

MIT Open Access Articles

*On Adaptive Choice of Shifts in Rational Krylov
Subspace Reduction of Evolutionary Problems*

The MIT Faculty has made this article openly available. **Please share** how this access benefits you. Your story matters.

Citation: Druskin, Vladimir, Chad Lieberman, and Mikhail Zaslavsky. "On Adaptive Choice of Shifts in Rational Krylov Subspace Reduction of Evolutionary Problems." SIAM Journal on Scientific Computing 32.5 (2010): 2485-2496. Print.

As Published: <http://dx.doi.org/10.1137/090774082>

Publisher: Society of Industrial and Applied Mathematics (SIAM)

Persistent URL: <http://hdl.handle.net/1721.1/60578>

Version: Final published version: final published article, as it appeared in a journal, conference proceedings, or other formally published context

Terms of Use: Article is made available in accordance with the publisher's policy and may be subject to US copyright law. Please refer to the publisher's site for terms of use.



ON ADAPTIVE CHOICE OF SHIFTS IN RATIONAL KRYLOV SUBSPACE REDUCTION OF EVOLUTIONARY PROBLEMS*

VLADIMIR DRUSKIN[†], CHAD LIEBERMAN[‡], AND MIKHAIL ZASLAVSKY[†]

Abstract. We compute $u(t) = \exp(-tA)\varphi$ using rational Krylov subspace reduction for $0 \leq t < \infty$, where $u(t), \varphi \in \mathbf{R}^N$ and $0 < A = A^* \in \mathbf{R}^{N \times N}$. A priori optimization of the rational Krylov subspaces for this problem was considered in [V. Druskin, L. Knizhnerman, and M. Zaslavsky, *SIAM J. Sci. Comput.*, 31 (2009), pp. 3760–3780]. There was suggested an algorithm generating sequences of equidistributed shifts, which are asymptotically optimal for the cases with uniform spectral distributions. Here we develop a recursive greedy algorithm for choice of shifts taking into account nonuniformity of the spectrum. The algorithm is based on an explicit formula for the residual in the frequency domain allowing adaptive shift optimization at negligible cost. The effectiveness of the developed approach is demonstrated in an example of the three-dimensional diffusion problem for Maxwell's equation arising in geophysical exploration. We compare our approach with the one using the above-mentioned equidistributed sequences of shifts. Numerical examples show that our algorithm is able to adapt to the spectral density of operator A . For examples with near-uniform spectral distributions, both algorithms show the same convergence rates, but the new algorithm produces superior convergence for cases with nonuniform spectra.

Key words. matrix function, matrix exponential, greedy algorithm, time-domain Maxwell system, rational approximation, model reduction

AMS subject classifications. 30C85, 30E10, 41A05, 41A20, 65M60, 86-08

DOI. 10.1137/090774082

1. Introduction. Many boundary value problems can be reduced to computation of

$$(1.1) \quad u = f(A)\varphi,$$

where $u, \varphi \in \mathbf{R}^N$ and $A = A^* \in \mathbf{R}^{N \times N}$. We shall assume throughout this paper that $\|\varphi\| = 1$.

We are particularly interested in the solution of the evolutionary equation

$$(1.2) \quad Au + u_t = 0, \quad u|_{t=0} = \varphi$$

for $0 \leq t < \infty$ with $A > 0$, and $f(A) = \exp(-At)$ in this case. In the frequency domain (due to Plancherel's equality) this problem is equivalent to computing (1.1) with

$$(1.3) \quad f(A) = (A + sI)^{-1}$$

for s on the entire imaginary axis. Depending on application, condition number of operator A , and time interval, different approaches can be applied for solving (1.2). A short summary is given in [5]. Our main application arises from geophysical deep hydrocarbon exploration, in which case (1.2) is the semidiscretized time-domain Maxwell

*Received by the editors October 16, 2009; accepted for publication (in revised form) June 11, 2010; published electronically August 19, 2010.

<http://www.siam.org/journals/sisc/32-5/77408.html>

[†]Schlumberger Doll Research, 1 Hampshire St., Cambridge, MA 02139 (druskin1@slb.com, mzaslavsky@slb.com).

[‡]Massachusetts Institute of Technology, Cambridge, MA 02139 (celieber@mit.edu).

system (in the diffusion approximation). In practice A is a large ill-conditioned matrix given by finite-difference or finite-element discretization of the spatial operator.

We will solve this problem by projecting it onto the rational Krylov subspaces (RKSs)

$$U_n = \text{span} \left\{ (A + s_1 I)^{-1} \varphi, (A + s_2 I)^{-1} (A + s_1 I)^{-1} \varphi, \dots, \prod_{j=1}^n (A + s_j I)^{-1} \varphi \right\}$$

originally introduced by Ruhe for the solution of eigenproblems [17]. We assume that the RKS is computed using iterative methods for which there are no computational advantages in solving multiple linear systems with the same shifts (because of extensive memory requirements for the discretization of large-scale geophysical electromagnetic problems). Therefore, we do not require that the shifts coincide.

The RKSs are widely used in model reduction, in particular, for computation of transfer functions of linear problems; see, e.g., [9]. A key question in the construction of the RKS is how to choose shifts s_j .

It can be done a priori based on estimates of A 's spectral interval and (or) specifics of $f(A)$. Optimization for a single shift was considered in [7, 14, 16], and optimal choice of multiple shifts was investigated in [3, 11, 5]. The two latter works dealt with computing (1.3) when s covers, respectively, an imaginary interval and the entire imaginary axis. They began with derivation of upper bounds for the rational Krylov subspace reduction (RKSR) error via the error of a special two-parameter rational interpolation (so-called skeleton approximation [19], which is closely related to the alternating direction implicit (ADI) approximation [2, 12]), when one parameter covers A 's spectral interval and the other the domain of s . Then they reduced the optimization of these bounds via the third Zolotarev problem that has closed-form solution in terms of elliptic integrals. The shifts can be chosen as the Zolotarev interpolation points adjusted to the spectral interval. However, when adaptive accuracy control is required, a better choice is a sequence distributed uniformly with respect to Zolotarev measure (a.k.a. equidistributed sequence), which yields nested Krylov subspaces, i.e., $U_n \subset U_{n+1}$ [5]. The obtained solutions give the optimal linear convergence rate for the upper bound. This bound, however, corresponds to the case when the spectrum is uniformly distributed in the spectral interval. For cases where the spectrum is otherwise distributed, the optimal convergence may be faster; see section 3.

Indeed, an effective adaptive choice of shifts optimized for the spectral measure of operator A would be preferable to a priori algorithms. A recently developed approach to the adaptive shift optimization is based on the theory of the \mathcal{H}_2 -optimality of reduced-order models [10]. It generates a sequence of Krylov subspaces with the shifts converging to those satisfying the first-order \mathcal{H}_2 -optimality conditions. If what only matters is the size of the final reduced model, such an approach can be very efficient indeed. However, this method requires the construction of many Krylov subspaces which will not be utilized in the final model—only the last subspace is used. In our application, for example, the majority of the computational cost is in constructing the Krylov subspace (i.e., solving shifted linear systems). Therefore, this method is not feasible for our purposes. Instead, we prefer to build a single Krylov subspace by adaptively choosing shifts s_j in a greedy fashion.

A natural approach is to choose a new shift where the error is largest. Such a greedy approach for adaptive shift placement for the RKS-based model was first

discussed in [9] and recently applied to parametric model reduction [4]. Its dependence on s may have multiple local maxima, and, as was mentioned in [9], in order to use this method, a good error estimator is required (which is an open problem). If $-s_i$ are well separated from the spectra of the true and approximate resolvent, then the norm of the residual can be used as an error estimate. As a function of s it may still have multiple local extrema, but it can be rewritten in a simple explicit form which is suitable for analysis of extremal points. Our approach is based on a representation of the RKSR residual as

$$(1.4) \quad \frac{r_n(A)\varphi}{r_n(-s)}, \quad r_n(z) = \prod_{j=1}^n \frac{z - \lambda_j^{(n)}}{z + s_j},$$

where $\lambda_j^{(n)}$ are the Ritz values of A on U_n . Formula (1.4) follows from the explicit representation of the RKSR solution as a special case of the skeleton approximation obtained in [11]. The shift s_{n+1} is chosen at the maximum of (1.4) as a function of s on the spectral interval of A . Our numerical results on the examples of the time-domain three-dimensional Maxwell equations arising from geophysical applications show that the adaptive algorithm has similar convergence to an algorithm of [5] for problems with near uniform media (i.e., near uniform spectral distributions) and significantly overperforms the latter for high-contrast media (i.e., strongly nonuniform spectral distribution).

2. Properties of the RKSR in the frequency domain and shift optimizations. In this section we discuss the construction of the RKSR and rational approximation. We recall a solution and bounds on the error based on Zolotarev's theory. It is established that we may restrict shifts to the spectral interval. We discuss a natural error estimate for the RKSR based on the skeleton approximation, which leads directly to the description of our algorithm.

2.1. The RKSR and rational approximation. We begin by considering n distinct *complex* shifts s_j outside $[-\lambda_{\max}, -\lambda_{\min}]$ symmetrically with respect to the real axis. Then we shall show that they can be restricted to $[\lambda_{\min}, \lambda_{\max}]$.

To approximate (1.1), we will use Galerkin projection on U_n . Let $G_n = [g_1, \dots, g_n]$ be the matrix of an orthonormal basis of U_n generated with the help of the rational Arnoldi method (see, e.g., [17, 8, 3]). Then the RKSR's approximant to u is defined as

$$(2.1) \quad u_n = G_n f(V_n) G_n^* \varphi,$$

where

$$(2.2) \quad V_n = G_n^* A G_n.$$

Let $\lambda_j^{(n)} \in \mathbf{R}$ and $z_j^{(n)} \in \mathbf{R}^N$, $j = 1, \dots, n$, be, respectively, the Ritz values and (normalized) Ritz vectors of A on U_n , i.e., $z_j^{(n)} = G_n h_j^{(n)}$, where $h_j^{(n)}$ is a normalized eigenvector of the problem

$$V_n h_j^{(n)} - \lambda_j^{(n)} h_j^{(n)} = 0.$$

Then for $f(A) = (A + sI)^{-1}$ we can write the RKSR solution

$$v_n = G_n (V_n + sI)^{-1} G_n^* \varphi$$

via spectral decomposition as

$$(2.3) \quad v_n = \sum_{i=1}^n \frac{1}{\lambda_i^{(n)} + s} z_j^{(n)} (z_j^{(n)})^* \varphi.$$

By construction $z_j^{(n)} \in U_n$, so $z_j^{(n)} = \tilde{z}_j^{(n)}(A)\varphi$, where $\tilde{z}_j^{(n)}(\lambda)$ is an $[n - 1/n]$ -rational function. Thus, we can write

$$v_n = \tilde{v}_n(A, s)\varphi,$$

where $\tilde{v}_n(\lambda, s)$ is an $[n - 1/n]$ -rational function with respect to λ and s separately. It is known that the RKSR solution interpolates (1.3) at s_j [9]. The interpolation properties of \tilde{v} as a function of λ and s are given by the following lemma (see [11] for the proof).

LEMMA 2.1. *We have*

$$\left(\tilde{v} - \frac{1}{\lambda + s} \right) \Big|_{s=s_l} = 0, \quad \lambda \geq 0, \quad l = 1, \dots, n,$$

and

$$\left(\tilde{v} - \frac{1}{\lambda + s} \right) \Big|_{\lambda=\lambda_l^{(n)}} = 0, \quad s > 0, \quad l = 1, \dots, n.$$

The skeleton approximation f_{skel} of the function of two variables $1/(\lambda + s)$ arises in many applications, such as the ADI (see, e.g., [2, 12]) or tensor-product approximations [15, 19]. Here we define f_{skel} according to the two latter works as

$$(2.4) \quad f_{\text{skel}}(\lambda, s) = \left(\frac{1}{\lambda+s_1}, \dots, \frac{1}{\lambda+s_n} \right) M^{-1} \begin{pmatrix} \frac{1}{s+\lambda_1^{(n)}} \\ \vdots \\ \frac{1}{s+\lambda_n^{(n)}} \end{pmatrix},$$

where $M = (M_{kl})$ is the n -by- n matrix with the entries $M_{kl} = 1/(\lambda_k^{(n)} + s_l)$.

The relative error of the skeleton approximation is given by

$$(2.5) \quad \delta(\lambda, s) = \left(\frac{1}{\lambda + s} - f_{\text{skel}}(\lambda, s) \right) / \frac{1}{\lambda + s} = \frac{r_n(\lambda)}{r_n(-s)};$$

i.e., $\lambda_j^{(n)}$ and s_j are interpolating points. Both \tilde{v} and f_{skel} are $[n - 1/n]$ -rational functions of λ and of s , so from Lemma 2.1 and (2.5) we obtain

$$(2.6) \quad \tilde{v} \equiv f_{\text{skel}}.$$

Connection of the RKSR and the skeleton approximation given by (2.6) was obtained in [11].

The following theorem provides the main tool for our algorithm.

THEOREM 2.2. *The RKSR residual is given by*

$$(2.7) \quad d_n = \varphi - (A + sI)v_n = \frac{r_n(A)\varphi}{r_n(-s)}.$$

Proof. From (2.6) we obtain

$$v_n \equiv f_{\text{skel}}(A, s)\varphi.$$

Then using (2.5) we obtain $\varphi - (A + sI)v_n = \delta(A, s)\varphi = \frac{r_n(A)\varphi}{r_n(-s)}$. \square

Theorem 2.2 gives a simple analytic formula for the residual d_n as a function of s . That allows us to consider the RKSR as a two-parameter rational interpolation problem and to use a concept similar to the one employed in rational ADI with Bagby points [1] by choosing s_{n+1} at a global maximum of $\|d_n\|$ on a subset of the complex plane. Below we shall determine this subset.

2.2. Zolotarev’s theory and restriction on s_j . As was pointed in [9], the popular choices for the approximation of (1.3) on the imaginary axis are the RKSs with *real* or *imaginary* shifts; the latter is the most intuitive option. It was shown in [11] that imaginary shifts yield optimal convergence rate for the *bounded* imaginary intervals of s and A being symmetric *unbounded* nonnegative self-adjoint operators with continuum spectrum on $[0, \infty)$. However, as was shown in [5], *real* shifts can also produce optimal geometric convergence rate for the approximation of (1.3) on *the entire* imaginary axis with *bounded* positive-definite self-adjoint A having uniform spectral distribution. The results of [5] can be summarized as follows.

The Plancherel identity connects time and frequency domain errors as

$$(2.8) \quad \int_0^\infty \|\exp(-tA)\varphi - G \exp(-tV)G^* \varphi\|^2 dt = \frac{1}{2\pi} \int_{-i\infty}^{+i\infty} \|(A + sI)^{-1}\varphi - G(V + sI)^{-1}G^* \varphi\|^2 \cdot |ds|.$$

The right-hand side of (2.8) can be estimated via the error of the skeleton approximation as

$$(2.9) \quad \sqrt{\int_{-i\infty}^{+i\infty} \|(A + sI)^{-1}\varphi - G_n(V_n + sI)^{-1}G_n^* \varphi\|^2 \cdot |ds|} \leq 2c \min_{\hat{\lambda}_1^{(n)}, \dots, \hat{\lambda}_n^{(n)}} \frac{\max_{\lambda \in [\lambda_{\min}, \lambda_{\max}]} |\hat{r}_n(\lambda)|}{\min_{s \in i\mathbf{R} \cup \{\infty\}} |\hat{r}_n(s)|},$$

where

$$(2.10) \quad c = \sqrt{\int_{-i\infty}^{+i\infty} \left| \frac{1}{\lambda_{\min} + s} \right|^2 \cdot |ds|},$$

$$\hat{r}_n(z) = \prod_{j=1}^n \frac{z - \hat{\lambda}_j^{(n)}}{z + s_j}.$$

Here we distinguish \hat{r} from r by using parameters $\hat{\lambda}_i^{(n)}$ (used for minimization) instead of $\lambda_i^{(n)}$ (which are Ritz values of A on U_n). The choice of shifts is targeted to minimization of the upper bound of the L_2 -error and is given by the solution of the minimax problem

$$(2.11) \quad \min_{s_1, \dots, s_n, \hat{\lambda}_1^{(n)}, \dots, \hat{\lambda}_n^{(n)}} \frac{\max_{\lambda \in [\lambda_{\min}, \lambda_{\max}]} |\hat{r}_n(\lambda)|}{\min_{s \in i\mathbf{R} \cup \{\infty\}} |\hat{r}_n(s)|}.$$

It was shown in [5] that if we restrict ourselves to

$$s_i = \hat{\lambda}_n^{(i)}, \quad i = 1, \dots, n,$$

the upper bound can be, at worst, a factor of two larger, i.e.,

$$2 \min_{s_1, \dots, s_n, \hat{\lambda}_1^{(n)}, \dots, \hat{\lambda}_n^{(n)}} \frac{\max_{\lambda \in [\lambda_{\min}, \lambda_{\max}]} |\hat{r}_n(\lambda)|}{\min_{s \in i\mathbf{R} \cup \{\infty\}} |\hat{r}_n(s)|} \geq \min_{\hat{\lambda}_1 = s_1, \dots, \hat{\lambda}_n = s_n} \frac{\max_{\lambda \in [\lambda_{\min}, \lambda_{\max}]} |\hat{r}_n(\lambda)|}{\min_{s \in i\mathbf{R} \cup \{\infty\}} |\hat{r}_n(s)|}.$$

In turn, the minimization of the right-hand side of this inequality is equivalent to the minimization of

$$\hat{\delta}_n = \min_{\hat{\lambda}_1^{(n)} = s_1, \dots, \hat{\lambda}_n^{(n)} = s_n} \max_{\lambda \in [\lambda_{\min}, \lambda_{\max}]} |\hat{r}_n(\lambda)|,$$

which is the classical Zolotarev problem that yields $s_j = s_j^{(n)} \in [\lambda_{\min}, \lambda_{\max}]$ expressed in terms of elliptic integrals. The minimum satisfies two-sided inequality

$$(2.12) \quad \rho^n \leq \hat{\delta}_n \leq 2\rho^n.$$

The geometric convergence factor ρ is also expressed in terms of elliptic integrals, and for large condition numbers $\frac{\lambda_{\max}}{\lambda_{\min}}$ it can be well-approximated by

$$\exp\left(-\frac{\pi^2}{2} / \log \frac{4\lambda_{\max}}{\lambda_{\min}}\right);$$

i.e., it shows logarithmic dependence on A 's condition number. These (dependent on n) shifts s_j^n or a nested sequence of equidistributed shifts (EDSs) $s_j, j = 1, \dots, \infty$, with the same limiting measure and geometric convergence rate as $n \rightarrow \infty$ are used in the a priori algorithms of [5].

For both of the above cases the shifts are located in $[\lambda_{\min}, \lambda_{\max}]$. Likewise, the shifts of the ADI approximation of the Lyapunov exponent satisfying necessary optimality conditions belong to $[\lambda_{\min}, \lambda_{\max}]$ [2]. The same is true for the shifts satisfying the first-order \mathcal{H}_2 -optimality condition

$$(2.13) \quad s_i = \lambda_i^{(n)}, \quad i = 1, \dots, n$$

[10, 13]. So, it is reasonable to restrict $s_j, j = 1, \dots, n$, to $[\lambda_{\min}, \lambda_{\max}]$. In this case $-s_j$ will be separated from the poles of the true and approximate resolvent by at least $2\lambda_{\min}$, so the residual can be used for the error estimate.

2.3. Properties of the objective function on the interval of optimization. The function $\frac{1}{r_n(-s)}$ has poles at $s = -\lambda_i \in [-\lambda_{\max}, -\lambda_{\min}]$, and, by construction, the interpolation points s_i lie in $[\lambda_{\min}, \lambda_{\max}]$ (see Figure 1). The following proposition is useful in the maximization of $|\frac{1}{r_n(-s)}|$ on $[\lambda_{\min}, \lambda_{\max}]$.

PROPOSITION 2.3. *Let us assume that all poles $-\lambda_i^{(n)}$ are on $[-\lambda_{\max}, -\lambda_{\min}]$ and that all interpolation points s_i are on $[\lambda_{\min}, \lambda_{\max}]$ and do not coincide. Then all the extrema of $\frac{1}{r_n(-s)}$ on $[0, \infty)$ are ripples (local maxima of $|\frac{1}{r_n(-s)}|$) located only between the interpolation points, such that there is one and only one ripple between two adjacent interpolation points.*

Proof. As was mentioned above, function $\frac{1}{r_n(-s)}$ is analytic for $s > 0$ and has zeros and poles at s_i and $-\lambda_i^{(n)}, i = 1, \dots, n$, respectively. Thus, according to Rolle's

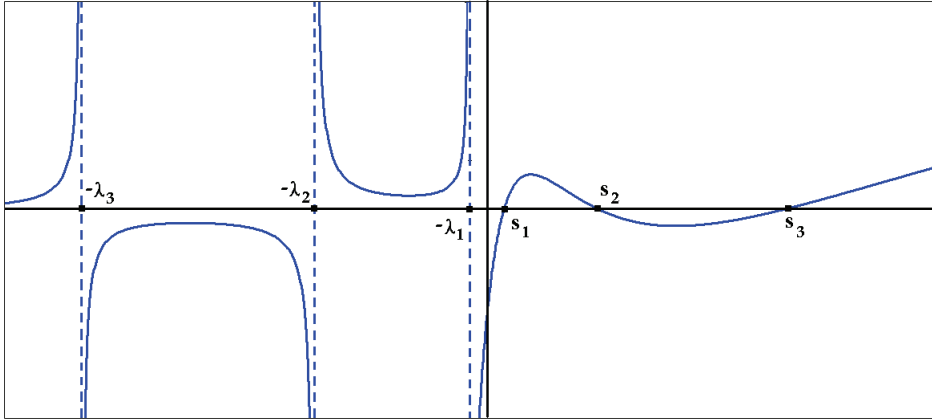


FIG. 1. An example of $1/r_n(-s)$ for $n = 3$. Three shifts $s_1, s_2,$ and s_3 on $s \in [\lambda_{min}, \lambda_{max}]$ and three Ritz values $\lambda_1, \lambda_2,$ and λ_3 .

theorem, $|\frac{1}{r_n(-s)}|$ has at least one local maximum on each interval between zeros, i.e., at least $n - 1$ local maxima in total. Moreover, since $\frac{1}{r_n(-s)}$ has no zeros for $s < 0$, it has at least one local extremum between each two nearest poles, i.e., $n - 1$ more local extrema. By construction $\frac{1}{r_n(-s)}$ is an $[n/n]$ rational function; i.e., it has $2n - 2$ extremal points. So, for $s > 0$, it has exactly $n - 1$ extremal points and has only one ripple between two adjacent zeros. \square

So, according to Proposition 2.3, in order to maximize $|\frac{1}{r_n(-s)}|$ on $[\lambda_{min}, \lambda_{max}]$, it is sufficient to choose the maximum of $|\frac{1}{r_n(-\lambda_{min})}|, |\frac{1}{r_n(-\lambda_{max})}|$ and the $n - 1$ local maxima between the interpolation points.

We now present the algorithm for constructing a reduced model based on adaptive shifts.

ALGORITHM 2.4.

1. Estimate $\lambda_{min}, \lambda_{max}$. Set $s_1 = \lambda_{min}, s_2 = \lambda_{max}, n = 2$; compute g_1 .
2. Compute g_n using rational Arnoldi; see, e.g., [3, 9].
3. Compute V_n using (2.2); compute Raleigh–Ritz eigenpairs $\lambda_i^{(n)}, z_i^{(n)}$ for $i = 1, \dots, n$; compute v_n using (2.3).
4. Find s_{n+1} satisfying

$$(2.14) \quad \left| \frac{1}{r_n(-s_{n+1})} \right| = \max_{[\lambda_{min}, \lambda_{max}]} \left| \frac{1}{r_n(-s)} \right|.$$

If there are several shifts for which the maximum is attained, it suffices to select any one of the corresponding shifts.

5. Set $n = n + 1$.

Repeat steps 1 through 3 until convergence, which can be controlled by stabilization of v_n or a tolerance on the residual for s on the entire imaginary axis.

3. Numerical experiments. Our application of interest is the time-domain forward electromagnetic problem arising in geophysical oil exploration. Under the assumption that displacement currents have negligible effect, the problem can be reduced to the diffusion time-domain Maxwell equations in \mathbf{R}^3 , i.e.,

$$(3.1) \quad \nabla \times (\mu\sigma)^{-1} \nabla \times \mathbf{H} + \frac{\partial \mathbf{H}}{\partial t} = 0, \quad \mathbf{H}|_{t=0} = \mathbf{H}_0,$$

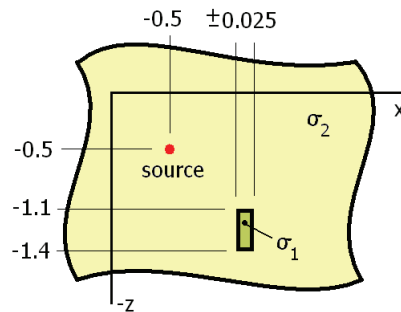


FIG. 2. Medium for test case. Uniform medium $\sigma_2 = 1$ with a resistive inclusion. The inclusion has conductivity $\sigma_1 = 3$ and $\sigma_1 = 10^5$ for the low-contrast and high-contrast test cases, respectively. The electrical source is placed at $(x, z) = (-0.5, -0.5)$, and the receiver is at $(0.5, -0.5)$. The inclusion is placed at $[-0.025, 0.025] \times [-1.4, -1.1]$.

with zero boundary conditions at infinity. Here $\mathbf{H}(t)$ is the time-dependent vector-valued magnetic field, μ is the magnetic permeability (assumed constant), and σ is a spatially varying but temporally constant uniformly bounded positive electrical conductivity distribution. The corresponding inverse problem seeks an estimate of the electrical conductivity given electromagnetic field data at a set of receivers. For the accurate solution of the inverse problem, we must have forward solutions on the time interval $[t_{\max}, t_{\min}]$ for very large t_{\max}/t_{\min} (10^4 or greater) to separate and resolve near and far zone inhomogeneities of geological formations.

Conservative finite-difference discretization of (3.1) yields nonnegative symmetric $A \approx \nabla \times (\mu\sigma)^{-1} \nabla \times$. The operator A has a null space, but its restriction to the divergence-free subspace is positive definite; therefore, we require $\nabla \cdot \mathbf{H}_0 = 0$. We use the preconditioned conjugate gradient method described in [20] to solve the linear systems with shifted matrices $A + s_i I$. The preconditioner is based on separable operators for layered media.

We test our adaptive algorithm on three different media: a uniform medium; one with low-contrast resistivity; and one with high-contrast resistivity. The uniform medium has a conductivity $\sigma_2 = 1$, and the latter two have a local inclusion. The conductivity in the inclusion is $\sigma_1 = 3$ for the low-contrast case and $\sigma_1 = 10^5$ for the high-contrast case. Gershgorin's theorem estimates the spectral interval $[\lambda_{\min}, \lambda_{\max}]$ as $[0.9 \cdot 10^{-3}; 1.5 \cdot 10^3]$, $[0.3 \cdot 10^{-3}; 1.5 \cdot 10^3]$, and $[0.9 \cdot 10^{-8}; 1.5 \cdot 10^3]$ for the homogeneous, low-contrast, and high-contrast media, respectively. The electromagnetic field is excited by a point (horizontal) electric dipole source (see Figure 2). We used $N = 0.5 \cdot 10^5$ for all the examples.

In each case, we are interested in convergence of the adaptive algorithm compared to the EDSs coming from the Zolotarev formulation [5]. As exact we consider the approximate solution given by the EDS for $n = 150$. In Figures 3, 4, and 5 we plot the results for the uniform medium, low-contrast, and high-contrast cases, respectively. The (b) figures show graphs of $\sqrt{\int_0^\infty \|u_n - u\|^2 dt}$ (which due to (2.8) is also the L_2 -error in the frequency domain) for the adaptive algorithm and EDS together with the lower bound from (2.12) that gives the asymptotical error bound for the EDS. The (a) figures show the shift distributions for both approaches and the Ritz values for the adaptive method corresponding to $n = 30$. In Figure 6 for $n = 20$ and $n = 40$ we also plotted the time distributions of $\|u_n - u\|$ on $[t_{\min}; t_{\max}] = [10^{-2} s; 10^2 s]$ for both

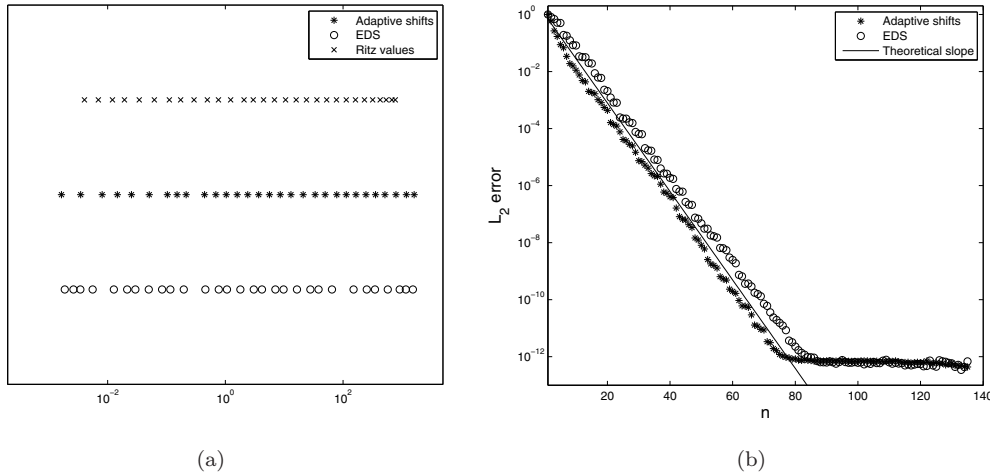


FIG. 3. Shift distribution (a) and convergence results (b) for our adaptive algorithm and the EDSs for the uniform medium test case. Shifts for both approaches are distributed uniformly, and both algorithms show comparable rates of convergence.

the algorithms and all three media. The error curves of the adaptive algorithm consistently exhibit better uniformity with more pronounced effects for higher contrasts.

The discrete problems come from the approximation of operators with continuum spectrum and $n \ll N$, so we see linear (geometric) convergence for all the cases. For the uniform medium the spectrum is close to uniform, and therefore the shift distributions are almost the same. Both algorithms converge at the same rate. The presence of the low-contrast inclusion leads to slight spectral adaptation. Therefore one can observe a slightly better convergence rate for the adaptive shift approach. However, for the high-contrast case, we see that our algorithm can adapt to the spectrum of operator, and all adaptive shifts are located at the upper half of spectral interval (in logarithmic scale), which results in higher accuracy at earlier times. In this case the adaptive shifts produce reduced models which converge at a geometric rate about a factor of two faster than the EDS-based models. It is interesting to see that in all cases the shift distribution of the adaptive algorithm tries to approximate the Ritz value distribution, i.e., in some sense tries to mimic \mathcal{H}_2 -optimality condition (2.13).

In summary, our adaptive algorithm matches the convergence of the a priori algorithms for uniform spectral distributions and overperforms in cases where the optimal convergence is faster, i.e., for nonuniform spectral distributions.

4. Concluding remarks.

Remark 1. It is known that the Ritz values $\lambda_i^{(n)}$ solve the minimization problem

$$\|r_n(A)\varphi\| = \min_{\lambda_1, \dots, \lambda_n} \left\| \prod_{j=1}^n (A - \lambda_j I) q_n(A)^{-1} \varphi \right\|,$$

where $q(z) = \prod_{j=1}^n (z + s_j)$. So, after choosing s_n maximizing $\left| \frac{1}{r_n(-s)} \right|$ on $(n - 1)$ st iteration ($n > 2$), the algorithm finds $\lambda_1 = \lambda_1^{(n)}, \dots, \lambda_n = \lambda_n^{(n)}$ minimizing

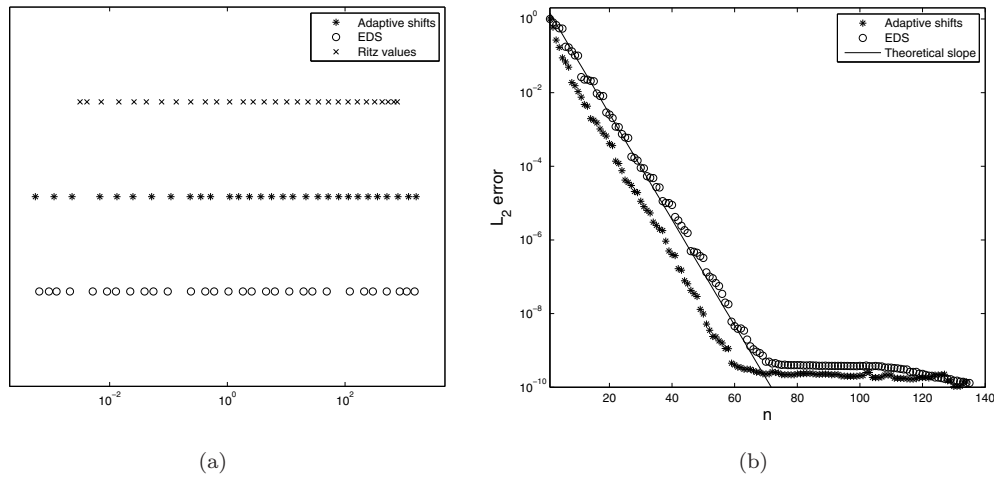


FIG. 4. Shift distribution (a) and convergence results (b) for our adaptive algorithm and the EDSs for the low-contrast test case. Shifts for our approach show minor adaptation, and therefore convergence rate is slightly better than for the EDS-based approach.

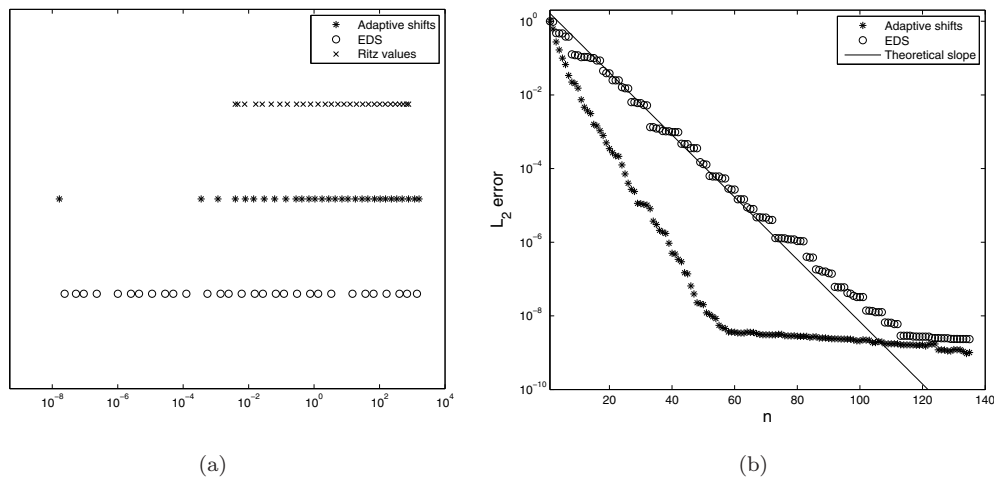


FIG. 5. Shift distribution (a) and convergence results (b) for our adaptive algorithm and the EDSs for the high-contrast test case. Shifts for our approach are essentially adapted to the spectrum of operator A and cover just a half of spectral interval (in logarithmic scale). Therefore our adaptive algorithm outperforms the EDS for the high-contrast case by factor 2 in convergence rate.

$$\left\| \prod_{j=1}^n (A - \lambda_j I) q_n(A)^{-1} \varphi \right\|$$

on the n th iteration. In the worst case scenario, its error seems to be bounded between the error of Zolotarev’s solution (which minimizes $\|d_n\|$ with respect to all

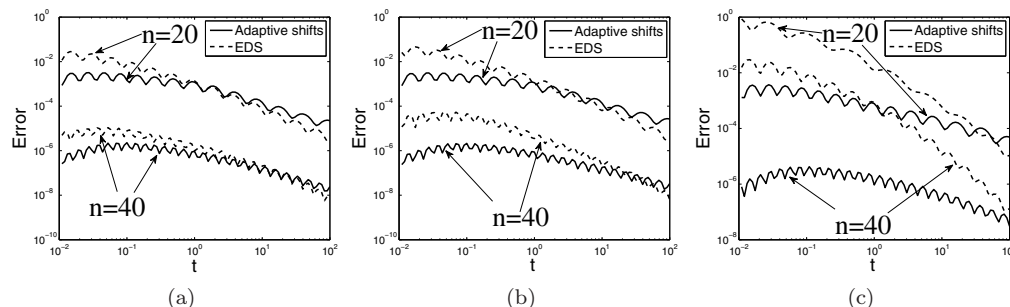


FIG. 6. Time distributions of the errors of the adaptive and a priori algorithms as a function of time for $n = 20$ and $n = 40$. For the uniform medium (a) both the algorithms show almost the same accuracy, but the error of the adaptive algorithm is slightly more uniform. For the low-contrast case (b) errors for both algorithms are still comparable, but the advantage of the adaptive algorithm becomes more pronounced. For the high-contrast case (c) the adaptive algorithm yields significantly better error distribution on the entire time interval (with the strongest effects at early times).

$s_i = \lambda_i^{(n)}$, $i = 1, \dots, n$, simultaneously) and the error of Bagby's solution (which is obtained through sequential minimization with respect to s_n and λ_n under all previous s_i , λ_i , $i = 1, \dots, n - 1$) (see [18]). But both the Zolotarev and Bagby points yield optimal geometric convergence rate for (2.12). We do not have a proof that our algorithm possesses such a property, i.e., yields the geometric rate not worse than the one given by the Zolotarev algorithm for uniformly distributed spectrum on $[\lambda_{\min}, \lambda_{\max}]$; however, our numerical experiments confirm validity of such a conjecture. They clearly show that the adaptive algorithm matches the convergence of the a priori algorithms for uniform spectral distributions and overperforms for nonuniform spectral distributions. Moreover, the shifts try to mirror the Ritz value in average, i.e., to approximate in a weak sense the first-order optimality conditions (2.13). Indeed, it would be interesting to estimate the geometric convergence rate of the adaptive algorithm based on the logarithmic potential theory.

Remark 2. We believe that the algorithm can be extended to non-Hermitian A if, instead of the spectral interval, the maximization of $|\frac{1}{r_n(-s)}|$ is performed at the boundary of the A 's field of value or, if its estimate is unavailable, at the estimated boundary of A 's spectrum. Likewise, for the so-called parameter-dependent Krylov subspace reduction of the high-order Hermitian problems [6], the residual maximization can be performed at the boundary of the so-called resonance domain, which gives an easily computable estimate of the nonlinear spectra of both the original and the reduced problems. Indeed such extensions must be thoroughly investigated.

Remark 3. In the current form, our adaptive algorithm needs an a priori estimate of A 's spectral interval, even though the estimate can be quite rough. In the future, we would like the algorithm to be fully adaptive, i.e., with the spectral estimate as part of the algorithm.

Acknowledgment. The authors are thankful to Leonid Knizhnerman for important comments.

REFERENCES

- [1] T. BAGBY, *On interpolation by rational functions*, Duke Math. J., 36 (1969), pp. 95–104.
- [2] B. LE BAILLY AND J. THIRAN, *Optimal rational functions for the generalized Zolotarev problem in the complex plane*, SIAM J. Numer. Anal., 38 (2000), pp. 1409–1424.

- [3] B. BECKERMANN AND L. REICHEL, *Error estimates and evaluation of matrix functions via the Faber transform*, SIAM J. Numer. Anal., 47 (2009), pp. 3849–3883.
- [4] T. BUI-THANH, K. WILLCOX, AND O. GHATTAS, *Model reduction for large-scale systems with high-dimensional parametric input space*, SIAM J. Sci. Comput., 30 (2008), pp. 3270–3288.
- [5] V. DRUSKIN, L. KNIZHNERMAN, AND M. ZASLAVSKY, *Solution of large scale evolutionary problems using rational Krylov subspaces with optimized shifts*, SIAM J. Sci. Comput., 31 (2009), pp. 3760–3780.
- [6] V. DRUSKIN AND M. ZASLAVSKY, *On convergence of Krylov subspace approximations of time-invariant Hermitian dynamical systems*, Linear Algebra Appl., to appear.
- [7] J. VAN DEN ESHOF AND M. HOCHBRUCK, *Preconditioning Lanczos approximations to the matrix exponential*, SIAM J. Sci. Comput., 27 (2006), pp. 1438–1457.
- [8] R. FREUND, *Model reduction methods based on Krylov subspaces*, Acta Numer., 12 (2003), pp. 267–319.
- [9] E. GRIMME, *Krylov Projection Methods for Model Reduction*, Ph.D. thesis, University of Illinois at Urbana-Champaign, Urbana, IL, 1997.
- [10] S. GUGERCIN, A. ANTOULAS, AND C. BEATTIE, *A rational Krylov iteration for optimal \mathcal{H}_2 model reduction*, in Proceedings of the 17th International Symposium on Math, Theory of Networks and Systems, Kyoto, Japan, 2006, pp. 1665–1667.
- [11] L. KNIZHNERMAN, V. DRUSKIN, AND M. ZASLAVSKY, *On optimal convergence rate of the rational Krylov subspace reduction for electromagnetic problems in unbounded domains*, SIAM J. Numer. Anal., 47 (2009), pp. 953–971.
- [12] V. LEBEDEV, *On Zolotarev problems in the alternating direction method. II*, in Trudy Semin. S. L. Sobolev, No. 1, Novosibirsk, Nauka, Moscow, 1976, pp. 51–59 (in Russian).
- [13] L. MEIER AND D. LUENBERGER, *Approximation of linear constant systems*, IEEE Trans. Automat. Control, 12 (1967), pp. 585–588.
- [14] I. MORET AND P. NOVATI, *RD-rational approximations of matrix exponential*, BIT, 43 (2003), pp. 1–18.
- [15] I. OSELEDETS, *Lower bounds for separable approximations of the Hilbert kernel*, Sb. Math., 198 (2007), pp. 425–432 (in Russian).
- [16] M. POPOLIZIO AND V. SIMONCINI, *Acceleration techniques for approximating the matrix exponential operator*, SIAM J. Matrix Anal. Appl., 30 (2008), pp. 657–683.
- [17] A. RUHE, *The rational Krylov algorithm for nonsymmetric eigenvalue problems. III: Complex shifts for real matrices*, BIT, 34 (1994), pp. 165–176.
- [18] E. SAFF AND V. TOTIK, *Logarithmic Potentials with External Fields*, Springer-Verlag, Berlin, 1997.
- [19] E. TYRTYSHNIKOV, *Mosaic-skeleton approximations*, Calcolo, 33 (1996), pp. 47–57.
- [20] M. ZASLAVSKY, S. DAVYDYCHEVA, V. DRUSKIN, A. ABUBAKAR, T. HABASHY, AND L. KNIZHNERMAN, *Finite-difference solution of the 3D electromagnetic problem using divergence-free preconditioners*, in Proceedings of the SEG Annual Meeting, New Orleans, LA, 2006, pp. 775–778.