Distance to singularity for (quadratic) matrix polynomials

SPEAKER: Miryam Gnazzo
JOINT WORK WITH: Nicola Guglielmi, GSSI

14 February 2022

GRAN SASSO SCIENCE INSTITUTE
SCHOOL OF ADVANCED STUDIES
Scuola Universitaria Superiore

Distance to singularity

A quadratic matrix polynomial $Q(\lambda)$ is a polynomial in the form:

$$Q(\lambda) = \lambda^2 A + \lambda B + C,$$

where $A, B, C \in \mathbb{C}^{n \times n}$. The matrix polynomial $Q(\lambda)$ is called regular if $\det(Q(\lambda)) \not\equiv 0$. Otherwise $Q(\lambda)$ is called singular.

Distance to singularity

A quadratic matrix polynomial $Q(\lambda)$ is a polynomial in the form:

$$Q(\lambda) = \lambda^2 A + \lambda B + C,$$

where $A, B, C \in \mathbb{C}^{n\times n}$. The matrix polynomial $Q(\lambda)$ is called regular if $\det(Q(\lambda)) \not\equiv 0$. Otherwise $Q(\lambda)$ is called singular.

Distance to singularity

Given a regular quadratic matrix polynomial $\lambda^2 A + \lambda B + C$, we look for the distance to singularity:

$$d(A, B, C) = \min \{ \| [\Delta A, \Delta B, \Delta C] \| \text{ such that } \lambda^2 (A + \Delta A) + \lambda (B + \Delta B) + (C + \Delta C) \text{ singular} \}.$$

Motivation

Second-order control system (Nichols, Kautsky, 2001):

$$J\ddot{\mathbf{z}} - D\dot{\mathbf{z}} - C\mathbf{z} = B\mathbf{u}, \qquad \mathbf{z}(0), \dot{\mathbf{z}}(0) \text{ given}$$

where $\mathbf{z}(t) \in \mathbb{R}^n$, $\mathbf{u}(t) \in \mathbb{R}^m$, $J, D, C \in \mathbb{R}^{n\times n}$ and $B \in \mathbb{R}^{n\times m}$.

Control problem: design a controller $\mathbf{u} = K_1\mathbf{z} + K_2\dot{\mathbf{z}} + \mathbf{r}$, where $K_1, K_2 \in \mathbb{R}^{m\times n}$, $\mathbf{r}(t) \in \mathbb{R}^m$ such that

$$J\ddot{\mathbf{z}} - (D + BK_2)\,\dot{\mathbf{z}} - (C + BK_1)\,\mathbf{z} = B\mathbf{r}$$

has desired properties.

Its behavior is governed by the eigenstructure of

$$Q(\lambda) = \lambda^2 J - \lambda\,(D + BK_2) - (C + BK_1)\,.$$

Motivation

Motivation: Ill conditioning of eigenvalues.

Example: consider the quadratic matrix polynomial, with α, β, γ small:

$$Q(\lambda) = \lambda^2 \begin{bmatrix} 1 & \alpha \\ \beta & \gamma \end{bmatrix} + \lambda \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ \gamma & 1 \end{bmatrix}.$$

For $\alpha, \beta = 0$ and $\gamma \neq 0$, the eigenvalues are: $\lambda = 0, 1, -\frac{1}{2} + \frac{\sqrt{3}}{2}, -\frac{1}{2} - \frac{\sqrt{3}}{2}$.

For $\gamma = 0$ and $\alpha, \beta \neq 0$, the eigenvalues are: $\lambda = 0, -\frac{\alpha+\beta}{\alpha\beta}$.

Motivation

Motivation: Ill conditioning of eigenvalues.

Example: consider the quadratic matrix polynomial, with α, β, γ small:

$$Q(\lambda) = \lambda^2 \begin{bmatrix} 1 & \alpha \\ \beta & \gamma \end{bmatrix} + \lambda \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ \gamma & 1 \end{bmatrix}.$$

For $\alpha, \beta = 0$ and $\gamma \neq 0$, the eigenvalues are: $\lambda = 0, 1, -\frac{1}{2} + \frac{\sqrt{3}}{2}, -\frac{1}{2} - \frac{\sqrt{3}}{2}$.

For $\gamma = 0$ and $\alpha, \beta \neq 0$, the eigenvalues are: $\lambda = 0, -\frac{\alpha+\beta}{\alpha\beta}$.

$Q(\lambda)$ is in a neighborhood of the singular matrix polynomial

$$\lambda^2 \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \lambda \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

A few references

- R. Byers, C. He, V. Mehrmann, (1998), Where is the nearest non-regular pencil?, *Linear Algebra and its Applications*.
- C. Mehl,V. Mehrmann, M. Wojtylak.(2015). On the distance to singularity via low rank perturbations. *Operators and Matrices*.
- N. Guglielmi, C. Lubich, V. Mehrmann, (2017) On the nearest singular matrix pencil, *SIAM Journal on Matrix Analysis and Applications*.
- M. Giesbrecht, J. Haraldson, G. Labahn, (2017). Computing the nearest rank-deficient matrix polynomial. *In: Proc. International Symposium on Symbolic and Algebraic Computation (ISSAC). ACM Press.*

Reformulation of the problem

$\lambda^2 (A + \Delta A) + \lambda (B + \Delta B) + (C + \Delta C)$ is singular $\Leftrightarrow$

$$\det \left(\mu_i^2 (A + \Delta A) + \mu_i (B + \Delta B) + (C + \Delta C) \right) = 0,$$

with distinct points $\mu_i \in \mathbb{C}$, for $i = 1, \ldots, d$ and $d \geq 2n + 1$.

Reformulation of the problem

$\lambda^2 (A + \Delta A) + \lambda (B + \Delta B) + (C + \Delta C)$ is singular $\Leftrightarrow$

$$\det\left(\mu_i^2 (A + \Delta A) + \mu_i (B + \Delta B) + (C + \Delta C)\right) = 0,$$

with distinct points $\mu_i \in \mathbb{C}$, for $i = 1, \ldots, d$ and $d \geq 2n + 1$.

Underlying optimization problem

$$(\Delta A_*, \Delta B_*, \Delta C_*) = \arg \min_{\Delta A, \Delta B, \Delta C \in \mathbb{C}^{n \times n}} \| [\Delta A, \Delta B, \Delta C] \|_F$$

subj to $\det\left(\mu_i^2(A + \Delta A) + \mu_i(B + \Delta B) + (C + \Delta C)\right) = 0$

for $i = 1, \ldots, d$.

Idea of the method

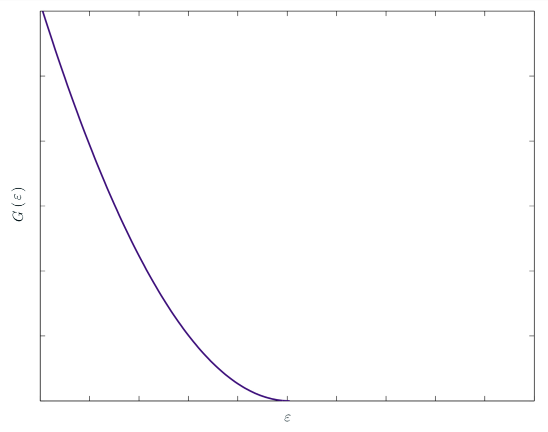
Consider $[\Delta A, \Delta B, \Delta C] = \varepsilon\, [\Delta, \Theta, \Gamma]$ of norm ε. Define the functional

$$G_\varepsilon\,(\Delta, \Theta, \Gamma) := \frac{1}{2} \sum_{i=1}^{d} \sigma_i^2(\Delta, \Theta, \Gamma),$$

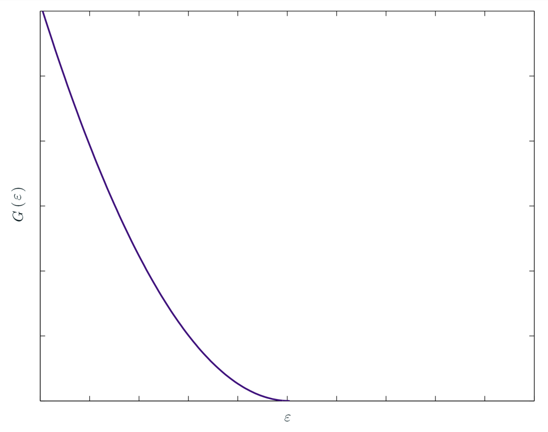
where $\sigma_i\,(\Delta, \Theta, \Gamma)$ is the smallest singular value of the matrix $\mu_i^2(A + \varepsilon\Delta) + \mu_i(B + \varepsilon\Theta) + (C + \varepsilon\Gamma)$.

- Compute $G(\varepsilon) = \min_{\Delta,\Theta,\Gamma} G_\varepsilon\,(\Delta, \Theta, \Gamma)$.
- Compute $\varepsilon^\star = \min\,\{\varepsilon \in \mathbb{R}^+ :\ G(\varepsilon) = 0\}$.

Idea of the method

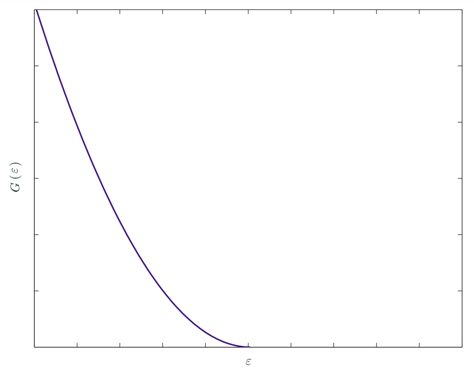


Idea of the method



- Fix the norm ε and solve $G(\varepsilon) = \min_{\Delta,\Theta,\Gamma} G_\varepsilon(\Delta, \Theta, \Gamma)$;
- Tune the value of ε in order to find the smallest zero of $G(\varepsilon)$.

Idea of the method



- Fix the norm ε and solve $G(\varepsilon) = \min_{\Delta,\Theta,\Gamma} G_{\varepsilon}(\Delta, \Theta, \Gamma)$;
  $\longrightarrow$ **Inner iteration**
- Tune the value of ε in order to find the smallest zero of $G(\varepsilon)$. $\longrightarrow$ **Outer iteration**

Inner iteration

Lemma: Let $\Delta(t), \Theta(t), \Gamma(t) \in \mathbb{C}^{n\times n}$ be a smooth path of matrices, with derivatives $\dot{\Delta}(t), \dot{\Theta}(t), \dot{\Gamma}(t)$. Then $G_\varepsilon\left(\Delta(t), \Theta(t), \Gamma(t)\right)$ is differentiable with

$$\frac{d}{dt}G_\varepsilon(\Delta, \Theta, \Gamma) = \varepsilon \mathrm{Re}\left\langle \left[M_2, M_1, M_0\right], \left[\dot{\Delta}, \dot{\Theta}, \dot{\Gamma}\right] \right\rangle,$$

where $\langle X, Y\rangle = \mathrm{trace}\left(X^H Y\right)$ and

$$M_2 = \sum_{i=1}^{d} \sigma_i \bar{\mu}_i^2 u_i v_i^H,\ M_1 = \sum_{i=1}^{d} \sigma_i \bar{\mu}_i u_i v_i^H,\ M_0 = \sum_{i=1}^{d} \sigma_i u_i v_i^H$$

with u_i, v_i left and right singular vectors associated with σ_i.

Inner iteration

Along the solutions of the system of ODEs

$$\begin{cases} \dot{\Delta} = -M_2 + \eta\Delta \\ \dot{\Theta} = -M_1 + \eta\Theta \\ \dot{\Gamma} = -M_0 + \eta\Gamma \end{cases}$$

where $\eta := \mathrm{Re}\left\langle [M_2, M_1, M_0], [\Delta, \Theta, \Gamma] \right\rangle$, we have that:

1. $\| [\Delta(t), \Theta(t), \Gamma(t)] \|_F = 1$ is conserved;
2. $\frac{d}{dt} G_\varepsilon \left(\Delta(t), \Theta(t), \Gamma(t)\right) \leq 0$.

This is a constrained gradient system.

Our goal

Computing the stationary points of the gradient system.

Outer iteration

Outer iteration: update the parameter ε up to the smallest root $\varepsilon^\star$ of $G(\varepsilon) = 0$.

Newton-like method

We approach the root $\varepsilon^\star$ from the left-hand side using

$$\varepsilon_{k+1} = \varepsilon_k - \frac{G(\varepsilon_k)}{G'(\varepsilon_k)},$$

where $G'(\varepsilon) = \frac{d}{d\varepsilon} G(\varepsilon) = -\| \left[M_2(\varepsilon), M_1(\varepsilon), M_0(\varepsilon) \right] \|_F$.

Example

Consider

$$P(\lambda) = \lambda^2 \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \lambda \begin{bmatrix} 0 & 1 \\ 0.5 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

Example

Consider

$$P(\lambda) = \lambda^2 \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \lambda \begin{bmatrix} 0 & 1 \\ 0.5 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

The computed distance to singularity is $d \approx 0.2794$ and the nearest singular polynomial is

$$\widetilde{P}(\lambda) = \lambda^2 \begin{bmatrix} 0.8645 & 0 \\ 0 & 0 \end{bmatrix} + \lambda \begin{bmatrix} 0 & 1.1058 \\ 0.6736 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0.8645 \end{bmatrix}.$$

Palindromic quadratic matrix polynomials

A palindromic quadratic matrix polynomials is a polynomial in the form

$$P(\lambda) := \lambda^2 A + \lambda B + A^H,$$

where $B = B^H$ and $A, B \in \mathbb{C}^{n\times n}$.

We allow perturbations which respect the structure

$$\lambda^2 (A + \varepsilon\Delta) + \lambda (B + \varepsilon\Theta) + (A + \varepsilon\Delta)^H, \quad \Delta, \Theta \in \mathbb{C}^{n\times n}.$$

Given a regular quadratic matrix polynomial $\lambda^2 A + \lambda B + C$, we look for:

$$d_P(A, B, C) = \min \left\{ \| [\Delta A, \Delta B, \Delta C] \|_F \ : \lambda^2 (A + \Delta A) + \right.$$
$$\left. + \lambda (B + \Delta B) + (C + \Delta C) \text{ singular and palindromic} \right\}.$$

Palindromic quadratic matrix polynomials

Projection onto the manifold w.r.t. Frobenius inner product

$$\Pi_{\mathcal{M}} : \mathbb{C}^{n\times 3n} \longmapsto \mathcal{M} := \left\{[\Delta,\Theta,\Gamma] \in \mathbb{C}^{n\times 3n} :\ \Theta = \Theta^H,\ \Gamma = \Delta^H\right\}$$

$$[\Delta,\Theta,\Gamma] \longmapsto \left[\frac{\Delta+\Gamma^H}{2}, \frac{\Theta+\Theta^H}{2}, \frac{\Delta^H+\Gamma}{2}\right].$$

The system of ODEs becomes

$$\begin{cases} \dot{\Delta} = -\dfrac{M_2+M_0^H}{2} + \eta\Delta, \\ \dot{\Theta} = -\dfrac{M_1+M_1^H}{2} + \eta\Theta, \end{cases}$$

where $\eta = \mathrm{Re}\left\langle \left[\frac{M_2+M_0^H}{2}, \frac{M_1+M_1^H}{2}\right], [\Delta,\Theta] \right\rangle$.

Example

Consider the palindromic polynomial

$$P(\lambda) = \lambda^2 \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + \lambda \begin{bmatrix} -1 & -1.5 \\ -1.5 & -1 \end{bmatrix} + \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

The nearest palindromic singular polynomial is

$$\lambda^2 \begin{bmatrix} 0.2335 & 0.1866 \\ 0.2807 & 0.2335 \end{bmatrix} + \lambda \begin{bmatrix} -1.3095 & -1.3077 \\ -1.3077 & -1.3095 \end{bmatrix} + \begin{bmatrix} 0.2335 & 0.2807 \\ 0.1866 & 0.2335 \end{bmatrix}.$$

Open issues and future work

- Choice of the set of complex points μ_i:
  - Number of points $d \geq 2n+1$;
  - Optimal choice of μ_i;
  - Influence on the numerical results;
  - Adding a set of test points $\widetilde{\mu}_i$.
- Efficient integration of the gradient system;
- Different additional structures on the matrix polynomials;
- Computational challenges of matrix polynomials of higher degree.

Numerical experiments

$$P(\lambda) = \lambda^2 \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \lambda \begin{bmatrix} 0 & 1 \\ 0.5 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

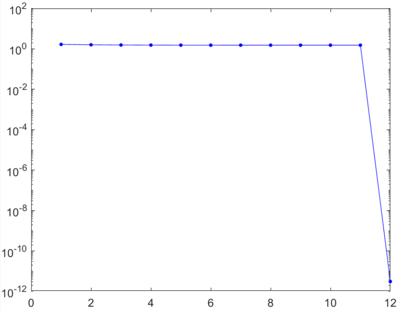


Figure 1: $\left|\varepsilon_{\text{up}} - \varepsilon_{\text{low}}\right|$

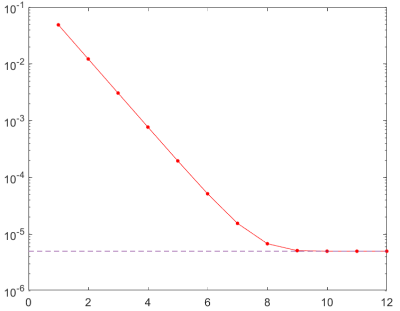


Figure 2: $G(\varepsilon)$

Iterations $= 12$; Tol$=5 \times 10^{-6}$; $d = 2n + 1 = 5$.

Numerical experiments

$$P(\lambda) = \lambda^2 \begin{bmatrix} 0 & 0 \\ 0 & -2 \end{bmatrix} + \lambda \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} -2 & 0 \\ -3 & 1 \end{bmatrix}.$$

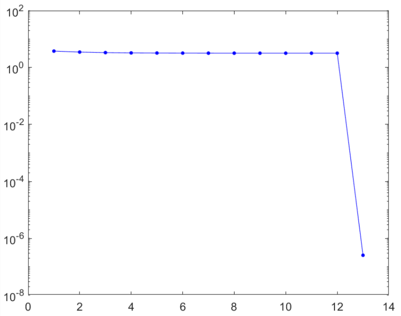


Figure 3: $\left|\varepsilon_{\text{up}} - \varepsilon_{\text{low}}\right|$

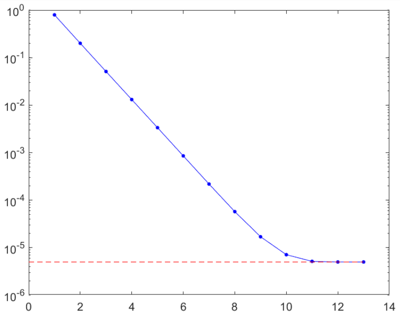


Figure 4: $G(\varepsilon)$

Iterations $= 13$; Tol$=5 \times 10^{-6}$; $d = 2n + 1 = 5$.

Numerical experiments $n = 2$ randn

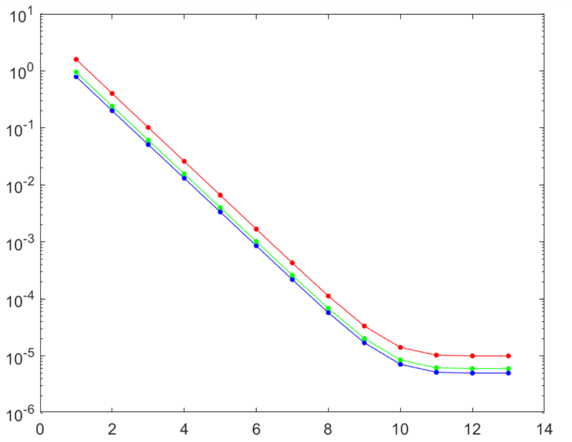


Figure 5: $G(\varepsilon)$

Three choices of d: $d = 5$ (blue), $d = 10$ (red) and $d = 3n = 6$ (green). Always 13 iterations.

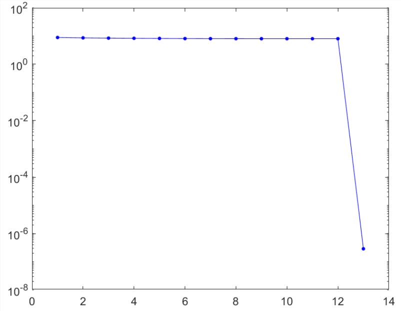


Figure 6: $\left|\varepsilon_{\text{up}} - \varepsilon_{\text{low}}\right|$, $d = 11$

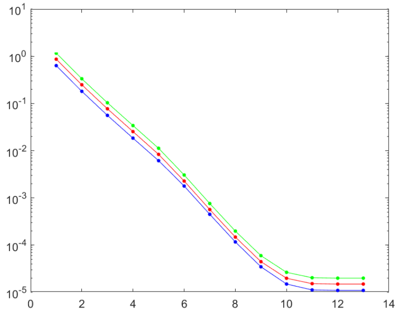


Figure 7: $G(\varepsilon)$

Iterations $= 13$; Tol$=d \times 10^{-6}$; $d = 11$(blue), $d = 15$(red) and $d = 20$(green).

Numerical experiments $n = 10$ real

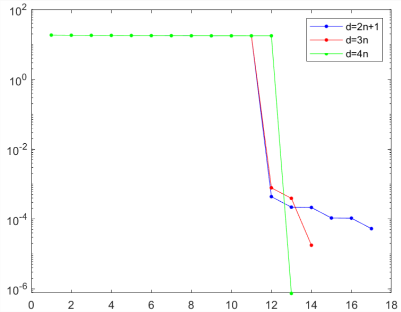


Figure 8: $\left|\varepsilon_{\text{up}} - \varepsilon_{\text{low}}\right|$

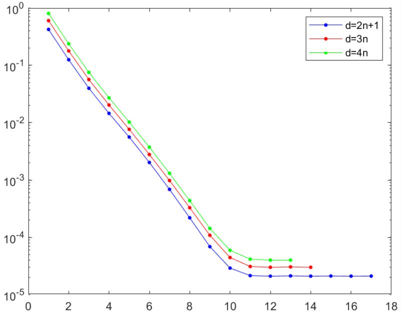


Figure 9: $G(\varepsilon)$

Tol=$d \times 10^{-6}$; $d = 21$(blue), $d = 30$(red) and $d = 40$(green).

Numerical experiments $n = 10$ real

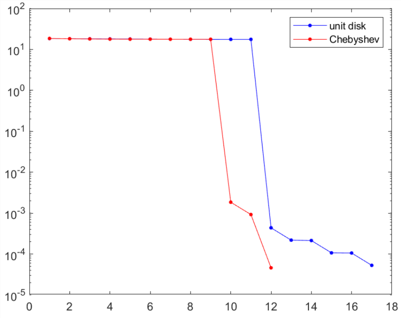


Figure 10: $|\varepsilon_{\text{up}} - \varepsilon_{\text{low}}|$

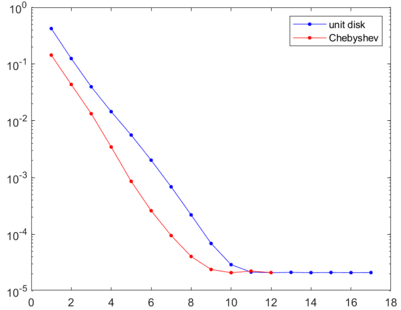


Figure 11: $G(\varepsilon)$

Tol=$d \times 10^{-6}$, $d = 21$, Points on unit disk and Chebyshev.

Numerical experiments $n = 10$ complex

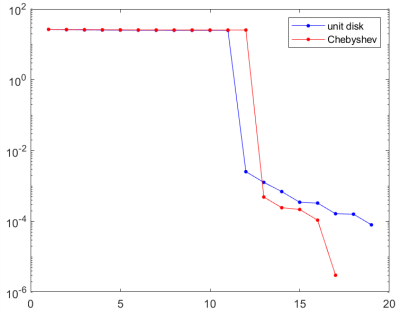


Figure 12: $\left|\varepsilon_{\text{up}} - \varepsilon_{\text{low}}\right|$

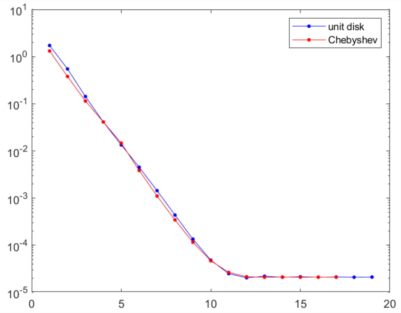


Figure 13: $G(\varepsilon)$

Tol=$d \times 10^{-6}$, $d = 21$, Points on unit disk and Chebyshev.