CONDITION NUMBER ESTIMATES FOR THE NONOVERLAPPING OPTIMIZED SCHWARZ METHOD AND THE 2-LAGRANGE MULTIPLIER METHOD FOR GENERAL DOMAINS AND CROSS POINTS*

SÉBASTIEN LOISEL†

Abstract. The optimized Schwarz method and the closely related 2-Lagrange multiplier method are domain decomposition methods which can be used to parallelize the solution of partial differential equations. Although these methods are known to work well in special cases (e.g., when the domain is a square and the two subdomains are rectangles), the problem has never been systematically stated nor analyzed for general domains with general subdomains. The problem of cross points (when three or more subdomains meet at a single vertex) has been particularly vexing. We introduce a 2-Lagrange multiplier method for domain decompositions with cross points. We estimate the condition number of the iteration and provide an optimized Robin parameter for general domains. We hope that this new systematic theory will allow broader utilization of optimized Schwarz and 2-Lagrange multiplier preconditioners.

Key words. Domain decomposition, Schwarz method, partial differential equation, parallel preconditioner, Krylov space, 2-Lagrange multiplier

1. Introduction. In mathematics, physics and engineering, it is useful to solve elliptic partial differential equations, such as the Laplace problem

$$\Delta u = f \text{ in } \Omega \text{ and } u = 0 \text{ on } \partial\Omega. \tag{1.1}$$

Such problems are often solved numerically. The discretized problem has the form

$$Au = f, \tag{1.2}$$

where A is a large invertible $n \times n$ matrix, f is a given n-dimensional vector, and u is the desired solution.

When A is large, it may be desirable to solve it iteratively, by breaking it up into smaller pieces using a domain decomposition method. Such methods are readily made parallel, since each subdomain can be assigned to a separate processor. At the geometric level, a nonoverlapping domain decomposition is a partition of the domain Ω into nonoverlapping parts $\Omega_1, \ldots, \Omega_p$. From this partition, we can further define the "artificial interface"

$$\Gamma = \Omega \cap \left(\bigcup_{i=1}^{p} \partial\Omega_i \right).$$

The set $\partial\Omega$ is called the natural boundary, since it is an intrinsic part of the original problem definition. The interface Γ is artificial in that it bears no relationship to the "physical" problem (1.1). Indeed, it is introduced purely for the purpose of calculation.

The partition $\Omega_1, \ldots, \Omega_p$ can be used to assign the n degrees of freedom of u to the various subdomains. Although the geometric domain decomposition is nonoverlapping, from the algebraic point of view there is a kind of overlap since degrees of

*THIS VERSION: HERIOT-WATT MATHEMATICS REPORT HWM11-5, 1 MAR 2011

†Department of Mathematics, Heriot-Watt University, Edinburgh, EH14 4AS (S.Loisel@hw.ac.uk).

freedom along interfaces may have to be shared between two or more adjacent subdomains.[1] The current article is concerned mainly with such methods, which are nonoverlapping in the geometric (or PDE) interpretation, while at the algebraic level, the degrees of freedom along Γ are shared between the adjacent subdomains.

The 2-Lagrange multiplier (2LM) method was introduced in [9]. The main idea is to replace the large system (1.2) with a related nonsingular system

$$A_{\mathrm{2LM}}\lambda = h \tag{1.3}$$

of much smaller dimension. The λ vector is the Lagrange multiplier vector. Although A_{2LM} is much smaller than A, we consider that A_{2LM} is still too large to be solved directly, and we are interested in finding a parallel iterative solver that exploits the decomposition $\Omega_1, \ldots, \Omega_p$ of the domain Ω. In the case of two nonoverlapping subdomains, there are two Lagrange multipliers per interface node in Γ, which explains the nomenclature. The terminology "Lagrange multiplier" is an analogy to the FETI methods [10], where the Lagrange multipliers genuinely arise as part of a relaxation of continuity constraints.

The 2LM method is known to be closely related to the nonoverlapping optimized Schwarz method (OSM). We now briefly outline the history of the OSM, and refer to [11] and [12] for details. The OSM was introduced in [3], [30], [31], under various names. Attempts have been made to find optimized transmission conditions for various differential equations and/or suitable domains; see e.g., [16], [19], [20], [21], [22].

For some simple problems and domains, optimized transmission conditions can be found by Fourier analysis. In addition to the Laplace and Helmholtz problems, this Fourier method can be used for various other canonical problems; see [2], [7], [14], [15], [27], [28] for convection-diffusion problems, [13], [17] for the wave equation, [1], [5] for Maxwell's equations, [6] for fluid dynamics, and [29], [32] for the shallow water equation.

Proofs of convergence for more general situations have been recently obtained [23], [26]; but the techniques used are not amenable to finding the optimal parameters. A proof of convergence for the nonoverlapping algorithm without cross points was provided in [24], using energy estimates (this proof does not provide condition number estimates or optimized parameter values). We mention in passing that one way of obtaining convergence of the nonoverlapping algorithm is to define a relaxation of the method [4]. There is an ongoing effort to estimate the spectral radius of the optimized Schwarz iteration with cross points, see [18].

All of these methods have been described for domain decompositions without cross points, but the presence of cross points poses a difficulty. OSMs have seen limited deployment in applications. We surmise that this is because of the poorly understood performance of the OSM when there are cross points. It is difficult for practitioners to use the method when it is not even clearly defined. Our major innovation in the present article is to introduce and analyze a systematic method to deal with cross

[1] We briefly mention that this has caused some confusion in the community. Some domain decompositions which are nonoverlapping at the geometric (or PDE) level (for example, the nonoverlapping optimized Schwarz method) end up using overlapping blocks of the matrix A. Conversely, a (classical) additive Schwarz preconditioner with nonoverlapping blocks of A actually corresponds (in the geometric or PDE interpretation) to an overlapping domain decomposition whose subdomains overlap by exactly one grid interval. In other words, the amounts of overlap at the geometric and algebraic levels are not necessarily related.

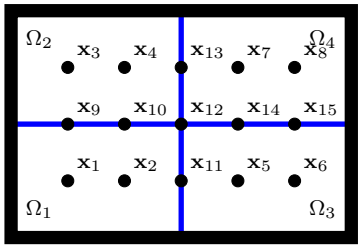

FIG. 2.1. *A discretized rectangular domain with 15 grid points. The artificial interface is* $\Gamma = \{\mathbf{x}_9, \ldots, \mathbf{x}_{15}\}$. *The interface vertex* $\mathbf{x}_{12}$ *is a cross point. The first subdomain is* $\bar{\Omega}_1 = \{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_9, \mathbf{x}_{10}, \mathbf{x}_{11}, \mathbf{x}_{12}\}$.

points. Our new methods generalize the nonoverlapping OSM or 2LM method to general domains and subdomains with cross points.

We have three main results. Our first main result is to give a nonoverlapping OSM, or a 2LM system, defined even when there are cross-points. We show that we can recover the solution to the system $Au = f$ from the unique solution to this nonoverlapping OSM. This main result is important because the 2LM method or the nonoverlapping OSM has never been formulated systematically in the presence of cross points with general subdomains.

Our second main result gives the optimized Robin parameter and a condition number estimate for the nonoverlapping OSM in terms of the spectral properties of local Schur complements.[2] We apply this estimate to a PDE to obtain our third main result.

Our paper is organized as follows. In Section 2, we introduce the 2LM method and show that (1.2) and (1.3) are equivalent, even if A is nonsymmetric. In Section 3, we provide condition number estimates for the 2LM method in terms of algebraic properties of A. In Section 4, we give condition number estimates for the case when A is the discretization of an elliptic PDE. We end with some conclusions in Section 5.

2. Solving $Au = f$ using Robin subproblems. We assume that the domain[3] Ω is in $\mathbb{R}^2$ or $\mathbb{R}^3$ and is discretized using some grid of points $\{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$, as in Figure 2.1. The domain is further partitioned into nonoverlapping subdomains $\Omega_1, \ldots, \Omega_p$ with artificial interface Γ.

To fix ideas, in the remainder of the present paper, we will assume that the vertices are arranged as follows:

$$\overbrace{\mathbf{x}_1, \ldots, \mathbf{x}_{n_{I1}}}^{\in \Omega_1}, \overbrace{\mathbf{x}_{n_{I1}+1}, \ldots, \mathbf{x}_{n_{I1}+n_{I2}}}^{\in \Omega_2}, \ldots, \overbrace{\mathbf{x}_{n_I - n_{Ip}+1}, \ldots, \mathbf{x}_{n_I}}^{\in \Omega_p}, \overbrace{\mathbf{x}_{n_I+1}, \ldots, \mathbf{x}_n}^{\in \Gamma},$$

see Figure 2.1.

Throughout the present article, we will use (1.1) as a model problem. Nevertheless, we do not use any of the special features of (1.1) until Section 3. For instance, our first main result (at the end of Section 2) does not require that the matrix A be symmetric or positive definite.

[2] The idea that the condition number could be estimated from the spectral properties of the local Schur complements occurred to us when we were reading [8].

[3] The ideas in the present paper can be applied to more general domains and even graphs. Our choice makes for a simple exposition.

2.1. Restriction matrices and traces. We now consider n-dimensional vectors. We interpret such a vector u as a function defined at each vertex $\mathbf{x}_j$. To each subdomain Ω_j, we may define a restriction matrix R_j, which restricts an arbitrary n-dimensional vector u to an n_j-dimensional vector $R_j u$, which contains only the components of u corresponding to Ω_j and its artificial boundary $\partial\Omega_j \cap \Gamma$.

EXAMPLE 2.1. *The restriction matrix R_1 corresponding to Ω_1 in Figure 2.1 is*

$$R_1 = \left[\begin{array}{cccccccccccccc} \mathbf{1} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \mathbf{1} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{1} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{1} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{1} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{1} & 0 & 0 \end{array}\right] = \left[\begin{array}{c} R_{I1} \\ \hline R_{\Gamma 1} \end{array}\right].$$

The top two rows, labeled R_{I1}, correspond to the restriction to the interior vertices $\{\mathbf{x}_1, \mathbf{x}_2\}$, while the bottom four rows, labeled $R_{\Gamma 1}$, correspond to the restriction to the interface vertices $\{\mathbf{x}_9, \mathbf{x}_{10}, \mathbf{x}_{11}, \mathbf{x}_{12}\}$.

We similarly partition $R_2, \ldots, R_p$ into interior parts R_{Ij} (top) and artificial interface parts $R_{\Gamma j}$ (bottom).

For $j = 1, \ldots, p$, let u_j be an arbitrary n_j-dimensional vector. We interpret such a vector as a function on Ω_j. We further partition u_j so that the top part u_{Ij} corresponds to the vertices of $\Omega_j \setminus \Gamma$, while the bottom part $u_{\Gamma j}$ corresponds to the vertices of $\partial\Omega_j \cap \Gamma$, i.e.: $u_j = \left[\begin{array}{c} u_{Ij} \\ u_{\Gamma j} \end{array}\right]$. We can think of the vector $(u_1^T, \ldots, u_p^T)$ as a function which is defined on Ω, and which is continuous inside of each Ω_j, but which has jump discontinuities across Γ. For such a function, we define the **multi-valued** or **many-sided trace**

$$u_G = \left[\begin{array}{c} u_{\Gamma 1} \\ \vdots \\ u_{\Gamma p} \end{array}\right] \tag{2.1}$$

of dimension n_G. The n_G degrees of freedom of u_G correspond to vertices $\{\mathbf{x}_{n_I+1}, \ldots, \mathbf{x}_n\}$ on Γ, but u_G contains multiple degrees of freedom for each $\mathbf{x}_j$ (one per subdomain adjacent to $\mathbf{x}_j$). For each interface $\mathbf{x}_j \in \Gamma$, we let m_j be the number of subdomains adjacent to $\mathbf{x}_j$. We say that $\mathbf{x}_j$ is a **regular interface vertex** if $m_j = 2$, and we say that $\mathbf{x}_j$ is a **cross point** if $m_j \geq 3$.

2.2. Continuous many-sided traces. Although in general the many-sided trace u_G corresponds to a discontinuous function, it may happen that u_G corresponds in fact to a continuous function. This occurs precisely when the degrees of freedom of u_G associated with the interface vertices $\mathbf{x}_j$ all agree, for each $j = n_I + 1, \ldots, n$. We now make this concept more explicit. To that end, we define $\breve{\Pi}$ to be the permutation matrix that reorders the entries of the many-sided trace u_G so that all degrees of freedom associated with the first interface vertex $\mathbf{x}_{n_I+1}$ appear first, followed by the degrees of freedom associated with the second interface vertex $\mathbf{x}_{n_I+2}$, and so on.

Then, for $k = n_I + 1, \ldots, n$, we let

$$\check{M}_k = \begin{bmatrix} 1 & -\frac{1}{m_k-1} & \cdots & -\frac{1}{m_k-1} \\ -\frac{1}{m_k-1} & 1 & \cdots & -\frac{1}{m_k-1} \\ \vdots & \vdots & \ddots & \vdots \\ -\frac{1}{m_k-1} & -\frac{1}{m_k-1} & \cdots & 1 \end{bmatrix} \in \mathbb{R}^{m_k \times m_k}. \tag{2.2}$$

Note that the kernel of $\check{M}_k$ is spanned by the vector of ones.[4] Finally, let

$$M = \check{\Pi}^T \operatorname{diag}\{\check{M}_{n_I+1}, \ldots, \check{M}_n\}\check{\Pi}. \tag{2.3}$$

By the construction of M, we have that u_G is a continuous many-sided trace if and only if

$$Mu_G = 0. \tag{2.4}$$

We also define the $n_G \times n_G$ symmetric matrix G of "interface interactions" by

$$G = \begin{bmatrix} R_{\Gamma 1} R_{\Gamma 1}^T & \cdots & R_{\Gamma 1} R_{\Gamma p}^T \\ \vdots & \ddots & \vdots \\ R_{\Gamma p} R_{\Gamma 1}^T & \cdots & R_{\Gamma p} R_{\Gamma p}^T \end{bmatrix} \tag{2.5}$$

The entries of G are all either 0 or 1. The rows (or columns) of G precisely span the space of continuous many-sided traces, and hence we have that $MG = GM = 0$. The matrix $\check{\Pi}$ can be used to block-diagonalize G and we obtain

$$G = \check{\Pi}^T \operatorname{diag}\{\mathbf{1}_{m_{n_I+1} \times m_{n_I+1}}, \ldots, \mathbf{1}_{m_n \times m_n}\}\check{\Pi}, \tag{2.6}$$

where $\mathbf{1}_{m_j \times m_j}$ denotes the $m_j \times m_j$ matrix of ones.

We define the orthogonal projection matrix K whose range is the space of continuous many-sided traces. It can be given succinctly using the matrix $\check{\Pi}$:

$$K = \check{\Pi}^T \operatorname{diag}\left\{ \frac{1}{m_{n_I+1}} \mathbf{1}_{m_{n_I+1} \times m_{n_I+1}}, \ldots, \frac{1}{m_n} \mathbf{1}_{m_n \times m_n} \right\} \check{\Pi}. \tag{2.7}$$

If $u_1, \ldots, u_p$ are given and if the many-sided trace u_G is continuous, then there is a unique u such that

$$u_k = R_k u, \quad k = 1, \ldots, p. \tag{2.8}$$

This u is given by "gluing together" the local functions $u_1, \ldots, u_p$.

2.3. Decomposition of the matrix A. We now consider the $n \times n$ linear system (1.2) where A and f are given and u is the unknown quantity. We assume that the "global stiffness matrix" A is decomposed into "local stiffness matrices" $A_{N1}, \ldots, A_{Np}$, one per subdomain, and likewise for the data f, such that

$$A = \sum_{j=1}^{p} R_j^T A_{Nj} R_j \quad \text{and} \quad f = \sum_{j=1}^{p} R_j^T f_j. \tag{2.9}$$

[4] The formula (2.2) is not the only possible choice, and indeed all of our results hold (mutatis mutandis) if we use any other normal matrix M whose kernel is the range of G.

For $j = 1, \ldots, p$, the matrix A_{Nj} acts on the subdomain Ω_j and hence can be partitioned into blocks that act on the interior of Ω_j and on artificial interface vertices of Ω_j. We can likewise partition $f_1, \ldots, f_p$ to obtain

$$A_{Nj} = \begin{bmatrix} A_{IIj} & A_{I\Gamma j} \\ A_{\Gamma Ij} & A_{\Gamma\Gamma j} \end{bmatrix} \quad \text{and} \quad f_j = \begin{bmatrix} f_{I1} \\ f_{\Gamma j} \end{bmatrix}.$$

In the case of the model problem (1.1), the local stiffness matrices correspond to Neumann problems on the subdomains, with bilinear forms

$$a_j(u, v) := \int_{\Omega_j} \nabla u \cdot \nabla v, \quad \text{for } j = 1, \ldots, p.$$

The vector f_j is obtained from the functional $v \mapsto \int_{\Omega_j} fv$.

2.4. Robin subproblems. Given "Robin data" $\lambda_1, \ldots, \lambda_p$ and "transmission condition" matrices $B_1, \ldots, B_p$, we can compute "local solutions" $u_1, \ldots, u_p$ using

$$\begin{bmatrix} A_{IIk} & A_{I\Gamma k} \\ A_{\Gamma Ik} & A_{\Gamma\Gamma k} + B_k \end{bmatrix} \overbrace{\begin{bmatrix} u_{Ik} \\ u_{\Gamma k} \end{bmatrix}}^{u_k} = \begin{bmatrix} f_{Ik} \\ f_{\Gamma k} + \lambda_k \end{bmatrix}, \text{ for } k = 1, \ldots, p. \tag{2.10}$$

We can eliminate interior nodes from (2.10) by using Schur complements. For each Neumann matrix A_{Nk}, we define the Schur complement and "condensed right-hand-side"

$$S_k = A_{\Gamma\Gamma k} - A_{\Gamma Ik} A_{IIk}^{-1} A_{I\Gamma k} \quad \text{and} \quad g_k = f_{\Gamma k} - A_{\Gamma Ik} A_{IIk}^{-1} f_{Ik}.$$

We define S and B to be the block-diagonal matrices $S = \operatorname{diag}\{S_1, \ldots, S_p\}$ and $B = \operatorname{diag}\{B_1, \ldots, B_p\}$, respectively, and we define g to be the column vector $g = \left[g_1^T, \ldots, g_p^T\right]^T$. The system (2.10) is then equivalent to the Schur relation

$$(S + B)u_G = g + \lambda. \tag{2.11}$$

Here and in the sequel, we assume that B and $S + B$ are invertible.

In the domain decomposition parlance, the Schur complements $S_1, \ldots, S_p$ are known as (discrete) Dirichlet-to-Neumann maps. For the model problem (1.1), it is known that each S_j is nonsingular if $\partial\Omega_j$ intersects the natural boundary $\partial\Omega$. If $\partial\Omega_j$ does not intersect $\partial\Omega$ then the kernel of S_j is spanned by the vector $\mathbf{1}$ of ones. We then say that Ω_j **floats**. This characterization of the kernel of S will be used in Section 3 and onwards.

2.5. The equivalence of (1.2) and (1.3). In the present subsection, we show that the $n_G \times n_G$ system (1.3), where

$$A_{\text{2LM}} = (BM - GB)(S + B)^{-1} + G \quad \text{and} \quad h = -(BM - GB)(S + B)^{-1} g, \tag{2.12}$$

is equivalent to (1.2). The solution λ of (1.3) is a many-sided trace $\lambda = \left[\lambda_1^T, \ldots, \lambda_p^T\right]^T$. We do not assume that (1.2) arises from the model problem (1.1). We begin with a technical result.

LEMMA 2.2. *Assume that A is invertible. Let R_Γ be the matrix which restricts u to its single-valued trace u_Γ: $R_\Gamma = \begin{bmatrix} 0 & I \end{bmatrix} \in \mathbb{R}^{(n-n_I)\times n}$. There is a unique solution $u_1, \ldots, u_p, \lambda_1, \ldots, \lambda_p$ to the simultaneous equations (2.10), (2.4) and the equation*

$$\sum_k R_\Gamma R_{\Gamma k}^T \lambda_k = \sum_k R_\Gamma R_{\Gamma k}^T B_k u_{\Gamma k}. \tag{2.13}$$

If v is the unique solution to the linear problem $Av = f$, then $u = v$, where u is the unique solution to (2.8).

Proof. Assume we have $\lambda_1, \ldots, \lambda_p$ as well as $u_1, \ldots, u_p$ satisfying (2.10), (2.4), (2.13). By (2.4), the local solutions $u_1, \ldots, u_p$ meet continuously and we obtain a u such that (2.8) is satisfied. We see that, for this u,

$$\begin{aligned} Au &\overset{(2.9)}{=} \sum_k R_k^T \begin{bmatrix} A_{Ik} & A_{I\Gamma k} \\ A_{\Gamma Ik} & A_{\Gamma\Gamma k} \end{bmatrix} R_k u \\ &= \begin{bmatrix} A_{II}u_I + A_{I\Gamma}u_\Gamma \\ \sum_k R_\Gamma R_{\Gamma k}^T A_{\Gamma Ik} u_{Ik} + \sum_k R_\Gamma R_{\Gamma k}^T A_{\Gamma\Gamma k} u_{\Gamma k} \end{bmatrix}, \end{aligned} \tag{2.14}$$

where we have used (2.8). Equation (2.10) further yields

$$Au = \begin{bmatrix} f_I \\ f_\Gamma + \sum_k R_\Gamma R_{\Gamma k}^T \lambda_k - \sum_k R_\Gamma R_{\Gamma k}^T B_k u_{\Gamma k} \end{bmatrix} \overset{(2.13)}{=} f,$$

which shows that indeed any solution $u_1, \ldots, u_p, \lambda_1, \ldots, \lambda_p$ to the system (2.10), (2.4), (2.13) yields u, via (2.8), which is the unique solution to $Au = f$, as required.

We now show the uniqueness of $u_1, \ldots, u_p, \lambda_1, \ldots, \lambda_p$. Assume that $u_1^*, \ldots, u_p^*$, $\lambda_1^*, \ldots, \lambda_p^*$ is a different solution of (2.10), (2.4), (2.13). If $u_i = u_i^*$ for $i = 1, \ldots, p$, then (2.10) gives that $\lambda_i = \lambda_i^*$ for $i = 1, \ldots, p$. Hence, we may assume that $u_i \neq u_i^*$ for some i. We then obtain u^* satisfying (2.8) for which again $Au^* = f$. Since A is invertible, it must be that $u^* = u$. Hence $u_i^* = R_i u^* = R_i u = u_i$, which contradicts $u_i^* \neq u_i$. Hence, the solution to the system (2.10), (2.4), (2.13) is unique. ∎

Using (2.11), the systems (2.4) and (2.13) (with (2.10) having been eliminated) become

$$M(S+B)\lambda = -M(S+B)g \quad \text{and} \tag{2.15}$$

$$\sum_{k=1}^p R_\Gamma R_{\Gamma k}^T (I - B_k(S_k + B_k)^{-1})\lambda_k = \sum_{k=1}^p R_\Gamma R_{\Gamma k}^T B_k (S_k + B_k)^{-1} g_k, \tag{2.16}$$

respectively. Since M is a square matrix, the system (2.15) is already square and hence the system (2.15), (2.16) is rectangular (taller than it is wide). By construction, any solution $\lambda_1, \ldots, \lambda_p$ can be turned into a solution u of $Au = f$ using (2.10) and then (2.8). It is more convenient to solve a square nonsingular system. This is achieved by picking some matrices C_1 and C_2, and finding the system $C_1(2.15) + C_2(2.16)$. We now make the choices $C_1 = B$ and

$$C_2 = \begin{bmatrix} R_{\Gamma 1} R_\Gamma^T \\ \vdots \\ R_{\Gamma p} R_\Gamma^T \end{bmatrix} \tag{2.17}$$

These choices C_1 and C_2 give (1.3) and (2.12). We are now able to show our first main result.

THEOREM 2.3. *Assume that A and $S+B$ are invertible, and B^{-1} is positive definite. The system (1.3) is equivalent to the system (2.15)–(2.16). In particular, it has a unique solution λ, from which the unique solution u to (1.2) can be recovered.*

Proof. It suffices to show that the rows of the left-hand-side of (2.15) and (2.16) lie in the linear span of the rows of (2.12). We begin by recovering the rows of (2.16).

We left-multiply $A_{2\mathrm{LM}}$ by GB^{-1}. By construction, the rows of G are continuous many-sided traces, and hence $GM = 0$. Therefore,

$$GB^{-1}A_{2\mathrm{LM}} = GB^{-1}G(I - B(S+B)^{-1}). \tag{2.18}$$

We will now show that the range of $GB^{-1}G$ is precisely the range of G. This will allow us to recover (2.16) from the rows of (2.18).

Clearly, the range of $GB^{-1}G$ is included in the range of G. Let $k = \operatorname{rank} G$, and choose a matrix U of width k such that the columns of GU form a basis for the column space of G. In fact, by using a QR decomposition, we may assume that the columns of GU are orthonormal. Since B^{-1} is positive definite, there is a real number $\alpha > 0$ such that $v^T B^{-1} v \geq \alpha v^T v$, for all vectors v. Therefore, since the columns of GU are orthonormal, and since G is symmetric, we get that

$$v^T U^T G B^{-1} G U v \geq \alpha v^T U^T G G U v = \alpha v^T v \text{ for any } v.$$

Hence, $X = U^T G B^{-1} G U$ is positive definite. Since X is a $k \times k$ matrix, we get that the rank of X is k. But,

$$k = \operatorname{rank} X = \operatorname{rank} U^T G B^{-1} G U \leq \operatorname{rank} G B^{-1} G \leq k.$$

Hence, $\operatorname{rank} GB^{-1}G = k = \operatorname{rank} G$ and the range of $GB^{-1}G$ is the entire range of G.

Therefore, there is a matrix Y such that $YGB^{-1}G = G$. Left-multiplying (2.18) by Y, we obtain

$$YGB^{-1}A_{2\mathrm{LM}} = G(I - B(S+B)^{-1}). \tag{2.19}$$

We can now recover (2.16) by selecting suitable rows of (2.19). We now show how to do this.

Note that each row of R_Γ coincides with some row of some $R_{\Gamma j}$. Therefore, there is a matrix V, which selects the appropriate rows of

$$G = \begin{bmatrix} R_{\Gamma 1} \\ \vdots \\ R_{\Gamma p} \end{bmatrix} \begin{bmatrix} R_{\Gamma 1}^T & \dots & R_{\Gamma p}^T \end{bmatrix},$$

such that $VG = R_\Gamma \begin{bmatrix} R_{\Gamma 1}^T & \dots & R_{\Gamma p}^T \end{bmatrix}$. For this matrix V, we have

$$VYGB^{-1}A_{2\mathrm{LM}} = R_\Gamma \begin{bmatrix} R_{\Gamma 1}^T & \dots & R_{\Gamma p}^T \end{bmatrix} (I - B(S+B)^{-1}),$$

which is the matrix on the left-hand-side of (2.16).

Now that we have recovered the matrix on the left-hand-side of (2.16), we may recover (2.15) via the relation $(1.3) = C_1(2.15) + C_2(2.16)$, as required. ∎

REMARK 2.4. *For Theorem 2.3, we have not assumed that A is symmetric nor positive definite.*

2.6. Motivation for the 2LM method. When the subdomains are arranged in a strip (and there are no cross points), it is known [33] that the Richardson iteration applied to (1.3) is equivalent to the OSM. This is interesting because the OSM is known to converge rapidly [25]. Existing convergence factor estimates[5] using Fourier transforms suggest that the condition number of A_{2LM} varies in the grid parameter h like $O(h^{-\frac{1}{2}})$. The remainder of the present article will show that this is true in general (cf. (4.3)), and further elucidate the dependence of the condition number on the number of subdomains.

3. The symmetric case; condition number estimates. Here and in the rest of the present paper, we assume that A arises from (1.1) or a similar problem, as follows. We assume that A symmetric and positive definite and S is symmetric and semidefinite. We also assume that the kernel of S is spanned by the indicating functions of the subdomains that float.

From (2.2), (2.6), (2.7), we see that $(M+G)K = MK + GK = G$. Since the range of G is precisely the kernel of M, we also conclude that $M+G$ is invertible. Hence,

$$(M+G)^{-1}G = K. \tag{3.1}$$

Therefore, we take $B = aI$ (where $a > 0$ is a parameter to be chosen) and we left-multiply (1.3) by $(M-G)^{-1}$ to obtain an equivalent symmetric system:

$$A_{\mathrm{S2LM}}\lambda = h_S, \text{ where } A_{\mathrm{S2LM}} = \overbrace{a(S+aI)^{-1}}^{Q} - K \text{ and } h_S = -\overbrace{a(S+aI)^{-1}}^{Q} g. \tag{3.2}$$

The matrix Q is interpreted as the Robin-to-Dirichlet map, scaled by the tuning parameter a. Since condition numbers are submultiplicative,

$$\mathcal{K}(A_{\mathrm{2LM}}) \leq M_0\mathcal{K}(A_{\mathrm{S2LM}}) \text{ and } \mathcal{K}(A_{\mathrm{S2LM}}) \leq M_0\mathcal{K}(A_{\mathrm{2LM}}), \tag{3.3}$$

where $M_0 = \mathcal{K}(M-G)$. We say that A_{2LM} and A_{S2LM} are **spectrally equivalent**.[6]

In order to estimate the condition number of A_{S2LM}, we prove a matrix analytical result which allows us to bound the modulus of the eigenvalues of A_{S2LM} below and above. We will make repeated use of the weighted Young inequality

$$|\xi\zeta| \leq \frac{s}{2}\xi^2 + \frac{1}{2s}\zeta^2. \tag{3.4}$$

LEMMA 3.1. *Let Q be a symmetric and positive definite matrix, with eigenvalues $0 < q_1 \leq q_2 \leq \ldots \leq q_n \leq 1$. Let K be an orthogonal projection and define $A_{\mathrm{S2LM}} = Q - K$. Let $P = I - Q$ and denote the k nonzero eigenvalues of P by $1 > p_1 \geq p_2 \geq \ldots \geq p_k > 0$. Let $\mathcal{N} = \ker P$ be the nullspace of P. Define E to be the orthogonal projection onto $\mathcal{N}$. Assume that there is a real number $0 \leq r < 1$ such that*

$$\|Ec\| \leq r\|c\| \tag{3.5}$$

for every c such that $Kc = c$. Assume that $\min\{q_1, p_k\} > 0$ is small and r is close to 1. Then, the spectrum of $A_{\mathrm{S2LM}} = Q - K$ satisfies

$$|\sigma(A_{\mathrm{S2LM}})| \subset [\min\{p_k, q_1\}(1-r), 1]. \tag{3.6}$$

[5] It bears repeating that such estimates are only known for special PDEs and special domains, divided into two special subdomains, while the present article considers the general case.

[6] We will see in Section 3.1 that, under some conditions, M_0 is not too large.

In particular, the condition number $\mathcal{K}(A_{\mathrm{S2LM}})$ is bounded by

$$\mathcal{K}(A_{\mathrm{S2LM}}) \leq (\min\{p_k, q_1\}(1-r))^{-1}. \tag{3.7}$$

Proof. The spectrum of Q can be given as a function of the spectrum of S:

$$\sigma(Q) = \left\{ \frac{a}{z+a} \ : \ z \in \sigma(S) \right\}.$$

Note that $0 < \frac{a}{z+a} \leq 1$ (since $z \geq 0$ by the semi-definite hypothesis on S) and hence $\sigma(Q) \subset [0,1]$. Since $\sigma(K) = \{0,1\}$, we have that $\sigma(A_{\mathrm{S2LM}}) = \sigma(Q-K) \subset [-1,1]$, which proves the upper bound of (3.6). We now estimate the eigenvalue of A_{S2LM} with the smallest magnitude.

For any given λ, define $c = K\lambda$ (the continuous part of λ) and $d = \lambda - c$ (the discontinuous part of λ). In this way, $\|\lambda\|^2 = \|c\|^2 + \|d\|^2$ and $A_{\mathrm{S2LM}}\lambda = -Pc + Qd$. By the spectral theorem, without loss of generality we may assume that P and Q are both diagonal matrices $P = \mathrm{diag}\{p_1, \ldots, p_k, 0, \ldots, 0\}$ and $Q = \mathrm{diag}\{q_1, \ldots, q_k, 1, \ldots, 1\}$, since P and Q are symmetric and $PQ = QP$. Therefore, we have

$$\|A_{\mathrm{S2LM}}\lambda\|^2 = \sum_{i=1}^{k} p_i^2 c_i^2 + \sum_{i=1}^{n} q_i^2 d_i^2 - 2\sum_{i=1}^{k} p_i c_i q_i d_i.$$

We introduce a parameter $t > 0$ (to be determined later) and use the hypothesis that $c^T d = 0$ to obtain

$$\begin{aligned}
\|A_{\mathrm{S2LM}}\lambda\|^2 &= \sum_{i=1}^{k} p_i^2 c_i^2 + \sum_{i=1}^{n} q_i^2 d_i^2 - 2\sum_{i=1}^{k} p_i c_i q_i d_i + 2\sum_{i=1}^{n} t c_i d_i \\
&= \sum_{i=1}^{k} p_i^2 c_i^2 + \sum_{i=1}^{n} q_i^2 d_i^2 - 2\sum_{i=1}^{k} \left(1 - \frac{t}{p_i q_i}\right) p_i c_i q_i d_i + 2\sum_{i=k+1}^{n} t c_i d_i.
\end{aligned}$$

We now use the weighted Young inequality (3.4) with the choices $\xi = p_i c_i$, $\zeta = q_i d_i$ and $s = s_i > 0$, for $i = 1, \ldots, k$, where we have introduced the positive parameters $s_1, \ldots, s_k$, which will be determined later. In addition, we also use the weighted Young inequality with the choices $\xi = c_i$, $\zeta = d_i$ and $s = \sigma$, for $i = k+1, \ldots, n$ where $\sigma > 0$ will be determined later. We obtain

$$\begin{aligned}
\|A_{\mathrm{S2LM}}\lambda\|^2 \geq &\sum_{i=1}^{k} p_i^2 \left(1 - \left(1 - \frac{t}{p_i q_i}\right) s_i\right) c_i^2 + \sum_{i=1}^{k} q_i^2 \left(1 - \left(1 - \frac{t}{p_i q_i}\right) s_i^{-1}\right) d_i^2 \\
&+ \sum_{i=k+1}^{n} (1 - t\sigma^{-1}) d_i^2 - \sum_{i=k+1}^{n} t\sigma c_i^2
\end{aligned}$$

Now we use the hypothesis that $\|Ec\| \leq r\|c\|$, which implies that $\|Ec\|^2 \leq \frac{r^2}{1-r^2}\|(I -$

$E)c\|^2$, and hence

$$\begin{aligned}\|A_{\mathrm{S2LM}}\lambda\|^2 &\geq \sum_{i=1}^{k} \overbrace{\left(p_i^2 - p_i^2\left(1-\frac{t}{p_iq_i}\right)s_i - \frac{t\sigma r^2}{1-r^2}\right)}^{\alpha_i} c_i^2 \\ &\quad + \sum_{i=1}^{k} \overbrace{q_i^2\left(1-\left(1-\frac{t}{p_iq_i}\right)s_i^{-1}\right)}^{\beta_i} d_i^2 + \sum_{i=k+1}^{n}(1-t\sigma^{-1})d_i^2 \qquad (3.8)\\ &\geq \min_i\{\alpha_i\}\|(I-E)c\|^2 + \min_i\{\beta_i\}\|(I-E)d\|^2 + (1-t\sigma^{-1})\|Ed\|^2.\end{aligned}$$

We again use that $\|Ec\| \leq r\|c\|$, which implies that $\|(I-E)c\|^2 \geq (1-r^2)\|c\|^2$, and hence

$$\|A_{\mathrm{S2LM}}\lambda\|^2 \geq \overbrace{\min\left\{(1-r^2)\min_i\{\alpha_i\}, \min_i\{\beta_i\}, 1-t\sigma^{-1}\right\}}^{\rho}\|\lambda\|^2.$$

There are now two cases to consider, namely $\min\{p_k, q_1\} = p_k$ and $\min\{p_k, q_1\} = q_1$. We treat the case $\min\{p_k, q_1\} = p_k$ in detail; the case $\min\{p_k, q_1\} = q_1$ is done in a similar fashion.

We now choose the parameters $s_1, \ldots, s_k, \sigma, t$ in such a way that $\rho \geq p_k^2(1-r)^2$. This can be achieved using the following procedure. First, we solve $1 - t\sigma^{-1} = p_k^2(1-r)^2$ for t. Then we solve $\beta_i = p_k^2(1-r)^2$ for s_i (cf. (3.8)). The resulting weights are

$$t = \sigma - \sigma\overbrace{(r-1)^2p_k^2}^{<1} > 0 \quad \text{and} \quad s_i = \frac{q_i\left(p_iq_i - \sigma + \sigma\,(r-1)^2\,{p_k}^2\right)}{p_i\underbrace{\left({q_i}^2 - {p_k}^2\,(r-1)^2\right)}_{>0}}. \qquad (3.9)$$

The coefficient t is positive, and so is the coefficient s_i, provided that σ is small enough. We make the choice

$$\sigma = p_k(1-r), \qquad (3.10)$$

and check (by substituting into (3.9)) that this value of σ is small enough that s_i is indeed positive, provided $p_k < r$.

We substitute the values of t, s_i, q_i and σ given by (3.9), (3.10) and $q_i = 1 - p_i$ into $(1-r^2)\alpha_i$ to obtain

$$\begin{aligned}&\phi(p_i) := (1-r^2)\alpha_i = \qquad (3.11)\\ &\Big((-1+r)\,p_k\left(-2\,p_i\,(-1+r)\,(r+1)\,(-1+p_i) - (-1+r)^5\,{p_k}^5\right.\\ &-(-1+r)\left(-2\,p_ir^2+1+{p_i}^2\right)p_k + 2\,p_i\,(r+1)\,(-1+r)^3\,(-1+p_i)\,{p_k}^2\\ &\left.+(-1+r)^3\left(-2\,p_ir^2+{p_i}^2r^2+2\right){p_k}^3\right)\Big) \Big/ \left((-1+p_i)^2 - {p_k}^2\,(-1+r)^2\right).\end{aligned}$$

The function $\phi(p_i)$ has no singularities in the interval $p_i \in [p_k, 1-p_k]$ and $\phi'(p_i) = 0$ at the roots

$$p^{(1)} = \frac{-1+{p_k}^2\,(-1+r)^2}{-1+(-1+r)\,p_k+{p_k}^2\,(-1+r)^2} \quad \text{and} \qquad (3.12)$$

$$p^{(2)} = -\,(p_kr-p_k-1)\,(p_kr-p_k+1)^2\,. \qquad (3.13)$$

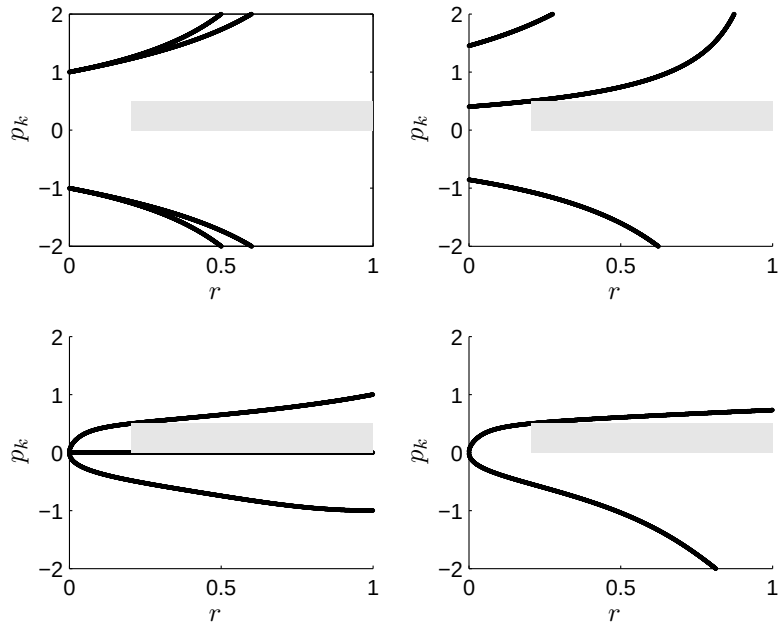

FIG. 3.1. *Solutions of* $\phi(p_i) = p_k^2(1-r)^2$ *for various values of* p_i, p_k *and* r. *Top-left:* $p_i = p^{(1)}$. *Top-right:* $p_i = p^{(2)}$. *Bottom-left:* $p_i = p_k$. *Bottom-right:* $p_i = 1 - p_k$. *The region* R *has been lightly shaded.*

We now find the sign of $\Phi_{p_i}(r, p_k) := \phi(p_i) - p_k^2(1-r)^2$. Since $\phi(p_i)$ is a differentiable function for $p_i \in [p_k, 1-p_k]$, we have that

$$\Phi_{p_i}(r, p_k) \geq \min_{p_i \in \{p^{(1)}, p^{(2)}, p_k, 1-p_k\}} \Phi_{p_i}(r, p_k).$$

We consider the case $p^{(1)}$ in detail, the other cases are similar.

To find the sign of $\Phi_{p^{(1)}}(r, p_k)$, we solve $\Phi_{p^{(1)}}(r, p_k) = 0$ for the unknown p_k, giving the roots

$$p_k^{1,\pm} = \pm\frac{\sqrt{1-r^2}}{r-1} \quad \text{and} \quad p_k^{2,\pm} = \frac{\pm 1}{r-1}.$$

We have plotted these roots in the (r, p_k) plane in Figure 3.1 (top-left). We note that the zeros of $\Phi_{p^{(1)}}(r, p_k)$ (as a function of p_k and r) do not intersect the rectangle $R = \{(r, p_k) \mid 0.204 < r < 1 \text{ and } 0 < p_k < 0.5\}$. We then compute $\Phi_{p^{(1)}}(0.5, 0.25) \approx 0.7229 > 0$, which proves that $\Phi_{p^{(1)}}(r, p_k)$ is positive throughout the rectangle R.

The cases $p_i = p^{(2)}$, $p_i = p_k$ and $p_i = 1 - p_k$ lead to the top-right, bottom-left and bottom-right parts (respectively) of Figure 3.1, and the result follows. □

REMARK 3.2. *Lemma 3.1 assumes that* $\min\{q_1, p_k\}$ *is small and* r *approaches* 1, *but we were not able to find a counter example. These "limiting" assumptions may not be necessary. In the case* $p_k = \min\{q_1, p_k\}$, *our current proof method shows that (3.7) holds provided that* $0 < p_k < 0.5$ *and* $0.5 < r < 1$. *It is possible that the choice*

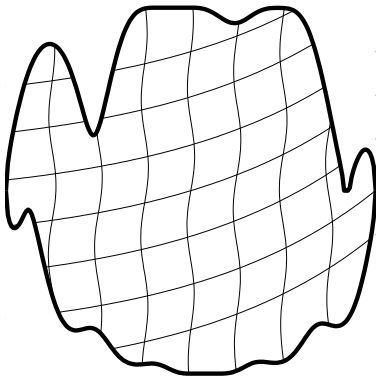

FIG. 3.2. *A gridlike domain decomposition of diameter* $\ell = 7$.

of σ *in (3.10) could be improved in such a way that the conclusion (3.6) holds*[7] *for all* $0 \le r < 1$.

We now give an example that shows that, for our purpose, (3.6) cannot be improved meaningfully without additional assumptions on A_{S2LM}.

EXAMPLE 3.3 ($\sigma_{\min}(A_{\mathrm{S2LM}})$ with $n = 2$). *For given parameters* $\theta \in \mathbb{R}$ *and* $0 < q_1 < 1$, *we set*

$$K = \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix} \begin{bmatrix} \cos\theta & \sin\theta \end{bmatrix} = \begin{bmatrix} \cos^2\theta & \cos\theta\sin\theta \\ \cos\theta\sin\theta & \sin^2\theta \end{bmatrix},$$
$$Q = \begin{bmatrix} q_1 & \\ & 1 \end{bmatrix} \quad \textit{and} \quad A_{\mathrm{S2LM}} = Q - K = \begin{bmatrix} q_1 - \cos^2\theta & -\cos\theta\sin\theta \\ -\cos\theta\sin\theta & \cos^2\theta \end{bmatrix}.$$

Note that we have chosen $n = 2$ *and* $k = 1$, *and we have that* $p_k = p_1 = 1 - q_1$. *We then have that* $r = \sin\theta$, *or* $\theta = \arcsin r$. *Direct calculations give*

$$\sigma_{\min}(A_{\mathrm{S2LM}}) = \frac{(p_k - 1) + \sqrt{(p_k+1)^2 - 4r^2 p_k}}{2}, \tag{3.14}$$

from which we find

$$1 \le \frac{\sigma_{\min}(A_{\mathrm{S2LM}})}{p_k(1-r)} \le 4 \quad \textit{for all} \quad 0 < p_k < \frac{1}{2} \quad \textit{and} \quad 0 < r < 1. \tag{3.15}$$

A series expansion further shows that $\sigma_{\min}(A_{\mathrm{S2LM}}) \approx 2p_k(1-r)$ *near* $p_k = 0$ *and* $r = 1$.

In dimensions $n \ge 3$, one can find examples of Q and K such that the condition number $\mathcal{K}(Q-K)$ is much smaller than $(p_k(1-r))^{-1}$, but Example 3.3 shows that the estimate (3.7) cannot be improved without further assumptions on A_{S2LM}.

3.1. Gridlike domain decomposition. We now estimate r in (3.7). For $j = 1, \ldots, p$, let $n_{\Gamma j}$ denote the number of vertices on the artificial interface $\partial\Omega_j \cap \Gamma$ of the jth subdomain. Let $\Omega_1, \ldots, \Omega_{p_I}$ be the subdomains that float (they are not

[7] We mention that the case $r = 0$ admits a simpler proof.

adjacent to the natural boundary $\partial\Omega$). Consider the piecewise constant many-sided trace $\lambda \in \mathcal{N}$ given by

$$\lambda = \begin{bmatrix} z_1 \frac{\mathbf{1}_{n_{\Gamma 1}}}{\sqrt{n_{\Gamma 1}}} \\ \vdots \\ z_{p_I} \frac{\mathbf{1}_{n_{\Gamma p_I}}}{\sqrt{n_{\Gamma p_I}}} \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \overbrace{\begin{bmatrix} \frac{\mathbf{1}_{n_{\Gamma 1}}}{\sqrt{n_{\Gamma 1}}} & & & \\ & \ddots & & \\ & & \frac{\mathbf{1}_{n_{\Gamma p_I}}}{\sqrt{n_{\Gamma p_I}}} & \\ 0 & \dots & 0 & \\ \vdots & \ddots & \vdots & \\ 0 & \dots & 0 & \end{bmatrix}}^{J} \overbrace{\begin{bmatrix} z_1 \\ \vdots \\ z_{p_I} \end{bmatrix}}^{z} \in \mathcal{N},$$

where we have denoted by $\mathbf{1}_j$ the j-dimensional column vector of ones. Note that J has orthonormal columns ($J^T J = I$) and hence $\|Jz\| = \|z\|$. We can project $\lambda = Jz$ onto the continuous many-sided traces by multiplying by K. We are therefore trying to estimate $\frac{\|KJz\|}{\|Jz\|} = \frac{\|KJz\|}{\|z\|}$. We are looking for the principal eigenvalue of $Z := J^T K^2 J = J^T K J$. We have that $Z = DWD$, where

$$D = \operatorname{diag}\left\{ \frac{1}{\sqrt{n_{\Gamma 1}}}, \dots, \frac{1}{\sqrt{n_{\Gamma p_I}}} \right\} \quad \text{and} \quad W = (W_{ij}) = \left(\sum_{\mathbf{x}_k \in \partial\Omega_i \cap \partial\Omega_j} \frac{1}{m_k} \right). \tag{3.16}$$

Since J has orthonormal columns and since K is an orthogonal projection, we have that $\rho(Z) \leq 1$ and $\rho(Z) = 1 - \lambda_{\min}(I - Z)$ and hence

$$\|KJz\| \leq \sqrt{1 - \lambda_{\min}(I - Z)} \|Jz\|. \tag{3.17}$$

It turns out that $I - Z$ is the stiffness matrix of a graph Laplacian. For simplicity of exposition, we treat a special case. We now assume that the domain decomposition is gridlike.

DEFINITION 3.4. *A domain decomposition is* **gridlike of diameter** ℓ *in 2d (resp. 3d) if the following properties hold:*

1) There is a homeomorphism Φ such that

$$\begin{aligned} \Phi(\Gamma) = \Phi(\Omega) \cap \{(x,y) \mid x \in \{1,\dots,\ell\} \text{ or } y \in \{1,\dots,\ell\} \} \text{ (2d case), or} \\ \Phi(\Gamma) = \Phi(\Omega) \cap \{(x,y,z) \mid x \in \{1,\dots,\ell\} \text{ or } y \in \{1,\dots,\ell\} \text{ or} \\ z \in \{1,\dots,\ell\} \} \text{ (3d case).} \end{aligned}$$

In particular, each subdomain Ω_i is either adjacent to the natural boundary $\partial\Omega$, or $\Phi(\Omega_i)$ is a square (resp. a cube).

2) The subdomain Ω_i **floats** *(i.e. S_i has kernel spanned by the vector $\mathbf{1}$ of ones) if and only if $\partial\Omega_i \cap \partial\Omega$ is empty. In other words, the kernel $\mathcal{N}$ of S is the set of many-valued traces that are constant on each subdomain, and that are zero on the subdomains that are adjacent to the natural boundary $\partial\Omega$.*

3) For subdomains that float, the number of degrees of freedom m_e is the same along each edge of the square (in 3d, we assume that there are m_e degrees of freedom per edge and m_e^2 per face). We further assume that m_e is the same for each floating subdomain. In 2d, each squarelike subdomain consists

of $4m_e$ edge degrees of freedom, and 4 cross points, hence there are $n_{\Gamma k} = 4m_e + 4$ degrees of freedom per floating subdomain. (In 3d, we get $n_{\Gamma k} = 6m_e^2 + 12m_e + 8$ degrees of freedom per floating subdomain.) Therefore, the trace $u_{\Gamma k}$ of u_k on the floating subdomain Ω_k is a vector in $\mathbb{R}^{4m_e+4}$ (resp. $u_{\Gamma k} \in \mathbb{R}^{6m_e^2+12m_e+8}$).

REMARK 3.5. *In a typical situation where Ω has unit diameter and all the subdomains are approximately the same size, then the Euclidian diameter H of the subdomains is roughly $1/\ell$. Furthermore, for a gridlike decomposition, the coefficient M_0 of (3.3) is $M_0 = 3$ in 2d, and $M_0 = 7$ in 3d, since $M - G$ block diagonalizes with $\tilde{\Pi}$ (cf. (2.3) and (2.6)) and the condition of $M - G$ is given by the 4×4 case in 2d, or the 8×8 case in 3d.*[8]

LEMMA 3.6. *Let $\Omega_1, \ldots, \Omega_p$ be a gridlike domain decomposition of diameter $\ell \geq 2$ of Ω. Let c be a continuous many-sided trace and let $\mathcal{N}$ be the space of piecewise constant many-sided traces that are zero everywhere except on the floating subdomains. Let E be the orthogonal projection onto $\mathcal{N}$. Then we have the inequality*

$$\|Ec\| \leq \overbrace{\sqrt{1 - C_{\mathcal{N}} \ell^{-2}}}^{r} \|c\| \tag{3.18}$$

where

$$C_{\mathcal{N}} = \frac{\pi^2}{4} \approx 2.47 \qquad \textit{(2d case), or} \qquad C_{\mathcal{N}} = \frac{3\pi^2}{16} \approx 1.85 \qquad \textit{(3d case).} \tag{3.19}$$

Proof. We give the proof for the 2d case, the 3d case is proven in a completely analogous manner. For a planar gridlike domain decomposition with $m = m_e$ degrees of freedom along each edge, we find from (3.16) that

$$(I - Z)_{ij} = \begin{cases} \dfrac{8m + 12}{16m + 16} & \text{if } i = j, \\ -\dfrac{2m + 2}{16m + 16} & \text{if } \Omega_i \text{ and } \Omega_j \text{ share an edge,} \\ -\dfrac{1}{16m + 16} & \text{if } \Omega_i \text{ and } \Omega_j \text{ share only a corner, and} \\ 0 & \text{otherwise.} \end{cases}$$

We write $(I - Z)$ as $X + Y$, where

$$X_{ij} = \begin{cases} \dfrac{8m + 8}{16m + 16} = \dfrac{1}{2} & \text{if } i = j, \\ -\dfrac{2m + 2}{16m + 16} = -\dfrac{1}{8} & \text{if } \Omega_i \text{ and } \Omega_j \text{ share an edge,} \\ 0 & \text{otherwise.} \end{cases}$$

[8] In the non-gridlike case, one can show from (2.2) and (2.6) that $M_0 = \max_k m_k - 1$, which is only large in degenerate cases where a cross point is adjacent to a large number of subdomains. We also note that even if the choice of $\check{M}_k$ is not the one given by (2.2), the only impact is to obtain a different value of M_0 and we still have spectral equivalence of A_{2LM} and A_{S2LM}, provided that one makes a "moderate" choice for $\check{M}_k$. This also shows that all moderate choices of $\check{M}_k$ lead to spectrally equivalent algorithms, since A_{S2LM} is independent of any choice of a symmetric M. A_{S2LM} is independent of the choice of M because (3.1) holds for any symmetric or normal choice of M whose kernel is the range of G.

and

$$Y_{ij} = \begin{cases} \dfrac{4}{16m+16} & \text{if } i = j, \\ -\dfrac{1}{16m+16} & \text{if } \Omega_i \text{ and } \Omega_j \text{ share only a corner, and} \\ 0 & \text{otherwise.} \end{cases}$$

Because Ω_i shares only a corner with precisely four adjacent subdomains, we see that Y is diagonally semidominant and hence semidefinite. Likewise, because Ω_i shares an edge with precisely four adjacent subdomains, X is also semidefinite (and, as we will see shortly, positive definite). Hence, $\lambda_{\min}(I - Z) \geq \lambda_{\min}(X)$. The matrix X is the graph Laplacian of the subdomain connectivity graph. To estimate the smallest eigenvalue of X, note that X is a submatrix of $X_\square$, the $(\ell-1)^2 \times (\ell-1)^2$ finite difference Laplacian for the square of side $\ell + 1$ with a 5-point scheme and with homogeneous Dirichlet conditions around the perimeter. The eigenvalues of $X_\square$ were computed in [34] and we obtain the desired result.

In 3d, the proof is completely analogous and we find that

$$C_{\mathcal{N}} \geq \frac{3(m_e^2 + 2m_e + 1)\pi^2}{4(3m_e^2 + 6m_e + 4)} \geq \frac{3\pi^2}{16},$$

as required. ∎

3.2. Condition numbers of A_{2LM} and A_{S2LM}. Our second main result is algebraic.

THEOREM 3.7 (Condition number estimates for A_{2LM} and A_{S2LM}, algebraic case). *Assume that A is symmetric and positive definite. Let Ω be a domain and $\Omega_1, \ldots, \Omega_p$ be a domain decomposition with restriction matrices $R_1, \ldots, R_p$. Assume that S is symmetric and semi-definite. Let $s_{\min} > 0$ be the smallest* **nonzero** *eigenvalue of S, and $s_{\max}$ be the largest eigenvalue of S, with "nonsingular" condition number $\mathcal{K}_0(S) = \frac{s_{\max}}{s_{\min}}$. Let $B = a_{\text{opt}} I$ where*

$$a_{\text{opt}} = \sqrt{s_{\max} s_{\min}}. \tag{3.20}$$

Define K by (2.7) and A_{S2LM} by (3.2). Assume that the domain decomposition is gridlike of diameter ℓ. For large values of $\mathcal{K}_0(S)$ and of ℓ, we have the following estimates:

$$\mathcal{K}(A_{\text{S2LM}}) \leq \frac{8}{\pi^2}\ell^2\sqrt{\mathcal{K}_0(S)} \text{ (2d)}, \qquad \mathcal{K}(A_{\text{S2LM}}) \leq \frac{32}{3\pi^2}\ell^2\sqrt{\mathcal{K}_0(S)} \text{ (3d)}, \tag{3.21}$$

$$\mathcal{K}(A_{\text{2LM}}) \leq \frac{24}{\pi^2}\ell^2\sqrt{\mathcal{K}_0(S)} \text{ (2d)}, \qquad \mathcal{K}(A_{\text{2LM}}) \leq \frac{224}{3\pi^2}\ell^2\sqrt{\mathcal{K}_0(S)} \text{ (3d)}. \tag{3.22}$$

Proof. We let $Q = a(S + aI)^{-1}$ (cf. (3.2)) with eigenvalues $0 < q_1 \leq \ldots \leq q_n \leq 1$ and $P = I - Q = S(S + aI)^{-1}$ with positive eigenvalues $1 > p_1 \geq \ldots \geq p_k > 0$, as per the statement of Lemma 3.1. We calculate that

$$p_k = \min_{z \in \{s_{\min}, \ldots, s_{\max}\}} \frac{z}{z + a_{\text{opt}}} = \frac{s_{\min}}{s_{\min} + \sqrt{s_{\min} s_{\max}}} = \frac{1}{1 + \sqrt{\mathcal{K}_0(S)}} \quad \text{and} \tag{3.23}$$

$$q_1 = \min_{z \in \{0, s_{\min}, \ldots, s_{\max}\}} \frac{a_{\text{opt}}}{z + a_{\text{opt}}} = \frac{\sqrt{s_{\min} s_{\max}}}{s_{\max} + \sqrt{s_{\min} s_{\max}}} = \frac{1}{1 + \sqrt{\mathcal{K}_0(S)}} = p_k. \tag{3.24}$$

We substitute the value of r given by (3.18) and the values of p_k and q_1 into (3.7) to find that

$$\mathcal{K}(A_{\mathrm{S2LM}}) \leq \frac{1+\sqrt{\mathcal{K}_0(S)}}{1-\sqrt{1-C_{\mathcal{N}}\ell^{-2}}} = \frac{2}{C_{\mathcal{N}}}\sqrt{\mathcal{K}_0(S)}\ell^2 + \text{ smaller terms.} \tag{3.25}$$

Using (3.19) with (3.25) yields (3.21), and (3.22) follows from (3.3). The value of M_0 is $M_0 = 3$ in 2d, or $M_0 = 7$ in 3d, cf. Remark 3.5. ∎

4. Estimates for the elliptic case. The main application is for elliptic problems.

LEMMA 4.1. *Let $h > 0$ be the fine grid parameter, and let $H > 0$. Assume that the following properties hold:*

1) *$\Omega_1, \ldots, \Omega_p$ are polygons or polyhedra of diameter $H_i < H$.*
2) *For $i = 1, \ldots, p$, either Ω_i floats, or the size of the intersection of $\partial\Omega_i$ with $\partial\Omega$ is comparable to $\partial\Omega_i$.*
3) *The triangulation T_h is* quasi-uniform *(cf. [35, Definition B.3]).*
4) *The matrix A is the finite element discretization of the bilinear form*

$$a(u,v) = \int_\Omega a(\mathbf{x})\nabla u(\mathbf{x}) \cdot \nabla v(\mathbf{x})\, d\mathbf{x},$$

with piecewise polynomial basis functions. The function $a(\mathbf{x})$ is assumed to be bounded $0 < a_{\min} \leq a(\mathbf{x}) \leq a_{\max}$ in such a way that $a(u,v)$ is equivalent to the seminorm $\int_\Omega \nabla u \cdot \nabla v$. We further assume that $a(\mathbf{x})$ is constant on each subdomain Ω_i.

Then, there is a constant C_{dd}, which depends on the shape of Ω and the subdomains, but not on the grid parameter h or on the size of the subdomains, such that the inequality

$$\mathcal{K}_0(S) \leq C_{dd}\frac{H}{h}, \tag{4.1}$$

is satisfied.

Proof. By replacing each subdomain Ω_i by $\frac{1}{H_i}\Omega_i$, we may assume without loss of generality that $H_i = 1$. Let $u_{\Gamma i}$ be a (finite element) trace on $\partial\Omega_i$. If Ω_i floats, further assume that the average of $u_{\Gamma i}$ is zero (so that $u_{\Gamma i}$ is orthogonal to the kernel of S_i). According to [35, Lemma 4.10], there are constants c and C such that

$$c|u_{\Gamma i}|^2_{H^{\frac{1}{2}}(\partial\Omega_i)} \leq u_i^T S_i u_i \leq C|u_{\Gamma i}|^2_{H^{\frac{1}{2}}(\partial\Omega_i)}. \tag{4.2}$$

According to [35], all constants appearing in the present proof depend on the regularity and shape of the domain decomposition and on the elliptic operator, but do not depend on the size or number of the subdomains, or on the finite element grid parameter h.

There is also a constant c' such that $\|u_{\Gamma i}\|_{L^2(\partial\Omega_i)} \leq c'|u_{\Gamma i}|_{H^{\frac{1}{2}}(\partial\Omega_i)}$ [35, Lemma A.17]. There is yet another constant C' such that $|u_{\Gamma i}|_{H^{\frac{1}{2}}(\partial\Omega_i)} \leq \frac{C'}{\sqrt{h}}\|u_{\Gamma i}\|_{L^2(\partial\Omega_i)}$ [35, Lemma B.5]. Hence, the estimate (4.2) becomes

$$\frac{c}{(c')^2}\|u_{\Gamma i}\|^2_{L^2(\partial\Omega_i)} \leq u_i^T S_i u_i \leq \frac{C(C')^2}{h}\|u_{\Gamma i}\|^2_{L^2(\partial\Omega_i)}.$$

The spectral equivalence of $u_{\Gamma i}^T u_{\Gamma i}$ and $\|u_{\Gamma i}\|^2_{L^2(\partial\Omega_i)}$ [35, Lemma B.5 or the end of the proof of Lemma 4.11] gives (4.1). ∎

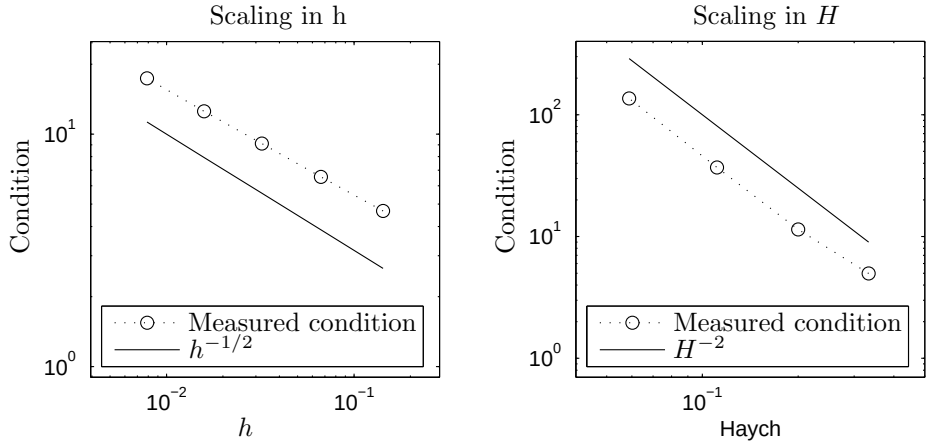

FIG. 4.1. *Scaling of the condition number of A_{S2LM}, as measured with MATLAB's* `eigs` *command, in the parameters ℓ and in h for the usual 5-point discrete Laplacian on the unit square with a regular grid and with homogeneous Dirichlet data. Subdomains are arranged as a grid of size ℓ, normalized to the unit square. Left: 3×3 subdomains ($\ell = 2$ and 4 cross points), varying the value of the finite element grid parameter $h = \frac{1}{7}, \frac{1}{15}, \frac{1}{31}, \frac{1}{63}, \frac{1}{127}$, verifying the estimate (4.3) $= O(h^{-\frac{1}{2}})$. Right: subdomains of width $H = \frac{1}{\ell+1}$ and $h = \frac{H}{5}$, varying the diameter $\ell = 2, 4, 8, 16$ (up to $17 \times 17 = 289$ subdomains and 256 cross points), verifying the estimate (4.3) $= O(H^{-2})$.*

Our third main result is the condition number estimates for the 2LM method for an elliptic problem.

THEOREM 4.2 (Condition number estimates for A_{2LM} and A_{S2LM} for elliptic PDEs). *Let Ω be a domain of unit diameter and $\Omega_1, \ldots, \Omega_p$ be a domain decomposition with restriction matrices $R_1, \ldots, R_p$. Consider an elliptic problem and assume the domain decomposition is gridlike of diameter ℓ with subdomains of diameter $H \approx 1/\ell$. Assume that the hypotheses of Lemma 4.1 are satisfied. Let $a = a_{\mathrm{opt}} > 0$ be the optimized parameter choice (3.20), and let $B = aI$. Define K by (2.7) and A_{S2LM} by (3.2). Define further M by (2.2), G by (2.5) and A_{2LM} by (2.12). Then, when h and H are sufficiently small, the condition numbers $\mathcal{K}(A_{\mathrm{S2LM}})$ and $\mathcal{K}(A_{\mathrm{2LM}})$ are bounded by*

$$\mathcal{K}(A_{\mathrm{S2LM}}), \mathcal{K}(A_{\mathrm{2LM}}) \leq C H^{-\frac{3}{2}} h^{-\frac{1}{2}}, \tag{4.3}$$

where the constant C depends on the regularity of the elliptic form $a(u, v)$ as well as the shape of Ω and the shapes of the subdomains, but not on the sizes or number of subdomains, nor on the parameter h of the triangulation T_h.

The proof is by substituting the estimate (4.1) into the condition number estimates of Theorem 3.7.

5. Conclusions. We have given a new optimized 2-Lagrange multiplier method and provided condition number estimates. The condition number estimates are consistent with the optimized Schwarz literature and are verified by numerical experiments.

Acknowledgements. We are grateful to our colleagues Martin J. Gander, Felix Kwok and Hui Zhang for useful discussions.

REFERENCES

[1] A. Alonso-Rodriguez and L. Gerardo-Giorda, *New non-overlapping domain decomposition methods for the time-harmonic Maxwell system*, SIAM Journal on Scientific Computing, 28 (2006), pp. 102–122.

[2] D. Bennequin, M. J. Gander, and L. Halpern, *A homographic best approximation problem with application to optimized Schwarz waveform relaxation*, 78 (2009), pp. 185–223.

[3] P. Chevalier and F. Nataf, *Symmetrized method with optimized second-order conditions for the Helmholtz equation*, Contemporary Mathematics, 218 (1998), pp. 400–407.

[4] Q. Deng, *An optimal parallel nonoverlapping domain decomposition iterative procedure*, SIAM Journal on Numerical Analysis, (2003), pp. 964–982.

[5] V. Dolean, M. J. Gander, and L. Gerardo-Giorda, *Optimized Schwarz methods for Maxwell's equations*, SIAM Journal on Scientific Computing, 31 (2009), pp. 2193–2213.

[6] V. Dolean, S. Lanteri, and F. Nataf, *Optimized interface conditions for domain decomposition methods in fluid dynamics*, International Journal on Numerical Methods in Fluids, 40 (2002), pp. 1539–1550.

[7] O. Dubois, *Optimized Schwarz methods for the advection-diffusion equation and for problems with discontinuous coefficients*, PhD thesis, Department of Mathematics and Statistics, McGill University, 2007.

[8] O. Dubois and S. H. Lui, *Convergence estimates for an optimized schwarz method for PDEs with discontinuous coefficients*, Numerical Algorithms, 51 (2009), pp. 115–131.

[9] Charbel Farhat, Antonini Macedo, Michel Lesoinne, Francois-Xavier Roux, Frédéric Magoulès, and Armel de La Bourdonnaie, *Two-level domain decomposition methods with lagrange multipliers for the fast iterative solution of acoustic scattering problems*, Computer Methods in Applied Mechanics and Engineering, 184 (2000), pp. 213–239.

[10] Charbel Farhat and François-Xavier Roux, *A method of finite element tearing and interconnecting and its parallel solution algorithm*, International Journal for Numerical Methods in Engineering, 32 (1991), pp. 1205–1227.

[11] M. J. Gander, *Optimized Schwarz methods*, SIAM Journal on Numerical Analysis, 44 (2006), pp. 699–731.

[12] Martin J. Gander, *Schwarz methods in the course of time*, Electronic Transactions on Numerical Analysis, 31 (2008), pp. 228–255.

[13] M. J. Gander and L. Halpern, *Absorbing boundary conditions for the wave equation and parallel computing*, Mathematics of Computations, 74 (2004), pp. 153–176.

[14] ———, *Optimized Schwarz waveform relaxation for advection reaction diffusion problems*, SIAM Journal on Numerical Analysis, 45 (2007), pp. 666–697.

[15] M. J. Gander, L. Halpern, and C. Japhet, *Optimized Schwarz algorithms for coupling convection and convection-diffusion problems*, in Proceedings of the Thirteenth International Conference on Domain Decomposition Methods, N. Debit, M. Garbey, R. Hoppe, J. Périaux, and Y. Kuznetsov D. Keyes, eds., ddm.org, 2001, pp. 253–260.

[16] M. J. Gander, L. Halpern, and F. Nataf, *Optimal convergence for overlapping and non-overlapping Schwarz waveform relaxation*, in Proceedings of the Eleventh International Conference on Domain Decomposition Methods, C.-H. Lai, P. E. Bjorstad, M. Cross, and O. Widlund, eds., ddm.org, 1999, pp. 27–36.

[17] ———, *Optimal Schwarz waveform relaxation for the one dimensional wave equation*, SIAM Journal on Numerical Analysis, 41 (2003), pp. 1643–1681.

[18] M. J. Gander and F. Kwok, *Best Robin parameters for optimized Schwarz methods at cross points*, in preparation, (2010).

[19] M. J. Gander, F. Magoules, and F. Nataf, *Optimized Schwarz methods without overlap for the Helmholtz equation*, SIAM Journal on Scientific Computing, 24 (2002), pp. 38–60.

[20] C. Japhet, *Conditions aux limites artificielles et décomposition de domaine: Méthode OO2 (optimisé d'ordre 2). application à la résolution de problèmes en mécanique des fluides*, CMAP (Ecole Polytechnique), 373 (1997).

[21] ———, *Optimized Krylov-Ventcell method. Application to convection-diffusion problems.*, in Proceedings of the Ninth International Conference on Domain Decomposition Methods, P. E. Bjørstad, M. S. Espedal, and D. E. Keyes, eds., ddm.org, 1998, pp. 382–389.

[22] C. Japhet, F. Nataf, and F. Rogier, *The optimized order 2 method. application to convection-diffusion problems.*, Future Generation Computer Systems, 17 (2001), pp. 17–30.

[23] J.-H. Kimn, *A convergence theory for an overlapping Schwarz algorithm using discontinuous iterates*, Numerische Mathematik, 100 (2005), pp. 117–139.

[24] P.-L. Lions, *On the Schwarz alternating method III: a variant for non-overlapping subdomains*, in Third international symposium on domain decomposition methods for partial differential equations, T. F. Chan, R. Glowinski, J. Périaux, and O. Widlund, eds., SIAM, Philadelphia, 1990, pp. 47–70.

[25] S. Loisel, J. Côté, M. J. Gander, L. Laayouni, and A. Qaddouri, *Optimized domain decomposition methods for the spherical Laplacian*, SIAM Journal on Numerical Analysis, (To appear, 2010).

[26] Sébastien Loisel and Daniel B. Szyld, *On the geometric convergence of optimized schwarz methods with applications to elliptic problems*, Numerische Mathematik, 114 (2009), pp. 697–728.

[27] G. Lube, L. Mueller, and F.-C. Otto, *A non-overlapping domain decomposition method for the advection-diffusion problem*, Computing, 64 (2000), pp. 49–68.

[28] V. Martin, *An optimized Schwarz waveform relaxation method for unsteady convection diffusion equation*, Applied Numerical Mathematics, 52 (2005), pp. 401–428.

[29] ———, *Schwarz waveform relaxation algorithms for the linear viscous equatorial shallow water equations*, SIAM Journal on Scientific Computing, 31 (2009), pp. 3595–3625.

[30] F. Nataf, *Absorbing boundary conditions in block Gauss-Seidel methods for convection problems*, Mathematical Models and Methods in Applied Sciences, 6 (1996), pp. 481–502.

[31] F. Nataf and F. Nier, *Convergence rate of some domain decomposition methods for overlapping and nonoverlapping subdomains*, Numerische Mathematik, 75 (1997), pp. 357–377.

[32] A. Qaddouri, L. Laayouni, S. Loisel, J. Côté, and M. J. Gander, *Optimized Schwarz methods with an overset grid for the shallow-water equations: preliminary results.*, Applied Numerical Mathematics, 58 (2008), pp. 459–471.

[33] François-Xavier Roux, *Optimization of interface operator based on algebraic approach*, in Domain Decomposition Methods in Science and Engineering, Ismael Herrera, David E. Keyes, Olof B. Widlund, and Robert Yates, eds., vol. 70 of Lecture Notes in Computational Science and Engineering, National Autonomous University of Mexico, 2002, pp. 297–304.

[34] D. E. Rutherford, *Some continuant determinants arising in physics and chemistry*, Proceedings of the Royal Society of Edinburgh Section A, 62 (1947), pp. 229–236.

[35] Andrea Toselli and Olof B. Widlund, *Domain Decomposition Methods – Algorithms and Theory*, vol. 34 of Springer Series in Computational Mathematics, Springer Berlin Heidelberg, 2005.