SIAM J. SCI. COMPUT.
Vol. 0, No. 0, pp. 000–000

IS THERE MORE THAN ONE DIRICHLET–NEUMANN ALGORITHM FOR THE BIHARMONIC PROBLEM?*

MARTIN J. GANDER[†] AND YONGXIANG LIU[‡]

Abstract. The biharmonic problem is a fourth order partial differential equation and thus requires two boundary conditions, and not just one like in the Laplace case, where the notion of Dirichlet and Neumann boundary conditions comes from. A variational formulation gives an indication of which two conditions one could consider as Neumann, and which two as Dirichlet, and this choice was made in the literature to define Dirichlet–Neumann (and other) domain decomposition algorithms for the biharmonic equation. We show here that if one chooses other sets of two boundary conditions as Dirichlet and Neumann, one can obtain other Dirichlet–Neumann algorithms, and we prove that the classical choice leads to an algorithm with much less favorable convergence characteristics than our new choice. Our proof is based on showing that even optimizing the relaxation matrices (not just scalars) arising in the Dirichlet–Neumann algorithm running on two boundary conditions, the classical choice of Dirichlet and Neumann cannot achieve contraction rates comparable to our new choice, even though mesh independent convergence is achieved. We illustrate our results with numerical experiments, also exploring situations not covered by our analysis, and a simulation of the Golden Gate Bridge.

Key words. Dirichlet–Neumann algorithms, biharmonic problem, definition of Dirichlet and Neumann conditions

AMS subject classification. 65N55

DOI. 10.1137/19M1297956

1. Introduction.

We are interested in understanding Dirichlet–Neumann domain decomposition algorithms for the biharmonic equation, and thus consider as our model problem

$$\begin{aligned} \Delta^2 u &= f \quad \text{in } \Omega, \\ u &= 0 \quad \text{on } \partial\Omega, \\ \partial_n u &= 0 \quad \text{on } \partial\Omega, \end{aligned} \tag{1.1}$$

where Ω is a bounded domain in $\mathbb{R}^2$, f is the source term, and ∂_n denotes the normal derivative along the boundary. Since the operator has four derivatives, we need two boundary conditions, which is quite different from the classical Laplace operator, where it suffices to impose one. We have chosen in (1.1) to impose homogeneous boundary conditions both on the traces and the normal derivatives, which in the Laplace case notation would mean to impose homogeneous Dirichlet and Neumann conditions simultaneously. So the notion of Dirichlet and Neumann conditions for the biharmonic equation is different from the Laplace case.

The biharmonic equation (1.1) is important for modeling vibrations, and the first successful numerical methods were the groundbreaking variational methods of Ritz [30, 31] (see also [21] for a historical introduction). In addition to modeling thin plate bending problems and vibrations, there is, however, another mathematical

*Submitted to the journal's Methods and Algorithms for Scientific Computing section November 5, 2019; accepted for publication (in revised form) February 11, 2021; published electronically DATE.
https://doi.org/10.1137/19M1297956

Funding: This work was supported by the National Science Foundation of China (11701536).

[†]Section of Mathematics, University of Geneva, 2-4 rue du Lièvre, CP 64 CH-1211 Genève, Switzerland (Martin.Gander@unige.ch).

[‡]Corresponding author. Peng Cheng Laboratory, Shenzhen 518055, China (liuyx@pcl.ac.cn).

interpretation of the biharmonic problem (1.1): one can relate it to the Stokes problem for an incompressible viscous fluid in two dimensions. More precisely, we have the following two meaningful mathematical interpretations of (1.1):

1. There exists a unique function $u \in H_0^2(\Omega)$ that solves the minimization problem

   $$\min_{v\in H_0^2(\Omega)} J_1(v), \quad J_1(v) := \frac{1}{2}\int_\Omega |\Delta v|^2 d\Omega - \int_\Omega fvd\Omega. \tag{1.2}$$

   The corresponding variational problem is finding $u \in H_0^2(\Omega)$, s.t.

   $$\int_\Omega \Delta u\Delta v d\Omega = \int_\Omega fvd\Omega \quad \forall v \in H_0^2(\Omega). \tag{1.3}$$

   According to Green's formula, $\int_\Omega \Delta^2 uvd\Omega = \int_\Omega \Delta u\Delta vd\Omega + \int_{\partial\Omega} \partial_n \Delta uvds - \int_{\partial\Omega} \Delta u\partial_n vds$, which shows that u is a weak solution of (1.1). This problem is related to the Stokes problem in two spatial dimensions by setting the velocity of the fluid to be $\phi = [\partial_y u, -\partial_x u]^T$ (see [4, p. 282] for more details).
2. There also exists a unique function $u \in H_0^2(\Omega)$ that solves the minimization problem

   $$\begin{aligned} &\min_{v\in H_0^2(\Omega)} J_2(v), \\ &J_2(v) := \frac{1}{2}\int_\Omega |\Delta v|^2 + 2(1-\sigma)((\partial_{xy}v)^2 - \partial_{xx}v\partial_{yy}v)d\Omega - \int_\Omega fvd\Omega. \end{aligned} \tag{1.4}$$

   Here J_2 is the total potential energy of a thin plate (see, e.g., [31, 17]) and σ is a material constant called the Poisson's ratio, which lies in $[0, \frac{1}{2}]$. For example, the value for rubber is 0.4999, for copper it is 0.33, for clay it lies in $[0.30, 0.45]$, steel is in $[0.27, 0.30]$, concrete is in $[0.1, 0.2]$, glass is in $[0.18, 0.3]$, and for cork it is 0. The equivalent variational problem to (1.4) is finding $u \in H_0^2(\Omega)$, s.t.

   $$\begin{aligned} &\int_\Omega \Delta u\Delta v + (1-\sigma)(2\partial_{xy}u\partial_{xy}v - \partial_{xx}u\partial_{yy}v - \partial_{yy}u\partial_{xx}v)d\Omega \\ &\qquad = \int_\Omega fvd\Omega \quad \forall v \in H_0^2(\Omega). \end{aligned} \tag{1.5}$$

   Using Green's formula

   $$\int_\Omega 2\partial_{xy}u\partial_{xy}v - \partial_{xx}u\partial_{yy}v - \partial_{yy}u\partial_{xx}vd\Omega = \int_{\partial\Omega} -\partial_{\tau\tau}u\partial_n v + \partial_{n\tau}u\partial_\tau vds,$$

   where ∂_τ is the tangential derivative along the boundary, we see that u also solves problem (1.1).

If $v \in H_0^2(\Omega)$, then $J_1(v) = J_2(v)$, because the additional term $2(1-\sigma)((\partial_{xy}v)^2 - \partial_{xx}v\partial_{yy}v)$ in $J_2(v)$ then vanishes when using Green's formula, which also implies that in that case $J_2(v)$ is independent of σ. However, for $v \in H^2(\Omega)$, $J_1(v)$ and $J_2(v)$ may not be equivalent any more. For the functional $J_2(v)$, one can enlarge the range of σ to $\sigma \in (-1, 1)$, and the variational form remains well posed, so we will consider this larger range in our analysis.

The classical Dirichlet boundary condition for problem (1.1) is imposing the value and normal derivative at the boundary, as we did in (1.1), and we thus introduce our first Dirichlet trace operator

$$\mathcal{D}_1(u) := \begin{bmatrix} u \\ \frac{\partial u}{\partial n} \end{bmatrix}. \tag{1.6}$$

In contrast to Dirichlet–Neumann domain decomposition algorithms, for Schwarz domain decomposition algorithms, only the Dirichlet traces are needed, and there are many studies of Schwarz algorithms where the classical Dirichlet condition (1.6) is imposed between subdomains: Zhang analyzed in [35] a two level additive Schwarz method for conforming C^1 finite elements, and if the overlap size $\delta = O(H)$, where H corresponds to the diameter of the subdomains, then the condition number is independent of the subdomain size H and mesh size h. Later, Brenner analyzed in [3] the additive Schwarz preconditioner for nonconforming elements and gave a condition number bound $C(1+\frac{H}{\delta})^4$ for large overlap and $C(1+\frac{H}{\delta})^3$ for small overlap. Similarly, Feng and Rahman proposed an additive average Schwarz method with Morley finite element discretizations in [12], and its condition number is also estimated to be $O(1+\frac{H}{h})^3$. A nonoverlapping Schwarz preconditioner for a discontinuous Galerkin discretization was introduced by Feng and Karakashian in [11], with a condition number estimate $O(1+\frac{H}{h})^3$. We note that in contrast to condition number estimates for Schwarz methods applied to Laplace problems, there is an additional exponent 3 present for the biharmonic equation, and it was shown in [20] that choosing two different traces as Dirichlet condition leads to Schwarz methods for the biharmonic problem with a performance like when applied to Laplace's equation.

For the Neumann boundary conditions corresponding to the classical Dirichlet condition (1.6), there are two possibilities, depending on which of the functionals J_1 or J_2 we consider: for J_1 defined in (1.2), the corresponding Neumann condition would be

$$\mathcal{N}_1(u) := \begin{bmatrix} \Delta u \\ -\partial_n \Delta u \end{bmatrix}, \tag{1.7}$$

and for $J_2(v)$ defined in (1.4), it would be

$$\mathcal{N}_2(u) := \begin{bmatrix} \Delta u - (1-\sigma)\partial_{\tau\tau} u \\ -\partial_n \Delta u - (1-\sigma)\partial_\tau(\partial_{n\tau} u) \end{bmatrix}. \tag{1.8}$$

Now condition (1.7) does not always lead to a well posed problem for the biharmonic equation, which can be checked by setting $w := -\Delta u$: then the biharmonic problem is equivalent to a Poisson equation $-\Delta w = f$ with both Dirichlet and Neumann boundary conditions imposed, e.g., $w = g_1$ and $\partial_n w = g_2$ on the boundary, which is overdetermined. However, condition (1.8) can be interpreted as the freely supported boundary condition for the thin plate problem, and this is always well posed up to a linear function.

For solving the thin plate problem, there are many domain decomposition methods in the literature that use the Neumann condition (1.8): Tallec, Mandel, and Vidrascu presented a Neumann–Neumann-type preconditioner in [32], where they proved the condition number of the preconditioned system to be $O(1+\log\frac{H}{h})^2$ by using the abstract Schwarz framework. Then Dohrmann presented a substructuring method in [6] with some constraints on the substructure boundary. The corresponding condition number was proved to be $O(1+\log\frac{H}{h})^2$ by Mandel and Dohrmann in [26].

The FETI method was proposed and studied by Farhat and Mandel in [10] and also by Mandel, Tezaur and Farhat in [27], where continuity of the transverse displacements is enforced at the substructure corners, and the condition number is $O(1+\log\frac{H}{h})^3$. Gervasio proposed in [23] a Dirichlet–Neumann algorithm for a biharmonic problem with two lower order terms, where the original problem was transformed into an equivalent system of two Poisson equations, and $\mathcal{D}_1$ was used as the Dirichlet condition and a condition similar to $\mathcal{N}_2$ was used as the Neumann condition; a contraction estimate was also derived, but without any optimization.

The classical clamped Dirichlet condition (1.6) is, however, not the only possible choice for a Dirichlet condition. Instead of (1.6) and (1.7), one could also consider

$$\mathcal{D}_3(u) := \begin{bmatrix} u \\ \Delta u \end{bmatrix} \tag{1.9}$$

as the Dirichlet condition, and then naturally the corresponding Neumann condition would be

$$\mathcal{N}_3(u) := \begin{bmatrix} \partial_n u \\ -\partial_n \Delta u \end{bmatrix} \tag{1.10}$$

(see, for example, [8]). Similarly, in the thin plate case, instead of (1.6) and (1.8), another choice for the Dirichlet condition would be

$$\mathcal{D}_4(u) := \begin{bmatrix} u \\ \Delta u - (1-\sigma)\partial_{\tau\tau} u \end{bmatrix}, \tag{1.11}$$

and then the corresponding Neumann condition would be

$$\mathcal{N}_4(u) := \begin{bmatrix} \partial_n u \\ -\partial_n \Delta u - (1-\sigma)\partial_\tau(\partial_{n\tau} u) \end{bmatrix}. \tag{1.12}$$

When the boundary is flat, conditions (1.9) and (1.11) are essentially equivalent, since imposing u also imposes $\partial_{\tau\tau}$. Similarly also conditions (1.10) and (1.12) are equivalent for flat boundaries. For curved boundaries, however, and as transmission conditions in domain decomposition methods, these conditions are different.

In addition to the domain decomposition methods we have mentioned above, there are also other iterative methods for solving the biharmonic equation (1.1), such as multigrid methods [29, 33], multilevel method [28, 34], and meshless methods [25]. We are interested here, however, specifically in the Dirichlet–Neumann domain decomposition method, because this method depends on what one chooses as the Dirichlet and then the corresponding Neumann condition. In section 2, we present several such Dirichlet–Neumann methods, of which only one has so far been studied in the literature. We then give a precise convergence analysis for these methods in section 3, which reveals that there are better choices than the classical one for the Dirichlet–Neumann method. In section 4, we show numerical results to illustrate our analysis.

2. Four Dirichlet–Neumann algorithms.

To simplify the description of the various Dirichlet–Neumann algorithms and the analysis that follows, we consider the biharmonic equation on an unbounded domain,

$$\Delta^2 u = f \quad \text{in } \Omega := \mathbb{R}^2, \tag{2.1}$$

and impose that the solution u decays at infinity. We assume that Ω is divided into two nonoverlapping subdomains $\Omega_1 = (-\infty, 0) \times \mathbb{R}$ and $\Omega_2 = (0, +\infty) \times \mathbb{R}$ with the interface at $x = 0$ denoted by $\Gamma := \overline{\Omega}_1 \cap \overline{\Omega}_2$. Let n_i, $i = 1, 2$ be the unit outward normal vector of Ω_i on Γ, and τ_i be the corresponding tangential vector along Γ, which implies that $n_1 = -n_2$ and $\tau_1 = -\tau_2$. Let $f_1 := f|_{\Omega_1}, f_2 := f|_{\Omega_2}$. To simplify the presentation of the Dirichlet–Neumann algorithms, we also introduce the redundant operator $\mathcal{D}_2 := \mathcal{D}_1$. We can then define four different Dirichlet–Neumann algorithms by just using the indices $j = 1, 2, 3, 4$:

Dirichlet–Neumann algorithm DN_j: for a given initial guess

$$\mathbf{g}_1^0 = \begin{bmatrix} g_{1A}^0 \\ g_{1B}^0 \end{bmatrix},$$

perform for iteration index $n = 0, 1, 2, \ldots$ the following steps:[1]

1. Compute in Ω_1 an approximate solution u_1^n by solving the Dirichlet problem

$$\begin{aligned} \Delta^2 u_1^n &= f_1 \quad \text{in } \Omega_1, \\ \mathcal{D}_j(u_1^n) &= \mathbf{g}_1^n \quad \text{on } \Gamma. \end{aligned} \tag{2.2}$$

2. Update the transmission condition for Ω_2 by setting

$$\mathbf{g}_2^n = \mathcal{N}_j(u_1^n).$$

3. Compute in Ω_2 an approximate solution u_2^n by solving the Neumann problem

$$\begin{aligned} \Delta^2 u_2^n &= f_2 \quad \text{in } \Omega_2, \\ \mathcal{N}_j(u_2^n) &= \mathbf{g}_2^n \quad \text{on } \Gamma. \end{aligned} \tag{2.3}$$

4. Update the transmission condition for Ω_1 by setting

$$\mathbf{g}_1^* = \mathcal{N}_j(u_2^n).$$

5. Relax the transmission condition

$$\begin{bmatrix} g_{1A}^{n+1} \\ g_{1B}^{n+1} \end{bmatrix} = \underbrace{\begin{bmatrix} \theta_{11} & \theta_{12} \\ \theta_{21} & \theta_{22} \end{bmatrix}}_{\Theta:=} \begin{bmatrix} g_{1A}^* \\ g_{1B}^* \end{bmatrix} + \begin{bmatrix} 1-\theta_{11} & -\theta_{12} \\ -\theta_{21} & 1-\theta_{22} \end{bmatrix} \begin{bmatrix} g_{1A}^n \\ g_{1B}^n \end{bmatrix}.$$

A first important observation is that for DN_1, the subproblem (2.3) is not always well posed, as we explained in the introduction, and so the Dirichlet–Neumann algorithm DN_1 cannot be used in practice. We therefore only study the remaining three Dirichlet–Neumann algorithms DN_j, $j = 2, 3, 4$, in what follows. Note that since there are two traces for the fourth order biharmonic problem, we have an entire relaxation matrix Θ containing four adjustable parameters to improve the convergence rate. In [23] only a scalar value θ was used for relaxation, like in the Dirichlet–Neumann method applied to Laplace's equation. We will prove in the next section that even though the relaxation matrix Θ seems more powerful, the scalar θ, i.e., $\Theta = \theta I$, where I denotes the identity matrix, is structurally already optimal. We next want to determine which of the Dirichlet–Neumann algorithms DN_j, $j = 2, 3, 4$, has the most favorable convergence properties.

[1] This iteration index n is not related to the notation ∂_n for the normal derivative.

3. Convergence behavior of the Dirichlet–Neumann methods. To compare the different Dirichlet–Neumann methods, we need to study their contraction properties, which we need to optimize using the four parameters in the relaxation matrix Θ, to obtain a fair comparison. From the algorithm definition, we see that the solution u of the model problem (2.1) is a fixed point in each case, and by linearity, it suffices to consider the error equations, i.e., $f = 0$, and to analyze convergence to zero. We use Fourier analysis to study the contraction properties of the Dirichlet–Neumann methods as was done in [5] for Dirichlet–Neumann applied to Laplace problems. Fourier analysis has become mainstream for the study of optimized Schwarz methods (see, for example, [16, 13, 15]) and proved useful also in the form of Laplace transforms for studying waveform relaxation variants of the Dirichlet–Neumann method [18, 19]. We thus take a Fourier transform in the y direction, and then the functions $u_i(x, y)$, $g_{iA}(y)$, $g_{iB}(y)$ become $\widehat{u}_i = \widehat{u}_i(x, k)$, $\widehat{g}_{iA} = \widehat{g}_{iA}(k)$, and $\widehat{g}_{iB} = \widehat{g}_{iB}(k)$. For each $k \neq 0$, we obtain an ordinary differential equation (ODE) on each subdomain,[2]

$$\frac{\partial^4 \widehat{u}_1}{\partial x^4} - 2k^2 \frac{\partial^2 \widehat{u}_1}{\partial x^2} + k^4 \widehat{u}_1 = 0, \quad x \leqslant 0, \qquad \frac{\partial^4 \widehat{u}_2}{\partial x^4} - 2k^2 \frac{\partial^2 \widehat{u}_2}{\partial x^2} + k^4 \widehat{u}_2 = 0, \quad x \geqslant 0. \tag{3.1}$$

In order to solve these ODEs, we need to solve the characteristic equation $\lambda^4 - 2k^2\lambda^2 + k^4 = 0$, from which we obtain $\lambda = \pm k$. The general solution on subdomain Ω_i, $i = 1, 2$, is thus given by

$$\widehat{u_i}(x, k) = C_1 e^{|k|x} + C_1^{'} x e^{|k|x} + C_2 e^{-|k|x} + C_2^{'} x e^{-|k|x}.$$

Since by assumption u_i decays at infinity, the constants corresponding to the growing solutions must be zero, and we obtain

$$\widehat{u}_1(x, k) = C_1 e^{|k|x} + C_1^{'} x e^{|k|x}, \quad x \leqslant 0, \quad \widehat{u}_2(x, k) = C_2 e^{-|k|x} + C_2^{'} x e^{-|k|x}, \quad x \geqslant 0. \tag{3.2}$$

This form of the subdomain solutions is the same for these Dirichlet–Neumann algorithms DN_j, $j = 2, 3, 4$; the convergence of each algorithm is only determined by how the constants C_1, C_1', C_2, and C_2' are affected by the iteration of DN_j.

3.1. Analysis of $\boldsymbol{DN_2}$. Introducing the subdomain solution (3.2) at iteration n into the Dirichlet condition of step 1 of DN_2 in (2.2), and recalling that the interface Γ is at $x = 0$, we obtain

$$\begin{bmatrix} \widehat{u}_1^n(0, k) \\ \partial_{n_1} \widehat{u}_1^n(0, k) \end{bmatrix} = \begin{bmatrix} C_1^n \\ C_1^n |k| + C_1^{'n} \end{bmatrix} = \begin{bmatrix} \widehat{g}_{1A}^n \\ \widehat{g}_{1B}^n \end{bmatrix} \quad \text{on } \Gamma. \tag{3.3}$$

This leads using step 2 to the transmission condition for Ω_2,

$$\begin{aligned} \begin{bmatrix} \widehat{g}_{2A}^n \\ \widehat{g}_{2B}^n \end{bmatrix} &= \begin{bmatrix} \partial_{xx} \widehat{u}_1^n(0, k) - |k|^2 \widehat{u}_1^n(0, k) + (1 - \sigma)|k|^2 \widehat{u}_1^n(0, k) \\ -\partial_{n_2}(\partial_{xx} \widehat{u}_1^n(0, k) - |k|^2 \widehat{u}_1^n(0, k)) + \partial_{n_2}((1 - \sigma)|k|^2 \widehat{u}_1^n(0, k)) \end{bmatrix} \\ &= \begin{bmatrix} C_1^n (1 - \sigma)|k|^2 + C_1^{'n} \cdot 2|k| \\ -C_1^n (1 - \sigma)|k|^3 - C_1^{'n}((1 - \sigma)|k|^2 - 2|k|^2) \end{bmatrix} \quad \text{on } \Gamma. \end{aligned} \tag{3.4}$$

To simplify the notation, we now introduce the vectors

$$\widetilde{\mathbf{C}}_i^n := \begin{bmatrix} C_i^n \\ C_i^{'n} \end{bmatrix}, \quad \widetilde{\mathbf{g}}_i^n := \begin{bmatrix} \widehat{g}_{iA}^n \\ \widehat{g}_{iB}^n \end{bmatrix}$$

[2] The case of the constant mode $k = 0$ is excluded because of our assumption that solutions decay to zero at infinity.

and the matrices

$$A_1 := \begin{bmatrix} 1 & 0 \\ |k| & 1 \end{bmatrix},$$
$$T_1 := \begin{bmatrix} (1-\sigma)|k|^2 & 2|k| \\ -(1-\sigma)|k|^3 & -(1-\sigma)|k|^2 + 2|k|^2 \end{bmatrix}.$$

We can then write (3.3) and (3.4) in compact form,

$$A_1 \widetilde{\mathbf{C}}_1^n = \widetilde{\mathbf{g}}_1^n, \text{ and } \quad T_1 \widetilde{\mathbf{C}}_1^n = \widetilde{\mathbf{g}}_2^n.$$

Similarly, we obtain for the subproblem in Ω_2 the matrices

$$A_2 := \begin{bmatrix} (1-\sigma)|k|^2 & -2|k| \\ (1-\sigma)|k|^3 & -(1-\sigma)|k|^2 + 2|k|^2 \end{bmatrix},$$
$$T_2 := \begin{bmatrix} 1 & 0 \\ -|k| & 1 \end{bmatrix}.$$

The Dirichlet–Neumann algorithm DN_2 can thus be written in Fourier in the compact form

$$\begin{aligned} A_1 \widetilde{\mathbf{C}}_1^n = \widetilde{\mathbf{g}}_1^n, \quad T_1 \widetilde{\mathbf{C}}_1^n = \widetilde{\mathbf{g}}_2^n, \\ A_2 \widetilde{\mathbf{C}}_2^n = \widetilde{\mathbf{g}}_2^n, \quad T_2 \widetilde{\mathbf{C}}_2^n = \widetilde{\mathbf{g}}_1^*, \\ \widetilde{\mathbf{g}}_1^{n+1} = \Theta \widetilde{\mathbf{g}}_1^* + (I - \Theta)\widetilde{\mathbf{g}}_1^n, \end{aligned}$$

where I denotes the identity matrix. By eliminating the intermediate variables $\widetilde{\mathbf{g}}_2^n$ and $\widetilde{\mathbf{g}}_1^*$, we obtain on the interface Γ the iteration

$$\widetilde{\mathbf{g}}_1^{n+1} = [I - \Theta(I - T_2 A_2^{-1} T_1 A_1^{-1})]\widetilde{\mathbf{g}}_1^n. \tag{3.5}$$

To analyze the convergence factor of the DN_2 algorithm, we need to study for each Fourier mode $k \neq 0$ the spectrum of the iteration matrix $I - \Theta(I - T_2 A_2^{-1} T_1 A_1^{-1})$ and determine the best choice of the relaxation matrix Θ over all $k \neq 0$ to see how fast DN_2 can converge. By a direct calculation, we obtain

$$T_2 A_2^{-1} T_1 A_1^{-1} = -\frac{1}{(1-\sigma)(\sigma+3)} \begin{bmatrix} |k| & 0 \\ 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1+\sigma & -2 \\ -2 & 1+\sigma \end{bmatrix}^2 \begin{bmatrix} |k| & 0 \\ 0 & 1 \end{bmatrix}. \tag{3.6}$$

We start with the simplest choice of using only a scalar relaxation parameter, $\Theta := \theta I$.

THEOREM 3.1. *If $\Theta = \theta I$, then the convergence factor of the corresponding DN_2 method is for all Fourier modes k bounded by*

$$\rho = \max\left\{ \left| \frac{\sigma + 3 - 4\theta}{\sigma + 3} \right|, \left| \frac{1 - \sigma - 4\theta}{1 - \sigma} \right| \right\}. \tag{3.7}$$

The optimal choice for $\sigma \in (-1, 1)$ of the scalar relaxation parameter and resulting convergence factor are

$$\theta = \frac{(1-\sigma)(\sigma+3)}{8} \quad \Longrightarrow \quad \rho^1_{DN_2} = \frac{1+\sigma}{2}, \tag{3.8}$$

where we use the superscript 1 in $\rho^1_{DN_2}$ to denote the first variant of DN_2 with a scalar relaxation parameter.

Proof. Using (3.6), we obtain by a direct computation for the iteration matrix

$$\begin{aligned}&I-\Theta(I-T_2A_2^{-1}T_1A_1^{-1})\\&=\begin{bmatrix}|k| & 0\\ 0 & 1\end{bmatrix}^{-1}\frac{1}{(1-\sigma)(\sigma+3)}\begin{bmatrix}(1-\sigma)(\sigma+3)-8\theta & 4(1+\sigma)\theta\\ 4(1+\sigma)\theta & (1-\sigma)(\sigma+3)-8\theta\end{bmatrix}\begin{bmatrix}|k| & 0\\ 0 & 1\end{bmatrix}.\end{aligned}$$

The corresponding characteristic polynomial is

$$\lambda^2-2\frac{(1-\sigma)(\sigma+3)-8\theta}{(1-\sigma)(\sigma+3)}+\frac{16\theta^2-16\theta+(1-\sigma)(\sigma+3)}{(1-\sigma)(\sigma+3)},$$

where the dependence on the Fourier mode k cancels, and we obtain for all $k\neq 0$

$$\lambda=\left\{\frac{\sigma+3-4\theta}{\sigma+3},\frac{1-\sigma-4\theta}{1-\sigma}\right\}.$$

Since $\sigma\in(-1,1)$, the optimal choice of θ is obtained by equioscillation, i.e.,

$$\frac{\sigma+3-4\theta}{\sigma+3}=-\frac{1-\sigma-4\theta}{1-\sigma}.$$

The solution of this equation gives the result of the theorem. □

Theorem 3.1 shows that the Dirichlet–Neumann algorithm DN_2 converges independently of the frequency k, which implies mesh independent convergence of DN_2: no matter how fine the mesh is, and how high the frequency signals it can carry, are the convergence factor does not depend on this. Theorem 3.1 also shows that for σ close to -1, the convergence is very fast, and for σ close to 1, it is very slow. For the physically relevant range of $\sigma\in[0,\frac{1}{2}]$ the convergence factor lies in $[\frac{1}{2},\frac{3}{4}]$, which gives quite robust convergence, but in contrast to the Dirichlet–Neumann algorithm applied to Laplace's equation in this symmetric geometry (see, for example, [5, 9, 7][3], there is no parameter choice that gives convergence in two iterations, so DN_2 for the fourth order problem is missing an essential feature of Dirichlet–Neumann methods for second order problems, but maybe this is due only to our simplified choice of just a scalar relaxation parameter we used, so we need to investigate more general relaxation matrices. We prove in the next theorem that introducing a general diagonal relaxation matrix does not improve the situation.

Theorem 3.2. *If* $\Theta=\operatorname{diag}(\theta_{11},\theta_{22})$*, a general diagonal matrix, then the optimized convergence factor of* DN_2 *is for* $\sigma\in(-1,1)$

$$\begin{aligned}\rho_{DN_2}^2=\min_{\theta_{11},\theta_{22}}\max&\left\{\frac{|(1-\sigma)(\sigma+3)-4(\theta_{11}+\theta_{22})+4\sqrt{(1+\sigma)^2\theta_{11}\theta_{22}+(\theta_{11}-\theta_{22})^2}|}{(1-\sigma)(\sigma+3)},\right.\\&\left.\frac{|(1-\sigma)(\sigma+3)-4(\theta_{11}+\theta_{22})-4\sqrt{(1+\sigma)^2\theta_{11}\theta_{22}+(\theta_{11}-\theta_{22})^2}|}{(1-\sigma)(\sigma+3)}\right\},\end{aligned}$$

and the minimum is obtained when $\theta_{11}=\theta_{22}$*. (We use the superscript* 2 *in* $\rho_{DN_2}^2$ *to denote the second variant of* DN_2 *with a diagonal relaxation matrix.)*

[3] This important property is well known in the domain decomposition community and was already mentioned by the inventors of the algorithm in [2], who stated the following for when the method is used as a preconditioner: "It is also interesting to note that if a symmetric region is cut in half, and treated fully symmetrically, then $S=2S(1)$ and the conjugate gradient iteration converges in one step." In order to obtain this property for a nonsymmetric problem, one needs to use Robin transmission conditions (see [1, 22]).

Proof. Using (3.6), we compute again the iteration matrix

$$\begin{aligned}&I - \Theta(I - T_2 A_2^{-1} T_1 A_1^{-1})\\ &= \begin{bmatrix}|k| & 0\\ 0 & 1\end{bmatrix}^{-1} \frac{1}{(1-\sigma)(\sigma+3)} \begin{bmatrix}(1-\sigma)(\sigma+3) - 8\theta_{11} & 4(1+\sigma)\theta_{11}\\ 4(1+\sigma)\theta_{22} & (1-\sigma)(\sigma+3) - 8\theta_{22}\end{bmatrix} \begin{bmatrix}|k| & 0\\ 0 & 1\end{bmatrix}\end{aligned}$$

and the corresponding characteristic polynomial

$$\lambda^2 - 2\frac{(1-\sigma)(\sigma+3) - 4(\theta_{11}+\theta_{22})}{(1-\sigma)(\sigma+3)} + \frac{16\theta_{11}\theta_{22} - 8(\theta_{11}+\theta_{22}) + (1-\sigma)(\sigma+3)}{(1-\sigma)(\sigma+3)}.$$

Then the eigenvalues are

$$\lambda = \frac{(1-\sigma)(\sigma+3) - 4(\theta_{11}+\theta_{22}) \pm 4\sqrt{(1+\sigma)^2\theta_{11}\theta_{22} + (\theta_{11}-\theta_{22})^2}}{(1-\sigma)(\sigma+3)},$$

so the corresponding convergence factor is

$$\begin{aligned}\rho = \max\Bigg\{&\frac{|(1-\sigma)(\sigma+3) - 4(\theta_{11}+\theta_{22}) + 4\sqrt{(1+\sigma)^2\theta_{11}\theta_{22} + (\theta_{11}-\theta_{22})^2}|}{(1-\sigma)(\sigma+3)},\\ &\frac{|(1-\sigma)(\sigma+3) - 4(\theta_{11}+\theta_{22}) - 4\sqrt{(1+\sigma)^2\theta_{11}\theta_{22} + (\theta_{11}-\theta_{22})^2}|}{(1-\sigma)(\sigma+3)}\Bigg\}.\end{aligned} \tag{3.9}$$

To study this convergence factor, we introduce the two quantities $x_1 := \theta_{11} + \theta_{22}$ and $x_2 := \theta_{11} - \theta_{22}$ and the auxiliary function

$$q(x_1, x_2) := \sqrt{\frac{1}{4}(1+\sigma)^2 x_1^2 + (1 - \frac{1}{4}(1+\sigma)^2)x_2^2}.$$

Then $q(x_1, x_2)$ is always real since $\sigma \in (-1, 1)$. We can thus rewrite the convergence factor ρ in (3.9) as

$$\rho(x_1, x_2) = \frac{|(1-\sigma)(\sigma+3) - 4x_1| + 4q(x_1, x_2)}{(1-\sigma)(\sigma+3)}. \tag{3.10}$$

A direct calculation shows that

$$\frac{(1-\sigma)(\sigma+3)}{4} \cdot \frac{\partial\rho}{\partial x_2} = \frac{\partial q}{\partial x_2} = \frac{(1 - \frac{1}{4}(1+\sigma)^2)x_2}{\sqrt{\frac{1}{4}(1+\sigma)^2 x_1^2 + (1 - \frac{1}{4}(1+\sigma)^2)x_2^2}}.$$

Furthermore,

$$\begin{aligned}&\frac{(1-\sigma)(\sigma+3)}{4} \cdot \frac{\partial^2\rho}{\partial x_2^2}\\ &= \frac{\partial^2 q}{\partial x_2^2}\\ &= \frac{(1 - \frac{1}{4}(1+\sigma)^2)}{(\frac{1}{4}(1+\sigma)^2 x_1^2 + (1 - \frac{1}{4}(1+\sigma)^2)x_2^2)^{\frac{1}{2}}} - \frac{(1 - \frac{1}{4}(1+\sigma)^2)^2 x_2^2}{(\frac{1}{4}(1+\sigma)^2 x_1^2 + (1 - \frac{1}{4}(1+\sigma)^2)x_2^2)^{\frac{3}{2}}}\\ &= \frac{(1 - \frac{1}{4}(1+\sigma)^2) \cdot \frac{1}{4}(1+\sigma)^2 x_1^2}{(\frac{1}{4}(1+\sigma)^2 x_1^2 + (1 - \frac{1}{4}(1+\sigma)^2)x_2^2)^{\frac{3}{2}}}\\ &> 0,\end{aligned}$$

so ρ attains its minimum at $x_2 = 0$, which implies $\theta_{11} = \theta_{22}$. □

So a diagonal relaxation matrix does not improve DN_2. We show in the next theorem that not even a fully general relaxation matrix Θ can recover the fundamental convergence property the Dirichlet–Neumann algorithm has for Laplace's equation. The proof of this result is technical and quite involved, but it is very important to know that DN_2 does not have the fundamental property Dirichlet–Neumann algorithms have for Laplace's equation, in view of what we will show for our new Dirichlet–Neumann algorithms DN_3 and DN_4 for the biharmonic problem.

To prove this technical result, we need some preparation. We define the function

$$l_0(y,a,b) := 2(1+\sigma)\left(ya + \frac{1}{y}b\right), \tag{3.11}$$

where $y \in [y_{min}, y_{max}]$ ($y_{min}, y_{max} > 0$), $a, b \in \mathbb{R}$, and $\sigma \in (-1,1)$ is a constant. The function l_0 will be important in a term in the representation of the eigenvalues of the iteration matrix. If we fix the value of a, b, then $l_0(y,a,b)$ is a continuous bounded function in y. For the following analysis, we also introduce the two quantities

$$m := \frac{1}{2}(l_0(y,a,b)_{max} + l_0(y,a,b)_{min}), \quad d := \frac{1}{2}(l_0(y,a,b)_{max} - l_0(y,a,b)_{min}), \tag{3.12}$$

where $l_0(y,a,b)_{max}$, $l_0(y,a,b)_{min}$ are the maximum and minimum of $l_0(y,a,b)$ for $y \in [y_{min}, y_{max}]$. Here m represents the mean value of $l_0(y,a,b)$ and d is the oscillation amplitude. According to this definition, we have

$$l_0(y,a,b)_{max} = m + d, \text{ and } \quad l_0(y,a,b)_{min} = m - d.$$

Taking a derivative of l_0 with respect to y, we get

$$\frac{\partial l_0(y,a,b)}{\partial y} := 2(1+\sigma)\left(a - \frac{1}{y^2}b\right).$$

To study how the bounds of $l_0(y,a,b)$, depend on the values of a, b and σ, we need to consider the following four cases, where we assume that $ab \neq 0$:

1. $\frac{\partial l_0(y,a,b)}{\partial y}|_{y=y_{min}} < 0$, $\frac{\partial l_0(y,a,b)}{\partial y}|_{y=y_{max}} < 0$. Then $b > y^2 a$ for $\forall y \in [y_{min}, y_{max}]$, and we obtain

$$l_0(y,a,b)_{min} = l_0(y_{max},a,b) = 2\left(1+\sigma)(y_{max}a + \frac{1}{y_{max}}b\right), \tag{3.13}$$

$$l_0(y,a,b)_{max} = l_0(y_{min},a,b) = 2(1+\sigma)\left(y_{min}a + \frac{1}{y_{min}}b\right). \tag{3.14}$$

2. $\frac{\partial l_0(y,a,b)}{\partial y}|_{y=y_{min}} > 0$, $\frac{\partial l_0(y,a,b)}{\partial y}|_{y=y_{max}} > 0$. Then $b < y^2 a$ for $\forall y \in [y_{min}, y_{max}]$, which yields

$$l_0(y,a,b)_{min} = l_0(y_{min},a,b) = 2\left(1+\sigma)(y_{min}a + \frac{1}{y_{min}}b\right), \tag{3.15}$$

$$l_0(y,a,b)_{max} = l_0(y_{max},a,b) = 2\left(1+\sigma)(y_{max}a + \frac{1}{y_{max}}b\right). \tag{3.16}$$

3. $\frac{\partial l_0(y,a,b)}{\partial y}|_{y=y_{min}} < 0$, $\frac{\partial l_0(y,a,b)}{\partial y}|_{y=y_{max}} > 0$. This implies that $a > 0$ and $b > 0$. So $y_0 := \sqrt{\frac{b}{a}} \in [y_{min}, y_{max}]$, and we get

$$l_0(y,a,b)_{min} = l_0(y_0,a,b) = 4(1+\sigma)\sqrt{ab}, \tag{3.17}$$

$$l_0(y,a,b)_{max} = \max\{l_0(y_{min},a,b), l_0(y_{max},a,b)\}. \tag{3.18}$$

4. $\frac{\partial l_0(y,a,b)}{\partial y}|_{y=y_{min}} > 0$, $\frac{\partial l_0(y,a,b)}{\partial y}|_{y=y_{max}} < 0$. This implies that $a < 0$ and $b < 0$. So $y_0 := \sqrt{\frac{b}{a}} \in [y_{min}, y_{max}]$, and we have

$$l_0(y,a,b)_{min} = \min\{l_0(y_{min},a,b), l_0(y_{max},a,b)\}, \tag{3.19}$$

$$l_0(y,a,b)_{max} = l_0(y_0,a,b) = -4(1+\sigma)\sqrt{ab}. \tag{3.20}$$

These four cases cover all possible choices of a, b, except when the derivative of l_0 vanishes. We will also need more detailed information about the mean value m and the oscillation amplitude d. Our goal is to find the best choice of a and b, which make $l_0(y,a,b)$ have the smallest oscillation amplitude d for the same mean value m.

For case (1) and (2), $l_0(y,a,b)$ is a monotonic function of y, so the mean value and the oscillation amplitude for both cases can be written in the same form, namely,

$$m_{12} = \frac{1}{2}(l_0(y,a,b)_{max} + l_0(y,a,b)_{min}) = (1+\sigma)(y_{max}+y_{min})\left(a + \frac{1}{y_{max}y_{min}}b\right), \tag{3.21}$$

$$d_{12} = \frac{1}{2}(l_0(y,a,b)_{max} - l_0(y,a,b)_{min}) = (1+\sigma)(y_{max}-y_{min})\left|a - \frac{1}{y_{max}y_{min}}b\right|. \tag{3.22}$$

By a direct verification, we see that the range of m_{12} covers $\mathbb{R}$ when a and b are varying. For case (3) and (4), we need the next lemma.

LEMMA 3.3. *If the mean value m_{34} is fixed in case* (3) *or* (4)*, the best choice of $y_0 = \sqrt{\frac{b}{a}}$ is $y_{0*}^2 = y_{max}y_{min}$, which makes the oscillation as small as possible, i.e.,*

$$d_{34} = \min_{y_{min}\leqslant y_0 \leqslant y_{max}} \left\{\frac{1}{2}(l_0(y,a,b)_{max} - l_0(y,a,b)_{min})\right\}.$$

Proof. We only prove case (3) here. The other case can be proved similarly. First, we obtain by a direct calculation for $y_{0*}^2 = \frac{b}{a} = y_{max}y_{min}$ that

$$l_0(y_{min},a,b) = l_0(y_{max},a,b).$$

Assume there exists another choice $\widetilde{a}$ and $\widetilde{b}$, which leads to the same mean value m_{34}. By (3.12), and (3.17), (3.18), and a direct computation, we obtain

$$\begin{aligned} m_{34} &= (1+\sigma)\frac{(y_{0*}+y_{min})^2}{y_{min}}a = (1+\sigma)\frac{(y_{0*}+y_{max})^2}{y_{max}}a \\ &= \max\left\{(1+\sigma)\frac{(\widetilde{y}_0+y_{min})^2}{y_{min}}\widetilde{a}, (1+\sigma)\frac{(\widetilde{y}_0+y_{max})^2}{y_{max}}\widetilde{a}\right\}, \end{aligned} \tag{3.23}$$

where $\widetilde{y}_0 := \frac{\widetilde{b}}{\widetilde{a}}$. For simplicity, we assume that

$$(1+\sigma)\frac{(\widetilde{y}_0+y_{min})^2}{y_{min}}\widetilde{a} > (1+\sigma)\frac{(\widetilde{y}_0+y_{max})^2}{y_{max}}\widetilde{a},$$

which means $\widetilde{y}_0^2 > y_{max}y_{min}$; the case where the reverse inequality holds can be treated similarly. Then from (3.23), we obtain

$$\widetilde{a} = \frac{(y_{0*}+y_{min})^2}{(\widetilde{y}_0+y_{min})^2}a,$$

and thus

$$\begin{aligned} d_{34} &= \frac{1}{2}(l_0(y_{min}, a, b) - l_0(y_{0*}, a, b)) = (1+\sigma)\frac{(y_{0*} - y_{min})^2}{y_{min}}a, \\ \widetilde{d}_{34} &= \frac{1}{2}(l_0(y_{min}, a, b) - l_0(\widetilde{y}_0, a, b)) = (1+\sigma)\frac{(\widetilde{y}_0 - y_{min})^2}{y_{min}}\widetilde{a} \\ &= (1+\sigma)\frac{(\widetilde{y}_0 - y_{min})^2}{y_{min}}\frac{(y_{0*} + y_{min})^2}{(\widetilde{y}_0 + y_{min})^2}a. \end{aligned}$$

The proof of the lemma is now obtained by concluding that $d_{34} < \widetilde{d}_{34}$ from the fact that $\widetilde{y}_0 > y_{0*}$. □

According to Lemma 3.3, the best choice is $y_{0*} = \sqrt{y_{max}y_{min}}$ with smallest oscillation

$$d_{34} = (1+\sigma)(\sqrt{y_{max}} - \sqrt{y_{min}})^2|a|, \tag{3.24}$$

and its corresponding mean value is

$$m_{34} = (1+\sigma)(\sqrt{y_{max}} + \sqrt{y_{min}})^2 a. \tag{3.25}$$

By a direct verification, we see that the range of m_{34} covers $\mathbb{R}$ except 0, when a is varying. Furthermore, it is also necessary to compare d_{12} and d_{34} for the same mean value m, and we need the following.

LEMMA 3.4. *If $m_{12}(a,b) = m_{34}(\widetilde{a})$ for some a, b, and $\widetilde{a}$, then $d_{12}(a,b) > d_{34}(\widetilde{a})$.*

Proof. If $a \neq 0$, since $m_{12}(a,b) = m_{34}(\widetilde{a})$, from (3.21) and (3.25), we have

$$\widetilde{a} = \frac{y_{max} + y_{min}}{(\sqrt{y_{max}} + \sqrt{y_{min}})^2}\left(1 + \frac{1}{y_{max}y_{min}}\frac{b}{a}\right)a.$$

Then by (3.24),

$$d_{34}(\widetilde{a}) = (1+\sigma)\frac{(y_{max} + y_{min})(\sqrt{y_{max}} - \sqrt{y_{min}})^2}{(\sqrt{y_{max}} + \sqrt{y_{min}})^2}\left|1 + \frac{1}{y_{max}y_{min}}\frac{b}{a}\right||a|,$$

and therefore we have

$$\frac{d_{12}}{d_{34}} = \frac{(\sqrt{y_{max}} + \sqrt{y_{min}})^3}{(\sqrt{y_{max}} - \sqrt{y_{min}})(y_{max} + y_{min})}\left|\frac{y_{max}y_{min} - \frac{b}{a}}{y_{max}y_{min} + \frac{b}{a}}\right|.$$

Notice that for the cases (1) and (2), a and b satisfy $\frac{b}{a} < y_{min}^2$ or $\frac{b}{a} > y_{max}^2$, which leads to $\frac{d_{12}}{d_{34}} > 1$.

If $a = 0$, by a direct calculation, we still have

$$\frac{d_{12}}{d_{34}} = \frac{(\sqrt{y_{max}} + \sqrt{y_{min}})^3}{(\sqrt{y_{max}} - \sqrt{y_{min}})(y_{max} + y_{min})} > 1.$$ □

We are now ready to prove the main technical result on the relaxation matrix in DN_2.

THEOREM 3.5. *If Θ is a fully general 2×2 relaxation matrix, then the optimized convergence factor for DN_2,*

$$\rho^3_{DN_2} = \min_{\theta_{11},\theta_{12},\theta_{21},\theta_{22}} \rho, \qquad \Theta = \begin{bmatrix} \theta_{11} & \theta_{12} \\ \theta_{21} & \theta_{22} \end{bmatrix}, \tag{3.26}$$

attains its minimum when $\theta_{12} = \theta_{21} = 0$, up to higher order terms $O(\frac{1}{\sqrt{|k|_{max}}})$.[4] *(We use the superscript 3 in $\rho^3_{DN_2}$ to denote the third variant of DN_2 with a full general relaxation matrix.)*

Proof. Following the same approach as in Theorems 3.1 and 3.2, we obtain the iteration matrix

$$I - \Theta(I - T_2A_2^{-1}T_1A_1^{-1}) = \frac{1}{(1-\sigma)(\sigma+3)} \cdot \begin{bmatrix} (1-\sigma)(\sigma+3) - 8\theta_{11} + 4(1+\sigma)|k|\theta_{12} & 4(1+\sigma)\frac{1}{|k|}\theta_{11} - 8\theta_{12} \\ -8\theta_{21} + 4(1+\sigma)|k|\theta_{22} & (1-\sigma)(\sigma+3) + 4(1+\sigma)\frac{1}{|k|}\theta_{21} - 8\theta_{22} \end{bmatrix}$$

and the corresponding characteristic polynomial

$$\lambda^2 - 2\frac{(1-\sigma)(\sigma+3) - 4(\theta_{11}+\theta_{22}) + 2(1+\sigma)(|k|\theta_{12} + \frac{1}{|k|}\theta_{21})}{(1-\sigma)(\sigma+3)} + \frac{16\theta_{11}\theta_{22} - 8(\theta_{11}+\theta_{22}) + (1-\sigma)(\sigma+3) - 16\theta_{12}\theta_{21} + 4(1+\sigma)(|k|\theta_{12} + \frac{1}{|k|}\theta_{21})}{(1-\sigma)(\sigma+3)}.$$

Then the eigenvalues are

$$\lambda = \frac{1}{(1-\sigma)(\sigma+3)}\Bigg(2(1+\sigma)(|k|\theta_{12} + \frac{1}{|k|}\theta_{21}\Bigg) - 4(\theta_{11}+\theta_{22}) + (1-\sigma)(\sigma+3) \pm 2\sqrt{((1+\sigma)\left(|k|\theta_{12} + \frac{1}{|k|}\theta_{21}\right) - 2(\theta_{11}+\theta_{22}))^2 + \ 4(1-\sigma)(\sigma+3)(\theta_{12}\theta_{21} - \theta_{11}\theta_{22})}\Bigg). \tag{3.27}$$

In contrast, however, to Theorem 3.2, the eigenvalues λ may be complex now. So for proving Theorem 3.5, we need to estimate the bounds of each term in (3.27). To simplify the notation, let $y = |k|, a = \theta_{12}, b = \theta_{21}$ in the function $l_0(y, a, b)$ we defined in (3.11), and then we define

$$l(|k|) := 2(1+\sigma)\left(|k|\theta_{12} + \frac{1}{|k|}\theta_{21}\right) = l_0(|k|, \theta_{12}, \theta_{21}),$$

which is a function of k for fixed θ_{12}, θ_{21} and σ. The range of k is $|k| \in [|k|_{min}, |k|_{max}]$, where $|k|_{min}$ and $|k|_{max} > 0$.

[4] Such terms correspond to terms of $O(h^{\frac{1}{2}})$, where h represents the mesh size along the interface, so the theorem holds when the mesh size becomes small.

Note that by (3.27), we have

$$
\begin{aligned}
&\max_{|k|_{min}\leqslant|k|\leqslant|k|_{max}} |\lambda| \\
&\geqslant \max_{|k|_{min}\leqslant|k|\leqslant|k|_{max}} \frac{1}{(1-\sigma)(\sigma+3)}|2(1+\sigma)(|k|\theta_{12}+\frac{1}{|k|}\theta_{21}) \\
&\qquad\qquad\qquad\qquad -4(\theta_{11}+\theta_{22})+(1-\sigma)(\sigma+3)| \\
&= \max_{|k|_{min}\leqslant|k|\leqslant|k|_{max}} \frac{1}{(1-\sigma)(\sigma+3)}|l(|k|)-4(\theta_{11}+\theta_{22})+(1-\sigma)(\sigma+3)|.
\end{aligned}
$$

Because $|k|_{max}$ is usually of order $O(h^{-1})$ in the discrete situation and $|k|_{min}$ is a small constant (see [13]), to preserve $|\lambda|<1$ for the convergence, θ_{12} and θ_{21} need to be small enough, i.e., $\theta_{12}=O(\frac{1}{|k|_{max}})$ and $\theta_{21}=O(|k|_{min})$, which leads to the term $\theta_{12}\theta_{21}$ in (3.27) to be very small.

Based on this observation, we obtain that $l(|k|)$ is the main element of λ compared to the other terms in (3.27). For θ_{11},θ_{22}, and σ fixed, to optimize λ through controlling $l(|k|)$, we should choose the best values of θ_{12} and θ_{21}, which make $l(|k|)$ have the smallest oscillation amplitude d for the same mean value m.

Combining Lemmas 3.3 and 3.4, to get the smallest oscillation for the same mean value, the parameters θ_{12}, θ_{21} should be chosen such that $|k|_{0*}=\sqrt{\frac{\theta_{21}}{\theta_{12}}}=\sqrt{|k|_{max}|k|_{min}}$, which leads to $m=m_{34}, d=d_{34}$. Denoting by $\alpha:=4(\theta_{11}+\theta_{22})$, the eigenvalues (3.27) can be simplified to

$$
\lambda=\frac{(1-\sigma)(\sigma+3)+l(|k|)-\alpha\pm\sqrt{(l(|k|)-\alpha)^2+16(1-\sigma)(\sigma+3)(\theta_{12}\theta_{21}-\theta_{11}\theta_{22})}}{(1-\sigma)(\sigma+3)}.
$$

The convergence factor ρ is then the larger of the two eigenvalues, i.e.,

$$
\rho=\begin{cases}
\dfrac{1}{(1-\sigma)(\sigma+3)}\Big(|(1-\sigma)(\sigma+3)+l(|k|)-\alpha| \\
\quad+\sqrt{(l(|k|)-\alpha)^2+16(1-\sigma)(\sigma+3)(\theta_{12}\theta_{21}-\theta_{11}\theta_{22})}\Big) & \text{if } \lambda \text{ is real,}\\
\dfrac{1}{(1-\sigma)(\sigma+3)}\Big(|(1-\sigma)(\sigma+3)+l(|k|)-\alpha|^2+|(l(|k|)-\alpha)^2 \\
\quad+16(1-\sigma)(\sigma+3)(\theta_{12}\theta_{21}-\theta_{11}\theta_{22})|\Big)^{1/2} & \text{if } \lambda \text{ is complex.}
\end{cases}
\tag{3.28}
$$

To finish the proof, we need to solve the min-max problem

$$
\min_{\theta_{11},\theta_{12},\theta_{21}\theta_{22}}\ \max_{|k|_{min}\leqslant|k|\leqslant|k|_{max}} \rho(\theta_{11},\theta_{12},\theta_{21},\theta_{22},k). \tag{3.29}
$$

According to the definitions of m and d, we have $m-d\leqslant l(|k|)\leqslant m+d$, so by (3.28)

$$
\begin{aligned}
\max_{|k|_{min}\leqslant|k|\leqslant|k|_{max}} \rho &\geqslant \max_{|k|_{min}\leqslant|k|\leqslant|k|_{max}} \left\{\frac{|(1-\sigma)(\sigma+3)+l(|k|)-\alpha|}{(1-\sigma)(\sigma+3)}\right\} \\
&= \frac{|(1-\sigma)(\sigma+3)+m-\alpha|+d}{(1-\sigma)(\sigma+3)} \\
&\geqslant \frac{d}{(1-\sigma)(\sigma+3)}.
\end{aligned}
$$

In order to have convergence, we need that

$$\max_{|k|_{min}\leqslant|k|\leqslant|k|_{max}} \rho < 1,$$

which implies the necessary condition

$$\frac{d}{(1-\sigma)(\sigma+3)} < 1.$$

Noting that $m = m_{34}$ and $d = d_{34}$, by (3.24), we obtain

$$\theta_{12} = O\left(\frac{1}{(\sqrt{|k|_{max}} - \sqrt{|k|_{min}})^2}\right) = O\left(\frac{1}{|k|_{max}}\right),$$

and thus we have $\theta_{21} = |k|_{0*}^2\theta_{12} = O(1)$. So

$$\theta_{12}\theta_{21} = O\left(\frac{1}{|k|_{max}}\right)$$

is a small value compared to the other terms, and we can thus ignore it in the following.

Next, since $\theta_{12}\theta_{21} = |k|_{0*}^2\theta_{12}^2 \geqslant 0$, if $\theta_{11}\theta_{22} \leqslant 0$, then we have $16(1-\sigma)(\sigma+3)(\theta_{12}\theta_{21} - \theta_{11}\theta_{22}) \geqslant 0$. The eigenvalue λ is then real and by (3.28), this results in

$$\rho \geqslant \frac{|(1-\sigma)(\sigma+3) + l(|k|) - \alpha| + |l(|k|) - \alpha|}{(1-\sigma)(\sigma+3)} \geqslant 1,$$

which means the method does not converge. So we obtain as another necessary condition that $\theta_{11}\theta_{22} > 0$ for convergence.

Third, through direct manipulation, we find that

$$|m| - d = |m_{34}| - d_{34} = 4(1+\sigma)\sqrt{|k|_{max}}\sqrt{|k|_{min}}|\theta_{12}| = O\left(\frac{1}{\sqrt{|k|_{max}}}\right),$$

which indicates that $|m|$ and d are nearly the same, and we can replace one by the other when ignoring higher order terms.

From (3.28), in order to analyze the convergence factor ρ for all $|k| \in [|k|_{min}, |k|_{max}]$, we analyze $l(|k|) \in [l(|k|)_{min}, l(|k|)_{max}]$, which is equivalent to $m-d \leqslant l(|k|) \leqslant m+d$, s.t.

$$(3.30) \qquad l(|k|) \in \begin{cases} [0, 2m] & \text{if } m \geqslant 0, \\ [2m, 0] & \text{if } m \leqslant 0. \end{cases}$$

Based on this observation, we see that $0 \in [l(|k|)_{min}, l(|k|)_{max}]$.

Because $0 \leqslant \theta_{11}\theta_{22} \leqslant \frac{(\theta_{11}+\theta_{22})^2}{4} = \frac{\alpha^2}{64}$ and $\sigma \in (-1,1)$, we have $\alpha^2 - 16(1-\sigma)(\sigma+3)\theta_{11}\theta_{22} \geqslant 0$. By taking $l(|k|) = 0$ in (3.28) and omitting the small terms, we define

$$(3.31) \qquad \rho_1(\alpha, \theta_{11}, \theta_{22}) = \frac{|(1-\sigma)(\sigma+3) - \alpha| + \sqrt{\alpha^2 - 16(1-\sigma)(\sigma+3)\theta_{11}\theta_{22}}}{(1-\sigma)(\sigma+3)}.$$

Then we have
(3.32)

$$\max_{|k|_{min}\leqslant|k|\leqslant|k|_{max}} \rho \geqslant \max_{\substack{|k|_{min} \leqslant |k| \leqslant |k|_{max},\\ l(|k|) = 0}} \rho + O\left(\frac{1}{\sqrt{|k|_{max}}}\right) = \rho_1 + O\left(\frac{1}{\sqrt{|k|_{max}}}\right)$$

for any fixed $\theta_{11}, \theta_{22}, \theta_{12}, \theta_{21}$. Notice that $\rho_1(\alpha, \theta_{11}, \theta_{22}) = \rho(x_1, x_2)$, where $\rho(x_1, x_2)$ is defined in (3.10), which we already analyzed for proving Theorem 3.2. So by (3.31) and (3.32), we know that the matrix Θ in Theorem 3.2 is a solution of (3.29), which finishes our proof. □

Remark 3.6. The solution to problem (3.29) is not unique. The choice of Θ in Theorem 3.5 is only one particular solution. According to Theorems 3.1, 3.2, and 3.5, one of the optimal choices is $\theta^*_{11} = \theta^*_{22} = \frac{(1-\sigma)(\sigma+3)}{8}$ and $\theta^*_{12} = \theta^*_{21} = 0$. Then we have $m^* = 0$, and $\alpha^* = (1-\sigma)(\sigma+3)$. To find the other solutions, we assume that there is another choice of the relaxation matrix Θ, say, $\theta'_{11}, \theta'_{22}, \theta'_{12}, \theta'_{21}$, which leads to new values m', α' that also solve the problem (3.29), and we thus have

$$\max_{|k|_{min} \leqslant |k| \leqslant |k|_{max}} \rho(\theta^*_{11}, \theta^*_{12}, \theta^*_{21}, \theta^*_{22}, k) = \max_{|k|_{min} \leqslant |k| \leqslant |k|_{max}} \rho(\theta'_{11}, \theta'_{12}, \theta'_{21}, \theta'_{22}, k). \tag{3.33}$$

However, by (3.31) and (3.32), we could have

$$\max_{|k|_{min} \leqslant |k| \leqslant |k|_{max}} \rho(\theta'_{11}, \theta'_{12}, \theta'_{21}, \theta'_{22}, k) \geqslant \rho_1(\alpha', \theta'_{11}, \theta'_{22}) + O\left(\frac{1}{\sqrt{|k|_{max}}}\right),$$

where the right part of this inequality only depends on $\theta'_{11}, \theta'_{22}$. As in the above discussion, $\rho_1(\alpha', \theta'_{11}, \theta'_{22}) = \rho(x_1, x_2)$, and also according to Theorem 3.2, the optimal choice of $\theta'_{11}, \theta'_{22}$ should satisfy (this solution is unique) $\theta'_{11} = \theta'_{22} = \frac{(1-\sigma)(\sigma+3)}{8} = \theta^*_{11} = \theta^*_{22}$, which leads to $\alpha' = \alpha^*$. So the other solution to (3.29) should satisfy $\theta'_{11} = \theta^*_{11}$, $\theta'_{22} = \theta^*_{22}$, which means that the difference only lies in θ'_{12} or θ'_{21}. Substituting these into (3.28), we obtain

$$\rho_2(l_0(|k|, \theta'_{12}, \theta'_{21}), \alpha^*, \theta^*_{11}, \theta^*_{22}) := \begin{cases} \dfrac{|l_0(|k|, \theta'_{12}, \theta'_{21})| + \sqrt{t}}{\alpha^*}, & t \geqslant 0, \\ \dfrac{\left(|l_0(|k|, \theta'_{12}, \theta'_{21})|^2 + |t|\right)^{1/2}}{\alpha^*}, & t < 0, \end{cases} \tag{3.34}$$

where $t := (l_0(|k|, \theta'_{12}, \theta'_{21}))^2 - 2\alpha^* l_0(|k|, \theta'_{12}, \theta'_{21}) + \frac{1}{4}(1+\sigma)^2 \alpha^{*2}$. Furthermore, in order to find $\theta'_{12}, \theta'_{21}$, by (3.33) and the above discussion, we have

$$\begin{aligned} \rho_1(\alpha', \theta'_{11}, \theta'_{22}) &= \rho_1(\alpha^*, \theta^*_{11}, \theta^*_{22}) \\ &= \max_{|k|_{min} \leqslant |k| \leqslant |k|_{max}} \rho(\theta^*_{11}, \theta^*_{12}, \theta^*_{21}, \theta^*_{22}, |k|) \\ &= \max_{|k|_{min} \leqslant |k| \leqslant |k|_{max}} \rho(\theta'_{11}, \theta'_{12}, \theta'_{21}, \theta'_{22}, |k|) \\ &= \max_{|k|_{min} \leqslant |k| \leqslant |k|_{max}} \rho_2(l_0(|k|, \theta'_{12}, \theta'_{21}), \alpha', \theta'_{11}, \theta'_{22}), \end{aligned}$$

which shows that

$$\begin{aligned} &\max_{\min\{0, 2m'\} \leqslant l_0(|k|, \theta'_{12}, \theta'_{21}) \leqslant \max\{0, 2m'\}} \rho_2(l_0(|k|, \theta'_{12}, \theta'_{21}), \alpha', \theta'_{11}, \theta'_{22}) \\ &\qquad = \rho_1(\alpha', \theta'_{11}, \theta'_{22}) = \rho_1(\alpha^*, \theta^*_{11}, \theta^*_{22}). \end{aligned}$$

Here, we also denote $l' := l_0(|k|, \theta'_{12}, \theta'_{21})$ for simplicity. Solving this equation, we obtain $0 \leqslant m' \leqslant \frac{1}{8}(1+\sigma)^2 \alpha^*$. So when l' lies in the region $[0, 2m'] = [0, \frac{1}{4}(1+\sigma)^2 \alpha^*]$, then we have $\rho_1(\alpha^*, \theta^*_{11}, \theta^*_{22}) \geqslant \rho_2(l', \alpha^*, \theta^*_{11}, \theta^*_{22})$, where the equality only holds at the points $l' = 0$ and $l' = \frac{1}{4}(1+\sigma)^2 \alpha^*$ (see Figure 1).

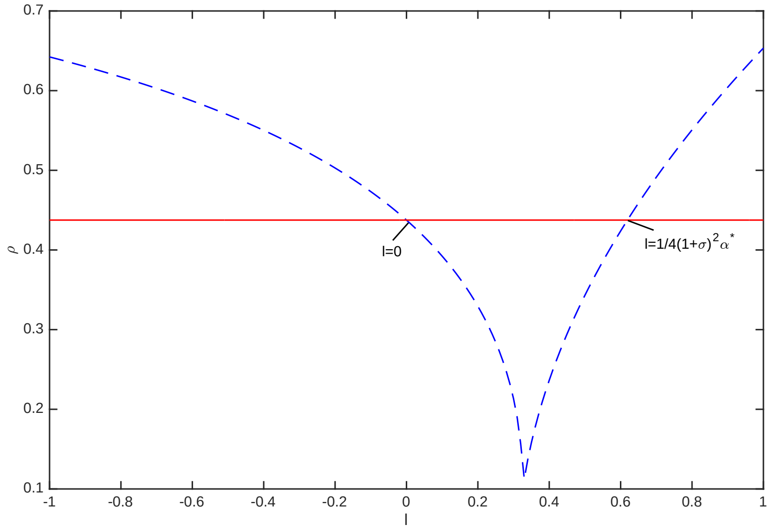

FIG. 1. *The two functions ρ_1 and ρ_2 of l. The solid line is $\rho_1(\alpha^*, \theta_{11}^*, \theta_{22}^*) = \frac{1+\sigma}{2}$ and the dashed line is $\rho_2(l, \alpha^*, \theta_{11}^*, \theta_{22}^*)$, where $\sigma = -\frac{1}{8}$.*

Having the range of m', together with (3.25), we obtain

$$0 \leqslant \theta_{12}' \leqslant \frac{(1+\sigma)\alpha^*}{8(\sqrt{|k|_{max}} + \sqrt{|k|_{min}})^2} = O\left(\frac{1}{|k|_{max}}\right),$$

and $\theta_{21}' = |k|_{max}|k|_{min}\theta_{12}'$, which implies

$$0 \leqslant \theta_{21}' \leqslant \frac{(1+\sigma)\alpha^*|k|_{max}|k|_{min}}{8(\sqrt{|k|_{max}} + \sqrt{|k|_{min}})^2} = O(1).$$

3.2. Analysis of $\boldsymbol{DN_3}$ and $\boldsymbol{DN_4}$. We study now the two Dirichlet–Neumann algorithms DN_3 and DN_4 and show that they become direct solvers for an appropriate choice of the relaxation matrix Θ, i.e., they have the same fundamental convergence property as the Dirichlet–Neumann algorithm when applied to Laplace's equation [5, 9, 7]. To obtain their error propagation equation similar to (3.5), we find after a short calculation the matrices

$$\bar{A}_1 := \begin{bmatrix} 1 & 0 \\ (1-\sigma)|k|^2 & 2|k| \end{bmatrix} \quad \bar{T}_1 := \begin{bmatrix} -|k| & -1 \\ -(1-\sigma)|k|^3 & -(1-\sigma)|k|^2 + 2|k|^2 \end{bmatrix}$$

and

$$\bar{A}_2 := \begin{bmatrix} |k| & -1 \\ (1-\sigma)|k|^3 & -(1-\sigma)|k|^2 + 2|k|^2 \end{bmatrix}, \quad \bar{T}_2 := \begin{bmatrix} 1 & 0 \\ (1-\sigma)|k|^2 & -2|k| \end{bmatrix}.$$

We thus obtain for DN_3 and DN_4 on the interface Γ the error iteration

$$\widetilde{\mathbf{g}}_1^{n+1} = [I - \Theta(I - \bar{T}_2\bar{A}_2^{-1}\bar{T}_1\bar{A}_1^{-1})]\widetilde{\mathbf{g}}_1^n, \tag{3.35}$$

for which we can prove the following convergence estimate.

THEOREM 3.7. *For a general relaxation matrix Θ as in* (3.26) *of Theorem* 3.5, *the convergence factor ρ of DN_3 and DN_4 is*

$$\rho = \max\{|1-\theta_{11}-\theta_{22}+\sqrt{(\theta_{11}-\theta_{22})^2+4\theta_{12}\theta_{21}}|, |1-\theta_{11}-\theta_{22}-\sqrt{(\theta_{11}-\theta_{22})^2+4\theta_{12}\theta_{21}}|\}.$$

The optimal choice of θ_{11}, θ_{12}, θ_{21}, and θ_{22} must satisfy the system of equations

$$\begin{aligned} \theta_{11}+\theta_{22} &= 1, \\ (\theta_{11}-\theta_{22})^2+4\theta_{12}\theta_{21} &= 0, \end{aligned} \tag{3.36}$$

and the corresponding convergence factor vanishes identically,

$$\rho_{DN_{3,4}} = 0.$$

In particular, the symmetric choice $\theta_{11} = \theta_{22} = \frac{1}{2}$ and $\theta_{12} = \theta_{21} = 0$ satisfies the optimality condition system (3.36) *and makes the convergence factor vanish.*

Proof. By a direct computation, we find that $\bar{T}_2\bar{A}_2^{-1}\bar{T}_1\bar{A}_1^{-1} = -I$, and therefore the interface iteration matrix of DN_3 and DN_4 becomes independent of $k \neq 0$,

$$I - \Theta(I - \bar{T}_2\bar{A}_2^{-1}\bar{T}_1\bar{A}_1^{-1}) = \begin{bmatrix} 1-2\theta_{11} & -2\theta_{12} \\ -2\theta_{21} & 1-2\theta_{22} \end{bmatrix}.$$

The corresponding eigenvalues are $\lambda_1 = 1-\theta_{11}-\theta_{22}+\sqrt{(\theta_{11}-\theta_{22})^2+4\theta_{12}\theta_{21}}$, and $\lambda_2 = 1-\theta_{11}-\theta_{22}-\sqrt{(\theta_{11}-\theta_{22})^2+4\theta_{12}\theta_{21}}$, and the result follows by equioscillation. □

Remark 3.8. We see that DN_3 and DN_4 for the biharmonic problem have for this symmetric configuration the same good convergence property as when the Dirichlet–Neumann algorithm is applied to Laplace's equation: we obtain a direct solver. Convergence is also independent of the frequency parameter k, and thus mesh independent, and in addition also independent of the problem parameter, in contrast to DN_2, which indicates that the Dirichlet–Neumann methods DN_3 and DN_4 are fundamentally better than DN_2, which we will next test numerically.

4. Numerical results.

We now present a numerical study of the Dirichlet–Neumann algorithms DN_j, $j = 2, 3, 4$, for the biharmonic equation in the rectangular domain $\Omega = (0,2)\times(0,1)$, i.e.,

$$\begin{aligned} \Delta^2 u = f &\quad \text{in } \Omega, \\ u = 0 &\quad \text{on } x = 0 \text{ and } x = 2, \\ \partial_n u = 0 &\quad \text{on } x = 0 \text{ and } x = 2, \\ u = 0 &\quad \text{on } y = 0 \text{ and } y = 1, \\ \Delta u - (1-\sigma)\partial_{\tau\tau} u = 0 &\quad \text{on } y = 0 \text{ and } y = 1, \end{aligned} \tag{4.1}$$

i.e., we impose the condition $\mathcal{D}_1(u) = \mathbf{0}$ on the left and right boundary, and the condition $\mathcal{D}_4(u) = \mathbf{0}$ on bottom and top boundary. We do not test DN_1 since one of the subproblems is not well posed in general (see the last paragraph before subsection 3.1). We choose for the right-hand side $f = 24y^3(1-y)^3 - 72x^2(2-x)^2(5y^2-5y+1) - 48(3x^2-6x+2)(5y^4-10y^3+6y^2-y)$, so that the solution is $u = (1+x)^2(1-x)^2y^3(1-y)^3$. We discretize (4.1) with the standard 13-point finite difference scheme (see, e.g., [17]). We then divide the domain into two subdomains Ω_1 and Ω_2 at $x = x_0$. We stop the Dirichlet–Neumann iterations when

$$\max_i \frac{\|u_i^n - u\|_{l^2}}{\|u\|_{l^2}} \leqslant 10^{-6}, \tag{4.2}$$

where u_i^n is the discrete subdomain solution on Ω_i at iteration n, and u is the discrete mono-domain solution which we compute by a direct solver. We also stop the iteration if the norm of the error vector becomes larger and larger in the first 50 iteration steps, and then we say it diverges, indicated by “div” in the tables below. We use a random initial guess to start the iteration; for the importance of this in testing, see [14, Figure 5.2].

4.1. Subdomains of the same size. We start with subdomains of the same size, the interface being at $x_0 = 1$, and thus work is perfectly load balanced. We test our Dirichlet–Neumann algorithms for the physical values of the parameter $\sigma = 0.4999$ (rubber), $\sigma = 0.225$ (glass), $\sigma = 0$ (cork), and also for the unphysical value $\sigma = -0.5$. In Table 1, we show a comparison of the iteration numbers needed by DN_j, $j = 2, 3, 4$, when the mesh size h is refined. We used the optimized choice $\theta_{12} = \theta_{21} = 0$ and $\theta_{11} = \theta_{22} = \frac{(1-\sigma)(\sigma+3)}{8}$ predicted by Theorem 3.1 for DN_2, and $\theta_{11} = \theta_{22} = \frac{1}{2}$ and $\theta_{12} = \theta_{21} = 0$ from Theorem 3.7 for DN_3 and DN_4. From Table 1, we see that while all three Dirichlet–Neumann algorithms have a mesh independent convergence behavior, DN_3 and DN_4 converge much faster than DN_2, which confirms that the new Dirichlet–Neumann algorithms DN_3 and DN_4 are much better solvers than DN_2, as expected from our analysis. We also show in parentheses the number of iterations when the algorithms are used as preconditioners for GMRES. We see that even when DN_2 is used as a preconditioner, which improves the iteration numbers, they still remain much larger than for DN_3 and DN_4, for which GMRES acceleration is clearly not needed in this symmetric case.

We next study numerically the behavior of the Dirichlet–Neumann algorithms for a fixed parameter σ and varying the relaxation matrix Θ. We first set $\theta_{12} = \theta_{21} = 0$ and vary the value $\theta_{11} = \theta_{22}$, which leads to the results shown in Table 2.

We see that the numerical results follow well the prediction in Theorem 3.1, with the best parameter choice close to $\theta_{11} = \theta_{22} = \frac{(1-\sigma)(\sigma+3)}{8}$. We see, however, also

Table 1

Iteration number comparison for the three Dirichlet–Neumann methods (in parentheses with GMRES acceleration) for different values of the parameter σ.

h		$\frac{1}{16}$	$\frac{1}{32}$	$\frac{1}{64}$	$\frac{1}{128}$	$\frac{1}{256}$	$\frac{1}{512}$	$\frac{1}{1024}$
DN_2	$\sigma = 0.4999$	111(14)	123(16)	126(16)	128(17)	128(17)	128(17)	129(17)
	$\sigma = 0.225$	52(14)	56(15)	57(15)	58(15)	57(16)	57(16)	57(17)
	$\sigma = 0$	34(13)	36(14)	37(14)	36(14)	37(15)	37(15)	37(15)
	$\sigma = -0.5$	17(10)	17(10)	17(10)	17(11)	17(11)	17(11)	17(11)
DN_3 and DN_4	$\sigma = 0.4999$	2(2)	2(2)	2(2)	2(2)	2(2)	2(2)	2(2)
	$\sigma = 0.225$	2(2)	2(2)	2(2)	2(2)	2(2)	2(2)	2(2)
	$\sigma = 0$	2(2)	2(2)	2(2)	2(2)	2(2)	2(2)	2(2)
	$\sigma = -0.5$	2(2)	2(2)	2(2)	2(2)	2(2)	2(2)	2(2)

Table 2

Iteration numbers for DN_2 (in parentheses with GMRES acceleration) with different choices of $\theta_{11} = \theta_{22} = \frac{(1-\sigma)(\sigma+3)}{m}$, where $\theta_{12} = \theta_{21} = 0$ ("div" means divergence).

m	2	4	8	10	12	16	32	64
$\sigma = 0.225$ $h = \frac{1}{256}$	div(16)	div(16)	57(16)	47(16)	58(16)	80(16)	167(16)	340(16)
$\sigma = 0$ $h = \frac{1}{128}$	div(14)	div(14)	36(14)	35(14)	43(14)	61(14)	129(14)	265(14)
$\sigma = -0.5$ $h = \frac{1}{512}$	div(11)	div(11)	17(11)	21(11)	27(11)	39(11)	86(11)	180(11)

Table 3

Iteration numbers for DN_2 (in parentheses with GMRES acceleration) with different choices of $\theta_{11} = \frac{(1-\sigma)(\sigma+3)}{n}$, $\theta_{22} = \frac{(1-\sigma)(\sigma+3)}{m}$, where $n = 8$ for the first table and $m = 8$ for the second table, and $\theta_{12} = \theta_{21} = 0$ ("div" means divergence).

m	2	4	8	12	16	32	64
$\sigma = 0.225$	div(16)	div(16)	57(16)	47(16)	58(16)	106(16)	203(15)
$\sigma = 0$	div(15)	div(15)	36(15)	35(15)	44(15)	83(15)	162(15)
$\sigma = -0.5$	div(10)	div(11)	17(11)	22(11)	29(11)	60(11)	122(10)

n	2	4	8	12	16	32	64
$\sigma = 0.225$	div(16)	div(16)	57(16)	48(16)	61(16)	116(16)	226(15)
$\sigma = 0$	div(15)	div(15)	36(15)	36(15)	47(15)	93(15)	185(14)
$\sigma = -0.5$	div(11)	div(11)	17(11)	23(11)	32(11)	69(11)	144(10)

that with GMRES, the parameter choice becomes much more robust. This is not difficult to understand: for an iteration of the form (3.5) corresponding to DN_2 used as an iterative solver, GMRES solves when using DN_2 as preconditioner systems with the system matrix of the form $\theta(I - T_2A_2^{-1}T_1A_1^{-1})$, and the coefficient θ thus does not affect the Krylov space; *this holds for any domain decomposition method using a scalar relaxation step* [24], *and the scalar relaxation step is irrelevant when using Krylov acceleration!*

We next vary $\theta_{11} \neq \theta_{22}$ to illustrate Theorem 3.2, which shows that for optimal performance of DN_2 with a diagonal relaxation matrix, the two parameters should be the same. The results are shown in Table 3, where we see that indeed the best performance is obtained when the two parameters are nearly the same. The slight difference is due to the fact that our analysis is based on unbounded domains, while our numerical simulation is using bounded domains, and we checked that it disappears when we run the simulation on the larger domain $\Omega = (0,4) \times (0,1)$ divided into two equal subdomains. We see also that when using DN_2 as a preconditioner then even with a diagonal relaxation matrix, not just a scalar, the iteration numbers depend only very little on the values chosen on the diagonal.

We finally vary now the off diagonal elements in the relaxation matrix Θ: we fix $\theta_{11} = \theta_{22} = \frac{(1-\sigma)(\sigma+3)}{8}$ and vary the parameters θ_{12} and θ_{21}. We show in Table 4 the results when we keep one of the two off diagonal parameters zero and vary the

Table 4

Iteration numbers for DN_2 with $\theta_{11} = \theta_{22} = \frac{(1-\sigma)(\sigma+3)}{8}$ and in the first table $\theta_{12} = \frac{h}{m}$ and $\theta_{21} = 0$, $h = \frac{1}{256}$ ($m = \infty$ means $\theta_{12} = 0$); in the second table we set $\theta_{12} = 0$ and $\theta_{21} = s$.

m	1	2	4	8	16	32	64	128	∞
$\sigma = 0.225$	div	div	62	56	56	57	57	57	57
$\sigma = 0$	div	89	36	36	36	37	36	36	36
$\sigma = -0.5$	div	46	19	17	17	17	17	17	17

s	4	2	1	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{32}$	$\frac{1}{64}$	$\frac{1}{128}$	0
$\sigma = 0.225$	div	299	28	30	32	38	46	51	54	55	57
$\sigma = 0$	div	71	20	22	22	28	32	34	35	36	36
$\sigma = -0.5$	111	30	17	12	12	15	16	17	17	17	17

Table 5

Iteration numbers for DN_3 and DN_4 with different choices of the entries in the relaxation matrix Θ. We set $\theta_{22} = 1 - \theta_{11}$, and $\theta_{12}\theta_{21} = -\frac{(1-2\theta_{11})^2}{4}$, and change θ_{11}, for $h = \frac{1}{256}$ and $\sigma = 0.225$.

θ_{11}	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
$\theta_{12} = -\theta_{21} = -\frac{(1-2\theta_{11})}{2}$	3	3	3	3	3	2	3	3	3	3	3
$-\theta_{12} = \theta_{21} = -\frac{(1-2\theta_{11})}{2}$	3	3	3	3	3	2	3	3	3	3	3
$\theta_{12} = 1 - 2\theta_{11}$, $\theta_{21} = -\frac{(1-2\theta_{11})}{4}$	3	3	3	3	3	2	3	3	3	3	3
$\theta_{12} = -\frac{(1-2\theta_{11})}{4}$, $\theta_{21} = 1 - 2\theta_{11}$	3	3	3	3	3	2	3	3	3	3	3

other. According to Remark 3.6, the order of θ_{12} should be $O(\frac{1}{|k|_{max}}) = O(h)$ and the order of θ_{21} should be $O(1)$, which is what we chose here. We see that when we keep $\theta_{21} = 0$, then numerically indeed the choice $\theta_{12} = 0$ works best as predicted by our unbounded domain analysis in Theorem 3.5. If we fix, however, $\theta_{12} = 0$ and vary θ_{21}, then there seems to be a small benefit for a nonzero θ_{21}, which is visibly not captured by our unbounded domain analysis.

In contrast to DN_2 the situation for DN_3 and DN_4 is much simpler: Table 5 illustrates that the diagonal choice from Theorem 3.7 is indeed the best one.

4.2. Numerical study of the influence of boundary conditions. Both our original problem (1.1) and the numerical model problem (4.1) contain boundary conditions, while we performed our analysis on unbounded domains. Our results for DN_3 and DN_4 on unbounded domains predicted well the performance of these algorithms in our numerical experiments on bounded domains. This is further illustrated in Table 6, which shows the iteration numbers for the original problem (1.1), and these iteration numbers are the same as the ones shown in Table 5 for problem (4.1), but for DN_2 we observed slight differences, which we now investigate further. If we choose for the material parameter $\sigma = 0.225, 0, -0.5$, and use the Dirichlet boundary condition $\mathcal{D}_1$ on top and bottom as in our original problem (1.1), we see in Figure 2 that there is a small but systematic difference between our theoretical results obtained for unbounded domains and numerical results measured on bounded domains, when σ diminishes. From some numerical tests of this difference, we estimate the numerically

Table 6

Iteration numbers for DN_3 and DN_4 with $\mathcal{D}_1$ on top and bottom boundary and different choices of the entries in the relaxation matrix Θ. We set $\theta_{22} = 1 - \theta_{11}$, and $\theta_{12}\theta_{21} = -\frac{(1-2\theta_{11})^2}{4}$, and change θ_{11}, for $h = \frac{1}{256}$ and $\sigma = 0.225$.

θ_{11}	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
$\theta_{12} = -\theta_{21} = -\frac{(1-2\theta_{11})}{2}$	3	3	3	3	3	2	3	3	3	3	3
$-\theta_{12} = \theta_{21} = -\frac{(1-2\theta_{11})}{2}$	3	3	3	3	3	2	3	3	3	3	3
$\theta_{12} = 1 - 2\theta_{11}$, $\theta_{21} = -\frac{(1-2\theta_{11})}{4}$	3	3	3	3	3	2	3	3	3	3	3
$\theta_{12} = -\frac{(1-2\theta_{11})}{4}$, $\theta_{21} = 1 - 2\theta_{11}$	3	3	3	3	3	2	3	3	3	3	3

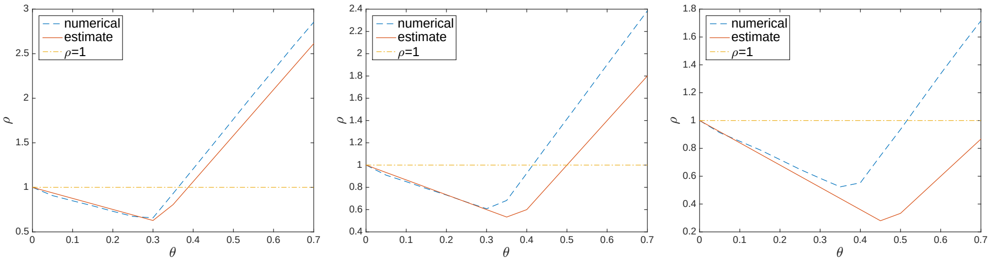

Fig. 2. *Theoretical and numerical contraction factors of DN_2 compared for scalar relaxation $\theta_{11} = \theta_{22} = \theta$ with Dirichlet boundary conditions $\mathcal{D}_1$ on top and bottom using* 50 *iterations for* $\sigma = 0.225, 0, -0.5$.

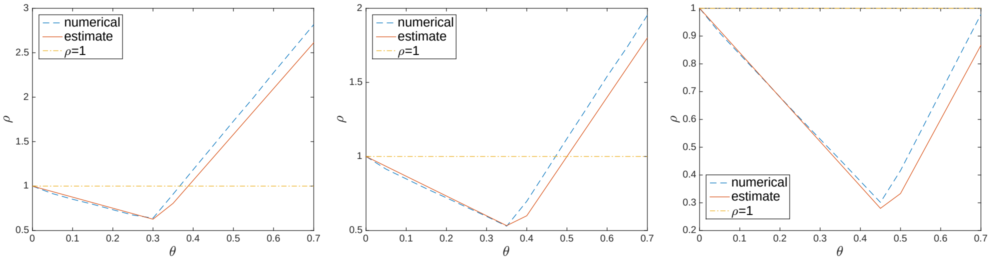

Fig. 3. *Results corresponding to Figure* 2 *but now with Dirichlet boundary conditions $\mathcal{D}_4$ on top and bottom.*

optimal θ on bounded domains to be $\frac{(1-\sigma)(\sigma+3)}{12}$ rather than $\frac{(1-\sigma)(\sigma+3)}{8}$. If we use, however, the Dirichlet condition $\mathcal{D}_4$ on top and bottom, for which the biharmonic operator is separable, and our Fourier analysis on the unbounded domain could thus also be performed on the bounded domain, we see in Figure 3 that indeed our theory from the unbounded domain analysis captures quite well the numerical behavior of the DN_2 method also on the bounded domain with the optimal θ close to $\frac{(1-\sigma)(\sigma+3)}{8}$.

4.3. Nonrectangular subdomains and subdomains of different size. We next test two cases with nonrectangular subdomains (see Figure 4). Case 1 is a parallelogram divided into two triangular subdomains, and Case 2 is a rectangle with a rectangular hole in it, which is divided into two symmetric subdomains. We use the Dirichlet condition $\mathcal{D}_1$ on the boundary. The results are shown in Table 7, and we see for both cases that DN_3 and DN_4 perform always better than DN_2 even in cases

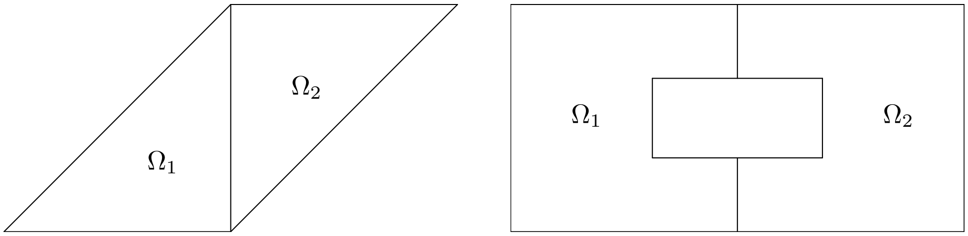

FIG. 4. *Two cases with nonrectangular subdomains. Left: Case* 1. *Right: Case* 2.

TABLE 7

Iteration numbers (in parentheses with GMRES acceleration) for different Dirichlet–Neumann methods with nonrectangular domain decomposition and mesh size h. First table: Case 1. *For convergence, we set* $\theta_{11} = \theta_{22} = \frac{(1-\sigma)(\sigma+3)}{32}$, $\theta_{12} = \theta_{21} = 0$, *and* $\sigma = 0.225$ *for the* DN_2 *method, and* $\theta_{11} = \theta_{22} = \frac{1}{4}$, $\theta_{12} = \theta_{21} = 0$ *for the* DN_3 *and* DN_4 *methods. Second table: Case* 2. *We set* $\theta_{11} = \theta_{22} = \frac{(1-\sigma)(\sigma+3)}{12}$, $\theta_{12} = \theta_{21} = 0$, *and* $\sigma = 0.225$ *for the* DN_2 *method, and* $\theta_{11} = \theta_{22} = \frac{1}{2}$, $\theta_{12} = \theta_{21} = 0$ *for the* DN_3 *and* DN_4 *methods.*

h	$\frac{1}{32}$	$\frac{1}{64}$	$\frac{1}{128}$	$\frac{1}{256}$	$\frac{1}{512}$	$\frac{1}{1024}$
DN_2	176(11)	184(12)	186(12)	187(13)	187(14)	188(15)
DN_3 and DN_4	34(6)	39(7)	42(8)	43(8)	45(9)	48(10)

h	$\frac{1}{32}$	$\frac{1}{64}$	$\frac{1}{128}$	$\frac{1}{256}$	$\frac{1}{512}$	$\frac{1}{1024}$
DN_2	57(10)	59(10)	60(11)	58(11)	58(10)	60(10)
DN_3 and DN_4	2(2)	2(2)	2(2)	2(2)	2(2)	2(2)

.

TABLE 8

Left: preconditioned GMRES iteration numbers for DN_2 *with different asymmetric domain decomposition and mesh size h. We set* $\theta_{11} = \theta_{22} = \frac{(1-\sigma)(\sigma+3)}{8}$, $\theta_{12} = \theta_{21} = 0$, *and* $\sigma = 0.225$. *Right: corresponding results for* DN_3 *and* DN_4 *with* $\theta_{11} = \theta_{22} = \frac{1}{2}$, $\theta_{12} = \theta_{21} = 0$.

h	$\frac{1}{32}$	$\frac{1}{64}$	$\frac{1}{128}$	$\frac{1}{256}$	$\frac{1}{512}$	$\frac{1}{1024}$	$\frac{1}{32}$	$\frac{1}{64}$	$\frac{1}{128}$	$\frac{1}{256}$	$\frac{1}{512}$	$\frac{1}{1024}$
$x_0 = \frac{1}{4}$	12	13	13	13	14	13	7	7	6	6	7	6
$x_0 = \frac{1}{2}$	11	11	12	12	12	13	5	5	5	5	5	4
$x_0 = 1$	10	10	11	11	11	11	2	2	2	2	2	2
$x_0 = \frac{3}{2}$	10	11	11	11	11	11	5	5	5	5	5	5
$x_0 = \frac{7}{4}$	12	12	13	13	13	13	7	7	6	7	6	5

not covered by our analysis. In addition, for Case 2, DN_3 and DN_4 is still optimal with convergence in two steps.

We also test cases which are not well load balanced, and thus less relevant in practice, to see if DN_2 could at least then have an advantage over DN_3 and DN_4. The results are shown in Table 8, and we see that load balancing is better for all the methods, but only DN_3 and DN_4 can take full advantage of this with fast convergence, while DN_2 cannot. In addition DN_3 and DN_4 perform always better than DN_2 even if the load is very unbalanced, so DN_2 should not be used for the biharmonic problem, since it never has any advantage over DN_3 and DN_4 and is much slower in the load balanced case.

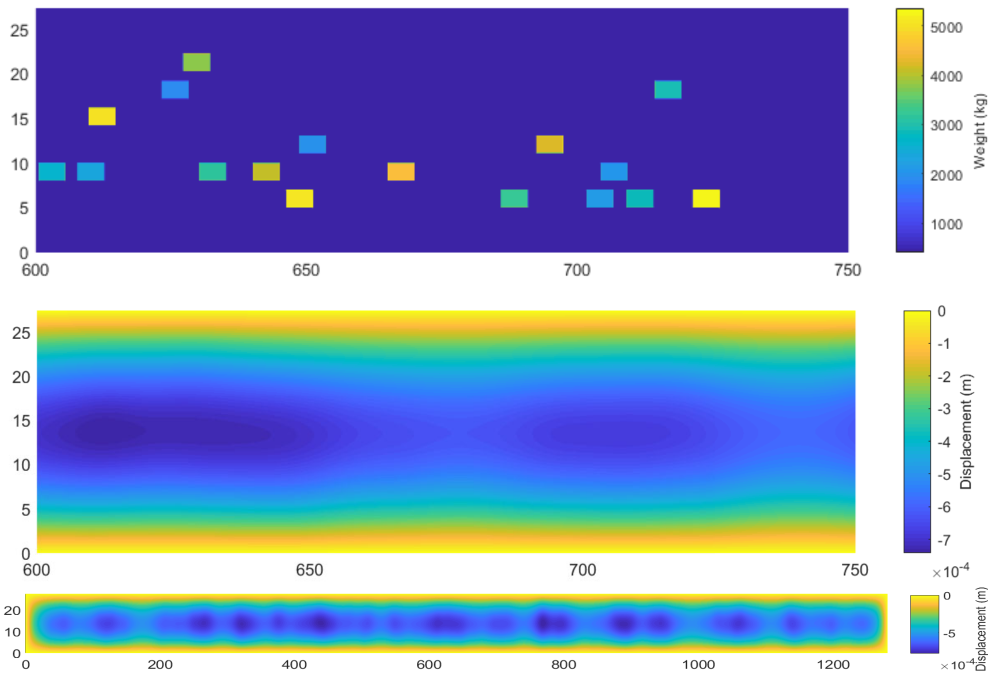

Fig. 5. *Top: illustration of the six-lane bridge interval* $[600, 750]$ *with different weights of cars, trucks, and vans randomly placed on it. Middle: corresponding displacement of the bridge under this load in the same interval* $[600, 750]$. *Bottom: displacement of the entire bridge under this load shown scaled* 4 *to* 1 *for better visibility.*

4.4. A many subdomain example. In this section, we show a numerical simulation for the Golden Gate Bridge with many subdomains. The Golden Gate Bridge is a typical suspension bridge supported by two towers. The entire bridge is about 2737 meters long, while the body between the two towers which we simulate is about 1280 meters. Since two sides of the body are supported by the towers and the two other sides are supported by the suspension cables, the boundary conditions for this example are $\mathcal{D}_4$. The width of bridge is 27.4 meters, and it is thus suitable to divide the domain into a sequence of 50 subdomains, where the size of each subdomain is 25.6×27.4. Because the bridge is made of steel, we set $\sigma = 0.28$. For the source terms of this problem, we randomly place cars, trucks, and vans on the bridge, as shown at the top of Figure 5.

We then simulate the displacement, which is the solution of our biharmonic problem. We see at the bottom of Figure 5 that the bridge is very stable under this load, with only a small displacement. We also use (4.2) as the iteration termination condition here. The iteration numbers (in parentheses with GMRES acceleration) used by our Dirichlet–Neumann algorithms to compute this solution are 44(29) for DN_2 and 7(6) for DN_3 and DN_4. This clearly shows that DN_3 and DN_4 are much better algorithms than DN_2 for solving the biharmonic problem also in this many subdomain application, and GMRES acceleration is not important for DN_3 and DN_4, similar to the two subdomain case.

5. Conclusions. We showed that in contrast to the classical Dirichlet–Neumann method, which was already well studied in the literature for the biharmonic problem, there are other quite natural Dirichlet–Neumann methods, and we proved that the convergence properties of two of them are very much superior for two subdomain decompositions and have the same convergence properties as Dirichlet–Neumann for Laplace's equation, while the classical one does not. We proved that there is indeed no relaxation matrix that could give the classical Dirichlet–Neumann method for the biharmonic problem such good convergence properties. It is therefore important when solving biharmonic problems with Dirichlet–Neumann methods to choose the good two traces for the Dirichlet step and the good remaining ones for the Neumann step; using clamped conditions for the Dirichlet step does not lead to a fast domain decomposition method, even though its convergence is still independent of the mesh size. Our work opens up a new direction for other domain decomposition methods for biharmonic problems, since they also use Dirichlet conditions, like the Schwarz methods, or Dirichlet and Neumann conditions, like the FETI and Neumann–Neumann methods. For all these cases, we expect the choice of what is Dirichlet and what is Neumann to have an important influence on the convergence behavior, and the classical clamped condition for Dirichlet seems to be less favorable, as we have already discovered for Schwarz methods in [20]. Corresponding results for optimized Schwarz methods and Neumann–Neumann and FETI methods will appear elsewhere.

Acknowledgment. We would like to thank two anonymous reviewers for their careful reading of our manuscript and their excellent suggestions on how to improve it.

REFERENCES

[1] Y. Achdou and F. Nataf, *A Robin-Robin preconditioner for an advection-diffusion problem*, C. R. Acad Sci., 325 (1996), pp. 1211–1216.

[2] P. E. Bjørstad and O. B. Widlund, *Iterative methods for the solution of elliptic problems on regions partitioned into substructures*, SIAM J. Numer. Anal., 23 (1986), pp. 1097–1120.

[3] S. Brenner, *A two-level additive Schwarz preconditioner for nonconforming plate elements*, Numer. Math., 72 (1996), pp. 419–447.

[4] P. Ciarlet, *The Finite Element Methods for Elliptic Problems*, North-Holland, Amsterdam, 1978.

[5] J. Côté, M. J. Gander, L. Laayouni, and S. Loisel, *Comparison of the Dirichlet-Neumann and optimal Schwarz method on the sphere*, in Domain Decomposition Methods in Science and Engineering XV, Lect. Notes. Comput. Sci. Eng., Springer-Verlag, Berlin, 2004, pp. 235–242.

[6] C. Dohrmann, *A preconditioner for substructuring based on constrained energy minimization*, SIAM J. Sci. Comput., 25 (2003), pp. 246–258.

[7] V. Dolean and M. J. Gander, *Multitrace formulations and Dirichlet-Neumann algorithms*, in Domain Decomposition Methods in Science and Engineering XXII, Lect. Notes. Comput. Sci. Eng., Springer-Verlag, Berlin, 2016, pp. 147–155.

[8] V. Dolean, F. Nataf, and G. Rapin, *How to use the Smith factorization for domain decomposition methods applied to the Stokes equations*, in Domain Decomposition Methods in Science and Engineering XVII, Lect. Notes. Comput. Sci. Eng., Springer-Verlag, Berlin, 2008, pp. 331–338.

[9] O. Dubois and M. J. Gander, *Optimized Schwarz methods for a diffusion problem with discontinuous coefficient*, Numer. Algorithms, 1 (2015), pp. 109–144.

[10] C. Farhat and J. Mandel, *The two-level FETI method for static and dynamic plate problems—Part* I: *An optimal iterative solver for biharmonic systems*, Comput. Methods Appl. Mech. Engrg., 155 (1998), pp. 129–152.

[11] X. Feng and O. Karakashian, *Two-level non-overlapping Schwarz preconditioners for a discontinuous Galerkin approximation of the biharmonic equation*, J. Sci. Comput., 22 (2005), pp. 289–314.

[12] X. Feng and T. Rahman, *An additive average Schwarz method for the plate bending problem*, J. Numer. Math., 10 (2002), pp. 109–125.

[13] M. J. Gander, *Optimized Schwarz Methods*, SIAM J. Sci. Comput., 44 (2006), pp. 699–731.

[14] M. J. Gander, *Schwarz methods over the course of time*, Electron. Trans. Numer. Anal., 31 (2008), pp. 228–255.

[15] M. J. Gander, L. Halpern, and F. Magoules, *An optimized Schwarz method with two-sided Robin transmission conditions for the Helmholtz equation*, Internat. J. Numer. Methods Fluids, 55 (2007), pp. 163–175.

[16] M. J. Gander, L. Halpern, and F. Nataf, *Optimized Schwarz methods*, in Proceedings of the 12th International Conference on Domain Decomposition Methods, T. Chan, T. Kako, H. Kawarada, and O. Pironneau, eds., Domain Decomposition Press, Chiba, Japan, 2001, pp. 15–18.

[17] M. J. Gander and F. Kwok, *Chladni figures and the Tacoma bridge: Motivating PDE eigenvalue problems via vibrating plates*, SIAM Rev., 54 (2012), pp. 573–596.

[18] M. J. Gander, F. Kwok, and B. Mandal, *Dirichlet-Neumann and Neumann-Neumann Waveform Relaxation for the Wave Equation*, in Domain Decomposition Methods in Science and Engineering XXII, Lect. Notes. Comput. Sci. Eng., Springer-Verlag, Berlin, 2016, pp. 501–509.

[19] M. J. Gander, F. Kwok, and B. Mandal, *Dirichlet-Neumann and Neumann-Neumann waveform relaxation algorithms for parabolic problems*, Electron. Trans. Numer. Anal., 45 (2016), pp. 424–456.

[20] M. J. Gander and Y. Liu, *On the definition of Dirichlet and Neumann conditions for the biharmonic equation and its impact on associated Schwarz methods*, in Domain Decomposition Methods in Science and Engineering XXIII, Lect. Notes. Comput. Sci. Eng., Springer-Verlag, Berlin, 2016, pp. 273–280.

[21] M. J. Gander and G. Wanner, *From Euler, Ritz and Galerkin to modern computing*, SIAM Rev., 54 (2013), pp. 627–666.

[22] L. Gerardo-Giorda, P. Le Tallec, and F. Nataf, *A Robin–Robin preconditioner for advection–diffusion equations with discontinuous coefficients*, Comput. Methods Appl. Mech. Engrg., 193 (2004), pp. 745–764.

[23] P. Gervasio, *Homogeneous and heterogeneous domain decomposition methods for plate bending problems*, Comput. Methods Appl. Mech. Engrg., 194 (2005), pp. 4321–4343.

[24] F. Kwok, *Dirichlet-Neumann and Neumann-Neumann Methods*, Summer School on DDM in Nice, University of Nice, Laboratoire Dieudonné, 2018.

[25] V. Leitão, *A meshless method for Kirchhoff plate bending problems*, Internat. J. Numer. Methods Engng., 52 (2001), pp. 1107–1130.

[26] J. Mandel and C. Dohrmann, *Convergence of a balancing domain decomposition by constraints and energy minimization*, Numer. Linear Algebra Appl., 10 (2003), pp. 639–659.

[27] J. Mandel, R. Tezaur, and C. Farhat, *A scalable substructuring method by Lagrange multipliers for plate bending problems*, SIAM J. Numer. Anal., 36 (1999), pp. 1370–1391.

[28] P. Peisker, *A multilevel algorithm for the biharmonic problem*, Numer. Math., 46 (1985), pp. 623–634.

[29] P. Peisker, W. Rust, and E. Stein, *Iterative solution methods for plate bending problems: Multigrid and preconditioned CG algorithm*, SIAM J. Numer. Anal., 27 (1990), pp. 1450–1465.

[30] W. Ritz, *Über eine neue Methode zur Lösung gewisser Variationsprobleme der mathematischen Physik*, J. für die reine und angewandte Mathematik (Crelle), 135 (1908), pp. 1–61.

[31] W. Ritz, *Theorie der Transversalschwingungen einer quadratischen Platte mit freien Rändern*, Ann. Phys., 18 (1909), pp. 737–807.

[32] P. Tallec, J. Mandel, and M. Vidrascu, *A Neumann–Neumann domain decomposition algorithm for solving plate and shell problems*, SIAM J. Numer. Anal., 35 (1998), pp. 836–867.

[33] P. Vaněk, J. Mandel, and M. Brezina, *Algebraic multigrid based on smoothed aggregation for second and fourth order problems*, Computing, 56 (1996), pp. 179–196.

[34] X. Zhang, *Multilevel Schwarz methods for the biharmonic Dirichlet problem*, SIAM J. Sci. Comput., 15 (1994), pp. 621–644.

[35] X. Zhang, *Two-level Schwarz methods for the biharmonic problem discretized by conforming $C1$ elements*, SIAM J. Numer. Anal., 33 (1996), pp. 555–570.