# EXTRACTING ACCURATE SINGULAR VALUES FROM APPROXIMATE SUBSPACES

LORENZO LAZZARINO, HUSSAM AL DAAS, YUJI NAKATSUKASA

*Mathematical Institute - University of Oxford*

*Computational Mathematics Theme - STFC UKRI*

DUE GIORNI DI ALGEBRA LINEARE NUMERICA E APPLICAZIONI,
21st January 2025

UNIVERSITY OF OXFORD
Mathematical Institute

Oxford Mathematics

Mathematical Institute

# PROBLEM SETTING AND CLASSICAL APPROACHES

1

## PROBLEM SETTING

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

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

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

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

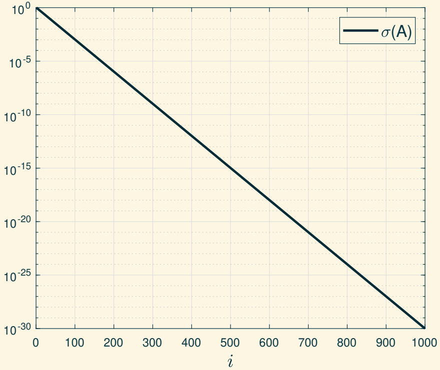
## PROBLEM SETTING

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

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

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

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



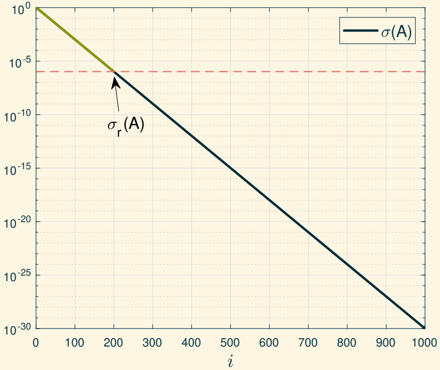
## PROBLEM SETTING

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

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

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

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

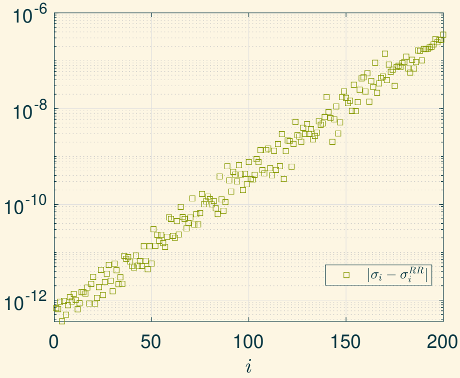


## CLASSICAL APPROACHES › *Rayleigh Ritz and (one-sided) SVD approximations*

### 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*



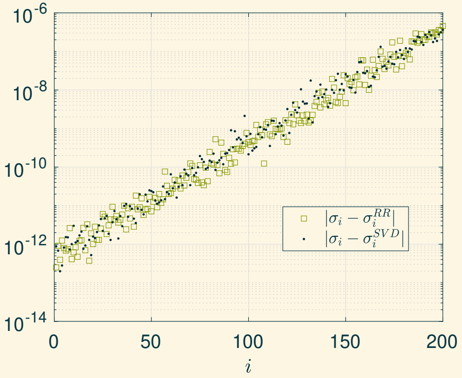
### Rayleigh Ritz (RR)

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

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

### (one-sided) SVD approximations

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



## GENERALIZED NYSTRÖM APPROXIMATION › *Motivational Comparison*

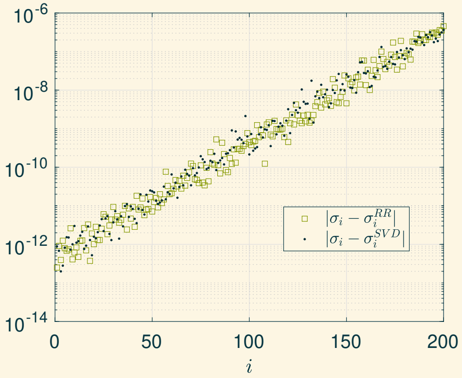
Mathematical Institute

**Generalized Nyström**

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

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

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



## GENERALIZED NYSTRÖM APPROXIMATION › *Motivational Comparison*

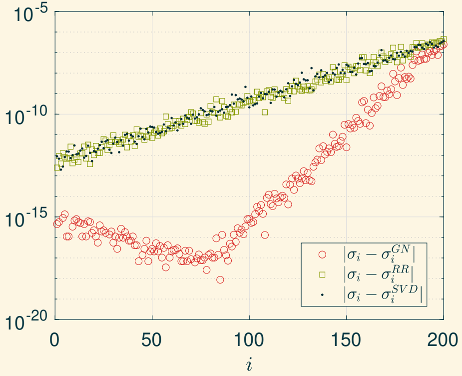
Mathematical Institute

**Generalized Nyström**

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

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

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



# GENERALIZED NYSTRÖM

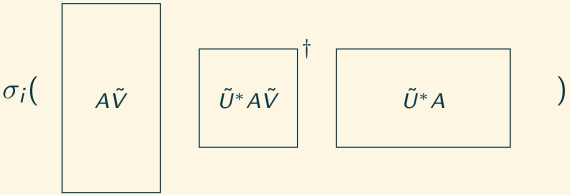
2

## GN APPROXIMATION AND EXTRACTING SINGULAR VALUES

**Generalized Nyström**

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

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



## GN APPROXIMATION AND EXTRACTING SINGULAR VALUES

Mathematical Institute

### Generalized Nyström

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

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

$$\sigma_i\Big( \boxed{Q_L} \; \boxed{R_L} \; \boxed{\tilde{U}^* A\tilde{V}}^{\dagger} \; \boxed{R_R^*} \; \boxed{Q_R^*} \Big)$$

## GN APPROXIMATION AND EXTRACTING SINGULAR VALUES



### Generalized Nyström

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

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

$$\sigma_i\left(\ \boxed{R_L}\ \ \boxed{\tilde{U}^*A\tilde{V}}^{\dagger}\ \ \boxed{R_R^*}\ \right)$$

## GN APPROXIMATION AND EXTRACTING SINGULAR VALUES

Oxford Mathematical Institute

### Generalized Nyström

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

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

$$\sigma_i\left(\ \boxed{R_L}\ \boxed{R_p^\dagger}\ \boxed{Q_p^*}\ \boxed{R_R^*}\ \right)$$

## GN AND MATRIX PERTURBATION THEORY

Oxford Mathematical Institute

### GN and Orthogonal Transformations

Consider $T_1$ and $T_2$ orthogonal matrices, then

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

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

1. Define $Q_1 = \begin{bmatrix} \tilde{U} & \tilde{U}_\perp \end{bmatrix} \quad Q_2 = \begin{bmatrix} \tilde{V} & \tilde{V}_\perp \end{bmatrix}$;
2. Consider the transformed matrix: $Q_1^* A Q_2$;
3. Consider the transformed GN approximation:

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

## GN AND MATRIX PERTURBATION THEORY

Mathematical Institute

### GN and Orthogonal Transformations

Consider $T_1$ and $T_2$ orthogonal matrices, then

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

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

1. Define $Q_1 = \begin{bmatrix}\tilde{U} & \tilde{U}_\perp\end{bmatrix} \quad Q_2 = \begin{bmatrix}\tilde{V} & \tilde{V}_\perp\end{bmatrix}$;
2. Consider the transformed matrix: $Q_1^* A Q_2$;
3. Consider the transformed GN approximation:

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

$$\rightarrow \quad |\sigma_i(A) - \sigma_i(A_{GN,\tilde{V},\tilde{U}})| = |\sigma_i(Q_1^* A Q_2) - \sigma_i((Q_1^* A Q_2)_{GN,\begin{bmatrix} I_r \\ 0 \end{bmatrix},\begin{bmatrix} I_{r+\ell} \\ 0 \end{bmatrix}})|$$

## GN AND MATRIX PERTURBATION THEORY › *Express $A_{GN}$ as a perturbation of the original matrix $A$*



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

(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 $A$*

Mathematical Institute

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

$$A_{GN,\tilde{V},\tilde{U}} = \left[\begin{array}{c} A_{11} \\ \hline A_{21} \end{array}\right] (\tilde{U}^* A \tilde{V})^\dagger \, \tilde{U}^* A$$

## GN AND MATRIX PERTURBATION THEORY › *Express $A_{GN}$ as a perturbation of the original matrix $A$*



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

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

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

Mathematical Institute

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

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

## GN AND MATRIX PERTURBATION THEORY › *Express $A_{GN}$ as a perturbation of the original matrix $A$*



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

$$\boxed{MM^{\dagger}M = M}$$

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

## GN AND MATRIX PERTURBATION THEORY › *Express $A_{GN}$ as a perturbation of the original matrix $A$*

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

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

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

## GN AND MATRIX PERTURBATION THEORY › *Express $A_{GN}$ as a perturbation of the original matrix $A$*

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

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

## GN AND MATRIX PERTURBATION THEORY › *Express $A_{GN}$ as a perturbation of the original matrix $A$*

Mathematical Institute

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

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

## GN AND MATRIX PERTURBATION THEORY › *Express $A_{GN}$ as a perturbation of the original matrix $A$*

Mathematical Institute

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

No-oversample ($\ell = 0$)
→ $A_{12} - A_{11}A_{11}^{\dagger}A_{12} = 0$, 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)

## GN AND MATRIX PERTURBATION THEORY › *Weyl's bound*

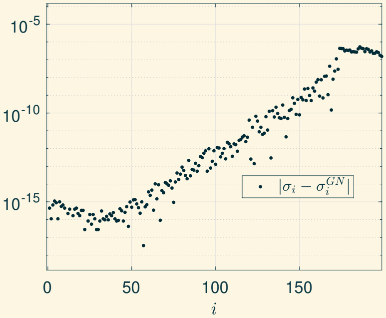
### Weyl's Theorem

For <u>any</u> 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)

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



## GN AND MATRIX PERTURBATION THEORY › *Weyl's bound*

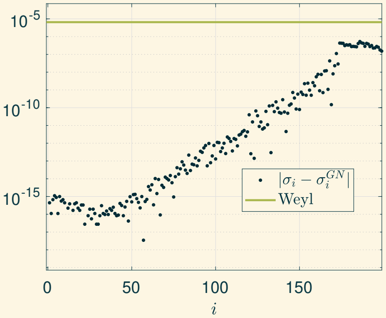
### Weyl's Theorem

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)

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



# MATRIX PERTURBATION THEORY RESULT

## 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

Oxford Mathematical Institute

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

Oxford Mathematical Institute

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 $\lambda_i$ 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 $\lambda_i$ 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 $2 \times 2$ 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

$$\begin{bmatrix} 0 & M \\ M^* & 0 \end{bmatrix} \tag{1}$$

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

## 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

$$\begin{bmatrix} 0 & M \\ M^* & 0 \end{bmatrix} \tag{1}$$

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

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

## FROM THE SYMMETRIC TO THE GENERAL RESULT

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 $G$

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

## FROM THE SYMMETRIC TO THE GENERAL RESULT

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 $G$

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

## FROM THE SYMMETRIC TO THE GENERAL RESULT

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 $G$

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

## FROM THE SYMMETRIC TO THE GENERAL RESULT

Mathematical Institute

General case

$\downarrow$

Transform to symmetric

$\downarrow$

**Obtain necessary structure**

$\downarrow$

Apply symmetric Result

$\downarrow$

Transform back

$\downarrow$

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

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 $G$

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

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

## FROM THE SYMMETRIC TO THE GENERAL RESULT

Mathematical Institute

General case

$\downarrow$

Transform to symmetric

$\downarrow$

**Obtain necessary structure**

$\downarrow$

Apply symmetric Result

$\downarrow$

Transform back

$\downarrow$

General Result

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

$$G_p = \begin{bmatrix} 0 & G_1 & | & 0 & B \\ G_1^* & 0 & | & C^* & 0 \\ - & - & - & - & - \\ 0 & C & | & 0 & G_2 \\ B^* & 0 & | & G_2^* & 0 \end{bmatrix}$$

$$\hat{G}_p = G_p + \begin{bmatrix} 0 & F_{11} & | & 0 & F_{12} \\ F_{11}^* & 0 & | & F_{21}^* & 0 \\ - & - & - & - & - \\ 0 & F_{21} & | & 0 & F_{22} \\ F_{12}^* & 0 & | & F_{22}^* & 0 \end{bmatrix} =: G_p + F_p.$$

## FROM THE SYMMETRIC TO THE GENERAL RESULT

General case

$\downarrow$

Transform to symmetric

$\downarrow$

Obtain necessary structure

$\downarrow$

**Apply symmetric Result**

$\downarrow$

Transform back

$\downarrow$

General Result

Define

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

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

General case

↓

Transform to symmetric

↓

Obtain necessary structure

↓

Apply symmetric Result

↓

**Transform back**

↓

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

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

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

- 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.

# APPLICATION TO GN AND COMPARISON

4

## BOUND ON GN APPROXIMATION ERROR > *Derivation*

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

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

## BOUND ON GN APPROXIMATION ERROR › *Derivation*

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

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

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

Define

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

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

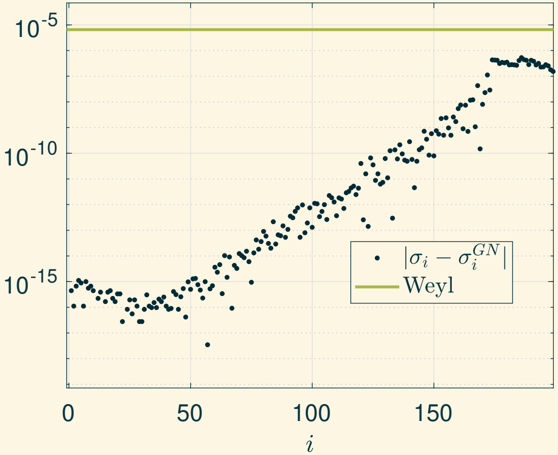
$$|\sigma_i(A) - \sigma_i(A_{GN})| = |\sigma_i(\bar{A}) - \sigma_i(\bar{A}_{GN})| \leq \left\|\bar{A}_{22} - \bar{A}_{21}\bar{A}_{11}^\dagger \bar{A}_{12}\right\|_2 \tau_i^2$$

- $\tau_i < 1$ 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
- $[\tilde{V}, \sim] = \mathtt{qr}(A^*\Omega, 0)$
- $[\tilde{U}, \sim] = \mathtt{qr}(A\Omega, 0)$
- $\tilde{V} \in \mathbb{R}^{1000 \times 200}$
- $\tilde{U} \in \mathbb{R}^{1000 \times 200}$
- Compute pseudoinverses by QR factorization

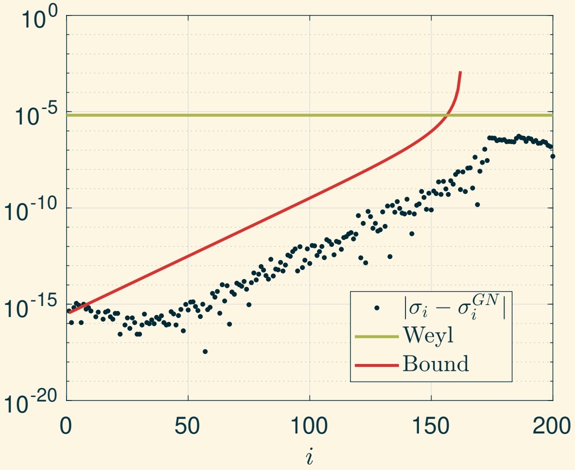
$$\sigma_i(A_{GN,\tilde{V},\tilde{U}}) = \sigma_i(A\tilde{V}(\tilde{U}^* A\tilde{V})^\dagger \tilde{U}^* A)$$



## BOUND ON GN APPROXIMATION ERROR > *Numerical illustration*

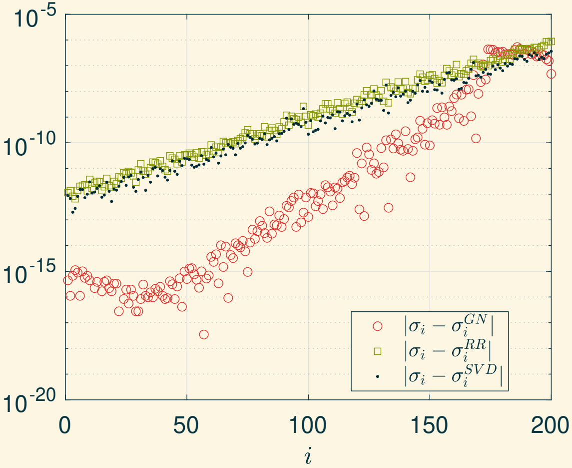
- $\ell = 0$
- $A \in \mathbb{R}^{1000 \times 1000}$
- $U_{ex}$, $V_{ex}$ Haar Matrices
- $\sigma_i(A)$ exponentially decaying
- $[\tilde{V}, \sim] = \mathtt{qr}(A^*\Omega, 0)$
- $[\tilde{U}, \sim] = \mathtt{qr}(A\Omega, 0)$
- $\tilde{V} \in \mathbb{R}^{1000 \times 200}$
- $\tilde{U} \in \mathbb{R}^{1000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

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



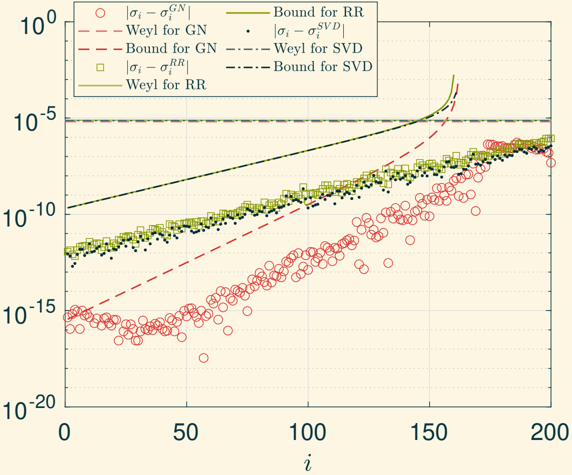
## COMPARISON OF METHODS > *Idea*

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



## COMPARISON OF METHODS > *Idea*

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



# THANK YOU!

EXTRACTING ACCURATE SINGULAR VALUES FROM APPROXIMATE SUBSPACES

Lorenzo Lazzarino, Hussam Al Daas, Yuji Nakatsukasa