

# Convergence rate analysis and improved iterations for numerical radius computation

Tim Mitchell\*

January 31st, 2020

## Abstract

We analyze existing methods for computing the numerical radius and introduce improved algorithms. Current methods include level-set and cutting-plane based approaches, but until now, no formal convergence rate results have been established for these techniques. We first introduce an improved level-set method that is often significantly faster than the existing approach of Mengi and Overton. We then establish the first rate of convergence results for any numerical radius method, showing how the convergence of Uhlig's cutting-plane method varies from superlinear to linear based on the normalized curvature at outermost points in the field of values. Moreover, we introduce a more efficient cutting-plane method whose convergence rate we also derive. Finally, as cutting-plane methods become very expensive when the field of values closely resembles a circular disk centered at the origin, we introduce a third algorithm combining both approaches to remain efficient in all cases.

**Notation:**  $\|\cdot\|$  denotes the spectral norm,  $\Lambda(\cdot)$  the spectrum (the set of eigenvalues) of a square matrix, and  $\lambda_{\max}(\cdot)$  and  $\lambda_{\min}(\cdot)$ , respectively, the largest and smallest eigenvalue of a Hermitian matrix.

## 1 Introduction

The trajectory of the discrete-time dynamical system

$$x_{k+1} = Ax_k, \quad (1)$$

where  $A \in \mathbb{C}^{n \times n}$  and  $x_k \in \mathbb{C}^n$ , is clearly tied to powers of  $A$ , since  $x_k = A^k x_0$  and so  $\|x_k\| \leq \|A^k\| \|x_0\|$ . As is well known, the asymptotic behavior of (1) is characterized by the moduli of the eigenvalues of  $A$ . Given the *spectral radius* of  $A$ ,

$$\rho(A) := \max\{|\lambda| : \lambda \in \Lambda(A)\}, \quad (2)$$

$\lim_{k \rightarrow \infty} \|x_k\| = 0$  for all  $x_0$  if and only if  $\rho(A) < 1$ , with the asymptotic decay rate being faster the closer  $\rho(A)$  is to zero. However, if  $A$  is non-normal, its eigenvalues alone do not reveal the nature of the transient behavior of (1). Indeed, a central theme of Trefethen's and Embree's treatise on pseudospectra [TE05] is addressing the very question of how large  $\|A^k\|$  gets before the asymptotic decay takes over.

One perspective is given by the *field of values* (*numerical range*) of  $A$ ,

$$W(A) := \{x^* Ax : x \in \mathbb{C}^n, \|x\| = 1\}, \quad (3)$$

---

\*Max Planck Institute for Dynamics of Complex Technical Systems, Magdeburg, 39106 Germany  
mitchell@mpi-magdeburg.mpg.de.

as it can impart information about the transient behavior of (1). More specifically, consider the maximum of the moduli of points in  $W(A)$ , i.e., the *numerical radius*

$$r(A) := \max\{|z| : z \in W(A)\}. \quad (4)$$

It is known that  $\frac{1}{2}\|A\| \leq r(A) \leq \|A\|$ ; see [HJ91, p. 44]. Combining the lower bound with the power inequality  $r(A^k) \leq r(A)^k$  [Ber65, Pea66] yields

$$\|A^k\| \leq 2r(A)^k. \quad (5)$$

As  $2r(A)^k \leq \|A\|^k$  if and only if  $r(A) \leq \sqrt[k]{0.5}\|A\|$ , and  $r(A) \leq \|A\|$  always holds, it follows that  $2r(A)^k$  is often a tighter upper bound for  $\|A^k\|$  than  $\|A\|^k$  is, and so the numerical radius can be particularly useful in estimating the transient behavior of (1).

*Remark 1.* As noted in [TE05], other quantities for assessing the transient behavior of (1) are the  $\varepsilon$ -*pseudospectral radius* and the *Kreiss constant* [Kre62]. For algorithms to compute these, see [MO05, BM19] and [Mit19a, Mit19b], respectively.

In 1996, Watson proposed two iterations akin to the power method for  $r(A)$ , the first of which at least converges to locally optimal approximations, and a second simpler iteration that is less computationally intensive [Wat96]. However, as these iterations are so closely related to the power method, both may exhibit very slow convergence, and the possibility of the cheaper iteration failing cannot be ruled out since it lacks any convergence guarantees. Shortly thereafter, inspired by Byers' breakthrough algorithm for the distance to instability [Bye88], He and Watson combined Watson's cheaper iteration with a certificate test that either asserts  $r(A)$  has been attained (to a given tolerance) or provides a way of restarting in order to aid convergence to  $r(A)$ , if not theoretically guarantee it [HW97]. Then in 2005, Mengi and Overton showed [MO05] that the certificate test of He and Watson actually enables an iteration guaranteed to converge to  $r(A)$ , namely an analogue of (the discrete-time variant of) the quadratically convergent level-set-based  $\mathcal{H}_\infty$  norm algorithm of Boyd, Balakrishnan, Bruinsma, and Steinbuch (BBBS) [BB90, BS90]. At each iteration, progress toward  $r(A)$  is made by performing the certificate test, which computes all unimodular eigenvalues of a certain  $2n \times 2n$  generalized eigenvalue problem. In practice, Mengi and Overton observed that their numerical radius method converged quadratically, although this does not appear to have ever been established formally.

In 2009, Uhlig proposed a geometric approach to computing  $r(A)$  [Uhl09], based on Johnson's tangential cutting-plane method for approximating the boundary of  $W(A)$  [Joh78], which stems from the much earlier Bendixson-Hirsch theorem [Ben02] and fundamental results of Kippenhahn [Kip51]. Indeed, Johnson even mentioned that his algorithm could be adapted to compute the numerical radius, but that a modified version might be more efficient [Joh78, Remark 3]. Uhlig's method converges to  $r(A)$  by computing the largest eigenvalue of a sequence of  $n \times n$  Hermitian matrices, and as these subproblems can be reliably solved via sparse eigensolvers, Uhlig's method is the only one that also is practical for computing the numerical radius of large and sparse matrices. All other scalable methods for the numerical radius at best only guarantee returning locally optimal approximations, which of course could be arbitrarily bad. Although Uhlig noted that the convergence of his method can sometimes be quite slow (it has been suspected to be linear), no rate of convergence analysis appears to have been established for his method. Nevertheless, Uhlig's experimental results showed that his algorithm being decisively faster in practice than the apparently quadratically convergent method of Mengi and Overton.

To shed light on this disparity, in this paper, we both examine the algorithms of Mengi and Overton and of Uhlig as well as propose three improved iterations. First, similar to our recent work with Benner on faster computation of the  $\mathcal{H}_\infty$  norm [BM18], we use local-optimization techniques to obtain an accelerated variant of the level-set-based  $r(A)$  algorithm of Mengi and Overton. We then establish certain convergence rate results for the iteration of Uhlig. Specifically, we derive the exact rate of convergence as a function of the amount of local curvature at outermost boundary points of  $W(A)$ . In fact, we show that this convergence is actually superlinear in the best case and linear otherwise, with the rate of convergence at most one quarter and

typically less, often appreciably so. To the best of our knowledge, this is the first formal convergence rate analysis given for Uhlig’s method. Moreover, we develop a second algorithm that combines our aforementioned local-optimization techniques with a new cutting-plane procedure to address some inefficiencies in Uhlig’s original approach.

An important application of our work here is when the numerical radius is the subject of optimization for some parametrized matrix, e.g., one can attempt to minimize the transient behavior of a dynamical system by minimizing its numerical radius. As recently established by Lewis and Overton [LO18], the numerical radius is “partly smooth” with respect to a certain manifold in matrix space, which implies among other things that although it is not differentiable at all matrices, it is locally Lipschitz and hence can be minimized over parametrized matrices using gradient-based nonsmooth optimization techniques. However, minimizing the numerical radius can be an expensive endeavor, and as we will explain, which method is most efficient actually varies depending on the shape of the field of values, which of course is changing as the numerical radius is optimized. In fact, this application motivates our third algorithm, which is a hybrid that dynamically chooses between our two improved methods mentioned above in order to remain efficient.

The paper is organized as follows. We first give necessary preliminaries on the field of values, the numerical radius, and earlier algorithms for  $r(A)$  in §2. Then in §3, we describe our improved level-set-based  $r(A)$  algorithm using local optimization. Our new rate of convergence results for Uhlig’s iteration are presented in §4. We present our second improved algorithm, which uses optimization and a new cutting-plane procedure, in §5, and analyze the exact convergence rate of this new cutting process. We describe our third hybrid algorithm in §6, validate our three improved methods experimentally in §7, and give concluding remarks in §8.

All experiments in this paper were done in MATLAB R2017b on a laptop with an Intel i7-6567U dual-core CPU laptop, 16GB of RAM, and macOS v10.14.

## 2 The field of values, the numerical radius, and earlier algorithms

We will need the following well-known facts about the field of values and the numerical radius; see [HJ91, Chapter 1]. We will also generally assume that  $A$  is non-normal, as otherwise  $r(A)$  can be computed just from the spectral radius of  $A$ , as seen in the following properties.

*Remark 2.* Given  $A \in \mathbb{C}^{n \times n}$ ,

- (A1)  $W(A) \subset \mathbb{C}$  is a compact, convex set,
- (A2) if  $A$  is real, then  $W(A)$  has real axis symmetry,
- (A3) if  $A$  is normal, then  $W(A)$  is the convex hull of  $\Lambda(A)$ ,
- (A4) the boundary of  $W(A)$ ,  $\partial W(A)$ , is a piecewise smooth algebraic curve,
- (A5) if  $v \in \partial W(A)$  is a point where  $\partial W(A)$  is not differentiable, i.e., a corner, then  $v \in \Lambda(A)$ .

**Definition 2.1.** Given a nonempty closed set  $\mathcal{D} \subset \mathbb{C}$ , a point  $\tilde{z} \in \mathcal{D}$  is (globally) *outermost* if  $|\tilde{z}| = \max\{|z| : z \in \mathcal{D}\}$  and *locally outermost* if  $\tilde{z}$  is an outermost point of  $\mathcal{D} \cap \mathcal{N}$ , for some neighborhood  $\mathcal{N}$  of  $\tilde{z}$ .

For the continuous-time analogue of (1),  $\dot{x} = Ax$ , we have the *numerical abscissa*

$$\alpha_W(A) := \max\{\operatorname{Re} z : z \in W(A)\}, \quad (6)$$

i.e., the maximal real part of all points in  $W(A)$ . Unlike the numerical radius, computing the numerical abscissa is straightforward. In fact, all one needs to do is compute the largest eigenvalue of a single Hermitian matrix [HJ91, p. 34]:

$$\alpha_W(A) = \lambda_{\max}\left(\frac{1}{2}(A + A^*)\right). \quad (7)$$

For  $\theta \geq 0$ ,  $W(e^{i\theta}A)$  is  $W(A)$  rotated counter-clockwise about the origin. Consider

$$H(\theta) := \frac{1}{2} (e^{i\theta}A + e^{-i\theta}A^*), \quad (8)$$

so  $\alpha_W(e^{i\theta}A) = \lambda_{\max}(H(\theta))$  and  $\alpha_W(A) = \lambda_{\max}(H(0))$ . For some angle  $\theta$ , let  $\lambda_\theta$  and  $x_\theta$  denote, respectively,  $\lambda_{\max}(H(\theta))$  and an associated (normalized) eigenvector. Furthermore, let  $L_\theta$  denote the line  $\{e^{-i\theta}(\lambda_\theta + it) : t \in \mathbb{R}\}$  and  $P_\theta$  the half-plane  $e^{-i\theta}\{z : \operatorname{Re} z \leq \lambda_\theta\}$  (of the two half-planes) determined by line  $L_\theta$ . Then  $P_\theta$  is a *supporting hyperplane* for  $W(A)$  with the following properties: [Joh78, p. 597]

(B1)  $W(A) \subseteq P_\theta$  for all  $\theta \in [0, 2\pi)$ ,

(B2)  $b_\theta = x_\theta^* A x_\theta \in \partial W(A) \cap L_\theta$ , i.e.,  $L_\theta$  is a tangent at  $b_\theta \in \mathbb{C}$ ,

(B3)  $W(A) = \bigcap_{\theta \in [0, 2\pi)} P_\theta$ .

Furthermore, as  $H(\theta + \pi) = -H(\theta)$ ,  $P_{\theta+\pi}$  also can be obtained via the smallest eigenvalue of  $H(\theta)$  and its associated eigenvector. The Bendixson-Hirsch theorem is a special case of these properties, i.e., the bounding box (with sides parallel to the axes) of  $W(A)$ , for  $\theta = 0$  and  $\theta = \frac{\pi}{2}$ .

Using (7) and (8), it is easy to see that computing the numerical radius is equivalent to solving the following one-variable maximization problem:

$$r(A) = \max_{\theta \in [0, 2\pi)} h(\theta) \quad \text{where} \quad h(\theta) := \lambda_{\max}(H(\theta)). \quad (9)$$

Furthermore, again using  $H(\theta + \pi) = -H(\theta)$ , it also follows that

$$r(A) = \max_{\theta \in [0, \pi)} \rho(H(\theta)). \quad (10)$$

However, as (9) and (10) may have multiple maxima (possibly infinite in fact, which happens if  $W(A)$  is a disk centered at the origin), it is not so straightforward to find a global maximizer of either in order to obtain  $r(A)$ .

While Watson's first iteration converges to local maximizers of (10), it was his second non-guaranteed iteration that He and Watson employed with their new certificate test to compute  $r(A)$ , due to it being significantly cheaper (though both require  $\mathcal{O}(n^2)$  work per iteration).<sup>1</sup> Whenever this so-called "simple iteration" terminates or stagnates (e.g., due to slow convergence), their certificate test either asserts that  $r(A)$  has in fact been attained (to a tolerance), or it restarts the simple iteration in a monotonic way, so that more progress toward  $r(A)$  hopefully can be made in the next round. As Mengi and Overton pointed out, given some estimate  $\gamma$ , the certificate test actually provides more information than just whether some  $\gamma$  is sufficiently close to  $r(A)$ . In fact, the test can provide all the  $\gamma$ -level set points (if any) of  $h(\theta)$ , i.e., the set of points for which  $h(\theta) = \gamma$ . These points can be used to obtain all the intervals with endpoints in the  $\gamma$ -level set of  $h(\theta)$  but otherwise lie under  $h(\theta)$ . Thus, as an analogue of the BBBS  $\mathcal{H}_\infty$ -norm algorithm, given an estimate  $\gamma_j$  for  $r(A)$ , the method of Mengi and Overton uses (9) and the certificate test to compute all the intervals under  $h(\theta)$  whose endpoints are in the  $\gamma_j$ -level set of  $h(\theta)$ . Then  $h(\theta)$  is evaluated at all the midpoints of these intervals, and  $\gamma_{j+1}$  is set to the highest of the corresponding function values. The sequence of  $\gamma_j$  estimates must converge to  $r(A)$ , and, since the BBBS algorithm converges quadratically, it is not so surprising that Mengi and Overton observed that their variant for  $r(A)$  converged quadratically in practice.

The certificate (or level-set) test is based on [MO05, Theorem 3.1], which is a slight restatement of [HW97, Theorem 2] from He and Watson. We omit the proof.

<sup>1</sup> In [HW97, Equation (3)], (10) is written as " $\max_{0 \leq \theta < \pi} \lambda_{\max \text{mod}}(H(\theta))$ ", where  $\lambda_{\max \text{mod}}$  denotes the maximum modulus eigenvalue", but either the absolute value bars are missing or this should be the maximum modulus *over the* eigenvalues, i.e., the spectral radius. Meanwhile in [Wat96, Equation (2.8)], (10) is written as  $\max_{0 \leq \theta < \pi} |\lambda_{\max}(H(\theta))|$ , but this too is a typo as an outermost eigenvalue is needed here, not  $\lambda_{\max}$ .

**Theorem 2.2.** *Given  $\gamma \in \mathbb{R}$ , the pencil  $R_\gamma - \lambda S$  has  $e^{i\theta}$  as an eigenvalue or is singular if and only if  $\gamma$  is an eigenvalue of  $H(\theta)$  defined in (8), where*

$$R_\gamma := \begin{bmatrix} 2\gamma I & -A^* \\ I & 0 \end{bmatrix} \quad \text{and} \quad S := \begin{bmatrix} A & 0 \\ 0 & I \end{bmatrix}. \quad (11)$$

Note that while all points in the  $\gamma$ -level set of  $h(\theta) = \lambda_{\max}(H(\theta))$  will have corresponding unimodular eigenvalues of  $R_\gamma - \lambda S$ , the converse is not necessarily true. It may be that for some of the unimodular eigenvalues of  $R_\gamma - \lambda S$ ,  $\gamma$  corresponds to eigenvalues of  $H(\theta)$  other than the largest one. Furthermore,  $R_\gamma - \lambda S$  is always nonsingular for all  $\gamma > h(\theta)$ , for any  $\theta \in [0, 2\pi)$ . This is because if  $W(A)$  and a disk centered at the origin have more than  $n$  shared boundary points, then  $W(A)$  is that disk; see [TY99, Lemma 6]. The downside to the BBBS-like iteration is that it is rather costly, cubic work per iteration with a signification constant factor, as all unimodular eigenvalues of  $R_\gamma - \lambda S$  must be computed for each  $\gamma = \gamma_j$ ,

Uhlig's algorithm, on the other hand, computes the numerical radius by following the ideas of Johnson [Joh78], i.e., by refining a convex polygonal approximation

$$\mathcal{P}_j := \bigcap_{\theta \in \{\theta_1, \dots, \theta_j\}} P_\theta \quad (12)$$

of  $W(A)$ , i.e., the intersection of the set of supporting hyperplanes for angles  $\{\bar{\theta}_1, \dots, \bar{\theta}_j\}$ ; for clarity, for the remainder of this paper, we will always use an overline accent on top of an angle, e.g.,  $\bar{\theta}$ , to indicate when an angle is specifying a supporting hyperplane, i.e.,  $P_{\bar{\theta}}$ . Any other angles denoted via some variant of  $\theta$  will not have an overline. Note that for any such  $\mathcal{P}_j$ , the following properties hold:

(C1)  $W(A) \subseteq \mathcal{P}_j$  (hence  $\mathcal{P}_j$  is an overapproximation of  $W(A)$ ),

(C2)  $|b| \leq r(A)$  holds for any tangential boundary point  $b \in \partial W(A) \cap \mathcal{P}_j$ ,

(C3)  $r(A) \leq |c_j|$ , where  $c_j$  is an outermost corner of the polygon determined by  $\mathcal{P}_j$ .

Since  $\mathcal{P}_j$  gives both lower and upper bounds on  $r(A)$ , by refining  $\mathcal{P}_j$ , these bounds can be brought sufficiently close to each other in order to compute  $r(A)$ . Uhlig's method achieves this via a greedy strategy. On each iteration, his algorithm attempts to chop off an outermost (farthest from the origin) corner  $c_j$  from  $\mathcal{P}_j$ , by adding the supporting hyperplane for  $\bar{\theta}_{j+1} = -\text{Arg}(c_j)$  to create  $\mathcal{P}_{j+1}$ . If  $c_j$  happens to be a corner of  $\partial W(A)$ , then the polygon determined by  $\mathcal{P}_{j+1}$  remains unchanged from that of  $\mathcal{P}_j$ , but this operation asserts that  $c_j$  is indeed a locally outermost point of  $\partial W(A)$  and possibly a candidate value where  $r(A)$  is attained. Otherwise,  $c_j \notin W(A)$  and the addition of the supporting hyperplane for  $\bar{\theta}_{j+1} = -\text{Arg}(c_j)$  prunes corner  $c_j$  from the polygon  $\mathcal{P}_j$ , thus yielding  $\mathcal{P}_{j+1}$  with a smaller polygonal region and reducing the current upper bound for  $r(A)$  (assuming there were no ties for the outermost corner of  $\mathcal{P}_j$ ). Furthermore, this cutting operation must also add a new tangential boundary point  $b_{j+1}$  on the newly added edge of  $\mathcal{P}_{j+1}$ , where  $b_{j+1} = x_{j+1}^* A x_{j+1}$  and  $x_{j+1}$  is the eigenvector for  $\lambda_{\max}(H(\bar{\theta}_{j+1}))$ . Hence, the current lower bound on  $r(A)$  may also be improved (increased). Figure 1 depicts Uhlig's method when a corner is cut.

*Remark 3.* Note that if both the largest and the smallest eigenvalues of  $H(\bar{\theta}_{j+1})$  are computed (with their associated eigenvectors), then  $\mathcal{P}_j$  can be refined not just for the supporting hyperplane for angle  $\bar{\theta}_{j+1} = -\text{Arg}(c_j)$  but also the parallel one with angle  $\pi - \text{Arg}(c_j)$ . Whether this second refinement will be useful is not necessarily clear *a priori*, but if the entire spectrum of  $H(\bar{\theta}_{j+1})$  is computed, e.g., by `eig` in MATLAB, there is little reason *not* to do this second refinement in addition to the one for corner  $c_j$ . Nevertheless, Uhlig's method generally only makes use of the single cut for  $c_j$ , as he noted that it is typically cheaper (and yet still reliable) to compute  $\lambda_{\max}(H(\theta))$  via a sparse method (specifically, by `eigs` in MATLAB) than via `eig`, at least when  $n$  is not too small. In fact, Uhlig's code automatically switches between `eig` and `eigs` depending

on how large  $n$  is (while the code of Mengi and Overton only uses `eig`). However, `eigs` also can be set to find eigenvalues from *both ends* of the spectrum, which Uhlig does not appear to have considered. With this in mind, in the next section, we will reexamine which of these options would be most efficient.

Compared to computing  $r(A)$  via Johnson’s  $\partial W(A)$  algorithm, Uhlig’s greedy method is generally more efficient as it only needs to sufficiently approximate  $\partial W(A)$  in the region where  $r(A)$  is attained. However, if  $W(A)$  is a disk centered at the origin, then  $r(A)$  is attained at every boundary point, and so Uhlig’s method, like Johnson’s, must approximate all of  $\partial W(A)$ . Such so-called *disk matrices*, i.e., those whose field of values are circular disks centered at the origin, are relatively rare, but they can arise from optimizing the numerical radius of certain parametrized matrices; see [LO18] for a thorough discussion on this topic. As we will illustrate in §4, these cutting-plane methods are extremely expensive for disk matrices or matrices very near to them, i.e., when  $W(A)$  is close to being a disk centered at the origin.

Finally, in the same paper [Uhl09], Uhlig also shows how Chebfun<sup>2</sup> [DHT14], a software package for approximating functions via interpolation, can be used to compute the numerical radius with just a handful of lines of code in MATLAB. However, in the numerical experiments, this Chebfun-based approach was typically orders of magnitude slower than both Uhlig’s cutting-plane algorithm and the level-set based algorithm of Mengi and Overton.

### 3 An enhanced level-set-based algorithm via local optimization

We begin with a slight modification to the BBBS-like method of Mengi and Overton. As Theorem 2.2 allows all points in any  $\gamma$ -level set of either  $h(\theta)$  or  $\rho(H(\theta))$  to be computed, it is also possible to do a BBBS-like iteration using (10). This has the potential to be a bit faster than Mengi and Overton’s  $h(\theta)$ -based method, as  $\rho(H(\theta)) \geq h(\theta)$  always holds,  $\rho(H(\theta)) \geq 0$  (unlike  $h(\theta)$ , which may be negative in places), and the optimization domain is reduced from  $[0, 2\pi)$  to  $[0, \pi)$ . Hence, by instead working with (10),  $\gamma_j \geq 0$  always holds, every update must be at least as good (and possibly better) as that of working with (9), and there may be fewer level-set intervals per iteration, all of which may reduce the overall cost in a meaningful way.

We now consider a second more substantial modification, namely to use local optimization on top of the BBBS-like step at every iteration, i.e., the BBBS-like step is used to initialize optimization in order to find a maximizer of (10), which must provide an update to  $\gamma_j$  at least as good (and often much better) than the BBBS-like step alone. The main benefit of this modification is speed, as it can greatly reduce the total number of expensive eigenvalue computations with  $R_\gamma - \lambda S$  incurred. This approach is analogous to that proposed in [BM18] for faster and more accurate computation of the  $\mathcal{H}_\infty$  norm and is even similar to [HW97] for the numerical radius. The secondary benefit of this optimization-enhanced strategy is that it can also significantly reduce the risk of numerical issues compared to solely working with  $R_\gamma - \lambda S$  to update  $\gamma_j$ . In their 1997 paper, He and Watson showed that the condition number of a unimodular eigenvalue of  $R_\gamma - \lambda S$  actually blows up as  $\theta$  approaches a critical value of  $h(\theta)$  or  $\rho(H(\theta))$  [HW97, Theorem 4]<sup>3</sup>, as this is when a pair of unimodular eigenvalues of  $R_\gamma - \lambda S$  coalesce into a double eigenvalue. In their algorithm, or Mengi and Overton’s BBBS-like iteration, a pair of unimodular eigenvalues of  $R_\gamma - \lambda S$  always begin to coalesce near convergence. If rounding errors prevent this pair of eigenvalues from being accurately detected, these two algorithms may stagnate before attaining  $r(A)$  to the desired accuracy; indeed, He and Watson expressed that their analytical result was “hardly encouraging” [HW97, p. 336], even though they did not observe this issue in their experiments. However, an example of such a deleterious effect is shown in [BM19, Figure 2], where analogous eigenvalue computations of matrix pencils are shown to greatly reduce numerical accuracy when computing the *pseudospectral abscissa* [BLO03].

<sup>2</sup>Available at <http://www.chebfun.org>.

<sup>3</sup>The exact statement appears in the last lines of the corresponding proof on p. 335.

In contrast,  $\rho(H(\theta))$  has rather nice properties, and so finding its local maximizers can be done quite reliably via standard optimization techniques. Although the spectral radius is not Lipschitz in general, it is Lipschitz for  $H(\theta)$ , since  $H(\theta)$  is Hermitian (see e.g., [Kat82, Theorem II.6.8]). Furthermore, since (10) is a max-max problem (as opposed to a min-max or max-min problem), in general we do not expect local maximizers of (10) to be points where  $\rho(H(\theta))$  is nonsmooth. Consequently, it should be typical that local optimizers of (10) can be obtained using Newton's method with only a handful of iterations. In fact, in their concluding remarks [HW97, p.341–2], He and Watson even mention that their use of Watson's simple iteration could be replaced with Newton's method, as it would have faster local convergence (quadratic). However, they seem to dismiss this idea due to their concerns of additional costs for computing the first and second derivatives of  $\rho(H(\theta))$  and the need for a line search. Actually, these additional costs are negligible for small-scale problems, and for large-scale problems, it would be still efficient if one forgoes computing the second derivative of  $\rho(H(\theta))$  and uses the secant method instead of Newton's method, which would still result in fast local convergence (superlinear). Furthermore, with either Newton's method or the secant method, steps of length one are eventually accepted, and so the need for a line search should not be a critical concern in terms of overall efficiency.

Let  $\lambda_j$  be an eigenvalue attaining  $\rho(H(\theta))$  with  $x_j$  its corresponding normalized eigenvector, and assume  $\lambda_j$  is unique (so  $j = 1$  or  $j = n$  when the spectrum of  $H(\theta)$  is sorted). By standard perturbation theory of eigenvalues,

$$\rho'(H(\theta)) = \operatorname{sgn}(\lambda_j) \cdot x_j^* H'(\theta) x_j = \operatorname{sgn}(\lambda_j) \cdot x_j^* \left( \frac{i}{2} (e^{i\theta} A - e^{-i\theta} A^*) \right) x_j. \quad (13)$$

Hence, given  $\lambda_j$  and  $x_j$ , the additional cost of obtaining  $\rho'(H(\theta))$  mostly amounts to the single matrix-vector product  $H'(\theta)x_j$ . To compute  $\rho''(H(\theta))$ , we will need the following result for second derivatives of eigenvalues; see [Lan64, OW95].

**Theorem 3.1.** *For  $t \in \mathbb{R}$ , let  $H(t)$  be a twice-differentiable  $n \times n$  Hermitian matrix family with, for  $t = 0$ , eigenvalues  $\lambda_1 \geq \dots \geq \lambda_n$  and associated eigenvectors  $x_1, \dots, x_n$ , with  $\|x_k\| = 1$  for all  $k$ . Then assuming  $\lambda_j$  is unique,*

$$\lambda_j''(t) \Big|_{t=0} = x_j^* H''(0) x_j + 2 \sum_{k \neq j} \frac{|x_k^* H'(0) x_j|^2}{\lambda_k - \lambda_j}.$$

Thus, still assuming  $\lambda_j$  is unique and also assuming that the full eigenvalue decomposition of  $H(\theta)$  is already available, obtaining  $\rho''(H(\theta))$  via Theorem 3.1 is also cheap, as  $H'(\theta)x_j$  would already be computed for  $\rho'(H(\theta))$  and

$$H''(\theta)x_j = -H(\theta)x_j = -\lambda_j x_j.$$

Of course, while computing all eigenvalues and eigenvectors of  $H(\theta)$  is not necessarily cheap (it requires cubic work),  $r(A)$  methods based on Theorem 2.2 already assume for the given  $n$  that all unimodular eigenvalues of pencil  $R_\gamma - \lambda S$  can be computed practically, which recall is a  $2n \times 2n$  generalized eigenvalue problem. Since  $H(\theta)$  is an  $n \times n$  Hermitian matrix, computing the eigenvalues and eigenvectors of  $H(\theta)$  will be many times less expensive than obtaining all unimodular eigenvalues of  $R_\gamma - \lambda S$ .

In Algorithm 1, we give pseudocode for our optimization-enhanced level-set-based algorithm to compute  $r(A)$ . Without the local optimization, it is just a BBBS-like iteration for (10), which, assuming it converges quadratically, we expect would typically require three to six iterations. By finding maximizers of (10) (with only a handful of relatively cheap  $n \times n$  eigenvalue computations), we should be able to forgo the expense of refining estimates of a global maximizer of (10) using a sequence of  $R_\gamma - \lambda S$  eigenvalue computations. For instance, if local optimization happens to find a global maximizer of (10) on the very first iteration, then only one eigenvalue computation with  $R_\gamma - \lambda S$  would be necessary, which thus could make the new optimization-enhanced algorithm, say, four to five times faster than the standard BBBS-like iteration alone.

---

**Algorithm 1** An Improved Level-Set-Based Algorithm

---

**Input:**  $A \in \mathbb{C}^{n \times n}$  and guesses  $\mathcal{M} = \{\theta_1, \dots, \theta_q\}$  where  $r(A)$  is obtained in (10).

**Output:**  $\gamma$  such that  $\gamma = r(A)$ .

```
1: if  $A$  is normal then
2:    $\gamma \leftarrow \rho(A)$ 
3:   return
4: end if
5:  $\theta_{\text{ub}} \leftarrow \pi$  ( $0.5\pi$  if either  $\text{Im } A = 0$  or  $\text{Re } A = 0$  holds, symmetric cases)
6:  $\theta_{\text{max}} \leftarrow$  angle of an eigenvalue of  $A$  attaining  $\rho(A)$  (0 if this eigenvalue is 0)
7:  $\mathcal{M} \leftarrow \mathcal{M} \cup 0 \cup \theta_{\text{max}}$ 
8: while  $\mathcal{M}$  is not empty do
9:    $\theta_{\text{BBBS}} \leftarrow \arg \max_{\theta \in \mathcal{M}} \rho(H(\theta))$ 
10:   $\gamma \leftarrow$  maximization of  $\rho(H(\theta))$  via local optimization initialized at  $\theta_{\text{BBBS}}$ 
11:   $\Lambda_{\text{RS}} \leftarrow \{\text{Arg } \lambda : R_\gamma - \lambda S = 0, |\lambda| = 1, \text{Arg } \lambda \in [0, \theta_{\text{ub}})\}$ 
12:   $\{\theta_1, \dots, \theta_q\} \leftarrow \Lambda_{\text{RS}}$  sorted in increasing order with any duplicates removed
13:   $\mathcal{M} \leftarrow \{\theta : \rho(H(\theta)) > \gamma \text{ where } \theta = 0.5(\theta_k + \theta_{k+1}), k = 1, \dots, q-1\}$ 
14: end while
```

---

NOTE: For simplicity, we assume here that all eigenvalues are obtained exactly and optimization always finds a local maximizer  $\theta_*$  of (10) also exactly, with  $\rho(H(\theta_*)) > \rho(H(\theta_{\text{BBBS}}))$ , assuming  $\theta_{\text{BBBS}}$  is not stationary and  $\theta_* = \theta_{\text{BBBS}}$  otherwise. The method reduces to a BBBS-like iteration using (10) if lines 9 and 10 are replaced with  $\gamma \leftarrow \max_{\theta \in \mathcal{M}} \rho(H(\theta))$  and is terminated once  $\gamma \approx r(A)$ . Running optimization from other angles in  $\mathcal{M}$  (in addition to  $\theta_{\text{BBBS}}$ ) every iteration may also be advantageous, particularly if this can be done via parallel processing. Adding zero to the initial set  $\mathcal{M}$  avoids having to deal with “wrap-around” intervals (since they are no longer possible for all computed values of  $\gamma$ ), while  $\theta_{\text{max}}$  is just a reasonable guess for a global maximizer of (10).

Note that the smoothness conditions needed at global maximizers of (10) for Newton’s method to have quadratic local convergence also would imply that the BBBS-like iteration converges quadratically (but it would still be more expensive).

We now address some practical concerns for implementing Algorithm 1. First, there is the question of how much cheaper it is to compute all eigenvalues and eigenvectors of  $H(\theta)$  compared to computing all unimodular eigenvalues of  $R_\gamma - \lambda S$ , which affects which optimization method should be used. This is addressed by Table 1, where we benchmark `eig` and `eigs` for dense random matrices of varying dimension. For comparison purposes only, we also tested computing eigenvalues of  $R_\gamma - \lambda S$  by using

$$S^{-1}R_\gamma = \begin{bmatrix} 2\gamma A^{-1} & -A^{-1}A^* \\ I & 0 \end{bmatrix},$$

which has the benefit of being much faster but is not recommended numerically since it requires forming  $A^{-1}$  explicitly (and  $A$  may not even be invertible). For reliability, eigenvalues of  $R_\gamma - \lambda S$  should be computed using its pencil form, and ideally, via first transforming it to have Hamiltonian structure in order to use a structure-preserving eigensolver, e.g., those of [BBMX02, BSV16]. The main takeaway from Table 1 is that using `eig` to compute all eigenvalues and eigenvectors of  $H(\theta)$  is 11-15 times faster than computing all eigenvalues of  $S^{-1}R_\gamma$  (which again is not recommended) and 27-140 times faster than computing all eigenvalues of pencil  $R_\gamma - \lambda S$ . In other words, we can afford to do many eigenvalue computations with  $H(\theta)$  before it becomes a noticeable cost relative to a single computation with  $R_\gamma - \lambda S$ . The second takeaway is that `eigs` is at best only twice as fast as `eig` for the dense examples tested, and this is generally only seen for  $n \geq 1000$  and when a single eigenvalue is requested. Requesting only one eigenvalue can be problematic, since the desired eigenvalue may not necessarily be the one

$n$	$\mathbf{eigs}(H(\theta), k, \text{'LM'})$		$\mathbf{eigs}(H(\theta), k, \text{'LR'})$		$\mathbf{eigs}(H(\theta), k, \text{'BE'})$			$\mathbf{eig}(\cdot)$	
	$k = 1$	$k = 6$	$k = 1$	$k = 6$	$k = 2$	$k = 4$	$k = 6$	$\mathbf{eig}(S^{-1}R_\gamma)$	$\mathbf{eig}(R_\gamma, S)$
200	1.5	3.2	1.5	1.5	1.2	1.3	1.7	11.5	26.8
400	0.8	1.3	0.6	1.1	1.2	1.1	1.2	14.9	63.5
600	0.9	1.5	0.8	1.4	1.6	1.6	1.5	15.6	77.3
800	0.8	1.5	0.7	1.4	1.5	1.2	1.4	14.7	101.8
1000	0.6	1.2	0.6	1.0	1.0	1.0	1.2	12.9	95.3
1200	0.6	1.1	0.5	0.9	1.0	1.0	1.0	13.1	108.8
1400	0.5	0.9	0.5	0.8	1.7	0.8	0.9	14.0	113.3
1600	0.5	0.9	0.5	0.8	1.0	0.9	0.9	15.2	139.8

Table 1: Running times normalized to  $\mathbf{eig}(H(\theta))$ , i.e., the running time of a given operation divided by the corresponding running time of  $\mathbf{eig}(H(\theta))$ , which ranged from 0.0108 seconds ( $n = 200$ ) to 1.35 seconds ( $n = 1600$ ). For each  $n$ , five  $A$  matrices were generated via  $\mathbf{randn}(n)+1i*\mathbf{randn}(n)$  to construct the corresponding five  $H(\theta)$ ,  $R_\gamma$ , and  $S$  matrices, where  $\theta = 1$  and  $\gamma = \rho(H(1))$ . To account for variability, the average time of a given operation over the five different examples was used to obtain the numbers below. Computing eigenvalues of  $H(\theta)$  via  $\mathbf{eigs}$  was done in multiple ways, where  $k$  is the number of eigenvalues requested, 'LM' (largest modulus) means request the outermost eigenvalues, 'LR' (largest real) means request the largest eigenvalues, and 'BE' (both ends) means request that half of the  $k$  eigenvalues be the smallest eigenvalues and the rest be the largest eigenvalues. For  $H(\theta)$ , eigenvectors were requested, as they are needed to obtain the tangential boundary points. For  $S^{-1}R_\gamma$  and  $(R_\gamma, S)$ , only eigenvalues were requested.

$\mathbf{eigs}$  first finds and returns. If more than one eigenvalue is requested, e.g., for robustness, then using  $\mathbf{eigs}$  has little to no benefit and in fact can be noticeably slower than  $\mathbf{eig}$ , at least for dense matrices up to  $n = 1600$ . When running the same tests with sparse examples generated via  $\mathbf{sprandn}(n, n, 0.05)+1i*\mathbf{sprandn}(n, n, 0.05)$ ,  $\mathbf{eigs}$  (in any of the configurations) was up to five times faster than  $\mathbf{eig}$  and already about two times faster for  $n = 400$ . However, in the context of optimization, resorting to  $\mathbf{eigs}$  means forgoing  $\rho''(H(\theta))$  and so the superlinearly convergent secant method must be used in lieu of Newton's method. This likely means additional iterations will be incurred during optimization, which may offset any gains from using  $\mathbf{eigs}$  and potentially be slower overall, particular if  $A$  is dense. Meanwhile,  $\mathbf{eigs}$  is less reliable than  $\mathbf{eig}$  and calling  $\mathbf{eig}$  for  $H(\theta)$  is already an insignificant cost compared to computing eigenvalues of  $R_\gamma - \lambda S$ . Hence, for Algorithm 1, using  $\mathbf{eig}(H(\theta))$  and doing Newton's method seems most appropriate. On the other hand, for Uhlig's cutting-plane-based method, which only needs the largest real eigenvalue of  $H(\theta)$ ,  $\mathbf{eigs}$  is perhaps the more efficient choice, particularly if  $A$  is sparse.

Finally, there are a few important details for implementing Algorithm 1 in practice, where computations will not be exact. First, at each iteration, the unimodular eigenvalues of  $R_\gamma - \lambda S$  should not be computed with the current value of estimate  $\gamma$  but rather  $\gamma(1 + \mathbf{tol})$ , for some relative termination tolerance  $\mathbf{tol} > 0$ . This is because maximizers of (10) generally will not be found exactly. If a global maximizer of (10) has been obtained to the desired accuracy, then  $\gamma(1 + \mathbf{tol}) > r(A)$  and so the algorithm will terminate. This helps to prevent computing eigenvalues of multiple instances of  $R_\gamma - \lambda S$  once  $\gamma \approx r(A)$ . Relatedly, Algorithm 1 should terminate if no unimodular eigenvalues of  $R_\gamma - \lambda S$  are obtained. Second, since  $\rho(H(\theta))$  may be nonsmooth and have stationary points that are not local maximizers, it is not certain that Newton's method will converge to a local maximizer. However, this is not such a problem since even if optimization makes no progress in increasing  $\gamma$ , Algorithm 1 still takes a BBBS-type step. Furthermore, we can mitigate against encountering slow convergence by simply setting a low limit, say, 15, for the maximum number of iterations of Newton's method. Third, when  $\theta_{\text{BBBS}}$  is a stationary point of (10), it is associated with a double unimodular eigenvalue of  $R_\gamma - \lambda S$ , where  $\gamma = \rho(H(\theta_{\text{BBBS}}))$ . If  $\theta_{\text{BBBS}}$  is not a maximizer and rounding errors prevent this double eigenvalue from being detected, the resulting level-set interval  $[\theta_k, \theta_{k+1}]$  will contain  $\theta_{\text{BBBS}}$ . If

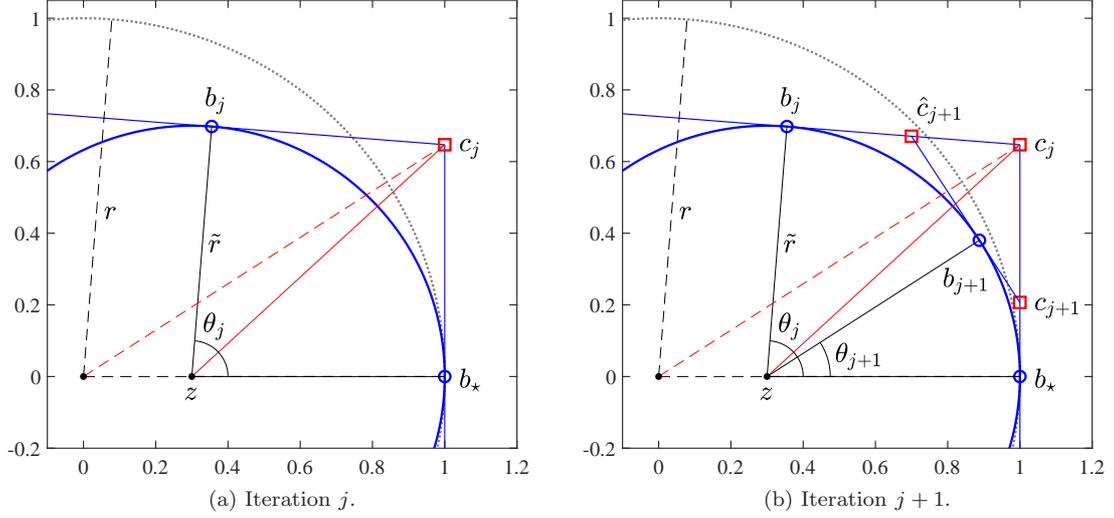


Figure 1: Depiction of Uhlig’s method on an example where  $W(A)$  is a circular disk of radius  $\tilde{r} = 0.7$  centered at  $z = 0.3$ , and so  $r = r(A) = 1$  is attained at the point  $b_* \in \partial W(A)$ . The dotted circle is the circle of radius  $r(A) = 1$  centered at the origin. See Example 4.2 for a complete description.

$\theta_{\text{BBBS}}$  also happens to be its midpoint (or near to it), then the algorithm may get stuck here, precisely because  $\gamma = \rho(H(\theta_{\text{BBBS}}))$ . Note that this issue can occur even in the standard BBBS iteration. As pointed out in [BLO03, p. 372-3] for the criss-cross algorithm for the pseudospectral abscissa, a robust fix is simple, namely to split any level-set interval  $[\theta_k, \theta_{k+1}]$  into  $[\theta_k, \theta_{\text{BBBS}}]$  and  $[\theta_{\text{BBBS}}, \theta_{k+1}]$  whenever  $\theta_{\text{BBBS}} \approx 0.5(\theta_k + \theta_{k+1})$ .

## 4 Convergence rate analysis of Uhlig’s iteration

To analyze the convergence rate of Uhlig’s method, we will consider how it behaves on a certain example, which will in fact inform how it behaves for general problems in terms of local curvature at outermost points of  $W(A)$ . We now define this precisely.

**Definition 4.1.** Let  $b_* \in \partial W(A)$  be a locally outermost point of  $W(A)$  and let  $r_* \geq 0$  be the radius of the osculating circle of  $\partial W(A)$  at  $b_*$ , where  $r_* = 0$  if  $b_*$  happens to be a corner of  $\partial W(A)$ . Then we define the *normalized curvature of  $\partial W(A)$  at  $b_*$*  as

$$\mu := \frac{r_*}{|b_*|} \in [0, 1].$$

When  $\mu = 0$ , the osculating circle actually just a point, i.e., the case when  $b_*$  is a corner of  $\partial W(A)$  and so  $b_* \in \Lambda(A)$ . As  $\mu \rightarrow 1$ ,  $A$  becomes closer to a disk matrix, with  $A$  being a disk matrix when  $\mu = 1$ .

**Example 4.2** (See Figure 1 for a diagrammatic description). For  $n \geq 2$ , consider

$$A = zI + \tilde{r}K_n, \quad \text{where} \quad K_2 = \begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix}, \quad K_3 = \begin{bmatrix} 0 & \sqrt{2} & 0 \\ 0 & 0 & \sqrt{2} \\ 0 & 0 & 0 \end{bmatrix}, \quad K_n = \begin{bmatrix} 0 & \sqrt{2} & & & \\ & 1 & & & \\ & & \ddots & & \\ & & & 1 & \\ & & & & \sqrt{2} \\ & & & & & 0 \end{bmatrix},$$

$\tilde{r} > 0$ , and  $z \geq 0$ . Since  $K_n$  is a disk matrix with  $r(K_n) = 1$  (see [LO18, Section 4]),  $W(A)$  is a circular disk with radius  $\tilde{r}$  centered at  $z$ . Hence,  $r(A) = z + \tilde{r}$  with  $r(A)$  being attained

at  $b_\star = z + \tilde{r} \in \partial W(A)$  on the real axis. Letting  $r$  be short for  $r(A)$  (which we will often use), the normalized curvature of  $\partial W(A)$  at  $b_\star$  is  $\mu = \frac{\tilde{r}}{r}$ . Suppose  $\mathcal{P}_j$  contains the supporting hyperplane  $\bar{\theta}_\star = \text{Arg}(b_\star) = 0$  and that the next tangential boundary point in  $\partial W(A) \cap \mathcal{P}_j$  in the counter-clockwise direction is  $b_j$ . Let  $c_j$  be the corner of  $\mathcal{P}_j$  between  $b_j$  and  $b_\star$ . If Uhlig's method now attempts to cut corner  $c_j$ ,  $\mathcal{P}_{j+1}$  will be produced by adding the supporting hyperplane for  $\bar{\theta}_j = -\text{Arg}(c_j)$  to  $\mathcal{P}_j$ , which will result in a new tangential boundary point  $b_{j+1}$  and two new corners,  $\hat{c}_{j+1}$  and  $c_{j+1}$ , which are, respectively, in the counter-clockwise and clockwise directions relative to  $b_{j+1}$ . Consider the sequence of corners  $\{c_k\}$  produced by Uhlig's cutting procedure if it is only applied to the clockwise-direction corners, i.e.,  $\{c_k\} = c_j, c_{j+1}, \dots$  all of which lie on the tangential line  $L_{\bar{\theta}_\star}$  passing through  $b_\star$ . Then for all  $k \geq j$ ,

$$(D1) \quad \theta_{k+1} = \text{Arg}(b_{k+1} - z) = \arctan\left(\mu \tan \frac{1}{2}\theta_k\right) \text{ where } \theta_k := \text{Arg}(b_k - z),$$

$$(D2) \quad r(A) < |c_k|$$

$$(D3) \quad \text{if } |\hat{c}_k| \leq r(A), \text{ then } |\hat{c}_{k+1}| < r(A) \text{ and } |c_{k+1}| < r(A),$$

noting that the angle associated with the supporting plane for  $b_k$  is  $\bar{\theta}_k = -\theta_k$ .

It is easy to see that the last two properties (D2) and (D3) must hold, while the first holds by the following geometric argument. First note that

$$\tilde{r} \tan \frac{1}{2}\theta_k = |c_k - b_\star| \quad \text{and} \quad \text{Arg}(c_k) = \arctan\left(\frac{|c_k - b_\star|}{r}\right).$$

Then  $\text{Arg}(c_k) = \arctan\left(\mu \tan \frac{1}{2}\theta_k\right)$  follows by substituting the first of these equations into the second, and so the first property (D1) is obtained by additionally noting that  $\text{Arg}(c_k) = \text{Arg}(b_{k+1} - z)$  also holds.

**Theorem 4.3.** *The sequence  $\{\theta_k\}$  produced by Uhlig's cutting procedure as described in Example 4.2 converges to zero Q-linearly with rate  $\frac{1}{2}\mu$ .*

*Proof.* First note that  $\lim_{k \rightarrow \infty} \theta_k = \theta_\star = 0$  and  $\theta_k > 0$  for all  $k \geq j$ . Then

$$\lim_{k \rightarrow \infty} \frac{|\theta_{k+1} - \theta_\star|}{|\theta_k - \theta_\star|} = \lim_{k \rightarrow \infty} \frac{\theta_{k+1}}{\theta_k} = \lim_{k \rightarrow \infty} \frac{\arctan\left(\mu \tan \frac{1}{2}\theta_k\right)}{\theta_k}.$$

Since the numerator and denominator both go to zero as  $k \rightarrow \infty$ , the result follows by considering the continuous version of the limit:

$$\lim_{\theta \rightarrow 0} \frac{\arctan\left(\mu \tan \frac{\theta}{2}\right)}{\theta} = \lim_{\theta \rightarrow 0} \frac{\mu \tan \frac{1}{2}\theta}{\theta} = \lim_{\theta \rightarrow 0} \frac{\frac{1}{2}\mu\theta}{\theta} = \frac{1}{2}\mu,$$

where the first and second equalities are obtained, respectively, using small-angle approximations  $\arctan x \approx x$  and  $\tan x \approx x$  for  $x \approx 0$ . Hence,  $\theta_k$  converges to zero Q-linearly with rate of  $\frac{1}{2}\mu$ .  $\square$

Although Theorem 4.3 gives us the exact rate of convergence of the sequence of angles produced by Uhlig's iteration, the angles themselves do not indicate when the method should terminate (to a tolerance). For that, we must look at when the upper bound given by (D2) is sufficiently close to  $r(A)$ , so we now consider how the moduli of these outermost corners  $c_k$  converge.

**Theorem 4.4.** *The sequence  $\{|c_k|\}$  produced by Uhlig's cutting procedure as described in Example 4.2 converges to  $r(A)$  Q-linearly with rate  $\frac{1}{4}\mu^2$ .*

*Proof.* Recalling that  $r = r(A)$ , first note that

$$\cos \theta_{k+1} = \frac{r}{|c_k|} \quad \text{and so} \quad |c_k| = r \sec \theta_{k+1} > r$$

for all  $k \geq j$ . Hence, we consider the limit

$$\lim_{k \rightarrow \infty} \frac{||c_{k+1}| - r|}{||c_k| - r|} = \lim_{k \rightarrow \infty} \frac{r \sec \theta_{k+2} - r}{r \sec \theta_{k+1} - r} = \lim_{k \rightarrow \infty} \frac{\sec \theta_{k+1} - 1}{\sec \theta_k - 1},$$

which when substituting in  $\theta_{k+1} = \arctan(\mu \tan \frac{1}{2}\theta_k)$  becomes

$$\lim_{k \rightarrow \infty} \frac{\sec(\arctan(\mu \tan \frac{1}{2}\theta_k)) - 1}{\sec \theta_k - 1}.$$

Since the numerator and denominator both go to zero as  $k \rightarrow \infty$ , we consider the continuous version of the limit, i.e.,

$$\lim_{\theta \rightarrow 0} \frac{\sec(\arctan(\mu \tan \frac{1}{2}\theta)) - 1}{\sec \theta - 1} = \lim_{\theta \rightarrow 0} \frac{\sec(\mu \tan \frac{1}{2}\theta) - 1}{\sec \theta - 1} = \lim_{\theta \rightarrow 0} \frac{\sec(\frac{1}{2}\mu\theta) - 1}{\sec \theta - 1},$$

again using the small-angle approximations for arctan and tan. As this is an indeterminate form, we will apply L'Hôpital's Rule. However, it will be convenient to first multiply by the following identity to get rid of the  $\sec \theta$  in the denominator:

$$\lim_{\theta \rightarrow 0} \frac{\cos \theta}{\cos \theta} \cdot \frac{\sec(\frac{1}{2}\mu\theta) - 1}{\sec \theta - 1} = \lim_{\theta \rightarrow 0} \frac{\cos \theta (\sec(\frac{1}{2}\mu\theta) - 1)}{1 - \cos \theta} = \lim_{\theta \rightarrow 0} \frac{f_1(\theta)}{f_2(\theta)}.$$

For  $f_2(\theta)$ ,  $f_2'(\theta) = \sin \theta$ , while for  $f_1(\theta)$ , we have

$$\begin{aligned} f_1'(\theta) &= -\sin \theta (\sec(\frac{1}{2}\mu\theta) - 1) + \frac{1}{2}\mu \cos \theta \cdot \sec(\frac{1}{2}\mu\theta) \cdot \tan(\frac{1}{2}\mu\theta) \\ &= -\sin \theta (\sec(\frac{1}{2}\mu\theta) - 1) + \frac{1}{2}\mu \cos \theta \cdot \frac{\sin(\frac{1}{2}\mu\theta)}{\cos^2(\frac{1}{2}\mu\theta)}, \end{aligned}$$

where the second equality uses that  $\sec x = \frac{1}{\cos x}$  and  $\tan x = \frac{\sin x}{\cos x}$ . As  $f_1'(0) = 0$  and  $f_2'(0) = 0$ , we still have an indeterminate form and so will again apply L'Hôpital's Rule. For the denominator, we have  $f_2''(\theta) = \cos \theta$  and so  $f_2''(0) = 1$ . For the numerator, it will be convenient to write  $f_1'(\theta) = -g_1(\theta) + \frac{1}{2}\mu g_2(\theta)$ , where

$$g_1(\theta) = \sin \theta (\sec(\frac{1}{2}\mu\theta) - 1) \quad \text{and} \quad g_2(\theta) = \cos \theta \cdot \frac{\sin(\frac{1}{2}\mu\theta)}{\cos^2(\frac{1}{2}\mu\theta)},$$

and differentiate the parts separately. For  $g_1(\theta)$ , we have that

$$g_1'(\theta) = \cos \theta (\sec(\frac{1}{2}\mu\theta) - 1) + \frac{1}{2}\mu \sin \theta \cdot \sec(\frac{1}{2}\mu\theta) \cdot \tan(\frac{1}{2}\mu\theta),$$

and so  $g_1'(0) = 0$ . For  $g_2(\theta)$ , we have that

$$g_2'(\theta) = -\sin \theta \cdot \frac{\sin(\frac{1}{2}\mu\theta)}{\cos^2(\frac{1}{2}\mu\theta)} + \cos \theta \cdot \frac{\frac{1}{2}\mu (\cos^3(\frac{1}{2}\mu\theta) + 2 \sin^2(\frac{1}{2}\mu\theta) \cdot \cos(\frac{1}{2}\mu\theta))}{\cos^4(\frac{1}{2}\mu\theta)},$$

and so  $g_2'(0) = \frac{1}{2}\mu$ . Hence,  $f_1''(0) = -g_1'(0) + \frac{1}{2}\mu g_2'(0) = \frac{1}{4}\mu^2$  and so  $\{|c_k|\}$  converges Q-linearly with rate  $\frac{1}{4}\mu^2$ .  $\square$

Theorem 4.4 allows us to estimate how many iterations will be needed until it is no longer necessary to refine corner  $c_k$  from Example 4.2, i.e., the value of  $k$  such that  $|c_k| \leq r(A) \cdot (1 + \mathbf{tol})$ . For simplicity, it will be more convenient to assume that  $c_j$  is  $c_0$ , with  $|c_0| = s_0 r(A)$  for some scalar  $s_0 > (1 + \mathbf{tol})$ . Using the Q-linear rate given by Theorem 4.4, we have that

$$|c_k| - r(A) \leq (|c_0| - r(A)) \cdot \left(\frac{1}{4}\mu^2\right)^k,$$

and so if

$$r(A) + (|c_0| - r(A)) \cdot \left(\frac{1}{4}\mu^2\right)^k \leq r(A) \cdot (1 + \mathbf{tol}),$$

then it follows that  $|c_k| \leq r(A) \cdot (1 + \mathbf{tol})$ . In other words, in this case, the modulus of corner  $c_k$  (of the polygonal approximation) is sufficiently close enough to  $r(A)$  such that it no longer needs to be considered for cutting. Hence, by first dividing the above equation by  $r(A)$  and doing some simple manipulations, we have that  $|c_k|$  is sufficiently close to  $r(A)$  if

$$k \geq \frac{\log(\mathbf{tol}) - \log(s_0 - 1)}{\log\left(\frac{1}{4}\mu^2\right)}. \quad (14)$$

So for Example 4.2, if  $s_0 = 100$  and  $\mathbf{tol} = 1\mathbf{e-14}$ , we have that  $k \approx 27, 14, 7$ , and 4 iterations are needed, respectively, for normalized curvatures  $\mu = 1, 0.5, 0.1$ , and 0.01. While this is linear convergence, it is rather fast linear convergence.

Theorems 4.3 and 4.4 are quite illuminating with regard to how Uhlig's method behaves in general. Assume that  $b_\star \in W(A)$  is an outermost point where  $r(A)$  is attained and the normalized curvature of  $\partial W(A)$  at  $b_\star$  is  $\mu$ . Without loss of generality, we can assume  $b_\star$  is on the positive real axis. In the limiting case of  $\mu \rightarrow 0$ ,  $b_\star$  is a corner, and the refinement of corners  $c_k$  (as described in Example 4.2) will converge superlinearly. If  $\mu = 1$ , the osculating circle of  $\partial W(A)$  at  $b_\star$  is  $\partial W(A)$  itself, as  $A$  must be a disk matrix. Since this case is covered by Example 4.2, Theorems 4.3 and 4.4 directly apply, and so  $\{|c_k|\}$  converges linearly with rate  $\frac{1}{4}$ . When  $\mu \in (0, 1)$ , we can apply Theorems 4.3 and 4.4 by considering the osculating circle of  $\partial W(A)$  at  $b_\star$  and additionally assuming that Uhlig's method has already obtained the supporting hyperplane for  $b_\star$ . Under this assumption,  $\{|c_k|\}$  then converges linearly with rate  $\frac{1}{4}\mu^2$ . While Uhlig's method may only obtain the supporting hyperplane for  $b_\star$  as it converges, note that it could be cheaply added to the polygonal approximation, which we will take advantage of later on for our second improved algorithm in §5. Per §3, local maximizers of (10) can be obtained with fast convergence using optimization, and given a good enough initial guess, such optimization would yield a global maximizer.

While Theorems 4.3 and 4.4 show that Uhlig's cutting procedure has rather fast linear convergence (and superlinear at corners), these analyses do not take into account the fact that other corners, besides the sequence  $\{c_k\}$ , may also need to be refined. As mentioned earlier, when  $A$  is a disk matrix, i.e.,  $\mu = 1$ , Uhlig's method must accurately approximate all of  $\partial W(A)$  in order to assert  $r(A)$  has been computed to sufficient accuracy, since  $|c| \leq r(A) \cdot (1 + \mathbf{tol})$  must hold for every corner  $c$  in the polygonal approximation to  $W(A)$ . In other words, when  $W(A)$  is a disk centered at the origin, the fast linear convergence explained by Theorems 4.3 and 4.4 actually must be incurred over and over again, precisely because there are infinitely many outermost points in  $W(A)$ . Even if  $\mu < 1$  but  $\mu \approx 1$  still holds,  $W(A)$  will still be very close to a disk centered at the origin, and so Uhlig's method will still need to approximate all of  $\partial W(A)$  to a high of level accuracy. To illustrate just how expensive this can get, when  $\mu = 1$  it is straightforward to determine the minimum number of supporting hyperplanes needed to ensure that  $|c| \leq r(A) \cdot (1 + \mathbf{tol})$  holds for all corners of approximation  $\mathcal{P}_j$ . The optimal solution is when  $\mathcal{P}_j$  is a  $k$ -vertex regular polygon (also centered at the origin), where

$$k \geq \left\lceil \frac{\pi}{\operatorname{arcsec}(1 + \mathbf{tol})} \right\rceil. \quad (15)$$

Hence,  $k$  must be at least 223, 22,215, 2,221,343, and 149,078,414, respectively, when  $\mathbf{tol}$  is set to  $1\mathbf{e-4}$ ,  $1\mathbf{e-8}$ ,  $1\mathbf{e-12}$ , and  $\mathbf{eps}$ , where  $\mathbf{eps} \approx 2.22 \times 10^{-16}$ . Furthermore, Uhlig's method may require up to double the number of supporting hyperplanes indicated by (15). If Uhlig's method happens to build a regular polygon with  $k - 1$  vertices, then  $|c| > r(A) \cdot (1 + \mathbf{tol})$  will hold for all of its corners, and so Uhlig's method will need to refine every single one, thus doubling the number of supporting hyperplanes. Clearly level-set-based methods, like ours described in §3, should be much faster than Uhlig's method when computing  $r(A)$  of disk matrices to more than just a few digits.

---

**Algorithm 2** An Improved Cutting-Plane-Based Algorithm

---

**Input:**  $A \in \mathbb{C}^{n \times n}$ .**Output:**  $\gamma$  such that  $\gamma = r(A)$ .

```
1: if  $A$  is normal then
2:    $\gamma \leftarrow \rho(A)$ 
3:   return
4: end if
5:  $\theta_{\max} \leftarrow$  angle of an eigenvalue of  $A$  attaining  $\rho(A)$  (0 if this eigenvalue is 0)
6:  $\mathcal{P}_0 \leftarrow$  the supporting hyperplanes given by angles  $[0, 0.5\pi, \pi, 1.5] - \theta_{\max}$ 
7: for  $k = 0, 1, 2, \dots$  do
8:    $\gamma \leftarrow \max |b|$  over all boundary points  $b \in \mathcal{P}_k$ 
9:    $\bar{\theta} \leftarrow$  angle of the supporting hyperplane with the outermost boundary point
10:  if  $\rho(H(\bar{\theta})) \geq \gamma$  and  $\bar{\theta}$  is not a stationary point then
11:     $\gamma \leftarrow$  maximization of  $\rho(H(\theta))$  via local optimization initialized at  $\bar{\theta}$ 
12:     $\mathcal{P}_k \leftarrow \mathcal{P}_k \cap P_{\bar{\theta}_1} \cap \dots \cap P_{\bar{\theta}_q}$  for all angles encountered during optimization
13:  end if
14:   $c \leftarrow$  outermost corner of  $\mathcal{P}_k$ 
15:  if the optimal cut can be applied to  $c$  per §5.2 then
16:     $\mathcal{P}_{k+1} \leftarrow \mathcal{P}_k \cap P_{\bar{\theta}}$  for angle  $\bar{\theta}$  given by (22)
17:  else
18:     $\mathcal{P}_{k+1} \leftarrow \mathcal{P}_k \cap P_{\bar{\theta}}$  for angle  $\bar{\theta} = -\text{Arg}(c)$  (Uhlig's cut)
19:  end if
20: end for
```

---

NOTE: For simplicity, we forgo giving pseudocode to describe optimizations that exploit symmetry of  $W(A)$  and the termination conditions that are needed in practice. Note that the rotation of initial rectangular bounding box  $\mathcal{P}_0$  for  $W(A)$ , as specified by  $\theta_{\max}$ , differs from Uhlig's approach of using the direction of the average of the eigenvalues of  $A$ .

## 5 An improved cutting-plane algorithm

In Algorithm 2, we give pseudocode for our second improved iteration, which addresses some inefficiencies in Uhlig's approach. The main two components of the algorithm are as follows. First, by also leveraging local optimization, like we have done for Algorithm 1, we can efficiently find locally outermost points in  $W(A)$ . Second, given such a boundary point which is known to be outermost, we can invoke a new cutting procedure which minimizes the total number of cuts needed to compute  $r(A)$  to the desired level of accuracy. When this new cut cannot be invoked, we will fall back to using Uhlig's cutting procedure. In the next three subsections, we describe our new cutting procedure, how it can be accurately estimated in practice, and then derive its exact Q-linear rate of convergence for Example 4.2.

### 5.1 An optimal-cutting strategy

Let  $b_j$  and  $b_\star$  be two consecutive tangential boundary points of  $W(A)$  on the polygonal approximation  $\mathcal{P}_j$ , and for concreteness, without loss of generality we can assume that points are ordered in the clockwise direction. Suppose that  $|b_j| < r$ ,  $|b_\star| = \gamma > 0$ , and  $b_\star$  is a locally outermost point in  $W(A)$ , hence  $\gamma$  is the best (largest) lower bound for  $r(A)$  known so far. Figure 1 shows a depiction using an instance of Example 4.2 where  $\gamma = r(A) = r$ . In Figure 1b, Uhlig's cut of the corner  $c_j$  between  $b_j$  and  $b_\star$  produces two new corners  $\hat{c}_{j+1}$  and  $c_{j+1}$ , but since  $|\hat{c}_{j+1}| < \gamma$  and  $|c_{j+1}| > \gamma$ , it is only necessary to subsequently refine  $c_{j+1}$ . However, in Figure 2a we show

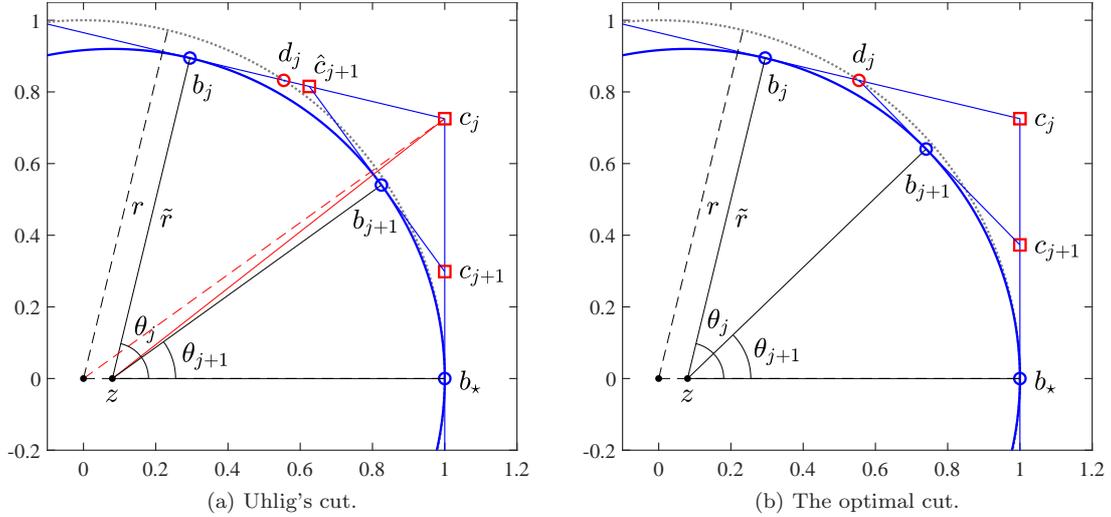


Figure 2: Depictions of corner  $c_j$  between tangential boundary points  $b_j$  and  $b_*$  being cut by Uhlig's procedure (left) and the optimal cut (right). In this scenario, the two new corners produced by Uhlig's cut must both be subsequently refined ( $\hat{c}_{j+1}$  and  $c_{j+1}$ ), whereas in Figure 1b, Uhlig's cut is shown when only one of the two new corners must be refined ( $c_{j+1}$ ). Meanwhile, the optimal cut always makes the largest possible reduction in  $\theta_j$  such that only corner  $c_{j+1}$  must be refined.

another scenario where both of the two corners produced by Uhlig's cut will require subsequent cutting as well. So while Theorems 4.3 and 4.4 indicate the number of iterations Uhlig's method will need to sufficiently refine the sequence  $\{c_k\}$ , the actual number of eigenvalue computations with  $H(\theta)$  could be significantly higher (roughly double), since the corners to the left, e.g.,  $\hat{c}_{j+1}$ , also may need to be refined. In other words, as we have just seen for the special case of disk matrices, the real cost of cutting-plane methods like Uhlig's is not so much how quickly they find outermost points of  $W(A)$  but rather how quickly they can trim outermost corners until none have moduli more than  $r(A) \cdot (1 + \epsilon)$ .

Comparing Figures 1b and 2a immediately suggests that an optimal cutting strategy would be to make the largest reduction in the angle  $\theta_j$  (as described in Example 4.2 and shown in these figures) such that  $|\hat{c}_{j+1}| = \gamma$ , recalling that  $\hat{c}_{j+1}$  is between  $b_j$  and  $c_j$  on the tangent line for  $b_j$ . In Figure 2a, this ideal corner is labeled  $d_j$ , while the corresponding optimal cut for this same example is shown in Figure 2b, where  $d_j$  coincides with  $\hat{c}_{j+1}$ , and so the latter is not labeled. Recalling the definitions of the tangential line  $L_\theta$  and supporting hyperplane  $P_\theta$  for a tangential boundary point  $b_\theta$  (from the beginning of §2),  $d_j$  is determined by

$$d_j = e^{-i\bar{\theta}_j}(\lambda_j + it) \quad \text{where} \quad t = -\sqrt{\gamma^2 - \lambda_j^2}, \quad (16)$$

$\bar{\theta}_j$  is the angle of the supporting hyperplane for  $b_j$ , and  $\lambda_j = \lambda_{\max}(H(\bar{\theta}_j))$ .

In §5.3, we will deduce the exact Q-linear rate of convergence of the optimal cut for Example 4.2, where the optimal cut can be described in an analytical recursive form and computed exactly. As we will see, the convergence rate of the optimal cut is often better than Uhlig's procedure, and also describes the convergence rate for our improved cutting-plane algorithm. However, before this analysis, we first discuss how optimal cuts can be accurately estimated for general numerical radius problems, i.e., where the field of values is not restricted to be a disk as in Example 4.2.

## 5.2 Computing the optimal cut

Let  $b_j$  and  $b_*$  be as before, and without loss of generality, assume that  $\text{Arg}(b_*) = 0$  (i.e.,  $b_*$  is on the real axis) and  $\text{Arg}(b_j) \in (0, \pi)$ . A necessary condition for the optimal cut to be applicable is that  $\bar{\theta}_j \in (-\frac{\pi}{2}, 0)$ , i.e., the tangential line for  $b_j$  is decreasing from left to right in the complex plane. The optimal cut for corner  $c_j$  will produce a new tangential boundary point  $b_{j+1}$  between  $b_j$  and  $b_*$  such that the new tangential line associated with the supporting hyperplane for  $b_{j+1}$  passes through  $d_j$ . Let angle  $\bar{\theta}_{\text{Opt}} \in (\bar{\theta}_j, 0)$  denote the angle of this optimal supporting hyperplane for  $b_{j+1}$ . Since we no longer assume that the field of values is a disk, we do not know its boundary between  $b_j$  and  $b_*$  and so will need to approximate it in order to estimate  $\bar{\theta}_{\text{Opt}}$ .

Consider the sideways quadratic (opening up to the left in the complex plane)

$$q(y) = q_2 y^2 + q_1 y + q_0,$$

fitted such that it passes through  $b_j$  for  $y = \text{Im } b_j$  and  $b_*$  for  $y = 0$ . Since these two conditions leave one remaining degree of freedom to determine  $q(y)$  uniquely and  $b_*$  is a locally outermost point of  $\partial W(A)$ , we also specify that  $q(y)$  should be tangent to  $\partial W(A)$  at  $b_*$ . Hence,  $q_0$ ,  $q_1$ , and  $q_2$  are determined by

$$q(0) = \gamma, \tag{17a}$$

$$q(\text{Im } b_j) = \text{Re } b_j, \tag{17b}$$

$$q'(0) = 0. \tag{17c}$$

Solving these equations yields

$$q_2 = \frac{\text{Re } b_j - \gamma}{(\text{Im } b_j)^2}, \quad q_1 = 0, \quad \text{and} \quad q_0 = \gamma. \tag{18}$$

We can assess whether  $q(y)$  is a good fit for  $\partial W(A)$  about  $b_*$  by checking how close  $q(y)$  is to being also tangent to  $\partial W(A)$  at  $b_j$ , i.e.,  $q(y)$  is a good fit if

$$q'(\text{Im } b_j) \approx \tan \bar{\theta}_j. \tag{19}$$

If these two values are not sufficiently close, then we will consider  $q(y)$  a poor local approximation of  $\partial W(A)$  at  $b_j$  (and  $b_*$ ) and will not attempt do an optimal cut; in this case, we will instead just use Uhlig's cutting procedure. Otherwise, when  $q(y)$  does accurately reflect the region of  $\partial W(A)$  between  $b_j$  and  $b_*$ , it can be used to estimate  $\bar{\theta}_{\text{Opt}}$  as follows. Given our good approximation  $q(y)$ , we need to determine the line

$$l(y) = l_1 y + l_0,$$

such that  $l(y)$  passes through  $d_j$  for  $y = \text{Im } d_j$  and is tangent to the quadratic  $q(y)$  for some  $\tilde{y} \in (0, \text{Im } d_j)$ . Hence, we have the following set of equations to solve in order to determine  $l_0$ ,  $l_1$  and  $\tilde{y}$ :

$$\text{Re } d_j = l(\text{Im } d_j) \iff \text{Re } d_j = l_1 \text{Im } d_j + l_0, \tag{20a}$$

$$q(\tilde{y}) = l(\tilde{y}) \iff q_2 \tilde{y}^2 + q_0 = l_1 \tilde{y} + l_0, \tag{20b}$$

$$q'(\tilde{y}) = l'(\tilde{y}) \iff 2q_2 \tilde{y} = l_1. \tag{20c}$$

Solving these equations yields

$$\tilde{y} = \text{Im } d_j - \sqrt{(\text{Im } d_j)^2 + \frac{q_0 - \text{Re } d_j}{q_2}}, \quad l_0 = -q_2 \tilde{y}^2 + q_0, \quad \text{and} \quad l_1 = \frac{\text{Re } d_j - l_0}{\text{Im } d_j}, \tag{21}$$

where  $l_1$  follows directly from (20a),  $l_0$  is obtained by substituting the value of  $l_1$  given in (20c) into (20b), and  $\tilde{y}$  follows from substituting the value of  $l_0$  given in (21) into  $l_1$  in (21) (so that  $l_1$

now only has  $\tilde{y}$  as an unknown), and then substituting this version of  $l_1$  into (20c), which results in a quadratic equation in  $\tilde{y}$ . Thus, if  $q(y)$  is a sufficiently accurate approximation of  $\partial W(A)$ , it follows that

$$\bar{\theta}_{\text{Opt}} \approx \arctan l_1. \quad (22)$$

Our fitted quadratic  $q(y)$  also gives us an estimate of the normalized curvature of  $\partial W(A)$  at  $b_*$ , namely

$$\mu_{\text{est}} := \frac{1}{2q_2\gamma}, \quad (23)$$

as the osculating circle of  $q(y)$  at  $y = 0$  has radius  $\frac{1}{2}q_2$ . This estimate will be useful in our third algorithm given in §6, which is a hybrid that begins by running Algorithm 2 but automatically switches to Algorithm 1 once it detects that the cutting-plane approach will be (much) less efficient.

In terms of implementation, if  $b_*$  is not on the positive real axis, we can simply first rotate all the points so that it then is, compute the rotated version of  $\bar{\theta}_{\text{Opt}}$ , and then rotate everything back. Similarly, if  $b_j$  is not in the upper half of the complex plane, we simply flip the problem, compute the flipped version of  $\bar{\theta}_{\text{Opt}}$ , and then flip back. Often both rotating and flipping will be needed. Lastly, in inexact arithmetic, we cannot expect that the optimal cut will pass through  $d_j$  exactly, and so rounding error may cause  $\hat{c}_{j+1}$  to be computed such that  $|\hat{c}_{j+1}| \approx \gamma$  but  $|\hat{c}_{j+1}| > \gamma(1 + \text{tol})$  still holds. As this rounding error would cause this second corner to also be subsequently refined, we first perturb  $d_j$  such that it is a tiny bit closer to  $b_j$ , i.e., we replace it with  $(1 - \delta)d_j + \delta b_j$  for some small value of  $\delta \in (0, 1)$ .

Finally, note that as  $b_j$  converges to  $b_*$ , the local approximation of  $\partial W(A)$  given by  $q(y)$  becomes better and better, and so the osculating circle of  $q(y)$  converges to the osculating circle of  $\partial W(A)$  at  $b_*$ . Hence, the exact convergence rate of optimal cuts applied to Example 4.2, which we derive next, should essentially describe the convergence rate of optimal cuts at locally outermost points of arbitrary problems, where  $W(A)$  may have any (permissible) shape.

### 5.3 Convergence analysis of the optimal cut

In order to analyze the convergence rate of the sequence of angles  $\{\theta_k\}$  produced by the optimal cutting strategy on Example 4.2, we will first need to derive the exact recursion formula for this sequence, which we do in the following two lemmas.

**Lemma 5.1.** *For Example 4.2 with  $\mu \in (0, 1)$ , consider corner  $c_j$  between tangential boundary points  $b_j$  and  $b_*$  with  $|b_j| < r$ ,  $|b_*| = r$ ,  $\theta_j = \text{Arg}(b_j - z) \in (0, \pi)$ , and  $\text{Arg}(b_*) = 0$ . The point  $d_j$  on the tangent line for  $b_j$  and closest to  $c_j$  with  $|d_j| = r$  is*

$$d_j = b_j - \mathbf{i}t_j e^{\mathbf{i}\theta_j}, \quad (24)$$

where  $b_j = z + \tilde{r}e^{\mathbf{i}\theta_j}$  and

$$t_j = -z \sin \theta_j + \sqrt{z^2 \sin^2 \theta_j - 2z\tilde{r}(\cos \theta_j - 1)} \geq 0. \quad (25)$$

*Proof.* It is easy to see that (24) must hold for some  $t \geq 0$ . To obtain (25), we use the fact that  $|d_j|^2 = r^2$  and solve for  $t$  using (24). First, we have

$$0 = |d_j|^2 - r^2 = (b_j - \mathbf{i}t e^{\mathbf{i}\theta_j})(\bar{b}_j + \mathbf{i}t e^{-\mathbf{i}\theta_j}) - r^2 = t^2 + \mathbf{i}t(b_j e^{-\mathbf{i}\theta_j} - \bar{b}_j e^{\mathbf{i}\theta_j}) + |b_j|^2 - r^2.$$

which by substituting in the following two equivalences

$$\begin{aligned} b_j e^{-\mathbf{i}\theta_j} - \bar{b}_j e^{\mathbf{i}\theta_j} &= (z + \tilde{r}e^{\mathbf{i}\theta_j})e^{-\mathbf{i}\theta_j} - (z + \tilde{r}e^{-\mathbf{i}\theta_j})e^{\mathbf{i}\theta_j} = z(e^{-\mathbf{i}\theta_j} - e^{\mathbf{i}\theta_j}) = -\mathbf{i}2z \sin \theta_j, \\ |b_j|^2 &= z^2 + \tilde{r}^2 + z\tilde{r}(e^{\mathbf{i}\theta_j} + e^{-\mathbf{i}\theta_j}) = z^2 + \tilde{r}^2 + 2z\tilde{r} \cos \theta_j, \end{aligned}$$

yields

$$0 = t^2 + t2z \sin \theta_j + (z^2 + \tilde{r}^2 - r^2 + 2z\tilde{r} \cos \theta_j).$$

By substituting in  $r^2 = (z + \tilde{r})^2 = z^2 + \tilde{r}^2 + 2z\tilde{r}$ , this simplifies further to

$$0 = t^2 + t2z \sin \theta_j + 2z\tilde{r}(\cos \theta_j - 1).$$

Hence, by the quadratic formula, we obtain (25).  $\square$

**Lemma 5.2.** *For Example 4.2 with  $\mu \in (0, 1)$ , given tangential boundary points  $b_j$  and  $b_\star$  with  $|b_j| < r$ ,  $|b_\star| = r$ ,  $\theta_j = \text{Arg}(b_j - z) \in (0, \frac{\pi}{2})$ , and  $\text{Arg}(b_\star) = 0$ , the optimal cut for corner  $c_j$  produces angle*

$$\theta_{j+1} = \text{Arg}(b_{j+1} - z) = -\theta_j + 2 \arctan \left( \frac{\tilde{r} \sin \theta_j - t_j \cos \theta_j}{\tilde{r} \cos \theta_j + t_j \sin \theta_j} \right). \quad (26)$$

where  $t_j > 0$  and is defined by (25).

*Proof.* Let  $\hat{\theta} = \theta_j - \text{Arg}(d_j - z)$ . Then it follows that

$$\theta_{j+1} = \theta_j - 2\hat{\theta} = -\theta_j + 2 \text{Arg}(d_j - z) = -\theta_j + 2 \arctan \left( \frac{\text{Im}(d_j - z)}{\text{Re}(d_j - z)} \right),$$

where the last equality follows because  $\text{Re}(d_j - z) > 0$ , as  $\text{Re} b_j > z$ . Using (24) and substituting in  $b_j = z + \tilde{r}e^{i\theta_j}$ , we have that

$$d_j - z = \tilde{r}e^{i\theta_j} - it_je^{i\theta_j} \quad \implies \quad \begin{aligned} \text{Re}(d_j - z) &= \tilde{r} \cos \theta_j + t_j \sin \theta_j, \\ \text{Im}(d_j - z) &= \tilde{r} \sin \theta_j - t_j \cos \theta_j, \end{aligned}$$

where  $t_j > 0$  is given by (25). Substituting these inside the arctan in the equation above completes the proof.  $\square$

Next, before we determine how fast  $\{\theta_k\}$  converges, we must first show that it indeed converges to zero.

**Lemma 5.3.** *For Example 4.2 with  $\mu \in (0, 1)$ , given tangential boundary points  $b_j$  and  $b_\star$  with  $|b_j| < r$ ,  $|b_\star| = r$ ,  $\theta_j = \text{Arg}(b_j - z) \in (0, \frac{\pi}{2})$ , and  $\text{Arg}(b_\star) = 0$ , the recursion (26) for optimal cuts is a sequence of angles  $\{\theta_k\}$  converging to zero.*

*Proof.* By construction, it is clear that  $\{\theta_k\}$  is monotone, i.e.,  $\theta_{k+1} < \theta_k$  for all  $k$ , and bounded below by zero, and so  $\{\theta_k\}$  converges to some limit  $L$ . By way of contradiction, assume that  $L > 0$  and so  $0 < L < \theta_j < \frac{\pi}{2}$ . Hence,

$$\lim_{k \rightarrow \infty} \theta_{k+1} = L = -L + 2 \arctan \left( \frac{\tilde{r} \sin L - t(L) \cos L}{\tilde{r} \cos L + t(L) \sin L} \right) \implies \tan L = \frac{\tilde{r} \sin L - t(L) \cos L}{\tilde{r} \cos L + t(L) \sin L}.$$

Multiplying both sides by  $\cos L(\tilde{r} \cos L + t(L) \sin L)$  yields

$$\tilde{r} \sin L + t(L) \frac{\sin^2 L}{\cos L} = \tilde{r} \sin L - t(L) \cos L \quad \implies \quad t(L)(\sin^2 L + \cos^2 L) = 0,$$

which implies  $t(L) = 0$ , which contradicts Lemma 5.2, and so  $L = 0$ .  $\square$

We now have the necessary pieces to derive the exact (linear) rate of convergence of the angles produced by optimal cuts on Example 4.2.

**Theorem 5.4.** *For Example 4.2 with  $\mu \in (0, 1)$ , given tangential boundary points  $b_j$  and  $b_\star$  with  $|b_j| < r$ ,  $|b_\star| = r$ ,  $\theta_j = \text{Arg}(b_j - z) \in (0, \frac{\pi}{2})$ , and  $\text{Arg}(b_\star) = 0$ , the sequence  $\{\theta_k\}$  produced by optimal cuts converges to zero  $Q$ -linearly with rate  $\frac{2(1-\sqrt{1-\mu})}{\mu} - 1$ .*

*Proof.* By Lemmas 5.2 and 5.3, (26) holds,  $\theta_k \rightarrow 0$ , and  $\theta_k \geq 0$  for all  $k$ , so

$$\lim_{k \rightarrow \infty} \frac{|\theta_{k+1} - \theta_*|}{|\theta_k - \theta_*|} = \lim_{k \rightarrow \infty} \frac{\theta_{k+1}}{\theta_k} = \lim_{k \rightarrow \infty} \frac{-\theta_k + 2 \arctan \left( \frac{\tilde{r} \sin \theta_k - t_k \cos \theta_k}{\tilde{r} \cos \theta_k + t_k \sin \theta_k} \right)}{\theta_k}.$$

Using the continuous version of  $t_k$

$$t(\theta) = -z \sin \theta + \sqrt{z^2 \sin^2 \theta - 2z\tilde{r}(\cos \theta - 1)},$$

we instead consider the entire limit in continuous form:

$$\lim_{\theta \rightarrow 0} \frac{-\theta + 2 \arctan \left( \frac{\tilde{r} \sin \theta - t(\theta) \cos \theta}{\tilde{r} \cos \theta + t(\theta) \sin \theta} \right)}{\theta} = -1 + \lim_{\theta \rightarrow 0} \frac{2}{\theta} \cdot \frac{\tilde{r}\theta - t(\theta) \cos \theta}{\tilde{r} \cos \theta + t(\theta)\theta}, \quad (27)$$

where the equality holds by using the small-angle approximations  $\arctan x \approx x$  (as the ratio inside the arctan above goes to zero as  $\theta \rightarrow 0$ ) and  $\sin x \approx x$ . Again using  $\sin x \approx x$  as well as the small-angle approximation  $1 - \cos x \approx \frac{x^2}{2}$ , we also have the small-angle approximation

$$t(\theta) \approx -z\theta + \sqrt{z^2\theta^2 - 2z\tilde{r}\left(\frac{-\theta^2}{2}\right)} = -\theta \left( z - \theta\sqrt{z^2 + z\tilde{r}} \right) = -\theta (z - \sqrt{zr}), \quad (28)$$

where the last equality holds by substituting in  $\tilde{r} = r - z$ . Via substituting in (28), the limit on the right-hand side of (27) is

$$\lim_{\theta \rightarrow 0} \frac{2}{\theta} \cdot \frac{\tilde{r}\theta + \theta(z - \sqrt{zr}) \cos \theta}{\tilde{r} \cos \theta - \theta(z - \sqrt{zr}) \theta} = \lim_{\theta \rightarrow 0} \frac{2(\tilde{r} + (z - \sqrt{zr}) \cos \theta)}{\tilde{r} \cos \theta - \theta^2(z - \sqrt{zr})} = \frac{2(\tilde{r} + z - \sqrt{zr})}{\tilde{r}}.$$

Recalling that  $\tilde{r} = \mu r$  and that  $z = r - \tilde{r} = r - \mu r$ , by substitutions we can rewrite the ratio above as

$$\frac{2 \left( \mu r + (r - \mu r) - \sqrt{(r - \mu r)r} \right)}{\mu r} = \frac{2(r - r\sqrt{1 - \mu})}{\mu r} = \frac{2(1 - \sqrt{1 - \mu})}{\mu}.$$

Subtracting one from the value above completes the proof.  $\square$

As we are about to see, there will be no need to derive the convergence rate of  $\{c_k\}$  for the corners  $c_k$  produced by the optimal cutting strategy (which would be an analogue of Theorem 4.4 for Uhlig's method). In Figure 3a, the convergence rates for the respective sequences of angles generated by Uhlig's procedure and the optimal cutting strategy (the rates given by Theorems 4.3 and 5.4) are plotted. The rate for Uhlig's method is worse for normalized curvatures up to  $\mu \approx 0.8283$ , after which the optimal cutting strategy's rate becomes higher as it increases to one.<sup>4</sup> In Figure 3b, via a simulation, we show plots of the total number of cuts needed by both approaches in order to sufficiently refine  $\partial W(A)$  for Example 4.2 near its outermost point. With this perspective, we see that the optimal cutting procedure becomes significantly better than Uhlig's method for normalized curvatures between  $\mu \approx 0.8384$  and  $\mu \approx 0.9610$ , with Uhlig's method requiring almost up to double the number of cuts. It is ironic that just as the rate of convergence of Uhlig's method becomes better than that of the optimal strategy, the optimal strategy nevertheless becomes significantly better in terms of total cost. For other normalized curvature values, the optimal strategy is slightly less costly than Uhlig's method and never more costly. However, as  $\mu \rightarrow 1$ , the total costs of both methods start to skyrocket, which is in line with our cost analysis for disk matrices given by (15).

<sup>4</sup>The values of  $\mu$  in this paragraph were determined by looking at zoomed in versions of the plots in Figure 3.

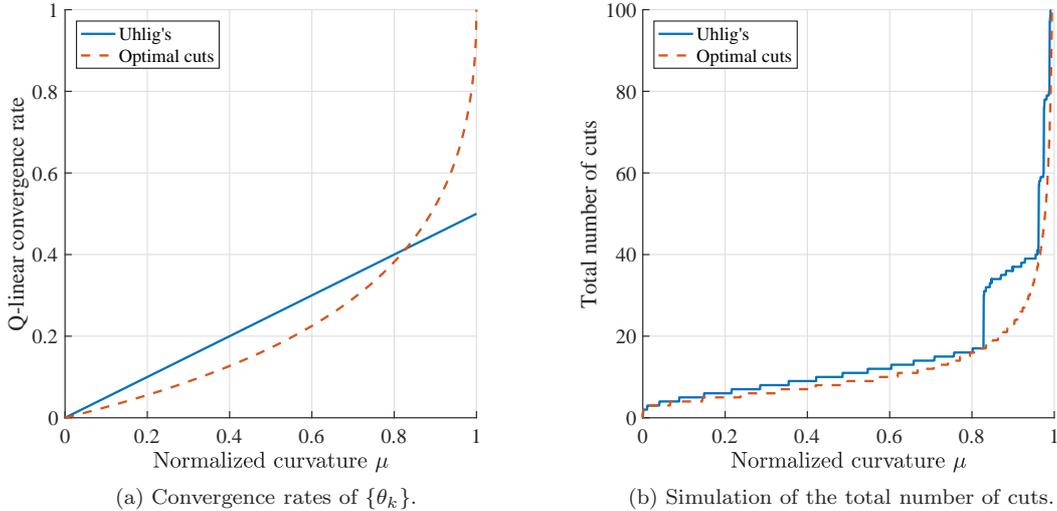


Figure 3: On the left, we plot the corresponding convergence rates (of the angles) for Uhlig’s procedure and optimal cuts given by Theorems 4.3 and 5.4, respectively. On the right, we plot the total number of cuts respectively required by Uhlig’s procedure and the optimal cutting strategy via simulating them on Example 4.2 (using analytic trigonometric descriptions rather than actually computing eigenvalues of  $H(\theta)$ ) in order to sufficiently refine the region of  $\partial W(A)$  between  $b_j$  and  $b_*$ , where  $\theta_j = \frac{\pi}{4}$  and  $\text{tol}=1\text{e-}14$ .

## 6 A hybrid algorithm

We now describe a hybrid algorithm, whose main ingredients are Algorithms 1 and 2; as such, we forgo providing pseudocode for this hybrid algorithm and just describe the basic operation here. Furthermore, we must assume that it is still feasible to run Algorithm 1 for the given dimension  $n$ . The key idea is that via two straightforward techniques, a) we can predict whether it would likely be more efficient to use Algorithm 1 over Algorithm 2 or vice-versa, and b) when the cutting-plane approach is chosen (Algorithm 2), we can also algorithmically determine on the fly when it would likely be better to switch back to our optimized level-set approach (Algorithm 1). We begin with the former.

With Algorithm 1, we know that all unimodular eigenvalues of  $R_\gamma - \lambda S$  (which recall is a  $2n \times 2n$  matrix pencil) must be computed for at least one value of  $\gamma$  and that this is the dominant cost in the level-set approach. However, given our new improvements to work with (10) instead of (9) and to use local optimization, it is reasonable to assume that Algorithm 1 may only require a single eigenvalue computation with  $R_\gamma - \lambda S$ , namely when a global maximizer of (10) is found in the very first iteration. Meanwhile, Algorithm 2 only requires extremal eigenvalue computations with  $H(\theta)$  (which recall is a  $n \times n$  Hermitian matrix, possibly sparse), and data like that shown in Figure 3b provides good estimates of the total number of cuts that will be needed to compute  $r(A)$  based on the normalized curvatures at outermost points in  $W(A)$ . By doing benchmarking as we have done in Table 1, we can construct look-up tables to determine for a given dimension how many eigenvalue computations we can do with  $H(\theta)$  for the price of a single eigenvalue computation with  $R_\gamma - \lambda S$ . As such, we can predict which method will be faster. For example, if  $\mu$  is, say, less than 0.83, then very few cuts will be needed by Algorithm 2 and so it is likely to be faster than Algorithm 1 for all but the smallest values of  $n$ . Furthermore, when  $n \geq 800$  (per Table 1), Algorithm 2 can afford to do about 100 cuts while still being faster than Algorithm 1 and so our cutting-plane approach will be faster for almost any normalized curvature value  $\mu$  not too close to one (where the number of required cuts blows up). However, since we do not know

the normalized curvature at an outermost point in  $W(A)$  *a priori*, this motivates our second technique of how to dynamically decide when to switch back to Algorithm 1.

Suppose that our hybrid method chooses to start the computation with Algorithm 2, based on using the look-up table data, the dimension of the problem, whether the matrix is sparse,  $W(A)$  has symmetry, the desired level of accuracy, etc. If the normalized curvature where  $r(A)$  is attained is one or very close to it, Algorithm 2 will become extremely expensive to use when computing  $r(A)$  to more than just a few digits. To avoid this blow-up in cost, one could just set a small fixed limit on the total number of cuts that can be performed before halting Algorithm 2 and switching to Algorithm 1. When switching, Algorithm 1 will be warm started by using the initial guess  $\theta_0 = -\text{Arg}(b)$ , where  $b$  is the outermost tangential boundary point in the polygonal approximation to  $W(A)$  built by Algorithm 2. If  $b$  is a globally outermost point, then Algorithm 1 will immediately verify this via a single eigenvalue computation with  $R_\gamma - \lambda S$  and terminate. However, our optimal cutting strategy provides a more sophisticated yet still simple way to determine when to switch (if at all). Once optimal cuts are applicable, via (23) we also have a good estimate for the actual normalized curvature at the corresponding locally outermost point in  $W(A)$ . Hence, if  $\mu_{\text{est}}$  is too close to one, we can immediately switch to Algorithm 1. Furthermore, inside Algorithm 2 the current value of  $\mu_{\text{est}}$  can be used to estimate how many more cuts will be needed before convergence is obtained. This prediction can then be compared with the look-up table data to dynamically set the limit on the total number of cuts allowed before switching.

This third hybrid algorithm of course requires a one-time offline tuning procedure, since the look-up table values are likely to be hardware (and software) dependent. Nevertheless, for applications where the numerical radius is being optimized, our hybrid algorithm can be particularly advantageous. In this setting, the numerical radius generally will be computed many many times and the shape of the field of values will not only be changing but very possibly converging to a disk centered at the origin; regarding this phenomenon, again see [LO18]. On such problems, our hybrid method can automatically get the best of both algorithms, while avoiding the worst of both. Initially, it will likely only use Algorithm 2, which can be substantially faster, and if the field of values becomes too close to being a disk, it will start automatically falling back to Algorithm 1, thus avoiding the cost blow-ups that can occur for cutting-plane approaches when  $\mu$  is too close to one.

## 7 Numerical validation

We implemented Algorithms 1 and 2 in MATLAB and performed the following three key experiments, to first validate our new methods against earlier methods, and then finally, to compare Algorithms 1 and 2 to each other. As implementing and evaluating our third hybrid algorithm requires running tuning procedures, we leave this to future work. However, implementations of all three of our new methods will be added to a future release of ROSTAPACK: RObust STA-bility PACKage [Mit], an open-source library implemented in MATLAB and licensed under the AGPL.

### 7.1 Validating Algorithm 1

We begin by comparing Algorithm 1 to the `numr` routine, which is Mengi and Overton’s MATLAB implementation of their earlier level-set-based method. To do this, we generated six dense complex random matrices of varying dimensions and recorded the costs of running `numr` and Algorithm 1 on these problems. Furthermore, we ran our implementation of Algorithm 1 in two configurations: with the local-optimization enhancement enabled and then again without it. Hence, this compares Algorithm 1 as stated in the pseudocode against a BBBS-like iteration on (10) (local optimization now disabled) and against a BBBS-like iteration on (9) (`numr`). The problems and results are given in Table 2. As can be seen, even our BBBS-like iteration on (10) is always faster than `numr`, up to 2.3 times faster (for  $n = 500$ ), as it always reduces the total

$n$	# of eig( $H(\theta)$ )			# of eig( $R_\gamma, S$ )			Time (sec.)		
	Alg. 1		MO	Alg. 1		MO	Alg. 1		MO
	Opt.	MP		Opt.	MP		Opt.	MP	
100	7	5	34	1	4	5	0.1	0.2	0.3
200	9	7	38	2	4	6	0.8	1.4	2.3
300	16	5	26	1	4	5	1.4	4.4	6.0
400	9	10	43	1	5	7	3.0	13.9	19.6
500	5	4	54	1	3	7	4.2	12.4	28.6
600	6	5	36	1	4	5	7.9	28.5	37.2

Table 2: On six dense examples, generated via  $A = \text{randn}(n) + 1i \cdot \text{randn}(n)$ , the costs of Algorithm 1 (Alg. 1) and the earlier BBBS-like method of Mengi and Overton (MO) are shown. We used our own implementation of Algorithm 1, with  $\text{tol} = 1e-14$ , and tested it both as described in §3 (Opt.) and then again with our local optimization disabled but still using a BBBS-iteration on (10) (MP, for midpoints only). We used Mengi and Overton’s implementation of their BBBS-like algorithm on (9), `numr`, which does not have a user-settable termination tolerance. We report the total number of eigenvalue computations with  $H(\theta)$  and  $R_\gamma - \lambda S$ , as well as the elapsed wall-clock running times in seconds (which is an average of three trials to account for variability). The computed values for  $r(A)$  all agreed to at least 14 digits on all six examples.

number of expensive eigenvalue computations with  $R_\gamma - \lambda S$  by a noticeable amount. Reenabling local optimization, i.e., Algorithm 1 as stated, leads to even better speedups, where now only one eigenvalue computation with  $R_\gamma - \lambda S$  is typically needed (with  $n = 200$  being the only exception, as it required two computations with  $R_\gamma - \lambda S$ ). In terms of improvement in running times, the speedup results are commensurate: Algorithm 1 ranged from 2.9 times faster ( $n = 200$ ) to 6.8 times faster ( $n = 500$ ) than `numr`.

## 7.2 Validating Algorithm 2

The simulation illustrated in Figure 3b indicates that our optimal-cutting strategy should be appreciably faster than Uhlig’s method for normalized curvature values  $\mu$  approximately between 0.8384 and 0.9610. However, this simulation was done using a trigonometric description of Example 4.2, where we have an exact formula for computing the optimal cut. As outlined in §5.2, in practice the optimal cut must be estimated by sufficiently approximating  $\partial W(A)$  at locally outermost points in  $W(A)$ . Hence, we now verify that our new procedure does in fact work by testing Algorithm 2 on actual instances of Example 4.2 for different normalized curvature values. For comparison purposes, we also tested Uhlig’s method on these problems, which we did by simply disabling the local-optimization and optimal-cutting enhancements in our implementation of Algorithm 2. The exact setup and results are given in Figure 4. In Figure 4a, where the instances have real-axis symmetry, we see that the actual total number of cuts for both methods are nearly identical to the simulation plots in Figure 3b, thus validating that our optimal-cutting strategy indeed works, with a near perfect correspondence with the simulation. When we rotate the problem instances, so that the symmetry in the field of values is no longer apparent to the codes, we see in Figure 4b that the total costs about double, as should be expected.

## 7.3 Comparing Algorithms 1 and 2

Finally, we compare Algorithms 1 and 2 against each other, in their intended configurations, i.e., as stated in their respective pseudocodes, without anything disabled. For simplicity, we chose to use `eig` for all computations with  $H(\theta)$ , even for Algorithm 2, because we observed that while `eigs` often works well on  $H(\theta)$ , it was not always reliable for computing the initial eigenvalues of  $A$ . Interestingly, Uhlig’s own MATLAB implementation of his method, `NumRadius`, also always

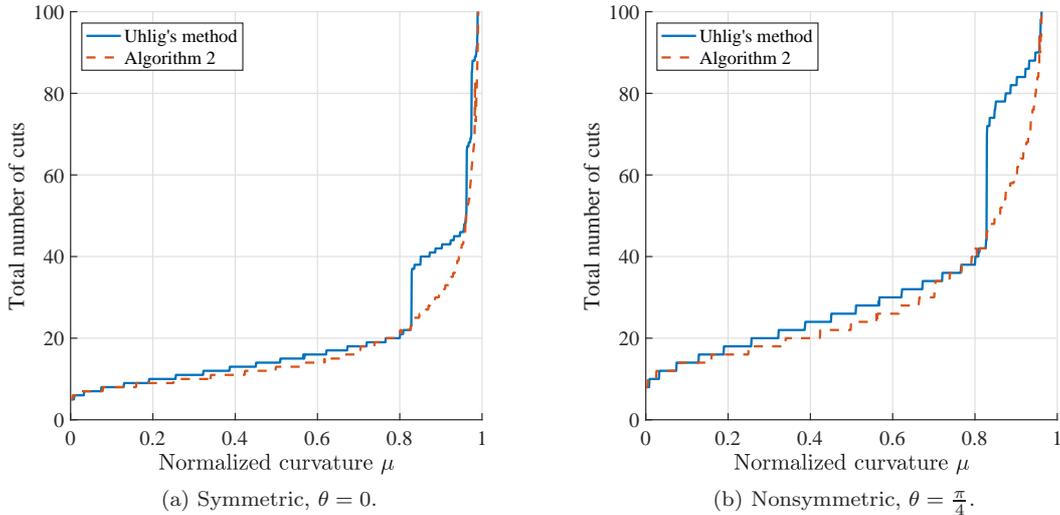


Figure 4: Total number of cuts needed by our own implementations of Uhlig’s method and Algorithm 2 to compute the numerical radius of instances of Example 4.2, where  $A = e^{i\theta}(1 - \mu)I + \mu K_5$  for  $\mu = 0.001, 0.002, \dots, 0.999$ . For this particular example,  $r(A) = 1$  always holds and  $\mu$  is the normalized curvature at the point in  $W(A)$  attaining  $r(A)$ . For the left plot,  $W(A)$  is a disk of radius  $\mu$  centered at  $(1 - \mu)$  and since  $A$  is real, the code automatically detects that only supporting hyperplanes with angles  $\bar{\theta} \in [0, \pi]$  are needed. In the right plot, the total number of cuts is about double as now this example has been rotated into the upper right quadrant. While our code can also detect arbitrary scalar rotations of otherwise symmetric problems, we disabled this feature for the purpose of this experiment, in order to do total cost comparisons for both the symmetric and nonsymmetric case. For this experiment, we set `tol` =  $1e-14$  and only used supporting hyperplanes corresponding to  $\lambda_{\max}$  of  $H(\theta)$  (as opposed to also using those for  $\lambda_{\min}$ , per Remark 3).

uses `eig` for the initial computation of eigenvalues of  $A$ , presumably for a similar reason. For large and sparse matrices, Uhlig’s code completely forgoes computing any eigenvalues of  $A$  and skips checking whether or not  $A$  is normal (as it would be cubic work). Skipping the normality check has little ill effect since if  $A$  is normal,  $W(A)$  will be the convex hull of the eigenvalues of  $A$ , and as we have shown, Uhlig’s method always has superlinear convergence when all outermost points of  $W(A)$  are corners. Skipping the computation of eigenvalues of  $A$  means that the rotation of the initial polygonal approximation and the guess for where  $r(A)$  is attained are likely to be uninformed guesses, but the local optimization used in Algorithm 2 would likely offset this.

Following Uhlig’s experiments in [Uhl09, Section 3], we compare our algorithms on some of the same problems, also with varying dimensions. The first three examples are ones Uhlig used and can be obtained via `gallery` in MATLAB: Gear (`gallery('gearmat', n)`), Grcar (`gallery('grcar', n)`), and FM (`gallery('fiedler', n) + 1i*gallery('moler', n)`). Gear and Grcar have few nonzero entries, respectively  $2n$  and  $5n - 7$ , while FM is completely dense. We used `randn(n) + 1i*randn(n)` as an additional dense example, which we call `randn`. Lastly, we also included an example which is nearly a disk matrix, namely the nonsymmetric example used in Figure 4b, but for  $K_{320}$ ,  $K_{640}$ , and  $K_{1280}$  and with  $\mu = 0.999$ ; we refer to this example as Nearly Disk.

In Table 3, we give the exact setup and results of our comparison. As can be seen, for all but the Nearly Disk example, the gap between Algorithm 1 and Algorithm 2 widens dramatically as  $n$  increases, again highlighting the high cost of computing all unimodular eigenvalues of  $R_\gamma - \lambda S$ . Furthermore, as Algorithm 1 only ever required one to two such computations, it is likely that

Problem	$n$	# of calls to <code>eig(·)</code>			Time (sec.)	
		Alg. 1		Alg. 2	Alg. 1	Alg. 2
		$H(\theta)$	$R_\gamma - \lambda S$	$H(\theta)$		
Gear	320	1	1	5	1.4	0.1
Gear	640	2	1	3	21.4	0.5
Gear	1280	2	1	4	343.2	3.1
Grcar	320	45	2	30	3.0	0.6
Grcar	640	82	2	31	38.2	3.5
Grcar	1280	161	2	30	576.4	22.8
FM	320	9	1	20	1.1	0.3
FM	640	6	1	18	10.0	1.0
FM	1280	9	1	20	88.0	7.2
<code>randn</code>	320	5	1	41	1.5	0.8
<code>randn</code>	640	17	2	73	26.7	7.6
<code>randn</code>	1280	18	2	61	232.5	46.6
Nearly Disk	320	2	1	511	2.0	10.6
Nearly Disk	640	2	1	511	17.2	58.5
Nearly Disk	1280	2	1	531	167.7	429.1

Table 3: On five different examples, each of three different sizes, the respective costs of Algorithms 1 and 2 are shown in terms of the total number of eigenvalue computations incurred and elapsed wall-clock running times in seconds (again as an average of three trials). Both algorithms used `tol = 1e-14` and the respective values computed for the numerical radius always agreed to at least 14 digits over all problems. For these experiments, Algorithm 2 was set to use `eig` instead of `eigs` and to add the supporting hyperplanes for both  $\lambda_{\max}$  and  $\lambda_{\min}$  of  $H(\theta)$  on every cut (again per Remark 3).

Mengi and Overton’s earlier method would have fared significantly worse in this comparison. However, even with our enhanced level-set approach, our improved cutting-plane iteration is nevertheless always fastest for these four problems, ranging from about 2 to 111 times faster than Algorithm 1, with the biggest gap for Gear with  $n = 1280$ . If we had used `eigs` instead of `eig` to do cuts, these performance gaps would likely be even wider.

The reason Algorithm 2 is fastest on Gear, Grcar, FM, and `randn` is because none have a field of values that is nearly a disk centered at the origin. The Gear matrix is particularly easy for the cutting-plane approach, as its field of values not only has real-axis symmetry but  $r(A)$  is attained on the real axis at a corner of  $\partial W(A)$ . By our convergence rate results, we hence know that the cutting-plane approach should have superlinear local convergence for this example, and this is indeed reflected in Table 3, where at most just five cuts were sufficient for any size of Gear tested. The respective field of values for Grcar, FM, and `randn` do not have corners, so the total number of cuts for Algorithm 2 increased, from 18 to 73 for these three examples, due to the slower linear local convergence of the cutting process at outermost points that are not corners of  $\partial W(A)$ . However, these totals all remain relatively small because the normalized curvatures at their respective outermost points are all sufficiently smaller than one; per our convergence rate analyses using Example 4.2, this means that the linear rate of convergence of the cutting process is actually still quite fast.

On the other hand, on Nearly Disk, the total number of cuts Algorithm 2 needed to compute  $r(A)$  severely increased, ranging from 511 to 531 cuts, which is inline with our earlier analysis from (15). Indeed, this blow-up in the number of cuts needed to compute  $r(A)$  completely turns things around, with our level-set approach now being 2.6 to 5.3 times faster than Algorithm 2 on Nearly Disk. This underscores the need for our third hybrid algorithm, which could have automatically shifted back to Algorithm 1 to avoid incurring so many cuts and thus been faster than Algorithm 2 alone. Finally, in contrast to Algorithm 2, note that Nearly Disk is actually a

relatively easy problem for Algorithm 1, as it only needs a single computation with  $R_\gamma - \lambda$  and just two with  $H(\theta)$ . This is not terribly surprising as for disk matrices, or ones that are nearly so, (10) will be nearly a constant function, i.e., an easy maximization problem.

## 8 Conclusion

In this paper, we have given the first formal analysis of the local convergence rate of Uhlig’s method for computing the numerical radius, while also introducing three improved algorithms. Our optimization-enhanced level-set approach is generally several times faster than the earlier method of Mengi and Overton. Meanwhile, our improved cutting-plane method eliminates inefficiencies in Uhlig’s cutting-plane method for problematic normalized curvature values, as predicted by our additional convergence rate analysis and simulation of this new approach. Furthermore, while cutting-plane approaches are generally faster than level-set-based ones, we have shown that the reverse becomes true, in a dramatic fashion, when computing the numerical radius of matrices whose fields of values are a circular disk centered at the origin, or nearly so. This effect motivated our third hybrid method, which can dynamically detect when to switch between our level-set and cutting-plane approaches in order to remain efficient across all numerical radii problems.

## References

- [BB90] S. Boyd and V. Balakrishnan. A regularity result for the singular values of a transfer matrix and a quadratically convergent algorithm for computing its  $L_\infty$ -norm. *Syst. Cont. Lett.*, 15:1–7, 1990.
- [BBMX02] P. Benner, R. Byers, V. Mehrmann, and H. Xu. Numerical computation of deflating subspaces of skew-Hamiltonian/Hamiltonian pencils. *SIAM J. Matrix Anal. Appl.*, 24(1):165–190, 2002.
- [Ben02] I. Bendixson. Sur les racines d’une équation fondamentale. *Acta Math.*, 25(1):359–365, 1902.
- [Ber65] C. A. Berger. A strange dilation theorem. *Notices Amer. Math. Soc.*, 12:590, 1965.
- [BLO03] J. V. Burke, A. S. Lewis, and M. L. Overton. Robust stability and a criss-cross algorithm for pseudospectra. *IMA J. Numer. Anal.*, 23(3):359–375, 2003.
- [BM18] P. Benner and T. Mitchell. Faster and more accurate computation of the  $\mathcal{H}_\infty$  norm via optimization. *SIAM J. Sci. Comput.*, 40(5):A3609–A3635, October 2018.
- [BM19] P. Benner and T. Mitchell. Extended and improved criss-cross algorithms for computing the spectral value set abscissa and radius. *SIAM J. Matrix Anal. Appl.*, 40(4):1325–1352, 2019.
- [BS90] N. A. Bruinsma and M. Steinbuch. A fast algorithm to compute the  $H_\infty$ -norm of a transfer function matrix. *Syst. Cont. Lett.*, 14(4):287–293, 1990.
- [BSV16] P. Benner, V. Sima, and M. Voigt. Algorithm 961: Fortran 77 subroutines for the solution of skew-Hamiltonian/Hamiltonian eigenproblems. *ACM Trans. Math. Software*, 42(3):Art. 24, 26, 2016.
- [Bye88] R. Byers. A bisection method for measuring the distance of a stable matrix to unstable matrices. *SIAM J. Sci. Statist. Comput.*, 9:875–881, 1988.
- [DHT14] T. A Driscoll, N. Hale, and L. N. Trefethen. *Chebfun Guide*. Pafnuty Publications, 2014. <http://www.chebfun.org/docs/guide/>.

- [HJ91] R. A. Horn and C. R. Johnson. *Topics in Matrix Analysis*. Cambridge University Press, Cambridge, 1991.
- [HW97] C. He and G. A. Watson. An algorithm for computing the numerical radius. *IMA Journal of Numerical Analysis*, 17(3):329–342, June 1997.
- [Joh78] C. R. Johnson. Numerical determination of the field of values of a general complex matrix. *SIAM J. Numer. Anal.*, 15(3):595–602, 1978.
- [Kat82] T. Kato. *A Short Introduction to Perturbation Theory for Linear Operators*. Springer-Verlag, New York - Berlin, 1982.
- [Kip51] R. Kippenhahn. Über den Wertevorrat einer Matrix. *Math. Nachr.*, 6(3-4):193–228, 1951.
- [Kre62] H.-O. Kreiss. Über die Stabilitätsdefinition für Differenzgleichungen die partielle Differentialgleichungen approximieren. *BIT Numerical Mathematics*, 2(3):153–181, 1962.
- [Lan64] P. Lancaster. On eigenvalues of matrices dependent on a parameter. *Numer. Math.*, 6:377–387, 1964.
- [LO18] A. S. Lewis and M. L. Overton. Partial smoothness of the numerical radius at matrices whose fields of values are disks. e-print 1901.00050, arXiv, December 2018. math.NA.
- [Mit] T. Mitchell. ROSTAPACK: RObust STAbility PACKage. <http://timitchell.com/software/ROSTAPACK>.
- [Mit19a] T. Mitchell. Computing the Kreiss constant of a matrix. e-print arXiv:1907.06537, arXiv, July 2019. math.OC.
- [Mit19b] T. Mitchell. Fast interpolation-based globality certificates for computing Kreiss constants and the distance to uncontrollability. e-print arXiv:1910.01069, arXiv, October 2019. math.OC.
- [MO05] E. Mengi and M. L. Overton. Algorithms for the computation of the pseudospectral radius and the numerical radius of a matrix. *IMA J. Numer. Anal.*, 25(4):648–669, 2005.
- [OW95] M. L. Overton and R. S. Womersley. Second derivatives for optimizing eigenvalues of symmetric matrices. *SIAM J. Matrix Anal. Appl.*, 16(3):697–718, 1995.
- [Pea66] C. Pearcy. An elementary proof of the power inequality for the numerical radius. *Michigan Math. J.*, 13:289–291, 1966.
- [TE05] L. N. Trefethen and M. Embree. *Spectra and pseudospectra: The behavior of non-normal matrices and operators*. Princeton University Press, Princeton, NJ, 2005.
- [TY99] B.-S. Tam and S. Yang. On matrices whose numerical ranges have circular or weak circular symmetry. *Linear Algebra Appl.*, 302–303:193–221, 1999. Special issue dedicated to Hans Schneider (Madison, WI, 1998).
- [Uhl09] F. Uhlig. Geometric computation of the numerical radius of a matrix. *Numer. Algorithms*, 52(3):335–353, 2009.
- [Wat96] G. A. Watson. Computing the numerical radius. *Linear Algebra Appl.*, 234:163–172, 1996.