SIAM J. NUMER. ANAL.
Vol. 55, No. 3, pp. 1330–1356

ANALYSIS OF THE PARALLEL SCHWARZ METHOD FOR GROWING CHAINS OF FIXED-SIZED SUBDOMAINS: PART I*

G. CIARAMELLA† AND M. J. GANDER†

Abstract. In implicit solvation models, the electrostatic contribution to the solvation energy can be estimated by solving a system of elliptic partial differential equations modeling the reaction potential. The domain of definition of such elliptic equations is the union of the van der Waals cavities corresponding to the atoms of the solute molecule. Therefore, the computations can naturally be performed using Schwarz methods, where each atom of the molecule corresponds to a subdomain. In contrast to classical Schwarz theory, it was observed numerically that the convergence of the Schwarz method in this case does not depend on the number of subdomains, even without coarse correction. We prove this observation by analyzing the Schwarz iteration matrices in Fourier space and evaluating corresponding norms in a simplified setting. In order to obtain our contraction results, we had to choose a specific iteration formulation, and we show that other formulations of the same algorithm can generate Schwarz iteration matrices with much larger norms leading to the failure of norm arguments, even though the spectral radii are identical. By introducing a new optimality concept for Schwarz iteration operators with respect to error estimation, we finally show how to find Schwarz iteration matrix formulations which permit such small norm estimates.

Key words. domain decomposition methods, Schwarz methods, chain of atoms, elliptic PDE, COSMO solvation model

AMS subject classifications. 65N55, 65F10, 65N22, 70-08, 35J57

DOI. 10.1137/16M1065215

1. Introduction.

Recent developments in physical and chemical applications are creating a large demand for numerical methods for the solution of complicated systems of equations, which are often used before rigorous numerical analysis results are available. As a particular example we study here the new methodology that was recently presented in [2] and supported by [20, 21]. Based on a physical model approximation of solvation phenomena, called COSMO [1, 16, 26], the authors introduced the so-called ddCOSMO, which is a new formulation of the Schwarz methods for the solution of solvation problems where large molecular systems, given by chains of atoms, are involved, and each atom corresponds to a subdomain. The Schwarz methods are written in a boundary element form (see, e.g., [22] for a review of the application of boundary element methods), and no theoretical analyses of the algorithm are performed in [2, 20, 21]. The authors observe, however, in their numerical experiments an unusual convergence behavior of the Schwarz methods used: the convergence of the iterative procedure without coarse correction is in many cases independent of the number of atoms and thus subdomains; see, e.g., [2, Figure 10]. Schwarz methods are a mature field (see, e.g., [25, 8] and references therein), and it is well known that, in general, for elliptic problems, the convergence of Schwarz methods without a coarse space component depends on the number of subdomains. We prove in what follows that in the specific case of [2, 20, 21] for an approximate geometrical setting, the convergence indeed does not depend on the number of atoms in the molecule, and is

*Received by the editors March 10, 2016; accepted for publication (in revised form) January 17, 2017; published electronically May 30, 2017.
http://www.siam.org/journals/sinum/55-3/M106521.html
Funding: The work of the authors was supported by the Swiss National Science Foundation (SNSF) funding the projects PASC-DIAPHANE and 200020_149107/1.
†Section de Mathématiques, University of Geneva, CH-1211 Geneva, Switzerland (gabriele.ciaramella@unige.ch, martin.gander@unige.ch).

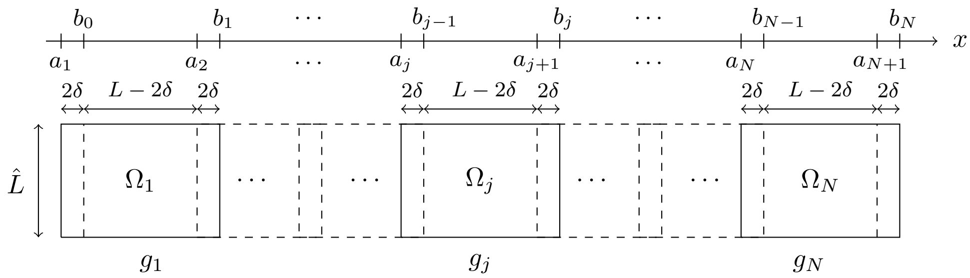

Fig. 1. *Grid of points used to define the geometry of the problem and our rectangular model for the chain of atoms.*

thus independent of the number of subdomains. This is our first main result, which we prove for an approximate 2-dimensional model that describes a chain of atoms whose domains are approximated by rectangles.

Our analysis leads us to estimate the spectral radius of iteration operators. This is in general a difficult task, and it is often easier to estimate norms. For a given method, different iteration operator formulations are possible that have the same spectral radius, but the iteration operator norms are in general different. This fact can strongly influence the convergence analysis, which in the worst case can even be inconclusive. To the best of our knowledge, there has been no attempt so far to characterize, describe, and compare such different iteration operator formulations in terms of error estimation. Our second main contribution is a precise characterization of the best Schwarz iteration operator formulation with respect to error estimation. Based on a new optimality concept for error estimation, this formulation allowed us to obtain our convergence result independent of the number of subdomains.

Our paper is organized as follows: in section 2, we introduce the approximate model that describes a chain of atoms whose 2-dimensional circular domains are approximated by rectangular domains, and we define the parallel Schwarz method for the solution to this approximate model. Section 3 is devoted to constructing Schwarz iteration matrices in Fourier space. In section 4, we prove convergence of the parallel Schwarz method independently of the number of subdomains by estimating norms of Schwarz iteration matrices. In section 5, we show that different iteration formulations are possible, which can lead to Schwarz iteration matrices having very different norms, even though the spectral radii are the same. We then introduce a new concept of optimality of transfer operators with respect to error estimation, and we prove optimality of the Schwarz iteration operator that we used in section 3. We illustrate our analysis with numerical experiments in section 6, and present our conclusions in section 7.

2. Parallel Schwarz method for a solvation model. We now define our simplified solvation model consisting of a chain of N rectangular atoms, and study a parallel one-level Schwarz method for computing its state. To define the domain of each atom, let $L \in \mathbb{R}^+$ and $\delta \in \mathbb{R}^+$, and define the grid points $a_j = (j-1)L - \delta$ for $j = 1, \dots, N+1$, and $b_j = jL + \delta$ for $j = 0, \dots, N$, as shown in Figure 1. The domain of the jth atom of the chain is a rectangle of dimension $\Omega_j := (a_j, b_j) \times (0, \hat{L})$. In addition, some forces f_j and g_j act in the interior and on the boundary of the jth atom. The chain of atoms is also shown in Figure 1. The state $u_j(x,y)$ of the jth atom is governed by the boundary value problem

$$\begin{aligned} &-\Delta u_j = f_j \quad \text{in } \Omega_j, \quad u_j(\cdot,0) = g_j(\cdot,0), \quad u_j(\cdot,\hat{L}) = g_j(\cdot,\hat{L}), \\ &\quad u_j = u_{j-1} \quad \text{in } [a_j, b_{j-1}] \times (0,\hat{L}), \quad u_j = u_{j+1} \quad \text{in } [a_{j+1}, b_j] \times (0,\hat{L}), \end{aligned} \tag{1}$$

where the last two conditions describe the interaction of the jth atom with atom $j-1$ and atom $j+1$, and the function f_j is equal to f_{j+1} or f_{j-1} in the overlap. We note that in the original model considered in [20], $f_j = 0$, but here we consider the more general case of possibly nonzero f_j. Notice that problem (1) is defined for $j = 2, \ldots, N-1$. The states of the first and the last atom are governed by

$$\begin{aligned} &-\Delta u_1 = f_1 \quad \text{in } \Omega_1, \\ &u_1(\cdot,0) = g_1(\cdot,0), \quad u_1(\cdot,\hat{L}) = g_1(\cdot,\hat{L}), \quad u_1(a_1,\cdot) = g_1(a_1,\cdot), \\ &u_1 = u_2 \quad \text{in } [a_2, b_1] \times (0,\hat{L}) \end{aligned} \tag{2}$$

and

$$\begin{aligned} &-\Delta u_N = f_N \quad \text{in } \Omega_N, \\ &u_N(\cdot,0) = g_N(\cdot,0), \quad u_N(\cdot,\hat{L}) = g_N(\cdot,\hat{L}), \quad u_N(b_N,\cdot) = g_N(b_N,\cdot), \\ &u_N = u_{N-1} \quad \text{in } [a_N, b_{N-1}] \times (0,\hat{L}). \end{aligned} \tag{3}$$

For a large number of atoms N, it is natural to apply a domain decomposition method to solve (1)–(3). In particular, we focus on the parallel Schwarz method,

$$\begin{aligned} &-\Delta u_j^n = f_j \quad \text{in } \Omega_j, \quad u_j^n(\cdot,0) = g_j(\cdot,0), \quad u_j^n(\cdot,\hat{L}) = g_j(\cdot,\hat{L}), \\ &u_j^n(a_j,\cdot) = u_{j-1}^{n-1}(a_j,\cdot), \quad u_j^n(b_j,\cdot) = u_{j+1}^{n-1}(b_j,\cdot) \end{aligned}$$

for $j = 2, \ldots, N-1$ and

$$\begin{aligned} -\Delta u_1^n &= f_1 \quad \text{in } \Omega_1, & -\Delta u_N^n &= f_N \quad \text{in } \Omega_N, \\ u_1^n(\cdot,0) &= g_1(\cdot,0), & u_N^n(\cdot,0) &= g_N(\cdot,0), \\ u_1^n(\cdot,\hat{L}) &= g_1(\cdot,\hat{L}), \qquad \text{and} & u_N^n(\cdot,\hat{L}) &= g_N(\cdot,\hat{L}), \\ u_1^n(a_1,\cdot) &= g_1(a_1,\cdot), & u_N^n(a_N,\cdot) &= u_{N-1}^{n-1}(a_N,\cdot), \\ u_1^n(b_1,\cdot) &= u_2^{n-1}(b_1,\cdot), & u_N^n(b_N,\cdot) &= g_N(b_N,\cdot). \end{aligned}$$

Notice that 2δ represents the overlap between two consecutive subdomains, and δ can assume values in the interval $(0, L/2)$. We remark that the Schwarz method above corresponds to the "Jacobi" version of the Schwarz method used in [2], but applied to a chain of atoms defined on rectangular domains.

In order to analyze the convergence of the parallel Schwarz method, we denote by e_j^n the error between the exact solution u_j and the approximate solution computed at the iteration n, that is, $e_j^n := u_j - u_j^n$. By linearity, the parallel Schwarz method for the errors is

$$\begin{aligned} &-\Delta e_j^n = 0 \ \text{ in } \Omega_j, \quad e_j^n(\cdot,0) = 0, \quad e_j^n(\cdot,\hat{L}) = 0, \\ &e_j^n(a_j,\cdot) = e_{j-1}^{n-1}(a_j,\cdot), \quad e_j^n(b_j,\cdot) = e_{j+1}^{n-1}(b_j,\cdot) \end{aligned} \tag{4}$$

for $j = 2, \ldots, N-1$ and

$$
(5) \qquad \begin{aligned} -\Delta e_1^n &= 0 \quad \text{in } \Omega_1,\\ e_1^n(\cdot,0) &= 0,\ e_1^n(\cdot,\hat{L}) = 0,\\ e_1^n(a_1,\cdot) &= 0,\\ e_1^n(b_1,\cdot) &= e_2^{n-1}(b_1,\cdot), \end{aligned} \qquad \text{and} \qquad \begin{aligned} -\Delta e_N^n &= 0 \quad \text{in } \Omega_N,\\ e_N^n(\cdot,0) &= 0,\ e_N^n(\cdot,\hat{L}) = 0,\\ e_N^n(a_N,\cdot) &= e_{N-1}^{n-1}(a_N,\cdot),\\ e_N^n(b_N,\cdot) &= 0. \end{aligned}
$$

In what follows, we study the convergence of the parallel Schwarz method, that is, $e_j^n \to 0$ as $n \to \infty$. Our convergence analysis is performed using a Fourier sine series to solve the elliptic problems (4)–(5). This technique allows us to study the convergence of the parallel Schwarz method for each coefficient of the Fourier series. In particular, for each mode we can construct a Schwarz iteration matrix T that is used to generate a sequence of errors in the Fourier coefficients $\boldsymbol{v}^n$ as $\boldsymbol{v}^n = T\boldsymbol{v}^{n-2}$. Notice that the step from $n-2$ to n is considered in order to analyze the decay of the error corresponding to the jth atom. This fact will become clear in the next section and is common for the analysis of Schwarz-type methods; see, e.g., [7].

3. Constructing the Schwarz iteration matrix. To construct the Schwarz iteration matrix corresponding to the parallel Schwarz method (4)–(5), we use a sine series expansion,

$$
(6) \qquad e_j^n(x,y) = \sum_{m=1}^{\infty} v_j^n(x,k)\sin(ky), \quad k = \frac{\pi m}{\hat{L}},
$$

where the Fourier coefficients $v_j^n(x,k)$ are given by

$$
(7) \qquad v_j^n(x,k) = \tilde{c}_j(k,\delta)e^{kx} + \tilde{d}_j(k,\delta)e^{-kx},
$$

and $\tilde{c}_j(k,\delta)$ and $\tilde{d}_j(k,\delta)$ are computed using the conditions $v_j^n(a_j,k) = v_{j-1}^{n-1}(a_j,k)$ and $v_j^n(b_j,k) = v_{j+1}^{n-1}(b_j,k)$, which are obtained by transforming the transmission conditions. Notice that k is parametrized by m; hence one should formally write k_m, but we drop the subscript m for simplicity. We obtain that

$$
(8) \qquad \begin{aligned} v_j^n(x,k) = e^{kx}e^{-jkL}&\Big[g_{A1}(k,\delta)v_{j+1}^{n-1}(b_j,k) - g_{A2}(k,\delta)v_{j-1}^{n-1}(a_j,k)\Big]\\ &+ e^{-kx}e^{jkL}\Big[g_{B1}(k,\delta)v_{j-1}^{n-1}(a_j,k) - g_{B2}(k,\delta)v_{j+1}^{n-1}(b_j,k)\Big], \end{aligned}
$$

with

$$
(9) \qquad \begin{aligned} g_{A1}(k,\delta) &:= \frac{e^{3k\delta+2kL}}{e^{4k\delta+2kL}-1}, \quad g_{A2}(k,\delta) := \frac{e^{k\delta+kL}}{e^{4k\delta+2kL}-1},\\ g_{B1}(k,\delta) &:= \frac{e^{3k\delta+kL}}{e^{4k\delta+2kL}-1}, \quad g_{B2}(k,\delta) := \frac{e^{k\delta}}{e^{4k\delta+2kL}-1}. \end{aligned}
$$

In a similar fashion we solve the problems (5) on the left and right and obtain

$$
v_1^n(x,k) = w_1(x,k;\delta)v_2^{n-1}(b_1,k),
$$

with

$$
w_1(x,k;\delta) := \frac{e^{k\delta+kL}}{1-e^{4k\delta+2kL}}\Big[e^{-kx} - e^{2k\delta+kx}\Big],
$$

and

$$
v_N^n(x,k) = z_N(x,k;\delta)v_{N-1}^{n-1}(a_N,k),
$$

with

$$z_N(x,k;\delta) := \frac{e^{k\delta+kL}}{1-e^{4k\delta+2kL}}\Big[e^{kx-kNL} - e^{kNL+2k\delta-kx}\Big].$$

Our convergence analysis focuses on the Fourier coefficients $v_j^n(x,k)$, and we thus rewrite (8) in the form

$$v_j^n(x,k) = w_j(x,k;\delta)v_{j+1}^{n-1}(b_j,k) + z_j(x,k;\delta)v_{j-1}^{n-1}(a_j,k), \tag{10}$$

where

$$w_j(x,k;\delta) := e^{kx}e^{-jkL}g_{A1}(k,\delta) - e^{-kx}e^{jkL}g_{B2}(k,\delta) \tag{11}$$

and

$$z_j(x,k;\delta) := e^{-kx}e^{jkL}g_{B1}(k,\delta) - e^{kx}e^{-jkL}g_{A2}(k,\delta). \tag{12}$$

By applying (10) recursively, we obtain

$$\begin{aligned} v_j^n(x,k) = {} & w_j(x,k;\delta)w_{j+1}(b_j,k;\delta)v_{j+2}^{n-2}(b_{j+1},k) \\ & + w_j(x,k;\delta)z_{j+1}(b_j,k;\delta)v_j^{n-2}(a_{j+1},k) \\ & + z_j(x,k;\delta)w_{j-1}(a_j,k;\delta)v_j^{n-2}(b_{j-1},k) \\ & + z_j(x,k;\delta)z_{j-1}(a_j,k;\delta)v_{j-2}^{n-2}(a_{j-1},k). \end{aligned} \tag{13}$$

Evaluating (13) at $x = a_j$ and $x = b_j$ yields

$$\begin{aligned} v_{j-1}^n(a_j,k) = {} & w_a(k,\delta)w_b(k,\delta)v_{j+1}^{n-2}(b_j,k) + w_a(k,\delta)z_b(k,\delta)v_{j-1}^{n-2}(a_j,k) \\ & + z_a(k,\delta)w_a(k,\delta)v_{j-1}^{n-2}(b_{j-2},k) + z_a(k,\delta)z_a(k,\delta)v_{j-3}^{n-2}(a_{j-2},k), \\ v_{j+1}^n(b_j,k) = {} & w_b(k,\delta)w_b(k,\delta)v_{j+3}^{n-2}(b_{j+2},k) + w_b(k,\delta)z_b(k,\delta)v_{j+1}^{n-2}(a_{j+2},k) \\ & + z_b(k,\delta)w_a(k,\delta)v_{j+1}^{n-2}(b_j,k) + z_b(k,\delta)z_a(k,\delta)v_{j-1}^{n-2}(a_j,k), \end{aligned} \tag{14}$$

where we used the fact that

$$\begin{aligned} & w_a(k,\delta) := w_j(a_{j+1},k;\delta) = w_{j-1}(a_j,k;\delta),\ w_b(k,\delta) := w_{j+1}(b_j,k;\delta) = w_j(b_{j-1},k;\delta), \\ & z_a(k,\delta) := z_j(a_{j+1},k;\delta) = z_{j-1}(a_j,k;\delta),\ z_b(k,\delta) := z_{j+1}(b_j,k;\delta) = z_j(b_{j-1},k;\delta) \end{aligned} \tag{15}$$

for $j = 2,\dots,N-1$. Similarly, since $w_1(a_2,k;\delta) = w_a(k,\delta)$ and $z_N(b_{N-1},k;\delta) = z_b(k,\delta)$, for the first and the last atom we get

$$\begin{aligned} v_1^n(a_2,k) &= w_a(k,\delta)w_b(k,\delta)v_3^{n-2}(b_2,k) + w_a(k,\delta)z_b(k,\delta)v_1^{n-2}(a_2,k), \\ v_N^n(b_{N-1},k) &= z_b(k,\delta)w_a(k,\delta)v_N^{n-2}(b_{N-1},k) + z_b(k,\delta)z_a(k,\delta)v_{N-2}^{n-2}(a_{N-1},k). \end{aligned} \tag{16}$$

For the atoms $j = 2$ and $j = N-1$ we have

$$\begin{aligned} v_2^n(a_3,k) = {} & w_a(k,\delta)z_a(k,\delta)v_2^{n-2}(b_1,k) \\ & + w_a(k,\delta)z_b(k,\delta)v_2^{n-2}(a_3,k) + w_a(k,\delta)w_b(k,\delta)v_4^{n-2}(b_3,k), \\ v_2^n(b_1,k) = {} & w_a(k,\delta)z_b(k,\delta)v_2^{n-2}(b_1,k) \\ & + w_b(k,\delta)z_b(k,\delta)v_2^{n-2}(a_3,k) + w_b(k,\delta)w_b(k,\delta)v_4^{n-2}(b_3,k) \end{aligned}$$

and

$$\begin{aligned} v_{N-1}^n(a_N,k) &= w_a(k,\delta)z_b(k,\delta)v_{N-1}^{n-2}(a_N,k) \\ &\qquad + z_a(k,\delta)w_a(k,\delta)v_{N-1}^{n-2}(b_{N-2},k) + z_a(k,\delta)z_a(k,\delta)v_{N-3}^{n-2}(a_{N-2},k), \\ v_{N-1}^n(b_{N-2},k) &= w_b(k,\delta)z_b(k,\delta)v_{N-1}^{n-2}(a_N,k) \\ &\qquad + z_b(k,\delta)w_a(k,\delta)v_{N-1}^{n-2}(b_{N-2},k) + z_b(k,\delta)z_a(k,\delta)v_{N-3}^{n-2}(a_{N-2},k). \end{aligned}$$

Before we present the Schwarz iteration matrix, we prove the following lemma, which is useful for obtaining a simpler representation of the matrix.

LEMMA 1. *For any* $(k,\delta) \in (0,\infty) \times [0,L]$, *the quantities defined in* (15) *satisfy* $w_a(k,\delta) \geq 0$ *and* $w_b(k,\delta) \geq 0$. *Moreover,* $w_a(k,\delta) = z_b(k,\delta)$ *and* $w_b(k,\delta) = z_a(k,\delta)$.

Proof. A direct calculation using (11) and (9) shows that

$$w_a(k,\delta) = \frac{e^{2k\delta+2kL} - e^{2k\delta}}{e^{4k\delta+2kL} - 1} \quad \text{and} \quad w_b(k,\delta) = \frac{e^{4k\delta+kL} - e^{kL}}{e^{4k\delta+2kL} - 1}. \tag{17}$$

The first statement follows now from (17). Then, a direct calculation involving (12) and (9) allows us to compute that $z_b(k,\delta) = \frac{e^{2k\delta+2kL}-e^{2k\delta}}{e^{4k\delta+2kL}-1}$ and $z_a(k,\delta) = \frac{e^{4k\delta+kL}-e^{kL}}{e^{4k\delta+2kL}-1}$. Comparing these with (17), the second statement follows. ∎

We are now ready to complete the construction of the Schwarz iteration matrix. By defining $\boldsymbol{v}^n(k) \in \mathbb{R}^{2N}$ as

$$\begin{aligned} \boldsymbol{v}^n(k) := \big(&0, v_2^n(b_1,k), v_1^n(a_2,k), v_3^n(b_2,k), \ldots, v_{j-1}^n(a_j,k), v_{j+1}^n(b_j,k), \ldots, \\ &v_{N-2}^n(a_{N-1},k), v_N^n(b_{N-1},k), v_{N-1}^n(a_N,k), 0\big)^\top, \end{aligned} \tag{18}$$

(14) and (16) can be written in the form

$$\boldsymbol{v}^n(k) = T(k,\delta)\boldsymbol{v}^{n-2}(k),$$

where (using Lemma 1) the Schwarz iteration matrix $T(k,\delta) \in \mathbb{R}^{2N\times 2N}$ is given by

$$T(k,\delta) = \left(\begin{array}{cc|cc|cc|cc|cc|cc|cc}
0 & 0 & & & 0 & 0 & & & & & & & & \\
0 & z_bz_b & & & w_bz_b & w_bw_b & & & & & & & & \\ \hline
 & & z_bz_b & z_bw_b & & & 0 & 0 & & & & & & \\
 & & w_bz_b & z_bz_b & & & w_bz_b & w_bw_b & & & & & & \\ \hline
0 & w_bz_b & & & z_bz_b & w_bz_b & & & \ddots & \ddots & & & & \\
0 & 0 & & & w_bz_b & z_bz_b & & & \ddots & \ddots & & & & \\ \hline
 & & w_bw_b & w_bz_b & & & \ddots & \ddots & & & 0 & 0 & & \\
 & & 0 & 0 & & & \ddots & \ddots & & & w_bz_b & w_bw_b & & \\ \hline
 & & & & \ddots & \ddots & & & z_bz_b & w_bz_b & & & 0 & 0 \\
 & & & & \ddots & \ddots & & & w_bz_b & z_bz_b & & & w_bz_b & 0 \\ \hline
 & & & & & & w_bw_b & w_bz_b & & & z_bz_b & w_bz_b & & \\
 & & & & & & 0 & 0 & & & z_bw_b & z_bz_b & & \\ \hline
 & & & & & & & & w_bw_b & w_bz_b & & & z_bz_b & 0 \\
 & & & & & & & & 0 & 0 & & & 0 & 0
\end{array}\right)$$

and we omitted the dependence on k and δ for simplicity. Notice that we added a first and a last zero entry in the vector $\boldsymbol{v}^n(k)$ leading to the first and last zero rows and columns to have the Schwarz iteration matrix $T(k,\delta)$, because this reveals a block structure corresponding to the atoms in the chain that will be useful later.

4. Convergence analysis of the parallel Schwarz method.

We now prove that the parallel Schwarz method (4)–(5) converges independently of the number of atoms N. We start by proving essential properties of the Schwarz iteration matrix $T(k,\delta)$, using that $k>0$, which always holds because of (6), i.e., the Dirichlet boundary conditions on the boundary of the atoms.

LEMMA 2. *The following statements hold:*

(a) *For any $\delta>0$, the map $k\in(0,\infty)\mapsto(z_b+w_b)(k,\delta)\in\mathbb{R}$ is strictly monotonically decreasing.*

(b) *For any $k>0$, the map $\delta\in(0,L/2)\mapsto(z_b+w_b)(k,\delta)\in\mathbb{R}$ is strictly monotonically decreasing.*

(c) *For any $k>0$, we have $(z_b+w_b)(k,0)=1$, $\frac{\partial(z_b+w_b)}{\partial k}(k,0)=0$, and $\frac{\partial(z_b+w_b)}{\partial \delta}(k,0)<0$.*

Proof. From Lemma 1 and using (17), we obtain

$$\begin{aligned}(z_b+w_b)(k,\delta)&=\frac{e^{2k\delta+2kL}-e^{2k\delta}+e^{4k\delta+kL}-e^{kL}}{e^{4k\delta+2kL}-1}\\&=\frac{(e^{2k\delta+kL}-1)e^{2k\delta}+(e^{2k\delta+kL}-1)e^{kL}}{(e^{2k\delta+kL}-1)(e^{2k\delta+kL}+1)}=\frac{e^{2k\delta}+e^{kL}}{e^{2k\delta+kL}+1}.\end{aligned}\tag{19}$$

Differentiating (19) with respect to k, we get

$$\begin{aligned}\frac{\partial(z_b+w_b)}{\partial k}(k,\delta)&=\frac{Le^{kL}+2\delta e^{2k\delta}}{e^{2k\delta+kL}+1}-\frac{(e^{kL}+e^{2k\delta})(L+2\delta)e^{kL+2k\delta}}{(e^{2k\delta+kL}+1)^2}\\&=-\frac{\left[Le^{4k\delta+kL}+2\delta e^{2k\delta+2kL}-2\delta e^{2k\delta}-Le^{kL}\right]}{(e^{2k\delta+kL}+1)^2}.\end{aligned}\tag{20}$$

Notice that $e^{4k\delta+kL}-e^{kL}>0$ and $e^{2k\delta+2kL}-e^{2k\delta}>0$ for any $k>0$ and $\delta>0$. Hence, we have that $\frac{\partial(z_b+w_b)}{\partial k}(k,\delta)<0$, and statement (a) follows.

Claim (b) can be proved in a similar way: we have that

$$\frac{\partial(z_b+w_b)}{\partial \delta}(k,\delta)=\frac{2ke^{2k\delta}}{e^{2k\delta+kL}+1}-\frac{2k(e^{kL}+e^{2k\delta})e^{kL+2k\delta}}{(e^{2k\delta+kL}+1)^2}=-\frac{2k\left[e^{2k\delta+2kL}-e^{2k\delta}\right]}{(e^{2k\delta+kL}+1)^2}.\tag{21}$$

Notice that $e^{2k\delta+2kL}-e^{2k\delta}>0$ for any $k>0$ and $\delta\in(0,L/2)$, which implies that $\frac{\partial(z_b+w_b)}{\partial \delta}(k,\delta)<0$, and statement (b) follows. Claim (c) follows by continuity of (19)–(21) with respect to δ and passing to the limit for $\delta\to0$. ◻

We can now prove that the parallel Schwarz method for the solution of (4)–(5) converges independently of the number of atoms N by proving that the spectral radius of the Schwarz iteration matrix $T(k,\delta)$ is bounded by a function of k and δ that does not depend on N.

THEOREM 3. *For any $(k,\delta)\in(0,\infty)\times(0,L)$ we have the bound*

$$\rho(T(k,\delta))\le\|T(k,\delta)\|_\infty\le\lambda(k,\delta)<1,$$

where $\rho(T(k,\delta))$ is the spectral radius of $T(k,\delta)$ and

$$\lambda(k,\delta):=\left(\frac{e^{2k\delta}+e^{kL}}{e^{2k\delta+kL}+1}\right)^2,$$

which is independent of the number N of atoms. Moreover, for $N \geq 5$ it holds that $\|T(k,\delta)\|_\infty = \lambda(k,\delta)$.

Proof. By Lemma 1, all the entries of $T(k,\delta)$ are positive. The sum of the entries of the jth row for $5 \leq j \leq 2N-5$ is

$$z_b w_b + z_b z_b + z_b w_b + w_b w_b = \big(z_b + w_b\big)^2,$$

where we omitted the dependence on k and δ for simplicity. Moreover, it follows from Lemma 1 that

$$\sum_{\ell} \big(T(k,\delta)\big)_{j,\ell} \leq \Big(z_b(k,\delta) + w_b(k,\delta)\Big)^2 \tag{22}$$

for $j = 1,\dots,2N$. Hence we obtain for the infinity norm of the Schwarz iteration matrix

$$\|T(k,\delta)\|_\infty \leq \big(z_b(k,\delta) + w_b(k,\delta)\big)^2, \tag{23}$$

and equality holds if $N \geq 5$. Now using (19) and Lemma 2, we get

$$z_b(k,\delta) + w_b(k,\delta) = \frac{e^{2k\delta} + e^{kL}}{e^{2k\delta+kL} + 1} < 1, \tag{24}$$

and combining (23) with (24) concludes the proof. ∎

We now use Theorem 3 to obtain convergence in the L^2 norm using Parseval's identity.

Corollary 4. *Under the assumptions of Theorem* 3 *we have that*

$$\|T(k,\delta)\|_2 \leq \|T(k,\delta)\|_\infty. \tag{25}$$

In addition, with $c := \max_k \|T(k,\delta)\|_2^2$, the following inequality holds:

$$\sum_j \Big(\|e_j^n(a_j,\cdot)\|_{L^2}^2 + \|e_j^n(b_j,\cdot)\|_{L^2}^2\Big) \leq c \sum_j \Big(\|e_j^{n-2}(a_j,\cdot)\|_{L^2}^2 + \|e_j^{n-2}(b_j,\cdot)\|_{L^2}^2\Big).$$

Proof. Since the Schwarz iteration matrix satisfies $\|T(k,\delta)\|_\infty = \|T(k,\delta)\|_1$, (25) follows by applying the standard estimate $\|T(k,\delta)\|_2 \leq \sqrt{\|T(k,\delta)\|_\infty \|T(k,\delta)\|_1}$. The second statement can be shown using Parseval's identity. ∎

Next, our aim is to obtain a sharper estimate of the spectral radius. According to the estimate (22), the effect of the inner atoms dominates the effect of the first and the last atom of the chain.

We first notice that $T(k,\delta)$ can be written as a sum,

$$T(k,\delta) = T_1(k,\delta) + T_2(k,\delta),$$

where $T_1(k,\delta)$ is a symmetric matrix given by

$$
T_1 := \left(\begin{array}{cc|cc|cc|cc|cc|cc}
0 & 0 & & & 0 & 0 & & & & & & \\
0 & z_b^2 & & & z_b w_b & 0 & & & & & & \\ \hline
& & z_b^2 & w_b z_b & & & 0 & 0 & & & & \\
& & w_b z_b & z_b^2 & & & z_b w_b & 0 & & & & \\ \hline
0 & z_b w_b & & & z_b^2 & w_b z_b & & & \ddots & \ddots & & \\
0 & 0 & & & w_b z_b & z_b^2 & & & \ddots & \ddots & & \\ \hline
& & 0 & z_b w_b & & & \ddots & \ddots & & & 0 & 0 \\
& & 0 & 0 & & & \ddots & \ddots & & & z_b w_b & 0 \\ \hline
& & & & \ddots & \ddots & & & z_b^2 & w_b z_b & & \\
& & & & \ddots & \ddots & & & w_b z_b & z_b^2 & & \\ \hline
& & & & & & 0 & z_b w_b & & & z_b^2 & 0 \\
& & & & & & 0 & 0 & & & 0 & 0
\end{array}\right)
$$

and $T_2(k, \delta)$ is given by

$$
T_2 := \left(\begin{array}{cc|cc|cc|cc|cc|cc|cc}
0 & 0 & & & 0 & 0 & & & & & & & & \\
0 & 0 & & & 0 & w_b^2 & & & & & & & & \\ \hline
& & 0 & 0 & & & 0 & 0 & & & & & & \\
& & 0 & 0 & & & 0 & w_b^2 & & & & & & \\ \hline
0 & 0 & & & 0 & 0 & & & \ddots & \ddots & & & & \\
0 & 0 & & & 0 & 0 & & & \ddots & \ddots & & & & \\ \hline
& & w_b^2 & 0 & & & \ddots & \ddots & & & 0 & 0 & & \\
& & 0 & 0 & & & \ddots & \ddots & & & 0 & w_b^2 & & \\ \hline
& & & & \ddots & \ddots & & & 0 & 0 & & & 0 & 0 \\
& & & & \ddots & \ddots & & & 0 & 0 & & & 0 & 0 \\ \hline
& & & & & & w_b^2 & 0 & & & 0 & 0 & & \\
& & & & & & 0 & 0 & & & 0 & 0 & & \\ \hline
& & & & & & & & & & & & 0 & 0 \\
& & & & & & & & & & & & 0 & 0
\end{array}\right).
$$

Under the action of an appropriate permutation matrix P, the matrix $T_1(k, \delta)$ can be transformed into $T_1'(k, \delta) = PT_1(k, \delta)P^\top$ with block-diagonal structure,

$$
T_1'(k, \delta) = \begin{pmatrix} T_{11}'(k, \delta) & 0 \\ 0 & T_{22}'(k, \delta) \end{pmatrix},
$$

where $T_{11}'(k, \delta)$ and $T_{22}'(k, \delta)$ are tridiagonal matrices of the form

$$
\begin{pmatrix}
0 & & & & & & \\
& z_b^2 & w_b z_b & & & & \\
& w_b z_b & z_b^2 & w_b z_b & & & \\
& & \ddots & \ddots & \ddots & & \\
& & & w_b z_b & z_b^2 & w_b z_b & \\
& & & & z_b w_b & z_b^2 & \\
& & & & & & 0
\end{pmatrix}
\quad \text{and} \quad
\begin{pmatrix}
z_b^2 & w_b z_b & & & \\
w_b z_b & z_b^2 & w_b z_b & & \\
& \ddots & \ddots & \ddots & \\
& & w_b z_b & z_b^2 & w_b z_b \\
& & & z_b w_b & z_b^2
\end{pmatrix}.
$$

The block structure of T_1 corresponds to the specific ordering of the vector $\boldsymbol{v}^n$ defined in (18). In particular, $\boldsymbol{v}^n = (\boldsymbol{v}_1^n, \boldsymbol{v}_2^n, \dots, \boldsymbol{v}_N^n)^\top$, where $\boldsymbol{v}_j^n = (v_{j-1}^n(a_j,k), v_{j+1}^n(b_j,k))$, with $\boldsymbol{v}_1^n = (0, v_2^n(b_1,k))$ and $\boldsymbol{v}_N^n = (v_{N-1}^n(a_N,k), 0)$. In order to obtain this particular block-tridiagonal structure, the permutation matrix P is

$$P\boldsymbol{v}^n = \big(\boldsymbol{v}_1^n, \boldsymbol{v}_3^n, \boldsymbol{v}_5^n, \dots, \boldsymbol{v}_2^n, \boldsymbol{v}_4^n, \boldsymbol{v}_6^n, \dots\big)^\top.$$

This transformation allows us to identify the two Schwarz subsequences $\big\{\boldsymbol{v}_1^n, \boldsymbol{v}_3^n, \boldsymbol{v}_5^n, \dots\big\}$ and $\big\{\boldsymbol{v}_2^n, \boldsymbol{v}_4^n, \boldsymbol{v}_6^n, \dots\big\}$ as pointed out in [19]. These subsequences correspond to a well-known red-black ordering of the subdomains; see, for example, [25, page 18].

Since $T_2(k,\delta)^\top T_2(k,\delta)$ is a diagonal matrix having entries equal to w_b^4 and zero, and denoting by μ the maximum between the dimension of the nonzero block of T'_{11} and the dimension of T'_{22}, we have

$$\|T_1(k,\delta)\|_2 = \max_{1\le j\le \mu}\left\{\left| z_b^2 + 2w_b z_b \cos\left(\frac{j\pi}{\mu+1}\right)\right|\right\}, \quad \|T_2(k,\delta)\|_2 = w_b^2. \tag{26}$$

THEOREM 5. *Let $N \ge 4$. Then, for any $k > 0$ and $\delta > 0$, the estimate*

$$\rho(T(k,\delta)) \le \gamma(k,\delta,\mu) \le \|T(k,\delta)\|_\infty < 1$$

holds with

$$\gamma(k,\delta,\mu) := \max_{1\le j\le \mu}\left\{\left| z_b^2(k,\delta) + 2w_b(k,\delta) z_b(k,\delta) \cos\left(\frac{j\pi}{\mu+1}\right)\right|\right\} + w_b^2(k,\delta), \tag{27}$$

where μ is the maximum between the dimension of the nonzero block of T'_{11} and the dimension of T'_{22}.

Proof. Using the triangle inequality and (26), we get

$$\|T(k,\delta)\|_2 \le \|T_1(k,\delta)\|_2 + \|T_2(k,\delta)\|_2 = \gamma(k,\delta,\mu).$$

The claim follows by noticing that

$$\gamma(k,\delta,\mu) \le z_b^2 + 2w_b z_b + w_b^2 = \big(z_b + w_b\big)^2 = \|T(k,\delta)\|_\infty$$

and using Theorem 3. ∎

Using Theorems 3 and 5 and Corollary 4, we thus obtain the estimate

$$\rho(T) \le \|T\|_2 \le \gamma(k,\delta,\mu) \le \|T\|_\infty \le \lambda(\delta,k) < 1. \tag{28}$$

Example 1. Consider a chain of $N = 5$ atoms. The corresponding matrix $T(k,\delta) \in \mathbb{R}^{10\times 10}$. Denoting by $P_{j,k} \in \mathbb{R}^{5\times 5}$ the matrix performing the permutation of the jth row with the kth row, and letting $P := (P_{3,4}P_{4,5}P_{2,3}) \otimes I_2$, where $\otimes$ denotes the Kronecker product and I_2 is the 2×2 identity, we get

$$P\big(\boldsymbol{v}_1^n, \boldsymbol{v}_2^n, \boldsymbol{v}_3^n, \boldsymbol{v}_4^n, \boldsymbol{v}_5^n\big)^\top = \big(\boldsymbol{v}_1^n, \boldsymbol{v}_3^n, \boldsymbol{v}_5^n, \boldsymbol{v}_2^n, \boldsymbol{v}_4^n\big)^\top,$$

and the matrix $T'(k,\delta)$ is obtained by $T'(k,\delta) = PT(k,\delta)P^\top$. An intuitive and pictorial representation of the decomposition of T_1 can be obtained with the help of graph theory [13, 5]. In particular, we can associate to T_1 an adjacency matrix A_d defined by

$$(A_d)_{jk} := \begin{cases} 1 & \text{if} \quad (T_1)_{jk} \neq 0 \quad \text{and} \quad j \neq k, \\ 0 & \text{otherwise.} \end{cases}$$

A well-known result says that the graph corresponding to an adjacency matrix $A_d \in \mathbb{R}^{m\times m}$ is disconnected if and only if the matrix $Y := A_d + A_d^2 + \cdots + A_d^{m-1}$ has at least one zero entry; see, e.g., [5, Corollary B, page 161]. A simple calculation shows that in our example the matrix Y has zero entries, which means that the corresponding graph is not connected. Moreover, as shown in Figure 2, a simple plot shows that the graph corresponding to A_d has only two connected components. These components correspond to T'_{11} and T'_{22}, respectively.

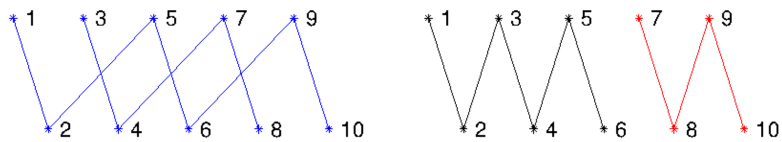

FIG. 2. *Graphs corresponding to the adjacency matrices obtained from T_1 (left) and from $T'_1 = PT_1P^\top$ (right). A comparison between the two pictures shows that the action of the matrix P corresponds only to a relabeling of the nodes revealing the red-black ordering (color available online).*

5. Optimality of the Schwarz iteration matrix for error estimation.

The Schwarz iteration matrix T presented in section 3 and also used, for example, in [10, 9] is not the only possible one. It is often easier to construct the Schwarz iteration matrix that acts on the constants of the formal solution of the differential equation considered; see, for example, [24, 6]. Doing this in our case, we obtain from (7)–(8) that a formal solution to (4) can be written as

$$v_j^n(x,k) = e^{k(x-jL)} c_j^n(k,\delta) + e^{-k(x-jL)} d_j^n(k,\delta). \tag{29}$$

By using the boundary conditions $v_j^n(a_j,k) = v_{j-1}^{n-1}(a_j,k)$ and $v_j^n(b_j,k) = v_{j+1}^{n-1}(b_j,k)$, we obtain the system of equations

$$D \begin{pmatrix} c_j^n(k,\delta) \\ d_j^n(k,\delta) \end{pmatrix} = L \begin{pmatrix} c_{j-1}^{n-1}(k,\delta) \\ d_{j-1}^{n-1}(k,\delta) \end{pmatrix} + R \begin{pmatrix} c_{j+1}^{n-1}(k,\delta) \\ d_{j+1}^{n-1}(k,\delta) \end{pmatrix}, \tag{30}$$

where

$$D := \begin{pmatrix} e^{-k(L+\delta)} & e^{k(L+\delta)} \\ e^{k\delta} & e^{-k\delta} \end{pmatrix}, \quad L := \begin{pmatrix} e^{-k\delta} & e^{k\delta} \\ 0 & 0 \end{pmatrix}, \quad R := \begin{pmatrix} 0 & 0 \\ e^{k(\delta-L)} & e^{-k(\delta-L)} \end{pmatrix}.$$

Equation (30) for all the subdomains reads

$$\begin{pmatrix} \ddots & & & \\ & D & & \\ & & D & \\ & & & \ddots \end{pmatrix} \boldsymbol{w}^n = \begin{pmatrix} \ddots & & & & \\ L & 0 & R & & \\ & L & 0 & R & \\ & & & & \ddots \end{pmatrix} \boldsymbol{w}^{n-1}, \tag{31}$$

where $\boldsymbol{w}^n := (\ldots, c_{j-1}^n, d_{j-1}^n, c_j^n, d_j^n, c_{j+1}^n, d_{j+1}^n, \ldots)^\top$. By defining the matrices

$$\tilde{T}_1 := D^{-1}L, \quad \tilde{T}_2 := D^{-1}R, \quad \tilde{T}_3 := \tilde{T}_1\tilde{T}_2 + \tilde{T}_2\tilde{T}_1, \tag{32}$$

we obtain the new Schwarz iteration matrix

$$\tilde{T} := \begin{pmatrix} & & \ddots & & & & \\ \tilde{T}_1^2 & 0 & \tilde{T}_3 & 0 & \tilde{T}_2^2 & & \\ & \tilde{T}_1^2 & 0 & \tilde{T}_3 & 0 & \tilde{T}_2^2 & \\ & & & & \ddots & & \end{pmatrix}, \tag{33}$$

and (31) can be written in compact form as

$$\boldsymbol{w}^n = \tilde{T}\boldsymbol{w}^{n-2}. \tag{34}$$

As in section 4, we can obtain from (33) the infinity norm

$$\|\tilde{T}(k,\delta)\|_\infty = \frac{\left(2e^{6\delta k} + e^{4\delta k} - 2e^{2\delta k}\right)e^{4kL} + \left(e^{8\delta k} - 2e^{4\delta k} - 1\right)e^{2kL} + e^{4\delta k}}{(e^{2kL+4\delta k} - 1)^2}.$$

For $\delta = 0$ the previous norm becomes $\|\tilde{T}(k,0)\|_\infty = 1$. Moreover, by fixing $L = 1$, $\hat{L} = 1$, $m = 1$ (hence $k = \pi$), and $\delta = 0.05$, we obtain that $\|\tilde{T}\|_\infty \approx 1.2$, which shows that the infinity norm of $\tilde{T}$ is not always bounded by 1. This shows that different Schwarz iteration matrices can have different norms, and only our first Schwarz iteration matrix allowed us to get the convergence estimate. Notice also that, as in the case of the matrix T (see Theorem 3), the norm $\|\tilde{T}\|_\infty$ is not affected by the components of the first and last subdomains for $N \geq 5$. For this reason, in what follows we exclude in our analysis the atoms $j = 1$ and $j = N$, and we remove from the Schwarz iteration matrices the corresponding rows and columns. Hence, we work with the Schwarz iteration matrices T_N and $\tilde{T}_N$ in $\mathbb{R}^{2(N-2)\times 2(N-2)}$ given by

$$T_N = \left(\begin{array}{cc|cc|cc|cc|cc} z_b^2 & w_b z_b & & & 0 & 0 & & & & \\ w_b z_b & z_b^2 & & & z_b w_b & w_b^2 & & & & \\ \hline & & z_b^2 & w_b z_b & & & \ddots & \ddots & & \\ & & w_b z_b & z_b^2 & & & \ddots & \ddots & & \\ \hline w_b^2 & z_b w_b & & & \ddots & \ddots & & & 0 & 0 \\ 0 & 0 & & & \ddots & \ddots & & & z_b w_b & w_b^2 \\ \hline & & \ddots & \ddots & & & z_b^2 & w_b z_b & & \\ & & \ddots & \ddots & & & w_b z_b & z_b^2 & & \\ \hline & & & & w_b^2 & z_b w_b & & & z_b^2 & w_b z_b \\ & & & & 0 & 0 & & & w_b z_b & z_b^2 \end{array}\right) \tag{35}$$

and

$$\tilde{T}_N = \begin{pmatrix} \tilde{T}_3 & 0 & \tilde{T}_2^2 & & & & \\ 0 & \tilde{T}_3 & 0 & \tilde{T}_2^2 & & & \\ \tilde{T}_1^2 & 0 & \tilde{T}_3 & 0 & \tilde{T}_2^2 & & \\ & & & \ddots & & & \\ & & & & \tilde{T}_1^2 & 0 & \tilde{T}_3 \end{pmatrix}.$$

We now construct a matrix G that allows us to transform $\tilde{T}_N$ into the Schwarz iteration matrix T_N. To do so we evaluate (29) at $x = a_j$ and $x = b_j$ and obtain

$$\begin{pmatrix} v_j^n(a_j,k) \\ v_j^n(b_j,k) \end{pmatrix} = D \begin{pmatrix} c_j^n(k,\delta) \\ d_j^n(k,\delta) \end{pmatrix}.$$

Similarly, we evaluate (8) at $x = a_j$ and $x = b_j$ and get

$$\begin{pmatrix} v_j^n(a_j,k) \\ v_j^n(b_j,k) \end{pmatrix} = D \begin{pmatrix} -g_{A2}(k,\delta) & g_{A1}(k,\delta) \\ g_{B1}(k,\delta) & -g_{B2}(k,\delta) \end{pmatrix} \begin{pmatrix} v_{j-1}^{n-1}(a_j,k) \\ v_{j+1}^{n-1}(b_j,k) \end{pmatrix}.$$

By combining the two previous inequalities and using the fact that D is invertible for $k > 0$, we obtain that

$$\begin{pmatrix} c_j^n(k,\delta) \\ d_j^n(k,\delta) \end{pmatrix} = \begin{pmatrix} -g_{A2}(k,\delta) & g_{A1}(k,\delta) \\ g_{B1}(k,\delta) & -g_{B2}(k,\delta) \end{pmatrix} \begin{pmatrix} v_{j-1}^{n-1}(a_j,k) \\ v_{j+1}^{n-1}(b_j,k) \end{pmatrix}.$$

Defining

$$G := \begin{pmatrix} \ddots & & \\ & G_2 & \\ & & \ddots \end{pmatrix}, \quad G_2 := \begin{pmatrix} -g_{A2}(k,\delta) & g_{A1}(k,\delta) \\ g_{B1}(k,\delta) & -g_{B2}(k,\delta) \end{pmatrix},$$

it follows from (34) that $G\boldsymbol{v}^n(k) = \tilde{T}_N G \boldsymbol{v}^{n-2}(k)$, which implies

$$\boldsymbol{v}^n(k) = G^{-1}\tilde{T}_N G \boldsymbol{v}^{n-2}(k), \tag{36}$$

and from (34) and (36) we obtain

$$T_N = G^{-1}\tilde{T}_N G. \tag{37}$$

Since (37) implies that $\rho(T_N) = \rho(\tilde{T}_N)$, the convergence of the parallel Schwarz method does not depend on the choice of T_N or $\tilde{T}_N$, but because T_N and $\tilde{T}_N$ have very different norms, the convergence analysis of the Schwarz method is strongly affected by the choice of the Schwarz iteration matrix.

This implies that a transform map of the type $\tilde{T}_N \mapsto G^{-1}\tilde{T}_N G$ is useful in generalizing the results presented in section 4, where we proved convergence of the sequence (31) defined on the interfaces, that is, for $v_j^n(a_j,k)$ and $v_j^n(b_j,k)$. Using (10), we can construct an invertible matrix G such that the matrix $\hat{T}_N := G^{-1}\tilde{T}_N G$ describes the convergence of the parallel Schwarz method in two arbitrary distinct points belonging to (a_j, b_{j-1}). In particular, by noticing that $\rho(\hat{T}_N) = \rho(\tilde{T}_N) = \rho(T_N)$, we observe that the convergence behavior is the same for any point x in (a_j, b_{j-1}).

Now since the convergence of Schwarz methods is often done by norm estimates, and we have seen that different formulations give different norm estimates, then which is the best Schwarz iteration matrix to get the sharpest possible error estimate?

5.1. Determining Schwarz iteration matrices for optimal error estimation. We introduce the following new concept of optimality of Schwarz iteration matrices with respect to error estimation.

DEFINITION 6. *Let $\tilde{T} \in \mathbb{R}^{m\times m}$ be a Schwarz iteration matrix. The Schwarz iteration matrix $T \in \mathbb{R}^{m\times m}$ is said to be optimal in the norm $\|\cdot\|$ if it solves the* *optimization problem*

$$\min_{T\in\mathcal{T}(\tilde{T})} \|T\|, \tag{38}$$

where $\mathcal{T}(\tilde{T})$ denotes the set of all admissible Schwarz iteration matrices,

$$\mathcal{T}(\tilde{T}) := \big\{V \in \mathbb{R}^{m\times m} | \exists G \in \mathbb{R}^{m\times m} \text{ invertible, such that } V = G^{-1}\tilde{T}G\big\}.$$

Definition 7. *A stationary point of* (38) *is called a candidate to be an optimal Schwarz iteration matrix in the norm* $\|\cdot\|$.

Problem (38) can admit more than one solution. In fact, in general the set $\mathcal{T}(\tilde{T})$ is not convex, and the norm $\|\cdot\|$ is not strictly convex. Moreover, the standard techniques used to prove the existence of a solution for such a problem require closedness of the constraint set $\mathcal{T}(\tilde{T})$, and, as shown in the following example, this does not hold in general.

Example 2. Consider the matrix $\tilde{T} = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$. Now define $Q_n := \left(\begin{smallmatrix} n & 0 \\ 0 & 1 \end{smallmatrix}\right)$ having the inverse $Q_n^{-1} = \left(\begin{smallmatrix} 1/n & 0 \\ 0 & 1 \end{smallmatrix}\right)$. Define the sequence $\{T_n\}_n$ as $T_n := Q_n^{-1}\tilde{T}Q_n = \left(\begin{smallmatrix} 0 & 1/n \\ 0 & 0 \end{smallmatrix}\right)$. This sequence converges to the zero matrix, $T_n \to 0$. Since the zero matrix is only similar to itself, $0 \notin \mathcal{T}(\tilde{T})$. This shows that $\mathcal{T}(\tilde{T})$ is not closed in $\mathbb{R}^{2\times 2}$.

Fortunately, the set $\mathcal{T}(\tilde{T})$ has already been studied in the field of control problems on manifolds, where it is usually called the similarity orbit. A short discussion on the closedness of $\mathcal{T}(\tilde{T})$ is provided in [14, Example 2.2, page 167], where it is stated that a similarity orbit is closed if and only if the matrix $\tilde{T}$ is diagonalizable. However, no proof is provided, and it seems difficult to find a reference. Recalling that closedness implies the existence of a minimizer of (38), we now prove the sufficiency of diagonalizability.

Proposition 8. *If $\tilde{T} \in \mathbb{R}^{q\times q}$ is diagonalizable, then the similarity orbit $\mathcal{T}(\tilde{T})$ is closed in $\mathbb{R}^{q\times q}$.*

Proof. For a given sequence $\{T_n\}_n$ in $\mathcal{T}(\tilde{T})$ such that $T_n \to \hat{T} \in \mathbb{R}^{q\times q}$, we have to show that $\hat{T} \in \mathcal{T}(\tilde{T})$. Every element T_n in the sequence is diagonalizable, since $\tilde{T}$ is, and similarity between T_n and $\tilde{T}$ implies that they have the same eigenvalues λ_ℓ with the same geometric multiplicity q_ℓ, and diagonalizability implies that $\sum_\ell q_\ell = q$.

Now continuity of the determinant implies that $\det(T_n - \lambda_\ell I) \to \det(\hat{T} - \lambda_\ell I)$ with $\det(\hat{T} - \lambda_\ell I) = 0$, which shows that $\hat{T}$ has the same eigenvalues as T_n, and hence as $\tilde{T}$. If the eigenvalues of $\tilde{T}$ are distinct, then also those of $\hat{T}$ are distinct, and thus $\hat{T}$ is diagonalizable and the claim easily follows. If the eigenvalues are not distinct, then since T_n is diagonalizable, we have that

$$(T_n - \lambda_\ell I) = (S_n D S_n^{-1} - \lambda_\ell I) = S_n(D - \lambda_\ell I)S_n^{-1},$$

which implies that $\ker(T_n - \lambda_\ell I)$ and $\ker(D - \lambda_\ell I)$ have the same dimension. This means that the geometric multiplicity of each eigenvalue λ_ℓ is constant in the sequence $\{T_n\}_n$, and hence $\text{nullity}(T_n - \lambda_\ell I) = q_\ell$ for any n, where nullity : $\mathbb{R}^{q\times q} \to \mathbb{N}$ maps from the space of matrices to the dimension of the corresponding kernel. Now, we can use the fact that the nullity map is upper-semicontinuous (see [17, Example 2.6.1]) to write

$$q_\ell = \limsup_{n\to\infty} \text{nullity}(T_n - \lambda_\ell I) \leq \text{nullity}(\hat{T} - \lambda_\ell I),$$

and this holds for any eigenvalue λ_ℓ. Hence the condition $\sum_\ell q_\ell = q$ holds also for the limit $\hat{T}$, which means that $\hat{T}$ is diagonalizable, and the proof is complete. ∎

In order to characterize minimizers of (38), it is suitable to consider the first-order optimality system, which is derived in the following theorem.

THEOREM 9. *Let $\tilde{T} \in \mathbb{R}^{m\times m}$ be a given Schwarz iteration matrix. If a Schwarz iteration matrix T is a local minimizer for* (38)*, then there exists a pair of matrices $(G,\Lambda) \in \mathbb{R}^{m\times m}\times\mathbb{R}^{m\times m}$, with G invertible, such that $(T,\tilde{T},G,\Lambda)$ satisfy the first-order optimality system*

$$T = G^{-1}\tilde{T}G, \tag{39a}$$

$$\left[\Lambda, T^\top\right] = 0, \tag{39b}$$

$$\mathrm{trace}(-\Lambda^\top T) = \|T\|, \tag{39c}$$

$$\|\Lambda\|_* \leq 1, \tag{39d}$$

where $\|\cdot\|_*$ is the dual norm of $\|\cdot\|$ and $[\cdot,\cdot]$ is the commutator operator, $[P,Q] := PQ - QP$.

Proof. To derive the optimality system, we use the Frobenius scalar product $\langle\cdot,\cdot\rangle : \mathbb{R}^{m\times m}\times\mathbb{R}^{m\times m} \to \mathbb{R}$ defined as $\langle A,B\rangle := \mathrm{trace}(A^\top B)$, and we rewrite (38) in the equivalent form

$$\min \|M\| \quad \text{s.t.} \quad M - X\tilde{T}Y = 0, \quad XY - I = 0, \tag{40}$$

where $M,X,Y \in \mathbb{R}^{m\times m}$. The Lagrange function corresponding to (40) is

$$\mathcal{L}(M,X,Y,\Lambda,\Upsilon) := \|M\| + \langle\Lambda, M - X\tilde{T}Y\rangle + \langle\Upsilon, XY - I\rangle. \tag{41}$$

Using [15] and, e.g., [3], a necessary optimality condition is

$$0 \in \partial\mathcal{L}(M,X,Y,\Lambda,\Upsilon), \tag{42}$$

where $\partial\mathcal{L}(M,X,Y,\Lambda,\Upsilon)$ denotes the subdifferential of $\mathcal{L}$ at (M,X,Y,Λ,Υ). Notice that the Lagrange function defined in (41) is Lipschitz continuous with respect to M and differentiable with respect to X, Y, Λ, and Υ, since the trace of matrices is differentiable; see, e.g., [23] and references therein.

Denoting $U := \left(\mathbb{R}^{m\times m}\right)^5$ and defining $\tilde{x} := (M,X,Y,\Lambda,\Upsilon) \in U$, we notice that the elements in $\partial\mathcal{L}(\tilde{x})$ are elements of the dual space U^*, i.e., linear functionals acting on U; see [4]. Hence, the condition (42) is equivalent to the existence of an $\tilde{S}(\tilde{x}) \in \partial\mathcal{L}(\tilde{x})$, defined as $\tilde{S}(\tilde{x}) : U \to \mathbb{R}$, such that $\tilde{S}(\tilde{x})(\delta\tilde{x}) = 0$ for all $\delta\tilde{x} \in U$.

Define $h_1(\tilde{x}) := \|M\|$ and $h_2(\tilde{x}) := \langle\Lambda, M - X\tilde{T}Y\rangle + \langle\Upsilon, XY - I\rangle$. Then the Lagrange function can be written as $\mathcal{L}(\tilde{x}) = h_1(\tilde{x}) + h_2(\tilde{x})$. Since h_1 is convex and Lipschitz continuous and h_2 is smooth, we obtain (see [4])

$$\partial\mathcal{L}(\tilde{x}) = \partial h_1(\tilde{x}) + \{h_2'(\tilde{x})\},$$

where $h_2'(\tilde{x})$ is the directional (Gâteaux) derivative of h_2 at $\tilde{x}$. Moreover, $\partial h_1(\tilde{x})$ is the subdifferential of h_1 at $\tilde{x}$ and coincides with the subdifferential in the sense of convex analysis [4], and therefore every element $\tilde{S}(\tilde{x})$ in $\partial\mathcal{L}(\tilde{x})$ is of the form

$$\tilde{S}(\tilde{x}) = S(\tilde{x}) + h_2'(\tilde{x}),$$

where $S(\tilde{x}) \in \partial h_1(\tilde{x})$ and the condition $0 \in \partial\mathcal{L}(\tilde{x})$ means $\tilde{S}(\tilde{x})(\delta\tilde{x}) = 0$ for all $\delta\tilde{x} \in U$, and equivalently,

$$S(\tilde{x})(\delta\tilde{x}) + h_2'(\tilde{x})(\delta\tilde{x}) = 0 \quad \forall \delta\tilde{x} \in U, \tag{43}$$

where the elements $\delta\tilde{x}$ are of the form $(\delta M, \delta X, \delta Y, \delta\Lambda, \delta\Upsilon)$.

Now h_1 depends only on M, and hence $\partial h_1(\tilde{x}) = \partial\|M\|$ and the action of $S(\tilde{x})$ on $\delta\tilde{x}$ is given by $S(\tilde{x})(\delta\tilde{x}) = \langle S, \delta M\rangle$, where $S \in \partial\|M\|$. Notice that the convexity of the norm $\|\cdot\|$ guarantees that the subdifferential $\partial\|M\|$ is nonempty; see, e.g., [4]. Next, we compute $h_2'(\tilde{x})(\delta\tilde{x})$. Since we are in finite dimensions and h_2 is differentiable, we have

$$h_2'(\tilde{x})(\delta\tilde{x}) = \langle (h_2)_M, \delta M\rangle + \langle (h_2)_X, \delta X\rangle + \langle (h_2)_Y, \delta Y\rangle + \langle (h_2)_\Lambda, \delta\Lambda\rangle + \langle (h_2)_\Upsilon, \delta\Upsilon\rangle,$$

where $(h_2)_M$, $(h_2)_X$, $(h_2)_Y$, $(h_2)_\Lambda$, and $(h_2)_\Upsilon$ are the partial derivatives of h_2 at $\tilde{x}$ with respect to M, X, Y, Λ, and Υ. It is straightforward to obtain that

$$\langle (h_2)_M, \delta M\rangle = \langle \Lambda, \delta M\rangle,\ \langle (h_2)_\Lambda, \delta\Lambda\rangle = \langle M - X\tilde{T}Y, \delta\Lambda\rangle,\ \langle (h_2)_\Upsilon, \delta\Upsilon\rangle = \langle XY - I, \delta\Upsilon\rangle.$$

Next, we compute the directional derivatives with respect to X along δX, and Y along δY and obtain

$$\begin{aligned}
&\lim_{\alpha\to 0}\frac{1}{\alpha}\big[h_2(M, X + \alpha\delta X, Y, \Lambda, \Upsilon) - h_2(M, X, Y, \Lambda, \Upsilon)\big]\\
&\quad = \lim_{\alpha\to 0}\frac{1}{\alpha}\big[\langle \Lambda, M - (X + \alpha\delta X)\tilde{T}Y\rangle + \langle \Upsilon, (X + \alpha\delta X)Y - I\rangle\\
&\qquad\qquad - \langle \Lambda, M - X\tilde{T}Y\rangle - \langle \Upsilon, XY - I\rangle\big]\\
&\quad = \langle \Lambda, -\delta X\tilde{T}Y\rangle + \langle \Upsilon, \delta XY\rangle = \operatorname{trace}(-\Lambda^\top\delta X\tilde{T}Y) + \operatorname{trace}(\Upsilon^\top\delta XY)\\
&\quad = \operatorname{trace}(-\tilde{T}Y\Lambda^\top\delta X) + \operatorname{trace}(Y\Upsilon^\top\delta X) = \langle -\Lambda Y^\top\tilde{T}^\top, \delta X\rangle + \langle \Upsilon Y^\top, \delta X\rangle\\
&\quad = \langle \Upsilon Y^\top - \Lambda Y^\top\tilde{T}^\top, \delta X\rangle.
\end{aligned}$$

Similarly, for Y we have

$$\begin{aligned}
&\lim_{\alpha\to 0}\frac{1}{\alpha}\big[h_2(M, X, Y + \alpha\delta Y, \Lambda, \Upsilon) - h_2(M, X, Y, \Lambda, \Upsilon)\big]\\
&\quad = \lim_{\alpha\to 0}\frac{1}{\alpha}\big[\langle \Lambda, M - X\tilde{T}(Y + \alpha\delta Y)\rangle + \langle \Upsilon, X(Y + \alpha\delta Y) - I\rangle\\
&\qquad\qquad - \langle \Lambda, M - X\tilde{T}Y\rangle - \langle \Upsilon, XY - I\rangle\big]\\
&\quad = \langle \Lambda, -X\tilde{T}\delta Y\rangle + \langle \Upsilon, X\delta Y\rangle = \langle X^\top\Upsilon - \tilde{T}^\top X^\top\Lambda, \delta Y\rangle.
\end{aligned}$$

In summary, we thus obtain

$$\begin{aligned}
S(\tilde{x})(\delta\tilde{x}) + h_2'(\tilde{x})(\delta\tilde{x}) &= \langle S + \Lambda, \delta M\rangle + \langle \Upsilon Y^\top - \Lambda Y^\top\tilde{T}^\top, \delta X\rangle\\
&\quad + \langle X^\top\Upsilon - \tilde{T}^\top X^\top\Lambda, \delta Y\rangle + \langle M - X\tilde{T}Y, \delta\Lambda\rangle + \langle XY - I, \delta\Upsilon\rangle.
\end{aligned} \tag{44}$$

Since (43) implies that (44) has to vanish for all $(\delta M, \delta X, \delta Y, \delta\Lambda, \delta\Upsilon)$, we must have

$$-\Lambda \in \partial\|M\|, \tag{45a}$$

$$\Upsilon Y^\top - \Lambda Y^\top\tilde{T}^\top = 0, \tag{45b}$$

$$X^\top\Upsilon - \tilde{T}^\top X^\top\Lambda = 0, \tag{45c}$$

$$M - X\tilde{T}Y = 0, \tag{45d}$$

$$XY - I = 0. \tag{45e}$$

The condition (45e) implies that there exists an invertible matrix G such that $Y = G$ with $X = G^{-1}$, and thus (45d) becomes

$$M = G^{-1}\tilde{T}G. \tag{46}$$

Furthermore, conditions (45b) and (45c) become

$$\Upsilon G^\top = \Lambda G^\top \tilde{T}^\top, \tag{47}$$

$$G^{-\top}\Upsilon = \tilde{T}^\top G^{-\top}\Lambda. \tag{48}$$

By multiplying (47) on the right with $G^{-\top}$ and (48) on the left with $G^\top$ and subtracting the two equalities obtained, we get

$$0 = \Lambda(G^{-1}\tilde{T}G)^\top - (G^{-1}\tilde{T}G)^\top\Lambda,$$

and using (46) leads to

$$0 = \left[\Lambda, M^\top\right]. \tag{49}$$

To conclude the proof, we recall from [27] that (45a) is equivalent to

$$\mathrm{trace}(-\Lambda^\top M) = \|M\| \quad \text{with} \quad \|\Lambda\|_* \le 1, \tag{50}$$

where $\|\cdot\|_*$ is the dual norm of $\|\cdot\|$. Hence (46), (49), and (50) are equal to (39) by denoting $M = T$. ◻

In the particular case of the Frobenius norm, the optimality system corresponding to (38) is simpler than (39). In fact, one can consider in (38) the square of the Frobenius norm, which is differentiable; see, e.g., [23]. Then the optimality system can be derived similarly as in Theorem 9, and we obtain the following corollary.

Corollary 10. *Let $\tilde{T} \in \mathbb{R}^{m\times m}$ be a given Schwarz iteration matrix, and consider the Frobenius norm $\|\cdot\|_F$. If a Schwarz iteration matrix T is a global minimizer for* (38) *with $\|\cdot\| = \|\cdot\|_F^2$, then there exists a pair of matrices $(G, \Lambda) \in \mathbb{R}^{m\times m} \times \mathbb{R}^{m\times m}$, with G invertible, such that $(T, \tilde{T}, G, \Lambda)$ satisfy the system*

$$T = G^{-1}\tilde{T}G, \tag{51a}$$

$$\left[\Lambda, T^\top\right] = 0, \tag{51b}$$

$$T + \Lambda = 0. \tag{51c}$$

Conditions (51b) and (51c) are equivalent to requiring that the matrix T be a normal matrix, a result which is obtained in [14, Example 2.2] with different theoretical arguments. For a complete discussion regarding normal matrices, we refer the reader to [12], which provides a very large number of results for characterizing normal matrices.

A necessary condition for the solvability of the optimality systems (39) and (51) is related to the kernel of the operator $\mathrm{ad}_{T^\top}(\cdot) := \left[\cdot, T^\top\right]$. Notice that if $\ker \mathrm{ad}_{T^\top} = \{0\}$ and $T \neq 0$, then (39) is not solvable because condition (39c) is not satisfied for $\Lambda = 0$. Similarly, if $\Lambda = 0$, then (51c) is not satisfied and also (51) is not solvable. For this reason, we prove the following result.

Proposition 11. *Let $T \neq 0$. Then* $\ker \mathrm{ad}_{T^\top} \neq \{0\}$.

Proof. By defining $A := -T^\top$ and $B := T^\top$, the equation $\mathrm{ad}_{T^\top}(X) = 0$ becomes

$$AX + XB = 0, \tag{52}$$

which is known as a Sylvester equation; see, e.g., [11, 18] and references therein. It is clear that a solution to (52) is $X = 0$. Now (52) has a unique solution X if and only if A and $-B$ have no eigenvalues in common; see [18]. Since $A = -T^\top = -B$, this condition is not satisfied, $X = 0$ is not the only element in $\ker \mathrm{ad}_{T^\top}$, and the claim follows. ∎

5.2. Optimality of the Schwarz iteration matrix of the atom chain. We now want to show that the Schwarz iteration matrix T_N given by (35) is optimal with respect to error estimation. To do so we need to study the optimization problem (38):

$$\min_{T_N \in \mathcal{T}(\tilde{T}_N)} \|T_N\|_\infty. \tag{53}$$

Recalling that $T_N = (I_{N-2} \otimes G_2)^{-1} \tilde{T}_N (I_{N-2} \otimes G_2)$, the optimization problem (53) can be written as

$$\min_{G_2} \|(I_{N-2} \otimes G_2)^{-1} \tilde{T}_N (I_{N-2} \otimes G_2)\|_\infty. \tag{54}$$

Now, the structure of T_N and $\tilde{T}_N$ (see also Theorem 3) allows us to write that

$$\|T_N\|_\infty = \|T_M\|_\infty \quad \text{for } N, M \geq 5,$$

and

$$\|\tilde{T}_N\|_\infty = \|\tilde{T}_M\|_\infty \quad \text{for } N, M \geq 7.$$

Hence, minimizing $\|(I_{N-2} \otimes G_2)^{-1} \tilde{T}_N (I_{N-2} \otimes G_2)\|_\infty$ in G_2 is equivalent to minimizing $\|(I_5 \otimes G_2)^{-1} \tilde{T}_7 (I_5 \otimes G_2)\|_\infty$, and problem (54) becomes

$$\min_{G_2} \|(I_5 \otimes G_2)^{-1} \tilde{T}_7 (I_5 \otimes G_2)\|_\infty. \tag{55}$$

Therefore, to show that T_N is optimal it is sufficient to study the case $N = 7$. We start with a lemma which shows that T_N for $N = 7$ has distinct eigenvalues. This fact guarantees, according to Proposition 8, that problem (38) is well-posed.

LEMMA 12. *The matrix T_N for $N = 7$ has distinct eigenvalues.*

Proof. Define $c := \frac{z_b}{w_b}$. By Lemma 1, we have that $w_b > 0$ if $k > 0$ and $\delta \in (0, L/2)$, and we also obtain that

$$\frac{w_b(k,\delta)}{z_b(k,\delta)} = \frac{e^{4k\delta + kL} - e^{kL}}{e^{2k\delta + 2kL} - e^{2k\delta}}.$$

Direct calculations show that $\frac{w_b(k,L/2)}{z_b(k,L/2)} = 1$ for any k and $\frac{\partial}{\partial \delta} \frac{w_b(k,\delta)}{z_b(k,\delta)} > 0$ for any (k, δ). Therefore, we have that $c \geq 1$, and in particular, we see that $c > 1$ for any positive δ and k. Next, for $N = 7$ the matrix T_N can be written as $T_N = w_b^2 T'$, where T' is given by

$$T' = P^\top \begin{pmatrix} B_1 & 0 \\ 0 & B_2 \end{pmatrix} P,$$

where $P = \big(P_{3,4}P_{4,5}P_{2,3}\big) \otimes I_2$ is a permutation matrix and

$$B_1 = \begin{pmatrix} c^2 & c & 0 & 0 \\ c & c^2 & c & 1 \\ 1 & c & c^2 & c \\ 0 & 0 & c & c^2 \end{pmatrix} \quad \text{and} \quad B_2 = \begin{pmatrix} c^2 & c & 0 & 0 & 0 & 0 \\ c & c^2 & c & 1 & 0 & 0 \\ 1 & c & c^2 & c & 0 & 0 \\ 0 & 0 & c & c^2 & c & 1 \\ 0 & 0 & 1 & c & c^2 & c \\ 0 & 0 & 0 & 0 & c & c^2 \end{pmatrix}.$$

Therefore, it suffices to study the eigenvalues of B_1 and B_2. Direct calculations show that the four eigenvalues of B_1 are

$$\lambda_1(c) = -\frac{\sqrt{5c^2-4c}-2c^2+c}{2}, \quad \lambda_2(c) = \frac{\sqrt{5c^2-4c}+2c^2-c}{2},$$
$$\lambda_3(c) = -\frac{\sqrt{5c^2-4c}-2c^2-c}{2}, \quad \lambda_4(c) = \frac{\sqrt{5c^2-4c}+2c^2+c}{2}.$$

Recalling that $c > 1$ for any $\delta \in (0, L/2)$, it is straightforward to see that $\lambda_4(c) > \lambda_2(c) > \lambda_3(c) > \lambda_1(c)$ for any $c > 1$. Next, we compute the characteristic polynomial of B_2:

$$\det\,(B_2 - \lambda I_6) = p_1(\lambda; c)\, p_2(\lambda; c),$$

where

$$p_1(\lambda; c) = (\lambda^3 - 3c^2\lambda^2 - c\lambda^2 + 3c^4\lambda + 2c^3\lambda - 2c^2\lambda - c\lambda - c^6 - c^5 + 2c^4 + 2c^3 - c^2 - c),$$
$$p_2(\lambda; c) = (\lambda^3 - 3c^2\lambda^2 + c\lambda^2 + 3c^4\lambda - 2c^3\lambda - 2c^2\lambda + c\lambda - c^6 + c^5 + 2c^4 - 2c^3 - c^2 + c).$$

Cumbersome calculations would allow us to compute explicitly the roots of $p_1(\lambda; c)$ and $p_2(\lambda; c)$. These roots have very complicated expressions, and we need only show that they are distinct. For this reason we proceed as follows. If we calculate the intersections of the two polynomials and can show that at all intersections their value is nonzero, then this shows that they do not have any common roots. This is easy to obtain, because in the difference $p_2(\lambda; c) - p_1(\lambda; c)$ the cubic term λ^3 cancels, and we get

$$p_3(\lambda; c) := p_2(\lambda; c) - p_1(\lambda; c) = 2c\lambda^2 - (4c^3 - 2c)\lambda + 2c^5 - 4c^3 - 2c,$$

the roots of which are $\lambda_\pm = \frac{(2c^2-1)\pm\sqrt{4c^2-3}}{2}$. Notice that λ_+ and λ_- are the only two points where the $p_1(\lambda; c)$ and $p_2(\lambda; c)$ intersect. Since a direct calculation shows that $p_1(\lambda_\pm; c) \neq 0$ and $p_2(\lambda_\pm; c) \neq 0$ for any $c > 1$, it follows that $p_1(\lambda; c)$ and $p_2(\lambda; c)$ have distinct roots. To conclude the proof it suffices to see by a direct evaluation that $p_1(\lambda_j; c) \neq 0$ and $p_2(\lambda_j; c) \neq 0$ for any $c > 1$ and $j = 1, 2, 3, 4$. ∎

We now present the optimality characterization of the Schwarz iteration matrix T_N for error estimation.

THEOREM 13. *Consider the Schwarz iteration matrices T_N and $\tilde{T}_N$ given by* (35) *and* (5). *Assume that $N \geq 7$, and that the overlap satisfies $\delta \leq L/2$. Then $T_N \in \mathcal{T}(\tilde{T}_N)$ is a candidate to be an optimal matrix for error estimation with respect to $\|\cdot\|_\infty$. Furthermore, it holds that $\|T_N\|_\infty < \|\tilde{T}_N\|_\infty$.*

Proof. By (37), $T_N = G^{-1}\tilde{T}G$, and since G is a block-diagonal matrix, and using (33), the Schwarz iteration matrix $T_N \in \mathbb{R}^{2(N-2)\times 2(N-2)}$ can be written as

$$T_N = \begin{pmatrix} & & \ddots & & & & \\ G_2^{-1}\tilde{T}_1^2 G_2 & 0 & G_2^{-1}\tilde{T}_3 G_2 & 0 & G_2^{-1}\tilde{T}_2^2 G & & \\ & G_2^{-1}\tilde{T}_1^2 G_2 & 0 & G_2^{-1}\tilde{T}_3 G_2 & 0 & G_2^{-1}\tilde{T}_2^2 G_2 \\ & & & & \ddots & \end{pmatrix}. \tag{56}$$

Denoting by I the $N-2 \times N-2$ identity, the set

$$\mathcal{S}(\tilde{T}_N) := \left\{ V \in \mathbb{R}^{2(N-2)\times 2(N-2)} | \exists A \in \mathbb{R}^{2\times 2} \text{ invertible, } G = I \otimes A \text{ s.t. } V = G^{-1}\tilde{T}_N G \right\}$$

is a subset of $\mathcal{T}(\tilde{T}_N)$, and it thus suffices to prove that T_N given by (56) is a stationary point for (38) over $\mathcal{S}(\tilde{T}_N)$. Furthermore, as we discussed at the beginning of this section, for $N \geq 7$ the specific structure of the elements of $\mathcal{S}(\tilde{T}_N)$ shows that the map $V_N \in \mathcal{S}(\tilde{T}_N) \mapsto \|V_N\|_\infty$ is constant with respect to the dimension N. It is therefore sufficient to work with $N = 7$ (over $\mathcal{S}(\tilde{T}_7)$) and to show that T_N is a stationary point for (55). Lemma 12 ensures for $N = 7$ that T_N has distinct eigenvalues and hence is diagonalizable. Since T_N is similar to $\tilde{T}_N$, $\tilde{T}_N$ is diagonalizable as well. Hence, Proposition 8 guarantees that the set $\mathcal{T}(\tilde{T}_N)$ is closed, and thus the existence of a solution to (38) follows by standard optimization arguments. Next, we recall that for $N = 7$ the matrix T_N is given by

$$T_N = \left(\begin{array}{cc|cc|cc|cc|cc}
z_b^2 & w_b z_b & & & 0 & 0 & & & & \\
w_b z_b & z_b^2 & & & z_b w_b & w_b^2 & & & & \\ \hline
& & z_b^2 & w_b z_b & & & 0 & 0 & & \\
& & w_b z_b & z_b^2 & & & z_b w_b & w_b^2 & & \\ \hline
w_b^2 & z_b w_b & & & z_b^2 & w_b z_b & & & 0 & 0 \\
0 & 0 & & & w_b z_b & z_b^2 & & & z_b w_b & w_b^2 \\ \hline
& & w_b^2 & z_b w_b & & & z_b^2 & w_b z_b & & \\
& & 0 & 0 & & & w_b z_b & z_b^2 & & \\ \hline
& & & & w_b^2 & z_b w_b & & & z_b^2 & w_b z_b \\
& & & & 0 & 0 & & & w_b z_b & z_b^2
\end{array}\right),$$

where $\|T_N\|_\infty = (z_b + w_b)^2$. Next, our aim is to construct a matrix Λ such that $(T_N, \tilde{T}_N, G_2 \otimes I, \Lambda)$ solves the optimality system (39). To do this, one can compute the kernel of $\mathrm{ad}_{T_N^\top}$ and construct the element $\Lambda \in \ker \mathrm{ad}_{T_N^\top}$ given by

$$\Lambda = \left(\begin{array}{cc|cc|cc|cc|cc}
& \frac{1}{5} & & & \frac{w_b}{5z_b} & & & & & \\
\frac{1}{5} & & & & \frac{1}{5} & & & & & \\ \hline
& & \frac{1}{8} & \frac{w_b}{4z_b} & & & \frac{1}{4} & & & \\
& & & \frac{3}{8} & & & & \frac{1}{4} & & \\ \hline
& \frac{1}{5} & & & & \frac{1}{5} & & & \frac{w_b}{5z_b} & \\
& \frac{w_b}{5z_b} & & & \frac{1}{5} & & & & \frac{1}{5} & \\ \hline
& & \frac{1}{4} & & & & \frac{3}{8} & & & \\
& & & \frac{1}{4} & & & \frac{w_b}{4z_b} & \frac{1}{8} & & \\ \hline
& & & & & \frac{1}{5} & & & & \frac{1}{5} \\
& & & & & \frac{w_b}{5z_b} & & & \frac{1}{5} &
\end{array}\right),$$

which satisfies $\mathrm{trace}(-\Lambda^\top T_N) = \|T_N\|_\infty$ and

$$\|\Lambda\|_1 = \max\left\{\frac{w_b}{5z_b} + \frac{2}{5}, \frac{w_b}{4z_b} + \frac{5}{8}\right\}.$$

Let k be a real number; similarly as in Lemma 12 we have that $\frac{w_b(k,L/2)}{z_b(k,L/2)} = 1$ for any k, and $\frac{\partial}{\partial\delta}\frac{w_b(k,\delta)}{z_b(k,\delta)} > 0$ for any (k,δ). Hence $\frac{w_b(k,\delta)}{z_b(k,\delta)} \leq 1$ for any k and $\delta \leq L/2$. This means that for $\delta \leq L/2$ we have

$$\|\Lambda\|_1 \leq 1,$$

and therefore $(T_N, \tilde{T}_N, G_2 \otimes I, \Lambda)$ solves the optimality system (39), i.e., T_N is a candidate to be an optimal Schwarz iteration matrix with respect to $\|\cdot\|_\infty$.

Next, we show that $\|\tilde{T}_N\|_\infty > \|T_N\|_\infty$. A straightforward calculation leads to

$$\|\tilde{T}_N\|_\infty = \frac{\left(2e^{6\delta k} + e^{4\delta k} - 2e^{2\delta k}\right)e^{4kL} + \left(e^{8\delta k} - 2e^{4\delta k} - 1\right)e^{2kL} + e^{4\delta k}}{(e^{2kL+4\delta k} - 1)^2}. \tag{57}$$

According to Theorem 3, we have that

$$\|T_N\|_\infty = \left(\frac{e^{2k\delta} + e^{kL}}{e^{2k\delta+kL} + 1}\right)^2. \tag{58}$$

Subtracting (58) from (57), we obtain

$$\begin{aligned}\|\tilde{T}_N\|_\infty - \|T_N\|_\infty &= \frac{\left(2e^{6\delta k} - 2e^{2\delta k}\right)e^{4kL} - \left(2e^{6\delta k} - 2e^{2\delta k}\right)e^{3kL}}{(e^{2kL+4\delta k} - 1)^2} \\ &\quad + \frac{\left(2e^{4\delta k} - 2\right)e^{2kL} + \left(2e^{6\delta k} - 2e^{2\delta k}\right)e^{kL}}{(e^{2kL+4\delta k} - 1)^2} \\ &\geq \frac{\left(2e^{6\delta k} - 2e^{2\delta k}\right)e^{4kL} - \left(2e^{6\delta k} - 2e^{2\delta k}\right)e^{3kL}}{(e^{2kL+4\delta k} - 1)^2} \\ &\quad + \frac{\left(2e^{6\delta k} - 2e^{2\delta k}\right)e^{kL}}{(e^{2kL+4\delta k} - 1)^2}.\end{aligned} \tag{59}$$

Letting $y(k)^n := e^{nkL}$, $c(k,\delta) := \frac{(2e^{6\delta k} - 2e^{2\delta k})}{(e^{2kL+4\delta k}-1)^2}$ and noticing that $c(k,\delta) > 0$ for any $\delta > 0$ and $k > 0$, we deduce from (59) that

$$\|\tilde{T}_N\|_\infty - \|T_N\|_\infty \geq c(k,\delta)\big(y(k)^3 - y(k)^2 + 1\big)y(k),$$

which is positive since $y(k) = e^{kL} > 1$ for any $k > 0$. ∎

The same arguments used to prove Theorem 13 can be applied to study the optimality of T_N with respect to the norm $\|\cdot\|_1$, which leads to the following corollary.

COROLLARY 14. *Consider the Schwarz iteration matrices T_N and $\tilde{T}_N$ given by* (37) *and* (33)*, and assume that $N \geq 7$ and the overlap satisfies $\delta \leq L/2$. Then T_N is a candidate to be an optimal Schwarz iteration matrix with respect to $\|\cdot\|_1$ in $\mathcal{T}(\tilde{T}_N)$, and $\|T_N\|_1 < \|\tilde{T}_N\|_1$.*

We now show that T_N is not optimal with respect to $\|\cdot\|_F^2$, but T_N provides a better estimate of the spectral radius also in the Frobenius norm, that is, $\|T_N\|_F < \|\tilde{T}_N\|_F$.

THEOREM 15. *For the Schwarz iteration matrices T_N and $\tilde{T}_N$ given by* (37) *and* (33)*, T_N is not a candidate to be an optimal Schwarz iteration matrix with respect to $\|\cdot\|_F^2$ in $\mathcal{T}(\tilde{T}_N)$. However, for $N \geq 7$ we have that $\|T_N\|_F < \|\tilde{T}_N\|_F$, with*

$$\|T_N\|_F^2 = 4\big(z_b^4 + z_b^2 w_b^2 + w_b^4\big) + 4\big(z_b^4 + 2z_b^2 w_b^2\big) + 2(N-6)\big(z_b^4 + 2z_b^2 w_b^2 + w_b^4\big), \tag{60}$$

where w_b and z_b are given in Lemma 1*, and*

$$\begin{aligned} \|\tilde{T}_N\|_F^2 = {} & 2\,\mathrm{trace}\big(\tilde{T}_3^\top \tilde{T}_3 + (\tilde{T}_1^2)^\top \tilde{T}_1^2\big) + 2\,\mathrm{trace}\big(\tilde{T}_3^\top \tilde{T}_3 + (\tilde{T}_2^2)^\top \tilde{T}_2^2\big) \\ & + (N-6)\,\mathrm{trace}\big(\tilde{T}_3^\top \tilde{T}_3 + (\tilde{T}_2^2)^\top \tilde{T}_2^2 + (\tilde{T}_1^2)^\top \tilde{T}_1^2\big), \end{aligned} \tag{61}$$

where $\tilde{T}_1$, $\tilde{T}_2$, and $\tilde{T}_3$ are defined in (32).

Proof. The first claim follows by noticing that T_N does not commute with its transpose $T_N^\top$. To see this, we decompose the matrix $T_N = H + S$, where $H = \frac{1}{2}(T_N + T_N^\top)$ is symmetric and $S = \frac{1}{2}(T_N - T_N^\top)$ is skew-symmetric, i.e.,

$$H = \begin{pmatrix} z_b^2 & z_b w_b & & & w_b^2/2 & 0 & & & & & \\ z_b w_b & z_b^2 & & & z_b w_b & w_b^2/2 & & & & & \\ & & z_b^2 & z_b w_b & & & w_b^2/2 & 0 & & & \\ & & z_b w_b & z_b^2 & & & z_b w_b & w_b^2/2 & & & \\ w_b^2/2 & z_b w_b & & & z_b^2 & z_b w_b & & & w_b^2/2 & 0 & \\ 0 & w_b^2/2 & & & z_b w_b & z_b^2 & & & z_b w_b & w_b^2/2 & \\ \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \end{pmatrix}$$

and

$$S = \frac{w_b^2}{2} \begin{pmatrix} 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ & & & & & \ddots & & & & \end{pmatrix}.$$

We then obtain

$$\big[T_N, T_N^\top\big] = \big[H+S, H-S\big] = \big[H,H\big] + \big[S,H\big] - \big[H,S\big] - \big[S,S\big] = 2\big[S,H\big],$$

and since the first entry of $\big[T_N, T_N^\top\big]$ is given by $\big(2\big[S,H\big]\big)_{1,1} = -w_b^4$, we obtain by Lemma 1 that T_N does not commute with its transpose.

Next, we show that $\|T_N\|_F^2 < \|\tilde{T}_N\|_F^2$. The results in (60) and (61) can be proved by induction working on their structure presented in section 4 and (33). Then using (60) and (61), a direct calculation shows

$$\|\tilde{T}_N\|_F^2 - \|T_N\|_F^2 = \phi(k,\delta)(N-6) + \varphi(k,\delta), \tag{62}$$

where

$$\phi(k,\delta) = \frac{2(e^{12k\delta} - 2e^{8k\delta} + e^{4k\delta})(e^{8kL} - 3e^{6kL} + 8e^{4kL} - 3e^{2kL} + 1)}{(e^{2kL+4k\delta} - 1)^4}$$

and

$$\varphi(k,\delta) = \frac{2(e^{12k\delta} - 2e^{8k\delta} + e^{4k\delta})(3e^{8kL} - 10e^{6kL} + 22e^{4kL} - 10e^{2kL} + 3)}{(e^{2kL+4k\delta} - 1)^4}.$$

Since for any positive δ and k we have the inequalities

$$\begin{aligned}&(e^{12k\delta} - 2e^{8k\delta} + e^{4k\delta}) > 0, \quad (e^{8kL} - 3e^{6kL} + 8e^{4kL} - 3e^{2kL} + 1) > 0,\\ &(3e^{8kL} - 10e^{6kL} + 22e^{4kL} - 10e^{2kL} + 3) > 0,\end{aligned}$$

we obtain that $\phi(k,\delta) > 0$ and $\varphi(k,\delta) > 0$, and then (62) yields $\|T_N\|_F < \|\tilde{T}_N\|_F$. $\square$

5.3. Further remarks. In this section, we study Schwarz iteration matrices T_N and $\tilde{T}_N$ for $N = 4$. In particular, in Example 3 we show that $\tilde{T}_N$ does not satisfy the optimality system (39). In Example 4, we provide an example that shows the validity of Theorem 13 also for a case with $N < 7$; that is, T_N for $N = 4$ satisfies (39).

Example 3. To show that $\tilde{T}_N$ with $N = 4$ and G the identity does not satisfy the optimality system (39), we use (32) and (33) to obtain

$$\tilde{T}_N = \left(\begin{array}{cc|cc} \tilde{f} & \tilde{a} & & \\ \tilde{b} & \tilde{f} & & \\ \hline & & \tilde{f} & \tilde{a} \\ & & \tilde{b} & \tilde{f} \end{array}\right), \qquad \begin{aligned} \tilde{f} &:= \frac{e^{4\delta k}\left(e^{kL}-1\right)^2\left(e^{kL}+1\right)^2}{(e^{kL+2\delta k}-1)^2(e^{kL+2\delta k}+1)^2},\\ \tilde{a} &:= \frac{\left(e^{6\delta k}-e^{2\delta k}\right)e^{4kL}+\left(e^{2\delta k}-e^{6\delta k}\right)e^{2kL}}{(e^{kL+2\delta k}-1)^2(e^{kL+2\delta k}+1)^2},\\ \tilde{b} &:= \frac{\left(e^{6\delta k}-e^{2\delta k}\right)e^{2kL}+\left(e^{2\delta k}-e^{6\delta k}\right)}{(e^{kL+2\delta k}-1)^2(e^{kL+2\delta k}+1)^2}. \end{aligned} \tag{63}$$

Since $\tilde{f}$, $\tilde{a}$ and $\tilde{b}$ are positive, $\|\tilde{T}_N\|_\infty = \max\{\tilde{f} + \tilde{a}, \tilde{f} + \tilde{b}\}$. Next, we define $H := \tilde{T}_N \otimes I_4 - I_4 \otimes \tilde{T}_N^\top$, where I_4 is the 4×4 identity, and after a lengthy calculation we obtain

$$\ker H = I_2 \otimes \begin{pmatrix} 0 & 1 & 0 & 0 \\ \frac{\tilde{a}}{\tilde{b}} & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & \frac{\tilde{a}}{\tilde{b}} & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

The columns of $\ker H$ are the vectorizations of matrices $H_j \in \mathbb{R}^{4\times 4}$, with $j = 1, \dots, 8$, belonging to the kernel of $\mathrm{ad}_{\tilde{T}}$. Now, we look for a linear combination of $\Lambda = \sum_j c_j H_j$ that satisfies the optimality system (39), that is, $\mathrm{trace}\big(-\Lambda^\top \tilde{T}_N\big) = \|\tilde{T}_N\|_\infty$ and $\|\Lambda\|_1 \leq 1$, and compute

$$\Big(\mathrm{trace}(-H_1^\top \tilde{T}_N), \dots, \mathrm{trace}(-H_8^\top \tilde{T}_N)\Big) = \Big(-2\tilde{a}, -2\tilde{f}, 0, 0, 0, 0, -2\tilde{a}, -2\tilde{f}\Big)^\top.$$

This says that the only linear combination that can satisfy $\mathrm{trace}(-\Lambda^\top \tilde{T}_N) = \|\tilde{T}_N\|_\infty = \max\{\tilde{f} + \tilde{a}, \tilde{f} + \tilde{b}\}$ is $\Lambda = \frac{1}{2}(H_1 + H_2)$ with the restriction $\tilde{a} \geq \tilde{b}$, and this linear combination has to satisfy $\|\Lambda\|_1 \leq 1$, which means $\frac{1}{2}(\frac{\tilde{a}}{\tilde{b}} + 1) \leq 1$, which is satisfied for $\tilde{a} \leq \tilde{b}$. Hence $\tilde{T}$ with $G = I_6$ satisfies the optimality system only for $\tilde{a} = \tilde{b}$; this condition is in general not satisfied, as one sees in (63).

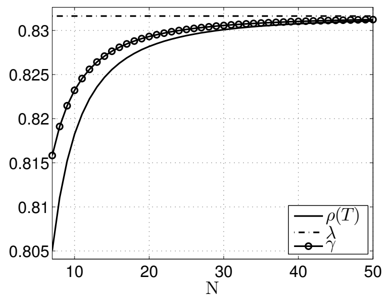

FIG. 3. *Behavior of $\rho(T)$, λ, and γ as a function of the number N of atoms.*

Example 4. In this example, we discuss the result obtained in Theorem 13 for the case $N = 4$. The matrix T_N for $N = 4$ is given by

$$T_N = \left(\begin{array}{cc|cc} z_b^2 & w_b z_b & & \\ w_b z_b & z_b^2 & & \\ \hline & & z_b^2 & w_b z_b \\ & & w_b z_b & z_b^2 \end{array}\right),$$

where $\|T_N\|_\infty = z_b^2 + z_b w_b$. Defining

$$\Lambda := \left(\begin{array}{cc|cc} -\frac{1}{2} & -\frac{1}{2} & & \\ -\frac{1}{2} & -\frac{1}{2} & & \\ \hline & & 0 & 0 \\ & & 0 & 0 \end{array}\right),$$

we see that $\Lambda \in \ker \mathrm{ad}_{\tilde{T}_N}$ and that $\mathrm{trace}(-\Lambda^\top T_N) = \|T_N\|_\infty$. Furthermore, we have that $\|T_N\|_1 = 1$, and hence $(T_N, \tilde{T}_N, G_2 \otimes I, \Lambda)$ solves the optimality system (39).

6. Numerical experiments. We first compute numerically the spectral radius of T and compare the results with the functions λ and γ from the spectral bounds proved in Theorems 3 and 5. Figure 3 shows the behavior of $\rho(T)$, λ, and γ for an increasing number of atoms N with $L = 1$, $\delta = 0.1$, and $k = 1$. We clearly see that λ does not depend on the number of atoms and is a global bound of $\rho(T)$ and γ. Notice that this figure shows exactly the estimate (28). The parallel Schwarz method therefore converges independently of the number of subdomains. In Figure 4 we study the decay of $\rho(T)$, λ, and γ as functions of the overlap δ and the Fourier modes k, for $L = 1$, $N = 10$, and $k = 1$ for the left panel and for $\delta = 0.05$ for the right panel. It is clear that the spectral radius $\rho(T)$ is bounded as $\rho(T) \leq \gamma \leq \lambda < 1$ in agreement with (28).

In the next experiment, we solve numerically problem (1)–(3) with $f_j = 0$ and $g_j = 0$ by applying the parallel Schwarz method. We generate the error sequence $\{e_j^n\}_n$ by solving (4)–(5) starting with an initial error $e_j^0 = 1$ in Ω_j. We choose atoms which are unit squares (including the overlap), and the Laplace operator is discretized by the classical 5-point finite difference stencil defined on a uniform grid with J interior points, with $J = 50$. The mesh size is $h = \frac{1}{J+1}$, $L - 2\delta = 1$, and the overlap is set to $2\delta = 10h$. The iterative procedure is stopped when

$$\|e^n\|_{\max} := \max_{1 \leq j \leq N} \max_{y \in [0,1]} \max\{|e^n(a_j, y)|, |e^n(b_j, y)|\} \leq \mathrm{tol},$$

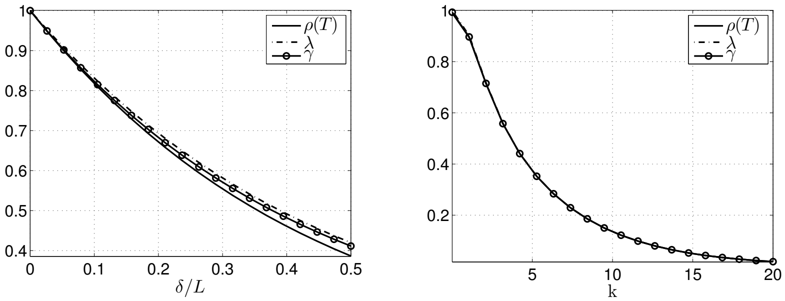

FIG. 4. *Decay of* $\rho(T)$*,* λ*, and* γ *with respect to the overlap ratio* δ/L *(left) and the Fourier modes* k *(right).*

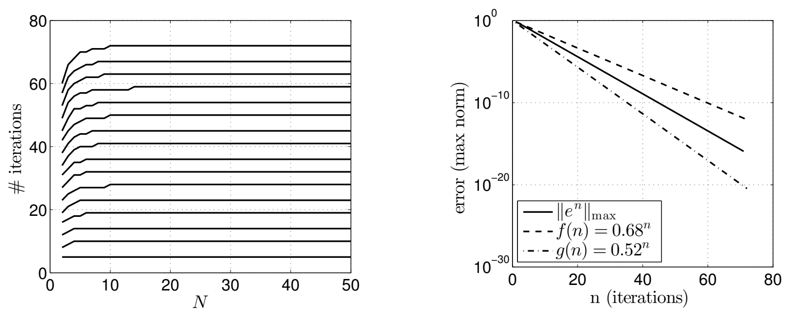

FIG. 5. *Left: Number of iterations performed as a function of the number of atoms* N*; each curve corresponds to a different fixed tolerance. Right: Rate of convergence of the Schwarz method for the solution with* $N = 50$ *atoms with* $\mathrm{tol} = 10^{-16}$*.*

where tol is a given tolerance. We repeat this experiment with different numbers of atoms N and different values of tol. The results are shown in Figure 5. The left figure shows the number of iterations performed as a function of the number of atoms N. Each of the 16 curves corresponds to a fixed value of the tolerance $\mathrm{tol} = 10^{-\alpha}$, where $\alpha = 1, \dots, 16$. The first curve from the bottom corresponds to $\alpha = 1$, the second to $\alpha = 2$, and so on until the top curve, which corresponds to $\alpha = 16$. Clearly, the number of iterations required increases when the tolerance is reduced, but all the curves have the same shape: the number of required iterations slightly increases for small N and then remains constant. This illustrates numerically what we proved in section 4: the convergence of the parallel Schwarz method is independent of the number N of atoms and hence subdomains. This behavior was also observed in Figure 10 of [2]; however, the authors considered only a tolerance $\mathrm{tol} = 10^{-2}$. In Figure 5 (right), we show the decay of the maximum norm $\|e^n\|_{\max}$ with respect to the iterations n. The result corresponds to $N = 50$ and $\mathrm{tol} = 10^{-16}$. To study the rate of convergence, the curve corresponding to $\|e^n\|_{\max}$ (solid line) is compared with the two functions $f(n) = 0.68^n$ (dashed line) and $g(n) = 0.52^n$ (dash-dotted line). Notice that for $\hat{L} = 1$, the lowest mode corresponds to $k = \pi$. Therefore, we can easily compute $\|T(\pi, 5h)\|_\infty \approx 0.354$. Recalling that $\boldsymbol{v}^n(k) = T(k,\delta)\boldsymbol{v}^{n-2}(k)$, we find $\|\boldsymbol{v}^n\|_{\max} \leq (\|T(\pi, 5h)\|_\infty^{1/2})^n \|\boldsymbol{v}^0\|_{\max}$, with $\|T(\pi, 5h)\|_\infty^{1/2} \approx 0.592$, which is the contraction factor shown in Figure 5 (right). Moreover, for a tolerance $\mathrm{tol} = 10^{-16}$ we estimate the number of iterations as $n \approx 2\frac{|\log(\mathrm{tol})|}{|\log(\|T(\pi,5h)\|_\infty)|} \approx 71$, which can be identified in the upper curve in Figure 5 (left).

7. Conclusions. We introduced an approximation of the physical model presented in [2, 20, 21] for a chain of atoms to prove that the convergence of the parallel one-level Schwarz method using one atom per subdomain is independent of the number of atoms and thus subdomains. Our convergence results presented for the 2-dimensional model can be trivially extended to a 3-dimensional model. We then showed that the choice of the Schwarz iteration matrix can strongly affect the convergence analysis of a Schwarz method. Motivated by this observation, we have defined a new concept of optimality for Schwarz iteration matrices for error estimation, and we have proved optimality of the Schwarz iteration matrix we used to obtain our convergence results. We illustrated our theoretical results with numerical experiments.

Acknowledgments. The authors would like to thank two anonymous referees for their helpful comments. A very special thanks goes to Bart Vandereycken for a very useful discussion regarding the set $\mathcal{T}(\tilde{T})$ and for pointing out the reference [14]. The authors would like to thank Julien Salomon for reading an earlier version of this manuscript and providing useful opinions and remarks. The first author would also like to thank Faycal Chaouqui for a discussion on Corollary 4.

REFERENCES

[1] V. Barone and M. Cossi, *Quantum calculation of molecular energies and energy gradients in solution by a conductor solvent model*, J. Phys. Chem. A, 102 (1998), pp. 1995–2001, https://doi.org/10.1021/jp9716997.

[2] E. Cancès, Y. Maday, and B. Stamm, *Domain decomposition for implicit solvation models*, J. Chem. Phys., 139 (2013), 054111, https://doi.org/10.1063/1.4816767.

[3] F. H. Clarke, *A new approach to Lagrange multipliers*, Math. Oper. Res., 1 (1976), pp. 165–174, https://doi.org/10.1287/moor.1.2.165.

[4] F. H. Clarke, *Optimization and Nonsmooth Analysis*, SIAM, Philadelphia, 1990, https://doi.org/10.1137/1.9781611971309.

[5] N. Deo, *Graph Theory with Applications to Engineering and Computer Science*, Prentice-Hall Series in Automatic Computation, Prentice-Hall, Englewood Cliffs, NJ, 1974.

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

[7] M. J. Gander, *Optimized Schwarz methods*, SIAM J. Numer. Anal., 44 (2006), pp. 699–731, https://doi.org/10.1137/S0036142903425409.

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

[9] M. J. Gander, L. Halpern, and F. Nataf, *Optimal Schwarz waveform relaxation for the one dimensional wave equation*, SIAM J. Numer. Anal., 41 (2003), pp. 1643–1681, https://doi.org/10.1137/S003614290139559X.

[10] M. J. Gander and A. M. Stuart, *Space-time continuous analysis of waveform relaxation for the heat equation*, SIAM J. Sci. Comput., 19 (1998), pp. 2014–2031, https://doi.org/10.1137/S1064827596305337.

[11] F. Gantmacher, *The Theory of Matrices*, Vol. 1, Chelsea Publishing, New York, 1959.

[12] R. Grone, C. R. Johnson, E. M. Sa, and H. Wolkowicz, *Normal matrices*, Linear Algebra Appl., 87 (1987), pp. 213–225, https://doi.org/10.1016/0024-3795(87)90168-6.

[13] F. Harary, *Graph Theory*, Addison-Wesley, Reading, MA, 1969.

[14] U. Helmke and J. B. Moore, *Optimization and Dynamical Systems*, Comm. Control Engrg. Ser., Springer-Verlag, London, 1994, https://doi.org/10.1007/978-1-4471-3467-1.

[15] J.-B. Hiriart-Urruty, *Refinements of necessary optimality conditions in nondifferentiable programming* I, Appl. Math. Optim., 5 (1979), pp. 63–82, https://doi.org/10.1007/BF01442544.

[16] A. Klamt and G. Schüürmann, *COSMO: A new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient*, J. Chem. Soc. Perkin Trans. 2, no. 5 (1993), pp. 799–805, https://doi.org/10.1039/P29930000799.

[17] K. Lange, *Optimization*, Springer Texts Statist. 95, 2nd ed., Springer-Verlag, New York, 2013, https://doi.org/10.1007/978-1-4614-5838-8.

[18] A. J. Laub, *Matrix Analysis for Scientists and Engineers*, SIAM, Philadelphia, 2005.
[19] P.-L. Lions, *On the Schwarz alternating method.* I, in Proceedings of the First International Symposium on Domain Decomposition Methods for Partial Differential Equations (Paris, 1987), SIAM, Philadelphia, 1988, pp. 1–42.
[20] F. Lipparini, G. Scalmani, L. Lagardère, B. Stamm, E. Cancès, Y. Maday, J.-P. Piquemal, M. J. Frisch, and B. Mennucci, *Quantum, classical, and hybrid QM/MM calculations in solution: General implementation of the ddCOSMO linear scaling strategy*, J. Chem. Phys., 141 (2014), 184108, https://doi.org/10.1063/1.4901304.
[21] F. Lipparini, B. Stamm, E. Cancès, Y. Maday, and B. Mennucci, *Fast domain decomposition algorithm for continuum solvation models: Energy and first derivatives*, J. Chem. Theory Comput., 9 (2013), pp. 3637–3648, https://doi.org/10.1021/ct400280b.
[22] J. H. C. Luke, *Convergence of a multiple reflection method for calculating Stokes flow in a suspension*, SIAM J. Appl. Math., 49 (1989), pp. 1635–1651, https://doi.org/10.1137/0149099.
[23] K. B. Petersen and M. S. Pedersen, *The Matrix Cookbook*, Technical University of Denmark, Kongens Lyngby, Denmark, 2012; available online at http://www2.imm.dtu.dk/pubdb/views/edoc_download.php/3274/pdf/imm3274.pdf.
[24] A. Qaddouri, L. Laayouni, S. Loisel, J. Côté, and M. J. Gander, *Optimized Schwarz methods with an overset grid for the shallow-water equations: Preliminary results*, Appl. Numer. Math., 58 (2008), pp. 459–471, https://doi.org/10.1016/j.apnum.2007.01.015.
[25] A. Toselli and O. Widlund, *Domain Decomposition Methods: Algorithms and Theory*, Springer Ser. Comput. Math. 34, Springer-Verlag, Berlin, 2005, https://doi.org/10.1007/b137868.
[26] T. N. Truong and E. V. Stefanovich, *A new method for incorporating solvent effect into the classical, ab initio molecular orbital and density functional theory frameworks for arbitrary shape cavity*, Chem. Phys. Lett., 240 (1995), pp. 253–260, https://doi.org/10.1016/0009-2614(95)00541-B.
[27] G. A. Watson, *Characterization of the subdifferential of some matrix norms*, Linear Algebra Appl., 170 (1992), pp. 33–45, https://doi.org/10.1016/0024-3795(92)90407-2.