \ell = 1.5r$
- $A \in \mathbb{R}^{1000 \times 1000}$
- U_{ex}, V_{ex} Haar Matrices
- $\sigma_i(A)$ exponentially decaying
- $[\tilde{V}, \sim] = \text{qr}(A^* \Omega, 0)$
- $[\tilde{U}, \sim] = \text{qr}(A \Omega, 0)$
- $\tilde{V} \in \mathbb{R}^{1000 \times 200}$
- $\tilde{U} \in \mathbb{R}^{1000 \times 300}$
- Compute pseudoinverses by QR factorization

$$\sigma_i(A_{GN, \tilde{V}, \tilde{U}}) = \sigma_i(A \tilde{V} (\tilde{U}^* A \tilde{V})^\dagger \tilde{U}^* A)$$



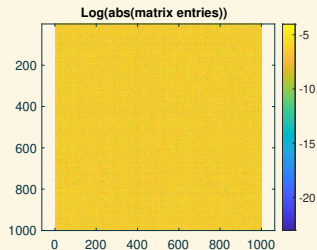
HEURISTIC BOUND FOR GN WITH OVERSAMPLE

A

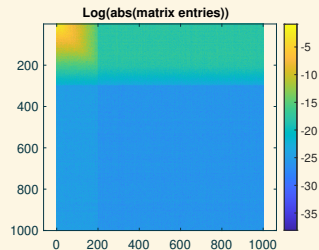


HEURISTIC BOUND FOR GN WITH OVERSAMPLE

A

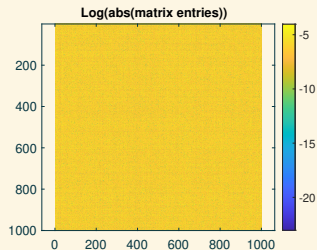


\bar{A}

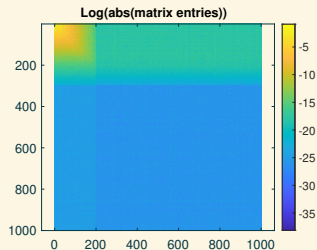


HEURISTIC BOUND FOR GN WITH OVERSAMPLE

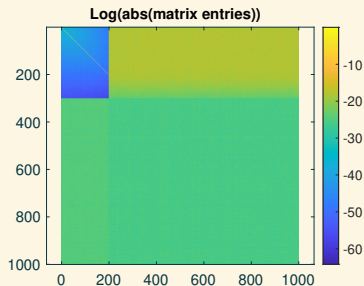
A



\bar{A}

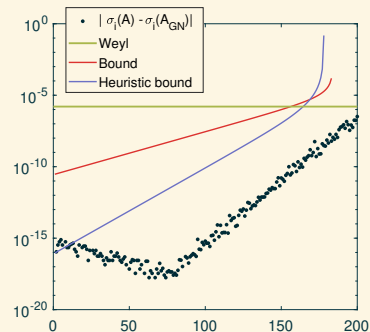
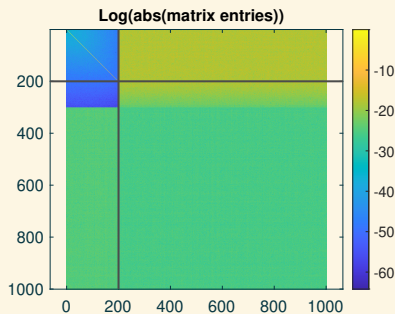


$\bar{\bar{A}}$



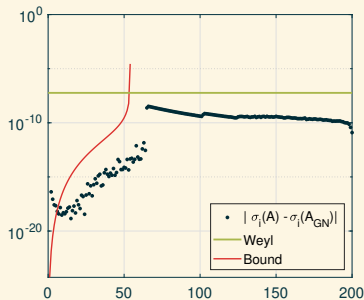
HEURISTIC BOUND FOR GN WITH OVERSAMPLE > Numerical illustration

$$\begin{aligned}\tilde{V} &\in \mathbb{R}^{1000 \times 200} \\ \tilde{U} &\in \mathbb{R}^{1000 \times 300} \\ \text{Size of } \tilde{A}_{11} &: 200 \times 200\end{aligned}$$

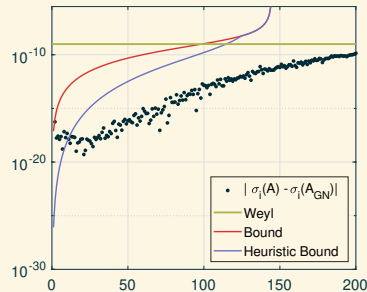


$$\sigma_i(A) = \left(\frac{1}{i}\right)^4$$

Without oversample ($\ell = 0$)



With oversample ($r + \ell = 1.5r$)



Provide ideas on how to make the bound computable in practice

$$\text{For } \tau_i > 0, \quad |\sigma_i(A) - \sigma_i(A_{GN})| \leq 2 \left\| \bar{A}_{12} - \bar{A}_{11} \bar{A}_{11}^\dagger \bar{A}_{12} \right\|_2 \tau_i + \left\| \bar{A}_{22} - \bar{A}_{21} \bar{A}_{11}^\dagger \bar{A}_{12} \right\|_2 \tau_i^2$$

$$\tau_i = \frac{\max\{\|\bar{A}_{12}\|_2, \|\bar{A}_{21}\|_2\} + \left\| \bar{A}_{12} - \bar{A}_{11} \bar{A}_{11}^\dagger \bar{A}_{12} \right\|_2}{\min_j |\sigma_i(\bar{A}) - \sigma_j(\bar{A}_{22})| - 2 \|E_{GN}\|_2}$$

Provide ideas on how to make the bound computable in practice

$$\text{For } \tau_i > 0, \quad |\sigma_i(A) - \sigma_i(A_{GN})| \leq 2 \left\| \bar{A}_{12} - \bar{A}_{11} \bar{A}_{11}^\dagger \bar{A}_{12} \right\|_2 \tau_i + \left\| \bar{A}_{22} - \bar{A}_{21} \bar{A}_{11}^\dagger \bar{A}_{12} \right\|_2 \tau_i^2$$

$$(\text{Forward Bound}) \quad \bar{A}_{GN} = \bar{A} - E_{GN} \implies \tau_i = \frac{\max\{\|\bar{A}_{12}\|_2, \|\bar{A}_{21}\|_2\} + \left\| \bar{A}_{12} - \bar{A}_{11} \bar{A}_{11}^\dagger \bar{A}_{12} \right\|_2}{\min_j |\sigma_i(\bar{A}) - \sigma_j(\bar{A}_{22})| - 2 \|E_{GN}\|_2}$$

Provide ideas on how to make the bound computable in practice

$$\text{For } \tau_i > 0, \quad |\sigma_i(A) - \sigma_i(A_{GN})| \leq 2 \left\| \bar{A}_{12} - \bar{A}_{11} \bar{A}_{11}^\dagger \bar{A}_{12} \right\|_2 \tau_i + \left\| \bar{A}_{22} - \bar{A}_{21} \bar{A}_{11}^\dagger \bar{A}_{12} \right\|_2 \tau_i^2$$

$$\text{(Forward Bound)} \quad \bar{A}_{GN} = \bar{A} - E_{GN} \implies \tau_i = \frac{\max\{\|\bar{A}_{12}\|_2, \|\bar{A}_{21}\|_2\} + \left\| \bar{A}_{12} - \bar{A}_{11} \bar{A}_{11}^\dagger \bar{A}_{12} \right\|_2}{\min_j |\sigma_i(\bar{A}) - \sigma_j(\bar{A}_{22})| - 2 \|E_{GN}\|_2}$$

$$\text{(Backward Bound)} \quad \bar{A} = \bar{A}_{GN} + E_{GN} \implies \tau_i = \frac{\max\{\|\bar{A}_{11} \bar{A}_{11}^\dagger \bar{A}_{12}\|_2, \|\bar{A}_{12}\|_2\} + \left\| \bar{A}_{12} - \bar{A}_{11} \bar{A}_{11}^\dagger \bar{A}_{12} \right\|_2}{\min_j |\sigma_i(\bar{A}_{GN}) - \sigma_j(\bar{A}_{21} \bar{A}_{11}^\dagger \bar{A}_{12})| - 2 \|E_{GN}\|_2}$$

Provide ideas on how to make the bound computable in practice

$$\text{For } \tau_i > 0, \quad |\sigma_i(A) - \sigma_i(A_{GN})| \leq 2 \left\| \bar{A}_{12} - \bar{A}_{11} \bar{A}_{11}^\dagger \bar{A}_{12} \right\|_2 \tau_i + \left\| \bar{A}_{22} - \bar{A}_{21} \bar{A}_{11}^\dagger \bar{A}_{12} \right\|_2 \tau_i^2$$

$$\tau_i = \frac{\overbrace{\max\{\|\bar{A}_{11} \bar{A}_{11}^\dagger \bar{A}_{12}\|_2, \|\bar{A}_{12}\|_2\}}^{=\|\bar{A}_{12}\|_2} + \overbrace{\|\bar{A}_{12} - \bar{A}_{11} \bar{A}_{11}^\dagger \bar{A}_{12}\|_2}^{\leq \|\bar{A}_{12}\|_2}}{\min_j |\sigma_i(\bar{A}_{GN}) - \sigma_j(\bar{A}_{21} \bar{A}_{11}^\dagger \bar{A}_{12})| - 2 \|E_{GN}\|_2}$$

Provide ideas on how to make the bound computable in practice

$$\text{For } \tau_i > 0, \quad |\sigma_i(A) - \sigma_i(A_{GN})| \leq 2 \left\| \bar{A}_{12} - \bar{A}_{11} \bar{A}_{11}^\dagger \bar{A}_{12} \right\|_2 \tau_i + \left\| \bar{A}_{22} - \bar{A}_{21} \bar{A}_{11}^\dagger \bar{A}_{12} \right\|_2 \tau_i^2$$

$$\tau_i = \frac{\overbrace{\max\{\|\bar{A}_{11} \bar{A}_{11}^\dagger \bar{A}_{12}\|_2, \|\bar{A}_{12}\|_2\}}^{=\|\bar{A}_{12}\|_2} + \overbrace{\|\bar{A}_{12} - \bar{A}_{11} \bar{A}_{11}^\dagger \bar{A}_{12}\|_2}^{\leq \|\bar{A}_{12}\|_2}}{\min_j |\sigma_i(\bar{A}_{GN}) - \sigma_j(\bar{A}_{21} \bar{A}_{11}^\dagger \bar{A}_{12})| - 2 \|E_{GN}\|_2}$$

