Proceedings of ICCAM 2004, submitted, 13 Sep 2004

A truncated-CG style method for symmetric generalized eigenvalue problems [1]

P.-A. Absil *, C. G. Baker, K. A. Gallivan

School of Computational Science, Florida State University, Tallahassee, FL 32306-4120, USA (http://www.csit.fsu.edu/{∼absil,∼cbaker,∼gallivan}).

Abstract

A numerical algorithm is proposed for computing an extreme eigenpair of a symmetric/positive-definite matrix pencil (A, B). The leftmost or the rightmost eigenvalue can be targeted. Knowledge of (A, B) is only required through a routine that performs matrix-vector products. The method has excellent global convergence properties and its local rate of convergence is superlinear. It is based on a constrained truncated-CG trust-region strategy to optimize the Rayleigh quotient, in the framework of a recently-proposed trust-region scheme on Riemannian manifolds.

Key words: Generalized eigenvalue problem, extreme eigenvalues, truncated conjugate gradient, Steihaug-Toint, trust-region, global convergence, superlinear convergence, matrix-free

1 Introduction

The generalized eigenvalue problem

$$Ax = \lambda Bx,$$

where A and B are $n \times n$ real symmetric matrices with B positive definite, arises in many scientific applications [Saa92]. The symmetric/positive-definite

* Corresponding author.

[1] This work was supported by the USA National Science Foundation under Grants ACI0324944 and CCR9912415, and by the School of Computational Science of Florida State University through a postdoctoral fellowship. This research was initiated while the first author was a Research Fellow with the Belgian National Fund for Scientific Research (FNRS) at the University of Liège.

pencil (A, B) is known to admit n real eigenvalues $\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_{n-1} \leq \lambda_n$ with associated B-orthonormal eigenvectors $v_1, \ldots, v_n$ (see [Ste01]). We call (λ_1, v_1) and (λ_n, v_n) the *leftmost* and *rightmost* eigenpairs, respectively.

Single vector iterations [Par80,BDDR00] are the simplest methods for the eigenproblem. It is worthwhile considering them briefly, as their advantages and drawbacks are ubiquitous in eigencomputation methods. If the matrix A is positive definite, the power method

$$Bx_{k+1} = Ax_k \tau_k, \tag{1}$$

where τ_k is a normalizing factor, converges to the principal eigenvector v_n of (A, B) from almost all initial points; but the rate of convergence is only linear and becomes very slow when the eigenvalues of (A, B) are not well separated. Similarly, an inverse iteration

$$(A - \mu B)x_{k+1} = Bx_k \tau_k, \tag{2}$$

with a shift μ that approximates λ_1, converges linearly to v_1 from almost all initial conditions. A higher rate of convergence can be obtained using a feedback-like process that makes the shift depend on the current iterate. When the shift is chosen as the Rayleigh quotient

$$\mathbb{R}_0^n \to \mathbb{R} : y \mapsto \frac{y^T A y}{y^T B y}, \tag{3}$$

where $\mathbb{R}_0^n$ denotes $\mathbb{R}^n$ without the origin, then a cubic rate of convergence is obtained, but global convergence is lost in the sense that the iteration converges to the "nearest" eigenvector; we refer e.g. to [Par80,BS89,ASVM04] for more details. If n is large, then only an approximate solution of (2) is sought, and the key question is to determine how crudely the solution can be approximated without tampering (too much) with the convergence of the exact iteration; for recent advances, see [SP99,GY00,SE02,Not03,KN03].

It is natural to think of combining the individual advantages of these simple methods and obtain an iteration for which iterates are cheap to compute, convergence holds globally and the rate of convergence is superlinear. There is evidence that such a method can come from an optimization approach; indeed, for the problem of finding a minimum of a smooth cost function the Euclidean space, the trust-region scheme proposed by Steihaug [Ste83] and Toint [Toi81], where the trust-region subproblems are approximately solved using a truncated CG inner iteration, possesses a similar combination of advantages.

It is well known (see for example [ST00]) that the leftmost and rightmost eigenvectors of (A, B) can be expressed as minimizers and maximizers of the

Rayleigh quotient (3)—which plays the role of a cost function. More precisely, assuming that $\lambda_1 < \lambda_2$ and $\lambda_{n-1} < \lambda_n$,

$$\frac{v_1^T A v_1}{v_1^T B v_1} < \frac{y^T A y}{y^T B y} < \frac{v_n^T A v_n}{v_n^T B v_n}$$

for all y that are collinear with neither v_1 nor v_n. The difficulty is that the optimizers of (3) are not isolated: all the points αv_1, $\alpha \in \mathbb{R}_0$, are minimizers, and all the points αv_n, $\alpha \in \mathbb{R}_0$, are maximizers. This is a cause of major difficulties of practical and theoretical nature; for example, applying the Newton method to the Rayleigh quotient (3) in $\mathbb{R}^n$ yields convergence to the origin in one step. A remedy to this difficulty is to impose some normalization condition on y that picks typically one or two allowed points in each (or almost each) line $\{\alpha y : \alpha \in \mathbb{R}^n\}$. This was recognized in the early work of Bradbury and Fletcher [BF66] where several normalization conditions were considered (such as $\|y\|_1 = 1$, $\|y\|_2 = 1$ and $\|y\|_\infty = 1$) and a nonlinear conjugate-gradient optimization approach was proposed. For the generalized eigenproblem, we propose to use the normalization $\|y\|_B = 1$, where $\|y\|_B := \sqrt{y^T B y}$; this particular normalization yields simplifications in the forthcoming developments. The optimization problem is thus to minimize or maximize the cost function

$$f : \{y \in \mathbb{R}^n : y^T B y = 1\} \to \mathbb{R} : y \mapsto \frac{y^T A y}{y^T B y}. \tag{4}$$

The minimizers are $\pm v_1$ and the maximizers are $\pm v_n$, i.e., the eigenvectors of (A, B) associated with the extreme eigenvalues.

The remaining issue is to adapt the classical (Euclidean) Steihaug approach to the constrained minimization of f. In recent work [ABG04b,ABG04a], we proposed a generalization of trust-region methods, and of the truncated CG algorithm in particular, to Riemannian manifolds. The algorithm we propose in this paper is nothing more than a particularization of this general algorithm to the optimization of f on the manifold $\{y : y^T B y = 1\}$.

We will see that since the algorithm is based on CG, it only requires a routine that returns Ax and Bx given x, along with storage space for a few n-vectors and a few scalars. Moreover, similar to the classical truncated-CG-based trust-region, the Riemannian algorithm of [ABG04a,ABG04b], with a suitably-chosen stopping criterion, converges superlinearly to local minimizers of the cost function; this means that the proposed algorithm converges locally superlinearly to the leftmost eigenvector $\pm v_1$.

The global behaviour of the proposed algorithm deserves careful discussion. As does the classical truncated CG, the general Riemannian truncated CG of [ABG04b] converges under mild assumptions to a set of stationary points of the cost function. In the case of the cost function f, the stationary points are the eigenvectors of (A, B). It can be concluded that, for any initial condition, the sequence generated by the algorithm converges to a set of eigenvectors of (A, B). More information can be obtained from the fact that f never increases between successive iterates. All the accumulation points of the sequence of iterates must share the same value of f. Since the function f at an eigenvector returns the corresponding eigenvalue, it follows that the sequence of iterates converges to an eigenspace of (A, B) associated to some eigenvalue. Moreover, the eigenspace of (A, B) associated to the leftmost eigenvalue λ_1—represented by $\pm v_1$—is a stable fixed point of the algorithm since it is the unique global minimum of the cost function f; the other eigenspaces are unstable because they are not local minima of f. Consequently, convergence to $\pm v_1$ is expected to occur in practice, although convergence to other eigenspaces—*unstable convergence* [Rut70]—is technically possible and generic convergence to $\pm v_1$ in exact arithmetic is not guaranteed by the theory. This is corroborated by our numerical experiments where the initial iterate was chosen from a random distribution and systematic convergence to $\pm v_1$ was observed, although carefully chosen initial iterates yield convergence to other eigenspaces. Note that this global behaviour is similar to that of the power method (1).

In summary, we propose a method based on truncated CG that computes the leftmost eigenpair of symmetric/positive-definite matrix pencils (A, B). Since the algorithm does not assume positive definiteness of A, it can also be applied to $(-A, B)$ and compute the rightmost eigenpair of (A, B) with the same convergence properties. It is also possible to compute a few extreme eigenvectors by using a block version of the algorithm [ABG04a] or by relying on deflation techniques [Par80]. The proposed method is "matrix-free" in the sense that A and B are only used through matrix-vector products. Moreover, it only needs memory storage for a few vectors of size n and a few scalars. Therefore, the method is particularly relevant for very large-scale problems. Finally, it has excellent and well-understood global and local convergence properties.

Of course, with $B = I$ the generalized eigenproblem reduces to the standard eigenproblem. However, in contrast to many methods that tackle the generalized eigenproblem by reducing it to a standard one, the proposed method deals naturally with the generalized eigenproblem; therefore, there is no interest in considering the case $B = I$ separately.

Finally, we point out that the link with the deflation-accelerated CG (DACG) algorithm of [GSF92,BGP97] is not as strong as it may seem. The DACG method minimizes the Rayleigh quotient using a nonlinear CG method, whereas the proposed algorithm uses linear CG as an inner iteration for approximately solving a Newton equation. In this respect, the proposed algorithm falls within the category of inexact Newton methods. The inexact scheme not only reduces the computational load while preserving superlinear convergence, but it also yields excellent global convergence properties that the exact Newton does not possess.

The paper is organized as follows. The algorithm is derived in Section 2. Its convergence properties are studied in Section 3. The algorithm has close ties with the Jacobi-Davidson method of Fokkema, Sleijpen and van der Vorst [FSvdV98] and the Tracemin method of Sameh, Wisniewski and Tong [SW82,ST00]; these connections are described briefly in Section 4. Promising numerical experiments are presented in Section 5. Conclusions are drawn in Section 6.

2 The Algorithm

The proposed method was initially derived from an algorithm for optimization on manifolds [ABG04a,ABG04b]. However, it can be presented with little if any reference to optimization and differential geometry, as discussed in this section. We return to the connection with the Riemannian Trust-Region method of [ABG04a,ABG04b] in Section 3 when we study the convergence properties of the algorithm.

Let (A, B) be a symmetric/positive-definite pencil, with (λ_1, v_1) the leftmost eigenpair. We consider the problem of computing the minimizer $\pm v_1$ of the Rayleigh quotient (4) constrained to the set $\{y : y^T By = 1\}$ by an iterative method evolving on $\{y : y^T By = 1\}$. Throughout the discussion, y denotes the current iterate. Consider the function

$$\widehat{f}_y(s) = \frac{(y+s)^T A(y+s)}{(y+s)^T B(y+s)}, \quad y^T Bs = 0, \tag{5}$$

where s has the value of an update vector tangent to the set $\{y : y^T By = 1\}$ and P is the orthogonal projector onto $\{s : y^T Bs = 0\}$, that is

$$P = I - By(y^T B^2 y)^{-1} y^T B. \tag{6}$$

One has

$$\begin{aligned}\widehat{f}_y(s) &= \frac{y^T Ay}{y^T By} + 2\frac{y^T As}{y^T By} + \frac{1}{y^T By}\left(s^T As - \frac{y^T Ay}{y^T By} s^T Bs\right) + O\left(\|s\|^3\right) \\ &= f(y) + 2\langle PAy, s\rangle + \frac{1}{2}\langle 2P(A - f(y)B)Ps, s\rangle + O\left(\|s\|^3\right),\end{aligned}$$

where $\langle u, v\rangle = u^T v$ denotes the inner product of the Euclidean space $\mathbb{R}^n$. Define

$$m_y(s) = f(y) + 2\langle PAy, s\rangle + \frac{1}{2}\langle P(A - f(y)B)Ps, s\rangle, \quad y^T Bs = 0, \tag{7}$$

to be the second order approximation of $\widehat{f}_y(s)$.

Assuming that the Hessian operator

$$\mathcal{H}_y : \{s : y^T Bs = 0\} \to \{s : y^T Bs = 0\} : s \mapsto 2P(A - f(y)B)Ps \quad (8)$$

is invertible, the quadratic model $m_y(s)$ admits one and only one stationary point s_*, solution of

$$PAy + P(A - f(y)B)Ps = 0, \quad y^T Bs = 0, \quad (9)$$

which, depending on whether the Hessian operator $\mathcal{H}_y$ is positive semidefinite, negative semidefinite, or neither, is a minimum, maximum, or saddle point of the model $m_y(s)$, respectively. The "pure" Newton approach [Smi94] consists in computing the update s_* and warping this update back onto the manifold, for example as $y_+ = (y + s_*)/\|y + s_*\|_B$. This development is also presented in [WSS98] as an application of Tapia's algorithm for constrained optimization [Tap74], and it is closely related to the rationale in [SW82,ST00] (with an fundamental difference explained in Section 4). It is also well known [Shu86,AMSV02] that this pure Newton method is equivalent to the Rayleigh quotient iteration, whose convergence behaviour is well understood [BS89].

This pure Newton approach, however, is limited by two difficulties. First, while our objective is to minimize the Rayleigh quotient (4), it is not guaranteed that the Newton iteration will converge to a minimizer; depending on the initial condition, it may converge to a saddle point (interior eigenvector) or a maximizer (rightmost eigenvector). Second, when n is large, only an approximate solution of the Newton equation (9) is sought, and it is not straightforward to determine when an approximate solution is sufficiently accurate, so that the fast convergence properties of the exact iteration are preserved while avoiding doing excessive work to compute an unnecessarily accurate approximation of s_*.

This paper innovates by proposing an inner iteration scheme for approximating s_* that addresses these two difficulties. The inner iteration directly stems from the truncated-CG trust-region method of Steihaug [Ste83]. The inner iteration proceeds as a classical CG enhanced with a dedicated stopping criterion. Steihaug's approach relies on the following observations. Consider the quadratic model $m_y(s)$ of (7) and assume for a moment that the Hessian operator $\mathcal{H}_y$ of (8) is positive-definite. Recall that CG (which can be viewed as an optimization algorithm for the quadratic model m_y [GV96]) builds a sequence $\{s_j\}$ of approximate minimizers of m_y, a sequence $\{d_j\}$ of search directions and a sequence $\{r_j\}$ of residuals. These search directions d_j are descent directions for the quadratic model $m_y(s)$ at s_j. The inner iterate s_{j+1} is the minimizer of $m_y(s)$ along the line $s_j + \alpha d_j$, hence $m_y(s_{j+1} \leq m_y(s_j)$. Finally, $\|s_{j+1}\|_I > \|s_j\|_I$.

Steihaug proposes three termination rules which work along the following lines.
(i) The raison d'être for the model $m_y(s)$ is to approximate $\widehat{f}_y(s)$ by a simpler function. As such, when $\|s\|$ gets large, the model loses its ability to closely match $\widehat{f}_y(s)$. Therefore, the CG process is terminated when it crosses the boundary of the *trust-region* $\{s : \|s\| \leq \Delta\}$, where Δ is the trust-region radius inherited from the previous outer iteration step. In other words, if $\|\arg\min_\alpha m_y(s_j+\alpha d_j)\| > \Delta$, then the inner iteration returns the approximate solution $\tilde{s} = s_j + \alpha d_j$ with $\alpha > 0$ and $\|\tilde{s}\| = \Delta$.
(ii) The Hessian operator $\mathcal{H}_y$ of (8) is positive-definite only when the current iterate y is sufficiently close to the minimizers $\pm v_1$. Consequently, it may happen that a search direction d_j is a direction of nonpositive curvature for the model $m_y(s)$, namely, $d_j^T \mathcal{H}_y d_j \leq 0$; then $\|\arg\min_\alpha m_y(s_j+\alpha d_j)\|$ is infinite. This case is considered separately in the iteration before α is computed.
(iii) Finally, the CG process is terminated when $\|r_i\|/\|r_0\| \leq \xi$ for some ξ. With a view on preserving the superlinear convergence of the exact algorithm, it is recommended in [CGT00] to use a stopping criterion of the form

$$\|r_j\| \leq \|r_0\| \min(\|r_0\|^\theta, \kappa) \tag{10}$$

for some $\theta > 0$ and $\kappa > 0$.

According to these termination criteria, the truncated CG process returns with an approximate minimizer $\tilde{s}$ of $m_y(s)$ constrained to the trust-region $\{s : \|s\| \leq \Delta\}$. A complete algorithm is obtained by embedding the inner process in a trust-region framework. The decision to accept or not the update $\tilde{s}$ and to modify the trust-region radius is based on the quotient

$$\rho = \frac{\widehat{f}_y(0) - \widehat{f}_y(\tilde{s})}{m_y(0) - m_y(\tilde{s})} \tag{11}$$

which compares the decrease predicted by the model with the decrease actually observed on $\widehat{f}_y$. If ρ is very small, then the model is very bad: the step is rejected and the trust-region radius is reduced. If ρ is small but less dramatically so, then the step is accepted but the trust-region radius is reduced. If ρ is close to 1, then there is a good agreement between the model and the function over the step, and the trust-region radius can be expanded.

These considerations yield the following algorithm.

Algorithm 1 (outer iteration – trust-region)
Data: symmetric $n \times n$ matrices A and B, with B positive definite.
Parameters: $\bar{\Delta} > 0$, $\Delta_0 \in (0, \bar{\Delta})$, and $\rho' \in (0, \frac{1}{4})$.
Input: initial iterate $x_0 \in \{y : y^T B y = 1\}$.
Output: sequence of iterates $\{x_k\}$ in $\{y : y^T B y = 1\}$.
for $k = 0, 1, 2, \ldots$

- *Obtain s_k using the Steihaug-Toint truncated conjugate-gradient method (Algorithm 2) to approximately solve the trust-region subproblem*

$$\min_{x_k^T s=0} m_{x_k}(s) \quad \text{s.t. } \|s\| \le \Delta_k, \tag{12}$$

where m is defined in (7)*.*
- *Evaluate*

$$\rho_k = \frac{\widehat{f}_{x_k}(0) - \widehat{f}_{x_k}(s_k)}{m_{x_k}(0) - m_{x_k}(s_k)} \tag{13}$$

where $\widehat{f}$ is defined in (5)*.*
- *Update the trust-region radius:*
  if $\rho_k < \frac{1}{4}$
  $\Delta_{k+1} = \frac{1}{4}\Delta_k$
  else if $\rho_k > \frac{3}{4}$ *and* $\|s_k\| = \Delta_k$
  $\Delta_{k+1} = \min(2\Delta_k, \bar{\Delta})$
  else
  $\Delta_{k+1} = \Delta_k$*;*
- *Update the iterate:*
  if $\rho_k > \rho'$

$$x_{k+1} = (x_k + s_k)/\|x_k + s_k\|_B \tag{14}$$

  else
  $x_{k+1} = x_k$*;*

end (for).

Algorithm 2 (inner iteration – truncated CG)
Set $s_0 = 0$, $r_0 = PAx_k = Ax_k - Bx_k(x_k^T B^2 x_k)^{-1} x_k^T BAx_k$, $\delta_0 = -r_0$*;*
for $j = 0, 1, 2, \dots$ *until a stopping criterion* (10) *is satisfied, perform the following operations, where* $\langle , \rangle$ *denotes the standard inner product and* $\mathcal{H}_{x_k}$ *denotes the Hessian operator defined in* (8)*.*
if $\langle \delta_j, \mathcal{H}_{x_k} \delta_j \rangle \le 0$
Compute τ *such that* $s = s_j + \tau\delta_j$ *minimizes* $m(s)$ *in* (7) *and satisfies* $\|s\| = \Delta$*;*
return s*;*
Set $\alpha_j = \langle r_j, r_j \rangle / \langle \delta_j, \mathcal{H}_{x_k} \delta_j \rangle$*;*
Set $s_{j+1} = s_j + \alpha_j \delta_j$*;*
if $\|s_{j+1}\| \ge \Delta$
Compute $\tau \ge 0$ *such that* $s = s_j + \tau\delta_j$ *satisfies* $\|s\| = \Delta$*;*
return s*;*
Set $r_{j+1} = r_j + \alpha \mathcal{H}_{x_k} \delta_j$*;*
Set $\beta_{j+1} = \langle r_{j+1}, r_{j+1} \rangle / \langle r_j, r_j \rangle$*;*
Set $\delta_{j+1} = -r_{j+1} + \beta_{j+1} \delta_j$*;*
end (for).

Finally, we mention that, as a CG process, the inner iteration nicely lends itself to preconditioning; actually, Steihaug's original paper [Ste83] deals with preconditioning. Let K be a preconditioner for $(A - f(y)B)$, i.e., some approximation of $(A - f(y)B)$ such that linear systems of the form $Ku = v$ are easily solved. Consider PKP as a preconditioner for the Hessian operator $P(A - f(s)B)P$ of (8). If this preconditioner is used in the CG process, the property that the length of the update vector increases becomes true in the K norm, i.e., $\|s_{j+1}\|_K > \|s_j\|_K$. In order to preserve the property that the inner iterates never re-enter the trust-region, the trust-region is defined as $\{s : \|s\|_K \leq \Delta\}$. The use of PKP as a preconditioner is made possible by the following result due to Olsen *et al.* [OJS90] (or see [SvdVM98]). Let u and v satisfy $PKPu = v$, $y^T Bu = 0 = y^T Bv$ and assume that $y^T BK^{-1}By \neq 0$. It follows that $u = P_{K^{-1}By,By} K^{-1} v = K^{-1} P_{By,K^{-T}By} v = \left(I - K^{-1} y (y^T K^{-1} y)^{-1} y^T\right) K^{-1} v$.

Note that some papers [SS98,vdE02] refer to reconditioning as replacing the Hessian in the correction equation (9) by some approximation. This is not what is meant here: without stopping criteria, the preconditioned CG would compute—in exact arithmetic—the *exact* solution of the Newton equation (9) in a finite number of steps. However, both approaches—solving exactly an inexact Newton equation (quasi-Newton approach) or solving approximately the exact Newton equation (inexact Newton approach)—are closely related [Căt04].

3 Convergence analysis

The global and local convergence properties of trust-region schemes, including the truncated CG variant of Steihaug and Toint, have been studied thoroughly in the literature; see [CGT00,NW99] and references therein. However, the method proposed in the previous section differs from a classical trust-region algorithm in order to accommodate the fact that the optimization problem is not defined on the Euclidean space but on the non-Euclidean set $\{y : y^T By = 1\}$. In particular, the "unwarped" cost function $\widehat{f}_y(s)$ depends on the current iterate, and the update defined by (14) is different from the classical additive update.

Fortunately, the proposed method is a particular application of the general Riemannian trust-region algorithm proposed and analyzed in [ABG04a,ABG04b]: the Riemannian manifold is here the set $\{y : y^T By = 1\}$ embedded in the Euclidean space $\mathbb{R}^n$; the cost function is the Rayleigh quotient (4); and the retraction (which defines how the manifold is locally unwarped onto the tangent space at the current iterate) is given by $R_y(s) = (y + s)/\|y + b\|_B$. Since the cost function and the retraction are smooth and the manifold is compact, it follows that all the assumptions in the convergence analysis [ABG04a] of the general algorithm are satisfied. This yields the following convergence results.

Theorem 3.1 *Let (A, B) be an $n \times n$ symmetric/positive-definite matrix pencil with eigenvalues $\lambda_1 < \lambda_2 \leq \ldots \leq \lambda_{n-1} \leq \lambda_n$ and an associated B-orthonormal basis of eigenvectors $(v_1, \ldots, v_n)$. Let $\mathcal{S}_i = \{y : Ay = \lambda_i By,\ y^T By = 1\}$ denote the intersection of the eigenspace of (A, B) associated to λ_i with the set $\{y : y^T By = 1\}$.*
(i) Let $\{x_k\}$ be a sequence of iterates generated by Algorithm 1. Then $\{x_k\}$ converges to the eigenspace of (A, B) associated to one of its eigenvalues. That is, there exists i such that $\lim_{k\to\infty} \operatorname{dist}(x_k, \mathcal{S}_i) = 0$.
(ii) Only the set $\mathcal{S}_1 = \pm v_1$ is stable. More precisely, given $i \in \{2, \ldots, n\}$ and $\epsilon > 0$, there exists x_0, $\|x_0\|_B = 1$, with $\operatorname{dist}(x_0, \mathcal{S}_i) < \epsilon$ such that the sequence $\{x_k\}$ generated by Algorithm 1 from the initial condition x_0, converges to an $\mathcal{S}_j$ with $\lambda_j < \lambda_i$.
(iii) There exists $c > 0$ such that, for all sequences $\{x_k\}$ generated by Algorithm 1 converging to $\mathcal{S}_1$, there exists $K > 0$ such that for all $k > K$,

$$\operatorname{dist}(x_{k+1}, \mathcal{S}_1) \leq c\,(\operatorname{dist}(x_k, \mathcal{S}_1))^{\min\{\theta+1,2\}} \tag{15}$$

with $\theta > 0$ as in (10).

Strictly speaking, $\operatorname{dist}(u, v)$ denotes the geodesic distance on $\{y : y^T By = 1\}$ between two points u and v, which is the length of the shortest curve on $\{y : y^T By = 1\}$ that joins u and v. However, this distance is asymptotically equivalent to the more classical Euclidean distance $\|u - v\|$ in the embedding space $\mathbb{R}^n$. That is, for all u with $\|u\|_B = 1$, there exist constants c_1, c_2 and ϵ such that, for all v that satisfies $\|v\|_B = 1$ and $\|v - u\| < \epsilon$, one has $c_1\|v - u\| \leq \operatorname{dist}(u, v) \leq c_2\|v - u\|$. Since all the statements involving "dist" in the convergence results are asymptotic, all the occurrences of dist can be replaced by the Euclidean distance.

4 Links with other methods

Not surprisingly, the proposed method relates to several Newton, CG or Krylov eigenvalues methods [ABG04a]. It can be anticipated that the strong convergence results presented in Section 3 will help understand the workings of several of these methods. In this section, we briefly consider the case of two well-known and successful methods whose workings are still the object of investigation in the literature.

4.1 Jacobi-Davidson

The Jacobi-Davidson (JD) method [FSvdV98,BDDR00] has the potential of possessing all the desired qualities mentioned in the introduction. However, its global and local convergence properties are not well understood, as they critically depend on the inner iteration utilized to compute an approximate solution of the Jacobi correction equation. It is known that JD has quadratic convergence when the Jacobi correction equation is solved exactly [SBFvdV96, Th 3.2], but it has also been observed that a moderate accuracy in the solution of the Jacobi correction equation is generally sufficient to ensure fast convergence. However, it is not straightforward to determine whether an approximate solution of the correction equation is sufficiently precise or needs additional refinement; for recent advances, see [Not02,vdE02].

Interestingly, in the case $B = I$, the Jacobi correction equation is identical to the Newton update equation (9) that the truncated CG scheme (Algorithm 2) solves approximately at each outer iteration. The truncated CG approach contains criteria that guarantee a sufficient accuracy. The criteria also make it possible to do without the subspace acceleration enhancement peculiar to JD and to use the approximate solution of the Jacobi correction equation returned by Algorithm 2 as an update vector (14); this yields the overall method (Algorithm 1-2) presented in Section 2, which has the desired guaranteed convergence behaviour described in Section 3.

In other words, while JD attempts to make up for inaccuracies in the solution of the Jacobi correction equation (9) by embedding it within a subspace acceleration scheme, the algorithm proposed here concentrates its efforts in squeezing just enough information out of the Jacobi correction equation to obtain global and superlinear convergence without resorting to a subspace acceleration technique à la Davidson. Furthermore, the method proposed here may benefit from enhancement by subspace acceleration. This enhancement is under investigation.

4.2 Tracemin

Sameh and Wisniewski [SW82] and Sameh and Tong [ST00] proposed and analyzed a trace minimization (Tracemin) algorithm for computing a few (p) minor eigenpairs of a symmetric positive definite matrix pencil (A, B). For simplicity, we consider the algorithm for the case $p = 1$; block versions of Algorithm 1-2 and the Tracemin algorithm will be considered elsewhere. The basic Tracemin method is derived as follows. Instead of (7), the Rayleigh quotient is approximated by the model

$$\begin{aligned} m_y^{TM}(s) &= y^T Ay + 2y^T As + s^T As \\ &= y^T Ay + 2\langle PAy, s\rangle + \frac{1}{2}\langle 2PAPs, s\rangle, \quad y^T Bs = 0. \end{aligned} \tag{16}$$

Comparing with the exact quadratic model (7), we see that there is a "missing term" in the second-order part. This indicates why the simple Tracemin method does not reach superlinear convergence. On the other hand, assuming that A is also positive definite, the model $m_y^{TM}(s)$ has an interesting beneficial feature: the exact minimizer $s_\star$ of (16) satisfies

$$\frac{(y+s_\star)^T}{\|y+s_\star\|_B} A \frac{(y+s_\star)}{\|y+s_\star\|_B} \leq \frac{y^T}{\|y\|_B} A \frac{y}{\|y\|_B},$$

and moreover, if CG is used to compute $s_\star$, then the above inequality is satisfied by all intermediate iterates of the CG process [ST00, Lemma 3.2]. Therefore, the basic Tracemin method is in fact a descent method for the Rayleigh quotient that is robust with respect to inexact solves.

To improve the speed of convergence of the iteration, Sameh and Wisniewski [SW82] and Sameh and Tong [ST00] proposed a dynamic shift technique that appears to be effective in practice but whose workings are not yet rigorously understood. The results of this paper may shed some light on this issue, since the "missing term" in (16) is simply a Rayleigh quotient shift.

5 Numerical experiments

In this section, we report on preliminary numerical experiments that show the strong potential of Algorithm 1-2 as a competitive method for computing extreme eigenpairs of symmetric/positive-definite matrix pencils.

The first set of experiments was conducted to illustrate the convergence properties presented in Section 3. The matrices A and B were chosen from random distributions and the initial condition x_0 was chosen from a normal distribution and B-normalized. More than 10^4 such experiments were conducted and convergence to the leftmost eigenvector v_1 was systematically observed. The θ parameter in the inner stopping criterion (10) was set to $\theta = 1.0$, and the observed results were compatible with the (at least) quadratic convergence proven in Section 3. In fact, due to the symmetry of the problem, it can be argued that the rate of convergence is actually $\min\{1+\theta, 3\}$, and this is supported by the numerical experiments. We refer to [ABG04a] for details.

A second set of experiments was conducted to compare Algorithm 1-2 with the Krylov subspace method for the generalized eigenproblem proposed by Golub and Ye [GY02, Alg. 1] (referred to as the *GY method*). *New* Note that the use of preconditioners is not considered here. These preliminary experiments were conducted on matrices of moderate size ($n = 100$); since the proposed algorithm is matrix-free, it is suitable for dealing with very-large-scale problems, but the influence of finite-precision arithmetic deserves further theoretical and numerical investigation.

In each experiment, a symmetric positive-definite matrix A was generated with specific eigenvalues. The symmetric positive definite matrix B was chosen as $B = SS^T + 1000I$, where S was a square matrix with elements chosen from a standard normal distribution. This choice allowed the eigenvalue distribution of the pencil to be essentially determined via A, while testing the ability of the method to operate on a non-trivial B. For each generated problem (A, B), the proposed method was applied using three different values of the θ parameter from criterion (10): $\theta = 0.5$, $\theta = 1.0$, and $\theta = 1.5$. The GY method was allowed to form a basis of size $m = 6$. This number was chosen so that both of the algorithms were allowed an equal amount of memory. The distance to the solution is measured by computing the angle between the current iterate and the leftmost eigenvector of the pencil.

Figure 1(a) shows the results of the first test, where the gap between the leftmost two eigenvalues is small ($\frac{\lambda_2(A,B)-\lambda_1(A,B)}{\lambda_{100}(A,B)-\lambda_1(A,B)} \approx .009$)(Figure 1(b)). The superlinear convergence of the proposed algorithm is clearly seen. Moreover, we see that in terms of the number of matrix-vector multiplications (which can be considered as a consistent measure of the computational cost of both algorithms), the proposed method outperforms the GY method, even for mild accuracy requirements.

Figure 1(c) shows the results of a second test, where the gap between the leftmost two eigenvalues was much larger ($\frac{\lambda_2(A,B)-\lambda_1(A,B)}{\lambda_{100}(A,B)-\lambda_1(A,B)} \approx .47$)(Figure 1(d)). The numerical performance, in term of matrix-vector multiplications, has improved for both algorithms. While the GY method experienced greater improvement in performance due to the larger gap, the proposed method performed comparably well.

Also note that while there is some variation in the performance of the proposed method for different values of θ, the performance is not dramatically sensitive to this parameter. This is important, because it suggests that the choice of θ be more easily made than is often the case with parameter-based methods in the literature in order to provide adequate performance of the algorithm across varying matrices A and B.

Note that the GY method has been shown to yield faster convergence when a preconditioner is used; future experiments will consider the relative performance of the preconditioned GY method against a preconditioned version of the proposed method.

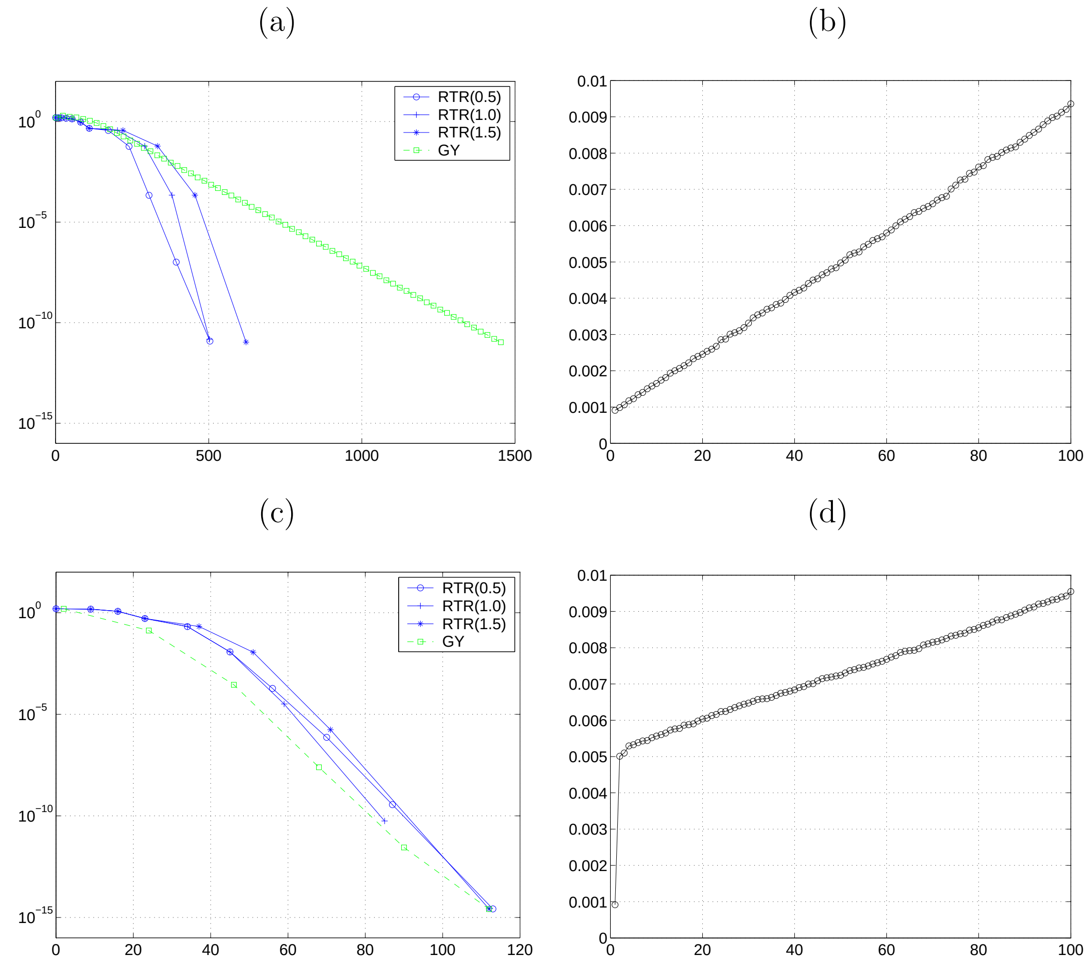

Fig. 1. Numerical efficiency of Algorithm 1-2 and the Krylov subspace method of [GY02, Alg. 1]. (a,c) plots the distance to the solution versus the number of matrix-vector products by A and B. (b,d) illustrates the spectrum of the pencil $A - \lambda B$.

6 Conclusion and future work

We have proposed a superlinearly convergent, practically globally convergent, matrix-free method for computing the leftmost eigenpair of a symmetric/positive-definite matrix pencil (A, B). The algorithm stems from a method of optimization on Riemannian manifolds [ABG04a,ABG04b]. It employs a trust-region strategy where the trust-region subproblems are solved approximately using a truncated conjugate-gradient method. The algorithm can be applied to $(-A, B)$ to compute the eigenvector corresponding to the rightmost eigenvalue (A, B). The algorithm is closely related to the Jacobi-Davidson method [FSvdV98] and the trace minimization method [SW82,ST00]; in particular, it suggests an efficient inner iteration for the Jacobi correction equation. Numerical experiments show that the proposed method is able to outperform a recently-proposed [GY02] Krylov subspace method for the generalized eigenproblem.

The current form of the proposed method is simply a straightforward application of the Riemannian trust-region method of [ABG04a,ABG04b] to the eigenproblem, but even in this simple form it demonstrates promising numerical results and sheds light on the behaviour of other well-known methods. In an upcoming paper, we report on improvements to the method that take into account properties specific to the eigenproblem. We will also report on a block version of the algorithm obtained by applying the Riemannian trust-region method on the Grassmann manifold to the trace minimization problem associated with the symmetric generalized eigenvalue problem.

Acknowledgements

Thanks to Alan Edelman, Paul Van Dooren and Gerard Sleijpen for useful discussions.

References

[ABG04a] P.-A. Absil, C. G. Baker, and K. A. Gallivan, *Trust-region methods on Riemannian manifolds*, Tech. Report FSU-CSIT-04-13, School of Computational Science, Florida State University, July 2004, www.csit.fsu.edu/∼absil/Publi/RTR.htm.

[ABG04b] ———, *Trust-region methods on Riemannian manifolds with applications in numerical linear algebra*, Proceedings of the 16th International Symposium on Mathematical Theory of Networks and Systems (MTNS2004), Leuven, Belgium, 5-9 July 2004, 2004.

[AMSV02] P.-A. Absil, R. Mahony, R. Sepulchre, and P. Van Dooren, *A Grassmann-Rayleigh quotient iteration for computing invariant subspaces*, SIAM Review **44** (2002), no. 1, 57–73.

[ASVM04] P.-A. Absil, R. Sepulchre, P. Van Dooren, and R. Mahony, *Cubically convergent iterations for invariant subspace computation*, SIAM J. Matrix. Anal. Appl. **26** (2004), no. 1, 70–96.

[BDDR00] Zhaojun Bai, James Demmel, Jack Dongarra, and Axel Ruhe (eds.), *Templates for the solution of algebraic eigenvalue problems*, Software, Environments, and Tools, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2000, A practical guide. MR 2001k:65003

[BF66] W. W. Bradbury and R. Fletcher, *New iterative methods for solution of the eigenproblem*, Numer. Math. **9** (1966), 259–267.

[BGP97] Luca Bergamaschi, Giuseppe Gambolati, and Giorgio Pini, *Asymptotic convergence of conjugate gradient methods for the partial symmetric eigenproblem*, Numer. Linear Algebra Appl. **4** (1997), no. 2, 69–84. MR 98b:65040

[BS89] S. Batterson and J. Smillie, *The dynamics of Rayleigh quotient iteration*, SIAM J. Numer. Anal. **26** (1989), no. 3, 624–636.

[Căt04] Emil Cătinaş, *The inexact, inexact perturbed and quasi-Newton methods are equivalent models*, Math. Comp. (2004), electronically published on March 23, 2004.

[CGT00] A. R. Conn, N. I. M. Gould, and Ph. L. Toint, *Trust-region methods*, MPS/SIAM Series on Optimization, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, and Mathematical Programming Society (MPS), Philadelphia, PA, 2000.

[FSvdV98] Diederik R. Fokkema, Gerard L. G. Sleijpen, and Henk A. van der Vorst, *Jacobi-Davidson style QR and QZ algorithms for the reduction of matrix pencils*, SIAM J. Sci. Comput. **20** (1998), no. 1, 94–125 (electronic). MR 99e:65061

[GSF92] Giuseppe Gambolati, Flavio Sartoretto, and Paolo Florian, *An orthogonal accelerated deflation technique for large symmetric eigenproblems*, Comput. Methods Appl. Mech. Engrg. **94** (1992), no. 1, 13–23. MR 92i:65072

[GV96] G. H. Golub and C. F. Van Loan, *Matrix computations, third edition*, Johns Hopkins Studies in the Mathematical Sciences, Johns Hopkins University Press, 1996.

[GY00] Gene H. Golub and Qiang Ye, *Inexact inverse iteration for generalized eigenvalue problems*, BIT **40** (2000), no. 4, 671–684. MR 2001j:65063

[GY02] G. H. Golub and Q. Ye, *An inverse free preconditioned Krylov subspace method for symmetric generalized eigenvalue problems*, SIAM J. Sci. Comput. **24** (2002), no. 1, 312–334 (electronic).

[KN03] Andrew V. Knyazev and Klaus Neymeyr, *A geometric theory for preconditioned inverse iteration. III. A short and sharp convergence estimate for generalized eigenvalue problems*, Linear Algebra Appl. **358** (2003), 95–114, Special issue on accurate solution of eigenvalue problems (Hagen, 2000). MR 2004c:65040

[Not02] Y. Notay, *Combination of Jacobi-Davidson and conjugate gradients for the partial symmetric eigenproblem*, Numer. Linear Algebra Appl. **9** (2002), no. 1, 21–44.

[Not03] ———, *Convergence analysis of inexact Rayleigh quotient iteration*, SIAM J. Matrix Anal. Appl. **24** (2003), no. 1, 627–644.

[NW99] J. Nocedal and S. J. Wright, *Numerical optimization*, Springer Series in Operations Research, Springer-Verlag, New York, 1999.

[OJS90] J. Olsen, P. Jørgensen, and J. Simons, *Passing the one-billion limit in full configuration-interaction (fci) calculations*, Chemical Physics Letters **169** (1990), 463–472.

[Par80] B. N. Parlett, *The symmetric eigenvalue problem*, Prentice-Hall, Inc., Englewood Cliffs, N.J. 07632, 1980, republished by SIAM, Philadelphia, 1998.

[Rut70] H. R. Rutishauser, *Simultaneous iteration method for symmetric matrices*, Numerische Mathematik **16** (1970), 205–223.

[Saa92] Y. Saad, *Numerical methods for large eigenvalue problems*, Manchester University Press, Manchester, 1992.

[SBFvdV96] Gerard L. G. Sleijpen, Albert G. L. Booten, Diederik R. Fokkema, and Henk A. van der Vorst, *Jacobi-Davidson type methods for generalized eigenproblems and polynomial eigenproblems*, BIT **36** (1996), no. 3, 595–633, International Linear Algebra Year (Toulouse, 1995). MR 97k:65095

[SE02] Valeria Simoncini and Lars Eldén, *Inexact Rayleigh quotient-type methods for eigenvalue computations*, BIT **42** (2002), no. 1, 159–182. MR 2003d:65033

[Shu86] M. Shub, *Some remarks on dynamical systems and numerical analysis*, Proc. VII ELAM. (L. Lara-Carrero and J. Lewowicz, eds.), Equinoccio, U. Simón Bolívar, Caracas, 1986, pp. 69–92.

[Smi94] S. T. Smith, *Optimization techniques on Riemannian manifolds*, Hamiltonian and gradient flows, algorithms and control, Fields Inst. Commun., vol. 3, Amer. Math. Soc., Providence, RI, 1994, pp. 113–136.

[SP99] P. Smit and M. H. C. Paardekooper, *The effects of inexact solvers in algorithms for symmetric eigenvalue problems*, Linear Algebra Appl. **287** (1999), no. 1-3, 337–357, Special issue celebrating the 60th birthday of Ludwig Elsner. MR 2000a:65049

[SS98] Andreas Stathopoulos and Yousef Saad, *Restarting techniques for the (Jacobi-)Davidson symmetric eigenvalue methods*, Electron. Trans. Numer. Anal. **7** (1998), 163–181 (electronic), Large scale eigenvalue problems (Argonne, IL, 1997). MR 99j:65062

[ST00] A. Sameh and Z. Tong, *The trace minimization method for the symmetric generlized eigenvalue problem*, J. Comput. Appl. Math. **123** (2000), 155–175.

[Ste83] T. Steihaug, *The conjugate gradient method and trust regions in large scale optimization*, SIAM J. Numer. Anal. **20** (1983), 626–637.

[Ste01] G. W. Stewart, *Matrix algorithms, vol II: Eigensystems*, Society for Industrial and Applied Mathematics, Philadelphia, 2001.

[SvdVM98] G. L. Sleijpen, H. A. van der Vorst, and E. Meijerink, *Efficient expansion of subspaces in the jacobi-davidson method for standard and generalized eigenproblems*, Electron. Trans. Numer. Anal. **7** (1998), 75–89.

[SW82] A. H. Sameh and J. A. Wisniewski, *A trace minimization algorithm for the generalized eigenvalue problem*, SIAM J. Numer. Anal. **19** (1982), no. 6, 1243–1259.

[Tap74] R. A. Tapia, *Newton's method for optimization problems with equality constraints*, SIAM J. Numer. Anal. **11** (1974), 874–886. MR 51 #9488

[Toi81] Ph. L. Toint, *Towards an efficient sparsity exploiting Newton method for minimization*, Sparse Matrices and Their Uses (I. S. Duff, ed.), Academic Press, London, 1981, pp. 57–88.

[vdE02] Jasper van den Eshof, *The convergence of Jacobi-Davidson iterations for Hermitian eigenproblems*, Numer. Linear Algebra Appl. **9** (2002), no. 2, 163–179. MR 2003b:65036

[WSS98] Kesheng Wu, Yousef Saad, and Andreas Stathopoulos, *Inexact Newton preconditioning techniques for large symmetric eigenvalue problems*, Electron. Trans. Numer. Anal. **7** (1998), 202–214 (electronic), Large scale eigenvalue problems (Argonne, IL, 1997). MR 99i:65042