SIAM J. SCI. COMPUT.
Vol. 42, No. 5, pp. A3180–A3209

MULTILEVEL OPTIMIZED SCHWARZ METHODS*

MARTIN J. GANDER† AND TOMMASO VANZAN†

Abstract. We define a new two-level optimized Schwarz method (OSM), and we provide a convergence analysis both for overlapping and nonoverlapping decompositions. The two-level analysis suggests how to choose the optimized parameters. We also discuss an optimization procedure which relies only on the already studied one-level min-max problems, and we show that these two approaches are asymptotically equivalent. The two-level OSM has mesh independent convergence and it is scalable. We then generalize the two-level method defining a multilevel domain decomposition method which uses the OSM as a smoother. The main advantage of the method consists of its robustness and generality with respect to the equations under study. Thanks to the smoothing properties of the OSM, both with and without overlap, we can define a unique algorithm which can be applied to several equations, both with homogeneous and heterogeneous coefficients. We present extensive numerical results to compare the multilevel OSM, the one-level OSM, and the multigrid scheme. The experiments show that the multilevel OSM inherits robustness from the one-level OSM for heterogeneous elliptic problems, wave problems, and heterogeneous couplings. Finally, we apply the method to design a two-level solver for the heterogeneous Stokes–Darcy system.

Key words. multilevel domain decomposition methods, optimized Schwarz methods, heterogeneous two-level solver, two-level Helmholtz solver, two-level Stokes–Darcy solver

AMS subject classifications. 65N55, 65N22, 65F10, 65F08

DOI. 10.1137/19M1259389

1. Introduction.

Domain decomposition methods are effective techniques to solve large scale problems by subdividing the original problem into smaller tasks which can then be solved in parallel. Among the domain decomposition methods, the optimized Schwarz method (OSM) has been proposed as an improvement over the classical Schwarz method because its more general transmission conditions between the subdomains can be adapted to the physics, and thus they lead to very efficient solvers. The one-level OSM was first introduced in a series of papers [34, 32, 4] before the first systematic study in [17]. Optimized transmission conditions have been obtained for homogeneous problems, such as Helmholtz equations [21, 18], Maxwell equations [12, 11], the advection diffusion equations [27], and for many other situations. The study of heterogeneous problems has only recently started, and OSMs have so far shown very positive properties such as mesh independence and faster convergence in case of discontinuous coefficients [25, 12].

One-level domain decomposition methods are however, in general, not weakly scalable, i.e., their rate of convergence deteriorates when the number of subdomains grows [38], except in specific geometries [5, 3], and two-level methods are needed to achieve scalability. A coarse grid correction for OSM was first studied numerically in [14], where the authors proposed considering a coarse mesh defined by a single mesh point for each subdomain. Variants of this idea were also discussed in [15], where a convergence analysis is carried out. The main goal of these approaches is to make the subdomains communicate among themselves to obtain a convergence independent of the possibly large number of subdomains. However, coarse corrections can do much

*Submitted to the journal's Methods and Algorithms for Scientific Computing section May 1, 2019; accepted for publication (in revised form) July 22, 2020; published electronically October 13, 2020.

https://doi.org/10.1137/19M1259389

†Section de Mathématiques, Université de Genève, 2-4 rue du Lièvre, Genève, CH-1211, Switzerland (martin.gander@unige.ch, tommaso.vanzan@unige.ch).

more than just providing scalability; indeed, they can actually be chosen to obtain a direct solver, i.e., the iteration becomes nilpotent [20, 22, 19].

In this paper we follow the same philosophy. We analyze the one-level OSM and identify in which coarse space lies the optimal coarse correction which would lead to a nilpotent method. Computing this coarse correction would, however, be as expensive as solving directly the original problem. We therefore approximate this coarse space geometrically, solving on a coarse grid the equation for the coarse correction like in multigrid. This idea leads to the definition of a multilevel domain decomposition method where, on each level, we use the OSM as a smoother that we call the multilevel optimized Schwarz method (MOSM). Multilevel domain decomposition methods trace back to the 90s; an introduction is available in chapter 3 of [35]. In the seminal contributions [40, 2], the authors framed several preconditioners, among them multilevel preconditioners, within the so-called Schwarz abstract theory, and they provided condition number estimates for the resulting preconditioned systems in very general settings. In this paper, we adopt a different perspective. We focus on the properties of the iterative method, carrying out a Fourier analysis which, although under more restrictive hypotheses, permits us to have a complete description of the method through the derivation of an iteration matrix which acts on the Fourier modes. In defining the method, we stress the idea of using domain decomposition methods as smoothers inside a multigrid scheme. In addition, we focus on optimized Schwarz methods, and thus, our goal is to study the dependence of the methods on the optimized transmission conditions. This is why we rely on a Fourier analysis.

The idea of using domain decomposition methods as smoothers in a multigrid scheme is not new. Some remarks pointing in this direction are available, for instance, in section 15.3.3 of [30], section 3.4 of [35], and section 6.5 of [39], but a theoretical study is, to the best of the authors' knowledge, still missing. A similar approach has been recently developed in [6], where the authors propose multilevel domain decomposition methods which work exclusively on the interfaces between subdomains. We choose to analyze the OSM because it has a very good smoothing property: in case of overlap, it inherits the smoothing property from the classical Schwarz method, which converges exponentially fast for high frequencies. This property can even be enhanced by an adapted choice of the transmission conditions in the OSM. However, the potentiality of OSM is remarkable in the case without overlap, essential for heterogeneous problems, in which the classical Schwarz method simply would not work, while the transmission conditions in OSM allow us to tune at will the OSM as a smoother or as a rougher. Thus, even though for homogeneous problems the classical parallel Schwarz method would do the job, for heterogeneous problems only the OSM has the desired properties. Furthermore, we show that there is no need to develop a complete new theory for the optimized transmission conditions in a multilevel setting. Indeed, we show that one can just choose the optimized parameters using the already available literature for the one-level OSM by just changing the range of frequencies in the min-max problems in order to optimize the smoothing property of OSM. We show that these two approaches are asymptotically equivalent as $h \to 0$. We also prove mesh independent convergence for the two-level OSM, recovering the well-known properties of multigrid schemes; see Chapter 2 of [39]. This is a significant improvement over the one-level Schwarz methods, which have a mesh dependent convergence, and therefore, their convergence deteriorates as the number of unknowns increases [17].

The strength of the approach presented in this paper lies in its generality and flexibility. In fact, even though the development of efficient smoothers has reached a certain maturity in the multigrid literature (see, for instance, chapters 5–8 of [39]), one

may have to use ad hoc solutions according to the specific equation under study, which might be very sophisticated and difficult to implement. Our approach is instead very general since the smoother, being a domain decomposition method, does not change according to the equation, and it is straightforward to implement as long as one has a routine for the one-level domain decomposition method.

This paper is organized as follows: in section 2, we present the two-level OSM for a nonoverlapping decomposition, and in section 3 we propose a convergence analysis based on Fourier expansion. Our analysis, although carried out in a simplified setting as usually done in the OSM literature, is able to provide estimates for the optimized parameter to use on each level. Section 4 defines the method for overlapping decompositions and studies its convergence properties. In section 5, we generalize the two-level OSM to a multilevel framework, discussing implementation details on how to modify the residual while moving from one grid to another to assure we add the right correction on the fine grid. Our analysis is completed by extensive numerical results in section 6, whose aim is to underline the robustness of the MOSM with respect to heterogeneity for elliptic problems, the efficiency of this method for the Helmholtz equation, and its scalability. Finally, we propose a larger simulation example where we show how our method can be used to develop a two-level solver for the Stokes–Darcy coupling.

2. Two-level OSM for a nonoverlapping decomposition. In this section we introduce the two-level OSM for a nonoverlapping decomposition. We consider a second order elliptic PDE,

$$\mathcal{L}u = f \quad \text{on} \quad \Omega, \tag{2.1}$$

where Ω is divided into two subdomains Ω_1 and Ω_2, with Lipschitz boundary, separated by an interface $\Gamma := \partial\Omega_1 \cap \partial\Omega_2$. We suppose that (2.1) has a unique solution $u \in H^1(\Omega)$. Given two initial guesses u_1^0, u_2^0, the one-level parallel OSM reads for $n \geq 1$,

$$\begin{array}{llll} \mathcal{L}u_1^n = f & \text{in} \quad \Omega_1, & \partial_x u_1^n + p u_1^n = \partial_x u_2^{n-1} + p u_2^{n-1} & \text{on } \Gamma, \\ \mathcal{L}u_2^n = f & \text{in} \quad \Omega_2, & -\partial_x u_2^n + p u_2^n = -\partial_x u_1^{n-1} + p u_1^{n-1} & \text{on } \Gamma. \end{array} \tag{2.2}$$

If we define two functions on the interface as

$$r_1 := -\partial_x u_1^n - p u_1^n + \partial_x u_2^n + p u_2^n \quad \text{and} \quad r_2 := -\partial_x u_1^n + p u_1^n + \partial_x u_2^n - p u_2^n, \tag{2.3}$$

and then we solve the coupled system

$$\begin{array}{llll} \mathcal{L}e_1 = 0 & \text{in} \quad \Omega_1, & \partial_x e_1 + p e_1 - \partial_x e_2 - p e_2 = r_1 & \text{on } \Gamma, \\ \mathcal{L}e_2 = 0 & \text{in} \quad \Omega_2, & -\partial_x e_2 + p e_2 + \partial_x e_1 - p e_1 = r_2 & \text{on } \Gamma, \end{array} \tag{2.4}$$

we have that $\widetilde{u}_1 := u_1^n + e_1$ and $\widetilde{u}_2 := u_2^n + e_2$ are solutions of problem (2.1). Indeed, from (2.4) we have that $\partial_x(e_1 - e_2) = -\partial_x(u_1^n - u_2^n)$ and $(e_1 - e_2) = -(u_1^n - u_2^n)$. Thus

$$\partial_x \widetilde{u}_1 = \partial_x u_1^n + \partial_x e_1 = \partial_x u_1^n - \partial_x(u_1^n - u_2^n) + \partial_x e_2 = \partial_x u_2^n + \partial_x e_2 = \partial_x \widetilde{u}_2^n.$$

Similarly, we have that $\widetilde{u}_1 = \widetilde{u}_2$ on Γ, and hence $\widetilde{u}_1, \widetilde{u}_2$ satisfy the PDE in the interior subdomains, the conditions on the continuity of the normal derivatives, and the Dirichlet traces on the interface Γ. Clearly, at the continuous level we have e such that $e_{|\Omega_j} = e_j$ lies in the complete infinite dimensional coarse space [22]

$$\mathcal{A} := \left\{ v \in H^{1,\text{disc}}(\Omega) : \mathcal{L}(v_{|\Omega_j}) = 0, j = 1, 2 \right\},$$

with $H^{1,\text{disc}}(\Omega) := \left\{v \in L^2(\Omega) : v_{|\Omega_j} \in H^1(\Omega_j), j = 1, 2\right\}$. With this observation, it has been proposed to construct a discrete coarse space $\mathcal{V}_h \subset \mathcal{A}$, a restriction matrix $\mathcal{R}_c$ to the coarse space, and to solve the linear system

$$(\mathcal{R}_c A \mathcal{R}_c^\top)^{-1} \mathcal{R}_c \mathbf{e} = \mathcal{R}_c(\mathbf{f} - A\mathbf{u}^n).$$

Following this strategy, it is possible to define a complete discrete coarse space $\mathcal{A}_h$ which leads to a direct method. However, the complete coarse space is too expensive to use and, therefore, it is usually approximated obtaining optimized coarse spaces which are subspaces of $\mathcal{A}_h$; see [22, 19, 20] for an overview.

In this paper we define a two-level OSM inspired by the multigrid method, which solves (2.4) on a coarse mesh. Our two-level method can be summarized as follows: we iterate algorithm (2.2) on a fine grid, define after n_1 iterations the functions r_1 and r_2, restrict them on a coarse grid where we solve directly system (2.4), interpolate the corrections e_1 and e_2 on the fine grid, and add them to the iterates.

To analyze the discrete version of this algorithm we set $\Omega := (-\frac{1}{2}, \frac{1}{2}) \times (0, 1)$, $\Omega_1 := (-\frac{1}{2}, 0) \times (0, 1)$, $\Omega_2 := (0, \frac{1}{2}) \times (0, 1)$, with an interface $\Gamma := \{0\} \times [0, 1]$. We discretize (2.2) on a fine mesh with mesh size $h := \frac{1}{2^\ell}$, and (2.4) on a coarser mesh with mesh size $H := \frac{1}{2^{\ell-1}}$. Thus the fine mesh has $N_y := 2^\ell - 1$ degrees of freedom in the y direction and $N_x := \frac{N_y+1}{2}$ in the x direction for each subdomain, while the coarser mesh has $N_y^c := \frac{N_y+1}{2} - 1$ and $N_x^c := \frac{N_y^c+1}{2}$. Therefore, each subdomain has $N := N_y N_x$ degrees of freedom on the fine mesh and $N^c := N_y^c N_x^c$ on the coarse one. In the following we use the index ℓ to indicate which mesh we are considering. In [36] the authors introduced the augmented system $\widetilde{A}_\ell \widetilde{\mathbf{u}}_\ell = \widetilde{\mathbf{f}}_\ell$, which twice contains the variables at the interface Γ, with

$$\widetilde{A}_\ell = \begin{pmatrix} A_{1,\ell} & -B_{12,\ell} \\ -B_{21,\ell} & A_{2,\ell} \end{pmatrix} \in \mathbb{R}^{2N,2N},$$

where $A_{j,\ell} \in \mathbb{R}^{N,N}$ is the discrete Laplacian in the domain Ω_j with Robin boundary conditions on Γ, $B_{ji,\ell}$ are interface operators, $\widetilde{\mathbf{f}}_\ell = [\mathbf{f}_{1,\ell}, \mathbf{f}_{2,\ell}]$ is the force vector, and $\widetilde{\mathbf{u}}_\ell = [\mathbf{u}_{1,\ell}, \mathbf{u}_{2,\ell}] \in \mathbb{R}^{2N}$ is the vector of the degrees of freedom on the mesh indexed by ℓ. They showed that the discrete version of (2.2), i.e.,

$$A_{j,\ell}\mathbf{u}_{j,\ell}^{n+1} = \mathbf{f}_{j,\ell} + \sum_{k \neq j} B_{jk,\ell}\mathbf{u}_{k,\ell}^n, \quad j = 1, 2,$$

is equivalent to the algebraic iterative method in the correction form

$$\widetilde{\mathbf{u}}_\ell^{n+1} = \mathcal{S}_\ell(\widetilde{A}_\ell, \widetilde{\mathbf{u}}_\ell^{n-1}, \widetilde{f}_\ell) := \widetilde{\mathbf{u}}^n + \sum_{j=1}^{2} R_{j,\ell}^T A_{j,\ell}^{-1} R_{j,\ell}(\widetilde{\mathbf{f}}_\ell - \widetilde{A}_\ell \widetilde{\mathbf{u}}_\ell^n), \tag{2.5}$$

where $R_{j,\ell} \in \mathbb{R}^{N,2N}$ are restriction operators on the domain Ω_j. The iterative method (2.5) can also be written in the standard form as

$$\widetilde{\mathbf{u}}_\ell^{n+1} = \mathcal{S}_\ell(\widetilde{A}_\ell, \widetilde{\mathbf{u}}_\ell^{n-1}, \widetilde{f}_\ell) = \begin{pmatrix} 0 & A_{1,\ell}^{-1} B_{12,\ell} \\ A_{2,\ell}^{-1} B_{21,\ell} & 0 \end{pmatrix} \widetilde{\mathbf{u}}_\ell^n + \begin{pmatrix} 0 & A_{1,\ell}^{-1} \\ A_{2,\ell}^{-1} & 0 \end{pmatrix} (\widetilde{\mathbf{f}}_\ell).$$

We emphasize that the residual $\mathbf{r}_\ell^n = \widetilde{\mathbf{f}}_\ell - \widetilde{A}_\ell \widetilde{\mathbf{u}}_\ell^n$ is a vector with zero entries except for the degrees of freedom associated to the interface, where it represents a discretization

of the functions in (2.3). Thus, we construct the restriction operator $\mathcal{R}_\ell$ so that it acts as the full weighting restriction operator $\mathcal{R}_{1D} \in \mathbb{R}^{N_y^c, N_y}$ for the points on the interface, and it has zero blocks corresponding to the interior degrees of freedom,

$$\mathcal{R}_\ell \mathbf{r}_\ell = \begin{pmatrix} 0 & & & \\ & \mathcal{R}_{1D} & & \\ & & \mathcal{R}_{1D} & \\ & & & 0 \end{pmatrix} \begin{pmatrix} 0 \\ \mathbf{r}_{1\Gamma,\ell} \\ \mathbf{r}_{2\Gamma,\ell} \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ \mathcal{R}_{1D}\mathbf{r}_{1\Gamma,\ell} \\ \mathcal{R}_{1D}\mathbf{r}_{2\Gamma,\ell} \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ \mathbf{r}_{1\Gamma,\ell-1} \\ \mathbf{r}_{2\Gamma,\ell-1} \\ 0 \end{pmatrix} = \mathbf{r}_{\ell-1}.$$

Other possible choices are available: one could replace the zero blocks with two-dimensional (2D) full weighting restriction operators $\mathcal{R}_{2D}$, or with straight injection operators. This change would not affect the method since they all map a zero function on the fine mesh to a zero function on the coarse mesh. Therefore, the properties of the restriction operators are uniquely defined once we characterized the action of the restriction operator on the interface. This is an advantage of the two-level OSM, and of a large class of two-level domain decomposition methods: they do not require one to restrict on the whole volume but only on the interfaces, which are one-dimensional (1D) curves for 2D problems, or 2D surfaces for three dimensional (3D) problems. We refer the reader to [6] for a new framework of two-level and multilevel domain decomposition methods defined directly on the interfaces. On the coarse mesh, we solve the restricted residual equation inverting the operator $\widetilde{A}_{\ell-1}$, which corresponds to a direct discretization of the original problem on the mesh indexed by $\ell-1$. Finally, concerning the interpolation operator, we define $\mathcal{I}_\ell = \text{diag}(\mathcal{I}_{2D,\ell}, \mathcal{I}_{2D,\ell})$, with $\mathcal{I}_{2D,\ell} \in \mathbb{R}^{N,N^c}$ being the standard linear interpolation operator from the coarse to the fine grid. Another possible choice is to define $\mathcal{I}_\ell^A$, which interpolates on the interface and then extends harmonically on the fine grid, i.e., $\mathcal{I}_\ell^A := \text{diag}(A_{1,\ell}^{-1}, A_{2,\ell}^{-1}) 2\mathcal{R}_\ell^\top$. With all of these ingredients, the algorithm we have described previously at the continuous level to solve the continuous problem (2.1) can be rewritten in the discrete form as follows.

Algorithm 2.1 Function two-level OSM($\widetilde{A}_\ell, \widetilde{\mathbf{u}}_\ell^0, \widetilde{\mathbf{f}}_\ell$).

- For $n = 1 : n_1$, $\widetilde{\mathbf{u}}_\ell^n \leftarrow \mathcal{S}_\ell(\widetilde{A}_\ell, \widetilde{\mathbf{u}}_\ell^{n-1}, \widetilde{\mathbf{f}}_\ell)$.
- $\mathbf{r}_\ell \leftarrow \widetilde{\mathbf{f}}_\ell - \widetilde{A}_\ell \widetilde{\mathbf{u}}_\ell^{n_1}$.
- $\mathbf{r}_{\ell-1} \leftarrow \mathcal{R}_\ell \mathbf{r}_\ell$.
- $\widetilde{\mathbf{e}}_{\ell-1} \leftarrow \widetilde{A}_{\ell-1}^{-1} \mathbf{r}_{\ell-1}$.
- $\widetilde{\mathbf{u}}_\ell^{n_1} \leftarrow \widetilde{\mathbf{u}}_\ell^{n_1} + \mathcal{I}_\ell \widetilde{\mathbf{e}}_{\ell-1}$.
- For $n = n_1 + 1 : n_2$, $\widetilde{\mathbf{u}}_\ell^n \leftarrow \mathcal{S}_\ell(\widetilde{A}_\ell, \widetilde{\mathbf{u}}_\ell^{n-1}, \widetilde{\mathbf{f}}_\ell)$.
- Return $\widetilde{\mathbf{u}}_\ell^{n_2}$.

Considering the error equation, i.e., $\widetilde{f}_\ell = 0$, Algorithm 2.1 leads to the classical iteration matrix of a two-level method,

$$\mathcal{S}_{2LOSM} \widetilde{\mathbf{u}}_\ell = \mathcal{S}_\ell^{n_2} \left(I - \mathcal{I}_\ell \widetilde{A}_{\ell-1} \mathcal{R}_\ell \widetilde{A}_\ell \right) \mathcal{S}_\ell^{n_1} \widetilde{\mathbf{u}}_\ell, \tag{2.6}$$

where $I \in \mathbb{R}^{2N,2N}$ is the identity matrix, and $\mathcal{S}_\ell^n$ corresponds to n iterations of the iterative method defined in (2.5). The convergence properties of the two-level method depend on an effective team play of the smoother and coarse correction. As in the multigrid literature, n_1 and n_2 are mainly chosen heuristically. The stronger the smoothing property of the smoother, the smaller n_1 and n_2 can be. We show in

sections 3.2 and 4.1 that OSMs are very efficient smoothers, so that common choices for n_1 and n_2 are 1 or 2. In our numerical experiments, we set $n_1 = n_2 = 2$.

Remark 2.1. We emphasize that the exact correction given by $\widetilde{\mathbf{e}}_\ell = \widetilde{A}_\ell^{-1}\mathbf{r}_\ell$ is a discrete harmonic function (in the PDE sense) in the interior of the subdomains. The coarse correction, being $\widetilde{\mathbf{e}}_{\ell-1} = \widetilde{A}_{\ell-1}^{-1}\mathcal{R}_\ell\mathbf{r}_\ell$, is still harmonic on the coarse mesh, but the interpolated correction $\mathcal{I}_\ell\widetilde{\mathbf{e}}_{\ell-1}$ is not harmonic on the fine grid. In other words, the linear interpolator destroys the harmonicity of the correction, and thus we conclude that with the linear interpolator $\mathcal{I}_\ell$ we cannot have a direct method! The interpolator $\mathcal{I}_\ell^A$ should therefore be preferred since it adds a correction which lies in the complete discrete coarse space. However, its use is more expensive since it requires solving subdomain problems. In the rest of this paper, we will always use the geometric interpolator $\mathcal{I}_\ell$ if not explicitly stated otherwise.

3. Convergence analysis for the two-level OSM. Our analysis is based on a semidiscrete study of Algorithm 2.1. We take into account the mesh properties in the y direction, while we consider a continuous problem in the x direction. For the sake of clarity, we carry out the calculation supposing that $\mathcal{L} = -\Delta$, but in Remark 3.1 we discuss how the analysis adapts to general second order operators. Furthermore, we assume n_1 and n_2 to be even numbers. Accordingly, the error equation of (2.1) can be written as

$$-\partial_{xx}u - \partial_{yy,h}u = 0 \quad \text{on} \quad \Omega_j,\ j = 1, 2,$$

where, using separation of variables, u is semidiscrete as well, i.e., $u = \phi(x)\psi(jh)$, where $j = 1, \dots, N_y$. Inserting this ansatz we obtain the eigenvalue equation in the y direction,

$$-\partial_{yy,h}\psi(jh) = \gamma^2\psi(jh),$$

whose solutions are given by $\psi_k(jh) := \sin(k\pi jh)$, $j = 1, \dots, N_y$, $k = 1, \dots, N_y$, and $\gamma^2(k) := \frac{4}{h^2}\sin^2(k\pi\frac{h}{2})$. Solving the equation in x we obtain $\phi_k(x) = A(k)e^{\lambda(k)x} + B(k)e^{-\lambda(k)x}$, with $\lambda(k) = \sqrt{\gamma^2(k)}$. To simplify the problem further, we suppose that the domain is unbounded in the x direction so that the general solution is given by

$$u_1 = \sum_{k=1}^{N_y} A(k)\psi_k e^{\lambda(k)x} \quad \text{and} \quad u_2 = \sum_{k=1}^{N_y} B(k)\psi_k e^{-\lambda(k)x}. \tag{3.1}$$

The initial guesses u_1^0 and u_2^0 can be written in the general form of (3.1) for a proper choice of $A(k)$ and $B(k)$. After an even number n_1 of iterations of the smoother, standard computations (see, for instance, [17]) show that

$$u_1^{n_1} = \sum_{k=1}^{N_y} \rho(k,p)^{n_1} A(k)\psi_k e^{\lambda(k)x} \quad \text{and} \quad u_2^{n_1} = \sum_{k=1}^{N_y} \rho(k,p)^{n_1} B(k)\psi_k e^{-\lambda(k)x},$$

where $\rho(k,p) = \left(\frac{\lambda(k)-p}{\lambda(k)+p}\right)$. If n_1 is not even, then the role of $A(k)$ and $B(k)$ is flipped:

$$u_1^{n_1} = \sum_{k=1}^{N_y} \rho(k,p)^{n_1} B(k)\psi_k e^{\lambda(k)x} \quad \text{and} \quad u_2^{n_1} = \sum_{k=1}^{N_y} \rho(k,p)^{n_1} A(k)\psi_k e^{-\lambda(k)x}.$$

The analysis follows the same calculations, but the notation to keep track of both cases becomes cumbersome. Thus, we prefer to assume that n_1 is even for the sake

of clarity. We compute now the residuals r_1 and r_2 in (2.3),

$$\begin{aligned} r_1 &= \textstyle\sum_{k=1}^{N_y} g_-(k)A(k)\rho(k,p)^{n_1}\psi_k + \sum_{k=1}^{N_y} g_+(k)B(k)\rho(k,p)^{n_1}\psi_k, \\ r_2 &= \textstyle\sum_{k=1}^{N_y} g_+(k)A(k)\rho(k,p)^{n_1}\psi_k + \sum_{k=1}^{N_y} g_-(k)B(k)\rho(k,p)^{n_1}\psi_k, \end{aligned} \tag{3.2}$$

where $g_-(k) := -\lambda(k) - p$ and $g_+(k) := -\lambda(k) + p$. We observe that r_1 and r_2 are 1D functions in the variable y, which are sums of the eigenfunctions of the discrete Laplacian. Well known results are available for the action of the full weighted restriction operator $\mathcal{R}_{1D}$ and the linear interpolation operator $\mathcal{I}_{1D} := 2\mathcal{R}_{1D}^{\top}$ on these functions; see, for instance, Chapter 2 of [30]. In particular, defining $\widetilde{k} := N_y + 1 - k$, we have

$$\mathcal{R}_{1D}(e_k\psi_k + e_{\widetilde{k}}\psi_{\widetilde{k}}) = (e_k c_k^2 - e_{\widetilde{k}} s_k^2)\phi_k, \tag{3.3}$$

where $c_k := \cos(k\pi\frac{h}{2})$, $s_k := \sin(k\pi\frac{h}{2})$, and $\phi_{k,j} := \sin(k\pi jH)$ with $k, j \in \mathcal{V} := \{1, 2, \ldots, N_y^c\}$ are the eigenvectors of the 1D discrete Laplacian on the coarse grid. The eigenfunction $\psi_{\frac{N_y+1}{2}}$ is actually mapped to zero by the restriction operator, i.e., $\mathcal{R}_{1D}\psi_{\frac{N_y+1}{2}} = 0$, and thus this frequency is not represented on the coarse level. Using these results we obtain

$$\begin{aligned} \mathcal{R}_{1D}r_1 &= \sum_{k=1}^{N_y^c} \phi_k \left[\rho(k)^{n_1}\left(g_-(k)A(k) + g_+(k)B(k)\right)c_k^2 - \rho(\widetilde{k})^{n_1}\left(g_-(\widetilde{k})A(\widetilde{k}) + g_+(\widetilde{k})B(\widetilde{k})\right)s_k^2\right], \\ \mathcal{R}_{1D}r_2 &= \sum_{k=1}^{N_y^c} \phi_k \left[\rho(k)^{n_1}\left(g_+(k)A(k) + g_-(k)B(k)\right)c_k^2 - \rho(\widetilde{k})^{n_1}\left(g_+(\widetilde{k})A(\widetilde{k}) + g_-(\widetilde{k})B(\widetilde{k})\right)s_k^2\right], \end{aligned} \tag{3.4}$$

where for the sake of brevity we omit the dependence of $\rho(k,p)$ on p. On the coarse mesh the general solution of the semidiscrete Laplace equation is again given by a formula similar to (3.1),

$$e_1 = \sum_{k=1}^{N_y^c} \overline{A}(k)\phi_k e^{\lambda_c(k)x} \quad \text{and} \quad e_2 = \sum_{k=1}^{N_y^c} \overline{B}(k)\phi_k e^{-\lambda_c(k)x}, \tag{3.5}$$

where $\lambda_c^2(k) := \frac{4}{H^2}\sin^2(k\pi\frac{H}{2})$ are the eigenvalues of the 1D Laplacian on the coarse mesh. Imposing the boundary conditions to solve the residual system (2.4), we obtain

$$\begin{aligned} (\lambda_c(k) + p)\bar{A}(k) + (\lambda_c(k) - p)\bar{B}(k) &= \mathcal{R}_{1D}r_1(k), \\ (\lambda_c(k) - p)\bar{A}(k) + (\lambda_c(k) + p)\bar{B}(k) &= \mathcal{R}_{1D}r_2(k), \end{aligned} \tag{3.6}$$

which leads to

$$\begin{aligned} \bar{A}(k) &= \frac{\mathcal{R}_{1D}r_1(k) + \mathcal{R}_{1D}r_2(k)}{4\lambda_c(k)} + \frac{\mathcal{R}_{1D}r_1(k) - \mathcal{R}_{1D}r_2(k)}{4p}, \\ \bar{B}(k) &= \frac{\mathcal{R}_{1D}r_1(k) + \mathcal{R}_{1D}r_2(k)}{4\lambda_c(k)} + \frac{\mathcal{R}_{1D}r_2(k) - \mathcal{R}_{1D}r_1(k)}{4p}. \end{aligned}$$

The last step is to interpolate the correction to the fine grid. Since we deal with a semidiscrete analysis, we can use the results on the interpolation of the eigenvectors of the Laplace operator [30]. In particular, we have that $\forall k \in \mathcal{V}$, $\mathcal{I}_{1D}\phi_k = c_k^2\psi_k - s_k^2\psi_{\widetilde{k}}$. It follows that

$$\begin{aligned} u_1^{n_1} + \mathcal{I}_\ell e_1 &= \textstyle\sum_{k=1}^{N_y} \left(\rho(k,p)^{n_1}A(k) + d_k^2\bar{A}(k)\right)e^{\lambda(k)x}\psi_k, \\ u_2^{n_1} + \mathcal{I}_\ell e_2 &= \textstyle\sum_{k=1}^{N_y} \left(\rho(k,p)^{n_1}B(k) + d_k^2\bar{B}(k)\right)e^{-\lambda(k)x}\psi_k, \end{aligned}$$

where $d_k^2 = c_k^2$ if $k \leq N_y^c$, $d_k^2 = -s_k^2$ if $k \geq N_y^c + 2$, and $d_k^2 = 0$ if $k = N_y^c + 1$. Algebraic calculations allow us to write a linear relation which maps the coefficients $A(k), B(k), A(\widetilde{k}), B(\widetilde{k})$ after one step of this two-level method. Denoting $\rho = \rho(k,p)$, $\widetilde{\rho} = \rho(\widetilde{k},p)$, and $\mathbf{v}_k^n = \left(A^n(k), B^n(k), A^n(\widetilde{k}), B^n(\widetilde{k})\right)^\top$, we obtain

$$\mathbf{v}_k^n = G_k^{n_2} \widetilde{D}_k G_k^{n_1} \mathbf{v}_k^{n-1} \quad \forall k \in \mathcal{V}, \tag{3.7}$$

where

$$\widetilde{D}_k := \begin{pmatrix} \left(1-\frac{c_k^4}{2}\left(1+\frac{\lambda(k)}{\lambda_c(k)}\right)\right) & \frac{c_k^4}{2}\left(1-\frac{\lambda(k)}{\lambda_c(k)}\right) & \frac{c_k^2 s_k^2}{2}\left(1+\frac{\lambda(\widetilde{k})}{\lambda_c(k)}\right) & \frac{c_k^2 s_k^2}{2}\left(\frac{\lambda(\widetilde{k})}{\lambda_c(k)}-1\right) \\ \frac{c_k^4}{2}\left(1-\frac{\lambda(k)}{\lambda_c(k)}\right) & \left(1-\frac{c_k^4}{2}\left(1+\frac{\lambda(k)}{\lambda_c(k)}\right)\right) & \frac{c_k^2 s_k^2}{2}\left(\frac{\lambda(\widetilde{k})}{\lambda_c(k)}-1\right) & \frac{c_k^2 s_k^2}{2}\left(1+\frac{\lambda(\widetilde{k})}{\lambda_c(k)}\right) \\ \frac{c_k^2 s_k^2}{2}\left(1+\frac{\lambda(k)}{\lambda_c(k)}\right) & \frac{c_k^2 s_k^2}{2}\left(\frac{\lambda(k)}{\lambda_c(k)}-1\right) & \left(1-\frac{s_k^4}{2}\left(1+\frac{\lambda(\widetilde{k})}{\lambda_c(k)}\right)\right) & \frac{s_k^4}{2}\left(1-\frac{\lambda(\widetilde{k})}{\lambda_c(k)}\right) \\ \frac{c_k^2 s_k^2}{2}\left(\frac{\lambda(k)}{\lambda_c(k)}-1\right) & \frac{c_k^2 s_k^2}{2}\left(1+\frac{\lambda(k)}{\lambda_c(k)}\right) & \frac{s_k^4}{2}\left(1-\frac{\lambda(\widetilde{k})}{\lambda_c(k)}\right) & \left(1-\frac{s_k^4}{2}\left(1+\frac{\lambda(\widetilde{k})}{\lambda_c(k)}\right)\right) \end{pmatrix}, \tag{3.8}$$

$$G_k^n := \begin{pmatrix} \rho(k)^n & & & \\ & \rho(k)^n & & \\ & & \rho(\widetilde{k})^n & \\ & & & \rho(\widetilde{k})^n \end{pmatrix}. \tag{3.9}$$

The action of the smoother is described by the matrix G_k^n, while $\widetilde{D}_k$ takes into account the coarse correction. Denoting $\mathbf{e}^n = \left(\mathbf{v}_1^n, \ldots, \mathbf{v}_{N_y^c}^n, A\left(\frac{N_y+1}{2}\right), B\left(\frac{N_y+1}{2}\right)\right)^\top$, we conclude that $\mathbf{e}^n = T\mathbf{e}^{n-1}$, where

$$T = \begin{pmatrix} G_1^{n_2}\widetilde{D}_1 G_1^{n_1} & & & & \\ & \ddots & & & \\ & & G_{N_c}^{n_2}\widetilde{D}_{N_c} G_{N_c}^{n_1} & & \\ & & & \rho\left(\frac{N_y+1}{2},p\right)^{n_1+n_2} & \\ & & & & \rho\left(\frac{N_y+1}{2},p\right)^{n_1+n_2} \end{pmatrix}. \tag{3.10}$$

Remark 3.1. [extension to more general differential operators] Equation (3.10) has been obtained supposing $\mathcal{L} = -\Delta$, but it can be readily extended to more general operators. The necessary hypothesis for the calculations are the assumptions on the geometry of the problem, on the use of a uniform mesh along the interface, and that $\psi_k(jh) = \sin(k\pi jh)$, so that we can characterize the action of the restriction and prolongation operators. As long as these assumptions are verified, one can consider a general equation $(-\nu\Delta + a_1\partial_x + a_2\partial_y + c)u = 0$. If $a_1 = a_2 = 0$ and $c \neq 0$, then (3.10) is still valid, replacing $\lambda(k)$ with $\lambda(k) = \sqrt{\gamma(k)^2 + \frac{c}{\nu}}$ and using the corresponding convergence factor [17]. If only $a_2 = 0$, then using the expansions

$$u_1 = \sum_{k=1}^{N_y} A(k)\psi_k e^{\lambda_+(k)x} \quad \text{and} \quad u_2 = \sum_{k=1}^{N_y} B(k)\psi_k e^{-\lambda_-(k)x}, \tag{3.11}$$

with $\lambda(k)_{+,-} = \frac{a_1 \pm \sqrt{a_1^2 + 4\nu(\frac{4}{h^2}\sin^2(k\pi\frac{h}{2})) + 4\nu c}}{2\nu}$, and carrying out the same calculations, one can derive a similar iteration matrix. The case $a_2 \neq 0$ cannot be treated in this framework because it leads to $\psi_k(jh) \neq \sin(k\pi jh)$; see section 3.3 of [25] for more details on tangential advection.

3.1. Optimization of the semidiscrete nonoverlapping two-level OSM. To optimize the parameter of the two-level method one would have to solve the minimization problem $\min_p \rho(T)$, where $\rho(T)$ is the spectral radius of the matrix T which, being block diagonal, is the maximum of the spectral radii of the matrices $G_k^{n_2}\widetilde{D}_k G_k^{n_1}$ and $\rho\left(\frac{N+1}{2},p\right)^{n_1+n_2}$. However, the eigenvalues of the matrices $G_k^{n_2}\widetilde{D}_k G_k^{n_1}$ are lengthy expressions. Thus, we look for a sharp upper bound of $\rho(T)$. We first prove the following lemma.

LEMMA 3.2. *Defining* $\Gamma(k,p) := 3\rho(k,p)^{n_1+n_2}s_k^2$, *we have*

$$\rho(T) \leq \max_{k\in[1,N_y]} \Gamma(k,p). \tag{3.12}$$

Proof. We define the matrix $\widetilde{T}$, which is obtained from T replacing the blocks $G_k^{n_2}\widetilde{D}_k G_k^{n_1}$ with $D_k := \widetilde{D}_k G_k^{n_1} G_k^{n_2} = \widetilde{D}_k G_k^{n}$, where $n := n_1 + n_2$. A classical property of the spectral radius states that $\rho(G_k^{n_2}\widetilde{D}_k G_k^{n_1}) = \rho(\widetilde{D}_k G_k^{n_1} G_k^{n_2}) = \rho(D_k)$. Therefore, we have $\rho(T) = \rho(\widetilde{T}) \leq \|\widetilde{T}\|_1$. We note that due to the diagonal structure of the matrix $\widetilde{T}$,

$$\|\widetilde{T}\|_1 = \max\left\{\max_{k\in\mathcal{V}}\|D_k\|_1, \rho\left(\frac{N_y+1}{2},p\right)^n\right\}.$$

Thus we focus on the term $\|D_k\|_1$. Using the trigonometric formula $\sin(2x) = 2\sin(x)\cos(x)$, we obtain

$$\frac{\lambda(k)}{\lambda_c(k)} = \frac{\frac{2}{h}\sin(k\pi\frac{h}{2})}{\frac{2}{H}\sin(k\pi\frac{H}{2})} = \frac{H\sin(k\pi\frac{h}{2})}{h\sin(k\pi h)} = \frac{1}{\cos(k\pi\frac{h}{2})} = \frac{1}{c_k} > 1 \quad \forall k\in\mathcal{V},$$
$$\frac{\lambda(\widetilde{k})}{\lambda_c(k)} = \frac{\frac{2}{h}\sin(\frac{\pi}{2}-k\pi\frac{h}{2})}{\frac{2}{H}\sin(k\pi\frac{H}{2})} = \frac{H\cos(k\pi\frac{h}{2})}{h\sin(k\pi h)} = \frac{1}{\sin(k\pi\frac{h}{2})} = \frac{1}{s_k} > 1 \quad \forall k\in\mathcal{V}.$$

Substituting these expressions into (3.8), direct calculations yield

$$\|D_k\|_1 = \max\{\rho^n\left(1-c_k^4+c_k s_k^2\right), \widetilde{\rho}^n\left(1-s_k^4+c_k^2 s_k\right)\}.$$

Exchanging the order of the max operations over a finite set we have

$$\begin{aligned}\max_{k\in\mathcal{V}}\|D_k\| &= \max_{k\in\mathcal{V}}\max\{\rho^n\left(1-c_k^4+c_k s_k^2\right), \widetilde{\rho}^n\left(1-s_k^4+c_k^2 s_k\right)\}\\ &= \max\{\max_{k\in\mathcal{V}}\rho^n\left(1-c_k^4+c_k s_k^2\right), \max_{k\in\mathcal{V}}\widetilde{\rho}^n\left(1-s_k^4+c_k^2 s_k\right)\}.\end{aligned}$$

Now we proceed with a change of variables in the second term in the curly brackets. Due to our hypothesis on the mesh, we have that $h = \frac{1}{N_y+1}$, so that $\rho(\widetilde{k},p) = \left(\frac{\frac{2}{h}\sin((N_y+1-k)\pi\frac{h}{2})-p}{\frac{2}{h}\sin((N_y+1-k)\pi\frac{h}{2})+p}\right) = \left(\frac{\frac{2}{h}c_k-p}{\frac{2}{h}c_k+p}\right)$, where we used the trigonometric identity $\sin(\frac{\pi}{2}-x) = \cos(x)$. Using again this relation and denoting $\mathcal{Z} := \{\frac{N_y+1}{2}+1,\dots,N_y\}$, we conclude that

$$\begin{aligned}\max_{k\in\mathcal{V}}\left(\frac{\frac{2}{h}c_k-p}{\frac{2}{h}c_k+p}\right)^n\left(1-s_k^4+c_k^2 s_k\right) &= \max_{k\in\mathcal{Z}}\left(\frac{\frac{2}{h}s_k-p}{\frac{2}{h}s_k+p}\right)^n\left(1-c_k^4+s_k^2 c_k\right)\\ &= \max_{k\in\mathcal{Z}}\left(\rho(k,p)^n\left(1-c_k^4+s_k^2 c_k\right)\right).\end{aligned}$$

Thus we obtain the equality

$$\|\widetilde{T}\|_1 = \max\left\{\max_{k\in\mathcal{V}\cup\mathcal{Z}}\left(\rho(k,p)^n\left(1-c_k^4+s_k^2c_k\right)\right),\rho\left(\frac{N+1}{2},p\right)^n\right\}.$$

Now we relax the discrete constraint, and we consider the continuous frequencies $k\in[1,\frac{N_y+1}{2})\cup(\frac{N_y+1}{2},N_y]$. Clearly, it holds that

$$\begin{aligned}&\max\left\{\max_{k\in\mathcal{V}\cup\mathcal{Z}}\left(\rho(k,p)^n\left(1-c_k^4+s_kc_k^2\right)\right),\rho\left(\frac{N_y+1}{2},p\right)^n\right\}\\&\le\max\left\{\max_{k\in[1,\frac{N_y+1}{2})\cup(\frac{N_y+1}{2},N_y]}\left(\rho(k,p)^n\left(1-c_k^4+s_kc_k^2\right)\right),\rho\left(\frac{N_y+1}{2},p\right)^n\right\}.\end{aligned}$$

We now use the key observation that

(3.13)

$$\begin{aligned}\lim_{k\to\frac{N_y+1}{2}}\rho(k,p)^n\left(1-c_k^4+s_k^2c_k\right)&=\rho\left(\frac{N_y+1}{2},p\right)^n\left(1-\left(\frac{\sqrt{2}}{2}\right)^4+\left(\frac{\sqrt{2}}{2}\right)^3\right)\\&>\rho\left(\frac{N_y+1}{2},p\right)^n,\end{aligned}\tag{3.14}$$

and hence we can bound $\|\widetilde{T}\|_1$ as

$$\|\widetilde{T}\|_1\le\max_{k\in[1,N_y]}\rho(k,p)^n\left(1-c_k^4+s_k^2c_k\right).$$

To further simply the right-hand side (RHS) of this inequality we use the relation

$$(1-c_k^4+s_k^2c_k)=s_k^2(1+c_k^2+c_k)\le s_k^2(3-s_k^2)\le s_k^2 3\quad\forall k\in[1,N_y],$$

and defining $\Gamma(k,p):=3\rho(k,p)^ns_k^2$ we obtain the desired bound. ∎

We now consider the problem $\min_p\|\widetilde{T}\|_1$. From now on we restrict our analysis to the case $n_1+n_2=2$. Due to Lemma 3.2, we study the simpler problem $\min_p\max_{k\in[0,N_y]}\Gamma(k,p)$, where we expanded the range of frequencies to $k\in[0,N_y]$.

THEOREM 3.3. *Assuming that $n_1+n_2=2$, the solution of the min-max problem*

$$\min_p\max_{k\in[0,N_y]}\Gamma(k,p)\tag{3.15}$$

is given by

$$p^*=\frac{(2\sqrt{6}+2\sqrt{3}-2\sqrt{2}-4)\sin(\frac{1}{2}hN_y\pi)}{h},\tag{3.16}$$

which is the unique root of the nonlinear equation

$$\Gamma(\widetilde{k},p)=\Gamma(N_y,p),$$

where $\widetilde{k}$ is the unique interior maximum of $\Gamma(k,p)$ in the interval $k\in[0,N_y]$.

Proof. First, we observe that $\Gamma(k,p) \geq 0\ \forall k,p$ and $\Gamma(k,p) = 0$ if and only if $k = \frac{2\arcsin(\frac{hp}{2})}{h\pi}$ or $k = 0$. Second, we compute the derivative of $\Gamma(k,p)$ with respect to p,

$$\operatorname{sign}\left(\frac{\partial\Gamma(k,p)}{\partial p}\right) = \operatorname{sign}(hp - 2s_k).$$

Therefore, at the optimum, p must lie inside the interval $[0, \frac{2}{h}s_{N_y}]$. We then look for the maximum with respect to k. We have that $\frac{\partial\Gamma(k,p)}{\partial k} = 0$ if and only if $k_1 = \frac{2\arcsin(\frac{hp}{2})}{h\pi}$, which therefore is a minimum and zero, and for $k_2 = 0$, $k_3 = \frac{1}{h}$, and $\widetilde{k} = \frac{2\arcsin(\frac{1}{2}(\sqrt{2}-1)ph)}{\pi h} < k_1$. We can conclude that the function for $p \in [0, \frac{2s_{N_y}}{h}]$ starts from zero at $k = 0$, it increases until it reaches an interior maximum at $\widetilde{k}$, then it decreases until the zero k_1 whereupon it is strictly increasing until k_3. We observe that $k_3 = \frac{1}{h} > N_y$. Therefore, the function has two local maxima, one located at $\widetilde{k}$ and the other at $k = N_y$. Moreover, varying $p \in [0, s_{N_y}]$, the zero $k_1(p)$ is mapped into the interval $[0, N_y]$. Suppose now that $\Gamma(\widetilde{k},p) > \Gamma(N_y,p)$; the other case is treated similarly. Since $\operatorname{sign}(\partial_p\Gamma) = \operatorname{sign}(k_1 - k)$, we have $\partial_p\Gamma(\widetilde{k},p) > 0$ and $\partial_p\Gamma(N_y,p) < 0$. Therefore, increasing p decreases the maximum of $\Gamma(k,p)$ until $\Gamma(\widetilde{k},p^*) = \Gamma(N_y,p^*)$. This is the optimal solution since varying the parameter p would increase the value of Γ either at $k = \widetilde{k}$ or $k = N_y$. The uniqueness follows from the strict monotonicity. Finally, solving the equation $\Gamma(\widetilde{k},p^*) = \Gamma(N_y,p^*)$ we get the expression for p^*. ∎

THEOREM 3.4 (mesh independent convergence). *Assuming that $n_1 + n_2 = 2$ and choosing p as in Theorem 3.3, the spectral radius of the two-level OSM iteration matrix T is bounded below 1 uniformly with respect to h,*

$$\rho(T(p^*)) \leq C < 1 \quad \text{as } h \to 0, \quad \text{with} \quad C = 0.0520. \tag{3.17}$$

Proof. Based on Lemma 3.2, we have

$$\rho(T(p)) \leq \|T(p)\|_1 \leq \max_{k\in[0,N_y]} \Gamma(k,p).$$

Taking the minimum with respect to p, the inequality still holds; thus

$$\min_p \rho(T(p)) \leq \min_p \max_{k\in[0,N_y]} \Gamma(k,p).$$

We denote with p^* the solution of the min-max problem studied in Theorem 3.3. Clearly, there is no reason why p^* would still be the solution of the min-max problem $\min_p \rho(T(p))$. Nevertheless, we have that

$$\min_p \rho(T(p)) \leq \rho(T(p^*)) \leq \Gamma(N_y,p^*) = \min_p \max_k \Gamma(k,p).$$

Substituting the expression of p^* we get that

$$\Gamma(N_y,p^*) = \frac{3(\sqrt{2}+3-\sqrt{6}-\sqrt{3})^2 \sin(\frac{hN_y\pi}{2})^2}{(\sqrt{2}+1-\sqrt{6}-\sqrt{3})^2}.$$

We now observe that $\Gamma(N_y,p^*)$ is a strictly decreasing function of h. Therefore, it has its maximum for $h \to 0$. We then compute the limit $\lim_{h\to 0}\Gamma(N_y,p^*) = 0.0520 =: C$. Hence we conclude that $\min_p \rho(T(p)) \leq \rho(T(p^*)) \leq C < 1$ as $h \to 0$. ∎

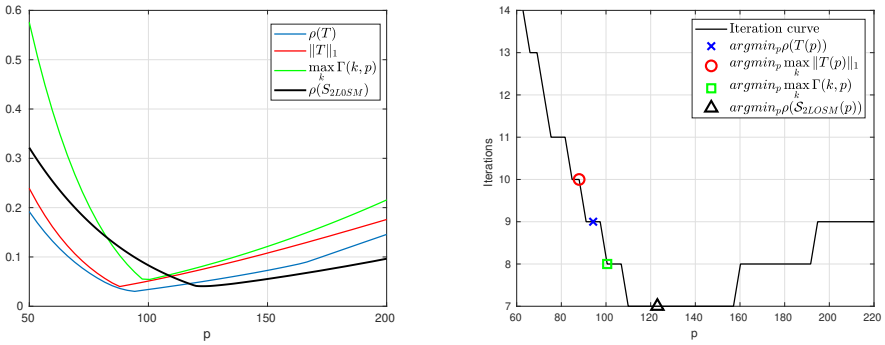

FIG. 1. *On the left, comparison between the spectral radius of $\mathcal{S}_{2LOSM}$ and T and various upper bounds. On the right, number of iterations required to reach convergence as a function of p and comparison between the predicted p obtained by solving different min-max problems involving the quantities presented in the left panel. The fine mesh corresponds to $\ell = 6$.*

Remark 3.5. The asymptotic performance of the one-level OSM has been a subject of intensive study. For straight interfaces, in [17] it has been shown that for zero order transmission conditions the spectral radius is bounded from above by $1-O(h^{\frac{1}{2}})$ in a nonoverlapping decomposition, and by $1-O(h^{\frac{1}{3}})$ in the overlapping case with overlap proportional to the mesh size. See also [33] for a generalization to arbitrary interfaces. For second order transmission conditions [17], we have, respectively, $1-O(h^{\frac{1}{4}})$ and $1-O(h^{\frac{1}{5}})$. Theorem 3.4 shows that the two-level OSM gains the same property of the multigrid scheme with a convergence independent of the mesh size because of the presence of the coarse correction. We emphasize that the same conclusion holds if one uses the classical parallel Schwarz method instead of OSM as smoother.

3.2. How to choose the optimized parameter in the nonoverlapping case. As we emphasized in the proof of Theorem 3.4, in general p^* is not a solution of the minimization problem $\min_p \rho(T(p))$. Thus we study numerically the behavior of the spectral radius and of the other bounds as functions of p. On the left of Figure 1, we plot the behavior of different quantities as p varies. From the right panel, we observe that the solutions of none of the min-max problems involving the different bounds or even $\rho(T(p))$ provide an optimized convergence. The reasons for this discrepancy lie in the several simplifications used in the literature for the derivation of the convergence factors for the one-level OSMs, which is mainly based on a continuous analysis. It therefore neglects the computation of the discrete derivative, and it approximates the eigenvalues of the discrete Laplacian with those of the continuous Laplacian.[1] In our analysis, we indeed take into account the eigenvalues of the discrete Laplacian, but we did not include the discrete derivative. We show that our small theoretical improvement actually worsens the approximation of the numerical convergence factor in the high frequencies regime. In Figure 2, we plot, for a fixed p, the numerical convergence factor, $\rho^2(k,p)$, and also the continuous analogue of $\rho^2(k,p)$, i.e., $\rho_c^2(k,p) = \left(\frac{\pi k - p}{\pi k + p}\right)^2$, which involves the continuous eigenvalues of the

[1] We remind the reader that usually the unbounded hypothesis is also made, but it has no significant impact in this case.

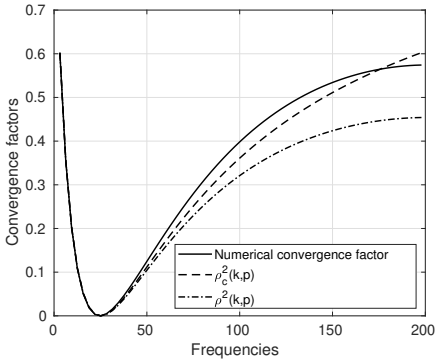

FIG. 2. *The solid line corresponds to the numerical convergence factor, the dashed line corresponds to* $\rho_c^2(k,p)$*, and the dash-dotted line to* $\rho^2(k,p)$.

Laplace operator in one dimension. It is evident that actually $\rho_c^2(k,p)$ is a better approximation of the numerical convergence factor. On the other hand, $\rho^2(k,p)$ is incorrectly faster for high frequencies, which is why our estimates for p are constantly lower than the optimal ones.

Guided by these observations we now consider an analysis which is less precise from the theoretical point of view than the one proposed in section 1, but that will provide a better approximation of the optimal parameter p. We carry out a complete continuous analysis by replacing the expansions (3.1) with

$$u_1 = \sum_{k=1}^{N_y} A(k)\psi_k e^{\pi k x} \quad \text{and} \quad u_2 = \sum_{k=1}^{N_y} B(k)\psi_k e^{-\pi k x}. \tag{3.18}$$

We insert this ansanz into the iterative method, and when dealing with the restriction and prolongation operators, we assume[2] that the same results as in the discrete case hold; see, for instance, (3.3). Repeating the same calculations, we obtain the recurrence relation $\mathbf{v}_k^n = \widetilde{D}_k^c \mathbf{v}_k^{n-1}$ similar to (3.8), where

$$\widetilde{D}_k^c := \begin{pmatrix} \left(1-c_k^4\right) & 0 & \frac{c_k^2 s_k^2}{2}\left(1+\frac{\widetilde{k}}{k}\right) & \frac{c_k^2 s_k^2}{2}\left(\frac{\widetilde{k}}{k}-1\right) \\ 0 & \left(1-c_k^4\right) & \frac{c_k^2 s_k^2}{2}\left(\frac{\widetilde{k}}{k}-1\right) & \frac{c_k^2 s_k^2}{2}\left(1+\frac{\widetilde{k}}{k}\right) \\ c_k^2 s_k^2 & 0 & 1-\frac{s_k^4\left(1+\frac{\widetilde{k}}{k}\right)}{2} & \frac{s_k^4\left(1-\frac{\widetilde{k}}{k}\right)}{2} \\ 0 & c_k^2 s_k^2 & \frac{s_k^4\left(1-\frac{\widetilde{k}}{k}\right)}{2} & 1-\frac{s_k^4\left(1+\frac{\widetilde{k}}{k}\right)}{2} \end{pmatrix}. \tag{3.19}$$

Defining $G_k^{c,n} = \mathrm{diag}(\rho_c^n(k,p), \rho_c^n(k,p), \rho_c^n(\widetilde{k},p), \rho_c^n(\widetilde{k},p))$, $\overline{D}_k := \widetilde{D}_k^c G_k^{c,n_1}$, and recall-

[2]It is a slight abuse of notation since under our hypothesis on the mesh, the eigenvectors of the discrete Laplacian correspond to the discretization on the mesh points of the eigenvectors of the continuous Laplacian.

ing $\mathbf{e}^n := \left(\mathbf{v}_1^n, \ldots, \mathbf{v}_{N_c}^n, A\left(\frac{N+1}{2}\right), B\left(\frac{N+1}{2}\right)\right)$, we conclude that $\mathbf{e}^n = \overline{T}\mathbf{e}^{n-1}$, where

$$\overline{T} = \begin{pmatrix} \overline{D}_1 & & & & & & \\ & \overline{D}_2 & & & & & \\ & & \ddots & & & & \\ & & & \overline{D}_{N_y^c} & & & \\ & & & & \rho_c^n\left(\frac{N_y+1}{2}, p\right) & & \\ & & & & & \rho_c^n\left(\frac{N_y+1}{2}, p\right) & \end{pmatrix}.$$

LEMMA 3.6. *Defining* $\overline{\Gamma}(k,p) := 3\rho_c^n(k,p)s_k^2$, *we have*

$$\rho(\overline{T}) \le \|\overline{T}\|_1 \le \max_{k\in[1,N_y]} \overline{\Gamma}(k,p).$$

Proof. The proof follows the steps of the proof of Lemma 3.2. Direct calculations show that

$$\|\overline{T}\|_1 = \max\left\{\max_{k\in\mathcal{V}} \rho_c^n(k,p)s_k^2(1+2c_k^2), \max_{k\in\mathcal{V}} \rho_c^n(\widetilde{k},p)c_k^2\left(1+s_k^2\frac{Ny+1}{k}\right), \rho_c^n\left(\frac{N_y+1}{2},p\right)\right\}.$$

Studying the second term in the brackets we conclude that $1+s_k^2\frac{Ny+1}{k} \le 2$ for $k \in \mathcal{V}$, so that we can consider the upper bound

$$\|\overline{T}\|_1 \le \max\left\{\max_{k\in\mathcal{V}} 3\rho_c^n(k,p)s_k^2, \max_{k\in\mathcal{V}} 3\rho_c^n(\widetilde{k},p)c_k^2, \rho_c^n\left(\frac{N_y+1}{2},p\right)\right\}.$$

Similarly to Lemma 3.2, we introduce the set $\mathcal{Z} := \left\{\frac{Ny+1}{2}+1, \ldots, Ny\right\}$, so that

$$\|T\|_1 \le \max\left\{\max_{k\in\mathcal{V}\cup\mathcal{Z}} 3\rho_c^n(k,p)s_k^2, \rho_c^n\left(\frac{N_y+1}{2},p\right)\right\}.$$

Finally, observing that $3\rho_c^n(\frac{Ny+1}{2},p)s_{\frac{Ny+1}{2}}^2 \ge \rho_c^n\left(\frac{N_y+1}{2},p\right)$ and considering a continous set of frequencies $k \in [1, N_y]$, we get the desired bound. ∎

We are ready to prove the following theorem.

THEOREM 3.7. *Assuming* $n = 2$, *the solution of the min-max problem*

$$\min_p \max_{k\in[0,N_y]} \overline{\Gamma}(k,p) \tag{3.20}$$

is given by $\overline{p}$, *which is the unique solution of the nonlinear equation*

$$\overline{\Gamma}(\widehat{k},p) = \overline{\Gamma}(N,p), \tag{3.21}$$

where $\overline{k}$ *is the unique interior maximum of* $\overline{\Gamma}(k,p)$.

Proof. The function has two zeros, one located at $k = 0$, the other at $k = \frac{p}{\pi}$. Analyzing the sign of the derivative with respect to p we obtain that

$$\operatorname{sign}\left(\frac{\partial\overline{\Gamma}}{\partial k}\right) = \operatorname{sign}(p - k\pi),$$

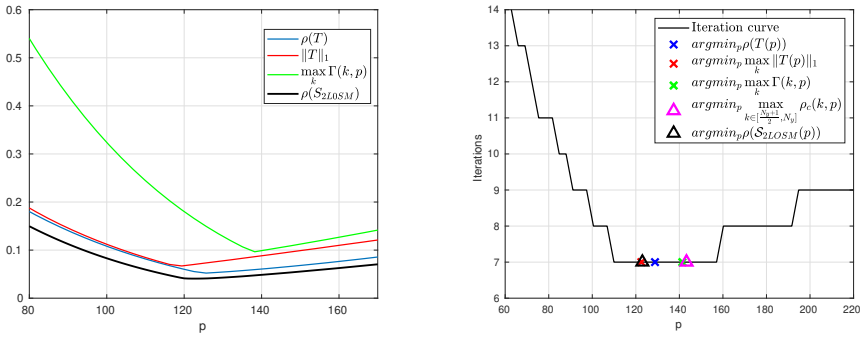

FIG. 3. *On the left, comparison between the spectral radius of $\mathcal{S}_{2LOSM}$ and of $\overline{T}$ and various upper bounds. On the right, number of iterations required to reach convergence for different values of p obtained by solving different min-max problems involving the quantities presented in the left panel. We also add a magenta triangle which represents the solution of* $\min_p \max_{k\in[\frac{N_y+1}{2},N_y]} \rho_c^2(k,p)$.

and we conclude that at the optimum $p \in [0, N_y\pi]$. The derivative with respect to k is given by

$$\frac{\partial \overline{\Gamma}}{\partial k} = 3\,\frac{(\pi\,k-p)\sin\left(\frac{1}{2}\,hk\pi\right)\pi\,\left(\cos\left(\frac{1}{2}\,hk\pi\right)hk^2\pi^2-\cos\left(\frac{1}{2}\,hk\pi\right)hp^2+4\,\sin\left(\frac{1}{2}\,hk\pi\right)p\right)}{(\pi\,k+p)^3}.$$

Therefore, the stationary points are located at $k = 0$, $k = \frac{p}{\pi}$, $k = \frac{1}{h}$, which is actually outside the interval $[0, N_y]$ and at $k = \widehat{k}$, which is the unique root of the equation $\cos\left(\frac{1}{2}\,hk\pi\right)hk^2\pi^2 - \cos\left(\frac{1}{2}\,hk\pi\right)hp^2 + 4\,\sin\left(\frac{1}{2}\,hk\pi\right) = 0$. Indeed, dividing by $\cos\left(\frac{1}{2}\,hk\pi\right) \neq 0\ \forall k \in [0, N_y]$, we get

$$hk^2\pi - hp^2 + 4\tan\left(\frac{1}{2}\pi hk\right)p = 0,$$

which is a strictly increasing function of k which for $k = 0$ is negative and for $k = N_y$ is positive. Moreover, we have that $\widehat{k} \leq \frac{p}{\pi}$. The function therefore has the following behavior: it starts from 0 at $k = 0$ and is strictly increasing until it reaches its local maximum at $k = \widehat{k}$. Then it decreases and reaches zero at $k = \frac{p}{\pi}$ but eventually it increases until the local maximum located on the boundary at $k = N_y$. Using the classical arguments of Theorem 3.3 we conclude that the solution is indeed given by equioscillation between the two local maxima. ∎

In conclusion we show in Figure 3 a comparison of the different optimized parameters that can be obtained minimizing the spectral radius, the 1-norm, or the upper bound $\overline{\Gamma}(k,p)$. We see that the unique solution of (3.21) leads to an optimized convergence.

Remark 3.8. In a one-level setting, one chooses the optimized parameter solving the min-max problem

$$\min_p \max_{[1,N_y]} \rho_c^2(k,p). \tag{3.22}$$

In the case without overlap, the parameter p solution of (3.22) does not lead to a smoother, since it tries to balance the convergence factor for low and high frequencies. However, similarly to the Jacobi smoother in a multigrid setting [30], one can choose

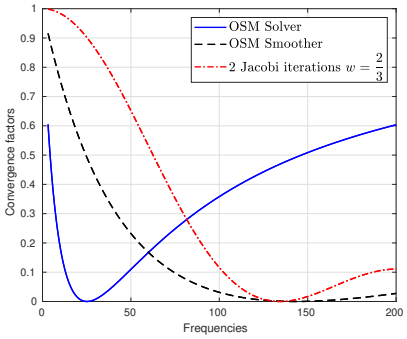

FIG. 4. *Comparison of the smoothing property for the Jacobi method with damping parameter* $w = \frac{2}{3}$*, OSM tuned as a solver, and OSM tuned as a smoother.*

p such that the OSM eliminates the high frequencies (see Figure 4) while the low ones are corrected on the coarser mesh. One obvious choice would then be to solve the min-max problem

$$\min_{p} \max_{k\in[\frac{N_y+1}{2},N_y]} \rho_c^2(k,p). \tag{3.23}$$

Figure 3 shows that this heuristic idea indeed leads to an excellent optimized parameter so that, instead of the min-max problems involving the new quantities $\Gamma, \overline{\Gamma}$, one could just use the same min-max solution involving the one-level convergence factor $\rho_c(k,p)$ by changing the interval for the variable k. The analytical solution of (3.23) is given by $p^* = \sqrt{\frac{N_y+1}{2}N_y} \approx Ch^{-1}$ as $h \to 0$, so that the asymptotic behavior of the optimized parameter obtained from (3.23) and (3.16) is the same.

4. Two-level OSM analysis for an overlapping decomposition. In this section, we present an analogous analysis for the overlapping two-level OSM. Given two initial guesses u_1^0, u_2^0 and an overlapping decomposition of Ω into Ω_1, Ω_2 with $\Gamma_j := \partial\Omega_j \setminus \partial\Omega$, $j = 1, 2$, the one-level parallel overlapping OSM reads, for $n \geq 1$,

$$\begin{array}{llll} \mathcal{L}u_1^n = f & \text{in} \quad \Omega_1, & \partial_x u_1^n + p u_1^n = \partial_x u_2^{n-1} + p u_2^{n-1} & \text{on } \Gamma_1, \\ \mathcal{L}u_2^n = f & \text{in} \quad \Omega_2, & -\partial_x u_2^n + p u_2^n = -\partial_x u_1^{n-1} + p u_1^{n-1} & \text{on } \Gamma_2. \end{array} \tag{4.1}$$

Defining two functions on the interfaces as

$$r_1 := -\partial_x u_1^n - p u_1^n + \partial_x u_2^n + p u_2^n \quad \text{on } \Gamma_1 \quad \text{and} \quad r_2 := -\partial_x u_1^n + p u_1^n + \partial_x u_2^n - p u_2^n \quad \text{on } \Gamma_2, \tag{4.2}$$

and then solving

$$\begin{array}{llll} \mathcal{L}e_1 = 0 & \text{in} \quad \Omega_1, & \partial_x e_1 + p e_1 - \partial_x e_2 - p e_2 = r_1 & \text{on } \Gamma_1, \\ \mathcal{L}e_2 = 0 & \text{in} \quad \Omega_2, & -\partial_x e_2 + p e_2 + \partial_x e_1 - p e_1 = r_2 & \text{on } \Gamma_2, \end{array} \tag{4.3}$$

we have that $\widetilde{u}_1 := u_1^n + e_1$ and $\widetilde{u}_2 := u_2^n + e_2$ are solutions of problem (2.1) in the sense that $\widetilde{u}_1 = u|_{\Omega_1}$ and $\widetilde{u}_2 = u|_{\Omega_2}$.

To analyze the method, we suppose that the two subdomains are $\Omega_1 = (-\frac{1}{2}, a) \times (0,1)$ and $\Omega_2 = (-a, \frac{1}{2}) \times (0,1)$, with two interfaces $\Gamma_1 = [a] \times [0,1]$ and $\Gamma_2 =$

$[-a]\times[0,1]$. Inserting the expansions (3.18) into the residual definition (4.2) we get

(4.4)
$$r_1 = \sum_{k=1}^{N_y} g_-(k)A(k)\rho_o^{n_1}(k,p)\psi_k e^{k\pi a} + \sum_{k=1}^{N_y} g_+(k)B(k)\rho_o^{n_1}(k,p)\psi_k e^{-k\pi a} \quad \text{on } \Gamma_1,$$
$$r_2 = \sum_{k=1}^{N_y} g_+(k)A(k)\rho_o^{n_1}(k,p)\psi_k e^{-k\pi a} + \sum_{k=1}^{N_y} g_-(k)B(k)\rho_o^{n_1}(k,p)\psi_k e^{k\pi a} \quad \text{on } \Gamma_2,$$

where $\rho_o(k,p) := \left(\frac{\pi k-p}{\pi k+p}\right)e^{-2a\pi k}$; see [17] for a derivation. Solving the corresponding version of (3.6), we obtain

$$\overline{A}(k) = \tfrac{\mathcal{R}_{1D}r_1(k)+\mathcal{R}_{1D}r_2(k)}{2F(k)} + \tfrac{\mathcal{R}_{1D}r_1(k)-\mathcal{R}_{1D}r_2(k)}{2G(k)},$$
$$\overline{B}(k) = \tfrac{\mathcal{R}_{1D}r_1(k)+\mathcal{R}_{1D}r_2(k)}{2F(k)} + \tfrac{\mathcal{R}_{1D}r_2(k)-\mathcal{R}_{1D}r_1(k)}{2G(k)},$$

where $F(k) := 2k\pi\cosh(\pi ka) + 2p\sinh(\pi ka)$, $G(k) := 2k\pi\sinh(\pi ka) + 2p\cosh(\pi ka)$. Then computing the updated approximations $u_1^{n_1} + \mathcal{I}e_1$ and $u_2^{n_1} + \mathcal{I}e_2$, we obtain that the vector $\mathbf{v}_k^n = \left(A^n(k), B^n(k), A^n(\widetilde{k}), B^n(\widetilde{k})\right)^t$ satisfies the recurrence relation $\mathbf{v}_k^n = \widetilde{D}_k^O\mathbf{v}_k^{n-1}$, with

$$\widetilde{D}_k^O := \begin{pmatrix} \rho_o^{n_1}\left(1-c_k^4\right) & 0 & \frac{\widetilde{\rho}_o^{n_1}c_k^2s_k^2}{2}\left(\frac{\widetilde{F}}{F}+\frac{\widetilde{G}}{G}\right) & \frac{\widetilde{\rho}_o^{n_1}c_k^2s_k^2}{2}\left(\frac{\widetilde{F}}{F}-\frac{\widetilde{G}}{G}\right) \\ 0 & \rho_o^{n_1}\left(1-c_k^4\right) & \frac{\widetilde{\rho}_o^{n_1}c_k^2s_k^2}{2}\left(\frac{\widetilde{F}}{F}-\frac{\widetilde{G}}{G}\right) & \frac{\widetilde{\rho}_o^{n_1}c_k^2s_k^2}{2}\left(\frac{\widetilde{F}}{F}+\frac{\widetilde{G}}{G}\right) \\ \rho_o^{n_1}c_k^2s_k^2 & 0 & \widetilde{\rho}_o^{n_1} - \frac{\widetilde{\rho}_o^{n_1}s_k^4\left(\frac{\widetilde{F}}{F}+\frac{\widetilde{G}}{G}\right)}{2} & \frac{\widetilde{\rho}_o^{n_1}s_k^4\left(\frac{\widetilde{F}}{F}-\frac{\widetilde{G}}{G}\right)}{2} \\ 0 & \rho_o^{n_1}c_k^2s_k^2 & \frac{\widetilde{\rho}_o^{n_1}s_k^4\left(\frac{\widetilde{F}}{F}-\frac{\widetilde{G}}{G}\right)}{2} & \widetilde{\rho}_o^{n_1} - \frac{\widetilde{\rho}_o^{n_1}2s_k^4\left(\frac{\widetilde{F}}{F}+\frac{\widetilde{G}}{G}\right)}{2} \end{pmatrix}, \tag{4.5}$$

where $F := F(k,p)$, $\widetilde{F} := F(\widetilde{k},p)$, and similarly for G and $\widetilde{G}$.

Remark 4.1. We note that the same calculations can be adapted to obtain an iteration matrix for a two-level method which uses the parallel Schwarz method as a smoother. We need to replace (4.1) with the classical parallel Schwarz method, the residuals are $r_1 = -u_1^n + u_2^n = -r_2$, and in the residual problem (4.3) we impose $e_1 - e_2 = r_1$ on Γ_1 and $e_2 - e_1 = r_2$ on Γ_2. Finally, we use the properties of the interpolation and restriction operators and the convergence factor $\rho_{PSM}(k) := e^{-2a\pi k}$.

Computing the 1-norm of $\widetilde{D}_k^O$ is delicate because the sign of the terms $\frac{\widetilde{F}}{F} - \frac{\widetilde{G}}{G}$ depends on p, and therefore many possible cases arise. Therefore, assuming $n_1 = 2$, we look for a proxy quantity to analyze. Inspired by section 3, we define $\overline{\Gamma}_{over}(k,p) := 3s_k\rho_o^2$. We consider the problem analogous to (3.20) for the overlapping case.

THEOREM 4.2. *The solution of the min-max problem*

$$\min_p \max_{k\in[0,+\infty]} \overline{\Gamma}_{over}(k,p) \tag{4.6}$$

is given by $\overline{p}$*, which is the unique solution of the nonlinear equation*

$$\overline{\Gamma}_{over}(\widehat{k},p) = \overline{\Gamma}_{over}(\widetilde{k},p),$$

where $\widehat{k}$ *and* $\widetilde{k}$ *are the interior maxima of* $\overline{\Gamma}_{over}(k,p)$ *for* $k\in[0,\infty]$.

Proof. We first observe that $\overline{\Gamma}_{over}(k,p) \geq 0 \, \forall k,p$ and $\overline{\Gamma}_{over}(k,p) = 0$ if and only if $k = 0$ or $k = \frac{p}{\pi}$. The sign of the derivative of $\overline{\Gamma}_{over}$ with respect to p is

$$\operatorname{sign}\left(\frac{\partial \overline{\Gamma}_{over}(k,p)}{\partial p}\right) = \operatorname{sign}(p - k\pi).$$

Therefore, we conclude that at the optimum $p \geq 0$. The zeros of the derivative with respect to k are located at $k = 0, k = \frac{p}{\pi}$, and at the only two zeros $\widetilde{k}, \widehat{k}$ of the nonlinear equation

$$\tan\left(\frac{hk\pi}{2}\right) = \frac{h\pi(p^2 - k^2\pi^2)}{4p\pi + 2\delta p^2 - 2\delta k^2\pi^2}. \tag{4.7}$$

Indeed, $g(k) := \tan\left(\frac{hk\pi}{2}\right)$ is a strictly increasing function in k, which is equal to zero for $k = 0$ and goes to infinity as $k \to +\infty$. The function $l(k,p,\delta) := \frac{h\pi(p^2 - k^2\pi^2)}{4p\pi + 2\delta p^2 - 2\delta k^2\pi^2}$ is positive for $k = 0$, it is strictly decreasing for every k, and equal to zero at $k = \frac{p}{\pi}$. Therefore, there exists a $\widehat{k}$ in the $[0, \frac{p}{\pi}]$ solution of (4.7). On the other hand, $l(k,p,\delta)$ has a vertical asymptote at $k_1 = \frac{\sqrt{\delta p(\delta p + 2\pi)}}{\delta\pi} > \frac{p}{\pi}$, and we have $\lim_{k \to k_1^+} l(k,p,\delta) = +\infty$ and $\lim_{k \to +\infty} l(k,p,\delta) = \frac{\pi h}{2\delta}$. Therefore, we conclude that there exists a $\widetilde{k} > \frac{p}{\pi}$ solution of (4.7). Hence, $\overline{\Gamma}_{over}(k,p)$ has two local maxima $\widehat{k} \leq \frac{p}{\pi} \leq \widetilde{k}$, and repeating the final argument of Theorem 3.3 we obtain that the solution of (4.6) is given by equioscillation. ∎

4.1. How to choose the optimized parameter in the overlapping case. The nonoverlapping OSM is not a natural smoother, and therefore the tuning of the transmission conditions is essential to achieve an efficient two-level method. In contrast, the overlapping OSM is a perfect smoother since it is exponentially fast for high frequencies, and thus we expect the tuning to be less important. Nevertheless, we want to study how close the solution of the optimization problem (4.6) involving $\overline{\Gamma}_{over}(k,p)$ is to the solution of $\min_p \rho(\mathcal{S}_{2LOSM})$. On the left panel of Figure 5 we plot the behavior of the iteration matrix (2.6) and of $\overline{\Gamma}_{over}(k,p)$ as a function of p. We denote with $\mathcal{S}_{2LOSM}$ the iteration matrix where we use the linear interpolator $\mathcal{I}_\ell$, and with $\mathcal{S}^A_{2LOSM}$ the one which uses $\mathcal{I}^A_\ell$. Concerning the choice of the optimized parameter, we deduce from Figure 5 that if we use the harmonic extension operator, then Theorem 4.6 provides a perfect choice for the optimized parameter. However, we observe that there is a significant difference in the spectral properties of $\mathcal{S}_{2LOSM}$ and $\mathcal{S}^A_{2LOSM}$. The explanation for this behavior is in Remark 2.1: If we use the linear interpolator $\mathcal{I}_\ell$, the corrections which we add to the iterates are no longer harmonic. Hence, we cannot assume that the expansions (3.18) hold in the interior of the subdomains, especially on Γ_j, $j = 1, 2$, where the smoother $\mathcal{S}_\ell$ acts. In Figure 6 we plot the first eigenvector of the iteration matrices $\mathcal{S}_{2LOSM}$ and $\mathcal{S}^A_{2LOSM}$ in the overlapping case. We can clearly observe that the eigenvector of $\mathcal{S}_{2LOSM}$ does not behave as an exponential along the x direction, as required by expansion (3.18).

On the right panel of Figure 5, we plot the spectral radius of $\mathcal{S}_{2LOSM}$ and $\mathcal{S}^A_{2LOSM}$ in the nonoverlapping case where the discrepancy is negligible, the $\min_p \mathcal{S}_{2LOSM}$ being attained at $p \approx 124$ and the $\min_p \mathcal{S}^A_{2LOSM}$ at $p \approx 118$. Hence, in the nonoverlapping case the use of $\mathcal{I}_\ell$ or of $\mathcal{I}^A_\ell$ does not influence the method significantly. We remark that using $\mathcal{I}_\ell$, the correction we add is not harmonic on the fine grid, but the nonoverlapping smoother takes values next to the interface and thus is less affected by the nonharmonicity inside the domain.

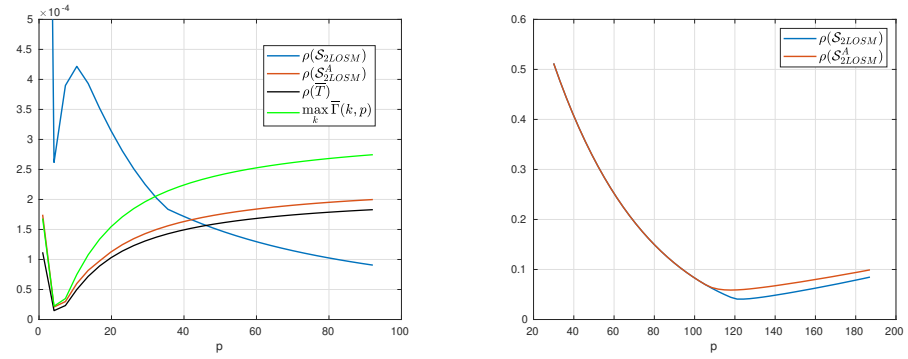

FIG. 5. *On the left, we plot the behavior of* $\mathcal{S}_{2LOSM}$, $\mathcal{S}^A_{2LOSM}$, $\rho(\overline{T})$, *and* $\max_k \overline{\Gamma}_{over}(k,p)$ *with respect to* p. *We remark that* $\mathcal{S}_{2LOSM}$ *does not behave as our analysis predicts. On the right, we plot* $\mathcal{S}_{2LOSM}$, $\mathcal{S}^A_{2LOSM}$ *in the nonoverlapping case in which the discrepancy is negligible. The fine mesh corresponds to* $\ell = 6$.

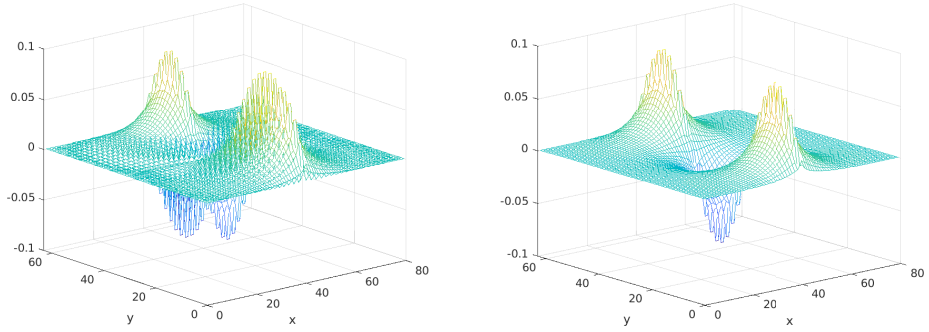

FIG. 6. *First eigenvector of* $\mathcal{S}_{2LOSM}$ *on the left and of* $\mathcal{S}^A_{2LOSM}$ *on the right.*

5. Multilevel generalization. In this section we generalize the two-level Algorithm 2.1 to a multilevel setting. As emphasized in the multigrid literature [30, 39], the coarse problem $\widetilde{A}_{\ell-1}\widetilde{\mathbf{e}}_{\ell-1} = \mathcal{R}_\ell \mathbf{r}_\ell^{n_1}$ may still be too large, and therefore one could use a two-level method to solve it. However, in the nonoverlapping case, the smoothing property of the OSM depends strongly on the transmission conditions, and therefore, moving from one grid to another, they need to be tuned according to the new mesh. This also implies that the residuals must be properly modified.

Suppose that at the continuous level we do some smoothing steps of the double-sided OSM (see [17]) with free parameters p and q. According to (2.3), the residual will be zero inside the domain, and on the interface Γ we have two functions,

$$r_1 := -\partial_x u_1^n - p u_1^n + \partial_x u_2^n + p u_2^n \quad \text{and} \quad r_2 := -\partial_x u_1^n + q u_1^n + \partial_x u_2^n - q u_2^n. \tag{5.1}$$

Suppose now that we want to change the parameters in the transmission conditions to a new couple (p_c, q_c). We are thus interested in the system

$$\begin{aligned} \mathcal{L}e_{1,c} &= 0 \quad \text{in} \quad \Omega_1, \qquad \partial_x e_1 + p_c e_1 - \partial_x e_2 - p_c e_2 = r_{1,c} \qquad \text{on } \Gamma, \\ \mathcal{L}e_{1,2} &= 0 \quad \text{in} \quad \Omega_2, \qquad -\partial_x e_2 + q_c e_2 + \partial_x e_1 - q_c e_1 = r_{2,c} \qquad \text{on } \Gamma \end{aligned} \tag{5.2}$$

for some choice of $r_{1,c}$ and $r_{2,c}$. We would like to choose $r_{1,c}$ and $r_{2,c}$ such that the

solutions of (5.2) and of (2.4) are identical. In this way the discrete solution of (5.2) with new parameters p_c, q_c on a coarse mesh would lead to a good coarse correction.

Therefore, we look at the expression of $r_{1,c}$ and $r_{2,c}$ such that $e_{1,c} = e_1$ and $e_{2,c} = e_2$. We observe that the transmission conditions in (5.1) with the two parameters p and q imply

$$(e_1 - e_2) = \frac{r_1 - r_2}{p+q}, \quad (\partial_x e_1 - \partial_x e_2) = \frac{qr_1 + pr_2}{p+q}.$$

Thus $r_{1,c}$ and $r_{2,c}$ should be such that

$$\begin{aligned} e_{1,c} - e_{2,c} &= \frac{r_{1,c} - r_{2,c}}{p_c + q_c} = \frac{r_1 - r_2}{p+q} = e_1 - e_2, \\ \partial_x e_{1,c} - \partial_x e_{2,c} &= \frac{p_c r_{2,c} + q_c r_{1,c}}{p_c + q_c} = \frac{pr_2 + qr_1}{p+q} = \partial_x e_1 - \partial_x e_2, \end{aligned} \tag{5.3}$$

so that $e_{j,c} \equiv e_j, j = 1, 2$ since they satisfy the same PDE in the interior of the subdomains and the same boundary conditions on the interface. Direct calculations from (5.3) lead to

$$r_{1,c} := r_{2,c} + \frac{p_c + q_c}{p+q}(r_1 - r_2), \quad r_{2,c} := r_1 \frac{q - q_c}{p+q} + r_2 \frac{p + q_c}{p+q}. \tag{5.4}$$

Moving to a discrete setting, we define $\mathbf{r}_\ell$ as the residual on the fine grid computed with parameters p_ℓ, q_ℓ, and with $\mathbf{r}_{l,c}$ the modified residual where the role of p_c and q_c in (5.4) is now played by $p_{\ell-1}, q_{\ell-1}$, i.e., the smoothing parameters we want to use on the coarse grid. We call $\mathcal{G}$ the operator which takes $\mathbf{r}_\ell$ and returns the modified residual according to (5.4), i.e., $\mathbf{r}_{l,c} = \mathcal{G}(\mathbf{r}_l, p_l, q_l, p_{l-1}, q_{l-1})$. Thanks to these observations, the multilevel optimized Schwarz method to solve the linear system $\widetilde{A}_{l_{\max}} \widetilde{\mathbf{u}}_{l_{\max}} = \mathbf{f}_{l_{\max}}$ consists of multiple calls of the MOSM function described by Algoritm 5.1 until convergence is reached. In the overlapping case, the smoothing property of the OSM is guaranteed by the overlap, and so there is no need to tune the parameters p_ℓ and q_ℓ on each mesh. We can always use the parameters solution of (4.6) without losing efficiency. Therefore, in the overlapping case, we just consider $\mathcal{G}$ as the identity operator.

Algorithm 5.1 Function MOSM($\widetilde{A}_\ell, \widetilde{u}_\ell^0, \widetilde{f}_\ell$).

- If $\ell = \ell_{min}$, then return $\widetilde{\mathbf{u}}_{\ell\min} \leftarrow \widetilde{A}_{\ell_{\min}}^{-1} \mathbf{f}_{\ell_{\min}}$.
- For $n = 1 : n_1$, $\widetilde{\mathbf{u}}_\ell^n \leftarrow \mathcal{S}_\ell(\widetilde{A}_\ell, \widetilde{\mathbf{u}}_\ell^{n-1}, \widetilde{\mathbf{f}}_\ell)$.
- $\mathbf{r}_{l,c} \leftarrow \mathcal{G}(\widetilde{\mathbf{f}}_\ell - \widetilde{A}_\ell \widetilde{\mathbf{u}}_\ell^{n_1}, p_\ell, q_\ell, p_{\ell-1}, q_{\ell-1})$.
- Set $\widetilde{\mathbf{e}}_{\ell-1} = 0$.
- Call γ times $\widetilde{\mathbf{e}}_{l-1} \leftarrow \text{MOSM}(\widetilde{A}_{\ell-1}, \widetilde{\mathbf{e}}_{l-1}, \mathcal{R}_l \mathbf{r}_{l,c})$.
- $\widetilde{\mathbf{u}}_\ell^{n_1} \leftarrow \widetilde{\mathbf{u}}_\ell^{n_1} + \mathcal{I}_\ell \widetilde{\mathbf{e}}_{l-1}$.
- For $n = n_1 + 1 : n_2$, $\widetilde{\mathbf{u}}_\ell^n \leftarrow \mathcal{S}_\ell(\widetilde{A}_\ell, \widetilde{\mathbf{u}}_\ell^{n-1}, \widetilde{\mathbf{f}}_\ell)$.
- Return $\widetilde{\mathbf{u}}_\ell^{n_2}$.

According to the value of γ we obtain a V-cycle ($\gamma = 1$) or W-cycle ($\gamma = 2$). In the numerical section we consider only the V-cycle, since the W-cycle shows a similar behavior.

TABLE 1

Top table: number of iterations necessary to reach the tolerance for a diffusion problem for the V-cycle OSMs and the multigrid scheme with a point Jacobi smoother. Bottom table: number of iterations needed to reach convergence as the number of levels increases in the multilevel methods for $\lambda = 1$.

λ	OSM(p)	OSM(p,q)	OSMo(p)	OSMV(p)	OSMV(p,q)	OSMoV(p)	MGV
1	164	57	13	4	4	2	11
10^5	6	5	-	1	1	-	11

# Levels	OSMV(p)	OSMV(p,q)	OSMoV(p)
2	4	4	2
4	4	4	2
6	4	4	2

6. Numerical results. Every experiment starts with a random initial guess with values between -1 and 1, and the RHS is equal to $f = 1$. We use the acronyms OSMo(p) and OSM(p) to indicate, respectively, a one-level OSM with a single sided optimized parameter p with and without overlap. OSMV(p,q) indicates a V-cycle OSM with two optimized parameters p, q. The optimized parameters are obtained by maximizing the smoothing property of the OSM scheme according to Remark 3.8. MGV stands for a multigrid V-cycle with a Jacobi smoother with damping parameter $w = \frac{2}{3}$. The number of pre- and postsmoothing steps is set equal to $n_1 = n_2 = 2$ on each level, except on the coarsest one, where the linear system is solved directly for all the multilevel schemes, i.e., for MGV, OSMV(p), OSMV(p,q), and OSMoV(p). For each equation we compute the exact solution $\widetilde{\mathbf{u}}_{\text{exact}}$ solving directly the augmented system, and we present a table containing the number of iterations required to reach a relative tolerance of $\text{Tol} := 10^{-6}$, i.e.,

$$\frac{\|\widetilde{\mathbf{u}}_l^n - \widetilde{\mathbf{u}}_{\text{exact}}\|_\infty}{\|\widetilde{\mathbf{u}}_{\text{exact}}\|_\infty} \leq \text{Tol}.$$

For heterogeneous problems, we do not consider overlapping methods.

6.1. Elliptic problems and scalability. We first consider the discrete setting described in sections 2 and 4.1 with overlap fixed to $a = 0.0625$. We study the heterogeneous diffusion equation

$$-\nabla \cdot \nu(x,y)\nabla u = f \quad \text{in} \quad \Omega, \qquad u = 0 \quad \text{on} \quad \partial\Omega, \tag{6.1}$$

where $\nu(x,y) = \nu_1$ in Ω_1 and $\nu(x,y) = \nu_2$ in Ω_2. We define the ratio $\lambda = \frac{\nu_1}{\nu_2}$ as a measure of the heterogeneity. The coarsest grid we use corresponds to $\ell = 3$ and the finest to $\ell = 9$, corresponding, respectively, to 56 and 261632 degrees of freedom. Table 1 shows the number of iterations needed to reach the tolerance for the different methods. For the homogeneous case, i.e., $\lambda = 1$, the V-cycle OSM is faster than both the one-level OSM and the multigrid method in terms of iteration counts. In the presence of heterogeneity, multigrid performance remains similar, while all the methods based on optimized Schwarz methods, both one-level and multilevel variants, become faster. The ability of OSMs to take advantage of heterogeneity is now well established; see, for instance, [25]. However, to have faster convergence, OSMs do require the jump in the diffusion coefficient to be aligned along the interfaces between the subdomains, and to properly rescale the transmission conditions according to the diffusivity constants of the adjacent subdomains. If this is not the case, OSMs

TABLE 2
Number of iterations to reach the tolerance for the anisotropic Laplace equation for the V-cycle OSM, the multigrid scheme with a point Jacobi smoother, and with a Line Jacobi smoother.

ϵ	OSMV(p)	MGV	MGV-Line Jacobi
10^{-1}	4	59	6
10^{-3}	5	4769	6

TABLE 3
Number of iterations to reach the tolerance for the advection-diffusion equation in different physical regimes.

ν	a_1	a_2	OSM(p)	OSM(p,q)	OSMV(p)	OSMV(p,q)	MGV
1	1	1	166	57	5	4	15
1	20	1	99	44	8	7	16
1	20	20	101	48	7	6	18

could even diverge [25]. Therefore, the method is not robust with respect to arbitrary decompositions into subdomains in the case of jumping diffusion coefficients. Some recent developments considering discontinuities across the interfaces are available in [29].

We then study the robustness of the methods with respect to the number of levels. We fix the finest grid to $\ell = 9$, and Table 1 shows that the number of iterations remains constant as the number of levels increases.

We then consider the anisotropic version of (6.1) where $\nu(x, y) = \mathrm{diag}(\epsilon, 1)$. If ϵ is small, then we have a higher diffusivity in the y direction than in the x direction, and multigrid performance deteriorates due to the inefficiency of classical smoothers; see Chapter 5 of [39]. Table 2 shows that the OSMV(p) does not suffer the anisotropy while multigrid becomes inefficient. A Jacobi line smoother fixes multigrid but it also makes each iteration much more expensive.

Next, we solve the advection diffusion equation

$$-\nu\Delta u + \mathbf{a}\cdot\nabla u = f \quad \text{in} \quad \Omega, \qquad u = 0 \quad \text{on} \quad \partial\Omega,$$

where $\nu \in \mathbb{R}$ and $\mathbf{a} = (a_1, a_2)^\top \in \mathbb{R}^2$. We refer to [27] for the analysis of the one-level OSM for advection-diffusion PDEs in bounded domains. The coarsest mesh is $\ell = 5$, equivalent to 992 degrees of freedom, so that on the coarsest level we still have a rough description of the boundary layer due to the advection. Table 3 shows the number of iterations required to reach convergence in different physical regimes. Note that being faster in terms of iteration numbers does not mean being faster in computational time; let us study the computational cost of a two-level optimized Schwarz method and of a multigrid scheme using point Jacobi. We denote with N, N_c, and M the number of degrees of freedom on the first level, on the second level and in each subdomain on the fine mesh. N_{sub} indicates the total number of subdomains, while N^{it} and N^{it}_{MG} are the number of iterations of the two-level optimized Schwarz method and of the multigrid scheme. Then, the computational cost (CC) of the two level optimized Schwarz method is $CC^{\text{MOSM}} = O(N^{\text{it}}((n_1+n_2)N_{sub}M^\gamma + N_c^\gamma))$, while for multigrid[3] it is $CC^{\text{MG}} = O(N^{\text{it}}_{\text{MG}}((n_1+n_2)N + N_c^\gamma))$, where γ is an exponent which depends on the structure of the matrix and on the linear solver used. It is clear that the

[3]We assume the use of a pointwise Jacobi smoother, which has a linear cost in N. However, in several situations, e.g., Table 2, one has to rely on more expensive smoothers.

N. subdomains	4	16	64	128
OSMoV(p)	3	3	4	4
RAS+Nicolaides	8	27	52	57

FIG. 7. *On the left, number of iterations needed to reach convergence as the number of subdomains increases. On the right, example of decomposition into* 16 *subdomains using Metis.*

subdomain solvers can represent a bottleneck due to the term M^γ. One solution is to increase N_{sub} so that M becomes smaller and one can then do the computations in parallel.

We now study the scalability properties of OSMoV(p) with two levels, i.e., a two-level method which uses an overlapping optimized Schwarz method with one optimized parameter p as a smoother on the fine level. We consider a square domain Ω divided into several subdomains by the partitioning tool Metis; see Figure 7 for an example of a decomposition. As we increase the number of subdomains, we keep the size of each subdomain approximately constant (around 400 degrees of freedom), so that the global problem becomes larger. In this setting, we solve (6.1) with $\nu(x, y) = 1$ and $f = 1$. The numbers of pre- and postsmoothing steps are equal to $n_1 = n_2 = 2$ and the overlap is constant, equal to four times the mesh size. Figure 7 shows that the two-level optimized Schwarz method is scalable and offers a comparison with the Restricted Additive Schwarz method (RAS) method equipped with the well-known Nicolaides coarse space; see section 4.2 of [13] for a detailed description. The OSMoV(p) requires far fewer iterations but, at least in its two-level variant, it requires solving a larger and more expensive coarse problem.

6.2. Helmholtz equation with a dispersion correction. We consider the Helmholtz equation in a square cavity open on the vertical edges with transparent Robin boundary conditions and with homogeneous Dirichlet conditions on the horizontal edges. Both OSMs and multigrid do not converge in general for Helmholtz problems when used iteratively; see [16, 21] for a detailed discussion. The oscillatory nature of the Helmholtz equation makes it difficult to design efficient two-level solvers. Some recent developments are available in [28, 8, 1, 37]. In this subsection we consider GMRES preconditioned by the one-level OSM [21], the V-cycle OSM, and the multigrid scheme. We show two numerical experiments. In the first, we test the two-level OSM with a fine mesh $\ell = 10$, approximately one million degrees of freedom, and a coarse mesh $\ell = 9$, and we compare the iterations required to converge for an increasing sequence of wave numbers ω. In Table 4, we see that the V-cycle OSM and multigrid schemes are extremely fast especially for low wave numbers. This is not surprising; in order for the coarse correction to be effective, we need a good representation of the error on the coarse mesh. Therefore, the lower the oscillations are, the better representation we have, and we are basically in an elliptic regime. As ω increases, the V-cycle OSM deteriorates and multigrid becomes highly ineffective.

In the second experiment we investigate the robustness of the multilevel methods with respect to the coarseness of the meshes. In order to provide a good coarse cor-

TABLE 4

Convergence behavior for the Helmholtz equation with different wavenumbers for a two-level method. Fine mesh is labeled $\ell = 10$ and coarse mesh $\ell = 9$.

ω	OSM(p)	OSMV(p)	MGV
5π	15	2	4
25π	25	4	6
50π	34	9	10
100π	60	40	129

TABLE 5

Convergence behavior of the V-cycle OSM and of the multigrid scheme for $\omega = 25\pi$ as the number of points per wavelength on the coarsest grid $G_{\ell_{\min}}$ is reduced. The right table refers to the dispersion correction. The finest grid corresponds to $\ell_{\max} = 8$ with $G_{\ell_{\max}} = 20.48$.

$\ell_{\min}$	$G_{\ell_{\min}}$	OSMV(p)	MGV
7	10.24	9	16
6	5.12	16	78
5	2.56	24	>200

$\ell_{\min}$	$G_{\ell_{\min}}$	OSMV(p)	MGV
7	10.24	9	8
6	5.12	16	20
5	2.56	26	>200

rection, the coarse mesh should have at least a resolution of approximately ten points per wavelength $G_\ell := \frac{2\pi}{h_\ell \omega} \approx 10$. Moreover, a coarse mesh amplifies the numerical dispersion; therefore, since this requirement sets a practical constraint on the use of multigrid for Helmholtz problems, some methods have been developed for dispersion correction such as optimized finite difference schemes; see [37, 7]. In the following, we do not use some specific new finite difference stencils to contain the numerical dispersion, but instead on each level, we modify the frequency ω of the Helmholtz equation. Indeed, in section 2 of [7], it is shown that choosing the Helmholtz frequency on each level such that

$$\omega_\ell(\theta) = \left| \sqrt{h_\ell^{-2}(4 - 2\cos(\omega h_\ell \cos(\theta)) - 2\cos(\omega h_\ell \sin(\theta)))} \right|, \tag{6.2}$$

reduces the numerical dispersion. Specifically, it removes the dispersion in the direction defined by the angle θ. Thus, supposing that on the finest grid $\ell_{\max}$ the numerical dispersion is negligible, on each coarser mesh we discretize the Helmholtz equation with a modify frequency $\omega_\ell(\theta)$. We choose the angle $\theta = \frac{\pi}{8}$ since it is very close to the value found numerically, which minimizes the maximum of the Euclidean distance between the points lying on the continuous dispersion relation $\{\xi \in \mathbb{R}^2 : \|\xi\| = \omega\}$ and the discrete one $\{\xi \in \mathbb{R}^2 : h^{-2}(4 - 2\cos(h_\ell \xi_1) - 2\cos(h_\ell \xi_2)) = \omega(\theta)\}$ for $\omega = 25\pi$.

Table 5 shows that the multigrid is very sensitive to the coarseness of the meshes, and that the dispersion correction improves its convergence behavior up to $G_\ell \approx 5$. The V-cycle OSM is instead more robust than multigrid, and it is unaffected by the correction of the frequency ω.

6.3. Helmholtz–Laplace heterogeneous coupling. We study the MOSM for the Helmholtz–Laplace coupling, which is a simplified model of a coupling between a hyperbolic and parabolic PDE,

$$-\Delta u + \omega^2 u = f \quad \text{in} \quad \Omega_1, \quad -\Delta u = f \quad \text{in} \quad \Omega_2, \qquad u = 0 \quad \text{on} \quad \partial\Omega.$$

In [24], the authors presented a one-level OSM, and they showed it converges as an iterative method despite the presence of the Helmholtz equation. We consider the finest grid $\ell = 9$ and the coarsest $\ell = 7$ such that we have more than 10 points

TABLE 6
Number of iterations necessary to reach the tolerance for the different methods in the Helmholtz-Laplace coupling.

ω	OSM(p)	OSM-V(p)	MG-V
5π	165	6	13
25π	80	9	div

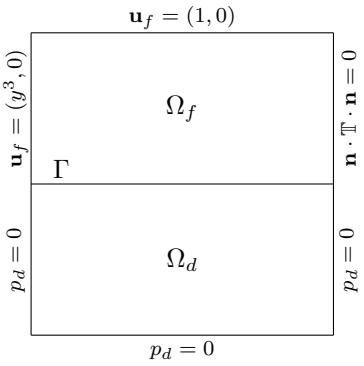

FIG. 8. *Geometry for the Stokes–Darcy problem.*

per wavelength. We see in Table 6 that the multigrid V-cycle diverges as an iterative method for a large wave number. The MOSM still converges but the coarse correction clearly becomes less effective. The one-level OSM instead improves its performance for increasing ω as long as the mesh size does not increase; see [24] for more details. If we choose the coarsest grid equal to $l = 8$, then also the multigrid V-cycle converges, but it requires 111 iterations, which illustrates the higher sensitivity of the multigrid scheme compared to the MOSM for wave problems; see also subsection 6.2.

6.4. Stokes–Darcy coupling. We consider the flow of a fluid in a domain $\Omega_f = (0,1) \times (0,1)$ which interacts through an interface $\Gamma = [0,1] \times \{0\}$ with a porous medium in a domain $\Omega_d = (0,1) \times (-1,0)$; see Figure 8. This physical phenomenon is commonly described as the Stokes–Darcy coupling [10], whose unknowns are the velocity field $\mathbf{u}_f = (u_{f,1}, u_{f,2})^\top$ and the fluid pressure p_f in Ω_f and the Darcy pressure p_d in Ω_d. In the Darcy domain the velocity is recovered through the relation $\mathbf{u}_d = (u_{d,1}, u_{d,2})^\top = -\mathbb{K}\nabla p_d$, where $\mathbb{K}$ is the permeability tensor. We suppose the fluid enters Ω_f from the left with a velocity profile $\mathbf{u}_f = (y^3, 0)^\top$; we impose $\mathbf{u}_f = (1,0)^\top$ on the top boundary, while we impose a zero normal stress condition on the right boundary, i.e., $-\mathbf{n} \cdot (2\mu_f \nabla^s \mathbf{u}_f - p_f \mathbf{I}) \cdot \mathbf{n} = 0$. The Stokes–Darcy system is

$$
\begin{aligned}
-\nabla \cdot (2\mu_f \nabla^s \mathbf{u}_f - p_f \mathbb{I}) &= f \quad \text{in } \Omega_f, \quad \mathcal{B}_f(\mathbf{u}_f, p_f) = 0 \text{ on } \partial\Omega_f \setminus \Gamma, \\
\nabla \cdot \mathbf{u}_f &= 0 \quad \text{in } \Omega_f, \\
-\nabla \cdot \mathbb{K}\nabla p_d &= -\nabla \cdot \mathbf{g}_d \quad \text{in } \Omega_d, \quad \mathcal{B}_d(p_d) = 0 \text{ on } \partial\Omega_d \setminus \Gamma, \\
\mathbf{u}_f \cdot \mathbf{n} &= -(\mathbb{K}_d \nabla p_d) \cdot \mathbf{n} + \mathbf{g}_d \cdot \mathbf{n}, \\
-\mathbf{n} \cdot (2\mu_f \nabla^s \mathbf{u}_f - p_f \mathbf{I}) \cdot \mathbf{n} &= p_d, \\
-\tau \cdot (2\mu_f \nabla^s \mathbf{u}_f - p_f \mathbf{I}) \cdot \mathbf{n} &= \chi_s(\mathbf{u}_f)_\tau,
\end{aligned}
\tag{6.3}
$$

where ∇^s is the symmetrized gradient, μ_f is the fluid viscosity, $\mathbf{g}_d$ is a body force vector, and $\mathcal{B}_f$ and $\mathcal{B}_d$ represent the boundary conditions to impose on the external

boundaries. The first two coupling conditions in (6.3) impose the continuity of the normal velocities and of normal stresses, while the last one is the Beaver–Joseph–Saffman condition. We refer the interested reader to [10] for a comprehensive introduction to the model.

In [9] the authors presented an optimized Schwarz method for the system (6.3) which computes for iterations $n = 1, 2, \ldots,$

$$
\begin{aligned}
-\nabla \cdot (2\mu_f \nabla^s \mathbf{u}_f^n - p_f^n \mathbb{I}) &= f \quad \text{in } \Omega_f, \\
\nabla \cdot \mathbf{u}_f^n &= 0 \quad \text{in } \Omega_f, \\
-\nabla \cdot \mathbb{K} \nabla p_d^n &= -\nabla \cdot \mathbf{g}_d \quad \text{in } \Omega_d, \\
p_d^n - s_1 \left(\mathbb{K} \nabla p_d^n \cdot \mathbf{n} - \mathbf{g}_d \cdot \mathbf{n}\right) &= -\mathbf{n} \cdot (2\mu_f \nabla^s \mathbf{u}_f^{n-1} - p_f^{n-1} \mathbf{I}) \cdot \mathbf{n} \\
&\quad + s_1 \mathbf{u}_f^{n-1} \cdot \mathbf{n} \quad \text{on } \Gamma, \\
-\mathbf{n} \cdot (2\mu_f \nabla^s \mathbf{u}_f^n - p_f^n \mathbf{I}) \cdot \mathbf{n} - s_2 \mathbf{u}_f^n \cdot \mathbf{n} &= p_d^{n-1} + s_2 \left(\mathbb{K} \nabla p_d^{n-1} \cdot \mathbf{n} - \mathbf{g}_d \cdot \mathbf{n}\right) \quad \text{on } \Gamma, \\
-\tau \cdot (2\mu_f \nabla^s \mathbf{u}_f^n - p_f^n \mathbf{I}) \cdot \mathbf{n} &= \chi_s (\mathbf{u}_f^n)_\tau \quad \text{on } \Gamma,
\end{aligned} \tag{6.4}
$$

and they obtained a convergence factor based on Fourier analysis. Even though it has been shown that the derived convergence factor is not accurate (see the discussion in [26]), we now show that algorithm (6.4) can be efficiently used as a smoother for a two-level OSM.

To obtain the enhanced matrix for the Stokes–Darcy coupling we consider the fixed point version of system (6.4) by letting $n \to \infty$, and we introduce the functional spaces $\mathbf{U}_f := \{\mathbf{u}_f \in (H^1(\Omega_s))^2 : \mathbf{u}_f = 0 \text{ on } \partial\Omega_f \setminus \Gamma\}$, $P_f := \{p_f \in L^2(\Omega_f) : \int_{\Omega_f} p_f = 0\}$, and $P_d := \{p_d \in H^1(\Omega_d) : p_d = 0 \text{ on } \partial\Omega_d \setminus \Gamma\}$. We set $f = 0$ and $\mathbf{g}_d = 0$, and we denote τ the tangential vector to Γ and $\mathbf{n}$ the normal vector to Γ pointing to the interior of Ω_d. Then the weak formulation of system (6.4) is

$$
\begin{aligned}
a_s(\mathbf{u}_f, \mathbf{v}_f) + b_f(\mathbf{v}_f, p_f) - b_{SD}(p_d, \mathbf{v}_f) &= \langle \bar{f}, \mathbf{v}_f \rangle && \forall \mathbf{v}_f \in \mathbf{U}_f, \\
b_f(\mathbf{u}_f, q_f) &= 0 && \forall q_f \in P_f, \\
a_d(p_d, q_d) - b_{DS}(\mathbf{u}_f, q_d) &= 0 && \forall q_d \in P_d,
\end{aligned} \tag{6.5}
$$

where

$$
\begin{aligned}
a_s(\mathbf{u}_f, \mathbf{v}_f) &:= \int_{\Omega_s} 2\mu_s \nabla^s \mathbf{u}_s : \nabla^s \mathbf{v}_s + \int_\Gamma \chi_s (\mathbf{u}_f)_\tau (\mathbf{v}_f)_\tau + \int_\Gamma s_2 (\mathbf{u}_f)_\mathbf{n} (\mathbf{v}_f)_\mathbf{n}, \\
b_f(\mathbf{v}_f, p_f) &:= \int_{\Omega_s} -p_f \nabla \cdot \mathbf{v}_f, \\
b_{SD}(p_d, \mathbf{v}_f) &:= -\int_\Gamma (p_d + s_2 \mathbb{K} \nabla p_d \cdot \mathbf{n}) \mathbf{v}_f \cdot \mathbf{n}, \\
a_d(p_d, q_d) &:= \int_{\Omega_d} \mathbb{K} \nabla p_d \cdot \nabla q_d + \int_\Gamma \frac{1}{s_1} p_d q_d, \\
b_{DS}(\mathbf{u}_f, q_d) &:= \int_\Gamma (\mathbf{u}_f \cdot \mathbf{n}) q_d - \frac{1}{s_1} \int_\Gamma \mathbf{n} \cdot (2\mu_f \nabla^s \mathbf{u}_f^n - p_f^n \mathbf{I}) \cdot \mathbf{n},
\end{aligned}
$$

and the functional $\bar{f}$ takes into account the nonhomogeneous Dirichlet conditions. A finite element discretization of (6.5) leads to the discretize system

$$
\begin{pmatrix} \begin{pmatrix} A_s & B_f \\ B_f^\top & 0 \end{pmatrix} & -B_{SD} \\ -B_{DS} & A_d \end{pmatrix} \begin{pmatrix} \begin{pmatrix} \mathbf{u}_f \\ \mathbf{p}_f \end{pmatrix} \\ \mathbf{p}_d \end{pmatrix} = \begin{pmatrix} \begin{pmatrix} \bar{\mathbf{f}} \\ \mathbf{0} \end{pmatrix} \\ \mathbf{0} \end{pmatrix}. \tag{6.6}
$$

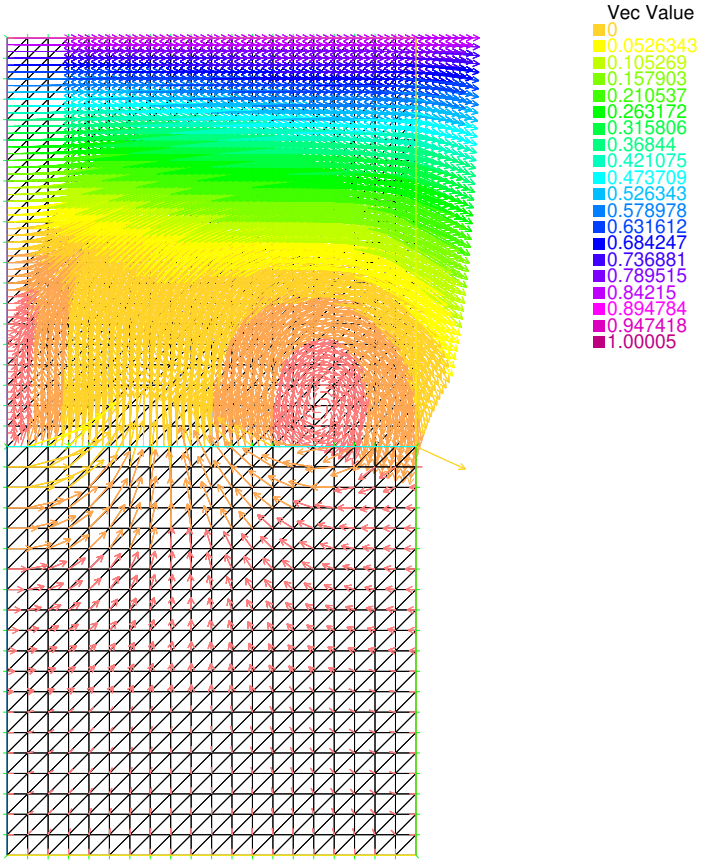

FIG. 9. *Plot of the velocity field solution to the problem described in Figure* 8.

We implement the two-level OSM using the finite element software FreeFem++ [31], and we choose $\mathcal{P}_1$-bubble elements for the Stokes velocity and for the Darcy pressure and $\mathcal{P}_1$ elements for the Stokes pressure. We have 25404 degrees of freedom on the fine mesh and 6454 on the coarse one. The optimized parameter p is chosen to maximizing the smoothing property of the convergence factor derived in [9]. The parameters are $h = 0.05$, $\mu_f = 0.1$, $\mathbb{K} = \mathrm{diag}(1, 1)$. We first compute the exact solution $(\bar{\mathbf{u}}_f, \bar{p}_f, \bar{p}_d)^\top$ by solving directly system (6.6), and then we count the number of iterations necessary for the one-level OSM and the MOSM to reach a tolerance of $\mathrm{Tol} = 10^{-6}$, i.e.,

$$\left(\|\mathbf{u}_f^n - \bar{\mathbf{u}}_f\|_{(H^1)^2} + \|p_f^n - \bar{p}_f\|_{L^2} + \|p_d^n - \bar{p}_d\|_{H^1}\right)^{\frac{1}{2}} \leq \mathrm{Tol}. \tag{6.7}$$

For the MOSM, we used two presmoothing steps and no postsmoothing. The one-level OSM requires 14 iterations, while the two-level OSM only 4. In Figure 9 we show the solution for the velocity fields in the two subdomains for the problem described by Figure 8.

7. Conclusions. We introduced a new two-level OSM and provided a convergence analysis both for overlapping and nonoverlapping decompositions. The analysis indicates how to choose the optimized transmission conditions; we have shown that it is possible to rely exclusively on the already available literature concerning the one-level OSM. The two level method can be generalized to obtain a multilevel domain

decomposition method which uses OSMs as smoothers on each level (MOSM). The MOSM exhibits mesh independent convergence and is weakly scalable. An extensive numerical section shows that MOSM inherits robustness from the underlying standard OSM, especially for heterogeneous and wave problems. The theoretical framework presented allows us to straightforwardly define two-level and multilevel methods for very general problems. As an example, we have discussed how our method can be readily applied to design a two-level domain decomposition solver for the Stokes–Darcy coupling. In this manuscript we have compared the MOSM with the one-level OSM and the multigrid scheme in terms of the number of iterations. A recent implementation in PETSc [23] has shown that coarse corrections can reduce the difference in time units between the OSM and the multigrid scheme. Hence, future developments may focus on a detailed time comparison of the MOSM and multigrid scheme, with an implementation in a compiled language.

REFERENCES

[1] M. Bonazzoli, V. Dolean, I. G. Graham, E. A. Spence, and P. Tournier, *Two-level preconditioners for the Helmholtz equation*, Domain Decomposition Methods in Science and Engineering XXIV, Lect. Notes Comput. Sci. Eng. 125, Springer, Cham, 2018, pp. 139–147.

[2] J. H. Bramble, J. E. Pasciak, and J. Xu, *Parallel multilevel preconditioners*, Math. Comp., 55 (1990), pp. 1–22.

[3] F. Chaouqui, G. Ciaramella, M. Gander, and T. Vanzan, *On the scalability of classical one-level domain-decomposition methods*, Vietnam J. Math., 46 (2018), pp. 1053–1088.

[4] P. Charton, F. Nataf, and F. Rogier, *Méthode de décomposition de domaine pour l'équation d'advection-diffusion*, C. R. Acad. Sci. Paris Sér. I Math., 313 (1991), pp. 623–626.

[5] G. Ciaramella and M. J. Gander, *Analysis of the parallel Schwarz method for growing chains of fixed-sized subdomains: Part* II, SIAM J. Numer. Anal., 56 (2018), pp. 1498–1524, https://doi.org/10.1137/17M1115885.

[6] G. Ciaramella and T. Vanzan, *Substructured Two-level and Multilevel Domain Decomposition Methods*, preprint, https://arxiv.org/abs/1908.05537, 2019.

[7] P. Cocquet, M. J. Gander, and X. Xiang, *A finite difference method with optimized dispersion correction for the Helmholtz equation*, Domain Decomposition Methods in Science and Engineering XXIV, Lect. Notes Comput. Sci. Eng. 125, Springer, Cham, 2018, pp. 205–213.

[8] L. Conen, V. Dolean, R. Krause, and F. Nataf, *A coarse space for heterogeneous Helmholtz problems based on the Dirichlet-to-Neumann operator*, J. Comput. Appl. Math., 271 (2014), pp. 83–99.

[9] M. Discacciati and L. Gerardo-Giorda, *Optimized Schwarz methods for the Stokes–Darcy coupling*, IMA J. Numer. Anal., 38 (2018), pp. 1959–1983.

[10] M. Discacciati and A. Quarteroni, *Navier-Stokes/Darcy coupling: modeling, analysis, and numerical approximation*, Rev. Mat. Complut. 22 (2009), pp. 315–426.

[11] V. Dolean, M. J. Gander, and L. Gerardo-Giorda, *Optimized Schwarz methods for Maxwell's equations*, SIAM J. Sci. Comput., 31 (2009), pp. 2193–2213, https://doi.org/10.1137/080728536.

[12] V. Dolean, M. J. Gander, and E. Veneros, *Optimized Schwarz methods for Maxwell equations with discontinuous coefficients*, in Domain Decomposition Methods in Science and Engineering XXI, Lect. Notes Comput. Sci. Eng. 98, Springer, Cham, 2014, pp. 517–525.

[13] V. Dolean, P. Jolivet, and F. Nataf, *An Introduction to Domain Decomposition Methods: Algorithms, Theory, and Parallel Implementation*, SIAM, Philadelphia, 2015, https://doi.org/10.1137/1.9781611974065.

[14] O. Dubois and M. J. Gander, *Convergence behavior of a two-level optimized Schwarz preconditioner*, in Domain Decomposition Methods in Science and Engineering XVIII, M. Bercovier, M. J. Gander, R. Kornhuber, and O. Widlund, eds., Lect. Notes Comput. Sci. Eng. 70, Springer, Berlin, Heidelberg, 2009, pp. 177–184.

[15] O. Dubois, M. J. Gander, S. Loisel, A. St-Cyr, and D. B. Szyld, *The optimized Schwarz method with a coarse grid correction*, SIAM J. Sci. Comput., 34 (2012), pp. A421–A458, https://doi.org/10.1137/090774434.

[16] O. G. Ernst and M. J. Gander, *Why it is difficult to solve Helmholtz problems with classical iterative methods*, Lect. Notes Comput. Sci. Eng. 83, Springer, Berlin, Heidelberg, 2012, pp. 325–363.
[17] M. J. Gander, *Optimized Schwarz methods*, SIAM J. Numer. Anal., 44 (2006), pp. 699–731, https://doi.org/10.1137/S0036142903425409.
[18] 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.
[19] M. J. Gander, L. Halpern, and K. Santugini-Repiquet, *Discontinuous coarse spaces for dd-methods with discontinuous iterates*, in Domain Decomposition Methods in Science and Engineering XXI, Lect. Notes Comput. Sci. Eng. 98, Springer, Cham, 2014, pp. 607–615.
[20] M. J. Gander and A. Loneland, *SHEM: An optimal coarse space for RAS and its multiscale approximation*, in Domain Decomposition Methods in Science and Engineering XXIII, Lect. Notes Comput. Sci. Eng. 116, C.-O. Lee, X.-C. Cai, D. E. Keyes, H. H. Kim, A. Klawonn, E.-J. Park, and O. B. Widlund, eds., 2017, Springer, Cham, pp. 313–321.
[21] M. J. Gander, F. Magoules, and F. Nataf, *Optimized Schwarz methods without overlap for the Helmholtz equation*, SIAM J. Sci. Comput., 24 (2002), pp. 38–60, https://doi.org/10.1137/S1064827501387012.
[22] M. J. Gander and B. Song, *Complete, optimal and optimized coarse spaces for additive Schwarz*, in Domain Decomposition Methods in Science and Engineering XXIV, Lect. Notes Comput. Sci. Eng. 125, Springer, Cham, 2018, pp. 301–309.
[23] M. J. Gander and S. Van Criekingen, *New coarse corrections for optimized restricted additive Schwarz using PETSc*, in Domain Decomposition Methods in Science and Engineering XXIV, Springer Verlag, submitted, 2019.
[24] M. J. Gander and T. Vanzan, *Heterogeneous optimized Schwarz methods for coupling Helmholtz and Laplace equations*, Domain Decomposition in Science end Engineering XXIV, Lect. Notes Comput. Sci. Eng. 125, Springer, Cham, 2018, pp. 311–320.
[25] M. J. Gander and T. Vanzan, *Heterogeneous optimized Schwarz methods for second order elliptic PDEs*, SIAM J. Sci. Comput., 41 (2019), pp. A2329–A2354, https://doi.org/10.1137/18M122114X.
[26] M. J. Gander and T. Vanzan, *On the derivation of optimized transmission conditions for the Stokes-Darcy coupling*, in Domain Decomposition in Science end Engineering XXV, accepted, 2019.
[27] M. J. Gander and T. Vanzan, *Optimized Schwarz methods for advection diffusion equations in bounded domains*, in Numerical Mathematics and Advanced Applications ENUMATH 2017, Lect. Notes Comput. Sci. Eng. 126, Springer, Cham, 2019, pp. 921–929.
[28] I. G. Graham, E. A. Spence, and E. Vainikko, *Domain decomposition preconditioning for high-frequency Helmholtz problems with absorption*, Math. Comput., 86 (2017), pp. 2089–2127.
[29] Y. Gu, *Nonlinear Optimized Schwarz Preconditioning for Heterogeneous Elliptic Problems*, Ph.D. thesis, Hong Kong Baptist University, Hong Kong, 2019.
[30] W. Hackbusch, *Multigrid Methods and Applications*, Springer Ser. Comput. Math. 4, Springer-Verlag, Berlin, 2013.
[31] F. Hecht, *New development in freefem++*, J. Numer. Math., 20 (2012), pp. 251–265.
[32] C. Japhet, *Optimized Krylov-Ventcell method. Application to convection-diffusion problems*, in Proceedings of the 9th International Conference on Domain Decomposition Methods, 1998, pp. 382–389.
[33] S. H. Lui, *A Lions non-overlapping domain decomposition method for domains with an arbitrary interface*, IMA J. Numer. Anal., 29 (2009), pp. 332–349.
[34] F. Nataf, F. Rogier, and E. De Sturler, *Optimal Interface Conditions for Domain Decomposition Methods*, Tech. report, École Polytech., Paris, 1994.
[35] B. Smith, P. Bjorstad, and W. Gropp, *Domain Decomposition: Parallel Multilevel Methods for Elliptic Partial Differential Equations*, Cambridge University Press, Cambridge, 2004.
[36] A. St-Cyr, M. J. Gander, and S. J. Thomas, *Optimized multiplicative, additive, and restricted additive Schwarz preconditioning*, SIAM J. Sci. Comput., 29 (2007), pp. 2402–2425, https://doi.org/10.1137/060652610.
[37] C. Stolk, A. Mostak, and K. B. Samir, *A multigrid method for the Helmholtz equation with optimized coarse grid corrections*, SIAM J. Sci. Comput., 36 (2014), pp. A2819–A2841, https://doi.org/10.1137/13092349X.

[38] A. Toselli and O. Widlund, *Domain Decomposition Methods—Algorithms and Theory*, Springer Ser. Comput. Math. 34, Springer-Verlag, Berlin, 2006.
[39] U. Trottenberg, C. W. Oosterlee, and A. Schuller, *Multigrid*, Academic Press, San Diego, 2001.
[40] J. Xu, *Iterative methods by space decomposition and subspace correction*, SIAM Rev., 34 (1992), pp. 581–613, https://doi.org/10.1137/1034116.