Received: 19 July 2018 | Revised: 5 July 2019 | Accepted: 23 September 2019

DOI: 10.1002/nla.2279

RESEARCH ARTICLE

WILEY

Synchronous and asynchronous optimized Schwarz methods for one-way subdivision of bounded domains

Mireille El Haddad[1,2] | **José C. Garay**[1,3] | **Frédéric Magoulès**[4] | **Daniel B. Szyld**[1]

[1]Department of Mathematics, Temple University, Philadelphia, Pennsylvania

[2]Department of Mathematics, St. Joseph University, Beirut, Lebanon

[3]Center for Computation and Technology, Louisiana State University, Baton Rouge, Louisiana

[4]Department of Applied Mathematics, CentraleSupélec, Châtenay-Malabry, France

Correspondence
Daniel B. Szyld, Department of Mathematics (038-16), Temple University, 1805 N. Broad Street, Philadelphia, PA 19122-6094.
Email: szyld@temple.edu

Funding information
U.S. Department of Energy, Grant/Award Number: DE-SC0016578

Summary

Convergence of both synchronous and asynchronous optimized Schwarz algorithms for the shifted Laplacian operator on a bounded rectangular domain, in a one-way subdivision of the computational domain, with overlap, is shown. Convergence results are obtained under very mild conditions on the size of the subdomains and on the amount of overlap. A couple of results are also given, relating the convergence rate of the asynchronous method to changes in the size of the domain. Numerical experiments illustrate the theoretical results.

KEYWORDS

asynchronous iterations, bounded domains, domain decomposition, optimized Schwarz methods, parallel computing

AMS CLASSIFICATION

65F10; 65N22; 15A06

1 | INTRODUCTION

Schwarz methods are a class of domain decomposition methods for the solution of partial differential equations (PDEs) in which the domain is divided into (usually overlapping) subdomains. Thus, additional artificial boundaries are created. In classical Schwarz methods, mimicking the original paper by Schwarz,[1] one uses Dirichlet boundary conditions on these artificial interfaces. The method proceeds iteratively by solving in each iteration the PDE in each subdomain using boundary data from the neighboring subdomains, until convergence. For full description; see, for example, the books of Dolean et al.[2] and Toselli et al.[3]

If instead of Dirichlet data one uses boundary conditions containing an operator of the form $\frac{\partial}{\partial \nu} + \Lambda$, where ν is the normal derivative pointing outwards and Λ is an approximation of the (optimal) Steklov–Poincaré operator by a linear combination of differential operators, one has one or more parameters that can be optimized for faster convergence. These methods are called optimized Schwarz; see, for example, other works.[2,4–7] There is more than one family of transmission conditions that can be used for a given PDE, each of these families consisting of a particular approximation of the optimal transmission conditions. For example, we have the $OO0$ and $OO2$ families. In the $OO0$ family, Λ is the zeroth-order approximation of the Steklov–Poincaré operator, that is, $\Lambda = \alpha$, where α is a constant. As for the $OO2$ family, one uses $\Lambda = \alpha + \beta\Delta$, where α and β are constants and Δ is the Laplace–Beltrami operator. The $OO0$ family corresponds to Robin boundary conditions and the $OO2$ to Ventcel boundary conditions.*

*In principle, these real constants α and β, can be of any sign. In practice, the optimal values are found to be for the cases $\alpha > 0$ and $\beta \leq 0$, and one could restrict their sign accordingly; cf. the work of Gander.[5]

Domain decomposition methods were designed specifically for parallel computing environments, because each local problem, that is, the problem in each subdomain, can be solved independently of the others, and only the appropriate boundary data need to be communicated to the other processors. As the number of processors in today's computing architectures grow dramatically, and the communication costs are overtaking the computing costs, there is a need to find methods, which lessen or eliminate the synchronization points in the programs.

Asynchronous methods refer to parallel iterative procedures, where each processor performs its task without waiting for other processors to be completed, that is, with whatever information it has locally available and with no synchronizations with other processes. Thus, asynchronous optimized Schwarz methods are those in which each local problem is solved with Robin-type data from neighboring subdomains, but in which there is no synchronization.

Convergence analysis and numerical experiments with asynchronous (classical) Schwarz methods can be found in other works.[8–11] Recently, in the works of Magoulès et al.[12] and Garay et al.,[13] convergence of synchronous and asynchronous optimal (and optimized) Schwarz has been proved for PDEs on an unbounded two-dimensional domain, in a one-way subdivision of the plane. While this was the first proof of the convergence of asynchronous optimized Schwarz, and this gave the numerical experiments a sound theoretical basis, the fact that it was for the whole plane, left open the problem of analyzing convergence for more practical bounded domains. More recently, in the work of Garay et al.,[14] a convergence analysis of the solution of Poisson's equation using optimized Schwarz on a rectangular domain has been presented, in which the domain is divided into $p \times q$ (overlapping) rectangles in a 2D configuration. That is, the artificial interfaces have cross-points. Unfortunately, the conditions for convergence were only applicable for a very small number of p and q.

In this paper, we consider a rectangular domain, but this time we consider a one-way subdivision. We are able to prove convergence for any number of subdomains. In a sense, we are able to offer a more practical convergence result than that of the whole plane in the work of Magoulès et al.,[12] and that of Garay et al.,[14] but for a less general partition of the subdomains. We use techniques taken from both of these papers, but tailored to our particular setup. In other words, here we fill a gap in the theory; we are able to prove convergence of asynchronous optimized Schwarz on a bounded domain for an arbitrary number of subdomains. We mention that the technique used in our proofs extend to the three-dimensional case, in a one-way subdivision, but we do not present any details on this.

In addition to our convergence results, we present an empirical formula for the Robin-type optimal parameter in terms of the overlap. We also present two results that relate the convergence rate of the asynchronous method to the size of the subdomain. In one case, we show that keeping the height constant, as well as the amount of overlap, then increasing the width, accelerates the convergence. Equivalently, for the same domain, increasing the number of subdomains (and thus decreasing the width of each subdomain), the convergence factor deteriorates. On the other hand, we show that maintaining each subdomain of the same size, changing how many subdomains we have, does not change the convergence factor.

The paper is structured as follows. In Section 2, we present the description of the synchronous optimized Schwarz algorithm for the shifted Laplacian operator in a rectangle $[0, L_1] \times [0, L_2]$ with Dirichlet boundary conditions on the physical boundaries. We prove its convergence using Fourier sine series and the maximum norm of an operator mapping the coefficients of these series from one step to the next. The theorems apply when either $OO0$ or $OO2$ artificial boundary conditions are used. They are proved under very mild conditions on the overlap between the subdomains. In Section 3, we prove the convergence of the asynchronous version of the algorithm. This requires also to prove that the Fourier sine series are uniformly convergent. In Section 4, we present two new theorems that deal with how the convergence factor relates to changes in the domain parameters. In Section 5, we show results of some numerical experiments on how the spectral radius of the abovementioned operator varies with different values of the Robin parameter for the $OO0$ case, the overlap, the number of subdomains, and the frequency parameter. Then, we produce an empirical formula for the optimal Robin parameter in the $OO0$ case in terms of the overlap. In Section 6, we present numerical experiments analyzing the thermal distribution on an air intake system. These experiments show that the asynchronous methods indeed converge and that they are faster than the synchronous counterpart.

2 | SYNCHRONOUS SCHWARZ WITH OPTIMIZED INTERFACE CONDITIONS

We consider the equation

$$\begin{cases} (\Delta - \eta)u &= f \ \text{ in } \Omega, \\ u &= g \ \text{ on } \partial\Omega, \end{cases} \tag{1}$$

with $\Omega = (0, L_1) \times (0, L_2)$ and $\eta \geq 0$.

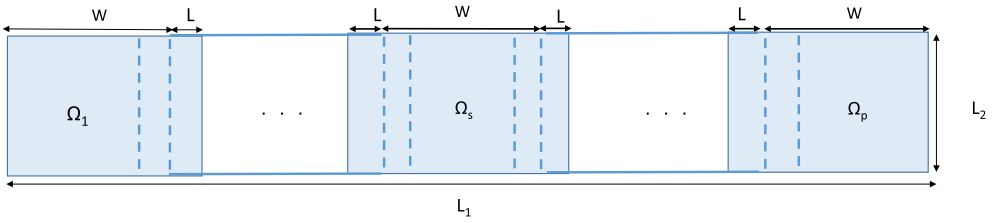

FIGURE 1 One-way domain decomposition

The domain Ω is divided into p overlapping subdomains, where the overlap are vertical strips. This means we have $p-1$ vertical lines, say, at coordinates $x = \ell_1, \ldots, \ell_{p-1}$, and we assume for simplicity that we have the same overlap $2L$ between the subdomains and that the width of each subdomain as if there were no overlap is $\ell_s - \ell_{s-1} = W = L_1/p$ for $s = 2, \ldots, p-1$, so that $\ell_s = \ell_1 + (s-1)W$. It follows then that the overlap satisfies $2L < W$. Thus, we have $\Omega_1 = [0, \ell_1 + L] \times [0, L_2]$, $\Omega_s = [\ell_{s-1} - L, \ell_s + L] \times [0, L_2]$, $s = 2, \ldots, p-1$, and $\Omega_p = [\ell_{p-1} - L, L_1] \times [0, L_2]$; see Figure 1. Our analysis would equally apply to situations in which the overlap varies from one subdomain to the next, including the special case where the first and last subdomains have the same width as the others, and when the subdomains have different widths. In addition, if the one-way subdivision was to be horizontal instead of vertical, our analysis would obviously apply as well.

Let $u^n = (u_1^n, \ldots, u_p^n)$, where u_s^n is the approximation of u_s^* (the restriction of the solution u^* to the subdomain Ω_s) in Ω_s at the nth iteration, and f_s the restriction of f to Ω_s, $s = 1, \ldots, p$. Thus, $u_s^n \in X_s$, a space of functions defined on Ω_s.

We consider transmission conditions (on the artificial interfaces) composed of local operators of the form $\frac{\partial}{\partial \nu} + \Lambda$, where Λ is an approximation of the Steklov–Poincaré operator by a linear combination of differential operators. In this context, the normal vector is along the x-direction (with the appropriate sign),

$$\Lambda = \alpha \tag{2}$$

for the $OO0$ case and

$$\Lambda = \alpha + \beta \frac{\partial^2}{\partial y^2} \tag{3}$$

for the $OO2$ case.

Starting with an initial approximation $u^0 = (u_1^0, \ldots, u_p^0)$, the *local problems* and the *synchronous* iteration process are described by the following equations:

$$\begin{cases}
\text{For } s = 1, \text{solve} & \\
(\Delta - \eta) u_1^{n+1} = f_1 & \text{on } \Omega_1, \\
u_1^{n+1} = g & \text{for } x = 0, \\
\frac{\partial u_1^{n+1}}{\partial x} + \Lambda\left(u_1^{n+1}\right) = \frac{\partial u_2^n}{\partial x} + \Lambda\left(u_2^n\right) & \text{for } x = \ell_1 + L, \\
u_s^{n+1} = g & \text{for } y = 0, \text{ and } y = L_2, \\
\text{For } s = 2, \ldots, p-1, \text{solve} & \\
(\Delta - \eta) u_s^{n+1} = f_s & \text{on } \Omega_s, \\
-\frac{\partial u_s^{n+1}}{\partial x} + \Lambda\left(u_s^{n+1}\right) = -\frac{\partial u_{s-1}^n}{\partial x} + \Lambda\left(u_{s-1}^n\right) & \text{for } x = \ell_{s-1} - L, \\
\frac{\partial u_s^{n+1}}{\partial x} + \Lambda\left(u_s^{n+1}\right) = \frac{\partial u_{s+1}^n}{\partial x} + \Lambda\left(u_{s+1}^n\right) & \text{for } x = \ell_s + L, \\
u_s^{n+1} = g & \text{for } y = 0, \text{ and } y = L_2, \\
\text{For } s = p, \text{solve} & \\
(\Delta - \eta) u_p^{n+1} = f_p & \text{on } \Omega_p \\
-\frac{\partial u_p^{n+1}}{x} + \Lambda\left(u_p^{n+1}\right) = -\frac{\partial u_{p-1}^n}{\partial x} + \Lambda\left(u_{p-1}^n\right) & \text{for } x = \ell_{p-1} - L, \\
u_p^{n+1} = g & \text{for } x = L_1, \\
u_s^{n+1} = g & \text{for } y = 0, \text{ and } y = L_2.
\end{cases} \tag{4}$$

Let $E_s^n = u^* - u_s^n$ be the error in subdomain Ω_s at iteration n. Using the definition of E_s^n and the linearity of the equations in (4), it follows that E_s^n satisfies the homogeneous case of (4), that is, E_s^n is the solution of (4) for the case $f = 0, g = 0$. That is, for $s = 2, \ldots, p$, the error solves the following local problem:

$$\begin{cases} (\Delta - \eta)E_s^{n+1} = 0 & \text{on } \Omega_s, \\ -\frac{\partial E_s^{n+1}}{\partial x} + \Lambda\left(E_s^{n+1}\right) = -\frac{\partial E_{s-1}^{n}}{\partial x} + \Lambda\left(E_{s-1}^{n}\right) & \text{for } x = \ell_{s-1} - L, \\ \frac{\partial E_s^{n+1}}{\partial x} + \Lambda\left(E_s^{n+1}\right) = \frac{\partial E_{s+1}^{n}}{\partial x} + \Lambda\left(E_{s+1}^{n}\right) & \text{for } x = \ell_s + L, \\ E_s^{n+1} = 0 & \text{for } y = 0, \text{ and } y = L_2. \end{cases} \tag{5}$$

To prove the convergence of the method, we shall show that $E_s^n \to 0$ as $n \to \infty$ for $s = 1, \ldots, p$. To that end, we use separation of variables and Fourier analysis to write the local errors as a Fourier sine series in variable y and then recast the equations (5) (and the corresponding ones for $s = 1$ and $s = p$) into a new fixed-point iteration. Then, we prove that the iteration operator of this new-fixed point iteration, which maps the coefficients of the series of all the local errors at iteration n to the coefficients at iteration $n + 1$, is contracting.

To that end, let us consider the eigenfunctions of the problem $(\Delta - \eta)\phi = 0$ on Ω_s with $\phi = 0$ for $y = 0$ and $y = L_2$. These are given by $\phi_m(y) = \sin(m\pi y/L_2)$, $m \in \mathbb{N}$. The expressions of these eigenfunctions can be obtained using separation of variables. In addition, note that $\{\phi_m\}_{m\in\mathbb{N}}$ is a complete orthogonal set that spans the space of piecewise continuous functions; see, for example, the work of Haberman.[15] Then, we can write the solution of (5) as

$$E_s(x, y) = \sum_{m=1}^{\infty} \hat{E}_s(x, m)\phi_m(y) = \sum_{m=1}^{\infty} \hat{E}_s(x, m) \sin\left(\frac{m\pi y}{L_2}\right), \tag{6}$$

where for a fixed x, $\{\hat{E}_s(x, m)\}_{m\in\mathbb{N}}$ is the set of coefficients of the expansion of $E_s(x, \cdot)$ in the basis $\{\phi_m\}_{m\in\mathbb{N}}$ and satisfies

$$\hat{E}_s(x, m) = \frac{2}{L_2} \int_0^{L_2} E_s(x, y) \sin\left(\frac{m\pi y}{L_2}\right) dy. \tag{7}$$

Note that for a fixed x, (6) is a Fourier sine series in the y variable. We remark that using these Fourier series to express the error is mentioned already in the work of Gander[16] for two subdomains; see also the work of Dolean et al.[2]

In the sequel, we develop explicit equations defining the coefficients $\hat{E}_s(x, m)$. To that end, we will use the orthogonality of the set $\{\phi_m\}_{m\in\mathbb{N}}$, and also the fact that, because $E_s(x, 0) = E_s(x, L_2)$=0, the order of integrals, first and second derivatives and summation commute when integral and differential operators are applied to the series in (6); see, for example, the work of Haberman.[15] In the interest of clarity, one or both of the arguments of the error E_s and the Fourier coefficient $\hat{E}_s$ are sometimes omitted in the presentation.

We begin by plugging (6) into the first line of (5), and interchanging the order of derivatives and summation, we obtain

$$\sum_{m=1}^{\infty} \frac{\partial^2 \hat{E}_s}{\partial x^2} \sin\left(\frac{m\pi y}{L_2}\right) + \sum_{m=1}^{\infty} \hat{E}_s \frac{\partial^2 \sin\left(\frac{m\pi y}{L_2}\right)}{\partial y^2} - \eta \sum_{m=1}^{\infty} \hat{E}_s \sin\left(\frac{m\pi y}{L_2}\right) = 0.$$

Evaluating the derivative in the second summation from the above expression and collecting terms, we can write

$$\sum_{m=1}^{\infty} \left(\frac{\partial^2 \hat{E}_s}{\partial x^2} - \left(\left(\frac{m\pi}{L_2}\right)^2 + \eta \right) \hat{E}_s \right) \sin\left(\frac{m\pi y}{L_2}\right) = 0.$$

Multiplying by $\sin(k\pi y/L_2)$, integrating between $y = 0$ and $y = L_2$, and using the orthogonality property of $\{\sin(m\pi y/L_2)\}_{m\in\mathbb{N}}$, we obtain

$$\frac{\partial^2 \hat{E}_s}{\partial x_2}\int_0^{L_2}\sin^2\left(\frac{k\pi y}{L_2}\right)dy - \left(\left(\frac{k\pi}{L_2}\right)^2+\eta\right)\hat{E}_s\int_0^{L_2}\sin^2\left(\frac{k\pi y}{L_2}\right)dy = 0.$$

Because the integral is not zero, we have that for each $k \in \mathbb{N}$, $\hat{E}_s(x,k)$ must satisfy the following ODE:

$$\frac{\partial^2 \hat{E}_s}{\partial x_2} - \left(\left(\frac{k\pi}{L_2}\right)^2+\eta\right)\hat{E}_s = 0. \tag{8}$$

We develop now boundary conditions for the ODE (8). We begin with the case $s = 1$. From the homogeneous Dirichlet boundary conditions at $x = 0$, we have that $E_1(0,y) = 0$. Then, using this fact in (7), it follows that

$$\hat{E}_1(0,k) = \frac{2}{L_2}\int_0^{L_2} E_1(0,y)\sin\left(\frac{k\pi y}{L_2}\right)dy = 0.$$

Similarly, for $s = p$, from (4), we have at $x = L_1$ that $E_p(L_1,y) = 0$. Then, from (7), we obtain

$$\hat{E}_p(L_1,k) = \frac{2}{L_2}\int_0^{L_2} E_p(L_1,y)\sin\left(\frac{k\pi y}{L_2}\right)dy = 0.$$

We write next the boundary conditions for $\hat{E}_s^{n+1}$ that arise from the artificial boundary conditions in (5). We will write the equations for the $OO2$ case, though one could easily do the same for other conditions Λ (the $OO0$ case corresponds to $\beta = 0$). Thus, using (3), we see that the artificial boundary conditions are given by

$$\begin{cases} \text{For } s = 2,\ \dots, p, \text{for } x = \ell_{s-1} - L, \\ -\frac{\partial E_s^{n+1}}{\partial x} + \alpha E_s^{n+1} + \beta\frac{\partial^2 E_s^{n+1}}{\partial y^2} = -\frac{\partial E_{s-1}^{n}}{\partial x} + \alpha E_{s-1}^{n} + \beta\frac{\partial^2 E_{s-1}^{n}}{\partial y^2}, \\ \text{For } s = 1,\ \dots, p-1, \text{for } x = \ell_s + L, \\ \frac{\partial E_s^{n+1}}{\partial x} + \alpha E_s^{n+1} + \beta\frac{\partial^2 E_s^{n+1}}{\partial y^2} = \frac{\partial E_{s+1}^{n}}{\partial x} + \alpha E_{s+1}^{n} + \beta\frac{\partial^2 E_{s+1}^{n}}{\partial y^2}. \end{cases}$$

Using the same argument as that used to obtain (8), we have that for each $k \in \mathbb{N}$, $\hat{E}_s^{n+1}(x,k)$ must satisfy

$$\begin{cases} \text{For } s = 2,\ \dots, p, \text{for } x = \ell_{s-1} - L, \\ -\frac{\partial \hat{E}_s^{n+1}}{\partial x} + \left(\alpha - \left(\frac{\pi k}{L_2}\right)^2\beta\right)\left(\hat{E}_s^{n+1}\right) = -\frac{\partial \hat{E}_{s-1}^{n}}{\partial x} + \left(\alpha - \left(\frac{\pi k}{L_2}\right)^2\beta\right)\left(\hat{E}_{s-1}^{n}\right) \\ \text{For } s = 1,\ \dots, p-1, \text{for } x = \ell_s + L, \\ \frac{\partial \hat{E}_s^{n+1}}{\partial x} + \left(\alpha - \left(\frac{\pi k}{L_2}\right)^2\beta\right)\left(\hat{E}_s^{n+1}\right) = \frac{\partial \hat{E}_{s+1}^{n}}{\partial x} + \left(\alpha - \left(\frac{\pi k}{L_2}\right)^2\beta\right)\left(\hat{E}_{s+1}^{n}\right). \end{cases} \tag{9}$$

We denote by $\lambda(k) = \alpha - \left(\frac{k\pi}{L_2}\right)^2\beta$, where $\alpha > 0, \beta \le 0$. The quantity λ depends on k; however, in the interest of clarity, the parameter k is sometimes omitted in our presentation. We also denote by $\theta(k)$ the positive root of the characteristic equation of the second-order ODE (8), namely, the root of the equation $\theta^2 - \left(\eta + \left(\frac{k\pi}{L_2}\right)^2\right) = 0$. Thus, $\theta(k) = \sqrt{\eta + \left(\frac{k\pi}{L_2}\right)^2}$, which is well defined for $\eta \ge 0$.

Collecting Equations (8)–(9), we obtain the following algorithm, which defines the coefficients $\hat{E}_s^{n+1}$ and represents a different realization of the algorithm described in (4):

$$
\begin{cases}
\text{For } s = 1, \text{solve} & \\
\frac{\partial^2 \hat{E}_1^{n+1}}{\partial x^2} - \left(\eta + \left(\frac{k\pi}{L_2}\right)^2\right)\hat{E}_1^{n+1} = 0 & \text{for } x \le \ell_1 + L, \\
\hat{E}_1^{n+1} = 0 & \text{for } x = 0, \\
\frac{\partial \hat{E}_1^{n+1}}{\partial x} + \lambda \hat{E}_1^{n+1} = \frac{\partial \hat{E}_2^{n}}{\partial x} + \lambda \hat{E}_2^{n} & \text{for } x = \ell_1 + L, \\
\text{For } s = 2, \ldots, p-1, \text{solve} & \\
-\frac{\partial \hat{E}_s^{n+1}}{\partial x} + \lambda \hat{E}_s^{n+1} = -\frac{\partial \hat{E}_s^{n}}{\partial x} + \lambda \hat{E}_s^{n} & \text{for } x = \ell_{s-1} - L, \\
\frac{\partial^2 \hat{E}_s^{n+1}}{\partial x^2} - \left(\eta + \left(\frac{k\pi}{L_2}\right)^2\right)\hat{E}_s^{n+1} = 0 & \text{for } \ell_{s-1} - L \le x \le \ell_s + L, \\
\frac{\partial \hat{E}_s^{n+1}}{\partial x} + \lambda \hat{E}_s^{n+1} = \frac{\partial \hat{E}_{s-1}^{n}}{\partial x} + \lambda \hat{E}_{s-1}^{n} & \text{for } x = \ell_s + L, \\
\text{For } s = p, \text{solve} & \\
-\frac{\partial \hat{E}_p^{n+1}}{\partial x} + \lambda \hat{E}_p^{n+1} = -\frac{\partial \hat{E}_{p-1}^{n}}{\partial x} + \lambda \hat{E}_{p-1}^{n} & \text{for } x = \ell_{p-1} - L, \\
\frac{\partial^2 \hat{E}_p^{n+1}}{\partial x^2} - \left(\eta + \left(\frac{k\pi}{L_2}\right)^2\right)\hat{E}_p^{n+1} = 0 & \text{for } \ell_{p-1} - L \le x, \\
\hat{E}_p^{n+1} = 0 & \text{for } x = L_1.
\end{cases}
\tag{10}
$$

We are ready to state and prove the convergence of the synchronous algorithm represented by (4) using its Fourier sine series coefficients counterpart (10).

Theorem 1. *Consider the shifted Laplacian problem (1) on a rectangular domain with a one-way domain decomposition with p vertical strips of width W, and for $s = 2, \ldots, p-1$, overlap 2L throughout. Consider transmission operators Λ defined on the interfaces, supposed to be the same on the left and right interfaces. Given an initial approximation $u^0 = (u_1^0, \ldots, u_p^0)$, we consider the synchronous optimized Schwarz algorithm described by (4). Let us assume that for each frequency k, the factor $\lambda(k)$, the width of the subdomain (as if there were no overlap) W, and the overlap L satisfy the following three conditions:*

$$\lambda(k) \neq -\theta(k), \tag{11}$$

$$L < W/3, \tag{12}$$

and

$$\left|\frac{\lambda(k) - \theta(k)}{\lambda(k) + \theta(k)}\right| e^{-\theta(k)(2L)} < \frac{1 - e^{-\theta(k)(2L)}}{1 + e^{-\theta(k)(W)}}. \tag{13}$$

Then, the synchronous algorithm described by (4) converges to the solution u^* for any initial approximation u^0.

We discuss here the hypotheses (11)–(13). In all our experiments, the approximate optimal conditions we used correspond to $\lambda = \alpha$ for $OO0$, or $\lambda = \alpha - \left(\frac{\pi k}{L_2}\right)^2 \beta$, with $\alpha > 0$ and $\beta < 0$, and thus, hypothesis (11) always hold. (See also Section 6 where it is shown that the optimal value of α is positive). Regarding the second hypothesis (12), usually in the numerical applications, we always have $L < W/3$. This is natural; the overlap (on each side) is not larger than a third of the width of the subdomain. Lastly, observing the hypothesis (13), we note that the left-hand side is always less than one. It is a product of two factors; the first is less than one, and the second is less than one with $L > 0$ and equal to one when $L = 0$. The term on the right of the condition is also less than one, and it is a monotonically increasing function of width W. Thus, for a fixed amount of overlap $2L$, for large-enough W, that is, for wide-enough subdomains, condition (13) is automatically satisfied. Alternatively, for fixed W, for small-enough overlap L, the conditions hold.

Proof of Theorem 1. Before we give the proof in full detail, we present a sketch of it.

1. We write the Fourier series coefficients of the error of the approximate solution in each subdomain at each iteration n, in terms of two exponentials, with the coefficients in front of these two exponentials, say, $A_s^n(k)$ and $B_s^n(k)$ to be determined.

2. We use the artificial boundary conditions on the interfaces to write relations between these coefficients at iteration $n+1$, as a function of those at iteration n. There are dependencies from the corresponding coefficient in the subdomains on the left and right of the subdomain.

3. This results in a linear map from the coefficients of the error series for all subdomains $s = 1, \dots, p$, at iteration n to all coefficients at iteration $n+1$. We explicitly write the matrix $\hat{\mathcal{T}}$ of this linear map and consider its max-norm.

4. We show that this norm is less than one, and thus, the map is a contraction implying that $\hat{E}^n \to 0$.

5. We relate the error E^n to the coefficients $\hat{E}^n$, showing that $E^n \to 0$, completing the proof.

Step 1. A general solution of the second-order ODE for a generic subdomain s in (10) has the form[†]

$$\hat{E}_s^n(x,k) = A_s^n(k)e^{\theta_1(x-(\ell_{s-1}-L))} + B_s^n(k)e^{\theta_2(x-(\ell_s+L))},\ s = 1, \dots, p, \tag{14}$$

where θ_i, $i = 1, 2$, are the two solutions of the characteristic equation of the mentioned second-order ODE, namely, $\theta_1 = -\theta(k)$, $\theta_2 = \theta(k)$.

Therefore, the solutions of the equations (10) have the following form:

$$\begin{cases} \hat{E}_1^n(x,k) = A_1^n(k)e^{-\theta(k)x} + B_1^n(k)e^{\theta(k)(x-(\ell_1+L))}, \\ \hat{E}_s^n(x,k) = A_s^n(k)e^{-\theta(k)(x-(\ell_{s-1}-L))} + B_s^n(k)e^{\theta(k)(x-(\ell_s+L))},\ s = 2, \dots, p-1, \\ \hat{E}_p^n(x,k) = A_p^n(k)e^{-\theta(k)(x-(\ell_{p-1}-L))} + B_p^n(k)e^{\theta(k)(x-L_1)}, \end{cases}$$

with the unknown coefficients $A_1^n(k), B_1^n(k), A_s^n(k), B_s^n(k), s = 2, \dots, p-1, A_p^n(k), B_p^n(k)$, to be determined. From now on, we drop parameter k in the expression of these coefficients. Direct computation of the derivatives and substitutions lead to the following equations:

$$\begin{cases} \frac{\partial \hat{E}_1^n(x,k)}{\partial x} = -\theta(k)A_1^n e^{-\theta(k)x} + \theta(k)B_1^n e^{\theta(k)(x-(\ell_1+L))}, \\ \frac{\partial \hat{E}_s^n(x,k)}{\partial x} = -\theta(k)A_s^n e^{-\theta(k)(x-(\ell_{s-1}-L))} + \theta(k)B_s^n e^{\theta(k)(x-(\ell_s+L))}, \quad s = 2, \dots, p-1, \\ \frac{\partial \hat{E}_p^n(x,k)}{\partial x} = -\theta(k)A_p^n e^{-\theta(k)(x-(\ell_{p-1}-L))} + \theta(k)B_p^n e^{\theta(k)(x-L_1)}. \end{cases}$$

Let us define

$$\kappa = (\lambda - \theta(k)), \quad \mu = (\lambda + \theta(k)) \text{ and } \gamma = \frac{\kappa}{\mu}.$$

Because by hypothesis $\lambda \neq -\theta(k)$, this leads to $\mu \neq 0$.

Step 2. The introduction of these expressions in the interface conditions implies that, for the domain Ω_1, we have the Dirichlet boundary condition on the left,

$$\hat{E}_1^{n+1}(0,k) = 0;$$

thus,

$$A_1^{n+1} = -B_1^{n+1}e^{-\theta(k)(\ell_1+L)}. \tag{15}$$

For subdomain Ω_1 receiving values from subdomain Ω_2, we have

$$\begin{aligned} \frac{\partial \hat{E}_1^{n+1}(\ell_1+L,k)}{\partial x} + \lambda \hat{E}_1^{n+1}(\ell_1+L,k) &= \frac{\partial \hat{E}_2^{n}(\ell_1+L,k)}{\partial x} + \lambda \hat{E}_2^{n}(\ell_1+L,k), \\ (\lambda - \theta(k))A_1^{n+1}e^{-\theta(k)(W+L)} + (\lambda + \theta(k))B_1^{n+1} &= (\lambda - \theta(k))A_2^{n}e^{-\theta(k)2L} + (\lambda + \theta(k))B_2^{n}e^{-\theta(k)W}, \\ \frac{\kappa}{\mu}A_1^{n+1}e^{-\theta(k)(W+L)} + B_1^{n+1} &= \frac{\kappa}{\mu}A_2^{n}e^{-\theta(k)2L} + B_2^{n}e^{-\theta(k)W}. \end{aligned} \tag{16}$$

[†]This form is general and could be used for a limiting case of an infinite domain. In the latter case, one would have $A_1^n(k) = 0$ and $B_p^n(k) = 0$.

We substitute (15) in (16) and obtain

$$B_1^{n+1} = \frac{\frac{\kappa}{\mu}A_2^n e^{-\theta(k)2L} + B_2^n e^{-\theta(k)W}}{\left(1 - \frac{\kappa}{\mu}e^{-\theta(k)2(W+L)}\right)}.$$

For subdomain Ω_p receiving values from subdomain Ω_{p-1}, we have

$$\begin{aligned}
-\frac{\partial \hat{E}_p^{n+1}(\ell_{p-1} - L, k)}{\partial x} + \lambda \hat{E}_p^{n+1}(\ell_{p-1} - L, k) &= -\frac{\partial \hat{E}_{p-1}^{n}(\ell_{p-1} - L, k)}{\partial x} + \lambda \hat{E}_{p-1}^{n}(\ell_{p-1} - L, k),\\
(\theta(k) + \lambda) A_p^{n+1} + B_p^{n+1}(-\theta(k) + \lambda) e^{-\theta(k)(\ell_{p-1} - L - L_1)} &= (\theta(k) + \lambda) A_{p-1}^n e^{-\theta(k)(\ell_{p-1} - \ell_{p-2})} + (-\theta(k) + \lambda) B_{p-1}^n e^{-\theta(k)2L}\\
A_p^{n+1} + \frac{\kappa}{\mu} B_p^{n+1} e^{-\theta(k)(W+L)} &= A_{p-1}^n e^{-\theta(k)W} + \frac{\kappa}{\mu} B_{p-1}^n e^{-\theta(k)2L}.
\end{aligned} \tag{17}$$

For the last subdomain, we have the Dirichlet boundary condition

$$\hat{E}_p^{n+1}(L_1, k) = 0;$$

thus,

$$B_p^{n+1} = -A_p^{n+1} e^{-\theta(k)(W+L)}. \tag{18}$$

We substitute (18) in (17) to obtain

$$A_p^{n+1} = \frac{A_{p-1}^n e^{-\theta(k)W} + \frac{\kappa}{\mu} B_{p-1}^n e^{-\theta(k)2L}}{\left(1 - \frac{\kappa}{\mu}e^{-\theta(k)2(W+L)}\right)}.$$

For the subdomain Ω_2 receiving values from subdomain Ω_1,

$$\begin{aligned}
-\frac{\partial \hat{E}_2^{n+1}(\ell_1 - L, k)}{\partial x} + \lambda \hat{E}_2^{n+1}(\ell_1 - L, k) &= -\frac{\partial \hat{E}_1^{n}(\ell_1 - L, k)}{\partial x} + \lambda \hat{E}_1^{n}(\ell_1 - L, k),\\
(\theta(k) + \lambda) A_2^{n+1} + (-\theta(k) + \lambda) B_2^{n+1} e^{-\theta(k)(\ell_2 - \ell_1 + 2L)} &= (\theta(k) + \lambda) e^{-\theta(k)(W-L)} A_1^n + (-\theta(k) + \lambda) B_1^n e^{-\theta(k)2L},\\
A_2^{n+1} + \frac{\kappa}{\mu} e^{-\theta(k)(W+2L)} B_2^{n+1} &= e^{-\theta(k)(W-L)} A_1^n + \frac{\kappa}{\mu} B_1^n e^{-\theta(k)2L}.
\end{aligned}$$

For the subdomain Ω_{p-1} receiving values from subdomain Ω_p,

$$\begin{aligned}
\frac{\partial \hat{E}_{p-1}^{n+1}(\ell_{p-1} + L, k))}{\partial x} + \lambda \hat{E}_{p-1}^{n+1}(\ell_{p-1} + L, k) &= \frac{\partial \hat{E}_p^{n}(\ell_{p-1} + L, k)}{\partial x} + \lambda \hat{E}_{p-1}^{n}(\ell_{p-1} + L, k),\\
(-\theta(k) + \lambda) A_{p-1}^{n+1} e^{-\theta(k)(\ell_{p-1} - \ell_{p-2} + 2L)} + (\theta(k) + \lambda) B_{p-1}^{n+1} &= (-\theta(k) + \lambda) A_p^n e^{-\theta(k)2L} + (\theta(k) + \lambda) B_p^n e^{-\theta(k)(l_{p-1} - L - L_1)},\\
\frac{\kappa}{\mu} e^{-\theta(k)(W+2L)} A_{p-1}^{n+1} + B_{p-1}^{n+1} &= \frac{\kappa}{\mu} A_p^n e^{-\theta(k)2L} + B_p^n e^{-\theta(k)(W-L)}.
\end{aligned}$$

For $s = 3, \ldots, p-1$, we have for the subdomain Ω_s receiving values from subdomain Ω_{s-1},

$$\begin{aligned}
-\frac{\partial \hat{E}_s^{n+1}(\ell_{s-1} - L, k)}{\partial x} + \lambda \hat{E}_s^{n+1}(\ell_{s-1} - L, k) &= -\frac{\partial \hat{E}_{s-1}^{n}(\ell_{s-1} - L, k)}{\partial x} + \lambda \hat{E}_{s-1}^{n}(\ell_{s-1} - L, k),\\
(\theta(k) + \lambda) A_s^{n+1} + (-\theta(k) + \lambda) B_s^{n+1} e^{-\theta(k)(\ell_s - \ell_{s-1} + 2L)} &= (\theta(k) + \lambda) A_{s-1}^n e^{-\theta(k)(\ell_{s-1} - \ell_{s-2})} + (-\theta(k) + \lambda) B_{s-1}^n e^{-\theta(k)2L},\\
A_s^{n+1} + \frac{\kappa}{\mu} e^{-\theta(k)(W+2L)} B_s^{n+1} &= A_{s-1}^n e^{-\theta(k)W} + \frac{\kappa}{\mu} B_{s-1}^n e^{-\theta(k)2L}.
\end{aligned}$$

In addition, for $s = 2, \ldots, p-2$, we have for the subdomain Ω_s receiving values from subdomain Ω_{s+1},

$$\frac{\partial \hat{E}_s^{n+1}(\ell_s + L, k)}{\partial x} + \lambda \hat{E}_s^{n+1}(\ell_s + L, k) = \frac{\partial \hat{E}_{s+1}^{n}(\ell_s + L, k)}{\partial x} + \lambda \hat{E}_{s+1}^{n+1}(\ell_s + L, k),$$

$$(-\theta(k) + \lambda) A_s^{n+1} e^{-\theta(k)(\ell_s - \ell_{s-1} + 2L)} + (\theta(k) + \lambda) B_s^{n+1} = (-\theta(k) + \lambda) A_{s+1}^{n} e^{-\theta(k)2L} + \quad (\theta(k) + \lambda) B_{s+1}^{n} e^{-\theta(k)(\ell_{s+1} - \ell_s)},$$

$$\frac{\kappa}{\mu} e^{-\theta(k)(W+2L)} A_s^{n+1} + B_s^{n+1} = \frac{\kappa}{\mu} A_{s+1}^{n} e^{-\theta(k)2L} + B_{s+1}^{n} e^{-\theta(k)W}.$$

Step 3. We now rewrite the 2×2 system formed by the two equations from above, which corresponds to the exchange between the two interfaces of the subdomains. Let $\gamma = \kappa/\mu = |\lambda - \theta(k)|/|\lambda + \theta(k)|$ as before, we then have

$$\begin{bmatrix} A_1^{n+1} \\ B_1^{n+1} \end{bmatrix} = \frac{1}{1 - \gamma e^{-2\theta(k)(W+L)}} \begin{bmatrix} -\gamma e^{-\theta(k)(W+3L)} & -e^{-\theta(k)(2W+L)} \\ \gamma e^{-\theta(k)(2L)} & e^{-\theta(k)(W)} \end{bmatrix} \begin{bmatrix} A_2^{n} \\ B_2^{n} \end{bmatrix}, \tag{19}$$

$$\begin{bmatrix} 1 & \gamma e^{-\theta(k)(W+2L)} \\ \gamma e^{-\theta(k)(W+2L)} & 1 \end{bmatrix} \begin{bmatrix} A_2^{n+1} \\ B_2^{n+1} \end{bmatrix} = \begin{bmatrix} e^{-\theta(k)(W-L)} & \gamma e^{-\theta(k)2L} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} A_1^{n} \\ B_1^{n} \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ \gamma e^{-\theta(k)2L} & e^{-\theta(k)W} \end{bmatrix} \begin{bmatrix} A_3^{n} \\ B_3^{n} \end{bmatrix}, \tag{20}$$

$$\begin{bmatrix} 1 & \gamma e^{-\theta(k)(W+2L)} \\ \gamma e^{-\theta(k)(W+2L)} & 1 \end{bmatrix} \begin{bmatrix} A_s^{n+1} \\ B_s^{n+1} \end{bmatrix} = \begin{bmatrix} e^{-\theta(k)(W)} & \gamma e^{-\theta(k)2L} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} A_{s-1}^{n} \\ B_{s-1}^{n} \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ \gamma e^{-\theta(k)2L} & e^{-\theta(k)W} \end{bmatrix} \begin{bmatrix} A_{s+1}^{n} \\ B_{s+1}^{n} \end{bmatrix}, \tag{21}$$

$$\begin{bmatrix} 1 & \gamma e^{-\theta(k)(W+2L)} \\ \gamma e^{-\theta(k)(W+2L)} & 1 \end{bmatrix} \begin{bmatrix} A_{p-1}^{n+1} \\ B_{p-1}^{n+1} \end{bmatrix} = \begin{bmatrix} e^{-\theta(k)(W)} & \gamma e^{-\theta(k)2L} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} A_{p-2}^{n} \\ B_{p-2}^{n} \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ \gamma e^{-\theta(k)2L} & e^{-\theta(k)(W-L)} \end{bmatrix} \begin{bmatrix} A_{p}^{n} \\ B_{p}^{n} \end{bmatrix}, \tag{22}$$

$$\begin{bmatrix} A_p^{n+1} \\ B_p^{n+1} \end{bmatrix} = \frac{1}{1 - \gamma e^{-2\theta(k)(W+L)}} \begin{bmatrix} e^{-\theta(k)W} & \gamma e^{-\theta(k)2L} \\ -\gamma e^{\theta(k)(2W+L)} & -e^{-\theta(k)(W+3L)} \end{bmatrix} \begin{bmatrix} A_{p-1}^{n} \\ B_{p-1}^{n} \end{bmatrix}. \tag{23}$$

Let us denote by S the matrix on the left-hand side in (20), by $R_2^{(l)}$ the matrix on the right-hand side in (19), by $R_1^{(l)}$ and R_2 the two matrices on the right-hand side in (20), by R_1 the first matrix on the right-hand side in (21), by $R_2^{(r)}$ the second matrix on the right-hand side in (22), and by $R_1^{(r)}$ the matrix on the right-hand side in (23).

Let us collect the $2p$ coefficients at step n in the vector

$$c^n = \left(c_1^n, c_2^n, \ldots, c_p^n\right)^T, \tag{24}$$

where $c_s^n = (A_s^n, B_s^n)$ are ordered pairs for $s = 1, \ldots, p$. For $s = 1$, we rewrite (19) as

$$c_s^{n+1} = R_2^{(l)} c_s^n.$$

For $s = 2$, we rewrite (20) as

$$Sc_s^{n+1} = R_1^{(l)} c_{s-1}^n + R_2 c_{s+1}^n.$$

For $s = 3, \ldots, p-2$, we rewrite (21) as

$$Sc_s^{n+1} = R_1 c_{s-1}^n + R_2 c_{s+1}^n.$$

For $s = p-1$, we rewrite (22) as

$$Sc_s^{n+1} = R_1 c_{s-1}^n + R_2^{(r)} c_{s+1}^n.$$

For $s = p$, we rewrite (23) as

$$c_s^{n+1} = R_1^{(r)} c_s^n.$$

Collecting these 2×2 systems into an $2p \times 2p$ system, we write it as

$$\hat{S} c^{n+1} = \hat{\mathcal{R}} c^n, \tag{25}$$

which is shown schematically in Figure 2, and where $\hat{S}$ is a symmetric positive definite block diagonal matrix with first and last diagonal blocks being a 2×2 identity matrix, the other $p-2$ diagonal blocks are 2×2 diagonal blocks being the matrix S, and $\hat{\mathcal{R}}$ is a block tridiagonal matrix with blocks commensurate with those of $\hat{S}$, and the typical block row is $[R_1|O|R_2]$.

$$\begin{bmatrix} [I] & & & & & \\ & [S] & & & & \\ & & \ddots & & & \\ & & & \ddots & & \\ & & & & [S] & \\ & & & & & [I] \end{bmatrix} \begin{bmatrix} A_1^{n+1} \\ B_1^{n+1} \\ A_2^{n+1} \\ B_2^{n+1} \\ A_3^{n+1} \\ B_3^{n+1} \\ \vdots \\ A_{p-1}^{n+1} \\ B_{p-1}^{n+1} \\ A_p^{n+1} \\ B_p^{n+1} \end{bmatrix} =$$

$$\begin{bmatrix} \begin{bmatrix} 0 \end{bmatrix} & \begin{bmatrix} R_2^{(l)} \end{bmatrix} & & & & & \\ \begin{bmatrix} R_1^{(l)} \end{bmatrix} & \begin{bmatrix} 0 \end{bmatrix} & \begin{bmatrix} R_2 \end{bmatrix} & & & & \\ & \begin{bmatrix} R_1 \end{bmatrix} & \begin{bmatrix} 0 \end{bmatrix} & \begin{bmatrix} R_2 \end{bmatrix} & & & \\ & & \ddots & \ddots & \ddots & & \\ & & & \ddots & \ddots & \begin{bmatrix} R_2 \end{bmatrix} & \\ & & & & \begin{bmatrix} R_1 \end{bmatrix} & \begin{bmatrix} 0 \end{bmatrix} & \begin{bmatrix} R_2^{(r)} \end{bmatrix} \\ & & & & & \begin{bmatrix} R_1^{(r)} \end{bmatrix} & \begin{bmatrix} 0 \end{bmatrix} \end{bmatrix} \begin{bmatrix} A_1^n \\ B_1^n \\ A_2^n \\ B_2^n \\ A_3^n \\ B_3^n \\ \vdots \\ A_{p-1}^n \\ B_{p-1}^n \\ A_p^n \\ B_p^n \end{bmatrix}.$$

FIGURE 2 Schematic representation of (25)

$$\hat{\mathcal{T}} = \begin{bmatrix} \begin{bmatrix} 0 \end{bmatrix} & \begin{bmatrix} R_2^{(l)} \end{bmatrix} & & & & & & & \\ \begin{bmatrix} S^{-1}R_1^{(l)} \end{bmatrix} & \begin{bmatrix} 0 \end{bmatrix} & \begin{bmatrix} S^{-1}R_2 \end{bmatrix} & & & & & & \\ & \begin{bmatrix} S^{-1}R_1 \end{bmatrix} & \begin{bmatrix} 0 \end{bmatrix} & \begin{bmatrix} S^{-1}R_2 \end{bmatrix} & & & & & \\ & & \ddots & \ddots & \ddots & & & & \\ & & & \ddots & \ddots & \begin{bmatrix} S^{-1}R_2 \end{bmatrix} & & & \\ & & & & \begin{bmatrix} S^{-1}R_1 \end{bmatrix} & \begin{bmatrix} 0 \end{bmatrix} & \begin{bmatrix} S^{-1}R_2^{(r)} \end{bmatrix} & \\ & & & & & \begin{bmatrix} R_1^{(r)} \end{bmatrix} & \begin{bmatrix} 0 \end{bmatrix} & \end{bmatrix}$$

FIGURE 3 Schematic representation of $\hat{\mathcal{T}}$

In order to analyze iteration (25), we rewrite it as

$$c^{n+1} = \hat{\mathcal{T}} c^n,$$

where $\hat{\mathcal{T}} = \hat{\mathcal{S}}^{-1}\hat{\mathcal{R}}$, and it is schematically represented in Figure 3. Because $\hat{\mathcal{S}}$ is block diagonal, so is its inverse. The first and last 2×2 blocks of $\hat{\mathcal{S}}$ are the 2×2 identity matrix, so their inverse is also the identity. The other diagonal blocks of $\hat{\mathcal{S}}^{-1}$ are all S^{-1}.

One can see by inspection that

$$S^{-1} = \frac{1}{1-\gamma^2 e^{-2\theta(k)(W+2L)}} \begin{bmatrix} 1 & -\gamma e^{-\theta(k)(W+2L)} \\ -\gamma e^{-\theta(k)(W+2L)} & 1 \end{bmatrix}.$$

We can then compute

$$S^{-1}R_1 = \frac{1}{1-\gamma^2 e^{-2\theta(k)(W+2L)}} \begin{bmatrix} e^{-\theta(k)W} & \gamma e^{-\theta(k)2L} \\ -\gamma e^{-\theta(k)W} e^{-\theta(k)(W+2L)} & -\gamma^2 e^{-\theta(k)2L} e^{-\theta(k)(W+2L)} \end{bmatrix}, \tag{26}$$

$$S^{-1}R_2 = \frac{1}{1-\gamma^2 e^{-2\theta(k)(W+2L)}} \begin{bmatrix} -\gamma^2 e^{-\theta(k)2L} e^{-\theta(k)(W+2L)} & -\gamma e^{\theta(k)W} e^{-\theta(k)(W+2L)} \\ \gamma e^{-\theta(k)2L} & e^{-\theta(k)W} \end{bmatrix}, \tag{27}$$

$$S^{-1}R_1^{(l)} = \frac{1}{1-\gamma^2 e^{-2\theta(k)(W+2L)}} \begin{bmatrix} e^{-\theta(k)(W-L)} & \gamma e^{\theta(k)2L} \\ -\gamma e^{-\theta(k)W} e^{-\theta(k)(W+L)} & -\gamma^2 e^{-\theta(k)2L} e^{-\theta(k)(W+2L)} \end{bmatrix}, \tag{28}$$

$$S^{-1}R_2^{(r)} = \frac{1}{1-\gamma^2 e^{-2\theta(k)(W+2L)}} \begin{bmatrix} -\gamma^2 e^{-\theta(k)2L} e^{-\theta(k)(W+2L)} & -\gamma e^{\theta(k)W} e^{-\theta(k)(W+L)} \\ \gamma e^{-\theta(k)2L} & e^{-\theta(k)(W-L)} \end{bmatrix}. \tag{29}$$

Step 4. Our goal is to show that $\hat{\mathcal{T}} = \hat{\mathcal{T}}(k)$ is contracting. To that end, we shall compute $\|\hat{\mathcal{T}}\|_\infty$, that is, its maximum row sum, with elements taken in absolute value. A typical block row of the block tridiagonal matrix $\hat{\mathcal{T}}$ is of the form $[S^{-1}R_1|O|S^{-1}R_2]$. Each of the two row sums of this typical block row is thus

$$\begin{aligned} \|[S^{-1}R_1 \mid O \mid S^{-1}R_2]\|_\infty &= \frac{1}{1-\gamma^2 e^{-2\theta(k)(W+2L)}} \cdot \left[e^{-\theta(k)W} + |\gamma| e^{-\theta(k)2L} + \gamma^2 e^{-\theta(k)2L} e^{-\theta(k)(W+2L)} \right. \\ &\quad \left. + |\gamma| e^{-\theta(k)W} e^{-\theta(k)(W+2L)} \right] \\ &= \frac{1}{1-\gamma^2 e^{-2\theta(k)(W+2L)}} \cdot \left[e^{-\theta(k)W}(1 + |\gamma| e^{-\theta(k)(W+2L)}) + |\gamma| e^{-\theta(k)2L}(1 + |\gamma| e^{-\theta(k)(W+2L)}) \right] \\ &= \frac{e^{-\theta(k)W} + |\gamma| e^{-\theta(k)2L}}{1 - |\gamma| e^{-\theta(k)(W+2L)}} \cdot \end{aligned} \tag{30}$$

For the second block of $\hat{\mathcal{T}}$, the maximum row sum is

$$\left\|\left[S^{-1}R_1^{(l)} \mid O \mid S^{-1}R_2\right]\right\|_\infty \frac{1}{1-\gamma^2 e^{-2\theta(k)(W+2L)}} \cdot \left[e^{-\theta(k)W}e^{-\theta(k)(-L)} + |\gamma|e^{-\theta(k)2L} + \gamma^2 e^{-\theta(k)2L}e^{-\theta(k)(W+2L)} + |\gamma|e^{-\theta(k)W}e^{-\theta(k)(W+2L)}\right]. \tag{31}$$

For the $p-1$ block of $\hat{\mathcal{T}}$, the maximum row sum is

$$\left\|\left[S^{-1}R_1 \mid O \mid S^{-1}R_2^{(r)}\right]\right\|_\infty = \frac{1}{1-\gamma^2 e^{-2\theta(k)(W+2L)}} \cdot \left[e^{-\theta(k)W} + |\gamma|e^{-\theta(k)2L} + \gamma^2 e^{-\theta(k)2L}e^{-\theta(k)(W+2L)} + |\gamma|e^{-\theta(k)W}e^{-\theta(k)(W+2L)}e^{-\theta(k)(-L)}\right]. \tag{32}$$

For the first and the last row of $\hat{\mathcal{T}}$, the maximum row sums are

$$\left|r_{21}^{(l)}\right| + \left|r_{22}^{(l)}\right| = \left|r_{11}^{(r)}\right| + \left|r_{12}^{(r)}\right| = \left(\frac{e^{-\theta(k)W} + |\gamma|e^{-\theta(k)2L}}{1-\gamma e^{-2\theta(k)(W+L)}}\right) e^{-\theta(k)(W+L)} \tag{33}$$

and

$$\left|r_{11}^{(l)}\right| + \left|r_{12}^{(l)}\right| = \left|r_{21}^{(r)}\right| + \left|r_{22}^{(r)}\right| = \frac{e^{-\theta(k)W} + |\gamma|e^{-\theta(k)2L}}{1-\gamma e^{-2\theta(k)(W+L)}}, \tag{34}$$

where $r_{ij}^{(l)} = (R_2^{(l)})_{ij}$ and $r_{ij}^{(r)} = (R_1^{(r)})_{ij}$.

In order to find the maximum, that is, $\|\hat{\mathcal{T}}(k)\|_\infty$, we will compare (30), (31), (32), (33), and (34).

First we compare (33) and (34). Because $e^{-\theta(k)(W+L)} < 1$, we have that

$$\left|r_{21}^{(l)}\right| + \left|r_{22}^{(l)}\right| = \left|r_{11}^{(r)}\right| + \left|r_{12}^{(r)}\right| < \left|r_{11}^{(l)}\right| + \left|r_{12}^{(l)}\right| = \left|r_{21}^{(r)}\right| + \left|r_{22}^{(r)}\right|.$$

Next, we compare (34) and (30). Since $W+2L < 2(W+L)$, then $e^{-\theta(k)(W+2L)} > e^{-\theta(k)2(W+L)}$. Consequently, we have that

$$\left|r_{11}^{(l)}\right| + \left|r_{12}^{(l)}\right| = \left|r_{21}^{(r)}\right| + \left|r_{22}^{(r)}\right| < \|[S^{-1}R_1 \mid O \mid S^{-1}R_2]\|_\infty.$$

If $\gamma > 0$, then $|\gamma| = \gamma$, and we have

$$|\gamma|e^{-\theta(k)(W+2L)} > \gamma e^{-\theta(k)2(W+L)}.$$

Then, we can write

$$1 - |\gamma|e^{-\theta(k)(W+2L)} < 1 - \gamma e^{-\theta(k)2(W+L)},$$

which implies

$$\frac{e^{-\theta(k)W} + |\gamma|e^{-\theta(k)2L}}{1-|\gamma|e^{-\theta(k)(W+2L)}} > \frac{e^{-\theta(k)W} + |\gamma|e^{-\theta(k)2L}}{1-\gamma e^{-\theta(k)2(W+L)}}.$$

However, if $\gamma < 0$ and $|\gamma| = -\gamma$, then

$$\gamma e^{-\theta(k)(W+2L)} < -\gamma e^{-\theta(k)2(W+L)}.$$

Then, we can write

$$1 + \gamma e^{-\theta(k)(W+2L)} < 1 - \gamma e^{-\theta(k)2(W+L)}.$$

Therefore, because here $1 - |\gamma| = 1 + \gamma$, we have

$$\frac{e^{-\theta(k)W} + |\gamma|e^{-\theta(k)2L}}{1+\gamma e^{-\theta(k)(W+2L)}} > \frac{e^{-\theta(k)W} + |\gamma|e^{-\theta(k)2L}}{1-\gamma e^{-\theta(k)2(W+L)}}.$$

Then, for all γ, we have

$$\frac{e^{-\theta(k)W}+|\gamma|e^{-\theta(k)2L}}{1-|\gamma|e^{-\theta(k)(W+2L)}}>\frac{e^{-\theta(k)W}+|\gamma|e^{-\theta(k)2L}}{1-\gamma e^{-\theta(k)2(W+L)}}.$$

Next, by comparing (30), (31), and (32), we show that

$$\|[S^{-1}R_1\,|\,O\,|\,S^{-1}R_2]\|_\infty<\left\|\left[S^{-1}R_1^{(l)}\,|\,O\,|\,S^{-1}R_2\right]\right\|_\infty$$
$$\text{and}\quad \|[S^{-1}R_1\,|\,O\,|\,S^{-1}R_2]\|_\infty<\left\|\left[S^{-1}R_1\,|\,O\,|\,S^{-1}R_2^{(r)}\right]\right\|_\infty.$$

This holds because the first term of $\|[S^{-1}R_1^{(l)}\,|\,O\,|\,S^{-1}R_2]\|_\infty$ is multiplied by $e^{-\theta(k)(-L)}>1$ and the last term of $\|[S^{-1}R_1\,|\,O\,|\,S^{-1}R_2^{(r)}]\|_\infty$ is multiplied by $e^{-\theta(k)(-L)}>1$. Therefore, for each k,

$$\|\hat{\mathcal{T}}(k)\|_\infty=\max\left\{\left\|\left[S^{-1}R_1^{(l)}\,|\,O\,|\,S^{-1}R_2\right]\right\|_\infty,\left\|\left[S^{-1}R_1\,|\,O\,|\,S^{-1}R_2^{(r)}\right]\right\|_\infty\right\}.$$

We compute now a bound for this maximum:

$$\begin{aligned}\left\|S^{-1}R_1^{(l)}\,|\,O\,|\,S^{-1}R_2\right\|_\infty=&\frac{1}{1-\gamma^2e^{-2\theta(k)(W+2L)}}\cdot\left[e^{-\theta(k)W}e^{-\theta(k)(-L)}+|\gamma|e^{-\theta(k)2L}\right.\\&\left.+\gamma^2e^{-\theta(k)2L}e^{-\theta(k)(W+2L)}+|\gamma|e^{-\theta(k)W}e^{-\theta(k)(W+2L)}\right]\\=&\frac{1}{1-\gamma^2e^{-2\theta(k)(W+2L)}}\cdot\left[e^{-\theta(k)W}(e^{-\theta(k)(-L)}+|\gamma|e^{-\theta(k)(W+2L)})\right.\\&\left.+|\gamma|e^{-\theta(k)2L}(1+|\gamma|e^{-\theta(k)(W+2L)})\right]\\<&\frac{1}{1-\gamma^2e^{-2\theta(k)(W+2L)}}\cdot\left[e^{-\theta(k)W}e^{-\theta(k)(-L)}(1+|\gamma|e^{-\theta(k)(W+2L)})\right.\\&\left.+|\gamma|e^{-\theta(k)2L}(1+|\gamma|e^{-\theta(k)(W+2L)})\right]\\=&\frac{e^{-\theta(k)(W-L)}+|\gamma|e^{-\theta(k)2L}}{1-|\gamma|e^{-\theta(k)(W+2L)}}<\frac{(1+|\gamma|)e^{-\theta(k)(2L)}}{1-|\gamma|e^{-\theta(k)(W+2L)}}<\frac{|\gamma|e^{-\theta(k)(2L)}}{1-|\gamma|e^{-\theta(k)(W+2L)}}.\end{aligned}$$

The first inequality holds because $e^{-\theta(k)(-L)}>1$, and the second holds if $e^{-\theta(k)(W-L)}<e^{-\theta(k)(2L)}$ is satisfied, which is true by our hypothesis that $L<W/3$.

$$\begin{aligned}\|S^{-1}R_1\,|\,O\,|\,S^{-1}R_2^{(r)}\|_\infty=&\frac{1}{1-\gamma^2e^{-2\theta(k)(W+2L)}}\cdot\left[e^{-\theta(k)W}+|\gamma|e^{-\theta(k)2L}+\gamma^2e^{-\theta(k)2L}e^{-\theta(k)(W+2L)}\right.\\&\left.+|\gamma|e^{-\theta(k)W}e^{-\theta(k)(W+2L)}e^{-\theta(k)(-L)}\right]\\=&\frac{1}{1-\gamma^2e^{-2\theta(k)(W+2L)}}\cdot\left[e^{-\theta(k)(W)}(1+|\gamma|e^{-\theta(k)(W+2L)}e^{-\theta(k)(-L)})\right.\\&\left.+|\gamma|e^{-\theta(k)2L}(1+|\gamma|e^{-\theta(k)(W+2L)})\right]\\=&\frac{1}{1-\gamma^2e^{-2\theta(k)(W+2L)}}\cdot\left[e^{-\theta(k)(W-L)}(e^{-\theta(k)(L)}+|\gamma|e^{-\theta(k)(W+2L)})\right.\\&\left.+|\gamma|e^{-\theta(k)2L}(1+|\gamma|e^{-\theta(k)(W+2L)})\right]\\<&\frac{1}{1-\gamma^2e^{-2\theta(k)(W+2L)}}\cdot\left[e^{-\theta(k)(W-L)}(1+|\gamma|e^{-\theta(k)(W+2L)})\right.\\&\left.+|\gamma|e^{-\theta(k)2L}(1+|\gamma|e^{-\theta(k)(W+2L)})\right]\\=&\frac{e^{-\theta(k)(W-L)}+|\gamma|e^{-\theta(k)2L}}{1-|\gamma|e^{-\theta(k)(W+2L)}}<\frac{(1+|\gamma|)e^{-\theta(k)(2L)}}{1-|\gamma|e^{-\theta(k)(W+2L)}}.\end{aligned}$$

The first inequality holds because $e^{-\theta(k)(L)} < 1$, and the second holds if $e^{-\theta(k)(W-L)} < e^{-\theta(k)(2L)}$ is satisfied, which is true by our hypothesis that $L < W/3$. All that remains to show is that this last bound is less than one. To that end, observe that the following are equivalent:

$$\begin{aligned}
\frac{e^{-\theta(k)(2L)} + |\gamma| e^{-\theta(k)(2L)}}{1 - |\gamma| e^{-\theta(k)(W+2L)}} &< 1 \\
e^{-\theta(k)(2L)} + |\gamma| e^{-\theta(k)(2L)} &< 1 - |\gamma| e^{-\theta(k)(W+2L)} \\
|\gamma|(1 + e^{-\theta(k)(W)}) e^{-\theta(k)(2L)} &< 1 - e^{-\theta(k)(2L)} \\
|\gamma| e^{-\theta(k)(2L)} &< \frac{1 - e^{-\theta(k)(2L)}}{1 + e^{-\theta(k)W}}.
\end{aligned}$$

The last inequality is precisely our hypothesis (13). Therefore, for each k, $\|\hat{\mathcal{T}}(k)\|_\infty < 1$.

Step 5. Our goal is to prove that the $E^n \to 0$, where $E^n = (E_1^n, E_2^n, \ldots, E_p^n)$. We showed that $\hat{\mathcal{T}}$ is contracting implying that $\hat{E}_s^n(x,k) \to 0$, for $s = 1, \ldots, p$.

By (6), we have

$$E_s^n(x,y) = \sum_{k=1}^{\infty} \hat{E}_s^n(x,k) \sin\left(\frac{k\pi y}{L_2}\right).$$

We apply the limit, and using the fact that we can interchange the limit and the summation (see, e.g., the work of Haberman[15]), we obtain

$$\lim_{n\to\infty} E_s^n(x,y) = \lim_{n\to\infty} \sum_{k=1}^{\infty} \hat{E}_s^n(x,k) \sin\left(\frac{k\pi y}{L_2}\right) = \sum_{k=1}^{\infty} \lim_{n\to\infty} \hat{E}_s^n(x,k) \sin\left(\frac{k\pi y}{L_2}\right) = 0,$$

thus completing the proof. □

3 | ASYNCHRONOUS SCHWARZ WITH OPTIMIZED INTERFACE CONDITIONS

We consider now an asynchronous version of the algorithm described by (4). We assume that we have p processors, one for each subdomain, each processor solves a local problem. We describe here a standard mathematical model of asynchronous iterations for this algorithm; see, for example, other works.[17,18,19] Let us define a *time stamp* as the instant of time at which at least one processor finishes its computation, that is, it completes the solution of its local problem, and produces a new update. Let $X_1, \ldots, X_p$ be given sets and X be their Cartesian product, that is, $X = X_1 \times \ldots \times X_p$. Thus, $u \in X$ implies $u = (u_1, \ldots, u_p)$ with $u_s \in X_s$ for $s \in \{1, \ldots, p\}$. Let $\mathcal{T}_s : X \to X_s$ where $s \in \{1, \ldots, p\}$, and let $\mathcal{T} : X \to X$ be a vector-valued map (iteration map) given by $\mathcal{T} = (\mathcal{T}_1, \ldots, \mathcal{T}_p)$ with a fixed point u^*, that is, $u^* = \mathcal{T}(u^*)$. Thus, the mathematical formulation of the synchronous iterative algorithm described in (4), given an initial approximation u^0, is

$$u^n = \mathcal{T}(u^{n-1}), for\ n = 1, 2, \ldots,$$

or equivalently

$$u_s^n = \mathcal{T}_s(u^{n-1}), \quad s = 1, \ldots, p, \quad n = 1, 2, \ldots$$

The resulting u^n at each iteration is an approximation of the fixed point u^*.

Let $\{t_n\}_{n\in\mathbb{N}}$ be the sequence of time stamps at which at least one processor updates its associated component. Let $\{\sigma(n)\}_{n\in\mathbb{N}}$ be a sequence with $\sigma(n) \subset \{1, \ldots, p\}$, $\forall n \in \mathbb{N}$. Set $\sigma(n)$ consists of labels (numbers) of the processors that update their associated component at the nth time stamp. Define for $s, q \in \{1, \ldots, p\}$, $\{\tau_q^s(n)\}_{n\in\mathbb{N}}$ a sequence of integers, representing the time stamp index of the update of the data coming from processor q and available in processor s before it

starts producing the local solution to be completed at the nth time stamp. Let $u^0 = \left(u_1^0, \ldots, u_p^0\right)$ be the initial approximation (of the fixed point u^*). Then, the new computed value updated by processor s at the nth time stamp is

$$u_s^n = \begin{cases} \mathcal{T}_s\left(u_1^{\tau_1^s(n)}, \ldots, u_p^{\tau_p^s(n)}\right), & s \in \sigma(n) \\ u_s^{n-1}, & s \notin \sigma(n)\cdot \end{cases} \tag{35}$$

We emphasize that u_s^n is the approximation to u_s^* at the time stamp t_n.

It is customary to assume that the three following conditions (necessary for convergence) are satisfied:

$$\tau_q^{(s)}(n) < n, \forall s, q \in \{1, \ldots, p\}, \forall n \in \mathbb{N}^*, \tag{36}$$

$$\text{card}\left\{n \in \mathbb{N}^* | s \in \sigma(n)\right\} = +\infty, \forall s \in \{1, \ldots, p\}, \tag{37}$$

$$\lim_{n \to +\infty} \tau_q^{(s)}(n) = +\infty, \forall s, q \in \{1, \ldots, p\}. \tag{38}$$

Condition (36) indicates that data used at the time t_n must have been produced before time t_n, that is, time does not flow backward. Condition (37) means that no process will ever stop updating its components. Condition (38) corresponds to the fact that new data will always be provided to the process. In other words, no process will have a piece of data that is never updated; see, for example, other works.[17,18,20,21]

In our context, when we want to solve the problem (1), the setup (35) is as follows (for s corresponding to an interior subdomain):

$$u_s^{n+1} = \begin{cases} \text{Solve} \begin{cases} (\Delta - \eta) u_s^{n+1} = f_s \text{ on } \Omega_s, \\ -\frac{\partial u_s^{n+1}}{\partial x} + \Lambda\left(u_s^{n+1}\right) = -\frac{\partial u_{s-1}^{\tau_{s-1}^{(s)}(n)}}{\partial x} + \Lambda\left(u_{s-1}^{\tau_{s-1}^{(s)}(n)}\right) \\ \frac{\partial u_s^{n+1}}{\partial x} + \Lambda\left(u_s^{n+1}\right) = \frac{\partial u_{s+1}^{\tau_{s+1}^{(s)}(n)}}{\partial x} + \Lambda\left(u_{s+1}^{\tau_{s+1}^{(s)}(n)}\right) \end{cases}, & \text{if } s \in \sigma(n+1) \\ u_s^n, & \text{if } s \notin \sigma(n+1), \end{cases} \tag{39}$$

and for $s = 1$ or $s = p$, one of the above boundary conditions is not present and we have the Dirichlet boundary condition instead as in (4).

Thus, $\mathcal{T}_s$ is given by (39), which is the same as in (4), for which Theorem 1 applies. Our goal is to show that the asynchronous optimized Schwarz method converges. As we did for the synchronous case, it suffices to show that the error of the solution converges to zero for $f = 0$ and $g = 0$.

Let $E_s^m(x, y)$ be the error of the local problem s at the time stamp t_m (of the asynchronous version of (4)), $s \in \{1, \ldots, p\}$, and $\hat{E}_s^m(x, k)$ be its corresponding Fourier series coefficients in the y-direction, as in (6) and (7). For any monotonically increasing sequence of time stamps $\{t_m\}_{m \in \mathbb{N}}$, we want to prove that the following holds:

$$\lim_{m \to \infty} E_s^m = 0, \quad \text{for } s = 1, \ldots, p.$$

Theorem 2. *For the shifted Laplacian problem (1), and a one-way domain decomposition with p vertical strips of width W, for $s = 2, \ldots, p-1$, overlap 2L throughout, and transmission operator Λ of the form (3) defined on the interfaces, supposed to be the same on the left and right interfaces, given an initial approximation $u^0 = (u_1^0, \ldots, u_p^0)$, consider the following asynchronous optimized Schwarz algorithm. In process s, corresponding to the subspace Ω_s, for any sequence of update indexes $\sigma(n)$ and time stamps of previous computations $\tau_q^{(s)}(n)$, $n = 1, \ldots$, satisfying the conditions (36)–(38), the asynchronous iteration (39), where, if $s = 1$, or $s = p$, one of the boundary conditions is not present and one uses Dirichlet boundary condition instead as in (4). Assume that $\lambda(k)$, W, and L, satisfy conditions (11)–(13). Let*

$$\Sigma = \{\ell_{s-1} - L : s = 2, \ldots, p\} \cup \{\ell_s + L : s = 1, \ldots, p-1\}$$

(i.e., Σ is the set of the x-coordinates of each of the artificial boundaries of each of the local problems.) Assume further that the initial approximation is such that the initial error $E_s^0(x, y)$ is uniformly bounded for $x \in \Sigma$ and $y \in [0, L_2]$. Then, for

all $s \in \{1, \dots, p\}$, the sequence of errors satisfies $\lim_{m\to\infty} E_s^m(x, y) = 0$ in Ω_s for any (monotonically increasing) sequence of time stamps $\{t_m\}_{m\in\mathbb{N}}$. In other words, the asynchronous optimized Schwarz algorithm (39) converges.

Proof. The goal is to prove that for any sequence of time stamps $\{t_m\}_{m\in\mathbb{N}}$ and for every $s \in \{1, \dots, p\}$, we have $\lim_{m\to\infty} E_s^m(x, y) = 0$ in $[\ell_{s-1}, \ell_s]$. We first note that, to prove this statement, it suffices to show that

$$\lim_{m\to\infty} |E_s^m(x, y)| = 0 \text{ uniformly in } \Sigma \cap [\ell_{s-1}, \ell_s] \times [0, L_2], \quad s = 1, \dots, p,$$

because this implies that the values of the boundary conditions of each local problem will converge to zero, and consequently so will the solution of each local problem in the interior of the corresponding subdomain.

We can say this because the derivatives of E_s^m exist and are continuous, and therefore, for $\ell = \ell_{s-1}$ or $\ell = \ell_s$, if E_s^m converges to zero uniformly in $\{\ell\} \times [0, L_2]$ as $m \to \infty$ and the first and derivatives of other orders of E_s^m are continuous, it also holds that $\lim_{m\to\infty} \left(\frac{\partial E_s^m}{\partial x} + \Lambda E_s^m \right)(x, y) = 0$ uniformly in $\{\ell\} \times [0, L_2]$.

Furthermore, our proof would reduce to showing that

$$\lim_{m\to\infty} |\hat{E}_s^m(x, y)| = 0 \text{ uniformly in } \Sigma \cap [\ell_{s-1}, \ell_s] \times [0, L_2], \quad s = 1, \dots, p. \tag{40}$$

This would follow from the following bound:

$$\begin{aligned} \lim_{m\to\infty} |E_s^m(x, y)| &= \lim_{m\to\infty} \left| \sum_{k=1}^{\infty} \hat{E}_s^m(x, k) \sin\left(\frac{k\pi y}{L_2}\right) \right| \\ &\leq \lim_{m\to\infty} \sum_{k=1}^{\infty} \left| \hat{E}_s^m(x, k) \right| = \sum_{k=1}^{\infty} \lim_{m\to\infty} \left| \hat{E}_s^m(x, k) \right| = 0 \end{aligned} \tag{41}$$

as long as the equality (41) holds.

Thus, we have a road map for the proof of the theorem, and we present its sketch next.

1. We first consider the coefficients c^0 of the initial error, defined in (24) (with $n = 0$), and prove that these are uniformly bounded, that is, $\|c^0(k)\|_\infty < C < \infty$, for some $C > 0$.
2. We use this to then prove that for all k, and $s = 1, \dots, p$, $\lim_{m\to\infty} |A_s^m(k)| = 0$ and $\lim_{m\to\infty} |B_s^m(k)| = 0$
3. We show that, for all $m \in \mathbb{N}$, the series $\sum_{k=1}^{\infty} \hat{E}_s^m(x, k)$ is uniformly convergent, which implies (41).

Step 1. We first prove that $\|c^0(k)\|_\infty < C < \infty$, for some $C > 0$, where c^0 is the initial value of c^n as in (24). To that end, we want to bound $A_s^0(k)$ and $B_s^0(k)$, $s = 1, \dots, p$, for all k. For ease of notation, for each subdomain s, let the left artificial Dirichlet boundary condition be $p_s(k)$ and the right artificial boundary condition $q_s(k)$, and let us define $\xi = \eta + \pi^2/L_2^2 > 0$. We will use the fact that for all k,

$$\theta(k) = \sqrt{\eta + \left(\frac{k\pi}{L_2}\right)^2} \geq \sqrt{\eta + \left(\frac{\pi}{L_2}\right)^2} = \sqrt{\xi}.$$

Thus, it follows from expression (14) that for $s = 2, \dots, p-1$ at $x = \ell_{s-1} - L$, we have

$$\hat{u}_s^0(\ell_{s-1} - L, k) = A_s^0(k) + B_s^0(k) e^{\theta(k)(\ell_{s-1} - \ell_s - 2L)} = A_s^0(k) + B_s^0(k) e^{-\theta(k)(W+2L)} = p_s(k) \tag{42}$$

and at $x = \ell_s + L$, we have

$$\hat{u}_s^0(\ell_s + L, k) = A_s^0(k) e^{-\theta(k)(\ell_s - \ell_{s-1} + 2L)} + B_s^0(k) = A_s^0(k) e^{-\theta(k)(W+2L)} + B_s^0(k) = q_s(k). \tag{43}$$

From (43), we have $B_s^0(k) = q_s(k) - A_s^0(k)e^{-\theta(k)(W+2L)}$. Then, plugging this expression of $B_s^0(k)$ into (42) gives

$$\begin{aligned} A_s^0(k) + \left[q_s(k) - A_s^0(k)e^{-\theta(k)(W+2L)}\right] e^{-\theta(k)(W+2L)} &= p_s(k), \\ A_s^0(k)\left[1 - e^{-2\theta(k)(W+2L)}\right] &= p_s(k) - q_s(k)e^{-\theta(k)(W+2L)}, \\ A_s^0(k) &= \frac{p_s(k) - q_s(k)e^{-\theta(k)(W+2L)}}{1 - e^{-2\theta(k)(W+2L)}}, \\ \left|A_s^0(k)\right| = \frac{|p_s(k) - q_s(k)e^{-\theta(k)(W+2L)}|}{|1 - e^{-2\theta(k)(W+2L)}|} &\leq \frac{|p_s(k)| + |q_s(k)|e^{-\theta(k)(W+2L)}}{1 - e^{-2\sqrt{\xi}(W+2L)}}. \end{aligned}$$

By a similar process, we obtain

$$\left|B_s^0(k)\right| \leq \frac{|q_s(k)| + |p_s(k)|e^{-\theta(k)(W+2L)}}{1 - e^{-2\sqrt{\xi}(W+2L)}}.$$

For $s = 1$, at $x = 0$, we have

$$\hat{u}_1^0(0,k) = A_1^0(k) + B_1^0(k)e^{\theta(k)(\ell_1+L)} = p_1(k), \tag{44}$$

and at $x = \ell_1 + L$, we have

$$\hat{u}_1^0(\ell_1 + L, k) = A_1^0(k)e^{-\theta(k)(\ell_1+L)} + B_1^0(k) = q_1(k). \tag{45}$$

From (45), we have $B_1^0(k) = q_1(k) - A_1^0(k)e^{-\theta(k)(\ell_1+L)}$. Then, plugging this expression of $B_1^0(k)$ into (44) gives

$$\begin{aligned} A_1^0(k) + \left[q_1(k) - A_1^0(k)e^{-\theta(k)(\ell_1+L)}\right] e^{-\theta(k)(\ell_1+L)} &= p_1(k), \\ A_1^0(k)\left[1 - e^{-2\theta(k)(\ell_1+L)}\right] &= p_1(k) - q_1(k)e^{-\theta(k)(\ell_1+L)}, \\ A_1^0(k) &= \frac{p_1(k) - q_1(k)e^{-\theta(k)(\ell_1+L)}}{1 - e^{-2\theta(k)(\ell_1+L)}}, \\ |A_1^0(k)| = \frac{|p_1(k) - q_1(k)e^{-\theta(k)(W+L)}|}{|1 - e^{-2\theta(k)(W+L)}|} &\leq \frac{|p_1(k)| + |q_1(k)|e^{-\theta(k)(W+L)}}{1 - e^{-2\sqrt{\xi}(W+L)}}. \end{aligned}$$

By a similar process, we obtain

$$|B_1^0(k)| \leq \frac{|q_1(k)| + |p_1(k)|e^{-\theta(k)(W+L)}}{1 - e^{-2\sqrt{\xi}(W+L)}}.$$

For $s = p$, at $x = \ell_{p-1} - L$, we have

$$\hat{u}_p^0(\ell_{p-1} - L, k) = A_p^0(k) + B_p^0(k)e^{-\theta(k)(W+L)} = q_p(k), \tag{46}$$

and at $x = L_2$, we have

$$\hat{u}_p^0(L_2, k) = A_p^0(k)e^{\theta(k)(W+L)} + B_p^0(k) = p_p(k). \tag{47}$$

From (46), we have $B_p^0(k) = q_p(k) - A_p^0(k)e^{-\theta(k)(W+L)}$. Then, plugging this expression of $B_p^0(k)$ into (47) gives

$$\begin{aligned} A_p^0(k) + \left[q_p(k) - A_p^0(k)e^{-\theta(k)(W+L)}\right] e^{-\theta(k)(W+L)} &= p_p(k), \\ A_p^0(k)\left[1 - e^{-2\theta(k)(W+L)}\right] &= p_p(k) - q_p(k)e^{-\theta(k)(W+L)}, \\ A_p^0(k) &= \frac{p_p(k) - q_p(k)e^{-\theta(k)(W+L)}}{1 - e^{-2\theta(k)(W+L)}}, \\ |A_p^0(k)| = \frac{|p_p(k) - q_p(k)e^{-\theta(k)(W+L)}|}{|1 - e^{-2\theta(k)(W+L)}|} &\leq \frac{|p_p(k)| + |q_p(k)|e^{-\theta(k)(W+L)}}{1 - e^{-2\sqrt{\xi}(W+L)}}. \end{aligned}$$

By a similar process, we obtain

$$|B_p^0(k)| \leq \frac{|q_p(k)| + |p_p(k)|e^{-\theta(k)(W+L)}}{1 - e^{-2\sqrt{\xi}(W+L)}}.$$

Let $p^*(k)$ and $q^*(k)$ be such that

$$\max_{s\in\{2,\dots,p-1\},s'=1,p}\left\{\max\left\{\frac{|p_s(k)|+|q_s(k)|e^{-\theta(k)(W+2L)}}{1-e^{-2\sqrt{\xi}(W+2L)}},\frac{|p_{s'}(k)|+|q_{s'}(k)|e^{-\theta(k)(W+L)}}{1-e^{-2\sqrt{\xi}(k)(W+L)}}\right.\right.$$
$$\left.\left.\frac{|q_s(k)|+|p_s(k)|e^{-\theta(k)(W+2L)}}{1-e^{-2\sqrt{\xi}(W+2L)}},\frac{|q_{s'}(k)|+|p_{s'}(k)|e^{-\theta(k)(W+L)}}{1-e^{-2\sqrt{\xi}(W+L)}}\right\}\right\}$$
$$=\frac{|p^*(k)|+|q^*(k)|e^{-\theta(k)(W+L)}}{1-e^{-2\sqrt{\xi}(W+L)}}.$$

Then, we have that

$$\|c^0(k)\|_\infty\le\frac{|p^*(k)|+|q^*(k)|e^{-\theta(k)(W+L)}}{1-e^{-2\sqrt{\xi}(W+L)}}\le M\frac{1+e^{-\theta(k)(W+L)}}{1-e^{-2\sqrt{\xi}(W+L)}},\tag{48}$$

where we have used the hypothesis that E_s^0 is uniformly bounded in $\Sigma\times L_2$, which implies that $\hat{E}_s^0$ is uniformly bounded in $k\in\mathbb{N}$ and $x\in\Sigma$, and thus, there exists a number $M>0$ such that $\hat{E}^0(x,k)\le M$ for all $k\in\mathbb{N}$ and $x\in\Sigma$. Then, we have $|p_s(k)|,|q_s(k)|\le M$ for all $k\in\mathbb{R}$ and $s\in\{1,...p\}$. Then, necessarily, $|p^*(k)|,|q^*(k)|\le M$. We conclude then, from (48), that $\|c^0(k)\|_\infty<C<\infty$, where

$$C=M\frac{1+e^{-\theta(k)(W+L)}}{1-e^{-2\sqrt{\xi}(W+L)}}.$$

Step 2. Let $\{t_m\}$ be a monotonically increasing sequence of time stamps. Let t_j be the time stamp at which processor s produces its first new update. Let $c_s^j(k)=(A_s^j(k),B_s^j(k))$ be the vector containing the coefficients corresponding to subdomain s, for $s=1,\dots,p$, at that time stamp t_j. We have, as in Theorem 1, that $\hat{\mathcal{T}}(k)=(\hat{\mathcal{T}}_1(k),\dots,\hat{\mathcal{T}}_p(k))$, where $\hat{\mathcal{T}}_s(k)$ is the local operator associated with the subdomain s; see, for example, Figure 3. Let us denote by $\rho_k=\|\hat{\mathcal{T}}(k)\|_\infty$, and recall that in Theorem 1 we showed that under our hypotheses $\rho_k<1$. Then, we have for $i=1,2$,

$$\left|c_s^j(k)_i\right|=\left|\left(\hat{\mathcal{T}}_s(k)c^0(k)\right)_i\right|\le\left\|\hat{\mathcal{T}}_s(k)c^0(k)\right\|_\infty\le\left\|\hat{\mathcal{T}}_s(k)\right\|_\infty\|c^0(k)\|_\infty$$
$$\le\left\|\hat{\mathcal{T}}(k)\right\|_\infty\|c^0(k)\|_\infty=\rho_k\|c^0(k)\|_\infty.$$

Then, each coefficient of the series representing the error of the approximation of the solution in subdomain s at the time stamp t_j satisfies $|A_s^j(k)|,|B_s^j(k)|\le\rho_k\|c^0(k)\|_\infty$. Now, for time stamps $t_m>t_j$, either this approximation is updated, in which case, again, the coefficients would continue to be bounded by $\rho_k\|c^0(k)\|_\infty$ (and most likely decrease), or it is not updated, in which case, it stays the same (see (39)), that is,

$$|A_s^m(k)|,|B_s^m(k)|\le\rho_k\|c^0(k)\|_\infty\text{ for }t_m>t_j.\tag{49}$$

Of course, we considered here one processor s arbitrary. Therefore, after every processor has produced its first update, say, at t_{j_1}, we have that (49) holds for all s for all $t_m>t_{j_1}$.

In the rest of the proof of Step 2, we will show that eventually the error of the approximation in every subdomain gets closer to the fixed point (which is zero in this context) by a fixed amount, that is, its max norm is reduced at least by ρ_k. Thus, as we reasoned for the first time stamp when every processor updated the approximation at least once, we consider the case where every processor updated the approximation at least twice. In other words, we can see that once every processor has produced a new update after t_{j_1}, say, at t_{j_2}, we have $|A_s^m(k)|,|B_s^m(k)|<\rho_k^2||c^0(k)||_\infty$ for all $t_m\ge t_{j_2}$. We know that every processor will eventually update its subdomain because of conditions (37)–(38). Continuing in this manner, we can say that

$$|A_s^m(k)|,|B_s^m(k)|<\rho_k^{\phi(m)}\|c^0(k)\|_\infty,$$

where, denoting by t_{j_i} the first time stamp at which all processors have updated their values at least i times, we have $\phi(m)=\max\{i\in\mathbb{N}:t_{j_i}\le t_m\}$, that is, $\phi(m)$ is the update number (at time t_m) of the processor that produced the least

number of updates among all processors until the instant of time t_m. Note also that with this definition, it holds that $\lim_{m\to\infty}\phi(m)=\infty$. This implies, in turn, that given a monotonically increasing sequence $\{t_m\}_{m\in\mathbb{N}}$, at time t_m, we have

$$|A_s^m|, |B_s^m| \le \rho_k^{\phi(m)} \|c^0(k)\|_\infty,$$

and that

$$\lim_{m\to\infty} |A_s^m(k)| \le \lim_{m\to\infty} \rho_k^{\phi(m)} \|c^0(k)\|_\infty = 0.$$

Similarly, $\lim_{m\to\infty}|B_s^m(k)|=0$.

Consequently,

$$\begin{aligned}
\lim_{m\to\infty} |\hat{E}_s^m(x,k)| &= \lim_{m\to\infty} \left| A_s^m(k) e^{-\theta(k)|x-(\ell_{s-1}-L)|} + B_s^m(k) e^{-\theta(k)|x-(\ell_s+L)|} \right| \\
&\le \lim_{m\to\infty} \left(|A_s^m(k)|\, e^{-\theta(k)|x-(\ell_{s-1}-L)|} + |B_s^m(k)|\, e^{-\theta(k)|x-(\ell_s+L)|} \right) \\
&= \left(\lim_{m\to\infty} |A_s^m(k)| \right) e^{-\theta(k)|x-(\ell_{s-1}-L)|} + \left(\lim_{m\to\infty} |B_s^m(k)| \right) e^{-\theta(k)|x-(\ell_s+L)|} \\
&= 0.
\end{aligned}$$

Step 3. To complete the proof, we need to show that (41) holds for $x \in \Sigma \cap [\ell_{s-1}, \ell_s]$ and $y \in \mathbb{R}$. To that end, We show now that, for all $m \in \mathbb{N}$, the series $\sum_{k=1}^{\infty} \hat{E}_s^{t_m}(x,k)$ is uniformly convergent. We prove this by showing that it satisfies the Weirstrass criterion for uniform convergence. We write the following bounds using (48), and the facts that $\theta(k) = \sqrt{\eta + (\frac{k\pi}{L_2})^2} \ge \frac{k\pi}{L_2}$, and $\rho_k < 1$,

$$\begin{aligned}
\left|\hat{E}_s^m(x,k)\right| &= \left| A_s^m(k) e^{-\theta(k)|x-(\ell_{s-1}-L)|} + B_s^m(k) e^{-\theta(k)|x-(\ell_s+L)|} \right| \\
&\le |A_s^m(k)| e^{-\theta(k)|x-(\ell_{s-1}-L)|} + |B_s^m(k)| e^{-\theta(k)|x-(\ell_s+L)|} \\
&\le \rho_k^{\phi(m)} \|c^0\|_\infty(k) \left(e^{-\theta(k)|x-(\ell_{s-1}-L)|} + e^{-\theta(k)|x-(\ell_s+L)|} \right) \\
&\le \frac{M}{1-e^{-2\sqrt{\xi}(W+L)}} (1 + e^{-\theta(k)(W+L)}) \left(e^{-\theta(k)|x-(\ell_{s-1}-L)|} + e^{-\theta(k)|x-(\ell_s+L)|} \right) \\
&\le \frac{M}{1-e^{-2\sqrt{\xi}(W+L)}} \left(1 + e^{-\frac{k\pi}{L_2}(W+L)} \right) \left(e^{-\frac{k\pi}{L_2}|x-(\ell_{s-1}-L)|} + e^{-\frac{k\pi}{L_2}|x-(\ell_s+L)|} \right). \\
&\le \frac{M}{1-e^{-2\sqrt{\eta}(W+L)}} \left(e^{-\left(\frac{k\pi}{L_2}\right)|x-(\ell_{s-1}-L)|} + e^{-\left(\frac{k\pi}{L_2}\right)|x-(\ell_s+L)|} \right. \\
&\quad \left. + e^{-\left(\frac{k\pi}{L_2}\right)[W+L+|x-(\ell_{s-1}-L)|]} + e^{-\left(\frac{k\pi}{L_2}\right)[W+L+|x-(\ell_s+L)|]} \right) =: g(x,k).
\end{aligned} \tag{50}$$

Thus, we have $|\hat{E}_s^m(x,k)| \le g(x,k)$ for any $m \in \mathbb{N}$. We show next that the series $\sum_{k=1}^{\infty} g(x,k)$ of nonnegative terms is convergent. Indeed, note that for $x \in \Sigma \cap [\ell_{s-1}, \ell_s]$, we have $|x-(\ell_{s-1}-L)|, |x-(\ell_s+L)| \ge 2L$. Therefore, each of the four terms in parenthesis in (50) is bounded either by $\left(e^{\frac{\pi}{L_2}2L}\right)^{-k}$ or $\left(e^{\frac{\pi}{L_2}(W+3L)}\right)^{-k}$, and therefore, the series is convergent. Thus, the proof is complete. □

The proof of Theorem 2, and especially Step 2, uses the idea that when a processor updates the information of the approximation to the solution of the corresponding subdomain, such approximation moves to a smaller neighborhood of the solution, that is, the fixed point of the opertator $\mathcal{T}$. This is conceptually similar to classical proofs of convergence of asynchronous iterations, such as those by Bertsekas[22] or El Tarazi.[23]

4 | HOW THE CONVERGENCE DEPENDS ON THE PARAMETERS

We present in this section two new results. In the first, we show that if we maintain the height of the domain and the overlap constant, the wider is each subdomain, the bounds on the convergence rates of the synchronous and asynchronous optimized Schwarz methods decrease, indicating a faster converge. This comment is consistent with the observation by

TABLE 1 Number of iterations or average and maximum updates, and computational time (in seconds)

# Subdomains	# Iter Synchronous	Time	# Updates avg Asynchronous	Max	Time
8	845	6,265	908	1,152	4,462
16	972	3,087	1,105	1,934	2,151
32	1,681	2,084	1,976	2,843	1,534

Xu[24] for two subdomains (for the synchronous case). Another interpretation of this result is that if we keep the domain fixed, and we increase the number of subdomains, then the convergence factor would deteriorate. This is in fact confirmed by our numerical experiments in Section 6; see Table 1. In the second result, we show that once we fix the width of each subdomain, the convergence bounds for the methods are independent of the number of subdomains.

Theorem 3. *The convergence bounds of the synchronous and asynchronous optimized Schwarz iterations given by (4) and (39) decrease as the ratio W/L_2 increases by increasing W and keeping L_2 constant.*

Proof. The convergence factor is given by

$$\rho := \sup_k \left\| \hat{\mathcal{T}}_k \right\|_\infty. \tag{51}$$

Note that $\|\hat{\mathcal{T}}_k\|_\infty$ is the maximum over the quantities on the right-hand side of (31) and (32). Fixing the values of all the parameters except for W in these quantities and taking the derivative over W, we obtain

$$\frac{\partial}{\partial W}\left(\|S^{-1}R_1^{(l)}|0|S^{-1}R_2\|_\infty\right) = -\frac{\theta(k)\left(2|\gamma|\left(\gamma^2+e^{4L\theta(k)}\right)e^{2\theta(k)(L+W)}+e^{\theta(k)W}\left(\gamma^2+e^{5L\theta(k)}\right)\left(\gamma^2+e^{2\theta(k)(2L+W)}\right)\right)}{\left(e^{2\theta(k)(2L+W)}-\gamma^2\right)^2} < 0,$$

$$\frac{\partial}{\partial W}\left(\|S^{-1}R_1|0|S^{-1}R_2^{(r)}\|_\infty\right) = -\frac{\theta(k)e^{\theta(k)W}\left(\gamma^2+e^{4L\theta(k)}\right)\left(2|\gamma|e^{\theta(k)(2L+W)}+\gamma^2+e^{2\theta(k)(2L+W)}\right)}{\left(e^{2\theta(k)(2L+W)}-\gamma^2\right)^2} < 0.$$

Thus, given that these derivatives are negative, the quantities on the right-hand side of (31) and (32) decrease as we increase the ratio W/L_2 by increasing W and keeping L_2 constant, for all $k \in \mathbb{N}$. Consequently, the maximum over these two quantities, $\|\hat{\mathcal{T}}_k\|_\infty$, also decreases for all $k \in \mathbb{N}$. Then, given that the supremum in (51) is attained, that is, there exists a $k_0 \in \mathbb{N}$ such that $\rho := \|\hat{\mathcal{T}}_{k_0}\|_\infty$, and given that $\|\hat{\mathcal{T}}_k\|_\infty$ decreases for all $k \in \mathbb{N}$, the convergence factor decreases when we increase the ratio W/L_2 by increasing W and keeping L_2 constant. □

We remark that because $H = W+2L$, then taking the derivative with respect to W is the same as taking the derivative with respect to H (for the interior subdomains). We can interpret this result intuitively as follows. If the width of the subdomain is larger, then the boundary data, which are Dirichlet (i.e., the physical boundaries for which exact data is known), are proportionally larger than boundary data coming from the neighboring subdomains, which is only an approximation to the values of the solution, and this fact leads to faster convergence.

Theorem 4. *Consider the synchronous and asynchronous optimized Schwarz iterations given by (4) and (39), respectively. For a fixed size of the subdomains, $W\times L_2$, and a fixed amount of overlap L, the convergence factors $\rho = \sup_k\|\hat{\mathcal{T}}_k\|_\infty$ do not change if one changes the number of subdomains p (as long as $p > 2$), whereas $\sup_k\rho(|\hat{\mathcal{T}}_k|)$ changes by a small amount, namely of the order of $e^{-\theta(k)(W-L)} \approx 0$.*

Proof. First observe that from (31) and (32), $\|\hat{\mathcal{T}}_k\|_\infty$ is independent of the number of subdomains p. Therefore, the convergence factor $\rho = \sup_k\|\hat{\mathcal{T}}_k\|_\infty$ is also independent of p.

For the result on the spectral radii, consider the matrices $\hat{\mathcal{T}}_k$ as in Figure 3, with entries as in (26)–(29). Observe then that

$$\theta(k)W = \sqrt{\eta+\left(\frac{k\pi}{L_2}\right)^2}W \geq \frac{k\pi W}{L_2} \text{ and } \theta(k)2L = \sqrt{\eta+\left(\frac{k\pi}{L_2}\right)^2}2L \geq \frac{k\pi 2L}{L_2}.$$

Note that $W \gg 2L$; therefore, $e^{-\theta(k)W} \ll e^{-\theta(k)2L}$. Moreover, assuming that $W > L_2$, we have that $e^{-\theta(k)W} \approx 0$, $e^{-\theta(k)(W+2L)} \approx 0$, and $e^{-\theta(k)(W-L)} \approx 0$. As a consequence, we can write from (26)–(29) that $\hat{\mathcal{T}}_k = \tilde{\mathcal{T}}_k + \delta_T$, where $\tilde{\mathcal{T}}_k$ is a

square tridiagonal matrix of order $N = 2(p-2)+2$ such that the entries on the main diagonal are zero and the entries on the first subdiagonal and superdiagonal are

$$\left(\tilde{\mathcal{T}}_k\right)_{i,i-1} = \begin{cases} 0, & \text{if } i \text{ is odd} \\ b, & \text{if } i \text{ is even,} \end{cases} \quad \text{and} \left(\tilde{\mathcal{T}}_k\right)_{i,i+1} = \begin{cases} b, & \text{if } i \text{ is odd} \\ 0, & \text{if } i \text{ is even,} \end{cases}$$

respectively, $b = \gamma e^{-\theta(k)2L}/(1-\gamma^2 e^{-2\theta(k)(W+2L)})$, and the nonzero entries of the matrix δ_T are very small. It follows from the structure of $\tilde{\mathcal{T}}_k$ that the spectral radii of $\tilde{\mathcal{T}}_k$ and $|\tilde{\mathcal{T}}_k|$ are equal to $|b|$. Therefore, we have that $\rho(|\hat{\mathcal{T}}_k|) \approx \rho(|\tilde{\mathcal{T}}_k|) = |b| = \gamma e^{-\theta(k)2L}/(1-\gamma^2 e^{-2\theta(k)(W+2L)})$. Because b is independent of the number of subdomains p, it follows that $\rho(|\hat{\mathcal{T}}_k|)$ is essentially independent of the number of subdomains. □

Another way to see these results is to observe that the spectral radius, as well as the max norm of the matrix in Figure 3, does not depend on how many blocks there are, only on the values of the entries in each block. This result is illustrated numerically in the next section in Figure 12.

5 | OPTIMAL VALUES OF THE ROBIN PARAMETER α

In Sections 2 and 3, we proved the convergence of the optimized Schwarz methods using the maximum norm assuming that $\lambda(k)$, W, and L satisfy conditions (11)–(13). In this section, the discussion is geared toward finding good values for the optimal or near-optimal Robin parameter α in the $OO0$ case for the synchronous and asynchronous algorithms. We restrict our discussion to the case $\eta = 0$.

Traditionally, the optimal value of the Robin parameter α in (2) is defined as the one that minimizes the spectral radius of the iteration operator, that is, the one that minimizes the asymptotic convergence factor of the iterative method. In our situation, because we have shown convergence using the maximum norm of the iteration operator for each frequency k, we can define the optimal value of α as the one that minimizes

$$\max_{1\le k\le\infty} \rho(\hat{\mathcal{T}}(k)), \tag{52}$$

or alternatively,

$$\max_{1\le k\le\infty} \|\hat{\mathcal{T}}(k)\|_\infty. \tag{53}$$

A third and important possibility is to consider

$$\max_{1\le k\le\infty} \rho\left(|\hat{\mathcal{T}}(k)|\right); \tag{54}$$

this is because the spectral radius of the absolute value of the iteration matrix is related to the asymptotic convergence of asynchronous iterations; see, for example, the work of Chazan et al.[21] As we shall see, the minimizers of (52), (53), and (54) are in fact very close to each other; therefore, any of these definitions can be used to obtain a practical value of the Robin parameter.

Remark . Note that in all our calculations, the maximum spectral radius of the iteration operators becomes monotonically decreasing with k, after certain value of k, say, k_0. Therefore, it suffices to compute the optimal parameter by considering only values of k in the range $1 \le k \le k_{\max}$, for some value of $k_{\max}$. In our experience, for our problem, a value of $k_{\max} = 20$ suffices.[‡] Consequently, finding the optimal α may not be an expensive operation.

This can be seen empirically in Figures 4 and 5, where we report the value of the spectral radii of $\hat{\mathcal{T}}(k)$ and $|\hat{\mathcal{T}}(k)|$, respectively, as a function of the frequency k, for the case of $p = 10$ subdomains and a value for the ratio $L/H = 0.01$, that is, when the overlap L is about 1% of the width of the subdomain, that is a total of 2% of the subdomain on each side. In these figures, and throughout this section, we consider for simplicity that $H = W + 2L$, and define the normalized Robin parameter by $\bar{\alpha} = \alpha H$ and the normalized overlap by $\bar{L} = L/H$. We note that it follows from the description of matrix $\hat{\mathcal{T}}(k)$ in Section 2 that its entries depend on $\bar{L}$, $\bar{\alpha}$, and k, whereas its block structure depends on p, the number of subdomains.[§]

[‡]For discrete problems, $k_{\max}$ might be different, depending on the discretization used, but never very large.
[§]We note that because we have explicit formulas for the entries of the matrices, for each k, $k = 1, \ldots, k_{\max}$, and the matrices are not so large, we can easily compute the spectral radii.

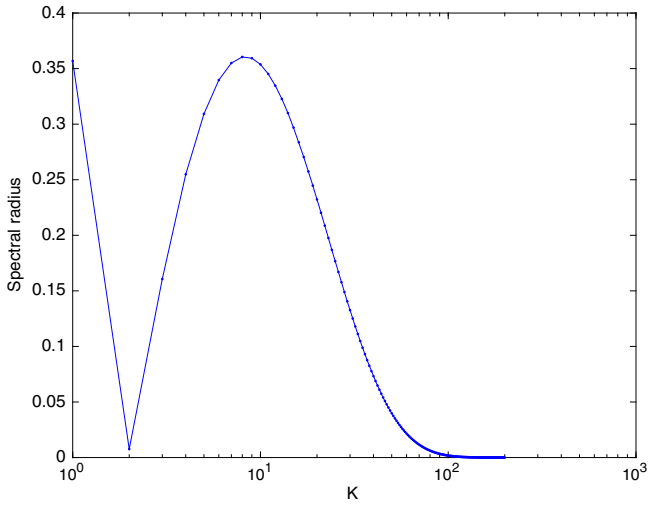

FIGURE 4 Spectral radius of $\hat{\mathcal{T}}(k)$ versus k for $p = 10$, $\bar{L} = \frac{L}{H} = 0.01$, and optimal $\bar{\alpha} = 6.3636$

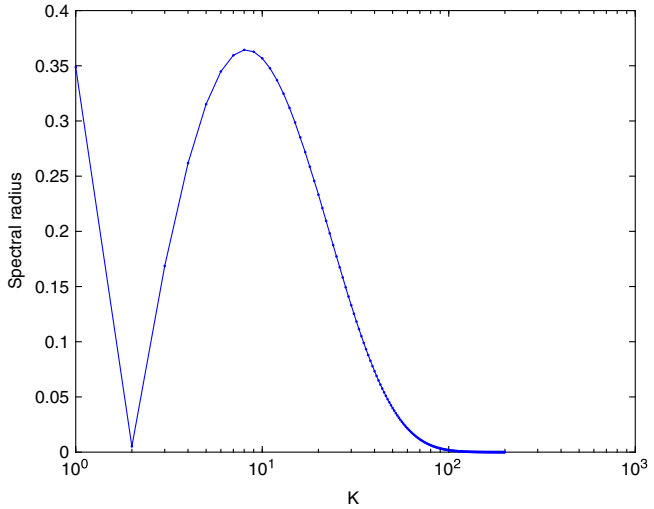

FIGURE 5 Spectral radius of $|\hat{\mathcal{T}}(k)|$ versus k for $p = 10$, $\bar{L} = \frac{L}{H} = 0.01$, and optimal $\bar{\alpha} = 6.2373$

The optimal values of the normalized parameter $\bar{\alpha}$ for the synchronous case is $\bar{\alpha} = 6.3636$, whereas that of the asynchronous case is $\bar{\alpha} = 6.2373$, that is, they differ by about 2% (cf. Figures 4 and 5). Furthermore, one can see in Figures 6–8 that not only the optimal values of the normalized Robin parameter are close to each other, but the value of the maximum norm and the spectral radius are also very close to each other for the range of values of interest. We also observe in these three figures that, for a fixed amount of overlap and a fixed number of subdomains, the convergence factors have only one minimum. Furthermore, for all positive values of $\bar{\alpha}$, the spectral radius of $\hat{\mathcal{T}}(k)$ is below one for all positive values of $\bar{\alpha}$, and the spectral radius of $|\hat{\mathcal{T}}(k)|$, and it is below one for all values of $\bar{\alpha}$ larger than a small constant. We mention that we have observed these properties for all other values of p and $\bar{L}$ for which we run our experiments; in other words, the graphs in Figures 4 and 5 are representative.

Our experiments include the computation of the spectral radius and the maximum norm of $\hat{\mathcal{T}}(k)$, and of its absolute value, for values of the normalized overlap $\bar{L}$ in the range from 0.001 to 0.06, that is, from 0.1% to 6% of the width of the subdomains; values of the normalized Robin parameter $\bar{\alpha}$ in the range from 0.01 to 300; frequency $k \in [1, 200]$; and number of subdomains p with values 5,10,20,30, and 40.

In Figure 9, we report the value of the spectral radius corresponding to the optimal parameter $\bar{\alpha}$, with varying (normalized) overlap from 0.1% to 6% of the width of the subdomains. We note that the spectral radius is monotonically decreasing as the overlap increases. This is consistent with our intuition as well as with the comparison theorems between two different splittings of a matrix (see, e.g., the work of Varga[25]): The larger the overlap, the faster the asymptotic convergence of the method.

In Figure 10, we show how the value of the optimal parameter $\bar{\alpha}$ varies with the normalized overlap. We found that this optimal value is monotonically decreasing with the increase in overlap. Here, the optimal value refers to that which minimizes $\rho(|\hat{\mathcal{T}}(k)|)$.

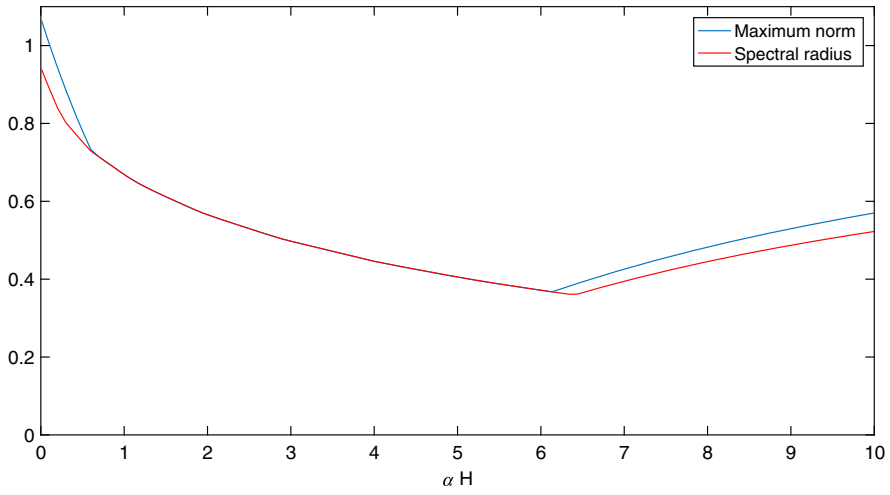

FIGURE 6 Maximum norm and spectral radius of $\hat{\mathcal{T}}(k)$ versus $\bar{\alpha}$ for $\bar{\alpha} \in [1, 10]$, $p = 10$, $\bar{L} = \frac{L}{H} = 0.01$, and $1 \leq k \leq k_{\max} = 20$

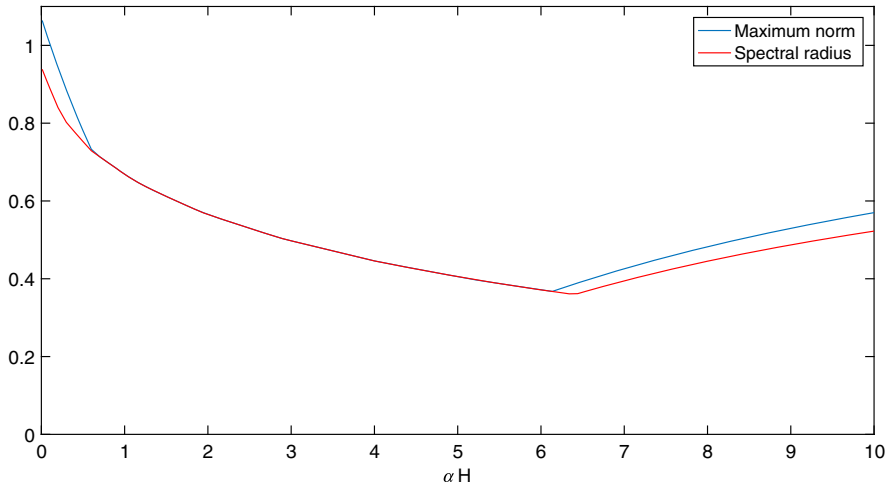

FIGURE 7 Maximum norm and spectral radius of $\hat{\mathcal{T}}(k)$ versus $\bar{\alpha}$ for $\bar{\alpha} \in [1, 10]$, $p = 10$, $\bar{L} = \frac{L}{H} = \frac{1}{30}$, and $1 \leq k \leq k_{\max} = 20$

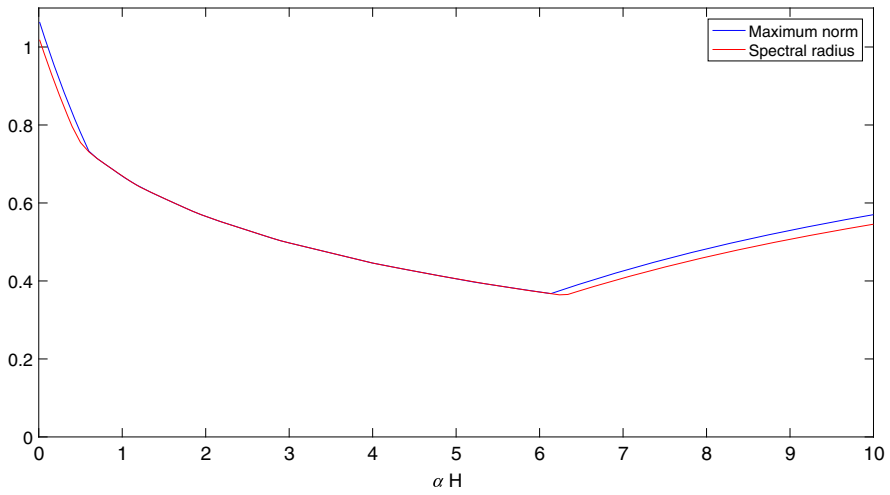

FIGURE 8 Maximum norm and spectral radius of $|\hat{\mathcal{T}}(k)|$ versus $\bar{\alpha}$ for $\bar{\alpha} \in [1, 10]$, $p = 10$, $\bar{L} = \frac{L}{H} = 0.01$, and $k_{\max} = 20$

Our next set of experiments show that both the optimal Robin parameter and the corresponding spectral radii do not change with the change of the number of subdomains, for a fixed value of $\bar{L}$; see Figures 11 and 12. Here again, the optimal value of the Robin parameter refers to that which minimizes $\rho(|\hat{\mathcal{T}}(k)|)$. We note that this behavior is consistent with the structure of the iteration matrix as shown in Figure 2, and the fact that the spectral radius (and the maximum norm) of such a matrix is independent of the number of blocks.

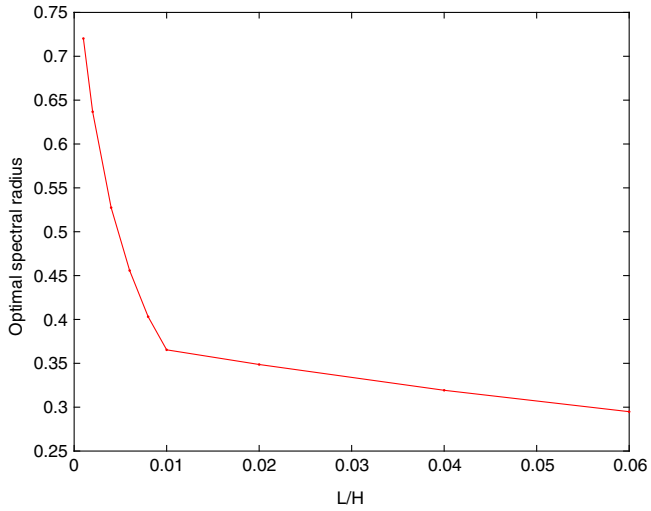

FIGURE 9 Optimal spectral radius of $|\hat{\mathcal{T}}(k)|$ versus $\bar{L} = \frac{L}{H}$ for $\bar{L} \in [0.001, 0.06]$, $p = 10$, $k_{\max} = 20$

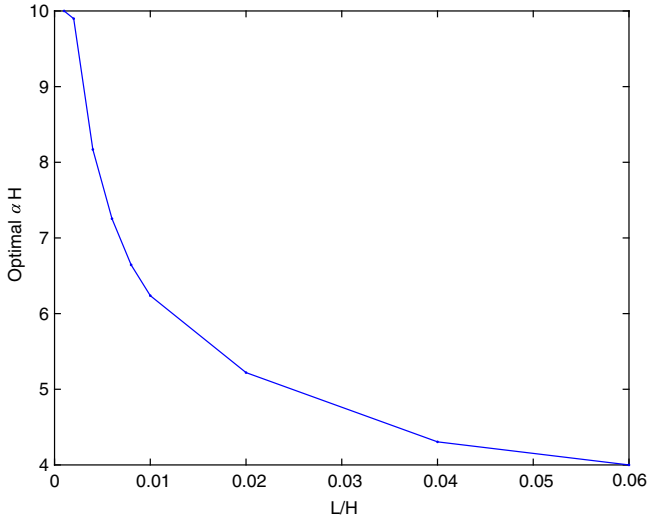

FIGURE 10 Optimal $\bar{\alpha}$ versus $\bar{L}$ for p=10, $k_{\max} = 20$

In summary, in our extensive experiments, we have observed the following.

1. There exist values of the normalized Robin parameter $\bar{\alpha}$ for which the maximum norm and the spectral radii of the iteration matrices are less than one.
2. For a given fixed $\bar{L}$ and the range of $\bar{\alpha}$ considered in the experiments, the maximum norm and the spectral radii of the iteration matrices have only one minimum and it approaches a constant less than one for large values of $\bar{\alpha}$.
3. Given fixed values of $\bar{L}$, $\bar{\alpha}$, p, and for values of k larger than some fixed value, the value of the maximum norm and the spectral radii of the iteration matrices are monotonically decreasing as k increases.
4. The optimal maximum norm and the spectral radii of the iteration matrices decrease monotonically, as $\bar{L}$ increases.
5. The optimal normalized Robin parameter $\bar{\alpha}$ decreases monotonically as $\bar{L}$ increases.
6. For a given value of the normalized overlap $\bar{L}$, and a fixed size of the subdomains, both the optimal normalized Robin parameter $\bar{\alpha}$, and the corresponding spectral radii remain constant as the number of subdomains increases.

In the rest of this section, we develop approximate formulas for the value of the optimal normalized Robin parameter $\bar{\alpha} = \alpha H$. As we have mentioned, the maximum norm and the spectral radii of the iteration operator depends on $\bar{\alpha}$, as well as of the normalized overlap $\bar{L} = L/H$, the number of subdomains p, and the parameter $k_{\max}$. Recall that the maximum norm and the spectral radii remain essentially constant for $k \geq k_{\max}$, for some value of $k_{\max}$. Thus, taking $k_{\max}$ large enough so that this holds, we observe that the values of $\bar{\alpha}_{\text{opt}}$ follows approximately a power law. Thus, we have for small normalized overlap that

$$\bar{\alpha}_{\text{opt}} \approx c\bar{L}^{\ell}, \tag{55}$$

where the constants c and ℓ are essentially independent of p, the number of subdomains.

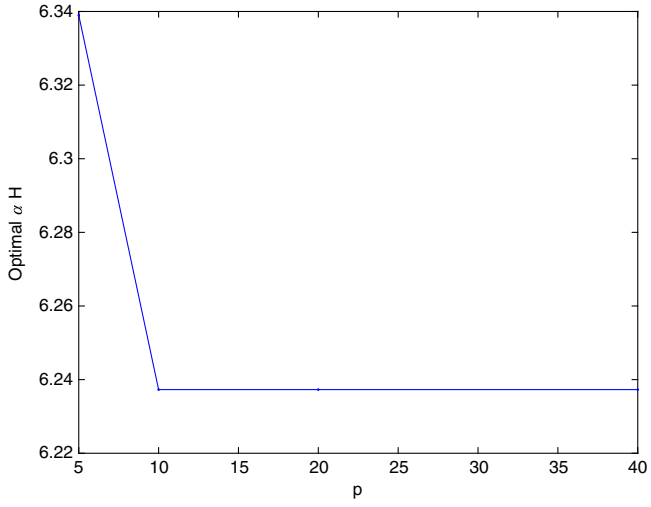

FIGURE 11 Optimal $\bar{\alpha}$ versus p for $\bar{L} = 0.01$, $k_{\max} = 20$

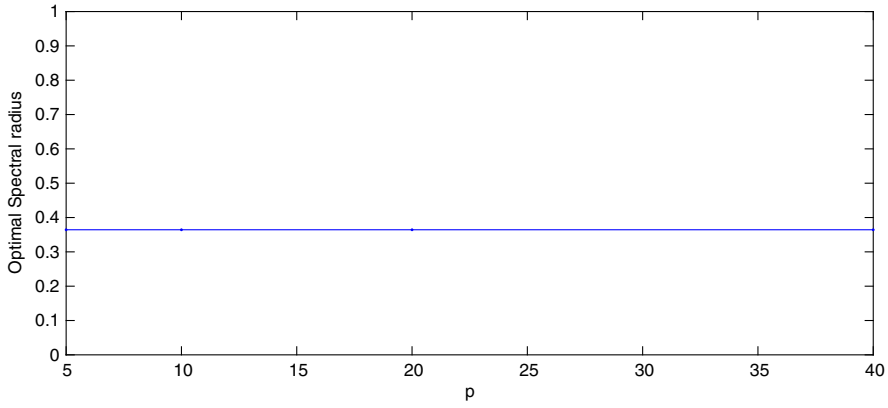

FIGURE 12 Optimal spectral radius of $|\hat{\mathcal{T}}(k)|$ versus p for $k_{\max} = 20$, $\bar{L} = \frac{L}{H} = 0.01$ and $\bar{\alpha} = 6.2373$

	$\mathbf{argmin}\rho(\hat{\mathcal{T}})$		$\mathbf{argmin}\rho(\lvert\hat{\mathcal{T}}\rvert)$	
p	c	ℓ	c	ℓ
5	1.7556	−0.2807	1.8322	−0.2664
10	1.7725	−0.2803	1.8434	−0.2690
20	1.7671	−0.2812	1.8841	−0.2605

TABLE 2 Coefficients of empirical formula of $\bar{\alpha}_{\text{opt}}$ for different numbers of subdomains p for the synchronous case (left) and the asynchronous case (right)

We report in Table 2 the values of the constant c and exponent ℓ for the case $\eta = 0$, for $p = 5, 10, 20$. These were computed using data fitting with a power function, that is, given the curve $\bar{\alpha}_{\text{opt}}$ versus $\bar{L}$ obtained by computing the optimal values of $\bar{\alpha}$, we take its log, and approximate this with a line. As it can be observed, the values for c and ℓ are essentially the same for different values of p. Furthermore, the values for the synchronous and asynchronous cases are not that different. In Figures 13 and 14, the comparison between the computed values of $\bar{\alpha}_{\text{opt}}$ and their approximation obtained using (55) are presented for different amounts of overlap for the synchronous and asynchronous case, respectively.

We also observed that the values of $\max_{k=1,\dots,k_{\max}} \rho(\hat{\mathcal{T}}(k))$ follows approximately a power law, and therefore, we can write a formula for the spectral radius, which indicates the asymptotic rate of convergence of the method. Thus, we have that

$$\max_{k=1,\dots,k_{\max}} \rho(\hat{\mathcal{T}}(k)) \approx c\bar{L}^{\ell}, \tag{56}$$

where the constants c and ℓ are essentially independent of p, the number of subdomains.

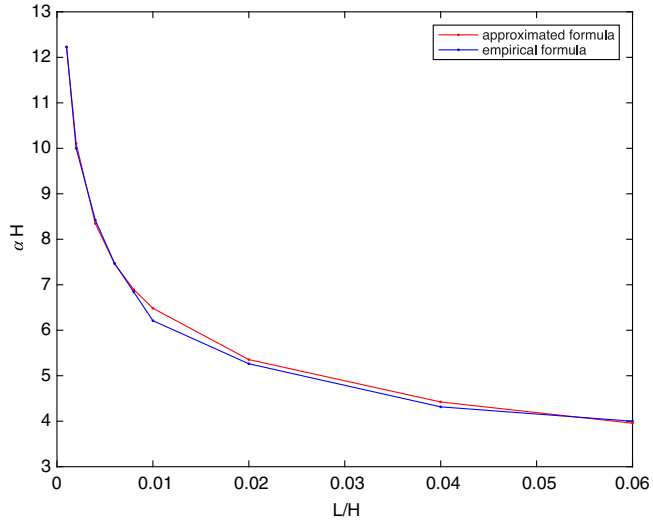

FIGURE 13 Computed values of $\bar{\alpha}_{\text{opt}}$ and values of $\bar{\alpha}_{\text{opt}}$ using (55) for different values of overlap $\bar{L}$ and $p = 10$, where $\rho(\hat{\mathcal{T}})$ is minimized

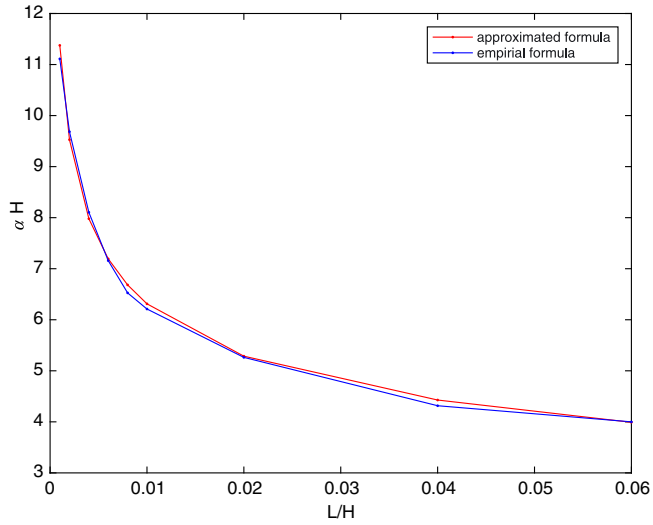

FIGURE 14 Computed values of $\bar{\alpha}_{\text{opt}}$ and values of $\bar{\alpha}_{\text{opt}}$ using (55) for different values of overlap $\bar{L}$ and $p = 10$, where $\rho(|\hat{\mathcal{T}}|)$ is minimized

TABLE 3 Coefficients of empirical formula of $\max_{k=1,\dots,k_{\max}} \rho(\hat{T}(k))$ for different numbers of subdomains p for the synchronous case (left) and the asynchronous case (right)

	$\max_k \rho(\hat{\mathcal{T}}(k))$		$\max_k \rho(\lvert\hat{\mathcal{T}}(k)\rvert)$	
p	c	ℓ	c	ℓ
5	0.0468	−0.3879	0.0525	−0.3721
10	0.0535	−0.3647	0.0531	−0.3702
20	0.0525	−0.3677	0.0576	−0.3501

We report in Table 3 the values of the constant c and exponent ℓ for $p = 5, 10, 20$. In Figures 15 and 16, the comparison between the computed values of $\max_{k=1,\dots,k_{\max}} \rho(\hat{T}(k))$ and their approximation obtained using (56) are presented for different amounts of overlap for the synchronous and asynchronous case, respectively. Note that the values of these constants are close to one another, for the various values of p and for the two cases considered here.

6 | NUMERICAL EXPERIMENTS

We present numerical experiments that illustrate the performance of the proposed asynchronous optimized Schwarz method on a three-dimensional bounded domain (as well as the synchronous counterpart). The experiments show the convergence of the methods, illustrating the results of our theorems, and show that the asynchronous version is faster in terms of execution time.

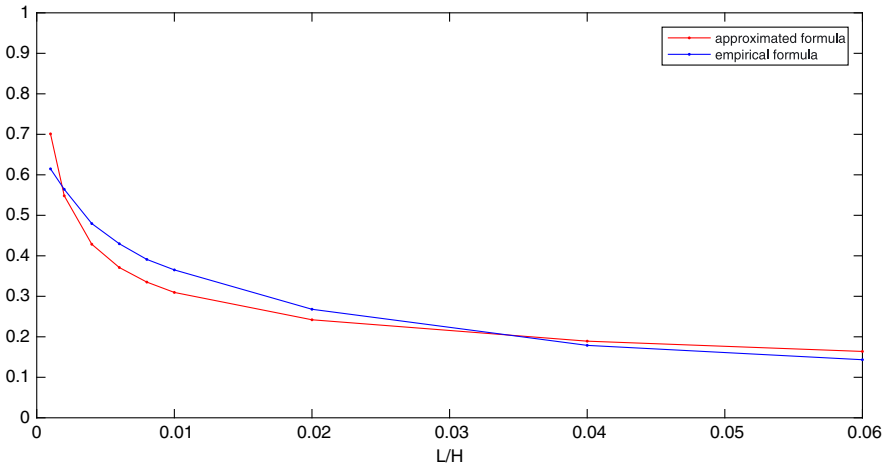

FIGURE 15 Computed values of $\max_{k=1,\dots,k_{\max}} \rho(\hat{T}(k))$ and values of $\max_{k=1,\dots,k_{\max}} \rho(\hat{T}(k))$ using (56) for different values of overlap $\bar{L}$ and $p = 10$

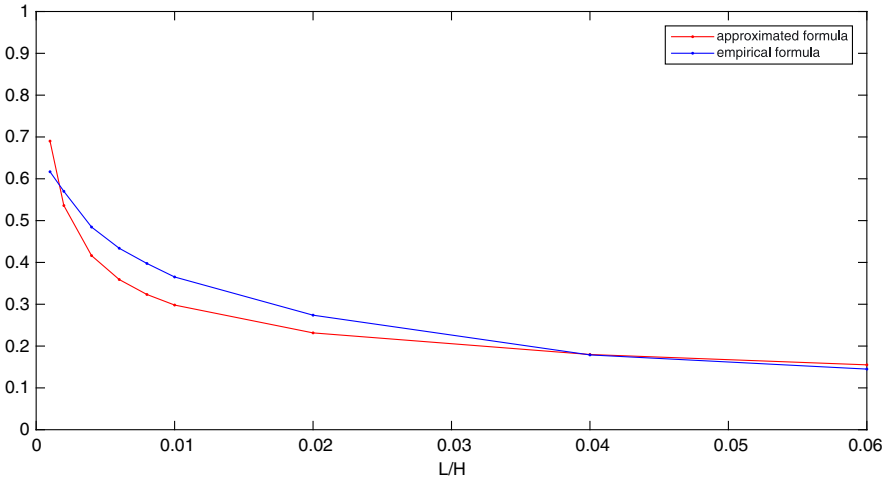

FIGURE 16 Computed values of $\max_{k=1,\dots,k_{\max}} \rho(|\hat{T}(k)|)$ and values of $\max_{k=1,\dots,k_{\max}} \rho(\hat{T}(k))$ using (56) for different values of overlap $\bar{L}$ and $p = 10$

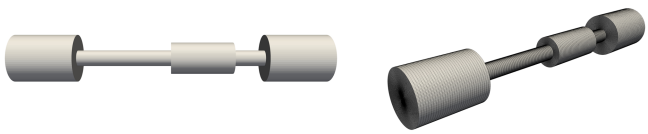

FIGURE 17 Computer-aided design (left) and finite element mesh (right) of the air intake system

The test cases are related to the study of the heat analysis in a three-dimensional air intake system. This problem is modeled as follows:

$$\rho \frac{\partial u}{\partial t} - \nabla(\kappa \nabla u) = f,$$

where u denotes the thermal field, κ is the thermal conductivity, and f is the heat-flux density of the source. Here, the steady-state heat equation is considered, that is, $\frac{\partial u}{\partial t} = 0$ which is by definition not time dependent, and which corresponds to the case where enough time has passed such that the thermal field no longer evolves in time. This leads to the following reduced equation $-\kappa \nabla^2 u = f$.

The numerical solution of the steady-state heat equation required for this study was performed on an air intake system of dimensions in the x-, y-, and z-directions equal to 1,401 mm, 210 mm, and 210 mm, respectively. The computer-aided design of the air intake is shown in the left part of Figure 17.

The model consists of imposed temperature on the inlet cavity, and zero temperature gradient boundary conditions are imposed on the outlet cavity. This means that the outlet cavity has reflection properties equivalent to open end.

The domain is meshed with unstructured hexahedra finite elements, and one example of a finite element mesh with 829,599 degrees of freedom and 801,440 finite element is shown in the right part of Figure 17. This mesh is split into 8, 16, and 32 subdomains in a one-way direction. These partitionings are illustrated in Figure 18. In all cases, the overlap used is

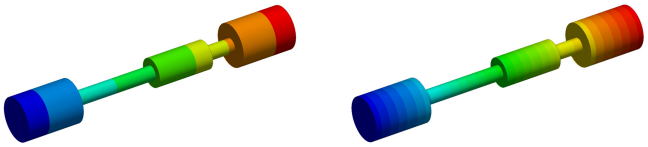

FIGURE 18 Mesh partitioning of the finite element mesh into eight subdomains (left) and 32 subdomains (right)

the minimum overlap, that is, one set of common nodes in the boundary between subdomains. The optimized coefficients of the Schwarz algorithms are obtained using the values of $\rho(|\hat{\mathcal{T}}_{k_{\max}}|)$ in terms of $\bar{\alpha} = \alpha H$, for the different values of p, in the same manner as was done in the previous section. Here, H is the average diameter of the subdomains.[¶] The resulting parameters were divided by H, leading to (nonnormalized) values of the synchronous optimized parameter α equal to 0.6300, 0.3868, 0.4310 for a partitioning into 8, 16, and 32 subdomains, respectively. The same value of the parameters is used for the synchronous and asynchronous version of the code.

To solve the resulting linear system, the synchronous and asynchronous optimized Schwarz methods with zeroth-order optimized interface conditions were implemented in the recently developed C++ library *Alinea*.[26] The parallel implementation of the asynchronous optimized Schwarz methods is quite similar to the synchronous implementation we have described in the work of Magoulès et al.,[27] except that the asynchronous iterations and asynchronous communications are managed by a new additional layer. This new additional layer, *JACK*,[28] a recently developed C++ library, is defined on top of the MPI library; the version of the MPI library used in the experiments is MPICH2.[29] This layer allows us to use asynchronous communications between the processors and to deal with continuous requests. This new layer also contains new functionality such as the detection of the asynchronous convergence of the algorithm for a given stopping criteria. Here, we use the stopping criterion developed in the work of Magoulès et al.[30] with a threshold equal to 10^{-8}; it is based on asynchronous iterations termination methods based on global residual, under various communication models.

The experiments are performed on a heterogeneous cluster composed of four nodes (Intel(R) Xeon(R) E5410, 2.33 GHz, 8 cores, RAM: 8 GB), four nodes (Intel(R) Core(TM) i7 2.80 GHz, 8 cores, RAM: 8 GB) each with graphics processing units accelerator (Tesla K20c 4799MB, GTX 570, 1,279 MB) and four nodes (Intel(R) Xeon(R) E5-2609, 2.10 GHz, 24 cores, RAM: 16 GB) each with graphics processing units accelerator (three of them with Quadro K4000 3 071 MB, and one with Quadro K600 1,023 MB), for a total of 160 cores. The interconnected network is a switched, star-shaped 10-Mb/s Ethernet network. We report our computational results in Table 1, where we vary the number of subdomains using the same discretization. Recall that in asynchronous iterations, we cannot talk about (global) iteration steps because each processor may update its approximation to the (local) solution at different moments, that is, on different time stamps. In fact, each processor (corresponding to each subdomain) would usually perform a different number of updates, that is, of local solutions. Thus, in Table 1, for the synchronous case, the number of iterations are reported, whereas for the asynchronous case, the average and the maximum number of updates among all processors are reported. In both cases, the total computational time (in seconds) is shown. It can be appreciated that the asynchronous optimized Schwarz performs better in terms of execution time than its synchronous counterpart. In fact, for each case, the time to converge for the asynchronous runs about 25% to 30% faster than in the synchronous case.

We mention that this experiment shows a situation for which it is natural to consider asynchronous methods. First, the domain of interest cannot be partitioned into identical subdomains. Second, the computational network is heterogeneous. Both of these practical situations indicate that load balancing would be an issue. We also mention that even in cases where the subdomains are identical, and the network is homogeneous, for very large problems, and very large networks, the communication among processors would generate the imbalance, and the asynchronous methods would be advantageous as well.

7 | CONCLUSION

In this paper, we filled the gap of the convergence theory of synchronous and asynchronous optimized Schwarz methods for a bounded (rectangular) domain. We showed convergence for a one-way subdivision of the domain, for an arbitrary number of subdomains. We showed that the convergence factor of the method is independent of the number of

[¶]For different number of subdomains, the average diameter changes, and this is why the values of the optimized parameter changes as well.

subdomains (for a fixed size of the subdomains). We provide a formula for the optimal Robin parameter, as a function of the overlap. We complete the paper with numerical experiments on a three-dimensional air intake problem, which illustrate our theoretical results.

ACKNOWLEDGEMENT

We are indebted to the two anonymous referees for their questions and comments, which helped us improve our presentation. This research is supported in part by the U.S. Department of Energy under grant DE-SC0016578.

ORCID

Frédéric Magoulès https://orcid.org/0000-0002-1198-7539
Daniel B. Szyld https://orcid.org/0000-0001-8010-0391

REFERENCES

1. Schwarz Hermann A. Ueber einen Grenzübergang durch alternierendes Verfahren. Vierteljahresschrift der Naturforschenden Gesellschaft in Zürich. 1870;15:272–286.
2. Dolean V, Jolivet P, Nataf F. An introduction to domain decomposition methods: Algorithms, theory, and parallel implementation. Philadelphia, PA: Society for Industrial and Applied Mathematics; 2015.
3. Toselli A, Widlund O. Domain decomposition methods - Algorithms and theory. Berlin, Germany: Springer; 2005. (Series in computational mathematics; No. 34).
4. Chevalier P, Nataf F. Symmetrized method with optimized second-order conditions for the Helmholtz equation. In: Mandel J, Farhat C, Cai X-C, editors. Domain Decomposition 10. Providence, RI: American Mathematical Society; 1998. p. 400–407. (Contemporary Mathematics vol. 218).
5. Gander MJ. Optimized Schwarz methods. SIAM J Numer Anal. 2006;44:699–731.
6. Japhet C. Optimized Krylov-Ventcell method: application to convection diffusion problems. In: Bjørstad PE, Espedal MS, Keyes DE, editors. Proceedings of the 9th International Conference on Domain Decomposition Methods in Science and Engineering; 1998; Bergen, Norway: ddm.org. p. 382–389.
7. Magoulès F, Roux F-X, Series L. Algebraic approximation of Dirichlet-to-Neumann maps for the equations of linear elasticity. Comput Methods Appl Mech Eng. 2006;195:3742–3759.
8. Chau M, Garcia T, Spiteri P. Asynchronous Schwarz methods applied to constrained mechanical structures in grid environment. Adv Eng Softw. 2014;74:1–15.
9. Frommer A, Schwandt H, Szyld DB. Asynchronous weighted additive Schwarz methods. Electron Trans Numer Anal. 1997;5:48–61.
10. Laitinen E, Lapin A, Pieska J. Asynchronous domain decomposition methods for continuous casting problem. J Comput Appl Math. 2003;194:393–413.
11. Magoulès F. Asynchronous Schwarz methods for peta and exascale computing. In: BHV Topping, P Iványi, editors. Developments in parallel, distributed, grid and cloud computing for engineering. Stirlingshire, UK: Saxe-Coburg; 2013. p. 229–248.
12. Magoulès F, Szyld D, Venet C. Asynchronous optimized Schwarz methods with and without overlap. Numerische Mathematik. 2017;137:199–227.
13. Garay JC, Magoulès F, Szyld DB. Convergence of asynchronous optimized Schwarz methods in the plane. In: Bjøstard PE, Brenner SC, Halpern L, Hyun Kim H, Kornhuber R, Rahman T, Widlund OB, editors. Domain decomposition methods in science and engineering XXIV. Berlin, Germany: Springer; 2018. (Lecture Notes in Computer Science and Engineering vol. 125).
14. Garay J, Magoulès F, Szyld D. Synchronous and asynchronous optimized Schwarz method for Poisson's equation in rectangular domains. Philadelphia, PA: Department of Mathematics, Temple University; 2018. Technical Report 17-10-18.
15. Haberman R. Applied partial differential equations with Fourier series and boundary value problems. Upper Saddle River, NJ: Prentice Hall; 2004.
16. Gander MJ. On the influence of geometry on optimized Schwarz methods. SeMA Journal. 2011;53:71–78.
17. Bertsekas DP, Tsitsiklis JN. Parallel and distributed computation: Numerical methods. Englewood Cliffs, NJ: Prentice Hall; 1989.
18. Frommer A, Szyld DB. On asynchronous iterations. J Comput Appl Math. 2000;123:201–216.
19. Szyld D B. Different models of parallel asynchronous iterations with overlapping blocks. Comput Appl Math. 1998;17:101–115.
20. Bahi JM, Contassot-Vivier S, Couturier R. Parallel iterative algorithms: From sequential to grid computing. Boca Raton, FL: CRC Press; 2007.
21. Chazan D, Miranker WD. Chaotic relaxations. Linear Algebra Appl. 1969;2:199–222.
22. Bertsekas D. Distributed asynchronous computation of fixed points. Mathematical Programming. 1983;27:107–120.
23. El Tarazi MN. Some convergence results for asynchronous algorithms. Numerische Mathematik. 1982;39:325–340.
24. Xu Y. The influence of domain truncation on the performance of optimized Schwarz methods. Electron Trans Numer Anal. 2018;49:182–209.
25. Varga RS. Matrix iterative analysis. 2nd ed. Berlin, Germany: Springer; 2000.

26. Magoulès F, Ahamed A-KC. Alinea: an advanced linear algebra library for massively parallel computations on graphics processing units. Int J High Perform Comput Appl. 2015;29(3):284–310.
27. Magoulès F, Ivanyi P, Topping BH. Non-overlapping Schwarz methods with optimized transmission conditions for the Helmholtz equation. Comput Methods Appl Mech Eng. 2004;193:4797–4818.
28. Magoulès F, Guillaume G-B. JACK: an asynchronous communication kernel library for iterative algorithms. J Supercomput. 2017;73:3468–3487.
29. Gropp W, Rajeev T, Lusk E. Using MPI-2: advanced features of the message-passing interface. Cambridge, MA: MIT Press; 1999.
30. Magoulès F, Gbikpi-Benissan G. Distributed convergence detection based on global residual error under asynchronous iterations. IEEE Trans Parallel Distrib Syst. 2018;29:819–829.

How to cite this article: El Haddad M, Garay JC, Magoulès F, Szyld DB. Synchronous and asynchronous optimized Schwarz methods for one-way subdivision of bounded domains. *Numer Linear Algebra Appl.* 2019;e2279. https://doi.org/10.1002/nla.2279