

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

GABRIELE CIARAMELLA[†] AND MARTIN J. GANDER[‡]

Abstract. In the ddCOSMO solvation model for the numerical simulation of molecules (chains of atoms), the unusual observation was made that the associated Schwarz domain-decomposition method converges independently of the number of subdomains (atoms) and this without coarse correction, i.e., the one-level Schwarz method is scalable. We analyzed this unusual property for the simplified case of a rectangular molecule and square subdomains using Fourier analysis, leading to robust convergence estimates in the L^2 -norm and later also for chains of subdomains represented by disks using maximum principle arguments, leading to robust convergence estimates in L^∞ . A convergence analysis in the more natural H^1 -setting proving convergence independently of the number of subdomains was, however, missing. We close this gap in this paper using tools from the theory of alternating projection methods and estimates introduced by P.-L. Lions for the study of domain decomposition methods. We prove that robust convergence independently of the number of subdomains is possible also in H^1 and show furthermore that even for certain two-dimensional domains with holes, Schwarz methods can be scalable without coarse-space corrections. As a by-product, we review some of the results of P.-L. Lions [On the Schwarz alternating method. I, in Domain Decomposition Methods for Partial Differential Equations, SIAM, Philadelphia, 1988, pp. 1–42] and in some cases provide simpler proofs.

Key words. domain decomposition methods, Schwarz methods, chain of subdomains, elliptic PDE, Laplace equation, COSMO solvation model

AMS subject classifications. 65N55, 65F10, 65N22, 70-08, 35J05, 35J57

1. Introduction. An integral equation-based implementation of the parallel Schwarz method (PSM) called ddCOSMO was introduced in [4, 28, 29] for the solution of an important class of solvation problems. In ddCOSMO, large molecular systems are solved, and each atom in the molecule corresponds to a subdomain representing the spherical van der Waals cavity of the corresponding atom. The physical model underlying these solvation processes is the COSMO model presented in [2, 23, 35]. It was observed in [4, 28, 29] that the ddCOSMO solver is scalable, i.e., the iteration number to achieve convergence of the PSM is independent of the number of subdomains (the number of atoms), which is very unusual for one-level PSMs. We first proved this behavior in [6] for the simplified two-dimensional geometrical setting of growing chains of fixed-sized rectangular subdomains using Fourier analysis, which led to robust L^2 -convergence estimates. We then studied the more appropriate geometric setting of circular subdomains using maximum principle arguments in [7], which led to robust convergence estimates in L^∞ for many important molecular chains, including ringed and branched chains. Convergence in the more natural H^1 -setting remained, however, open so far.

The purpose of the present manuscript is to close this gap by using a variational approach for the study of ddCOSMO. This technique of convergence analysis is very different from the Fourier and maximum principle techniques we used previously and has led to the powerful abstract Schwarz framework for the analysis of the additive and multiplicative Schwarz preconditioners; see [34] and the references therein. We introduce in Section 2 a variational formulation for the study of ddCOSMO and show how ddCOSMO is related to the alternating projection method in Hilbert spaces. To do so and to make this paper self-contained, we review some of the results of P.-L. Lions [26] and in some cases provide simpler proofs. This also allows us to show how P.-L. Lions's convergence analysis of the PSM is intricately

*Received May 11, 2018. Accepted August, 2018. Published online on October 16, 2018. Recommended by Daniel Szyld.

[†]Universität Konstanz, Germany (gabriele.ciaramella@uni-konstanz.de).

[‡]University of Geneva, Switzerland (martin.gander@unige.ch).

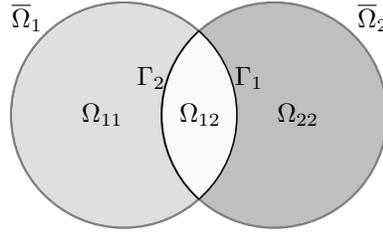


FIG. 2.1. Geometry of the sets defined in (2.1) corresponding to the intersection of two discs Ω_1 and Ω_2 . Notice that $\bar{\Omega} = \bar{\Omega}_{11} \cup \bar{\Omega}_{12} \cup \bar{\Omega}_{22}$. The two black curves correspond to Γ_1 and Γ_2 , and it holds that $\partial\Omega_{12} = \Gamma_1 \cup \Gamma_2$.

related to the alternating projection method by carefully retracing his convergence proof, which eventually led after many discussions to the overview given in Figure 2.4. We add concrete examples to illustrate the abstract concepts. This allows us in Section 3 to give concrete convergence estimates for ddCOSMO in H^1 and to prove that the convergence is robust with respect to the length of the molecules. In Section 4, we finally show that more general domains can also have associated one-level PSMs that converge independently of the number of subdomains, in particular domains with holes, which constitutes an interesting result in the context of multiscale problems and homogenization.

2. Variational approach for the parallel Schwarz method. In this section, we analyze the convergence of the parallel Schwarz method by means of a variational analysis. This approach was proposed by P.-L. Lions in [26]. In order to properly use the results provided in [26], we describe the proposed variational approach, prove the presented facts that are useful for our purposes, and provide the proofs and/or references that were not given in [26].

We start discussing the case of only two subdomains. As we will see, the problem of many subdomains can be treated by extending the results proved for two subdomains. Let Ω be a domain in \mathbb{R}^m with $m \in \{1, 2, 3\}$ that can be decomposed as $\Omega = \Omega_1 \cup \Omega_2$, where Ω_1 and Ω_2 are also open domains in \mathbb{R}^m . We assume that Ω , Ω_1 , and Ω_2 have Lipschitz boundaries. Associated with the subdomains Ω_1 and Ω_2 , we define

$$(2.1) \quad \begin{aligned} \Gamma_1 &:= \partial\Omega_1 \cap \bar{\Omega}_2, & \Gamma_2 &:= \partial\Omega_2 \cap \bar{\Omega}_1, \\ \Omega_{11} &:= \Omega_1 \setminus \bar{\Omega}_2, & \Omega_{22} &:= \Omega_2 \setminus \bar{\Omega}_1, & \Omega_{12} &:= \Omega_1 \cap \Omega_2. \end{aligned}$$

For any two functions $v, w \in H_0^1(\Omega)$, we denote by $\langle v, w \rangle := \int_{\Omega} \nabla v(\mathbf{x}) \cdot \nabla w(\mathbf{x}) \, d\mathbf{x}$ the inner product for $H_0^1(\Omega)$ and by $\|\cdot\|$ the corresponding induced norm. In the context of COSMO models, the Poisson equation has to be solved,

$$(2.2) \quad \begin{aligned} -\Delta u &= f & \text{in } \Omega, \\ u &= g & \text{on } \partial\Omega, \end{aligned}$$

where f is assumed in $H^{-1}(\Omega)$ (or in $L^2(\Omega)$) and $g \in H^{1/2}(\partial\Omega)$ such that (2.2) is uniquely solvable by $u \in H^1(\Omega)$; see, e.g., [13, Section 6.2]. The solution to (2.2) can be obtained by means of the alternating Schwarz method (AltSM). Given any initial guess $u^0 \in H^1(\Omega)$ such that $u^0 = g$ on $\partial\Omega$, this method generates a sequence $\{u^i\}_{i \in \mathbb{N}} \subset H^1(\Omega)$ whose odd elements u^{2n+1} , for $n = 0, 1, 2, \dots$, are defined as follows: the restriction of u^{2n+1} to Ω_1 is given by the solution of

$$(2.3) \quad \begin{aligned} -\Delta u^{2n+1} &= f & \text{in } \Omega_1, \\ u^{2n+1} &= g & \text{on } \partial\Omega_1 \setminus \Gamma_1, \\ u^{2n+1} &= u^{2n} & \text{on } \Gamma_1, \end{aligned}$$

which is then extended to $\overline{\Omega}_{22}$ by u^{2n} , that is, $u^{2n+1} := u^{2n}$ in $\overline{\Omega}_{22}$. Similarly, the even elements are denoted by u^{2n} , for $n = 1, 2, \dots$, whose restriction to Ω_2 is given by the solution of

$$(2.4) \quad \begin{aligned} -\Delta u^{2n} &= f && \text{in } \Omega_2, \\ u^{2n} &= g && \text{on } \partial\Omega_2 \setminus \Gamma_2, \\ u^{2n} &= u^{2n-1} && \text{on } \Gamma_2, \end{aligned}$$

which is then extended on $\overline{\Omega}_{11}$ by u^{2n-1} , that is, $u^{2n} := u^{2n-1}$ in $\overline{\Omega}_{11}$. So in summary, we have that

$$(2.5) \quad u^{2n+1} := \begin{cases} \text{solution of (2.3)} & \text{in } \overline{\Omega}_1, \\ u^{2n} & \text{in } \overline{\Omega}_{22}, \end{cases} \quad \text{and} \quad u^{2n} := \begin{cases} \text{solution of (2.4)} & \text{in } \overline{\Omega}_2, \\ u^{2n-1} & \text{in } \overline{\Omega}_{11}. \end{cases}$$

Lemma 2.1 below guarantees that u , u^{2n} , and u^{2n+1} are functions in $H^1(\Omega)$, and their traces on $\partial\Omega$ equal g .

To reformulate the AltSM as a projection method, we need the two spaces $H_0^1(\Omega_1)$ and $H_0^1(\Omega_2)$ and their extensions by zero to all of Ω , which we denote by V_1 and V_2 . These are closed subspaces of $V := H_0^1(\Omega)$ (see [26]), and they contain the differences of iterates of the AltSM as the following Lemma shows:

LEMMA 2.1. *For any u^{2n+1} and u^{2n} given by the AltSM (2.5), it holds that*

$$u^{2n}, u^{2n+1} \in H^1(\Omega), \quad u^{2n+1} - u^{2n} \in V_1, \quad u^{2n} - u^{2n-1} \in V_2.$$

Proof. To see that $u^{2n+1} \in H^1(\Omega)$, we need to show that the function

$$\mathbf{z} := \begin{cases} \nabla u^{2n+1} & \text{in } \Omega_1, \\ \nabla u^{2n} & \text{in } \Omega_{22}, \end{cases}$$

is the weak derivative of u^{2n+1} and that $\|\mathbf{z}\|_{L^2(\Omega)} < \infty$. To do so, let x_i be the i th component of \mathbf{x} . Recalling that $u^{2n+1}|_{\Omega_1} \in H^1(\Omega_1)$ and $u^{2n}|_{\Omega_{22}} \in H^1(\Omega_{22})$, we obtain for any smooth test function v with compact support that

$$\begin{aligned} \int_{\Omega} u^{2n+1} \partial_{x_j} v \, d\mathbf{x} &= \int_{\Omega_1} u^{2n+1} \partial_{x_j} v \, d\mathbf{x} + \int_{\Omega_{22}} u^{2n} \partial_{x_j} v \, d\mathbf{x} \\ &= - \int_{\Omega_1} \partial_{x_i} u^{2n+1} v \, d\mathbf{x} + \int_{\partial\Omega_1} u^{2n+1} v n_{1,j} \, ds_1 \\ &\quad - \int_{\Omega_{22}} \partial_{x_j} u^{2n} v \, d\mathbf{x} + \int_{\partial\Omega_{22}} u^{2n} v n_{22,j} \, ds_{22}, \end{aligned}$$

where we integrated by parts and $n_{1,j}$ denotes the j th component of the unit outward normal \mathbf{n}_1 on $\partial\Omega_1$, $n_{22,j}$ the j th component of \mathbf{n}_{22} on $\partial\Omega_{22}$, and ds_1 and ds_{22} the corresponding surface elements.

Since v has compact support in Ω , we have

$$\begin{aligned} \int_{\Omega} u^{2n+1} \partial_{x_j} v \, d\mathbf{x} &= - \int_{\Omega_1} \partial_{x_i} u^{2n+1} v \, d\mathbf{x} + \int_{\Gamma_1} u^{2n+1} v n_{1,j} \, ds_1 \\ &\quad - \int_{\Omega_{22}} \partial_{x_j} u^{2n} v \, d\mathbf{x} + \int_{\Gamma_1} u^{2n} v n_{22,j} \, ds_{22} \\ &= - \int_{\Omega_1} \partial_{x_j} u^{2n+1} v \, d\mathbf{x} - \int_{\Omega_{22}} \partial_{x_j} u^{2n} v \, d\mathbf{x} = - \int_{\Omega} z_j v \, d\mathbf{x}, \end{aligned}$$

where we used the fact that $\mathbf{n}_1 = -\mathbf{n}_{22}$ on Γ_1 and denoted by z_j the j th component of \mathbf{z} . This shows that \mathbf{z} is the weak derivative of u^{2n+1} . Now, recalling again that $u^{2n+1}|_{\Omega_1} \in H^1(\Omega_1)$ and $u^{2n}|_{\Omega_{22}} \in H^1(\Omega_{22})$, we obtain that

$$\|z_j\|_{L^2(\Omega)}^2 = \|\partial_{x_j} u^{2n+1}\|_{L^2(\Omega_1)}^2 + \|\partial_{x_j} u^{2n}\|_{L^2(\Omega_{22})}^2 < \infty,$$

and hence, $u^{2n+1} \in H^1(\Omega)$. Similarly, one can also show that $u^{2n} \in H^1(\Omega)$.

We now consider the difference $u^{2n+1} - u^{2n}$: first note that $u^{2n+1}|_{\Omega_1} \in H^1(\Omega_1)$ and $u^{2n}|_{\Omega_2} \in H^1(\Omega_2)$ since they solve (2.3) and (2.4) and that

$$\begin{aligned} u^{2n+1}|_{\overline{\Omega_{22}}} &= u^{2n}|_{\overline{\Omega_{22}}}, & u^{2n}|_{\overline{\Omega_{11}}} &= u^{2n-1}|_{\overline{\Omega_{11}}}, \\ u^{2n+1} &= u^{2n} \text{ on } \partial\Omega_1, & u^{2n} &= u^{2n-1} \text{ on } \partial\Omega_2. \end{aligned}$$

Hence, we obtain for the difference

$$u^{2n+1} - u^{2n} = \begin{cases} 0 & \text{in } \overline{\Omega} \setminus \Omega_1, \\ u^{2n+1}|_{\Omega_1} - u^{2n}|_{\Omega_1} & \text{in } \Omega_1, \end{cases}$$

and thus the difference satisfies $u^{2n+1} - u^{2n} \in V_1$ because

$$\|u^{2n+1}|_{\Omega_1} - u^{2n}|_{\Omega_1}\|_{H^1(\Omega_1)} \leq \|u^{2n+1}\|_{H^1(\Omega)} + \|u^{2n}\|_{H^1(\Omega)} < \infty.$$

Similarly, one can prove that $u^{2n} - u^{2n-1} \in V_2$. \square

The underlying Poisson equation (2.2) in weak form is

$$(2.6) \quad \langle u, v \rangle = \int_{\Omega} f v \, d\mathbf{x} \quad \forall v \in V, \quad u \in \{v \in H^1(\Omega) : v = g \text{ on } \partial\Omega\},$$

and the subdomain problems (2.3)–(2.4) of the AltSM in weak form are

$$(2.7) \quad \int_{\Omega_1} \nabla u^{2n+1} \cdot \nabla v_1 \, d\mathbf{x} = \int_{\Omega_1} f v_1 \, d\mathbf{x} \quad \forall v_1 \in V_1,$$

with $u^{2n+1} \in \{w \in H^1(\Omega) : w = u^{2n} \text{ in } \overline{\Omega_{22}} \cup \partial\Omega_1\}$, for $n = 0, 1, 2, \dots$, and

$$\int_{\Omega_2} \nabla u^{2n} \cdot \nabla v_2 \, d\mathbf{x} = \int_{\Omega_2} f v_2 \, d\mathbf{x} \quad \forall v_2 \in V_2,$$

with $u^{2n} \in \{w \in H^1(\Omega) : w = u^{2n-1} \text{ in } \overline{\Omega_{11}} \cup \partial\Omega_2\}$, for $n = 1, 2, \dots$. Since any function $v_1 \in V_1$ is zero on Ω_{22} , we can extend the integral in the weak subdomain problem (2.7) to Ω , and subtract from it the weak original problem (2.6) tested by any $v_1 \in V_1$ to obtain

$$(2.8) \quad \langle u^{2n+1} - u, v_1 \rangle = 0, \quad \forall v_1 \in V_1, \quad u^{2n+1} - u^{2n} \in V_1,$$

where we used Lemma 2.1. Similarly, we also obtain

$$(2.9) \quad \langle u^{2n} - u, v_2 \rangle = 0, \quad \forall v_2 \in V_2, \quad u^{2n} - u^{2n-1} \in V_2.$$

This is the alternating Schwarz method in variational form presented in [26]. Using the projection theorem [9, Theorem 4.3-1], the iteration (2.8)–(2.9) can also be written in terms of the orthogonal projections \mathbb{P}_{V_1} onto V_1 and \mathbb{P}_{V_2} onto V_2 ,

$$(2.10) \quad u^{2n+1} - u^{2n} = \mathbb{P}_{V_1}(u - u^{2n}),$$

for all $n \geq 0$, and

$$(2.11) \quad u^{2n} - u^{2n-1} = \mathbb{P}_{V_2}(u - u^{2n-1}),$$

for all $n \geq 1$. It is also possible to write the iteration in projection form in terms of the orthogonal complements of V_1 and V_2 denoted by V_1^\perp and V_2^\perp and defined with respect to the H_0^1 -inner product $\langle \cdot, \cdot \rangle$: using that $\mathbb{P}_{V_1} = I - \mathbb{P}_{V_1^\perp}$ and $\mathbb{P}_{V_2} = I - \mathbb{P}_{V_2^\perp}$, where $\mathbb{P}_{V_1^\perp}$ is the orthogonal projection onto V_1^\perp and $\mathbb{P}_{V_2^\perp}$ the orthogonal projection onto V_2^\perp , we obtain from identity (2.10) that

$$(2.12) \quad u - u^{2n+1} = \mathbb{P}_{V_1^\perp}(u - u^{2n}),$$

and from (2.11) that

$$(2.13) \quad u - u^{2n} = \mathbb{P}_{V_2^\perp}(u - u^{2n-1}).$$

This shows that the AltSM can be regarded as an alternating projection method (APM): by defining the error $e^{2n+1} := u - u^{2n+1}$ and inserting (2.13) into (2.12) we obtain¹

$$(2.14) \quad e^{2n+1} = \mathbb{P}_{V_1^\perp} \mathbb{P}_{V_2^\perp} e^{2n-1}$$

and similarly,

$$(2.15) \quad e^{2n} = \mathbb{P}_{V_2^\perp} \mathbb{P}_{V_1^\perp} e^{2n-2}.$$

By induction, we obtain from (2.14) and (2.15) that

$$(2.16) \quad e^{2n+1} = \left(\mathbb{P}_{V_1^\perp} \mathbb{P}_{V_2^\perp} \right)^k e^{2n-2k+1} = \left(\mathbb{P}_{V_1^\perp} \mathbb{P}_{V_2^\perp} \right)^n e^1$$

and

$$(2.17) \quad e^{2n} = \left(\mathbb{P}_{V_2^\perp} \mathbb{P}_{V_1^\perp} \right)^k e^{2n-2k} = \left(\mathbb{P}_{V_2^\perp} \mathbb{P}_{V_1^\perp} \right)^n e^0,$$

for $k = 0, 1, \dots, n$. We see that the convergence of the sequence $\{e^i\}_{i \in \mathbb{N}}$ is related to the properties of the composed projection operators $\mathbb{P}_{V_1^\perp} \mathbb{P}_{V_2^\perp}$ and $\mathbb{P}_{V_2^\perp} \mathbb{P}_{V_1^\perp}$. A standard result (see, e.g., [9, Theorem 4.3-1]) is that an orthogonal projection operator \mathbb{P} from a Hilbert space V onto a closed subspace of V is Lipschitz continuous with Lipschitz constant $L = 1$. This is, however, not enough to guarantee convergence of the sequence $\{e^i\}_{i \in \mathbb{N}}$, and other arguments are needed.

In Sections 2.1 and 2.2 (see Lemma 2.9 and Theorem 2.11) we show that the convergence of the APM (and thus the AltSM) depends on the inclination (the angle) between the subspace V_1^\perp and V_2^\perp . Therefore, the equalities (2.16) and (2.17), together with the property $\text{incl}(V_1^\perp, V_2^\perp) = \text{incl}(V_2^\perp, V_1^\perp)$, where incl denotes the inclination between two subspaces (see Definition 2.7 and Lemma 2.9 below) show that the convergence analysis does not depend on the choice between V_1 and V_2 for the first projection.

¹Notice that (2.14) is a stationary method: in a finite-dimensional setting, let u be the solution to $Au = f$. Then (2.14) can be rewritten as $u^{n+1} = (I - \mathbb{P}_{V_1})(I - \mathbb{P}_{V_2})u^n - (\tilde{\mathbb{P}}_{V_1} + \tilde{\mathbb{P}}_{V_2} - \tilde{\mathbb{P}}_{V_1} A \tilde{\mathbb{P}}_{V_2})f$, where $\tilde{\mathbb{P}}_{V_j} := \mathbb{P}_{V_j} A^{-1}$. This is the classical multiplicative Schwarz method; see, e.g., [34, page 24].

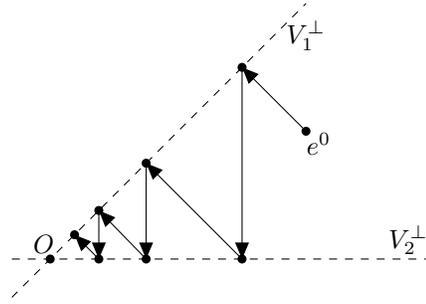


FIG. 2.2. Alternating projection method for two one-dimensional subspaces V_1^\perp and V_2^\perp (dashed lines). The APM procedure begins with the point $e^0 \in V_1^\perp + V_2^\perp$ and approaches asymptotically the point O that is the intersection $V_1^\perp \cap V_2^\perp$.

2.1. The alternating projection method. We now discuss the alternating projection method (APM) in a general Hilbert space framework. This method was studied first by von Neumann in [36], and we refer to [11] for a complete survey. In [11] the author provides a general discussion on the APM, listing a large number of results and references but omitting all the proofs. To the best of our knowledge, this is one of the first references, together with the work of P.-L. Lions [26], that connects explicitly the AltSM with APM; see [11, Section 7.3 on page 110 and also the Addendum on page 118]. A more recent reference that contains a short survey of the APM (without proofs) in connection with the method of reflections is [25]; see also [8, 31].

Let V be a Hilbert space endowed with the inner product (\cdot, \cdot) that induces the norm $\|\cdot\|$. Consider two closed subspaces V_1 and V_2 of V and denote by \mathbb{P}_{V_1} and \mathbb{P}_{V_2} the orthogonal projections of V onto V_1 and V_2 . For a given $e^0 \in V$, we define the sequence $\{e^i\}_{i \in \mathbb{N}}$ by setting

$$(2.18) \quad \begin{aligned} e^{2n+1} &:= e^{2n} - p^n, & \text{for } n = 0, 1, 2, \dots, \\ e^{2n} &:= e^{2n-1} - \tilde{p}^n, & \text{for } n = 1, 2, \dots, \end{aligned}$$

where

$$(2.19) \quad p^n := \mathbb{P}_{V_1} e^{2n}, \quad \tilde{p}^n := \mathbb{P}_{V_2} e^{2n-1}.$$

Note that (2.18)–(2.19) is equivalent to (2.10)–(2.11) by setting $e^{2n+1} = u - u^{2n+1}$ and $e^{2n} = u - u^{2n}$. The fact that $\mathbb{P}_{V_j^\perp} = I - \mathbb{P}_{V_j}$ allows us to rewrite (2.18)–(2.19) in the equivalent form

$$(2.20) \quad e^{2n+1} = \mathbb{P}_{V_1^\perp} e^{2n}, \quad e^{2n} = \mathbb{P}_{V_2^\perp} e^{2n-1},$$

which shows again the connection with the AltSM in (2.12)–(2.13). A graphical illustration of the APM is given in Figure 2.2.

Next, we want to study the convergence of the sequences $\{e^i\}_{i \in \mathbb{N}}$, $\{p^n\}_{n \geq 0}$, and $\{\tilde{p}^n\}_{n \geq 1}$. To do so, we prove Lemma 2.3, which represents a fundamental intermediate result to prove the convergence of the APM and in the next section to prove the convergence of the AltSM. Lemma 2.3 corresponds to Lemma 2.1 in [16], but we provide a more detailed and slightly

different proof. To prove Lemma 2.3, we need Kronecker's Lemma (see, e.g., [33, page 390]), and for completeness, we give an elementary proof of this fundamental result.

LEMMA 2.2 (Kronecker's Lemma). *Given a convergent series $\sum_{k=0}^{\infty} \mu_k = \tilde{\ell}$, and a monotone non-decreasing positive sequence $\{\lambda_n\}_{n \in \mathbb{N}}$ such that $\lambda_n \rightarrow \infty$. Then we have*

$$\lim_{n \rightarrow \infty} \frac{1}{\lambda_n} \sum_{k=0}^n \lambda_k \mu_k = 0.$$

Proof. Defining $s_n := \sum_{k=0}^n \mu_k$ and using summation by parts², we get

$$(2.21) \quad \frac{1}{\lambda_n} \sum_{k=0}^n \lambda_k \mu_k = s_n - \frac{1}{\lambda_n} \sum_{k=0}^{n-1} (\lambda_{k+1} - \lambda_k) s_k.$$

Since the series $\sum_{k=0}^{\infty} \mu_k$ is convergent, the corresponding partial sums converge, $s_n \rightarrow \tilde{\ell}$. Hence, for any $\epsilon > 0$, we can find an $N > 0$ such that $|s_n - \tilde{\ell}| < \epsilon$ for all $n > N$. Assuming that $n > N$, equation (2.21) becomes

$$\begin{aligned} \frac{1}{\lambda_n} \sum_{k=0}^n \lambda_k \mu_k &= s_n - \frac{1}{\lambda_n} \sum_{k=0}^{N-1} (\lambda_{k+1} - \lambda_k) s_k - \frac{1}{\lambda_n} \sum_{k=N}^{n-1} (\lambda_{k+1} - \lambda_k) s_k \\ &= s_n - \frac{1}{\lambda_n} \sum_{k=0}^{N-1} (\lambda_{k+1} - \lambda_k) s_k - \frac{1}{\lambda_n} \sum_{k=N}^{n-1} (\lambda_{k+1} - \lambda_k) (s_k - \tilde{\ell}) - \frac{1}{\lambda_n} \sum_{k=N}^{n-1} (\lambda_{k+1} - \lambda_k) \tilde{\ell} \\ &= s_n - \frac{1}{\lambda_n} \sum_{k=0}^{N-1} (\lambda_{k+1} - \lambda_k) s_k - \frac{1}{\lambda_n} \sum_{k=N}^{n-1} (\lambda_{k+1} - \lambda_k) (s_k - \tilde{\ell}) - \frac{\lambda_n - \lambda_N}{\lambda_n} \tilde{\ell}, \end{aligned}$$

where we used that the last sum in the second equality is a telescoping sum. The result follows by letting $n \rightarrow \infty$: the first term converges to $\tilde{\ell}$ and cancels with the last term. The second term converges to zero since the summation is over a fixed number N of terms. For the third term, since $\{\lambda_n\}_n$ is monotone and non-decreasing, it can be bounded by $\frac{\epsilon}{\lambda_n} (\lambda_n - \lambda_N) \leq \epsilon$. \square

LEMMA 2.3. *For the APM (2.18)–(2.19), we have that*

(a) *The sequences $\{\|e^i\|\}_{i \in \mathbb{N}}$, $\{\|p^n\|\}_{n \geq 0}$, $\{\|\tilde{p}^n\|\}_{n \geq 1}$ are non-increasing, and the last two converge to zero.*

(b) *The series $\sum_{k=0}^{\infty} \|p^k\|^2$ and $\sum_{k=1}^{\infty} \|\tilde{p}^k\|^2$ are convergent.*

(c) *The sequences $\{\|p^n\| \sum_{k=0}^n \|p^k\|\}_{n \geq 0}$ and $\{\|\tilde{p}^n\| \sum_{k=1}^n \|\tilde{p}^k\|\}_{n \geq 1}$ converge to zero.*

Proof. Part (a). By using (2.20) and the fact that projectors onto subspaces are non-expansive [9, Theorem 4.3-1(d)], we obtain

$$\|e^{2n+1}\| = \|\mathbb{P}_{V_1^\perp} e^{2n}\| \leq \|e^{2n}\| = \|\mathbb{P}_{V_2^\perp} e^{2n-1}\| \leq \|e^{2n-1}\|,$$

which means that $\{\|e^i\|\}_{i \in \mathbb{N}}$ is monotone and non-increasing. Since it is also non-negative and bounded from below ($\|e^i\| \geq 0$), this sequence converges to some non-negative limit ℓ .

Next, we show that $\{\|p^n\|\}_{n \geq 0}$ and $\{\|\tilde{p}^n\|\}_{n \geq 1}$ are non-increasing and converge to zero. To do so, notice that (2.19) and (2.20) imply that $p^n \in V_1$ and $e^{2n+1}, e^{2n-1} \in V_1^\perp$, and observe that

$$-\tilde{p}^n = e^{2n} - e^{2n-1} = e^{2n+1} + p^n - e^{2n-1}.$$

²We use a summation by parts formula in the form of a Newton series, that is, $\sum_{k=0}^n a_k b_k = a_n \sum_{k=0}^n b_k - \sum_{j=0}^{n-1} ((a_{j+1} - a_j) \sum_{k=0}^j b_k)$.

Therefore, we have

$$(2.22) \quad \|p^n\|^2 = (p^n, -\tilde{p}^n) - (p^n, e^{2n+1} - e^{2n-1}) = -(p^n, \tilde{p}^n).$$

Similarly we get $\|\tilde{p}^n\|^2 = -(\tilde{p}^n, p^{n-1})$. From this equation, using the Cauchy-Schwarz inequality, we obtain that $\|\tilde{p}^n\| \leq \|p^{n-1}\|$. Similarly, from (2.22), we have $\|p^n\| \leq \|\tilde{p}^n\|$. By applying these two inequalities recursively, we get

$$(2.23) \quad \|p^0\| \geq \|\tilde{p}^1\| \geq \|p^1\| \geq \|\tilde{p}^2\| \geq \|p^2\| \geq \dots$$

Hence, $\{\|p^n\|\}_{n \geq 0}$ and $\{\|\tilde{p}^n\|\}_{n \geq 1}$ are non-increasing. To show that these sequences converge to zero, we use (2.18), (2.19), and (2.20) to obtain that

$$(2.24) \quad \begin{aligned} \|p^k\|^2 + \|\tilde{p}^k\|^2 &= \|e^{2k}\|^2 - \|e^{2k+1}\|^2 + \|e^{2k-1}\|^2 - \|e^{2k}\|^2 \\ &= \|e^{2k-1}\|^2 - \|e^{2k+1}\|^2. \end{aligned}$$

Equation (2.23) and (2.24) imply that $0 \leq \|p^k\| \leq \|\tilde{p}^k\| \leq \sqrt{\|e^{2k-1}\|^2 - \|e^{2k+1}\|^2}$. Recalling that $\{\|e^i\|\}_{i \in \mathbb{N}}$ converges to some non-negative limit ℓ , we have for the difference $\|e^{2k-1}\|^2 - \|e^{2k+1}\|^2 \rightarrow \ell^2 - \ell^2 = 0$ as $k \rightarrow \infty$. Hence $\|p^k\| \rightarrow 0$ and $\|\tilde{p}^k\| \rightarrow 0$ as $k \rightarrow \infty$.

Part (b). Consider the partial sums $s_n := \sum_{k=0}^n \|p^k\|^2$ and $\tilde{s}_n := \sum_{k=1}^n \|\tilde{p}^k\|^2$. They are two monotone non-decreasing sequences. Hence, to show that they converge, it suffices to show that they are bounded from above. To do so, we sum (2.24) over k to get

$$\sum_{k=1}^n (\|p^k\|^2 + \|\tilde{p}^k\|^2) = \sum_{k=1}^n (\|e^{2k-1}\|^2 - \|e^{2k+1}\|^2).$$

Using this equation and noticing that its right-hand side is a telescoping sum, we get

$$s_n + \tilde{s}_n = \|p^0\|^2 + \sum_{k=1}^n (\|e^{2k-1}\|^2 - \|e^{2k+1}\|^2) \leq 2\|e^0\|^2 - \|e^{2n+1}\|^2 \leq 2\|e^0\|^2 < \infty,$$

where we used that $\|p^0\| = \|\mathbb{P}_{V_1} e^0\| \leq \|e^0\|$. Hence the claim follows.

Part (c). To prove the claim we use Lemma 2.2. We show the statement only for $\|p^n\| \sum_{k=0}^n \|p^k\|$ as the proof is the same for $\|\tilde{p}^n\| \sum_{k=1}^n \|\tilde{p}^k\|$. Define the sequences $\mu_n := \|p^n\|^2$ and $\lambda_n := \frac{1}{\|p^n\|}$. Now, we have that $\sum_{k=0}^{\infty} \mu_k = \sum_{k=0}^{\infty} \|p^k\|^2$, which converges according to Part (b). Hence, by Kronecker's Lemma 2.2, we have that

$$0 = \lim_{n \rightarrow \infty} \frac{1}{\lambda_n} \sum_{k=0}^n \lambda_k \mu_k = \lim_{n \rightarrow \infty} \|p^n\| \sum_{k=0}^n \|p^k\|,$$

which concludes our proof. \square

We are now ready to prove that the APM converges. This is done in Theorem 2.4, which shows convergence of the APM in the same sense as used by von Neumann in [36, Theorem 13.7 page 55]. We follow the proof given in [16] providing extra details; see also [11].

THEOREM 2.4. *For any $e^0 \in V$, the sequence $\{e^i\}_{i \in \mathbb{N}}$ generated by (2.18) and (2.19) (or equivalently (2.20)) converges strongly to $e^0 - q$, where q is the best approximation to e^0 in $\overline{V_1 + V_2}$ with respect to (\cdot, \cdot) .*

Proof. We define $\bar{p}^n := \sum_{k=0}^n p^k$, $\tilde{\bar{p}}^n := \sum_{k=1}^n \tilde{p}^k$ and proceed in three steps. In Step 1, we show that $e^{2n+1} = e^0 - \bar{p}^n - \tilde{\bar{p}}^n$. In Step 2, we demonstrate that $\|e^i\| \geq \text{dist}(e^0, \overline{V_1 + V_2})$. In Step 3, we prove that $e^i \rightarrow e^0 - q$ strongly in V , where q is the best approximation to e^0 in $\overline{V_1 + V_2}$.

Step 1. Using (2.18) and (2.19), we have

$$\begin{aligned} e^0 - \bar{p}^n - \bar{\bar{p}}^n &= e^0 - p^0 - \sum_{k=1}^n (p^k + \bar{p}^k) = e^0 - p^0 - \sum_{k=1}^n (e^{2k} - e^{2k+1} + e^{2k-1} - e^{2k}) \\ &= e^0 - p^0 + \sum_{k=1}^n (e^{2k+1} - e^{2k-1}), \end{aligned}$$

where the sum on the right-hand side is a telescoping sum. Therefore, recalling from (2.18) that $e^1 = e^0 - p^0$, we obtain

$$(2.25) \quad e^0 - \bar{p}^n - \bar{\bar{p}}^n = e^0 - p^0 + e^{2n+1} - e^1 = e^{2n+1}.$$

Step 2. By (2.25) in Step 1, we have that $e^0 - e^{2n+1} = \bar{p}^n + \bar{\bar{p}}^n$. Since $\bar{p}^n \in V_1$ and $\bar{\bar{p}}^n \in V_2$, we deduce that $e^0 - e^{2n+1} \in V_1 + V_2 \subset \overline{V_1 + V_2}$. Now, recalling that

$$\text{dist}(e^0, \overline{V_1 + V_2}) \leq \|e^0 - v\| \quad \forall v \in \overline{V_1 + V_2},$$

we choose $v = e^0 - e^{2n+1}$ to get

$$(2.26) \quad \text{dist}(e^0, \overline{V_1 + V_2}) \leq \|e^{2n+1}\|.$$

Since the sequence is non-increasing ($\|e^{2n+1}\| \leq \|e^{2n}\|$ for any $n \geq 0$ by Lemma 2.3), the claim follows.

Step 3. By Lemma 2.3 the sequence $\{\|e^i\|\}_{i \in \mathbb{N}}$ converges to some limit $\ell \geq 0$. We distinguish now the two cases $\ell > 0$ and $\ell = 0$. Assume that $\ell > 0$. Take an element $v^* \in V_1 + V_2$ such that $\|e^0 - v^*\| \leq \text{dist}(e^0, \overline{V_1 + V_2}) + \epsilon/3$, where ϵ is an arbitrary positive constant, and set $v^* = v_1^* + v_2^*$, where $v_1^* \in V_1$ and $v_2^* \in V_2$. Now, we recall that $\bar{p}^n = \sum_{k=1}^n p^k$, $\bar{\bar{p}}^n = \sum_{k=0}^n \bar{p}^k$ and notice that

$$|(p^n, v_2^*)| \leq \|p^n\| \|v_2^*\| \quad \text{and} \quad |(p^n, \bar{\bar{p}}^n)| \leq \|p^n\| \|\bar{\bar{p}}^n\| \leq \|\bar{p}^n\| \sum_{k=1}^n \|\bar{p}^k\|,$$

where we used that $\|p^n\| \leq \|\bar{p}^n\|$ because of (2.23). By Lemma 2.3, both sequences $\{(p^n, v_2^*)\}_n$ and $\{(p^n, \bar{\bar{p}}^n)\}_n$ converge to zero as $n \rightarrow \infty$. Therefore, since $\ell > 0$, there exists an $N \in \mathbb{N}^+$ such that for all $n \geq N$, we have

$$(2.27) \quad |(p^n, v_2^*)| < \frac{\epsilon \|e^{2n+1}\|}{3} \quad \text{and} \quad |(p^n, \bar{\bar{p}}^n)| < \frac{\epsilon \|e^{2n+1}\|}{3}.$$

Using (2.25) from Step 1, we have

$$\|e^{2n+1}\|^2 = (e^{2n+1}, e^{2n+1}) = (e^{2n+1}, e^0 - \bar{p}^n - \bar{\bar{p}}^n),$$

and recalling that $e^{2n+1} \in V_1^\perp$, $v_1^* \in V_1$, and $\bar{p}^n \in V_1$, we obtain

$$\begin{aligned} \|e^{2n+1}\|^2 &= (e^{2n+1}, e^0 - v^*) + (e^{2n+1}, v^* - \bar{p}^n - \bar{\bar{p}}^n) \\ &\leq \|e^{2n+1}\| \|e^0 - v^*\| + (e^{2n+1}, v_2^* + v_1^* - \bar{p}^n - \bar{\bar{p}}^n) \\ &= \|e^{2n+1}\| \|e^0 - v^*\| + (e^{2n} - p^n, v_2^* - \bar{\bar{p}}^n). \end{aligned}$$

Since $e^{2n} \in V_2^\perp$, $v_2^* \in V_2$, and $\bar{\bar{p}}^n \in V_2$, we get

$$\begin{aligned} \|e^{2n+1}\|^2 &\leq \|e^{2n+1}\| \|e^0 - v^*\| + (-p^n, v_2^* - \bar{\bar{p}}^n) \\ &\leq \|e^{2n+1}\| \|e^0 - v^*\| + |(p^n, v_2^*)| + |(p^n, \bar{\bar{p}}^n)|. \end{aligned}$$

Recalling that $\|e^0 - v^*\| \leq \text{dist}(e^0, \overline{V_1 + V_2}) + \epsilon/3$, dividing by $\|e^{2n+1}\|$, and using (2.27), we obtain

$$\|e^{2n+1}\| \leq \text{dist}(e^0, \overline{V_1 + V_2}) + \frac{\epsilon}{3} + \frac{|(p^n, v_2^*)|}{\|e^{2n+1}\|} + \frac{|(p^n, \tilde{p}^n)|}{\|e^{2n+1}\|} < \text{dist}(e^0, \overline{V_1 + V_2}) + \epsilon,$$

which holds for all $n > N$. Combining this estimate with (2.26) from Step 2 and the fact that $\{\|e^i\|\}_{i \in \mathbb{N}}$ is non-increasing, we have

$$\text{dist}(e^0, \overline{V_1 + V_2}) \leq \|e^i\| < \text{dist}(e^0, \overline{V_1 + V_2}) + \epsilon,$$

where $\epsilon > 0$ is arbitrary. Therefore, we have that $\|e^i\| \rightarrow \text{dist}(e^0, \overline{V_1 + V_2})$.

Since $\text{dist}(e^0, \overline{V_1 + V_2}) = \inf_{v \in \overline{V_1 + V_2}} \|e^0 - v\|$, the sequence $\{\|e^i\|\}_{i \in \mathbb{N}}$ is a minimizing sequence for the map $v \mapsto \|e^0 - v\|$. Therefore, the strong convergence $e^n \rightarrow e^0 - q$ follows from the uniqueness of the point $q \in \overline{V_1 + V_2}$; see, e.g., [30]. In fact, since $\overline{V_1 + V_2}$ is a closed subspace of the Hilbert space V , by the projection theorem, q is the unique solution to

$$\underset{v \in \overline{V_1 + V_2}}{\text{minimize}} \quad \|e^0 - v\|,$$

for a given $e^0 \in V$; see, e.g., [9, Theorem 4.3-1].

Consider now the case $\ell = 0$. Since $\text{dist}(e^0, \overline{V_1 + V_2}) \leq \|e^i\|$, we conclude that $\text{dist}(e^0, \overline{V_1 + V_2}) = 0$, which implies that $e^0 \in \overline{V_1 + V_2}$. Therefore, the above proof holds noticing that v^* can be chosen such that $\|e^0 - v^*\| < \epsilon/3$ and deleting $\|e^{2n+1}\|$ in (2.27). In this case, we have that $q = e^0$ and $e^i \rightarrow e^0 - q = 0$. \square

The following corollary will be used in the next section to prove convergence of the Schwarz method.

COROLLARY 2.5. *For any $e^0 \in V$ we have that*

$$\lim_{n \rightarrow \infty} \left(\mathbb{P}_{V_1^\perp} \mathbb{P}_{V_2^\perp} \right)^n e^0 = \mathbb{P}_{V_1^\perp \cap V_2^\perp} e^0.$$

Proof. By (2.20) and Theorem 2.4 we know that

$$\lim_{n \rightarrow \infty} e^n = \lim_{n \rightarrow \infty} \left(\mathbb{P}_{V_1^\perp} \mathbb{P}_{V_2^\perp} \right)^n e^0 = e^0 - q,$$

where $q = \mathbb{P}_{\overline{V_1 + V_2}} e^0$. Hence, using that

$$I - \mathbb{P}_{\overline{V_1 + V_2}} = \mathbb{P}_{\overline{V_1 + V_2}^\perp} \quad \text{and} \quad \overline{V_1 + V_2}^\perp = (V_1 + V_2)^\perp = V_1^\perp \cap V_2^\perp,$$

cf. [21, Section 4.2, page 221], we obtain

$$e^0 - q = e^0 - \mathbb{P}_{\overline{V_1 + V_2}} e^0 = \mathbb{P}_{\overline{V_1 + V_2}^\perp} e^0 = \mathbb{P}_{V_1^\perp \cap V_2^\perp} e^0,$$

which is our claim. \square

In an infinite-dimensional framework, the sum of two closed subspaces is not necessarily closed; see, e.g., [18, page 29]. Moreover, as we will see in Theorem 2.11, the fact that the sum $V_1 + V_2$ is closed allows us to prove geometric convergence of the APM. If $V_1 + V_2$ is not closed, then it is shown in [16, pages 312-313 and Theorem 4.2] that the APM can converge arbitrarily slowly; see also [11, 25] and the references therein for further discussion. We now recall the following result that gives equivalent conditions for $V_1 + V_2$ to be closed [15, 20]:

THEOREM 2.6. *Let V_1 and V_2 be two closed subspaces of a Hilbert space V . Then the following statements are equivalent:*

(a) $V_1 + V_2$ is closed.

(b) There exists a constant $C > 0$ such that for every $v \in V_1 + V_2$ there is a representation $v = v_1 + v_2$ with $v_1 \in V_1$ and $v_2 \in V_2$ such that

$$(2.28) \quad \max\{\|v_1\|, \|v_2\|\} \leq C\|v\|.$$

(c) $V_1^\perp + V_2^\perp$ is closed.

Proof. (a) \Rightarrow (b): The proof is a standard result that follows from the open mapping theorem; see, e.g., [32, Theorem 5.20 and Corollary 2.12] and [9, Theorem 5.6-2,4]: since the sum $V_1 + V_2$ is closed, the space $\tilde{V} := V_1 + V_2$ is a Banach space. Define a vector space $W := V_1 \times V_2$ that is the set of all ordered pairs (v_1, v_2) with $v_1 \in V_1$ and $v_2 \in V_2$ endowed with component-wise addition and scalar multiplication. The space W is also endowed with the norm $\|\cdot\|$ defined by $\|(v_1, v_2)\| := \|v_1\| + \|v_2\|$. Since V_1 and V_2 are closed subspaces of a Banach space, they are complete. Therefore W is a Banach space as well. Now, we consider the linear map $\Lambda : W \rightarrow \tilde{V}$ defined by $\Lambda(v_1, v_2) = v_1 + v_2$. To apply the open mapping theorem in the form given by [32, Corollary 2.12 (c), page 50], we need to show that Λ is continuous and bijective. The map Λ is continuous because $\|\Lambda(v_1, v_2)\| = \|v_1 + v_2\| \leq \|(v_1, v_2)\|$ and maps W onto \tilde{V} . In the case that $V_1 \cap V_2 = \{0\}$, that is, V_1 and V_2 are linearly independent, it holds that $\Lambda(v_1, v_2) = 0$ if and only if $(v_1, v_2) = (0, 0)$, which implies the injectivity of Λ . However, if $V_1 \cap V_2 \neq \{0\}$, then the map Λ is not injective. In this case, we define $K := \ker \Lambda = \{(v, -v) : v \in V_1 \cap V_2\}$, which is a closed subspace of W . We introduce the map $\tilde{\Lambda} : W/K \rightarrow \tilde{V}$, which is bijective and continuous. Notice that the quotient space W/K is still a Banach space; see, e.g., [9, Theorem 3.6-5]. Hence, by the open mapping theorem [32, Corollary 2.12 (c), page 50], we have that there exists a $C > 0$ such that $\|v_1\| + \|v_2\| = \|(v_1, v_2)\| \leq C\|v_1 + v_2\|$ for all $(v_1, v_2) \in W/K$.

(b) \Rightarrow (a): The statement is less standard. It is claimed in [16, Lemma 3.1], but the proof is not given. To the best of our knowledge, the only references where the proof is given are [20, pages 218-219] and [37]³. In particular, the result proved in [37] is a generalization of the result presented in [24]. We follow the proof in [37]: let $\{w_n\}_n$ be a sequence in $V_1 + V_2$ such that $w_n \rightarrow w \in V$ as $n \rightarrow \infty$. Since V is a Banach space, the sequence $\{w_n\}_n$ is a Cauchy sequence. Therefore, we can choose a subsequence $\{w_k\}_k$ (with $k = k(n)$) such that

$$(2.29) \quad \|w_{k+1} - w_k\| \leq \frac{1}{C2^k},$$

for $k = 1, 2, \dots$, where C is the constant in (2.28). Now, since $w_1 - w_2 \in V_1 + V_2$, there exist $z_1 \in V_1$ and $\tilde{z}_1 \in V_2$ such that $w_1 - w_2 = z_1 + \tilde{z}_1$. We can then use (2.28) for $w_1 - w_2$ and (2.29) to write

$$\max\{\|z_1\|, \|\tilde{z}_1\|\} \leq C\|w_1 - w_2\| \leq \frac{1}{2}.$$

Similarly, since $w_2 - w_3 \in V_1 + V_2$, there exist vectors $z_2 \in V_1$ and $\tilde{z}_2 \in V_2$ such that $w_2 - w_3 = z_2 + \tilde{z}_2$. Again one can deduce from (2.28) and (2.29) that

$$\max\{\|z_2\|, \|\tilde{z}_2\|\} \leq C\|w_2 - w_3\| \leq \frac{1}{2^2}.$$

³One of the referees pointed out that this result can be obtained by combining Theorem 2.10, Corollary 2.11, Theorem 2.16, and elaborating Remark 7 and Exercise 2.16 in [3].

Continuing in this way, we obtain two sequences $\{z_k\}_k$ and $\{\tilde{z}_k\}_k$ such that

$$w_{k+1} - w_k = z_k + \tilde{z}_k \text{ and } \max\{\|z_k\|, \|\tilde{z}_k\|\} \leq \frac{1}{2^k},$$

for $k = 1, 2, \dots$. Define now the partial sums $m_j := \sum_{k=1}^j \|z_k\|$ and $n_j := \sum_{k=1}^j \|\tilde{z}_k\|$. They form two sequences in j that are monotonically non-decreasing and bounded as

$$\max\{m_j, n_j\} = \max\left\{\sum_{k=1}^j \|z_k\|, \sum_{k=1}^j \|\tilde{z}_k\|\right\} \leq \sum_{k=1}^j \frac{1}{2^k} \leq \sum_{k=0}^{\infty} \frac{1}{2^k} = 2.$$

Therefore, we obtain the absolute convergence $\sum_{k=1}^{\infty} \|z_k\| < \infty$ and $\sum_{k=1}^{\infty} \|\tilde{z}_k\| < \infty$, which ensures the existence of $z, \tilde{z} \in V$ such that $z = \sum_{k=1}^{\infty} z_k$ and $\tilde{z} = \sum_{k=1}^{\infty} \tilde{z}_k$. Since $\sum_{k=1}^j z_k \in V_1$ and $\sum_{k=1}^j \tilde{z}_k \in V_2$ and the two subspaces V_1 and V_2 are closed, we have that $z \in V_1$ and $\tilde{z} \in V_2$. Now, we recall that $w_k \rightarrow w$ and write that

$$w - w_1 = \sum_{k=1}^{\infty} (w_{k+1} - w_k) = \sum_{k=1}^{\infty} (z_k + \tilde{z}_k) = z + \tilde{z}.$$

Finally, we conclude by noticing that $w_1 \in V_1 + V_2$ and $z + \tilde{z} \in V_1 + V_2$ imply that $w = w_1 + z + \tilde{z} \in V_1 + V_2$.

(a) \Leftrightarrow (c): see [12, Lemma 11]. \square

Now, we are going to prove that the APM (2.18)–(2.19) converges geometrically. To do so, we recall the notions of angles between subspaces of a Hilbert space; see, e.g., [12]. Notice that the cosine of such angles is also called inclination [16].

DEFINITION 2.7. *Let V_1 and V_2 be two closed subspaces of a Hilbert space V with inner product (\cdot, \cdot) and norm $\|\cdot\|$. The angle between V_1 and V_2 is the angle $\alpha(V_1, V_2) \in [0, \pi/2]$ whose cosine is defined by*

$$\text{incl}(V_1, V_2) := \sup\{(v_1, v_2) : v_1 \in V_1 \cap (V_1 \cap V_2)^\perp, v_2 \in V_2 \cap (V_1 \cap V_2)^\perp, \|v_1\| \leq 1, \|v_2\| \leq 1\}.$$

The minimal angle between V_1 and V_2 is the angle $\alpha_0(V_1, V_2) \in [0, \pi/2]$ whose cosine is defined by

$$\text{incl}_0(V_1, V_2) := \sup\{(v_1, v_2) : v_1 \in V_1, v_2 \in V_2, \|v_1\| \leq 1, \|v_2\| \leq 1\}.$$

Notice that angle and minimal angle of V_1 and V_2 coincide if $V_1 \cap V_2 = \{0\}$. However, the two angles do not coincide in general. This is clarified by the following example.

EXAMPLE 2.8. Let V be a finite-dimensional real Hilbert space defined as $V := \text{span}\{b_1, b_2, b_3\}$, where the linear-independent vectors b_j satisfy $b_j \perp b_k$ for $k \neq j$. Define the two subspaces $V_1 := \text{span}\{b_1, b_2\}$ and $V_2 := \text{span}\{b_2, b_3\}$. The vectors b_j and the subspaces V_1 and V_2 are depicted in Figure 2.3. To compute $\text{incl}_0(V_1, V_2)$, we take two arbitrary vectors $b_{V_1} \in V_1$ and $b_{V_2} \in V_2$. Obviously, we have $b_{V_1} = \alpha_1 b_1 + \alpha_2 b_2$ and $b_{V_2} = \beta_2 b_2 + \beta_3 b_3$, where α_j and β_j are real coefficients. Hence, we get for the scalar product $(b_{V_1}, b_{V_2}) = \alpha_2 \beta_2 \|b_2\|^2$. Therefore, to compute $\text{incl}_0(V_1, V_2)$ it suffices to maximize the product $\alpha_2 \beta_2$ under the constraints $\alpha_2 \leq 1/\|b_2\|$ and $\beta_2 \leq 1/\|b_2\|$. This is solved by $\alpha_2 = \beta_2 = 1/\|b_2\|$, which leads to $\text{incl}_0(V_1, V_2) = 1$. The corresponding angle is $\alpha_0(V_1, V_2) = 0$.

To compute $\text{incl}(V_1, V_2)$, observe that $V_1 \cap V_2 = \text{span}\{b_2\}$ and $(V_1 \cap V_2)^\perp = \text{span}\{b_1, b_3\}$. Hence, $V_1 \cap (V_1 \cap V_2)^\perp = \text{span}\{b_1\}$ and $V_2 \cap (V_1 \cap V_2)^\perp = \text{span}\{b_3\}$. Since $b_1 \perp b_3$, we have that $V_1 \cap (V_1 \cap V_2)^\perp \perp V_2 \cap (V_1 \cap V_2)^\perp$ as well, and hence $(b_{V_1}, b_{V_2}) = 0$ for any

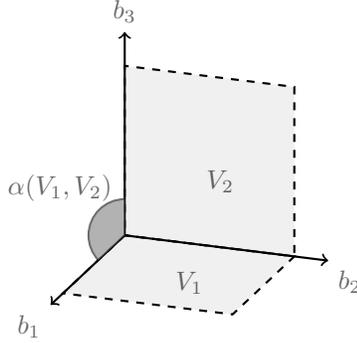


FIG. 2.3. The subspaces $V_1 = \text{span}\{b_1, b_2\}$ and $V_2 = \text{span}\{b_2, b_3\}$ and the angle $\alpha(V_1, V_2) = \pi/2$.

$b_{V_1} \in V_1 \cap (V_1 \cap V_2)^\perp$ and $b_{V_2} \in V_2 \cap (V_1 \cap V_2)^\perp$. This implies that $\text{incl}(V_1, V_2) = 0$ with the corresponding angle $\alpha(V_1, V_2) = \pi/2$. We see that in this example $\text{incl}(V_1, V_2) < \text{incl}_0(V_1, V_2)$ and $\alpha_0(V_1, V_2) < \alpha(V_1, V_2)$.

Lemma 2.9 summarizes several results characterizing incl and incl_0 and their relationship with the projection operators onto V_1 and V_2 . The proofs of the statements given in Lemma 2.9 can be found in [12] and the references therein.

LEMMA 2.9. In the context of Definition 2.7, denoting by $\|\cdot\|'$ the norm of all linear operators from V to V , we have

- (a) $0 \leq \text{incl}(V_1, V_2) \leq \text{incl}_0(V_1, V_2) \leq 1$.
- (b) $\text{incl}(V_1, V_2) = \text{incl}(V_2, V_1)$ and $\text{incl}_0(V_1, V_2) = \text{incl}_0(V_2, V_1)$.
- (c) If $V_1 \cap V_2 = \{0\}$, then $\text{incl}(V_1, V_2) = \text{incl}_0(V_1, V_2)$. On the other hand, if $V_1 \cap V_2 \neq \{0\}$, then $\text{incl}_0(V_1, V_2) = 1$.
- (d) $\text{incl}_0(V_1, V_2) = \|\mathbb{P}_{V_1} \mathbb{P}_{V_2}\|'$ and $\text{incl}(V_1, V_2) = \|\mathbb{P}_{V_1} \mathbb{P}_{V_2} - \mathbb{P}_{V_1 \cap V_2}\|'$.
- (e) $\text{incl}_0(V_1, V_2) = 0$ if and only if $V_1 \perp V_2$.
- (f) $\text{incl}(V_1, V_2) = 0$ if and only if \mathbb{P}_{V_1} and \mathbb{P}_{V_2} commute.
- (g) $\text{incl}_0(V_1, V_2) < 1$ if and only if $V_1 \cap V_2 = \{0\}$ and $V_1 + V_2$ is closed.
- (h) $\text{incl}(V_1, V_2) < 1$ if and only if $V_1 + V_2$ is closed (or $V_1^\perp + V_2^\perp$ is closed).
- (i) $\text{incl}(V_1, V_2) = \text{incl}(V_1^\perp, V_2^\perp)$, and if $V_1 \cap V_2 = \{0\}$ and $V_1 + V_2 = V$, then $\text{incl}_0(V_1, V_2) = \text{incl}_0(V_1^\perp, V_2^\perp)$ ⁴.
- (j) $\|(\mathbb{P}_{V_1} \mathbb{P}_{V_2})^n - \mathbb{P}_{V_1 \cap V_2}\|' = \text{incl}(V_1, V_2)^{2n-1}$.

Some of the results of Lemma 2.9 are illustrated by the following example.

EXAMPLE 2.10. Consider the same spaces V , V_1 , and V_2 as in Example 2.8. We have already seen that $\text{incl}(V_1, V_2) \leq \text{incl}_0(V_1, V_2)$ in agreement with Lemma 2.9(a). The orthogonal complements of V_1 and V_2 are $V_1^\perp = \text{span}\{b_3\}$ and $V_2^\perp = \text{span}\{b_1\}$. Hence, the angle between V_1^\perp and V_2^\perp is $\alpha(V_1^\perp, V_2^\perp) = \frac{\pi}{2}$, whose cosine is $\text{incl}(V_1^\perp, V_2^\perp) = 0$. Recalling from Example 2.8 that $\text{incl}(V_1, V_2) = 0$, we get $\text{incl}(V_1, V_2) = \text{incl}(V_1^\perp, V_2^\perp)$. Recalling the definition of incl_0 , we see that $\text{incl}_0(V_1^\perp, V_2^\perp) = 0$, which is also obtained by Lemma 2.9(e). Further, by Example 2.8 we have that $0 = \text{incl}_0(V_1^\perp, V_2^\perp) \leq \text{incl}_0(V_1, V_2) = 1$ in agreement with Lemma 2.9(e).

The geometric convergence of the APM is proved in the following theorem.

THEOREM 2.11. Assume that $V_1 + V_2$ is closed. Then for any $e^0 \in V$ we have

- (a) $\|e^{2n+1} - \mathbb{P}_{V_1^\perp \cap V_2^\perp} e^0\| \leq \text{incl}(V_1, V_2)^{2n-1} \|e^0\|$ with $\text{incl}(V_1, V_2) < 1$.

⁴The statement $\text{incl}_0(V_1, V_2) = \text{incl}_0(V_1^\perp, V_2^\perp)$ is not in general true for any V_1, V_2 closed subspaces of V ; see Theorem 2.15 in [12] and the remark thereafter.

(b) If $V_1 + V_2 = V$, then $\|e^{2n+1} - \mathbb{P}_{V_1^\perp \cap V_2^\perp} e^0\| \leq \text{incl}_0(V_1, V_2)^{2n-1} \|e^0\|$ with $\text{incl}_0(V_1, V_2) < 1$.

(c) If $V_1 \perp V_2$ or if \mathbb{P}_{V_1} and \mathbb{P}_{V_2} commute, then the APM converges in two iterations, that is, $\|e^2 - \mathbb{P}_{V_1^\perp \cap V_2^\perp} e^0\| = 0$.

Proof. Since V_1 and V_2 are closed subspaces of V , Theorems 2.4 and Corollary 2.5 guarantee convergence of the APM in the sense that $\|e^{2n+1} - \mathbb{P}_{V_1^\perp \cap V_2^\perp} e^0\| \rightarrow 0$.

To prove (a) we use Lemma 2.9(j) to obtain

$$\begin{aligned} \|e^{2n+1} - \mathbb{P}_{V_1^\perp \cap V_2^\perp} e^0\| &= \|(\mathbb{P}_{V_1^\perp} \mathbb{P}_{V_2^\perp})^n e^0 - \mathbb{P}_{V_1^\perp \cap V_2^\perp} e^0\| \\ &\leq \|(\mathbb{P}_{V_1^\perp} \mathbb{P}_{V_2^\perp})^n - \mathbb{P}_{V_1^\perp \cap V_2^\perp}\| \|e^0\| \\ &= \text{incl}(V_1^\perp, V_2^\perp)^{2n-1} \|e^0\|. \end{aligned}$$

Since $V_1 + V_2$ is closed, the statement (a) follows by Lemma 2.9(i) and (h) and Theorem 2.6(c).

To prove (b), since $V_1 + V_2 = V$, we have $V_1 + V_2 = \overline{V_1 + V_2}$, and hence, Theorem 2.6 guarantees that $V_1^\perp + V_2^\perp$ is closed. Moreover, we have

$$(2.30) \quad V_1^\perp \cap V_2^\perp = \overline{V_1 + V_2}^\perp = (V_1 + V_2)^\perp = V^\perp = \{0\}.$$

Then, the statement follows by (a) together with Lemma 2.9(a) and (g).

Statement (c) follows from (a), using Lemma 2.9(f) together with the fact that \mathbb{P}_{V_1} and \mathbb{P}_{V_2} commute. Alternatively, it follows from (b) using Lemma 2.9(e) together with the fact that $V_1 \perp V_2$. \square

Notice that because of Lemma 2.9(j), the estimate given in Theorem 2.11(a) is sharp. However, the computation of incl and incl_0 is not easy in practice. For this reason, an alternative proof of convergence of the APM is discussed in Section 2.2. In the particular case of domain decomposition methods, such an alternative proof provides a practical approach to estimate incl by exploiting the geometric setting of the domain decomposition and the corresponding functional properties of the subspaces V_1 and V_2 .

We conclude this section remarking that most of the presented convergence results can be extended to the case of more than two subspaces; see, e.g., [11, 12, 19]. Moreover, also the case that V_1 and V_2 are convex and closed subsets of V is studied; see, e.g., [5, 11] and the references therein.

2.2. P.-L. Lions's proof of convergence. The results presented in Section 2.1 can be successfully applied to the alternating Schwarz method introduced at the beginning of Section 2. However, even if one can show that the sum of the two subspaces V_1 and V_2 is closed, by using for example Theorem 2.6, it is difficult to estimate the inclination $\text{incl}(V_1^\perp, V_2^\perp)$ that represents, according to Theorem 2.11 and Lemma 2.9, the contraction factor governing the geometric convergence of the APM.

In the field of domain-decomposition and Schwarz type methods, an answer to this problem was provided by P.-L. Lions in [26]; see also the graphic representation in Figure 2.4. In fact P.-L. Lions uses the hypothesis $V_1 + V_2 = V$ to prove an intermediate result, namely Lemma I.1 in [26], in which the existence of a positive constant c_0 is shown that influences the contraction constant obtained in the main convergence result, that is, Theorem I.1 in [26]. After the proofs of these two results, the author works in the functional analytic setting related to elliptic PDEs and provides arguments that allow to estimate c_0 and the contraction factor. At this point, we recall Lemma 2.9(h), which relates the condition $V_1 + V_2 = V$ with the fact that $\text{incl}(V_1, V_2) < 1$, and we remark that, even if not explicitly written in [26], the author is avoiding a direct estimate of the inclination $\text{incl}(V_1, V_2)$ by exploiting the functional structure of the domain decomposition setting. At this point, we emphasize that P.-L. Lions uses

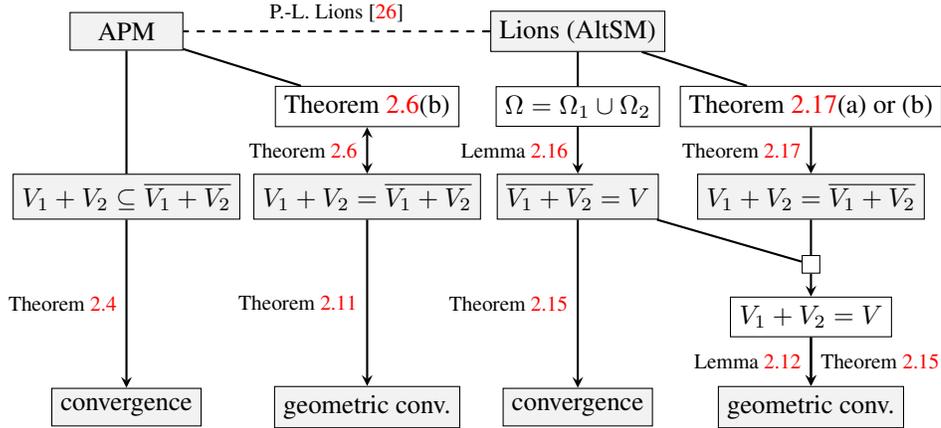


FIG. 2.4. Flow-chart describing the convergence analysis of the APM (left part) and the AltSM performed by P.-L. Lions in [26] (right part). Note that arrows of the types “ \rightarrow ” and “ \leftrightarrow ” correspond to implication and equivalence. The connection between APM and AltSM was made by P.-L. Lions in [26]. Despite this fact, P.-L. Lions’s convergence results of the AltSM are independent of the convergence results given in the APM literature. The only known APM result that P.-L. Lions uses (to motivate the assumptions of Theorem 2.15) is Corollary 2.5 proved for the first time by von Neumann in [36, Theorem 13.7, page 55].

Theorem 2.6 only in the direction “(a) \Rightarrow (b)” in his proofs. Then, instead of using Theorem 2.6 in the direction “(b) \Rightarrow (a)”, he derives by direct calculations a sufficient condition for $V_1 + V_2$ to be closed based on the particular geometric setting of the domain-decomposition problem. These calculations will allow us to obtain a concrete estimate of $\text{incl}(V_1, V_2)$.

In what follows, we prove Lemma I.1, Theorem I.1 and retrace P.-L. Lions’s arguments for the estimate of the contraction constant. We begin with Lemma I.1 which corresponds to the following lemma.

LEMMA 2.12. *Let V be a Hilbert space endowed with the inner product (\cdot, \cdot) , which induces the norm $\|\cdot\|$. Let V_1 and V_2 be two closed subspaces of V such that $V_1 + V_2 = V$. Then there exists a constant $c_0 \geq 1$ such that*

$$\|v\| \leq c_0 \left(\|\mathbb{P}_{V_1} v\|^2 + \|\mathbb{P}_{V_2} v\|^2 \right)^{1/2} \quad \forall v \in V.$$

Proof. The proof presented in [26] makes use of the implication “(a) \Rightarrow (b)” in Theorem 2.6: since $V_1 + V_2 = V$, there exists a constant $c_0 > 0$ such that⁵

$$(2.31) \quad \forall v \in V : \exists (v_1, v_2) \in V_1 \times V_2 \text{ with } \begin{matrix} v = v_1 + v_2 & \text{and} \\ \left(\|v_1\|^2 + \|v_2\|^2 \right)^{1/2} \leq c_0 \|v\|. \end{matrix}$$

Notice that $c_0 \geq 1$: to see this, choose, e.g., $v \in V_1$. Then $v_1 = v$ and $v_2 = 0$, hence, c_0 cannot be smaller than 1. Now, given any $a, b, c, d \in \mathbb{R}$, the Cauchy-Schwarz inequality allows us to write

$$(2.32) \quad ab + cd = \begin{bmatrix} a \\ c \end{bmatrix}^\top \begin{bmatrix} b \\ d \end{bmatrix} \leq \sqrt{a^2 + c^2} \sqrt{b^2 + d^2}.$$

⁵Notice that (2.31) is the infinite-dimensional counterpart of the ‘Stable Decomposition’ assumption considered in [34, page 40] and used to prove convergence results for additive and multiplicative Schwarz methods. Sufficient conditions related to the geometry of the domain decomposition are given in Theorem 2.17; see also Figure 2.4, which illustrates how these conditions are related to the convergence of the AltSM.

Since V_1 and V_2 are closed subspaces of V , the decompositions $V = V_1 \oplus V_1^\perp$ and $V = V_2 \oplus V_2^\perp$ hold; see, e.g., [9, Theorem 4.2-5]. Hence, we have that $v = \tilde{v}_1 + \tilde{v}_1^\perp$ with $\tilde{v}_1 \in V_1$ and $\tilde{v}_1^\perp \in V_1^\perp$, and $v = \tilde{v}_2 + \tilde{v}_2^\perp$ with $\tilde{v}_2 \in V_2$ and $\tilde{v}_2^\perp \in V_2^\perp$. Hence, we obtain

$$(2.33) \quad \begin{aligned} \|v\|^2 &= (v, v) = (v, v_1) + (v, v_2) = (\tilde{v}_1, v_1) + (\tilde{v}_2, v_2) \\ &= (\mathbb{P}_{V_1} v, v_1) + (\mathbb{P}_{V_2} v, v_2) \leq \|\mathbb{P}_{V_1} v\| \|v_1\| + \|\mathbb{P}_{V_2} v\| \|v_2\|. \end{aligned}$$

Using (2.32), the estimate (2.33) becomes

$$(2.34) \quad \|v\|^2 \leq (\|\mathbb{P}_{V_1} v\|^2 + \|\mathbb{P}_{V_2} v\|^2)^{1/2} (\|v_1\|^2 + \|v_2\|^2)^{1/2}.$$

Inserting (2.31) into (2.34), we obtain $\|v\|^2 \leq c_0 \|v\| (\|\mathbb{P}_{V_1} v\|^2 + \|\mathbb{P}_{V_2} v\|^2)^{1/2}$, which implies the claim. \square

Next, we show P.-L. Lions's convergence result. To do so, notice that Corollary 2.5 guarantees that $e^i \rightarrow \mathbb{P}_{V_1^\perp \cap V_2^\perp} e^0$. Hence, if we denote by $\tilde{e}^0 := e^0 - \mathbb{P}_{V_1^\perp \cap V_2^\perp} e^0$, then the sequence generated by the APM starting from \tilde{e}^0 (instead of e^0) is given by $\tilde{e}^i = e^i - \mathbb{P}_{V_1^\perp \cap V_2^\perp} e^0$. Hence, without loss of generality, we can assume that $V_1^\perp \cap V_2^\perp = \{0\}$. To clarify the idea behind this assumption, we provide the following two examples.

EXAMPLE 2.13. Consider $V = \mathbb{R}^3$, and assume that V_1 and V_2 are two planes in V passing through the origin such that $V_1 \neq V_2$. Then one has that $V = V_1 + V_2$. The orthogonal complements V_1^\perp and V_2^\perp are the two lines passing through the origin and orthogonal to V_1 and V_2 . It is also clear that $V_1^\perp \cap V_2^\perp = \{0\}$. Now consider any point $e^0 \in V$. At the first iteration of the APM, we have that $e^1 = \mathbb{P}_{V_1^\perp} e^0$. Then the sequence starts to “jump” between V_1^\perp and V_2^\perp and converges to $V_1^\perp \cap V_2^\perp = \{0\}$; see, e.g., Figure 2.2. In this case, the first step of the APM pushes e^0 into the subspace $V_1^\perp + V_2^\perp$ and then the sequence $\{e^i\}_{i \in \mathbb{N}}$, for $i \geq 1$, remains there and converges to the origin. This means that the APM for any $e^0 \in V$ generates a sequence converging to the origin.

EXAMPLE 2.14. Consider $V = \mathbb{R}^3$, and assume that V_1 and V_2 are two lines in V passing through the origin and lying in the x - z plane such that $V_1 \neq V_2$. Then one has that $V_1 + V_2$ is the x - z plane. The orthogonal complements V_1^\perp and V_2^\perp are the two planes passing through the origin, orthogonal to V_1 and V_2 ; see Figure 2.5. In this case $V_1^\perp \cap V_2^\perp$ coincides with the y -axis. Now consider any point $e^0 \in V$ lying outside the x - z plane. Denote by V_0 the plane passing through e_0 and parallel to the x - z plane. At the first iteration of the APM, we have that $e^1 = \mathbb{P}_{V_1^\perp} e^0$. Then the sequence starts to “jump” between the two planes V_1^\perp and V_2^\perp . In particular, one can observe that $\{e^i\}_i \subset V_0$ for $i \geq 0$. Hence e^n converges to $V_1^\perp \cap V_2^\perp \cap V_0$, which is the point \tilde{e} representing the intersection of V_0 with the y axis. In this case the APM generates a sequence that lies in V_0 and converges to \tilde{e} . It is clear that, in contrast to Example 2.13, the limit point of the sequence depends on the initial point e^0 . However, if we define the starting point $\tilde{e}^0 = e^0 - \mathbb{P}_{V_1^\perp \cap V_2^\perp} e^0$, then the APM sequence converges to the origin.

We are now ready to prove convergence of the AltSM. Recall the definition of u^{2n+1} and u^{2n} given in (2.5) and that the errors are $e^{2n+1} := u - u^{2n+1}$ and $e^{2n} := u - u^{2n}$, where u is the (weak) solution to (2.2).

THEOREM 2.15. *If $V_1^\perp \cap V_2^\perp = \{0\}$ or equivalently, if $V = \overline{V_1 + V_2}$ (recall (2.30)), then the sequence e^i converges to 0 strongly in V . If $V_1 + V_2 = V$, then there exists a positive constant $C_L(c_0) \in [0, 1)$, where c_0 is any constant satisfying (2.31) in Lemma 2.12, such that*

$$(2.35) \quad \|\mathbb{P}_{V_1^\perp} \mathbb{P}_{V_2^\perp}\|' \leq C_L(c_0),$$

where $\|\cdot\|'$ denotes the operator norm induced by $\|\cdot\|$, and therefore,

$$\|e^{2n+1}\| \leq (C_L(c_0))^n \|e^1\|, \quad \|e^{2n}\| \leq (C_L(c_0))^n \|e^0\|,$$

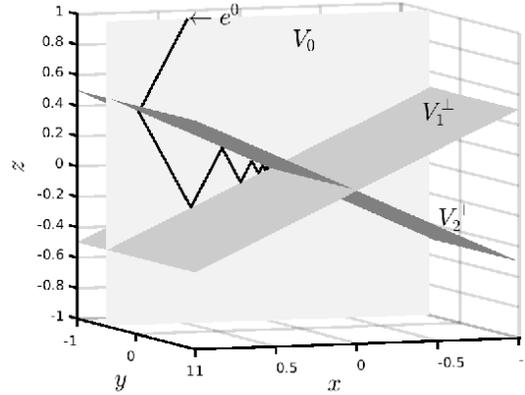


FIG. 2.5. Geometry of Example 2.14. The gray and the dark gray plane represent V_1^\perp and V_2^\perp . The light gray plane is V_0 . The black line represents the alternating projection procedure. The sequence $\{e^i\}_{i \in \mathbb{N}}$ starts in e^0 (point outside V_1^\perp and V_2^\perp and indicated by the arrow) and evolves “jumping” between V_1^\perp and V_2^\perp and remaining in V_0 .

for all $n \geq 0$. In particular, we have that⁶

$$C_L(c_0) = \sqrt{1 - \frac{1}{c_0^2}}.$$

Proof. Denote by $u^0 \in V$ the initial guess of the Schwarz method and by e^0 the corresponding error. Since $e^{2n+1} - e^{2n} = u^{2n} - u^{2n+1} \in V_1$ with $e^{2n+1} \in V_1^\perp$, and $e^{2n} - e^{2n-1} = u^{2n-1} - u^{2n} \in V_2$ with $e^{2n} \in V_2^\perp$, the Pythagorean identity implies that

$$(2.36) \quad \|e^{i+1} - e^i\|^2 = \|e^i\|^2 - \|e^{i+1}\|^2,$$

and we thus must have $\|e^i\|^2 \geq \|e^{i+1}\|^2$, i.e., the sequence $\{\|e^i\|\}_{i \in \mathbb{N}}$ is non-increasing. Therefore $\|e^i\| \rightarrow \ell$ for some $\ell \geq 0$, and (2.36) implies also that $e^{i+1} - e^i \rightarrow 0$ (and obviously $e^{i+1} - e^i \rightarrow 0$ in V). Using again the fact that $\{\|e^i\|\}_{i \in \mathbb{N}}$ is non-increasing, we obtain that the sequence $\{e^i\}_{i \in \mathbb{N}} \subset V$ is globally bounded, $\|e^i\| \leq \|e^0\|$. Since V is a Hilbert space (hence reflexive), $\{e^i\}_{i \in \mathbb{N}}$ contains weakly-convergent subsequences (in V). If for some subsequence $\{e^{i_k}\}_k$, $e^{i_k} \rightharpoonup \widehat{e}$ for some $\widehat{e} \in V$, then $e^{i_{k+1}} \rightharpoonup \widehat{e}$ as well, and thus $\widehat{e} \in V_1^\perp \cap V_2^\perp$, that is, $\widehat{e} = 0$. Since $\widehat{e} = 0$ is the unique element in $V_1^\perp \cap V_2^\perp$, all the weakly convergent subsequences converge to $\widehat{e} = 0$, and hence, the sequence $\{e^i\}_{i \in \mathbb{N}}$ itself is weakly convergent, that is, $e^i \rightharpoonup 0$. In order to show strong convergence, recalling (2.36), we write

$$\|e^{i+1}\|^2 = \|e^i\|^2 - \|e^{i+1} - e^i\|^2 = -\|e^{i+1}\|^2 + 2\langle e^{i+1}, e^i \rangle,$$

which implies that $\|e^{i+1}\|^2 = \langle e^i, e^{i+1} \rangle$. Next, using the fact that $\mathbb{P}_{V_1^\perp}$ and $\mathbb{P}_{V_2^\perp}$ are symmetric operators and recalling (2.12) and (2.13), we write (assuming without loss of generality that i is even)

$$\begin{aligned} \|e^{i+1}\|^2 &= \langle e^i, e^{i+1} \rangle = \langle \mathbb{P}_{V_2^\perp} e^{i-1}, e^{i+1} \rangle = \langle e^{i-1}, \mathbb{P}_{V_2^\perp} e^{i+1} \rangle = \langle e^{i-1}, e^{i+2} \rangle \\ &= \langle \mathbb{P}_{V_1^\perp} e^{i-2}, e^{i+2} \rangle = \langle e^{i-2}, \mathbb{P}_{V_1^\perp} e^{i+2} \rangle = \langle e^{i-2}, e^{i+3} \rangle \\ &= \dots = \langle e^{i-j}, e^{i+j+1} \rangle, \end{aligned}$$

⁶Notice how the estimate $C_L(c_0)$ of the contraction factor is similar to the one given in [34, Theorem 2.9], which corresponds to a (discrete) finite-dimensional problem.

for $j = 0, 1, \dots, n$, which for $j = i$ gives $\|e^{i+1}\|^2 = \langle e^0, e^{2i+1} \rangle$. Hence, the weak convergence of $e^{2i+1} \rightharpoonup 0$ implies that $\|e^{i+1}\| \rightarrow 0$, that is, $e^{i+1} \rightarrow 0$ (strongly in V).

Next, assume that $V_1 + V_2 = V$. Then Lemma 2.12 ensures that there exists a $c_0 > 0$ such that $\|v\| \leq c_0 (\|\mathbb{P}_{V_1} v\|^2 + \|\mathbb{P}_{V_2} v\|^2)^{1/2}$ for all $v \in V$. Hence, we have

$$(2.37) \quad \|\mathbb{P}_{V_1^\perp} v\| \leq c_0 (\|\mathbb{P}_{V_1} \mathbb{P}_{V_1^\perp} v\|^2 + \|\mathbb{P}_{V_2} \mathbb{P}_{V_1^\perp} v\|^2)^{1/2} = c_0 \|\mathbb{P}_{V_2} \mathbb{P}_{V_1^\perp} v\|,$$

where we used the fact that $\mathbb{P}_{V_1^\perp} \mathbb{P}_{V_1} v = 0$ for any $v \in V$. Moreover, the following holds:

$$(2.38) \quad \begin{aligned} \|\mathbb{P}_{V_1^\perp} v\|^2 &= \|\mathbb{P}_{V_2} \mathbb{P}_{V_1^\perp} v + \mathbb{P}_{V_2^\perp} \mathbb{P}_{V_1^\perp} v\|^2 \\ &= \|\mathbb{P}_{V_2} \mathbb{P}_{V_1^\perp} v\|^2 + \|\mathbb{P}_{V_2^\perp} \mathbb{P}_{V_1^\perp} v\|^2 + 2\langle \mathbb{P}_{V_2} \mathbb{P}_{V_1^\perp} v, \mathbb{P}_{V_2^\perp} \mathbb{P}_{V_1^\perp} v \rangle \\ &= \|\mathbb{P}_{V_2} \mathbb{P}_{V_1^\perp} v\|^2 + \|\mathbb{P}_{V_2^\perp} \mathbb{P}_{V_1^\perp} v\|^2. \end{aligned}$$

By using (2.38) and (2.37) and recalling that $c_0 \geq 1$, we get

$$\begin{aligned} \|\mathbb{P}_{V_2^\perp} \mathbb{P}_{V_1^\perp} v\|^2 &= \|\mathbb{P}_{V_1^\perp} v\|^2 - \|\mathbb{P}_{V_2} \mathbb{P}_{V_1^\perp} v\|^2 \\ &\leq \|\mathbb{P}_{V_1^\perp} v\|^2 - \frac{1}{c_0^2} \|\mathbb{P}_{V_1^\perp} v\|^2 = \left(1 - \frac{1}{c_0^2}\right) \|\mathbb{P}_{V_1^\perp} v\|^2. \end{aligned}$$

This implies that

$$\|\mathbb{P}_{V_2^\perp} \mathbb{P}_{V_1^\perp} v\| \leq C_L(c_0) \|\mathbb{P}_{V_1^\perp} v\| \leq C_L(c_0) \|v\|,$$

which gives (2.35). We can conclude the proof now by recalling (2.16) and (2.17), namely $e^{2n+1} = (\mathbb{P}_{V_1^\perp} \mathbb{P}_{V_2^\perp})^n e^1$ and $e^{2n} = (\mathbb{P}_{V_2^\perp} \mathbb{P}_{V_1^\perp})^n e^0$. \square

Note that we have seen in Theorem 2.11 (and Lemma 2.9(j)) that the contraction factor of the APM is the inclination $\text{incl}(V_1, V_2)$. The fact that P.-L. Lions is estimating incl becomes clear if we recall that $V_1^\perp \cap V_2^\perp = (V_1 + V_2)^\perp = (\overline{V_1 + V_2})^\perp = V^\perp = \{0\}$ (recall also [21, Section 4.2, page 221]) and then use (2.35) with Lemma 2.9(a), (d), and (i) to obtain that

$$(2.39) \quad \text{incl}(V_1, V_2) = \text{incl}(V_1^\perp, V_2^\perp) \leq \text{incl}_0(V_1^\perp, V_2^\perp) = \|\mathbb{P}_{V_2^\perp} \mathbb{P}_{V_1^\perp}\|' \leq C_L(c_0).$$

Note also that C_L decreases when c_0 decreases.

As discussed in the proof of Lemma 2.12, the existence of c_0 is guaranteed by the fact that $V_1 + V_2 = V$. We now discuss this condition: recall that $V = H_0^1(\Omega)$ and V_1 and V_2 are the spaces of functions in $H_0^1(\Omega_1)$ and $H_0^1(\Omega_2)$ extended to Ω by 0. In the following lemma we prove that $V = \overline{V_1 + V_2}$; see [26, pages 7 and 8].

LEMMA 2.16. *Let Ω be a domain in \mathbb{R}^m that can be decomposed as $\Omega = \Omega_1 \cup \Omega_2$, where Ω_1 and Ω_2 are two domains in \mathbb{R}^m . Then we have that $V = \overline{V_1 + V_2}$.*

Proof. Denote by $\mathcal{D}(\Omega)$ the space of test functions over Ω :

$$\mathcal{D}(\Omega) := \{v \in C^\infty(\Omega) : \text{supp } v \text{ is a compact subset of } \Omega\}.$$

Similarly, denote by $\mathcal{D}(\Omega_1)$ and $\mathcal{D}(\Omega_2)$ the spaces of test functions over Ω_1 and Ω_2 , extended to Ω by 0. A sufficient condition to obtain the claim is that

$$(2.40) \quad \forall \phi \in \mathcal{D}(\Omega), \exists \phi_1 \in \mathcal{D}(\Omega_1), \exists \phi_2 \in \mathcal{D}(\Omega_2) \quad \text{such that} \quad \phi = \phi_1 + \phi_2.$$

In fact, recalling that H_0^1 is the closure of \mathcal{D} in the H^1 -norm (see, e.g., [9, 13]) (2.40) implies that

$$\mathcal{D}(\Omega) = \mathcal{D}(\Omega_1) + \mathcal{D}(\Omega_2) \subset \overline{\mathcal{D}(\Omega_1)} + \overline{\mathcal{D}(\Omega_2)} = V_1 + V_2,$$

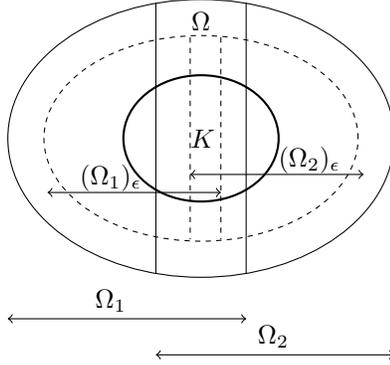


FIG. 2.6. Example of a decomposition of Ω , and of the sets K (boundary shown by the bold solid curve), $(\Omega_1)_\epsilon$ (boundary shown by the dashed left curve), and $(\Omega_2)_\epsilon$ (boundary shown by the dashed right curve) used in the proof of Lemma 2.16.

and hence $V = \overline{\mathcal{D}(\Omega)} \subset \overline{V_1 + V_2} \subset V$.

Next, we show that if $\Omega = \Omega_1 \cup \Omega_2$, then (2.40) holds. We define the sets

$$(\Omega_1)_\delta := \{x \in \Omega_1 : \text{dist}(x, \Omega_1^c) > \delta\}, \quad (\Omega_2)_\delta := \{x \in \Omega_2 : \text{dist}(x, \Omega_2^c) > \delta\},$$

where δ is an arbitrary positive constant. Next, let K be a compact set in Ω , then there exists an $\epsilon > 0$ such that $K \subset (\Omega_1)_\epsilon \cup (\Omega_2)_\epsilon$; see, e.g., Figure 2.6. Take an element $\phi \in \mathcal{D}(\Omega)$ such that $K = \text{supp}(\phi)$. Then there exist $\psi_1 \in \mathcal{D}(\Omega_1)$ and $\psi_2 \in \mathcal{D}(\Omega_2)$ such that $0 \leq \psi_j \leq 1$, and $\psi_j = 1$ in a neighborhood of $K \cap (\Omega_j)_\epsilon$, for $j = 1, 2$. We define

$$\phi_1 := \psi_1(\psi_1 + \psi_2)^{-1}\phi \quad \text{and} \quad \phi_2 := \psi_2(\psi_1 + \psi_2)^{-1}\phi,$$

and by noticing that

$$\phi_1 + \phi_2 = \psi_1(\psi_1 + \psi_2)^{-1}\phi + \psi_2(\psi_1 + \psi_2)^{-1}\phi = \phi,$$

we conclude our proof. \square

Once $\overline{V_1 + V_2} = V$ is proved, we study under which conditions the sum $V_1 + V_2$ is closed, which then means that $V = V_1 + V_2$. This is done in the following theorem.

THEOREM 2.17. *Recall the definition of the sets Ω_{11} and Ω_{22} given in (2.1). The equality $V_1 + V_2 = \overline{V_1 + V_2}$ holds if one of the following conditions is satisfied.*

- (a) *There exist nonnegative functions $\chi_1, \chi_2 \in W^{1,\infty}(\Omega)$ such that χ_1 vanishes on Ω_{22} , χ_2 vanishes on Ω_{11} , and $\chi_1 + \chi_2 = 1$ in Ω .*
- (b) *The domain Ω is Lipschitz, and there exist nonnegative functions $\chi_1, \chi_2 \in W_{loc}^{1,\infty}(\Omega)$, vanishing on Ω_{22} and Ω_{11} , such that $\chi_1 + \chi_2 = 1$ in Ω , and for some $C \geq 0$ it holds that*

$$|\nabla \chi_j(\mathbf{x})| \leq C \text{dist}(\mathbf{x}, \partial\Omega)^{-1}, \quad \text{a.e. in } \Omega, \quad j = 1, 2.$$

Proof. Let $v \in V = H_0^1(\Omega)$ and set $v_1 := \chi_1 v$ and $v_2 := \chi_2 v$. Assume (a) holds. Since $\chi_j \in W^{1,\infty}(\Omega)$, then $v_j = \chi_j v \in V_j$ for $j = 1, 2$. Assume now that (b) holds. Since $\chi_j \in L_{loc}^\infty(\Omega)$, it holds that $v_j \in L^2(\Omega)$, and we need to verify that $\nabla v_j \in L^2(\Omega)$. To do so, we compute

$$\begin{aligned}
 (2.41) \quad \int_{\Omega} |\nabla(\chi_j v)|^2 dx &= \int_{\Omega} |v \nabla \chi_j + \chi_j \nabla v|^2 dx \leq 2 \int_{\Omega} |v|^2 |\nabla \chi_j|^2 + |\nabla v|^2 dx \\
 &\leq 2C \int_{\Omega} \frac{|v|^2}{\text{dist}(\mathbf{x}, \partial\Omega)^2} dx + 2 \int_{\Omega} |\nabla v|^2 dx \leq 2(Cc^* + 1) \|v\|^2,
 \end{aligned}$$

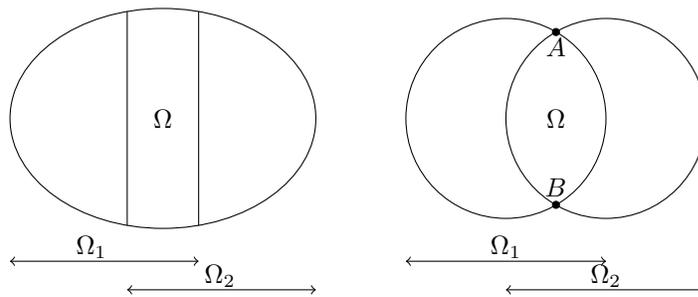


FIG. 2.7. Left: example of a decomposition with uniform overlap. Right: example of a non-uniform overlap (notice that $\bar{\Omega}_{11} \cap \bar{\Omega}_{22} = \{A, B\} \neq \emptyset$).

where the fact that Ω is a Lipschitz domain allowed us to use Hardy's inequality with constant c^* ; see [22, Corollary 3.11, together with Example 3.6 (3) and Remark 3.18]. Since $v \in H_0^1(\Omega)$, the estimate (2.41) shows that $\nabla v_j \in L^2(\Omega)$. \square

As remarked in [26], assumption (a) in Theorem 2.17 is satisfied provided that there is some uniform overlap of Ω_1 and Ω_2 , that is, $\bar{\Omega}_{11} \cap \bar{\Omega}_{22} = \emptyset$; see, e.g., [26, Figure 1a and 1b on page 3]. The assumption (a) fails in general if $\bar{\Omega}_{11} \cap \bar{\Omega}_{22} \neq \emptyset$. See Figure 2.7 for examples of these cases. Other examples are provided in [26, page 8–9]. We remark that our problem of chain of subdomains (Section 3) is such that $\bar{\Omega}_{11} \cap \bar{\Omega}_{22} \neq \emptyset$, hence the use of (a) is not possible.

Notice that the requirement that the domain is Lipschitz (cf. assumption (b) of Theorem 2.17) is due to the fact that we used Hardy's inequality to deal with the term containing $\text{dist}(\mathbf{x}, \partial\Omega)$. The Lipschitz assumption guarantees that Hardy's inequality holds for functions in $H_0^1(\Omega)$, and we refer to [22], as remarked in the proof, and to [1], where, even if a different approach is used, the Lipschitz property of the domain is sufficient to guarantee that the inequality holds. To conclude convergence of the Schwarz method (or APM) the existence of such a constant is enough. However, since we are going to explicitly compute the constant C_L used in Theorem 2.15, it will be necessary to have an explicit estimate for c^* . This is in general not an easy task, and we refer to, e.g., [14] and the references therein, where an optimization of c^* is discussed. We remark that Hardy's constant c^* depends in general on the domain: let $\Omega \subset \mathbb{R}^m$, then c^* depends on m ; see, e.g., [10, 22]. The dimension m plays a crucial role in Hardy-type inequalities. In fact, most of the work in the literature focuses on problems with $m \geq 3$. We refer to [1, 10] for some explicit estimates of c^* valid for $m \geq 2$. In the case $m = 1$, the assumption (a) in Theorem 2.17 is in general satisfied, hence, there is no need to invoke a Hardy-type inequality. Nevertheless, the reader can find an estimate of c^* for $m = 1$ in [1].

We summarize the results obtained in Lemma 2.16 and Theorems 2.15 and 2.17 in the following corollary; see also [26, page 9].

COROLLARY 2.18. *If $\Omega = \Omega_1 \cup \Omega_2$, then the alternating Schwarz method converges. If in addition (a) or (b) in Theorem 2.17 holds, then the Schwarz method converges geometrically.*

Next, we provide an estimate for the constant c_0 in Theorem 2.15. To do so, we recall from the proof of Lemma 2.12 that c_0 in (2.31) has to satisfy $(\|v_1\|^2 + \|v_2\|^2)^{1/2} \leq c_0 \|v\|$ for $v = v_1 + v_2$ with $v \in V$, $v_1 \in V_1$, and $v_2 \in V_2$. Now, assuming that (a) or (b) in Theorem 2.17

holds, we set $v_1 = \chi_1 v$ and $v_2 = \chi_2 v$ and compute

$$\begin{aligned}
 \|v_1\|^2 + \|v_2\|^2 &= \|\chi_1 v\|^2 + \|\chi_2 v\|^2 \\
 &= \int_{\Omega} |\nabla(\chi_1 v)|^2 + |\nabla(\chi_2 v)|^2 d\mathbf{x} \\
 &= \int_{\Omega} |v \nabla \chi_1 + \chi_1 \nabla v|^2 + |v \nabla \chi_2 + \chi_2 \nabla v|^2 d\mathbf{x} \\
 &= \int_{\Omega} |v|^2 (|\nabla \chi_1|^2 + |\nabla \chi_2|^2) + (\chi_1^2 + \chi_2^2) |\nabla v|^2 \\
 &\quad + 2v \chi_1 (\nabla \chi_1 \cdot \nabla v) + 2v \chi_2 (\nabla \chi_2 \cdot \nabla v) d\mathbf{x} \\
 &= \int_{\Omega} |v|^2 (|\nabla \chi_1|^2 + |\nabla \chi_2|^2) + (\chi_1 + \chi_2)^2 |\nabla v|^2 - 2\chi_1 \chi_2 |\nabla v|^2 \\
 &\quad + 2v \chi_1 (\nabla \chi_1 \cdot \nabla v) + 2v \chi_2 (\nabla \chi_2 \cdot \nabla v) d\mathbf{x}.
 \end{aligned}$$

Now, we recall that $\chi_1 + \chi_2 = 1$, notice that $\nabla \chi_1 = -\nabla \chi_2$, and obtain that

$$\|v_1\|^2 + \|v_2\|^2 = \int_{\Omega} 2|v|^2 |\nabla \chi_1|^2 + |\nabla v|^2 - 2\chi_1 \chi_2 |\nabla v|^2 + 2v(\chi_1 - \chi_2)(\nabla \chi_1 \cdot \nabla v) d\mathbf{x}.$$

Noticing that $-2\chi_1 \chi_2 |\nabla v|^2 \leq 0$ and using the Cauchy-Schwarz and Cauchy's inequality [13, Appendix B.2] ($2ab \leq \delta a^2 + b^2/\delta$), we can write for an arbitrary $\delta > 0$ that

$$\begin{aligned}
 \|v_1\|^2 + \|v_2\|^2 &\leq \int_{\Omega} 2|v|^2 |\nabla \chi_1|^2 + |\nabla v|^2 + 2|v| |\chi_1 - \chi_2| |\nabla \chi_1| |\nabla v| d\mathbf{x} \\
 &\leq \int_{\Omega} 2|v|^2 |\nabla \chi_1|^2 + (1 + \delta) |\nabla v|^2 + \frac{1}{\delta} |v|^2 |\chi_1 - \chi_2|^2 |\nabla \chi_1|^2 d\mathbf{x},
 \end{aligned}$$

and by the fact that $|\chi_1 - \chi_2| \leq 1$, we obtain

$$(2.42) \quad \|v_1\|^2 + \|v_2\|^2 \leq \left(2 + \frac{1}{\delta}\right) \int_{\Omega} |v|^2 |\nabla \chi_1|^2 d\mathbf{x} + (1 + \delta) \int_{\Omega} |\nabla v|^2 d\mathbf{x},$$

for all $\delta > 0$.

Now, we need to distinguish between the cases in which (a) or (b) in Theorem 2.17 holds. Assume that (a) holds. Then, (2.42) becomes⁷

$$\begin{aligned}
 (2.43) \quad \|v_1\|^2 + \|v_2\|^2 &\leq \left(2 + \frac{1}{\delta}\right) \|\nabla \chi_1\|_{L^\infty}^2 \int_{\Omega} |v|^2 d\mathbf{x} + (1 + \delta) \int_{\Omega} |\nabla v|^2 d\mathbf{x} \\
 &\leq \left(2 + \frac{1}{\delta}\right) \frac{\kappa^2}{\lambda_1} \int_{\Omega} |\nabla v|^2 d\mathbf{x} + (1 + \delta) \int_{\Omega} |\nabla v|^2 d\mathbf{x},
 \end{aligned}$$

where $\kappa = \|\nabla \chi_1\|_{L^\infty}$ and λ_1 is the first eigenvalue of the Laplace operator in $H_0^1(\Omega)$, that is, $\lambda_1 = \min_{w \in H_0^1(\Omega)} \frac{\langle \nabla w, \nabla w \rangle}{\langle w, w \rangle}$. We thus obtain

$$(2.44) \quad c_0^2 \leq 1 + 2 \frac{\kappa^2}{\lambda_1} + \frac{\kappa^2}{\lambda_1} \frac{1}{\delta} + \delta,$$

⁷Notice that in case of assumption (a) in Theorem 2.17, it is also possible to use a Poincaré inequality [34, 9] in (2.43): the domain decomposition of Ω allows us to bound the integral of $|v|^2$ over Ω by the sums of integrals over Ω_j . This is possible because there is a subset with non-zero measure of each boundary $\partial\Omega_j$ where v vanishes. This has an interesting interpretation in case of a decomposition of Ω into many subdomains as we will see later.

for any $\delta > 0$ as stated by P.-L. Lions on page 10 in [26].

Next, we assume (b), which implies that for some constant C , $|\nabla \chi_i(\mathbf{x})| \leq C \frac{1}{\text{dist}(\mathbf{x}, \partial\Omega)}$, and use (2.42) to estimate

$$\begin{aligned}
 \|v_1\|^2 + \|v_2\|^2 &\leq \left(2 + \frac{1}{\delta}\right) C^2 \int_{\Omega} \frac{|v|^2}{|\text{dist}(\mathbf{x}, \partial\Omega)|^2} d\mathbf{x} + (1 + \delta) \int_{\Omega} |\nabla v|^2 d\mathbf{x} \\
 &\leq \left(2 + \frac{1}{\delta}\right) C^2 c^* \int_{\Omega} |\nabla v|^2 d\mathbf{x} + (1 + \delta) \int_{\Omega} |\nabla v|^2 d\mathbf{x} \\
 &= \left(1 + 2C^2 c^* + \delta + \frac{1}{\delta} C^2 c^*\right) \|v\|^2,
 \end{aligned}$$

where we used Hardy's inequality (tacitly used by P.-L. Lions) with a corresponding constant c^* . Hence, we obtain

$$(2.45) \quad c_0^2 \leq 1 + 2C^2 c^* + \frac{1}{\delta} C^2 c^* + \delta,$$

for any $\delta > 0$. We remark that the use of the first eigenvalue λ_1 to get (2.44) and of the Hardy inequality to get (2.45) are very important to show scalability of the PSM. On the other hand, their use needs some extra care, as we will see in the proof of Theorem 3.4 below and in Section 4.

Notice that in both cases the constant c_0 is a function of δ of the form $f(\delta) = a + \frac{b}{\delta} + \delta$, with $a > 1$ and $b > 0$. This function has a minimum at $\delta = \sqrt{b}$. This means that the best choice of δ is $\delta = \frac{\kappa}{\lambda_1^{1/2}}$ for (2.44) and $\delta = C\sqrt{c^*}$ for (2.45). We summarize the above results in the following theorem.

THEOREM 2.19. *The constant c_0 in (2.31) is bounded by*

$$c_0^2 \leq \begin{cases} 1 + 2\frac{\kappa^2}{\lambda_1} + 2\frac{\kappa}{\sqrt{\lambda_1}} & \text{if (a) in Theorem 2.17 holds,} \\ 1 + 2C^2 c^* + 2C\sqrt{c^*} & \text{if (b) in Theorem 2.17 holds,} \end{cases}$$

where $\kappa = \|\nabla \chi_1\|_{L^\infty}$, λ_1 is the smallest eigenvalue of the Laplace operator in $H_0^1(\Omega)$, C is a constant such that $|\nabla \chi_i(\mathbf{x})| \leq C \frac{1}{\text{dist}(\mathbf{x}, \partial\Omega)}$, and c^* is the Hardy's inequality constant which depends on Ω .

3. Parallel Schwarz method for chains of subdomains. We consider the problem of a chain of N subdomains and denote by u_j the solution on the j th subdomain. Each subdomain Ω_j is a ball in \mathbb{R}^2 that intersects other subdomains Ω_k such that the entire problem is defined over $\Omega = \bigcup_{j=1}^N \Omega_j$. In particular, we focus on “chains” or “ensembles” of subdomains that were defined in [7]. Notice that our analysis is defined for simplicity in \mathbb{R}^2 . However, the same arguments hold also in \mathbb{R}^3 , where each spherical subdomain would represent the van der Waals cavity of an atom. We recall here the definitions and some of the figures presented in [7].

DEFINITION 3.1 (Chain of subdomains). *A set of a finite number N of subdomains, where each subdomain is an open ball $\Omega_j \subset \mathbb{R}^2$, for $j = 1, \dots, N$, is said to be a chain, denoted by \mathcal{C} , if and only if the domain of the chain, defined as $\Omega := \bigcup_{j=1}^N \Omega_j$, is a path connected set. We say that the chain \mathcal{C} is a*

- “linear chain” (see Figure 3.1 (left)) if $\Omega_j \cap \Omega_k \neq \emptyset$ if and only if $k \in \{j-1, j+1\}$, for $j = 2, \dots, N-1$, and $\Omega_1 \cap \Omega_j \neq \emptyset$ if and only if $j = 2$ and $\Omega_N \cap \Omega_j \neq \emptyset$ if and only if $j = N-1$;
- “ringed chain” (see Figure 3.1 (right)) if $\Omega_j \cap \Omega_{j\pm 1} \neq \emptyset$, $\bar{\Omega}_j \cap \bar{\Omega}_{j\pm k} = \emptyset$ for $k > 1$, and $\Omega_1 \cap \Omega_N \neq \emptyset$;

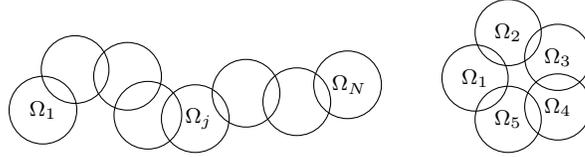


FIG. 3.1. Example of a linear chain (left) and a ringed chain (right).

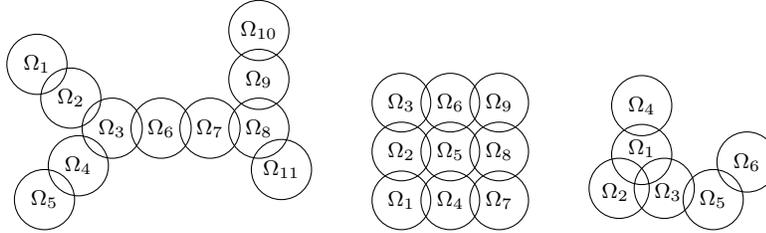


FIG. 3.2. Examples of a branched chain (left), a meshed chain (middle), and a branched chain (right) that does not satisfy (3.1).

- “branched chain” (see Figure 3.2 (left)) if \mathcal{C} can be decomposed into subchains \mathcal{C}_ℓ such that $\bigcup_\ell \mathcal{C}_\ell = \mathcal{C}$, where each subchain \mathcal{C}_ℓ is a linear chain or a ringed chain and the intersection of any two subchains is an empty set or a singleton. In the latter case, the unique element in the intersection is said to be a “connection subdomain”.
- “meshed chain” (see Figure 3.2 (middle)) if the centers of the subdomains Ω_j in \mathcal{C} are distributed on a grid.

Linear chains satisfy the following two conditions: for all distinct $j, k, \ell \in \{1, \dots, N\}$, we have

$$(3.1a) \quad \overline{\Omega}_j \cap \overline{\Omega}_k \cap \overline{\Omega}_\ell = \emptyset,$$

$$(3.1b) \quad \text{if } \Omega_j \cap \Omega_k \neq \emptyset \quad \text{and} \quad \Omega_k \cap \Omega_\ell \neq \emptyset, \quad \text{then } \Omega_j \cap \Omega_\ell = \emptyset.$$

These conditions were also assumed by P.-L. Lions in [26] to ensure a variational interpretation of the Schwarz method; for more details, see [7, 26].

DEFINITION 3.2 (Equidistant chain of subdomains). *Consider a chain \mathcal{C} of N subdomains such that the Ω_j are unit balls. Let Ω_j and Ω_k be any two subdomains in \mathcal{C} such that $\Omega_j \cap \Omega_k \neq \emptyset$ and denote by $d_{j,k} \in \mathbb{R}^+$ the distance between the center of Ω_j and Ω_k . We say that the chain \mathcal{C} is equidistant if $d_{j,k} = d_{n,m}$ for any $j, k, n, m \in \{1, \dots, N\}$ such that $\Omega_j \cap \Omega_k \neq \emptyset$ and $\Omega_n \cap \Omega_m \neq \emptyset$.*

Now, for $j = 1, \dots, N$, we define the index sets

$$\mathcal{I}_j := \{k : 1 \leq k \leq N; k \neq j, \text{ and } \Omega_j \cap \Omega_k \neq \emptyset\}.$$

With this notation, the solution u_j of the j th subdomain satisfies the Laplace problem

$$(3.2) \quad \begin{aligned} -\Delta u_j &= f_j && \text{in } \Omega_j, \\ u_j &= u_k && \text{in } \Omega_j \cap \Omega_k \quad \forall k \in \mathcal{I}_j, \\ u_j &= g_j && \text{on } \Gamma_{j,0} := \partial\Omega_j \setminus \bigcup_{k \in \mathcal{I}_j} (\Omega_k \cap \partial\Omega_j). \end{aligned}$$

Notice that the global problem is

$$(3.3) \quad \begin{aligned} -\Delta u &= f && \text{in } \Omega, \\ u &= g && \text{on } \partial\Omega. \end{aligned}$$

The functions u_j , f_j , and g_j in (3.2) are the restrictions of u , f , and g on Ω_j and $\partial\Omega_j$, for $j = 1, \dots, N$. The functions g and f are assumed to be in $H^{1/2}(\partial\Omega)$ and $L^2(\Omega)$. This implies that the unique solution u to (3.3) is in $H^1(\Omega)$; see, e.g., [13, 17].

As in [4, 6, 7], the parallel Schwarz method (PSM) for the solution of (3.2)–(3.3) is

$$(3.4) \quad \begin{aligned} -\Delta u_j^i &= f_j && \text{in } \Omega_j, \\ u_j^i &= g_j^i && \text{on } \partial\Omega_j, \end{aligned}$$

where

$$g_j^i = \begin{cases} g_j & \text{on } \Gamma_{j,0}, \\ u_k^{i-1} & \text{on } \partial\Omega_j \cap \bar{\Omega}_k, \quad \forall k \in \mathcal{I}_j. \end{cases}$$

We assume that $u_j^0 \in H^1(\Omega_j)$ with $u_j^0 = g_j$ on $\Gamma_{j,0}$ for all j . Recalling that $g \in H^{1/2}(\partial\Omega)$, it is easy to see that for $i = 1$ the boundary condition for the j th problem is

$$g_j^1 = \begin{cases} g_j & \text{on } \Gamma_{j,0}, \\ u_k^0 & \text{on } \partial\Omega_j \cap \bar{\Omega}_k, \quad \forall k \in \mathcal{I}_1, \end{cases}$$

which is in $H^{1/2}(\partial\Omega_j)$. This implies that $u_j^1 \in H^1(\Omega_j)$. Therefore, one can repeat this argument for $i = 2, 3, \dots$, to obtain that $g_j^i \in H^{1/2}(\partial\Omega_j)$ and hence that $u_j^i \in H^1(\Omega_j)$ for any j .

In what follows, we analyze the convergence of the parallel Schwarz method (3.4): our goal is to prove that the error sequence converges to zero geometrically and independently of the number of subdomains N . Moreover, we are interested in a direct estimate of the contraction constant governing the convergence behavior.

3.1. Convergence analysis. In this section, we show how to extend the convergence results presented in Sections 2.1 and 2.2 to the PSM introduced in (3.4). In the geometric setting of the subdomain chain problem, we follow similar arguments as in [26] and consider the variational interpretation of PSM to prove our convergence results.

Consider a chain \mathcal{C} having domain Ω such that its elements Ω_j satisfy the properties (3.1). Denote again by $u \in H_0^1(\Omega)$ the solution to (3.3). Then, for a given initial guess $u^0 \in H_0^1(\Omega)$, the PSM generates the sequences $\{u_j^i\}_{i \in \mathbb{N}}$, for $j = 1, \dots, N$, where the element u_j^i (defined on $\bar{\Omega}_j$) solves (3.4).

In order to obtain a variational interpretation of (3.4), we assume that the chain \mathcal{C} can be decomposed as

$$\mathcal{C} = \mathcal{C}_{\text{odd}} \cup \mathcal{C}_{\text{even}} \quad \text{where} \quad \mathcal{C}_{\text{odd}} := \{\Omega_j : j \in \mathcal{I}_{\text{odd}}\} \quad \text{and} \quad \mathcal{C}_{\text{even}} := \{\Omega_j : j \in \mathcal{I}_{\text{even}}\},$$

and $\mathcal{I}_{\text{odd}} \subset \mathcal{I} := \{1, \dots, N\}$ is defined as follows: \mathcal{I}_{odd} contains 1 and an integer $j \in \mathcal{I}$ belongs to \mathcal{I}_{odd} if and only if $\Omega_j \cap \Omega_k = \emptyset$ for all $k \in \mathcal{I}_{\text{odd}}$ with $k \neq j$. The set $\mathcal{I}_{\text{even}}$ is then defined as $\mathcal{I}_{\text{even}} := \{1, \dots, N\} \setminus \mathcal{I}_{\text{odd}}$. Obviously, we have that $\Omega_j \cap \Omega_k = \emptyset$ for any $j, k \in \mathcal{I}_{\text{odd}}$ and also $\mathcal{I}_{\text{even}} = \mathcal{I} \setminus \mathcal{I}_{\text{odd}}$. Notice that if (3.1) holds and if a ringed subchain is formed by an even number of subdomains, we can always obtain such a decomposition of

the chain. This decomposition is necessary to get a variational interpretation of the PSM; see also [26] Case 3.2 (Figure 3.c) on page 19 and the related discussion on pages 20–21. If it occurs that a ringed subchain has an odd number of subdomains, then one can nevertheless “join” two subdomains and obtain that the subchain has an even number of elements, but then the two “joined subdomains” have to be treated together in the Schwarz method as one subdomain. Notice that this modification is not necessary in practice because the convergence of the PSM can also be proved in the cases in which it has not a variational interpretation; see, e.g., [7, 27].

Once the sets \mathcal{I}_{even} and \mathcal{I}_{odd} are given, we define a “red” sequence $\{u_{red}^i\}_{i \in \mathbb{N}}$ such that its odd elements are given by

$$(3.5) \quad u_{red}^{2n-1} := \begin{cases} u_j^{2n-2} & \text{on } \bar{\Omega} \setminus \cup_{\ell=1, \ell \neq j}^N \bar{\Omega}_\ell, \text{ for } j \in \mathcal{I}_{even}, \\ u_j^{2n-1} & \text{on } \bar{\Omega}_j, \text{ for } j \in \mathcal{I}_{odd}, \end{cases}$$

while the even elements are defined as

$$(3.6) \quad u_{red}^{2n} := \begin{cases} u_j^{2n-1} & \text{on } \bar{\Omega}_j, \text{ for } j \in \mathcal{I}_{even}, \\ u_j^{2n} & \text{on } \bar{\Omega} \setminus \cup_{\ell=1, \ell \neq j}^N \bar{\Omega}_\ell, \text{ for } j \in \mathcal{I}_{odd}. \end{cases}$$

Similarly, one can define a “black” sequence $\{u_{black}^i\}_{i \in \mathbb{N}}$ as

$$u_{black}^{2n-1} := \begin{cases} u_j^{2n-2} & \text{on } \bar{\Omega} \setminus \cup_{\ell=1, \ell \neq j}^N \bar{\Omega}_\ell, \text{ for } j \in \mathcal{I}_{odd}, \\ u_j^{2n-1} & \text{on } \bar{\Omega}_j, \text{ for } j \in \mathcal{I}_{even}, \end{cases}$$

$$u_{black}^{2n} := \begin{cases} u_j^{2n-1} & \text{on } \bar{\Omega}_j, \text{ for } j \in \mathcal{I}_{odd}, \\ u_j^{2n} & \text{on } \bar{\Omega} \setminus \cup_{\ell=1, \ell \neq j}^N \bar{\Omega}_\ell, \text{ for } j \in \mathcal{I}_{even}. \end{cases}$$

Notice that, for a given initial guess u^0 , the black and red sequences are independent and have the same behavior [26]. Hence, without loss of generality, in what follows we focus on the red sequence.

We now follow the same arguments as in Section 2, in particular, Section 2.2. Consider the Hilbert spaces $H_0^1(\Omega_j)$, for $j = 1, \dots, N$, and notice that their elements can be extended to Ω by 0. We denote by \tilde{V}_j the extension of $H_0^1(\Omega_j)$, and we notice that \tilde{V}_j are closed subspaces of $V := H_0^1(\Omega)$. We define the two sets $V_1 = \sum_{j \in \mathcal{I}_{odd}} \tilde{V}_j$ and $V_2 = \sum_{j \in \mathcal{I}_{even}} \tilde{V}_j$. Moreover, recalling assumption (3.1), we notice that $\tilde{V}_j \perp \tilde{V}_k$ for any distinct $j, k \in \mathcal{I}_{odd}$ or $j, k \in \mathcal{I}_{even}$. Using the results presented in [24, 37], one can show that V_1 and V_2 are closed subspaces of V . The PSM (3.4) with respect to the red sequence given by (3.5)–(3.6) is then equivalent to

$$(3.7) \quad \langle u_{red}^{2n+1} - u, v_1 \rangle = 0, \quad \forall v_1 \in V_1, \quad u_{red}^{2n+1} - u_{red}^{2n} \in V_1,$$

and

$$(3.8) \quad \langle u_{red}^{2n} - u, v_2 \rangle = 0, \quad \forall v_2 \in V_2, \quad u_{red}^{2n} - u_{red}^{2n-1} \in V_2.$$

Now, we define the sequence $\{e^i\}_{i \in \mathbb{N}}$ such that $e^i := u - u_{red}^i$. Then (3.7) and (3.8) are equivalent to

$$e^{2n+1} = \mathbb{P}_{V_1^\perp} e^{2n} \quad \text{and} \quad e^{2n} = \mathbb{P}_{V_2^\perp} e^{2n-1},$$

where we have that $V_1^\perp = \bigcap_j \tilde{V}_{2j-1}^\perp$ for $j \in \mathcal{I}_{odd}$ and $V_2^\perp = \bigcap_j \tilde{V}_{2j}^\perp$ for $j \in \mathcal{I}_{even}$. In this way we see that the PSM can be regarded also as an APM and therefore all the convergence results and analysis provided in Sections 2.1 and 2.2 remain valid.

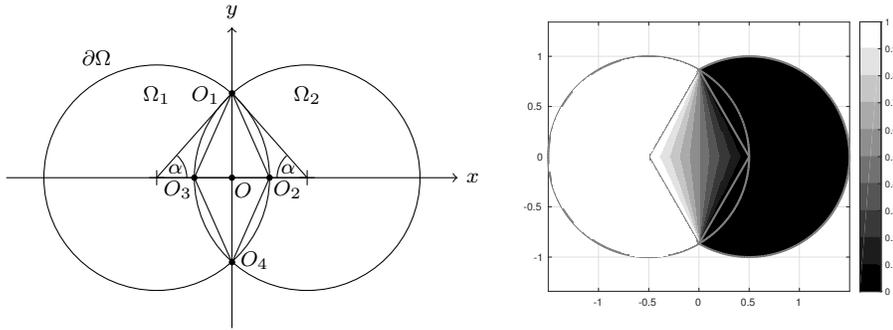


FIG. 3.3. Geometry of the domain of the function χ_1 (left) and plot of χ_1 (right).

Next, we want to apply the convergence results obtained in Sections 2.1 and 2.2 to prove that the PSM, for the solution of chains of subdomains, converges geometrically in the H^1 -norm and independently of the number N of subdomains in the chain. In particular, we apply Theorem 2.15 and Corollary 2.18, together with Theorems 2.17 and 2.19, to prove our result. To use these theorems, an estimate of the constant c_0 is needed, that is, we have to find two functions χ_1 and χ_2 such that the hypotheses of Theorem 2.17 are satisfied. To do so, we begin with the analysis of a case of only two subdomains Ω_1 and Ω_2 that are two unit balls intersecting as in Figure 3.3. Then we will extend the analysis to the case of N subdomains.

We need to introduce a partition of unity, that is, a pair of functions χ_1 and χ_2 that satisfy the assumptions of Theorem 2.17. To do so, we refer to the geometric setting in Figure 3.3 and define the function $\chi_1 : \Omega_1 \cup \Omega_2 \rightarrow \mathbb{R}$ by

$$(3.9) \quad \chi_1(x, y) := \begin{cases} 1 & \text{if } (x, y) \in \overline{\Omega}_{11} \cap (\mathbb{R}^- \times \mathbb{R}), \\ 0 & \text{if } (x, y) \in \overline{\Omega}_{22}, \\ \frac{1}{2} - \frac{1}{2} \frac{x}{\widehat{x}(|y|)} & \text{if } (x, y) \in \mathcal{T} \cup \widehat{\mathcal{T}}, \end{cases}$$

where \mathcal{T} and $\widehat{\mathcal{T}}$ are the sets of all points in the triangles of vertices O_1 - O_2 - O_3 and O_4 - O_2 - O_3 , and $\widehat{x}(y)$ is a function of y that represents the distance between the point $(0, y)$ and the corresponding point (x, y) on the segment O_1O_2 (and O_4O_2). In particular, we notice that the length of the segment $\overline{OO_1}$ (and $\overline{OO_4}$) is $\widetilde{y} := \sin \alpha$ and the length of $\overline{OO_2}$ is $\widetilde{x} := 1 - \cos \alpha$. Hence, one has that $\widehat{x}(y) = \frac{\widetilde{x}}{\widetilde{y}}(\widetilde{y} - y)$. Notice that $\chi_1(x, y)$ is symmetric with respect to the x -axis, and it is linear in x for any fixed y and $(x, y) \in (\Omega_1 \cap \Omega_2)$. We define also $\chi_2 := 1 - \chi_1$.

Next, we show that $|\nabla \chi_1(x, y)|$ is bounded in the sense required by Theorem 2.17(b). This is done in Lemma 3.3.

LEMMA 3.3. *Consider the function χ_1 defined in (3.9) and assume that the angle α is as in Figure 3.3 and $\alpha \in (0, \pi/3)$. Then we have the estimate*

$$|\nabla \chi_1(x, y)| \leq C(\alpha) \frac{1}{\text{dist}((x, y), O_1)}$$

with $C(\alpha) = \frac{|\sin \alpha|}{|1 - \cos \alpha|}$.

Proof. Recall that χ_1 is symmetric with respect to the x -axis. Hence, it suffices to study it restricted to the domain \mathcal{T} inside the triangle O_1 - O_2 - O_3 .

Since $\alpha \in (0, \pi/3)$, recalling $\widetilde{x} := 1 - \cos \alpha$ and $\widetilde{y} := \sin \alpha$ (defined above), we have that $\widetilde{x} - \widetilde{y} = 1 - \cos \alpha - \sin \alpha < 0$, and hence, $|\widetilde{x}| \leq |\widetilde{y}|$. Now, for any point $(x, y) \in \mathcal{T}$, we

consider the corresponding point (\tilde{x}, y) on the segment $\overline{O_1 O_2}$ and estimate

$$|x| \leq |\tilde{x}| = \frac{|\tilde{x}|}{|\tilde{y}|} |\tilde{y} - y| \leq |\tilde{y} - y|.$$

Using this bound, we compute

$$\begin{aligned} |\nabla \chi_1(x, y)|^2 &= |\partial_x \chi_1(x, y)|^2 + |\partial_y \chi_1(x, y)|^2 \leq \frac{|\tilde{y}|^2}{4|\tilde{x}|^2} \left(\frac{1}{|\tilde{y} - y|^2} + \frac{|x|^2}{|\tilde{y} - y|^4} \right) \\ &\leq \frac{|\tilde{y}|^2}{4|\tilde{x}|^2} \left(\frac{1}{|\tilde{y} - y|^2} + \frac{|\tilde{y} - y|^2}{|\tilde{y} - y|^4} \right) = \frac{|\tilde{y}|^2}{2|\tilde{x}|^2} \frac{1}{|\tilde{y} - y|^2}. \end{aligned}$$

Since $|x| \leq |\tilde{y} - y|$, the previous estimate implies that

$$|\nabla \chi_1(x, y)| \leq \frac{|\tilde{y}|}{\sqrt{2}|\tilde{x}|} \frac{1}{\max\{|x|, |\tilde{y} - y|\}} = \frac{|\tilde{y}|}{\sqrt{2}|\tilde{x}|} \frac{1}{|(x, y) - O_1|_\infty},$$

where we recall that $O_1 = (0, \tilde{y})$. Now, notice that $|(x, y) - O_1|_\infty \geq \frac{1}{\sqrt{2}} |(x, y) - O_1|$. Hence, we obtain that

$$|\nabla \chi_1(x, y)| \leq \frac{|\tilde{y}|}{\sqrt{2}|\tilde{x}|} \frac{1}{|(x, y) - O_1|_\infty} \leq \frac{|\tilde{y}|}{|\tilde{x}|} \frac{1}{|(x, y) - O_1|} = \frac{|\tilde{y}|}{|\tilde{x}|} \frac{1}{\text{dist}((x, y), O_1)},$$

where $\text{dist}((x, y), O_1) = |(x, y) - O_1|$. Hence, the claim follows with $C = \frac{|\tilde{y}|}{|\tilde{x}|} = \frac{|\sin \alpha|}{|1 - \cos \alpha|}$. \square

Notice that the angle α that has been introduced in Lemma 3.3 and Figure 3.3 (left) parametrizes the overlap. For this reason the estimate of the contraction factor introduced in the next theorem, denoted by K , is expressed as a function of α , which is a function of the overlap.

THEOREM 3.4. *Consider an equidistant chain \mathcal{C} of N subdomains such that Ω_j is a unit disk for $j = 1, \dots, N$, and assume that (3.1) holds (notice that \mathcal{C} can be, e.g., a linear, ringed, branched, or meshed chain). Denote by Ω the domain of the chain. Then the parallel Schwarz method (3.5)–(3.6) converges geometrically in the sense that*

$$\|e^i\|_{H^1(\Omega)} \leq \text{incl}(V_1, V_2) \|e^{i-1}\|_{H^1(\Omega)},$$

for $i = 1, 2, \dots$, with

$$\text{incl}(V_1, V_2) \leq K(\alpha) < 1,$$

where e^i is the error at the i th iteration given by $e^i := u - u^i$ and α is the angle defined as in Figure 3.3 and $K(\alpha)$ is given by

$$(3.10) \quad K(\alpha) = \sqrt{1 - \frac{1}{1 + 4C(\alpha) + 8C(\alpha)^2}},$$

and $C(\alpha) = \frac{|\sin \alpha|}{|1 - \cos \alpha|}$. The constant K as a function of α is displayed in Figure 3.4.

Moreover, consider another unit disk Ω_{N+1} such that the chain $\mathcal{C}' := \mathcal{C} \cup \Omega_{N+1}$ is equidistant and (3.1) holds. Denote by Ω' the domain of \mathcal{C}' . Then the parallel Schwarz method converges geometrically in the sense that

$$\|e^i\|_{H^1(\Omega')} \leq \text{incl}(V'_1, V'_2) \|e^{i-1}\|_{H^1(\Omega')},$$

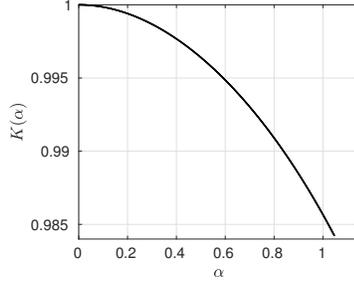


FIG. 3.4. Values of $K(\alpha) < 1$. Notice that $K(\alpha) \rightarrow 1$ as $\alpha \rightarrow 0$. We remark that $\text{incl}(V_1, V_2) \leq K(\alpha)$, hence the convergence factor is bounded by $K(\alpha)$, which is independent of the number N of subdomains.

where V'_1 and V'_2 are defined on Ω' similarly as V_1 and V_2 , $\text{incl}(V'_1, V'_2) \leq K(\alpha) < 1$, and $K(\alpha)$ is exactly given by (3.10).

Proof. Recall the definitions of $\mathcal{I}_{\text{even}}$, \mathcal{I}_{odd} , V_1 , and V_2 as in Section 3.1. Notice that the hypotheses on the Ω_j are such that Lemma 2.16 holds, hence, $V = \overline{V_1 + V_2}$. Next, in order to use Theorem 2.17, we want to find two functions $\tilde{\chi}_1$ and $\tilde{\chi}_2$ such that (b) holds. To this end, we first notice that assumption (3.1) ensures that $\partial\Omega_j \cap \overline{\Omega}_k \cap \overline{\Omega}_\ell = \emptyset$ for all distinct $j, k, \ell \in \mathcal{I}$. We define $\tilde{\chi}_1$ as

$$\tilde{\chi}_1(x, y) := \begin{cases} 1 & \text{if } (x, y) \in \Omega \setminus \bigcup_{j \in \mathcal{I}_{\text{even}}} \Omega_j, \\ 0 & \text{if } (x, y) \in \Omega \setminus \bigcup_{j \in \mathcal{I}_{\text{odd}}} \Omega_j, \\ \chi_1(\xi_{j,k}(x, y), \eta_{j,k}(x, y)) & \text{if } (x, y) \in \Omega_j \cap \Omega_k, j \in \mathcal{I}_{\text{odd}}, k \in \mathcal{I}_j, \end{cases}$$

where χ_1 is defined in (3.9), and we recall $\mathcal{I}_j = \{k : 1 \leq k \leq N, k \neq j, \text{ and } \Omega_j \cap \Omega_k \neq \emptyset\}$. The maps $(x, y) \mapsto (\xi_{j,k}, \eta_{j,k})$ are adequate changes of coordinates defined as

$$\begin{bmatrix} \xi_{j,k}(x, y) \\ \eta_{j,k}(x, y) \end{bmatrix} = \begin{bmatrix} \cos \theta_{j,k} & -\sin \theta_{j,k} \\ \sin \theta_{j,k} & \cos \theta_{j,k} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} x_{j,k} \\ y_{j,k} \end{bmatrix},$$

where $\theta_{j,k}$ is an angle defining the rotation and $[x_{j,k} \ y_{j,k}]^\top$ is a vector corresponding to a translation. The geometric details of such a change of coordinates are given in Figure 3.5, and an example of the resulting function $\tilde{\chi}_1$ for a linear equidistant chain of $N = 6$ subdomains is given in Figure 3.6.

It is easy to see that

$$\nabla \tilde{\chi}_1(x, y) = R(\theta_{j,k}) \nabla \chi_1(\xi_{j,k}, \eta_{j,k}),$$

where $R(\theta_{j,k}) = \begin{bmatrix} \cos \theta_{j,k} & -\sin \theta_{j,k} \\ \sin \theta_{j,k} & \cos \theta_{j,k} \end{bmatrix}$. Since $R(\theta_{j,k})$ is a rotation matrix, it holds that $|\nabla \tilde{\chi}_1(x, y)| = |\nabla \chi_1(\xi_{j,k}, \eta_{j,k})|$. Hence, by Lemma 3.3, we have that

$$\begin{aligned} |\nabla \tilde{\chi}_1(x, y)| &\leq C(\alpha) \frac{1}{\min(\text{dist}((x, y), \tilde{O}_1), \text{dist}((x, y), \tilde{O}_4))} \\ (3.11) \quad &\leq C(\alpha) \frac{1}{\text{dist}((x, y), \partial(\Omega_j \cap \Omega_k))} \end{aligned}$$

for any j, k , where $C(\alpha)$ is given in Lemma 3.3 and \tilde{O}_1 and \tilde{O}_4 are obtained by transforming O_1 and O_4 (see, e.g., Figure 3.5). Now, we define $\tilde{\chi}_2 = 1 - \tilde{\chi}_1$ and invoke Theorem 2.17 and

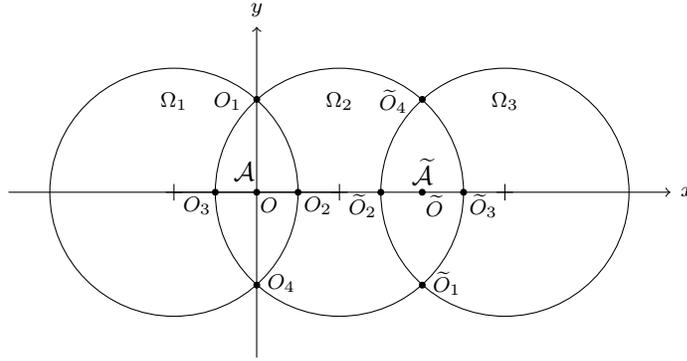


FIG. 3.5. *Geometry of the change of coordinates considered in the proof of Theorem 3.4. Consider the function $\chi_1(x, y)$ in $\Omega_1 \cap \Omega_2$. The change of coordinates to define $\tilde{\chi}_1$ in $\Omega_2 \cap \Omega_3$ is given by a rotation angle $\theta_{2,3} = \pi$ and a translation vector $(x_{j,k}, y_{j,k})^\top = (0, 0)^\top$. Notice that according to this transformation the set $\tilde{\mathcal{A}} = \Omega_2 \cap \Omega_3$ is mapped into $\mathcal{A} = \Omega_1 \cap \Omega_2$, and, in particular, $\tilde{O} \mapsto O$ and $\tilde{O}_j \mapsto O_j$ for $j = 1, 2, 3, 4$.*

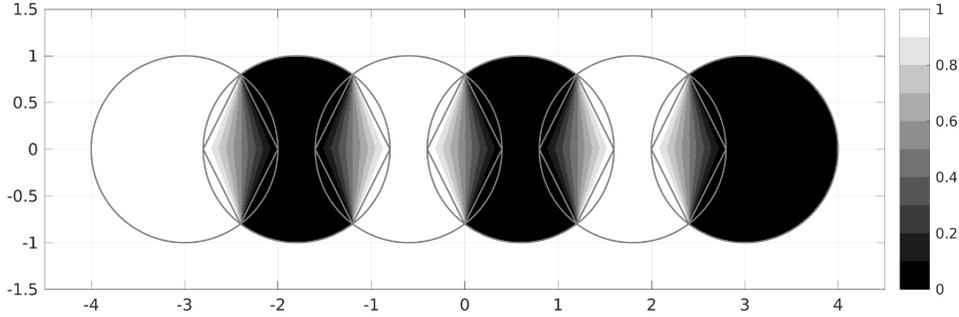


FIG. 3.6. *Plot of an example of the function $\tilde{\chi}_1$.*

Corollary 2.18 to obtain geometric convergence of the PSM in the H^1 -norm. Notice that in order to prove Theorem 2.17 and Corollary 2.18, the Hardy inequality is used and the Hardy constant c^* appears explicitly in the estimate of the contraction factor. An explicit formula of c^* is given in [10, Theorem 12] in case of convex domains. Notice, however, that in our case the domain Ω is not convex, hence, in order to apply [10, Theorem 12] we need to modify the proof of Theorem 2.19 exploiting the specific structure of our problem. In particular, the first integral in (2.42) can be estimated by

$$\begin{aligned}
 \int_{\Omega} |v(x, y)|^2 |\nabla \tilde{\chi}_1(x, y)|^2 dx dy &= \sum_{\substack{j \in \mathcal{I}_{odd} \\ k \in \mathcal{I}_j}} \int_{\Omega_j \cap \Omega_k} |v(x, y)|^2 |\nabla \tilde{\chi}_1(x, y)|^2 dx dy \\
 &\leq C(\alpha)^2 \sum_{\substack{j \in \mathcal{I}_{odd} \\ k \in \mathcal{I}_j}} \int_{\Omega_j \cap \Omega_k} \frac{|v(x, y)|^2}{\text{dist}((x, y), \partial(\Omega_j \cap \Omega_k))^2} dx dy \\
 &\leq C(\alpha)^2 c^* \sum_{\substack{j \in \mathcal{I}_{odd} \\ k \in \mathcal{I}_j}} \int_{\Omega_j \cap \Omega_k} |\nabla v(x, y)|^2 dx dy
 \end{aligned}$$

$$\leq C(\alpha)^2 c^* \int_{\Omega} |\nabla v(x, y)|^2 dx dy,$$

where we used (3.11) and Hardy’s inequality for the integral defined on $\Omega_j \cap \Omega_k$, which is convex since Ω_j and Ω_k are convex. Hence we can use Theorem 12 in [10] to get $c^* = 4$. We can now estimate the convergence factor. Recalling Theorem 2.15 (and the P.-L. Lions constant $C_L(c_0)$), estimate (2.39), and Theorem 2.19, we obtain that

$$\begin{aligned} \text{incl}(V_1, V_2) \leq C_L(c_0(\alpha)) &= \sqrt{1 - \frac{1}{(c_0(\alpha))^2}} \leq \sqrt{1 - \frac{1}{1 + 2C(\alpha)^2 c^* + 2C(\alpha)\sqrt{c^*}}} \\ &= \sqrt{1 - \frac{1}{1 + 8C(\alpha)^2 + 4C(\alpha)}} =: K(\alpha) < 1. \end{aligned}$$

The second statement, which gives the estimate when the number of subdomains grows from N to $N + 1$, can be obtained as follows: since the chain \mathcal{C} is linear, ringed, or branched and Ω_{N+1} is such that (3.1) holds, the chain \mathcal{C}' is also linear, ringed, or branched. Therefore, Lemmas 2.16 and 3.3, Theorems 2.15, 2.17, and 2.19, and all the arguments in the first part of this proof hold. Hence, the claim follows. \square

We conclude this section with further remarks on Theorem 2.17 and 2.19 and the two assumptions (a) and (b) considered there since they represent the main arguments for the estimate of the constant K . Notice that if assumption (a) holds, which means that χ_1 and χ_2 are more regular (and probably easier to find in some given configuration), then the constant c_0 depends on the first eigenvalue λ_1 of the Laplace operator in $H_0^1(\Omega)$. Hence, the constant c_0 is a “global” constant depending on the overall domain Ω . This fact prevents the use of assumption (a) to show that the convergence of the PSM is independent of the number N of subdomains. Moreover, as remarked also by P.-L. Lions in [26], assumption (a) is satisfied, in general, only for uniformly overlapping subdomains. Therefore, assumption (b), which is harder to use, represents a more powerful tool for the convergence analysis of the PSM. The fact that assumptions (a) and (b) can lead to different convergence estimates for the PSM is clarified in the next section.

4. Other growing chains of fixed-sized subdomains. The convergence analysis of the PSM presented in Section 3 is based on the main observation that the convergence of classical Schwarz methods (AltSM and PSM) is governed by the geometry of the overlaps between subdomains. In particular, the scalability result proved in Theorem 3.4 exploits the fact that the different subdomains have a common geometric overlapping structure that does not change when the number N of subdomains increases. This result remains valid (up to an estimate of the contraction factor $C_L(c_0)$ depending on the specific geometry) for all growing chains of fixed-sized subdomains having this property. Another important observation that allowed us to obtain our scalability result is that “each subdomain has to touch the boundary $\partial\Omega$ ”, that is, the measure of $\partial\Omega_j \cap \partial\Omega$ has to be nonzero for $j = 1, \dots, N$. In case that assumption (a) in Theorem 2.17 holds, the nonzero measure of $\partial\Omega_j \cap \partial\Omega$ would allow us to use a local Poincaré inequality (on each subdomain) in (2.43) to control the L^2 -norm of any function v . This is possible because v vanishes on $\partial\Omega_j \cap \partial\Omega$ for any $j = 1, \dots, N$. Estimating in this way will eventually lead to the same scalability result. In this section, we show two other examples of growing chains of fixed-sized subdomains for which the PSM is scalable. The first example is a linear chain of rectangular subdomains, which has already been studied in [6] using an analysis based on the maximum principle. The second example consists of a meshed chain of square-like subdomains: the domain is a square with many “holes”; when the number of subdomains grows, so does the number of “holes” and the size of the square. This

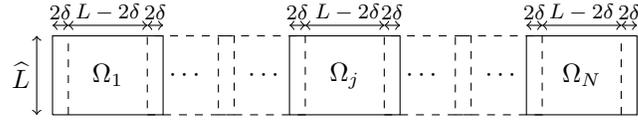


FIG. 4.1. Linear chain of rectangular subdomains.

example is important because it shows that the PSM for the solution of the Laplace equation on perforated domains is scalable without coarse correction and can be an efficient competitor for the methods of reflections; see, e.g., [25, 31].

4.1. Linear chain of rectangular subdomains. In this section, we consider a linear chain \mathcal{C} whose N elements (subdomains) are rectangular domains as shown in Figure 4.1. The domain Ω of the chain \mathcal{C} is a rectangle of height \widehat{L} and width $NL + 2\delta$. The overlap is represented by δ . Notice that this problem has already been studied in [6] using a different type of analysis.

In order to apply the convergence results presented in this section, we need to estimate the constant $C_L(c_0)$, where in this case c_0 depends on the overlap δ . To do so, we define a function χ_1 by

$$\chi_1(x, y) := \begin{cases} 1 & \text{if } (x, y) \in \Omega \setminus \bigcup_{j=1,3,5,\dots} \Omega_j, \\ 0 & \text{if } (x, y) \in \Omega \setminus \bigcup_{j=2,4,6,\dots} \Omega_j, \\ 1 - \frac{\xi_j(x)}{2\delta} & \text{if } (x, y) \in \Omega_j \cap \Omega_{j+1}, j = 1, 3, 5, \dots, \\ 1 - \frac{\xi_j(x)}{2\delta} & \text{if } (x, y) \in \Omega_j \cap \Omega_{j+1}, j = 2, 4, 6, \dots, \end{cases}$$

where ξ_j are changes of coordinates performing a translation in x . Notice that $\chi_1(x, y)$ is continuous, constant in y , and satisfies $|\nabla \chi_1(x, y)| = \frac{1}{2\delta}$ for any $(x, y) \in \Omega_j \cap \Omega_{j+1}$ and $|\nabla \chi_1(x, y)| = 0$ elsewhere. Therefore, assumption (a) in Theorem 2.17 is satisfied and we can use Theorem 2.19 to estimate c_0 and then $C_L(c_0)$. The first eigenvalue of the Laplace operator in $H_0^1(\Omega)$ over Ω is

$$(4.1) \quad \lambda_1 = \left(\frac{2\pi}{NL + 2\delta} \right)^2 + \left(\frac{2\pi}{\widehat{L}} \right)^2 \geq \left(\frac{2\pi}{\widehat{L}} \right)^2.$$

Hence, we have that $c_0^2 \leq 1 + \frac{1}{2\delta^2 \lambda_1} + \frac{1}{\delta \sqrt{\lambda_1}} \leq 1 + \frac{\widehat{L}^2}{8\pi^2 \delta^2} + \frac{\widehat{L}}{2\pi \delta}$, and then

$$C_L(c_0) \leq K_\lambda(\delta) := \sqrt{1 - \frac{1}{1 + \frac{\widehat{L}^2}{8\pi^2 \delta^2} + \frac{\widehat{L}}{2\pi \delta}}},$$

where $K_\lambda(\delta)$ is independent of N . Next, we estimate c_0 using (b) in Theorem 2.17. To do so, we notice that $|\nabla \chi_1(x, y)| \leq \frac{\widehat{L}}{4\delta} \frac{1}{\text{dist}((x, y), \partial\Omega)}$. Therefore, we compute $c_0^2 \leq 1 + \frac{\widehat{L}^2}{2\delta^2} + \frac{\widehat{L}}{\delta}$, and hence, $C_L(c_0) \leq K_\chi(\delta) := C_L(c_0) = \sqrt{1 - \frac{1}{1 + \frac{\widehat{L}^2}{2\delta^2} + \frac{\widehat{L}}{\delta}}}$, where $K_\chi(\delta)$ is independent of N (we used that $c^* = 4$ according to [10]). The two bounds $K_\lambda(\delta)$ and $K_\chi(\delta)$ as functions of δ are displayed in Figure 4.2.

Finally, we remark that, by evaluating c_0 as a function of the eigenvalue λ_1 , it results clearly from (4.1) that c_0 is smaller in two dimensions than in one dimension. This suggests a faster convergence of the PSM for the solution of the 2D problem. Moreover, in one

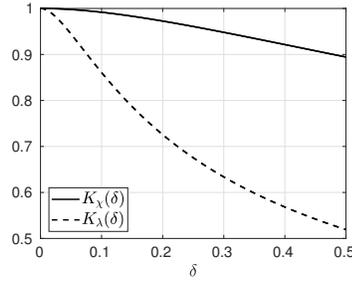


FIG. 4.2. Values of $K(\delta)$ corresponding to $L = \widehat{L} = 1$. We remark that $\text{incl}(V_1, V_2) \leq C_L(c_0) \leq K_\lambda(\delta) \leq K_\lambda(\delta) < 1$, hence the convergence factor is bounded by a constant (strictly smaller than 1) that is independent of the number N of subdomains.

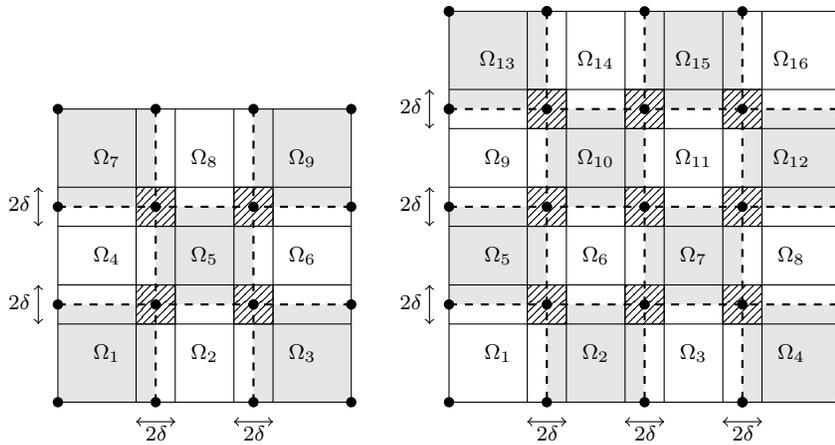


FIG. 4.3. Examples of meshed chains of $N = 9$ (left) and $N = 16$ (right) square-like subdomains. The hatch-filled area represents the “holes”. The white and gray areas are used to construct the partition of unity function χ_1 .

dimension, the domain Ω is an interval (a, b) , hence, the first eigenvalue is $\lambda_1 = \left(\frac{2\pi}{NL+2\delta}\right)^2$, which cannot be bounded from below (as before) to remove the dependence on N . Hence, using assumption (a) in Theorem 2.17 leads to an estimate of $C_L(c_0)$ by a factor that is strictly smaller than 1 but dependent on N . Moreover, the distance of any point x from the boundary is $\min\{|x - a|, |x - b|\}$, which obviously depends on the number N of subdomains. Hence, using assumption (b) in Theorem 2.17 also leads to an estimate of $C_L(c_0)$ by a factor that depends on N . For this reason, in 1D one cannot expect that the PSM converges independently of the number of subdomains. Notice also that $K_\lambda(\delta)$ is a better estimate of the contraction factor. Moreover, a simple comparison of $K_\lambda(\delta)$ with the contraction factor given in [6, Figure 4 (left)] shows that $K_\lambda(\delta)$ is a good estimate of the contraction factor.

4.2. Meshed chains of square-like subdomains. In this example, we consider a meshed chains of square-like subdomains. In particular, the domain Ω is a square with many “holes” that are uniformly distributed in it (so Ω is not simply connected). When the number of subdomains grows, so does the number of “holes” and the size of the square; see Figure 4.1. The overlap is represented by δ . To extend the scalability result presented in Theorem 3.4 to this case, we need to define a partition of unity function χ_1 (and set $\chi_2 = 1 - \chi_1$). This can

be easily done by constructing χ_1 similarly as for the linear chain discussed in Section 4.1. In particular, χ_1 is defined to assume the value 1 on the odd subdomains (gray ones in Figure 4.3) and the value 0 on the even subdomains (white ones in Figure 4.3). In the overlap, χ_1 varies linearly from 0 to 1 in such a way that it is constant in x (e.g., in the overlap $\Omega_1 \cap \Omega_2$ in Figure 4.3 (left)) or in y (e.g., in the overlap $\Omega_1 \cap \Omega_4$ in Figure 4.3 (left)). Notice that this partition of unity leads to the estimate of the same factor $K_\chi(\delta) < 1$ given in Section 4.1. This factor is independent of the number of subdomains if the geometry of the overlaps remains the same when N grows.

5. Conclusions. We conclude our extensive study of the parallel scalability of ddCOSMO with this third study which proved also scalability in the most natural setting for ddCOSMO, namely H^1 . This third contribution is technically the most difficult one, and we tried to be as complete as possible, relating ddCOSMO and the underlying parallel Schwarz method to the alternating projection method in Hilbert spaces and the groundbreaking work of P.-L. Lions. This allowed us to see that the parallel Schwarz method can also be scalable without a coarse space on more general domains than initially thought having the molecules in mind, in particular, on domains with enough holes.

Acknowledgments. We thank Prof. Carlo Franchetti for pointing out the reference [20] and an anonymous referee for a very detailed review (8 pages!) containing many useful suggestions to improve this manuscript.

REFERENCES

- [1] A. ANCONA, *On strong barriers and an inequality of Hardy for domains in \mathbf{R}^n* , J. London Math. Soc. (2), 34 (1986), pp. 274–290.
- [2] 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.
- [3] H. BREZIS, *Functional Analysis, Sobolev Spaces and Partial Differential Equations*, Springer, New York, 2011.
- [4] E. CANCÈS, Y. MADAY, AND B. STAMM, *Domain decomposition for implicit solvation models*, J. Chem. Phys., 139 (2013), Art. 054111, 14 pages.
- [5] W. CHENEY AND A. A. GOLDSTEIN, *Proximity maps for convex sets*, Proc. Amer. Math. Soc., 10 (1959), pp. 448–450.
- [6] G. CIARAMELLA AND M. J. GANDER, *Analysis of the parallel Schwarz method for growing chains of fixed-sized subdomains: Part I*, SIAM J. Numer. Anal., 55 (2017), pp. 1330–1356.
- [7] ———, *Analysis of the parallel Schwarz method for growing chains of fixed-sized subdomains: Part II*, SIAM J. Numer. Anal., 56 (2018), pp. 1498–1524.
- [8] G. CIARAMELLA, M. GANDER, L. HALPERN, AND J. SALOMON, *Review of the methods of reflections*, Tech. Rep., Oberwolfach Preprints OWP-2017-27, Mathematisches Forschungsinstitut Oberwolfach, Oberwolfach, 2017.
- [9] P. G. CIARLET, *Linear and Nonlinear Functional Analysis with Applications*, SIAM, Philadelphia, 2013.
- [10] E. B. DAVIES, *A review of Hardy inequalities*, in The Maz'ya Anniversary Collection, Vol. 2 (Rostock, 1998), J. Rossmann, P. Takáč, and G. Wildenhain, eds., vol. 110 of Oper. Theory Adv. Appl., Birkhäuser, Basel, 1999, pp. 55–67.
- [11] F. DEUTSCH, *The method of alternating orthogonal projections*, in Approximation Theory, Spline Functions and Applications (Maratea, 1991), S. P. Singh, ed., vol. 356 of NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci., Kluwer Acad. Publ., Dordrecht, 1992, pp. 105–121.
- [12] ———, *The angle between subspaces of a Hilbert space*, in Approximation Theory, Wavelets and Applications (Maratea, 1994), S. P. Singh, ed., vol. 454 of NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci., Kluwer Acad. Publ., Dordrecht, 1995, pp. 107–130.
- [13] L. C. EVANS, *Partial Differential Equations*, American Mathematical Society, Providence, 2002.
- [14] S. FILIPPAS AND A. TERTIKAS, *Optimizing improved Hardy inequalities*, J. Funct. Anal., 192 (2002), pp. 186–233.
- [15] C. FRANCHETTI AND W. LIGHT, *The alternating algorithm in uniformly convex spaces*, J. London Math. Soc. (2), 29 (1984), pp. 545–555.

- [16] ———, *On the von Neumann alternating algorithm in Hilbert space*, J. Math. Anal. Appl., 114 (1986), pp. 305–314.
- [17] D. GILBARG AND N. S. TRUDINGER, *Elliptic Partial Differential Equations of Second Order*, 2nd ed., Springer, Berlin, 1983.
- [18] P. R. HALMOS, *A Hilbert space problem book*, 2nd ed., Springer, New York, 1982.
- [19] I. HALPERIN, *The product of projection operators*, Acta Sci. Math. (Szeged), 23 (1962), pp. 96–99.
- [20] G. J. O. JAMESON, *Topology and Normed Spaces*, Chapman and Hall, London, 1974.
- [21] T. KATO, *Perturbation Theory for Linear Operators*, Springer, Berlin, 1995.
- [22] J. KINNUNEN AND O. MARTIO, *Hardy's inequalities for Sobolev functions*, Math. Res. Lett., 4 (1997), pp. 489–500.
- [23] A. KLAMT AND G. SCHÜÜRMAN, *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, (1993), pp. 799–805.
- [24] H. KOBER, *A theorem on Banach spaces*, Compositio Math., 7 (1939), pp. 135–140.
- [25] P. LAURENT, AND J. SALOMON, *On the method of reflections*, Preprint, HAL archive, <https://hal.archives-ouvertes.fr/hal-01439871v2>, submitted, 2017.
- [26] P.-L. LIONS, *On the Schwarz alternating method. I*, in First International Symposium on Domain Decomposition Methods for Partial Differential Equations (Paris, 1987), R. Glowinski, G. H. Golub, G. A. Meurant, and J. Périaux, eds., SIAM, Philadelphia, 1988, pp. 1–42.
- [27] ———, *On the Schwarz alternating method. II. Stochastic Interpretation and Order Properties*, in Domain Decomposition Methods (Los Angeles, 1988), T. F. Chan, R. Glowinski, J. Périaux, and O. B. Widlund, eds., SIAM, Philadelphia, 1989, pp. 47–70.
- [28] 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), Art. 184108, 13 pages.
- [29] 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.
- [30] C. G. LOONEY, *Locally uniformly quasi-convex programming*, SIAM J. Appl. Math., 28 (1975), pp. 881–884.
- [31] 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.
- [32] W. RUDIN, *Functional Analysis*, 2nd ed., McGraw-Hill, New York, 1991.
- [33] A. SHIRYAEV, *Probability*, Springer, New York, 1995.
- [34] A. TOSELLI AND O. WIDLUND, *Domain Decomposition Methods—Algorithms and Theory*, Springer, Berlin, 2005.
- [35] 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.
- [36] J. VON NEumann, *Functional Operators. II. The Geometry of Orthogonal Spaces*, Princeton University Press, Princeton, 1950.
- [37] Z.-M. ZHENG AND H.-S. DING, *A note on closedness of the sum of two closed subspaces in a Banach space*, Commun. Math. Anal., 19 (2016), pp. 62–67.