

## OPTIMAL DOMAIN SPLITTING FOR INTERPOLATION BY CHEBYSHEV POLYNOMIALS\*

TOBIN A. DRISCOLL<sup>†</sup> AND J. A. C. WEIDEMAN<sup>‡</sup>

**Abstract.** Polynomial interpolants defined using Chebyshev extreme points as nodes converge uniformly at a geometric rate when sampling a function that is analytic on an interval. However, the convergence rate can be arbitrarily close to unity if the function has a singularity close to the interval when extended to the complex plane. In such cases, splitting the interval and doing piecewise interpolation may be more efficient in the total number of nodes than the global interpolant. Because the convergence rate is determined by Bernstein ellipses obtained through a Joukowski conformal map, relative efficiency of splitting at any point in the interval can be calculated and then optimized over the interval. The optimal splitting may be applied recursively. The Chebfun software project uses a simple rule of thumb without prior singularity information to create a binary search that can be shown to do an excellent job of finding the optimal splitting in most cases. However, the process can use a large number of intermediate function evaluations that are not needed in the final approximant. A Chebyshev–Padé technique generates approximate singularity locations that are good enough to get close to the optimal splitting more directly in test cases. The technique is applied to a singularly perturbed boundary-value problem.

**Key words.** Chebyshev polynomial interpolation, domain decomposition, potential theory, spectral collocation, Chebfun

**AMS subject classifications.** 65D05, 65L50

**DOI.** 10.1137/130919428

**1. Introduction.** Approximation of a function by polynomial interpolation at Chebyshev points is a classical idea. It is a powerful computational tool as well, thanks to significant algorithmic and theoretical development [17] and the Chebfun software system [3, 6]. At the heart of its usefulness is exponential convergence: for a polynomial interpolant of  $f$  on  $N$  Chebyshev points in the standard interval  $[-1, 1]$ , the error can be expected to decay at a rate  $O(\alpha^{-N})$  for a constant  $\alpha > 1$  depending on the analyticity of  $f$  as extended into a complex region containing the interval. This key fact is reviewed in more detail in section 2. Because of this exponential convergence, Chebfun can obtain polynomial approximations accurate to machine double precision for smooth functions on any modern computer. For functions that lack analyticity on the interval  $[-1, 1]$  itself, however, the convergence rate is reduced from exponential to algebraic, in a manner analogous to the Gibbs phenomenon (and its generalization to jumps in a derivative) for trigonometric approximations [17]. Global polynomial approximations are often infeasible in such cases. In response, Chebfun includes a facility for detecting singularities on the interval and subdividing it accordingly to create piecewise approximations [14].

Functions like those in Figure 1 qualify as borderline cases. They are technically analytic in an appropriate region containing the interval, but the region is so small that the exponential convergence constant  $\alpha$  is very close to 1, and a large value of

---

\*Received by the editors May 2, 2013; accepted for publication (in revised form) May 22, 2014; published electronically August 7, 2014.

<http://www.siam.org/journals/sinum/52-4/91942.html>

<sup>†</sup>Department of Mathematical Sciences, University of Delaware, Newark DE 19176 (driscoll@udel.edu).

<sup>‡</sup>Department of Mathematical Sciences, University of Stellenbosch, Matieland 7602, South Africa (weideman@sun.ac.za). This author’s research was supported by the National Research Foundation of South Africa.

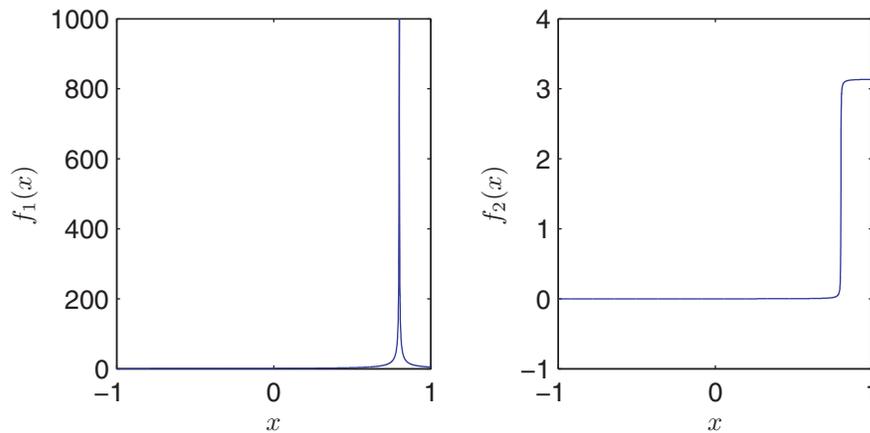


FIG. 1. Functions with singularities very close to the approximation interval:  $f_1(x) = [(x - \xi)(x - \bar{\xi})]^{-1/2}$  and  $f_2(x) = \text{Im}[\log(1 - x/\xi)]$ , where  $\xi = 0.8 + 0.001i$ . From the standpoint of polynomial approximation, both behave similarly and require far more nodes in a global approximation than in a piecewise representation.

$N$  is therefore required to get an accurate approximation. Chebfun offers a *splitting* mode that creates piecewise approximations in this situation as well. In this mode, if a polynomial approximation on the interval reaches a predetermined maximum degree while still leaving too large an error (as measured by Chebyshev series coefficients), the interval is bisected. The same criterion is applied recursively to the two half-intervals. Once a set of subintervals and accurate piecewise polynomial approximations have been found, adjacent representations are merged until the total representation length no longer decreases.

This divide-and-conquer process is robust, uses only information available from the point values of  $f$ , and results in compact representations. In fact, in section 4 we will give evidence that the final representations are essentially as compact as possible. However, while the recursive bisection process probes for the correct subintervals, it may evaluate  $f$  at numerous different sets of nodes and require many more point values of  $f$  than the final representation needs. For example, the functions in Figure 1 both end up with interval breakpoints at  $3/4$ ,  $51/64$ , and  $103/128$  and have final piecewise representations using totals of 380 and 363 point values, respectively. This compares very favorably with the 16381 and 16139 point values required of their global polynomial approximations. However, arriving at the piecewise representations required 1903 and 1893 total unique point values. Such excess is inconsequential when  $f$  is not expensive to evaluate and when the values of  $f$  on different subintervals are independent of one another. When, however,  $f$  is the solution of a differential equation, its values on all subintervals are globally linked, and the effort to evaluate  $f$  at  $N$  nodes costs nominally  $O(N^3)$  work. In such circumstances there is a strong incentive to minimize the number of different point sets at which  $f$  is evaluated, and to make the sets as small as possible.

In section 2 we show that the standard potential theory description of polynomial interpolation convergence is easily extended to the case of splitting the approximation interval at an arbitrary location. The resulting framework allows the computation of an optimal splitting location for any given rate-determining singularity location in the complex plane, and the resulting improvement in the effective convergence rate. One can then identify a region in which such singularities must lie for splitting to be bene-

ficial. In section 3, the behavior of optimal split location and rate is examined asymptotically as the complex singularity location approaches the interpolation interval.

While Chebfun does not attempt to optimize splitting locations, recursive bisection can be expected to eventually choose a location close to the optimal one. In section 4 we verify experimentally that Chebfun's algorithm does result in a convergence rate very close to recursive application of the optimal split. Conversely, recursive optimal splitting can reproduce Chebfun's compact representations without first creating a possibly much larger intermediate tree of splits.

The disadvantage of optimal splitting in practice is the need to know the location of the rate-determining singularity. This problem has appeared in the literature many times for nonperiodic and periodic functions (see [2, 5, 7, 8, 12]). In section 5 we give some experiments which indicate that low-degree rational expansions do an adequate job at this task in a few experimental cases.

**2. Convergence theory and the effect of splitting.** We limit the discussion here to approximation of a function on the standard interval  $[-1, 1]$ ; approximation on other intervals follows easily via affine mapping. A thorough review of the convergence rate of polynomial interpolation is beyond the scope of this paper; see [17] for details. In this paper we will be concerned only with the case of approximating a function  $f$  that is analytic when extended into a complex region containing  $[-1, 1]$ . The key theorem we will exploit, which is a specialized form of Theorem 5 in [16], is as follows.

**THEOREM 2.1.** *Suppose  $f$  is analytic on and inside the ellipse with foci  $\pm 1$  and semimajor axis  $a > 1$ . Let  $p_n$  be the polynomial of degree  $\leq n$  that interpolates  $f$  at the  $n + 1$  second-kind Chebyshev points (extrema of  $T_n$ ) in  $[-1, 1]$ . Then there exists a constant  $C > 0$  such that for all  $n > 0$  and all  $x \in [-1, 1]$ ,*

$$(2.1) \quad |f(x) - p_n(x)| \leq C\alpha^{-n},$$

where  $\alpha = \phi(a) := a + \sqrt{a^2 - 1} > 1$ .

Following [17], we refer to the ellipse mentioned in the theorem as a *Bernstein ellipse*. The source of the relationship between  $a$  and  $\alpha$  is easily described. The Joukowski transformation  $\psi(z) = (z + z^{-1})/2$  maps the exterior of the unit circle in the complex  $z$ -plane conformally to the complement of the interval  $[-1, 1]$  in the complex  $x$ -plane. It also maps the circle of radius  $\alpha$  to the Bernstein ellipse described by the theorem, and in fact,  $\alpha = \psi^{-1}(a) = \phi(a)$ . For reference, we note that if  $\phi$  is expressed as

$$(2.2) \quad \phi(x) = x + x\sqrt{1 - \frac{1}{x^2}},$$

where the usual computational (principal) branch of the square root function is used, then  $\phi$  is continuous in the complement of the real interval  $(-1, 1)$ . This proves to be the most convenient branch for our computational work. We also note that by symmetry, we are free to pay attention only to singularities of  $f$  in the first quadrant of the  $x$ -plane; a function with a pole only at  $\xi$  in that quadrant behaves no differently, from the standpoint of Theorem 2.1, from one that also has poles at  $-\xi$ ,  $\bar{\xi}$ , and  $-\bar{\xi}$ .

Suppose that  $f$  is analytic everywhere in the complex plane except at the point  $\xi$  and its symmetric counterparts. Theorem 2.1 implies that polynomial interpolants to  $f$  at Chebyshev points converge geometrically at a rate  $\alpha(\xi) = |\phi(\xi)|$ . Now suppose

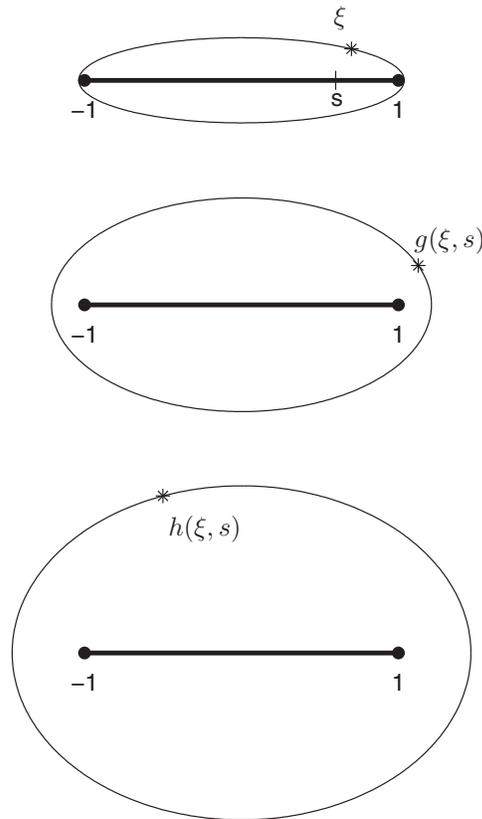


FIG. 2. *Effect of splitting on the convergence rate. Top: Original interval, showing singularity location  $\xi$ , Bernstein ellipse as in Theorem 2.1, and splitting location  $s$ . Bottom: Mapping the two subintervals defined by  $s$  moves the singularity away from the interval and creates larger Bernstein ellipses, thereby improving the convergence rate of polynomial interpolation.*

we split the interval  $[-1, 1]$  at the location  $x = s$  and consider separate polynomial interpolants on  $[-1, s]$  and  $[s, 1]$ . To understand the convergence of these interpolation subproblems, we have to remap each of the two subintervals back to  $[-1, 1]$  and apply Theorem 2.1 to the images of  $\xi$  under those transformations, as illustrated in Figure 2. Specifically, we define

$$(2.3) \quad g(x, s) = 2\frac{x+1}{s+1} - 1, \quad h(x, s) = 2\frac{x-s}{1-s} - 1$$

to map the left and right subintervals, respectively, and compute

$$(2.4) \quad \alpha_L(s) = \left| \phi(g(\xi, s)) \right| \quad \text{and} \quad \alpha_R(s) = \left| \phi(h(\xi, s)) \right|$$

as rates of convergence on the two subintervals. For convenience we have suppressed the dependence on  $\xi$ , which remains fixed for the discussion of this section.

However, if we want to compare the subinterval rates to the original rate  $|\phi(\xi)|$  on the unsplit interval, it is important to compare based on equal total numbers of interpolation nodes in both cases. If we ignore the role of the leading constant  $C$

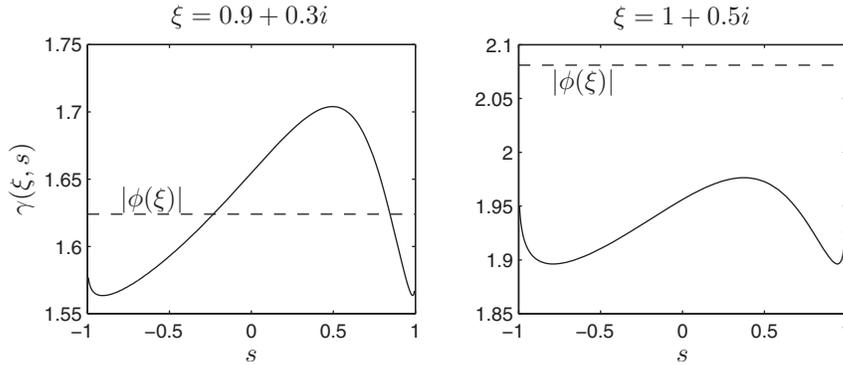


FIG. 3. Effective convergence rate after splitting, as a function of split location. Left: For a singularity  $\xi$  close enough to the interval  $[-1, 1]$ , some splitting locations improve the convergence rate compared to the unsplit rate  $|\phi(\xi)|$ . Right: If the singularity is too far from the interval, splitting never produces a faster rate.

in (2.1) (which can be different for the unsplit interval and the subintervals), then to achieve an accuracy  $\epsilon$  using the global polynomial approximation, we should expect to need about  $-\log(\epsilon)/\log \alpha$  nodes. Similarly, a piecewise approximation on the split interval requires

$$-\log(\epsilon) \left( \frac{1}{\log \alpha_L} + \frac{1}{\log \alpha_R} \right)$$

total nodes. By comparison, we define the effective splitting convergence rate  $\gamma(\xi, s)$  by

$$(2.5) \quad \gamma(\xi, s) = \exp \left[ \frac{\log \alpha_L \log \alpha_R}{\log(\alpha_L \alpha_R)} \right].$$

That is, given the best allocation of  $N$  nodes between  $[-1, s]$  and  $[s, 1]$ , we can expect geometric convergence of  $\gamma^{-N}$  in place of  $\alpha^{-N}$  in (2.1). Figure 3 shows  $\gamma(\xi, s)$  as a function of  $s$  for two singularity locations  $\xi$ . As can be seen in the figure, a singularity  $\xi$  close enough (in the Bernstein ellipse sense) to the interval  $[-1, 1]$  will lead to  $\gamma(\xi, s) > \alpha(\xi)$  for some values of the split location  $s$ , and splitting would be advantageous. For sufficiently distant singularities, however, splitting always leads to an inferior convergence rate and is not advised.

The optimal split location, say,  $s = \sigma(\xi)$ , is defined by the condition that the two convergence rates (2.4) are equal, i.e.,

$$(2.6) \quad \alpha_L(\xi, \sigma(\xi)) = \alpha_R(\xi, \sigma(\xi)).$$

It can be shown that this is equivalent to maximizing the effective convergence rate (2.5). Equation (2.6) can be solved explicitly for  $\sigma(\xi)$ , as shown in Appendix A. With  $\xi = x + iy$  in the first quadrant, one finds that

$$(2.7) \quad \sigma(\xi) = x - 2^{-1/2} \sqrt{(x^2 - y^2 - 1) + \sqrt{(1 + y^2 - x^2)^2 + 4x^2y^2}}.$$

Figure 4 shows the optimal splitting location over the first quadrant of the complex plane. Globally,  $\sigma(\xi) < \text{Re}(\xi)$ .

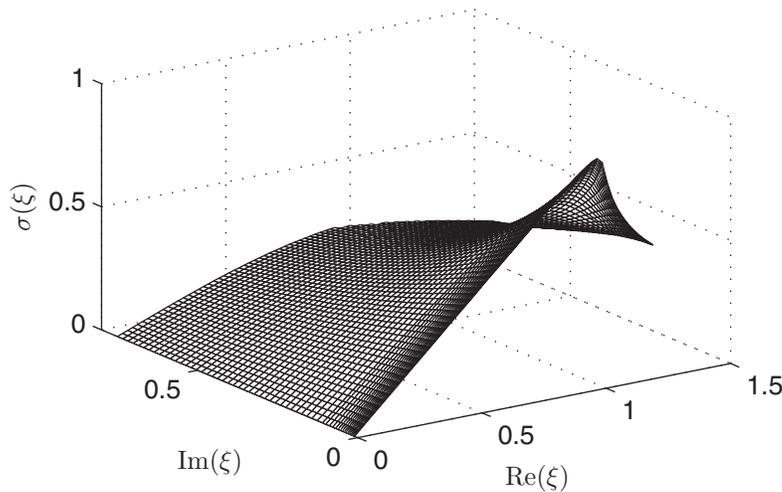


FIG. 4. Optimal splitting location  $\sigma(\xi)$  as a function of singularity location. The first quadrant of the complex plane is shown; the other quadrants are obtained by symmetry.

To quantify the improvement of a split versus an unsplit interval, define the optimal effective convergence rate by

$$(2.8) \quad \rho(\xi) = \gamma(\xi, \sigma(\xi)) = \sqrt{\alpha_S},$$

where  $\alpha_S$  is the common value of  $\alpha_L$  and  $\alpha_R$  in (2.6). This may be computed from  $\sigma(\xi)$  as described in Appendix A. Roughly speaking, we can expect a reduction in the number of total nodes needed to approximate to fixed accuracy by the factor

$$(2.9) \quad r(\xi) = \frac{\log \rho(\xi)}{\log |\phi(\xi)|}.$$

We refer to  $r$  as the splitting efficiency factor; splitting is advised only when  $r > 1$ .

Figure 5 shows the splitting efficiency factor over the first quadrant. We see from this figure that  $r \rightarrow 1$  as  $\xi$  approaches some finite curve containing the interval  $[-1, 1]$ ; for singularities outside of this boundary, splitting is contraindicated. Using the analysis of Appendix A, it is possible to show that this critical curve is in fact an ellipse, with foci at  $\pm 1$  and approximately defined by  $0.7044x^2 + 2.3830y^2 = 1$ . We can also see that near  $\xi = 1$ , there is a dramatic difference in efficiency to the left and right.

Figure 6 shows the efficiency factor along three horizontal lines close to the real axis. For singularities interior to the interval, the speedup due to splitting can easily go into the hundreds and is highly sensitive to  $\text{Im}(\xi)$ , but it rapidly drops off at  $\text{Re}(\xi) = 1$  and becomes less than 10 and insensitive to  $\text{Im}(\xi)$ .

**3. Asymptotic behavior of optimal splitting.** As  $\xi \rightarrow [-1, 1]$ , we identify two crucial asymptotic regimes,

1. *Vertical approach to a point in  $(-1, 1)$ .* Let  $\xi = p + i\epsilon$  for some  $0 \leq p < 1$ . (The case  $p < 0$  is found by symmetry.) Expansion of (2.7) reveals

$$(3.1) \quad \sigma = p - \frac{p}{(1-p^2)^{1/2}} \epsilon + O(\epsilon^3).$$

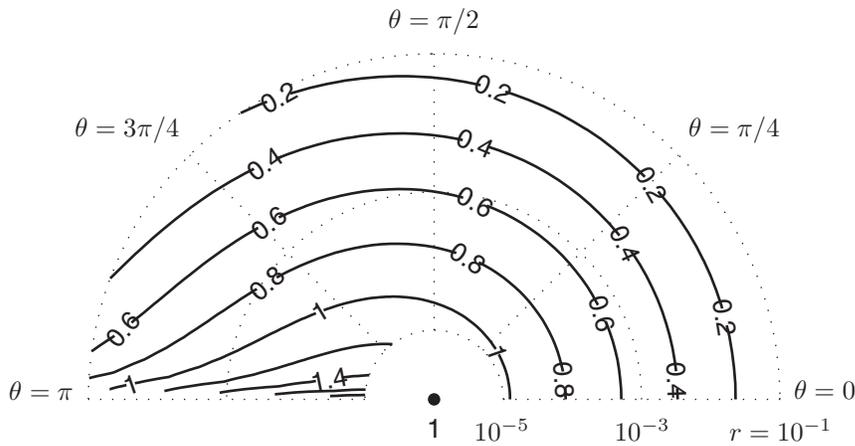
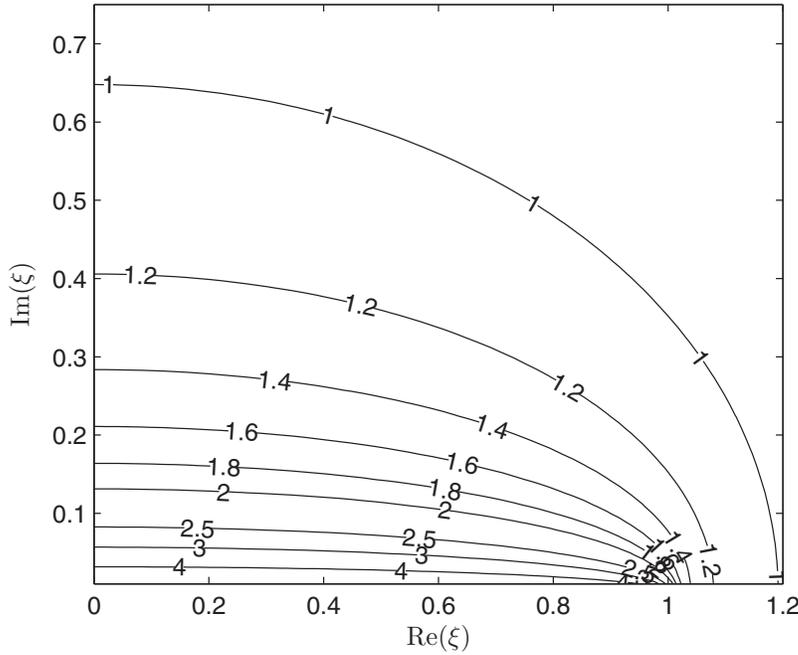


FIG. 5. Optimal splitting efficiency factor  $r(\xi)$  as a function of singularity location. The first quadrant of the complex plane is shown; the other quadrants are obtained by symmetry. Top: Outside of the largest contour shown, splitting is not recommended. The factor increases as  $(\text{Im } \xi)^{-1/2}$  as  $\text{Im } \xi \rightarrow 0$ . Bottom: Close-up view of the efficiency factor near the endpoint  $x = 1$ . Contour levels show  $\log_{10} r(\xi)$  for  $\xi$  arranged near 1 on a polar grid with logarithmic spacing in radius. The factor increases as  $|1 - \xi|^{-1/4}$  as  $\xi \rightarrow 1$  along any ray through  $\xi = 1$ , though the exponent is again  $-1/2$  along any vertical line intersecting the interval  $(0, 1)$ .

We have  $\sigma \leq p$  for all  $\epsilon$ , with equality only when  $p = 0$ . Using the calculations suggested in Appendix A for the optimal equivalent rate  $\rho(\xi)$ , we obtain

$$(3.2) \quad \log \rho = \frac{1}{\sqrt{2}(1-p^2)^{1/4}} \epsilon^{1/2} + \frac{1}{6\sqrt{2}(1-p^2)^{3/4}} \epsilon^{3/2} + O(\epsilon^{5/2}).$$

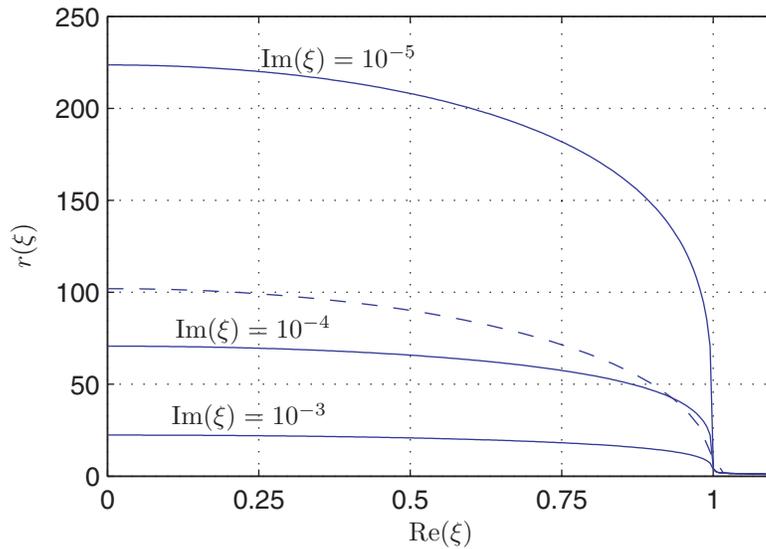


FIG. 6. Optimal splitting efficiency factor  $r(\xi)$  along three horizontal lines very close to the real axis. (The dashed line is explained in section 4.) For such singularities just above the approximation interval, the speedup due to splitting can be orders of magnitude larger than for a singularity that lies just beyond the endpoint.

We remark that the unsplit convergence rate  $\alpha(p + i\epsilon)$  is easily seen to be  $1 + O(\epsilon)$ , so the splitting efficiency factor  $r$  in (2.9) grows unboundedly like  $O(\epsilon^{-1/2})$  as  $\epsilon \rightarrow 0$ .

2. *Radial approach to the interval endpoint.* Inspection of (3.1)–(3.2) shows that the expansions become useless as the point  $p$  approaches 1. Thus, now suppose that the complex singularity is at  $\xi = 1 + \epsilon e^{i\theta}$  for some  $0 \leq \theta < \pi$ . Now we find

$$(3.3) \quad \sigma = 1 - (1 + \cos \theta)^{1/2} \sqrt{\epsilon} + (\cos \theta)\epsilon + O(\epsilon^{3/2}),$$

which in turn implies

$$(3.4) \quad \log \rho = \frac{(1 + \cos \theta)^{1/4}}{\sqrt{2}} \epsilon^{1/4} + \frac{\sqrt{2} \cos^4(\theta/2)}{3(1 + \cos \theta)^{5/4}} \epsilon^{3/4} + O(\epsilon^{5/4}).$$

Here the unsplit rate obeys  $\log \alpha = O(\epsilon^{1/2})$ , so that the splitting efficiency factor  $r$  is  $O(\epsilon^{-1/4})$ , which grows more slowly than in the interior-singularity case. It comes as no surprise that splitting is less effective very close to the boundary, considering the benefit already conferred by the crowding of Chebyshev nodes at the endpoints.

**4. Recursive splitting.** One appealing aspect of a splitting approach to adaptive approximation is that it is easily applied recursively. Recursive splitting is obviously important for functions with multiple singularities near the approximation interval. However, it can be beneficial even for functions with one singularity. With recursive optimal splitting, one can accelerate convergence beyond what is possible using a single split. Chebfun also uses recursion to compensate for the fact that its splits are made for convenience rather than optimality. For the remainder of this section we continue to let  $f$  have a single singularity at  $\xi$  in the complex plane.

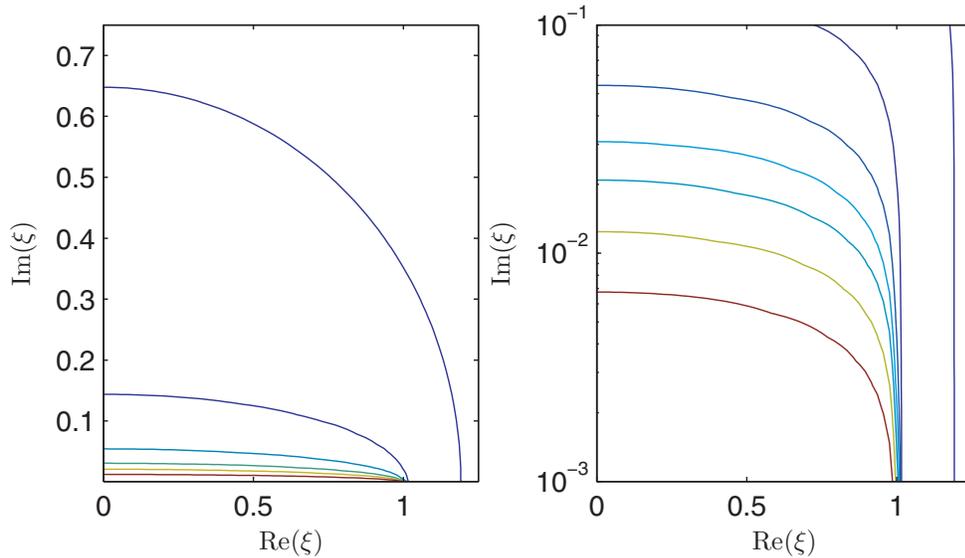


FIG. 7. Splitting efficiency factor  $\tilde{r}(\xi)$  for recursive optimal splitting. The picture on the right is a closer look for singularities close to the real axis. In both pictures, contour lines are drawn at the levels  $[1, 2, 4, 6, 8, 12, 20]$ , from right to left.

In the notation of section 2, we choose  $s = \sigma(\xi)$  for the optimal splitting location. If the equivalent split rate  $\gamma(\xi, \sigma)$  is greater than the unsplit rate  $|\phi(\xi)|$ , we accept the split. Let  $S$  be the set of all complex  $\xi$  for which this is the case. After the split, we have two approximation subproblems identical to the original but with singularity locations  $g(\xi, \sigma)$  and  $h(\xi, \sigma)$ . If  $g(\xi, \sigma) \in S$ , for example, further splitting of the left subinterval is advised, and so on. Recalling the definition of the equivalent split rate in (2.5), we define the recursive optimal splitting rate as

$$(4.1) \quad \tilde{\rho}(\xi) = \begin{cases} |\phi(\xi)|, & \xi \notin S, \\ \exp \left\{ [\log \tilde{\rho}(g(\xi, \sigma(\xi)))]^{-1} + [\log \tilde{\rho}(h(\xi, \sigma(\xi)))]^{-1} \right\}, & \xi \in S. \end{cases}$$

As with the single-split case, we define the splitting efficiency factor as

$$(4.2) \quad \tilde{r}(\xi) = \frac{\log \tilde{\rho}(\xi)}{\log |\phi(\xi)|},$$

which approximates the factor by which the total number of nodes can be reduced to achieve a fixed accuracy.

Figure 7 shows the splitting efficiency factor as a function of  $\xi$  in the first quadrant. It should be compared to Figures 5 and 6 for a single split. In Figure 6, for example, the dashed line represents the recursive splitting efficiency factor along the line  $\text{Im}(\xi) = 10^{-3}$ . Recursion gives improvement by a factor of 3–4 over a single split, at least for  $\text{Re}(\xi) < 0.75$  along this line. This factor is significant, but not nearly as large as the improvement from the first split. Quite simply, when the original singularity is near the interval, the first optimal split tends to map it close to the endpoints of its two subintervals, where additional splits have less dramatic benefits. However,

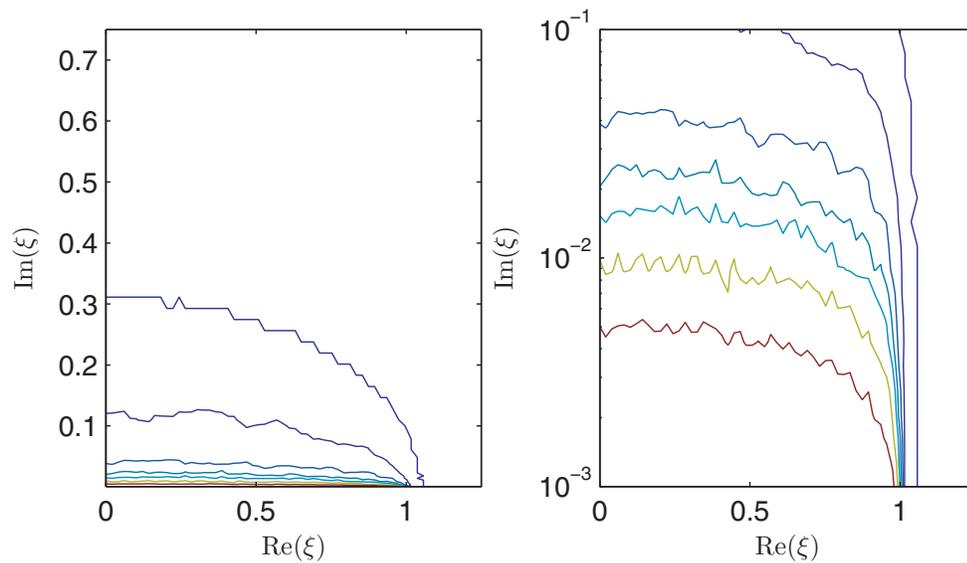


FIG. 8. Splitting efficiency factor for Chebfun's recursive bisection algorithm. The picture on the right is a closer look for singularities close to the real axis. In both pictures, contour lines are drawn at the levels  $[1, 2, 4, 6, 8, 12, 20]$ , from right to left. The contours and regions are the same as in Figure 7.

the benefit of recursion does continue to increase as the original  $\xi$  approaches the approximation interval.

As described in the introduction, Chebfun uses slow global convergence to justify bisection of the approximation interval, which is then treated piecewise recursively. We have conducted an experiment to compare the effectiveness of this practical approach with the efficiency factors computed from the convergence theory. For each complex  $\xi$ , Chebfun is used to find Chebyshev interpolants to  $f(x) = (x^2 + \xi^2)^{-1}$ , with the “splitting” option first set to “on” and then to “off.” (All other settings are given factory defaults.) The ratio of the total numbers of interpolation nodes in the two cases give an observation that should be comparable to the splitting efficiency factors  $r(\xi)$  or  $\tilde{r}(\xi)$ . The results of this experiment are given in Figure 8. The results should be compared to Figure 7 for recursive optimal splitting. The Chebfun algorithm is somewhat less effective than optimal splitting for singularities far from the interval, where splits other than optimal are not useful. However, as the singularity gets close to the interval, the difference between recursive bisection and recursive optimal splitting is insignificant.

**5. Determination of singularity location through Chebyshev–Padé (CP) approximation.** The goal of applying the optimal splitting machinery in place of Chebfun's empirical bisection approach is to reduce the total number and maximum size of the node sets on which the approximated function  $f$  is called, by predicting an optimal split location rather than finding it through binary search. In order accomplish this in practice, the location of the most significant singularity  $\xi$  must be estimated only using point values of  $f$ . This problem is related to a large literature of techniques for overcoming the Gibbs phenomenon for trigonometric and polynomial approximations; see [10, 11]. In this section we focus on the CP technique, which is widely known [1, 4] and has been used for coordinate mappings for polynomial approximation in PDEs [15].

If  $f(x)$  is given as a Chebyshev series,  $f(x) = \sum_{k=0}^{\infty} c_k T_k(x)$ , then the type  $(n, m)$  CP approximation (in the sense of Clenshaw–Lord) is defined via

$$(5.1) \quad R(x) = \frac{P(x)}{Q(x)} = \frac{\sum_{k=0}^n a_k T_k(x)}{\sum_{k=0}^m b_k T_k(x)},$$

where the coefficients are defined uniquely up to global scaling via the requirement that

$$\frac{P(x)}{Q(x)} - f(x) = \sum_{k=n+m+1}^{\infty} d_k T_k(x),$$

that is, the Chebyshev series of  $R$  and  $f$  match in the first  $n + m + 1$  coefficients. It is often observed that CP approximants reveal the singularity structure of  $f$ , in the sense that uncanceled roots of the denominator  $Q$  are close to poles and branch points of  $f$ . Note that to construct the CP approximation from point values of  $f$ , we incur aliasing error in calculating the Chebyshev series coefficients for  $f$ , which we might expect to degrade the accuracy of the singularities revealed by  $Q$ . In addition, traditional algorithms for finding  $P$  and  $Q$  are often confounded by spurious roots in those polynomials, but a robust algorithm has recently become available [9].

Chebfun includes functions for finding CP approximations and for finding roots of chebfuns quickly. An outline for a modification to Chebfun's recursive bisection algorithm is as follows:

1. Initialize with the interval  $[-1, 1]$ , and flag it for checking.
2. For each flagged subinterval  $[t_k, t_{k+1}]$ , do the following:
  - (a) Compute the Chebyshev series coefficients  $c_j$  for an  $N$ -point discretization on  $[t_k, t_{k+1}]$ .
  - (b) If the  $|c_j|$  are sufficiently small, remove the flag. Otherwise, do the following:
    - i. Find an  $(m, n)$  CP approximation  $P/Q$  on an extended interval  $[a, b] \supseteq [t_k, t_{k+1}]$ .
    - ii. Apply  $\phi$  to the roots  $q_i$  of  $Q$ , and let  $\xi = q_I$  minimize  $|\phi(q_i)|$ .
    - iii. Find the optimal split location  $\sigma(\xi)$  and the splitting efficiency factor  $r(\xi)$ .
    - iv. If  $r(\xi) > \theta \geq 1$ , accept the split and let  $s = \sigma(\xi)$ . Otherwise, let  $s = (t_k + t_{k+1})/2$ .
    - v. Replace  $[t_k, t_{k+1}]$  with  $[t_k, s]$  and  $[s, t_{k+1}]$ , and flag them for checking.
3. If some intervals are flagged, go to step 2. Otherwise, merge piecewise representations opportunistically and return the result.

We chose  $m = N/2$  and  $n = 2$  in the CP approximation, and we used the threshold  $\theta = 1.5$  to avoid making splitting decisions based on fairly weak evidence of singularities. The use of an extended interval in step 2(b)i was found to improve the quality of the singularity locations. Each endpoint was moved outward 10% of the subinterval width or to the next interval breakpoint, whichever came first. The value  $N = 128$  was used for the experiments in this section, except where noted below.

We first report on the results of two experiments on functions of the form

$$(5.2) \quad f(x) = \sin(e^{3x}) \operatorname{Im}[\log(1 - 1/\xi_1)] (x - \xi_2)^{-1/2}.$$

The branch singularity at the complex  $\xi_1$  gives a step-like behavior at  $x \approx \operatorname{Re}(\xi_1)$ , and the singularity at  $\xi_2 > 1$  gives a fairly weak blowup behavior near the right

endpoint. This singularity is difficult to detect accurately using CP approximation, both because of its type and its location.

In the first test, we let  $\xi_1 = -0.3 + 0.01i$  and  $\xi_2 = 1 + 10^{-5}$ . Since  $|\phi(\xi_1)| = 1.0105$  and  $|\phi(\xi_2)| = 1.0045$ , the weak blowup singularity is more significant than the interior step. In “splitting off” mode, Chebfun uses 5884 points to define a global interpolating polynomial. CP splitting found 6 interior breakpoints, leading to an 7-interval piecewise approximation requiring a total of 491 nodes. By comparison, Chebfun with “splitting on” needs 488 nodes on 5 intervals; however, finding the final form first required breaking into 12 subintervals.

In the second test, we reverse the roles of the singularities to make the interior step more significant, by setting  $\xi_1 = -0.3 + 10^{-5}i$  and  $\xi_2 = 1.005$ . Without splitting, Chebfun cannot resolve this function using  $2^{16}$  points. The `chebpadesplit` function chooses 8 interior breakpoints, leading to 9 subintervals. The piecewise interpolant required 588 nodes altogether. The Chebfun splitting mode resolves the function on 18 subintervals before compressing them to 8 subintervals, using a total of 673 function values.

Our final test concerns the original situation that motivated this look at optimal splitting, that of solving a differential equation. Consider the linear boundary-value problem (BVP)

$$(5.3) \quad -\epsilon^2 u'' + u = \sin(16e^x), \quad u(-1) = u(1) = 0.$$

For  $\epsilon \ll 1$ , the solution to this singularly perturbed problem has boundary layers of width  $O(\epsilon \log(1/\epsilon))$  at both ends [13], while the forcing function drives a need for some interior resolution. Note that to find a split location within  $10^{-6}$  of the boundary, for example, more than 20 recursive bisections are required, with a global solution of the BVP required each time.

Here “sampling” the function at a small number of points means solving the BVP on a coarse grid, so the function values themselves can be very inaccurate. When the domain has multiple subintervals, we use a domain-decomposition formulation of the problem with  $N$  nodes on each subinterval. (In principle, the numbers of nodes on fully converged subintervals might be reduced, but this reduction was not applied for the experiments here.) We employ a direct solution with enforced  $C^1$  continuity. The mechanics of singularity location and splitting are otherwise unchanged, though we found the optimal split acceptance threshold  $\theta = 1$  to work best when  $N = 64$ . Once the subintervals and discrete solution have been found satisfactorily, the Chebfun representations on each smooth piece are compressed to reflect a relative accuracy of  $10^{-8}$ . Higher accuracy is difficult to achieve consistently due to ill conditioning of the problem.

Table 1 summarizes the results of solutions for  $\epsilon^2$  between  $10^{-6}$  and  $10^{-12}$  and for  $N = 64$  and  $N = 128$  maximum points per subinterval. Overall, the numbers of subintervals and nodes grow as the boundary layers shrink, although the relationship is far from clear. In the most extreme case with  $N = 64$ , the CP splitting algorithm failed to give convergence. Aside from this case, however, the CP method led to very reasonable discretizations. For comparison, a manual splitting with internal breakpoints at  $\pm(1 - 100\epsilon)$  was also tested; the number of nodes in this discretization held constant at around 130, although in the most extreme case the accuracy could not be improved beyond about  $10^{-7}$ .

The CP splitting algorithm is far from ideal. It entails a number of parameters and choices that, when optimized for a few examples, can cause problems for others.

TABLE 1

Summary of splitting results for the singularly perturbed BVP (5.3). The number of subintervals and the total number of nodes are shown for automatic splitting by CP approximation with  $N = 64$  and  $N = 128$  as the maximum number of nodes per subinterval, and for a manual splitting based on known asymptotics of the boundary layer. In one case (not shown) the CP algorithm failed to converge in 30 subintervals.

$\epsilon^2$	CP splitting			Manual splitting	
	$N$	# of intervals	Total nodes	# of intervals	Total nodes
$10^{-6}$	64	4	119	3	131
$10^{-6}$	128	3	115		
$10^{-8}$	64	6	117	3	134
$10^{-8}$	128	3	104		
$10^{-10}$	64	10	138	3	136
$10^{-10}$	128	7	115		
$10^{-12}$	64	23	205	3	134
$10^{-12}$	128	9	150		

Even so, it clearly enables more efficient domain splittings than repeated bisection for functions that have singularities very close to the approximation interval.

**6. Conclusion.** Potential theory clarifies how to split an interval for optimal approximation by Chebyshev polynomial interpolation. The splitting mode of Chebfun does a very good job of applying binary search to find optimal splitting points, at least when splitting offers any significant benefit. However, the search process may require many more function evaluations than are eventually needed to represent the function accurately.

Improvement to binary splitting requires knowledge of the singularity structure of the approximated function in the complex plane. Given only information about function values on the real interval, one way to derive at least approximate singularity information is to use rational approximation. Even though the singularity locations are not always very accurate, such a strategy can be demonstrated to be effective in a few cases. It remains to be seen whether the approach is robust enough to be worthwhile in regular practice.

There may be alternative ways to extract singularity information from the available data. The rate-limiting singularity is reflected in the (decaying) amplitude and phase of the Chebyshev polynomial coefficients of the interpolant, so the tail of the coefficients might be fit to find it. Another possibility is to use subinterval interpolation information after one bisection, finding singularities by checking for intersections of Bernstein ellipses for different representations. As with rational approximation, the robustness of these approaches remain open to question.

**Appendix A. Derivation of optimal split location.** Here we derive (2.7). Consider the Bernstein ellipse as parameterized by

$$(A.1) \quad z = \frac{1}{2}(\alpha e^{i\theta} + \alpha^{-1} e^{-i\theta}), \quad 0 \leq \theta \leq 2\pi,$$

with semimajor axis  $a = (\alpha + \alpha^{-1})/2$ , corresponding to convergence rate  $\alpha = a + \sqrt{a^2 - 1}$ . If a point  $z = x + iy$  on a particular ellipse is given, then by comparing real and imaginary parts in (A.1) it follows that

$$x = \frac{1}{2}(\alpha + \alpha^{-1}) \cos \theta, \quad y = \frac{1}{2}(\alpha - \alpha^{-1}) \sin \theta,$$

and if  $\theta$  is eliminated,

$$\frac{4x^2}{(\alpha + \alpha^{-1})^2} + \frac{4y^2}{(\alpha - \alpha^{-1})^2} = 1.$$

Simplifying and putting  $\tau = \alpha^2$  yields

$$\tau^2 - 4(x^2 + y^2)\tau - 2(1 - 4x^2 + 4y^2) - 4(x^2 + y^2)\tau^{-1} - \tau^{-2} = 0.$$

Now put  $T = \tau + \tau^{-1}$  and use completion of squares to obtain the quadratic

$$(A.2) \quad T^2 - 4(x^2 + y^2)T - 4(1 - 2x^2 + 2y^2) = 0.$$

The value of the semimajor axis  $a$  can therefore be computed by solving two quadratics in succession: first solve (A.2) for  $T$  and then solve  $\tau^2 - T\tau + 1 = 0$  for  $\tau = \alpha^2$ .

Consider now the situation where the interval is split into  $[-1, s]$  and  $[s, 1]$ . To map  $[-1, s] \rightarrow [-1, 1]$  and  $[s, 1] \rightarrow [-1, 1]$  consider, respectively,

$$g(x, s) = 2\frac{x+1}{s+1} - 1, \quad h(x, s) = 2\frac{x-s}{1-s} - 1,$$

as defined in (2.3). Assume  $\xi = x + iy$  is a point on both ellipses. We now apply the condition that the value of  $T$  is the same in each case. On the left part of the interval, let

$$X = \operatorname{Re}\{g(x + iy, s)\} = 2\frac{x+1}{s+1} - 1, \quad Y = \operatorname{Im}\{g(x + iy, s)\} = \frac{2y}{s+1}$$

and substitute  $X \rightarrow x, Y \rightarrow y$  into (A.2) to obtain

$$(A.3) \quad (s+1)^2 T^2 - 4((2x+1-s)^2 + 4y^2)T - 4(s+1)^2 + 8(2x+1-s)^2 - 32y^2 = 0.$$

Likewise, on the right part of the interval, let

$$X = \operatorname{Re}\{h(x + iy, s)\} = 2\frac{x-s}{1-s} - 1, \quad Y = \operatorname{Im}\{h(x + iy, s)\} = \frac{2y}{1-s},$$

and substitution into (A.2) yields

$$(A.4) \quad (s-1)^2 T^2 - 4((2x-1-s)^2 + 4y^2)T - 4(s-1)^2 + 8(2x-1-s)^2 - 32y^2 = 0.$$

With some help from a computer algebra system,  $T$  can be eliminated from (A.3) and (A.4) to give a single equation for  $s$  in terms of  $\xi = x + iy$ , namely,

$$(A.5) \quad s^4 - 4xs^3 + (1 + y^2 + 5x^2)s^2 - 2x(1 + x^2 + y^2)s + x^2 = 0.$$

We conclude with an analysis of the roots, leading to the formula (2.7) for the optimal split location.

LEMMA A.1. *For  $\xi = x + iy$  in the first quadrant, (A.5) has a single real root in the interval  $(0, x)$ . It is given by*

$$(A.6) \quad s = x - \frac{1}{\sqrt{2}} \sqrt{(x^2 - y^2 - 1) + \sqrt{(1 + y^2 - x^2)^2 + 4x^2y^2}}.$$

*Proof.* Equation (A.5) simplifies if we define  $S = (s - x)^2$ :

$$S^2 + (1 + y^2 - x^2)S - x^2y^2 = 0.$$

This quadratic in  $S$  has one positive and one negative root. The negative root can be discarded because it gives rise to complex values of  $s$ . The positive root is

$$S = \frac{1}{2} \left( (x^2 - y^2 - 1) + \sqrt{(1 + y^2 - x^2)^2 + 4x^2y^2} \right) < x^2,$$

where the inequality on the right follows from

$$(1 + y^2 - x^2)^2 + 4x^2y^2 = (1 + x^2 + y^2)^2 - 4x^2 < (1 + x^2 + y^2)^2.$$

From this (A.6) follows, as does the fact that  $s \in (0, x)$ .  $\square$

Once the optimal splitting location  $s = \sigma(\xi)$  is known, the corresponding optimal convergence rate (2.8) can be computed as follows. It is possible to show from (A.3) and (A.4) that

$$T = \frac{8x}{s} - 6.$$

From here, the splitting rate  $\rho(\xi) = \alpha_S^{1/2} = \tau^{1/4}$  follows from solving  $\tau^2 - T\tau + 1 = 0$ .

#### REFERENCES

- [1] G. A. BAKER AND P. GRAVES-MORRIS, *Padé Approximants*, Encyclopedia Math. Appl. 59, Cambridge University Press, Cambridge, 1996.
- [2] N. S. BANERJEE AND J. F. GEER, *Exponential Approximations Using Fourier Series Partial Sums*, Technical report, ICASE, NASA Langley Research Center, Hampton, VA, 1997.
- [3] Z. BATTLES AND L. N. TREFETHEN, *An extension of MATLAB to continuous functions and operators*, SIAM J. Sci. Comput., 25 (2004), pp. 1743–1770.
- [4] D. BRAESS, *Nonlinear Approximation Theory*, Springer, New York, 1986.
- [5] T. A. DRISCOLL AND B. FORNBERG, *A Padé-based algorithm for overcoming the Gibbs phenomenon*, Numer. Algorithms, 26 (2001), pp. 77–92.
- [6] T. A. DRISCOLL, N. HALE, AND L. N. TREFETHEN, *Chebfun Guide*, Pafnuty Publications, Oxford, UK, 2014.
- [7] K. S. ECKHOFF, *On a high order numerical method for functions with singularities*, Math. Comp., 67 (1998), pp. 1063–1087.
- [8] A. GELB AND E. TADMOR, *Detection of edges in spectral data*, Appl. Comput. Harmon. Anal., 7 (1999), pp. 101–135.
- [9] P. GONNET, R. PACHÓN, AND L. N. TREFETHEN, *Robust rational interpolation and least-squares*, Electron. Trans. Numer. Anal., 38 (2011), pp. 146–167.
- [10] D. GOTTLIEB AND C.-W. SHU, *On the Gibbs phenomenon and its resolution*, SIAM Rev., 39 (1997), pp. 644–668.
- [11] A. J. JERRI, ED., *Advances in the Gibbs Phenomenon*, Sampling Publishing, Potsdam, NY, 2011.
- [12] G. KVERNADZE, T. HAGSTROM, AND H. SHAPIRO, *Locating discontinuities of a bounded function by the partial sums of its Fourier series*, J. Sci. Comput., 14 (1999), pp. 301–327.
- [13] N. MADDEN AND M. STYNES, *A uniformly convergent numerical method for a coupled system of two singularly perturbed linear reaction-diffusion problems*, IMA J. Numer. Anal., 23 (2003), pp. 627–644.
- [14] R. PACHÓN, R. PLATTE, AND L. N. TREFETHEN, *Piecewise-smooth Chebfuns*, IMA J. Numer. Anal., 30 (2010), pp. 898–916.
- [15] T. W. TEE AND L. N. TREFETHEN, *A rational spectral collocation method with adaptively transformed Chebyshev grid points*, SIAM J. Sci. Comput., 28 (2006), pp. 1798–1811.
- [16] L. N. TREFETHEN, *Spectral Methods in MATLAB*, SIAM, Philadelphia, 2000.
- [17] L. N. TREFETHEN, *Approximation Theory and Approximation Practice*, SIAM, Philadelphia, 2013.