ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

MATHICSE

Mathematics Institute of Computational Science and Engineering

School of Basic Sciences - Section of Mathematics

MATHICSE Technical Report

Nr. 06.2015

March 2015

An algorithm for computing minimal Geršgorin sets

V. R. Kostić, A. Miedlar, LJ. Cvetković

http://mathicse.epfl.ch

Phone: +41 21 69 37648

Address:
EPFL - SB - MATHICSE (Bâtiment MA)
Station 8 - CH-1015 - Lausanne - Switzerland

Fax: +41 21 69 32545

AN ALGORITHM FOR COMPUTING MINIMAL GERŠGORIN SETS

V. R. KOSTIĆ*, A. MIĘDLAR†, AND LJ. CVETKOVIĆ *

Abstract. The first algorithms for computing the minimal Geršgorn set were developed by Varga et all. in [17] for the use on small and medium size (dense) matrices. Here, we first discuss the existing methods and present a new approach based on the modified Newton's method to find zeros of the parameter dependent left-most eigenvalue of a Z-matrix. Additionally, sampling technique used in the original work is replaced by a special curve tracing procedure. The advantages of the new approach are presented on several test examples that arise in practical applications.

Keywords: minimal Geršgorin set, eigenvalue localization, curve tracing

1. Introduction. A well known localization technique applicable to the general square matrices is by means of the *Geršgorin set* and its many generalizations known as *Geršgorin-type sets*, cf. [4, 5, 6, 9, 13, 14, 15, 16]. Among all Geršgorin-type sets, minimal Geršgorin set, in general, gives the sharpest localization of eigenvalues of a given matrix, i.e., is the smallest set among all Geršgorin-type sets containing the eigenvalues, cf. [9]. Besides this, it owes its name to the fact that it is "completely filled" with eigenvalues of matrices belonging to the family of matrices, called extended equimodular, cf. [13, 16]. For these reasons, minimal Geršgorin set is of particular interest and a computational method for its approximation has been proposed in [17]. However, several drawbacks of the algorithm proposed in [17] have been noticed, and in this paper we address them in detail.

First, in Section 2, we present some preliminary information on the minimal Geršgorin set, review two the algorithmic approaches from [17] and present them in more elaborate form. Namely, these two approaches are formulated as *bisection Minimal Geršgorin Set algorithm* (bMGS) and *griding Minimal Geršgorin Set algorithm* (gMGS) for the sake of their comparison and in order to presents their drawbacks. Then, using the formulas for the derivative of a simple eigenvalue of a given matrix A, in Section 3 we develop the main contribution of this paper - a new algorithm called *explicit Minimal Geršgorin Set* (eMGS). Instead of using bisection as in bMGS, we derive the Newton's method with stabilized iteration steps to obtain a faster convergence of the inner iterations. Since the algorithm eMGS depends heavily on eigenvalue computations, its applicability is practically limited to the small or medium size matrices. Finally, in Section 4, we illustrate the performance of algorithm eMGS comparing to bMGS and gMGS for several test examples and conclude with final remarks in Section 5.

2. Preliminaries. Given an arbitrary matrix $A = [a_{ij}] \in \mathbb{C}^{n,n}$, its *i-th Geršgorin disk* is defined by

(2.1)

and the union of all these disks, denoted by

$$\Gamma(A) := \bigcup_i \Gamma_i(A), \quad i = 1, \ldots, n \tag{2.2}$$

*Department of Mathematics and Informatics, Faculty of Science, University of Novi Sad, Trg D. Obradovića 4, 21000 Novi Sad, Serbia. vkostic@dmi.uns.ac.rs, lila@dmi.uns.ac.rs

†Institut für Mathematik, MA 4-5, Technische Universität Berlin, Strasse des 17. Juni,10625, Berlin. miedlar@math.tu-berlin.de

is called the *Geršgorin set* of matrix A.

A well-known result due to Geršgorin [8] states that $\Gamma(A)$ contains the spectrum $\sigma(A)$ of matrix A, i.e.,

$$\sigma(A) := \{\lambda \in \mathbb{C} : \det(\lambda I - A) = 0\} \subseteq \Gamma(A). \tag{2.3}$$

As one can readily check, while the spectrum is invariant under similarity transformations, the Geršgorin set is not. Moreover,

$$\sigma(A) = \bigcap_{\det(S)\neq 0} \Gamma(S^{-1}AS), \tag{2.4}$$

holds for any square matrix A. Hence, obtaining a localization set for the spectrum of matrix A can be seen as restricting the similarity transformation S in (2.4) to belong to a certain subfamily of nonsingular matrices.

In the special case, when S is nonsingular diagonal matrix, (2.4) becomes the minimal Geršgorin set introduced in [13, 16]. Namely, for any $x = [x_1, x_2, \ldots, x_n]^T > 0$ in $\mathbb{R}^n$, i.e., $x_i > 0$ for all $i = 1, \ldots, n$, let $X := diag[x_1, x_2, \ldots, x_n]$ denotes the associated nonsingular diagonal matrix. Then, with the Geršgorin disks for $X^{-1}AX$ given by

$$\Gamma_i^{r^x}(A) := \left\{ z \in \mathbb{C} : |z - a_{ii}| \leq r_i^x(A) := \sum_{j\neq i} \frac{|a_{ij}|x_j}{x_i} \right\}, \quad i, j = 1, \ldots, n, \tag{2.5}$$

and with the associated Geršgorin set,

$$\Gamma^{r^x}(A) := \bigcup_i \Gamma_i^{r^x}(A), \quad i = 1, \ldots, n \tag{2.6}$$

we have that

$$\sigma(A) \subseteq \Gamma^{\mathcal{R}}(A) := \bigcap_{x\in\mathbb{R}^n,\, x>0} \Gamma^{r^x}(A). \tag{2.7}$$

The set $\Gamma^{\mathcal{R}}(A)$ is called the *minimal Geršgorin set* and it gives the sharpest inclusion set for $\sigma(A)$, with respect to *all* positive diagonal similarity transformations $X^{-1}AX$ of A. This fact has also an interesting interpretation in terms of diagonally dominant matrices, for detailed explanations see [9]. Namely, the minimal Geršgorin set corresponds to the class of H-matrices, or, equivalently, the class of generalized diagonally dominant (GDD) matrices, while the original Geršgorin set (2.2) corresponds to the strictly diagonally dominant (SDD) matrices. Due to the monotonicity principle between the classes of matrices and the corresponding localization sets, cf. [9, Theorem 8], $\Gamma^{\mathcal{R}}(A)$ can be seen as the minimal of all Geršgorin-type sets. The sharpness of the minimal Geršgorin set can also be expressed in the following way, cf. [16, Theorem 4.5].

Given an arbitrary matrix $A = [a_{ij}] \in \mathbb{C}^{n,n}$, the family of matrices

$$\widehat{\Omega}(A) := \{B = [b_{ij}] \in \mathbb{C}^{n,n} : b_{ii} = a_{ii} \text{ and } |b_{ij}| \leq |a_{ij}| \text{ for } i \neq j,\ i, j = 1, \ldots, n\} \tag{2.8}$$

is called the *extended equimodular family* of matrix A. Obviously, for every $B \in \widehat{\Omega}(A)$, $\sigma(B) \subseteq \Gamma^{\mathcal{R}}(B) \subseteq \Gamma^{\mathcal{R}}(A)$. Moreover,

$$\sigma\big(\widehat{\Omega}(A)\big) := \bigcup_{B\in\widehat{\Omega}(A)} \sigma(B) = \Gamma^{\mathcal{R}}(A), \tag{2.9}$$

i.e., each point of $\Gamma^{\mathcal{R}}(A)$ is an eigenvalue of *some* matrix B in $\widehat{\Omega}(A)$.

For this reason, the minimal Geršgorin set can be seen as a kind of *"pseudospectrum"* of a matrix. Term pseudospectrum usually stands for ε-pseudospectrum, cf. [12], i.e., the set of all eigenvalues of matrices $A+\Delta$, where $\Delta \in \mathbb{C}^{n,n}$ and $\|\Delta\| \leq \varepsilon$, which is the union of all spectra of the additively perturbed matrices. On the other hand, the minimal Geršgorin set is the union of all the spectra of *frequency perturbed* matrices (here frequency is a synonym for the argument of a complex number). Therefore, we can consider the minimal Geršgorin set as a *"frequency-pseudospectra"*.

As an interesting application of this concept, let us observe a time varying linear dynamical system that arises in models of oscillatory systems.

$$\dot{x}(t) = A(t)x(t),\ t \geq 0, \tag{2.10}$$

where the system matrix has time invariant diagonal entries $a_{ii}(t) = \alpha_{ii} \in \mathbb{C}$, $1 \leq i \leq n$, while the its off-diagonal entires $a_{ij}(t)$ are complex valued analytic functions bounded by $\alpha_{i,j} > 0$, $1 \leq i \neq j \leq n$. Using the concept of the matrix measure in infinity norm and Coppel inequality, cf. [7], it is known that the dynamical system (2.10) is exponentially asymptotically stable if there exists $\mu > 0$ such that

$$\max_{1\leq i\leq n} \left\{ Re(\alpha_{ii}) + \sum_{j\neq i} \alpha_{ij} \right\} < -\mu, \text{ i.e.,}$$

if the Geršgorin set is situated in the open left half-plane of $\mathbb{C}$. Similarly, using the vector norms $\|X(\cdot)\|_\infty$, for $X := diag[x_1, x_2, \ldots, x_n] > 0$, we obtain the same conclusion if the minimal Geršgorin set is situated in the open left half-plane of $\mathbb{C}$. But, in this case, since $A(t) \in \widehat{\Omega}([\alpha_{ij}])$, the later result is a tight one.

However, unlike the Geršgorin set $\Gamma(A)$ or scaled Geršgorin set $\Gamma^{r^x}(A)$, the minimal Geršgorin set $\Gamma^{\mathcal{R}}(A)$ is, in general, hard to determine numerically. As mentioned above, the only two computational methods that provide numerical approximation of $\Gamma^{\mathcal{R}}(A)$ with a finite number of calculations are given in [17]. First, we provide their short description, and then use them as the starting point for our analysis.

The methods discussed in [17] are based on the following characterization of the minimal Geršgorin set.

THEOREM 1. *([16, Proposition 4.3]) Given a matrix $A = [a_{ij}] \in \mathbb{C}^{n,n}$, and $z \in \mathbb{C}$, define the real valued function*

$$\nu_A(z) := \inf_{x>0} \max_{i\in N} (r_i^x(A) - |z - a_{ii}|). \tag{2.11}$$

Then, $z \in \Gamma^{\mathcal{R}}(A)$ if and only if $\nu_A(z) \geq 0$, and if $\nu_A(z) = 0$, then $z \in \partial\Gamma^{\mathcal{R}}(A)$. Moreover, $\nu_A(z)$ can be obtained as the right-most eigenvalue of an essentially nonnegative matrix $Q_A(z) = [q_{ij}(z)]_{i,j=1,\ldots,n}$ defined by

$$q_{ij}(z) := |a_{ij}| \text{ and } q_{ii}(z) := -|z - a_{ii}|, \quad \text{for } i \neq j,\ i,j = 1,\ldots,n. \tag{2.12}$$

Namely, if

$$\mu_A(z) := \max_i |z - a_{ii}|, \quad i = 1,\ldots,n \tag{2.13}$$

then matrix $B_A(z) = Q(z) + \mu(z)I_n$ is a nonnegative matrix in $\mathbb{R}^{n,n}$. Furthermore, from the Perron-Frobenius theory of nonnegative matrices [2], matrix $B_A(z)$ possesses

a nonnegative real eigenvalue, $\rho(B_A(z))$, called the *Perron root*. Moreover, if A is an irreducible matrix, then $B_A(z)$ is also irreducible for all $z \in \mathbb{C}$, and $\rho(B_A(z))$ can be characterized using Wieland's formula, cf. [2], to obtain

$$Q_A(z)x = \nu_A(z)x. \tag{2.14}$$

It follows from $\nu_A(z) = \rho(B_A(z)) - \mu_A(z)$, that that $\nu_A(z)$ is the right-most eigenvalue of $Q_A(z)$.

Given a complex number $z \in \mathbb{C}$, this provides us with a method to compute the value $\nu_A(z)$ with a reasonable accuracy. In [17], assuming irreducibility of the matrix A, authors suggest to use power method on the matrix $B_A(z)$ and then derive $\nu_A(z)$. In our implementation, instead, we will compute $\nu_A(z)$ as a right-most eigenvalue of $Q_A(z)$ using the MATLAB function *eigs*. So, the brute-force algorithm for computing the Minimal Geršgorin set $\Gamma^{\mathcal{R}}(A)$ can be constructed simply by computing the function surface $\nu_A(z)$ on the rectangular grid. Then, the approximation of the boundary $\partial\Gamma^{\mathcal{R}}(A)$ is obtained as its zero-level curve. On the other hand, the boundaries of the rectangular grid can be easily determined since the Minimal Geršgorin set, as well as the Geršgorin set, belongs to the region $[l_{re}, u_{re}] \times [l_{im}, u_{im}]$, where

$$l_{re} := \max_i \left(Re(a_{ii}) - r_i(A) \right), \quad u_{re} := \max_i \left(Re(a_{ii}) + r_i(A) \right), \tag{2.15}$$

$$l_{im} := \max_i \left(Im(a_{ii}) - r_i(A) \right), \quad u_{im} := \max_i \left(Im(a_{ii}) + r_i(A) \right). \tag{2.16}$$

Therefore, we summarize this method in the following algorithm called *griding MGS* (gMGS). The parameters are the matrix A and the number n_g of the grid points per vertical and per horizontal axis.

Algorithm gMGS

Input: A, n_g

1: Compute l_{re}, u_{re}, l_{im} and u_{im} using (2.15) and (2.16);
2: Set $\delta_x = \frac{u_{re} - l_{re}}{n_g}$ and $\delta_y = \frac{u_{im} - l_{im}}{n_g}$;
3: **for** $k_x = 0 : n_g$ **do**
4: **for** $k_y = 0 : n_g$ **do**
5: Set $z := (l_{re} + k_x\delta_x) + \imath(l_{im} + k_y\delta_y)$;
6: Compute $G(k_x, k_y) := \nu_A(z)$ as the r.m.e. of $Q_A(z)$;
7: **end for**
8: **end for**
9: Compute the zero level set $\mathcal{C}$ using the matrix G;

Output: $\mathcal{C}$

As expected, the algorithm demands for a large number of eigenvalue computations due to the fine gridding necessary to localize all the eigenvalues. To avoid this, the authors of [17] have developed another numerical procedure that is based on the following theorem.

THEOREM 2. *[16, Theorem 4.6] Given an arbitrary irreducible matrix* $A = [a_{ij}]_{i,j=1,\ldots,n} \in \mathbb{C}^{n,n}$, $n \geq 2$, *for every* $i = 1, \ldots, n$, $\nu_A(a_{ii}) > 0$. *Moreover, for each* $i = 1, \ldots, n$ *and each real* θ, $0 \leq \theta \leq 2\pi$, *there exists a largest number* $\widehat{\varrho}_i(\theta) > 0$ *such that*

$$\nu_A(a_{ii} + \widehat{\varrho}_i(\theta)e^{\imath\theta}) = 0 \ \text{ and } \ \nu_A(a_{ii} + te^{\imath\theta}) \geq 0, \ \text{ for all } \ 0 \leq t < \widehat{\varrho}_i(\theta), \tag{2.17}$$

i.e., the entire complex interval $[a_{ii} + te^{\imath\theta}]_{t=0}^{\widehat{\varrho}_i(\theta)}$ *is contained* $\Gamma^{\mathcal{R}}(A)$. This theorem implies that for each $i = 1, \ldots, n$,

$$\bigcup_{\theta=0}^{2\pi} [a_{ii} + te^{\imath\theta}]_{t=0}^{\widehat{\varrho}_i(\theta)} \tag{2.18}$$

is a *star-shaped* subset of $\Gamma^{\mathcal{R}}(A)$ with

$$a_{ii} + \widehat{\varrho}_i(\theta)e^{\imath\theta} \in \partial\Gamma^{\mathcal{R}}(A). \tag{2.19}$$

Thus, we can obtain the approximation of the boundary of the minimal Geršgorin set starting from each diagonal entry of matrix A and computing the points on its boundary for several, e.g. m, angles θ.

To that end, first observe that, [16, Exercise 7, p.108],

$$|\nu_A(z) - \nu_A(z')| \leq |z - z'|, \text{ for every } z, z' \in \mathbb{C}, \tag{2.20}$$

which implies that

$$\widehat{\varrho}_i(\theta) \geq \nu_A(a_{ii}) > 0, \text{ for } i = 1, \ldots, n. \tag{2.21}$$

Then, for a fixed step $\delta > 0$ and a sufficiently large $\ell_{ik} \in \mathbb{N}$, $i = 1, \ldots, n, 1 \leq k \leq m$

$$a_{ii} + (\nu_A(a_{ii}) + \ell_{ik}\delta)e^{\imath\theta} \notin \Gamma^{\mathcal{R}}(A),$$

and we have that

$$\widehat{\varrho}_i(\theta) \in [\nu_A(a_{ii}) + (\ell_{ik} - 1)\delta, \nu_A(a_{ii}) + \ell_{ik}\delta].$$

Since (2.20) implies that ν_A is uniformly continuous, a bisection method to obtain such $\widehat{\varrho}_i(\theta)$ is proposed in [17]. We present this method in a procedure **bSearch**.

Finally, we recall the famous result of Olga Taussky [11], on a sharpening of the Geršgorin Circle Theorem, which serves as a final point in approximating the minimal Geršgorin set.

THEOREM 3. *[11] Let* $A = [a_{ij}]_{i,j=1,\ldots,n}$ *be an irreducible matrix in* $\mathbb{C}^{n,n}$. *If, for each* $i = 1, \ldots, n$, $\lambda \in \sigma(A)$ *is such that* $\lambda \notin int\, \Gamma_i(A)$, *i.e.,*

$$|\lambda - a_{ii}| \geq r_i(A),$$

then

$$|\lambda - a_{ii}| = r_i(A), \text{ for each } i = 1, \ldots, n, \tag{2.22}$$

i.e., each *Geršgorin circle* $\{z \in \mathbb{C} : |z - a_{ii}| = r_i(A)\}$ *passes through* λ. Let ω_{jk}, $j = 1, \ldots, n$ and $1 \leq k \leq m$ (m is the number of angular directions), be a boundary point obtained as described above. Then, according to Theorem 3, the associated Geršgorin set consisting of the union of n Geršgorin disks can be expressed as

$$\Gamma^{\omega_{jk}}(A) := \bigcup_i \{z \in \mathbb{C} : g_i^A(\omega_{jk}, z) \leq 0\}, \quad i = 1, \ldots, n, \tag{2.23}$$

where

$$g_i^A(\omega_{jk}, z) := |z - a_{ii}| - |\omega_{jk} - a_{ii}|. \tag{2.24}$$

bSearch

Input: A, ξ, θ, tol

Set $z = \xi$ and compute $f := \nu_A(z)$ as the r.m.e. of $Q_A(z)$;
Set $\delta := f$ and $\ell := 0$;
while $f > 0$ **do**
  $\ell := \ell + 1$;
  Set $z = \xi + \ell\delta e^{\imath\theta}$ and compute $f = \nu_A(z)$ as the r.m.e. of $Q_A(z)$;
end while
Set $a := (\ell - 1)\delta$ and $b := \ell\delta$;
while $b - a > tol$ **do**
  Set $z := \xi + \frac{a+b}{2} e^{\imath\theta}$ and compute $f := \nu_A(z)$ as the r.m.e. of $Q_A(z)$;
  if $f > 0$ **then**
    $a := \frac{a+b}{2}$;
  else
    $b := \frac{a+b}{2}$;
  end if
end while
Set $\omega = \xi + \frac{a+b}{2} e^{\imath\theta}$;

Output: ω

Hence, the finite intersection

$$\widehat{\Gamma}^{\mathcal{R}}(A) := \Gamma(A) \cap \bigcap_{j=1}^{n} \bigcap_{k=1}^{m} \Gamma^{\omega_{jk}}(A), \tag{2.25}$$

gives an approximation of $\Gamma^{\mathcal{R}}(A)$, such that $\Gamma^{\mathcal{R}}(A) \subseteq \widehat{\Gamma}^{\mathcal{R}}(A)$ and for which $n \cdot m$ points on the boundary of $\widehat{\Gamma}^{\mathcal{R}}(A)$ are the *boundary points* of $\Gamma^{\mathcal{R}}(A)$.

Based on this, we can construct the procedure **gerApprox**, essentially introduced in [17], that uses the sample $\{\omega_{jk}\}$, $j = 1, \ldots, n$ and $1 \leq k \leq m$, of the boundary points to construct $\widehat{\Gamma}^{\mathcal{R}}(A)$ - an approximation of the Minimal Geršgorin set based on Geršgorin sets. The approximation is constructed on an equally spaced rectangular grid of n_g^2 points where parameter n_g is provided.

This allows us now to present the **bMGS** algorithm essentially given in [17]. Given an irreducible matrix $A = [a_{ij}] \in \mathbb{C}^{n,n}$, some tolerance $tol > 0$, an integer $m \geq 1$, and the number n_g of the grid points (vertical and horizontal), the algorithm bMGS for approximating the minimal Geršgorin set $\Gamma^{\mathcal{R}}(A)$ computes $n \cdot m$ points $\{\omega_{jk}\}_{1 \leq j \leq n,\, 1 \leq k \leq m}$ on the boundary of $\Gamma^{\mathcal{R}}(A)$. For each diagonal entry $a_{ii}, i = 1, \ldots, n$ of matrix A, m such boundary points are computed using the **bSearch** procedure that performs bisection on the intervals $[(\ell_{ik} - 1)\,\nu_A(a_{ii}),\ \ell_{ik}\,\nu_A(a_{ii})]$ for $i = 1, \ldots, n$ and $1 \leq k \leq m$. Then, for the set of such boundary points $\{\omega_{jk}\}$, the boundary $\partial\widehat{\Gamma}^{\mathcal{R}}(A)$ of the intersection (2.25) is determined via Geršgorin sets.

The main drawback of this algorithm is the bisection method which needs to be performed $n \cdot m$ times. It is not only the slow convergence of the method, but also the fact that we need several (ℓ_{ik}) pre-computations of the right-most eigenvalue ν_A, in order to determine the upper limit of the starting bisection interval. Therefore, the whole procedure suffers from repeated eigenvalue computations which are especially

gerApprox

Input: A, $\{\omega_{jk}\}_{\substack{1\le j\le n\\ 1\le k\le m}}$, n_g

Compute l_{re}, u_{re}, l_{im} and u_{im} using (2.15) and (2.16);
Set $\delta_x = \frac{u_{re}-l_{re}}{n_g}$ and $\delta_y = \frac{u_{im}-l_{im}}{n_g}$;
for $k_x = 0 : n_g$ **do**
  for $k_y = 0 : n_g$ **do**
    Set $z := (l_{re} + k_x\delta_x) + \imath(l_{im} + k_y\delta_y)$;
    Compute $G(k_x, k_y) := \max_{i=1\ldots n}\{|z - a_{ii}| - r_i(A)\}$;
    for $j = 1 : n$ and $k = 1 : m$ **do**
      Set $G_{jk}(k_x, k_y) = g_1^A(\omega_{jk}, z)$ using (2.24);
      for $i = 2 : n$ **do**
        Update $G_{jk}(k_x, k_y) \leftarrow \min\left\{G_{jk}(k_x, k_y), g_i^A(\omega_{jk}, z)\right\}$ using (2.24);
      end for
      Update $G(k_x, k_y) \leftarrow \max\{G(k_x, k_y), G_{jk}(k_x, k_y)\}$;
    end for
  end for
end for
Compute the zero level set $\mathcal{C}$ using the matrix G;

Output: $\mathcal{C}$

Algorithm bMGS

Input: A, m, tol, n_g

for $j = 1 : n$ and $k = 1 : m$ **do**
  Set $\theta_k = k\frac{2\pi}{m}$
  Run procedure **bSearch**(A, a_{jj}, θ_k, tol) to compute ω_{jk};
end for
Run procedure **gApprox**(A, $\{\omega_{jk}\}_{\substack{1\le j\le n\\ 1\le k\le m}}$, n_g) to compute $\mathcal{C}$;

Output: $\{\omega_{jk}\}_{\substack{1\le j\le n\\ 1\le k\le m}}$, $\mathcal{C}$

challenging for even medium size matrices. Furthermore, using a small step $\delta > 0$ while searching for the upper limit of the bisection interval $\nu_A(a_{jj}) + \ell_{jk}\delta$ may result in skipping the first zero of ν_A on the ray $\{a_{jj} + te^{\imath\theta_k}\}_{t>0}$ and ending up in the point of the boundary which does not define a star-shaped set (2.19). Finally, the algorithm was developed only for irreducible matrices, and therefore not applicable for a large class of problems that arise in practice, like, for example, block diagonal matrices.

In order to avoid these two drawbacks, in the following, we develop new algorithm called the *explicit* MGS algorithm, (eMGS), also intended for small and medium size matrices.

3. Explicit algorithm for computing the minimal Geršgorin set. First, let us start with the observation that the minimal Geršgorin set of a general complex matrix can be expressed as the (finite) union of the minimal Geršgorin sets of irreducible matrices of the smaller size.

Namely, given an arbitrary $A \in \mathbb{C}^{n,n}$, $n \ge 2$, there always exists its *normal*

reduced form, cf. [16, Equation (1.20)],

$$PAP^T = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1m} \\ & A_{22} & \cdots & A_{1m} \\ & & \ddots & \vdots \\ & & & A_{mm} \end{bmatrix}, \tag{3.1}$$

where P is a suitable permutation matrix and diagonal blocks $A_{ii} \in \mathbb{C}^{n_i,n_i}$, $1 \leq i \leq m$, are either 1×1 or irreducible $n_i \times n_i$, $n_i \geq 2$, matrices. Then, the following lemma holds.

LEMMA 1. *Given an arbitrary matrix $A = [a_{ij}] \in \mathbb{C}^{n,n}$, let $A_{ii} \in \mathbb{C}^{n_i,n_i}$, $n_i \geq 1$, $1 \leq i \leq m$, be the diagonal blocks of its normal reduced form (3.1). Then,*

$$\Gamma^{\mathcal{R}}(A) = \bigcup_{i=1}^{m} \Gamma^{\mathcal{R}}(A_{ii}).$$

Proof. First, observe that the minimal Geršgorin set is invariant under simultaneous permutations of rows and columns. Therefore, we may assume that A is given in its normal reduced form (3.1). Next, given $k \in \mathbb{N}$, define diagonal matrix $X_k = diag[x_1^{(k)}, x_2^{(k)}, \ldots, x_n^{(k)}]$ such that $x_i^{(k)} := j^{-k}$, $i \in N$, whenever index i belongs to the set of indices corresponding to the matrix A_{jj}. Then, $\lim_{k\to\infty} X_k^{-1} A X_k = diag[A_{11}, A_{22}, \ldots, A_{mm}]$, and, consequently,

$$\bigcap_{k\in\mathbb{N}} \Gamma(X_k^{-1} A X_k) = \bigcup_{i=1}^{m} \Gamma(A_{ii}),$$

which completes the proof. □

Therefore, in the rest of the paper, without loss of generality, we assume that A is an irreducible matrix and derive the algorithm for its numerical computation. To that end, first, we introduce an explicit (**eSearch**) procedure, to replace the bisection method implemented in **bSearch**. The main reason for the previous use of bisection was the fact that the ν_A is a uniformly continuous function. This property is, however, not sufficient for the construction of a faster method based on Newton's iterations. Nevertheless, since $\nu_A(z)$ can be characterized as the shifted Perron root of a nonnegative irreducible matrix, we observe that $\nu_A(z)$ is a simple eigenvalue of $Q_A(z)$, cf. [2], and, as such, it is a differentiable function of the entries of matrix $Q_A(z)$, cf. [10, Theorem 5 and 9]. Therefore, we can easily obtain the first order derivative of $\nu_A(z)$ and construct a modified Newton's method to compute the boundary points of the minimal Geršgorin set, as follows.

LEMMA 2. *Given an arbitrary irreducible matrix $A = [a_{ij}] \in \mathbb{C}^{n,n}$, a complex number ξ and a real θ, $0 \leq \theta < 2\pi$, let us define a function $f_A^{\xi,\theta} : \mathbb{R}_0^+ \to \mathbb{R}$ such that*

$$f_A^{\xi,\theta}(t) := \nu_A(\xi + t\, e^{\imath\theta}).$$

Then, $f_A^{\xi,\theta}$ is ∞-differentiable at $t \notin \{(\xi - a_{ii})e^{-\imath(\pi-\theta)} \ : \ i = 1, 2, \ldots, n\}$, and its first derivative is given as

$$\frac{\partial}{\partial t} f_A^{\xi,\theta}(t) = -\frac{y(t)^T\, D_A^{\xi,\theta}(t)\, x(t)}{y(t)^T x(t)}, \tag{3.2}$$

where $x(t)$ and $y(t)$ are the right and left eigenvectors of $Q_A(\xi+t\,e^{\imath\theta})$ corresponding to the eigenvalue $f_A^{\xi,\theta}(t)$, respectively, and the diagonal matrix $D_A^{\xi,\theta}(t) := diag[d_1(t), d_2(t), \ldots, d_n(t)]$ is given by

$$d_i(t) := \frac{Re[(\xi - a_{ii})\, e^{-\imath\theta}] + t}{|(\xi - a_{ii})\, e^{-\imath\theta} + t|}, \quad i = 1, 2, \ldots, n. \tag{3.3}$$

Proof. Assuming that $t \notin \{(\xi - a_{ii})e^{-\imath(\pi-\theta)} \ : \ i = 1, 2, \ldots, n\}$, we have that for all $i = 1, 2 \ldots n$, $|\xi + te^{\imath\theta} - a_{ii}| = |(\xi - a_{ii})\, e^{-\imath\theta} + t| \neq 0$, which according to (2.12), assures that the entries of $Q_A(\xi + t\, e^{\imath\theta})$ are ∞-differentiable functions in t, i.e.,

$$\frac{\partial}{\partial t} q_{ii}(\xi + te^{\imath\theta}) = -\frac{Re[(\xi - a_{ii})\, e^{-\imath\theta}] + t}{|(\xi - a_{ii})\, e^{-\imath\theta} + t|} \ \text{ and } \ \frac{\partial}{\partial t} q_{ij}(\xi + te^{\imath\theta}) = 0,$$

for all $i, j = 1, 2, \ldots, n$ and $i \neq j$, implying that $\frac{\partial}{\partial t} Q_A(\xi + t\, e^{\imath\theta}) = -D_A^{\xi,\theta}(t)$. Next, the fact that matrix A is irreducible implies irreducibility of matrix $Q_A(\xi + t\, e^{\imath\theta})$, and, therefore, assures that $f_A^{\xi,\theta}(t)$ is its simple eigenvalue. Now, denoting the corresponding right and left eigenvectors by $x(t)$ and $y(t)$, respectively, we can use a well known result on differentiability of simple eigenvalues, cf. [10, Theorem 5], and obtain that $f_A^{\xi,\theta}$ is ∞-differentiable at t, i.e.,

$$\frac{\partial}{\partial t} f_A^{\xi,\theta}(t) = \frac{y(t)^T \, \frac{\partial}{\partial t} Q_A(\xi + t\, e^{\imath\theta})\, x(t)}{y(t)^T x(t)} = -\frac{y(t)^T \, D_A^{\xi,\theta}(t)\, x(t)}{y(t)^T x(t)}.$$

□

REMARK 1. *Note here, that $f_A^{\xi,\theta}(t)$ is not differentiable function only if $\xi + te^{\imath\theta}$ is diagonal entry of the matrix A.*

This theoretical result allows us to formulate the modified Newton's method for computing zeros of the function $f_A^{\xi,\theta}$. Namely, let $t_0 := f_A^{\xi,\theta}(\xi) > 0$, and define the sequence $\{t_k\}_{k=1,2,\ldots}$ with

$$t_{k+1} := t_k + \gamma_k \Delta_k, \quad k = 1, 2, \ldots, \tag{3.4}$$

where Δ_k is defined as

$$\Delta_k := \begin{cases} -\dfrac{f_A^{\xi,\theta}(t_k)}{\frac{\partial}{\partial t} f_A^{\xi,\theta}(t_k)}, & \text{if } \frac{\partial}{\partial t} f_A^{\xi,\theta}(t_k) < 0, \\ \\ f_A^{\xi,\theta}(t_k), & \text{otherwise,} \end{cases} \tag{3.5}$$

and $\gamma_k = 1$, when $f_A^{\xi,\theta}(t_{k+1}) \geq 0$, or $\gamma_k = \tau^{q_k}$, otherwise, with $\tau \in (0, 1)$ and the smallest $q_k \in \mathbb{N}$ such that

$$f_A^{\xi,\theta}(t_k + \tau^{q_k}\, \Delta_k) > 0 \quad \text{and} \quad f_A^{\xi,\theta}(t_k + \tau^{q_k - 1}\, \Delta_k) < 0. \tag{3.6}$$

LEMMA 3. *Given an arbitrary irreducible matrix $A = [a_{ij}] \in \mathbb{C}^{n,n}$, a complex number ξ and a real θ, $0 \leq \theta < 2\pi$, a sequence $\{t_k\}_{k=1,2,\ldots}$ is well defined by (3.4), and it converges to $\widehat{t} > 0$ such that $f_A^{\xi,\theta}(\widehat{t}\,) = 0$.*

Proof. First, observe that the uniform continuity of ν_A expressed in (2.20) implies that for every $t > 0$ and every $\varepsilon \geq 0$

$$|f_A^{\xi,\theta}(t+\varepsilon) - f_A^{\xi,\theta}(t)| \leq \varepsilon, \tag{3.7}$$

and, consequently, whenever $f_A^{\xi,\theta}(t)$ is differentiable,

$$\left| \frac{\partial}{\partial t} f_A^{\xi,\theta}(t) \right| \leq 1. \tag{3.8}$$

Therefore, for all $k = 1, 2, \ldots$, $\Delta_k \geq f_A^{\xi,\theta}(t_k)$.
Now, we show that the sequence $\{t_k\}_{k=1,2,\ldots}$, is well defined. First, since $f_A^{\xi,\theta}(t_0) = \nu_A(\xi + \nu_A(\xi)\, e^{\imath\theta})$, (2.21) implies that $f_A^{\xi,\theta}(t_0) \geq 0$, and consequently $\Delta_0 \geq 0$. Thus, $t_0 + \Delta_0 \geq t_0 > 0$, and $f_A^{\xi,\theta}(t_0+\Delta_0)$ is well defined. If $f_A^{\xi,\theta}(t_0+\Delta_0) < 0$, the continuity of ν_A, i.e., $f_A^{\xi,\theta}$, implies that there exists $q_0 = 1, 2, \ldots$, such that $\gamma_0 = \tau^{q_0}$. Otherwise, if $f_A^{\xi,\theta}(t_0+\Delta_0) \geq 0$, $\gamma_0 = 1$. By this, we have obtained $t_1 := t_0 + \gamma_0 \Delta_0 \geq t_0 > 0$ such that $f_A^{\xi,\theta}(t_1) \geq 0$. Continuing in the same way, we obtain that the sequence $\{t_k\}_{k=1,2,\ldots}$ is well defined and that $f_A^{\xi,\theta}(t_k) \geq 0$ with $t_k \geq t_{k-1} > 0$ for all $k = 1, 2 \ldots$. In order to prove the convergence of this monotonically nondecreasing sequence $\{t_k\}_{k=1,2,\ldots}$, we will show that it is bounded.
Let us assume, that $\{t_k\}_{k=1,2\ldots}$ is unbounded. Then, for some $\ell \in \mathbb{N}$, there exists a subsequence $\{t_{k_\ell}\}$, such that $\lim_{\ell\to\infty} t_{k_\ell} = +\infty$. From the construction of the sequence we have that $f_A^{\xi,\theta}(t_{k_\ell}) > 0$, implying that $z_\ell := \xi + t_{k_\ell}\, e^{\imath\theta} \in \Gamma^{\mathcal{R}}(A)$, for all $\ell \in \mathbb{N}$, see Theorem 1. However, since $|z_\ell| \to +\infty$, this contradicts the fact that the minimal Geršgorin set is a compact set in $\mathbb{C}$ and proof the boundedness of the sequence. As a result, the sequence is convergent, and we denote its limit as $\widehat{t} = \lim_{k\to\infty} t_k$. Now, we distinguish two cases.
First, let $\liminf_{k=1,2,\ldots} \gamma_k > 0$. Then

$$0 \leq f_A^{\xi,\theta}(t_k) \leq \Delta_k = \frac{t_{k+1} - t_k}{\gamma_k},$$

and taking the limit for $k \to \infty$ yields $f_A^{\xi,\theta}(\widehat{t}\,) = 0$.
Second, let $\lim_{\ell\to\infty} \gamma_{k_\ell} = 0$. Then, obviously, $\gamma_{k_\ell} = \tau^{q_{k_\ell}} < 1$, for $\ell = 1, 2, \ldots$, and $\lim_{\ell\to\infty} q_{k_\ell} = +\infty$. Then (3.6) implies that $0 \leq f_A^{\xi,\theta}(\widehat{t}\,) \leq 0$, which completes the proof. □ Using the results of Lemma 2 and Lemma 3, we introduce the procedure **eSearch** which defines the explicit MGS algorithm and state the main result of this section.

THEOREM 4. *Given an arbitrary irreducible matrix $A = [a_{ij}] \in \mathbb{C}^{n,n}$, a complex number ξ and real numbers θ and τ, $0 \leq \theta < 2\pi$, $\tau \in (0,1)$, the procedure **eSearch** produces $\omega = \xi + \widehat{t}\, e^{\imath\theta} \in \mathbb{C}$ such that $\omega \in \partial\Gamma^{\mathcal{R}}(A)$. Furthermore, if $\frac{\partial^2}{\partial t^2} f_A^{\xi,\theta}(\widehat{t}\,) > 0$, then the convergence rate of the **eSearch** procedure is locally quadratic, otherwise, the convergence is linear with the convergence rate* $\lim\limits_{k\to\infty} \sup(1 - \tau^{q_k})$.

Proof. According to Lemma 3, lines 1–15 of the procedure **eSearch**, see Appendix A.3, generate a sequence $\{t_k\}_{k=1,2,\ldots}$, such that $f_A^{\xi,\theta}(\widehat{t}\,) = 0$, with $\widehat{t} = \lim_{k\to\infty} t_k$. If the local minimum of the function $f_A^{\xi,\theta}$ is obtained in $\widehat{t}$, then $\xi + \widehat{t}\, e^{\imath\theta}$ is not necessarily a point on the boundary of $\Gamma^{\mathcal{R}}(A)$. For that reason, lines 16–19 are added to check if $f_A^{\xi,\theta}$ changes it sign in $\widehat{t}$. If this is not the case, a restart is performed as long as this condition is not satisfied. Finally, we obtain $\widehat{t} > 0$, as in Theorem 2, such that

eSearch

Input: A, ξ, θ, τ, tol

1: Set $z = \xi$ and compute r.m.e.-triplet (f, x, y) of $Q_A(z)$;
2: Set $t := f$ and $\Delta = 2 \cdot tol$
3: **while** $\Delta > tol$ **do**
4: Compute $df := -x^T D_A^{\xi}(t)\, y$, for $D_A^{\xi,\theta}(t)$ given in (3.3)
5: **if** $df > 0$ **then**
6: Set $df := -1$
7: **end if**
8: Set $\Delta := -f/df$ and $\gamma := 1$;
9: Set $z = \xi + (t + \gamma\,\Delta)e^{\imath\theta}$ and compute r.m.e.-triplet (f, x, y) of $Q_A(z)$;
10: **while** $f < 0$ **do**
11: $\gamma := \gamma \cdot \tau$;
12: Set $z = \xi + (t + \gamma\,\Delta)e^{\imath\theta}$ and compute r.m.e.-triplet (f, x, y) of $Q_A(z)$;
13: **end while**
14: Set $t := t + \gamma\,\Delta$;
15: **end while**
16: Set $z = \xi + (t + 2\,tol)e^{\imath\theta}$ and compute r.m.e.-triplet (f, x, y) of $Q_A(z)$;
17: **if** $f > 0$ **then**
18: Set $t := t + 2\,tol$ and $\Delta := 2\,tol$
19: **go to 3**
20: **end if**
21: Set $\omega = \xi + t\, e^{\imath\theta}$;

Output: ω, x

$\omega = \xi + \widehat{t}\, e^{\imath\theta} \in \mathbb{C}$. Notice here, that a restart is not likely to happen too often. Namely, in the case when we approach the double zero of $f_A^{\xi,\theta}$, the Newton's iteration produces a point in which both $f_A^{\xi,\theta}$ and $\frac{\partial}{\partial t} f_A^{\xi,\theta}$ are positive, and, according to (3.5), it will continue to converge non-decreasingly far away from the local minimum. Therefore, it remains to prove the local quadratic convergence rate of the introduced scheme.

Let us assume that $\widehat{t} = \lim_{k\to\infty} t_k$, $f_A^{\xi,\theta}(\widehat{t}) = 0$, and there exists $\varepsilon > 0$ such that $f_A^{\xi,\theta}$ is monotonically decreasing in the $\varepsilon-$neighborhood of $\widehat{t}$. Thus, $\frac{\partial}{\partial t} f_A^{\xi,\theta}(t) < 0$, for $|t - \widehat{t}| < \varepsilon$, and, consequently, there exists $k_0 \in \mathbb{N}$ such that for all $k \geq k_0$, $\Delta_k = -\frac{f_A^{\xi,\theta}(t_k)}{\frac{\partial}{\partial t} f_A^{\xi,\theta}(t_k)}$, i.e., the sequence $\{t_k\}_{k\geq k_0}$ is produced by the damped Newton's method.
Since $\frac{\partial^2}{\partial t^2} f_A^{\xi,\theta}(t_k) > 0$ for $k \geq k_0$, $f_A^{\xi,\theta}$ is a locally convex function and $\gamma_k = 1$. Therefore, the Newton's method is locally undamped, which implies the quadratic convergence. Otherwise, let $\frac{\partial^2}{\partial t^2} f_A^{\xi,\theta}(t) \leq 0$ for $|t - \widehat{t}| < \varepsilon$. Then, it is easy to verify that for some t_k^*, $\widehat{t} - t_k < t_k^* < \widehat{t}$,

$$\frac{\widehat{t} - t_{k+1}}{\widehat{t} - t_k} = (1 - \tau^{q_k}) + \frac{1}{2}\frac{\partial^2}{\partial t^2} f_A^{\xi,\theta}(t_k^*)(\widehat{t} - t_k),$$

and, as a consequence,

$$\lim_{k\to\infty} \sup \frac{\widehat{t} - t_{k+1}}{\widehat{t} - t_k} = \lim_{k\to\infty} \sup(1 - \tau^{q_k}),$$

which completes the proof. □

Now, having developed eSearch to improve bSearch, we continue with reducing the number of eigenvalue computations by avoiding unnecessary computations of the star-shaped subsets. Namely, we need to compute only one star-shaped subset (2.18) for each disjoint component of the minimal Geršgorin set, so the number of the overall **eSearch** calls in most of the cases can be reduced from $n \cdot m$ used by bMGS. But, since there is no simple way to test if the already approximated star-shaped subset is disjoint from another possibly existing component, we use the centers of the star-shaped subsets, i.e., diagonal entries of the matrix, to reduce the number of computations whenever it is possible.

The idea is the following. Starting with a set of all nonequal diagonal entries of the matrix A sorted such that the leftmost diagonal entries come first

$$\mathcal{L} := \{a_{i_1 i_1}, a_{i_2 i_2}, \ldots, a_{i_{\widetilde{n}} i_{\widetilde{n}}}\}, \tag{3.9}$$

and denoting the number of the components of $\Gamma^{\mathcal{R}}(A)$ by $s \leq \widetilde{n} \leq n$, the algorithm will, for each $i = 1, \ldots, \widetilde{n}$ construct a polygon of points (monotone x-wise and y-wise) in $\mathbb{C}$, denoted by $\{\omega_{i,j}\}_{j=1}^{m_i}$, where $\omega_{i,m_i+1} := \omega_{i,1}$, such that $dist(\omega_{i,j}, \partial\Gamma^{\mathcal{R}}(A)) < \varepsilon_1$ and $|\omega_{i,j+1} - \omega_{i,j}| < \varepsilon_2$, $(j = 1, 2, \ldots, m_i)$, where $\varepsilon_1 > 0$ is the maximum distance allowed between the polygon nodes and points on the boundary of the minimal Geršgorin set $\Gamma^{\mathcal{R}}(A)$ and $\varepsilon_2 > 0$ is the maximum distance allowed between successive nodes of the same polygon.

Since in practice we don't know the exact number of existing disjoint components, our method starts first with the leftmost entry of $\mathcal{L}$, i.e., $a_{i_1 i_1}$ and constructs a polygon $\{\omega_{1,j}\}_{j=1}^{m_1}$ that approximates the boundary of the disjoint component of $\Gamma^{\mathcal{R}}(A)$ that contains $a_{i_1 i_1}$. Secondly, using a winding number algorithm, cf. [1], we test which entries from $\mathcal{L}$ are contained in the constructed polygon. The set of such entries we denote by $\mathcal{S}_1 \subseteq \mathcal{L}$. Then, if $\mathcal{L} \backslash \mathcal{S}_1 \neq \emptyset$ we take the existing entry and construct the new polygon $\{\omega_{2,j}\}_{j=1}^{m_2}$ that approximates the boundary of the next disjoint component that contains diagonal entries of A that are in $\mathcal{S}_2 \subseteq \mathcal{L} \backslash \mathcal{S}_1$. This procedure we continue until all entries of $\mathcal{L}$ have been considered, i.e., $\mathcal{L} = \mathcal{S}_1 \dot{\cup} \mathcal{S}_2 \dot{\cup} \ldots \dot{\cup} \mathcal{S}_s$.

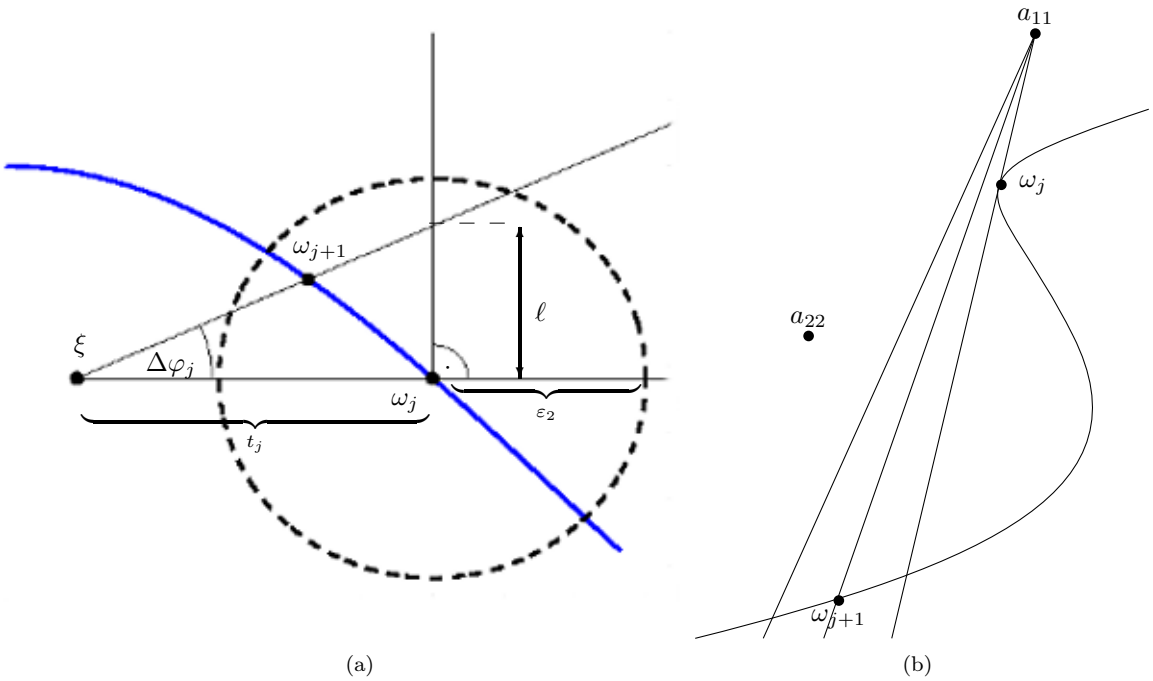
In the following, we describe how to construct a polygon that approximates one disjoint component. Without loss of generality, and for the sake of notational simplicity, let us assume that the $\Gamma^{\mathcal{R}}(A)$ has only one component and all distinct diagonal entries are ordered such that $Re(a_{11}) \leq Re(a_{22}) \leq \ldots \leq Re(a_{nn})$, so we produce a polygon $\{\omega_j\}_{j=1}^{m}$.

Starting with $\xi := a_{11}$ and with $\varphi_1 := -\pi$, using t_1=eSearch(A,ξ,φ_1,ε_1) we construct the first boundary point $\omega_1 := \xi + t_1 e^{i\varphi_1}$. Then, we increase the angle φ_1 by

$$\Delta\varphi_1 := \arctan \frac{\ell}{t_1}$$

to $\varphi_2 := \varphi_1 + \Delta\varphi_1$, where $\ell := \tau\varepsilon_2$ and $0 < \tau < 1$ is the predetermined parameter of the method, c.f. Figure 1(a), and construct a new point $\omega_2 := \xi + t_2 e^{i\varphi_2}$ using t_2 =eSearch(A,ξ,φ_2,ε_1). Proceeding in the same way, as long as the angle $\varphi_m \leq \pi$, we generate a finite sequence of points ω_j, $j = 1, 2, \ldots, m$ such that $|\omega_{j+1} - \omega_j| < \varepsilon_2$, and the set $\Gamma^{\mathcal{R}}(A)$ is approximated by the polygon $\{\omega_j\}_{j=1}^{m}$.

However, due to large variety of the geometrical properties of $\partial\Gamma^{\mathcal{R}}(A)$, finding a new point, even for infinitely small $\ell > 0$ is not always possible, c.f., Figure 1(b). By the following lemma, we can overcome this difficulty by changing ξ to be a different diagonal entry of the matrix A.



(a) (b)

Fig. 1: (a) illustration for parameters $\Delta\varphi_j, t_j$, ℓ and ε_2, (b) changing $\xi = a_{11}$ to $\xi = a_{22}$.

LEMMA 4. *Given an irreducible matrix $A \in \mathbb{C}^{n,n}$, for every point $\omega \in \partial\Gamma^{\mathcal{R}}(A)$ there exists a sufficiently small $\varepsilon > 0$ and an index $1 \leq i \leq n$ such that for all $\alpha \in [0,1]$ and all $z \in \mathbb{C}$ satisfying $|z - \omega| < \varepsilon$, $arg(z - a_{ii}) > arg(\omega - a_{ii})$ and $z \in \partial\Gamma^{\mathcal{R}}(A)$, it holds that $\alpha z + (1-\alpha)a_{ii} \in \Gamma^{\mathcal{R}}(A)$.*

Proof. First, without the loss of generality, assume that all diagonal entries of the matrix A are distinct, and define the family of matrices $A(t) := D - tB$ for $t \in [0,1]$, where $D = diag(a_{11}, a_{22}, \ldots, a_{nn})$ and $A = D - B$. Furthermore, given a $z \in \mathbb{C}$, let $x(t)$ be an eigenvector of an eigenvalue $\nu_{A(t)}(z)$ of $Q_{A(t)}(z)$ such that $||x(t)||_1 = 1$. Then for the arbitrary $t_1, t_2 \in [0,1]$ we obtain that

$$|\nu_{A(t_2)}(z) - \nu_{A(t_1)}(z)| = |r_i^{x(t_2)}(A(t_2)) - r_i^{x(t_1)}(A(t_1))| = |t_2 r_i^{x(t_2)}(A) - t_1 r_i^{x(t_1)}(A)|,$$

for every $i \in N$, and consequently

$$|\nu_{A(t_2)}(z) - \nu_{A(t_1)}(z)| \leq |t_2 - t_1| \max_{i \in N}\{r_i^{x(t_2)}(A), r_i^{x(t_1)}(A)\} \leq |t_2 - t_1| M,$$

where $M := \max_{i \in N} \max_{||x||_1 = 1, x > 0} r_i^x(A) > 0$ due to the irreducibility of the matrix A. Therefore, $t \mapsto \nu_{A(t)}(z)$ is nondecreasing uniformly continuous function for every $z \in \mathbb{C}$. So, $\Gamma^{\mathcal{R}}(A(t))$ continuously expands in $\mathbb{C}$ from the n points, $\{a_{ii} : i \in N\}$, for $t = 0$, to $\Gamma^{\mathcal{R}}(A)$ for $t = 1$. Since, $\Gamma^{\mathcal{R}}(A(t))$ consists of the n disjoint disks around diagonal entries of the matrix A, for the small enough $t > 0$ claim of the lemma obviously holds for $A(t)$. So, in order to prove it for the original matrix A, it suffices to see that the observed property is preserved when the coalescence of the disjointed components occur. □

Assuming that for a point ω_j on the boundary $\partial\Gamma^{\mathcal{R}}(A)$ we have updated the angle to φ_{j+1} and obtained a new point ω_{j+1} such that $|\omega_{j+1}-\omega_j| \geq \varepsilon_2$, we will update ξ to be the diagonal entry from Lemma 4. If such a diagonal entry can not be found for a given ℓ, then we need to decrease the value of ℓ until such diagonal entry a_{ii} can be provided and set $\widehat{\xi} = a_{ii}$. Then, we continue to generate points of the polygon around $\widehat{\xi}$ starting from ω_j and $\varphi_j := \frac{\omega_j - \widehat{\xi}}{|\omega_j - \widehat{\xi}|}$. If this case occurs, we need to specify when to stop producing the new points of the polygon. Since we have changed the center of the star-shaped subset, angle φ_j is not any more a valid indicator. Instead, we introduce an angle

$$\theta_j := \frac{\omega_j - \xi}{|\omega_j - \xi|} \in [-\pi, \pi], \text{ for } j = 1, 2, \ldots, \tag{3.10}$$

and stop at $j = m$ such that $|\theta_{m+1} - \theta_m| > \pi$. In our numerical implementation, we have exploited the fact that when $|\omega_{j+1} - \omega_j| \geq \varepsilon_2$, the new diagonal entry from Lemma 4 is often the one closest to the point $\frac{\omega_{j+1}+\omega_j}{2}$. Finally, according to Theorem 4, eMGS produces numerical approximation of the minimal Geršgorin set $\Gamma^{\mathcal{R}}(A)$. So, in the following section we compare its performance to the gMGS and bMGS algorithms.

Algorithm eMGS

Input: A, ε_1, ε_2, τ

Set $\overline{\mathcal{S}} = \{a_{i_1 i_1}, a_{i_2 i_2}, \ldots, a_{i_{\tilde{n}} i_{\tilde{n}}}\}$ as in (3.9) and initialize $i = 1$;
while $\overline{\mathcal{S}} \neq \emptyset$ **do**
  Set $\xi_0 = \overline{\mathcal{S}}(1)$, and $\mathcal{S}_i = \{\xi_0\}$;
  Initialize $\xi = \xi_0$, $\varphi = -\pi, \theta_0 = -\pi$, $\theta_1 = -\pi$, $\ell = \tau\varepsilon_2$ and $j = 1$;
  Run **eSearch**(ξ, φ) to compute $\omega_{i,j} \in \mathbb{C}$ and $t_j > 0$;
  while $|\theta_j - \theta_{j-1}| < \pi$ **do**
    Update $\varphi \leftarrow \varphi + \arctan(\frac{\ell}{t_j})$ and $j \leftarrow j + 1$;
    Run **eSearch**(ξ, φ) to compute $\omega_{i,j}$ and t_j, and set θ_j using (3.10);
    while $|\omega_{i,j} - \omega_{i,j-1}| < \varepsilon_2$ and $|\theta_j - \theta_{j-1}| < \pi$ **do**
      Update $\varphi \leftarrow \varphi + \arctan(\frac{\ell}{t_j})$ and $j \leftarrow j + 1$;
      Run **eSearch**(ξ, φ) to compute $\omega_{i,j}$ and t_j, and set θ_j using (3.10);
    end while
    if $|\theta_j - \theta_{j-1}| < \pi$ **then**
      Set $k = 1$ and $\widetilde{\mathcal{S}}$ to $\overline{\mathcal{S}}$ ordered w.r.t. the distance to $\frac{\omega_{i,j}+\omega_{i,j-1}}{2}$;
      repeat
        if $\widetilde{\mathcal{S}}(k) \neq \xi$ **then**
          Run **eSearch**(ξ, $\frac{\omega_{i,j-1}-\widetilde{\mathcal{S}}(k)}{|\omega_{i,j-1}-\widetilde{\mathcal{S}}(k)|}$) to compute ω_{old} and t_{old};
          if $|\omega_{old} - \omega_{i,j-1}| < \varepsilon_1$ **then**
            Run **eSearch**(ξ, $\frac{\omega_{i,j}-\widetilde{\mathcal{S}}(k)}{|\omega_{i,j}-\widetilde{\mathcal{S}}(k)|}$) to compute ω_{new} and t_{new};
          end if
        end if
        Update $k \leftarrow k + 1$;
      until $\big(|\omega_{old} - \omega_{i,j-1}| < \varepsilon_1$ and $|\omega_{old} - \omega_{new}| < \varepsilon_2\big)$ or $k > length(\widetilde{\mathcal{S}})$
      if $\big(|\omega_{old} - \omega_{i,j-1}| < \varepsilon_1$ and $|\omega_{old} - \omega_{new}| < \varepsilon_2\big)$ **then**
        Set $\xi = \widetilde{\mathcal{S}}(k)$ and update $\mathcal{S}_i \leftarrow \mathcal{S}_i \cup \{\xi\}$;
        Set $\omega_{i,j} = \omega_{new}$, $t_j = t_{new}$ and set θ as in (3.10);
      else
        Update $\ell \leftarrow \ell\tau$ and $j \leftarrow j - 1$ and set $\varphi = \frac{\omega_{i,j}-\xi}{|\omega_{i,j}-\xi|}$;
      end if
    end if
  end while
  Update $\overline{\mathcal{S}} \leftarrow \overline{\mathcal{S}} \setminus \mathcal{S}_i$ and $i \leftarrow i + 1$;
  Update $\overline{\mathcal{S}}$ to exclude all elements inside of the polygon $\{\omega_{i,j}\}_{1 \le j \le m_i}$
end while

Output: $\{\{\omega_{1,j}\}_{1 \le j \le m_1}, \{\omega_{2,j}\}_{1 \le j \le m_2}, \ldots, \{\omega_{s,j}\}_{1 \le j \le m_s}\}$

4. Numerical examples. In this section, we illustrate the behavior of the introduced algorithm eMGS and compare it with the state of the art approaches, i.e., gMGS and bMGS on several test matrices of varying size (small and medium). All algorithms were implemented in MATLAB version 8.3.0.532 (R2014a) and tested on an 3.3 GHz Intel® Core™ i5-4590 machine.

EXAMPLE 5. *The first example is the cyclic matrix*

$$A = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & i & 1 \\ 1 & 0 & 0 & -i \end{bmatrix}.$$

The initial values are chosen as $\varepsilon_1 = 10^{-10}, \varepsilon_2 = 0.05, \tau = 3$ *for eMGS algorithm,* $n_g = 100$ *for gMGS algorithm and tol* $= 10^{-10}, m = 80, n_g = 100$ *for bMGS algorithm.*

The total number of eigenvalue computations (#eval) is $\#\text{eval}_{eMGS} = 3946$, $\#\text{eval}_{gMGS} = 10000$ and $\#\text{eval}_{bMGS} = 8000$ while the corresponding computational time (CPU) is $0.426982s$, $0.396512s$ and $0.636345s$, respectively. Figure 2 presents the geometry of obtained minimal Geršgorin sets, more precisely, their boundaries. The bMGS algorithm, at first sight seems to be approximating the boundary correctly, but when zoomed near the origin, see Figure 2(b), we notice that it is not revealing the boundary completely. Moreover, it requires many more iterations and therefore is not really efficient. In comparison, the gMGS method is computationally less demanding, however it still does not resolve correctly the boundary near the origin, see Figure 2(d). Our new eMGS approach, although slightly more expensive in this small size example, is much more reliable and provides a very good approximation of the boundary of the minimal Geršgorin set, see Figure 2(f).

EXAMPLE 6. *Our second example is the Tolosa matrix* `tols340.mtx` *of size* $n = 340$ *from the* MATRIX MARKET *repository [3]. It is a highly nonnormal, sparse matrix of medium size used in the stability analysis of a flying airplane.*

The initial values are chosen as $\varepsilon_1 = 10^{-6}, \varepsilon_2 = 300, \tau = 3$ *for the eMGS algorithm, resulting in 155 eigenvalue computations in* $16.58s$ *of CPU time, see Figure 3(a).*

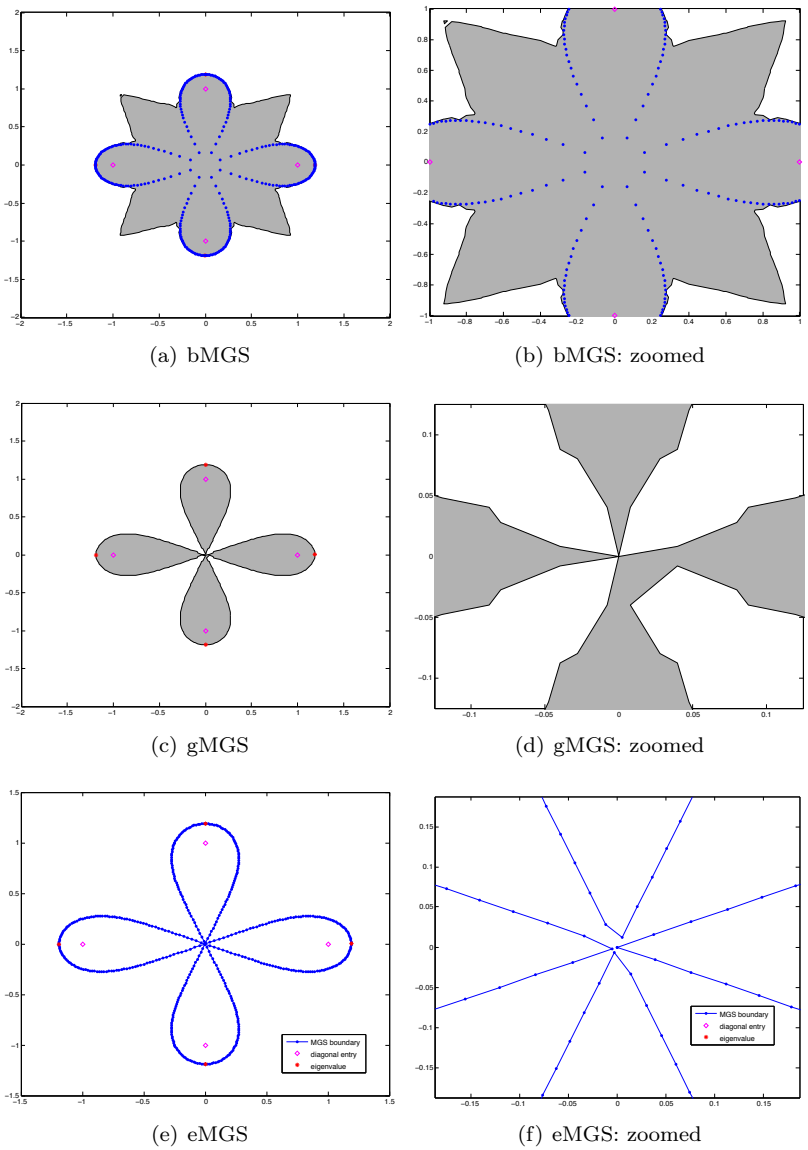
On the other hand, setting tol $= 10^{-6}$*, and* $m = 36, n_g = 100$ *for bMGS algorithm results in 3600 eigenvalue computations in around* $9h15min$ *of CPU time. Furthermore, bMGS in this case doesn't approximate minimal Geršgorin set of Tolosa matrix due to small value of* m *and* n_g *parameters, cf. Figure 3(b). Note, increasing the values* m *and* n_g *would prolong the computations even more!*

Finally, gMGS algorithm does not provide any meaningful results for an arbitrary choice of initial parameters. Thus, in this case of the middle size, sparse and nonnormal matrix, the only practically available algorithm would be eMGS.

EXAMPLE 7. *The last example is a triangular matrix*

$$A_\mu = \begin{bmatrix} \mu & 1 & 0 & \cdots & 0 \\ 1 & 2\mu & 1 & \ddots & \vdots \\ 0 & 1 & 3\mu & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & 1 \\ 0 & \cdots & 0 & 1 & 20\mu \end{bmatrix}.$$

of size $n = 20$*, where parameter* μ *takes values* 1, 2.3 *and* 2.7*. For the parameter* $\mu = 1$ *and* $\mu = 2.7$ *all three algorithms approximate the boundary of the minimal Geršgorin set quite accurate, see Figures 4–6, with a slight advantage of the eMGS approach regarding the time/iteration ratio, see Table 1. However, for* $\mu = 2.3$ *the eMGS algorithm significantly outperforms the two other approaches. Although, the*



(a) bMGS (b) bMGS: zoomed

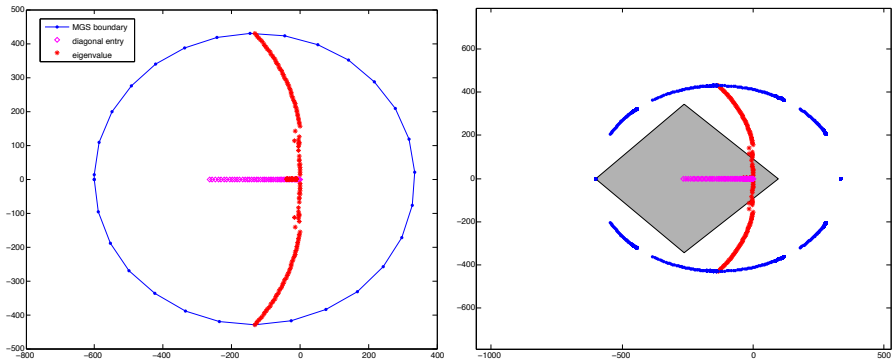
(c) gMGS (d) gMGS: zoomed

(e) eMGS (f) eMGS: zoomed

Fig. 2: The results of the bMGS, gMGS and eMGS algorithm for the cyclic matrix, cf. Example 5.

gMGS algorithm approximate the boundary of the MGS correctly its computational cost is much higher than the one of the new approach. On the other hand the bMGS algorithm do not recover the disconnected components of the minimal Geršgorin set, see Figure 5(c).

5. Concluding remarks. In this paper we give an elaborate treatment of the known methods for computing the minimal Geršgorin set, and develop new method called explicit minimal Geršgorin set algorithm that is mainly suitable for medium size



(a) eMGS (b) bMGS

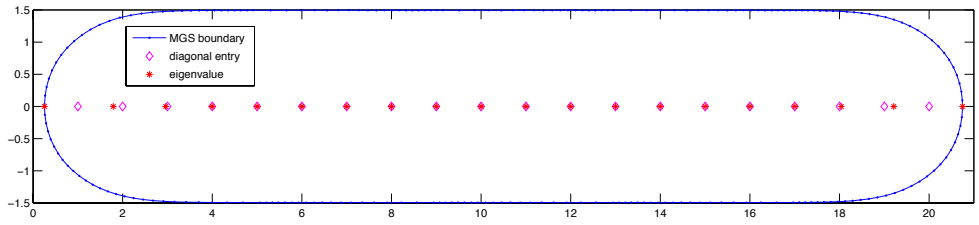
Fig. 3: The results of the eMGS and bMGS algorithm for the Tolosa matrix of size $n = 340$, cf. Example 6.

$\mu = 1$			
	eMGS $\varepsilon_1 = 10^{-6}$, $\varepsilon_2 = 0.4$, $\tau = 3$	**gMGS** $n_g = 100$	**bMGS** $tol = 10^{-6}$, $m = 40$, $n_g = 100$
#eval	2393	10000	4000
CPU [s]	0.663115	0.763675	2.650294
$\mu = 2.3$			
	eMGS $\varepsilon_1 = 10^{-6}$, $\varepsilon_2 = 0.3$, $\tau = 3$	**gMGS** $n_g = 100$	**bMGS** $tol = 10^{-6}$, $m = 120$, $n_g = 100$
#eval	6895	16000	24000
CPU [s]	1.360519	11.618339	12.837484
$\mu = 2.7$			
	eMGS $\varepsilon_1 = 10^{-6}$, $\varepsilon_2 = 0.3$, $\tau = 3$	**gMGS** $n_g = 100$	**bMGS** $tol = 10^{-6}$, $m = 60$, $n_g = 200$
#eval	6893	10000	6000
CPU [s]	1.301505	0.939988	3.501672

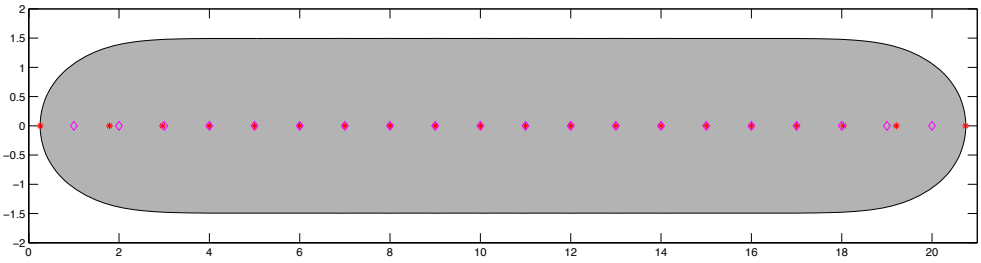
Table 1: Number of eigenvalue computations (#eval) and computational time (CPU) for the eMGS, gMGS and bMGS algorithm for A_μ with $\mu = 1, 2.3$ and 2.7, cf. Example 7.

irreducible matrices, or large sparse matrices that can be reduced to block triangular form (3.1) with the medium size irreducible diagonal blocks. While for the small size generic matrices new algorithm behaves similarly to the simple grinding technique, in more sophisticated (small size) cases that occur in practice it is more reliable one. Furthermore, for the medium size matrices in general it is practically the only applicable one.

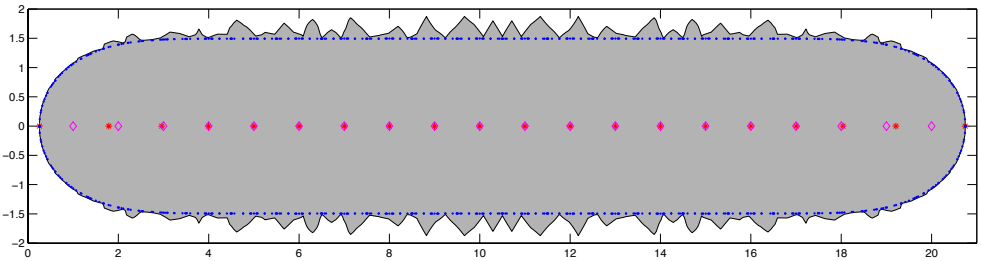
Due to the interesting applications of the minimal Geršgorin set (as in the stability of frequency time dependant dynamical systems), an interesting open problem is to derive the algorithm for the case of large and sparse matrices. The main difficulties to be overcome lies in the fact that construction of this set relies on the eigenvalue



(a) eMGS



(b) gMGS



(c) bMGS

Fig. 4: The results of the eMGS, gMGS and bMGS algorithm for A_μ with $\mu = 1$, cf. Example 7.

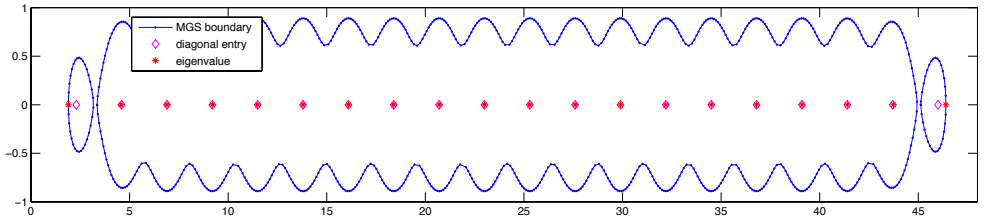
computations in a similar way as the ε-pseudospectra does.

6. Acknowledgments. The work of V. R. Kostić and Lj. Cvetković has been partially supported by the Ministry of Science, Research Grant 174019, Provincial Secretariat for Science of Vojvodina, Research Grants 1136, 1850 and 2010. The work of A. Międlar has been supported by the DFG Research Fellowship under the DFG Project *Adaptive methods for nonlinear eigenvalue problems with parameters* and Chair of Numerical Algorithms and High-Performance Computing (ANCHP), Mathematics Institute of Computational Science and Engineering (MATHICSE), École Polytechnique Fédérale de Lausanne (EPFL).

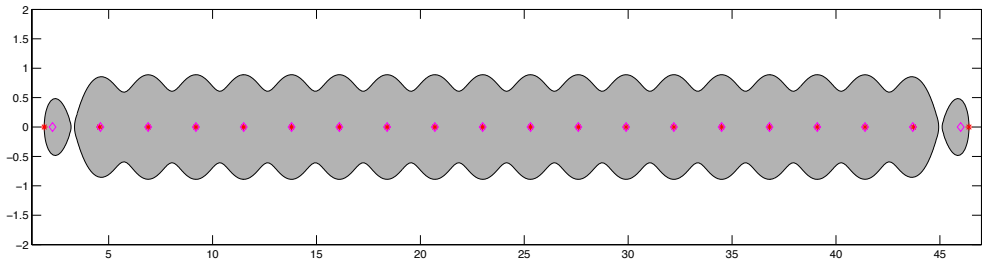
REFERENCES

[1] T. Akenine-Möller, E. Haines, and N. Hoffman. *Real-Time Rendering 3rd Edition*. A. K. Peters, Ltd., Natick, MA, USA, 2008.

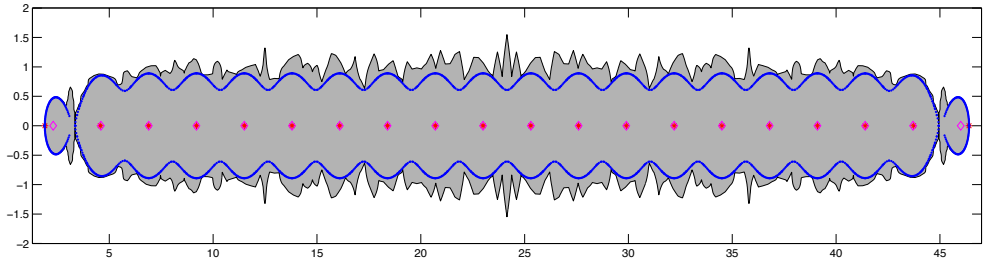
[2] A. Berman and R. Plemmons. *Nonnegative matrices in the mathematical sciences*, volume 9 of



(a) eMGS



(b) gMGS



(c) bMGS

Fig. 5: The results of the eMGS, gMGS and bMGS algorithm for A_μ with $\mu = 2.3$, cf. Example 7.

Classics in Applied Mathematics. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1994. Revised reprint of the 1979 original.

[3] B. Boisvert, R. Pozo, K. Remington, B. Miller, and R. Lipman. Matrix Market repository. `http://math.nist.gov/MatrixMarket/`.

[4] L. Cvetković and V. Kostić. Between Geršgorin and the the minimal Geršgorin set. *J. Comput Appl. Math.*, 196(2):452–458, 2006.

[5] L. Cvetković, V. Kostić, R. Bru, and F. Pedroche. A simple generalization of Geršgorin's theorem. *Adv. Comput. Math.*, 35(2-4):271–280, 2011.

[6] L. Cvetković, V. Kostić, and R. S. Varga. A new Geršgorin-type eigenvalue inclusion set. *ETNA (Elec. Trans. on Numer. An.)*, 18:73–80, 2004.

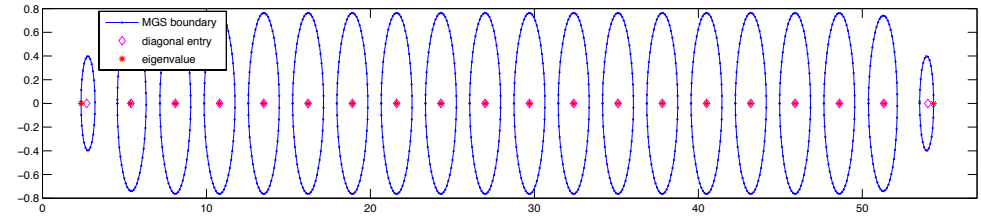
[7] C. Desoer. *Feedback Systems: Input-output Properties*. Elsevier Science, 2012.

[8] S. Geršgorin. Über die Abgrenzung der Eigenwerte einer Matrix. *Izv. Akad. Nauk SSSR Ser. Mat.*, 1:749–754, 1931.

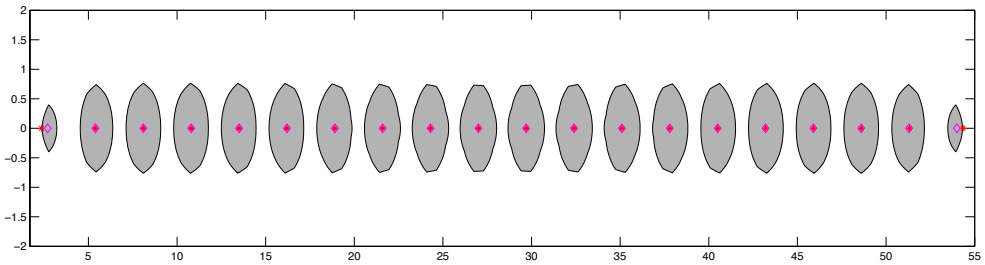
[9] V. Kostić. On general principles of eigenvalue localizations via diagonal dominance. *Adv. Comput. Math.*, 41(1):55–75, 2015.

[10] P. Lancaster. On eigenvalues of matrices dependent on a parameter. *Numer. Math.*, 6:377–387, 1964.

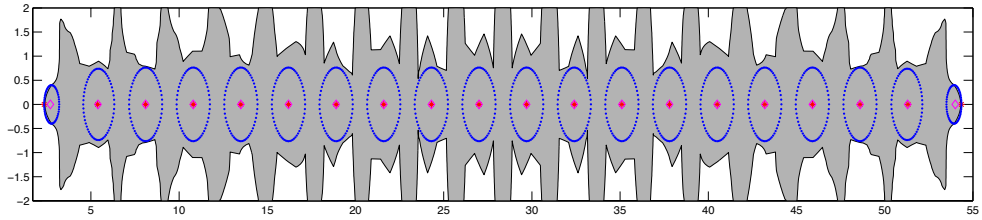
[11] O. Taussky. Bounds for the characteristic roots of matrices. *Duke Math. J.*, 15:1043–1044,



(a) eMGS



(b) gMGS



(c) bMGS

Fig. 6: The results of the eMGS, gMGS and bMGS algorithm for A_μ with $\mu = 2.7$, cf. Example 7.

1948.

[12] L. N. Trefethen and M. Embree. *Spectra and pseudospectra.* Princeton University Press, Princeton, NJ, 2005. The behavior of nonnormal matrices and operators.

[13] R. S. Varga. Minimal Gerschgorin sets. *Pacific J. Math.*, 15:719–729, 1965.

[14] R. S. Varga. Gerschgorin disks, Brauer ovals of Cassini (a vindication), and Brualdi sets. *Information*, 4(2):171–178, 2001.

[15] R. S. Varga. Geršgorin-type eigenvalue inclusion theorems and their sharpness. *Electron. Trans. Numer. Anal.*, 12:113–133 (electronic), 2001.

[16] R. S. Varga. *Geršgorin and His Circles*, volume 36 of *Springer Series in Computational Mathematics.* Springer-Verlag, Berlin, New York, 2004.

[17] R. S. Varga, L. Cvetković, and V. Kostić. Aproximation of the minimal Geršgorin set of a squre complex matrix. *Electron. Trans. Numer. Anal.*, 30:398–405, 2008.

Recent publications:

MATHEMATICS INSTITUTE OF COMPUTATIONAL SCIENCE AND ENGINEERING
Section of Mathematics
Ecole Polytechnique Fédérale
CH-1015 Lausanne

42.2014 ASSYR ABDULLE, PATRICK HENNING:
A reduced basis localized orthogonal decomposition

43.2014 PAOLA F. ANTONIETTI, ILARIO MAZZIERI, ALFIO QUARTERONI:
Improving seismic risk protection through mathemactical modeling

44.2014 LUCA DEDÈ, ALFIO QUARTERONI, SEHNGFENG ZHU:
Isogeometric analysis and proper orthogonal decomposition for parabolic problems

45.2014 ZVONIMIR BUJANOVIC, DANIEL KRESSNER:
A block algorithm for computing antitriangular factorizations of symmetric matrices

46.2014 ASSYR ABDULLE:
The role of numerical integration in numerical in numerical homogenization

47.2014 ANDREA MANZONI, STEFANO PAGANI, TONI LASSILA:
Accurate solution of Bayesian inverse uncertainty quantification problems using model and error reduction methods

48.2014 MARCO PICASSO:
From the free surface flow of a viscoelastic fluid towards the elastic deformation of a solid

49.2014 FABIO NOBILE, FRANCESCO TESEI:
A multi level Monte Carlo method with control variate for elliptic PDEs with log-normal coefficients

***

01.2015 PENG CHEN, ALFIO QUARTERONI, GIANLUIGI ROZZA:
Reduced order methods for uncertainty quantification problems

02.2015 FEDERICO NEGRI, ANDREA MANZONI, DAVID AMSALLEM:
Efficient model reduction of parametrized systems by matrix discrete empirical interpolation

03.2015 GIOVANNI MIGLIORATI, FABIO NOBILE, RAÚL TEMPONE:
Convergence estimate in probability and in expectation for discrete least squares with noisy evaluations at random points

04.2015 FABIO NOBILE, LORENZO TAMELLINI, FRANCESCO TESEI, RAÚL TEMPONE:
An adaptive sparse grid alorithm for elliptic PDEs with lognormal diffusion coefficent

05.2015 MICHAEL STEINLECHNER:
Riemannian optimization for high-dimensional tensor completion

06.2015 V. R. KOSTIĆ, A. MIEDLAR, LJ. CVETKOVIĆ:
An algorithm for computing minimal Geršgorin sets