A Jacobi–Davidson type method for a right definite two-parameter eigenvalue problem

Bor Plestenjak
Department of mathematics, University of Ljubljana, Slovenia

Joint work with Michiel E. Hochstenbach (Utrecht University)

Two-parameter eigenvalue problem. We consider

$$\begin{aligned} A_1 x &= \lambda B_1 x + \mu C_1 x \\ A_2 y &= \lambda B_2 y + \mu C_2 y, \end{aligned} \tag{1}$$

where A_i, B_i, C_i are $n_i \times n_i$ real matrices for $i = 1, 2$. A pair (λ, μ) is an *eigenvalue* if it satisfies (1) for nonzero x, y. The tensor product $x \otimes y$ is the corresponding *eigenvector*.

Problem: find (approximate) eigenvalue (λ, μ) and (approximate) eigenvector $x \otimes y$.

Right definite problem.

- Matrices A_i, B_i, C_i are symmetric.
- For nonzero x, y we have $\begin{vmatrix} x^T B_1 x & x^T C_1 x \\ y^T B_2 y & y^T C_2 y \end{vmatrix} > 0.$

A right definite problem has real eigenpairs and there exist $n_1 n_2$ linearly independent eigenvectors.

Tensor product approach. On the tensor product space $\mathbb{R}^{n_1} \otimes \mathbb{R}^{n_2}$ we define

$$\begin{aligned} \Delta_0 &= B_1 \otimes C_2 - C_1 \otimes B_2 \\ \Delta_1 &= A_1 \otimes C_2 - C_1 \otimes A_2 \\ \Delta_2 &= B_1 \otimes A_2 - A_1 \otimes B_2. \end{aligned}$$

Problem (1) can be then expressed as two coupled generalized eigenvalue problems (GEP)

$$\begin{aligned} \Delta_1 z &= \lambda \Delta_0 z \\ \Delta_2 z &= \mu \Delta_0 z. \end{aligned} \tag{2}$$

If a problem is right definite then

- Δ_i is symmetric,
- Δ_0 is symmetric positive definite,
- $\Delta_0^{-1}\Delta_1$ and $\Delta_0^{-1}\Delta_2$ commute.

We can solve problem (2) using standard methods for GEP, but the dimension is $N := n_1 n_2$.

Jacobi-Davidson type method. As in each subspace method we start with a given search subspace from which approximations to eigenpairs are computed (*extraction*). In the extraction we solve the same type of eigenvalue problem as the original one, but of a smaller dimension. After each step we expand the subspace by a new direction (*expansion*). As the search subspace grows, the eigenpair approximations should converge to an eigenpair of the original problem.

Extraction. Let the k-dimensional search subspaces $\mathcal{U}_k$ of $\mathbb{R}^{n1}$ and $\mathcal{V}_k$ of $\mathbb{R}^{n2}$ be represented by matrices $U_k \in \mathbb{R}^{n1\times k}$ and $V_k \in \mathbb{R}^{n2\times k}$ with orthonormal columns, respectively. The Ritz–Galerkin conditions

$$\begin{aligned}(A_1 - \sigma B_1 - \tau C_1)u &\perp \mathcal{U}_k\\ (A_2 - \sigma B_2 - \tau C_2)v &\perp \mathcal{V}_k,\end{aligned}$$

lead to the smaller projected right definite two-parameter problem

$$\begin{aligned} U_k^T A_1 U_k c &= \sigma U_k^T B_1 U_k c + \tau U_k^T C_1 U_k c,\\ V_k^T A_2 V_k d &= \sigma V_k^T B_2 V_k d + \tau V_k^T C_2 V_k d,\end{aligned} \tag{3}$$

An eigenvalue (σ, τ) of (3) is a *Ritz value*. If (σ, τ) is an eigenvalue of and $c \otimes d$ is the corresponding eigenvector, then $u \otimes v$ is a *Ritz vector*, where $u = U_k c$ and $v = V_k d$. Altogether we obtain k^2 *Ritz pairs* that are approximations to the eigenpairs of (1).

Residuals are defined as

$$\begin{aligned} r_1 &= (A_1 - \sigma B_1 - \tau C_1)u\\ r_2 &= (A_2 - \sigma B_2 - \tau C_2)v.\end{aligned}$$

Expansion. We are looking for improvements $s \perp u$, $t \perp v$, such that

$$\begin{aligned} A_1(u+s) &= \lambda B_1(u+s) + \mu C_1(u+s)\\ A_2(v+t) &= \lambda B_2(v+t) + \mu C_2(v+t).\end{aligned} \tag{4}$$

Theorem: $((\sigma, \tau), u\otimes v)$ Ritz pair, $u = x - s$, $v = y - t$ $\Rightarrow$ $\sqrt{(\lambda-\sigma)^2 + (\mu-\tau)^2} = \mathcal{O}(\|s\|^2 + \|t\|^2)$

If we rewrite first equation in (4) as

$$(A_1 - \sigma B_1 - \tau C_1)s = -r_1 + (\lambda-\sigma)B_1 u + (\mu-\tau)C_1 u + (\lambda-\sigma)B_1 s + (\mu-\tau)C_1 s$$

then

- $(\lambda-\sigma)B_1 s + (\mu-\tau)C_1 s$ is a "mixed" 3rd order term and we neglect it
- $(\lambda-\sigma)B_1 u + (\mu-\tau)C_1 u$ is a "mixed" 2nd order term and we:
  - **I** neglect it
  - **II** project it away

I. Correction equations with orthogonal projections. When we neglect 2nd and 3rd order terms we obtain

$$(A_1 - \sigma B_1 - \tau C_1)s = -r_1.$$

As the right-hand side is orthogonal to u, so is the left-hand side. We get two separate correction equations

$$\boxed{\begin{aligned} (I - uu^T)(A_1 - \sigma B_1 - \tau C_1)(I - uu^T)s &= -r_1 \\ (I - vv^T)(A_2 - \sigma B_2 - \tau C_2)(I - vv^T)t &= -r_2. \end{aligned}} \tag{5}$$

Orthogonal projections preserve the symmetry of the matrices. The equations (5) for s and t are not of full rank but they are consistent. We solve them only approximately with a Krylov subspace method with initial guess 0, for instance by a few steps of MINRES.

II. Correction equation with oblique projection. When we neglect 3rd order terms we obtain

$$\begin{aligned} (A_1 - \sigma B_1 - \tau C_1)s &= -r_1 + (\lambda - \sigma)B_1u + (\mu - \tau)C_1u \\ (A_2 - \sigma B_2 - \tau C_2)t &= -r_2 + (\lambda - \sigma)B_2v + (\mu - \tau)C_2v. \end{aligned}$$

If we define

$$\begin{aligned} M &= \begin{bmatrix} A_1 - \sigma B_1 - \tau C_1 & 0 \\ 0 & A_2 - \sigma B_2 - \tau C_2 \end{bmatrix} \\ r &= \begin{bmatrix} r_1 \\ r_2 \end{bmatrix} \end{aligned}$$

then we can write

$$M \begin{bmatrix} s \\ t \end{bmatrix} = -r + (\lambda - \sigma)\begin{bmatrix} B_1u \\ B_2v \end{bmatrix} + (\mu - \tau)\begin{bmatrix} C_1u \\ C_2v \end{bmatrix}.$$

We project 2nd order terms away using the oblique 2D-projection $P = I - V(W^TV)^{-1}W^T$, where

$$\begin{aligned} \mathrm{span}(V) &= \mathrm{span}\left(\begin{bmatrix} B_1u \\ B_2v \end{bmatrix}, \begin{bmatrix} C_1u \\ C_2v \end{bmatrix}\right) \\ W &= \begin{bmatrix} u & 0 \\ 0 & v \end{bmatrix}, \end{aligned}$$

and obtain the correction equation

$$\boxed{PMP \begin{bmatrix} s \\ t \end{bmatrix} = -r.} \tag{6}$$

As before we solve (6) only approximately with a Krylov subspace method with initial guess 0, for instance by a few steps of GMRES.

The JD-type method with the correction equation (6) is a Newton scheme, accelerated by the projection of (1) onto the subspace of all previous approximations. Therefore, we can expect locally at least quadratic convergence when the correction equations are solved exactly.

Selection of Ritz pairs.

- If we are interested in exterior eigenvalues, for instance in the one with maximal λ, then in every step we select Ritz pair $((\sigma,\tau), u\otimes v)$ with maximal σ. In this case we have monotonic convergence $\sigma_k \uparrow \lambda$.
  Of course, like in many subspace methods misconvergence is possible and we can obtain $\lambda \neq \lambda_{\max}$.
- If we are interested in interior eigenvalues, for instance in the one closest to $(0,0)$, then in every step we select Ritz pair $((\sigma,\tau), u\otimes v)$ with minimum $\sigma^2+\tau^2$. The convergence is erratic, but numerical results show that the method can be used for interior eigenvalues as well.
  The question remains if it is possible to generalize harmonic Ritz values to a right definite two-parameter eigenvalue problem. Any progress on this subject might lead to better methods for interior eigenvalues.

Computing more eigenpairs. Standard deflation techniques can not be applied for two reasons:

- $(x\otimes y)^{\perp_{\Delta_0}}$ can not be written as $\mathcal{U}\otimes\mathcal{V}$, where $\mathcal{U}\subset\mathbb{R}^{n1}$ and $\mathcal{V}\subset\mathbb{R}^{n2}$.
- There can exist eigenvalues (λ,μ) and (λ',μ') with eigenvectors $x\otimes y$ and $x'\otimes y'$, respectively, such that $(\lambda,\mu)\neq(\lambda',\mu')$ and $x=x'$.

Our approach: As eigenvectors with different eigevalues are Δ_0-orthogonal, i.e: $(x_1\otimes y_1)^T\Delta_0(x_2\otimes y_2)=0$, we consider in selection only Ritz vectors that are Δ_0-orthogonal to already computed eigenvectors.

Summary of the algorithm.

1. $s=u_1$ and $t=v_1$ (starting vectors)
   for $k=1,2,\ldots$ (outer loop)
2. MGS$(U_{k-1},s)\to U_k$ (expansion)
   MGS$(V_{k-1},t)\to V_k$
3. Compute rightmost Ritz pair $((\sigma,\tau), c\otimes d)$ of (extraction)
$$\begin{aligned} U_k^TA_1U_kc &= \sigma U_k^TB_1U_kc+\tau U_k^TC_1U_kc\\ V_k^TA_2V_kd &= \sigma V_k^TB_2V_kd+\tau V_k^TC_2V_kd\end{aligned}$$
4. Compute residuals
$$\begin{aligned} r_1 &= (A_1-\sigma B_1-\tau C_1)u\\ r_2 &= (A_2-\sigma B_2-\tau C_2)v\end{aligned}$$
5. Stop if $(\|r_1\|^2+\|r_2\|^2)^{1/2}\le\varepsilon$ (stopping criteria)
6. Solve approximately an $s\perp u, t\perp v$ from correction equation(s)

JD-like algorithm for rightmost eigenvalue of a right definite two-parameter eigenvalue problem

Restarts. As the existing methods are able to solve only low-dimensional two-parameter problems in a reasonable time, we expand search spaces up to the preselected dimension $l_{\max}$ and then restart the algorithm using the most promising $l_{\min}$ eigenvector approximations as a basis for the initial search space.

Time complexity. Suppose that $n=n_1=n_2$ and let m be the number of GMRES steps. The time complexity of one outer step is:

- dense matrices: $\mathcal{O}(mn^2)$
- sparse matrices: $\mathcal{O}(mMV)$, where MV denotes matrix-vector multiplication with $n\times n$ matrix.

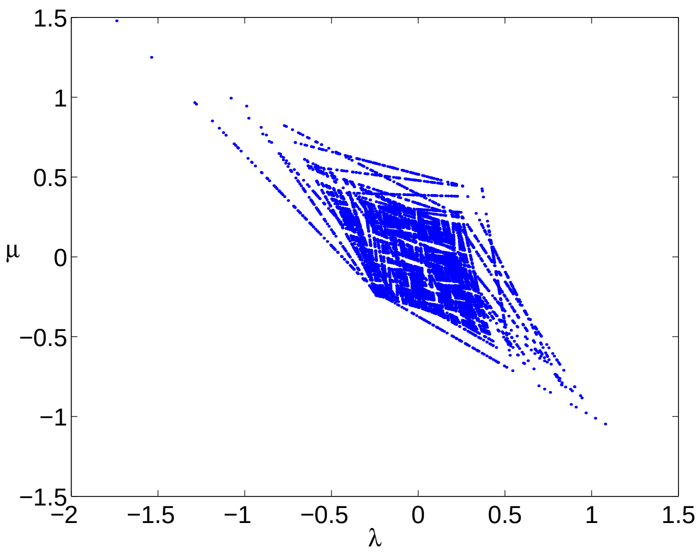
Numerical experiments. For the test examples we take

$$A_i = Q_i F_i Q_i^T$$
$$B_i = Q_i G_i Q_i^T$$
$$C_i = Q_i H_i Q_i^T,$$

where

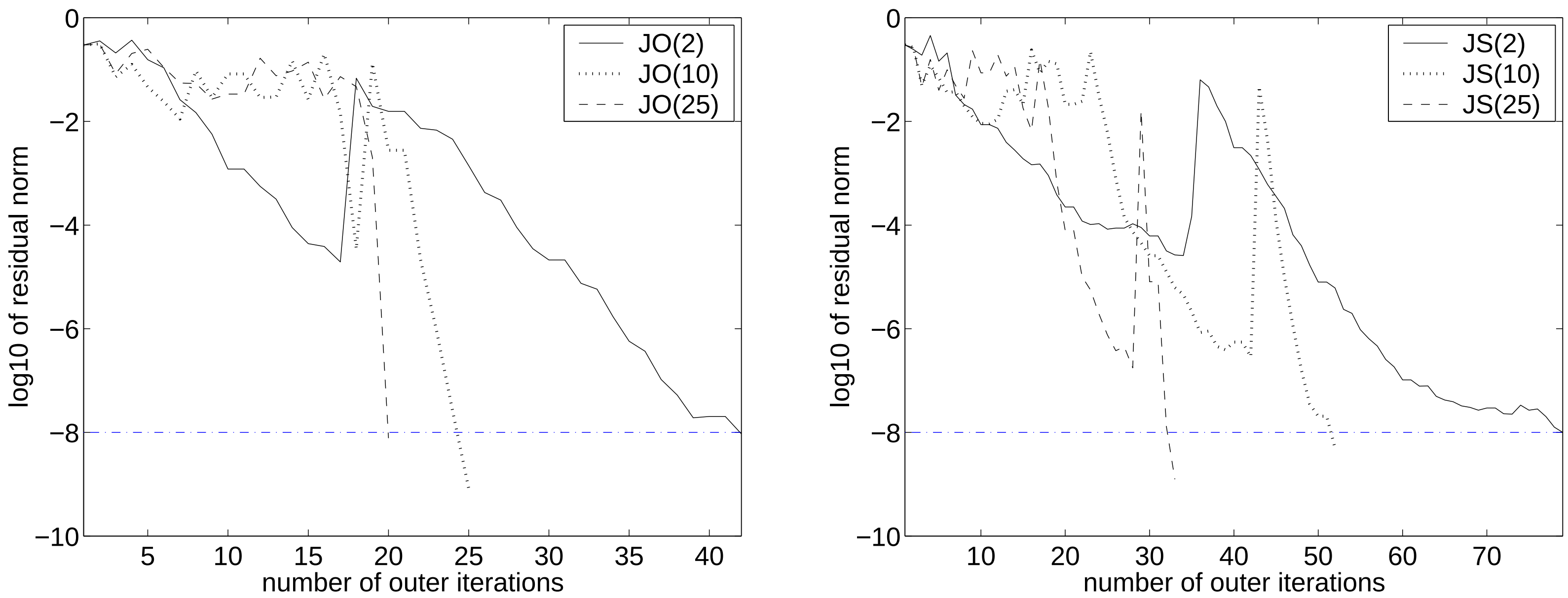
- F_1, F_2, G_2, H_1 are random diagonal matrices $\in (0,1)$,
- G_1, H_2 are random diagonal matrices $\in (1,2)$,
- Q_i is a random orthogonal matrix.

We shift the obtained right definite problem so that the arithmetic mean of the eigenvalues is $(0,0)$.



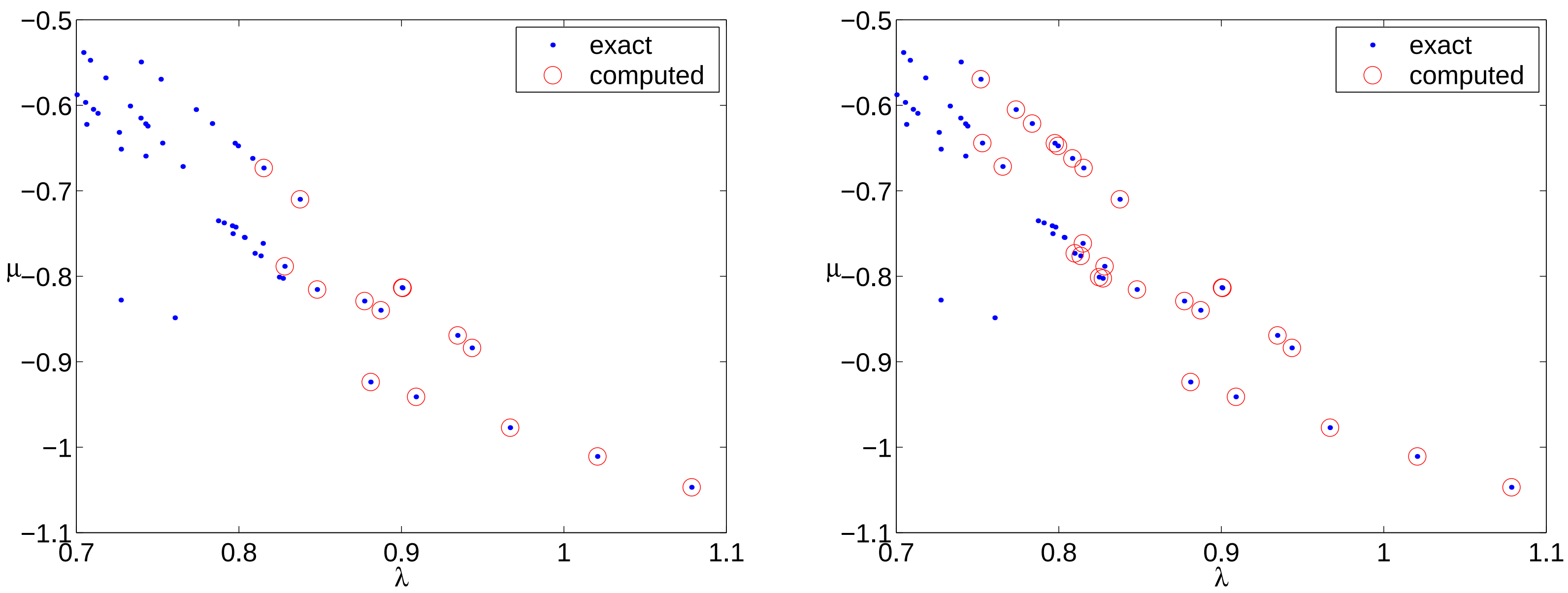
Distribution of eigenvalues for $n = 100$

Convergence history.



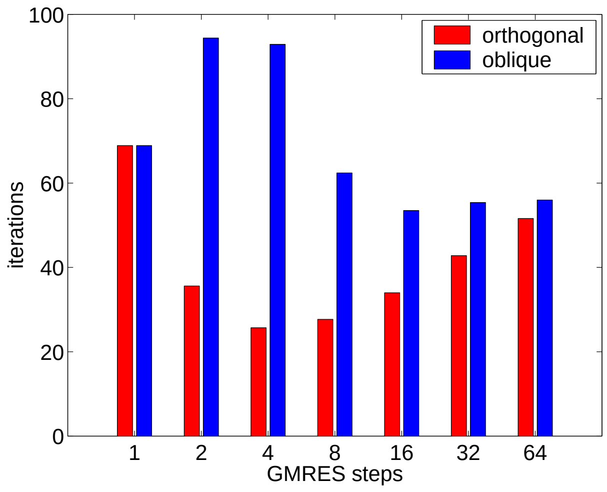
Convergence plot for eigenvalue with the maximal λ for $n = 100$ and $u = v = [1 \ \cdots \ 1]^T$. The plots show the $\log_{10}$ of the residual norm $\rho := (\|r_1\|^2 + \|r_2\|)^{1/2}$ versus the outer iteration number k using 2 (solid line), 10 (dotted line), and 25 (dashed line) GMRES steps to solve the correction equation with orthogonal projections (left plot) and oblique projections (right plot), respectively.

Computing more eigenpairs.

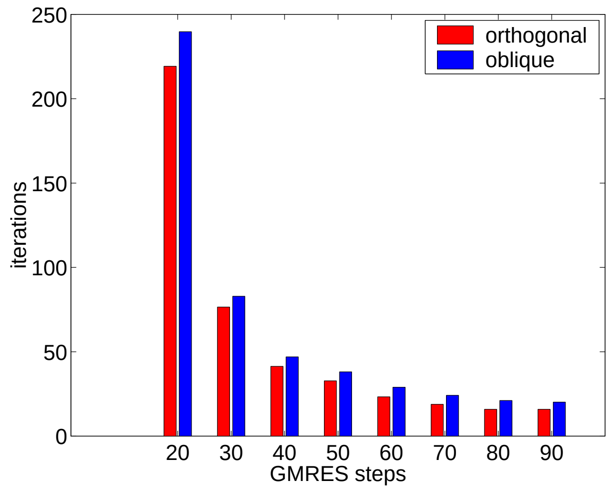


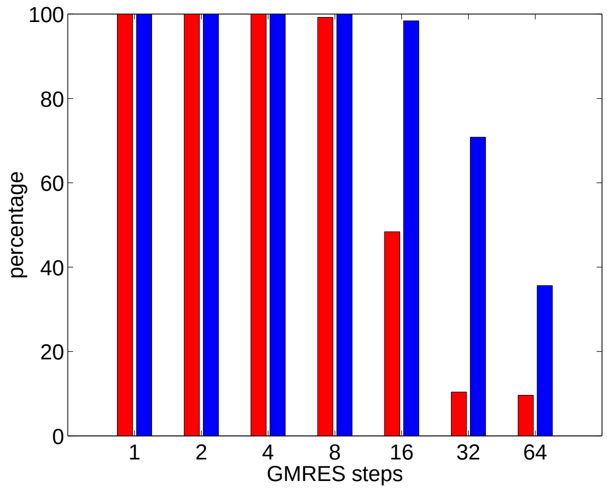
First 15 (left plot) and first 30 (right plot) computed eigenvalues with maximal λ for $n = 100$ computed using selection for Ritz vectors. The JD-type method used 5 GMRES steps for the correction equation with orthogonal projections.

Statistics. We tested the method using different correction equations and different inner iteration processes. The folowing plots present the average number of iterations, percentage of convergence to the chosen eigenvalue, and average number of flops over 250 trials with different random initial vectors.



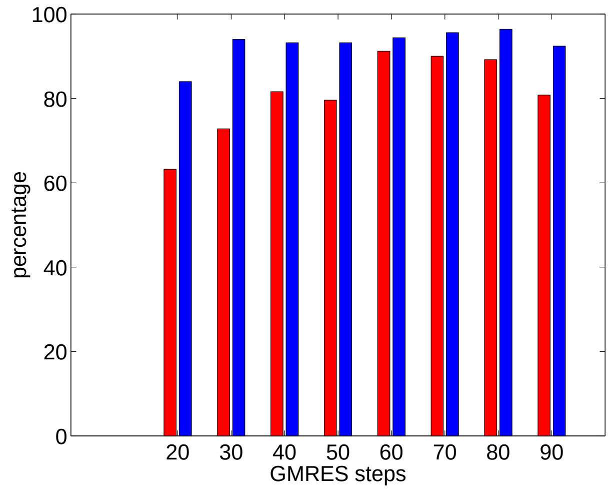
The average number of iterations (the number of outer steps)

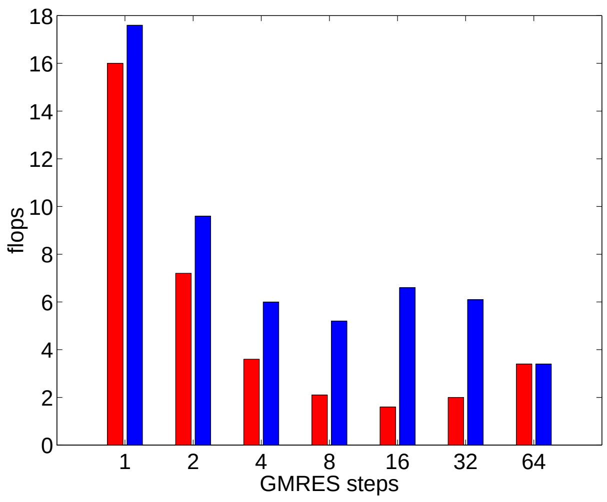




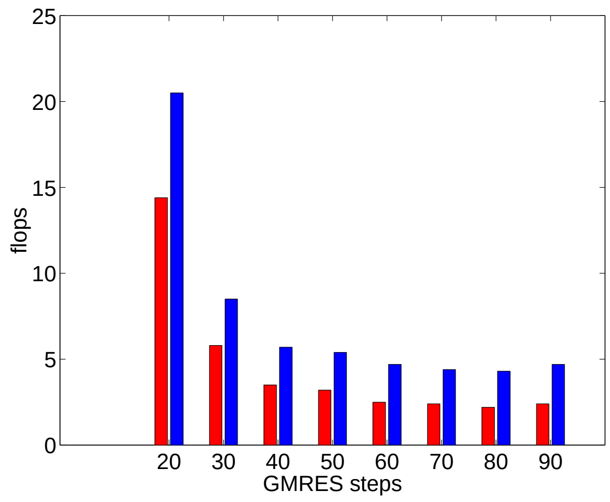
Percentage of convergence to the goal, i.e.

- left: eigenvalue with the maximal λ,
- right: eigenvalue closest to $(0,0)$.





Average number of flops ($\times 10^8$)



Conclusion. New JD-like method for two-parameter eigenvalue problem

- first method for large matrices
- choice of orthogonal or oblique (Newton-like) correction equation: 2 orthogonal 1-D or 1 oblique 2-D projector
- orthogonal variant is less expensive for exterior eigenvalues
- oblique variant is more expensive but more reliable for interior eigenvalues
- harmonic approach is not obvious
- straightforward generalization to $>$ 2-parameters
- work in progress: Ritz theory, JD for general two-parameter eigenvalue problem

Preprint is available at **www-lp.fmf.uni-lj.si/plestenjak/bor.htm**.