CUBICALLY CONVERGENT ITERATIONS FOR INVARIANT SUBSPACE COMPUTATION*

P.-A. ABSIL†, R. SEPULCHRE‡, P. VAN DOOREN§, AND R. MAHONY¶

Abstract. We propose a Newton-like iteration that evolves on the set of fixed dimensional subspaces of $\mathbb{R}^n$ and converges locally cubically to the invariant subspaces of a symmetric matrix. This iteration is compared in terms of numerical cost and global behaviour with three other methods that display the same property of cubic convergence. Moreover, we consider heuristics that greatly improve the global behaviour of the iterations.

Key words. Invariant subspace, Grassmann manifold, cubic convergence, symmetric eigenproblem, inverse iteration, Rayleigh quotient, Newton method, global convergence

AMS subject classifications. 65F15

1. Introduction. The problem of computing a p-dimensional eigenspace (i.e. invariant subspace) of an n-by-n matrix $A = A^T$ is ubiquitous in applied mathematics, with applications in control theory, pattern recognition, data compression and coding, antenna array processing, and a multitude of other domains.

Several methods for subspace estimation were proposed in the late seventies and early eighties. Demmel [Dem87] provides a joint analysis of three of the early methods that refine initial estimates of arbitrary p-dimensional eigenspaces of a (possibly nonsymmetric) $n \times n$ data matrix A. The early methods depend on the various numerical solutions of a common Riccati equation. These methods converge at best quadratically (Chatelin's Newton-based method [Cha84]) even when A is symmetric and involve the solution of a Sylvester equation at each iteration step. Moreover, the iterations defined depend on a choice of normalization condition used to generate the Riccati equation as well as the present iterative estimate of the eigenspace. More recently, iterations have been proposed that operate "intrinsically" on the Grassmann manifold, the set of p-planes in $\mathbb{R}^n$. Watkins and Elsner [WE91] have studied a multishifted QR algorithm that, as we will show, conceals a Grassmannian generalization of the Rayleigh quotient iteration (RQI). Edelman, Arias and Smith [EAS98] derived a Newton iteration directly on the Grassmann manifold to find critical points of a generalized Rayleigh quotient. A practical implementation of this method was investigated by Lundström and Eldén [LE02]. In a recent paper [AMSV02], the authors proposed a generalization of the RQI posed directly on the Grassmann manifold where scalar shifts are replaced by a matrix shift. All these algorithms are intrinsically defined on the Grassmann manifold (i.e. the next iterate only depends on A and the current iterate) and converge locally cubically to the isolated p-dimensional eigenspaces of $A = A^T$.

*An abridged version of this article will appear in the proceedings of the 42nd IEEE Conference on Decision and Control.

†School of Computational Science and Information Technology, Florida State University, Tallahassee FL 32306-4120. Part of this work was done while the author was a Research Fellow with the Belgian National Fund for Scientific Research (Aspirant du F.N.R.S.) at the University of Liège (`www.montefiore.ulg.ac.be/~absil`).

‡Department of Electrical Engineering and Computer Science, Université de Liège, Institut Montefiore (B28), Grande Traverse 10, 4000 Liège, Belgium (`www.montefiore.ulg.ac.be/systems`).

§Department of Mathematical Engineering, Université Catholique de Louvain, Bâtiment Euler (A.119), Avenue Georges Lemaître 4, 1348 Louvain-la-Neuve, Belgium.

¶Department of Engineering, Australian National University, ACT, 0200, Australia.

In the present paper, we compare the three recently proposed cubically convergent iterations [WE91, EAS98, LE02, AMSV02] and propose a fourth cubically convergent method inspired by the multi-homogeneous Newton methods considered by Dedieu and Shub [DS00]. The first goal of this paper is to compare the four iterations in terms of numerical cost and global behaviour. The global behaviour of these iterations is of particular interest as existing analytical results focus on the local convergence rates. In the case where $p = 1$ and only a single eigenvector is computed the three recently proposed methods degenerate to the same iteration, the classical Rayleigh quotient iteration, for which the global behaviour is well understood [PK69, Par80, BS89, PS95]. In contrast, almost no global analysis has been undertaken for the various iterations when $p > 1$. In this paper, we show that although the local performance of the methods is comparable, the global performance differs appreciably. In particular, we study for each method how the shape of the basin of attraction of an eigenspace deteriorates when some eigenvalues of A are clustered.

The second goal of this paper is to propose modifications to the methods that improve the global performance of the iterations without compromising the local performance. The purpose of the modifications is to ensure that each given eigenspace is surrounded by a large basin of attraction. This guarantees that the iteration converges to the targeted eigenspace even when started rather far away from it. For the Grassmannian Rayleigh quotient iteration of [AMSV02] we propose a simple threshold on the distance between successive iterates that improves the shape of the basins of attraction. For the two Newton based methods, we introduce a deformation parameter τ that achieves a continuous transition between the original iteration and a gradient flow with large basins of attraction. This deformation technique is related to line search methods and trust region methods in optimization. We propose a simple choice for τ that dramatically enlarges the basins of attraction around the attractors while preserving cubic convergence. In the case of the new Newton-like iteration proposed in this paper, the resulting algorithm (Algorithm 5.2) displays an excellent global behaviour, combined with a cubic rate of convergence and a numerical cost of $O(np^2)$ flops per iteration when A is suitably condensed.

This paper is organized as follows. After a short review of subspaces, eigenspaces and their representations (Section 2), we state four cubically convergent iterative algorithms for eigenspace computation (Section 3). These iterations are compared in terms of numerical cost and global behaviour in Section 4. In Section 5, we propose ways of improving the global behaviour of the iterations. The main results are summarized in the concluding Section 6.

2. Subspaces and eigenspaces. In the present section, we introduce concepts and notation pertaining to subspaces and eigenspaces.

Unless otherwise stated, all scalars, vectors and matrices are real. The superscript ^{T} denotes the transpose. Following conventions in [HM94], we use $\mathrm{Grass}(p, n)$ to denote the *Grassmann manifold* of the p-dimensional subspaces of $\mathbb{R}^n$, $\mathbb{RP}^{n-1} = \mathrm{Grass}(1, n)$ to denote the real projective space, and $\mathrm{ST}(p, n)$ to denote the *noncompact Stiefel manifold*, i.e. the set of n-by-p matrices with full rank. The columns space of $Y \in \mathrm{ST}(p, n)$ is denoted by $\mathrm{span}(Y)$. The "span" mapping is an application on $\mathrm{ST}(p, n)$ onto $\mathrm{Grass}(p, n)$ that is nowhere invertible. Given a matrix Y in $\mathrm{ST}(p, n)$, the set of matrix representations of the subspace $\mathrm{span}(Y)$ is

$$\mathrm{span}^{-1}(\mathrm{span}(Y)) = Y\,\mathrm{GL}_p := \{YM : M \in \mathrm{GL}_p\}$$

where GL_p denotes the set of p-by-p invertible matrices. This identifies $\mathrm{Grass}(p, n)$

with $\mathrm{ST}(p,n)/\mathrm{GL}_p := \{Y\,\mathrm{GL}_p : Y \in \mathrm{ST}(p,n)\}$. More details on the Grassmann manifold and matrix representations can be found in [FGP94, AMS02, Abs03].

Let A be an $n \times n$ matrix. Let $\mathcal{X}$ be a p-dimensional subspace of $\mathbb{R}^n$ and let $Q = [X|X_\perp]$ be an orthogonal $n \times n$ matrix such that X spans $\mathcal{X}$. Then Q^TAQ may be partitioned in the form $Q^TAQ = \left(\begin{smallmatrix} A_{11} & A_{12} \\ A_{21} & A_{22}\end{smallmatrix}\right)$ where $A_{11} \in \mathbb{R}^{p\times p}$. The subspace $\mathcal{X}$ is an *eigenspace* (i.e. invariant subspace) of A if and only if $A_{21} = 0$. By *spectrum* of $\mathcal{X}$, we mean the set of eigenvalues of A_{11}. The *external gap* of the eigenspace $\mathcal{X}$ of A is the shortest distance between the eigenvalues of A_{11} and the eigenvalues of A_{22}. The *internal gap* of $\mathcal{X}$ is the shortest distance between two eigenvalues of A_{11}. We say that $\mathcal{X}$ is a *nondefective* eigenspace of A if A_{11} is nondefective. The eigenspace $\mathcal{X}$ is termed *spectral* [RR02] if A_{11} and A_{22} have no eigenvalue in common (i.e. nonvanishing external gap). When $A = A^T$, an eigenspace is spectral if and only if it is *isolated*, i.e. there exists a ball in $\mathrm{Grass}(p,n)$ centered on $\mathcal{V}$ that does not contain any eigenspace of A other than $\mathcal{V}$. The span of a full-rank $n \times p$ matrix Y is an eigenspace of A if and only if there exists a matrix L such that $AY = YL$, in which case Y is called an *eigenbasis* and L the corresponding *eigenblock* [JS01].

From now on, we assume that $A = A^T$ unless otherwise specified.

3. Four iterations for eigenspace computation. In this section, we define four iterations that evolve on the Grassmann manifold of p-planes in $\mathbb{R}^n$ and converge locally cubically to the spectral eigenspaces of a symmetric $n \times n$ matrix A.

3.1. Shifted inverse iterations. Inverse iteration is a widely used method for computing eigenvectors of A corresponding to selected eigenvalues for which an approximation is available [Ips97]. Let $\hat\lambda$ be an approximation to an eigenvalue of A. Inverse iteration generates a sequence of vectors x_k starting from an initial vector x_0 by solving the systems of linear equations

$$(A - \hat\lambda I)z = x_k \tag{3.1}$$

and usually normalizing the result $x_{k+1} := z/\|z\|$. From a theoretical point of view, the norm of x_k is irrelevant: the iteration (3.1) induces an iteration on the projective space, i.e. the set of one-dimensional subspaces of $\mathbb{R}^n$. Except in some nongeneric cases, the iteration converges to an eigenvector of A with an eigenvalue closest to $\hat\lambda$, and the rate of convergence is linear. However, a higher rate of convergence can be achieved by adapting $\hat\lambda$ "online" using the information given by the current iterate x_k. For $A = A^T$, the choice of the feedback law $\hat\lambda := \rho(x_k)$, where ρ denotes the Rayleigh quotient

$$\rho(y) := \tfrac{y^TAy}{y^Ty}, \tag{3.2}$$

yields the well known Rayleigh quotient iteration (RQI)

$$(A - \rho(x_k)I)z = x_k,\ x_{k+1} = z/\|z\|. \tag{3.3}$$

The fact that ρ provides a quadratic approximation of an eigenvalue around the corresponding eigenvector explains why the rate of convergence is lifted from linear to cubic [Par80, AMSV02].

In the present paper, we consider two ways of generalizing the RQI to the Grassmann manifold of p-planes in $\mathbb{R}^n$. The first possibility is to use multiple scalar shifts, where the shifts are the Ritz values computed from the current subspace.

ALGORITHM 3.1 (RSQR). *Iterate the mapping* $\mathrm{Grass}(p,n) \ni \mathcal{Y} \mapsto \mathcal{Y}_+ \in \mathrm{Grass}(p,n)$ *defined by*
1. Pick an orthonormal $n \times p$ *matrix* Y *that spans* $\mathcal{Y}$.
2. Solve for $Z \in \mathbb{R}^{n\times p}$ *the equation*

$$(A - \rho_1 I)\dots(A - \rho_p I)Z = Y \tag{3.4}$$

where $\rho_1, \dots, \rho_p$ *are the eigenvalues of* $Y^T AY$ *repeated according to their multiplicity.*
3. Define $\mathcal{Y}_+$ *as the span of* Z. We call this iteration *RSQR* because of its link with the generalized Rayleigh-Shifted QR algorithm studied in [WE91]. It comes as a corollary from the results of [WE91] that RSQR converges locally cubically to the spectral eigenspaces of $A = A^T$, as we now explain.

The RQI algorithm is related to the Rayleigh-shifted QR algorithm, as shown e.g. in the enlightening paper by Watkins [Wat82]. The QR algorithm on the matrix A with Rayleigh quotient shift can be written as a QR decomposition

$$(A - \sigma_k I)Q_k = Q_{k+1}R_{k+1} \tag{3.5}$$

where σ_k is the lower right element of $A_k = Q_k^T AQ_k$. Taking the inverse transpose of (3.5) yields, assuming $A = A^T$,

$$(A - \sigma_k I)^{-1}Q_k = Q_{k+1}R_{k+1}^{-T}, \tag{3.6}$$

where R_{k+1}^{-T} is now a lower triangular matrix. The last column of (3.6) yields

$$(A - \sigma_k I)^{-1}x_k = r_{k+1}^{-1}x_{k+1},$$

where x_k denotes the last column of Q_k and r_k denotes the lower right element of R_k. And this is RQI (3.3). In [WE91], Watkins and Elsner study a *generalized Rayleigh-quotient shift strategy* for the QR algorithm. It consists in replacing $(A - \sigma_k I)$ by $\mathcal{P}(A)$, where $\mathcal{P}(\lambda)$ is the characteristic polynomial of the $p \times p$ lower right submatrix of $Q_k^T AQ_k$. In this case, equation (3.5) becomes

$$\mathcal{P}(A)Q_k = Q_{k+1}R_{k+1} \tag{3.7}$$

or equivalently, taking the inverse transpose,

$$\mathcal{P}(A)^{-1}Q_k = Q_{k+1}R_{k+1}^{-T}$$

whose last p columns yield

$$\mathcal{P}(A)^{-1}X_k = X_{k+1}L_{k+1}.$$

Here X_k denotes the last p columns of Q_k and $\mathcal{P}(A) := (A - \rho_1 I)\dots(A - \rho_p I)$ where $\rho_1, \dots, \rho_p$ denote the eigenvalues of $X_k^T AX_k$. This iteration maps the span of X_k to the span of X_{k+1}, and this is the above-defined RSQR (Algorithm 3.1). The developments in [WE91] show that this iteration converges locally cubically to the spectral eigenspaces of A. That is, for each spectral eigenspace $\mathcal{V}$ of A, there exist a scalar c and a neighbourhood B such that $\mathrm{dist}(\mathcal{Y}_+, \mathcal{V}) \le c\,\mathrm{dist}(\mathcal{Y}, \mathcal{V})^3$ for all $\mathcal{Y}$ in B. The distance between two subspaces can be e.g. defined by the projection 2-norm $\mathrm{dist}_{p2}(\mathcal{Y}, \mathcal{V}) = \|P_\mathcal{Y} - P_\mathcal{V}\|_2$, where $P_\mathcal{Y}$ and $P_\mathcal{V}$ denote the orthogonal projectors onto $\mathcal{Y}$ and $\mathcal{V}$, respectively [GV96]. Any compatible definition, such as the geodesic distance on the Grassmann manifold, can be used [EAS98].

Another Grassmannian generalization of the RQI, which uses a matrix shift instead of multiple scalar shifts, has been proposed in [AMSV02]. This iteration has been called Grassmann-RQI (GRQI)[1].

ALGORITHM 3.2 (GRQI). *Iterate the mapping* $\mathrm{Grass}(p,n) \ni \mathcal{Y} \mapsto \mathcal{Y}_+ \in \mathrm{Grass}(p,n)$ *defined by*
1. Pick a basis $Y \in \mathbb{R}^{n\times p}$ *that spans* $\mathcal{Y}$.
2. Solve

$$\boldsymbol{T}_Y Z := AZ - Z \underbrace{(Y^TY)^{-1}Y^TAY}_{R_A(Y)} = Y \tag{3.8}$$

for $Z \in \mathbb{R}^{n\times p}$.
3. Define $\mathcal{Y}_{k+1}$ *as the span of* Z. The matrix $R_A(Y)$ can be interpreted as a block shift that reduces to the scalar Rayleigh quotient (3.2) in the case $p = 1$. The computations in Algorithm 3.2 are done in terms of $n\times p$ matrices, but they induce an iteration on the Grassmann manifold. That is, $\mathcal{Y}_{k+1}$ does not depend on the choice of the representative Y of $\mathcal{Y}_k$ chosen in (3.8). The GRQI method converges locally cubically to the spectral eigenspaces of A [Smi97, AMSV02].

Like the classical RQI mapping, which is ill-defined by (3.3) when $\rho(x)$ is an eigenvalue of A, the two iterations RSQR (Algorithm 3.1) and GRQI (Algorithm 3.2) are defined *almost* everywhere on $\mathrm{Grass}(p,n)$, i.e. there are points of singularity. In order to characterize these singularities, we introduce notations that will be used throughout the text. Let X denote an $n\times p$ orthonormal matrix (i.e. $X^TX = I$) that spans the current iterate, and let $[X|X_\perp]$ be an orthogonal $n\times n$ matrix. Define $A_{11} := X^TAX$, $A_{12} := X^TAX_\perp$, $A_{21} := X_\perp^TAX$, $A_{22} := X_\perp^TAX_\perp$. Let $\rho_1,\dots,\rho_p$ denote the eigenvalues of A_{11} enumerated with their multiplicity. Then the RSQR and GRQI methods map the span of X to the span of an $n\times p$ matrix

$$X_+ = ZM, \tag{3.9}$$

where M is any invertible $p\times p$ matrix chosen so that $X_+^TX_+ = I$, and Z verifies

$$RSQR : (A-\rho_1 I)\dots(A-\rho_p I)Z = X, \tag{3.10}$$
$$GRQI : AZ - ZA_{11} = X. \tag{3.11}$$

In RSQR, the matrices $(A-\rho_i I)$ are invertible if and only if the ρ_i's are not eigenvalues of A, in which case Z is well defined by the RSQR equation (3.10) and is full rank. In GRQI, a Sylvester equation (3.11) has to be solved. The solution Z exists and is unique if and only if the spectra of A and of A_{11} are disjoint. Indeed, rotating $X \mapsto XQ$ so that $A_{11} = \mathrm{diag}(\rho_1,\dots,\rho_p)$ decouples the Sylvester equation (3.11) into p linear systems of equations

$$(A-\rho_i)z_i = x_i$$

where x_i and z_i denote the ith column of the rotated X and Z respectively. So, the conditions for existence and uniqueness of Z are the same in both inverse iterations. An additional subtlety of GRQI is that the computed Z may *a priori* be rank deficient. However, numerical experiments suggest that if Z is the unique solution of the GRQI equation (3.11), then it is full rank (see [AH02] for details).

[1]During the final preperation of this manuscript, the authors became aware of an independent derivation of the GRQI method in [Smi97].

If the span of X is close to $\mathcal{V}$, then the eigenvalues of A_{11} are close to the eigenvalues of $A|_{\mathcal{V}}$, which are obviously eigenvalues of A. Therefore, equations (3.10) and (3.11) are intrinsically *ill-conditioned* when the span of X is close to an eigenspace $\mathcal{Y}$. This ill-conditioning is essential for the fast convergence of the shifted iterations, and does *not* mean that the span of the computed Z is ill-conditioned as a function of X. This fact was already emphasized in the case $p = 1$ by Peters and Wilkinson [PW79]. The proof of cubic convergence of RSQR and GRQI shows that the span of Z is well conditioned when the span of X is "sufficiently close" to the target eigenspace $\mathcal{V}$. We shall see later (Section 4.2) that the notion of "sufficiently close" depends on the structure of A.

3.2. Newton iterations. It comes as a direct consequence from the definitions in Section 2 that the p-dimensional eigenbases of A are the full-rank $n \times p$ solutions of the matrix equation

$$F(Y) := \Pi_{Y_\perp} AY = 0 \tag{3.12}$$

where $\Pi_{Y_\perp} := I - Y(Y^TY)^{-1}Y^T$ is the orthogonal projector onto the orthogonal complement of span(Y). This formulation of eigenbasis computation as a zero finding problem calls for the utilization of the Newton iteration (see e.g. [DS83, NW99]) in the Euclidean space $\mathbb{R}^{n\times p}$, which consists in solving the Newton equation

$$F(Y) + \mathrm{D}F(Y)[\Delta] = 0 \tag{3.13}$$

where $\mathrm{D}F(Y)[\Delta]$ denotes the directional derivative of F at Y in the direction of Δ, and performing the update

$$Y_+ = Y + \Delta. \tag{3.14}$$

However, the solutions of (3.12) are not isolated in $\mathbb{R}^{n\times p}$, namely, if Y is a solution, then all the elements of the equivalence class $Y\,\mathrm{GL}_p$ are solutions, too. In fact, since F is homogeneous of degree one, i.e. $F(YM) = F(Y)M$, the solution of the Newton equation (3.13), when unique, is $\Delta = -Y$. So any point Y is mapped to $Y_+ = 0$. This is clearly a solution of $F(Y) = 0$, but it spans the trivial zero-dimensional subspace.

A remedy consists in constraining Δ to belong to the *horizontal space*

$$H_Y := \{\Delta \in \mathbb{R}^{n\times p} : Y^T\Delta = 0\}, \tag{3.15}$$

orthogonal to the equivalence class $Y\,\mathrm{GL}_p$. With this constraint on Δ, the solutions Δ of $F(Y+\Delta) = 0$ become isolated. However, the Newton equation (3.13) has in general no solution Δ in H_Y, so the Newton equation (3.13) must be relaxed. We will consider two approaches. The first one consists in projecting the Newton equation (3.13) onto H_Y

$$\Pi_{Y_\perp}(F(Y) + \mathrm{D}F(Y)[\Delta]) = 0,\ Y^T\Delta = 0. \tag{3.16}$$

The second approach consists in solving the Newton equation (3.13) in the least squares sense, that is

$$\Delta = \arg\min_{Y^T\Delta=0} \|F(Y) + \mathrm{D}F(Y)[\Delta]\|^2. \tag{3.17}$$

In the remainder of the present section, we develop the ideas (3.16) and (3.17) and show how they relate to methods proposed in the literature.

Define a map $J_Y : H_Y \to H_Y$ by projecting the Fréchet derivative of F in a direction $\Delta \in H_Y$ back onto H_Y,

$$J_Y : H_Y \to H_Y : \Delta \mapsto \Pi_{Y_\perp} \mathrm{D}F(Y)[\Delta] = \Pi A \Pi \Delta - \Delta (Y^T Y)^{-1} Y^T A Y \tag{3.18}$$

where Π is a shorthand notation for $\Pi_{Y_\perp}$. Using this notation, equation (3.16) may be written

$$J_Y[\Delta] = -F(Y). \tag{3.19}$$

The Newton-Grassman (NG) algorithm is formally stated as follows:

ALGORITHM 3.3 (NG). *Iterate the mapping* $\mathrm{Grass}(p,n) \ni \mathcal{Y} \mapsto \mathcal{Y}_+ \in \mathrm{Grass}(p,n)$ *defined by*

1. Pick a basis $Y \in \mathbb{R}^{n\times p}$ *that spans* $\mathcal{Y}$ *and solve the equation*

$$\Pi A \Pi \Delta - \Delta (Y^T Y)^{-1} Y^T A Y = -\Pi A Y \tag{3.20}$$

under the constraint $Y^T \Delta = 0$*, where* $\Pi := I - Y(Y^T Y)^{-1} Y^T$.

2. Perform the update

$$\mathcal{Y}_+ = \mathrm{span}(Y + \Delta). \tag{3.21}$$

One checks that $\mathcal{Y}_+$ does not depend on the Y chosen in step 1. Indeed, if Y yields the solution Δ of (3.20), then YM produces the solution ΔM for any $M \in \mathrm{GL}_p$, and $\mathrm{span}(Y+\Delta) = \mathrm{span}((Y+\Delta)M)$.

Algorithm NG admits the following geometric interpretation, valid for arbitrary A. The Grassmann manifold, endowed with the essentially unique Riemannian metric invariant by the action of the group of rotations, is a Riemannian manifold. In [Smi94], Smith proposes a Newton iteration on abstract Riemannian manifolds. This iteration, applied on the Grassmann manifold in order to solve (3.12), yields the search direction Δ given by (3.20), where Δ is interpreted as an element of the tangent space $T_Y \mathrm{Grass}(p,n)$; see [AMS02] for details. The update (3.21) is a simplification of the Riemannian updating procedure

$$\mathcal{Y}_+ = \mathrm{Exp}_{\mathcal{Y}} \Delta \tag{3.22}$$

consisting in following geodesics on the Grassmann manifold. Assuming $A = A^T$, Algorithm NG —but with geodesic update (3.22) instead of (3.21)— is also obtained by applying the Riemannian Newton method on $\mathrm{Grass}(p,n)$ for finding a stationary point of a generalized Rayleigh quotient [EAS98].

Algorithm NG was previously proposed for the case $A = A^T$ in [LST98], where quadratic convergence (at least) was proven. In [AMS02], it is shown that for *arbitrary* A, NG with either geodesic update (3.22) or projected update (3.21) converges locally quadratically to the spectral p-dimensional eigenspaces of A. When $A = A^T$ (which is assumed to hold in the present paper) the rate of convergence of NG is shown to be *cubic*.

Now we turn to the least squares approach (3.17). As shown in Appendix A, the solution Δ of the minimization problem (3.17) verifies

$$J^T \circ J[\Delta] + \Pi A Y (Y^T Y)^{-1} Y^T A^T \Pi \Delta = -J^T[F(Y)], \tag{3.23}$$

where J^T denotes the adjoint of the operator J (3.18) defined with respect to the inner product $\langle \Omega_1, \Omega_2 \rangle_X = \mathrm{trace}((X^T X)^{-1} \Omega_1^T \Omega_2)$. Assuming $A = A^T$, the operator J is self-adjoint and we obtain the following algorithm.

ALGORITHM 3.4 (NH). *Iterate the mapping* $\mathrm{Grass}(p,n) \ni \mathcal{Y} \mapsto \mathcal{Y}_+ \in \mathrm{Grass}(p,n)$ *defined by*
1. Pick a basis $Y \in \mathbb{R}^{n\times p}$ *that spans* $\mathcal{Y}$ *and solve the equation*

$$\begin{aligned}\Pi A^2 \Pi\Delta - 2\Pi A\Pi\Delta(Y^TY)^{-1}Y^TAY + \Delta(Y^TY)^{-1}Y^TAY(Y^TY)^{-1}Y^TAY \\ = -\Pi A\Pi AY + \Pi AY(Y^TY)^{-1}Y^TAY \end{aligned} \quad (3.24)$$

for the unknown Δ *under the constraint* $Y^T\Delta = 0$.
2. Perform the update

$$\mathcal{Y}_+ = \mathrm{span}(Y + \Delta). \quad (3.25)$$

Here again, it is checked that $\mathcal{Y}_+$ does not depend on the Y chosen in step 1. This least squares approach can be interpreted as a matrix generalization of the homogeneous Newton method proposed by Dedieu and Shub [DS00].

Algorithm NH converges locally cubically to the spectral eigenspaces of A. This property can be deduced from the corresponding property in NG. Applying the operator J on the NG equation (3.19) yields

$$J^T \circ J[\Delta] = -J^T[F(Y)] \quad (3.26)$$

which only differs from the NH equation (3.23) by the term $\Pi AY(Y^TY)^{-1}Y^TA^T\Pi\Delta$. Since ΠAY is zero at the solution and smooth, the operators in the left-hand side of (3.26) and (3.19) differ only at the second order. Since the right-hand side is of first order, the discrepancy between the solutions Δ of the NH equation (3.24) and the NG equation (3.20) is cubic, whence cubic convergence of NG is preserved in NH.

Like the inverse iterations (Section 3.1), the two Newton methods NG (Algorithm 3.3) and NH (Algorithm 3.4) have points of singularity. Let us rewrite the key equations in a slightly more compact form, using the notations of Section 3.1. The two Newton iterations map the span of an orthonormal X to the span of

$$X_+ = (X + \Delta)M = (X + X_\perp H)M, \quad (3.27)$$

where M is chosen to orthonormalize X_+ (M can e.g. be obtained by a QR factorization), and Δ or H verify

$$\begin{aligned} NG: &\ \Pi A\Pi\Delta - \Delta A_{11} = -\Pi AX,\ X^T\Delta = 0, && (3.28)\\ &\text{or } A_{22}H - HA_{11} = -A_{21}. && (3.29)\\ NH: &\ \Pi A^2\Pi\Delta - 2\Pi A\Pi\Delta A_{11} + \Delta A_{11}^2 = -\Pi A\Pi AX + \Pi AXA_{11},\ X^T\Delta = 0 && (3.30)\\ &\text{or } (A_{21}A_{12} + A_{22}A_{22})H - 2A_{22}HA_{11} + HA_{11}^2 = -A_{22}A_{21} + A_{21}A_{11} && (3.31)\end{aligned}$$

The inverse iterations (RSQR and GRQI) and the Newton iterations (NG and NH) are built on very different principles. In the inverse iterations, a new basis Z appears directly as the solution of a linear system of equations that becomes more and more *ill-conditioned* (i.e. almost singular) as the iterate X approaches an eigenspace. In the Newton methods, a correction Δ, verifying the horizontality constraints, is computed and added to the current iterate X. It is thus not surprising that the two approaches involve different singularities. In NG (3.29), H exists and is unique if and only if the spectra of A_{22} and A_{11} are disjoint. Note the difference with inverse iterations: the matrix A is replaced by the *projected* matrix A_{22}. In NH (3.31), H

exists and is unique if and only if the eigenvalues of the quadratic eigenvalue problem $(A_{21}A_{12} + A_{22}A_{22} - 2A_{22}\lambda + \lambda^2 I)x \equiv (A_{21}A_{12} + (A_{22} - \lambda I)^2)x = 0$ are distinct from the eigenvalues of A_{11}, see (4.6). When the span of X is close to $\mathcal{V}$, the residual matrix A_{21} has small norm, and the Sylvester operator on the left-hand side of (3.29) and (3.31) is *well-conditioned*. Indeed, the eigenvalues of A_{22} are close to those of $A|_{\mathcal{V}_\perp}$, the eigenvalues of A_{11} are close to those of $A_{\mathcal{V}}$, and the spectra of $A|_{\mathcal{V}_\perp}$ and $A_{\mathcal{V}}$ are separated since by hypothesis $\mathcal{V}$ is a spectral eigenspace.

4. Comparison of methods. In the previous section, we have formulated four iterations —two shifted inverse iterations (RSQR and GRQI) and two Newton methods (NG and NH)— that evolve on the Grassmann manifold of p-planes in $\mathbb{R}^n$ and converge locally cubically to the spectral p-dimensional eigenspaces of a symmetric $n \times n$ matrix A. Surprisingly, and in spite of different underlying approaches, RSQR and GRQI coincide with NG in the particular case $p = 1$, as pointed out by several authors [Shu86, Smi94, ADM+02, MA03]. When $p > 1$, however, the four methods differ.

In the present section, we compare the iterations in terms of numerical cost and global behaviour. Low numerical cost and large basins of attraction are two desirable features for methods that compute invariant subspaces from a first estimate.

4.1. Practical implementation. Comparing the implementation of the four different techniques depends to a large extent on the structure of the matrix A. If we assume first that A is dense, then all four methods have a comparable complexity, namely $O(pn^3)$, which mainly accounts for the p matrix factorizations that each requires. The RSQR solves

$$RSQR : (A - \rho_1 I) \cdots (A - \rho_p I)Z = X, \tag{4.1}$$

which involves p symmetric matrices $(A - \rho_i I)$. In the case of the three other methods, the first thing to do is to reduce A_{11} to a diagonal form. This is cheap since A_{11} is a $p \times p$ matrix and p is in practical applications typically much smaller than n. Moreover, the diagonalization always exists since A_{11} is symmetric. This diagonalization decouples equations (3.11), (3.28) or (3.30) into p independent systems of linear equations of the form

$$
\begin{aligned}
GRQI : & \ (A - \rho_i I)z = x, && (4.2)\\
NG : & \ \Pi(A - \rho_i I)\Pi\delta = -\Pi A x, \ X^T\delta = 0 && (4.3)\\
& \ \text{or } (A_{22} - \rho I)h = -A_{21}e, && (4.4)\\
NH : & \ \Pi(A - \rho_i I)^2\Pi\delta = -g, \ X^T\delta = 0 && (4.5)\\
& \ \text{or } ((A_{21}A_{12} + A_{22}A_{22}) - 2\rho A_{22} + \rho^2 I)h = -(A_{22}A_{21} - A_{21}A_{11})e, && (4.6)
\end{aligned}
$$

where $e \in \mathbb{R}^p$ is the eigenvector defined by $A_{11}e = \rho_i e$ and $x := Xe$, $z := Ze$, $\delta := \Delta e$, $h := He$.

Clearly $O(pn^3)$ seems excessive since most eigenvalue solvers require only $O(n^3)$ floating point operations (i.e. flops). A significant improvement is obtained by proceeding in three phases as follows: (i) reduce the matrix A to a tridiagonal form in $O(n^3)$ flops, (ii) compute an eigenspace of the tridiagonal matrix, (iii) compute the corresponding eigenspace of the original A in $O(n^2p)$ flops. We now focus on the second phase and assume that A is already in tridiagonal form. For RSQR the solution of (4.1) requires now $O(np^2)$ flops, while for GRQI (4.2) this is $O(np)$; the subsequent reorthogonalization of Z requires $O(np^2)$ for both methods. For the Newton updates

NG and NH, we use an idea from [PW79] which shows that the direction of the solution z of (4.2) is also given by the direction of $x+\delta$ where

$$\begin{bmatrix} A-\rho_i I & x \\ x^T & 0 \end{bmatrix} \begin{bmatrix} \delta \\ m \end{bmatrix} = \begin{bmatrix} -Ax \\ 0 \end{bmatrix}.$$

In a similar fashion, one can rewrite the Newton methods NG and NH as $(n+p)\times(n+p)$ symmetric problems :

$$\begin{bmatrix} A-\rho_i I & X \\ X^T & 0 \end{bmatrix} \begin{bmatrix} \delta \\ m \end{bmatrix} = \begin{bmatrix} -Ax \\ 0 \end{bmatrix} \tag{4.7}$$

and

$$\begin{bmatrix} (A-\rho_i I)^2 & X \\ X^T & 0 \end{bmatrix} \begin{bmatrix} \delta \\ m \end{bmatrix} = \begin{bmatrix} -g \\ 0 \end{bmatrix}, \tag{4.8}$$

respectively, rather than solving the dense problems (4.4) and (4.6). When $(A-\rho_i I)$ is tridiagonal, (4.7) and (4.8) can be solved in $O(np^2)$ flops each. The LDL^T decomposition of $(A-\rho_i I)$ and the QR decomposition of $(A-\rho_i I)$ both require $O(n)$ flops. The above problems are then replaced by

$$\begin{bmatrix} LDL^T & X \\ X^T & 0 \end{bmatrix} \begin{bmatrix} \delta \\ m \end{bmatrix} = \begin{bmatrix} -Ax \\ 0 \end{bmatrix} \tag{4.9}$$

and

$$\begin{bmatrix} R^T R & X \\ X^T & 0 \end{bmatrix} \begin{bmatrix} \delta \\ m \end{bmatrix} = \begin{bmatrix} -g \\ 0 \end{bmatrix}, \tag{4.10}$$

respectively, where L has only two diagonals and R only three. Solving the systems (4.9) and (4.10) (possibly with iterative refinement to ensure stability) requires $O(np^2)$ flops each. For a tridiagonal matrix A, the complexity for all four methods is thus $O(np^2)$ in addition to the cost of phases (i) and (iii). We point out however that there exist very efficient numerical methods for computing all the eigenvectors of tridiagonal matrices such that the computed eigenvectors are orthogonal to working precision [DP03]. Moreover, the MRRR algorithm announced in [DP03] would compute p eigenvectors of a tridiagonal matrix with lower order of complexity, $O(np)$, than the one reported above.

If the matrix A is sparse or banded, say with bandwith $2q+1$, then the numerical cost per iterate of GRQI, NG or NH is $O(nq^2p)+O(np^2)$ assuming $p, q << n$. If the bandwidth is sufficiently narrow, namely $q^2 \approx p$, then the numerical cost remains $O(np^2)$. For RSQR, assuming that the linear system (4.1) is solved by Gauss elimination and back-substitution, the numerical cost per iterate is $O(nq^2p)+O(nqp^2)$, hence the complexity of the algorithm essentially increases by a factor q at most as long as $q \leq p$. Another possibility, explained in Section 3.1, is to implement RSQR as a multishift QR algorithm [BD89]. Chasing a $p\times p$ bulge down a tridiagonal matrix can be done with approximately n Householder reflections of dimension $p \times p$ and applying those to X will yield the solution Z of (32). The numerical cost is thus $O(np^2)$ but this implicit method has to be implemented with care [BBM02a, BBM02b] in order to work properly.

Finally, if the matrix A is very large but sparse, one could consider alternative sparse matrix techniques such as reordering methods that reduce the bandwidth of

A or even iterative methods. If an approximate solution is sought using an iterative solver, a stopping criterion has also to be chosen for the inner iteration. Computing the first iterates with high precision may be unnecessary [EW96]. Iterative solvers are considered for the case $p = 1$ in [SE02], including a comparison between the RQI equation (4.2) and the Newton equation (4.3).

4.2. Basins of attraction. The four subspace methods under investigation in this paper, i.e. the two inverse iterations RSQR (Algorithm 3.1) and GRQI (Algorithm 3.2) and the two Newton methods NG (Algorithm 3.3) and NH (Algorithm 3.4), display *local* cubic convergence to the spectral eigenspaces of the symmetric matrix A. By "local convergence", it is meant that around each p-dimensional eigenspace $\mathcal{V}$, there exists a ball B in the Grassmann manifold $\mathrm{Grass}(p, n)$ such that the iteration converges to $\mathcal{V}$ for all initial point in B. But nothing has been said yet about the size of these balls. This is however an important issue, since a large ball means that the iteration will converge to the target eigenspace even if the initial estimate is not very precise.

It has been shown for previously available methods that the basins of attraction are prone to deteriorate when some eigenvalues are clustered. Batterson and Smillie [BS89] have drawn the basins of attraction of the RQI for $n = 3$ and have shown that they deteriorate when two eigenvalues are clustered. The bounds involved in the convergence results of the methods analyzed in [Dem87] blow up when the external gap vanishes.

In the present section, we illustrate properties of the basins of attraction on three examples. The first two examples are low-dimensional problems ($n = 3$ and $p = 1, 2$) for which faithful two-dimensional pictures of the basins of attraction can be drawn (the dimension of $\mathrm{Grass}(1, 3)$ and $\mathrm{Grass}(2, 3)$ is two). The third example is a higher dimensional case. In these examples, the matrices A are chosen to illustrate the influence of the eigenvalue gaps on the basins of attraction.

In order to graphically represent basins of attraction, we take advantage of the following facts. Let $\mathcal{F}_A$ denote one of the four iteration mappings mentioned above. The mappings are invariant by orthogonal changes of coordinates, i.e. $Q\mathcal{F}_A(\mathcal{Y}) = \mathcal{F}_{QAQ^T}(Q\mathcal{Y})$ for all Q orthogonal. Therefore, we work without loss of generality in a coordinate system in which A is diagonal. Moreover, once A is diagonal, the mappings are invariant by multiplication by a sign matrix. To show this, note that sign matrices are orthogonal, replace Q above by a sign matrix S and use the relation $SAS = A$. Consequently, it is sufficient to represent the basins of attraction in the first orthant. The other orthants are deduced by symmetry. Note also that the matrices A, $-A$ and $A - \sigma I$ yield the same sequences of iterates for all σ.

Example 1 (Dependence on external gap). We consider the case $n = 3$ and $p = 1$ (iterates are one-dimensional subspaces of $\mathbb{R}^3$). Then the two inverse iterations (RSQR and GRQI) reduce to the RQI, which is equivalent to NG (see e.g. [Smi94]). Figure 4.1 represents the basins of attraction of the RQI for $A = \mathrm{diag}(1, 2 - \gamma, 2)$. On the left-hand side of the figure $\gamma = .2$, and on the right-hand side γ is reduced to 0.01 in order to illustrate the effect of a small eigenvalue gap. Figure 4.1 should be read as follows. Displayed is the simplex $\{x \in \mathbb{R}^n : x_1 + x_2 + x_3 = 1, x_i > 0\}$. The iterates —one-dimensional subspaces of $\mathbb{R}^3$— are represented by their intersections with the simplex. The three vertices correspond to the three eigendirections of A, and the corresponding eigenvalues are indicated. The three colors indicate the three basins of attraction. It is seen that the basin of attraction of the upper vertex shrinks as its external gap is reduced. The basins of attraction of NH are qualitatively similar

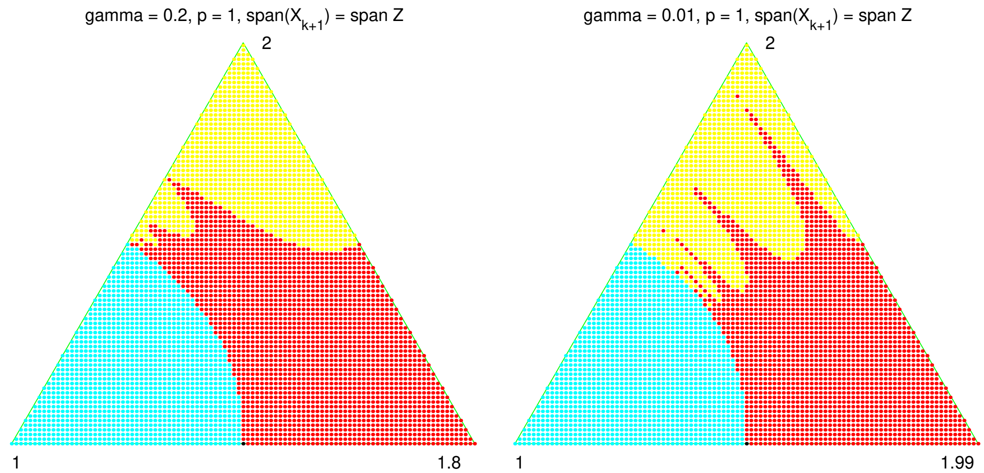

FIG. 4.1. *Basins of attraction for RSQR, GRQI and NG ($n = 3$, $p = 1$). The three vertices correspond to the three eigenspaces. A similar figure appears in [BS89]. This figure also applies to NG with $n = 3$, $p = 2$ (see Example 2 in Section 4.2).*

to the RQI-NG case. In conclusion, this simple example shows the dependence on external gap in all methods.

Example 2 (Dependence on internal gap in GRQI). We now investigate the case $n = 3$, $p = 2$ (iterates are 2-planes in $\mathbb{R}^3$) using the same two matrices A as above. Let us first consider the case of RSQR. Its basins of attraction are shown on Figure 4.2, where 2-planes are represented by the intersection of their normal vector with the simplex. The three vertices correspond to the three 2-dimensional eigenspaces of A. For example, the upper vertex corresponds to the minor eigenspace. On the right-hand plot of Figure 4.2 and the ones that follow, the eigenspace represented by the lower left vertex has a small internal gap and a large external gap, while the two other vertices correspond to eigenspaces with a large internal gap and a small external gap. Figure 4.2 shows that the basins of attraction for RSQR collapse when the *external* gap is small. On this low-dimensional example, a small *internal* gap does not affect the basin of attraction.

The basins of attraction of GRQI are shown on Figure 4.3, with the same conventions as for the RSQR plot. One notices a peak growing towards the eigenspace with small internal gap. The tip of the peak is very close to the eigenspace, but this can hardly be seen on the figure because the peak is very narrow. This shows that for GRQI the basins of attraction may deteriorate around the eigenspaces with small internal gap. We will explain this feature analytically in Section 4.3.

The Newton iteration NG displays the following *duality property*: If $\mathcal{X}^k$ is a sequence of iterates generated by NG, then $\mathcal{X}^k_\perp$ also forms a sequence of iterates of NG. To see this, let H verify the NG equation (3.29), note that $X_\perp - XH^T$ is orthogonal to $X + X_\perp H$, and $(-H^T)$ verifies $A_{11}(-H^T) - (-H^T)A_{22} = -A_{12}$ which is just the NG at the iterate $X_\perp$. By this duality property, the orthogonal complements of the iterates of NG ($p = 2$, $n = 3$) are one-dimensional iterates of NG ($p = 1$, $n = 3$). Representing 2-planes by the intersection of their normal vector with the simplex, the picture for NG in the case $p = 2$, $n = 3$ is the same as for the case $p = 1$, $n = 3$ illustrated on Figure 4.1.

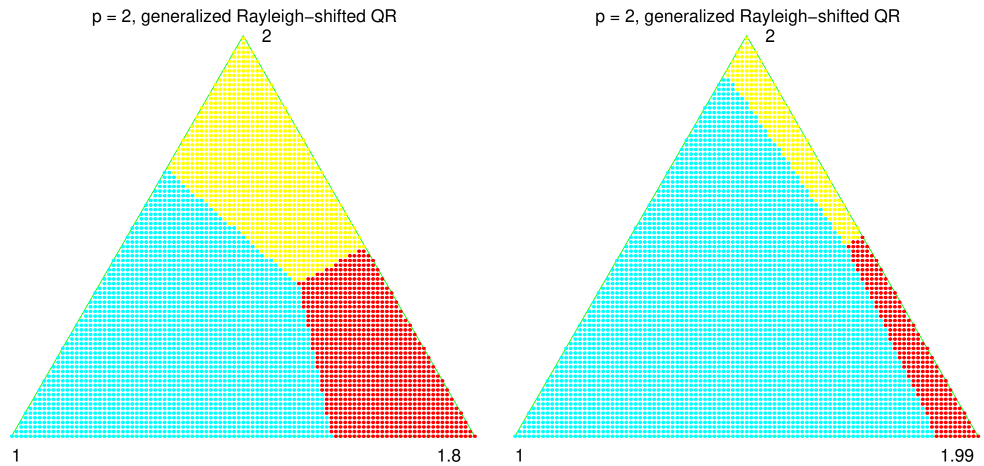

FIG. 4.2. *Basins of attraction for RSQR (case $p = 2$, $n = 3$). The elements of* $\mathrm{Grass}(2, 3)$ *(i.e. 2-planes in $\mathbb{R}^3$) are represented by the intersection of their normal vector with the simplex.*

The basins of attraction of the Newton iteration NH are shown on Figure 4.4, with the conventions explained above. The basins of attraction do not collapse in this low-dimensional example. One however should not conclude that everything goes well in higher dimensions, as we will show shortly.

In conclusion, this second example illustrates that the four methods are different when $p > 1$. It also reveals a dependence on internal gap occurring in GRQI.

Example 3 (Higher-dimensional case). The principal interest of the low-dimensional example studied above lies in the two dimensional representation of the basins of attraction. We now consider an example in $\mathrm{Grass}(3, 7)$, with $\dim \mathrm{Grass}(3, 7) = 12$, in order to further investigate the influence of the eigenvalue gaps on the basins of attraction. We use the matrix

$$A = \mathrm{diag}(1, 2, 2 + \gamma, 2 + 2\gamma, 3, 4, 5)$$

where γ is a small number (we choose $\gamma = 10^{-2}$). We select three different eigenspaces in order to illustrate the influence of internal and external gaps. In each case, we pick 10^4 initial points randomly at three given distances of the targeted eigenspace and we count how often the sequence of iterates fails to converge to the target. We declare that the sequence converges if $\mathrm{dist}(\mathcal{X}^k, \mathcal{V}_{target}) < 10^{-6}$ with $k = 100$, where dist denotes the largest principal angle between the two arguments. The condition is usually already verified for very small k (see Figure 4.5), but if the iteration is started close to the boundary of the basin of attraction then the condition may be verified after arbitrarily many steps.

Here are the results of our experiments:

(i) Convergence to the eigenspace $\mathcal{V}_{leli}$ with eigenvalues 1, 3 and 4. This eigenspace has a large external gap and a large internal gap. The ratios of sequences that failed to converge to the targeted eigenspace are shown on Figure 4.5a page 17. As predicted by the theory, the four methods (RSQR, GRQI, NG and NH) invariably converge to the targeted eigenspace when the initial error is small. When the initial error is large, the methods sometimes fail, and RSQR fails much more often than the other methods.

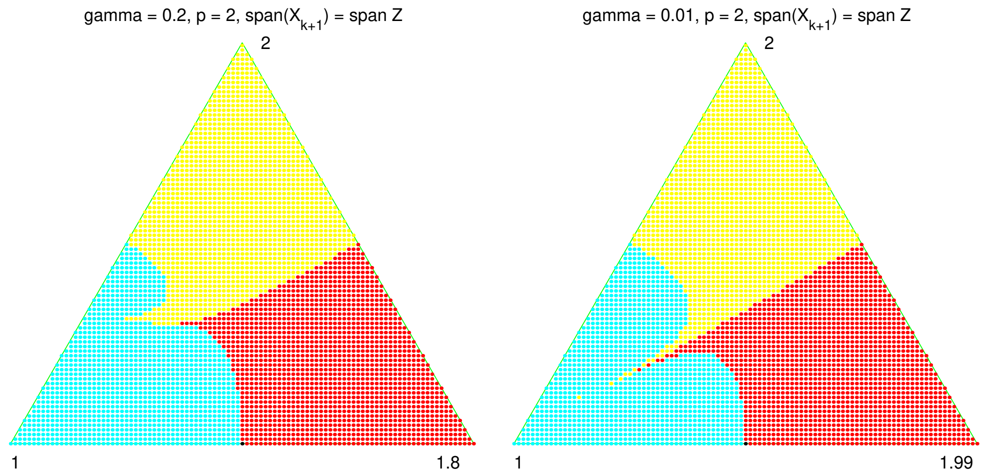

FIG. 4.3. *Basins of attraction for GRQI (case $p = 2$, $n = 3$).*

(ii) Convergence to $\mathcal{V}_{lesi}$ with eigenvalues 2, $2+\gamma$ and $2+2\gamma$ (Figure 4.5b). This illustrates the influence of a small internal gap. All methods except GRQI have a large basin of attraction around $\mathcal{V}_{lesi}$. This confirms the information obtained in the lower dimensional case that the basins of attraction of eigenspaces with small internal gap are collapsed in GRQI (see the peak observed for GRQI on Figure 4.3).
In a different experiment not reported on the figure, we also considered initial points situated at the distance $\frac{2}{3}\frac{\pi}{2}$ of $\mathcal{V}_{lesi}$. At such a distance, $\mathcal{V}_{lesi}$ is seldom the closest eigenspace, so convergence to $\mathcal{V}_{lesi}$ is not expected. And indeed, the iterates of GRQI, NG and NH seldom converged to $\mathcal{V}_{lesi}$ (probability of convergence around 0.02). However, the iterates of RSQR did very often converge to $\mathcal{V}_{lesi}$, with probability 0.95. This means that the basin of attraction of $\mathcal{V}_{lesi}$ has a very large area under RSQR. It suggests that the eigenspaces with clustered eigenvalues have an oversized basin of attraction under RSQR, to the detriment of the other basins of attraction.

(iii) Convergence to $\mathcal{V}_{seli}$ with eigenvalues 2, 3 and 4 (Figure 4.5c). This eigenspace has a large internal gap but a small external gap. The number of failures of RSQR is about 10 times worse than for the other methods, and all the methods sometimes fail to converge to $\mathcal{V}_{seli}$ unless they are started very close to it. This means that the basin of attraction of $\mathcal{V}_{seli}$ is small for each method. Therefore, one usually tries to avoid small external gaps by enlarging the targeted eigenspace to include whole clusters of eigenvalues. However, this approach requires a priori information on the eigenvalues. In Section 5 we will propose modified Newton methods that display large basins of attraction around eigenspaces like $\mathcal{V}_{seli}$.

4.3. Dependence on eigenvalue gaps. The numerical experiments reported in the previous section have led to the following observations. For the four methods under investigation, collapsed basins of attraction are observed around eigenspaces with small external eigenvalue gap. The basins of attraction of GRQI also deteriorate when the internal gap between eigenvalues is small. Under RSQR, the eigenspaces corresponding to clusters of eigenvalues have a particularly large basin of attraction. In the present section, we justify these observations analytically. As an aside, we obtain an alternative proof of cubic convergence for the Newton methods.

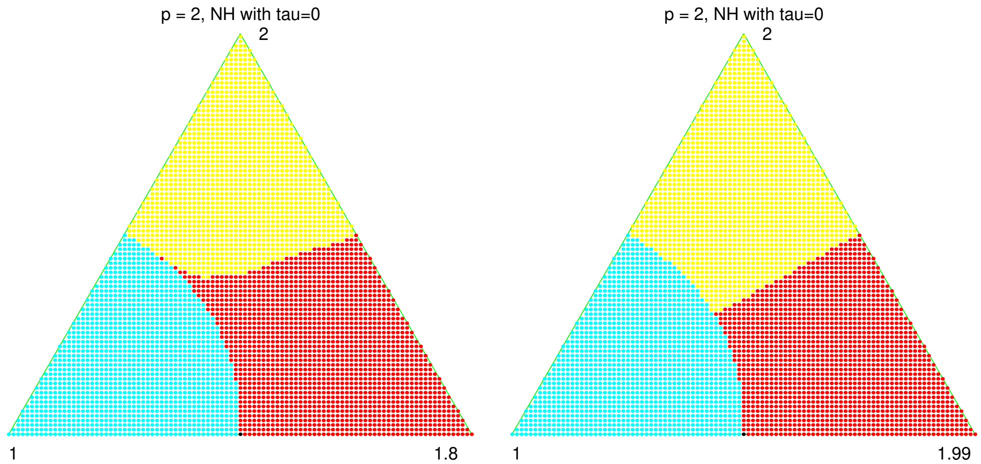

FIG. 4.4. *Basins of attraction for NH (case $p = 2$, $n = 3$). For the basins of attraction of NG ($p = 2$, $n = 3$), see Figure 4.1.*

RSQR. For simplicity of the argument, consider $A = \mathrm{diag}(1, 1+\gamma, 2)$ with γ small. Let $\mathcal{V}$ be an eigenspace of A with small external gap, e.g. $\mathcal{V} = \mathrm{span}(e_2, e_3)$ corresponding to the eigenvalues $1+\gamma$ and 2. We now exhibit a subspace $\mathcal{X}^0$ close to $\mathcal{V}$ that is mapped by RSQR to a subspace close to $\mathrm{span}(e_1, e_2)$. Let $\mathcal{X}^0 = \mathrm{span}(e_2 + \alpha e_1, e_3 + \beta e_1)$ with $|\alpha|, |\beta| << 1$. Then $\mathcal{X}^0$ is close to $\mathcal{V}$. The Ritz values of $(A, \mathcal{X}^0)$ are $\rho_1 = 1 + \gamma - \alpha^2\gamma + O(\alpha^4) + O(\alpha^2\beta^2)$ and $\rho_2 = 2 - \beta^2 + O(\beta^4) + O(\alpha^2\beta^2)$, and one obtains for the new iterate computed by RSQR from $\mathcal{X}^0$

$$\mathcal{X}^1 = \mathrm{span}\left((A - \rho_1 I)^{-1}(A - \rho_2 I)^{-1}\begin{bmatrix}\alpha & \beta\\ 1 & 0\\ 0 & 1\end{bmatrix}\right) \simeq \mathrm{span}\begin{bmatrix}-\alpha^3 & 1\\ 1 & 0\\ 0 & \gamma/\beta^3\end{bmatrix}.$$

If $\gamma << \beta^3$, then $\mathcal{X}^1$ is close to $\mathrm{span}(e_1, e_2)$. In other words, given a $\mathcal{X}^0$ that is close to $\mathcal{V}$ but does not contain e_3, if the cluster is sufficiently tight then $\mathcal{X}^1$ is close to the eigenspace corresponding to the cluster. This shows that the basin of attraction of $\mathrm{span}(e_1, e_2)$ contains points close to $\mathcal{V}$.

The behaviour we have just observed can be interpreted as a "cooperation" between clustered eigenvalues. If a Ritz value is a good shift for one eigenvalue in a cluster, it is also a good shift for all the eigenvalues in the cluster. Moreover, Ritz values of a randomly chosen subspace are more likely to be close to a cluster than to an isolated eigenvalue. This explains the oversized basins of attraction observed around eigenspaces with clustered eigenvalues.

GRQI. For GRQI, both a small external gap and a small internal gap may affect the quality of the basin of attraction of $\mathcal{V}$, as we now show.

GRQI maps the basis $Y = V + V_\perp K$ to the $Z = VZ_1 + V_\perp Z_2$ where

$$\Lambda_1 Z_1 + Z_1(I_p + K^T K)^{-1}(\Lambda_1 + K^T\Lambda_2 K) = I_p \tag{4.11}$$

$$\Lambda_2 Z_2 + Z_2(I_p + K^T K)^{-1}(\Lambda_1 + K^T\Lambda_2 K) = K, \tag{4.12}$$

see [AMSV02, Abs03]. Define $K_+ = Z_2 Z_1^{-1}$ so that the span of Z is the same as the span of $V + V_\perp K_+$.

Let us first consider equation (4.12). This is a Sylvester equation. It is well conditioned when K is small, therefore $Z_2 = O(K)$ due to the right-hand side. If the external gap of $\mathcal{V}$ is small, i.e. $\text{gap}[\Lambda_1, \Lambda_2]$ is small, then the Sylvester operator is arbitrarily ill-conditioned for small K, so Z_2 and K_+ may be large.

Now consider equation (4.11). This Sylvester equation is ill-conditioned when K is small. Since Λ_1 is diagonal, the lines of (4.11) are decoupled. Without loss of generality, let us consider the first line. Put $Z_1 = \begin{pmatrix} \zeta_{11} & \zeta_{12} \\ \zeta_{21} & \zeta_{22} \end{pmatrix}$, $\Lambda_1 = \begin{pmatrix} \sigma & \\ & \Sigma \end{pmatrix}$, $E = \Lambda_1 - (I_p + K^T K)^{-1}(\Lambda_1 + K^T \Lambda_2 K) = \begin{pmatrix} E_{11} & E_{12} \\ E_{21} & E_{22} \end{pmatrix}$. The first line of (4.11) yields (see [AMSV02])

$$\zeta_{11} = [E_{11} - E_{12}(\sigma I_p - \Sigma + E_{22})^{-1} E_{21}]^{-1} \tag{4.13}$$

$$\zeta_{12} = -\zeta_{11} E_{12}(\sigma I_p - \Sigma + E_{22})^{-1} \tag{4.14}$$

One obtains that $Z_1^{-1} = O(E) = O(K^2)$ and concludes that $K_+ = O(K^3)$, so the algorithm converges locally cubically [AMSV02]. However, if σ is close to an eigenvalue of Σ (i.e. if the internal gap is small), then $(\sigma I - \Sigma - E_{22})^{-1}$ is large for some small E (i.e. small K). This suggests that if the internal gap is small, there are some small Ks for which Z_1^{-1} is large, whence K_+ is large.

Newton methods. We show here that NG converges locally cubically to $\mathcal{V}$ and that the basin of attraction collapses when the external gap is small, but not when the internal gap is small. A similar development for NH leads to the same conclusions.

Let V be an orthonormal basis of the eigenspace $\mathcal{V}$ such that $V^T A V = \Lambda_1$ is diagonal, and let $V_\perp$ be an orthonormal basis of $\mathcal{V}_\perp$ such that $V_\perp^T A V_\perp = \Lambda_2$ is diagonal. The external gap of $\mathcal{V}$ is $\text{gap}[\Lambda_1, \Lambda_2]$. After some manipulations, one obtains that under NG (Algorithm 3.3 with projective update), the span of $V + V_\perp K$ is mapped to the span of $V + V_\perp K_+$ where K_+ verifies

$$K_+ = (K + (I + KK^T)^{-1}(L - K))(I - K^T(I + KK^T)^{-1}(L - K))^{-1}, \tag{4.15}$$

in which L solves

$$\begin{aligned}(\Lambda_2 + K\Lambda_1 K^T)(I + KK^T)^{-1} L - L\,(I + K^T K)^{-1}(\Lambda_1 + K^T \Lambda_2 K) \\ = [K\Lambda_1 K^T (I + KK^T)^{-1} + \Lambda_2((I + KK^T)^{-1} - I)]K \\ - K[(I + KK^T)^{-1} K^T \Lambda_2 K + ((I + K^T K)^{-1} - I)\Lambda_1].\end{aligned} \tag{4.16}$$

One deduces from (4.16) that $L = O(K^3)$ and then $K_+ = O(K^3)$, which means that the Newton iteration NG converges cubically; the reader is referred to [AMS02] for a detailed proof of cubic convergence. If the $\text{gap}[\Lambda_1, \Lambda_2]$ is small, then the Sylvester operator on the left-hand side of (4.16) becomes arbitrarily ill-conditioned for small K's (remember that the eigenvalues of a Sylvester operator are the differences between the eigenvalues of the two matrices involved in the equation [Ste73]), whence K_+ can be large even if K is small. This reasoning suggests that if the external gap of $\mathcal{V}$ is small, then some initial points close to $\mathcal{V}$ do not yield convergence to $\mathcal{V}$. On the other hand, the conditioning of the Sylvester operator in (4.16) is not affected by the internal gap of $\mathcal{V}$.

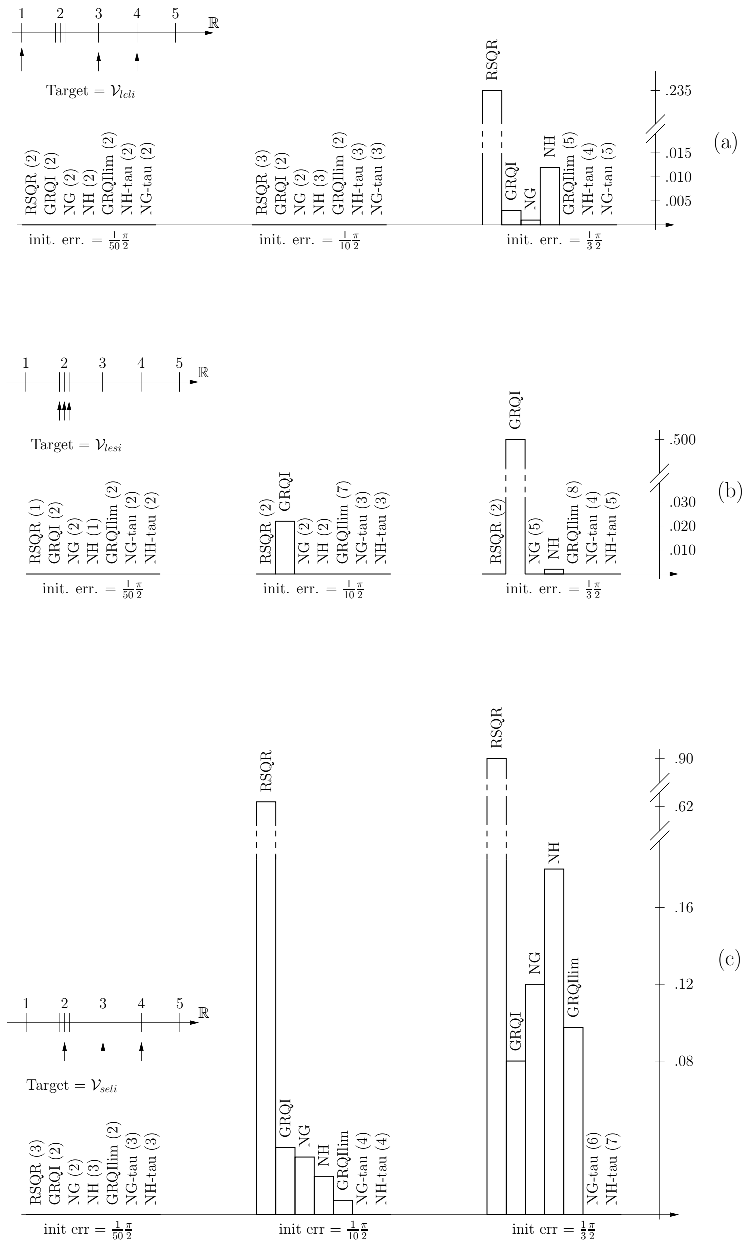

FIG. 4.5. *Ratio of sequences that failed to converge to the targeted eigenspace in Example 3 (Section 4). 'Init. err" gives the value of the largest principal angle between the initial subspace and the targeted eigenspace. Each ratio has been estimated using* 10^4 *randomly chosen starting points in each case. The absence of bar means that the sequence converged for all trials. We declare that a sequence converges if the largest principal angle between the 100th iterate and the target is smaller than* 10^{-6}. *The numbers between parentheses indicate the maximal number of iterates (evaluated on the* 10^4 *trials) necessery for the convergence condition to be satisfied.*

5. Improving the basins of attraction. Large basins of attraction are desirable as they ensure that the iteration will converge to the targeted eigenspace even if the initial subspace is a relatively poor estimate. The analysis in Section 4 has shown that a small external gap, and in the case of GRQI a small internal gap, produces a degradation of the basins of attraction of the iterations defined in Section 3. For this reason, we now discuss ways of improving the shape of the basins of attraction.

5.1. GRQI with limited variations. By experimenting with GRQI, we noticed that the sequences of iterates that diverge from the target eigenspace start with a big jump, i.e. the distance between the initial and second iterates is large. In an attempt to prevent this behaviour, we apply a threshold value on the distance between two successive iterates.

This can be implemented in the following way. Let $\mathcal{X}$ be the current iterate and let X be an orthonormal $n \times p$ matrix that spans $\mathcal{X}$. Let θ_{max} be a threshold value on the principal angles between $\mathcal{X}$ and $\mathcal{X}_+$. Compute Z, the solution of the GRQI equation (3.11). Orthonormalize Z, e.g. by a Gram-Schmidt process. Then, by the CS decomposition theorem [PW94, GV96], there exist orthogonal matrices U_1 and V_1 and an orthonormal matrix Y with $Y^T X = 0$ such that

$$ZV_1 = XU_1C + YS$$

where $C = \operatorname{diag}(\cos(\theta_1), \ldots, \cos(\theta_p))$, $S = \operatorname{diag}(\sin(\theta_1), \ldots, \sin(\theta_p))$, with $0 \leq \theta_1 \leq \ldots \leq \theta_p \leq \frac{\pi}{2}$. The θ_i's are the principal angles between $\operatorname{span}(X)$ and $\operatorname{span}(Z)$, and the columns of XU_1 and ZV_1 are the corresponding principal vectors. Define $\theta_i^{new} = \min\{\theta_i, \theta_{max}\}$. Then define $C^{new} = \operatorname{diag}(\cos(\theta_1^{new}), \ldots, \cos(\theta_p^{new}))$, $S^{new} = \operatorname{diag}(\sin(\theta_1^{new}), \ldots, \sin(\theta_p^{new}))$, and let the new iterate $\mathcal{X}_+$ be the span of $Z^{new} = XU_1C^{new} + YS^{new}$.

The matrix Z^{new} is obtained from the original Z in $O(np^2)$ flops by computing the singular value decomposition $X^T Z = U_1 C V_1^T$, then $S = \sin(\arccos C)$, and solving $YS = ZV_1 - XU_1C$. In fact, only the last columns of U_1, S and Y corresponding to the θ_i's larger than the threshold θ_{max} have to be computed; the other columns of ZV_1 are unmodified in Z^{new}.

We chose $\theta_{max} = \frac{\pi}{10}$ in numerical experiments. The basins of attraction of this modified GRQI are displayed on Figure 5.1 for the low-dimensional case ($n = 3$, $p = 2$) investigated in the previous section (Example 2). Compare Figure 4.3 (GRQI) and Figure 5.1: the peak has been removed. Experimental results for the higher-dimensional case (Example 3 in the previous section) are displayed on Figure 4.5 (see columns labelled "GRQIlim"). They illustrate that this heuristic effectively suppresses the problem of dependence on the internal gap.

Another (arguably more natural) modification of GRQI consists in taking $\mathcal{X}_+$ on the Grassmann geodesic [EAS98, AMS02] between $\mathcal{X}$ and $\operatorname{span}(Z)$ with $\theta_p(\mathcal{X}, \mathcal{X}_+) = \theta_{max}$. This amounts to defining $\theta_i^{new} = \lambda\theta_i$ with $\lambda = \frac{\theta_{max}}{\theta_p}$. However, the previously described technique works slightly better in experiments.

5.2. Modified Newton methods.

Deformation parameter τ. As explained in Section 3.2, the NG iteration attempts to find a p-plane $\mathcal{Y}$ such that each basis Y of $\mathcal{Y}$ verifies

$$F(Y) := \Pi_{Y_\perp} AY = 0 \tag{5.1}$$

where $\Pi_{Y_\perp} := I - Y(Y^TY)^{-1}Y^T$. Equation (5.1) holds if and only if $\mathcal{Y}$ is an invariant subspace of A.

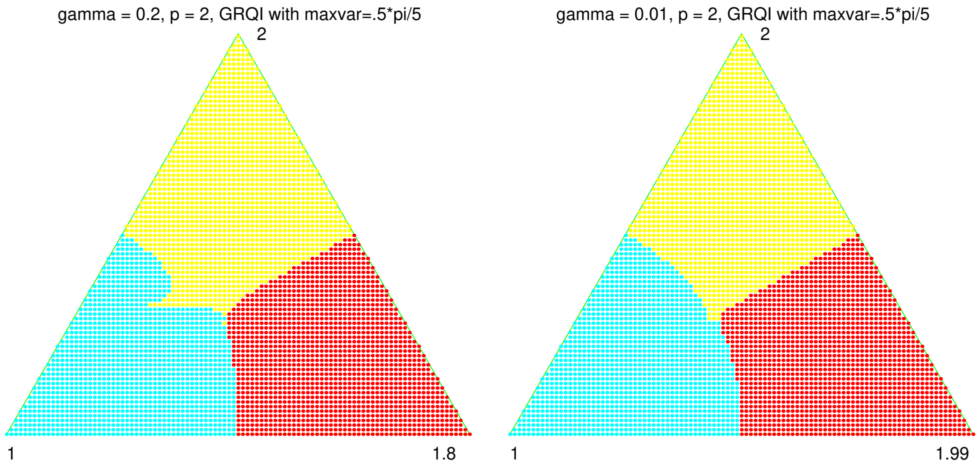

FIG. 5.1. *Basins of attraction for GRQI with limited steps (Section 5.1) in the case $p = 2$, $n = 3$. Compare with the original GRQI (Figure 4.3).*

Let us define a *cost function*

$$f(Y) := \frac{1}{2}\mathrm{trace}((Y^TY)^{-1}F(Y)^TF(Y)). \tag{5.2}$$

It is easily checked that $f(Y)$ depends only on the span of Y, and not on the basis Y itself [AMS02]. So, the cost function f defines a scalar field on the Grassmann manifold. This scalar field is zero at the eigenspaces of A and strictly positive everywhere else. An illustration of the level curves of f is shown on Figure 5.2. We stress that f reaches its minimum value (zero) at all the eigenspaces of A, and not only at an extremal eigenspace. This is a fundamental difference with the more familiar Rayleigh quotient.

Section 4 has shown that the basins of attraction of the two Newton methods (NG and NH) deteriorate in the presence of a small external gap. On the other hand, Figure 5.2 suggests that the basins of attraction of the steepest descent flow of the cost function f remain broad even when the eigenvalue gap shrinks. A numerical simulation of the steepest descent flow of f in Example 3 of Section 4 shows that the distance between each eigenspace and the boundary of its basin of attraction is large (greater than $\frac{1}{3}\frac{\pi}{2}$) in all cases.

This prompts us to follow the steepest descent of f when the solution is far away from a solution and use the Newton method in the neighbourhood of a solution. It is however difficult to decide when the commutation between the two behaviours should occur. If the Newton iteration takes over too soon, the basins of attraction may be collapsed. If the transition occurs late in the iterative process, then more steps will be necessary before obtaining a good approximation of the eigenspace. A remedy is to implement a smooth transition between the two behaviours by means of a deformation parameter, an idea which connects with trust region methods (see e.g. [DS83] or chap. 11 in [NW99]). We now show how this deformation approach works out in our case.

Let Y be an basis for the current subspace, let H_Y be the horizontal space defined as in (3.15), and let $J : H_Y \to H_Y : \Pi \mathrm{D}F(Y)[\Delta]$ be as in (3.18). The derivative of

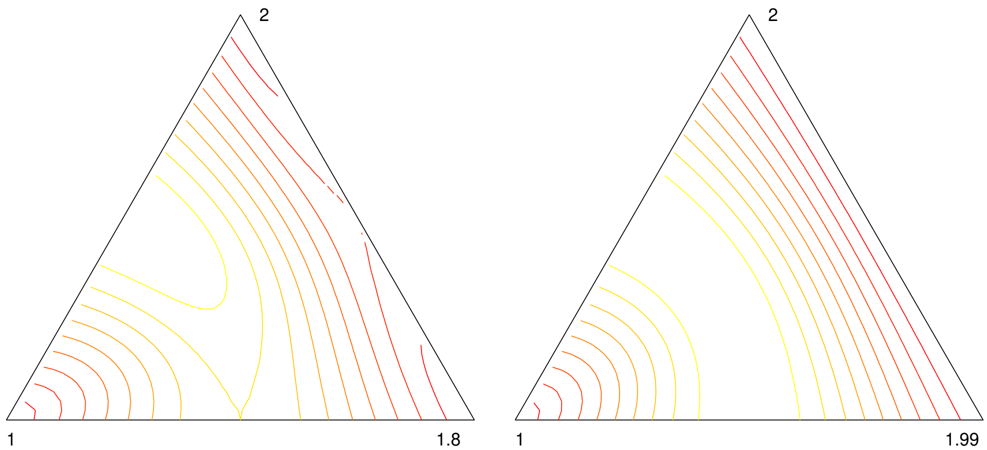

Fig. 5.2. *Level curves of the cost function f defined in* (5.2). *The cost function vanishes at the three eigenspaces (represented by the three vertices) and is strictly positive everywhere else. The gradient descent flow for f consists in following the steepest descent path of these level curves.*

the cost function f in the direction of Δ, with $Y^T\Delta = 0$, is given by

$$\begin{aligned} Df(Y)[\Delta] &= \mathrm{trace}((Y^TY)^{-1}F(Y)^TDF(Y)[\Delta]) \\ &= \mathrm{trace}((Y^TY)^{-1}F(Y)^TJ[\Delta]) \\ &= \mathrm{trace}((Y^TY)^{-1}(J^T[F(Y)])^T\Delta) \\ &= \mathrm{trace}(\Delta^TJ^T[F(Y)](Y^TY)^{-1}), \end{aligned}$$

where J^T denotes the adjoint of the operator J defined with respect to the inner product $\langle\Omega_1,\Omega_2\rangle_Y = \mathrm{trace}((Y^TY)^{-1}\Omega_1^T\Omega_2)$. Then a formula in [AMS02] directly yields

$$\mathrm{grad}\ f(Y) = J^T[F(Y)]. \tag{5.3}$$

On the other hand, the NG equation (3.20) reads $J[\Delta] = -F(Y)$, or equivalently

$$J^T \circ J[\Delta] = -J^T[F(Y)].$$

A continuous deformation between the gradient descent flow of f and the Newton method NG is thus given by

$$(J^T \circ J + \tau\,\mathrm{Id})[\Delta] = -J^T[F(Y)].$$

If τ is small, then Δ is close to the NG-vector given by the NG equation (3.28), and the iteration is close to pure NG. If τ is large, then the direction of Δ is close to the negative gradient of f, and the iteration is similar to a Euler integration of the gradient descent flow of f. Because we assume $A = A^T$, the operator J is self-adjoint and the modified NG algorithm can be expressed as follows.

Algorithm 5.1 (NG-tau). *Iterate the mapping $\mathcal{Y} \mapsto \mathcal{Y}_+$ defined by*
1. Pick an orthonormal basis Y that spans $\mathcal{Y}$ and solve the equation

$$\Pi A\Pi A\Pi\Delta + \Delta Y^TAYY^TAY - 2\Pi A\Pi\Delta Y^TAY + \tau\Delta = -(\Pi A\Pi AY - \Pi AYY^TAY), \tag{5.4}$$

where $\Pi := (I - YY^T)$*, under the constraint* $Y^T\Delta = 0$*.*
2. Perform the update $\mathcal{Y}_+ = \mathrm{span}(Y + \Delta)$*.*

We now introduce a τ deformation parameter in the NH iteration such that the limiting cases $\tau = 0$ and $\tau = \infty$ correspond to pure NH and gradient descent for f, respectively. This is easily done because the right-hand side of the NH equation (3.30) is precisely $-\,\mathrm{grad}\, f$ (compare (3.23) and (5.3)).

ALGORITHM 5.2 (NH-tau). *Iterate the mapping* $\mathcal{Y} \mapsto \mathcal{Y}_+$ *defined by*
1. Pick an orthonormal basis Y *that spans* $\mathcal{Y}$ *and solve the equation*

$$\Pi A^2 \Pi\Delta + \Delta Y^T AYY^T AY - 2\Pi A\Pi\Delta Y^T AY + \tau\Delta = -(\Pi A\Pi AY - \Pi AYY^T AY), \quad (5.5)$$

where $\Pi := (I - YY^T)$*, under the constraint* $Y^T\Delta = 0$*.*
2. Perform the update $\mathcal{Y}_+ = \mathrm{span}(Y + \Delta)$*.*

Note that the only difference between NG-tau and NH-tau is in the first term of (5.4) and (5.5).

Practical implementation. The major computational work in NG-tau (Algorithm 5.1) or NH-tau (Algorithm 5.2) is solving (5.4) or (5.5) for Δ. Like in the case of the original NG and NH iterations (see Section 4.1), the first thing to do is to diagonalize the small $p \times p$ matrix $A_{11} := Y^T AY$. This decouples equations (5.4) or (5.5) into p individual systems of linear equations of the form

$$((\Pi A\Pi - \rho_i I)^2 - \tau I)\delta = -g, \; Y^T\delta = 0 \quad (5.6)$$

$$\Pi((A - \rho_i I)^2 - \tau I)\Pi\delta = -g, \; Y^T\delta = 0 \quad (5.7)$$

for NG-tau and NH-tau, respectively. In the case of NH-tau (5.7), the projectors are outside the matrix, which allows for the utilization of the techniques described in Section 4.1. It is possible to obtain the Cholesky decomposition of $(A-\rho_i)^2 - \tau I = R_\tau^T R_\tau$ from that of $(A - \rho_i)^2 = R^T R$ in $O(n)$ flops when R has only three diagonals [Par80]. The algorithm NH-tau will thus again require $O(np^2)$ flops per iteration. In the case of NG-tau, in the absence of an efficient algorithm for solving (5.6), the cost for producing a new iterate involves $O(n^3)$ flops, even if A is tridiagonal. Thus, NH-tau has a serious advantage over NG-tau in terms of numerical cost.

Choosing the deformation parameter. There exist many strategies for tuning the τ parameter in order to improve the global behaviour of the algorithm while preserving the ultimate rate of convergence of the Newton method. In a line search approach, one selects τ so that the direction of K remains in a sector around the negative gradient of f and then perform a line search along the direction of the K computed from (5.4). Equation (5.4) is also helpful in trust region methods. A large τ corresponds to a small trust region, while $\tau = 0$ corresponds to a trust region that contains the exact next Newton iterate.

Classical strategies for choosing τ involve several parameters that the user can choose at his convenience [DS83, NW99]. In the present case, the very simple choice $\tau := f$ preserves the local cubic convergence and considerably enlarges the basins of attraction around the eigenspaces, both for NG-tau and NH-tau.

Local cubic convergence of NG-tau and NH-tau with $\tau = f$ is direct: τ is quadratic in the distance between the current iterate $\mathcal{Y}$ and the target eigenspace $\mathcal{V}$, while the right-hand side of (5.4) or (5.5) is linear in the distance. Consequently, the perturbation on the solution K of (5.4) induced by $\tau = f$ is cubic.

The global behaviour of NG-tau and NH-tau is illustrated on Figures 5.3 and 5.4 in our low-dimensional example utilized in Section 4.2. Comparison with Figure 4.1

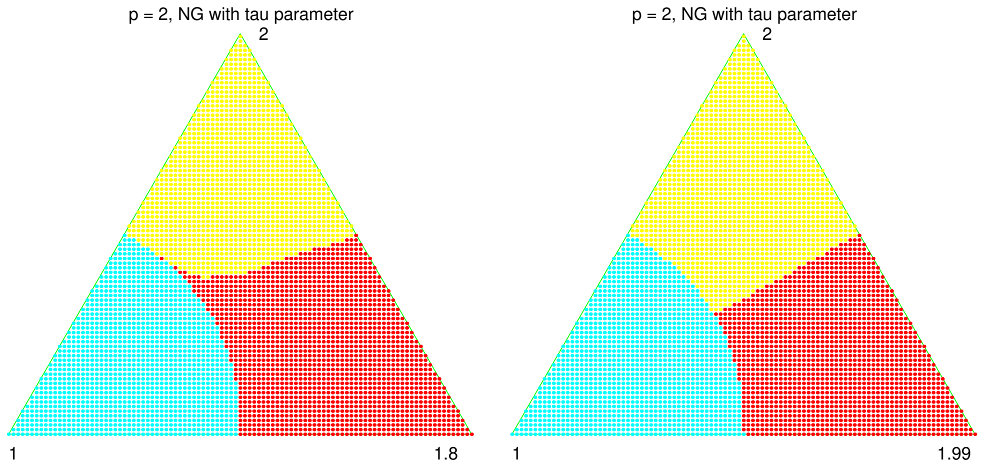

FIG. 5.3. *Attraction basins for NG with $\tau := f$ (5.4) in the case $p = 2$, $n = 3$. Compare with original NG on Figure 4.1. Local cubic convergence is preserved.*

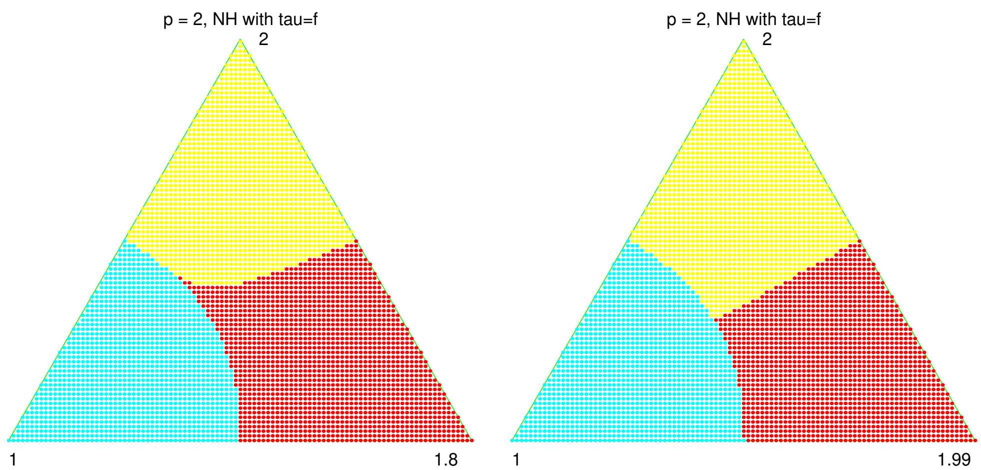

FIG. 5.4. *Attraction basins for NH-tau with $\tau := f$ (5.4) in the case $p = 2$, $n = 3$.*

shows that the basins of attraction have been considerably enlarged around the three eigenspaces. The improvement is even more spectacular in the larger dimensional case (Example 3 in Section 4.2), see Figure 4.5. Both NG-tau and NH-tau invariably converged to the targeted eigenspace. We had to choose the largest principal angle between the first iterate and the target greater than $\frac{1}{2.2}\frac{\pi}{2}$ in order to observe convergence to a wrong eigenspace.

Note that the balls centered on the eigenspaces of A overlap if their radius (measured in arc length on Grassmann [AMS02] or by means of the largest principal angle) is larger than $\pi/4$. So there is a geometrical limitation on the size of the basins of attraction. Our results show that in the NG-tau and NH-tau, the basins of attraction are so large that the geometrical limit is almost reached.

6. Conclusion. We have compared four iterative methods, i.e. RSQR (Algorithm 3.1), GRQI (Algorithm 3.2), NG (Algorithm 3.3) and NH (Algorithm 3.4), that operate on the set of p-dimensional subspaces of $\mathbb{R}^n$ and refine initial estimates of invariant subspaces of a symmetric n-by-n matrix A with cubic rate of convergence. Methods RSQR and GRQI are formulated as shifted inverse iterations. The former uses multiple scalar shifts while the latter involves a matrix shift. Algorithms NG and NH are derived from a Newton argument. Algorithm RSQR can be traced back to [PK69, PP73] and its proof of cubic local convergence is implicitly contained in [WE91]. GRQI is studied in [Smi97, AMSV02]. NG appears in [LST98, EAS98, LE02] and is connected to [Ste73, DMW83, Cha84, Dem87, Fat98, DF01]. Its local rate of convergence is studied in [AMS02]; see also Sections 3.2 and 4.3. To our knowledge, NH was never mentioned before in the literature.

We have shown that although these four iterations converge locally cubically to the spectral (i.e. isolated) eigenspaces of A, they appreciably differ in their global behaviour. The basin of attraction of an eigenspace $\mathcal{V}$ collapses when the eigenvalues of A relative to $\mathcal{V}$ are not well separated from the other eigenvalues of A. Moreover, in the case of GRQI, the basin of attraction of $\mathcal{V}$ also deteriorates if the eigenvalues relative to $\mathcal{V}$ are clustered. This dependence on eigenvalue gaps means that the sequence of iterates may diverge from $\mathcal{V}$ even if the initial point is a good approximation of $\mathcal{V}$.

For three of the methods, we have proposed ways of improving the shape of the basins of attraction. In the GRQI case, our numerical experiments suggest that a simple heuristic imposing a limitation on the distance between successive iterates removes the bad influence of clustered eigenvalues in the target eigenspace. In the Newton case, we have introduced a deformation parameter that achieves a continuous deformation between the pure Newton case (NG or NH) and the gradient descent flow of a cost function. Our experiments show that a simple choice of the deformation parameter spectacularly improves the shape of the basins of attraction while preserving the ultimate cubic convergence rate.

We also commented on the practical implementation of the various iterations. With the exception of the deformed NG iteration, a new iterate of each method can be computed in $O(np^2)$ flops when A has bandwidth $2q+1$ and $q = O(p^{1/2})$. When $q = 1$ there exist very efficient methods that compute all eigenvectors; see [DP03]. When A is sparse but not banded the computational cost of one iteration step will depend on the type of sparsity, but the complexity is essentially that of p sparse solves and therefore likely to be only linear in n.

In the Newton methods presented here, it is essential to compute the updates with high accuracy in order to take advantage of the cubic rate of convergence. Another approach consists in using acceleration techniques that exploit the useful information given by the previous updates in order to improve the current approximate solution. This allows for lower accuracy solves of the Newton equations, e.g. using iterative solvers; see [FSV98, Kny01] for more details. In the $p = 1$ case, this approach yields e.g. the celebrated Jacobi-Davidson method [SV96] for which the use of iterative solvers as inner solution process is well understood [Not02, Not03]. As an aside, the Jacobi-Davidson method is equivalent to RQI with $p = 1$ when the Newton equations are solved exactly (this rejoins our remark on the $p = 1$ case in Section 4). In the $p > 1$ case, we obtain a "block Jacobi-Davidson" that was touched upon in recent references [LE02, Bra03].

Among the algorithms considered here, our study suggests the NH algorithm with deformation parameter (Algorithm 5.2) as the method of choice for its remarkable combination of advantages: excellent global behaviour, cubic rate of convergence and low numerical cost $O(np^2)$ when A is suitably condensed.

7. Acknowledgements. We thank M. Petre for his contribution to the Matlab script that generates pictures of basins of attraction.

This paper presents research partially supported by the Belgian Programme on Inter-university Poles of Attraction, initiated by the Belgian State, Prime Minister's Office for Science, Technology and Culture. Part of this work was performed while the first author was a guest at the Mathematisches Institut der Universität Würzburg under a grant from the European Nonlinear Control Network. The hospitality of the members of the department is gratefully acknowledged. The work was completed while the first and second authors were visiting the department of Mechanical and Aerospace Engineering at Princeton University. The hospitality of the members of the department, especially Prof. N. Leonard, is gratefully acknowledged. The second author thanks N. Leonard and E. Sontag for partial financial support under US Air Force Grants F49620-01-1-0063 and F49620-01-1-0382.

Appendix A. Derivation of Algorithm NH.

In this section, we explain how the NH equation, i.e. (3.23) or (3.24), is derived from the minimization problem (3.17).

Let F be defined as in (3.12), $F(Y) := \Pi_{Y_\perp} AY$ and let H_Y denote the horizontal space (3.15), $H_Y := \{Y^T\Delta = 0\}$. Let $\mathbb{J}$ denote the operator $DF(Y)$ restricted to act on H_Y,

$$\mathbb{J}[\Delta] = \Pi A \Pi \Delta - \Delta(Y^TY)^{-1}Y^TAY - Y(Y^TY)^{-1}\Delta^T AY = J[\Delta] - Y(Y^TY)^{-1}\Delta^T AY$$

where J denotes the operator $\Pi DF(Y)$ defined in (3.18) restricted to act on H_Y. Let

$$m_Y(\Delta) := \frac{1}{2}\|F(Y) + \mathbb{J}[\Delta]\|^2 = \frac{1}{2}\mathrm{trace}((Y^TY)^{-1}(F(Y)+\mathbb{J}[\Delta])^T(F(Y)+\mathbb{J}[\Delta]))$$

where the $(Y^TY)^{-1}$ factor is introduced so that $m_{YM}(\Delta M) = m_Y(\Delta)$ for all $M \in \mathrm{GL}_p$ (this allows us to take Y not necessarily orthonormal).

The minimization problem (3.17) is to compute $\Delta^* = \arg\min_{\Delta \in H_Y} m_Y(\Delta)$. To this end, define $\mathbb{J}^T$, the adjoint of $\mathbb{J}$, by requiring that $\mathbb{J}^T$ is on $\mathbb{R}^{n\times p}$ into H_Y and verifies $\mathrm{trace}((Y^TY)^{-1}\Omega^T\mathbb{J}[\Delta]) = \mathrm{trace}((Y^TY)^{-1}(\mathbb{J}^T[\Omega])^T\Delta)$ for all $\Omega \in \mathbb{R}^{n\times p}$ and all $\Delta \in H_Y$. One obtains

$$\mathbb{J}^T[\Omega] = J^T[\Pi\Omega] - \Pi AY(Y^TY)^{-1}\Omega^T Y$$

and

$$J^T : H_Y \to H_Y : \Delta \mapsto \Pi A^T \Pi \Delta - \Delta(Y^TY)^{-1}YA^TY^T.$$

Then one readily obtains

$$\mathrm{D}m_Y(\Delta)[\tilde{\Delta}] = \mathrm{trace}((Y^TY)^{-1}(\mathbb{J}^T[F(Y)] + \mathbb{J}^T \circ \mathbb{J}[\Delta])^T\tilde{\Delta})$$

hence the solution Δ^* of the minimization problem (3.17) verifies the normal equations $\mathbb{J}^T \circ \mathbb{J}[\Delta^*] = -\mathbb{J}^T[F(Y)]$ that is

$$J^T \circ J[\Delta^*] + \Pi AY(Y^TY)^{-1}Y^TA^T\Delta^* = -J[F(Y)]. \tag{A.1}$$

If $A = A^T$, then J is self-adjoint and the latter equation develops into the NH equation (3.24).

REFERENCES

[Abs03] P.-A. Absil, *Invariant subspace computation: A geometric approach*, Ph.D. thesis, Faculté des Sciences Appliquées, Université de Liège, Secrétariat de la FSA, Chemin des Chevreuils 1 (Bât. B52), 4000 Liège, Belgium, 2003.

[ADM+02] R. L. Adler, J.-P. Dedieu, J. Y. Margulies, M. Martens, and M. Shub, *Newton's method on Riemannian manifolds and a geometric model for the human spine*, IMA J. Numer. Anal. **22** (2002), no. 3, 359–390.

[AH02] P.-A. Absil and K. Hueper, *Regularized Grassmann-Rayleigh quotient iteration*, technical note, 2002.

[AMS02] P.-A. Absil, R. Mahony, and R. Sepulchre, *Riemannian geometry of Grassmann manifolds with a view on algorithmic computation*, submitted to Acta Appl. Math. (manuscript available from http://www.montefiore.ulg.ac.be/systems/Publi/Grass_geom.htm), November 2002.

[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.

[BBM02a] K. Braman, R. Byers, and R. Mathias, *The multishift QR algorithm. Part I: maintaining well-focused shifts and level 3 performance*, SIAM J. Matrix Anal. Appl. **23** (2002), no. 4, 929–947.

[BBM02b] ———, *The multishift QR algorithm. Part II: aggressive early deflation*, SIAM J. Matrix Anal. Appl. **23** (2002), no. 4, 948–973.

[BD89] Z. Bai and J. Demmel, *On a block implementation of Hessenberg QR iteration*, Intl. J. of High Speed Comput. **1** (1989), 97–112, also available online as LAPACK Working Note 8 from http://www.netlib.org/lapack/lawns/lawn08.ps and http://www.netlib.org/lapack/lawnspdf/lawn08.pdf.

[Bra03] J. Brandts, *The Riccati algorithm for eigenvalues and invariant subspaces of matrices with inexpensive action*, Linear Algebra Appl. **358** (2003), 335–365.

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

[Cha84] F. Chatelin, *Simultaneous Newton's iteration for the eigenproblem*, Computing, Suppl. **5** (1984), 67–74.

[Dem87] J. W. Demmel, *Three methods for refining estimates of invariant subspaces*, Computing **38** (1987), 43–57.

[DF01] L. Dieci and M. J. Friedman, *Continuation of invariant subspaces*, Numer. Linear Algebra Appl. **8** (2001), no. 5, 317–327.

[DMW83] J. J. Dongarra, C. B. Moler, and J. H. Wilkinson, *Improving the accuracy of computed eigenvalues and eigenvectors*, SIAM J. Numer. Anal. **20** (1983), no. 1, 23–45.

[DP03] I. S. Dhillon and B. N. Parlett, *Orthogonal eigenvectors and relative gaps*, to appear in SIAM J. Matrix Anal. Appl., 2003.

[DS83] J. E. Dennis and R. B. Schnabel, *Numerical methods for unconstrained optimization and nonlinear equations*, Prentice Hall Series in Computational Mathematics, Prentice Hall, Englewood Cliffs, NJ, 1983.

[DS00] J.-P. Dedieu and M. Shub, *Multihomogeneous Newton method*, Math. Comp. **69** (2000), no. 231, 1071–1098.

[EAS98] A. Edelman, T. A. Arias, and S. T. Smith, *The geometry of algorithms with orthogonality constraints*, SIAM J. Matrix Anal. Appl. **20** (1998), no. 2, 303–353.

[EW96] S. Eisenstat and H. Walker, *Choosing the forcing terms in an inexact Newton method*, SIAM J. Scientific Comp. **17** (1996), no. 1, 16–32.

[Fat98] J.-L. Fattebert, *A block Rayleigh quotient iteration with local quadratic convergence*, Electron. Trans. Numer. Anal. **7** (1998), 56–74.

[FGP94] J. Ferrer, M. I. García, and F. Puerta, *Differentiable families of subspaces*, Linear Algebra Appl. **199** (1994), 229–252.

[FSV98] D. R. Fokkema, G. L. G. Sleijpen, and H. A. Van der Vorst, *Accelerated inexact Newton schemes for large systems of nonlinear equations*, SIAM J. Sci. Comput. **19** (1998), no. 2, 657–674.

[GV96] G. H. Golub and C. F. Van Loan, *Matrix computations, third edition*, The John Hopkins University Press, 1996.

[HM94] U. Helmke and J. B. Moore, *Optimization and dynamical systems*, Springer, 1994.

[Ips97] I. C. F. Ipsen, *Computing an eigenvector with inverse iteration*, SIAM Rev. **39** (1997), no. 2, 254–291.

[JS01] Z. Jia and G. W. Stewart, *An analysis of the Rayleigh-Ritz method for approximating eigenspaces*, Math. Comp. **70** (2001), no. 234, 637–647.

[Kny01] A. V. Knyazev, *Toward the optimal preconditioned eigensolver: locally optimal block preconditioned conjugate gradient method*, SIAM J. Sci. Comput. **23** (2001), no. 2, 517–451.

[LE02] E. Lundström and L. Eldén, *Adaptive eigenvalue computations using Newton's method on the Grassmann manifold*, SIAM J. Matrix Anal. Appl. **23** (2002), no. 3, 819–839.

[LST98] R. Lösche, H. Schwetlick, and G. Timmerman, *A modified block Newton iteration for approximating an invariant subspace of a symmetric matrix*, Linear Algebra Appl. **275-276** (1998), 381–400.

[MA03] R. Mahony and P.-A. Absil, *The continuous-time Rayleigh quotient flow on the sphere*, Linear Algebra Appl. **368C** (2003), 343–357.

[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, 21–44.

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

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

[PK69] B. N. Parlett and W. Kahan, *On the convergence of a practical QR algorithm*, Information Processing 68 (Proc. IFIP Congress, Edinburgh, 1968), Vol. 1: Mathematics, Software (Amsterdam), North-Holland, 1969, pp. 114–118.

[PP73] B. N. Parlett and W. G. Poole, *A geometric theory for the QR, LU and Power Iteration*, SIAM J. Numerical Analysis **10** (1973), no. 2, 389–412.

[PS95] R. D. Pantazis and D. B. Szyld, *Regions of convergence of the Rayleigh quotient iteration method*, Numer. Linear Algebra Appl. **2** (1995), no. 3, 251–269.

[PW79] G. Peters and J. H. Wilkinson, *Inverse iteration, ill-conditioned equations and Newton's method*, SIAM Review **21** (1979), no. 3, 339–360.

[PW94] C. C. Paige and M. Wei, *History and generality of the CS decomposition*, Linear Algebra Appl. **208/209** (1994), 303–326.

[RR02] A. C. M. Ran and L. Rodman, *A class of robustness problems in matrix analysis*, Interpolation Theory, Systems Theory and Related Topics, The Harry Dym Anniversary Volume (D. Alpay, I. Gohberg, and V. Vinnikov, eds.), Operator Theory: Advances and Applications, vol. 134, Brikhäuser, 2002, pp. 337–383.

[SE02] V. Simoncini and L. Eldén, *Inexact Rayleigh quotient-type methods for eigenvalue computations*, BIT **42** (2002), no. 1, 159–182.

[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 (Anthony Bloch, ed.), Fields Institute Communications, vol. 3, American Mathematical Society, 1994, pp. 113–136.

[Smi97] P. Smit, *Numerical analysis of eigenvalue algorithms based on subspace iterations*, Ph.D. thesis, CentER, Tilburg University, P.O. Box 90153, 5000 LE Tilburg, The Netherlands, 1997.

[Ste73] G. W. Stewart, *Error and perturbation bounds for subspaces associated with certain eigenvalue problems*, SIAM Review **15** (1973), no. 4, 727–764.

[SV96] G. L. G. Sleijpen and H. A. Van der Vorst, *A Jacobi-Davidson iteration method for linear eigenvalue problems*, SIAM J. Matrix Anal. Appl. **17** (1996), no. 2, 401–425.

[Wat82] D. S. Watkins, *Understanding the QR algorithm*, SIAM Review **24** (1982), no. 4, 427–440.

[WE91] D. S. Watkins and L. Elsner, *Convergence of algorithms of decomposition type for the eigenvalue problem*, Linear Algebra Appl. **143** (1991), 19–47.