

# RAYLEIGH QUOTIENT ITERATION AND SIMPLIFIED JACOBI-DAVIDSON WITH PRECONDITIONED ITERATIVE SOLVES FOR GENERALISED EIGENVALUE PROBLEMS

MELINA A. FREITAG\*, ALASTAIR SPENCE†, AND EERO VAINIKKO ‡

**Abstract.** The computation of a right eigenvector and corresponding finite eigenvalue of a large sparse generalised eigenproblem  $\mathbf{Ax} = \lambda\mathbf{Mx}$  using preconditioned Rayleigh quotient iteration and the simplified Jacobi-Davidson method is considered. Both methods are inner-outer iterative methods and we consider GMRES and FOM as iterative algorithms for the (inexact) solution of the inner systems that arise. The performance of the methods is measured in terms of the number of inner iterations needed at each outer solve. For inexact Rayleigh quotient iteration we present a detailed analysis of both unpreconditioned and preconditioned GMRES and it is shown that the number of inner iterations increases as the outer iteration proceeds. We discuss a tuning strategy for the generalised eigenproblem, and show how a rank one change to the preconditioner produces significant savings in overall costs for the Rayleigh quotient iteration. Furthermore for a specially tuned preconditioner we prove an equivalence result between inexact Rayleigh quotient iteration and simplified Jacobi-Davidson method. The theory is illustrated by several examples, including the calculation of a complex eigenvalue arising from a stability analysis of the linearised Navier-Stokes equations.

**Key words.** Eigenvalue approximation, Rayleigh quotient iteration, Jacobi-Davidson method, iterative methods, preconditioning.

**AMS subject classifications.** 65F10, 65F15.

**1. Introduction.** In this paper we consider the computation of a simple eigenvalue and corresponding eigenvector of the generalised eigenproblem  $\mathbf{Ax} = \lambda\mathbf{Mx}$  where  $\mathbf{A}$  and  $\mathbf{M}$  are large sparse nonsymmetric matrices. We examine the inexact Rayleigh quotient iteration (RQI) and the simplified Jacobi-Davidson (JD) method. The RQI requires the repeated solution of shifted linear systems of the form

$$(\mathbf{A} - \rho\mathbf{M})\mathbf{y} = \mathbf{Mx} \tag{1.1}$$

where  $\rho$  is a generalised Rayleigh quotient and we analyse in detail the performance of preconditioned GMRES applied to (1.1). We shall show that a simple modification of a standard preconditioner produces significant savings in costs.

The convergence theory for inverse iteration with inexact solves has been considered in [11] for several shift strategies. This theory covers the most general setting, where  $\mathbf{A}$  and  $\mathbf{M}$  are nonsymmetric with  $\mathbf{M}$  allowed to be singular (for other special cases, for example, symmetric problems or only fixed shifts, see [18, 25, 36, 27, 1, 2]). It was shown that, for a fixed shift strategy, a decreasing tolerance provides linear convergence, whereas for a generalised Rayleigh Quotient shift a constant solve tolerance gives linear convergence and a decreasing tolerance achieves quadratic convergence.

For inexact inverse iteration the costs of the inner solves using Krylov methods has been investigated in [1] and [14] for the symmetric solvers CG/MINRES and in [3] for the nonsymmetric solver GMRES. In these papers it was shown that, for the standard eigenvalue problem, the number of inner iterations remains approximately constant as the outer iteration proceeds if no preconditioner is used but increases if a standard preconditioner is applied. A so-called tuned preconditioner has been introduced in [14] for the standard Hermitian eigenproblem and in [10] for the generalised eigenproblem (though no analysis of the inner solves was provided in [10]). In [44] a detailed analysis of the convergence of MINRES for the tuned preconditioner applied to the standard Hermitian eigenproblem was given. The idea of modifying the

---

\*Department of Mathematical Sciences, University of Bath, Claverton Down, BA2 7AY, United Kingdom (email: [m.freitag@maths.bath.ac.uk](mailto:m.freitag@maths.bath.ac.uk)). This author was supported by EPSRC, UK.

†Department of Mathematical Sciences, University of Bath, Claverton Down, BA2 7AY, United Kingdom (email: [as@maths.bath.ac.uk](mailto:as@maths.bath.ac.uk))

‡Faculty of Mathematics and Computer Science, University of Tartu, Liivi 2, Tartu 50409, Estonia (email: [eero.vainikko@ut.ee](mailto:eero.vainikko@ut.ee)).

preconditioner with small rank changes has been extended to subspace versions of inexact inverse iteration: Robbé et al [29] consider tuning for inexact inverse subspace iteration and [13] extends the idea to the shift-and-invert Arnoldi method.

This paper has several aims which distinguish it from earlier reports on tuning. First, we extend the results from [14] for standard Hermitian eigenproblems to the generalised nonsymmetric eigenproblem  $\mathbf{Ax} = \lambda\mathbf{Mx}$  and give a detailed analysis of the costs of the inner solves as the outer iteration proceeds. Second, we extend the analysis to the choice of Rayleigh quotient shifts and note that the analysis readily extends to other variable shift strategies. For the generalised eigenproblem it turns out that for both unpreconditioned and preconditioned GMRES a tuning strategy gives significant improvement. This analysis suggests that tuning provides a powerful computational tool for reducing the costs in the computations of eigenvalues of large nonsymmetric problems. Third, we present numerical results for an eigenvalue problem arising from a mixed finite element discretisation of the linearised steady Navier-Stokes equation. Specifically, we calculate a complex eigenvalue near the imaginary axis which suggests a nearby Hopf bifurcation in the corresponding time-dependent problem. Fourth, this paper provides a theoretical tool to compare the numerical performance of preconditioned inexact RQI and the preconditioned JD method and hence analyse the latter. Various experiments on iterative solves for the inner linear systems arising in inexact RQI and inexact JD [28] have shown that, if standard preconditioners are used, JD gives some advantage. In [12] it was shown that for the standard eigenvalue problem, the simplified JD method, that is the JD method without subspace expansion, used with a standard preconditioner is equivalent to applying inexact inverse iteration with a tuned preconditioner if the same Galerkin-Krylov method is used for the inner solves. In this paper we extend this theory to the generalised nonsymmetric eigenproblem. This shows that inexact RQI with a tuned preconditioner gives identical results to those obtained using the preconditioned simplified JD method. Numerical results are given comparing preconditioned solves in simplified JD and inexact RQI where GMRES is used as solver.

The paper is organised as follows. Section 2 gives some background results including the convergence theory of inexact RQI and a convergence theory of GMRES, where the right hand side is given special consideration. Section 3 presents the concept of tuning and its implementation for the generalised eigenproblem. We present theoretical results for RQI and right preconditioning, though we note that the analysis extends to other variable shift choices and left preconditioning. A numerical example taken from the stability analysis of the linearised steady Navier-Stokes equations is given in Section 4. Section 5 provides a result on the equivalence between the preconditioned JD method and inexact RQI with a tuned preconditioner for a generalised eigenproblem and a Galerkin-Krylov solver. Numerical results confirm this equivalence. We also present numerical results when GMRES is used as the iterative solver. These results show a very close relationship between the RQI and the simplified JD method.

Throughout this paper we use  $\|\cdot\| = \|\cdot\|_2$ .

**2. Preliminaries.** Consider the general nonsymmetric eigenproblem

$$\mathbf{Ax} = \lambda\mathbf{Mx}, \quad \mathbf{x} \neq \mathbf{0}, \quad (2.1)$$

with  $\mathbf{A} \in \mathbb{C}^{n \times n}$  and  $\mathbf{M} \in \mathbb{C}^{n \times n}$ , where we assume that  $(\lambda_1, \mathbf{x}_1)$  is a simple, well-separated finite eigenpair with corresponding left eigenvector  $\mathbf{u}_1^H$ , that is  $\mathbf{Ax}_1 = \lambda_1\mathbf{Mx}_1$  and  $\mathbf{u}_1^H\mathbf{A} = \lambda_1\mathbf{u}_1^H\mathbf{M}$ , where  $\mathbf{u}_1^H\mathbf{Mx}_1 \neq 0$ . We consider inexact RQI (see [11]) to compute  $\lambda_1$ . Let  $\mathbf{x}^{(i)}$  be an approximation to  $\mathbf{x}_1$ . Then, at each step  $i$ , a solve of the system

$$(\mathbf{A} - \rho(\mathbf{x}^{(i)})\mathbf{M})\mathbf{y}^{(i)} = \mathbf{Mx}^{(i)} \quad (2.2)$$

is required, where (a normed)  $\mathbf{y}^{(i)}$  is a better approximation to  $\mathbf{x}_1$  and  $\rho(\mathbf{x}^{(i)})$  is the generalised Rayleigh quotient defined by

$$\rho(\mathbf{x}^{(i)}) = \frac{\mathbf{x}^{(i)H}\mathbf{M}^H\mathbf{Ax}^{(i)}}{\mathbf{x}^{(i)H}\mathbf{M}^H\mathbf{Mx}^{(i)}}. \quad (2.3)$$

This has the desirable property that, for a given  $\mathbf{x}^{(i)}$ , the eigenvalue residual

$$\mathbf{r}^{(i)} = \mathbf{A}\mathbf{x}^{(i)} - \rho(\mathbf{x}^{(i)})\mathbf{M}\mathbf{x}^{(i)} \quad (2.4)$$

is minimised with respect to  $\|\cdot\|$ . Suppose that  $\mathbf{x}^{(i)}$  is normalised such that  $\|\mathbf{M}\mathbf{x}^{(i)}\| = 1$ . This scaling is possible if  $\mathbf{x}^{(i)}$  is close to  $\mathbf{x}_1$ , an eigenvector belonging to a finite eigenvalue of (2.1). Assuming  $\rho(\mathbf{x}^{(i)})$  is closer to  $\lambda_1$  than to any other eigenvalue and  $\mathbf{x}^{(0)}$  is close enough to  $\mathbf{x}_1$ , we have the following convergence result. Let

$$\mathbf{d}^{(i)} = \mathbf{M}\mathbf{x}^{(i)} - (\mathbf{A} - \rho(\mathbf{x}^{(i)})\mathbf{M})\mathbf{y}^{(i)}, \quad \text{with} \quad \|\mathbf{d}^{(i)}\| \leq \tau^{(i)}$$

with  $\tau^{(0)}$  small enough. For a Rayleigh quotient shift (2.3) and a decreasing tolerance  $\tau^{(i)} = \delta\|\mathbf{r}^{(i)}\|$  quadratic convergence is obtained, for a fixed  $\tau^{(i)} := \tau^{(0)}$  the convergence is linear (see Remark 3.6 in [11]).

A popular method to solve the linear system (2.2) iteratively is GMRES. We now give a convergence result for GMRES applied to the system  $\mathbf{B}\mathbf{z} = \mathbf{b}$ , where  $\mathbf{B}$  has a well-separated simple eigenvalue near zero and the form of the right hand side,  $\mathbf{b}$  is taken into consideration. This theory is general in the sense that it does not need the system matrix  $\mathbf{B}$  to be diagonalisable. In the theory in Section 3 we shall choose  $\mathbf{B}$  to be  $\mathbf{A} - \rho(\mathbf{x}^{(i)})\mathbf{M}$  or  $(\mathbf{A} - \rho(\mathbf{x}^{(i)})\mathbf{M})\mathbf{P}^{-1}$  where  $\mathbf{P}$  is some preconditioner, and  $\mathbf{b}$  will be  $\mathbf{M}\mathbf{x}^{(i)}$ .

We summarise some theoretical results assuming that  $\mathbf{B}$  has an algebraically simple eigenpair  $(\mu_1, \mathbf{w}_1)$  with  $\mu_1$  near zero and well-separated from the rest of the spectrum. Schur's theorem [17, page 313] ensures the existence of a unitary matrix  $[\mathbf{w}_1, \mathbf{W}_1^\perp]$  such that

$$\mathbf{B} = \begin{bmatrix} \mathbf{w}_1 & \mathbf{W}_1^\perp \end{bmatrix} \begin{bmatrix} \mu_1 & \mathbf{n}_{12}^H \\ \mathbf{0} & \mathbf{N}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{w}_1 & \mathbf{W}_1^\perp \end{bmatrix}^H, \quad (2.5)$$

where  $\mathbf{w}_1 \in \mathbb{C}^{n \times 1}$ ,  $\mathbf{W}_1^\perp \in \mathbb{C}^{n \times (n-1)}$ ,  $\mathbf{n}_{12} \in \mathbb{C}^{(n-1) \times 1}$  and  $\mathbf{N}_{22} \in \mathbb{C}^{(n-1) \times (n-1)}$ . Since  $\mu_1$  is not contained in the spectrum of  $\mathbf{N}_{22}$  the equation

$$\mathbf{f}^H \mathbf{N}_{22} - \mu_1 \mathbf{f}^H = \mathbf{n}_{12}^H \quad (2.6)$$

has a unique solution  $\mathbf{f} \in \mathbb{C}^{(n-1) \times 1}$  (see [17, Lemma 7.1.5]) and with

$$\mathbf{W}_2 = (-\mathbf{w}_1 \mathbf{f}^H + \mathbf{W}_1^\perp)(\mathbf{I} + \mathbf{f}\mathbf{f}^H)^{-\frac{1}{2}} \quad (2.7)$$

and  $\mathbf{C} = (\mathbf{I} + \mathbf{f}\mathbf{f}^H)^{-\frac{1}{2}} \mathbf{N}_{22} (\mathbf{I} + \mathbf{f}\mathbf{f}^H)^{\frac{1}{2}}$  we obtain the block-diagonalisation of  $\mathbf{B}$ :

$$\mathbf{B} = \begin{bmatrix} \mathbf{w}_1 & \mathbf{W}_2 \end{bmatrix} \begin{bmatrix} \mu_1 & \mathbf{0}^H \\ \mathbf{0} & \mathbf{C} \end{bmatrix} \begin{bmatrix} \mathbf{v}_1^H \\ \mathbf{V}_2^H \end{bmatrix}, \quad (2.8)$$

where  $[\mathbf{v}_1, \mathbf{V}_2]^H$  is the inverse of the nonsingular matrix  $[\mathbf{w}_1, \mathbf{W}_2]$ , see [39, Theorem 1.18]. Note that  $\mathbf{C}$  and  $\mathbf{N}_{22}$  have the same spectrum. We have

$$\mathbf{v}_1 = \mathbf{w}_1 + \mathbf{W}_1^\perp \mathbf{f} \quad \text{and} \quad \mathbf{V}_2 = \mathbf{W}_1^\perp (\mathbf{I} + \mathbf{f}\mathbf{f}^H)^{\frac{1}{2}}.$$

Furthermore,  $\mathbf{B}\mathbf{w}_1 = \mu_1 \mathbf{w}_1$  and  $\mathbf{v}_1^H \mathbf{B} = \mu_1 \mathbf{v}_1^H$ , that is,  $\mathbf{v}_1$  is the left eigenvector of  $\mathbf{B}$  corresponding to  $\mu_1$ . Note that  $\mathbf{v}_1^H \mathbf{w}_1 = 1$ ,  $\mathbf{V}_2^H \mathbf{W}_2 = \mathbf{I}$ ,  $\mathbf{v}_1^H \mathbf{W}_2 = \mathbf{0}^H$  and  $\mathbf{V}_2^H \mathbf{w}_1 = \mathbf{0}$ . Also  $\|\mathbf{w}_1\| = 1$ , and  $\mathbf{W}_2$  has orthonormal columns.

Further, introduce the separation function  $\text{sep}(\mu_1, \mathbf{C})$  (see, for example [38, 40]) defined by

$$\text{sep}(\mu_1, \mathbf{C}) := \begin{cases} \|(\mu_1 \mathbf{I} - \mathbf{C})^{-1}\|_2^{-1}, & \mu_1 \notin \Lambda(\mathbf{C}) \\ 0, & \mu_1 \in \Lambda(\mathbf{C}) \end{cases}$$

and note that by definition  $\text{sep}(\mu_1, \mathbf{C}) = \sigma_{\min}(\mu_1 \mathbf{I} - \mathbf{C})$ , where  $\sigma_{\min}$  is the minimum singular value and the quantities  $\text{sep}(\mu_1, \mathbf{C})$  and  $\text{sep}(\mu_1, \mathbf{N}_{22})$  are related by (see [40])

$$\frac{\text{sep}(\mu_1, \mathbf{N}_{22})}{\kappa} \leq \text{sep}(\mu_1, \mathbf{C}) \leq \kappa \text{sep}(\mu_1, \mathbf{N}_{22}), \quad \text{where } \kappa = \sqrt{1 + \mathbf{f}^H \mathbf{f}}. \quad (2.9)$$

Further, let

$$\mathcal{P} = \mathbf{I} - \mathbf{w}_1 \mathbf{v}_1^H, \quad (2.10)$$

be an oblique projector (see, for example, [30, 41] and [40]) onto  $\mathcal{R}(\mathbf{W}_2)$  along  $\mathcal{R}(\mathbf{w}_1)$  and  $\mathbf{I} - \mathcal{P}$  projects onto  $\mathcal{R}(\mathbf{w}_1)$  along the orthogonal complement  $\mathcal{R}(\mathbf{W}_2)$  of  $\mathcal{R}(\mathbf{v}_1)$ . With the results after (2.8) we have  $\|\mathcal{P}\| = \sqrt{1 + \|\mathbf{f}\|^2} = \kappa$ . Before analysing GMRES we state a proposition which follows from the perturbation theory of eigenvectors belonging to simple eigenvalues (see [39], [40] and [38]).

**PROPOSITION 2.1.** *Let  $\mu_1$  be a simple eigenvalue of  $\mathbf{B}$  with corresponding right eigenvector  $\mathbf{w}_1$  and let  $\mathbf{W} = [\mathbf{w}_1, \mathbf{W}_1^\perp]$  be unitary such that (2.5) holds. Let*

$$\mathbf{B} \hat{\mathbf{w}} = \xi \hat{\mathbf{w}} + \mathbf{e}, \quad \text{with } \|\hat{\mathbf{w}}\| = 1, \quad (2.11)$$

that is  $(\hat{\mathbf{w}}, \xi)$  is an approximate eigenpair of  $\mathbf{B}$  with residual  $\mathbf{e}$ . If  $\mathbf{e}$  is small enough such that  $\|\mathbf{e}\| < \frac{1}{2} \text{sep}(\mu_1, \mathbf{N}_{22})$  and  $\frac{\|\mathbf{e}\|(\|\mathbf{e}\| + \|\mathbf{n}_{12}\|)}{(\text{sep}(\mu_1, \mathbf{N}_{22}) - 2\|\mathbf{e}\|)^2} < \frac{1}{4}$ , then there exists a unique vector  $\mathbf{p}$  satisfying

$$\|\mathbf{p}\| \leq 2 \frac{\|\mathbf{e}\|}{\text{sep}(\mu_1, \mathbf{N}_{22}) - 2\|\mathbf{e}\|} \quad \text{such that } \hat{\mathbf{w}} = \frac{\mathbf{w}_1 + \mathbf{W}_1^\perp \mathbf{p}}{\sqrt{1 + \mathbf{p}^H \mathbf{p}}}.$$

*Proof.* Write (2.11) as a perturbed eigenvalue problem  $(\mathbf{B} - \mathbf{e} \hat{\mathbf{w}}^H) \hat{\mathbf{w}} = \xi \hat{\mathbf{w}}$  and apply [40, Theorem 2.7, Chapter 5].  $\square$

Proposition 2.1 shows that the eigenvector  $\hat{\mathbf{w}}$  of the perturbed problem (2.11) compared to the exact problem  $\mathbf{B} \mathbf{w}_1 = \mu_1 \mathbf{w}_1$  depends on the size of the norm of the perturbation  $\mathbf{e}$  and on the separation of the eigenvalue  $\mu_1$  from the rest of the spectrum.

Using the block factorisation (2.8) and the oblique projector (2.10) we have the following convergence result for GMRES applied to the linear system  $\mathbf{B} \mathbf{z} = \mathbf{b}$ .

**THEOREM 2.2 (GMRES convergence).** *Suppose the nonsymmetric matrix  $\mathbf{B} \in \mathbb{C}^{n \times n}$  has a simple eigenpair  $(\mu_1, \mathbf{w}_1)$  with block diagonalisation (2.8). and set  $\mathcal{P} = \mathbf{I} - \mathbf{w}_1 \mathbf{v}_1^H$ . Let  $\mathbf{z}_k$  be the result of applying  $k$  steps of GMRES to  $\mathbf{B} \mathbf{z} = \mathbf{b}$  with starting value  $\mathbf{z}_0 = \mathbf{0}$ . Then*

$$\|\mathbf{b} - \mathbf{B} \mathbf{z}_k\| \leq \min_{p_{k-1} \in \Pi_{k-1}} \|p_{k-1}(\mathbf{C})\| \|\mathbf{I} - \mu_1^{-1} \mathbf{C}\| \|\mathbf{V}_2\| \|\mathcal{P} \mathbf{b}\|, \quad (2.12)$$

where  $\Pi_{k-1}$  is the set of complex polynomials of degree  $k-1$  normalised such that  $p(0) = 1$  and  $\|\mathbf{V}_2\| = \sqrt{1 + \|\mathbf{f}\|^2} = \|\mathbf{v}_1\| = \kappa$  where  $\mathbf{f}$  is given by (2.6) and  $\kappa$  is given by (2.9).

*Proof.* The residual for GMRES satisfies (see [20])

$$\|\mathbf{b} - \mathbf{B} \mathbf{z}_k\|_2 = \min_{p_k \in \Pi_k} \|p_k(\mathbf{B}) \mathbf{b}\|,$$

where  $\Pi_k$  is the set of polynomials of degree  $k$  with  $p(0) = 1$ . Introduce special polynomials  $\hat{p}_k \in \Pi_k$ , given by

$$\hat{p}_k(z) = p_{k-1}(z) (1 - \mu_1^{-1} z),$$

where  $p_{k-1} \in \Pi_{k-1}$ . Note that similar polynomials were introduced by Campbell et al. [5] (see also [43]). Then we can write

$$\begin{aligned} \|\mathbf{b} - \mathbf{B} \mathbf{z}_k\|_2 &= \min_{p_k \in \Pi_k} \|p_k(\mathbf{B}) \mathcal{P} \mathbf{b} + p_k(\mathbf{B}) (\mathbf{I} - \mathcal{P}) \mathbf{b}\| \leq \min_{\hat{p}_k \in \Pi_k} \|\hat{p}_k(\mathbf{B}) \mathcal{P} \mathbf{b} + \hat{p}_k(\mathbf{B}) (\mathbf{I} - \mathcal{P}) \mathbf{b}\| \\ &\leq \min_{p_{k-1} \in \Pi_{k-1}} \|p_{k-1}(\mathbf{B}) (\mathbf{I} - \mu_1^{-1} \mathbf{B}) \mathcal{P} \mathbf{b} + p_{k-1}(\mathbf{B}) (\mathbf{I} - \mu_1^{-1} \mathbf{B}) (\mathbf{I} - \mathcal{P}) \mathbf{b}\|. \end{aligned}$$

For the second term we have  $(\mathbf{I} - \mu_1^{-1}\mathbf{B})(\mathbf{I} - \mathcal{P})\mathbf{b} = (\mathbf{I} - \mu_1^{-1}\mathbf{B})\mathbf{w}_1\mathbf{v}_1^H\mathbf{b} = \mathbf{0}$ , using (2.10) and  $\mathbf{B}\mathbf{w}_1 = \mu_1\mathbf{w}_1$ . Therefore

$$\|\mathbf{b} - \mathbf{B}\mathbf{z}_k\| \leq \min_{p_{k-1} \in \Pi_{k-1}} \|p_{k-1}(\mathbf{B})(\mathbf{I} - \mu_1^{-1}\mathbf{B})\mathcal{P}\mathbf{b}\|. \quad (2.13)$$

With  $\mathcal{P}^2 = \mathcal{P}$  and  $\mathcal{P}\mathbf{B} = \mathbf{B}\mathcal{P}$  we have

$$\begin{aligned} p_{k-1}(\mathbf{B})(\mathbf{I} - \mu_1^{-1}\mathbf{B})\mathcal{P}\mathbf{b} &= p_{k-1}(\mathbf{W}_2\mathbf{C}\mathbf{V}_2^H)(\mathbf{I} - \mu_1^{-1}\mathbf{B})\mathcal{P}\mathbf{b} \\ &= \mathbf{W}_2 p_{k-1}(\mathbf{C})(\mathbf{I} - \mu_1^{-1}\mathbf{C})\mathbf{V}_2^H\mathcal{P}\mathbf{b}, \end{aligned}$$

and hence the result (2.12) follows, since  $\mathbf{W}_2$  has orthonormal columns.  $\square$

Note that the minimum in (2.12) is taken with respect to the smaller matrix  $\mathbf{C}$  instead of  $\mathbf{B}$ . To bound the first term on the right hand side of (2.12) we use the  $\varepsilon$ -pseudospectrum  $\Lambda_\varepsilon(\mathbf{C})$  of a matrix  $\mathbf{C}$ , defined by

$$\Lambda_\varepsilon(\mathbf{C}) := \{z \in \mathbb{C} : \|(z\mathbf{I} - \mathbf{C})^{-1}\|_2 \geq \varepsilon^{-1}\}. \quad (2.14)$$

(see, for example [9]) and results from complex analysis.

**PROPOSITION 2.3.** *Suppose there exists a convex closed bounded set  $E$  in the complex plane satisfying  $0 \notin E$ , containing the  $\varepsilon$ -pseudospectrum  $\Lambda_\varepsilon(\mathbf{C})$ . Let  $\Psi$  be the conformal mapping that carries the exterior of  $E$  onto the exterior of the unit circle  $\{|w| > 1\}$  and that takes infinity to infinity. Then*

$$\min_{p_{k-1} \in \Pi_{k-1}} \|p_{k-1}(\mathbf{C})\| \leq S \left( \frac{1}{|\Psi(0)|} \right)^{k-1}, \quad \text{where } S = \frac{3\mathcal{L}(\Gamma_\varepsilon)}{2\pi\varepsilon} \quad (2.15)$$

and  $|\Psi(0)| > 1$ . Hence

$$\|\mathbf{b} - \mathbf{B}\mathbf{z}_k\| \leq S \left( \frac{1}{|\Psi(0)|} \right)^{k-1} \|\mathbf{V}_2\| \|(\mathbf{I} - \mu_1^{-1}\mathbf{C})\| \|\mathcal{P}\mathbf{b}\|, \quad (2.16)$$

for any choice of the parameter  $\varepsilon$ , where  $\mathcal{L}(\Gamma_\varepsilon)$  is the contour length of  $\Gamma_\varepsilon$  and  $\Gamma_\varepsilon$  is the contour or union of contours enclosing  $\Lambda_\varepsilon(\mathbf{C})$ .

*Proof.* Using the pseudospectral bound in [42, Theorem 26.2] we get

$$\|p_{k-1}(\mathbf{C})\| \leq \frac{\mathcal{L}(\Gamma_\varepsilon)}{2\pi\varepsilon} \max_{z \in \Lambda_\varepsilon(\mathbf{C})} |p_{k-1}(z)|,$$

where  $\mathcal{L}(\Gamma_\varepsilon)$  is the contour length of  $\Gamma_\varepsilon$ , and

$$\min_{p_{k-1} \in \Pi_{k-1}} \max_{z \in \Lambda_\varepsilon(\mathbf{C})} |p_{k-1}(z)| \leq \frac{3}{|\Psi(0)|^{k-1}} \quad (2.17)$$

follows from [22, Proof of Lemma 1]. As  $0 \notin E$  and as  $\Psi$  maps the exterior of  $E$  onto the exterior of a unit disc we have  $|\Psi(0)| > 1$  and hence, with  $\Lambda_\varepsilon(\mathbf{C}) \subset E$  and (2.15) we obtain (2.16) from (2.12).  $\square$

The following corollary is immediately obtained from Proposition 2.3.

**COROLLARY 2.4.** *Let  $\mathbf{C}$  be perturbed to  $\mathbf{C} + \delta\mathbf{C}$ , where  $\|\delta\mathbf{C}\| < \varepsilon$ . Then*

$$\min_{p_{k-1} \in \Pi_{k-1}} \|p_{k-1}(\mathbf{C} + \delta\mathbf{C})\| \leq S_\delta \left( \frac{1}{|\Psi(0)|} \right)^{k-1}, \quad (2.18)$$

where  $S_\delta = 3\mathcal{L}(\Gamma_\varepsilon)(2\pi(\varepsilon - \|\delta\mathbf{C}\|))^{-1}$ .

Note that the bound in (2.16) describes the convergence behaviour in the worst-case sense and is by no means sharp, see [26]. Simpler bounds using Chebychev polynomials can be derived if  $\mathbf{B}$  is diagonalisable and the eigenvalues are located in an ellipse or circle (see Saad [32]).

PROPOSITION 2.5 (Number of inner iterations). *Let the assumptions of Theorem 2.2 hold and let  $\mathbf{z}_k$  be the approximate solution of  $\mathbf{B}\mathbf{z} = \mathbf{b}$  obtained after  $k$  iterations of GMRES with starting value  $\mathbf{z}_0 = \mathbf{0}$ . If the number of inner iterations satisfies*

$$k \geq 1 + \frac{1}{\log|\Psi(0)|} \left( \log(S\|\mu_1\mathbf{I} - \mathbf{C}\|\|\mathbf{V}_2\|) + \log \frac{\|\mathcal{P}\mathbf{b}\|}{|\mu_1|\tau} \right), \quad (2.19)$$

then  $\|\mathbf{b} - \mathbf{B}\mathbf{z}_k\| \leq \tau$ .

*Proof.* Taking log's in (2.16) gives the required result.  $\square$

The bound in (2.19) is only a sufficient condition, the desired accuracy might be reached for a much smaller value of  $k$ . We shall see in the following section that the important term in (2.19) is

$$\log \frac{\|\mathcal{P}\mathbf{b}\|}{|\mu_1|\tau}. \quad (2.20)$$

Since  $\tau$  may be small and  $|\mu_1|$  will be small (since we assume  $\mu_1$  is near zero) (2.20) will be large unless  $\|\mathcal{P}\mathbf{b}\|$  is also small. We shall make this more precise in the next section, but first let us analyse further the term  $\mathcal{P}\mathbf{b} = (\mathbf{I} - \mathbf{w}_1\mathbf{v}_1^H)\mathbf{b}$  for a general  $\mathbf{b}$ .

PROPOSITION 2.6. *Consider the linear system  $\mathbf{B}\mathbf{z} = \mathbf{b}$ . Let  $\mathcal{P} = \mathbf{I} - \mathbf{w}_1\mathbf{v}_1^H$  where  $\mathbf{v}_1$  and  $\mathbf{w}_1$  are left and right eigenvectors of  $\mathbf{B}$ . Furthermore, let  $\mathbf{b}$  be decomposed as  $\mathbf{b} = b_1\mathbf{w}_1 + \mathbf{W}_2\mathbf{b}_2$ , where  $\mathbf{w}_1$  and  $\mathbf{W}_2$  are as in (2.8),  $b_1 \in \mathbb{C}$  and  $\mathbf{b}_2 \in \mathbb{C}^{(n-1) \times 1}$ . Then*

$$\|\mathcal{P}\mathbf{b}\| = \|\mathbf{b}_2\|. \quad (2.21)$$

*Proof.* For general  $\mathbf{b}$  we have  $\mathcal{P}\mathbf{b} = (\mathbf{I} - \mathbf{w}_1\mathbf{v}_1^H)\mathbf{b} = (\mathbf{b} - \mathbf{v}_1^H\mathbf{b}\mathbf{w}_1)$ . Using the decomposition  $\mathbf{b} = b_1\mathbf{w}_1 + \mathbf{W}_2\mathbf{b}_2$  we have

$$\mathcal{P}\mathbf{b} = b_1\mathbf{w}_1 + \mathbf{W}_2\mathbf{b}_2 - \mathbf{v}_1^H(b_1\mathbf{w}_1 + \mathbf{W}_2\mathbf{b}_2)\mathbf{w}_1 = \mathbf{W}_2\mathbf{b}_2,$$

where we have used  $\mathbf{v}_1^H\mathbf{w}_1 = 1$  and  $\mathbf{v}_1^H\mathbf{W}_2 = \mathbf{0}^H$  (see remarks after (2.8)). Hence  $\|\mathcal{P}\mathbf{b}\| = \|\mathbf{W}_2\|\|\mathbf{b}_2\| = \|\mathbf{b}_2\|$ , since  $\mathbf{W}_2$  has orthonormal columns.  $\square$

Hence, for any vector  $\mathbf{b}$ , which is not parallel to  $\mathbf{w}_1$  (that is,  $\|\mathbf{b}_2\| \neq 0$ ) we have that  $\|\mathcal{P}\mathbf{b}\|$  is nonzero. However, if  $\mathbf{b}_2$  is small, then  $\|\mathcal{P}\mathbf{b}\|$  will be small and this may help counter the small terms  $|\mu_1|$  and  $\tau$  in the numerator of (2.20). This is a key point and is discussed more concretely in the next section.

**3. Tuning for the nonsymmetric generalised eigenproblem.** This section contains one of the main results in this paper. Specifically we discuss the use of a tuned preconditioner to obtain an efficient solution procedure for the linear systems arising in inexact RQI. The main theoretical results are given in Theorem 3.5 which relies on Proposition 2.5.

In [14] it is shown that for the standard Hermitian eigenproblem inexact inverse iteration with a fixed shift leads to a constant number of inner iterations as the outer iteration proceeds, even though the solve tolerance is decreased in every step. This somehow surprising outcome is a result of the fact that the right hand side of the linear system is an increasingly better approximation to the eigendirection of the system matrix. As we shall see this result does not hold for the unpreconditioned (or preconditioned) linear system (2.2), but can be achieved by an appropriate rank-one change to the standard preconditioner.

**3.1. Unpreconditioned solution of the linear system (2.2).** Inexact RQI involves the iterative solution of the linear system (2.2). We assume that this is done using GMRES, so that, for fixed  $i$ , the theory in Section 2 applies with

$$\mathbf{B} = \mathbf{A} - \rho(\mathbf{x}^{(i)})\mathbf{M} \quad \text{and} \quad \mathbf{b} = \mathbf{M}\mathbf{x}^{(i)},$$

where we normalise  $\mathbf{x}^{(i)}$  as  $\|\mathbf{M}\mathbf{x}^{(i)}\| = 1$ , so that  $\|\mathbf{b}\| = 1$ . In this case, with  $\mathbf{w}_1$ ,  $\mathbf{W}_2$ ,  $\mathbf{v}_1$  and  $\mathbf{V}_2$  defined as in (2.8) there is no reason for the norm of  $\mathcal{P}\mathbf{b} = \mathcal{P}\mathbf{M}\mathbf{x}^{(i)} = (\mathbf{I} - \mathbf{w}_1\mathbf{v}_1^H)\mathbf{M}\mathbf{x}^{(i)}$  to be small since  $\mathbf{M}\mathbf{x}^{(i)}$  will not be close to  $\mathbf{w}_1$ , the eigenvector of  $\mathbf{A} - \rho(\mathbf{x}^{(i)})\mathbf{M}$  corresponding to the eigenvalue  $\mu_1^{(i)}$  nearest zero. Thus  $\|\mathcal{P}\mathbf{M}\mathbf{x}^{(i)}\| \not\rightarrow 0$  as  $i \rightarrow \infty$ . Proposition 2.5 states that not more than  $k^{(i)}$  inner iterations per outer iteration  $i$  are required to solve (2.2) to a tolerance  $\tau^{(i)}$ , where  $k^{(i)}$  is bounded by

$$k^{(i)} \geq 1 + \frac{1}{\log|\Psi(0)|} \left( \log S \|\mu_1\mathbf{I} - \mathbf{C}\| \|\mathbf{V}_2\| + \log \frac{\|\mathcal{P}\mathbf{M}\mathbf{x}^{(i)}\|}{|\mu_1|\tau^{(i)}} \right). \quad (3.1)$$

If we choose a fixed shift and a decreasing tolerance  $\tau^{(i)}$  the lower bound on the inner iterations will increase with  $i$  as  $i \rightarrow \infty$ . If we choose a fixed tolerance and a variable shift  $\rho(\mathbf{x}^{(i)})$ , then  $\mu_1^{(i)}$ , the eigenvalue of  $\mathbf{A} - \rho(\mathbf{x}^{(i)})\mathbf{M}$  will tend to zero as  $i \rightarrow \infty$ , and so, again, the lower bound on the number of inner iterations will increase with  $i$ . If  $\tau^{(i)}$  also decreases, say with  $\tau^{(i)} = C\|\mathbf{r}^{(i)}\|$  for some  $C$ , then the lower bound will further increase with  $i$  as  $\|\mathbf{r}^{(i)}\| \rightarrow 0$ . This behaviour can be observed in practice (see, for example [2, 3]).

We note that both a right preconditioned system, given by

$$(\mathbf{A} - \rho(\mathbf{x}^{(i)})\mathbf{M})\mathbf{P}^{-1}\tilde{\mathbf{y}}^{(i)} = \mathbf{M}\mathbf{x}^{(i)}, \quad \mathbf{P}^{-1}\tilde{\mathbf{y}}^{(i)} = \mathbf{y}^{(i)},$$

and a left preconditioned system, given by

$$\mathbf{P}^{-1}(\mathbf{A} - \rho(\mathbf{x}^{(i)})\mathbf{M})\mathbf{y}^{(i)} = \mathbf{P}^{-1}\mathbf{M}\mathbf{x}^{(i)},$$

where  $\mathbf{P}$  is a preconditioner for  $\mathbf{A} - \rho(\mathbf{x}^{(i)})\mathbf{M}$  yield similar results, since  $\mathbf{M}\mathbf{x}^{(i)}$  is not an eigenvector (or close to an eigenvector) of  $(\mathbf{A} - \rho(\mathbf{x}^{(i)})\mathbf{M})\mathbf{P}^{-1}$  and  $\mathbf{P}^{-1}\mathbf{M}\mathbf{x}^{(i)}$  is also not close to an eigendirection of  $\mathbf{P}^{-1}(\mathbf{A} - \rho(\mathbf{x}^{(i)})\mathbf{M})$ . This behaviour occurs for both fixed and variable shifts and is indeed observed in the example in Section 4.

In the next two subsections we shall show that when preconditioned solves are used this situation can be significantly improved with a simple rank-one change to the preconditioner. Note that, in this paper we only treat the case of solving (2.2) with a right preconditioner. The tuned left preconditioner and the tuning operator for an unpreconditioned system are straightforward consequences of this approach.

In the following subsections we shall assume that a good preconditioner for  $\mathbf{A}$  is also a good preconditioner for  $\mathbf{A} - \sigma\mathbf{M}$ . This is the approach taken in [34] and it is likely to be the case if  $\mathbf{A}$  arises from a discretised partial differential equation where a tailor-made preconditioner for  $\mathbf{A}$  may be available. Further we restrict ourselves to right preconditioners, all results readily extend to left preconditioners.

**3.2. The ideal preconditioner.** In this subsection we discuss a rather theoretical case. Suppose  $\mathbf{x}^{(i)} = \mathbf{x}_1$  (that is, convergence has occurred) and consider the question of applying GMRES to a preconditioned version of the linear system

$$(\mathbf{A} - \rho(\mathbf{x}_1)\mathbf{M})\mathbf{y} = \mathbf{M}\mathbf{x}_1, \quad (3.2)$$

which is the limiting case of (2.2) with  $\rho(\mathbf{x}_1) = \lambda_1$ . Set  $\hat{\mathbf{x}}_1 = \mathbf{x}_1/\|\mathbf{x}_1\|$  and suppose the preconditioner

$$\mathbb{P} = \mathbf{P}(\mathbf{I} - \hat{\mathbf{x}}_1\hat{\mathbf{x}}_1^H) + \mathbf{A}\hat{\mathbf{x}}_1\hat{\mathbf{x}}_1^H, \quad (3.3)$$

is used, which satisfies  $\mathbb{P}\mathbf{x}_1 = \mathbf{A}\mathbf{x}_1 = \lambda_1\mathbf{M}\mathbf{x}_1$ , and the preconditioned form of (3.2) is

$$(\mathbf{A} - \rho(\mathbf{x}_1)\mathbf{M})\mathbb{P}^{-1}\mathbf{y} = \mathbf{M}\mathbf{x}_1. \quad (3.4)$$

The action of  $\mathbb{P}$  projects out the  $\hat{\mathbf{x}}_1$  component in  $\mathbf{P}$  and replaces it by the  $\hat{\mathbf{x}}_1$  component in  $\mathbf{A}$ . Note that a similar idea was presented in [33] for standard Hermitian eigenproblems. Now,

$(\mathbf{A} - \rho(\mathbf{x}_1)\mathbf{M})\mathbb{P}^{-1}\mathbf{M}\mathbf{x}_1 = \mathbf{0}$ , that is  $\mathbf{M}\mathbf{x}_1$  (the right hand side of (3.4)) is an exact eigenvector of  $(\mathbf{A} - \rho(\mathbf{x}_1)\mathbf{M})\mathbb{P}^{-1}$  with corresponding eigenvalue zero. In that case GMRES applied to (3.4) would converge in just one step (cf. (2.16) in Section 2, where  $\|\mathcal{P}\mathbf{b}\| = \|(\mathbf{I} - \mathbf{M}\mathbf{x}_1\mathbf{v}_1^H)\mathbf{M}\mathbf{x}_1\| = 0$  holds in this situation since  $\mathbf{v}_1^H\mathbf{M}\mathbf{x}_1 = 1$  and hence  $\|\mathbf{b} - \mathbf{B}\mathbf{z}_1\| = 0$  in exact arithmetic). We call (3.3) the ideal preconditioner.

The next theorem shows that modifying  $\mathbf{P}$  by a rank-one change to produce  $\mathbb{P}$  does not adversely affect the conditioning of the preconditioner. More precisely, Theorem 3.1 shows that the spectrum of  $\mathbf{A}\mathbb{P}^{-1}$  is close to the spectrum of  $\mathbf{A}\mathbf{P}^{-1}$ . Hence, if  $\mathbf{P}$  is a good preconditioner for  $\mathbf{A}$ , then so is  $\mathbb{P}$ .

**THEOREM 3.1.** *Let  $\mathbb{P}$  be given by (3.3) and assume that the block diagonalisation*

$$\mathbf{U}^{-1}\mathbf{A}\mathbf{X} = \begin{bmatrix} t_{11} & \mathbf{0}^H \\ \mathbf{0} & \mathbf{T}_{22} \end{bmatrix}, \quad \mathbf{U}^{-1}\mathbf{M}\mathbf{X} = \begin{bmatrix} s_{11} & \mathbf{0}^H \\ \mathbf{0} & \mathbf{S}_{22} \end{bmatrix}, \quad (3.5)$$

where  $\lambda_1 = t_{11}/s_{11}$  is valid. Hence  $\mathbf{x}_1 = \mathbf{X}\mathbf{e}_1$  and  $\mathbf{u}_1 = \mathbf{U}^{-H}\mathbf{e}_1$  are the right and left eigenvectors corresponding to  $\lambda_1$ . Then the matrix  $\mathbf{A}\mathbb{P}^{-1}$  has the same eigenvalues as the matrix

$$\begin{bmatrix} 1 & \frac{1}{s_{11}}\mathbf{e}_1^T\mathbf{U}^{-1}\mathbf{A}\mathbb{P}^{-1}\mathbf{U}\bar{\mathbf{I}}_{n-1} \\ \mathbf{0} & (\bar{\mathbf{I}}_{n-1}^T\mathbf{U}^{-1}\mathbf{A}\mathbf{X}\bar{\mathbf{I}}_{n-1})(\bar{\mathbf{I}}_{n-1}^T\mathbf{U}^{-1}\mathbf{P}\mathbf{X}\bar{\mathbf{I}}_{n-1})^{-1} \end{bmatrix}.$$

*Proof.* With  $\mathbf{U}^{-1}\mathbf{M}\mathbf{x}_1 = s_{11}\mathbf{e}_1$  and  $\mathbf{A}\mathbb{P}^{-1}\mathbf{M}\mathbf{x}_1 = \mathbf{M}\mathbf{x}_1$  we have that  $\mathbf{A}\mathbb{P}^{-1}$  has the same eigenvalues as

$$[\mathbf{U}^{-1}\mathbf{M}\mathbf{x}_1 \ \bar{\mathbf{I}}_{n-1}]^{-1}\mathbf{U}^{-1}\mathbf{A}\mathbb{P}^{-1}\mathbf{U}[\mathbf{U}^{-1}\mathbf{M}\mathbf{x}_1 \ \bar{\mathbf{I}}_{n-1}] = \begin{bmatrix} 1 & \frac{1}{s_{11}}\mathbf{e}_1^T\mathbf{U}^{-1}\mathbf{A}\mathbb{P}^{-1}\mathbf{U}\bar{\mathbf{I}}_{n-1} \\ \mathbf{0} & \bar{\mathbf{I}}_{n-1}^T\mathbf{U}^{-1}\mathbf{A}\mathbb{P}^{-1}\mathbf{U}\bar{\mathbf{I}}_{n-1} \end{bmatrix}.$$

Introducing  $\mathbf{U}^{-1}\mathbf{A}\mathbf{X} = \text{diag}(t_{11}, \mathbf{T}_{22})$  we can write

$$[\mathbf{U}^{-1}\mathbf{M}\mathbf{x}_1 \ \bar{\mathbf{I}}_{n-1}]^{-1}\mathbf{U}^{-1}\mathbf{A}\mathbb{P}^{-1}\mathbf{U}[\mathbf{U}^{-1}\mathbf{M}\mathbf{x}_1 \ \bar{\mathbf{I}}_{n-1}] = \begin{bmatrix} 1 & \frac{1}{s_{11}}\mathbf{e}_1^T\mathbf{U}^{-1}\mathbf{A}\mathbb{P}^{-1}\mathbf{U}\bar{\mathbf{I}}_{n-1} \\ \mathbf{0} & \mathbf{T}_{22}\bar{\mathbf{I}}_{n-1}^T\mathbf{X}^{-1}\mathbb{P}^{-1}\mathbf{U}\bar{\mathbf{I}}_{n-1} \end{bmatrix}.$$

Finally, observe that  $\mathbf{P}\mathbf{X}\bar{\mathbf{I}}_{n-1} = \mathbf{P}\mathbf{X}_2 = \mathbf{P}\mathbf{X}_2$  and hence

$$\bar{\mathbf{I}}_{n-1}^T\mathbf{X}^{-1}\mathbb{P}^{-1}\mathbf{U}\bar{\mathbf{I}}_{n-1} = \bar{\mathbf{I}}_{n-1}^T(\mathbf{U}^{-1}\mathbf{P}\mathbf{X})^{-1}\bar{\mathbf{I}}_{n-1}.$$

We have  $\mathbf{U}^{-1}\mathbf{P}\mathbf{X} = \mathbf{U}^{-1}\mathbb{P}[\mathbf{x}_1 \ \mathbf{X}_2] = \mathbf{U}^{-1}[\lambda_1\mathbf{M}\mathbf{x}_1 \ \mathbf{P}\mathbf{X}_2] = [\lambda_1 s_{11}\mathbf{e}_1 \ \mathbf{U}^{-1}\mathbf{P}\mathbf{X}_2]$ . Taking the inverse using the block structure of  $\mathbf{U}^{-1}\mathbf{P}\mathbf{X}$  and using  $\mathbf{T}_{22} = \bar{\mathbf{I}}_{n-1}^T\mathbf{U}^{-1}\mathbf{A}\mathbf{X}\bar{\mathbf{I}}_{n-1}$  gives the result.  $\square$

Note that the block diagonalisation (3.5) is possible, since we assume that  $(\mathbf{A}, \mathbf{M})$  has a simple eigenvalue (see [11, Corollary 2.4])

Theorem 3.1 shows that one eigenvalue of  $\mathbf{A}\mathbb{P}^{-1}$  is equal to one and all the other eigenvalues are equal to eigenvalues of  $(\bar{\mathbf{I}}_{n-1}^T\mathbf{U}^{-1}\mathbf{A}\mathbf{X}\bar{\mathbf{I}}_{n-1})(\bar{\mathbf{I}}_{n-1}^T\mathbf{U}^{-1}\mathbf{P}\mathbf{X}\bar{\mathbf{I}}_{n-1})^{-1}$ . Therefore if  $\mathbf{P}$  is a good preconditioner for  $\mathbf{A}$ , then  $\bar{\mathbf{I}}_{n-1}^T\mathbf{U}^{-1}\mathbf{P}\mathbf{X}\bar{\mathbf{I}}_{n-1}$  will be a good approximation to  $\bar{\mathbf{I}}_{n-1}^T\mathbf{U}^{-1}\mathbf{A}\mathbf{X}\bar{\mathbf{I}}_{n-1}$  and hence the eigenvalues of  $\mathbf{A}\mathbb{P}^{-1}$  should be clustered around one.

Now, since  $\mathbf{M}\mathbf{x}_1$  is a simple eigenvector with corresponding eigenvalue zero of  $(\mathbf{A} - \rho(\mathbf{x}_1)\mathbf{M})\mathbb{P}^{-1}$ , the block-factorisation (cf. (2.5))

$$(\mathbf{A} - \rho(\mathbf{x}_1)\mathbf{M})\mathbb{P}^{-1} = \begin{bmatrix} \bar{\mathbf{w}}_1 & \bar{\mathbf{W}}_1^\perp \\ \mathbf{0} & \bar{\mathbf{N}}_{22} \end{bmatrix} \begin{bmatrix} \bar{\mathbf{w}}_1 & \bar{\mathbf{W}}_1^\perp \end{bmatrix} \quad (3.6)$$

and block-diagonalisation (cf. (2.8))

$$(\mathbf{A} - \rho(\mathbf{x}_1)\mathbf{M})\mathbb{P}^{-1} = \begin{bmatrix} \bar{\mathbf{w}}_1 & \bar{\mathbf{W}}_2 \end{bmatrix} \begin{bmatrix} 0 & \mathbf{0}^H \\ \mathbf{0} & \bar{\mathbf{C}} \end{bmatrix} \begin{bmatrix} \bar{\mathbf{v}}_1^H \\ \bar{\mathbf{v}}_2^H \end{bmatrix}$$

exist, where  $\bar{\mathbf{w}}_1 = \mathbf{M}\mathbf{x}_1$ .

The perfect preconditioner introduced in this section is only of theoretical concern. In the next section we introduce a practical preconditioner, but the ideal preconditioner is used to prove our main result about the independence of  $k^{(i)}$  on  $i$  (Theorem 3.5).

**3.3. The tuned preconditioner.** As a practical preconditioner we propose to use

$$\mathbb{P}_i = \mathbf{P}(\mathbf{I} - \hat{\mathbf{x}}^{(i)}\hat{\mathbf{x}}^{(i)H}) + \mathbf{A}\hat{\mathbf{x}}^{(i)}\hat{\mathbf{x}}^{(i)H}, \quad \text{where } \hat{\mathbf{x}}^{(i)} = \mathbf{x}^{(i)} / \|\mathbf{x}^{(i)}\|, \quad (3.7)$$

(cf. (3.3)) which satisfies

$$\mathbb{P}_i \mathbf{x}^{(i)} = \mathbf{A}\mathbf{x}^{(i)}. \quad (3.8)$$

the same condition as used in [14] and [10]. Clearly, as  $\mathbf{x}^{(i)} \rightarrow \mathbf{x}_1$  the tuned preconditioner  $\mathbb{P}_i$  will tend to the ideal preconditioner  $\mathbb{P}$  (and  $\mathbf{M}\mathbf{x}_1$ , the right hand side of the system, will be an exact eigenvector for  $(\mathbf{A} - \rho(\mathbf{x}_1)\mathbf{M})\mathbb{P}^{-1}$ ) and we expect that, in the limit, the performance of GMRES applied to  $(\mathbf{A} - \rho(\mathbf{x}^{(i)})\mathbf{M})\mathbb{P}_i^{-1}\tilde{\mathbf{y}}^{(i)} = \mathbf{M}\mathbf{x}^{(i)}$  will be superior to its performance applied to  $(\mathbf{A} - \rho(\mathbf{x}^{(i)})\mathbf{M})\mathbf{P}^{-1}\tilde{\mathbf{y}}^{(i)} = \mathbf{M}\mathbf{x}^{(i)}$ . We measure the deviation of  $\mathbf{x}^{(i)}$  from  $\mathbf{x}_1$  using the decomposition

$$\mathbf{x}^{(i)} = \alpha^{(i)}(\mathbf{x}_1 q^{(i)} + \mathbf{X}_2 \mathbf{p}^{(i)}), \quad (3.9)$$

where  $\alpha^{(i)} := \|\mathbf{U}^{-1}\mathbf{M}\mathbf{x}^{(i)}\|$ ,  $q^{(i)} \in \mathbb{C}$ ,  $\mathbf{p}^{(i)} \in \mathbb{C}^{(n-1) \times 1}$ ,  $\mathbf{X}_2 = \mathbf{X}\bar{\mathbf{I}}_{n-1}$  and  $\bar{\mathbf{I}}_{n-1} = \begin{bmatrix} \mathbf{0}^H \\ \mathbf{I}_{n-1} \end{bmatrix} \in \mathbb{C}^{n \times (n-1)}$  with  $\mathbf{I}_{n-1}$  being the identity matrix of size  $n-1$ . Clearly  $q^{(i)}$  and  $\mathbf{p}^{(i)}$  measure how close the approximate eigenvector  $\mathbf{x}^{(i)}$  is to the sought eigenvector  $\mathbf{x}_1$ . From [11] we have

$$|\rho(\mathbf{x}^{(i)}) - \lambda_1| \leq C_1 \|\mathbf{p}^{(i)}\| \quad (3.10)$$

and

$$\frac{1}{C_1} \|\mathbf{r}^{(i)}\| \leq \|\mathbf{p}^{(i)}\| \leq C_2 \|\mathbf{r}^{(i)}\|, \quad (3.11)$$

for some constants  $C_1$  and  $C_2$ .

LEMMA 3.2. *Let (3.9) hold. Then*

$$\hat{\mathbf{x}}^{(i)}\hat{\mathbf{x}}^{(i)H} - \hat{\mathbf{x}}_1\hat{\mathbf{x}}_1^H = \mathcal{E}^{(i)}, \quad (3.12)$$

where  $\|\mathcal{E}^{(i)}\| \leq C_3 \|\mathbf{p}^{(i)}\|$  with  $C_3 := \|\mathbf{M}\|\|\mathbf{X}\|\|\mathbf{U}^{-1}\|$ . Furthermore we have

$$\|(\mathbf{A} - \rho(\mathbf{x}^{(i)})\mathbf{M})\mathbb{P}_i^{-1} - (\mathbf{A} - \rho(\mathbf{x}_1)\mathbf{M})\mathbb{P}^{-1}\| \leq \beta_1 \|\mathbf{p}^{(i)}\|, \quad (3.13)$$

where  $\beta_1$  is independent of  $i$  for large enough  $i$ .

*Proof.* We use the sine of the largest canonical angle and have (see [16, p. 76])

$$\sin \angle(\hat{\mathbf{x}}^{(i)}, \hat{\mathbf{x}}_1) = \|\mathbf{X}_1^\perp \hat{\mathbf{x}}^{(i)}\| = \|\hat{\mathbf{x}}^{(i)}\hat{\mathbf{x}}^{(i)H} - \hat{\mathbf{x}}_1\hat{\mathbf{x}}_1^H\| = \|\mathcal{E}^{(i)}\|,$$

where  $\mathbf{X}_1^\perp \in \mathbb{C}^{n \times (n-1)}$  is a matrix with orthonormal columns which are orthogonal to  $\mathbf{x}_1$  (see [17, p. 69]). Hence

$$\|\mathcal{E}^{(i)}\| = \|\mathbf{X}_1^\perp \hat{\mathbf{x}}^{(i)}\| = \left\| \mathbf{X}_1^\perp \left( \frac{\mathbf{x}^{(i)} - \alpha^{(i)} q^{(i)} \mathbf{x}_1}{\|\mathbf{x}^{(i)}\|} \right) \right\|,$$

and using (3.9) as well as  $\|\mathbf{X}_1^\perp\| = 1$ ,  $\|\mathbf{X}_2\| \leq \|\mathbf{X}\|$ ,  $\alpha^{(i)} \leq \|\mathbf{U}^{-1}\|\|\mathbf{M}\|\|\mathbf{x}^{(i)}\|$  this yields

$$\|\mathcal{E}^{(i)}\| = \left\| \mathbf{X}_1^\perp \left( \frac{\alpha^{(i)} \mathbf{X}_2 \mathbf{p}^{(i)}}{\|\mathbf{x}^{(i)}\|} \right) \right\| \leq \|\mathbf{M}\|\|\mathbf{X}\|\|\mathbf{U}^{-1}\|\|\mathbf{p}^{(i)}\| =: C_3 \|\mathbf{p}^{(i)}\|. \quad (3.14)$$

Furthermore, we have  $\mathbb{P}_i = \mathbb{P} + (\mathbf{A} - \mathbf{P})\mathcal{E}^{(i)}$  and therefore, with  $\rho(\mathbf{x}_1) = \lambda_1$ , we have

$$\begin{aligned} (\mathbf{A} - \rho(\mathbf{x}_1)\mathbf{M})\mathbb{P}^{-1} - (\mathbf{A} - \rho(\mathbf{x}^{(i)})\mathbf{M})\mathbb{P}_i^{-1} &= (\mathbf{A} - \rho(\mathbf{x}^{(i)})\mathbf{M})(\mathbb{P}^{-1} - \mathbb{P}_i^{-1}) \\ &\quad + (\rho(\mathbf{x}^{(i)}) - \lambda_1)\mathbf{M}\mathbb{P}^{-1} \\ &= (\mathbf{A} - \rho(\mathbf{x}^{(i)})\mathbf{M})\mathbb{P}^{-1}(\mathbb{P}_i - \mathbb{P})\mathbb{P}_i^{-1} \\ &\quad + (\rho(\mathbf{x}^{(i)}) - \lambda_1)\mathbf{M}\mathbb{P}^{-1} \\ &= (\mathbf{A} - \rho(\mathbf{x}^{(i)})\mathbf{M})\mathbb{P}^{-1}(\mathbf{A} - \mathbf{P})\mathcal{E}^{(i)}(\mathbb{P} + (\mathbf{A} - \mathbf{P})\mathcal{E}^{(i)})^{-1} \\ &\quad + (\rho(\mathbf{x}^{(i)}) - \lambda_1)\mathbf{M}\mathbb{P}^{-1} \end{aligned}$$

Using (3.14) we have that

$$\begin{aligned} \|\mathcal{E}^{(i)}(\mathbb{P} + (\mathbf{A} - \mathbf{P})\mathcal{E}^{(i)})^{-1}\| &\leq \frac{\|\mathbb{P}^{-1}\|}{1 - \|\mathbb{P}^{-1}(\mathbf{A} - \mathbf{P})\|} \|\mathcal{E}^{(i)}\| \\ &\leq \frac{\|\mathbb{P}^{-1}\|}{1 - \|\mathbb{P}^{-1}(\mathbf{A} - \mathbf{P})\|C_3\|\mathbf{p}^{(i)}\|} C_3\|\mathbf{p}^{(i)}\| \end{aligned}$$

can be bounded by a constant independent of  $i$  since  $\|\mathbf{p}^{(i)}\|$  is decreasing. With (3.10) and (3.11) we obtain the result for some  $\beta_1 > 0$ .  $\square$

Now, assume that  $\bar{\mathbf{w}}_1$  is a simple eigenvector of  $(\mathbf{A} - \rho(\mathbf{x}_1)\mathbf{M})\mathbb{P}^{-1}$ . Then for  $i$  large enough,  $(\mathbf{A} - \rho(\mathbf{x}^{(i)})\mathbf{M})\mathbb{P}_i^{-1}$  can be block diagonalised as

$$(\mathbf{A} - \rho(\mathbf{x}^{(i)})\mathbf{M})\mathbb{P}_i^{-1} = \begin{bmatrix} \bar{\mathbf{w}}_1^{(i)} & \bar{\mathbf{W}}_2^{(i)} \end{bmatrix} \begin{bmatrix} \bar{\mu}_1^{(i)} & \mathbf{0}^H \\ \mathbf{0} & \bar{\mathbf{C}}^{(i)} \end{bmatrix} \begin{bmatrix} \bar{\mathbf{v}}_1^{(i)} & \bar{\mathbf{V}}_2^{(i)} \end{bmatrix}^H, \quad (3.15)$$

and (3.13) with [40, Theorem 2.7, Chapter 5] gives

$$\begin{aligned} c_0\|\mathbf{p}^{(i)}\| &\leq |\bar{\mu}_1^{(i)}| \leq c_1\|\mathbf{p}^{(i)}\|, \\ \|\bar{\mathbf{C}} - \bar{\mathbf{C}}^{(i)}\| &\leq c_2\|\mathbf{p}^{(i)}\|, \\ \|\bar{\mathbf{V}}_2 - \bar{\mathbf{V}}_2^{(i)}\| &\leq c_3\|\mathbf{p}^{(i)}\|, \end{aligned}$$

where  $c_0, c_1, c_2$  and  $c_3$  are positive constants independent of  $i$ . We have the following theorem, which shows that if the tuned preconditioner is used, then the projected right hand side  $\mathcal{P}\mathbf{b}$  (see (2.16)), used in the linear system residual behaves like the eigenvalue residual (which is not the case if the standard preconditioner  $\mathbf{P}$  were used).

**THEOREM 3.3** (Right tuned preconditioner for nonsymmetric eigenproblem). *Let the assumptions of Theorem 2.2 hold and consider the solution of*

$$(\mathbf{A} - \rho(\mathbf{x}^{(i)})\mathbf{M})\mathbb{P}_i^{-1}\tilde{\mathbf{y}}^{(i)} = \mathbf{M}\mathbf{x}^{(i)}, \quad \text{where } \mathbf{y}^{(i)} = \mathbb{P}_i^{-1}\tilde{\mathbf{y}}^{(i)}. \quad (3.16)$$

with  $\mathbb{P}_i$  chosen as in (3.8). Let zero be a simple eigenvalue of  $(\mathbf{A} - \rho(\mathbf{x}^{(i)})\mathbf{M})\mathbb{P}^{-1}$  such that (3.6) holds. Furthermore, let  $|\rho(\mathbf{x}^{(i)})| > K$  for all  $i$  and some positive constant  $K$ . Then

$$\|\mathcal{P}^{(i)}\mathbf{M}\mathbf{x}^{(i)}\| \leq C_4\|\mathcal{P}^{(i)}\|\|\mathbf{r}^{(i)}\| \quad (3.17)$$

for some positive constant  $C_4$  independent of  $i$  for large enough  $i$ , where  $\mathcal{P}^{(i)} = \mathbf{I} - \bar{\mathbf{w}}_1^{(i)}\bar{\mathbf{v}}_1^{(i)H}$  and  $\bar{\mathbf{v}}_1^{(i)}$  and  $\bar{\mathbf{w}}_1^{(i)}$  are left and right eigenvectors of  $(\mathbf{A} - \rho(\mathbf{x}^{(i)})\mathbf{M})\mathbb{P}_i^{-1}$ .

*Proof.* Using  $\mathcal{P}^{(i)}\bar{\mathbf{w}}_1^{(i)} = 0$  we obtain

$$\mathcal{P}^{(i)}\mathbf{M}\mathbf{x}^{(i)} = \mathcal{P}^{(i)}(\mathbf{M}\mathbf{x}^{(i)} - \alpha\bar{\mathbf{w}}_1^{(i)}) \quad \forall \alpha \in \mathbb{C}. \quad (3.18)$$

Then, with  $\mathbf{r}^{(i)}$  given by (2.4) and condition (3.8) as well as  $\rho(\mathbf{x}^{(i)}) \neq 0$  we get

$$(\mathbf{A} - \rho(\mathbf{x}^{(i)})\mathbf{M})\mathbb{P}_i^{-1} \left( \mathbf{M}\mathbf{x}^{(i)} + \frac{\mathbf{r}^{(i)}}{\rho(\mathbf{x}^{(i)})} \right) = \frac{1}{\rho(\mathbf{x}^{(i)})}\mathbf{r}^{(i)}. \quad (3.19)$$

Hence,  $\mathbf{M}\mathbf{x}^{(i)} + \mathbf{r}^{(i)}/\rho(\mathbf{x}^{(i)})$  is an approximate eigenvector of  $(\mathbf{A} - \rho(\mathbf{x}^{(i)})\mathbf{M})\mathbb{P}_i^{-1}$  with approximate eigenvalue zero. Using (2.4) again and normalising this approximate eigenvector yields

$$(\mathbf{A} - \rho(\mathbf{x}^{(i)})\mathbf{M})\mathbb{P}_i^{-1} \frac{\mathbf{A}\mathbf{x}^{(i)}}{\|\mathbf{A}\mathbf{x}^{(i)}\|} = \frac{1}{\|\mathbf{A}\mathbf{x}^{(i)}\|} \mathbf{r}^{(i)}.$$

For  $i$  large enough (and hence  $\|\mathbf{p}^{(i)}\|$  as well as  $\|\mathbf{r}^{(i)}\|$  small enough) we can apply Proposition 2.1 to  $(\mathbf{A} - \rho(\mathbf{x}^{(i)})\mathbf{M})\mathbb{P}_i^{-1}$  with  $\hat{\mathbf{w}} = \mathbf{A}\mathbf{x}^{(i)}/\|\mathbf{A}\mathbf{x}^{(i)}\|$  to get

$$\left\| \frac{\mathbf{A}\mathbf{x}^{(i)}}{\|\mathbf{A}\mathbf{x}^{(i)}\|} - \frac{\bar{\mathbf{w}}_1^{(i)}}{\sqrt{1 + \bar{\mathbf{p}}_i^H \bar{\mathbf{p}}_i}} \right\| \leq \frac{2e^{(i)}}{\text{sep}(\bar{\mu}_1^{(i)}, \bar{\mathbf{N}}_{22}^{(i)}) - 2e^{(i)}}, \quad (3.20)$$

where  $e^{(i)} = \|\mathbf{r}^{(i)}\|/\|\mathbf{A}\mathbf{x}^{(i)}\| \leq C_5\|\mathbf{r}^{(i)}\|$  since  $\|\mathbf{A}\mathbf{x}^{(i)}\|$  is assumed to be bounded. Multiplying (3.20) by  $\|\mathbf{A}\mathbf{x}^{(i)}\|/|\rho(\mathbf{x}^{(i)})|$  we have

$$\left\| \mathbf{M}\mathbf{x}^{(i)} + \frac{\mathbf{r}^{(i)}}{|\rho(\mathbf{x}^{(i)})|} - \frac{\|\mathbf{A}\mathbf{x}^{(i)}\|}{|\rho(\mathbf{x}^{(i)})|} \frac{\bar{\mathbf{w}}_1^{(i)}}{\sqrt{1 + \bar{\mathbf{p}}_i^H \bar{\mathbf{p}}_i}} \right\| \leq \frac{2f^{(i)}}{\text{sep}(\bar{\mu}_1^{(i)}, \bar{\mathbf{N}}_{22}^{(i)}) - 2e^{(i)}}, \quad (3.21)$$

where, with  $|\rho(\mathbf{x}^{(i)})| > K$ ,

$$f^{(i)} = \frac{\|\mathbf{A}\mathbf{x}^{(i)}\|}{|\rho(\mathbf{x}^{(i)})|} e^{(i)} \leq \frac{1}{|\rho(\mathbf{x}^{(i)})|} \|\mathbf{r}^{(i)}\| \leq \frac{1}{K} \|\mathbf{r}^{(i)}\|. \quad (3.22)$$

Furthermore (3.21) yields

$$\left\| \mathbf{M}\mathbf{x}^{(i)} - \frac{\|\mathbf{A}\mathbf{x}^{(i)}\|}{|\rho(\mathbf{x}^{(i)})|} \frac{\bar{\mathbf{w}}_1^{(i)}}{\sqrt{1 + \bar{\mathbf{p}}_i^H \bar{\mathbf{p}}_i}} \right\| - \frac{\|\mathbf{r}^{(i)}\|}{|\rho(\mathbf{x}^{(i)})|} \leq \frac{2f^{(i)}}{\text{sep}(\bar{\mu}_1^{(i)}, \bar{\mathbf{N}}_{22}^{(i)}) - 2e^{(i)}}.$$

Setting  $\alpha := \frac{\|\mathbf{A}\mathbf{x}^{(i)}\|}{|\rho(\mathbf{x}^{(i)})|\sqrt{1 + \bar{\mathbf{p}}_i^H \bar{\mathbf{p}}_i}}$  in (3.18) we use this bound to obtain

$$\|\mathcal{P}^{(i)}\mathbf{M}\mathbf{x}^{(i)}\| \leq \|\mathcal{P}^{(i)}(\mathbf{M}\mathbf{x}^{(i)} - \alpha\bar{\mathbf{w}}_1^{(i)})\| \leq \|\mathcal{P}^{(i)}\| \left( \frac{2f^{(i)}}{\text{sep}(\bar{\mu}_1^{(i)}, \bar{\mathbf{N}}_{22}^{(i)}) - 2e^{(i)}} + \frac{\|\mathbf{r}^{(i)}\|}{|\rho(\mathbf{x}^{(i)})|} \right). \quad (3.23)$$

Finally, with (3.22) and  $|\rho(\mathbf{x}^{(i)})| > K$  we obtain

$$\|\mathcal{P}^{(i)}\mathbf{M}\mathbf{x}^{(i)}\| \leq \|\mathcal{P}^{(i)}\| \left( \frac{1}{K} \frac{2}{\text{sep}(\bar{\mu}_1^{(i)}, \bar{\mathbf{N}}_{22}^{(i)}) - 2C_5\|\mathbf{r}^{(i)}\|} \|\mathbf{r}^{(i)}\| + \frac{\|\mathbf{r}^{(i)}\|}{K} \right).$$

Using (3.11) as well as  $|\bar{\mu}_1^{(i)}| \leq c_1\|\mathbf{p}^{(i)}\|$  and  $\|\bar{\mathbf{N}}_{22}^{(i)} - \bar{\mathbf{N}}_{22}\| \leq c_4\|\mathbf{p}^{(i)}\|$  for appropriately chosen constants  $c_1$  and  $c_4$  and  $\|\mathbf{p}^{(i)}\|$  small enough (see [40, p. 234] and comments after (3.15)), the term

$$\frac{2}{\text{sep}(\bar{\mu}_1^{(i)}, \bar{\mathbf{N}}_{22}^{(i)}) - 2C_5\|\mathbf{r}^{(i)}\|} \leq \frac{2}{\text{sep}(\bar{\mu}_1, \bar{\mathbf{N}}_{22}) - c_1\|\mathbf{p}^{(i)}\| - c_4\|\mathbf{p}^{(i)}\| - 2C_1C_5\|\mathbf{p}^{(i)}\|},$$

can be bounded by a constant independent of  $i$  for large enough  $i$ . Hence the result (3.17) is obtained for an appropriately chosen constant  $C_4$ .  $\square$

Before proving the main result of this section we need another lemma.

**LEMMA 3.4.** *Let  $\mathbf{B}_1$  and  $\mathbf{B}_2$  be two matrices of the same dimensions and let  $\mathcal{P}_\gamma(\mathbf{B}_1)$  and  $\mathcal{P}_\gamma(\mathbf{B}_2)$  be the spectral projections onto the eigenvectors of  $\mathbf{B}_1$  and  $\mathbf{B}_2$  corresponding to*

the eigenvalues inside a closed contour  $\gamma$ . Assume that  $\|\mathbf{B}_1 - \mathbf{B}_2\| \leq \xi$  and let  $m_\gamma(\mathbf{B}_1) = \max_{\lambda \in \gamma} \|(\lambda \mathbf{I} - \mathbf{B}_1)^{-1}\|$ . If  $\xi m_\gamma(\mathbf{B}_1) < 1$  then

$$\|\mathcal{P}_\gamma(\mathbf{B}_1) - \mathcal{P}_\gamma(\mathbf{B}_2)\| \leq \frac{1}{2\pi} \mathcal{L}(\gamma) \frac{\xi m_\gamma^2(\mathbf{B}_1)}{1 - \xi m_\gamma(\mathbf{B}_1)},$$

where  $\mathcal{L}(\gamma)$  is the length of  $\gamma$ .

*Proof.* See [24] and [15, Section 8.2].  $\square$

We can finally prove the following Theorem which provides the main result of this section.

**THEOREM 3.5.** *Let the assumptions of Theorem 3.3 be satisfied. Compute  $\tilde{\mathbf{y}}_{k^{(i)}}$  satisfying the stopping criterion*

$$\|(\mathbf{A} - \rho(\mathbf{x}^{(i)})\mathbf{M})\mathbb{P}_i^{-1}\tilde{\mathbf{y}}_{k^{(i)}} - \mathbf{M}\mathbf{x}^{(i)}\| \leq \tau^{(i)} = \delta \|\mathbf{r}^{(i)}\|^\zeta \quad \delta < 1,$$

where  $\rho(\mathbf{x}^{(i)})$  is the generalised Rayleigh quotient (2.3),  $\mathbb{P}_i^{-1}$  is the tuned preconditioner and

(a)  $\zeta = 0$  is used for solves with a fixed tolerance and

(b)  $\zeta = 1$  is used for solves with a decreasing tolerance.

Then, for large enough  $i$ ,  $k^{(i)}$ , the bound on the number of inner iterations used by GMRES to compute  $\tilde{\mathbf{y}}_{k^{(i)}}$  satisfying this stopping criterion, is

(1a) bounded independently of  $i$  for  $\zeta = 0$ ,

(1b) increasing with order  $\log(\|\mathbf{r}^{(i)}\|^{-1})$  for  $\zeta = 1$ ,

where  $\mathbf{r}^{(i)}$  is the eigenvalue residual. In contrast the bound on the number of inner iterations used by GMRES to compute  $\tilde{\mathbf{y}}_{k^{(i)}}$  satisfying the stopping criterion

$$\|(\mathbf{A} - \rho(\mathbf{x}^{(i)})\mathbf{M})\mathbf{P}^{-1}\tilde{\mathbf{y}}_{k^{(i)}} - \mathbf{M}\mathbf{x}^{(i)}\| \leq \tau^{(i)} = \delta \|\mathbf{r}^{(i)}\|^\zeta \quad \delta < 1,$$

where  $\mathbf{P}$  is the standard preconditioner, is

(2a) increasing with order  $\log(\|\mathbf{r}^{(i)}\|^{-1})$  for  $\zeta = 0$ ,

(2b) increasing with order  $2 \log(\|\mathbf{r}^{(i)}\|^{-1})$  for  $\zeta = 1$ .

*Proof.* Let  $\Psi$  and  $\mathbf{E}$  be given by Proposition 2.3 applied to  $\bar{\mathbf{C}}$ . For large enough  $i$  (and hence small enough  $\|\mathbf{p}^{(i)}\|$  and  $\|\mathbf{r}^{(i)}\|$  which we use interchangeably, cf. (3.11)) decomposition (3.15) exists. By Proposition 2.5 the residual obtained after  $k^{(i)}$  iterations of GMRES starting with 0 is less than  $\tau^{(i)} = \delta \|\mathbf{r}^{(i)}\|^\zeta$  if

$$k^{(i)} \geq 1 + \frac{1}{\log |\Psi(0)|} \left( \log S_\delta \|\bar{\mu}_1^{(i)} \mathbf{I} - \bar{\mathbf{C}}^{(i)}\| \|\bar{\mathbf{V}}_2^{(i)}\| + \log \frac{\|\mathcal{P}^{(i)}\mathbf{M}\mathbf{x}^{(i)}\|}{\delta \|\bar{\mu}_1^{(i)}\| \|\mathbf{r}^{(i)}\|^\zeta} \right). \quad (3.24)$$

Using the bounds after (3.15) the argument of the first log term in the brackets can be bounded by

$$\|\bar{\mu}_1^{(i)} \mathbf{I} - \bar{\mathbf{C}}^{(i)}\| \|\bar{\mathbf{V}}_2^{(i)}\| \leq (\|\bar{\mathbf{C}}\| + c_1 \|\mathbf{p}^{(i)}\| + c_2 \|\mathbf{p}^{(i)}\|) (\|\bar{\mathbf{V}}_2\| + c_3 \|\mathbf{p}^{(i)}\|), \quad (3.25)$$

Since  $\|\mathbf{p}^{(i)}\|$  is decreasing (3.25) can be bounded independently of  $i$  for small enough  $\|\mathbf{p}^{(i)}\|$ .

For the second log term in the brackets we use Theorem 3.3 and obtain

$$\|\mathcal{P}^{(i)}\mathbf{M}\mathbf{x}^{(i)}\| \leq C_4 \|\mathcal{P}^{(i)}\| \|\mathbf{r}^{(i)}\| \quad (3.26)$$

The term  $\|\mathcal{P}^{(i)}\|$  can be bounded as follows

$$\|\mathcal{P}^{(i)}\| \leq \|\mathcal{P}^{(i)} - \mathcal{P}\| + \|\mathcal{P}\|, \quad (3.27)$$

where  $\mathcal{P} = \mathbf{I} - \bar{\mathbf{w}}_1 \bar{\mathbf{v}}_1^H$  is the spectral projection of  $(\mathbf{A} - \lambda_1 \mathbf{M})\mathbb{P}^{-1}$  onto  $\bar{\mathbf{W}}_2$ . For small enough  $\|\mathbf{p}^{(i)}\|$  we apply Lemma 3.4 with  $\mathbf{B}_1 = (\mathbf{A} - \lambda_1 \mathbf{M})\mathbb{P}^{-1}$  and  $\mathbf{B}_2 = (\mathbf{A} - \rho(\mathbf{x}^{(i)})\mathbf{M})\mathbb{P}_i^{-1}$  and use (3.13). Taking  $\gamma$  as a circle of centre zero and radius  $\|\mathbf{p}^{(i)}\|$ , we obtain

$$\|\mathcal{P}^{(i)} - \mathcal{P}\| \leq \frac{\beta_1 m_\gamma^2((\mathbf{A} - \lambda_1 \mathbf{M})\mathbb{P}^{-1}) \|\mathbf{p}^{(i)}\|}{1 - \beta_1 m_\gamma((\mathbf{A} - \lambda_1 \mathbf{M})\mathbb{P}^{-1}) \|\mathbf{p}^{(i)}\|} \|\mathbf{p}^{(i)}\|.$$

Since  $\|\mathbf{p}^{(i)}\|$  is decreasing we have for a small enough  $\|\mathbf{p}^{(i)}\|$

$$\frac{\beta_1 m_\gamma^2 ((\mathbf{A} - \lambda_1 \mathbf{M}) \mathbb{P}^{-1}) \|\mathbf{p}^{(i)}\|}{1 - \beta_1 m_\gamma ((\mathbf{A} - \lambda_1 \mathbf{M}) \mathbb{P}^{-1}) \|\mathbf{p}^{(i)}\|},$$

can be bounded independent of  $i$ . Hence (3.26) and (3.27) imply  $\|\mathcal{P}^{(i)} \mathbf{M} \mathbf{x}^{(i)}\| \leq C_6 \|\mathcal{P}\| \|\mathbf{r}^{(i)}\|$  for some constant  $C_6$ . Together with the bounds on  $|\bar{\mu}_1^{(i)}|$  and equivalence of  $\|\mathbf{p}^{(i)}\|$  and  $\|\mathbf{r}^{(i)}\|$  we obtain the results for  $\zeta = 0$  and  $\zeta = 1$  (parts (1a) and (1b)), respectively. For the standard preconditioner we have that  $\|\mathcal{P}^{(i)} \mathbf{M} \mathbf{x}^{(i)}\|$  is bounded by a constant independent of  $\|\mathbf{r}^{(i)}\|$  and hence from (3.24) we obtain parts (2a) and (2b), respectively.  $\square$

Theorem 3.5 gives upper bounds on the iteration numbers only and for that reason the results are qualitative rather than quantitative. For quantitative results we refer to Section 4.

The following theorem provides a method to implement the tuning concept efficiently.

**LEMMA 3.6** (Implementation of  $\mathbb{P}_i^{-1}$ ). *Let  $\mathbf{x}^{(i)}$  be the approximate eigenvector obtained from the  $i$ th iteration of inexact inverse iteration. Then (3.7) satisfies (3.8) and, assuming  $\mathbf{x}^{(i)H} \mathbf{P}^{-1} \mathbf{A} \mathbf{x}^{(i)} \neq 0$ , we have*

$$\mathbb{P}_i^{-1} = \left( \mathbf{I} - \frac{(\mathbf{P}^{-1} \mathbf{A} \mathbf{x}^{(i)} - \mathbf{x}^{(i)}) \mathbf{x}^{(i)H}}{\mathbf{x}^{(i)H} \mathbf{P}^{-1} \mathbf{A} \mathbf{x}^{(i)}} \right) \mathbf{P}^{-1}.$$

*Proof.* Sherman-Morrison Formula.  $\square$

Note that only one extra back solve  $\mathbf{P}^{-1} \mathbf{A} \mathbf{x}^{(i)}$  per outer iteration is necessary for the implementation of the tuned preconditioner, which can be computed before the actual inner iteration. All further extra costs are inner products.

**REMARK 3.7** (Left tuned preconditioner). *For left preconditioning, namely*

$$\mathbb{P}_i^{-1} (\mathbf{A} - \rho(\mathbf{x}^{(i)}) \mathbf{M}) \mathbf{y}^{(i)} = \mathbb{P}_i^{-1} \mathbf{M} \mathbf{x}^{(i)} \quad (3.28)$$

*the tuning works similarly. For  $\rho(\mathbf{x}^{(i)}) \neq 0$  we have*

$$\mathbb{P}_i^{-1} (\mathbf{A} - \rho(\mathbf{x}^{(i)}) \mathbf{M}) \left( \mathbb{P}_i^{-1} \mathbf{M} \mathbf{x}^{(i)} + \frac{\mathbb{P}_i^{-1} \mathbf{r}^{(i)}}{\rho(\mathbf{x}^{(i)})} \right) = \frac{\mathbb{P}_i^{-1} \mathbf{r}^{(i)}}{\rho(\mathbf{x}^{(i)})},$$

so that  $\mathbb{P}_i^{-1} \mathbf{M} \mathbf{x}^{(i)} + \frac{\mathbb{P}_i^{-1} \mathbf{r}^{(i)}}{\rho(\mathbf{x}^{(i)})}$  is an approximate eigenvector of  $\mathbb{P}_i^{-1} (\mathbf{A} - \rho(\mathbf{x}^{(i)}) \mathbf{M})$ . Hence, as  $\|\mathbf{r}^{(i)}\| \rightarrow 0$ , the right hand side of (3.28) tends to an eigenvector of the iteration matrix; a similar situation to the right preconditioned case.

**REMARK 3.8.** *Note that as a consequence of (3.8) we have  $(\mathbf{A} \mathbb{P}_i^{-1}) \mathbf{A} \mathbf{x}^{(i)} = \mathbf{A} \mathbf{x}^{(i)}$ , that is,  $\mathbf{A} \mathbf{x}^{(i)}$  is an eigenvector of  $\mathbf{A} \mathbb{P}_i^{-1}$  corresponding to eigenvalue 1, which reproduces the behaviour of the perfect preconditioner (see Theorem 3.1)*

**REMARK 3.9** (Fixed shifts). *The above theory also holds for fixed shifts. In that case we have to decrease the solve tolerance and the number of inner iteration increases logarithmically if a standard preconditioner is applied and is bounded independently of  $i$  if the tuned preconditioner is applied. The proof is similar to the one for Theorem 3.5.*

**4. Numerical examples.** We investigate the linearised stability of fluid flows governed by the steady-state Navier-Stokes equations. Here we merely summarise the main points and refer to [19] for further details.

Suppose that a velocity field  $\mathbf{w}$  has been computed for some particular parameter value. To assess its stability the PDE eigenproblem

$$\left. \begin{aligned} -\epsilon \Delta \mathbf{u} + \mathbf{w} \cdot \nabla \mathbf{u} + \mathbf{u} \cdot \nabla \mathbf{w} + \nabla p &= \lambda \mathbf{u} \\ \nabla \cdot \mathbf{u} &= 0, \end{aligned} \right\} \quad (4.1)$$

for some eigenvalue  $\lambda \in \mathbb{C}$  and nontrivial eigenfunction  $(\mathbf{u}, p)$ , satisfying suitable homogeneous boundary conditions needs to be solved. The parameter  $\epsilon$  is the viscosity, which is inversely proportional to the Reynolds number  $Re$ . The eigenfunction  $(\mathbf{u}, p)$  consists of the velocity  $\mathbf{u}$  and the pressure  $p$  both defined on a 2D computational domain. Discretisation of (4.1) with mixed finite elements yields the finite dimensional generalised eigenvalue problem (2.1) where  $\mathbf{x} = (\mathbf{U}_1^T, \mathbf{U}_2^T, \mathbf{P}^T)^T$  is a vector of  $n$  degrees of freedom approximating  $(u_1, u_2, p)^T$ , and the matrices  $\mathbf{A}$  and  $\mathbf{M}$  take the form:

$$\mathbf{A} = \begin{bmatrix} \mathbf{F}_{11} & \mathbf{F}_{12} & \mathbf{B}_1^T \\ \mathbf{F}_{21} & \mathbf{F}_{22} & \mathbf{B}_2^T \\ \mathbf{B}_1 & \mathbf{B}_2 & 0 \end{bmatrix}, \quad \mathbf{M} = \begin{bmatrix} \mathbf{M}_u & 0 & 0 \\ 0 & \mathbf{M}_u & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad (4.2)$$

where  $\mathbf{A}$  is nonsymmetric,  $\mathbf{M}$  is positive semi-definite, both are large and sparse, see [21, 19].

We seek the rightmost eigenvalue which, for the case considered here with  $Re = 25.0$ , is complex and near the imaginary axis. As iterative solver we use right preconditioned FGMRES [31] where the preconditioner is given by the block preconditioner suggested by Elman [7, 8].

We use systems with 6734, 27294 or 61678 degrees of freedom and Reynolds number  $Re = 25.0$ . For the inner solve we test three different approaches:

- solves with variable shift and decreasing tolerance,
- solves with fixed shift and decreasing tolerance,
- solves with variable shift and fixed tolerance.

For all three methods we apply both the standard and the tuned version of the block preconditioner.

TABLE 4.1

Number of inner iterations for problem with  $Re = 25.0$  and 6734 DOF if Rayleigh quotient shift and decreasing tolerance is applied.

Var. shift Decr. tol.	STANDARD PRECONDITIONER		TUNED PRECONDITIONER	
	Inner it. $k^{(i)}$	Eigenvalue residual $\ \mathbf{r}^{(i)}\ $	Inner it. $k^{(i)}$	Eigenvalue residual $\ \mathbf{r}^{(i)}\ $
Outer it. $i$				
1	236	0.2963e+00	206	0.3003e+00
2	268	0.2892e-02	167	0.3059e-02
3	388	0.8455e-04	285	0.1021e-03
4	487	0.3707e-07	342	0.5524e-07
5	569	0.3253e-11	337	0.3639e-11
Total it.	1948		1351	
Solve time	188.35		118.29	

Table 4.1 shows the iteration numbers if a Rayleigh quotient shift and a decreasing tolerance is applied. This leads to quadratic convergence of the overall algorithm, which can be seen from the 3rd and 5th column of the table. For the standard preconditioner as well as for the tuned preconditioner the number of inner iterations  $k^{(i)}$  increases per outer iteration  $i$ , however, for the tuned preconditioner the increase is less rapid, namely (from Theorem 3.5) of order  $2 \log(\|\mathbf{p}^{(i)}\|^{-1})$  for the standard preconditioner and of order  $\log(\|\mathbf{p}^{(i)}\|^{-1})$  for the tuned preconditioner, where  $\|\mathbf{p}^{(i)}\|$  is proportional to the norm of the eigenvalue residual  $\|\mathbf{r}^{(i)}\|$ . Hence the total number of iterations for inexact RQI with the tuned preconditioner is less than with the standard preconditioner. Here the saving is over 30 per cent (1351 versus 1948 iterations with a corresponding saving in computation time).

Table 4.2 shows the iteration numbers for fixed shift solves with a decreasing tolerance. The overall linear convergence can clearly be seen from the columns 3 and 5. According to Remark 3.9, the number of inner iterations for the standard preconditioner increases logarithmically whereas for the tuned preconditioner the number of inner iterations per outer iteration remains approximately constant. This behaviour can indeed be observed from columns 2 and 4 in Table 4.2. The total savings in iteration numbers is over 50 per cent if the tuned preconditioner is applied.

TABLE 4.2

Number of inner iterations for problem with  $Re = 25.0$  and 6734 DOF if fixed shift and decreasing tolerance is applied.

Fixed shift Decr. tol.	STANDARD PRECONDITIONER		TUNED PRECONDITIONER	
	Outer it. $i$	Inner it. $k^{(i)}$	Eigenvalue residual $\ \mathbf{r}^{(i)}\ $	Inner it. $k^{(i)}$
1	236	0.2963e+00	206	0.3003e+00
2	252	0.5041e-02	157	0.5145e-02
3	333	0.1254e-02	230	0.1225e-02
4	360	0.2343e-03	199	0.2260e-03
5	392	0.3065e-04	189	0.2929e-04
6	421	0.3131e-05	184	0.2934e-05
7	447	0.2696e-06	180	0.2349e-06
8	483	0.2040e-07	179	0.1490e-07
9	516	0.1341e-08	182	0.7313e-09
10	543	0.8150e-10	187	0.3434e-10
Total it.	3983		1903	
Solve time	360.38		157.65	

TABLE 4.3

Number of inner iterations for problem with  $Re = 25.0$  and 6734 DOF if Rayleigh quotient shift and fixed tolerance is applied.

Var. shift Fixed tol.	STANDARD PRECONDITIONER		TUNED PRECONDITIONER	
	Outer it. $i$	Inner it. $k^{(i)}$	Eigenvalue residual $\ \mathbf{r}^{(i)}\ $	Inner it. $k^{(i)}$
1	209	0.2956e+00	183	0.2984e+00
2	252	0.3962e-02	170	0.3955e-02
3	292	0.7056e-03	138	0.8659e-03
4	377	0.7723e-05	185	0.3534e-04
5	424	0.3327e-06	215	0.6833e-06
6	463	0.1483e-07	187	0.2681e-07
7	514	0.3302e-09	198	0.7628e-09
8	548	0.1155e-10	200	0.1903e-10
Total it.	3079		1484	
Solve time	291.77		105.84	

The iteration numbers for Rayleigh quotient shift solves with fixed tolerance are shown in Table 4.3. Overall linear convergence of the eigenvalue residuals can readily be observed. For the standard preconditioner the number of inner iterations per outer iteration increases logarithmically whereas for the tuned preconditioner the number of inner iterations remains approximately constant. This behaviour is expected from Theorem 3.5. Applying the tuned preconditioner saves more than 50 per cent in total number of iterations and time to solve the problem.

Table 4.4 summarises the total number of iterations for different problem sizes when the standard or the tuned preconditioner is applied within the inner solves. The outer iteration is stopped once the eigenvalue residual is smaller than  $10^{-10}$ . All problems show a significant improvement both in terms of iteration numbers and computation time when the tuned preconditioner is used. For 61678 degrees of freedom (last two columns) inexact inverse iteration with the standard preconditioner stagnates at an eigenvalue residual of order  $10^{-9}$  whereas inexact inverse iteration with a tuned preconditioner converges to a smaller eigenvalue residual and, moreover, the convergence is achieved with many fewer iterations (that is fewer matrix-vector products).

### 5. Equivalence between preconditioned JD and preconditioned inexact RQI.

In this section we compare inexact RQI with a special tuned preconditioning to the simplified preconditioned JD method and show that they are in fact equivalent. This is a significant

TABLE 4.4

Total number of inner iterations until eigenvalue residual  $\|\mathbf{r}^{(i)}\|$  is smaller than  $10^{-10}$  for  $Re = 25.0$  and different degrees of freedom if standard or tuned preconditioner are applied. (\*) represents stagnation at eigenvalue residual  $\|\mathbf{r}^{(i)}\| = 10^{-9}$

Degrees of freedom		6734		27294		61678	
		no tuning	tuning	no tuning	tuning	no tuning	tuning
<i>Var. shift</i>	total it.	1948	1351	2300	1561	>2742	1921
	<i>Decr. tol.</i> solve time	188.35	118.29	365.91	251.47	*	542.73
<i>Fixed shift</i>	total it.	3983	1903	4727	2119	>4777	2653
	<i>Decr. tol.</i> solve time	360.38	157.65	808.17	303.48	*	640.41
<i>Var. shift</i>	total it.	3079	1484	3039	1735	>2792	2412
	<i>Fixed tol.</i> solve time	291.77	105.64	532.14	231.05	*	564.07

extension to the theory for the standard eigenproblem in [12] and helps to provide a convergence theory for the preconditioned JD method. The proofs use ideas from [37, 23, 12]. In this section we drop the index  $i$  since we only consider the inner iteration.

For the generalised eigenproblem, Sleijpen et al. [35] introduced a JD type method which we describe briefly. Assume  $(\rho(\mathbf{x}), \mathbf{x})$  is an approximation to  $(\lambda_1, \mathbf{x}_1)$  and introduce the orthogonal projections

$$\Pi_1 = \mathbf{I} - \frac{\mathbf{M}\mathbf{x}\mathbf{w}^H}{\mathbf{w}^H\mathbf{M}\mathbf{x}} \quad \text{and} \quad \Pi_2 = \mathbf{I} - \frac{\mathbf{x}\mathbf{u}^H}{\mathbf{u}^H\mathbf{x}}, \quad (5.1)$$

where  $\mathbf{u}^H\mathbf{x} \neq 0$  and  $\mathbf{w}^H\mathbf{M}\mathbf{x} \neq 0$ . With  $\mathbf{r} = \mathbf{A}\mathbf{x} - \rho(\mathbf{x})\mathbf{M}\mathbf{x}$  solve the correction equation

$$\Pi_1(\mathbf{A} - \rho(\mathbf{x})\mathbf{M})\Pi_2\mathbf{s} = -\mathbf{r}, \quad \text{where } \mathbf{s} \perp \mathbf{u}, \quad (5.2)$$

for  $\mathbf{s}$ . This is the JD correction equation which maps  $\text{span}\{\mathbf{u}\}^\perp$  onto  $\text{span}\{\mathbf{w}\}^\perp$ . An improved guess for the eigenvector is given by a suitably normalised  $\mathbf{x} + \mathbf{s}$ . Sleijpen et al. [35, Theorem 3.2] have shown that if (5.2) is solved exactly then  $\mathbf{x}^{(i)}$  converges quadratically to the right eigenvector  $\mathbf{x}_1$ .

Several choices for the projectors  $\Pi_1$  and  $\Pi_2$  are possible, depending on the choice of  $\mathbf{w}$  and  $\mathbf{u}$ . We show that if a certain tuned preconditioner is used in inexact RQI applied to the generalised eigenproblem then this method is equivalent to the simple JD method with correction equation (5.2) and a standard preconditioner. From now on we assume without loss of generality that  $\mathbf{x}$  is normalised such that  $\mathbf{x}^H\mathbf{u} = 1$ .

Let  $\mathbf{P}$  be any preconditioner for  $\mathbf{A} - \rho(\mathbf{x})\mathbf{M}$ , then a system of the form

$$(\mathbf{A} - \rho(\mathbf{x})\mathbf{M})\mathbf{P}^{-1}\tilde{\mathbf{y}} = \mathbf{M}\mathbf{x}, \quad \text{with } \mathbf{y} = \mathbf{P}^{-1}\tilde{\mathbf{y}} \quad (5.3)$$

has to be solved at each inner iteration for inexact RQI whilst a system of the form

$$\left(\mathbf{I} - \frac{\mathbf{M}\mathbf{x}\mathbf{w}^H}{\mathbf{w}^H\mathbf{M}\mathbf{x}}\right)(\mathbf{A} - \rho(\mathbf{x})\mathbf{M})(\mathbf{I} - \mathbf{x}\mathbf{u}^H)\tilde{\mathbf{P}}^\dagger\tilde{\mathbf{s}} = -\mathbf{r}, \quad \text{with } \mathbf{s} = \mathbf{P}^{-1}\tilde{\mathbf{s}} \quad (5.4)$$

needs to be solved at each inner iteration of the simplified JD method, where the preconditioner is restricted such that

$$\tilde{\mathbf{P}} = \left(\mathbf{I} - \frac{\mathbf{M}\mathbf{x}\mathbf{w}^H}{\mathbf{w}^H\mathbf{M}\mathbf{x}}\right)\mathbf{P}(\mathbf{I} - \mathbf{x}\mathbf{u}^H).$$

Following a similar analysis as in Section [35, Proposition 7.2] and introducing the projector  $\Pi_2^{\mathbf{P}}$  given by

$$\Pi_2^{\mathbf{P}} = \mathbf{I} - \frac{\mathbf{P}^{-1}\mathbf{M}\mathbf{x}\mathbf{u}^H}{\mathbf{u}^H\mathbf{P}^{-1}\mathbf{M}\mathbf{x}}, \quad (5.5)$$

we have that a Krylov solve applied to (5.4) generates the subspace

$$\text{span}\{\mathbf{r}, \Pi_1(\mathbf{A} - \rho(\mathbf{x})\mathbf{M})\Pi_2^{\mathbf{P}}\mathbf{P}^{-1}\mathbf{r}, (\Pi_1(\mathbf{A} - \rho(\mathbf{x})\mathbf{M})\Pi_2^{\mathbf{P}}\mathbf{P}^{-1})^2\mathbf{r}, \dots\},$$

whereas the solution  $\tilde{\mathbf{y}}$  to (5.3) lies in the Krylov subspace

$$\text{span}\{\mathbf{M}\mathbf{x}, (\mathbf{A} - \rho(\mathbf{x})\mathbf{M})\mathbf{P}^{-1}\mathbf{M}\mathbf{x}, ((\mathbf{A} - \rho(\mathbf{x})\mathbf{M})\mathbf{P}^{-1})\mathbf{M}\mathbf{x}, \dots\}.$$

These subspaces are not equal, but if a tuned version of the preconditioner is applied within the inner solve arising for inverse iteration then we can show an equivalence between the inexact simplified JD method and inexact RQI, as follows.

Ideally we would like to compare the JD method with a standard preconditioner to inexact RQI with a tuned preconditioner which satisfies  $\mathbb{P}\mathbf{x} = \mathbf{A}\mathbf{x}$ , however, equivalence can only be shown for a slightly different choice of  $\mathbb{P}$ . Numerical experiments will show that the difference is minor. Instead of tuning using  $\mathbb{P}\mathbf{x} = \mathbf{A}\mathbf{x}$  we use

$$\mathbb{P}\mathbf{x} = \mathbf{M}\mathbf{x}, \quad (5.6)$$

which, assuming the normalisation  $\mathbf{u}^H\mathbf{x} = 1$ , is achieved by  $\mathbb{P} = \mathbf{P} + (\mathbf{M} - \mathbf{P})\mathbf{x}\mathbf{u}^H$ . Using the Sherman-Morrison formula and assuming  $\mathbf{u}^H\mathbf{P}^{-1}\mathbf{M}\mathbf{x} \neq 0$  its inverse  $\mathbb{P}^{-1}$  is given by

$$\mathbb{P}^{-1} = \mathbf{P}^{-1} - \frac{(\mathbf{P}^{-1}\mathbf{M}\mathbf{x} - \mathbf{x})\mathbf{u}^H\mathbf{P}^{-1}}{\mathbf{u}^H\mathbf{P}^{-1}\mathbf{M}\mathbf{x}}. \quad (5.7)$$

We can then generalise results in [12, Lemmata 1 and 3] as follows.

LEMMA 5.1. *Consider vectors  $\mathbf{w}$  and  $\mathbf{u}$  for which  $\mathbf{u}^H\mathbf{x} \neq 0$  and  $\mathbf{w}^H\mathbf{M}\mathbf{x} \neq 0$ . Let  $\mathbf{x}$  be a vector normalised such that  $\mathbf{x}^H\mathbf{u} = 1$  and let  $\rho(\mathbf{x}) = \frac{\mathbf{w}^H\mathbf{A}\mathbf{x}}{\mathbf{w}^H\mathbf{M}\mathbf{x}}$  be the generalised Rayleigh quotient. Let  $\mathbf{P}$  be a preconditioner for  $\mathbf{A}$  and let  $\Pi_1$  be defined as in (5.5). Further, let the tuned preconditioner  $\mathbb{P}$  satisfy (5.6) and let  $\mathbf{r} = \mathbf{A}\mathbf{x} - \rho(\mathbf{x})\mathbf{M}\mathbf{x} = \Pi_1\mathbf{r}$ . Introduce the subspaces*

$$\mathcal{K}_k = \text{span}\{\mathbf{M}\mathbf{x}, \mathbf{A}\mathbb{P}^{-1}\mathbf{M}\mathbf{x}, (\mathbf{A}\mathbb{P}^{-1})^2\mathbf{M}\mathbf{x}, \dots, (\mathbf{A}\mathbb{P}^{-1})^k\mathbf{M}\mathbf{x}\},$$

$$\mathcal{L}_k = \text{span}\{\mathbf{M}\mathbf{x}, \mathbf{r}, \Pi_1\mathbf{A}\Pi_2^{\mathbf{P}}\mathbf{P}^{-1}\mathbf{r}, \dots, (\Pi_1\mathbf{A}\Pi_2^{\mathbf{P}}\mathbf{P}^{-1})^{k-1}\mathbf{r}\}$$

and

$$\mathcal{M}_k = \text{span}\{\mathbf{M}\mathbf{x}, \mathbf{r}, \Pi_1\mathbf{A}\Pi_2^{\mathbf{P}}\mathbf{P}^{-1}\mathbf{r}, \dots, (\Pi_1\mathbf{A}\Pi_2^{\mathbf{P}}\mathbf{P}^{-1})^{k-1}\mathbf{r}\}$$

Then, for every  $k \geq 1$ , we have  $\mathcal{K}_k = \mathcal{L}_k = \mathcal{M}_k$ .

*Proof.* The proof is very similar to the proofs of Lemmata 1 and 3 in [12].  $\square$

Finally, [12, Theorem 4] can be generalised to the following result:

THEOREM 5.2. *Let the assumptions of Lemma 5.1 hold with  $\mathbf{w} = \mathbf{M}\mathbf{x}$ . Let  $\mathbf{y}_{k+1}^{RQ}$  and  $\mathbf{s}_k^{JD}$  be the approximate solutions to*

$$(\mathbf{A} - \rho(\mathbf{x})\mathbf{M})\mathbb{P}^{-1}\tilde{\mathbf{y}} = \mathbf{M}\mathbf{x}, \quad \text{with } \mathbf{y} = \mathbb{P}^{-1}\tilde{\mathbf{y}}, \quad (5.8)$$

and

$$\left(\mathbf{I} - \frac{\mathbf{M}\mathbf{x}\mathbf{x}^H\mathbf{M}^H}{\mathbf{x}^H\mathbf{M}^H\mathbf{M}\mathbf{x}}\right)(\mathbf{A} - \rho(\mathbf{x})\mathbf{M})(\mathbf{I} - \mathbf{x}\mathbf{u}^H)\tilde{\mathbf{P}}^\dagger\tilde{\mathbf{s}} = -\mathbf{r}, \quad \text{with } \mathbf{s} = \tilde{\mathbf{P}}^\dagger\tilde{\mathbf{s}}, \quad (5.9)$$

respectively, obtained by  $k+1$  ( $k$ , respectively) steps of the same Galerkin-Krylov method with starting vector zero. Then there exists a constant  $c \in \mathbb{C}$  such that

$$\mathbf{y}_{k+1}^{RQ} = c(\mathbf{x} + \mathbf{s}_k^{JD}). \quad (5.10)$$

*Proof.* The argument is similar to the one for [12, Theorem 4].  $\square$

Note that there is no restriction on the choice of  $\mathbf{u}$  used to normalise  $\mathbf{x}$ . Indeed, we give results for two different choices in the following example.

EXAMPLE 5.3. Consider a generalised eigenproblem  $\mathbf{Ax} = \lambda\mathbf{Mx}$ , where the matrix  $\mathbf{A}$  is given by the matrix `sherman5.mtx` from the Matrix Market library [4]. The matrix  $\mathbf{M}$  is given by a tridiagonal matrix with entries  $2/3$  on the diagonal and entries  $1/6$  on the sub- and superdiagonal. We seek the eigenvector belonging to the smallest eigenvalue, use a fixed shift  $\sigma = 0$  and an initial starting guess of all ones. We compare inexact RQI with simplified inexact JD method and investigate the following approaches to preconditioning:

- (a) no preconditioner is used for the inner iteration.
- (b) a standard preconditioner is used for the inner iteration.
- (c) a tuned preconditioner with  $\mathbb{P}\mathbf{x} = \mathbf{Mx}$  is used for the inner iteration.

We use FOM as a solver with incomplete LU factorisation with drop tolerance 0.005 as preconditioner where appropriate. Furthermore, we carry out exactly 10 steps of preconditioned FOM for the inner solve in the simplified JD method, while precisely 11 steps of preconditioned FOM are taken for each inner solve in inexact RQI. We do this in order to verify (5.10). We also restrict the number of total outer solves to 20. Furthermore, we use two different choices for  $\mathbf{u}$ , namely

- (i) a constant  $\mathbf{u}$  given by a vector of all ones,
- (ii) a variable  $\mathbf{u}^{(i)}$  given by  $\mathbf{u}^{(i)} = \mathbf{M}^H \mathbf{Mx}^{(i)}$ , which changes at each outer iteration.

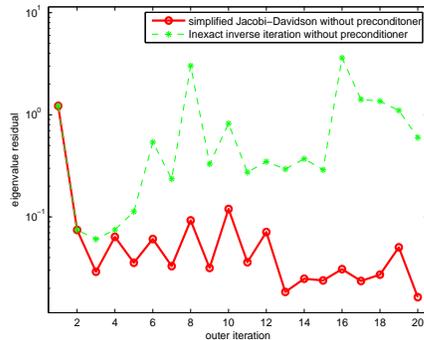


FIG. 5.1. Convergence history of the eigenvalue residuals for Example 5.3, case (a) and a constant  $\mathbf{u}$

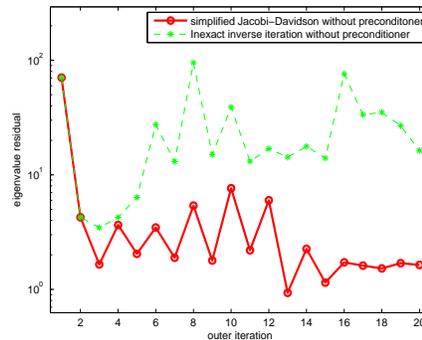


FIG. 5.2. Convergence history of the eigenvalue residuals for Example 5.3, case (a) and a variable  $\mathbf{u}^{(i)} = \mathbf{M}^H \mathbf{Mx}^{(i)}$

Figures 5.1 to 5.6 show the results for Example 5.3. We can make two observations: first of all, we see that only for case (c), when the tuned preconditioner is applied to inexact RQI and a standard preconditioner is used with a simplified JD method, the convergence history of the eigenvalue residuals is the same (see Figures 5.5 and 5.6), as we would expect from Theorem 5.2. If no preconditioner is used (see Figures 5.1 and 5.2) or a standard preconditioner is applied (see Figures 5.3 and 5.4), then inexact RQI and the simplified JD are not equivalent. Secondly, we can use any vector  $\mathbf{u}$  within the JD method (see Figures on the left compared to Figures on the right) and will get the same results. In particular, the choice of  $\mathbf{u}^{(i)} = \mathbf{x}^{(i)}$ , which is not presented here, leads to similar outcome.

This generalisation has two practical implications: Firstly, if inexact RQI is used with a tuned preconditioner we obtain the same results as in the inexact simplified JD method with a standard preconditioner. Hence, if we use inexact RQI with a tuned preconditioner the choice of  $\Pi_1$  and  $\Pi_2$  does not have to be taken care of, whereas for the simplified JD method we have to consider choices for  $\Pi_1$  and  $\Pi_2$  (and hence choices for  $\mathbf{w}$  and  $\mathbf{u}$ ) and the implication on the overall convergence rate of the algorithm ([35] indicate that only particular choices of  $\mathbf{w}$  and

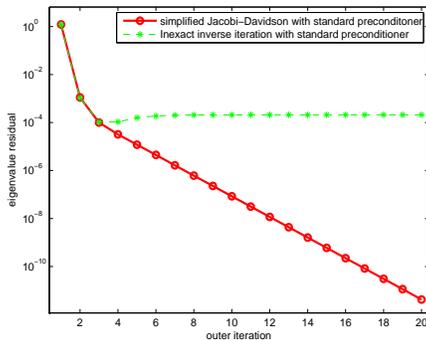


FIG. 5.3. Convergence history of the eigenvalue residuals for Example 5.3, case (b) and a constant  $\mathbf{u}$

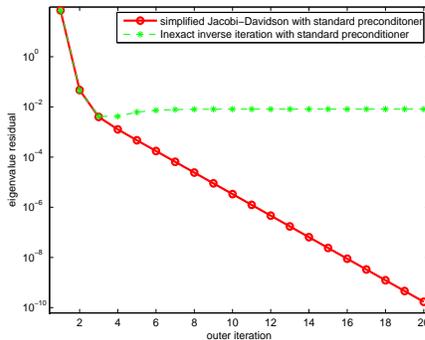


FIG. 5.4. Convergence history of the eigenvalue residuals for Example 5.3, case (b) and a variable  $\mathbf{u}^{(i)} = \mathbf{M}^H \mathbf{M} \mathbf{x}^{(i)}$

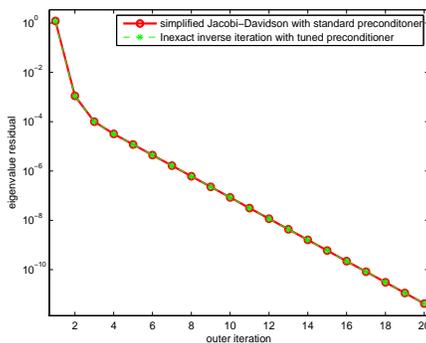


FIG. 5.5. Convergence history of the eigenvalue residuals for Example 5.3, case (c) and a constant  $\mathbf{u}$

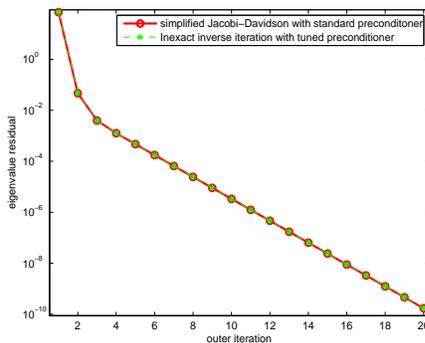


FIG. 5.6. Convergence history of the eigenvalue residuals for Example 5.3, case (c) and a variable  $\mathbf{u}^{(i)} = \mathbf{M}^H \mathbf{M} \mathbf{x}^{(i)}$

$\mathbf{u}$  lead to quadratic convergence of the eigenvectors). For inexact RQI we simply use (3.7) or (5.7) with any value for  $\mathbf{u}$  and obtain quadratic convergence for an appropriately chosen solve tolerance. Another implication is that tuning the preconditioner does not have any effect on the JD method.

Finally, we show a numerical example where we use GMRES instead of FOM, which supports the result that simplified JD with a standard preconditioner and inexact RQI with a tuned preconditioner are in fact very closely related. This is not surprising since using [6, Theorem 2], the residuals for FOM and GMRES are related to each other in the sense that the FOM residual norm and the GMRES residual norm will be approximately equal to each other if the GMRES residual norm is reduced at each step. Hence, similar results to the ones in Example 5.3, where FOM was applied are expected for GMRES, although exact equivalence of inexact RQI with a tuned preconditioner and inexact simplified JD method with a standard preconditioner is only shown for a Galerkin-Krylov method such as FOM.

EXAMPLE 5.4. Consider the generalised eigenproblem  $\mathbf{A} \mathbf{x} = \lambda \mathbf{M} \mathbf{x}$  arising from the Galerkin-FEM discretisation on regular triangular elements with piecewise linear functions of the convection-diffusion operator

$$-\Delta u + 5u_x + 5u_y = \lambda u \quad \text{on } (0, 1)^2.$$

We use a  $32 \times 32$  grid leading to 961 degrees of freedom. We seek the smallest eigenvalue, which in this case is given by  $\lambda_1 \approx 32.15825765$ .

We apply *inexact RQI* as well as *simplified JD with Rayleigh quotient shift* and a fixed solve tolerance  $\tau = 0.2$  to this problem. For both methods we use the same starting guess and the overall computation stops once  $\|\mathbf{r}^{(i)}\| < 10^{-12}$ . We apply preconditioned GMRES (instead of FOM) within the inner solve of each method. Both the simplified JD approach and the *inexact RQI* are tested with a standard preconditioner and a tuned preconditioner, where here, tuning is applied with  $\mathbb{P}\mathbf{x} = \mathbf{A}\mathbf{x}$  (instead of  $\mathbb{P}\mathbf{x} = \mathbf{M}\mathbf{x}$  as in Theorem 5.2).

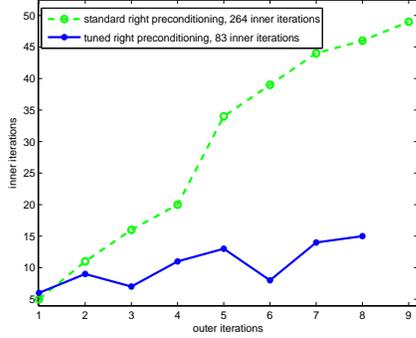


FIG. 5.7. Number of inner iterations against outer iterations with standard and tuned preconditioning (Example 5.4) when using *inexact RQ* iteration

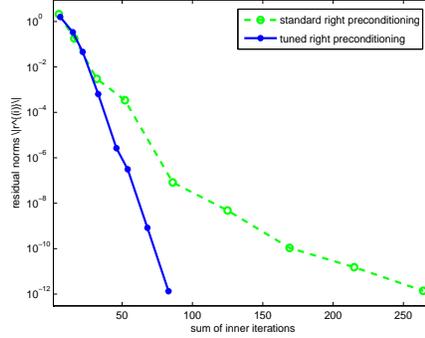


FIG. 5.8. Residual norms vs total number of inner iterations with standard and tuned preconditioning (Example 5.4) when using *inexact RQ* iteration

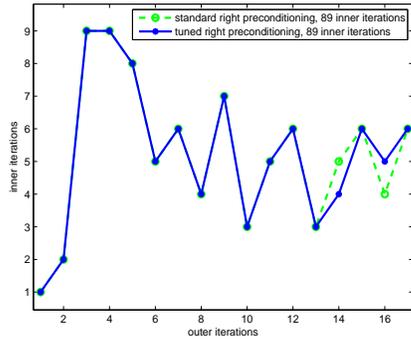


FIG. 5.9. Number of inner iterations against outer iterations with standard and tuned preconditioning (Example 5.4) when using *inexact simplified JD* method

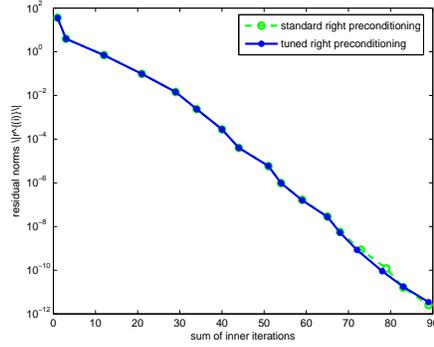


FIG. 5.10. Residual norms vs total number of inner iterations with standard and tuned preconditioning (Example 5.4) when using *inexact simplified JD* method

The results for Example 5.4 are plotted in Figures 5.7 to 5.10. First of all, we can see that for the simplified JD method tuning the preconditioner has no effect, the results for the standard and the tuned preconditioner are very similar (see Figures 5.9 and 5.10), the total number of iterations is the same.

For *inexact RQI* tuning the preconditioner reduces the number of iterations from 264 to 83 (see Figures 5.7 and 5.8) and the total number of iterations is even smaller than the one for *inexact simplified JD* (83 versus 89) iterations. If we had used FOM and tuning with  $\mathbb{P}\mathbf{x} = \mathbf{M}\mathbf{x}$  (instead of GMRES and tuning with  $\mathbb{P}\mathbf{x} = \mathbf{A}\mathbf{x}$ ) those numbers would be equal.

In this example we used the projectors  $\Pi_1$  and  $\Pi_2$  from (5.1) with  $\mathbf{w} = \mathbf{M}\mathbf{x}$  and  $\mathbf{u} = \mathbf{M}^H \mathbf{M}\mathbf{x}$  for the simplified JD method. In general the choice of the projectors for JD for the generalised eigenproblem has to be taken care of, whereas, if we just use the tuned precon-

ditioner in conjunction with inexact RQI we do not have to worry about the choice of the projectors.

**6. Conclusions and further thoughts.** We have developed and analysed a tuned preconditioner for inexact Rayleigh quotient iteration applied to a generalised nonsymmetric eigenproblem. We have applied this tuning strategy to a practical eigenvalue problem arising from the mixed FE discretisation of the linearised steady Navier-Stokes equations and found it to be very successful in reducing the total number of inner iterations. Furthermore we gave an extension of the results in [12] to generalised eigenproblems. Several numerical examples support our theory and show that tuning yields an improvement over the standard preconditioning strategy. These results also help to understand the Jacobi-Davidson method.

Both theory and results suggest that the full preconditioned Jacobi-Davidson method with subspace acceleration is equivalent to subspace accelerated inverse iteration with a tuned preconditioner.

## REFERENCES

- [1] J. BERNS-MÜLLER, I. G. GRAHAM, AND A. SPENCE, *Inexact inverse iteration for symmetric matrices*, Linear Algebra Appl., 416 (2006), pp. 389–413.
- [2] J. BERNS-MÜLLER AND A. SPENCE, *Inexact inverse iteration with variable shift for nonsymmetric generalized eigenvalue problems*, SIAM J. Matrix Anal. Appl., 28 (2006), pp. 1069–1082.
- [3] ———, *Inexact inverse iteration and GMRES*, 2008. Submitted.
- [4] B. BOISVERT, R. POZO, K. REMINGTON, B. MILLER, AND R. LIPMAN, *Matrix market*. available online at <http://math.nist.gov/MatrixMarket/>, 2004.
- [5] S. CAMPBELL, I. IPSEN, C. KELLEY, AND C. MEYER, *GMRES and the minimal polynomial*, BIT, 36 (1996), pp. 664–675.
- [6] J. CULLUM AND A. GREENBAUM, *Relations between Galerkin and norm-minimizing iterative methods for solving linear systems*, SIAM J. Matrix Anal. Appl., 17 (1996), pp. 223–247.
- [7] H. C. ELMAN, *Preconditioning for the steady-state Navier-Stokes equations with low viscosity*, SIAM J. Sci. Comput., 20 (1999), pp. 1299–1316 (electronic).
- [8] H. C. ELMAN AND D. J. SILVESTER, *Fast nonsymmetric iterations and preconditioning for navier-stokes equations*, SIAM J. Sci. Comput., 17 (1996), pp. 33–46.
- [9] M. EMBREE, *How descriptive are GMRES convergence bounds?*, 1999. Numerical Analysis Report 99/08, Oxford University Computing Laboratory.
- [10] M. A. FREITAG AND A. SPENCE, *Convergence rates for inexact inverse iteration with application to preconditioned iterative solves*, BIT, 47 (2007), pp. 27–44.
- [11] ———, *Convergence theory for inexact inverse iteration applied to the generalised nonsymmetric eigenproblem*, Electron. Trans. Numer. Anal., 28 (2007), pp. 40–64.
- [12] ———, *Rayleigh quotient iteration and simplified Jacobi-Davidson method with preconditioned iterative solves*, Linear Algebra Appl., 428 (2008), pp. 2049–2060.
- [13] ———, *Shift-invert Arnoldi’s method with preconditioned iterative solves*, 2008. Submitted to SIAM J. Matrix Anal. Appl.
- [14] ———, *A tuned preconditioner for inexact inverse iteration applied to Hermitian eigenvalue problems*, IMA J. Numer. Anal., 28 (2008), pp. 522–551.
- [15] S. K. GODUNOV, *Modern Aspects of Linear Algebra*, Providence, RI, 1998. Translation of Mathematical Monographs, 175. Amer. Math. Soc.
- [16] G. H. GOLUB AND C. F. VAN LOAN, *Matrix Computations*, John Hopkins University Press, Baltimore, 2nd ed., 1989.
- [17] ———, *Matrix Computations*, John Hopkins University Press, Baltimore, 3rd ed., 1996.
- [18] G. H. GOLUB AND Q. YE, *Inexact inverse iteration for generalized eigenvalue problems*, BIT, 40 (2000), pp. 671–684.
- [19] I. G. GRAHAM, A. SPENCE, AND E. VAINIKKO, *Parallel iterative methods for Navier-Stokes equations and application to eigenvalue computation*, Concurrency and Computation: Practice and Experience, 15 (2003), pp. 1151–1168.
- [20] A. GREENBAUM, *Iterative methods for solving linear systems*, vol. 17 of Frontiers in Applied Mathematics, SIAM, Philadelphia, PA, 1997.
- [21] P. M. GRESHO AND R. L. SANI, *Advection-diffusion and isothermal laminar flow*, in Incompressible Flow and the Finite Element Method, John Wiley & Sons, Ltd., Chichester, 1998.
- [22] M. HOCHBRUCK AND C. LUBICH, *On Krylov subspace approximations to the matrix exponential operator*, SIAM J. Numer. Anal., 34 (1997), pp. 1911–1925.
- [23] M. E. HOCHSTENBACH AND G. L. G. SLEIJPEN, *Two-sided and alternating Jacobi-Davidson*, Linear Algebra Appl., 358 (2003), pp. 145–172. Special Issue on accurate solution of eigenvalue problems (Hagen, 2000).

- [24] T. KATO, *Perturbation Theory for Linear Operators*, Springer Verlag, 1966.
- [25] Y.-L. LAI, K.-Y. LIN, AND W.-W. LIN, *An inexact inverse iteration for large sparse eigenvalue problems*, Numer. Linear Algebra Appl., 1 (1997), pp. 1–13.
- [26] J. LIESEN AND P. TICHÝ, *Convergence analysis of Krylov subspace methods*, GAMM Mitt. Ges. Angew. Math. Mech., 27 (2004), pp. 153–173 (2005).
- [27] Y. NOTAY, *Convergence analysis of inexact Rayleigh quotient iteration*, SIAM J. Matrix Anal. Appl., 24 (2003), pp. 627–644.
- [28] ———, *Inner iterations in eigenvalue solvers*, 2005. Report GANMN 05-01, Université Libre de Bruxelles, Brussels, Belgium, <http://homepages.ulb.ac.be/~ynotay>.
- [29] M. ROBBÉ, M. SADKANE, AND A. SPENCE, *Inexact inverse subspace iteration with preconditioning applied to non-Hermitian eigenvalue problems*, 2007. To appear in SIAM J. Matrix Anal. Appl.
- [30] Y. SAAD, *Numerical Methods for Large Eigenvalue Problems*, Halsted Press, New York, 1992.
- [31] ———, *A flexible inner-outer preconditioned GMRES algorithm*, SIAM J. Sci. Comput., 14 (1993), pp. 461–469.
- [32] ———, *Iterative Methods for Sparse Linear Systems*, SIAM, Philadelphia, PA, 2nd ed., 2003.
- [33] R. SHEPARD, A. F. WAGNER, J. L. TILSON, AND M. MINKOFF, *The subspace projected approximate matrix (SPAM) modification of the Davidson method*, J. Comput. Phys., 172 (2001), pp. 472–514.
- [34] V. SIMONCINI AND L. ELDÉN, *Inexact Rayleigh quotient-type methods for eigenvalue computations*, BIT, 42 (2002), pp. 159–182.
- [35] G. L. G. SLEIJPEN, A. G. L. BOOTEN, D. R. FOKKEMA, AND H. A. VAN DER VORST, *Jacobi-Davidson type methods for generalized eigenproblems and polynomial eigenproblems*, BIT, 36 (1996), pp. 595–633. International Linear Algebra Year (Toulouse, 1995).
- [36] P. SMIT AND M. H. C. PAARDEKOOPER, *The effects of inexact solvers in algorithms for symmetric eigenvalue problems*, Lin. Alg. Appl, 287 (1999), pp. 337–357.
- [37] A. STATHOPOULOS AND Y. SAAD, *Restarting techniques for the (Jacobi-)Davidson symmetric eigenvalue methods*, Electron. Trans. Numer. Anal., 7 (1998), pp. 163–181.
- [38] G. W. STEWART, *Error and perturbation bounds for subspaces associated with certain eigenvalue problems*, SIAM Rev., 15 (1973), pp. 727–764.
- [39] ———, *Matrix Algorithms II: Eigensystems*, SIAM, Philadelphia, PA, 2001.
- [40] G. W. STEWART AND J.-G. SUN, *Matrix Perturbation Theory*, Academic Press, Boston, 1990.
- [41] D. B. SZYLD, *On recurring theorems on the norm of oblique projections*, 2005. Research Report 05-12-30, Department of Mathematics, Temple University, December 2005.
- [42] L. N. TREFETHEN AND M. EMBREE, *Spectra and Pseudospectra: The Behavior of Nonnormal Matrices and Operators*, Princeton University Press, 2005.
- [43] H. A. VAN DER VORST AND C. VUIK, *The superlinear convergence behaviour of GMRES*, J. Comput. Appl. Math., 48 (1993), pp. 327–341.
- [44] F. XUE AND H. C. ELMAN, *Convergence analysis of iterative solvers in inexact Rayleigh quotient iteration*, Preprint UMCP-CSD:CS-TR-4902, University of Maryland, Department of Computer Science and Institute for Advanced Computer Studies, 2008.