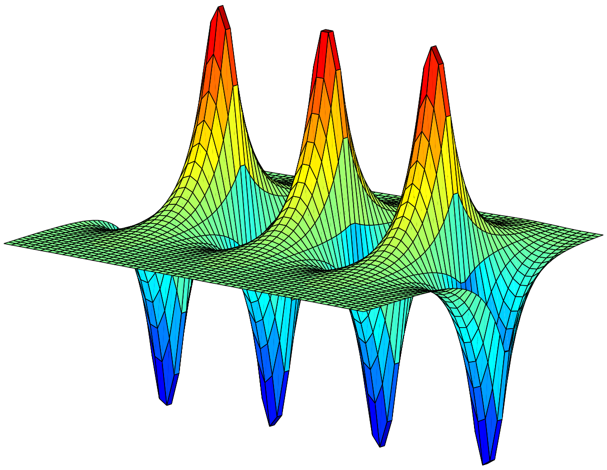
Universiteit Utrecht

Mathematical Institute



Using domain decomposition in the Jacobi-Davidson method

by

Menno Genseberger, Gerard Sleijpen, and Henk Van der Vorst

Preprint

No. 1164

October, 2000

Also available at URL `http://www.math.uu.nl/publications/`

Using domain decomposition in the Jacobi-Davidson method

Menno Genseberger[§], Gerard Sleijpen[*], and Henk Van der Vorst[*]

October, 2000

Abstract

The Jacobi-Davidson method is suitable for computing solutions of large n-dimensional eigenvalue problems. It needs (approximate) solutions of specific n-dimensional linear systems. Here we propose a strategy based on a nonoverlapping domain decomposition technique in order to reduce the wall clock time and local memory requirements. For a model eigenvalue problem we derive optimal coupling parameters. Numerical experiments show the effect of this approach on the overall Jacobi-Davidson process. The implementation of the eventual process on a parallel computer is beyond the scope of this paper.

Keywords: Eigenvalue problems, domain decomposition, Jacobi-Davidson, Schwarz method, nonoverlapping, iterative methods.

2000 Mathematics Subject Classification: 65F15, 65N25, 65N55.

1 Introduction

The Jacobi-Davidson method [17] is a valuable approach for the solution of large (generalized) linear eigenvalue problems. The method reduces the large problem to a small one by projecting it on an appropriate low dimensional subspace. Approximate solutions for eigenpairs of the large problem are obtained from the small problem by means of a Rayleigh-Ritz principle. The heart of the Jacobi-Davidson method is how the subspace is expanded. To keep the dimension of the subspace, and consequently the size of the small problem, low it is essential that all necessary information of the wanted eigenpair(s) is collected in the subspace after a small number of iterations. Therefore, the subspace should be expanded with a vector that contains important information not already present in the subspace. The correction equation of the Jacobi-Davidson method aims to prescribe such a vector.

But in itself, the correction equation poses a large linear problem, with size equal to the size of the originating large eigenvalue problem. Because of this, most of the computational work of the Jacobi-Davidson method arises from solving the correction equation. In practice the eigenvalue problem is often so large that an accurate solution of the correction equation is too expensive. However, often approximate solutions of the correction equation suffice to obtain sufficiently fast convergence of the Jacobi-Davidson

[§]Mathematical Institute, Utrecht University, and CWI, Amsterdam, The Netherlands.
[*]Mathematical Institute, Utrecht University, P.O. Box 80.010, 3508 TA Utrecht, The Netherlands.
E-mail: genseber@math.uu.nl, sleijpen@math.uu.nl, vorst@math.uu.nl

method. The speed of this convergence depends on the accuracy of the approximate solution. Jacobi-Davidson lends itself to be used in combination with a preconditioned iterative solver for the correction equation. In such a case the quality of the preconditioner is critical.

Nonoverlapping domain decomposition methods for *linear systems* have been studied well in literature. Because of the absence of overlapping regions they have computational advantages compared to domain decomposition methods with overlap. But much depends on the coupling that should be chosen carefully.

In this paper we will show how a nonoverlapping domain decomposition technique can be incorporated in the correction equation of Jacobi-Davidson, when applied to PDE type of eigenvalue problems. The technique is based on work by Tang and by Tan and Borsboom for linear systems.

For a linear system Tang [20] proposed to enhance the system with duplicates in order to enable an additive Schwarz method with minimal overlap (for more recent publications, see for example [7], [12] and [10]). Tan and Borsboom [19, 18] refined this idea by introducing more flexibility for the unknowns near the interfaces between the subdomains. In this way additional degrees of freedom are created, reflected by coupling equations for the unknowns near the interfaces and their virtual counterparts. Now, the key point is to tune these interface conditions for the given problem in order to improve the speed of convergence of the iterative solution method. This approach is very effective for classes of linear systems stemming from advection-diffusion problems [19, 18].

The operator in the correction equation involves the matrix of the large eigenvalue problem shifted by an approximate eigenvalue. In the computational process, this shift will become arbitrarily close to the desired eigenvalue. This is a situation that requires special attention when applying the domain decomposition technique.

An eigenvalue problem imposes a mildly nonlinear problem. Therefore, for the computation of solutions to the eigenvalue problem one needs a nonlinear solver, for instance, a Newton method. In fact, Jacobi-Davidson can be seen as an accelerated inexact Newton method [16]. Here, we shall, as explained above, combine the Jacobi-Davidson method with a Krylov solver for the correction equation. A preconditioner for the Krylov solver is constructed with domain decomposition. A similar type of nesting, but for general nonlinear systems, can be found in the Newton-Krylov-Schwarz algorithms by Cai, Gropp, Keyes et al. in [4] and [5]. In these two papers the subdomains have overlap, therefore there is no analysis for the tuning of the coupling between subdomains. Furthermore, the eigenvalue problem is nonlinear but with a specific structure; we will exploit this structure.

Our paper is organized as follows. First, we recall the enhancement technique for domain decomposition in §2. Then, in §3 we discuss the Jacobi-Davidson method. We outline how the technique can be applied to the correction equation and how the projections in the correction equation should be handled. For a model eigenvalue problem we investigate, in §4, in detail how the coupling equations should be chosen for optimal performance. It will turn out that the shift plays a critical role. Section §5 gives a number of illustrative numerical examples.

2 Domain decomposition

2.1 Canonical enhancement of a linear system

Tang [20] has proposed the concept of matrix enhancement, which gives elegant possibilities for the formulation of effective domain decomposition of the underlying PDE problem. The idea is to decompose the grid into nonoverlapping subgrids and to expand the subgrids by introducing additional gridpoints and additional unknows along the interfaces of the decomposition. This approach artificially creates some overlap on gridpoint level and the overlap is minimal. For hyperbolic systems of PDEs, this approach was further refined by Tan in [18] and by Tan and Borsboom in [19]. Discretization of the PDE leads to a linear system of equations. Tang duplicates and adjusts those equations in the system that couple across the interfaces. Tan and Borsboom introduce a double set of additional gridpoints along the interfaces in order to keep each equation confined to one expanded subgrid. As a consequence, none of the equations has to be adjusted. Then they enhanced the linear system by 'new' equations that can be viewed as discretized boundary conditions for the internal boundaries (along the interfaces). Since the last approach offers more flexibility, this is the one we follow.

We start with the linear nonsingular system

$$\mathbf{B}\,\mathbf{y} = \mathbf{d}, \tag{1}$$

that results from discretization of a given PDE over some domain. Now, we partition the matrix $\mathbf{B}$, and the vectors $\mathbf{y}$ and $\mathbf{d}$ correspondingly,

$$\begin{bmatrix} \mathbf{B}_{11} & \mathbf{B}_{1\ell} & \mathbf{B}_{1r} & \mathbf{B}_{12} \\ \mathbf{B}_{\ell 1} & B_{\ell\ell} & B_{\ell r} & \mathbf{B}_{\ell 2} \\ \mathbf{B}_{r1} & B_{r\ell} & B_{rr} & \mathbf{B}_{r2} \\ \mathbf{B}_{21} & \mathbf{B}_{2\ell} & \mathbf{B}_{2r} & \mathbf{B}_{22} \end{bmatrix}, \quad \begin{bmatrix} \mathbf{y}_1 \\ y_\ell \\ y_r \\ \mathbf{y}_2 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} \mathbf{d}_1 \\ d_\ell \\ d_r \\ \mathbf{d}_2 \end{bmatrix}.$$

The labels are not chosen arbitrarily: we associate with label 1 (and 2, respectively) elements/operations of the linear system corresponding to subdomain 1 (2, respectively) and with label ℓ (resp. r) elements/operations corresponding to the left (resp. right) of the interface between the two subdomains. The central blocks $B_{\ell\ell}$, $B_{\ell r}$, $B_{r\ell}$ and B_{rr} are square matrices of equal size, say, n_i by n_i. They correspond to the unknowns along the interface. Since the number of unknowns along the interface will typically be much smaller than the total number of unknows, n_i will be much smaller than n, the size of $\mathbf{B}$.

For a typical discretization, the matrix $\mathbf{B}$ is banded and the unknowns are only locally coupled. Therefore it is not unreasonable to assume that $\mathbf{B}_{r1}, \mathbf{B}_{21}, \mathbf{B}_{12}$ and $\mathbf{B}_{\ell 2}$ are zero. For this situation, we define the 'canonical enhancement' $\mathbf{B}_I$ of $\mathbf{B}$, $\underset{\sim}{\mathbf{y}}$ of $\mathbf{y}$, and $\underline{\mathbf{d}}$ of $\mathbf{d}$, by

$$\mathbf{B}_I \equiv \left[\begin{array}{ccc|ccc} \mathbf{B}_{11} & \mathbf{B}_{1\ell} & \mathbf{B}_{1r} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{B}_{\ell 1} & B_{\ell\ell} & B_{\ell r} & 0 & 0 & \mathbf{0} \\ \mathbf{0} & I & 0 & -I & 0 & \mathbf{0} \\ \hline \mathbf{0} & 0 & -I & 0 & I & \mathbf{0} \\ \mathbf{0} & 0 & 0 & B_{r\ell} & B_{rr} & \mathbf{B}_{r2} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{B}_{2\ell} & \mathbf{B}_{2r} & \mathbf{B}_{22} \end{array}\right], \quad \underset{\sim}{\mathbf{y}} \equiv \begin{bmatrix} \mathbf{y}_1 \\ y_\ell \\ \tilde{y}_r \\ \tilde{y}_\ell \\ y_r \\ \mathbf{y}_2 \end{bmatrix}, \text{ and } \quad \underline{\mathbf{d}} \equiv \begin{bmatrix} \mathbf{d}_1 \\ d_\ell \\ 0 \\ 0 \\ d_r \\ \mathbf{d}_2 \end{bmatrix}. \tag{2}$$

One easily verifies that $\mathbf{B}_I$ is also nonsingular and that $\underline{\mathbf{y}}$ is the unique solution of

$$\mathbf{B}_I\,\underset{\sim}{\mathbf{y}} = \underline{\mathbf{d}}, \tag{3}$$

with $\underline{\mathbf{y}} \equiv (\mathbf{y}_1^T, y_\ell^T, y_r^T, y_\ell^T, y_r^T, \mathbf{y}_2^T)^T$.

With this linear system we can associate a simple iterative scheme for the two coupled subblocks:

$$\begin{bmatrix} \mathbf{B}_{11} & \mathbf{B}_{1\ell} & \mathbf{B}_{1r} \\ \mathbf{B}_{\ell 1} & B_{\ell\ell} & B_{\ell r} \\ \mathbf{0} & I & 0 \end{bmatrix} \begin{bmatrix} \mathbf{y}_1^{(i+1)} \\ y_\ell^{(i+1)} \\ \tilde{y}_r^{(i+1)} \end{bmatrix} = \begin{bmatrix} \mathbf{d}_1 \\ d_\ell \\ \tilde{y}_\ell^{(i)} \end{bmatrix},$$

$$\begin{bmatrix} 0 & I & \mathbf{0} \\ B_{r\ell} & B_{rr} & \mathbf{B}_{r2} \\ \mathbf{B}_{2\ell} & \mathbf{B}_{2r} & \mathbf{B}_{22} \end{bmatrix} \begin{bmatrix} \widetilde{y}_\ell^{(i+1)} \\ y_r^{(i+1)} \\ \mathbf{y}_2^{(i+1)} \end{bmatrix} = \begin{bmatrix} \widetilde{y}_r^{(i)} \\ d_r \\ \mathbf{d}_2 \end{bmatrix}. \tag{4}$$

These systems can be solved in parallel and we can view this as a simple additive Schwarz iteration (with no overlap and Dirichlet-Dirichlet coupling). The extra unknowns $\widetilde{y}_\ell$ and $\widetilde{y}_r$, in the enhanced vector $\underset{\sim}{\mathbf{y}}$, will serve for communication between the subdomains during the iterative solution process of the linear system. After termination of the iterative process, we have to undo the enhancement. We could simply skip the values of the additional elements, but since these carry also information one of the alternatives could be the following one.

With an approximate solution

$$\underset{\sim}{\mathbf{y}}^{(i)} = (\mathbf{y}_1^{(i)\,T}, y_\ell^{(i)\,T}, \widetilde{y}_r^{(i)\,T}, \widetilde{y}_\ell^{(i)\,T}, y_r^{(i)\,T}, \mathbf{y}_2^{(i)\,T})^T$$

of (3), we may associate the approximate solution $\mathbf{R}\underset{\sim}{\mathbf{y}}$ of (1) given by

$$\mathbf{R}\underset{\sim}{\mathbf{y}} \equiv (\mathbf{y}_1^{(i)\,T}, \tfrac{1}{2}(y_\ell^{(i)} + \widetilde{y}_\ell^{(i)})^T, \tfrac{1}{2}(y_r^{(i)} + \widetilde{y}_r^{(i)})^T, \mathbf{y}_2^{(i)\,T})^T,$$

that is, we simply average the two sets of unknowns that should have been equal to each other at full convergence.

2.2 Interface coupling matrix

From (2) we see that the interface unknowns and the additional interface unknowns are coupled in a straightforward way by

$$\begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix} \begin{bmatrix} y_\ell \\ \widetilde{y}_r \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix} \begin{bmatrix} \widetilde{y}_\ell \\ y_r \end{bmatrix}, \tag{5}$$

but, of course, we may replace the coupling matrix by any other nonsingular interface coupling matrix C:

$$C \equiv \begin{bmatrix} C_{\ell\ell} & C_{\ell r} \\ -C_{r\ell} & -C_{rr} \end{bmatrix}. \tag{6}$$

This leads to the following block system

$$\mathbf{B}_C\underset{\sim}{\mathbf{y}} = \left[\begin{array}{ccc|ccc} \mathbf{B}_{11} & \mathbf{B}_{1\ell} & \mathbf{B}_{1r} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{B}_{\ell 1} & B_{\ell\ell} & B_{\ell r} & 0 & 0 & \mathbf{0} \\ \mathbf{0} & C_{\ell\ell} & C_{\ell r} & -C_{\ell\ell} & -C_{\ell r} & \mathbf{0} \\ \hline \mathbf{0} & -C_{r\ell} & -C_{rr} & C_{r\ell} & C_{rr} & \mathbf{0} \\ \mathbf{0} & 0 & 0 & B_{r\ell} & B_{rr} & \mathbf{B}_{r2} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{B}_{2\ell} & \mathbf{B}_{2r} & \mathbf{B}_{22} \end{array}\right] \begin{bmatrix} \mathbf{y}_1 \\ y_\ell \\ \widetilde{y}_r \\ \widetilde{y}_\ell \\ y_r \\ \mathbf{y}_2 \end{bmatrix} = \underline{\mathbf{d}}. \tag{7}$$

In a domain decomposition context, we will have for the approximate solution $\underset{\sim}{\mathbf{y}}$ that $\widetilde{y}_r \approx y_r$ and $\widetilde{y}_\ell \approx y_\ell$. If we know some analytic properties about the local behavior of the true solution $\mathbf{y}$ across the interface, for instance, smoothness up to some degree, then we may try to identify a convenient coupling matrix C that takes advantage of this knowledge. We want preferably a C so that

$$\begin{aligned} -C_{\ell\ell}\widetilde{y}_\ell - C_{\ell r}y_r &\approx -C_{\ell\ell}y_\ell - C_{\ell r}y_r \approx 0 \\ \text{and} \quad -C_{r\ell}y_\ell - C_{rr}\widetilde{y}_r &\approx -C_{r\ell}y_\ell - C_{rr}y_r \approx 0. \end{aligned}$$

In that case (7) is almost decoupled into two independent smaller linear systems (identified by the two boxes). We may expect fast convergence for the corresponding additive Schwarz iteration.

2.3 Solution of the coupled subproblems

The goal of the enhancement of the matrix of a given linear system, together with a convenient coupling matrix C, is to get two smaller mildly coupled subsystems that can be solved in parallel.

Additive Schwarz for the linear system (7) leads to the following iterative scheme

$$\begin{bmatrix} \mathbf{B}_{11} & \mathbf{B}_{1\ell} & \mathbf{B}_{1r} \\ \mathbf{B}_{\ell 1} & B_{\ell\ell} & B_{\ell r} \\ \mathbf{0} & C_{\ell\ell} & C_{\ell r} \end{bmatrix} \begin{bmatrix} \mathbf{y}_1^{(i+1)} \\ y_\ell^{(i+1)} \\ \widetilde{y}_r^{(i+1)} \end{bmatrix} = \begin{bmatrix} \mathbf{d}_1 \\ d_r \\ g_r^{(i)} \end{bmatrix},$$
$$\begin{bmatrix} C_{r\ell} & C_{rr} & \mathbf{0} \\ B_{r\ell} & B_{rr} & \mathbf{B}_{r2} \\ \mathbf{B}_{2\ell} & \mathbf{B}_{2r} & \mathbf{B}_{22} \end{bmatrix} \begin{bmatrix} \widetilde{y}_\ell^{(i+1)} \\ y_r^{(i+1)} \\ \mathbf{y}_2^{(i+1)} \end{bmatrix} = \begin{bmatrix} g_\ell^{(i)} \\ d_\ell \\ \mathbf{d}_2 \end{bmatrix}, \tag{8}$$

and
$$g_r^{(i)} = C_{\ell\ell}\,\widetilde{y}_\ell^{(i)} + C_{\ell r}\,y_r^{(i)}, \quad g_\ell^{(i)} = C_{r\ell}\,y_\ell^{(i)} + C_{rr}\widetilde{y}_r^{(i)}. \tag{9}$$

The additive Schwarz method can be represented as a block Jacobi iteration method. To see this, consider the matrix splitting $\mathbf{B}_C = \mathrm{M}_C - \mathrm{N}$, where

$$\mathrm{M}_C \equiv \begin{bmatrix} \mathrm{M}_1 & \mathbf{0} \\ \mathbf{0} & \mathrm{M}_2 \end{bmatrix},$$

with M_1 the matrix at the top in (8) and M_2 the matrix at the bottom. We assume that C *is such that* M_C *is nonsingular*. The approximate solution $\underset{\sim}{\mathbf{y}}^{(i+1)}$ of (7) at step $i+1$ of the block Jacobi method,

$$\underset{\sim}{\mathbf{y}}^{(i+1)} = \underset{\sim}{\mathbf{y}}^{(i)} + \mathrm{M}_C^{-1}\,\underset{\sim}{\mathbf{r}}^{(i)} \quad \text{with} \quad \underset{\sim}{\mathbf{r}}^{(i)} \equiv \underline{\mathbf{d}} - \mathbf{B}_C\underset{\sim}{\mathbf{y}}^{(i)}, \tag{10}$$

corresponds to the approximate solutions at step $i+1$ of the additive Schwarz method. In view of the fact that one wants to have $g_r^{(i)}$ and $g_\ell^{(i)}$ as small as possible in norm, the starting value $\underset{\sim}{\mathbf{y}}^{(0)} \equiv \mathbf{0}$ is convenient, but it is conceivable to construct other starting values for which the two vectors are small in norm (for instance, after a restart of some acceleration scheme).

Jacobi is a one step method and the updates from previous steps are discarded. The updates can also be stored in a space $\mathcal{V}_m$ and be used to obtain more accurate approximations. This leads to a subspace method that, at step m, searches for the approximate solution in the space $\mathcal{V}_m$, which is precisely equal to the Krylov subspace $\mathcal{K}_m(\mathrm{M}_C^{-1}\mathbf{B}_C, \mathrm{M}_C^{-1}\underline{\mathbf{d}})$. For instance, GMRES [14] finds the approximation in $\mathcal{V}_m$ with the smallest residual, and may be useful if only a few iterations are to be expected.

Krylov subspace methods can be interpreted as accelerators of the domain decomposition method (10). The resulting method can also be seen as a preconditioned Krylov subspace method where, in this case, the preconditioner is based on domain decomposition: the matrix M_C. This preconditioning approach where a system of the form $\mathrm{M}_C^{-1}\mathbf{B}_C\underset{\sim}{\mathbf{x}} = \underset{\sim}{\mathbf{r}}^{(0)}$ is solved, is referred to as left preconditioning. Here $\underset{\sim}{\mathbf{r}}^{(0)} \equiv \mathrm{M}_C^{-1}(\underline{\mathbf{d}} - \mathbf{B}_C\underset{\sim}{\mathbf{y}}^{(0)})$ and $\underline{\mathbf{y}} = \underset{\sim}{\mathbf{y}}^{(0)} + \underset{\sim}{\mathbf{x}}$,

Since $\mathrm{M}_C^{-1}\mathbf{B}_C = \mathbf{I} - \mathrm{M}_C^{-1}\mathrm{N}$, the search subspace $\mathcal{V}_m$ coincides with the Krylov subspace $\mathcal{K}_m(\mathrm{M}_C^{-1}\mathrm{N}, \mathrm{M}_C^{-1}\underline{\mathbf{d}})$. The rank of both N and $\mathrm{M}_C^{-1}\mathrm{N}$ is equal to the dimension of C which, in this case where C is nonsingular, is $2n_i$. This shows that the dimension of $\mathcal{V}_m$ is at most $2n_i$. Therefore, the exact solution $\underline{\mathbf{y}}$ of (7) belongs to $\mathcal{V}_m$ for $m \geq 2n_i$ and GMRES finds $\underline{\mathbf{y}}$ in at most $2n_i$ steps. (For further discussion see, for instance, [3, §3.2], [22, §2], and [2].)

2.4 Right preconditioning

We can also use M_C as a right preconditioner. In that case solution $\underline{\mathbf{y}}$ of (7) is obtained as $\underline{\mathbf{y}} = \underset{\sim}{\mathbf{y}}^{(0)} + \mathrm{M}_C^{-1}\underset{\sim}{\mathbf{x}}$ where $\underset{\sim}{\mathbf{x}}$ is solved from

$$\mathbf{B}_C \mathrm{M}_C^{-1} \underset{\sim}{\mathbf{x}} = \underset{\sim}{\mathbf{r}}^{(0)} \quad \text{with} \quad \underset{\sim}{\mathbf{r}}^{(0)} \equiv \underline{\mathbf{d}} - \mathbf{B}_C \underset{\sim}{\mathbf{y}}^{(0)}. \tag{11}$$

Right preconditioning has some advantages for domain decomposition. To see this, first note that any vector of the form $\mathrm{N}\underset{\sim}{\mathbf{v}}$ '*vanishes outside the artificial boundary*', that is, only the $\tilde{\cdot}_r$ and $\tilde{\cdot}_\ell$ component of this vector are nonzero. Since $\mathbf{B}_C \mathrm{M}_C^{-1} = \mathbf{I} - \mathrm{N}\mathrm{M}_C^{-1}$, multiplication by this operator preserves the property of vanishing outside the artificial boundary. Moreover, if $\underset{\sim}{\mathbf{y}}^{(0)} \equiv \mathrm{M}_C^{-1}\underline{\mathbf{d}}$, then $\underset{\sim}{\mathbf{r}}^{(0)} = \underline{\mathbf{d}} - \mathbf{B}_C \underset{\sim}{\mathbf{y}}^{(0)} = \mathrm{N}\mathrm{M}_C^{-1}\underline{\mathbf{d}}$ vanishes outside the artificial boundary.

Therefore, if, for $\underset{\sim}{\mathbf{y}}^{(0)} \equiv \mathrm{M}_C^{-1}\underline{\mathbf{d}}$, equation (11) is solved with a Krylov subspace method with an initial guess that vanishes outside the artificial boundary, for instance $\underset{\sim}{\mathbf{x}}^{(0)} = \mathbf{0}$, then all the intermediate vectors also vanish outside the artificial boundary. Consequently, only vectors of size $2n_i$ have to be stored and the vector updates and dot products are $2n_i$ dimensional operations.

For appropriate $\underset{\sim}{\mathbf{y}}^{(0)}$, the left preconditioned equation can also be formulated in a $2n_i$ dimensional subspace. However, with respect to the standard basis, it is not so easy to identify the corresponding subspace. We will use the $2n_i$ dimensional subspace, characterized by right preconditioning as corresponding to the artificial boundary, for the derivation of properties of the eigensystem of the iteration matrix.

2.5 Convergence analysis

As a consequence of (10), the errors $\mathbf{e}^{(i)} \equiv \underline{\mathbf{y}} - \underset{\sim}{\mathbf{y}}^{(i)}$ in the block Jacobi method satisfy:

$$\underset{\sim}{\mathbf{e}}^{(i+1)} = (\mathbf{I} - \mathrm{M}_C^{-1}\mathbf{B}_C)\underset{\sim}{\mathbf{e}}^{(i)} = \mathrm{M}_C^{-1}\mathrm{N}\underset{\sim}{\mathbf{e}}^{(i)}. \tag{12}$$

Therefore, the convergence rate of Jacobi depends on the spectral properties of the '*error propagation matrix*' $\mathrm{M}_C^{-1}\mathrm{N}$. These properties also determine the convergence behavior of other Krylov subspace methods. With right preconditioning, we have to work with $\underset{\sim}{\mathbf{x}} - \underset{\sim}{\mathbf{x}}^{(i)}$, which would lead to the error propagation matrix $\mathrm{N}\mathrm{M}_C^{-1}$, but this matrix has the same eigenvalues as the previous one, so we can analyse either of them with the same result.
For the Jacobi iteration, the spectral radius of $\mathrm{M}_C^{-1}\mathrm{N}$ (or of $\mathrm{N}\mathrm{M}_C^{-1}$ in the right preconditioned situation) should be strictly less than 1. For other methods, as GMRES, clustering of the eigenvalues of the error propagation matrix around 0 is a desirable property for fast convergence.

The kernel of N forms the space of eigenvectors of $\mathrm{M}_C^{-1}\mathrm{N}$ that are associated with eigenvalue 0. Consider an eigenvalue $\sigma \neq 0$ of $\mathrm{M}_C^{-1}\mathrm{N}$ with eigenvector $\underset{\sim}{\mathbf{z}} \equiv (\mathbf{z}_1^T, z_\ell^T, \tilde{z}_r^T, \tilde{z}_\ell^T, z_r^T, \mathbf{z}_2^T)^T$:

$$\mathrm{M}_C^{-1}\mathrm{N}\underset{\sim}{\mathbf{z}} = \sigma \underset{\sim}{\mathbf{z}}. \tag{13}$$

Since N maps all components, except for the $\tilde{\cdot}_\ell$ and $\tilde{\cdot}_r$ ones, to zero, we have that all components of $\mathrm{M}_C\underset{\sim}{\mathbf{z}}$, except for the $\tilde{\cdot}_\ell$ and $\tilde{\cdot}_r$ components, are zero. The eigenvalue problem $\sigma\mathrm{M}_C\underset{\sim}{\mathbf{z}} = \mathrm{N}\underset{\sim}{\mathbf{z}}$ can be decomposed into two coupled problems:

$$\sigma \begin{bmatrix} \mathbf{B}_{11} & \mathbf{B}_{1\ell} & \mathbf{B}_{1r} \\ \mathbf{B}_{\ell 1} & B_{\ell\ell} & B_{\ell r} \\ \mathbf{0} & C_{\ell\ell} & C_{\ell r} \end{bmatrix} \begin{bmatrix} \mathbf{z}_1 \\ z_\ell \\ \tilde{z}_r \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ 0 \\ g_r \end{bmatrix}, \quad \sigma \begin{bmatrix} C_{r\ell} & C_{rr} & \mathbf{0} \\ B_{r\ell} & B_{rr} & \mathbf{B}_{r2} \\ \mathbf{B}_{2\ell} & \mathbf{B}_{2r} & \mathbf{B}_{22} \end{bmatrix} \begin{bmatrix} \tilde{z}_\ell \\ z_r \\ \mathbf{z}_2 \end{bmatrix} = \begin{bmatrix} g_\ell \\ 0 \\ \mathbf{0} \end{bmatrix}, \tag{14}$$

with

$$g_r \equiv C_{\ell\ell}\,\widetilde{z}_\ell + C_{\ell r}\,z_r, \quad g_\ell \equiv C_{r\ell}\,z_\ell + C_{rr}\widetilde{z}_r. \tag{15}$$

In the context of PDEs, the systems in (14) can be interpreted as representing homogeneous partial differential equations with inhomogeneous boundary conditions along the artificial boundary: the left system for domain 1, the right system for domain 2. The values g_r and g_ℓ at the artificial boundaries are defined by (15): the value g_r for domain 1 is determined by the solution of the PDE at domain 2, while the solution of the PDE at domain 1 determines the value at the internal boundary of domain 2.

We have the following properties, that help to identify the relevant part of the eigensystem:

(i) N is an $n + 2n_i$ by $n + 2n_i$ matrix. Since C is nonsingular, we have that rank$(\mathrm{N}) = 2n_i$, and it follows that dim(ker$(\mathrm{N})) = n$. Hence, $\sigma = 0$ is an eigenvalue with geometric multiplicity n.

(ii) Since rank$(\mathrm{N}) = 2n_i$, there are at most $2n_i$ nonzero eigenvalues σ, counted according to algebraic multiplicity.

(iii) If σ is a nonzero eigenvalue then the corresponding components g_r and g_ℓ are non-zero. To see this, take $g_r = 0$. Then from (14) we have that $(\mathbf{z}_1^T, z_\ell^T, \widetilde{z}_r^T)^T = 0$. Hence, $g_\ell = 0$, so that $\underset{\sim}{\mathbf{z}}$ would be zero.

(iv) If σ is an eigenvalue with corresponding nonzero components g_r and g_ℓ then $-\sigma$ is an eigenvalue with eigenvector with components g_r and $-g_\ell$ (use (14) and (15)).

(v) The vector $\widetilde{\widetilde{z}}_\ell \equiv (z_\ell^T, \widetilde{z}_r^T)^T$ is linearly independent of $\widetilde{\widetilde{z}}_r \equiv (\widetilde{z}_\ell^T, z_r^T)^T$. To prove this, suppose that $\alpha\,\widetilde{\widetilde{z}}_\ell = \beta\,\widetilde{\widetilde{z}}_r$ for some $\alpha, \beta \neq 0$. Then, from (14) it follows that $\mathbf{B}\widetilde{\mathbf{z}} = 0$ where

$$\widetilde{\mathbf{z}} \equiv (\alpha\,\mathbf{z}_1^T, \alpha\,z_\ell^T, \alpha\,\widetilde{z}_r^T, \beta\,\mathbf{z}_2^T)^T = (\alpha\,\mathbf{z}_1^T, \beta\,\widetilde{z}_\ell^T, \beta\,z_r^T, \beta\,\mathbf{z}_2^T)^T.$$

As $\mathbf{B}$ is nonsingular, we have $\widetilde{\mathbf{z}} = 0$. Hence, $\underset{\sim}{\mathbf{z}} = \mathbf{0}$ and $\underset{\sim}{\mathbf{z}}$ is not an eigenvector.

Consequently the value of σ cannot be equal to ± 1. To prove this, suppose that $\sigma = 1$. Then by combining the last row of the left part and the first row of the right part of (14) with (15), we find that $C(\widetilde{\widetilde{z}}_\ell - \widetilde{\widetilde{z}}_r) = 0$. Since C is nonsingular, this implies that $\widetilde{\widetilde{z}}_\ell = \widetilde{\widetilde{z}}_r$, i.e. the vectors are linearly dependent. The value -1 for σ is then excluded on account of property (iv).

The magnitude of σ dictates the error reduction. From (14) and (15) it follows that

$$\begin{aligned} \sigma(C_{\ell\ell}z_\ell + C_{\ell r}\widetilde{z}_r) = g_r = C_{\ell\ell}\widetilde{z}_\ell + C_{\ell r}\,z_r \\ \sigma(C_{r\ell}\widetilde{z}_\ell + C_{rr}z_r) = g_\ell = C_{r\ell}z_\ell + C_{rr}\widetilde{z}_r, \end{aligned} \tag{16}$$

which leads to

$$|\sigma|^2 = \frac{(C_{\ell\ell}\widetilde{z}_\ell + C_{\ell r}\,z_r)^*(C_{r\ell}z_\ell + C_{rr}\widetilde{z}_r)}{(C_{\ell\ell}z_\ell + C_{\ell r}\,\widetilde{z}_r)^*(C_{r\ell}\widetilde{z}_\ell + C_{rr}\,z_r)}. \tag{17}$$

From (16) we conclude that multiplying both $C_{\ell\ell}$ and $C_{\ell r}$ by a nonsingular matrix does not affect the value of σ. Likewise, both $C_{r\ell}$ and C_{rr} may be multiplied by (another) singular matrix with no effect to σ. This can be exploited to bring the C matrices to some convenient form.

<u>The one-dimensional case</u>. We first study the one-dimensional case, because this will not only give some insight in how to reduce σ, but it will also be useful to control local situations in the two-dimensional case.

In this situation the problem simplifies: the matrices $C_{\ell\ell}$, $C_{\ell r}$, $C_{r\ell}$, and C_{rr} are scalars, and so are the vector parts z_ℓ, z_r, $\widetilde{z}_\ell$, and $\widetilde{z}_r$. Because of the freedom to scale the matrices (scalars), we may take C as

$$C = \begin{bmatrix} C_{\ell\ell} & C_{\ell r} \\ -C_{r\ell} & -C_{rr} \end{bmatrix} = \begin{bmatrix} 1 & \alpha_\ell \\ -\alpha_r & -1 \end{bmatrix}. \tag{18}$$

With $\mu_\ell \equiv \widetilde{z}_r / z_\ell$, $\mu_r \equiv \widetilde{z}_\ell / z_r$, we have from (17) that

$$|\sigma|^2 = \left| \frac{\mu_r + \alpha_\ell}{1 + \alpha_\ell \mu_\ell} \cdot \frac{\alpha_r + \mu_\ell}{\alpha_r \mu_r + 1} \right|. \tag{19}$$

The μ-values will be interpreted as local growth factors at the artificial boundary: μ_ℓ shows how $\underset{\sim}{\mathbf{z}}$ changes at the artificial boundary of the left domain; μ_r shows the same for the right domain.
Note that $\widetilde{\widetilde{z}}_\ell$ depends linearly on $\widetilde{\widetilde{z}}_r$ if $\mu_r \mu_\ell = 1$. Since this situation is excluded on account of property (v), we have that $\mu_r \mu_\ell \neq 1$. The best choice for the minimization of σ in (19) is obviously $\alpha_\ell = -\mu_r$ and $\alpha_r = -\mu_\ell$, leading to $\sigma = 0$, which gives optimal damping.

The optimal choice for α_ℓ and α_r results in a coupling that annihilates the 'outflow' g_r and g_ℓ of the two domains. This leads effectively to two uncoupled subdomains: an ideal situation.

<u>More dimensions</u>. In the realistic case of a more dimensional overlap ($n_i > 1$), there is no choice for α_ℓ and α_r (i.e., $C_{\ell\ell} = I$, $C_{\ell r} = \alpha_\ell I$, etc.) that leads to an error reduction matrix with only trivial eigenvalues. But, the conclusion that the outflow should be minimized in some average sense for the best error reduction is here also correct. In our application in §4, we will identify coupling matrices C that lead to satisfactory clustering of most of the eigenvalues σ, of the error propagation matrix, around 0. We will do so by selecting the α_r and α_ℓ as suitable averages of the local growth factors μ_r and μ_ℓ.

3 The eigenvalue problem

3.1 The Jacobi-Davidson method

For the computation of a solution to an eigenvalue problem the Jacobi-Davidson method [17], is an iterative method that in each iteration:

1. computes an approximation for an eigenpair from a given subspace, using a Rayleigh-Ritz principle,
2. computes a correction for the eigenvector from a so-called correction equation,
3. expands the subspace with the computed correction.

The correction equation mentioned in step 2 is characteristic for the Jacobi-Davidson method, for example, the Arnoldi method [1, 13] simply expands the subspace with the residual for the approximated eigenpair, and the Davidson method [6] expands the subspace with a preconditioned residual. The success of the Jacobi-Davidson method depends on how fast good approximations for the correction equation can be obtained and it is for that purpose that we will try to exploit the enhancement techniques discussed in the previous section.

Therefore, we will consider this correction equation in some more detail. We will do this for the standard eigenvalue problem

$$\mathbf{A}\,\mathbf{x} = \lambda\,\mathbf{x}. \tag{20}$$

Given an approximate eigenpair $(\theta, \mathbf{u})$ (with residual $\mathbf{r} \equiv \theta\mathbf{u} - \mathbf{A}\mathbf{u}$) that is close to some wanted eigenpair $(\lambda, \mathbf{x})$, a correction $\mathbf{t}$ for the normalized $\mathbf{u}$ is computed from the correction equation:

$$\mathbf{t} \perp \mathbf{u}, \qquad (\mathbf{I} - \mathbf{u}\mathbf{u}^*)(\mathbf{A} - \theta\,\mathbf{I})(\mathbf{I} - \mathbf{u}\mathbf{u}^*)\,\mathbf{t} = \mathbf{r}, \tag{21}$$

or in augmented formulation ([15, §3.4])

$$\begin{bmatrix} \mathbf{A} - \theta\,\mathbf{I} & \mathbf{u} \\ \mathbf{u}^* & 0 \end{bmatrix} \begin{bmatrix} \mathbf{t} \\ \varepsilon \end{bmatrix} = \begin{bmatrix} \mathbf{r} \\ 0 \end{bmatrix}. \tag{22}$$

In many situations it is quite expensive to solve this correction equation accurately and fortunately it is also not always necessary to do so. A common technique is to compute an approximation for $\mathbf{t}$ by a few steps of a preconditioned iterative method, such as GMRES or Bi-CGSTAB.

When a preconditioner M for $\mathbf{A} - \theta\,\mathbf{I}$ is available, then $(\mathbf{I} - \mathbf{u}\,\mathbf{u}^*)\mathrm{M}(\mathbf{I} - \mathbf{u}\,\mathbf{u}^*)$ can be used as left preconditioner for (21). This leads to the linear system (see, [17, §4])

$$\mathbf{P}\,\mathrm{M}^{-1}(\mathbf{A} - \theta\,\mathbf{I})\,\mathbf{P}\,\mathbf{t} = \mathbf{P}\,\mathrm{M}^{-1}\,\mathbf{r} \quad \text{where} \quad \mathbf{P} \equiv \mathbf{I} - \frac{\mathrm{M}^{-1}\,\mathbf{u}\,\mathbf{u}^*}{\mathbf{u}^*\,\mathrm{M}^{-1}\,\mathbf{u}}. \tag{23}$$

The operator at the left hand side in (23) involves two (skew) projectors $\mathbf{P}$. However, when we start the iterative solution process for (23) with initial guess $\mathbf{0}$, then $\mathbf{P}\mathbf{t}$ may be replaced with $\mathbf{t}$ at each iteration of a Krylov iteration method: projection at the right can be skipped in each step of the Krylov subspace solver.

Right preconditioning, which has advantages in the domain decomposition approach, can be carried out in a similar way, with similar reductions in the application of $\mathbf{P}$, as we will see in §3.3 below. However, because the formulas with right preconditioning look slightly more complicated, we will present our arguments mainly for left preconditioning.

3.2 Enhancement of the correction equation

We use the domain decomposition approach as presented in §2 to solve the correction equation (21). Again, we will assume that we have two subdomains and we will use the same notations for the enhanced vectors. With $\mathbf{B} \equiv \mathbf{A} - \theta\,\mathbf{I}$ this leads to the enhanced Jacobi-Davidson correction equation

$$\underline{\mathbf{t}} \perp \underline{\mathbf{u}}, \quad (\mathbf{I} - \underline{\mathbf{u}}\,\underline{\mathbf{u}}^*)\,\mathbf{B}_C\,(\mathbf{I} - \underline{\mathbf{u}}\,\underline{\mathbf{u}}^*)\,\underline{\mathbf{t}} = \underline{\mathbf{r}} \tag{24}$$

with $\underline{\mathbf{u}} \equiv (\mathbf{u}_1^T, u_\ell^T, 0^T, 0^T, u_r^T, \mathbf{u}_2^T)^T$, and likewise $\underline{\mathbf{r}} \equiv (\mathbf{r}_1^T, r_\ell^T, 0^T, 0^T, r_r^T, \mathbf{r}_2^T)^T$. The dimension of the zero parts, indicated by 0, is assumed to be the same as the dimension of u_ℓ (and u_r).
To see why this is correct, apply the enhancements of §2 to the augmented formulation (22) of the correction equation, and use the fact that the augmented and the projected form are equivalent. We assume $\mathbf{u}$ to be normalized. Then $\underline{\mathbf{u}}$ is normalized as well.

With

$$(\mathbf{I} - \underline{\mathbf{u}}\,\underline{\mathbf{u}}^*)\mathrm{M}_C(\mathbf{I} - \underline{\mathbf{u}}\,\underline{\mathbf{u}}^*) \tag{25}$$

as the left preconditioner, we obtain

$$\mathbf{P}\,\mathrm{M}_C^{-1}\mathbf{B}_C\,\mathbf{P}\,\underline{\mathbf{t}} = \mathbf{P}\,\mathrm{M}_C^{-1}\underline{\mathbf{r}} \quad \text{with} \quad \mathbf{P} \equiv \mathbf{I} - \frac{\mathrm{M}_C^{-1}\underline{\mathbf{u}}\,\underline{\mathbf{u}}^*}{\underline{\mathbf{u}}^*\mathrm{M}_C^{-1}\underline{\mathbf{u}}}. \tag{26}$$

In comparison with the error propagation (12) of the block Jacobi method for ordinary linear systems, the error propagation matrix $\mathrm{M}_C^{-1}\mathrm{N}$ is now embedded by the projections $\mathbf{P}$. These projections prevent the operator in the correction equation from getting (nearly) singular: as θ approximates the wanted eigenvalue λ, in the asymptotic case θ is even equal to λ, $\mathbf{B}$ gets close to singular in the direction of the wanted eigenvector $\mathbf{x}$. For ordinary linear systems this possibility is excluded by imposing $\mathbf{B}$ to be nonsingular (see remark (v) in §2.5). Here we have to allow a singular $\mathbf{B}$. In our analysis of the propagation matrix of the correction equation, for the model problem in §4.3, in first instance we will ignore the projections. Afterwards, we will justify this (both analytically (§4.3) as well as numerically (§5.2)).

Note. We have enhanced the correction equation. Another option is to start with an enhancement of the eigenvalue problem itself. However, this does not result in essential differences ([9]). If the correction equations for these two different approaches are solved exactly, then the approaches are even equivalent.

3.3 Right preconditioning

In §2.4 we have showed that, without projections, right preconditioning for domain decomposition leads to an equation that is defined by its behavior on the artificial boundary only. Although the projections slightly complicate matters, the computations for the projected equation can also be restricted to vectors corresponding to the artificial boundary, as we will see below. Moreover, similar to the situation for left preconditioning, right preconditioning requires only one projection per iteration of a Krylov subspace method. In this section, we will use the underscore notation for vectors in order to emphasize that they are defined in the enhanced space.

First we analyze the action of the right preconditioned matrix.
The inverse on $\underline{\mathbf{u}}^{\perp}$ of the projected preconditioner in (25) is equal to (cf. [15, §7.1.1] and [8])

$$\mathbf{P}\mathrm{M}_C^{-1} = \left(\mathbf{I} - \frac{\mathrm{M}_C^{-1}\underline{\mathbf{u}}\,\underline{\mathbf{u}}^*}{\underline{\mathbf{u}}^*\mathrm{M}_C^{-1}\underline{\mathbf{u}}}\right)\mathrm{M}_C^{-1} = \mathrm{M}_C^{-1}\left(\mathbf{I} - \frac{\underline{\mathbf{u}}\,\underline{\mathbf{u}}^*\mathrm{M}_C^{-1}}{\underline{\mathbf{u}}^*\mathrm{M}_C^{-1}\underline{\mathbf{u}}}\right), \tag{27}$$

with $\mathbf{P}$ as in (26). This expression represents the Moore–Penrose inverse of the operator in (25), on the entire space. Note that $\underline{\mathbf{u}}^*\mathbf{P} = \mathbf{0}$ (by definition of $\mathbf{P}$) and $\underline{\mathbf{u}}^*\mathrm{N} = \mathbf{0}$ (by definition of $\underline{\mathbf{u}}$ and N). Therefore, for the operator that is involved in right preconditioning (cf. (11)), we have that

$$\begin{aligned}
&(\mathbf{I} - \underline{\mathbf{u}}\,\underline{\mathbf{u}}^*)\mathrm{B}_C(\mathbf{I} - \underline{\mathbf{u}}\,\underline{\mathbf{u}}^*)\mathbf{P}\mathrm{M}_C^{-1} \\
&\qquad = (\mathbf{I} - \underline{\mathbf{u}}\,\underline{\mathbf{u}}^*)\mathrm{B}_C\mathbf{P}\mathrm{M}_C^{-1} \\
&\qquad = (\mathbf{I} - \underline{\mathbf{u}}\,\underline{\mathbf{u}}^*)\mathrm{B}_C\mathrm{M}_C^{-1}\left(\mathbf{I} - \frac{\underline{\mathbf{u}}\,\underline{\mathbf{u}}^*\mathrm{M}_C^{-1}}{\underline{\mathbf{u}}^*\mathrm{M}_C^{-1}\underline{\mathbf{u}}}\right), \\
&\qquad = \mathbf{I} - \underline{\mathbf{u}}\,\underline{\mathbf{u}}^* - (\mathbf{I} - \underline{\mathbf{u}}\,\underline{\mathbf{u}}^*)\mathrm{N}\mathbf{P}\mathrm{M}_C^{-1} \\
&\qquad = \mathbf{I} - \underline{\mathbf{u}}\,\underline{\mathbf{u}}^* - \mathrm{N}\mathbf{P}\mathrm{M}_C^{-1}.
\end{aligned} \tag{28}$$

Hence, this operator maps a vector $\underset{\sim}{\mathbf{v}}$ that is orthogonal to $\underline{\mathbf{u}}$ to the vector

$$(\mathbf{I} - \underline{\mathbf{u}}\,\underline{\mathbf{u}}^*)\mathrm{B}_C(\mathbf{I} - \underline{\mathbf{u}}\,\underline{\mathbf{u}}^*)\mathbf{P}\mathrm{M}_C^{-1}\underset{\sim}{\mathbf{v}} = \underset{\sim}{\mathbf{v}} - \mathrm{N}\mathbf{P}\mathrm{M}_C^{-1}\underset{\sim}{\mathbf{v}}$$

that is also orthogonal to $\underline{\mathbf{u}}$.

Therefore, right preconditioning for (24) can be carried out in the following steps (cf. §2.4):

1. Compute $\quad \underset{\sim}{\mathbf{t}}^{(0)} \equiv \mathbf{P}\mathrm{M}_C^{-1}\underline{\mathbf{r}} \quad$ and $\quad \underset{\sim}{\mathbf{r}}^{(0)} \equiv \mathrm{N}\underset{\sim}{\mathbf{t}}^{(0)}$.

2. Compute an (approximate) solution $\underset{\sim}{\mathbf{s}}^{(m)}$ of

$$(\mathbf{I} - \mathrm{N}\mathbf{P}\mathrm{M}_C^{-1})\underset{\sim}{\mathbf{s}} = \underset{\sim}{\mathbf{r}}^{(0)},$$

   with (m steps of) a Krylov subspace method with initial guess $\mathbf{0}$.

3. Update $\underset{\sim}{\mathbf{t}}^{(0)}$ to the (approximate) solution $\underline{\mathbf{t}}$ of (24):

$$\underline{\mathbf{t}} = \underset{\sim}{\mathbf{t}}^{(0)} + \mathbf{P}\mathbf{M}_C^{-1}\underset{\sim}{\mathbf{s}}^{(m)}.$$

As in §2.4, the intermediate vectors in the solution process for the equation in step 2 vanish outside the artificial boundary. Therefore, for the solution of the right preconditioned enhanced correction equation, only $2n_i$-dimensional vectors have to be stored, and the vector updates and dot products are also for vectors of length $2n_i$.

4 Tuning of the coupling matrix for a model problem

Now we will address the problem whether it is possible to reduce the computing time for the Jacobi-Davidson process, by an appropriate choice of the coupling matrix C. We have, in §2, introduced the decomposition of a linear system, into two coupled subsystems, in an algebraic way. In this section we will demonstrate how knowledge of the physical equations from which the linear system originates can be used for tuning of the coupling parameters.

4.1 The model problem

As a model problem we will consider the two-dimensional advection-diffusion operator:

$$\mathcal{L}(\widehat{\varphi}) \equiv a\frac{\partial^2}{\partial x^2}\,\widehat{\varphi} + b\frac{\partial^2}{\partial y^2}\,\widehat{\varphi} + u\frac{\partial}{\partial x}\,\widehat{\varphi} + v\frac{\partial}{\partial y}\,\widehat{\varphi} + c\widehat{\varphi}, \tag{29}$$

that is defined on the open domain $\Omega = (0,\omega_x) \times (0,\omega_y)$ in $\mathbb{R}^2$, with constants $a > 0$, $b \geq 0$, c, u and v. We will further assume Dirichlet boundary conditions: $\widehat{\varphi} = 0$ on $\partial\Omega$ of Ω. We are interested in some eigenvalue $\widehat{\lambda} \in \mathbb{C}$ and corresponding eigenfunction $\widehat{\varphi}$ of $\mathcal{L}$:

$$\begin{cases} \mathcal{L}(\widehat{\varphi}) = \widehat{\lambda}\,\widehat{\varphi} & \text{on} \quad \Omega, \\ \widehat{\varphi} = 0 & \text{on} \quad \partial\Omega. \end{cases} \tag{30}$$

We will use the insights, obtained with this simple model problem, for the construction of couplings for more complicated partial differential operators.

Discretization. We discretize $\mathcal{L}$ with central differences with stepsize $h = (h_x, h_y) = (\frac{\omega_x}{n_x+1}, \frac{\omega_y}{n_y+1})$ for the second order part and stepsize $2h = (2h_x, 2h_y)$ for the first order part, where n_x and n_y are positive integers:

$$\widehat{L}(\widehat{\varphi}) \equiv a\frac{\delta_x^2}{h_x^2}\,\widehat{\varphi} + b\frac{\delta_y^2}{h_y^2}\,\widehat{\varphi} + u\frac{\delta_x}{2h_x}\,\widehat{\varphi} + v\frac{\delta_y}{2h_y}\,\widehat{\varphi} + c\widehat{\varphi}. \tag{31}$$

The operator $\dfrac{\delta_x}{h_x}$ denotes the central difference operator, defined as

$$\frac{\delta_x}{h_x}\,\widehat{\psi}(x,y) \equiv \frac{\widehat{\psi}(x+\frac{1}{2}h_x, y) - \widehat{\psi}(x-\frac{1}{2}h_x, y)}{h_x},$$

and $\dfrac{\delta_y}{h_y}$ is defined similar. This leads to the discretized eigenvalue problem

$$\begin{cases} L(\varphi) = \lambda\varphi & \text{on} \quad \Omega_h, \\ \varphi = 0 & \text{on} \quad \partial\Omega_h, \end{cases} \tag{32}$$

where Ω_h and $\partial\Omega_h$ is the uniform rectangular grid of points $(j_x h_x, j_y h_y)$ in Ω and in $\partial\Omega$, respectively. We have skipped the hat $\widehat{\cdot}$ in order to indicate that the functions are restricted to the appropriate grid, and that the operator L is restricted to grid functions. The vector φ is defined on $\Omega_h \cup \partial\Omega_h$.

We use the boundary conditions $\varphi = 0$ at $\partial\Omega_h$ for the elimination of these values of φ from $L(\varphi) = \lambda\varphi$.

Identification of grid functions with vectors and of operators on grid functions with matrices leads to an eigenvalue problem as in (20) of dimension $n \equiv n_x \cdot n_y$: the eigenvector $\mathbf{x}$ corresponds to the eigenfunction φ restricted to Ω_h. The matrix $\mathbf{A}$ corresponds to the operator L from which the boundary conditions have been eliminated. In our application, we obtain the corresponding vectors by enumeration of the grid points from bottom to top first (i.e., the y-coordinates first) and then from left to right ([21, §6.3]). In our further analysis, we will switch from one representation to another (grid function or vector), selecting the representation that is the most convenient at that moment.

4.2 Decomposition of the physical domain

For some $0 < \omega_{x1} < \omega_x$ we decompose the domain Ω in two subdomains $\Omega_1 \equiv (0, \omega_{x1}] \times (0, \omega_y)$ and $\Omega_2 \equiv (\omega_{x1}, \omega_x) \times (0, \omega_y)$.

Let n_{x1} be the number of grid points in the x direction in Ω_1. Then $\Omega_1 \cap \Omega_h$ and $\Omega_2 \cap \Omega_h$ is an $n_{x1} \times n_y$ and $n_{x2} \times n_y$ grid respectively with $n_{x1} + n_{x2} = n_x$. To number the grid points in the x direction, we use local indices j_{x1}, $1 \leq j_{x1} \leq n_{x1}$, and j_{x2}, $1 \leq j_{x2} \leq n_{x2}$, in Ω_1 and Ω_2 respectively.

Because of the 5 point star discretization, the unknowns at the last row of grid points ($j_{x1} = n_{x1}$) in the y direction in Ω_1 are coupled with those at the first row of grid points ($j_{x2} = 1$) in the y direction in Ω_2, and vice versa. The unknowns for $j_{x1} = n_{x1}$ are denoted by the vector y_ℓ, and the unknowns for $j_{x2} = 1$ are denoted by y_r, just as in §2. Now we enhance the system with the unknowns $\widetilde{y}_r$ and $\widetilde{y}_\ell$, which, in grid terminology, correspond to a virtual new row of gridpoints to the right of Ω_1, and the left of Ω_2, respectively. These new virtual gridpoints serve as boundary points for the domains Ω_1 and Ω_2. See Fig. 1 for an illustration.

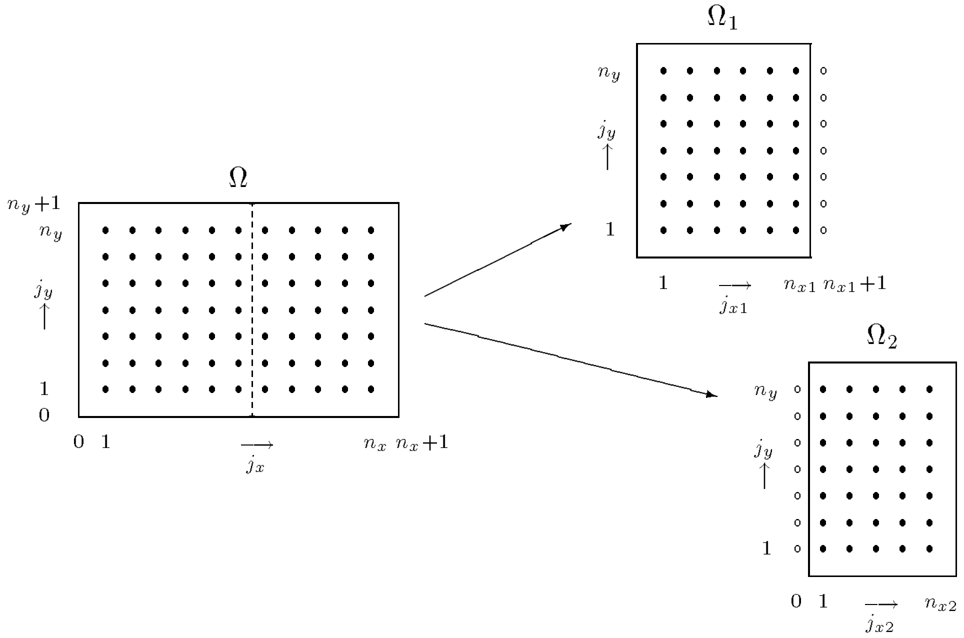
The vectors y_ℓ, y_r, $\widetilde{y}_\ell$, and $\widetilde{y}_r$ are n_y dimensional (the n_i in §2.1 is now equal to n_y). The $2n_y$ by $2n_y$ matrix C, that couples y_ℓ, $\widetilde{y}_r$, $\widetilde{y}_\ell$, and y_r can be interpreted as discretized boundary conditions of the differential operator at the internal newly created boundary between Ω_1 and Ω_2 [19, 18].
Note that the internal boundary conditions are explicitly expressed in the total system matrix $\mathbf{B}_C$, through C, whereas the external boundary conditions have been used to eliminate the values at the external boundary (see §4.1).

4.3 Eigenvectors of the error propagation matrix

We will now analyze the eigensystem of the error reduction matrix $\mathbf{M}_C^{-1}\mathbf{N}$ (see §2.5) and discuss appropriate coupling conditions (that is, the internal boundary conditions) as represented by the matrix C. Here, the matrices $\mathbf{M}_C$ and $\mathbf{N}$ are defined for $\mathbf{B} \equiv \mathbf{A} - \theta\mathbf{I}$, as explained in §§2.2-2.3, for some approximate eigenvalue θ (cf., §§3.1-3.2). The matrix $\mathbf{A}$ corresponds to L, as explained in §4.1.

First, we will discuss in section §4.3.1 the case of one spatial dimension (i.e., no y variable). The results for the one-dimensional case are easy to interpret. Moreover, since the two-dimensional eigenvalue problem in (30) is a tensor product of two one-dimensional problems, the results for the one-dimensional case can conveniently be used for the analysis in §4.3.2 of the two-dimensional problem.

FIGURE 1. *Decomposition of the domain Ω into two subdomains Ω_1 and Ω_2.*
The bullets (•) represent the grid points of the original grid. The circles (o) represent the extra grid points at the internal boundary. The indices j_x and j_y refer to numbering in the x direction and y direction respectively of the grid points in the grids: the pair (j_x, j_y) corresponds to point $(j_x h_x, j_y h_y)$ in Ω. For the numbering of the grid points in the x direction in the two subdomains a local index is used: $j_{x1} = j_x$ in Ω_1 ($0 \leq j_{x1} \leq n_{x1}+1$) and $j_{x2} = j_x - n_{x1}$ in Ω_2 ($0 \leq j_{x2} \leq n_{x2}+2$).



4.3.1 The one-dimensional case

In this section, we will discuss the case of one spatial dimension: there is no y variable. To simplify notations, we will skip the index x for this case.

Suppose that we have an approximate eigenvalue θ for some eigenvalue θ of $\mathbf{B}$.
To simplify formulas, we shift the approximate eigenvalue by c. The matrix $\mathbf{B}$ in §2.5 corresponds to the three point stencil of the finite difference operator

$$a\,\frac{\delta^2}{h^2} + u\,\frac{\delta}{2h} - \theta.$$

For the eigensystem of $\mathrm{M}_C^{-1}\mathrm{N}$, we have to solve the systems in (14) for an $\widetilde{x}_r \neq 0$ and $\widetilde{x}_\ell \neq 0$, that is, we have to compute solutions ψ_1 and ψ_2 for the discretized PDE on domain 1 and domain 2, respectively (cf. §2.5). The functions ψ_1 and ψ_2 should satisfy

$$\left[a\,\frac{\delta^2}{h^2} + u\,\frac{\delta}{2h} - \theta\right]\psi_p(j_p h) = 0 \quad \text{for} \quad 1 \leq j_p \leq n_p \quad \text{and} \quad p = 1, 2. \tag{33}$$

The conditions on the external boundaries imply that

$$\psi_1(0) = 0 \quad \text{and} \quad \psi_2(n_2 h + h) = 0.$$

For the solutions of (33), we try functions of the form $\psi(jh) = \zeta^j$. Then ζ satisfies

$$\left(1 + \tfrac{uh}{2a}\right)\zeta - 2D + \left(1 - \tfrac{uh}{2a}\right)\zeta^{-1} = 0 \quad \text{with} \quad D \equiv 1 + \frac{h^2}{2a}\theta. \tag{34}$$

Let ζ_+ and ζ_- denote the roots of this equation, such that $|\zeta_+| \geq |\zeta_-|$. In the regular case where $\zeta_+ \neq \zeta_-$, the solutions ψ_1 and ψ_2 are, apart from scaling, given by

$$\psi_1(j_1 h) = \zeta_+^{j_1} - \zeta_-^{j_1} \quad \text{and} \quad \psi_2(j_2 h) = \zeta_-^{j_2-n_2-1} - \zeta_+^{j_2-n_2-1}.$$

We distinguish three different situations:

(i) *Harmonic behavior:* $\zeta_- = \bar{\zeta}_+ \notin \mathbb{R}$.
If $\zeta_0 \in \mathbb{R}$ and $\tau \in [0, 2\pi)$ are such that $\zeta_+ = \zeta_0 \exp(i\tau)$. Then, up from scaling factors,

$$\psi_1(j_1 h) = \zeta_0^{j_1} \sin(\tau j_1) \quad \text{and} \quad \psi_2(j_2 h) = \zeta_0^{j_2} \sin(\tau(j_2 - n_2 - 1)).$$

(ii) *Degenerated harmonic behavior:* $\zeta_+ = \zeta_-$.
In this case we have, apart from scaling factors,

$$\psi_1(j_1 h) = j_1 \zeta_0^{j_1} \quad \text{and} \quad \psi_2(j_2 h) = (n_2 + 1 - j_2)\zeta_0^{j_2}.$$

(iii) *Dominating behavior:* $|\zeta_+| > |\zeta_-|$.
Near the artificial boundary, that is for $j_1 \approx n_1$ and $j_2 \approx 1$, we have apart from scaling factors that

$$\psi_1(j_1 h) = \zeta_+^{j_1} \left(1 - \left(\frac{\zeta_-}{\zeta_+}\right)^{j_1}\right) \approx \zeta_+^{j_1}$$

and

$$\psi_2(j_2 h) = \zeta_-^{j_2-n_2-1} \left(1 - \left(\frac{\zeta_-}{\zeta_+}\right)^{n_2+1-j_2}\right) \approx c\zeta_-^{j_2},$$

so that, apart from a scaling factor again, $\psi_2(j_2 h) \approx \zeta_-^{j_2}$.
How accurate the approximation is depends on the ratio $|\zeta_-|/|\zeta_+|$ and on the size of n_1 and n_2.

The coupling matrix C is 2 by 2 ($n_i = 1$). We consider a C as in (18). Then, according to (19), the absolute value of the eigenvalue σ is given by

$$|\sigma|^2 = \left|\frac{\alpha_\ell + \mu_r}{1 + \alpha_r \mu_r}\right| \left|\frac{\alpha_r + \mu_\ell}{1 + \alpha_\ell \mu_\ell}\right|, \tag{35}$$

where $\mu_\ell = \psi_1(n_1 h + h)/\psi_1(n_1 h)$ and $\mu_r = \psi_2(0)/\psi_2(h)$: z_ℓ in (14) corresponds to $\psi_1(n_1 h)$, $\tilde{z}_r$ to $\psi_1(n_1 h + h)$, etcetera.

In the case of dominating behavior (cf. (iii)), we have that $\mu_\ell \approx \zeta_+$ and $\mu_r \approx 1/\zeta_-$. As observed in (iii), the accuracy of the approximation depends on the ratio $|\zeta_-|/|\zeta_+|$ and on the values of n_1 and n_2. But already for modest (and realistic) values of these quantities, we obtain useful estimates, and we may expect a good error reduction for the choice $\alpha_\ell = -1/\zeta_-$ and $\alpha_r = -\zeta_+$. The parameters ζ_+ and ζ_- would also appear in a local mode analysis: they do not depend on the external boundary condition nor on the position of the artificial boundary.

The value for $|\sigma|$ in (35) is equal to one when $\mu_r = 1/\bar{\mu}_\ell$, regardless α_ℓ and α_r (assuming these are real). If we would follow the local mode approach for the situations (i) and (ii), that is, if we would estimate μ_ℓ by ζ_+ and μ_r by $1/\zeta_-$, then we would encounter such values for μ_ℓ and μ_r. In specific situations, we may do better by using the expressions for ψ_1 and ψ_2 in (i) and (ii), that is, we may find coupling parameters α_ℓ and α_r that lead to an eigenvalue σ with $|\sigma| < 1$. However, then we need information on the external boundary conditions and the position of the artificial boundary. Certainly in the case of a higher spatial dimension, this is undesirable. Moreover, if θ is an exact eigenvalue of $\mathbf{A}$ then we are in the situation in (i): the functions ψ_1 and ψ_2 are multiples of the components on domain 1 and domain 2, respectively, of the eigenfunction and $\sigma = 1$ (see (v) in §2.5 and the remark in §3.2). In this case there is no value of α_ℓ and α_r for which $|\sigma| < 1$.

We define $\nu \equiv (2a + uh)/(2a - uh)$. In order to simplify the forthcoming discussion for two spatial dimensions, observe that, in the case of dominating growth (iii), that is, $\mu_\ell \approx \zeta_+$ and $\mu_r \approx 1/\zeta_-$, (35) implies that

$$|\sigma|^2 \approx \left|\frac{\widetilde{\alpha}_\ell + \widetilde{\zeta}}{1 + \widetilde{\alpha}_\ell\, \widetilde{\zeta}}\right| \left|\frac{\widetilde{\alpha}_r + \widetilde{\zeta}}{1 + \widetilde{\alpha}_r\, \widetilde{\zeta}}\right|, \text{ where } \widetilde{\alpha}_\ell \equiv \frac{\alpha_\ell}{\sqrt{\nu}},\ \widetilde{\alpha}_r \equiv \sqrt{\nu}\alpha_r,\ \widetilde{\zeta} \equiv \sqrt{\nu}\zeta_+. \tag{36}$$

Here we have used that $\zeta_+ \cdot \zeta_- = 1/\nu$, which follows from (34).

If, for the Laplace operator (where $u = 0$ and $c = 0$), we use Ritz values for the approximate eigenvalues θ, then θ takes values between $\lambda^{(n)}$ and $\lambda^{(0)}$. Hence, $\theta \in (-4a/h^2, 0)$, and the roots ζ_+ and ζ_- are always complex conjugates. We will see in the next subsections that, for two spatial dimensions, the Ritz values that are of interest lead to a dominant root, also for the Laplace operator, and we will see that local mode analysis is then a convenient tool for the identification of effective coupling parameters.

4.3.2 Two dimensions

Similar to the one-dimensional case we are interested in functions χ_1 and χ_2 such that,

$$L(\chi_p) = 0 \quad \text{on} \quad \Omega_h \cap \Omega_p, \quad p = 1, 2, \tag{37}$$

and that satisfy the external boundary conditions. But now χ_1 and χ_2 are functions that depend on both the x- and y direction whereas the operator L (here L is introduced in §4.1) acts in these two directions. Since the finite difference operator $\frac{\delta_x}{h_x}$ acts only in the x direction and $\frac{\delta_y}{h_y}$ acts only in the y direction, their actions are independent of each other. Therefore, in this case of constant coefficients[1], we can write the operator L in equation (37) as a sum of tensor product of one-dimensional operators:

$$L = L_x \otimes \mathbf{I} + \mathbf{I} \otimes L_y, \tag{38}$$

where

$$L_x \equiv a\frac{\delta_x^2}{h_x^2} + u\frac{\delta_x}{2h_x} \quad \text{and} \quad L_y \equiv b\frac{\delta_y^2}{h_y^2} + v\frac{\delta_y}{2h_y} + c - \theta. \tag{39}$$

L_x and L_y incorporate the action of L in the x direction and y direction respectively.

Since the domain Ω is rectangular and since on each of the four boundary sides of Ω we have the same boundary conditions, the tensor product decomposition of L corresponds to a tensor product decomposition of the matrix $\mathbf{A}$.

We try to construct solutions of (37) by tensor product functions, that is by functions χ_p of the form

$$\chi_p(j_{xp}h_x, j_y h_y) = \psi_p(j_{xp}h_x) \otimes \varphi(j_y h_y) = \psi_p(j_{xp}h_x) \cdot \varphi(j_y h_y).$$

For φ we select eigenfunctions $\varphi^{(l)}$ of the operator L_y that satisfy the boundary conditions for the y direction. Then

$$L(\chi_p) = (L_x\psi_p) \otimes \varphi^{(l)} + \psi_p \otimes \lambda^{(l)}\varphi^{(l)} = (L_x + \lambda^{(l)})(\psi_p) \otimes \varphi^{(l)},$$

[1] It is sufficient if a and u are constants as functions of y, b and v are constants as function of x, and c is a product of a function in x and a function in y.

where $\lambda^{(l)}$ is the eigenvalue of L_y that corresponds to $\varphi^{(l)}$. Apparently, for each eigensolution of the 'y-operator' L_y, the problem of finding solutions of (37) reduces to a one-dimensional problem as discussed in the previous subsection: find ψ_p such that

$$(L_x + \lambda^{(l)})(\psi_p) = \left[a\frac{\delta_x^2}{h_x^2} + u\frac{\delta_x}{2h_x} + \lambda^{(l)} \right] \psi_p = 0, \tag{40}$$

and that satisfy the external boundary conditions in the x direction. To express the dependency of the solutions ψ_p on the selected eigenfunction of L_y, we denote the solution as $\psi_p^{(l)}$.

Now, consider matrixpairs $(C_{\ell r}, C_{\ell\ell})$ and $(C_{r\ell}, C_{rr})$ for which the eigenfunctions $\varphi^{(l)}$ of L_y are also eigenfunctions:

$$C_{\ell r}\varphi^{(l)} = \alpha_\ell^{(l)} C_{\ell\ell}\varphi^{(l)} \quad \text{and} \quad C_{r\ell}\varphi^{(l)} = \alpha_r^{(l)} C_{rr}\varphi^{(l)}. \tag{41}$$

Examples of such matrices are scalar multiples of the identity matrix (for instance, $C_{\ell r} = \alpha_\ell^{(l)} I$ and $C_{\ell\ell} = I$), but there are others as well, as we will see in §4.4. For such a C there is a 1–1 correspondence for each function $\varphi^{(l)}$ on the two subdomains: a component in the direction of $\psi_1^{(l)} \otimes \varphi^{(l)}$ on subdomain 1 is transferred by $\mathrm{M}_C^{-1}\mathrm{N}$ to a component in the direction of $\psi_2^{(l)} \otimes \varphi^{(l)}$ on subdomain 2 and vice versa. More precisely, if C is such that (41) holds and if $\psi^{(l)} \equiv (c_l\psi_1^{(l)}, \psi_2^{(l)})^T$ for some scalar c_l then, by construction of $\psi^{(l)}$, M_C maps $\psi^{(l)} \otimes \varphi^{(l)}$ onto a vector that is zero except for the $\tilde{\cdot}_\ell$ and $\tilde{\cdot}_r$ components (cf. (14)) which are equal to

$$c_l \left(\psi_1^{(l)}(n_{1x}h_x) + \alpha_\ell^{(l)} \psi_1^{(l)}(n_{1x}h_x + h_x) \right) C_{\ell\ell}\varphi^{(l)} \tag{42}$$

and

$$\left(\alpha_r^{(l)} \psi_2^{(l)}(0) + \psi_2^{(l)}(h_x) \right) C_{rr}\varphi^{(l)}, \tag{43}$$

respectively. In its turn, N maps $\psi^{(l)} \otimes \varphi^{(l)}$ onto a vector that is zero except for the $\tilde{\cdot}_\ell$ and $\tilde{\cdot}_r$ components (cf. (14) and (15)) which are equal to

$$\left(\psi_2^{(l)}(0) + \alpha_\ell^{(l)} \psi_2^{(l)}(h_x) \right) C_{\ell\ell}\varphi^{(l)} \tag{44}$$

and

$$c_l \left(\alpha_r^{(l)} \psi_1^{(l)}(n_{1x}h_x) + \psi_1^{(l)}(n_{1x}h_x + h_x) \right) C_{rr}\varphi^{(l)}, \tag{45}$$

respectively. By a combination of (42) and (44), and (43) and (45), respectively, one can check that, for an appropriate scalar c_l, $\psi^{(l)} \otimes \varphi^{(l)}$ is an eigenvector of $\mathrm{M}_C^{-1}\mathrm{N}$ with corresponding eigenvalue $\sigma^{(l)}$ such that

$$|\sigma^{(l)}|^2 = \left| \frac{\alpha_\ell^{(l)} + \mu_r^{(l)}}{1 + \alpha_r^{(l)}\mu_r^{(l)}} \right| \left| \frac{\alpha_r^{(l)} + \mu_\ell^{(l)}}{1 + \alpha_\ell^{(l)}\mu_\ell^{(l)}} \right|, \tag{46}$$

where (here we assumed that $\psi_1^{(l)}(n_{1x}h_x) \neq 0$ and $\psi_2^{(l)}(h_x) \neq 0$)

$$\mu_\ell^{(l)} \equiv \psi_1^{(l)}(n_{1x}h_x + h_x)/\psi_1^{(l)}(n_{1x}h_x) \quad \text{and} \quad \mu_r^{(l)} \equiv \psi_2^{(l)}(0)/\psi_2^{(l)}(h_x).$$

Note that the expression for $\sigma^{(l)}$ does not involve the value of c_l. From property (iv) in §2.5 we know that $\psi_-^{(l)} \otimes \varphi^{(l)}$ where $\psi_-^{(l)} \equiv (c_l\psi_1^{(l)}, -\psi_2^{(l)})^T$ is also an eigenvector with eigenvalue $-\sigma^{(l)}$.
As span$\{\psi^{(l)}, \psi_-^{(l)}\}$ = span$\{(\psi_1^{(l)}, 0)^T, (0, \psi_2^{(l)})^T\}$ the functions $\psi^{(l)} \otimes \varphi^{(l)}$ and $\psi_-^{(l)} \otimes \varphi^{(l)}$ are linearly independent and

$$\text{span}\{\psi^{(1)} \otimes \varphi^{(1)}, \psi_-^{(1)} \otimes \varphi^{(1)}, \ldots, \psi^{(n_y)} \otimes \varphi^{(n_y)}, \psi_-^{(n_y)} \otimes \varphi^{(n_y)}\} =$$
$$\text{span}\{ \begin{pmatrix} \psi_1^{(1)} \otimes \varphi^{(1)} \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ \psi_2^{(1)} \otimes \varphi^{(1)} \end{pmatrix}, \ldots, \begin{pmatrix} \psi_1^{(n_y)} \otimes \varphi^{(n_y)} \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ \psi_2^{(n_y)} \otimes \varphi^{(n_y)} \end{pmatrix} \}.$$

From this it follows that the total number of linear independently eigenfunctions of the form $\psi^{(l)} \otimes \varphi^{(l)}$ is equal to $2\, n_y$. Note that our approach with tensorproduct functions leads to the required result: once we know the n_y functions $\varphi^{(1)}, \ldots, \varphi^{(n_y)}$, we can, up to scalars, construct all eigenvectors of $\mathrm{M}_C^{-1}\mathrm{N}$ that correspond to the case (ii) in §2.5, i.e. the eigenvectors with, in general, nonzero eigenvalues.[2]

Apparently, the problem of finding the two times n_y nontrivial eigensolutions of $\mathrm{M}_C^{-1}\mathrm{N}$ breaks up into n_y 'one'-dimensional problems. For each l, the matrix $\mathrm{M}_C^{-1}\mathrm{N}$ has two eigenvectors $\sigma^{(l)}$ and $-\sigma^{(l)}$ with components that, on domain p, correspond to a scalar multiple of $\psi_p^{(l)} \otimes \varphi^{(l)}$ $(p = 1, 2)$.

Errors will be transferred in the iterative solution process of (7) from one subdomain to the other. These errors can be decomposed in eigenvectors of $\mathrm{M}_C^{-1}\mathrm{N}$, that is, they can be expressed on subdomain $p\,(p = 1, 2)$ as linear combination of the functions $\psi_p^{(l)} \otimes \varphi^{(l)}$. The component of the error on domain p in the direction of $\psi_p^{(l)} \otimes \varphi^{(l)}$ is transferred in each step of the iteration process precisely to the component in the direction of $\psi_{3-p}^{(l)} \otimes \varphi^{(l)}$ on domain $3 - p$. In case of the block Jacobi method, transference damps this component by a factor $|\sigma^{(l)}|$.

Here, as in the case of one spatial dimension (§4.3.1), the size of the eigenvalues $\sigma^{(l)}$ is determined by the growth factor $\mu_\ell^{(l)}$ of $\psi_1^{(l)}$ and $\mu_r^{(l)}$ of $\psi_2^{(l)}$ in (46).
In case of dominated behavior, these factors can adequately be estimated by the dominating root of the appropriate characteristic equation (cf. (34)). The scalars, that is, the matrices $C_{\ell r}$ and $C_{r\ell}$ can be tuned to minimize the $|\sigma^{(l)}|$. This will be the subject of our next section.
As we explained in §4.3.1, we see no practical way to tune our coefficients in case of harmonic behavior. However, in our applications the number of eigenvalues that can not be controlled is limited as we will see in our next subsection. Except for a few eigenvalues, the eigenvalues of the error reduction matrix $\mathrm{M}_C^{-1}\mathrm{N}$ will be small in absolute value: the eigenvalues cluster around 0. If θ is equal to an eigenvalue λ of $\mathbf{A}$, then 1 is an eigenvalue of $\mathrm{M}_C^{-1}\mathrm{N}$ (see (v) in §2.5 and §3.2) and $\mathrm{M}_C^{-1}\mathbf{B}_C$ is singular. However, the projections that have been discussed in §3.2, will remove this singularity. An accurate approximation θ of λ (a desirable situation) corresponds to a near singular matrix $\mathrm{M}_C^{-1}\mathbf{B}_C$, and here, the projection will also improve the conditioning of the matrix.

4.4 Optimizing the coupling

In this section, we will discuss the construction of a coupling matrix C that leads to a clustering of eigenvalues $\sigma^{(l)}$ of $\mathrm{M}_C^{-1}\mathrm{N}$ round 0. We give details for the Laplace operator. We will concentrate on the error modes $\psi_p^{(l)} \otimes \varphi^{(l)}$ on domain p with dominated growth in the x direction, that is, modes for which $\psi_p^{(l)}$ exhibits the dominated behavior as described in (iii) of §4.3.1. For these modes and for C as in (18) and (41), we have that (cf., (36) and (46))

$$|\sigma^{(l)}|^2 \approx \left|\frac{\widetilde{\alpha}_\ell^{(l)} + \widetilde{\zeta}^{(l)}}{1 + \widetilde{\alpha}_\ell^{(l)}\widetilde{\zeta}^{(l)}}\right| \left|\frac{\widetilde{\alpha}_r^{(l)} + \widetilde{\zeta}^{(l)}}{1 + \widetilde{\alpha}_r^{(l)}\widetilde{\zeta}^{(l)}}\right|. \tag{47}$$

Here, for $\nu \equiv (2a + uh_x)/(2a - uh_x)$, the quantities $\widetilde{\alpha}_\ell^{(l)}$, $\widetilde{\alpha}_r^{(l)}$ and $\widetilde{\zeta}^{(l)}$ are defined as in (36): $\widetilde{\alpha}_\ell^{(l)} \equiv \alpha_\ell^{(l)}/\sqrt{\nu}$, $\widetilde{\alpha}_r^{(l)} \equiv \sqrt{\nu}\,\alpha_r^{(l)}$, $\widetilde{\zeta}^{(l)} \equiv \sqrt{\nu}\,\zeta_+^{(l)}$, where here $\zeta_+^{(l)}$ is the dominant root of (34) for $\lambda' = \lambda^{(l)}$. Note that, in view of the symmetry in the expression for $|\sigma^{(l)}|^2$, it suffices to study a C for which $\widetilde{\alpha}_\ell^{(l)} = \widetilde{\alpha}_r^{(l)}$.

Let E be the set of l's in $\{1, \ldots, n_y\}$ for which the $\psi_p^{(l)}$ exhibit dominated growth, or, equivalently, for which the characteristic equation associated with the operator $L_x + \lambda^{(l)}$ in (40) (cf., (34)) has a dominant

[2] For $\alpha_\ell^{(l)} \to -\mu_r^{(l)}$ or $\alpha_r^{(l)} \to -\mu_\ell^{(l)}$ one of the nonzero eigenvalues degenerates to a defective zero eigenvalue. But then still this construction yields all nonzero eigenvalues. To avoid a technical discussion we give no details here.

root $\zeta_+^{(l)}$: $E \equiv \{l = 1, \ldots, n_y \mid |\zeta_+^{(l)}| > |\zeta_-^{(l)}|\}$. We are interested in $\alpha^{(l)} \equiv \widetilde{\alpha}_\ell^{(l)} = \widetilde{\alpha}_r^{(l)}$ for which

$$\sigma_{\text{opt}} \equiv \max\left\{ \left| \frac{\alpha^{(l)} + \zeta}{1 + \alpha^{(l)}\zeta} \right| \,\middle|\, \zeta \in \widehat{E} \right\} \quad \text{with} \quad \widehat{E} \equiv \{\sqrt{\nu}\,\zeta_+^{(l)} \mid l \in E\} \tag{48}$$

is 'as small as possible'.

Simple coupling. For the choice $C_{\ell r} = \sqrt{\nu}\,\alpha I$ and $C_{r\ell} = (\alpha/\sqrt{\nu})I$, we can easily analyze the situation.
Then $\alpha^{(l)} = \alpha$ for all l and we should find the $\alpha = \alpha_{\text{opt}}$ that minimizes $\max |(\alpha + \zeta)/(1 + \alpha\zeta)|$. We assume that $|uh_x| < 2a$. Note that then $\sqrt{\nu}$ times the dominant characteristic roots are real and > 1. Therefore, the two extremal values

$$\mu \equiv \min \widehat{E} \quad \text{and} \quad M \equiv \max \widehat{E} \tag{49}$$

determine the size of the maximum. This leads to

$$-\alpha_{\text{opt}} = 1 + \frac{\sqrt{(\mu^2 - 1)(M^2 - 1)}}{\mu + M} + \frac{(\mu - 1)(M - 1)}{\mu + M} > 1 \tag{50}$$

and

$$\sigma_{\text{opt}} = \frac{\sqrt{M^2 - 1} - \sqrt{\mu^2 - 1}}{M\sqrt{\mu^2 - 1} + \mu\sqrt{M^2 - 1}} > 0. \tag{51}$$

Laplace operator. To get a feeling for what we can expect, we interpret and discuss the results for the Laplace operator, that is, we now take $u = v = c = 0$. Further, we concentrate on the computation of (one of) the largest eigenvalue of L and we assume that θ is close to the target eigenvalue. Then

$$\lambda^{(l)} = -\frac{2b}{h_y^2}\left(1 - \cos\left(\pi\frac{l}{n_y + 1}\right)\right) - \theta. \tag{52}$$

First we derive a lower bound for μ and an upper bound for M.
For $D^{(l)} \equiv 1 - \frac{h_x^2}{2a}\lambda^{(l)}$ (cf., (34)), we have that $|D^{(l)}| > 1$, or, equivalently, $|\zeta_+^{(l)}| > |\zeta_-^{(l)}|$, if and only if $\lambda^{(l)} < 0$. Hence $l_e \equiv \min E$ is the smallest integer l for which $\lambda^{(l)} < 0$ and

$$l_e \equiv \lfloor \widetilde{l}_e \rfloor + 1 \quad \text{where} \quad \widetilde{l}_e \equiv \frac{2}{\pi}(n_y + 1)\arcsin\left(\frac{h_y}{2}\sqrt{\frac{-\theta}{b}}\right).$$

(The noninteger value $l = \widetilde{l}_e$ is the 'solution' of $\lambda^{(l)} = 0$.) For $h_y \ll 1$, $\widetilde{l}_e \approx \dfrac{\omega_y}{\pi}\sqrt{\dfrac{-\theta}{b}}$.

For an impression on the error reduction that can be achieved with a suitable coupling, we are interested in lower bounds for $\mu - 1$ that are as large as possible. With $\delta \equiv D^{(l_e)} - 1$ we have that $\mu - 1 = \delta + \sqrt{2\delta + \delta^2} \geq \sqrt{2\delta}$. Therefore, we are interested in positive lower bounds for δ:

$$\begin{aligned}
\delta &= \rho^2\left(\cos\left(\pi\frac{\widetilde{l}_e}{n_y + 1}\right) - \cos\left(\pi\frac{l_e}{n_y + 1}\right)\right) \geq \pi\rho^2\frac{l_e - \widetilde{l}_e}{n_y + 1}\sin\left(\pi\frac{\widetilde{l}_e}{n_y + 1}\right) \\
&\geq 2\pi\frac{b}{a}\widetilde{l}_e(l_e - \widetilde{l}_e)\left(\frac{h_x}{\omega_y}\right)^2 \qquad \text{where} \quad \rho \equiv \frac{h_x}{h_y}\sqrt{\frac{b}{a}}.
\end{aligned}$$

The bound for δ depends on the distance of $\widetilde{l}_e$ to the integers, which can be arbitrarily small. This means that, even for the optimal coupling parameters, the (absolute value of the) eigenvalue $\sigma^{(l_e)}$ can be arbitrarily close to one. Since, for optimal coupling, the damping that we achieve for the smallest l in E is the same as for the largest, it seems to be undesirable to concentrate on damping the error modes associated with l_e as much as possible. Therefore, we remove l_e from the set E and concentrate on damping the error modes associated with l in $E' \equiv E\backslash\{l_e\}$. For the δ and μ associated with this slightly reduced set E' we have that

$$\mu - 1 \geq \sqrt{2\delta} \geq 2\kappa h_x \quad \text{where} \quad \kappa \equiv \frac{1}{\omega_y}\sqrt{\pi l_e \frac{b}{a}}. \tag{53}$$

The lower bound for $\mu - 1$ is sharp for $h \to 0$ with ρ fixed, i.e., for given ρ, $h = (h_x, h_y)$ is such that $h_x = h_y \rho \sqrt{a/b}$.

An upper bound for M follows from the observations that $\theta < 0$ and that the cosine takes values between -1 and 1: we have that $D^{(l)} \leq 1 + 2\rho^2$ and

$$M - 1 \leq 2\rho^2 + \sqrt{4\rho^2 + 4\rho^4}.$$

Put

$$M' \equiv \sqrt{\frac{M-1}{M+1}} \leq \sqrt[4]{\frac{\rho^2}{1+\rho^2}}.$$

Then, for $h \to (0,0)$ such that ρ is fixed, we have that

$$-\alpha_{\text{opt}} = 1 + 2M'\sqrt{\kappa h_x} + \mathcal{O}(h_x) \quad \text{and} \quad 1 - \sigma_{\text{opt}} = 2\frac{\sqrt{\kappa h_x}}{M'} + \mathcal{O}(h_x).$$

Here we used that

$$-\alpha_{\text{opt}} = 1 + \sqrt{2(\mu-1)}M' + \mathcal{O}(\mu - 1) \quad \text{and} \quad 1 - \sigma_{\text{opt}} = \sqrt{2(\mu-1)}/M' + \mathcal{O}(\mu - 1)$$

for $\mu \to 1$ (see (50) and (51)).

So, for small stepsizes h, the 'best' 'asymptotic error reduction factor' σ_{opt} is less than one with a difference from one that is proportional to the square root of h_y.

We tried to cluster the eigenvalues of $\mathbf{M}_C^{-1}\mathbf{B}$ around one as much as possible. With $\alpha = \alpha_{\text{opt}}$, at most l_e eigenvalues may be located outside the disk with radius σ_{opt} and center one. After an initial l_e steps we may expect the convergence of GMRES to be determinated by σ_{opt} (provided that the basis of eigenvectors is not too skew). Therefore, as long as l_e is a modest integer, we expect GMRES to converge well in this situation. We will now argue that, in realistic situations, l_e will be modest as compared to the index of the eigenvalue of $\mathbf{A}$ where we are interested in. For clearness of arguments, we assume the stepsizes to be small: $h \to (0,0)$ with ρ fixed: $\lambda^{(l_e)} \approx -b\pi^2(l_e/\omega_y)^2 - \theta$.

Suppose that, for some $\tau > 0$, we are interested in the smallest eigenvalue λ of $\mathbf{A}$ that is larger than $-\tau$. Since, in the Jacobi-Davidson process, θ converges to λ, θ will eventually be larger than $-\tau$. We concentrate on this 'asymptotic' situation.[3]

[3] The Jacobi-Davidson process can often be started in practice with an approximate eigenvector that is already close to the wanted eigenvector. Then θ will be close to λ. For instance, if one is interested in a number of eigenvalues close to some target value, then the search for the second and following eigenvectors will be started with a search subspace that has been constructed for the first eigenvector. This search subspace will be 'rich' with components in the direction of the eigenvectors that are wanted next (see [8, §3.4]).

Then, $l_e \leq C_1(\tau') + 1$, where

$$C_1(\tau') \equiv \#\{l \in \mathbb{N} \mid l^2 \leq \tau'\} \quad \text{and} \quad \tau' \equiv \tau \frac{\omega_y^2}{b\pi^2}.$$

The number of eigenvalues $\lambda^{(m_x,m_y)} \approx -a\pi^2(m_x/\omega_x)^2 - b\pi^2(m_y/\omega_y)^2$ of $\mathbf{A}$ that are larger than $-\tau$ is approximately equal to

$$C_2(\tau') \equiv \#\{(m_x, m_y) \in \mathbb{N}^2 \mid \; m_y^2 + \frac{a}{b}\frac{\omega_y^2}{\omega_x^2} m_x^2 \leq \tau'\}.$$

Since $C_1(\tau')^2 \lesssim 2\frac{\omega_y}{\omega_x}\sqrt{\frac{a}{b}}\, C_2(\tau')$, the number $l_e + 1$ of error modes that we do not try to control with appropriate coupling coefficients is proportional to the *square root* of the index number of the wanted eigenvalue (if the eigenvalues have been increasingly ordered). For instance, if $a = b$, $\omega_x = \omega_y$, and $\tau' = 15$, then eight eigenvalues of $\mathbf{A}$ are larger than $-\tau$, and we do not 'control' four modes. One of these modes corresponds with the wanted eigenvalue and is 'controlled' by the projections in the correction equation of the Jacobi-Davidson process.

In practice, deflation will be used for the computation of the, say, eight eigenvalue of $\mathbf{A}$. The first seven eigenvalues will be computed first and will be deflated from $\mathbf{A}$. In such an approach, the three modes that we did not try to control in our coupling, will be controlled by the projection on the space orthogonal to the detected eigenvectors. See §5.2.2 for a numerical example.

We analyzed the situation where the domain has been decomposed into two subdomains. Of course, in practice, we will interested in a decomposition of more subdomains. In these situations, the number of modes that we did not try to control by the coupling, will be proportional to the number of artificial boundaries. For numerical results, see §5.4. Deflation will be more important if the number subdomains is larger. Note that the observations in the §§4.3.1 and 4.3.2 on the error modes that exhibit dominated behavior also apply to the situation of more than two subdomains: the essential observation in case of dominated growth is that, on one subdomain, the influence of the 'dominated' component (as represented by $\zeta_-^{(l)}$) is negligible at the artificial boundary regardless the boundary condition at the other end of the subdomain.

Stronger couplings. In §4.3.2, we considered coupling matrices C with eigenvectors related to ones of L_y, the y-component of the finite difference operator L. Instances of such matrices can easily be formed by using L_y itself.

For ease of notation we consider the Laplace operator. Inclusion of first order terms only results in extra factors ν (cf. (36) in §4.3.1). Consider the matrices

$$C_{\ell\ell} = C_{rr} = 1 + \gamma L_y \quad \text{and} \quad C_{\ell r} = C_{r\ell} = \alpha + \beta L_y, \tag{54}$$

where α, β, and γ are appropriate scalars. With β and γ, we introduce interaction parallel to the interface in the coupling. Then $\alpha_\ell^{(l)}$ in (41) is equal to

$$\alpha_\ell^{(l)} = q_\ell(\lambda^{(l)}) \quad \text{where} \quad q_\ell(\lambda) \equiv \frac{\alpha + \beta\lambda}{1 + \gamma\lambda}. \tag{55}$$

Note that the dominant root $\zeta_+^{(l)}$ (cf. (34) with $\lambda' = \lambda^{(l)}$) depends on $\lambda^{(l)}$: $\zeta_+^{(l)} = w_\ell(\lambda^{(l)})$ for some function w_ℓ. Hence, we are interested in finding scalars α, β, and γ for which

$$\sigma'_{\text{opt}} \equiv \max_\lambda \left| \frac{q_\ell(\lambda) + w_\ell(\lambda)}{1 + q_\ell(\lambda) w_\ell(\lambda)} \right| \tag{56}$$

TABLE 1. *The table shows the values that can be achieved for the damping σ'_{opt} in* (56) *for the Laplace equation on the the unit square by optimizing the coupling in* (54) *with respect to some of the parameters α, β and γ. For explanation see the example in §4.4.*

	1	2	3	4
optimized w.r.t.	α	α, β	α, γ	α, β, γ
σ'_{opt}	0.696	0.157	0.376	0.093

is as small as possible. Here λ ranges over the set of eigenvalues $\lambda^{(l)}$ of L_y that lead to a dominant root $\zeta_+^{(l)} = w_\ell(\lambda^{(l)})$. For $\beta = \gamma = 0$ we have the 'simple coupling' as discussed above. For the coupling at the right side of the artificial boundary, we have similar expressions. Finding the minimum of (56) is a non-linear problem (in α, β and γ; q_ℓ is rational and q_ℓ is in the denominator) and can not analytically be solved. But a numerical solution can be obtained with, for instance, a modified Rémès algorithm. We discuss our results for a simple example in order to illustrate how much can be gained by including interactions parallel to the artificial boundary in the coupling.

Example. Table 1 shows values for σ'_{opt} for the Laplace operator on the unit square ($a = b = 1$, $u = v = c = 0$, $\Omega = (0,1) \times (0,1)$), with $\theta = -34\pi^2$ (then $l_e = 6$ and 24 eigenvalues are larger than θ), $n_x = 180$, $n_y = 120$ and $\omega_{x1} = \frac{1}{3}$. In case 1 in the table, we took $\beta = \gamma = 0$ and we optimized with respect to α. This case corresponds to the 'simple coupling' as discussed above. We learn from column 2 of Table 1 that an additional parameter β allows a considerable reduction of the damping factor.

With $\beta = \gamma = 0$ the explicit coupling is in the x direction only, this corresponds to a two point stencil for the boundary conditions on the artificial boundary. The parameter β introduces a coupling in the y directions which corresponds to a four point stencil for the artificial boundary conditions. If in addition $\gamma \neq 0$, the coupling corresponds to a six point stencil. Extension from a two to a four point stencil appears to be more effective than the extension from a four to a six point stencil (a reduction of σ'_{opt} from 0.696 to 0.157 as compared to a reduction from 0.157 to 0.093 in Table 1). The parameter $\beta \neq 0$ gives a coupling of the internal boundary conditions on the artificial interface (the $\circ$'s in Fig. 1), while γ gives a coupling of the internal boundary conditions on points of the original domain (the $\bullet$'s in Fig. 1 closest to the cut). Note that an optimal β (with $\gamma = 0$) gives better values than an optimal γ (with $\beta = 0$).

Experimentally we verified that the values for σ'_{opt} obtained with a 'local mode analysis' (where we neglected ζ_- terms) correspond rather well with the actual radius of the cluster of eigenvalues of $\mathrm{M}_C^{-1}\mathrm{N}$: except for the first l_e+1 eigenvalues, in all cases all eigenvalues of $\mathrm{M}_C^{-1}\mathrm{N}$ are in the disc with center 0 and radius σ'_{opt}. Since we did not optimize for the first l_e eigenvalues, it is no surprise that these eigenvalues are not in the disc. The $l_e + 1$th eigenvalue corresponds to the situation where $|\zeta_+^{(l)}|$ is closest to $|\zeta_-^{(l)}|$ and then the predictions of the local mode analysis may expected to be the least reliable. For an experiment with larger stepsize see §5.2.3.

5 Numerical experiments

The experiments presented in this section illustrate the numerical behavior of the Jacobi-Davidson method in combination with the domain decomposition method, as described in §3 and §4. We will focus on some characteristic properties. All experiments are performed with MATLAB 5.3.0 on a Sun Sparc Ultra 5 work-station.

In §5.1 we will discuss the circumstances under which experiments have been performed. Because Jacobi-Davidson is a nested iterative method, an inexact solution of the correction equation affects the

TABLE 2. *Convergence of Jacobi-Davidson, with accurate solution of the correction equation, towards the eigenvalue of smallest absolute value (=largest eigenmode) of the discretized ($n = 99$, $h = 0.01$) eigenvalue problem for the one-dimensional Laplace operator.*

step	selected Ritz value	residual selected Ritz pair	number of correct digits selected Ritz value
1	-3992.4322622	9.74e+03	-3.6
2	-1487.8343933	3.99e+03	-3.2
3	-581.73159839	1.62e+03	-2.8
4	-283.84104294	7.22e+02	-2.4
5	-123.01979659	3.23e+02	-2.1
6	-42.762088608	1.15e+02	-1.5
7	-17.253205686	4.49e+01	-0.87
8	-9.8982441731	7.41e+00	1.5
9	-9.8687926855	5.15e-04	9.8
10	-9.8687926854	6.26e-12	12

outerloop. Therefore, we will also check how the exact process behaves and which stage of the process is most sensitive to inexact solution.

Then, in §5.2, we consider the spectrum of the error propagator for the asymptotic situation $\theta = \lambda$. This spectrum contains all information for understanding the convergence behavior of the Jacobi iteration method. The predictions of §4.4 on the optimized coupling are verified and we investigate the effect of deflation.

The next question is how the Jacobi-Davidson method behaves when inexact solutions for the correction equation are obtained with Jacobi iterations. In §5.3 we compare different types of coupling, and left and right preconditioning. Furthermore, we consider GMRES as an accelerator of the Jacobi iterative method.

We conclude, in §5.4, with an experiment that shows what happens when we have more than two subdomains.

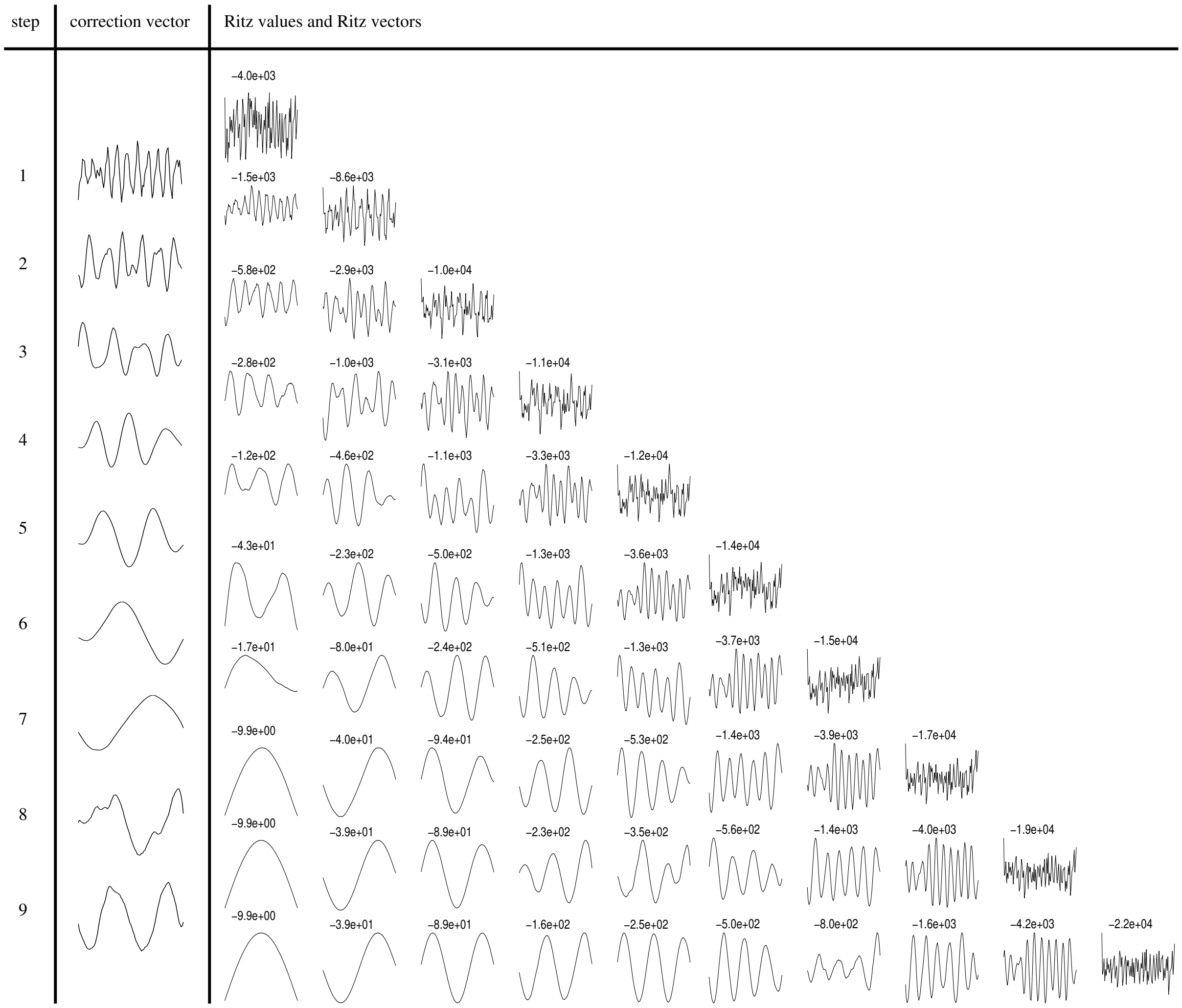
5.1 Reference process

We first consider the standard Jacobi-Davidson method, when applied to the discretized eigenvalue problem for the Laplace operator. No domain is decomposed and correction vectors are obtained by accurate solution of the correction equation.

The first experiment gives a global impression of the speed of convergence. For that purpose we confine ourselves to the one-dimensional case, described in §4.3.1. We take $n = 99$, $h = 0.01$. For the starting vector of the Jacobi-Davidson process, we take a random vector generated in MATLAB (with seed equal to 226). We want to compute the eigenvalue of smallest absolute value ($\lambda_1 = -\left(200 \sin \frac{\pi}{200}\right)^2 = -9.86879268536\ldots$). The corresponding eigenvector describes the largest eigenmode of the discretized PDE.

Table 2 and Fig. 2 show what happens in the iteration process. The second column of Table 2 gives the selected Ritz value θ for the correction equation, the third column gives the 2-norm of the residual $\mathbf{r} \equiv \mathbf{A}\mathbf{u} - \theta\mathbf{u}$ of the corresponding Ritz pair $(\theta, \mathbf{u})$, and the fourth column lists the number of correct

FIGURE 2. *Convergence behavior of Jacobi-Davidson with accurate solution of the correction equation, when applied to the discretized ($n = 99$, $h = 0.01$) eigenvalue problem for the one-dimensional Laplace operator. The process is started with one random vector. In each step a correction vector is computed (second column) by which the search subspace is expanded. In the third column all Ritz values of the search subspaces before/after expansion are printed. Right below this number the corresponding Ritz vector is graphically displayed.*



digits of the Ritz value: $-\log^{10}|\lambda - \theta|$.

From Table 2 we observe that Jacobi-Davidson needs about 8 steps before the (theoretically cubic) convergence to the desired eigenvalue sets in. This might have been expected: as the startvector is random it is likely that the components of all eigenmodes are about equally represented in the startvector. Therefore, in the beginning the eigenvalues with larger absolute value will dominate for a while. In Fig. 2 we display the Ritz vectors after each iteration of the Jacobi-Davidson process. The corresponding eigenmodes are of high frequency, which explains the order of appearance of Ritz vectors (high frequencies dominate initially).

A proper target value in the correction equation (21), instead of the Ritz value, may help to overcome the initial phase of slow convergence, but this is beyond the scope of this paper. Our concern is the question how much the process is affected when the correction equation is solved approximately by performing accurate solves on the subdomains only and by tuning the interface conditions. A less accurate solution of the correction equation will, in general, result in more steps of Jacobi-Davidson (outer iterations) for the same precision for the approximate eigenpair. In particular, we do not want to extend the 'slow phase' by destroying the 'fast phase' with too inaccurate solution steps. We take the 'exact' Jacobi-Davidson process in Table 2 as our reference. In order to see what happens in the final, potentially fast phase, we select a parabola shaped startvector.

In the next subsections we will mainly consider the more interesting two-dimensional case, with physical sizes $\omega_x = 2$ and $\omega_y = 1$. The number of grid points in x- and y direction are $n_x = 63$ and $n_y = 31$, so $h_x : h_y = 1 : 1$. The eigenvalue corresponding to the largest eigenmode of the discretized Laplace operator is equal to $-12.328585\ldots$. In Table 3 the convergence history for Jacobi-Davidson to this eigenpair is presented when starting with the parabolic vector

$$\{(\frac{j_x}{n_x+1}(1-\frac{j_x}{n_x+1}), \frac{j_y}{n_y+1}(1-\frac{j_y}{n_y+1})) \mid 1 \le j_x \le n_x, 1 \le j_y \le n_y\}, \tag{57}$$

and with accurate solutions of the correction equation. The second column of this table shows the selected

TABLE 3. *Convergence history of Jacobi-Davidson applied to the discretized eigenvalue problem of the two-dimensional Laplace operator ($n_x = 63, n_y = 31, \omega_x = 2$ and $\omega_y = 1$) with accurate solutions of the correction equation.*

step	θ	$\theta - \lambda$	$\|\|\mathbf{r}\|\|_2$	$\|\|\mathbf{r}'\|\|_2$
1	-12.4896	-1.61e-01	4.19e+00	4.19e+00
2	-12.3286	-9.65e-07	8.55e-03	6.10e-03
3	-12.3286	-1.55e-13	1.76e-10	1.19e-10
4	-12.3286	-1.33e-13	7.71e-14	3.90e-14

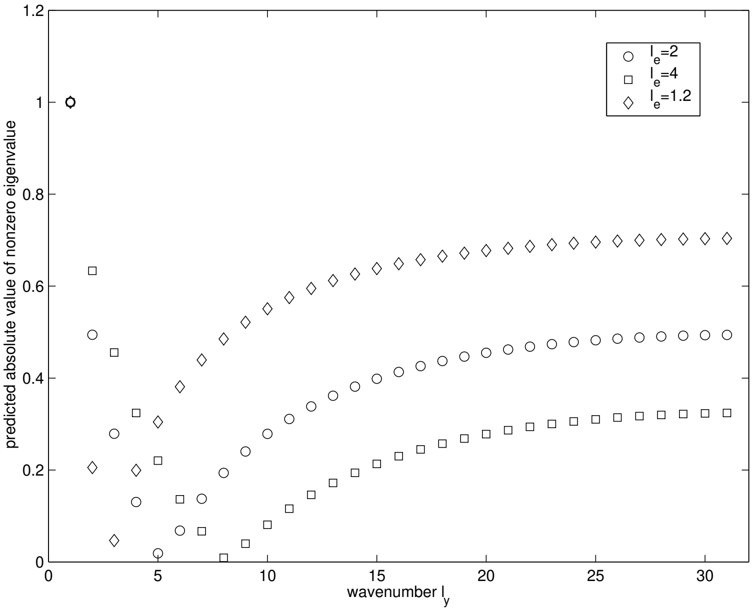
Ritz value for the correction equation, the third column the error $\theta - \lambda$ for this Ritz value, and the fourth column gives the 2-norm of the residual $\mathbf{r}$ for the corresponding normalized Ritz pair. Jia and Stewart [11] have pointed out that for θ, and given the information in the subspace $\mathcal{V}$, a better, in residual sense, approximate eigenvector can be computed; the norm of the residual of this so-called refined Ritz vector is given by the quantity

$$||\mathbf{r}'||_2 = \min_{\mathbf{u}\in\mathcal{V}} ||\mathbf{A}\mathbf{u} - \theta\mathbf{u}||,$$

represented in the fifth column in Table 3.

These experiments set the stage for the domain decomposition experiments.

FIGURE 3. *Predicted amplification of the error propagator* $\mathrm{M}_C^{-1}\mathrm{N}$ *with simple optimized coupling for the largest eigenvalue* $\lambda^{(1,1)}$ *of the Laplace operator for* $l_e = 2$, $l_e = 4$, *and* $l_e = 1.2$. *For explanation, see* §*5.2.1.*



5.2 Spectrum of the error propagator

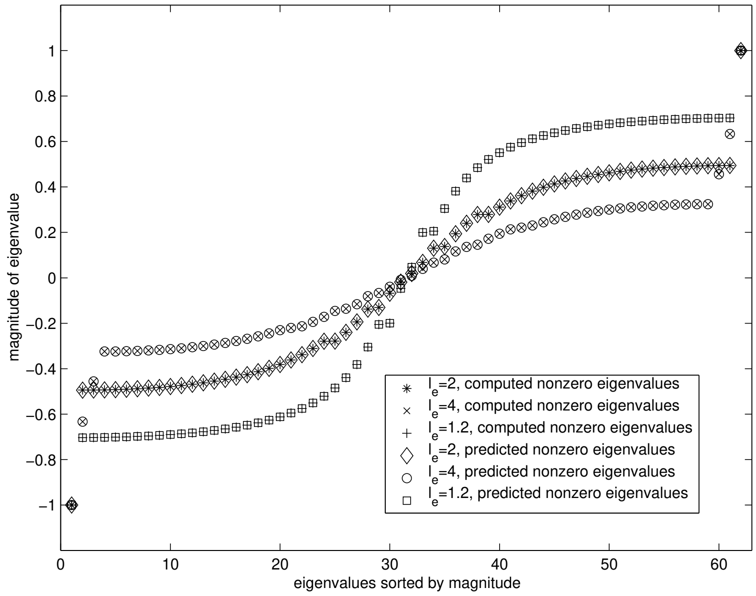
From §2.5 we know that the convergence properties of the Jacobi iterative method depend on the spectrum of the error propagator $\mathrm{M}_C^{-1}\mathrm{N}$. Therefore, we will investigate these spectra for some typical situations. We consider the asymptotic case $\theta = \lambda$. Although θ approximates λ in practice, during the iteration process θ becomes very close to λ, and that is the reason we think that the asymptotic case gives a good indication.

5.2.1 Predicted and computed spectra

First we consider the determination of the parameter α_{opt} (50) for the simple optimized coupling. The value of α_{opt} depends on the extremal values μ and M of the collection of dominant roots $\widehat{E}$ (48) for which α_{opt} is optimized. The value μ depends amongst others on θ, and M only depends on h_x, h_y, and on the coefficients a and b.

We illustrate the sensitivity of α_{opt} w.r.t. the lower bound μ, for θ equal to the largest eigenvalue $\lambda^{(1,1)}$ of the Laplace operator, with $\omega_x = 2, \omega_y = 1, n_x = 63, n_y = 31$ and $n_{x1} = 26$. For a dominant root $\zeta_+^{(l)}$, $\lambda^{(l)}$ in (52) should be smaller than 0. Then $\dfrac{4b}{h_y^2}\sin^2\left(\dfrac{\pi}{2}\dfrac{l_e}{n_y+1}\right) > \theta$. Since $\theta \approx \dfrac{5}{4}\pi^2$ and $\dfrac{4b}{h_y^2}\sin^2\left(\dfrac{\pi}{2}\dfrac{l_e}{n_y+1}\right) \approx l_e^2\pi^2$, we have approximately that $l_e^2 > \dfrac{5}{4}$. The smallest such integer l_e is $l_e = 2$. In order to show that this is a sharp value for l_e and thus a sharp lower bound for the μ (53), we shall compare the case $l_e = 2$ with the case for the smaller value $l_e = 1.2$. We also included the case $l_e = 4$, where apart from the mode $l_y = 1$, the modes $l_y = 2$ and $l_y = 3$ are excluded from the optimization process (i.e. for the computation of an optimal α).

FIGURE 4. *Predicted and computed nonzero eigenvalues of the error propagator* $\mathrm{M}_C^{-1}\mathrm{N}$ *with simple optimized coupling for the largest eigenvalue* $\lambda^{(1,1)}$ *of the Laplace operator for* $l_e = 2$, $l_e = 4$, *and* $l_e = 1.2$. *For explanation, see §5.2.1.*



For these three cases ($l_e = 2, l_e = 4$, and $l_e = 1.2$) we have computed the corresponding α ($\alpha = -1.6287\ldots$, $\alpha = -2.1279\ldots$, and $\alpha = -1.2800\ldots$, respectively). In Fig. 3 the predicted amplification of the error propagator $\mathrm{M}_C^{-1}\mathrm{N}$ for these values of α are shown. Here we calculated for each mode (with wavenumber l_y) the expected amplification $|\sigma^{(l_y)}|$ with expression (46). Indeed, we see that (for $l_e = 2$) the second leftmost circle ($l_y = 2$) in Fig. 3 represents the same value as for the rightmost circle ($l_y = 31$), which was our goal. If l_e is close to 1, then because the mode $l_y = 1$ can not be damped at all, the overall damping for $l_e = 1.2$ is predicted to be less, whereas $l_e = 4$ should lead to a better damping of the remaining modes $l_y = 4, \ldots, 31$ that are taken into account, which is confirmed in Fig. 3 for different values of α.

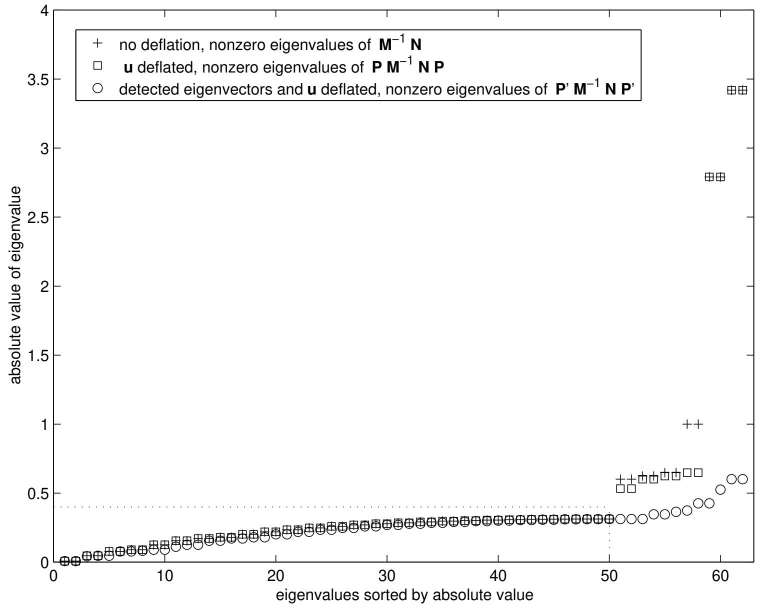
Fig. 4 shows the *exact* nonzero eigenvalues σ of $\mathrm{M}_C^{-1}\mathrm{N}$ sorted by magnitude for different values of α. We also plotted in this figure the *predicted* nonzero eigenvalues sorted by magnitude. We see that the predictions are very accurate.

In Fig. 4 we see also the effect of the value l_e on the eigenvalues. Again, we see that it is better to overestimate l_e than underestimate. The point symmetry in Fig. 4 is due to the fact that if σ is an eigenvalue of $\mathrm{M}_C^{-1}\mathrm{N}$ then $-\sigma$ is also an eigenvalue (remark (iv) of §2.5). Furthermore, note that for each process one eigenvalue is equal to 1, independent of α. By a combination of remark (v) of §2.5 and the discussion at the end of §3.2, we see that the corresponding eigenvector is of the form $\underline{\mathbf{y}}$ that corresponds to the eigenvector $\mathbf{y}$ that we are looking for with our Jacobi-Davidson process. Hence the occurrence of 1 in the spectrum is not a problem: the projections in the correction equation take care of this, as we will show now.

5.2.2 Deflation

Now we show, by means of an example, how deflation improves the condition of the preconditioned correction equation (26). For the discretized Laplace operator we take $\omega_x = \omega_y = 1$, $n_x = n_y = 31$, $n_{x1} =$

FIGURE 5. *The effect of deflation on the nonzero eigenvalues of the error propagator with simple optimized coupling. For explanation, see §5.2.2. The dotted lines indicate the area of Fig. 6.*



15 and $\theta = \lambda^{(4,4)}$. There are 19 eigenvalues larger than $\lambda^{(4,4)}$. If we determine the α_{opt} for the simple optimized coupling, then $\widetilde{l}_e \approx 5.6944$. So the modes $l_y = 1, \ldots, 6$ are not taken into account for the optimization of α, since they do not show dominant behavior. Hence we do not necessarily damp these modes with the resulting α_{opt}.

One of them, more precisely the mode $l_y = 4$, is connected to the y-component of the eigenvector $\varphi^{(4,4)}$ corresponding to $\lambda^{(4,4)}$: this mode can not be controlled at all with α because the operator $\mathbf{A}$ is shifted by $\lambda^{(4,4)}$ and therefore singular in the direction of $\varphi^{(4,4)}$. In the correction equation (26) the operator stays well-conditioned due to the projection $\mathbf{P}$ that deflates exactly the direction $\mathbf{u} = \varphi^{(4,4)}$. Since the error propagator originates from the enhanced operator in the correction equation, this projection is actually incorporated in the error propagator (§3.2): $\mathbf{P}\mathrm{M}_C^{-1}\mathrm{N}\mathbf{P}$.

The other non-dominant modes $l_y = 1, 2, 3, 5, 6$, can not be controlled by α_{opt}. But, as remarked in §4.4, in practice one starts the computation with the largest eigenvalues and when arrived at $\lambda^{(4,4)}$, the 19 largest eigenvalues with corresponding eigenvectors are already computed and will be deflated from the operator $\mathbf{B}$. Deflation in the enhanced correction equation is performed by the projection

$$\mathbf{P}' \equiv \mathbf{I} - \mathrm{M}_C^{-1}\underline{\mathbf{X}}\left(\underline{\mathbf{X}}^*\mathrm{M}_C^{-1}\underline{\mathbf{X}}\right)^{-1}\underline{\mathbf{X}}^*.$$

Here $\underline{\mathbf{X}} \equiv (\mathbf{X}_1^T, X_\ell^T, 0^T, 0^T, X_r^T, \mathbf{X}_2^T)^T$, where $\mathbf{X} \equiv (\mathbf{X}_1^T, X_\ell^T, X_r^T, \mathbf{X}_2^T)^T$ is a matrix of which the columns form an orthonormal basis for the space spanned by the 19 already computed eigenvectors and the approximate 20th eigenvector. This implies that we are dealing with the error propagator $\mathbf{P}'\mathrm{M}_C^{-1}\mathrm{N}\mathbf{P}'$.

For α_{opt} we computed the nonzero eigenvalues of $\mathrm{M}_C^{-1}\mathrm{N}$, $\mathbf{P}\mathrm{M}_C^{-1}\mathrm{N}\mathbf{P}$ and $\mathbf{P}'\mathrm{M}_C^{-1}\mathrm{N}\mathbf{P}'$. In Fig. 5 their absolute values are plotted. The '$+$'-s (no deflation) indicate that the most right 12 eigenvalues have not been controlled by α_{opt}. This is in agreement with the fact that the modes $l_y = 1, \ldots, 6$ have not been taken into account for the determination of α_{opt}: to each mode l_y there correspond exactly two eigenvalues

TABLE 4. *Values of coupling parameters and predicted amplification σ'_{opt} for four types of optimized coupling. For explanation, see §5.2.3.*

type no.	1	2	3	4
optimized w.r.t.	α	α, β	α, γ	α, β, γ
α	-2.138	-0.4988	-1.373	-0.2080
β		0.001375		0.001959
γ			0.0002230	-0.0001352
predicted σ'_{opt}	0.3128	0.01875	0.1196	0.007686

$-\sigma^{(l_y)}$ and $+\sigma^{(l_y)}$. Two eigenvalues have absolute value 1 (position 57 and 58 on the horizontal axis). They correspond to the eigenvector $\varphi^{(4,4)}$ of $\mathbf{A}$.
The '$\square$'-s show that deflation with $\mathbf{u}$ makes these absolute values become less than 1. But, with deflation by $\mathbf{u}$, the other uncontrolled eigenvalues stay where they were without deflation; four absolute values are even larger than 2.5. Fortunately, deflation with the 19 already computed eigenvectors drastically reduces these absolute values, as the '$\circ$'-s show.

From this example we learned that deflation may help to cluster the part of the spectrum that we can not control with the coupling parameters, and therefore improves the conditioning of the preconditioned correction equation. The remaining part of the spectrum, that is the eigenvalues that are in control (indicated by the dotted lines in Fig. 5), may be damped even more. This will be subject of the next section.

5.2.3 Stronger coupling

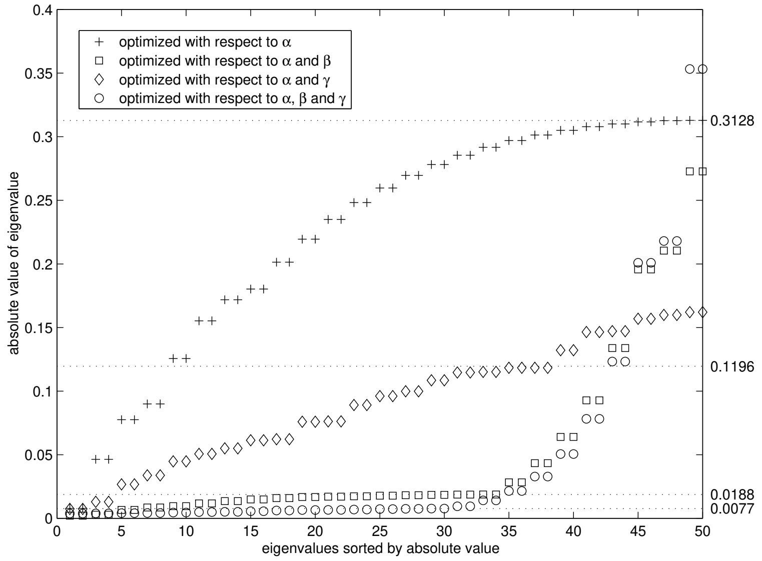
At the end of §4.4, it was illustrated that the inclusion of interactions parallel to the artificial boundary provides more coupling parameters by which a better coupling can be realized. We will apply this now to the example in §5.2.2 in order to investigate how much we can improve the spectrum of the error propagator and how accurate the value of the predicted amplification σ'_{opt} is for the different types of coupling.

Table 4 contains the values of the coupling parameters and the predicted amplification σ'_{opt} for the different types of coupling when $l_e = 7$, as in §5.2.2. These values are obtained by application of a Rémès algorithm to expression (56). As in the final example of §4.4, we see that be the best coupling is predicted to be of type 4, followed by type 2, and then type 3. But, the question remains what the exact spectrum may be for these types op coupling.

We computed the exact nonzero eigenvalues of the error propagator $\mathrm{M}_C^{-1}\mathrm{N}$ for the four types of coupling from Table 4. From §5.2.2, we know that with the coupling parameters we only control the $2n_y - 12 = 50$ nonzero eigenvalues of the error propagator with lowest absolute value. Therefore, we exclude the 12 other nonzero eigenvalues from our further discussion. In Fig. 6 the 50 eigenvalues with lowest absolute value are plotted. The corresponding predicted values of σ'_{opt} are indicated by dotted lines in Fig. 6. From inspection of the eigenvectors, we have verified that for the four different types of coupling, the 12 eigenvalues with highest absolute value that are excluded correspond to the modes $l_y = 1, \ldots, 6$. (Computation of the eigenvectors is rather time consuming. Therefore, we restricted ourselves here to a grid that is coarser than the one in the example at the end of §4.4.)

Indeed, as predicted, it pays off to include more coupling parameters. For type 1 the predicted value of σ'_{opt} is almost exact. The value for type 3 seems to be accurate for the eigenvalues at positions $1, \ldots, 38$. For types 2 and 4, the value becomes less accurate after position 34. We believe that this is because of neglecting the ζ_- terms in the expression for σ'_{opt}: for types 2 and 4 the eigenvectors, that correspond to the

FIGURE 6. *The effect of different types of optimized coupling on the nonzero eigenvalues of the error propagator. The values of the coupling parameters are given in Table 4. The corresponding predicted values of σ'_{opt} are indicated by dotted lines. For explanation, see §5.2.3.*



eigenvalues with position larger than 34, have a low value of l_y. In our quest for optimizing the spectral radius of the error propagator, we have now arrived at a level where we can no longer ignore the contributions of the terms ζ_-. This is confirmed by inspecting the eigenvectors: the eigenvalues that deviate from the predicted σ'_{opt} have eigenvectors that correspond to low values of l_y. But still, the predicted σ'_{opt} gives a good indication for the quality of the coupling and will be better for finer grids.

5.3 Effect on the overall process

In §5.2 spectra of the error propagator have been studied. These spectra provide information on the convergence behavior of the Jacobi iterative method. Now we turn our attention to the overall Jacobi-Davidson method itself. We are interested in how approximate solutions of the correction equation, obtained with a linear solver ('the innerloop'), affects the Jacobi-Davidson process ('the outerloop').

Here we consider two types of coupling:

1. the simple optimized coupling with one coupling parameter α,

2. the Neumann-Dirichlet coupling.

Although we have seen in §4.4 and §5.2.3, that there exist better choices for the coupling, we believe that the overall process with the simple optimized coupling gives a good indication of what we may expect for the stronger optimized couplings. The choice for the Neumann-Dirichlet coupling is motivated by the fact that it is commonly used in domain decomposition methods.

The testproblem will be the same as the one in §5.2.1. First we discuss the Jacobi iterative method as a solver for the correction equation. We do this for both the left and right preconditioned variant. Then we compare the results with those obtained by the GMRES method.

TABLE 5. *Convergence history of Jacobi-Davidson applied to the discretized eigenvalue problem of the two-dimensional Laplace operator for approximate solutions to the correction equation obtained with left (left) and right (right) preconditioned Jacobi iterations on two subdomains and simple optimized coupling. For explanation see §5.3.1.*

optimized coupling, $l_e = 2$

	left DD-preconditioned				right DD-preconditioned			
step	$\theta - \lambda$	$\|\|\mathbf{r}\|\|_2$	$\|\|\mathbf{r}'\|\|_2$	α	$\theta - \lambda$	$\|\|\mathbf{r}\|\|_2$	$\|\|\mathbf{r}'\|\|_2$	α
	3 Jacobi inner iterations				2 Jacobi inner iterations			
1	-1.61e-01	4.19e+00	4.19e+00	-1.6275	-1.61e-01	4.19e+00	4.19e+00	-1.6275
2	-4.98e-03	3.14e+00	2.55e+00	-1.6287	-4.98e-03	3.14e+00	2.55e+00	-1.6287
3	-2.20e-04	1.90e-01	1.81e-01	-1.6287	-2.20e-04	1.90e-01	1.81e-01	-1.6287
4	-1.62e-07	7.12e-03	6.74e-03	-1.6287	-1.62e-07	7.12e-03	6.74e-03	-1.6287
5	-2.13e-12	4.16e-05	3.91e-05	-1.6287	-2.09e-12	4.16e-05	3.91e-05	-1.6287
6	-1.53e-13	1.36e-06	9.37e-07	-1.6287	-1.47e-13	1.36e-06	9.37e-07	-1.6287
7	-1.62e-13	8.43e-09	5.78e-09	-1.6287	-1.81e-13	8.43e-09	5.78e-09	-1.6287
8	-1.39e-13	1.19e-10	8.84e-11		-1.44e-13	1.19e-10	8.84e-11	
	4 Jacobi inner iterations				3 Jacobi inner iterations			
1	-1.61e-01	4.19e+00	4.19e+00	-1.6275	-1.61e-01	4.19e+00	4.19e+00	-1.6275
2	-4.23e-03	2.89e+00	2.43e+00	-1.6287	-4.23e-03	2.89e+00	2.43e+00	-1.6287
3	-2.70e-05	6.42e-02	6.20e-02	-1.6287	-2.70e-05	6.42e-02	6.20e-02	-1.6287
4	-5.95e-09	1.02e-03	7.36e-04	-1.6287	-5.95e-09	1.02e-03	7.36e-04	-1.6287
5	-1.53e-13	2.84e-06	2.61e-06	-1.6287	-1.58e-13	2.84e-06	2.61e-06	-1.6287
6	-1.76e-13	2.81e-08	1.54e-08	-1.6287	-9.95e-14	2.81e-08	1.54e-08	-1.6287
7	-1.44e-13	8.33e-12	8.30e-12		-1.42e-13	8.34e-12	8.28e-12	

5.3.1 The Jacobi iterative process

In §5.2.1 we have computed the spectra of the error propagator $\mathrm{M}_C^{-1}\mathrm{N}$, for α_{opt} and two other near optimal values of α. We further investigate these three cases for the Jacobi iterative process.

Table 5 shows the convergence behavior of Jacobi-Davidson, when the correction equation is solved with the Jacobi iterative method and with coupling parameter α_{opt}, obtained for $l_e = 2$. The left (on the left) and right (on the right) preconditioned variant are presented. Moreover, we have varied the number of Jacobi inner iterations.

When we compare the top part of Table 5 with the bottom part, then we see that more Jacobi inner iterations lead to less outer iterations for the same precision. More Jacobi iterations yields a better approximation of the correction vector and a better approximation of the correction vector results in fewer Jacobi-Davidson steps. When we compare the left part with the right part in Table 5, then we see that m steps with right preconditioned Jacobi iterations produces exactly the same results as with $m + 1$ left preconditioned Jacobi iterations. This is explained by stage 1 in §3.3 of right preconditioning: one extra preconditioning step is performed.

From §5.2.1 we know that the spectra of the error propagator are less optimal for $l_e = 4$ and $l_e = 1.2$, and therefore Jacobi will perform not as good as for $l_e = 2$. How does this affect the Jacobi-Davidson process? In Table 6 data are presented for three left preconditioned Jacobi iterations in each outer iteration, for $l_e = 4$ (left) and $l_e = 1.2$ (right). We should compare this with the top left part of Table 5. From this we see, that also Jacobi-Davidson performs less well for less optimal couplings.

Now we consider the Neumann-Dirichlet coupling. In our enhancement terminology (cf. §2.2) this

TABLE 6. *Convergence history of Jacobi-Davidson applied to the discretized eigenvalue problem for the two-dimensional Laplace operator for approximate solutions to the correction equation obtained with 3 left preconditioned Jacobi iterations on two subdomains and two almost optimal simple couplings. For explanation see §5.3.1.*

step	$\theta-\lambda$	$\|\|\mathbf{r}\|\|_2$	$\|\|\mathbf{r}'\|\|_2$	α	$\theta-\lambda$	$\|\|\mathbf{r}\|\|_2$	$\|\|\mathbf{r}'\|\|_2$	α
	$l_e = 4$				$l_e = 1.2$			
1	-1.61e-01	4.19e+00	4.19e+00	-2.1274	-1.61e-01	4.19e+00	4.19e+00	-1.2729
2	-2.93e-03	2.27e+00	2.00e+00	-2.1279	-1.33e-02	5.03e+00	3.31e+00	-1.2794
3	-1.12e-03	5.92e-01	4.62e-01	-2.1279	-1.92e-06	2.96e-02	2.94e-02	-1.2800
4	-1.46e-05	6.50e-02	5.83e-02	-2.1279	-4.11e-10	6.69e-04	5.57e-04	-1.2800
5	-4.02e-10	5.91e-04	5.71e-04	-2.1279	-1.18e-12	5.35e-05	3.97e-05	-1.2800
6	-2.47e-12	6.71e-05	4.05e-05	-2.1279	-1.24e-13	1.45e-06	1.21e-06	-1.2800
7	-1.47e-13	1.82e-07	1.14e-07	-2.1279	-3.13e-13	9.31e-08	5.82e-08	-1.2800
8	-1.67e-13	2.84e-10	2.82e-10		-1.46e-13	2.83e-09	2.09e-09	-1.2800
9					-1.72e-13	1.24e-10	1.09e-10	

can be interpreted as a Neumann boundary condition on the left: $C_{\ell\ell} = I$ and $C_{\ell r} = -I$, and a Dirichlet boundary condition on the right: $C_{r\ell} = I$ and $C_{rr} = I$. For dominated behavior (cf. §4.3.1 (iii), and §4.4 (48)) and for two subdomains it follows from (16) that

$$\sigma^2 \approx \frac{\overline{(\zeta-1)}\,(1+\zeta)}{\overline{(1-\zeta)}\,(\zeta+1)} = -1.$$

From this we see that for $\theta = \lambda^{(1,1)}$, the error propagator has, besides -1 and $+1$, only eigenvalues near $-\sqrt{-1}$ and $\sqrt{-1}$. Hence, the eigenvectors of $\mathbf{M}_C^{-1}\mathbf{N}$ will hardly be damped. Therefore, the Jacobi iteration will not perform well with Neumann-Dirichlet coupling. From Table 7 we see that Jacobi-Davidson clearly suffers from this effect.

5.3.2 GMRES

At the end of §2.3 we noted that Krylov subspace methods can be viewed as accelerators of the Jacobi iterative method. If we apply GMRES for the solution of the correction equation, instead of Jacobi iterations as in §5.3.1, then we should expect at least the same speed of convergence in the inner iteration. As a consequence, the speed of convergence of the Jacobi-Davidson (outer) iteration should be not worse but presumably better.

Our expectations are confirmed by the results in Table 8, for the simple optimized coupling and in Table 9 for the Neumann-Dirichlet coupling. For the same type of coupling one should compare the data for GMRES(m) with m Jacobi iterations: GMRES optimizes over the Krylov subspace spanned by powers of the (preconditioned) operator, whereas Jacobi uses only the last iteration vector for the computation of a solution to the linear system.

Note that with left preconditioned GMRES(4) and with Neumann-Dirichlet coupling, we have almost recovered the exact Jacobi-Davidson process from §5.1. This can be explained as follows. The eigenvalue distribution of the error propagator has besides -1 and $+1$, all other eigenvalues clustered around $\pm\sqrt{-1}$ for two subdomains. However, for four distinct eigenvalues, GMRES needs four steps at most for convergence. So the spectral properties of the error propagator for two subdomains with Neumann-Dirichlet coupling are worse for the Jacobi iterative method but ideal for the acceleration part of GMRES. This is

TABLE 7. *Convergence history of Jacobi-Davidson applied to the discretized eigenvalue problem for the two-dimensional Laplace operator for approximate solutions to the correction equation obtained with left (left) and right (right) preconditioned Jacobi iterations on two subdomains and Neumann-Dirichlet coupling. For explanation see §5.3.1.*

Neumann-Dirichlet coupling

	left DD-preconditioned			right DD-preconditioned		
step	$\theta - \lambda$	$\|\|\mathbf{r}\|\|_2$	$\|\|\mathbf{r}'\|\|_2$	$\theta - \lambda$	$\|\|\mathbf{r}\|\|_2$	$\|\|\mathbf{r}'\|\|_2$
	4 Jacobi inner iterations			3 Jacobi inner iterations		
1	-1.61e-01	4.19e+00	4.19e+00	-1.61e-01	4.19e+00	4.19e+00
2	-5.07e-02	8.72e+00	3.98e+00	-5.07e-02	8.72e+00	3.98e+00
3	-1.79e-02	4.85e+00	3.29e+00	-1.79e-02	4.85e+00	3.29e+00
4	-1.20e-02	2.40e+00	2.03e+00	-1.20e-02	2.40e+00	2.03e+00
5	-4.55e-03	2.69e+00	1.68e+00	-4.55e-03	2.69e+00	1.68e+00
6	-2.93e-04	6.90e-01	6.13e-01	-2.93e-04	6.90e-01	6.13e-01
7	-1.40e-04	3.74e-01	3.29e-01	-1.40e-04	3.74e-01	3.29e-01
8	-2.00e-05	2.10e-01	1.74e-01	-2.00e-05	2.10e-01	1.74e-01
9	-4.11e-06	7.32e-02	6.63e-02	-4.11e-06	7.32e-02	6.63e-02
10	-8.12e-07	3.88e-02	3.49e-02	-8.12e-07	3.88e-02	3.49e-02
11	-1.54e-07	1.41e-02	1.12e-02	-1.54e-07	1.41e-02	1.12e-02
12	-1.50e-08	5.84e-03	5.28e-03	-1.50e-08	5.84e-03	5.28e-03
13	-3.20e-09	2.62e-03	1.59e-03	-3.19e-09	2.62e-03	1.58e-03
14	-7.27e-10	1.22e-03	1.01e-03	-3.68e-10	9.02e-04	8.00e-04
15	-1.31e-10	5.86e-04	5.38e-04	-1.30e-10	5.82e-04	5.35e-04
16	-2.34e-11	2.63e-04	1.72e-04	-2.35e-11	2.63e-04	1.72e-04
17	-2.26e-12	5.03e-05	4.78e-05	-4.16e-13	5.03e-05	4.78e-05
18	-7.46e-13	2.08e-05	1.65e-05	-5.68e-14	2.08e-05	1.65e-05
19	-1.63e-13	3.90e-06	3.21e-06	-7.53e-13	3.88e-06	3.19e-06
20	4.12e-13	1.49e-06	1.25e-06	1.14e-13	1.27e-06	1.04e-06
21	9.95e-13	8.53e-07	7.63e-07	6.25e-13	3.60e-07	2.54e-07
22	-6.79e-13	2.55e-07	1.30e-07	-3.91e-13	2.30e-07	1.25e-07
23	4.01e-13	3.81e-08	3.56e-08	-7.11e-14	3.81e-08	3.56e-08
24	7.11e-14	1.18e-08	8.40e-09	-5.47e-13	1.18e-08	8.39e-09
25	4.90e-13	1.45e-09	1.41e-09	2.98e-13	1.19e-09	1.16e-09
26	6.98e-13	6.58e-10	6.30e-10	-5.90e-13	5.02e-10	4.80e-10

TABLE 8. *Convergence history of Jacobi-Davidson applied to the discretized eigenvalue problem for the two-dimensional Laplace operator for approximate solutions to the correction equation obtained with left (left) and right (right) preconditioned GMRES on two subdomains and simple optimized coupling. For explanation see §5.3.2.*

optimized coupling, $l_e = 2$

	left DD-preconditioned				right DD-preconditioned			
step	$\theta - \lambda$	$\|\|\mathbf{r}\|\|_2$	$\|\|\mathbf{r}'\|\|_2$	α	$\theta - \lambda$	$\|\|\mathbf{r}\|\|_2$	$\|\|\mathbf{r}'\|\|_2$	α
	GMRES(3)				GMRES(2)			
1	-1.61e-01	4.19e+00	4.19e+00	-1.6275	-1.61e-01	4.19e+00	4.19e+00	-1.6275
2	-2.72e-05	1.67e-01	1.67e-01	-1.6287	-3.74e-05	1.16e-01	1.16e-01	-1.6287
3	-3.05e-08	6.68e-03	6.23e-03	-1.6287	-5.89e-08	6.63e-03	5.43e-03	-1.6287
4	-3.06e-11	2.72e-04	2.71e-04	-1.6287	-1.46e-11	1.19e-04	1.13e-04	-1.6287
5	1.78e-15	1.72e-06	1.66e-06	-1.6287	-1.56e-13	1.46e-06	1.26e-06	-1.6287
6	-2.59e-13	1.34e-08	1.03e-08	-1.6287	-1.69e-13	6.81e-09	5.71e-09	-1.6287
7	-1.26e-13	7.94e-10	6.71e-10		-7.28e-14	4.38e-11	4.03e-11	
	GMRES(4)				GMRES(3)			
1	-1.61e-01	4.19e+00	4.19e+00	-1.6275	-1.61e-01	4.19e+00	4.19e+00	-1.6275
2	-1.52e-06	3.07e-02	3.02e-02	-1.6287	-1.34e-06	2.76e-02	2.71e-02	-1.6287
3	-1.39e-12	3.35e-05	3.32e-05	-1.6287	-4.85e-12	4.30e-05	4.13e-05	-1.6287
4	-1.42e-13	1.87e-07	1.76e-07	-1.6287	-1.42e-13	7.62e-07	7.31e-07	-1.6287
5	-1.79e-13	1.21e-09	1.17e-09	-1.6287	-1.19e-13	3.20e-09	3.19e-09	-1.6287
6	-1.85e-13	4.64e-12	4.09e-12		-1.28e-13	1.10e-11	1.05e-11	

TABLE 9. *Convergence history of Jacobi-Davidson applied to the discretized eigenvalue problem for the two-dimensional Laplace operator for approximate solutions to the correction equation obtained with left (left) and right (right) preconditioned GMRES on two subdomains and Neumann-Dirichlet coupling. For explanation see §5.3.2.*

Neumann-Dirichlet coupling

	left DD-preconditioned			right DD-preconditioned		
step	$\theta - \lambda$	$\|\|\mathbf{r}\|\|_2$	$\|\|\mathbf{r}'\|\|_2$	$\theta - \lambda$	$\|\|\mathbf{r}\|\|_2$	$\|\|\mathbf{r}'\|\|_2$
	GMRES(3)			GMRES(2)		
1	-1.61e-01	4.19e+00	4.19e+00	-1.61e-01	4.19e+00	4.19e+00
2	-1.20e-04	3.80e-01	3.80e-01	-5.87e-05	8.67e-02	8.48e-02
3	-5.48e-05	2.00e-01	1.96e-01	-7.21e-09	2.19e-03	2.18e-03
4	-1.13e-06	2.78e-02	1.73e-02	-1.71e-13	1.57e-06	1.22e-06
5	-1.99e-08	5.95e-03	4.43e-03	-1.49e-13	3.25e-08	3.09e-08
6	-8.17e-12	7.64e-05	7.48e-05	-1.74e-13	3.10e-12	2.98e-12
7	-1.79e-13	3.88e-06	3.83e-06			
8	-1.99e-13	1.41e-07	1.32e-07			
9	-1.14e-13	1.90e-09	1.61e-09			
10	-1.71e-13	4.80e-11	2.58e-11			
	GMRES(4)			GMRES(3)		
1	-1.61e-01	4.19e+00	4.19e+00	-1.61e-01	4.19e+00	4.19e+00
2	-9.65e-07	8.55e-03	6.10e-03	-9.65e-07	8.55e-03	6.10e-03
3	-1.44e-13	5.84e-10	5.79e-10	-1.55e-13	5.35e-10	5.30e-10
4	-1.21e-13	8.56e-14	1.01e-14	-1.49e-13	3.92e-14	4.12e-14

not a typical situation. In §5.4 we will see how the picture changes for more subdomains and with less accurate preconditioners.

5.4 More subdomains

We describe an experiment that illustrates what happens when the number of subdomains is increased. For each number of subdomains we keep the preconditioner fixed.

Our modelproblem is a channel that is made larger by extending new subdomains. We compute the largest eigenvalue and corresponding eigenvector of the Laplace operator on this channel. After adding a subdomain, this results in a different eigenvalue problem. For p subdomains the physical size and number of gridpoints in the y direction are taken to be fixed: $\omega_y = 1$ and $n_y = 63$, whereas in the x direction they increase: $\omega_x = p$ and $n_x = 63 + (p-1)\cdot 64$ for $1 \leq p \leq 6$.

Now, the idea is that the DD-preconditioner consists of block matrices defined on the enhanced subdomain grids. For the channel this results in one block matrix of size $(63+1)\times(63+1)$ (corresponding to the first subdomain on the left), $p-2$ block matrices of size $(64+2)\times(64+2)$ (corresponding to the $p-2$ intermediate subdomains) and one block matrix of size $(64+1)\times(64+1)$ (corresponding to the last subdomain on the right). If we select the same coupling between all subdomains, then we need to know the inverse action of 3 blocks (corresponding to the left, right, and a single intermediate subdomain). Furthermore, we construct the preconditioner only for the value of θ_1 of the first Jacobi-Davidson step. This fixed preconditioner is used for all iteration steps.

In order to be able to interpret the results properly, we have checked how Jacobi-Davidson with accurate solutions to the correction equation on the undecomposed domain (the 'exact' process) behaves. In Fig. 7 and Fig. 8 this is represented by the solid line.

We consider simple optimized (type 1), strong optimized (type 4), and Neumann-Dirichlet couplings. In each Jacobi-Davidson step we solve the correction equation approximately by right preconditioned GMRES(3). The number of nonzero eigenvalues of the error propagator is proportional to the number of subdomains. Because of this, it is reasonable that with a fixed number of inner iterations the accuracy will deteriorate for more subdomains.

Fig. 7 represents the convergence history of Jacobi-Davidson for the 'exact process' and for the inexact processes with different types of coupling, when starting with the vector (57). The 'exact process' does not change significantly for increasing values of p. For the inexact processes, the number of outer iterations increases when the number of subdomains increases (as expected). For the simple optimized coupling one can roughly say that convergence on p subdomains requires $5 + p$ outer iterations. The strong optimized coupling needs about $1-2$ iterations less. But for the Neumann-Dirichlet coupling the results do not show such a linear relationship: when increasing from 2 to 3 or from 3 to 4 subdomains, the number of outer iterations almost doubles.

When we compare the right bottom part of Table 9 with the two subdomain case in Fig. 7, then we see what happens when the preconditioner is less accurate for Neumann-Dirichlet coupling: the exact Jacobi-Davidson process can not longer be reproduced. Because the shift θ_1 in M_C is not equal to the shift θ in B_C, the eigenvalues of the error propagator that were close to $\pm\sqrt{-1}$ (cf. §5.3) start to deviate. This results in worse circumstances for GMRES.

From these results we conclude that the optimized couplings outperform the Neumann-Dirichlet coupling for more than 2 subdomains and a less accurate preconditioner

So far we have only considered the eigenvalue problem for the Laplace operator. The analysis of §4 also accommodates problems with first order operators. To illustrate that this does not give essential

FIGURE 7. *Convergence history of Jacobi-Davidson applied to the discretized eigenvalue problem for the two-dimensional Laplace operator for accurate solutions to the correction equation and increasing values of ω_x and n_x (solid lines) versus approximate solutions to the correction equation obtained from right preconditioned GMRES(3) with strong optimized (type 4) coupling (dashed lines with '$\circ$'), simple optimized (type 1) coupling (dash-dotted lines with '$\square$') and Neumann-Dirichlet coupling (dotted lines with '$*$') on an increasing number of subdomains. For explanation see §5.4.*

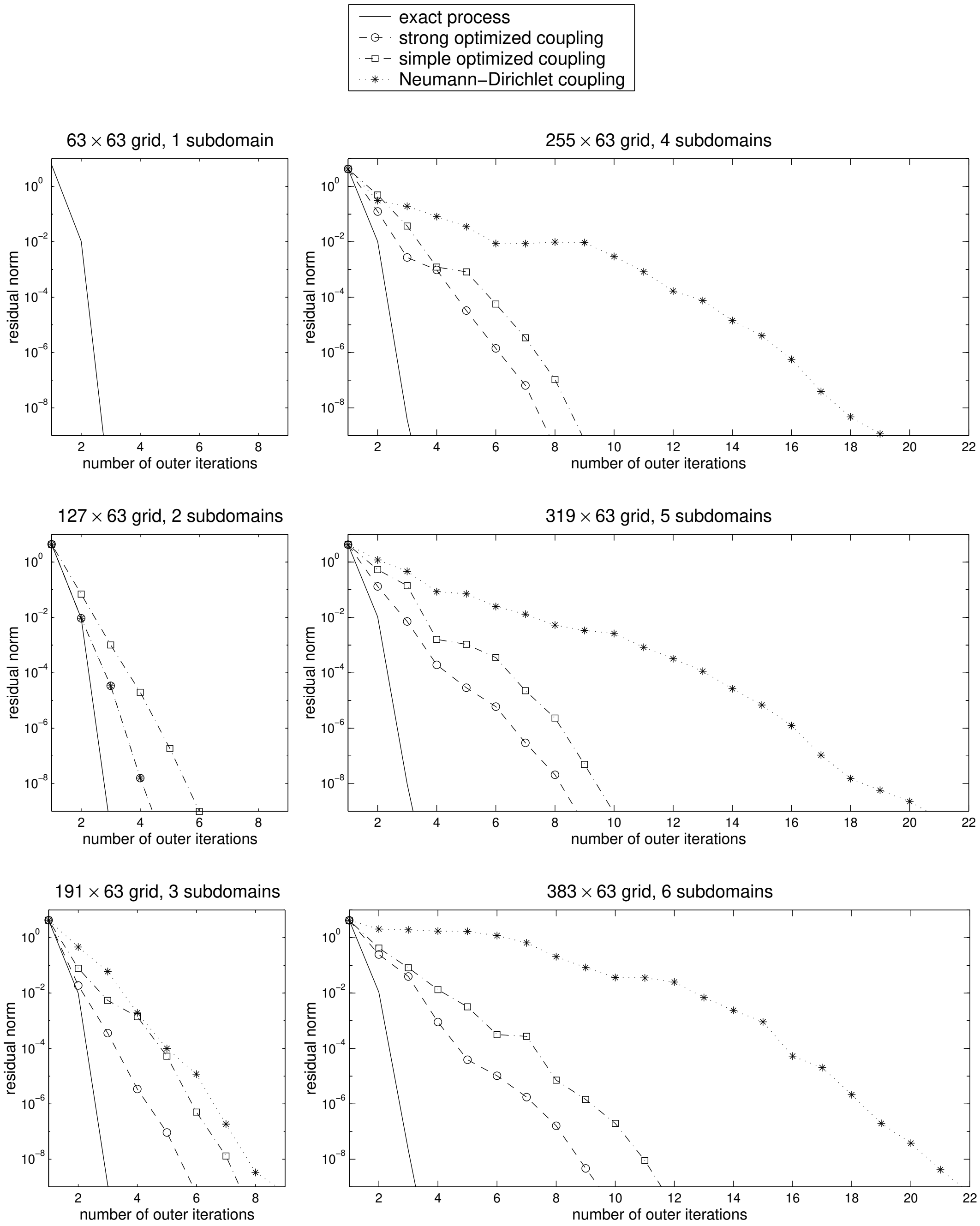
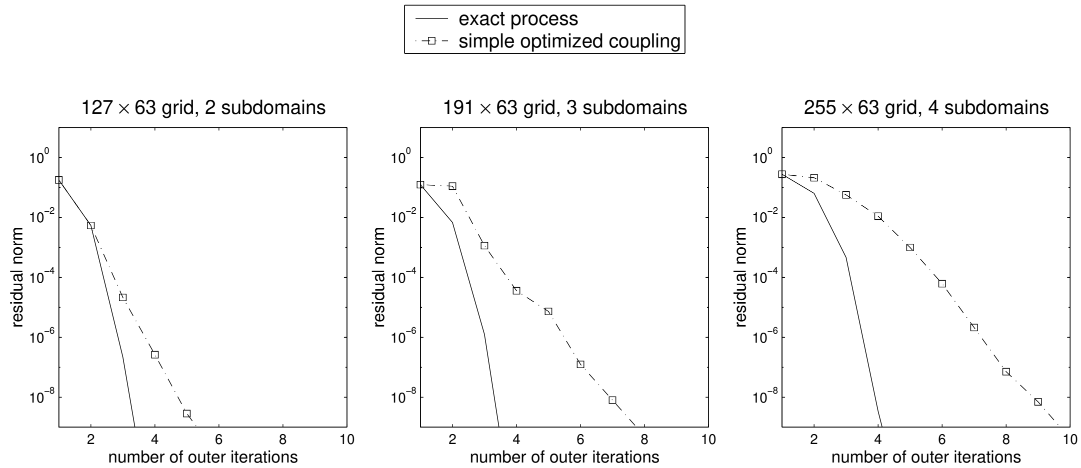


FIGURE 8. *Convergence history of Jacobi-Davidson applied to the discretized eigenvalue problem for the two-dimensional advection-diffusion operator (58) for accurate solutions to the correction equation and increasing values of ω_x and n_x (solid lines) versus approximate solutions to the correction equation obtained from right preconditioned GMRES(3) with simple optimized (type 1) coupling (dash-dotted lines with '□') on an increasing number of subdomains. For explanation see §5.4.*



differences, we consider

$$\frac{\partial^2}{\partial x^2} + \frac{2}{p}\frac{\partial}{\partial x} + \frac{\partial^2}{\partial y^2} + 5\frac{\partial}{\partial y} \tag{58}$$

on a domain with physical sizes $\omega_x = \frac{5}{4}p$ and $\omega_y = \frac{3}{4}$. Here $p \in \{2, 3, 4\}$ is the number of subdomains. With Jacobi-Davidson we compute the largest eigenvalue. In order to be in the convergence region of interest, Jacobi-Davidson is started with a vector equal to $(\mathbf{A} - 25\mathbf{I})^{-1}$ times the vector (57) (25 is close to the largest eigenvalue). All other settings are the same as in the previous experiment of this section.

Fig. 8 shows the convergence history of Jacobi-Davidson for accurate solutions and for approximate solutions of the correction equation. The approximate solutions are obtained from right preconditioned GMRES(3) with simple optimized (type 1) coupling. As in the previous experiment, the preconditioner is constructed only once at the first Jacobi-Davidson step. We see that the pictures in Fig. 8 are similar to those in Fig. 7.

6 Conclusions

In this paper we have outlined and analyzed how a nonoverlapping domain decomposition technique can be incorporated in the Jacobi-Davidson method. For large eigenvalue problems the solution of correction equations may become too expensive in terms of CPU time or/and memory. Domain decomposition may be attractive in a parallel computing environment.

For a model eigenvalue problem with constant coefficients we have analyzed how the coupling equations should be tuned. By numerical experiments we have verified our analysis. Indeed, further experiments showed that tuning of the coupling results in faster convergence of the Jacobi-Davidson process.

In realistical problems, the coefficient functions will not be constant and the domain will have a complicated geometry. For the determination of suitable coupling matrices, we intend to locally apply the approach that we discussed here. This 'local' approach is the subject of our next study.

Acknowledgements

The first author's contribution was sponsored by the Netherlands Organization for Scientific Research (NWO) under Grant No. 611-302-100.

References

[1] W. E. ARNOLDI, *The principle of minimized iteration in the solution of the matrix eigenvalue problem*, Quart. Appl. Math., 9 (1951), pp. 17–29.

[2] E. BRAKKEE, *Domain decomposition for the incompressible Navier-Stokes Equations*, Ph.D. thesis, Technische Universiteit Delft, Delft, The Netherlands, 1996.

[3] E. BRAKKEE, AND P. WILDERS, *The influence of interface conditions on convergence of Krylov-Schwarz domain decomposition for the advection-diffusion equation*, J. Sci. Comput., 12 (1997), pp. 11–30.

[4] X.-C. CAI, W. D. GROPP, D. E. KEYES, R. G. MELVIN, AND D. P. YOUNG, *Parallel Newton-Krylov-Schwarz algorithms for the transonic full potential equation*, SIAM J. Sci. Comput., 19:246–265, 1998.

[5] X.-C. CAI, AND D. E. KEYES, *Nonlinearly preconditioned inexact Newton algorithms*, submitted to SIAM J. Sci. Comput.

[6] E. R. DAVIDSON, *The iterative calculation of a few of the lowest eigenvalues and corresponding eigenvectors of large real-symmetric matrices*, J. Comput. Phys., 17 (1975), pp. 87–94.

[7] Q. DENG, *An analysis for a nonoverlapping domain decomposition iterative procedure*, SIAM J. Sci. Comput., 18 (1997), pp. 1517–1525.

[8] D. R. FOKKEMA, G. L. G. SLEIJPEN, AND H. A. VAN DER VORST, *Jacobi-Davidson style QR and QZ algorithms for the reduction of matrix pencils*, SIAM J. Sci. Comput., 20 (1999), pp. 94–125 (electronic).

[9] M. GENSEBERGER, *Domain decomposition on different levels of the Jacobi-Davidson method*, in preparation.

[10] A. HADJIDIMOS, D. NOUTSOS, AND M. TZOUMAS, *Nonoverlapping domain decomposition: a linear algebra viewpoint*, Math. Comput. Simulation, 51 (2000), pp. 597–625.

[11] Z. JIA, AND G. W. STEWART, *An analysis of the Rayleigh-Ritz method for approximating eigenspaces*, Technical Report TR-99-24/TR-4015, Department of Computer Science, University of Maryland, USA, 1999.

[12] G. LUBE, L. MÜLLER, AND F. C. OTTO, *A non-overlapping domain decomposition method for the advection-diffusion problem*, Computing, 64 (2000), pp. 49–68.

[13] Y. SAAD, *Numerical methods for large eigenvalue problems*, Manchester University Press, Manchester, UK, 1992.

[14] Y. SAAD AND M. H. SCHULTZ, *GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems*, SIAM J. Sci. Stat. Comput., 7 (1986), pp. 856–869.

[15] 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).

[16] G. L. G. SLEIJPEN AND H. A. VAN DER VORST, *The Jacobi-Davidson method for eigenvalue problems and its relation with accelerated inexact Newton scheme*. in Iterative Methods in Linear Algebra II, S. D. Margenov, and P. S. Vassilevski, eds, IMACS Ann. Comput. Appl. Math., 3:377–389, 1996.

[17] G. L. G. SLEIJPEN AND H. A. VAN DER VORST, *A Jacobi-Davidson iteration method for linear eigenvalue problems*, SIAM J. Matrix Anal. Appl., 17 (1996), pp. 401–425.

[18] K. H. TAN, *Local coupling in domain decomposition*, Ph.D. thesis, Utrecht University, Utrecht, The Netherlands, 1995.

[19] K. H. TAN AND M. J. A. BORSBOOM, *On generalized Schwarz coupling applied to advection-dominated problems*, in Domain decomposition methods in scientific and engineering computing (University Park, PA, 1993), Amer. Math. Soc., Providence, RI, 1994, pp. 125–130.

[20] W. P. TANG, *Generalized Schwarz splittings*, SIAM J. Sci. Stat. Comput., 13 (1992), pp. 573–595.

[21] R. S. VARGA, *Matrix iterative analysis*, Prentice-Hall Inc., Englewood Cliffs, N.J., 1962.

[22] P. WILDERS AND E. BRAKKEE, *Schwarz and Schur: an algebraic note on equivalence properties*, SIAM J. Sci. Comput., 20 (1999), pp. 2297–2303 (electronic).