The closest normal structured matrix

Erna Begović Kovač

University of Zagreb

ebegovic@fkit.hr

Joint work with Heike Faßbender and Philip Saltenberger

(TU Braunschweig)

GAMM ANLA18

10 October 2018

This work has been supported in part by Croatian Science Foundation under the project 3670.

OUTLINE

- Introduction
- Structured matrices
  and structure-preserving transformations
- Jacobi-type algorithm
  for the reduction to the canonical form
- Finding the closest normal matrix with a given structure
- Numerical examples

E. Begović Kovač, H. Faßbender, P. Saltenberger: On normal and structured matrices under unitary structure-preserving transformations. arXiv:1810.03369 [math.NA]

INTRODUCTION

- Set of normal matrices: $\mathcal{N} = \{X \ : \ XX^H = X^H X\}$
- X is normal if and only if there is unitary U such that

$$U^H X U = \begin{bmatrix} \diagdown \end{bmatrix}.$$

- A. Ruhe: *Closest normal matrix finally found!*
  BIT 27 (4) (1987) 585–598.

  Does NOT preserve given matrix structure.

INTRODUCTION

- Set of normal matrices: $\mathcal{N} = \{X \; : \; XX^H = X^H X\}$
- X is normal if and only if there is unitary U such that

$$U^H X U = \begin{bmatrix} \ddots \end{bmatrix}.$$

- A. Ruhe: *Closest normal matrix finally found!*
  BIT 27 (4) (1987) 585–598.

  Does NOT preserve given matrix structure.

Suppose that A has a structure $\mathcal{S}$, $A \in \mathcal{S}$.

Minimization problem:

$$\min \left\{ \|A - X\|_F^2 \; : \; X \in \mathcal{N} \cap \mathcal{S} \right\}$$

MAXIMIZATION PROBLEM

Theorem (Causey 1964, Gabriel 1979)

Let $A \in \mathbb{C}^{n \times n}$ and let $X = ZDZ^H$, where Z is unitary and D is diagonal. Then X is a nearest normal matrix to A in the Frobenius norm if and only if

(a) $\|\mathrm{diag}(Z^H A Z)\|_F = \max_{QQ^H = I} \|\mathrm{diag}(Q^H A Q)\|_F$, and

(b) $D = \mathrm{diag}(Z^H A Z)$.

→ Finding the closest normal matrix is equivalent to finding an unitary transformation that maximizes Frobenius norm of the diagonal.

→ This theorem has to be modified to fulfill structure-preserving requirements.

- N. J. Higham: *Matrix nearness problem and applications*. In Applications of Matrix theory 22 (1989) 1–27.

STRUCTURED MATRICES

- **Hamiltonian** A (J-Hermitian):

$$(JA)^H = JA, \quad \text{that is } A^H = JAJ, \quad \text{where} \quad J = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix}.$$

We can write it as

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & -A_{11}^H \end{bmatrix}, \qquad A_{12}^H = A_{12},\ A_{21}^H = A_{21}.$$

STRUCTURED MATRICES

- **Hamiltonian** A (J-Hermitian):

$$(JA)^H = JA, \quad \text{that is } A^H = JAJ, \quad \text{where} \quad J = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix}.$$

We can write it as

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & -A_{11}^H \end{bmatrix}, \qquad A_{12}^H = A_{12},\ A_{21}^H = A_{21}.$$

- **Skew-Hamiltonian** A (J-skew-Hermitian):

$$(JA)^H = -JA, \quad \text{that is } A^H = -JAJ.$$

We can write it as

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{11}^H \end{bmatrix}, \qquad A_{12}^H = -A_{12},\ A_{21}^H = -A_{21}.$$

- For every skew-Hamiltonian W there is Hamiltonian H (and viceversa) such that $W = \imath H$.

STRUCTURED MATRICES–cont.

- **Per-Hermitian** A (F-Hermitian):

$$(FA)^H = FA, \quad \text{that is } A^H = FAF,$$

$$\text{where} \quad F = \begin{bmatrix} 0 & \cdots & 0 & 1 \\ \vdots & & \iddots & 0 \\ 0 & \iddots & & \vdots \\ 1 & 0 & \cdots & 0 \end{bmatrix}.$$

$\rightarrow$ Hermitian about its anti-diagonal

- **Perskew-Hermitian** A (F-skew-Hermitian):

$$(FA)^H = -FA, \quad \text{that is } A^H = -FAF.$$

$\rightarrow$ Skew-Hermitian about its anti-diagonal

- For every perskew-Hermitian K there is per-Hermitian M (and viceversa) such that $K = \imath M$.

STRUCTURE-PRESERVING TRANSFORMATIONS

- For Hamiltonian and skew-Hamiltonian

$$M \text{ is } \textbf{symplectic} \text{ if } M^H J M = J.$$

- For per-Hermitian and perskew-Hermitian

$$M \text{ is } \textbf{perplectic} \text{ if } M^H F M = F.$$

CANONICAL FORM — HAMILTONIAN

Theorem (BK, Faßbender, Saltenberger)

For every normal Hamiltonian $A \in \mathbb{C}^{2n\times 2n}$ there is unitary symplectic U such that

$$U^H A U = \begin{bmatrix} D_1 & 0 & 0 & 0 \\ 0 & D_2 & 0 & D_3 \\ 0 & 0 & -D_1^H & 0 \\ 0 & -D_3 & 0 & D_2 \end{bmatrix},$$

where D_j, $j = 1, 2, 3$ diagonal matrices,
$D_1 \in \mathbb{C}^{n_1\times n_1}$, $D_2 \in \imath\mathbb{R}^{n_2\times n_2}$, $D_3 \in \mathbb{R}^{n_2\times n_2}$, $n_1 + n_2 = n$.

$$U^H A U = \begin{bmatrix} \Lambda_1 & \Lambda_2 \\ -\Lambda_2 & -\Lambda_1^H \end{bmatrix} = \begin{bmatrix} \diagdown & \diagdown \\ \diagdown & \diagdown \end{bmatrix} =: \Lambda_{\mathcal{H}}$$

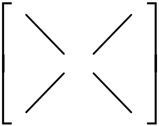
Theorem (BK, Faßbender, Saltenberger)

For every normal per-Hermitian $A \in \mathbb{C}^{2n \times 2n}$ there is unitary perplectic U such that

$$U^H A U = \begin{bmatrix} D_1 & 0 & 0 & 0 \\ 0 & D_2 & D_3 & 0 \\ 0 & FD_3F & FD_2F & 0 \\ 0 & 0 & 0 & FD_1F \end{bmatrix},$$

where D_1 i D_2 are diagonal, and D_3 is antidiagonal matrix, $D_1 \in \mathbb{C}^{n_1 \times n_1}$, $D_2 \in \mathbb{R}^{n_2 \times n_2}$, $D_3 \in \mathbb{R}^{n_2 \times n_2}$, $n_1 + n_2 = n$.

$$U^H A U = \begin{bmatrix} \Lambda_1 & \Lambda_2 F \\ F\Lambda_2 & F\Lambda_1^H F \end{bmatrix} =$$



$$=: \Lambda_{\mathcal{P}}$$

MAXIMIZATION ALGORITHM

$$\max_{ZZ^H=I,\ Z\in Sp_{2n}(\mathbb{C})} \left\{ f_{\mathcal{H}}(Z) := \|\mathrm{diag}(Z^H A Z)\|_F^2 + \|\mathrm{diag}(J Z^H A Z)\|_F^2 \right\}$$

- Iterative algorithm of the form

$$A^{(k+1)} = R_k^H A^{(k)} R_k, \quad k \geq 0.$$

- Transformations R_k are structure-preserving rotations obtained by embedding **two Jacobi rotations**

$$\begin{bmatrix} c & -s \\ s & c \end{bmatrix} := \begin{bmatrix} \cos\phi & -e^{\imath\alpha}\sin\phi \\ e^{-\imath\alpha}\sin\phi & \cos\phi \end{bmatrix} \qquad \text{in } I_{2n}.$$

  They are chosen to maximize

$$\|\mathrm{diag}(A^{(k+1)})\|_F^2 + \|\mathrm{diag}(JA^{(k+1)})\|_F^2.$$

- D. S. Mackey, N. Mackey, F. Tisseur: *Structured tools for structured matrices*. Electron. J. Linear Al. 10 (2003) 106–145.

MAXIMIZATION ALGORITHM

$$\max_{ZZ^H=I,\ Z\in Pp_{2n}(\mathbb{C})} \left\{ f_{\mathcal{P}}(Z) := \|\mathrm{diag}(Z^H A Z)\|_F^2 + \|\mathrm{diag}(F Z^H A Z)\|_F^2 \right\}$$

- Iterative algorithm of the form

$$A^{(k+1)} = R_k^H A^{(k)} R_k, \quad k \geq 0.$$

- Transformations R_k are structure-preserving rotations obtained by embedding **two Jacobi rotations**

$$\begin{bmatrix} c & -s \\ s & c \end{bmatrix} := \begin{bmatrix} \cos\phi & -e^{\imath\alpha}\sin\phi \\ e^{-\imath\alpha}\sin\phi & \cos\phi \end{bmatrix} \qquad \text{in } I_{2n}.$$

  They are chosen to maximize

$$\|\mathrm{diag}(A^{(k+1)})\|_F^2 + \|\mathrm{diag}(F A^{(k+1)})\|_F^2.$$

- D. S. Mackey, N. Mackey, F. Tisseur: *Structured tools for structured matrices*. Electron. J. Linear Al. 10 (2003) 106–145.

SYMPLECTIC ROTATIONS

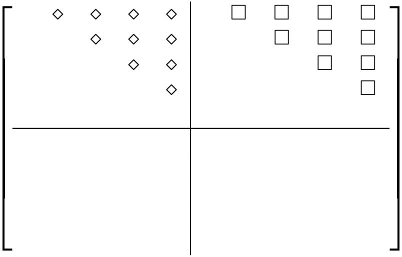
$$R(i,j,\phi,\alpha) = \left[\begin{array}{cc|cc} c & -s & & \\ \bar{s} & c & & \\ \hline & & c & -s \\ & & \bar{s} & c \end{array}\right] \begin{array}{c} i \\ j \\ \\ n+i \\ n+j \end{array}$$

$$R(i,j,\phi,\alpha) = \left[\begin{array}{cc|cc} c & & & -s \\ & c & -\bar{s} & \\ \hline & s & c & \\ \bar{s} & & & c \end{array}\right] \begin{array}{c} i \\ j-n \\ \\ n+i \\ j \end{array}$$

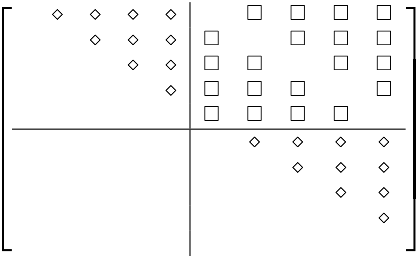
PIVOT POSITIONS (SYMPLECTIC)

$$
\left[\begin{array}{cccc|cccc}
\diamond & \diamond & \diamond & \diamond & \Box & \Box & \Box & \Box \\
 & \diamond & \diamond & \diamond & & \Box & \Box & \Box \\
 & & \diamond & \diamond & & & \Box & \Box \\
 & & & \diamond & & & & \Box \\
\hline
 & & & & & & & \\
 & & & & & & & \\
 & & & & & & & \\
 & & & & & & &
\end{array}\right]
$$

PIVOT POSITIONS (SYMPLECTIC)



considering double rotations
$\rightsquigarrow$



PERPLECTIC ROTATIONS

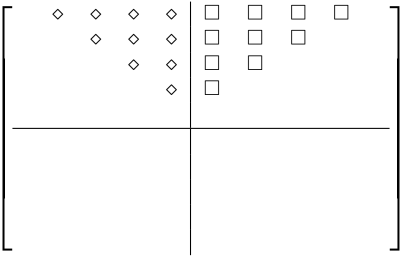
$$R(i,j,\phi,\alpha) = \left[\begin{array}{cc|cc} c & -s & & \\ \bar{s} & c & & \\ \hline & & c & \bar{s} \\ & & -s & c \end{array}\right] \begin{array}{c} i \\ j \\ 2n-j+1 \\ 2n-i+1 \end{array}$$

$$R(i,j,\phi,\alpha) = \left[\begin{array}{cc|cc} c & & -s & \\ & c & & \bar{s} \\ \hline \bar{s} & & c & \\ & -s & & c \end{array}\right] \begin{array}{c} i \\ 2n-j+1 \\ j \\ 2n-i+1 \end{array}$$

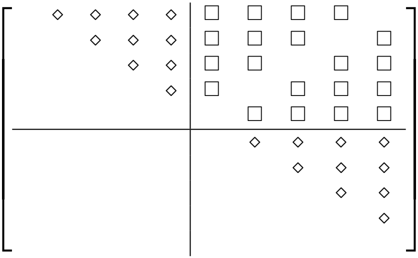
PIVOT POSITIONS (PERPLECTIC)

$$
\left[\begin{array}{cccc|cccc}
\diamond & \diamond & \diamond & \diamond & \square & \square & \square & \square \\
 & \diamond & \diamond & \diamond & \square & \square & \square & \\
 & & \diamond & \diamond & \square & \square & & \\
 & & & \diamond & \square & & & \\
 & & & & & & & \\
\hline
 & & & & & & & \\
 & & & & & & & \\
 & & & & & & & \\
 & & & & & & &
\end{array}\right]
$$

PIVOT POSITIONS (PERPLECTIC)



considering double rotations
$\rightsquigarrow$



REDUCTION TO CANONICAL FORM

Jacobi-type algorithm 1

Input: $A \in \mathbb{C}^{2n \times 2n} \in \mathcal{S}$, $Z_0 = I$
Output: structure-preserving unitary Z
REPEAT
  Select (i_k, j_k).
  Find ϕ_k and α_k for $R(i_k, j_k, \phi_k, \alpha_k)$.
  $A^{(k+1)} = R_k^H A^{(k)} R_k$
  $Z_{k+1} = Z_k R_k$
UNTIL convergence

REDUCTION TO CANONICAL FORM

Jacobi-type algorithm 1

Input: $A \in \mathbb{C}^{2n \times 2n} \in \mathcal{S}$, $Z_0 = I$
Output: structure-preserving unitary Z
REPEAT
  Select (i_k, j_k).
  Find ϕ_k and α_k for $R(i_k, j_k, \phi_k, \alpha_k)$.
  $A^{(k+1)} = R_k^H A^{(k)} R_k$
  $Z_{k+1} = Z_k R_k$
UNTIL convergence

- Cyclic pivot strategy
- Convergence condition:

$$|\langle \text{grad} f(Z), Z\dot{R}(i_k, j_k, 0, \alpha_k) \rangle| \geq \eta \|\text{grad} f(Z)\|_F,$$

where $\dot{R}(i,j,\phi,\alpha) = \frac{\partial}{\partial \phi} R(i,j,\phi,\alpha)$ and $f = f_{\mathcal{H}}$ or $f = f_{\mathcal{P}}$.

CONVERGENCE

Theorem (BK, Faßbender, Saltenberger)

Let A be Hamiltonian (or skew-Hamiltonian) and let $(Z_k)_k$ be a sequence of unitary symplectic matrices generated by the Jacobi algorithm. Every accumulation point of $(Z_k)_k$ is a stationary point of function $f_{\mathcal{H}}$.

Theorem (BK, Faßbender, Saltenberger)

Let A be per-Hermitian (or perskew-Hermitian) and let $(Z_k)_k$ be a sequence of unitary perplectic matrices generated by the Jacobi algorithm. Every accumulation point of $(Z_k)_k$ is a stationary point of function $f_{\mathcal{P}}$.

CONVERGENCE

Theorem (BK, Faßbender, Saltenberger)

Let A be Hamiltonian (or skew-Hamiltonian) and let $(Z_k)_k$ be a sequence of unitary symplectic matrices generated by the Jacobi algorithm. Every accumulation point of $(Z_k)_k$ is a stationary point of function $f_{\mathcal{H}}$.

Theorem (BK, Faßbender, Saltenberger)

Let A be per-Hermitian (or perskew-Hermitian) and let $(Z_k)_k$ be a sequence of unitary perplectic matrices generated by the Jacobi algorithm. Every accumulation point of $(Z_k)_k$ is a stationary point of function $f_{\mathcal{P}}$.

- M. Ishteva, P.-A. Absil, P. Van Dooren: *Jacobi algorithm for the best low multilinear rank approximation of symmetric tensors.*
  SIAM J. Matrix Anal. Appl. 34(2) (2013) 651–672.

THE CLOSEST NORMAL MATRIX

- Let A be Hamiltonian. Analogy with unstructured case:
  1. (i) Find Z that maximizes
     $f_{\mathcal{H}}(Z) = \|\mathrm{diag}(Z^H A Z)\|_F^2 + \|\mathrm{diag}(J Z^H A Z)\|_F^2$,
  2. (ii) Extract the canonical form,
  3. (iii) Solution is given by $X = Z \begin{bmatrix} \diagdown & \diagdown \\ \diagdown & \diagdown \end{bmatrix} Z^H$.

$\rightarrow$ But this can produce a matrix that is not normal!

THE CLOSEST NORMAL MATRIX

- Let A be Hamiltonian. Analogy with unstructured case:
  - (i) Find Z that maximizes
    $f_{\mathcal{H}}(Z) = \|\mathrm{diag}(Z^H AZ)\|_F^2 + \|\mathrm{diag}(JZ^H AZ)\|_F^2$,
  - (ii) Extract the canonical form,
  - (iii) Solution is given by $X = Z \begin{bmatrix} \ddots & \ddots \\ \ddots & \ddots \end{bmatrix} Z^H$.

  $\rightarrow$ But this can produce a matrix that is not normal!

- We set

$$f_{\mathcal{D}}(Z) = \|\mathrm{diag}(Z^H AZ)\|_F^2.$$

  - (i) Find Z that maximizes $f_{\mathcal{D}}$.
  - (ii) Extract the diagonal.
  - (iii) Solution is given by $X = Z \begin{bmatrix} \ddots & \\ & \ddots \end{bmatrix} Z^H$.

ADDITIONAL ROTATIONS

$\rightarrow$ To find Z that maximizes $f_{\mathcal{D}}$ we add new rotations to the Jacobi algorithm.

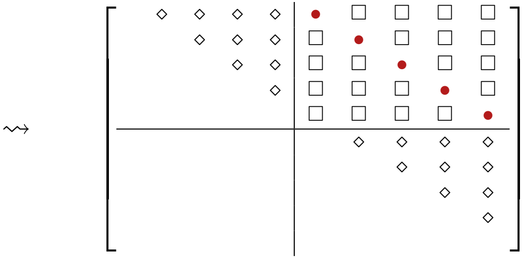
- Symplectic rotations

$$R(i, n+i, \phi, 0) = \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix} \begin{matrix} i \\ n+i \end{matrix}$$

ADDITIONAL ROTATIONS

$\rightarrow$ To find Z that maximizes $f_{\mathcal{D}}$ we add new rotations to the Jacobi algorithm.

- Symplectic rotations



ADDITIONAL ROTATIONS

$\rightarrow$ To find Z that maximizes $f_{\mathcal{D}}$ we add new rotations to the Jacobi algorithm.

- Symplectic rotations

$$R(i, n+i, \phi, 0) = \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix} \begin{matrix} i \\ n+i \end{matrix}$$

- Perplectic rotations

$$R(i, 2n-i+1, \phi, -\frac{\pi}{2}) = \begin{bmatrix} \cos\phi & \imath\sin\phi \\ \imath\sin\phi & \cos\phi \end{bmatrix} \begin{matrix} i \\ 2n-i+1 \end{matrix}$$

ADDITIONAL ROTATIONS

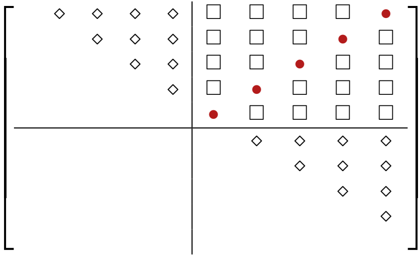
$\rightarrow$ To find Z that maximizes $f_{\mathcal{D}}$ we add new rotations to the Jacobi algorithm.

- Symplectic rotations

$$R(i, n+i, \phi, 0) = \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix} \begin{matrix} i \\ n+i \end{matrix}$$

- Perplectic rotations

$\rightsquigarrow$



ADDITIONAL ROTATIONS

→ To find Z that maximizes $f_{\mathcal{D}}$ we add new rotations to the Jacobi algorithm.

- Symplectic rotations

$$R(i, n+i, \phi, 0) = \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix} \begin{matrix} i \\ n+i \end{matrix}$$

- Perplectic rotations

$$R(i, 2n-i+1, \phi, -\frac{\pi}{2}) = \begin{bmatrix} \cos\phi & \imath\sin\phi \\ \imath\sin\phi & \cos\phi \end{bmatrix} \begin{matrix} i \\ 2n-i+1 \end{matrix}$$

DIAGONALIZATION ALGORITHM

Jacobi-type algorithm 2

Input: $A \in \mathbb{C}^{2n \times 2n} \in \mathcal{S}$, $Z_0 = I$
Output: structure-preserving unitary Z
REPEAT
    Select (i_k, j_k). (additional pivot positions are included)
    Find ϕ_k and α_k for $R(i_k, j_k, \phi_k, \alpha_k)$.
    $A^{(k+1)} = R_k^H A^{(k)} R_k$
    $Z_{k+1} = Z_k R_k$
UNTIL convergence

DIAGONALIZATION ALGORITHM

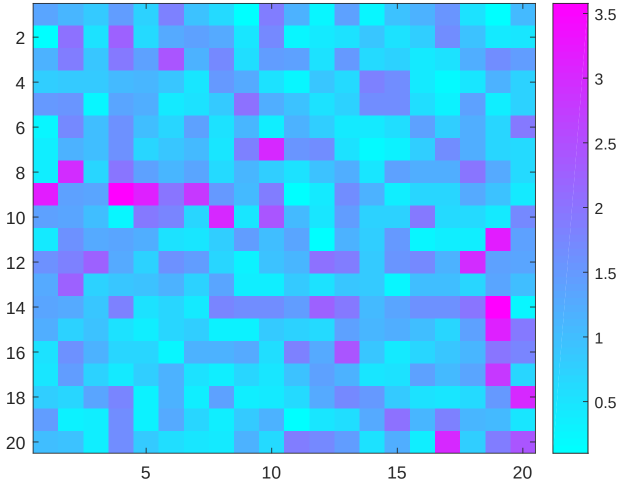
Jacobi-type algorithm 2

Input: $A \in \mathbb{C}^{2n \times 2n} \in \mathcal{S}$, $Z_0 = I$
Output: structure-preserving unitary Z
REPEAT
Select (i_k, j_k). (additional pivot positions are included)
Find ϕ_k and α_k for $R(i_k, j_k, \phi_k, \alpha_k)$.
$A^{(k+1)} = R_k^H A^{(k)} R_k$
$Z_{k+1} = Z_k R_k$
UNTIL convergence

Theorem (BK, Faßbender, Saltenberger)

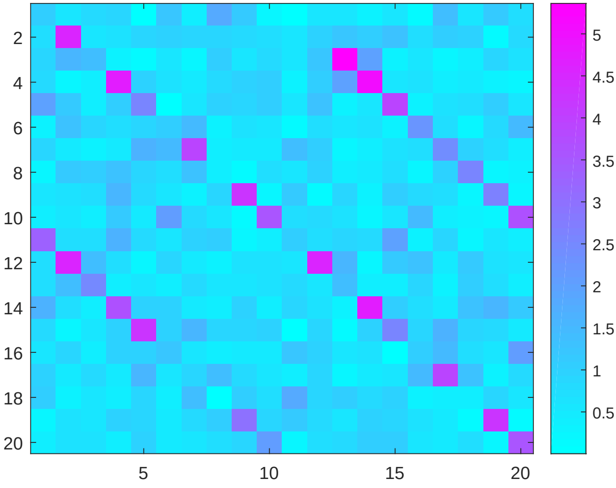
Let A be Hamiltonian and let $(Z_k)_k$ be a sequence of unitary symplectic matrices generated by the Jacobi algorithm with additional rotations. Every accumulation point of $(Z_k)_k$ is a stationary point of function $f_{\mathcal{D}}$.

NUMERICAL EXAMPLES — Canonical form



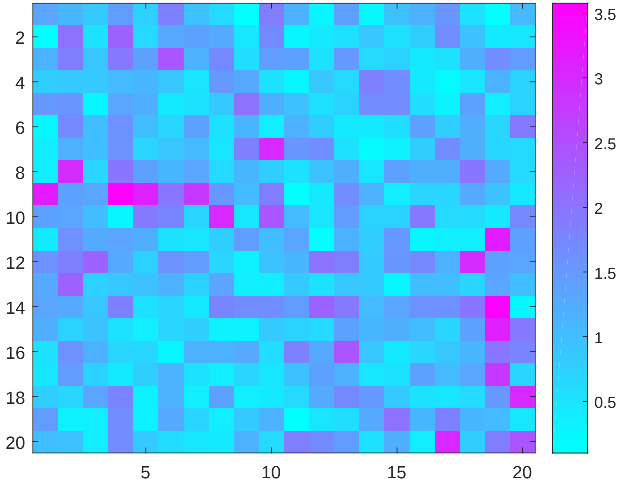
Random Hamiltonian 20×20 matrix.

NUMERICAL EXAMPLES — Canonical form



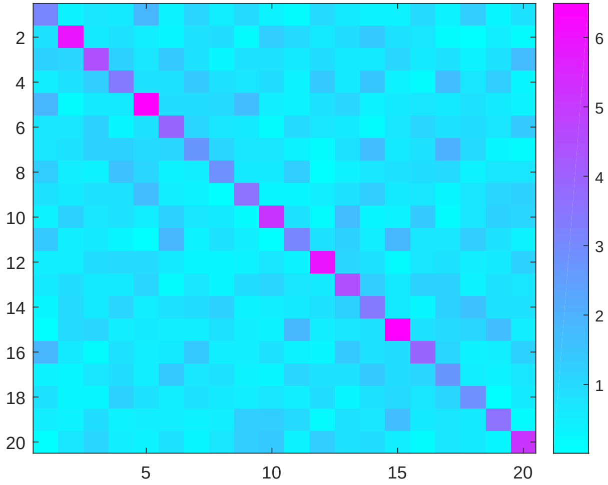
Reduction to the canonical form (Algorithm 1) after 10 cycles.

NUMERICAL EXAMPLES — Diagonalization



The same Hamiltonian 20×20 matrix.

NUMERICAL EXAMPLES — Diagonalization

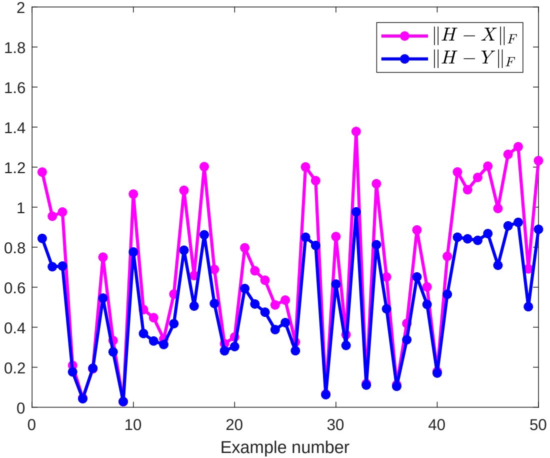


Diagonalization (Algorithm 2) after 10 cycles.

NUMERICAL EXAMPLES — Distance from normal matrix

We take normal Hamiltonian X and set $H = X + E$, such that H is Hamiltonian, but not normal.

Algorithm 2 on H gives its closest normal Y.



NUMERICAL EXAMPLES — Departure from normality

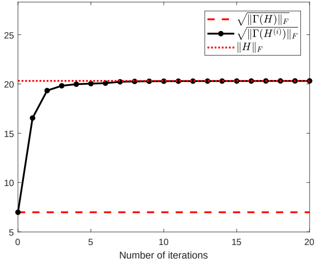
For any matrix A its Schur form

$$U^H A U = T = D + N$$

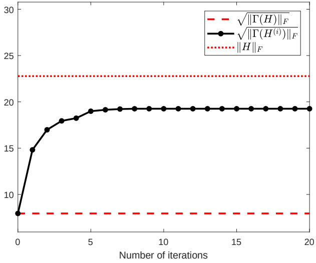
exists, where U is unitary, $D = \mathrm{diag}(T)$ and N is strictly upper triangular. The quantity $\Delta(A) = \|N\|_F$ is referred to as A's departure from normality.

We compare $\Delta(H)$ and $\mathrm{off}(H^{(20)})$ where $H^{(20)}$ is obtained by 20 iterations of Algorithm 2 and $\mathrm{off}(A) = \|A - \mathrm{diag}(A)\|_F^2$.

Example i	Size of H_i	$\Delta(H_i)$	$\mathrm{off}(H_i^{(20)})$
1	10	$7.1 \cdot 10^{+0}$	$6.4 \cdot 10^{+0}$
2	10	$4.0 \cdot 10^{-3}$	$3.1 \cdot 10^{-3}$
3	20	$3.5 \cdot 10^{-5}$	$3.1 \cdot 10^{-5}$
4	20	$5.3 \cdot 10^{+2}$	$4.4 \cdot 10^{+2}$
5	30	$7.7 \cdot 10^{+0}$	$6.7 \cdot 10^{+0}$
6	30	$1.0 \cdot 10^{-1}$	$9.0 \cdot 10^{-2}$
7	40	$7.9 \cdot 10^{-7}$	$6.6 \cdot 10^{-7}$
8	40	$3.1 \cdot 10^{+3}$	$2.7 \cdot 10^{+3}$
9	50	$1.1 \cdot 10^{-2}$	$9.5 \cdot 10^{-3}$
10	100	$7.8 \cdot 10^{-7}$	$6.8 \cdot 10^{-7}$



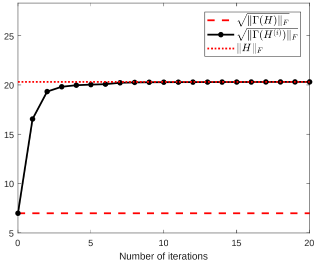
Normal Hamiltonian 20×20



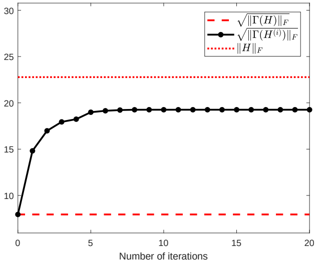
Random Hamiltonian 20×20

$$\Gamma(A) := \|\mathrm{diag}(Z^H A Z)\|_F^2 + \|\mathrm{diag}(J Z^H A Z)\|_F^2$$

NUMERICAL EXAMPLES — Convergence of Algorithm 1



Normal Hamiltonian 20×20



Random Hamiltonian 20×20

$$\Gamma(A) := \|\text{diag}(Z^H A Z)\|_F^2 + \|\text{diag}(J Z^H A Z)\|_F^2$$

THANK YOU!