Méthodes d'inversion de type one-shot et décomposition de domaine

Tuan-Anh Vu

INSTITUT POLYTECHNIQUE DE PARIS

NNT : 2024IPPAE009

Thèse de doctorat

One-shot inversion methods and domain decomposition

Thèse de doctorat de l'Institut Polytechnique de Paris
préparée à l'École nationale supérieure de techniques avancées

École doctorale n°574 Mathématiques Hadamard (EDMH)
Spécialité de doctorat: Mathématiques Appliquées

Thèse présentée et soutenue à Palaiseau, France, le 11/07/2024, par

Tuan-Anh Vu

Composition du Jury:

Marc Bonnet Directeur de Recherche, CNRS, France.	Président
Christophe Geuzaine Professeur, University of Liège, Belgium.	Rapporteur
Dinh Nho Hào Professeur, Hanoi Institute of Mathematics, VAST, Vietnam.	Rapporteur
Faker Ben Belgacem Professeur, Université de technologie de Compiègne, France.	Examinateur
Slim Chaabane Professeur, Sfax University, Tunisia.	Examinateur
Houssem Haddar Directeur de recherche, INRIA, UMA, ENSTA Paris, France.	Directeur de thèse
Marcella Bonazzoli Chargée de recherche, INRIA, UMA, ENSTA Paris, France.	Co-directrice de thèse

Acknowledgments

Throughout the writing of this thesis, I have received great support from my supervisors, my family, my colleagues, and my friends.

First of all, I would like to thank Houssem and Marcella, my supervisor and co-supervisor, for teaching me and creating the best conditions for me to complete the thesis. I would have never improved so far my writing and presentation style without your valuable corrections and advice. Besides, you are the kind ones who always listen to me, try to understand me, and guide me when my mind gets lost. It is such a great pleasure in my life to meet and work with both of you. I am always proud of being one of your students. I also want to address my thanks to Prof. Christophe Geuzaine and Prof. Dinh Nho Hào for having accepted to be reporters of my thesis, as well as to all other jury members, Prof. Marc Bonnet, Prof. Faker Ben Belgacem and Prof. Slim Chaabane. I am very honored that each of you agreed to participate in the evaluation of this thesis.

I sincerely thank all the members of my family for their immense support and encouragement. My grandparents, my parents, and the family of my sister always motivate me to pursue my passion for mathematics. I have to especially thank my beloved aunt, Madam Lan, who can be understood as my "second mother". She is among those who understand me best, and always gives me positive emotional support. The process of writing this thesis also took place at the same time as the maturity of my wonderful love. I fell in love with Bien-Thuy on December 16th, 2019 when I was in France and she was in Vietnam. Then the darkness of the epidemic spread over the world. Not only did Bien-Thuy wait patiently from the end of 2019 to the end of 2021 for our first physical touch, but she also sympathized and shared with me many things in everyday life. She became the infinite source of inspiration for important proofs in the thesis. Our love has a perfect ending with a wedding full of happiness on January 28th, 2024.

I would like to express my gratefulness to all the professors and colleagues that I have met, mainly at the Sorbonne University, CMAP & CMLS (École polytechnique), Inria Saclay Île-de-France, and UMA (ENSTA Paris). In particular, I would like to thank Prof. Frédéric Jean (UMA) for suggesting references for the Jury-Marden Criterion, and Rahma (ISP-EDP Associate Team) for suggesting references for the opening chapter of the thesis. I would like to thank Prof. Frank Pacard, who gave me many valuable educational experiences during my time as a teaching assistant in his course at École polytechnique. Marie and Corinne are both kind and enthusiastic secretaries who guided me through a plenty of administrative procedures. Florent, Gabriel, Giuseppe, Roxane, Rutger, Suney, and Théo have been my trusted friends since the master's course at the Sorbonne University, with whom I can share my difficulties. Dorian, Fabien, Hadrien, Morgane, Zheyi, and many members of the IDEFIX Inria Project Team as well as the people at UMA helped me a lot with work and life issues, sometimes even with French.

I am lucky to have so many Vietnamese friends in France, such as Dinh-Vu, Duc-Quang, Manh-Linh, Ha-Quang, the couple of Hai-Yen and Tu-Anh, Hong-Quan, Huu-Nhan, Minh-Hang, Minh-Hieu, Ngoc-Ky, Quynh-Nga, Quynh-Trang, Thanh-Loan, Thanh-Vuong, Tien-Tai, Tran-Trung, and Tuan-Hung. In particular, Tien-Tai was the one who helped me during my first days in France. I would like to thank Madam Ewa for her accommodation during my thesis. I had many beautiful memories with the members of Ewa's family, Ewa, Philippe, Sophia, Suzanne, and Veronika. My two friendly co-tenants, Max and Shahrouz, are truly like brothers to me.

Once again, I would like to sincerely thank all of you for your contributions to the success of this thesis.

Contents

Notations

Throughout the thesis, we use the following list of notations.

$\langle \cdot, \cdot \rangle$:	the usual Hermitian scalar product, defined by $\langle x, y \rangle \coloneqq \overline{y}^\intercal x, \forall x, y \in \mathbb{C}^n$.
$\|\cdot\|$:	the vector/matrix norms induced by $\langle \cdot, \cdot \rangle$.
$A^* = \overline{A}^\intercal$:	the adjoint operator of a matrix $A \in \mathbb{C}^{m \times n}$.
$z^* = \overline{z}$:	the conjugate of a complex number z. This notation is chosen for the sake of consistence with the notation of adjoint operator.
I:	the identity matrix, whose size is understood from context.
$\rho(T)$:	the spectral radius of a matrix $T \in \mathbb{C}^{n \times n}$.
$s(T)$:	$s(T) \coloneqq \sup_{z \in \mathbb{C}, \lvert z \rvert \geq 1} \|(I - T/z)^{-1}\|$, $\forall T \in \mathbb{C}^{n \times n}, \rho(T) < 1$. See Appendix 3.A.

Chapitre 0

Méthodes itératives pour les problèmes inverses

Contenu

Notre objectif principal est d'analyser la convergence de méthodes d'optimisation basées sur le gradient appliquées à la résolution de problèmes inverses dont les problèmes directs et adjoints doivent être résolus de manière itérative. On peut s'attendre à ce que les problèmes directs et adjoints soient résolus avec une haute précision puisque plus le gradient de la fonctionnelle coût est bien approximé, plus tôt l'inconnue du problème inverse est retrouvée. En conséquence, dans chaque itération externe de mise à jour de l'inconnue du problème inverse à l'aide de méthodes de descente de gradient, un nombre suffisamment grand d'itérations internes serait utilisé pour résoudre les problèmes directs et adjoints. Cependant, de nombreuses expériences numériques ont montré qu'un très petit nombre d'itérations internes peut encore conduire à une bonne convergence pour le problème inverse. En plus de réduire le nombre d'itérations internes, le processus de calcul peut être amélioré en choisissant des méthodes itératives efficaces, telles que les méthodes de décomposition de domaine, et en parallélisant les calculs des solutions des problèmes directes et adjointes . De plus, pour les problèmes directs et adjoints de grande taille, les solveurs directs, qui produiraient un gradient exact, ne sont pas pratiques en raison de leur coût élevé en mémoire. Cela nous a motivé à étudier *les méthodes d'inversion de type one-shot* et *les méthodes de décomposition de domaine*, qui sont présentées dans ce chapitre d'ouverture.

Le Chapitre 0 est organisé comme suit. La Section 0.1 est consacrée à la motivation de la thèse et à l'état de l'art de notre direction de recherche. Ensuite, dans la Section 0.2, nous formulons une classe générique de problèmes inverses linéaires dont les problèmes directs et adjoints sont résolus par des itérations de point fixe. Nous introduisons ensuite les méthodes d'inversion de type one-shot qui itèrent en même temps sur les solutions du problème direct et adjoint. Comme nous aimerions avoir des solveurs itératifs plus efficaces et pratiques que des itérations génériques de point fixe, nous présenterons les méthodes de décomposition de domaine dans la Section 0.3. Une application aux problèmes inverses de conductivité est présentée dans la Section 0.4. Nous terminons le chapitre dans la Section 0.5 en faisant une synthèse des contributions de la thèse.

0.1 Motivation et état de l'art

Pour les problèmes inverses de grande taille, qui surviennent souvent dans des applications réelles, la solution des problèmes directs et adjoints correspondants est généralement calculée à l'aide d'un solveur itératif, tel que les méthodes de point fixe préconditionnées ou de sous-espace de Krylov, plutôt qu'exactement par un solveur direct, comme les solveurs de type LU (voir par exemple [49, 1]). En effet, les systèmes linéaires correspondants pourraient être trop grands pour être traités avec des solveurs directs en raison de leur coût en mémoire élevé. De plus, les solveurs itératifs sont plus faciles à paralléliser sur de nombreux cœurs pour accélérer le temps de calcul. En couplant un solveur itératif avec une méthode d'optimisation basée sur le gradient, l'idée des *méthodes de type one-shot à un pas* est d'itérer en même temps sur la solution du problème direct (la variable d'état), la solution du problème adjoint (l'état adjoint) et sur l'inconnue du problème inverse (le paramètre ou la variable d'optimisation). Si deux itérations internes ou plus sont effectuées sur l'état et l'état adjoint avant de mettre à jour le paramètre (en partant des itérations précédentes comme tentative initiale pour l'état et l'état adjoint), on parle de *méthodes de type one-shot à pas multiples*. Notre objectif est d'analyser rigoureusement la convergence de telles méthodes d'inversion. En particulier, nous nous intéressons aux schémas dans lesquels les itérations internes sur les problèmes directs et adjoints sont incomplètes, c.-à-d. arrêtées avant d'atteindre la convergence. En effet, résoudre les problèmes directs et adjoints exactement par des solveurs directs ou très précisément par des solveurs itératifs pourrait prendre beaucoup de temps avec peu d'amélioration de la précision de la solution du problème inverse.

Le concept de méthodes de type one-shot a été introduit pour la première fois par Ta'asan [46] pour des problèmes de contrôle optimal. Sur la base de cette idée, diverses méthodes connexes, telles que les méthodes all-at-once, où l'équation d'état est incluse dans la fonctionnelle coût, ont été développées pour l'optimisation de forme aérodynamique, voir par exemple [47, 45, 25, 42, 41] et la revue de la littérature dans l'introduction de [42]. Des approches all-at-once pour les problèmes inverses pour l'identification des paramètres ont été étudiées, par exemple, dans [19, 5, 33, 31, 32, 39, 40]. Une méthode alternative, appelée Wavefield Reconstruction Inversion (WRI), a été introduite pour l'imagerie sismique dans [50], comme une amélioration de la Full Waveform Inversion (FWI) classique [48]. WRI est une méthode de pénalisation qui combine les avantages de l'approche all-at-once avec ceux de l'approche réduite (où l'équation d'état représente une contrainte et est imposée à chaque itération, comme dans FWI), et a été étendue à des problèmes inverses plus généraux dans [51].

Peu de preuves de convergence, notamment pour les méthodes de type one-shot à

pas multiples, sont disponibles dans la littérature. En particulier, pour les problèmes de conception optimale non-linéaires, Griewank [17] a proposé une version de méthodes de type one-shot à un pas dans laquelle un préconditionneur basé sur la Hessienne est utilisé dans l'itération sur les variables d'optimisation. L'auteur a prouvé des conditions pour garantir que les valeurs propres réelles du Jacobien des itérations couplées sont inférieures à 1, mais ce sont juste des conditions nécessaires et non suffisantes pour exclure les valeurs propres réelles inférieures à -1. De plus, aucune condition permettant d'assurer que également les valeurs propres complexes soient inférieures à 1 en module n'a été trouvée, et les méthodes de type one-shot à pas multiples n'ont pas été étudiées. Dans [22, 23, 14] une fonction de pénalisation exacte de type Lagrangien doublement augmenté a été présentée pour coordonner les itérations couplées, et la convergence globale de l'approche d'optimisation proposée a été prouvée sous certaines hypothèses. Cette approche particulière de méthode de type one-shot à un pas a ensuite été étendue aux problèmes dépendant du temps dans [18]. Plus récemment, pour des problèmes inverses géométriques, des méthodes d'inversion de type one-shot combinées à la décomposition de domaine ont été analysées dans [7, 28] dans le cas de géométries cylindriquement invariantes.

0.2 Les méthodes d'inversion de type one-shot

Nous présenterons des méthodes d'inversion de type one-shot (à un pas ou à pas multiples) pour une classe générique de problèmes inverses linéaires, pour lesquels nous avons développé une théorie de convergence rigoureuse. Notre analyse se place directement dans le cadre discret de dimension finie.

0.2.1 Problème inverse linéaire

Nous nous concentrons sur des problèmes inverses linéaires (discrétisés), qui correspondent à un *problème direct* de la forme : trouver $u \equiv u(\sigma)$ tel que

$$u = Bu + M\sigma + F \tag{1}$$

où $u \in \mathbb{R}^{n_u}$, $\sigma \in \mathbb{R}^{n_\sigma}$, $B \in \mathbb{R}^{n_u \times n_u}$, $M \in \mathbb{R}^{n_u \times n_\sigma}$ et $F \in \mathbb{R}^{n_u}$. Ici $I - B$ est la matrice inversible associée au problème direct (par exemple obtenue après discrétisation d'un modèle EDP et application d'un préconditionneur au système linéaire), avec le paramètre σ. L'équation (1) est également appelée *équation d'état* et u est la *variable d'état*. Pour σ donné, nous supposons que l'on résout u par une itération de point fixe

$$u_{\ell+1} = Bu_\ell + M\sigma + F, \quad \ell = 0, 1, \ldots \tag{2}$$

Nous supposons $\rho(B) < 1$ afin que l'itération de point fixe (2) converge pour toutes les tentatives initiale u_0 (voir par exemple [16, Théorème 2.1.1]) . En mesurant $g = Hu(\sigma)$, où $H \in \mathbb{R}^{n_g \times n_u}$, nous considérons le *problème inverse linéaire* de reconstruction de σ à partir de la connaissance de g . Posons $A \coloneqq H(I - B)^{-1}M$. Le problème inverse peut s'écrire synthétiquement sous la forme $A\sigma = g - H(I - B)^{-1}F$, ce qui revient à inverser la matrice mal-conditionnée A. Nous supposerons dans la suite l'unicité de la solution de ce problème inverse, qui équivaut à l'injectivité de A. En résumé, nous définissons

$$\begin{array}{ll} \text{problème direct :} & u = Bu + M\sigma + F, \\ \text{problème inverse :} & \text{mesurer } g = Hu(\sigma), \text{ retrouver } \sigma \end{array} \tag{3}$$

avec les hypothèses :

$$\rho(B) < 1, \quad H(I-B)^{-1}M \text{ est injectif.} \tag{4}$$

Remarque 0.2.1. Considérer les matrices B et M à valeurs réelles n'est pas une hypothèse restrictive. En effet, le cas des matrices à valeurs complexes peut être réécrit comme un système d'équations à valeurs réelles en doublant la taille du système linéaire (voir Section 4.1).

Pour résoudre le problème inverse, nous écrivons sa formulation de moindres carrés régularisés : étant donné quelques mesures g, nous retrouvons la *solution régularisée* $\sigma_\alpha^{\text{ex}}$ définie par

$$\sigma_\alpha^{\text{ex}} = \text{argmin}_{\sigma\in\mathbb{R}^{n_\sigma}} J(\sigma) \quad \text{où } J(\sigma) \coloneqq \frac{1}{2}\|Hu(\sigma)-g\|^2 + \frac{\alpha}{2}\|\sigma\|^2,\ \alpha \geq 0. \tag{5}$$

Dans le cas où nous avons des mesures exactes, c.-à-d. $g = Hu(\sigma^{\text{ex}})$, on peut prendre $\alpha = 0$ afin que $\sigma_\alpha^{\text{ex}} = \sigma^{\text{ex}}$. Cependant, en pratique, les mesures sont souvent affectées par du bruit, donc en général on ne peut pas reconstruire σ^{ex} mais juste son approximation $\sigma_\alpha^{\text{ex}}$. Nous résolvons maintenant le problème de minimisation (5). En utilisant la technique classique du Lagrangien avec des produits scalaires réels, nous introduisons l'*état adjoint* $p \equiv p(\sigma)$, qui est la solution de

$$p = B^*p + H^*(Hu-g)$$

et nous permet de calculer le gradient de la fonctionnelle coût comme

$$\nabla J(\sigma) = M^*p(\sigma) + \alpha\sigma.$$

L'algorithme classique de descente de gradient donc s'écrit

$$\textbf{Descente de gradient usuelle :} \quad \begin{cases} \sigma^{n+1} = \sigma^n - \tau M^*p^n - \tau\alpha\sigma^n, \\ u^n = Bu^n + M\sigma^n + F, \\ p^n = B^*p^n + H^*(Hu^n - g), \end{cases} \tag{6}$$

où $\tau > 0$ est le pas de descente, et les équations d'état et d'état adjoint sont résolues exactement par un solveur direct à chaque itération pour σ. Notons que lorsque $F = 0$, (6) équivaut à $\sigma^{n+1} = \sigma^n - \tau A^*(A\sigma^n - g) - \tau\alpha\sigma^n$. On peut également envisager une version légèrement différente où un schéma implicite est appliqué au terme de régularisation conduisant au schéma de gradient semi-implicite suivant

$$\textbf{Descente de gradient semi-implicite :} \quad \begin{cases} \sigma^{n+1} = \sigma^n - \tau M^*p^n - \tau\alpha\sigma^{n+1}, \\ u^n = Bu^n + M\sigma^n + F, \\ p^n = B^*p^n + H^*(Hu^n - g). \end{cases} \tag{7}$$

On peut montrer que les deux algorithmes convergent pour $\tau > 0$ suffisamment petit : pour toute tentative initiale, (6) converge si et seulement si $\tau < \frac{2}{\rho(A^*A)+\alpha}$ et (7) converge si et seulement si $(\rho(A^*A)-\alpha)\tau < 2$. Ces résultats indiquent en particulier que l'on gagne plus de stabilité avec le schéma semi-implicite.

0.2.2 Principe des méthodes d'inversion de type one-shot

Nous nous intéressons à des méthodes où les problèmes directs et adjoints sont plutôt résolus de manière itérative comme dans (2), et où l'on itère en même temps sur la solution du problème direct et sur l'inconnue du problème inverse : de telles méthodes sont appelées *méthodes de type one-shot*. Plus précisément, nous nous intéressons à deux variantes de *méthodes de type one-shot à pas multiples*, définies comme suit. Soit n l'indice de l'itération (externe) sur σ. Nous mettons à jour $\sigma^{n+1} = \sigma^n - \tau M^* p^n - \tau\alpha\sigma^n$ comme dans les méthodes de descente de gradient (ou respectivement, $\sigma^{n+1} = \sigma^n - \tau M^* p^n - \tau\alpha\sigma^{n+1}$ comme dans les méthodes de descente de gradient semi-implicites), mais les équations d'état et d'état adjoint sont maintenant résolues par une méthode d'itération de point fixe, en utilisant uniquement k *itérations internes* et il est important de noter que pour la tentative initiale nous choisissons les informations de l'étape (externe) précédente. Nous obtenons alors les deux variantes suivantes d'algorithmes de type one-shot à pas multiples

$$\textbf{one-shot à } k \textbf{ pas :} \quad \begin{cases} \sigma^{n+1} = \sigma^n - \tau M^* p^n - \tau\alpha\sigma^n, \\ u_0^{n+1} = u^n, p_0^{n+1} = p^n, \\ \text{for } \ell = 0, 1, \dots, k-1 : \\ \quad \left| \begin{array}{l} u_{\ell+1}^{n+1} = B u_\ell^{n+1} + M\sigma^{n+1} + F, \\ p_{\ell+1}^{n+1} = B^* p_\ell^{n+1} + H^*(H u_\ell^{n+1} - g), \end{array} \right. \\ u^{n+1} := u_k^{n+1}, p^{n+1} := p_k^{n+1} \end{cases} \tag{8}$$

et

$$\textbf{one-shot semi-implicite à } k \textbf{ pas :} \quad \begin{cases} \sigma^{n+1} = \sigma^n - \tau M^* p^n - \tau\alpha\sigma^{n+1}, \\ u_0^{n+1} = u^n, p_0^{n+1} = p^n, \\ \text{for } \ell = 0, 1, \dots, k-1 : \\ \quad \left| \begin{array}{l} u_{\ell+1}^{n+1} = B u_\ell^{n+1} + M\sigma^{n+1} + F, \\ p_{\ell+1}^{n+1} = B^* p_\ell^{n+1} + H^*(H u_\ell^{n+1} - g), \end{array} \right. \\ u^{n+1} := u_k^{n+1}, p^{n+1} := p_k^{n+1}. \end{cases} \tag{9}$$

En particulier, lorsque $k = 1$, on obtient les deux algorithmes suivants

$$\textbf{one-shot à un pas :} \quad \begin{cases} \sigma^{n+1} = \sigma^n - \tau M^* p^n - \tau\alpha\sigma^n, \\ u^{n+1} = B u^n + M\sigma^{n+1} + F, \\ p^{n+1} = B^* p^n + H^*(H u^n - g) \end{cases} \tag{10}$$

et

$$\textbf{one-shot semi-implicite à un pas :} \quad \begin{cases} \sigma^{n+1} = \sigma^n - \tau M^* p^n - \tau\alpha\sigma^{n+1}, \\ u^{n+1} = B u^n + M\sigma^{n+1} + F, \\ p^{n+1} = B^* p^n + H^*(H u^n - g). \end{cases} \tag{11}$$

Notons que lorsque $k \to \infty$, la méthode de type one-shot à k pas (8) converge formellement vers la descente de gradient usuelle (6), tandis que la méthode de type one-shot à k pas semi-implicite (9) converge formellement vers la descente de gradient semi-implicite (7). Puisque l'analyse des deux schémas (8) et (9) peut se faire à partir d'arguments similaires, nous choisissons de nous concentrer sur un seul d'entre eux, à savoir le schéma semi-implicite. Nous analysons d'abord la méthode de type one-shot à un pas (11) ($k = 1$) dans

le Chapitre 2 puis celle à pas multiples (9) ($k \geq 1$) dans le Chapitre 3. De plus, nous nous référons à la Section 4.2 pour l'analyse dans le cas $n_u = n_\sigma = n_g = 1$ (pour lequel on peut appliquer des outils fondamentaux) et $\alpha = 0$ (pour lequel les deux schémas (8) et (9) coïncident).

0.3 Méthodes de décomposition de domaine

Les *méthodes de décomposition de domaine (Domain decomposition methods, DDMs, en anglais)* sont une famille de méthodes itératives permettant de résoudre efficacement des problèmes directs de grande taille. L'idée clé derrière ces méthodes est de décomposer le domaine de calcul en plusieurs sousdomaines, dans lesquels les sousproblèmes sont définis de manière appropriée et résolus en parallèle à l'aide d'un solveur direct (robuste). La construction des sousproblèmes joue un rôle fondamental dans la DDM ; essentiellement, cela implique comment construire des sousdomaines avec ou sans recouvrement, quelles conditions aux limites sont imposées aux interfaces et comment échanger des informations entre les sousdomaines. Après cela, les solutions des sousproblèmes doivent également être concaténées de manière correcte, généralement à l'aide d'une partition d'unité, pour garantir la convergence vers la solution globale. Alors que la version originale de la DDM de Schwarz[1] [44] et ses variantes classiques utilisent des conditions de Dirichlet sur les interfaces, les versions successives, introduites pour la première fois par P. L. Lions[2] [34, 36], les *méthodes de Schwarz optimisées (Opimized Schwarz method, OSM, en anglais)* utilisent des conditions de transmission de Robin et cela s'est avéré plus efficace, voir [12, Chapitre 2]. P. L. Lions a fait une preuve de convergence pour le cas elliptique [35], puis Després[3] l'a étendue au cas de l'équation de Helmholtz [11] et des équations de Maxwell en régime harmonique [10]. Ici, nous nous concentrerons sur les OSMs ; cependant, notre introduction inclut également la méthode classique de Schwarz afin que nous puissions décrire les avantages des OSMs par rapport aux méthodes classiques.

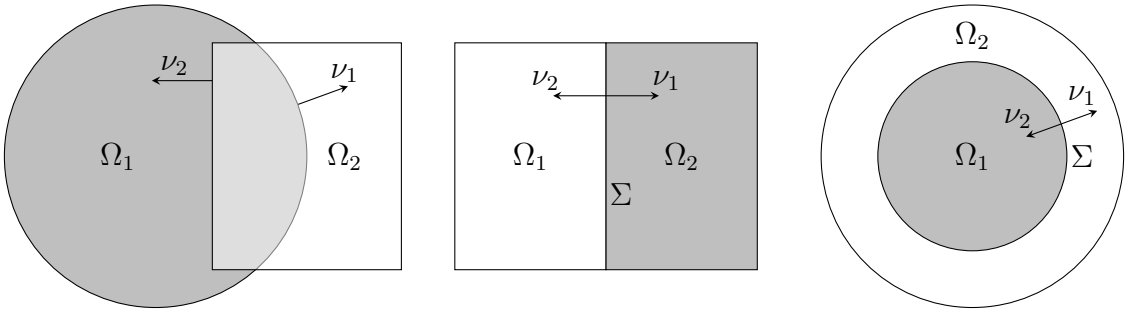
0.3.1 Méthodes de Schwarz optimisées

On considère un problème direct bien posé de la forme

$$\begin{cases} \mathcal{A}(u) = b & \text{l'équation dans } \Omega, \\ \mathcal{B}(u) = f & \text{condition aux limites sur } \partial\Omega \end{cases}$$

et nous supposons que la solution u est suffisamment régulière, et que la trace de u et la dérivée conormale $\mathcal{N}u$ sont continues. Ici, nous considérerons l'équation $-\triangle u + \eta u = 0$ où $\mathcal{N} = \frac{\partial}{\partial \nu}$ et la version légèrement plus générale $-\operatorname{div}(\sigma \nabla u) + \eta u = 0$ où $\mathcal{N} = \sigma \frac{\partial}{\partial \nu}$.

[1] Hermann Amandus Schwarz (1843-1921), mathématicien allemand.
[2] Pierre-Louis Lions (1956), mathématicien français.
[3] Bruno Després (1965), mathématicien français.



(a) Sousdomaines avec recouvrement.

(b) Sousdomaines sans recouvrement.

(c) Sousdomaines sans recouvrement.

FIGURE 1 : Quelques exemples typiques de décompositions en deux sousdomaines.

Nous décomposons $\overline{\Omega}$ en $\overline{\Omega} = \overline{\Omega}_1 \cup \overline{\Omega}_2$ (Ω_1 et Ω_2 peuvent se recouvrir ou non) comme indiqué dans la Figure 1. Pour $i = 1, 2$, on définit l'interface

$$\Sigma_i := \begin{cases} \partial\Omega_i \cap \Omega_{3-i} & \text{si } \Omega_1 \text{ et } \Omega_2 \text{ avec recouvrement,} \\ \partial\Omega_1 \cap \partial\Omega_2 =: \Sigma & \text{si } \Omega_1 \text{ et } \Omega_2 \text{ sans recouvrement,} \end{cases} \tag{12}$$

où on impose la condition de transmission ; et nous désignons également par ν_i le vecteur normal externe de Ω_i et $\mathcal{N}_i$ l'opérateur de trace conormale correspondant à ν_i.

La *méthode classique de Schwarz (avec conditions de transmission de Dirichlet)* et la *méthodes de Schwarz optimisée OSM (avec conditions de transmission de Robin)* s'écrivent alors respectivement

Algorithme 0.1: Méthode classique de Schwarz pour deux sousdomaines

Prendre des tentatives initiales $u_{i;0}$ dans Ω_i, $i = 1, 2$;
pour $\ell = 0, 1, ...$ **faire**
  pour $i = 1, 2$ **faire**
    Résoudre

$$\begin{cases} \mathcal{A}(u_{i;\ell+1}) = b & \text{dans } \Omega_i, \\ \mathcal{B}(u_{i;\ell+1}) = f & \text{sur } \partial\Omega_i \cap \partial\Omega, \\ u_{i;\ell+1} = u_{3-i;\ell} & \text{sur } \Sigma_i \end{cases}$$

    pour calculer $u_{i;\ell+1}$;
  fin
fin
Coller les solutions des sousproblèmes de manière appropriée

et

Algorithme 0.2: Méthode de Schwarz optimisée (OSM) pour deux sousdomaines

Prendre des tentatives initiales $u_{i;0}$ dans Ω_i, $i = 1, 2$;
pour $\ell = 0, 1, ...$ **faire**
  pour $i = 1, 2$ **faire**
    Résoudre

$$\begin{cases} \mathcal{A}(u_{i;\ell+1}) = b & \text{dans } \Omega_i, \\ \mathcal{B}(u_{i;\ell+1}) = f & \text{sur } \partial\Omega_i \cap \partial\Omega, \\ (\mathcal{N}_i + \beta)\, u_{i;\ell+1} = (\mathcal{N}_i + \beta)\, u_{3-i;\ell} & \text{sur } \Sigma_i \end{cases}$$

    pour calculer $u_{i;\ell+1}$;
  fin
fin
Coller les solutions des sousproblèmes de manière appropriée.

Dans le cas de la OSM avec des sousdomaines sans recouvrement (voir les Figures 1b et 1c), nous aimerions avoir une notation unique pour la dérivée conormale sur l'interface Σ nous réécrivons donc la condition de transmission du sous-problème dans Ω_2 comme suit. Pour $x \in \Sigma$, soit $\nu(x) = \nu_1(x)$ le vecteur normal externe de Ω_1 en x, donc $\nu_2(x) = -\nu(x)$ est le vecteur normal externe de Ω_2 en x. On note $\mathcal{N} = \mathcal{N}_1$ puis $\mathcal{N}_2 = -\mathcal{N}$. Avec la nouvelle notation, la condition de transmission

$$(\mathcal{N}_2 + \beta)\, u_{2;\ell+1} = (\mathcal{N}_2 + \beta)\, u_{1;\ell} \quad \text{sur } \Sigma$$

devient

$$(-\mathcal{N} + \beta)\, u_{2;\ell+1} = (-\mathcal{N} + \beta)\, u_{1;\ell} \quad \text{sur } \Sigma,$$

ce qui équivaut à

$$(\mathcal{N} - \beta)\, u_{2;\ell+1} = (\mathcal{N} - \beta)\, u_{1;\ell} \quad \text{sur } \Sigma.$$

La OSM dans le cas de sousdomaines sans recouvrement est donc réécrite comme

Algorithme 0.3: OSM pour deux sousdomaines sans recouvrement

Prendre des tentatives initiales $u_{i;0}$ dans Ω_i, $i = 1, 2$;
pour $\ell = 0, 1, ...$ **faire**
  Résoudre

$$\begin{cases} \mathcal{A}(u_{1;\ell+1}) = b & \text{dans } \Omega_1, \\ \mathcal{B}(u_{1;\ell+1}) = f & \text{sur } \partial\Omega_1 \cap \partial\Omega, \\ (\mathcal{N} + \beta)\, u_{1;\ell+1} = (\mathcal{N} + \beta)\, u_{2;\ell} & \text{sur } \Sigma \end{cases}$$

  pour calculer $u_{1;\ell+1}$;
  Résoudre

$$\begin{cases} \mathcal{A}(u_{2;\ell+1}) = b & \text{dans } \Omega_2, \\ \mathcal{B}(u_{2;\ell+1}) = f & \text{sur } \partial\Omega_2 \cap \partial\Omega, \\ (\mathcal{N} - \beta)\, u_{2;\ell+1} = (\mathcal{N} - \beta)\, u_{1;\ell} & \text{sur } \Sigma \end{cases}$$

  pour calculer $u_{2;\ell+1}$;
fin
Coller les solutions des sousproblèmes de manière appropriée.

Cependant, l'implémentation numérique des dérivées normales du membre de droite est problèmatique dans la pratique. Par exemple, dans la formulation faible du sousproblème dans Ω_1, nous devrions peut-être discrétiser $\int_\Sigma \frac{\partial u_{2;\ell}}{\partial \nu} \phi \, ds$ où $u_{2;\ell}$ est définie sur un maillage dans Ω_2 mais ϕ est une fonction test définie sur un autre maillage dans Ω_1. Une erreur

numérique se produira lorsque nous tenterons d'interpoler la dérivée normale entre deux maillages de côtés opposés, ce qui a un impact significatif sur la convergence (voir aussi [12, Section 2.3]). Par conséquent, il est plus pratique d'éviter de travailler avec une dérivée normale comme suit. Nous introduisons les *valeurs d'impédance* (ou *variables d'impédance*)

$$\lambda_{1;\ell+1} := (\mathcal{N}+\beta)\,u_{2;\ell} \text{ et } \lambda_{2;\ell+1} := (\mathcal{N}-\beta)\,u_{1;\ell} \text{ sur } \Sigma, \quad \ell = 0,1,... \tag{13}$$

alors les conditions de transmission sur Σ dans l'Algorithme 0.3 donnent

$$(\mathcal{N}+\beta)u_{1;\ell} = \lambda_{1;\ell} \text{ et } (\mathcal{N}-\beta)u_{2;\ell} = \lambda_{2;\ell} \text{ sur } \Sigma, \quad \ell = 0,1,... \tag{14}$$

Pour $\ell = 0,1,...$, on a

$$\lambda_{1;\ell+1} \overset{(13)}{=} (\mathcal{N}+\beta)\,u_{2;\ell} = (\mathcal{N}-\beta)\,u_{2;\ell} + 2\beta u_{2;\ell} \overset{(14)}{=} \lambda_{2;\ell} + 2\beta u_{2;\ell}$$

et

$$\lambda_{2;\ell+1} \overset{(13)}{=} (\mathcal{N}-\beta)\,u_{1;\ell} = (\mathcal{N}+\beta)\,u_{1;\ell} - 2\beta u_{1;\ell} \overset{(14)}{=} \lambda_{1;\ell} - 2\beta u_{1;\ell}.$$

sur Σ. La différence est que dans l' implémentation, nous sommes désormais capables de définir un maillage sur Σ pour les deux suites $(\lambda_{1;\ell})_{\ell\geq 0}$ et $(\lambda_{2;\ell})_{\ell\geq 0}$, et il suffit d'interpoler les solutions de sousproblèmes sur ce maillage. Nous réécrivons donc la OSM dans l'Algorithme 0.3 comme une itération sur les valeurs d'impédance λ_1 et λ_2, c.-à-d. comme une *méthode de sous-structuration* :

Algorithme 0.4: Version sous-structurée de la OSM pour deux sousdomaines sans recouvrement

Prendre des tentatives initiales $\lambda_{i;0}$, $i = 1,2$;
pour $\ell = 0,1,...$ **faire**
  Résoudre

$$\begin{cases} \mathcal{A}(u_{1;\ell}) = b & \text{in } \Omega_1, \\ \mathcal{B}(u_{1;\ell}) = f & \text{on } \partial\Omega_1 \cap \partial\Omega, \\ (\mathcal{N}+\beta)\,u_{1;\ell} = \lambda_{1;\ell} & \text{on } \Sigma \end{cases}$$

  pour calculer $u_{1;\ell}$;
  Résoudre

$$\begin{cases} \mathcal{A}(u_{2;\ell}) = b & \text{in } \Omega_2, \\ \mathcal{B}(u_{2;\ell}) = f & \text{on } \partial\Omega_2 \cap \partial\Omega, \\ (\mathcal{N}-\beta)\,u_{2;\ell} = \lambda_{2;\ell} & \text{on } \Sigma \end{cases}$$

  pour calculer $u_{2;\ell}$;
  $\lambda_{1;\ell+1} = \lambda_{2;\ell} + 2\beta u_{2;\ell}$, $\lambda_{2;\ell+1} = \lambda_{1;\ell} - 2\beta u_{1;\ell}$;
fin
Coller les solutions des sousproblèmes de manière appropriée.

Dans la section suivante, nous discuterons la convergence des Algorithmes 0.1–0.4.

0.3.2 Sur la convergence des méthodes de décomposition de domaine

Un aperçu de la convergence des méthodes de Schwarz classiques et des OSM peut être trouvé respectivement dans [12, Chapitre 1] et [12, Chapitre 2]. Ici, nous aimerions discuter en détail la convergence de deux modèles typiques. Le premier, également présenté dans [12], est l'équation

$$(-\triangle + \eta)u = \tilde{f} \text{ dans } \mathbb{R}^2 \quad (\eta \in \mathbb{R}, \tilde{f} \text{ est le terme de source})$$

avec une condition convenable à l'infini (selon le signe de η) qui garantit l'unicité du problème. On décompose le plan $\mathbb{R}^2$ en deux demi-plans

$$\Omega_1 = (-\infty, \delta) \times \mathbb{R} \text{ et } \Omega_2 = (0, \infty) \times \mathbb{R} \tag{15}$$

où $\delta \geq 0$ est la taille du recouvrement, et $\delta = 0$ signifie qu'il n'y a pas de recouvrement. Pour les interfaces où l'on impose les conditions aux limites de transmission, on continue d'utiliser les notations Σ_i définies dans (12) ; dans ce cas on a

$$\Sigma_1 = \{(x, y) \in \mathbb{R}^2 : x = \delta\} \text{ et } \Sigma_2 = \{(x, y) \in \mathbb{R}^2 : x = 0\}. \tag{16}$$

Ce modèle à deux demi-plans est typique puisqu'après avoir effectué la transformée de Fourier dans la direction y, les erreurs dans les sousdomaines ont des expressions explicites (qui dépendent du signe de η). Nous nous intéressons également à un deuxième modèle typique avec un domaine borné constitué de deux cercles concentriques (voir par exemple [7, Section 3.1]). Soit la fonction constante par morceaux σ définie dans le cercle $B(0, R_2) \subset \mathbb{R}^2$, qui prend la valeur $\sigma_1 \in \mathbb{R}$ dans le petit cercle $B(0, R_1)$, $R_1 < R_2$, et la valeur $\sigma_2 \in \mathbb{R}$ dans l'anneau $R_1 < \sqrt{x^2 + y^2} < R_2$, on considère l'équation

$$-\operatorname{div}(\sigma \nabla u) = 0 \text{ dans } B(0, R_2) \subset \mathbb{R}^2$$

avec la condition aux limites de Dirichlet ou de Neumann. Pour les DDM sans recouvrement, nous utilisons directement le cercle $B(0, R_1)$ et l'anneau $R_1 < \sqrt{x^2 + y^2} < R_2$ comme sousdomaines (voir Figure 1c). Dans ce modèle, les erreurs du sous-problème ont des expressions explicites (sous forme de série de Fourier). L'étude des erreurs de sous-problèmes dans ces deux modèles montre que la convergence des DDM concerne certaines suites géométriques et est donc caractérisée par ce que l'on appelle le *facteur de convergence* associé aux ratios de ces suites. Les preuves de convergence pour des configurations géométriques plus générales sont activement étudiées dans la littérature. Une preuve de convergence plus générale pour la OSM sans recouvrement avec des conditions de transmission du second ordre pour les équations elliptiques est présentée dans [12, Section 2.1.2]. Récemment, Gong et al [15] ont prouvé la convergence des OSMs pour l'équation de Helmholtz avec des décompositions de domaines par bandes. De plus, dans le cas sans recouvrement, de nombreux articles récents, par exemple [8], se concentrent sur le problème de convergence avec la présence de points de croisement, c.-à-d. des points où trois sousdomaines ou plus sont adjacents.

Pour donner aux lecteurs une idée de la preuve de convergence, nous présentons brièvement l'analyse de convergence pour le premier modèle dans deux cas : le cas d'une équation elliptique ($\eta > 0$), que l'on retrouve pour la méthode classique de Schwarz dans [12, Section 1.5.2] et pour la OSM dans [12, Section 2.1.1] ; et le cas de l'équation de Helmholtz ($\eta < 0$), que l'on retrouve pour la méthode classique de Schwarz dans [12, Section 2.2.1] et pour la OSM dans [12, Section 2.2.2]. Cela permettra de mieux comprendre les avantages de l'OSM par rapport à la méthode Schwarz classique.

Le cas d'une équation elliptique

Soit $\eta > 0$, nous considérons le problème global

$$\begin{cases} (-\triangle + \eta)u = \tilde{f} & \text{dans } \mathbb{R}^2, \\ u \text{ est bornée} & \text{à l'infini} \end{cases}$$

pour lequel la méthode classique de Schwarz et la OSM s'écrivent respectivement sous la forme

$$i=1,2: \quad \begin{cases} (-\triangle+\eta)u_{i;\ell+1}=\tilde{f} & \text{dans } \Omega_i, \\ u_{i;\ell+1} \text{ est bornée} & \text{à l'infini}, \\ u_{i;\ell+1}=u_{3-i;\ell} & \text{sur } \Sigma_i \end{cases}$$

et

$$i=1,2: \quad \begin{cases} (-\triangle+\eta)u_{i;\ell+1}=\tilde{f} & \text{dans } \Omega_i, \\ u_{i;\ell+1} \text{ est bornée} & \text{à l'infini}, \\ \left(\frac{\partial}{\partial \nu_i}+\beta\right)u_{i;\ell+1}=\left(\frac{\partial}{\partial \nu_i}+\beta\right)u_{3-i;\ell} & \text{sur } \Sigma_i \end{cases}$$

où $\Omega_i, \Sigma_i, i=1,2$, sont respectivement définis dans (15) et (16). Nous définissons les erreurs $e_{i;\ell} \coloneqq u_i^\ell - u|_{\Omega_i}$ et nous nous intéressons à leurs transformées de Fourier dans la direction y notées $\hat{e}_{i;\ell}=\hat{e}_{i;\ell}(x,\xi)$, qui sont définies par

$$\hat{a}(x,\xi)=(\mathcal{F}a)(x,\xi) \coloneqq \int_{\mathbb{R}} a(x,y)\mathrm{e}^{-\mathrm{i}\xi y}\,\mathrm{d}y.$$

Il est facile de vérifier que $\hat{e}_{i;\ell}, i=1,2$, satisfont

$$\begin{cases} \left(-\frac{\partial^2}{\partial x^2}+\xi^2+\eta\right)\hat{e}_{1;\ell+1}=0, & x<\delta, \xi\in\mathbb{R}, \\ \hat{e}_{1;\ell+1}(\delta,\xi)=\hat{e}_{2;\ell}(\delta,\xi), & \forall \xi\in\mathbb{R}, \\ \left(-\frac{\partial^2}{\partial x^2}+\xi^2+\eta\right)\hat{e}_{2;\ell+1}=0, & x>0, \xi\in\mathbb{R}, \\ \hat{e}_{2;\ell+1}(0,\xi)=\hat{e}_{1;\ell}(0,\xi), & \forall \xi\in\mathbb{R} \end{cases} \tag{17}$$

pour la méthode classique de Schwarz, et

$$\begin{cases} \left(-\frac{\partial^2}{\partial x^2}+\xi^2+\eta\right)\hat{e}_{1;\ell+1}=0, & x<\delta, \xi\in\mathbb{R}, \\ \left(\frac{\partial}{\partial x}+\beta\right)(\hat{e}_{1;\ell+1})(\delta,\xi)=\left(\frac{\partial}{\partial x}+\beta\right)(\hat{e}_{2;\ell})(\delta,\xi), & \forall \xi\in\mathbb{R}, \\ \left(-\frac{\partial^2}{\partial x^2}+\xi^2+\eta\right)\hat{e}_{2;\ell+1}=0, & x>0, \xi\in\mathbb{R}, \\ \left(-\frac{\partial}{\partial x}+\beta\right)(\hat{e}_{2;\ell+1})(0,\xi)=\left(-\frac{\partial}{\partial x}+\beta\right)(\hat{e}_{1;\ell})(0,\xi), & \forall \xi\in\mathbb{R} \end{cases} \tag{18}$$

pour la OSM. Puisque les solutions sont également bornées à l'infini, les solutions fondamentales de (17) et (18) pour $\xi\in\mathbb{R}$ fixe sont de la forme

$$\hat{e}_{1;\ell}(x,\xi)=\hat{e}_{1;\ell}(\delta,\xi)\mathrm{e}^{\lambda(\xi)(x-\delta)}, x<\delta \quad \text{et} \quad \hat{e}_{2;\ell}(x,\xi)=\hat{e}_{2;\ell}(0,\xi)\mathrm{e}^{-\lambda(\xi)x}, x>0 \tag{19}$$

où

$$\lambda(\xi) \coloneqq \sqrt{\xi^2+\eta}. \tag{20}$$

Pour la convergence, il suffit de se concentrer sur $\hat{e}_{1;\ell}(\delta,\xi)$ et $\hat{e}_{2;\ell}(0,\xi)$. Des conditions aux limites de transmission, on déduit que $\hat{e}_{1;\ell}(\delta,\xi)$, $\ell\geq 0$ et $\hat{e}_{2;\ell}(0,\xi)$, $\ell\geq 0$ se comportent comme deux suites géométriques de même ratio $\tilde{r}$:

$$\hat{e}_{1;\ell+1}(\delta,\xi)=\tilde{r}^2\hat{e}_{1;\ell-1}(\delta,\xi) \quad \text{et} \quad \hat{e}_{2;\ell+1}(0,\xi)=\tilde{r}^2\hat{e}_{2;\ell-1}(0,\xi), \quad \forall \ell\geq 1,$$

où

$$\tilde{r}=\begin{cases} \tilde{r}(\xi,\delta) \coloneqq \mathrm{e}^{-\lambda(\xi)\delta} & \text{pour la méthode classique de Schwarz}, \\ \tilde{r}(\xi,\delta;\beta) \coloneqq \dfrac{\lambda(\xi)-\beta}{\lambda(\xi)+\beta}\mathrm{e}^{-\lambda(\xi)\delta} & \text{pour la OSM}. \end{cases}$$

En utilisant la définition de $\lambda(\xi)$ dans (20), on obtient ainsi le *facteur de convergence* $\rho = |\tilde{r}|$ donné par

$$\rho = \begin{cases} \rho(\xi,\delta) := \mathrm{e}^{-\delta\sqrt{\xi^2+\eta}} & \text{pour la méthode classique de Schwarz,} \\ \rho(\xi,\delta;\beta) := \left|\dfrac{\sqrt{\xi^2+\eta}-\beta}{\sqrt{\xi^2+\eta}+\beta}\right| \mathrm{e}^{-\delta\sqrt{\xi^2+\eta}} & \text{pour l' OSM.} \end{cases} \tag{21}$$

La condition générale de convergence est $\rho < 1$. À partir de la formule (21) du facteur de convergence, nous avons les conclusions suivantes sur le modèle ($\eta > 0$).

- La méthode classique de Schwarz converge lorsqu'il y a du recouvrement ($\delta > 0$) et en plus, on a même une convergence uniforme puisque $\rho < \mathrm{e}^{-\delta\sqrt{\eta}} < 1$, $\forall \xi \in \mathbb{R}$.
- La méthode classique de Schwarz diverge lorsqu'il n'y a pas de recouvrement ($\delta = 0$).
- La OSM converge toujours, avec ou sans recouvrement. Lorsqu'il y a du recouvrement, grâce au coefficient de Robin β, la OSM converge uniformément avec un facteur de convergence plus petit et donc plus rapide que la méthode classique de Schwarz.

Le cas de l'équation de Helmholtz

On prend maintenant en compte le cas de l'équation de Helmholtz : $\eta = -\omega^2 < 0$ pour $\omega > 0$, ω est appelé *le nombre d'onde*. Avant d'aborder le modèle à deux demi-plans, il faut remarquer que la méthode classique de Schwarz n'est pas toujours applicable à l'équation de Helmholtz : pour certains nombres d'onde ω, les sousproblèmes peuvent être mal-posés même lorsque le problème global est bien-posé. Par exemple, considérons le problème global

$$\begin{cases} (-\triangle - \omega^2)u = \tilde{f} & \text{dans un domaine } \tilde{\Omega}, \\ u = \tilde{g} & \text{sur } \partial\tilde{\Omega} \end{cases} \tag{22}$$

pour un nombre d'onde ω donné tel que ce problème est bien posé. Nous décomposons $\tilde{\Omega}$ comme $\overline{\tilde{\Omega}} = \overline{\tilde{\Omega}}_1 \cup \overline{\tilde{\Omega}}_2$ où $\tilde{\Omega}_1$ et $\tilde{\Omega}_2$ sont des sousdomaines avec recouvrement comme dans la Figure 2. Pour $i = 1, 2$, nous définissons l'interface $\tilde{\Sigma}_i := \partial\tilde{\Omega}_i \cap \tilde{\Omega}_{3-i}$ où nous imposons la condition de transmission, et nous désignons par $\tilde{\nu}_i$ le vecteur normal externe de $\tilde{\Omega}_i$.

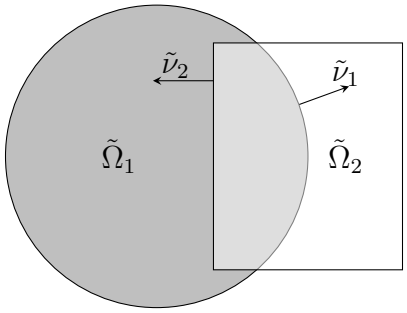


FIGURE 2 : Sousdomaines $\tilde{\Omega}_1$ et $\tilde{\Omega}_2$ avec recouvrement.

La méthode classique de Schwarz et la OSM s'écrivent respectivement sous la forme

$$i = 1, 2 : \quad \begin{cases} (-\triangle - \omega^2)u_{i;\ell+1} = \tilde{f} & \text{dans } \tilde{\Omega}_i, \\ u_{i;\ell+1} = \tilde{g} & \text{sur } \partial\tilde{\Omega}_i \cap \partial\tilde{\Omega}, \\ u_{i;\ell+1} = u_{3-i;\ell} & \text{sur } \tilde{\Sigma}_i \end{cases} \tag{23}$$

et

$$i=1,2: \quad \begin{cases} (-\triangle-\omega^2)u_{i;\ell+1}=\tilde{f} & \text{dans } \tilde{\Omega}_i, \\ u_{i;\ell+1}=\tilde{g} & \text{sur } \partial\tilde{\Omega}_i\cap\partial\tilde{\Omega}, \\ \left(\frac{\partial}{\partial\nu_i}+\beta\right)u_{i;\ell+1}=\left(\frac{\partial}{\partial\nu_i}+\beta\right)u_{3-i;\ell} & \text{sur } \tilde{\Sigma}_i. \end{cases} \tag{24}$$

Nous nous concentrerons sur le sousproblème sur $\tilde{\Omega}_1$ dans les schémas (23) et (24). L'unicité du sousproblème sur $\tilde{\Omega}_1$ défini dans (23) équivaut à l'unicité du problème homogène

$$\begin{cases} (-\triangle-\omega^2)v=0 & \text{dans } \tilde{\Omega}_1, \\ v=0 & \text{sur } \partial\tilde{\Omega}_1. \end{cases} \tag{25}$$

Si ω est tel qu'il existe $v\neq 0$ satisfaisant (25), c.-à-d. que ω^2 est une valeur propre de Dirichlet de l'opérateur laplacien, le sousproblème sur $\tilde{\Omega}_1$ défini dans (23) n'est pas bien-posé. Nous remarquons que pour toute géométrie de $\tilde{\Omega}_1$, il existe toujours des valeurs propres de Dirichlet, connues sous le nom de résonances pour $\tilde{\Omega}_1$. En revanche, l'unicité du sousproblème sur $\tilde{\Omega}_1$ défini dans (24) équivaut à l'unicité du problème homogène

$$\begin{cases} (-\triangle-\omega^2)v=0 & \text{in } \tilde{\Omega}_1, \\ v=0 & \text{on } \partial\tilde{\Omega}_1\cap\partial\tilde{\Omega}, \\ \left(\frac{\partial}{\partial\tilde{\nu}}+\beta\right)v=0 & \text{on } \tilde{\Sigma}_1. \end{cases} \tag{26}$$

Avec les conditions de Robin, pour tout $\omega>0$, nous avons que $v=0$ est la solution unique à (26). En effet, si on multiplie la première équation dans (26) par $\overline{v}$ et on intègre par parties en utilisant les conditions aux limites, on obtient

$$\begin{aligned} 0 &= \int_{\tilde{\Omega}_1}(-\triangle-\omega^2)v\overline{v}\,\mathrm{dx} \\ &= -\int_{\partial\tilde{\Omega}_1}\frac{\partial v}{\partial\tilde{\nu}}\overline{v}\,\mathrm{ds}+\int_{\tilde{\Omega}_1}|\nabla v|^2\,\mathrm{dx}-\int_{\tilde{\Omega}_1}\omega^2|v|^2\,\mathrm{dx} \\ &= \int_{\tilde{\Sigma}_1}\beta|v|^2\,\mathrm{ds}+\int_{\tilde{\Omega}_1}|\nabla v|^2\,\mathrm{dx}-\int_{\tilde{\Omega}_1}\omega^2|v|^2\,\mathrm{dx}. \end{aligned}$$

La partie imaginaire de la dernière ligne de l'équation implique $v=0$ sur $\tilde{\Sigma}_1$, si le coefficient Robin β est choisi comme $\beta=\mathrm{i}\tilde{\beta}$ pour certains $\tilde{\beta}>0$. En utilisant la condition de Robin, cela donne $\frac{\partial v}{\partial\tilde{\nu}}=0$ sur $\tilde{\Sigma}_1$. Par le Théorème 1.A.2, une application du principe de continuation unique, nous devons avoir $v=0$ dans $\tilde{\Omega}_1$. Par conséquent, les méthodes classiques de Schwarz ne sont pas toujours bien définies pour l'équation de Helmholtz alors que la OSM corrige cet inconvénient tant que l'on choisit un coefficient Robin β approprié.

Revenons maintenant à l'analyse de convergence pour le modèle à deux demi-plans, considérons le problème global

$$\begin{cases} (-\triangle-\omega^2)u=\tilde{f} & \text{dans } \mathbb{R}^2, \\ \lim\limits_{r\to\infty}\sqrt{r}\left(\frac{\partial u}{\partial r}+\mathrm{i}\omega u\right)=0 & \text{condition de Sommerfeld, } r:=\sqrt{x^2+y^2}. \end{cases}$$

La méthode classique de Schwarz s'écrit alors

$$i=1,2: \quad \begin{cases} (-\triangle-\omega^2)u_{i;\ell+1}=\tilde{f} & \text{dans } \Omega_i, \\ \lim\limits_{r\to\infty}\sqrt{r}\left(\frac{\partial u_{i;\ell+1}}{\partial r}+\mathrm{i}\omega u_{i;\ell+1}\right)=0 & \text{condition de Sommerfeld,} \\ u_{i;\ell+1}=u_{3-i;\ell} & \text{sur } \Sigma_i \end{cases}$$

et la OSM avec le coefficient de Robin $\beta = \mathrm{i}\tilde{\beta}$ ($\tilde{\beta} > 0$), qui est connue sous le nom de *méthode de Després*, s'écrit

$$i = 1, 2 : \quad \begin{cases} (-\triangle - \omega^2) u_{i;\ell+1} = \tilde{f} & \text{dans } \Omega_i, \\ \lim\limits_{r\to\infty} \sqrt{r} \left(\frac{\partial u_{i;\ell+1}}{\partial r} + \mathrm{i}\omega u_{i;\ell+1} \right) = 0 & \text{condition de Sommerfeld,} \\ \left(\frac{\partial}{\partial \nu_i} + \mathrm{i}\tilde{\beta} \right) u_{i;\ell+1} = \left(\frac{\partial}{\partial \nu_i} + \mathrm{i}\tilde{\beta} \right) u_{3-i;\ell} & \text{sur } \Sigma_i, \end{cases}$$

où Ω_i, Σ_i, $i = 1, 2$ sont respectivement définis dans (15) et (16). Dans ce cas, tous les sousproblèmes des deux méthodes sont bien définis. En répétant les mêmes arguments que dans la section précédente, nous nous concentrons sur $\hat{e}_{i;\ell}, i = 1, 2$, la transformée de Fourier dans la direction y des erreurs, qui satisfont

$$\begin{cases} \left(-\frac{\partial^2}{\partial x^2} + \xi^2 - \omega^2 \right) \hat{e}_{1;\ell+1} = 0, & x < \delta, \xi \in \mathbb{R}, \\ \hat{e}_{1;\ell+1}(\delta, \xi) = \hat{e}_{2;\ell}(\delta, \xi), & \forall \xi \in \mathbb{R}, \\ \left(-\frac{\partial^2}{\partial x^2} + \xi^2 - \omega^2 \right) \hat{e}_{2;\ell+1} = 0, & x > 0, \xi \in \mathbb{R}, \\ \hat{e}_{2;\ell+1}(0, \xi) = \hat{e}_{1;\ell}(0, \xi), & \forall \xi \in \mathbb{R} \end{cases} \tag{27}$$

pour la méthode classique de Schwarz, et

$$\begin{cases} \left(-\frac{\partial^2}{\partial x^2} + \xi^2 - \omega^2 \right) \hat{e}_{1;\ell+1} = 0, & x < \delta, \xi \in \mathbb{R}, \\ \left(\frac{\partial}{\partial x} + \mathrm{i}\tilde{\beta} \right) (\hat{e}_{1;\ell+1})(\delta, \xi) = \left(\frac{\partial}{\partial x} + \mathrm{i}\tilde{\beta} \right) (\hat{e}_{2;\ell})(\delta, \xi), & \forall \xi \in \mathbb{R}, \\ \left(-\frac{\partial^2}{\partial x^2} + \xi^2 - \omega^2 \right) \hat{e}_{2;\ell+1} = 0, & x > 0, \xi \in \mathbb{R}, \\ \left(-\frac{\partial}{\partial x} + \mathrm{i}\tilde{\beta} \right) (\hat{e}_{2;\ell+1})(0, \xi) = \left(-\frac{\partial}{\partial x} + \mathrm{i}\tilde{\beta} \right) (\hat{e}_{1;\ell})(0, \xi), & \forall \xi \in \mathbb{R} \end{cases} \tag{28}$$

pour la OSM. Par la condition de Sommerfeld, les solutions fondamentales de (27) et (28) ont la même forme que dans (19), c'est pour $\xi \in \mathbb{R}$ fixe,

$$\hat{e}_{1;\ell}(x, \xi) = \hat{e}_{1;\ell}(\delta, \xi) \mathrm{e}^{\lambda(\xi)(x-\delta)}, x < \delta \quad \text{et} \quad \hat{e}_{2;\ell}(x, \xi) = \hat{e}_{2;\ell}(0, \xi) \mathrm{e}^{-\lambda(\xi)x}, x > 0$$

mais dans ce cas,

$$\lambda(\xi) := \begin{cases} \mathrm{i}\sqrt{\omega^2 - \xi^2} & \text{si } |\xi| < \omega, \\ \sqrt{\xi^2 - \omega^2} & \text{si } |\xi| \geq \omega. \end{cases} \tag{29}$$

On observe que $\lambda(\xi)$ défini dans (29) peut être complexe alors que $\lambda(\xi)$ défini dans (20) est toujours réel. Encore une fois, à partir des conditions aux limites de transmission, nous déduisons que $\hat{e}_{1;\ell}(\delta, \xi)$ et $\hat{e}_{2;\ell}(0,\ xi)$, $\ell \geq 0$ se comportent comme deux suites géométriques de même ratio $\tilde{r}$:

$$\hat{e}_{1;\ell+1}(\delta, \xi) = \tilde{r}^2 \hat{e}_{1;\ell-1}(\delta, \xi) \quad \text{et} \quad \hat{e}_{2;\ell+1}(0, \xi) = \tilde{r}^2 \hat{e}_{2;\ell-1}(0, \xi), \quad \forall \ell \geq 1$$

où

$$\tilde{r} = \begin{cases} \tilde{r}(\xi, \delta) := \mathrm{e}^{-\lambda(\xi)\delta} & \text{pour la méthode Schwarz classique,} \\ \tilde{r}(\xi, \delta) := \dfrac{\lambda(\xi) - \mathrm{i}\tilde{\beta}}{\lambda(\xi) + \mathrm{i}\tilde{\beta}} \mathrm{e}^{-\lambda(\xi)\delta} & \text{pour la OSM.} \end{cases}$$

Le facteur de convergence $\rho = |\tilde{r}|$ est donc donné par

$$\boxed{\rho = \rho(\xi, \delta) := \left| \mathrm{e}^{-\lambda(\xi)\delta} \right| = \begin{cases} 1 & \text{si } |\xi| \leq \omega, \\ \mathrm{e}^{-\delta\sqrt{\xi^2 - \omega^2}} & \text{si } |\xi| > \omega \end{cases}} \tag{30}$$

pour la méthode classique de Schwarz, et

$$\rho = \rho(\xi, \delta; \tilde{\beta}) := \left| \frac{\lambda(\xi) - \mathrm{i}\tilde{\beta}}{\lambda(\xi) + \mathrm{i}\tilde{\beta}} \mathrm{e}^{-\lambda(\xi)\delta} \right| = \begin{cases} \left| \dfrac{\tilde{\beta} - \sqrt{\omega^2 - \xi^2}}{\tilde{\beta} + \sqrt{\omega^2 - \xi^2}} \right| & \text{si } |\xi| \leq \omega, \\ \mathrm{e}^{-\delta\sqrt{\xi^2 - \omega^2}} & \text{si } |\xi| > \omega \end{cases} \tag{31}$$

pour la OSM, où nous avons utilisé la définition de $\lambda(\xi)$ dans (29). À partir des formules (30) et (31) du facteur de convergence dans chaque méthode, nous avons les conclusions suivantes sur le modèle ($\eta = -\omega^2 < 0$).

- La méthode classique de Schwarz avec du recouvrement ($\delta > 0$) diverge pour les modes propagatifs $|\xi| \leq \omega$ et converge pour les modes évanescents $|\xi| > \omega$.
- La méthode classique de Schwarz sans recouvrement ($\delta = 0$) diverge pour tous les modes.
- La OSM avec du recouvrement converge pour tous les modes sauf pour $\xi = \pm\omega$.
- La OSM sans recouvrement converge pour les modes propagatifs $|\xi| < \omega$ et diverge pour les modes évanescents $|\xi| \geq \omega$.

0.4 Application au problème inverse de conductivité

0.4.1 Problème inverse de conductivité non-linéaire

En physique, le *potentiel électrique (ou la tension)* u satisfait l'équation

$$-\operatorname{div}(\sigma \nabla u) = 0 \text{ dans le domaine } \Omega \tag{32}$$

où la *conductivité* σ appartient à l'ensemble admissible

$$V(\Omega) := \{\sigma \in L^\infty(\Omega) : \underline{\sigma} < \sigma < \overline{\sigma}\}$$

pour certaines constantes $\underline{\sigma}, \overline{\sigma} > 0$ données. L'équation (32) est appelée l'*équation de conductivité*. Étant donné une condition aux limites de Dirichlet

$$u = f \text{ sur } \partial\Omega,$$

le potentiel électrique $u \in H^1(\Omega)$ est déterminé de manière unique à partir de $f \in H^{\frac{1}{2}}(\partial\Omega)$ si nous avons $\partial\Omega$ lipschitzien. De plus, u dépend continuement de f dans les normes de ces espaces donc nous avons un problème bien-posé. On suppose souvent par ailleurs que σ est constant près de $\partial\Omega$. Ensuite, $\frac{\partial u}{\partial \nu} \in H^{-\frac{1}{2}}(\partial\Omega)$, et la trace conormale

$$\sigma \frac{\partial u}{\partial \nu} = g \text{ sur } \partial\Omega$$

est bien definie. Le *problème de conductivité inverse* consiste à reconstruire σ à partir de la connaissance d'une ou plusieurs mesure(s) g. Un bref historique du problème de conductivité inverse est présenté dans [27, Section 1.2].

Dans nos travaux, nous nous intéressons à l'équation un peu plus générale

$$-\operatorname{div}(\sigma \nabla u) + \eta u = 0 \text{ dans } \Omega \tag{33}$$

où η est une fonction dans $L^\infty(\Omega)$. Comme mentionné dans la section précédente, si $\eta = -\omega^2 < 0$ dans Ω, l'équation (33) est appelée l'*équation de Helmholtz* et ω est appelé *le nombre d'onde*. Si $\eta > 0$ dans Ω, l'équation (33) est elliptique et appelée l'*équation de Laplace screening* (ou l'*équation de Poisson screening* lorsqu'il y a un terme de source, ou même appelée l'*équation de Helmholtz avec le bon signe*, voir par exemple [13]). La solution u de l'équation (33) a une trace u et une trace conormale $\sigma\frac{\partial u}{\partial \nu}$ continues, ce qui est important dans les algorithmes DDM presentés dans la section précédente.

Remarque 0.4.1. Ci-dessus, nous avons décrit le problème direct avec une condition aux limites de Dirichlet et le problème inverse avec des mesures de Neumann. De plus, nous considérons également le problème direct avec une condition aux limites de Neumann et le problème inverse avec des mesures de Dirichlet. Dans ce dernier cas, lorsque $\eta = 0$, nous devons veiller à ce que le problème direct soit bien-posé. Pour l'existence, on sait que le terme source (qui est nul dans notre cas) et la trace conormale g doivent satisfaire *la condition de compatibilité*

$$\int_{\partial\Omega} g \,\mathrm{ds} = 0.$$

Pour l'unicité, encore une fois lorsque $\eta = 0$, nous avons besoin d'une contrainte supplémentaire $\int_\Omega u \,\mathrm{dx} = 0$. Dans les sections suivantes, bien que cela ne soit pas explicitement mentionné, la condition et la contrainte de compatibilité ci-dessus seront implicites lorsque $\eta = 0$.

0.4.2 Linéarisation par approximation de Born

Il est possible de linéariser le problème inverse de conductivité comme suit. La solution u de l'équation (33) est également appelée le *champ total* et maintenant renommée u^{tot} :

$$\begin{cases} -\operatorname{div}(\sigma\nabla u^{\mathrm{tot}}) + \eta u^{\mathrm{tot}} = 0 & \text{dans } \Omega, \\ u^{\mathrm{tot}} = f & \text{sur } \partial\Omega. \end{cases} \tag{34}$$

Soit $\sigma_0 \in L^\infty(\Omega)$ telle que $\sigma_0 = \sigma$ dans un voisinage de $\partial\Omega$. Nous définissons le *champ incident* u^{inc} tel que

$$\begin{cases} -\operatorname{div}(\sigma_0\nabla u^{\mathrm{inc}}) + \eta u^{\mathrm{inc}} = 0 & \text{dans } \Omega, \\ u^{\mathrm{inc}} = f & \text{sur } \partial\Omega. \end{cases} \tag{35}$$

Le *champ diffracté* u^{scat} défini par la différence entre le champ total dans (34) et le champ incident dans (35), c.-à-d. $u^{\mathrm{scat}} \coloneqq u^{\mathrm{tot}} - u^{\mathrm{inc}}$, satisfait

$$\begin{cases} -\operatorname{div}(\sigma_0\nabla u^{\mathrm{scat}}) + \eta u^{\mathrm{scat}} = -\operatorname{div}(\tilde{\sigma}\nabla u^{\mathrm{inc}}) - \operatorname{div}(\tilde{\sigma}\nabla u^{\mathrm{scat}}) & \text{dans } \Omega, \\ u^{\mathrm{scat}} = 0 & \text{sur } \partial\Omega, \end{cases} \tag{36}$$

où $\tilde{\sigma} \coloneqq \sigma_0 - \sigma$ est le *contraste de conductivité*. On remarque que $\tilde{\sigma} = 0$ dans un voisinage de $\partial\Omega$. L'*approximation de Born* dit que nous pouvons ignorer $\operatorname{div}(\tilde{\sigma}\nabla u^{\mathrm{scat}})$ dans l'équation (36) lorsque $\tilde{\sigma} \ll 1$. On obtient ainsi le *champ linéarisé* u^{lin} :

$$\begin{cases} -\operatorname{div}(\sigma_0\nabla u^{\mathrm{lin}}) + \eta u^{\mathrm{lin}} = -\operatorname{div}(\tilde{\sigma}\nabla u^{\mathrm{inc}}) & \text{dans } \Omega, \\ u^{\mathrm{lin}} = 0 & \text{sur } \partial\Omega. \end{cases} \tag{37}$$

Notons que la solution u^{lin} a une trace continue u^{lin} et une trace conormale $\sigma_0\frac{\partial u^{\mathrm{lin}}}{\partial \nu}$. Maintenant, avec la mesure de Neumann $g \coloneqq \sigma_0\frac{\partial u^{\mathrm{lin}}}{\partial \nu}\big|_{\partial\Omega}$, l'application $\tilde{\sigma} \mapsto g$ est linéaire. La procédure de linéarisation est similaire si l'on remplace la condition de Dirichlet par la condition de Neumann.

0.4.3 Un aperçu de l'application

En résumé, nous formulons les problèmes inverses de conductivité comme suit, où, par abus de notations, dans le problème linéarisé nous notons $(u^{\text{lin}}, \tilde{\sigma})$ par (u, σ), qui sont utilisés dans le cadre abstrait présenté dans la Section 0.2. Le problème inverse de conductivité non linéaire est

$$\boxed{\begin{array}{ll} \text{problème direct :} & \begin{cases} -\operatorname{div}(\sigma\nabla u) + \eta u = 0 \text{ dans } \Omega, \\ u = f \text{ sur } \partial\Omega \quad (\text{ou } \sigma\frac{\partial u(\sigma)}{\partial \nu} = f \text{ sur } \partial\Omega), \end{cases} \\ \text{problème inverse :} & \eta \text{ donné,} \\ & \text{input } f, \\ & \text{mesurer } g = \sigma\frac{\partial u(\sigma)}{\partial \nu}\Big|_{\partial\Omega} \text{ (ou respectivement } g = u(\sigma)|_{\partial\Omega}), \\ & \text{retrouver } \sigma \end{array}} \tag{38}$$

et le problème inverse de conductivité linéarisé est

$$\boxed{\begin{array}{ll} \text{problème direct :} & \begin{cases} -\operatorname{div}(\sigma_0\nabla u) + \eta u = -\operatorname{div}(\sigma\nabla u_0) \text{ dans } \Omega, \\ u = 0 \text{ sur } \partial\Omega \quad (\text{ou } \sigma_0\frac{\partial u(\sigma)}{\partial \nu} = 0 \text{ sur } \partial\Omega), \end{cases} \\ \text{où} & \sigma = 0 \text{ dans un voisinage de } \partial\Omega \\ \text{et} & \begin{cases} -\operatorname{div}(\sigma_0\nabla u_0) + \eta u_0 = 0 \text{ dans } \Omega, \\ u_0 = f \text{ sur } \partial\Omega \quad (\text{ou respectivement } \sigma_0\frac{\partial u_0}{\partial \nu} = f \text{ sur } \partial\Omega), \end{cases} \\ \text{problème inverse :} & \sigma_0 \text{ et } \eta \text{ donnés,} \\ & \text{input } f, \\ & \text{mesurer } g = \sigma_0\frac{\partial u(\sigma)}{\partial \nu}\Big|_{\partial\Omega} \text{ (ou respectivement } g = u(\sigma)|_{\partial\Omega}), \\ & \text{retrouver } \sigma. \end{array}} \tag{39}$$

Malheureusement, le problème inverse de conductivité est très mal-posé dans de nombreux cas, même lorsque le problème est linéarisé. En effet, la solution du problème direct dans (39), lorsque Ω est un cercle dans $\mathbb{R}^2$, est de la forme $u = A\sigma$ où A est un opérateur linéaire inversible, et on peut montrer que la suite des valeurs singulières de A décroît de façon exponentielle, ce qui entraîne des difficultés significatives pour inverser A. La relation entre le nombre de mesures et les inconnues est également un problème : on sait qu'on ne peut pas retrouver σ avec juste un nombre fini de mesures. Des discussions sur ce type de problème peuvent être trouvées dans [20], [21], [6] et [28]. Dans nos expériences numériques, nous considérerons un cas simplifié où la conductivité σ est une fonction $\mathbb{P}^0$ avec un support connu et le nombre de mesures est convenablement conçu pour garantir l'unicité du problème inverse. L'application des méthodes de type one-shot à pas multiples (8) au problème inverse de conductivité linéarisée avec cette configuration est présentée dans la Section 3.2. Ensuite, la combinaison avec les DDMs sera étudiée au Chapitre 5 pour le problème inverse linéarisé (39) et au Chapitre 6 pour le problème inverse non-linéaire (38).

0.5 Résumé de la thèse et contributions

Les contributions de cette thèse sont détaillées dans chaque chapitre comme suit.

Méthodes de type one-shot à un pas pour les problèmes inverses linéaires. Dans le Chapitre 2, nous analysons la convergence de la méthode semi-implicite de type one-shot à un pas (11)

$$\begin{cases} \sigma^{n+1} = \sigma^n - \tau M^* p^n - \tau\alpha\sigma^{n+1}, \\ u^{n+1} = Bu^n + M\sigma^{n+1} + F, \\ p^{n+1} = B^* p^n + H^*(Hu^n - g). \end{cases} \qquad \text{(11 rappelée)}$$

Il est facile d'observer que le système d'erreurs de cet algorithme est en fait une itération de point fixe donnée par $\mathbf{x}^{n+1} = \mathbb{A}(\tau)\mathbf{x}^n$ où $\mathbf{x} = (\sigma, u, p)$ et $\mathbb{A}(\tau)$ est ce qu'on appelle la matrice d'itération. Analyser la convergence de (11) revient donc à étudier la localisation des valeurs propres de $\mathbb{A}(\tau)$. En effet, nous allons prouver que pour $\tau > 0$ suffisamment petit, toutes les valeurs propres de $\mathbb{A}(\tau)$ se trouvent à l'intérieur du cercle unité dans le plan complexe, ce qui donne la convergence de (1.11). Pour ce faire, nous donnons dans la Section 2.1 une forme appropriée pour l'équation aux valeurs propres de $\mathbb{A}(\tau)$, voir (2.9) pour l'expression de cette équation. On remarque que (2.9) n'est pas sous forme polynomiale et inclut un produit scalaire en $\mathbb{C}^{n_\sigma}$ avec la présence des matrices B, M, H et le pas de descente τ. Ensuite, dans la Section 2.2, nous étudions l'emplacement des racines de (2.9). Dans le cas de valeurs propres λ de $\mathbb{A}(\tau)$ réelles, (2.9) contient des opérateurs semi-définis positifs. Cependant, dans le cas où λ est complexe mais non réelle, nous perdons la propriété semi-définie positive en général. Nous aborderons cela en combinant les parties réelle et imaginaire de l'équation aux valeurs propres (2.9) afin d'obtenir de nouveaux opérateurs semi-définis positifs. L'étude est ensuite complétée grâce à plusieurs lemmes techniques. Ce que nous obtenons est une condition suffisante sur τ avec une borne explicite qui assure la convergence de la méthode one-shot semi-implicite (11). Ceci sera résumé dans la Section 2.3. En particulier, pour $\|B\| < 1$, la borne de τ dépend uniquement de $\|B\|, \|M\|, \|H\|$ et du paramètre de régularisation α ; cette borne ne dépend pas des dimensions de σ, u, p. Remarquons qu'en comparaison avec le cas scalaire qui sera étudié dans la Section 4.2, on observe que les conditions suffisantes obtenues dans le cas général ne sont pas optimales. Le contenu du Chapitre 2 est extrait de l'article :

> [4] M. Bonazzoli, H. Haddar, and T.A. Vu (2023, preprint). On the convergence analysis of one-shot inversion methods. hal-04151014

Méthodes de type one-shot à pas multiples pour les problèmes inverses linéaires. En développant la technique présentée dans le chapitre précédent pour la méthode semi-implicite de type one-shot à un pas, nous analyserons dans le Chapitre 3 la convergence de la méthode semi-implicite de type one-shot à k pas (9)

$$\begin{cases} \sigma^{n+1} = \sigma^n - \tau M^* p^n - \tau\alpha\sigma^{n+1}, \\ u_0^{n+1} = u^n, p_0^{n+1} = p^n, \\ \text{for } \ell = 0, 1, \dots, k-1 : \\ \quad \left| \begin{array}{l} u_{\ell+1}^{n+1} = Bu_\ell^{n+1} + M\sigma^{n+1} + F, \\ p_{\ell+1}^{n+1} = B^* p_\ell^{n+1} + H^*(Hu_\ell^{n+1} - g), \end{array} \right. \\ u^{n+1} := u_k^{n+1}, p^{n+1} := p_k^{n+1} \end{cases} \qquad \text{(9 rappelée)}$$

pour $k \geq 1$. Cette technique légèrement modifiée est également applicable au cas $k = 1$, ce qui explique pourquoi le cas $k = 1$ est inclus dans ce chapitre même s'il a été étudié dans le chapitre précédent. En effet, le système d'erreurs de l'algorithme (9) est une itération de

point fixe donnée par $\mathbf{x}^{n+1} = \mathbb{A}(k,\tau)\mathbf{x}^n$ où $\mathbf{x} = (\sigma, u, p)$ et $\mathbb{A}(k,\tau)$ est ce qu'on appelle la matrice d'itération. Analyser la convergence de (11) revient donc à étudier la localisation des valeurs propres de $\mathbb{A}(k,\tau)$. Étant donné $k \geq 1$, nous prouverons dans la Section 3.1 que pour $\tau > 0$ suffisamment petit, toutes les valeurs propres de $\mathbb{A}(k,\tau)$ se trouvent à l'intérieur du cercle unité dans le plan complexe, ce qui donne la convergence de (9). Pour ce faire, nous donnons d'abord une forme appropriée pour l'équation aux valeurs propres de $\mathbb{A}(k,\tau)$, voir (3.13) pour l'expression de cette équation. On remarque que (3.13) n'est pas sous forme polynomiale et inclut un produit scalaire dans $\mathbb{C}^{n_\sigma}$ avec la présence des matrices B, M, H et le pas de descente τ. L'équation des valeurs propres (3.13) pour $k \geq 1$, qui est une version plus générale et compliquée de (2.9) pour $k = 1$, peut être considéré comme la somme de plusieurs composantes et contrôler leurs signes reste l'idée clé de notre preuve. Dans le cas où des valeurs propres λ de $\mathbb{A}(k,\tau)$ réelles, (2.9) contient des opérateurs semi-définis positifs. Cependant, dans le cas où λ est complexe mais non réel, nous perdons la propriété de l'operateur semi-défini positif en général. Nous allons résoudre ce problème en combinant les parties réelle et imaginaire de l'équation aux valeurs propres (3.13) afin d'obtenir de nouveaux opérateurs semi-définis positifs. L'étude est ensuite complétée grâce à plusieurs lemmes techniques. Ce que nous obtenons est une condition suffisante sur τ avec une borne explicite qui assure la convergence de la méthode one-shot semi-implicite k ($k \geq 1$). En particulier, pour $\|B\| < 1$, la borne de τ dépend uniquement de $\|B\|, \|M\|, \|H\|$, le nombre d'itérations internes k et le paramètre de régularisation α ; cette borne ne dépend pas des dimensions de σ, u, p. La principale différence entre les deux techniques présentées respectivement au Chapitre 2 et au Chapitre 3 réside sur le choix d'un paramètre technique θ_0 dans le Lemme 3.A.4 : au Chapitre 2 on peut prendre $0 < \theta_0 \leq \frac{\pi}{4}$ mais au Chapitre 3 nous devons prendre $0 < \theta_0 < \frac{\pi}{4}$ puisque nos nouvelles estimations contiennent une division par $\cos 2\theta_0$. Après avoir terminé la preuve de convergence pour les méthodes de type one-shot à pas multiples, nous effectuons plusieurs expériences numériques sur un problème inverse jouet de Helmholtz en 2D et présentons leurs résultats dans la Section 3.2. Tout d'abord, nous considérons dans la Section 3.2.1 le cas de données sans bruit, sans régularisation ($\alpha = 0$). Deux séries d'expériences sont réalisées pour étudier la dépendance des méthodes de type one-shot à pas multiples par rapport au pas de descente τ et au nombre d'itérations internes k, pour un pas de descente fixe τ. Ensuite, nous considérons dans la Section 3.2.2 le cas où les mesures sont affectées par différents niveaux de bruit : $\varepsilon = 1\%, 3\%$ et 5%. Le paramètre de régularisation α est ensuite choisi par essais et erreurs. En particulier, nous observons que les méthodes semi-implicites de type one-shot à k pas, même avec k petit, nécessitent un nombre similaire d'itérations externes à l'algorithme semi-implicite de descente de gradient pour obtenir la même précision. Cela prouve le potentiel de ces méthodes puisque seules quelques itérations internes sont suffisantes. De plus, nous étudions dans la Section 3.2.3 la robustesse par rapport à la taille du problème discrétisé. Nos expériences montrent en effet que le nombre d'itérations externes n'est pas très affecté par la taille du système discret. Enfin, nous considérons dans la Section 3.2.4 la dépendance du nombre d'itérations externes par rapport à la norme de B. Bien que nous n'ayons pas analysé le taux de convergence, nous observons à partir des expériences numériques que plus la norme de B est petite, plus la convergence est rapide.
Le contenu du Chapitre 3 est extrait de l'article :

> [4] M. Bonazzoli, H. Haddar, and T.A. Vu (2023, preprint). On the convergence analysis of one-shot inversion methods. hal-04151014

Analyse de convergence dans certains cas particuliers. Dans le Chapitre 4, nous

analysons la convergence des méthodes semi-implicites de type one-shot à pas multiples (9) dans deux cas particuliers. Le premier cas particulier que nous présenterons dans la Section 4.1 est une extension aux problèmes inverses où le paramètre inverse est réel mais l'état et l'état adjoint sont complexes, appelons-les *problèmes inverses complexes*. Le but est de montrer que le problème inverse complexe peut être reconduit au cadre précédent du *problème inverse réel* dans la Section 0.2. En effet, nous étudions la même équation d'état $u = Bu + M\sigma + F$ que (1) où σ est toujours dans $\mathbb{R}^{n_\sigma}$ mais u est maintenant dans $\mathbb{C}^{n_u}$ au lieu de $\mathbb{R}^{n_u}$. La mesure $Hu(\sigma)$ où $H \in \mathbb{C}^{n_g \times n_u}$ est également désormais dans un espace complexe. Les deux hypothèses pour le problème inverse complexe sont définies comme dans (4) : $\rho(B) < 1$ et $H(I - B)^{-1}M$ est injectif. Avec la technique du Lagrangien pour l'équation complexe, nous définissons l'état adjoint p de $u(\sigma)$. Ensuite, en doublant les tailles de u et p (en fait en prenant leurs parties réelles et imaginaires), nous obtenons un problème inverse réel qui a la même structure de la Section 0.2. De plus, nous montrons que ce nouveau problème satisfait les hypothèses (4), donc le cas des problèmes inverses complexes est également couvert par notre théorie. Le reste du chapitre est dédié au cas dit *scalaire* (c.-à-d. $n_u = n_\sigma = n_g = 1$) sans paramètre de régularisation (c.-à-d. $\alpha = 0$), qui est la configuration la plus simple. Nous étudierons non seulement les algorithmes usuelles de descente de gradient et de type one-shot à pas multiples présentés précédemment, mais également les algorithmes décalés nommés la *méthode de descente de gradient décalée* et la *méthode de type one-shot à pas multiples décalée*. Chaque système d'erreurs de l'un de ces algorithmes se présente en effet sous la forme d'une itération de point fixe régie par une matrice 3×3. Par conséquent, l'équation aux valeurs propres associée à chaque algorithme n'est qu'un polynôme du troisième ordre. Grâce au critère de Jury-Marden (qui fournit une condition nécessaire et suffisante sur les coefficients d'un polynôme à coefficients réels pour que tous ses zéros se trouvent strictement à l'intérieur du cercle unité dans le plan complexe), nous établissons des conditions suffisantes et même nécessaires sur le pas de descente τ qui assurent la convergence de ces algorithmes. Les intervalles des τ admissibles sont tracées afin que nous puissions mieux comprendre les performances des algorithmes de type one-shot lorsque le nombre d'itérations internes tend vers l'infini.
Le contenu de ce chapitre est extrait de l'article :

> [3] M. Bonazzoli, H. Haddar, and T.A. Vu (2022). Convergence analysis of multi-step one-shot methods for linear inverse problems. Research Report RR-9477, Inria Saclay, ENSTA ParisTech. hal-03727759

Combinaison des méthodes de type one-shot à pas multiples avec les méthodes de décomposition de domaine pour les problèmes inverses linéaires. Les méthodes d'inversion de type one-shot à pas multiples présentées dans le Chapitre 0 peuvent être appliquées au cas où le solveur itératif du problème direct est une méthode de décomposition de domaine. Le but du chapitre 5 est d'étudier la combinaison des méthodes de type one-shot à pas multiples avec les méthodes décomposition de domaine pour le problème inverse de conductivité linéarisé. Le cas non-linéaire sera étudié après, dans le Chapitre 6. Pour commencer le Chapitre 5, nous rappelons dans la Section 5.1 le problème inverse de conductivité linéarisé dont le problème direct est de type Neumann, qui a été presenté dans la Section 0.4.2, et, en utilisant la technique classique du Lagrangien, nous définissons l'état adjoint correspondant. Ensuite, nous appliquons une méthode de décomposition de domaine, ou plus spécifiquement, la OSM sans recouvrement pour résoudre les problèmes directs et adjoints. Par souci de simplification, nous considérons le cas où le domaine est

divisé en deux sousdomaines sur lesquels nous appliquerons la OSM, et le contraste de conductivité inconnu est situé dans le sousdomaine interne. En appliquant l'OSM pour résoudre les problèmes directs et adjoints, nous trouvons un algorithme qui calcule simultanément les états direct et adjoint par rapport à un contraste de conductivité donné. De plus, comme dans l'algorithme 0.4, cet algorithme peut être réécrit de manière compacte en termes des variables d'impédance d'interface qui viennent des problèmes direct et adjoint. Nous analysons ensuite la convergence de cet algorithme sous certaines hypothèses raisonnables, et montrons que, en particulier, celles-ci sont satisfaites dans le cas du domaine circulaire présenté dans l'Annexe 5.A. En combinant cet algorithme OSM pour le calcul simultané des états direct et adjoint avec la mise à jour des paramètres de descente de gradient, nous obtenons un nouveau algorithme de décomposition de domaine qui résout le problème de conductivité inverse linéarisé. La Section 5.2 est dédiée aux versions discrétisées des algorithmes étudiés dans la section précédente, obtenues en discrétisant les problèmes directs et adjoints par la méthode des éléments finis. Nous observons que la version discrétisée de l'algorithme combiné est proche mais ne correspond pas exactement au cadre abstrait presenté dans la Section 0.2. Nous proposons donc un nouveau schéma comme suit. Nous construisons un problème inverse discret alternatif en termes de variable d'impédance de l'état en travaillant directement sur la version discrétisée de la OSM sans recouvrement pour le problème direct. Cette méthode nous permet de dériver un schéma alternatif où l'état adjoint est remplacé par l'adjoint numérique pour la reformulation du problème en termes de variables d'impédance. Ce deuxième schéma respecte en effet le cadre abstrait de la Section 0.2, et est présenté dans la Section 5.2.3.

Combinaison des méthodes de type one-shot à pas multiples avec les méthodes de décomposition de domaine pour les problèmes inverses non-linéaires. Le Chapitre 6, le dernier chapitre de la thèse, est dédié à l'application des méthodes d'inversion de type one-shot à pas multiples et des méthodes de décomposition de domaine au problème inverse de conductivité non-linéaire. Ce chapitre suit la structure du précédent. Nous traitons toujours le cas du problème direct de Neumann, et considérons le cas où le domaine est divisé en deux sousdomaines sur lesquels nous appliquerons la méthode de décomposition de domaine et où la conductivité inconnue est située dans le sousdomaine interne. Pour commencer le Chapitre 6, nous rappelons dans la Section 6.1 le problème inverse de conductivité non-linéaire dont le problème direct est de type Neumann, qui a été presenté dans la Section 0.4.1, et en utilisant la technique classique du Lagrangien, nous définissons l'état adjoint correspondant. Ensuite, nous appliquons la OSM sans recouvrement pour résoudre les problèmes direct et adjoint. Nous trouvons un algorithme qui calcule simultanément les états direct et adjoint par rapport à une conductivité donnée, et de plus, cet algorithme peut être réécrit de manière compacte en termes des variables d'impédance provenant des problèmes direct et adjoint. Nous analysons ensuite la convergence de l'algorithme sous certaines hypothèses raisonnables, et montrons que, en particulier, celles-ci sont satisfaites dans le cas d'un domaine circulaire présenté dans l'Annexe 6.A. En combinant cet algorithme OSM pour le calcul simultané des états direct et adjoint avec la mise à jour des paramètres de descente de gradient, nous obtenons un nouveau algorithme de décomposition de domaine qui résout le problème de conductivité inverse non linéaire. La Section 6.2 est ensuite dédiée aux versions discrétisées des algorithmes étudiés dans la section précédente, obtenues en discrétisant les problèmes directs et adjoints à l'aide des méthodes des éléments finis. Enfin, nous fournissons dans la Section 6.3 plusieurs expériences numériques pour comparer les performances de l'algorithme classique de descente de gradient (avec des solveurs directs pour les problèmes direct et adjoint) avec les algorithmes de décomposition

de domaine à un pas et à pas multiples. Comme dans la Section 3.2, nous observons que très peu d'itérations internes suffisent pour donner une bonne convergence de l'algorithme combiné.

Pour terminer ce manuscrit, nous présentons dans le Chapitre 7 des conclusions et des pistes d'extension de nos méthodes.

Chapter 1

Iterative methods for inverse problems

Contents

Our main goal is to analyze the convergence of gradient-based optimization methods applied to solve inverse problems whose forward and adjoint problems are solved iteratively. One may expect that the forward and adjoint problems should be solved with high accuracy since the better the gradient of the cost functional is approximated, the sooner the inverse problem unknown is retrieved. As a consequence, in each outer iteration of updating the inverse problem unknown using gradient descent methods, a sufficiently large number of inner iterations would be employed for solving the forward and adjoint problems. However, many numerical experiments showed that a very few number of inner iterations may still lead to a good convergence of the inverse problem. Besides reducing the number of inner iterations, the calculation process can be improved by choosing efficient iterative methods, such as domain decomposition methods, and parallelizing the computations of the forward and adjoint problem solutions. Moreover, for large-scale forward and adjoint problems, direct solvers, which would yield an exact gradient, are not practical due to their high memory cost. These facts motivated us to study *one-shot inversion methods* and *domain decomposition methods*, which are both introduced in this opening chapter.

Chapter 1 is organized as follows. Section 1.1 is dedicated to the thesis motivation and the state of the art in our research direction. Next, in Section 1.2, we formulate a generic class of linear inverse problems whose forward and adjoint problems are solved by fixed point iterations. We then introduce the one-shot inversion methods that iterate at the same time on the forward and adjoint problem solutions. As we would like to have more efficient and practical iterative solvers than the generic fixed point iterations, we shall introduce the domain decomposition methods in Section 1.3. An application to the inverse conductivity problems is presented in Section 1.4. We end the chapter in Section 1.5 by making a summary of the contributions of the thesis.

1.1 Motivation and state of the art

For large-scale inverse problems, which often arise in real life applications, the solution of the corresponding forward and adjoint problems is generally computed using an iterative solver, such as preconditioned fixed point or Krylov subspace methods, rather than exactly by a direct solver, such as LU-type solvers (see e.g. [49, 1]). Indeed, the corresponding linear systems could be too large to be handled with direct solvers because of their high memory requirement. In addition, iterative solvers are easier to parallelize on many cores for time speed-up. By coupling the iterative solver with a gradient-based optimization iteration, the idea of *one-step one-shot methods* is to iterate at the same time on the forward problem solution (the state variable), the adjoint problem solution (the adjoint state) and on the inverse problem unknown (the parameter or design variable). If two or more inner iterations are performed on the state and adjoint state before updating the parameter (by starting from the previous iterates as initial guess for the state and adjoint state), we speak of *multi-step one-shot methods*. Our goal is to rigorously analyze the convergence of such inversion methods. In particular, we are interested in those schemes where the inner iterations on the forward and adjoint problems are incomplete, i.e. stopped before achieving convergence. Indeed, solving the forward and adjoint problems exactly by direct solvers or very accurately by iterative solvers could be very time-consuming with little improvement in the accuracy of the inverse problem solution.

The concept of one-shot methods was first introduced by Ta'asan [46] for optimal control problems. Based on this idea, a variety of related methods, such as the all-at-once methods, where the state equation is included in the misfit functional, were developed for aerodynamic shape optimization, see for instance [47, 45, 25, 42, 41] and the literature review in the introduction of [42]. All-at-once approaches to inverse problems for parameter identification were studied in, e.g., [19, 5, 33, 31, 32, 39, 40]. An alternative method, called Wavefield Reconstruction Inversion (WRI), was introduced for seismic imaging in [50], as an improvement of the classical Full Waveform Inversion (FWI) [48]. WRI is a penalty method which combines the advantages of the all-at-once approach with those of the reduced approach (where the state equation represents a constraint and is enforced at each iteration, as in FWI), and was extended to more general inverse problems in [51].

Few convergence proofs, especially for the multi-step one-shot methods, are available in the literature. In particular, for non-linear design optimization problems, Griewank [17] proposed a version of one-step one-shot methods where a Hessian-based preconditioner is used in the design variable iteration. The author proved conditions to ensure that the real eigenvalues of the Jacobian of the coupled iterations are smaller than 1, but these are just necessary and not sufficient conditions to exclude real eigenvalues smaller than -1. In addition, no condition to also bound complex eigenvalues below 1 in modulus was

found, and multi-step methods were not investigated. In [22, 23, 14] an exact penalty function of doubly augmented Lagrangian type was introduced to coordinate the coupled iterations, and global convergence of the proposed optimization approach was proved under some assumptions. This particular one-step one-shot approach was later extended to time-dependent problems in [18]. More recently, for geometric inverse problems, one-shot inversion methods combined with domain decomposition were analyzed in [7, 28] in the case of cylindrically invariant geometries.

1.2 One-shot inversion methods

We shall present (multi-step) one-shot inversion methods for a generic class of linear inverse problems, for which we have developed a rigorous convergence theory. Our analysis is directly placed in the finite-dimensional discrete setting.

1.2.1 Linear inverse problem

We focus on (discretized) linear inverse problems, which correspond to a *direct (or forward) problem* of the form: seek $u \equiv u(\sigma)$ such that

$$u = Bu + M\sigma + F \tag{1.1}$$

where $u \in \mathbb{R}^{n_u}$, $\sigma \in \mathbb{R}^{n_\sigma}$, $B \in \mathbb{R}^{n_u \times n_u}$, $M \in \mathbb{R}^{n_u \times n_\sigma}$ and $F \in \mathbb{R}^{n_u}$. Here $I - B$ is the invertible matrix associated with the direct problem (e.g. obtained after discretization of a PDE model and applying a preconditioner to the linear system), with parameter σ. Equation (1.1) is also referred to as the *state equation* and u is the *state variable*. Given σ, we assume that one solves for u by a fixed point iteration

$$u_{\ell+1} = Bu_\ell + M\sigma + F, \quad \ell = 0, 1, \dots \tag{1.2}$$

We indeed assume $\rho(B) < 1$ so that the fixed point iteration (1.2) converges for any initial guess u_0 (see e.g. [16, Theorem 2.1.1]). Measuring $g = Hu(\sigma)$, where $H \in \mathbb{R}^{n_g \times n_u}$, we consider the *linear inverse problem* of retrieving σ from the knowledge of g. Let us set $A := H(I-B)^{-1}M$. The inverse problem can be synthetically written as $A\sigma = g - H(I-B)^{-1}F$, which amounts to inverting the ill-conditioned matrix A. We shall assume in the following the uniqueness of the solution for this inverse problem, which is equivalent to the injectivity of A. In summary, we set

$$\begin{array}{ll} \text{direct problem:} & u = Bu + M\sigma + F, \\ \text{inverse problem:} & \text{measure } g = Hu(\sigma), \text{ retrieve } \sigma \end{array} \tag{1.3}$$

with the assumptions:

$$\rho(B) < 1, \quad H(I - B)^{-1}M \text{ is injective.} \tag{1.4}$$

Remark 1.2.1. Considering real-valued matrices B and M is not a restrictive assumption. Indeed, the case of complex-valued matrices can be rewritten as a system of real-valued equations by doubling the size of the linear system (see Section 4.1).

To solve the inverse problem we write its regularized least squares formulation: given some measurements g, we seek the *regularized solution* $\sigma_\alpha^{\text{ex}}$ defined by

$$\sigma_\alpha^{\text{ex}} = \operatorname{argmin}_{\sigma \in \mathbb{R}^{n_\sigma}} J(\sigma) \quad \text{where } J(\sigma) := \frac{1}{2}\|Hu(\sigma) - g\|^2 + \frac{\alpha}{2}\|\sigma\|^2, \ \alpha \geq 0. \tag{1.5}$$

In case we have exact measurements, i.e. $g = Hu(\sigma^{\text{ex}})$, one can take $\alpha = 0$ so that $\sigma_\alpha^{\text{ex}} = \sigma^{\text{ex}}$. However, in practice, the measurements are often affected by noise, hence in general we cannot reconstruct σ^{ex} but its approximation $\sigma_\alpha^{\text{ex}}$. Now we solve the minimization problem in (1.5). Using the classical Lagrangian technique with real scalar products, we introduce the *adjoint state* $p \equiv p(\sigma)$, which is the solution of

$$p = B^* p + H^*(Hu - g)$$

and allows us to compute the gradient of the cost functional as

$$\nabla J(\sigma) = M^* p(\sigma) + \alpha\sigma.$$

The classical gradient descent algorithm then reads

$$\textbf{usual gradient descent:} \quad \begin{cases} \sigma^{n+1} = \sigma^n - \tau M^* p^n - \tau\alpha\sigma^n, \\ u^n = Bu^n + M\sigma^n + F, \\ p^n = B^* p^n + H^*(Hu^n - g), \end{cases} \tag{1.6}$$

where $\tau > 0$ is the descent step size, and the state and adjoint state equations are solved exactly by a direct solver at each iteration step for σ. Notice that when $F = 0$, (1.6) is equivalent to $\sigma^{n+1} = \sigma^n - \tau A^*(A\sigma^n - g) - \tau\alpha\sigma^n$. One can also consider a slightly different version where an implicit scheme is applied to the regularization term leading to the following semi-implicit gradient scheme

$$\textbf{semi-implicit gradient descent:} \quad \begin{cases} \sigma^{n+1} = \sigma^n - \tau M^* p^n - \tau\alpha\sigma^{n+1}, \\ u^n = Bu^n + M\sigma^n + F, \\ p^n = B^* p^n + H^*(Hu^n - g). \end{cases} \tag{1.7}$$

It can be shown that both algorithms converge for sufficiently small $\tau > 0$: for any initial guess, (1.6) converges if and only if $\tau < \frac{2}{\rho(A^*A)+\alpha}$ and (1.7) converges if and only if $(\rho(A^*A) - \alpha)\tau < 2$. These results indicate in particular that we gain more stability with the semi-implicit scheme.

1.2.2 Principle of one-shot inversion methods

We are interested in methods where the direct and adjoint problems are rather solved iteratively as in (1.2), and where we iterate at the same time on the forward problem solution and the inverse problem unknown: such methods are called *one-shot methods*. More precisely, we are interested in two variants of *multi-step one-shot methods*, defined as follows. Let n be the index of the (outer) iteration on σ. We update $\sigma^{n+1} = \sigma^n - \tau M^* p^n - \tau\alpha\sigma^n$ as in gradient descent methods (or respectively, $\sigma^{n+1} = \sigma^n - \tau M^* p^n - \tau\alpha\sigma^{n+1}$ as in semi-implicit gradient descent methods), but the state and adjoint state equations are now solved by a fixed point iteration method, using just k *inner iterations* and it is important to note that as initial guess we choose the information from the previous (outer) step. We then get the two following variants of multi-step one-shot algorithms

$$k\textbf{-step one-shot:} \quad \begin{cases} \sigma^{n+1} = \sigma^n - \tau M^* p^n - \tau\alpha\sigma^n, \\ u_0^{n+1} = u^n, p_0^{n+1} = p^n, \\ \text{for } \ell = 0, 1, \dots, k-1 : \\ \quad \left| \begin{array}{l} u_{\ell+1}^{n+1} = Bu_\ell^{n+1} + M\sigma^{n+1} + F, \\ p_{\ell+1}^{n+1} = B^* p_\ell^{n+1} + H^*(Hu_\ell^{n+1} - g), \end{array} \right. \\ u^{n+1} \coloneqq u_k^{n+1}, p^{n+1} \coloneqq p_k^{n+1} \end{cases} \tag{1.8}$$

and

$$\textbf{semi-implicit } k\textbf{-step one-shot:} \quad \begin{cases} \sigma^{n+1} = \sigma^n - \tau M^* p^n - \tau\alpha\sigma^{n+1}, \\ u_0^{n+1} = u^n, p_0^{n+1} = p^n, \\ \text{for } \ell = 0, 1, \dots, k-1 : \\ \quad \left| \begin{array}{l} u_{\ell+1}^{n+1} = B u_\ell^{n+1} + M\sigma^{n+1} + F, \\ p_{\ell+1}^{n+1} = B^* p_\ell^{n+1} + H^*(H u_\ell^{n+1} - g), \end{array} \right. \\ u^{n+1} := u_k^{n+1}, p^{n+1} := p_k^{n+1}. \end{cases} \tag{1.9}$$

In particular, when $k = 1$, we obtain the two following algorithms

$$\textbf{one-step one-shot:} \quad \begin{cases} \sigma^{n+1} = \sigma^n - \tau M^* p^n - \tau\alpha\sigma^n, \\ u^{n+1} = B u^n + M\sigma^{n+1} + F, \\ p^{n+1} = B^* p^n + H^*(H u^n - g) \end{cases} \tag{1.10}$$

and

$$\textbf{semi-implicit one-step one-shot:} \quad \begin{cases} \sigma^{n+1} = \sigma^n - \tau M^* p^n - \tau\alpha\sigma^{n+1}, \\ u^{n+1} = B u^n + M\sigma^{n+1} + F, \\ p^{n+1} = B^* p^n + H^*(H u^n - g). \end{cases} \tag{1.11}$$

Note that when $k \to \infty$, the k-step one-shot method (1.8) formally converges to the usual gradient descent (1.6), while the semi-implicit k-step one-shot method (1.9) formally converges to the semi-implicit gradient descent (1.7). Since the analysis of the two schemes (1.8) and (1.9) can be done following similar arguments, we choose to concentrate on only one of them, namely the semi-implicit scheme. We first analyze the one-step one-shot method (1.11) ($k = 1$) in Chapter 2 and then the multi-step one-shot method (1.9) ($k \geq 1$) in Chapter 3. Furthermore, we refer to Section 4.2 for the analysis in the case $n_u = n_\sigma = n_g = 1$ (for which we can apply basic tools) and $\alpha = 0$ (for which the two schemes (1.8) and (1.9) coincide).

1.3 Domain decomposition methods

Domain decomposition methods (DDMs) are a family of iterative methods to solve efficiently large-scale forward problems. The key idea behind those methods is to decompose the computational domain into many subdomains, in which subproblems are appropriately defined and solved in parallel using a (robust) direct solver. Building subproblems plays the main role in DDM; basically this involves how to design overlapping/nonoverlapping subdomains, which boundary conditions are imposed on the interfaces and how to exchange information between subdomains. After that, subproblem solutions also need to be concatenated in a correct way, usually with the help of a partition of unity, to guarantee the convergence to the global solution. While the original version of DDM by Schwarz[1] [44] and its classical variants use Dirichlet conditions on the interfaces, the later versions, first introduced by P. L. Lions[2] [34, 36], named *optimized Schwarz methods (OSMs)* use Robin

[1] Hermann Amandus Schwarz (1843-1921), German mathematician.
[2] Pierre-Louis Lions (1956), French mathematician.

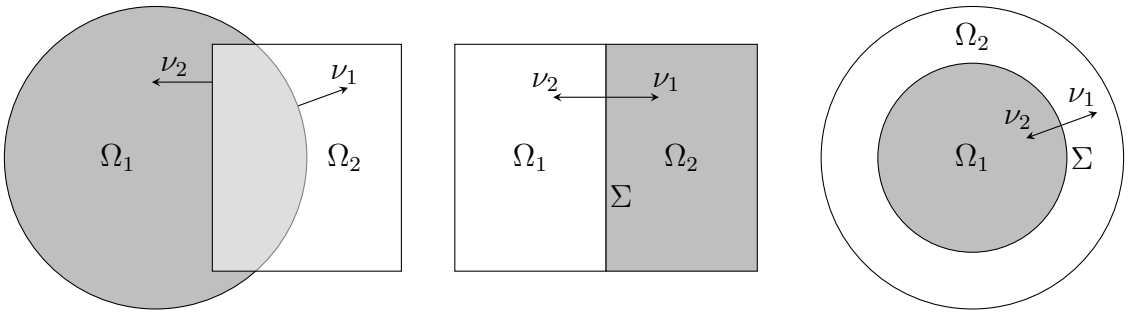
transmission conditions instead and this was proved to be more efficient, see [12, Chapter 2]. P. L. Lions made a convergence proof for the elliptic case [35] , then Després[3] extended it to the case of the Helmholtz equation [11] and time-harmonic Maxwell equations [10]. Here, we shall focus on OSMs; however our introduction also includes the classical Schwarz method so that we can describe the advantages of OSMs compared to the classical methods.

1.3.1 Optimized Schwarz methods

We consider a well-posed direct problem of the form

$$\begin{cases} \mathcal{A}(u) = b & \text{equation in } \Omega, \\ \mathcal{B}(u) = f & \text{boundary condition on } \partial\Omega \end{cases}$$

and we assume that the solution u is regular enough, and that the trace of u and conormal derivative $\mathcal{N}u$ are continuous. In our framework, we shall consider the equation $-\triangle u + \eta u = 0$ where $\mathcal{N} = \frac{\partial}{\partial \nu}$ and the slightly more general version $-\operatorname{div}(\sigma \nabla u) + \eta u = 0$ where $\mathcal{N} = \sigma \frac{\partial}{\partial \nu}$.



(a) Overlapping subdomains.

(b) Non-overlapping subdomains.

(c) Non-overlapping subdomains.

Figure 1.1: Some typical examples of two-subdomain decompositions.

We decompose $\overline{\Omega}$ into $\overline{\Omega} = \overline{\Omega}_1 \cup \overline{\Omega}_2$ (Ω_1 and Ω_2 can be overlapping or not) as sketched in Figure 1.1. For $i = 1, 2$, we define the interface

$$\Sigma_i := \begin{cases} \partial\Omega_i \cap \Omega_{3-i} & \text{if } \Omega_1 \text{ and } \Omega_2 \text{ are overlapping,} \\ \partial\Omega_1 \cap \partial\Omega_2 =: \Sigma & \text{if } \Omega_1 \text{ and } \Omega_2 \text{ are nonoverlapping,} \end{cases} \tag{1.12}$$

where we impose the transmission condition; also we denote by ν_i the outer normal vector of Ω_i and $\mathcal{N}_i$ the conormal trace operator corresponding to ν_i.

The *classical Schwarz method (with Dirichlet transmission conditions)* and the *OSM (with Robin transmission conditions)* then respectively read

[3] Bruno Després (1965), French mathematician.

Algorithm 1.1: Classical Schwarz method for two subdomains

Initialize $u_{i;0}$ in Ω_i, $i = 1, 2$;
for $\ell = 0, 1, ...$ **do**
  for $i = 1, 2$ **do**
    Solve

$$\begin{cases} \mathcal{A}(u_{i;\ell+1}) = b & \text{in } \Omega_i, \\ \mathcal{B}(u_{i;\ell+1}) = f & \text{on } \partial\Omega_i \cap \partial\Omega, \\ u_{i;\ell+1} = u_{3-i;\ell} & \text{on } \Sigma_i \end{cases}$$

    for $u_{i;\ell+1}$;
  end
end
Glue subproblem solutions appropriately

and

Algorithm 1.2: OSM for two subdomains

Initialize $u_{i;0}$ in Ω_i, $i = 1, 2$;
for $\ell = 0, 1, ...$ **do**
  for $i = 1, 2$ **do**
    Solve

$$\begin{cases} \mathcal{A}(u_{i;\ell+1}) = b & \text{in } \Omega_i, \\ \mathcal{B}(u_{i;\ell+1}) = f & \text{on } \partial\Omega_i \cap \partial\Omega, \\ (\mathcal{N}_i + \beta)\, u_{i;\ell+1} = (\mathcal{N}_i + \beta)\, u_{3-i;\ell} & \text{on } \Sigma_i \end{cases}$$

    for $u_{i;\ell+1}$;
  end
end
Glue subproblem solutions appropriately.

In the case of OSM with nonoverlapping subdomains (see Figures 1.1b and 1.1c), we would like to have one notation for the conormal derivative on the interface Σ so we rewrite the transmission condition for the subproblem in Ω_2 as follows. For $x \in \Sigma$, let $\nu(x) = \nu_1(x)$ be the outer normal vector of Ω_1 at x then $\nu_2(x) = -\nu(x)$ is the outer normal vector of Ω_2 at x. We denote $\mathcal{N} = \mathcal{N}_1$ then $\mathcal{N}_2 = -\mathcal{N}$. With the new notation, the interface boundary condition

$$(\mathcal{N}_2 + \beta)\, u_{2;\ell+1} = (\mathcal{N}_2 + \beta)\, u_{1;\ell} \quad \text{on } \Sigma$$

becomes

$$(-\mathcal{N} + \beta)\, u_{2;\ell+1} = (-\mathcal{N} + \beta)\, u_{1;\ell} \quad \text{on } \Sigma,$$

which is equivalent to

$$(\mathcal{N} - \beta)\, u_{2;\ell+1} = (\mathcal{N} - \beta)\, u_{1;\ell} \quad \text{on } \Sigma.$$

The OSM in the case of nonoverlapping subdomains is thus rewritten as

Algorithm 1.3: OSM for two nonoverlapping subdomains

Initialize $u_{i;0}$ in Ω_i, $i = 1, 2$;
for $\ell = 0, 1, ...$ **do**
  Solve

$$\begin{cases} \mathcal{A}(u_{1;\ell+1}) = b & \text{in } \Omega_1, \\ \mathcal{B}(u_{1;\ell+1}) = f & \text{on } \partial\Omega_1 \cap \partial\Omega, \\ (\mathcal{N} + \beta)\, u_{1;\ell+1} = (\mathcal{N} + \beta)\, u_{2;\ell} & \text{on } \Sigma \end{cases}$$

  for $u_{1;\ell+1}$;
  Solve

$$\begin{cases} \mathcal{A}(u_{2;\ell+1}) = b & \text{in } \Omega_2, \\ \mathcal{B}(u_{2;\ell+1}) = f & \text{on } \partial\Omega_2 \cap \partial\Omega, \\ (\mathcal{N} - \beta)\, u_{2;\ell+1} = (\mathcal{N} - \beta)\, u_{1;\ell} & \text{on } \Sigma \end{cases}$$

  for $u_{2;\ell+1}$;
end
Glue subproblem solutions appropriately.

However, there is an issue about the numerical implementation of the normal derivatives in the right-hand side in practice. For instance, in the weak formulation of the subproblem in Ω_1, we may have to discretize $\int_\Sigma \frac{\partial u_{2;\ell}}{\partial \nu}\phi \, \mathrm{ds}$ where $u_{2;\ell}$ is defined on a mesh in Ω_2 but ϕ is a test function defined on another mesh in Ω_1. A numerical error will occur when we try to interpolate the normal derivative between two opposite-side meshes, which has a significant impact on the convergence (see also [12, Section 2.3]). Therefore, it is more convenient to avoid working with normal derivative as follows. We introduce the *impedance values* (or *impedance variables*)

$$\lambda_{1;\ell+1} \coloneqq (\mathcal{N} + \beta)\, u_{2;\ell} \text{ and } \lambda_{2;\ell+1} \coloneqq (\mathcal{N} - \beta)\, u_{1;\ell} \text{ on } \Sigma, \quad \ell = 0, 1, ... \tag{1.13}$$

then the transmission conditions on Σ in Algorithm 1.3 gives

$$(\mathcal{N} + \beta)u_{1;\ell} = \lambda_{1;\ell} \text{ and } (\mathcal{N} - \beta)u_{2;\ell} = \lambda_{2;\ell} \text{ on } \Sigma, \quad \ell = 0, 1, ... \tag{1.14}$$

For $\ell = 0, 1, ...$, we have

$$\lambda_{1;\ell+1} \overset{(1.13)}{=} (\mathcal{N} + \beta)\, u_{2;\ell} = (\mathcal{N} - \beta)\, u_{2;\ell} + 2\beta u_{2;\ell} \overset{(1.14)}{=} \lambda_{2;\ell} + 2\beta u_{2;\ell}$$

and

$$\lambda_{2;\ell+1} \overset{(1.13)}{=} (\mathcal{N} - \beta)\, u_{1;\ell} = (\mathcal{N} + \beta)\, u_{1;\ell} - 2\beta u_{1;\ell} \overset{(1.14)}{=} \lambda_{1;\ell} - 2\beta u_{1;\ell}.$$

on Σ. The difference is that in implementation, we are now able to define a mesh on Σ for both sequences $(\lambda_{1;\ell})_{\ell\geq 0}$ and $(\lambda_{2;\ell})_{\ell\geq 0}$, and we only need to interpolate the subproblem solutions onto this mesh. We thus rewrite the OSM in Algorithm 1.3 as an iteration over the impedance values λ_1 and λ_2, that is as a *substructuring method*:

Algorithm 1.4: Substructuring version of OSM for two nonoverlapping subdomains

Initialize $\lambda_{i;0}$, $i = 1, 2$;
for $\ell = 0, 1, \ldots$ **do**
  Solve

$$\begin{cases} \mathcal{A}(u_{1;\ell}) = b & \text{in } \Omega_1, \\ \mathcal{B}(u_{1;\ell}) = f & \text{on } \partial\Omega_1 \cap \partial\Omega, \\ (\mathcal{N} + \beta)\, u_{1;\ell} = \lambda_{1;\ell} & \text{on } \Sigma \end{cases}$$

  for $u_{1;\ell}$;
  Solve

$$\begin{cases} \mathcal{A}(u_{2;\ell}) = b & \text{in } \Omega_2, \\ \mathcal{B}(u_{2;\ell}) = f & \text{on } \partial\Omega_2 \cap \partial\Omega, \\ (\mathcal{N} - \beta)\, u_{2;\ell} = \lambda_{2;\ell} & \text{on } \Sigma \end{cases}$$

  for $u_{2;\ell}$;
  $\lambda_{1;\ell+1} = \lambda_{2;\ell} + 2\beta u_{2;\ell}$, $\lambda_{2;\ell+1} = \lambda_{1;\ell} - 2\beta u_{1;\ell}$;
end
Glue subproblem solutions appropriately.

In the next section, we will discuss the convergence of Algorithms 1.1–1.4.

1.3.2 About the convergence of domain decomposition methods

An overview about the convergence of classical Schwarz methods and OSMs can be found respectively in [12, Chapter 1] and [12, Chapter 2]. Here, we would like to discuss in detail the convergence for two typical models. The first one, also presented in [12], is the equation

$$(-\triangle + \eta)u = \tilde{f} \text{ in } \mathbb{R}^2 \quad (\eta \in \mathbb{R}, \tilde{f} \text{ is a source term})$$

with a suitable condition at infinity (depending on the sign of η) that guarantees the uniqueness of the problem. We decompose the plane $\mathbb{R}^2$ into two half planes

$$\Omega_1 = (-\infty, \delta) \times \mathbb{R} \text{ and } \Omega_2 = (0, \infty) \times \mathbb{R} \tag{1.15}$$

where $\delta \geq 0$ is the overlap size, and $\delta = 0$ means there is no overlap. For the interfaces where we impose the transmission boundary conditions, we keep using the notations Σ_i defined in (1.12); in this case

$$\Sigma_1 = \{(x, y) \in \mathbb{R}^2 : x = \delta\} \text{ and } \Sigma_2 = \{(x, y) \in \mathbb{R}^2 : x = 0\}. \tag{1.16}$$

This two-half-plane model is typical since after taking the Fourier transform in the y-direction, the errors in the subdomains have explicit expressions (which depend on the sign of η). We are also interested in a second typical model with a bounded domain made of two concentric circles (see e.g. [7, Section 3.1]). Given the piecewise scalar function σ defined in the circle $B(0, R_2) \subset \mathbb{R}^2$ that takes the value $\sigma_1 \in \mathbb{R}$ in the smaller circle $B(0, R_1)$, $R_1 < R_2$, and the value $\sigma_2 \in \mathbb{R}$ in the annulus $R_1 < \sqrt{x^2 + y^2} < R_2$, we consider the equation

$$-\operatorname{div}(\sigma \nabla u) = 0 \text{ in } B(0, R_2) \subset \mathbb{R}^2$$

with either Dirichlet or Neumann boundary condition. For DDMs without overlap, we directly use the circle $B(0, R_1)$ and the annulus $R_1 < \sqrt{x^2 + y^2} < R_2$ as the subdomains (see

Figure 1.1c). In this model, the subproblem errors have explicit expressions (in Fourier series form). Studying the subproblem errors in those two models shows that the convergence of DDMs relates to some geometric sequences and is hence characterized by the so-called *convergence factor* associated with the ratios of these sequences. Convergence proofs for more general geometric configurations are actively studied in the literature. A more general convergence proof of nonoverlapping OSM with second order transmission conditions for elliptic equations is exhibited in [12, Section 2.1.2]. Recently, Gong et al [15] have proved the convergence of overlapping OSM for the Helmholtz equation for strip-wise domain decompositions. Moreover, in the nonoverlapping case, many recent articles, e.g. [8], focus on the convergence issue with the presence of cross-points, i.e. points where three or more subdomains are adjacent.

To give readers an idea of the convergence proof, we briefly present the convergence analysis for the first model in two cases: the case of an elliptic equation ($\eta > 0$), which can be found for classical Schwarz method in [12, Section 1.5.2] and for OSM in [12, Section 2.1.1]; and the case of the Helmholtz equation ($\eta < 0$), which can be found for classical Schwarz method in [12, Section 2.2.1] and for OSM in [12, Section 2.2.2]. This will provide better understanding of the advantages of OSM compared to the classical Schwarz method.

The case of an elliptic equation

Given $\eta > 0$, we consider the global problem

$$\begin{cases} (-\triangle + \eta)u = \tilde{f} & \text{in } \mathbb{R}^2, \\ u \text{ is bounded} & \text{at infinity} \end{cases}$$

for which the classical Schwarz method and the OSM are respectively written as

$$i = 1, 2: \quad \begin{cases} (-\triangle + \eta)u_{i;\ell+1} = \tilde{f} & \text{in } \Omega_i, \\ u_{i;\ell+1} \text{ is bounded} & \text{at infinity}, \\ u_{i;\ell+1} = u_{3-i;\ell} & \text{on } \Sigma_i \end{cases}$$

and

$$i = 1, 2: \quad \begin{cases} (-\triangle + \eta)u_{i;\ell+1} = \tilde{f} & \text{in } \Omega_i, \\ u_{i;\ell+1} \text{ is bounded} & \text{at infinity}, \\ \left(\frac{\partial}{\partial \nu_i} + \beta\right) u_{i;\ell+1} = \left(\frac{\partial}{\partial \nu_i} + \beta\right) u_{3-i;\ell} & \text{on } \Sigma_i \end{cases}$$

where Ω_i, Σ_i, $i = 1, 2$, are respectively defined in (1.15) and (1.16). We define the errors $e_{i;\ell} := u_i^\ell - u|_{\Omega_i}$ and we are interested in their Fourier transform in the y-direction denoted by $\hat{e}_{i;\ell} = \hat{e}_{i;\ell}(x, \xi)$, which is defined by

$$\hat{a}(x, \xi) = (\mathcal{F}a)(x, \xi) := \int_{\mathbb{R}} a(x, y) \mathrm{e}^{-\mathrm{i}\xi y} \, \mathrm{d}y.$$

It is easy to check that $\hat{e}_{i;\ell}, i = 1, 2$, satisfy

$$\begin{cases} \left(-\frac{\partial^2}{\partial x^2} + \xi^2 + \eta\right) \hat{e}_{1;\ell+1} = 0, & x < \delta, \xi \in \mathbb{R}, \\ \hat{e}_{1;\ell+1}(\delta, \xi) = \hat{e}_{2;\ell}(\delta, \xi), & \forall \xi \in \mathbb{R}, \\ \left(-\frac{\partial^2}{\partial x^2} + \xi^2 + \eta\right) \hat{e}_{2;\ell+1} = 0, & x > 0, \xi \in \mathbb{R}, \\ \hat{e}_{2;\ell+1}(0, \xi) = \hat{e}_{1;\ell}(0, \xi), & \forall \xi \in \mathbb{R} \end{cases} \tag{1.17}$$

for the classical Schwarz method, and

$$\begin{cases} \left(-\frac{\partial^2}{\partial x^2}+\xi^2+\eta\right)\hat{e}_{1;\ell+1}=0, & x<\delta,\xi\in\mathbb{R}, \\ \left(\frac{\partial}{\partial x}+\beta\right)(\hat{e}_{1;\ell+1})(\delta,\xi)=\left(\frac{\partial}{\partial x}+\beta\right)(\hat{e}_{2;\ell})(\delta,\xi), & \forall\xi\in\mathbb{R}, \\ \left(-\frac{\partial^2}{\partial x^2}+\xi^2+\eta\right)\hat{e}_{2;\ell+1}=0, & x>0,\xi\in\mathbb{R}, \\ \left(-\frac{\partial}{\partial x}+\beta\right)(\hat{e}_{2;\ell+1})(0,\xi)=\left(-\frac{\partial}{\partial x}+\beta\right)(\hat{e}_{1;\ell})(0,\xi), & \forall\xi\in\mathbb{R} \end{cases} \tag{1.18}$$

for the OSM. Since the solutions are also bounded at infinity, the fundamental solutions to (1.17) and (1.18) for fixed $\xi\in\mathbb{R}$ are of the form

$$\hat{e}_{1;\ell}(x,\xi)=\hat{e}_{1;\ell}(\delta,\xi)\mathrm{e}^{\lambda(\xi)(x-\delta)},x<\delta \quad \text{and} \quad \hat{e}_{2;\ell}(x,\xi)=\hat{e}_{2;\ell}(0,\xi)\mathrm{e}^{-\lambda(\xi)x},x>0 \tag{1.19}$$

where

$$\lambda(\xi):=\sqrt{\xi^2+\eta}. \tag{1.20}$$

For the convergence, it suffices to focus on $\hat{e}_{1;\ell}(\delta,\xi)$ and $\hat{e}_{2;\ell}(0,\xi)$. From the transmission boundary conditions, we deduce that $\hat{e}_{1;\ell}(\delta,\xi)$, $\ell\geq 0$ and $\hat{e}_{2;\ell}(0,\xi)$, $\ell\geq 0$ behave like two geometric sequences with the same ratio $\tilde{r}$:

$$\hat{e}_{1;\ell+1}(\delta,\xi)=\tilde{r}^2\hat{e}_{1;\ell-1}(\delta,\xi) \quad \text{and} \quad \hat{e}_{2;\ell+1}(0,\xi)=\tilde{r}^2\hat{e}_{2;\ell-1}(0,\xi), \quad \forall\ell\geq 1,$$

where

$$\tilde{r}=\begin{cases} \tilde{r}(\xi,\delta):=\mathrm{e}^{-\lambda(\xi)\delta} & \text{for the classical Schwarz method,} \\ \tilde{r}(\xi,\delta;\beta):=\dfrac{\lambda(\xi)-\beta}{\lambda(\xi)+\beta}\mathrm{e}^{-\lambda(\xi)\delta} & \text{for the OSM.} \end{cases}$$

Using the definition of $\lambda(\xi)$ in (1.20), we thus obtain the *convergence factor* $\rho=|\tilde{r}|$ given by

$$\boxed{\rho=\begin{cases} \rho(\xi,\delta):=\mathrm{e}^{-\delta\sqrt{\xi^2+\eta}} & \text{for the classical Schwarz method,} \\ \rho(\xi,\delta;\beta):=\left|\dfrac{\sqrt{\xi^2+\eta}-\beta}{\sqrt{\xi^2+\eta}+\beta}\right|\mathrm{e}^{-\delta\sqrt{\xi^2+\eta}} & \text{for the OSM.} \end{cases}} \tag{1.21}$$

The general condition for convergence is $\rho<1$. Based on formula (1.21) of the convergence factor, we have the following conclusions about the model ($\eta>0$).

- The classical Schwarz method converges when there is overlap ($\delta>0$) and in addition, we even have uniform convergence since $\rho<\mathrm{e}^{-\delta\sqrt{\eta}}<1$, $\forall\xi\in\mathbb{R}$.
- The classical Schwarz method diverges when there is no overlap ($\delta=0$).
- The OSM always converges, with or without overlap. When there is overlap, thanks to the Robin coefficient β, the OSM uniformly converges with a smaller convergence factor and thus faster than the classical Schwarz method.

The case of the Helmholtz equation

We now take into account the case of the Helmholtz equation: $\eta=-\omega^2<0$ for some $\omega>0$, ω is called *the wavenumber*. Before discussing the two-half-plane model, we must remark that the classical Schwarz method is not always applicable for the Helmholtz equation: for

certain wavenumbers ω, the subproblems may be ill-posed even when the global problem is well-posed. For example, consider the global problem

$$\begin{cases} (-\triangle - \omega^2)u = \tilde{f} & \text{in a domain } \tilde{\Omega}, \\ u = \tilde{g} & \text{on } \partial\tilde{\Omega} \end{cases} \tag{1.22}$$

for some given wavenumber ω such that this problem is well-posed. We decompose $\tilde{\Omega}$ as $\overline{\tilde{\Omega}} = \overline{\tilde{\Omega}}_1 \cup \overline{\tilde{\Omega}}_2$ where $\tilde{\Omega}_1$ and $\tilde{\Omega}_2$ are overlapping subdomains as sketched in Figure 1.2. For $i = 1, 2$, we define the interface $\tilde{\Sigma}_i := \partial\tilde{\Omega}_i \cap \tilde{\Omega}_{3-i}$ where we impose the transmission condition, also we denote by $\tilde{\nu}_i$ the outer normal vector of $\tilde{\Omega}_i$.

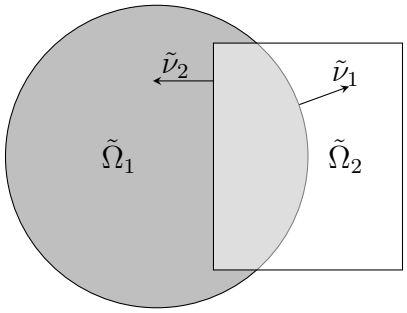


Figure 1.2: Overlapping subdomains $\tilde{\Omega}_1$ and $\tilde{\Omega}_2$.

The classical Schwarz method and the OSM are respectively written as

$$i = 1, 2 : \quad \begin{cases} (-\triangle - \omega^2)u_{i;\ell+1} = \tilde{f} & \text{in } \tilde{\Omega}_i, \\ u_{i;\ell+1} = \tilde{g} & \text{on } \partial\tilde{\Omega}_i \cap \partial\tilde{\Omega}, \\ u_{i;\ell+1} = u_{3-i;\ell} & \text{on } \tilde{\Sigma}_i \end{cases} \tag{1.23}$$

and

$$i = 1, 2 : \quad \begin{cases} (-\triangle - \omega^2)u_{i;\ell+1} = \tilde{f} & \text{in } \tilde{\Omega}_i, \\ u_{i;\ell+1} = \tilde{g} & \text{on } \partial\tilde{\Omega}_i \cap \partial\tilde{\Omega}, \\ \left(\frac{\partial}{\partial\nu_i} + \beta\right) u_{i;\ell+1} = \left(\frac{\partial}{\partial\nu_i} + \beta\right) u_{3-i;\ell} & \text{on } \tilde{\Sigma}_i. \end{cases} \tag{1.24}$$

We shall focus on the subproblem on $\tilde{\Omega}_1$ the schemes (1.23) and (1.24). The uniqueness of the subproblem on $\tilde{\Omega}_1$ defined in (1.23) is equivalent to the uniqueness of the homogeneous problem

$$\begin{cases} (-\triangle - \omega^2)v = 0 & \text{in } \tilde{\Omega}_1, \\ v = 0 & \text{on } \partial\tilde{\Omega}_1. \end{cases} \tag{1.25}$$

If ω is such that there exists $v \neq 0$ satisfying (1.25), that is, ω^2 is a Dirichlet eigenvalue of the Laplacian operator, the subproblem on $\tilde{\Omega}_1$ defined in (1.23) is not well-posed. We remark that for every geometry of $\tilde{\Omega}_1$, there always exist Dirichlet eigenvalues, known as the resonances for $\tilde{\Omega}_1$. On the other hand, the uniqueness of the subproblem on $\tilde{\Omega}_1$ defined in (1.24) is equivalent to the uniqueness of the homogeneous problem

$$\begin{cases} (-\triangle - \omega^2)v = 0 & \text{in } \tilde{\Omega}_1, \\ v = 0 & \text{on } \partial\tilde{\Omega}_1 \cap \partial\tilde{\Omega}, \\ \left(\frac{\partial}{\partial\tilde{\nu}} + \beta\right) v = 0 & \text{on } \tilde{\Sigma}_1. \end{cases} \tag{1.26}$$

With Robin conditions, for any $\omega > 0$, we have that $v = 0$ is the only solution to (1.26). Indeed, multiplying the first equation in (1.26) with $\overline{v}$ and integrating by parts using

boundary conditions gives

$$\begin{aligned} 0 = & \int_{\tilde{\Omega}_1} (-\triangle - \omega^2) v \overline{v} \,\mathrm{dx} \\ = & -\int_{\partial\tilde{\Omega}_1} \frac{\partial v}{\partial \tilde{\nu}} \overline{v} \,\mathrm{ds} + \int_{\tilde{\Omega}_1} |\nabla v|^2 \,\mathrm{dx} - \int_{\tilde{\Omega}_1} \omega^2 |v|^2 \,\mathrm{dx} \\ = & \int_{\tilde{\Sigma}_1} \beta |v|^2 \,\mathrm{ds} + \int_{\tilde{\Omega}_1} |\nabla v|^2 \,\mathrm{dx} - \int_{\tilde{\Omega}_1} \omega^2 |v|^2 \,\mathrm{dx}. \end{aligned}$$

The imaginary part of the last equation line implies $v = 0$ on $\tilde{\Sigma}_1$, if the Robin coefficient β is chosen as $\beta = \mathrm{i}\tilde{\beta}$ for some $\tilde{\beta} > 0$. Using the Robin condition, this yields $\frac{\partial v}{\partial \tilde{\nu}} = 0$ on $\tilde{\Sigma}_1$. By Theorem 1.A.2, an application of unique continuation principle, we must have $v = 0$ in $\tilde{\Omega}_1$. Hence, the classical Schwarz methods is not always well-defined for the Helmholtz equation while the OSM fixes this drawback as long as we choose a suitable Robin coefficient β.

Now back to the convergence analysis for the two-half-plane model, we consider the global problem

$$\begin{cases} (-\triangle - \omega^2) u = \tilde{f} & \text{in } \mathbb{R}^2, \\ \lim\limits_{r\to\infty} \sqrt{r} \left(\frac{\partial u}{\partial r} + \mathrm{i}\omega u \right) = 0 & \text{Sommerfeld condition, } r := \sqrt{x^2 + y^2}. \end{cases}$$

The classical Schwarz method then reads

$$i = 1, 2: \quad \begin{cases} (-\triangle - \omega^2) u_{i;\ell+1} = \tilde{f} & \text{in } \Omega_i, \\ \lim\limits_{r\to\infty} \sqrt{r} \left(\frac{\partial u_{i;\ell+1}}{\partial r} + \mathrm{i}\omega u_{i;\ell+1} \right) = 0 & \text{Sommerfeld condition,} \\ u_{i;\ell+1} = u_{3-i;\ell} & \text{on } \Sigma_i \end{cases}$$

and the OSM with Robin coefficient $\beta = \mathrm{i}\tilde{\beta}$ ($\tilde{\beta} > 0$), which is known as *Després method*, is written as

$$i = 1, 2: \quad \begin{cases} (-\triangle - \omega^2) u_{i;\ell+1} = \tilde{f} & \text{in } \Omega_i, \\ \lim\limits_{r\to\infty} \sqrt{r} \left(\frac{\partial u_{i;\ell+1}}{\partial r} + \mathrm{i}\omega u_{i;\ell+1} \right) = 0 & \text{Sommerfeld condition,} \\ \left(\frac{\partial}{\partial \nu_i} + \mathrm{i}\tilde{\beta} \right) u_{i;\ell+1} = \left(\frac{\partial}{\partial \nu_i} + \mathrm{i}\tilde{\beta} \right) u_{3-i;\ell} & \text{on } \Sigma_i, \end{cases}$$

where Ω_i, Σ_i, $i = 1, 2$ are respectively defined in (1.15) and (1.16). In this case, all subproblems in both methods are well-defined. By repeating the same arguments as in the previous section, we focus on $\hat{e}_{i;\ell}, i = 1, 2$, the Fourier transform in the y-direction of the errors, which satisfy

$$\begin{cases} \left(-\frac{\partial^2}{\partial x^2} + \xi^2 - \omega^2 \right) \hat{e}_{1;\ell+1} = 0, & x < \delta, \xi \in \mathbb{R}, \\ \hat{e}_{1;\ell+1}(\delta, \xi) = \hat{e}_{2;\ell}(\delta, \xi), & \forall \xi \in \mathbb{R}, \\ \left(-\frac{\partial^2}{\partial x^2} + \xi^2 - \omega^2 \right) \hat{e}_{2;\ell+1} = 0, & x > 0, \xi \in \mathbb{R}, \\ \hat{e}_{2;\ell+1}(0, \xi) = \hat{e}_{1;\ell}(0, \xi), & \forall \xi \in \mathbb{R} \end{cases} \tag{1.27}$$

for the classical Schwarz method, and

$$\begin{cases} \left(-\frac{\partial^2}{\partial x^2} + \xi^2 - \omega^2 \right) \hat{e}_{1;\ell+1} = 0, & x < \delta, \xi \in \mathbb{R}, \\ \left(\frac{\partial}{\partial x} + \mathrm{i}\tilde{\beta} \right) (\hat{e}_{1;\ell+1})(\delta, \xi) = \left(\frac{\partial}{\partial x} + \mathrm{i}\tilde{\beta} \right) (\hat{e}_{2;\ell})(\delta, \xi), & \forall \xi \in \mathbb{R}, \\ \left(-\frac{\partial^2}{\partial x^2} + \xi^2 - \omega^2 \right) \hat{e}_{2;\ell+1} = 0, & x > 0, \xi \in \mathbb{R}, \\ \left(-\frac{\partial}{\partial x} + \mathrm{i}\tilde{\beta} \right) (\hat{e}_{2;\ell+1})(0, \xi) = \left(-\frac{\partial}{\partial x} + \mathrm{i}\tilde{\beta} \right) (\hat{e}_{1;\ell})(0, \xi), & \forall \xi \in \mathbb{R} \end{cases} \tag{1.28}$$

for the OSM. By Sommerfeld condition, the fundamental solutions to (1.27) and (1.28) have the same form as in (1.19), that is for fixed $\xi \in \mathbb{R}$,

$$\hat{e}_{1;\ell}(x,\xi) = \hat{e}_{1;\ell}(\delta,\xi)\mathrm{e}^{\lambda(\xi)(x-\delta)}, x < \delta \quad \text{and} \quad \hat{e}_{2;\ell}(x,\xi) = \hat{e}_{2;\ell}(0,\xi)\mathrm{e}^{-\lambda(\xi)x}, x > 0$$

but in this case,

$$\lambda(\xi) := \begin{cases} \mathrm{i}\sqrt{\omega^2 - \xi^2} & \text{if } |\xi| < \omega, \\ \sqrt{\xi^2 - \omega^2} & \text{if } |\xi| \geq \omega. \end{cases} \tag{1.29}$$

We observe that $\lambda(\xi)$ defined in (1.29) can be complex while $\lambda(\xi)$ defined in (1.20) is always real. Again, from the transmission boundary conditions, we deduce that $\hat{e}_{1;\ell}(\delta,\xi)$ and $\hat{e}_{2;\ell}(0,\xi)$, $\ell \geq 0$ behave like two geometric sequences with the same ratio $\tilde{r}$:

$$\hat{e}_{1;\ell+1}(\delta,\xi) = \tilde{r}^2\hat{e}_{1;\ell-1}(\delta,\xi) \quad \text{and} \quad \hat{e}_{2;\ell+1}(0,\xi) = \tilde{r}^2\hat{e}_{2;\ell-1}(0,\xi), \quad \forall \ell \geq 1$$

where

$$\tilde{r} = \begin{cases} \tilde{r}(\xi,\delta) := \mathrm{e}^{-\lambda(\xi)\delta} & \text{for the classical Schwarz method,} \\ \tilde{r}(\xi,\delta) := \dfrac{\lambda(\xi) - \mathrm{i}\tilde{\beta}}{\lambda(\xi) + \mathrm{i}\tilde{\beta}}\mathrm{e}^{-\lambda(\xi)\delta} & \text{for the OSM.} \end{cases}$$

The convergence factor $\rho = |\tilde{r}|$ is thus given by

$$\boxed{\rho = \rho(\xi,\delta) := \left|\mathrm{e}^{-\lambda(\xi)\delta}\right| = \begin{cases} 1 & \text{if } |\xi| \leq \omega, \\ \mathrm{e}^{-\delta\sqrt{\xi^2-\omega^2}} & \text{if } |\xi| > \omega \end{cases}} \tag{1.30}$$

for the classical Schwarz method, and

$$\boxed{\rho = \rho(\xi,\delta;\tilde{\beta}) := \left|\frac{\lambda(\xi) - \mathrm{i}\tilde{\beta}}{\lambda(\xi) + \mathrm{i}\tilde{\beta}}\mathrm{e}^{-\lambda(\xi)\delta}\right| = \begin{cases} \left|\dfrac{\tilde{\beta} - \sqrt{\omega^2 - \xi^2}}{\tilde{\beta} + \sqrt{\omega^2 - \xi^2}}\right| & \text{if } |\xi| \leq \omega, \\ \mathrm{e}^{-\delta\sqrt{\xi^2-\omega^2}} & \text{if } |\xi| > \omega \end{cases}} \tag{1.31}$$

for the OSM, where we have used the definition of $\lambda(\xi)$ in (1.29). Based on the formulas (1.30) and (1.31) of the convergence factor in each method, we have the following conclusions about the model ($\eta = -\omega^2 < 0$).

- The classical Schwarz method with overlap ($\delta > 0$) diverges for propagative modes $|\xi| \leq \omega$ and converges for evanescent modes $|\xi| > \omega$.
- The classical Schwarz method without overlap ($\delta = 0$) diverges for any modes.
- The OSM with overlap converges for all modes except for $\xi = \pm\omega$.
- The OSM without overlap converges for propagative modes $|\xi| < \omega$ and diverges for evanescent modes $|\xi| \geq \omega$.

1.4 Application to the inverse conductivity problem

1.4.1 Non-linear inverse conductivity problem

In physics, the *electric (voltage) potential* u satisfies the equation

$$-\operatorname{div}(\sigma\nabla u) = 0 \text{ in the domain } \Omega \tag{1.32}$$

where the *conductivity* σ belongs to the admissible set

$$V(\Omega) \coloneqq \{\sigma \in L^\infty(\Omega) : \underline{\sigma} < \sigma < \overline{\sigma}\}$$

for some given constants $\underline{\sigma}, \overline{\sigma} > 0$. Equation (1.32) is called the *conductivity equation*. Given a Dirichlet boundary condition

$$u = f \text{ on } \partial\Omega,$$

the electric potential $u \in H^1(\Omega)$ is uniquely determined from $f \in H^{\frac{1}{2}}(\partial\Omega)$ if we have Lipschitz $\partial\Omega$. Also, u depends continuously on f in the norms of those spaces so we have a well-posed problem. We often assume moreover that σ is constant near $\partial\Omega$. Then, $\frac{\partial u}{\partial \nu} \in H^{-\frac{1}{2}}(\partial\Omega)$, and the conormal trace

$$\sigma \frac{\partial u}{\partial \nu} = g \text{ on } \partial\Omega$$

is well-defined. The *inverse conductivity problem* is to retrieve σ from the knowledge of one or several measurement(s) g. A brief history of the inverse conductivity problem is presented in [27, Section 1.2].

In our works, we are interested in the slightly more general equation

$$-\operatorname{div}(\sigma \nabla u) + \eta u = 0 \text{ in } \Omega \tag{1.33}$$

where η is a function in $L^\infty(\Omega)$. As mentioned in the previous section, if $\eta = -\omega^2 < 0$ in Ω, equation (1.33) is called *the Helmholtz equation* and ω is called *the wavenumber*. If $\eta > 0$ in Ω, equation (1.33) is elliptic and called *the screened Laplacian equation* (or *the screened Poisson equation* when there is a source term, or even called *the Helmholtz equation with good sign*, see e.g. [13]). The solution u of equation (1.33) has continuous trace u and conormal trace $\sigma \frac{\partial u}{\partial \nu}$, which is important in the DDM algorithms introduced in the previous section.

Remark 1.4.1. Above we have described the direct problem with a Dirichlet boundary condition and the inverse problem with Neumann measurements. In addition, we also consider the direct problem with a Neumann boundary condition and the inverse problem with Dirichlet measurements. In the latter case, when $\eta = 0$, we need to take care of the direct problem well-posedness. For the existence, it is known that the source term (which is zero in our case) and the conormal trace g must satisfy *the compatibility condition*

$$\int_{\partial\Omega} g \, \mathrm{ds} = 0.$$

For the uniqueness, again when $\eta = 0$, we require an extra constraint $\int_\Omega u \, \mathrm{dx} = 0$. In the following sections, though not explicitly mentioned, the above compatibility condition and constraint will be implied when $\eta = 0$.

1.4.2 Linearization using Born approximation

It is possible to linearize the inverse conductivity problem as follows. The solution u of equation (1.33) is also called the *total field* and now renamed by u^{tot}:

$$\begin{cases} -\operatorname{div}(\sigma \nabla u^{\text{tot}}) + \eta u^{\text{tot}} = 0 & \text{in } \Omega, \\ u^{\text{tot}} = f & \text{on } \partial\Omega. \end{cases} \tag{1.34}$$

Let $\sigma_0 \in L^\infty(\Omega)$ such that $\sigma_0 = \sigma$ in a neighborhood of $\partial\Omega$. We define the *incident field* u^{inc} such that

$$\begin{cases} -\operatorname{div}(\sigma_0 \nabla u^{\text{inc}}) + \eta u^{\text{inc}} = 0 & \text{in } \Omega, \\ u^{\text{inc}} = f & \text{on } \partial\Omega. \end{cases} \tag{1.35}$$

The *scattered field* u^{scat} defined by the difference between the total field in (1.34) and the incident field in (1.35), i.e. $u^{\text{scat}} := u^{\text{tot}} - u^{\text{inc}}$, satisfies

$$\begin{cases} -\operatorname{div}(\sigma_0 \nabla u^{\text{scat}}) + \eta u^{\text{scat}} = -\operatorname{div}(\tilde{\sigma} \nabla u^{\text{inc}}) - \operatorname{div}(\tilde{\sigma} \nabla u^{\text{scat}}) & \text{in } \Omega, \\ u^{\text{scat}} = 0 & \text{on } \partial\Omega, \end{cases} \tag{1.36}$$

where $\tilde{\sigma} := \sigma_0 - \sigma$ is the *conductivity contrast*. We remark that $\tilde{\sigma} = 0$ in a neighborhood of $\partial\Omega$. The *Born approximation* says that we can ignore $\operatorname{div}(\tilde{\sigma} \nabla u^{\text{scat}})$ in equation (1.36) when $\tilde{\sigma} \ll 1$. We hereby obtain the *linearized field* u^{lin}:

$$\begin{cases} -\operatorname{div}(\sigma_0 \nabla u^{\text{lin}}) + \eta u^{\text{lin}} = -\operatorname{div}(\tilde{\sigma} \nabla u^{\text{inc}}) & \text{in } \Omega, \\ u^{\text{lin}} = 0 & \text{on } \partial\Omega. \end{cases} \tag{1.37}$$

Note that the solution u^{lin} has continuous trace u^{lin} and conormal trace $\sigma_0 \frac{\partial u^{\text{lin}}}{\partial \nu}$. Now, with the Neumann measurement $g := \sigma_0 \frac{\partial u^{\text{lin}}}{\partial \nu}\Big|_{\partial\Omega}$, the map $\tilde{\sigma} \mapsto g$ is linear. The idea of linearization is similar if we replace the Dirichlet condition by the Neumann condition.

1.4.3 An overview about the application

In summary, we formulate the inverse conductivity problems as follows, where, by abuse of notations, in the linearized problem we denote $(u^{\text{lin}}, \tilde{\sigma})$ by (u, σ), which are used in the abstract framework presented in Section 1.2. The non-linear inverse conductivity problem is

$$\boxed{\begin{array}{ll} \text{direct problem:} & \begin{cases} -\operatorname{div}(\sigma \nabla u) + \eta u = 0 \text{ in } \Omega, \\ u = f \text{ on } \partial\Omega \quad (\text{or } \sigma \frac{\partial u(\sigma)}{\partial \nu} = f \text{ on } \partial\Omega), \end{cases} \\ \text{inverse problem:} & \text{given } \eta, \\ & \text{input } f, \\ & \text{measure } g = \sigma \frac{\partial u(\sigma)}{\partial \nu}\Big|_{\partial\Omega} \text{ (or respectively } g = u(\sigma)|_{\partial\Omega}), \\ & \text{retrieve } \sigma \end{array}} \tag{1.38}$$

and the linearized inverse conductivity problem is

$$\boxed{\begin{array}{ll} \text{direct problem:} & \begin{cases} -\operatorname{div}(\sigma_0 \nabla u) + \eta u = -\operatorname{div}(\sigma \nabla u_0) \text{ in } \Omega, \\ u = 0 \text{ on } \partial\Omega \quad (\text{or } \sigma_0 \frac{\partial u(\sigma)}{\partial \nu} = 0 \text{ on } \partial\Omega), \end{cases} \\ \text{where} & \sigma = 0 \text{ in a neighborhood of } \partial\Omega \\ \text{and} & \begin{cases} -\operatorname{div}(\sigma_0 \nabla u_0) + \eta u_0 = 0 \text{ in } \Omega, \\ u_0 = f \text{ on } \partial\Omega \quad (\text{or respectively } \sigma_0 \frac{\partial u_0}{\partial \nu} = f \text{ on } \partial\Omega), \end{cases} \\ \text{inverse problem:} & \text{given } \sigma_0 \text{ and } \eta, \\ & \text{input } f, \\ & \text{measure } g = \sigma_0 \frac{\partial u(\sigma)}{\partial \nu}\Big|_{\partial\Omega} \text{ (or respectively } g = u(\sigma)|_{\partial\Omega}), \\ & \text{retrieve } \sigma. \end{array}} \tag{1.39}$$

Unfortunately, the conductivity inverse problem is severely ill-posed in many cases, even when the problem is linearized. Indeed, the forward problem solution in (1.39), when Ω is a circle in $\mathbb{R}^2$, is of the form $u = A\sigma$ where A is an invertible linear operator, and one can show that the sequence of singular values of A decay exponentially, which causes significant difficulty in inverting A. The relation between the number of measurements and unknowns is also an issue: it is known that one can not retrieve σ with just a finite number of boundary measurements. Some discussions about this kind of issue can be found in [20], [21], [6] and [28]. In our numerical experiments, we shall consider a simplified case where the conductivity σ is a $\mathbb{P}^0$ function with known support and the number of measurements is suitably designed to ensure the uniqueness of the inverse problem. The application of multi-step one-shot methods (1.8) to the linearized conductivity inverse problem with this setting is presented in Section 3.2. Then, the combination with DDMs will be studied in Chapter 5 for the linearized inverse problem (1.39) and Chapter 6 for the non-linear inverse problem (1.38).

1.5 Thesis summary and contributions

The contributions of this thesis are detailed in each Chapter as follows.

One-step one-shot methods for linear inverse problems. In Chapter 2, we analyze the convergence of the semi-implicit one-step one-shot method (1.11)

$$\begin{cases} \sigma^{n+1} = \sigma^n - \tau M^* p^n - \tau\alpha\sigma^{n+1}, \\ u^{n+1} = Bu^n + M\sigma^{n+1} + F, \\ p^{n+1} = B^* p^n + H^*(Hu^n - g). \end{cases} \qquad \text{(1.11 recalled)}$$

It is easy to observe that the error system of this algorithm is actually a fixed point iteration given by $\mathbf{x}^{n+1} = \mathbb{A}(\tau)\mathbf{x}^n$ where $\mathbf{x} = (\sigma, u, p)$ and $\mathbb{A}(\tau)$ is the so-called iteration matrix. Analyzing the convergence of (1.11) is thus equivalent to studying the location of the eigenvalues of $\mathbb{A}(\tau)$. Indeed, we shall prove that for sufficiently small $\tau > 0$, all the eigenvalues of $\mathbb{A}(\tau)$ lie inside the unit circle in the complex plane, which yields the convergence of (1.11). To do that, we give in Section 2.1 a suitable form for the eigenvalue equation of $\mathbb{A}(\tau)$, see (2.9) for the expression of this equation. We remark that (2.9) is not in polynomial form and includes a scalar product in $\mathbb{C}^{n_\sigma}$ with the presence of the matrices B, M, H and the descent step τ. Next, in Section 2.2 we study the location of the roots of (2.9). In case an eigenvalue λ of $\mathbb{A}(\tau)$ is real, (2.9) contains some semi positive-definite operators. However, in case λ is complex but not real, we lose the semi positive-definite property in general. We shall tackle this by combining the real and the imaginary parts of the eigenvalue equation (2.9) so that we obtain some new semi positive-definite operators. The study is then completed thanks to several technical lemmas. What we obtain is a sufficient condition on τ with explicit bound that ensures the convergence of the semi-implicit one-step one-shot method (1.11). This will be summarized in Section 2.3. In particular, for $\|B\| < 1$, the bound of τ depends only on $\|B\|, \|M\|, \|H\|$ and the regularization parameter α; this bound does not depend on the dimensions of σ, u, p. Let us remark that comparing with the scalar case that will be studied in Section 4.2, we observe that the obtained sufficient conditions in the general case are not optimal.
The content of Chapter 2 is extracted from the article:

[4] M. Bonazzoli, H. Haddar, and T.A. Vu (2023, preprint). On the convergence analysis of one-shot inversion methods. hal-04151014

Multi-step one-shot methods for linear inverse problems. By developing the technique presented in the previous chapter for semi-implicit one-step one-shot method, we shall analyze in Chapter 3 the convergence of the semi-implicit k-step one-shot method (1.9)

$$\begin{cases} \sigma^{n+1} = \sigma^n - \tau M^* p^n - \tau\alpha\sigma^{n+1}, \\ u_0^{n+1} = u^n, p_0^{n+1} = p^n, \\ \text{for } \ell = 0, 1, \dots, k-1 : \\ \quad \left| \begin{array}{l} u_{\ell+1}^{n+1} = B u_\ell^{n+1} + M\sigma^{n+1} + F, \\ p_{\ell+1}^{n+1} = B^* p_\ell^{n+1} + H^*(H u_\ell^{n+1} - g), \end{array} \right. \\ u^{n+1} := u_k^{n+1}, p^{n+1} := p_k^{n+1} \end{cases} \qquad \text{(1.9 recalled)}$$

for $k \geq 1$. This slightly-modified technique is also applicable to the case $k = 1$, that explains why it is included in this chapter even though it was studied in the previous chapter. Indeed, the error system of the algorithm (1.9) is a fixed point iteration given by $\mathbf{x}^{n+1} = \mathbb{A}(k,\tau)\mathbf{x}^n$ where $\mathbf{x} = (\sigma, u, p)$ and $\mathbb{A}(k,\tau)$ is the so-called iteration matrix. Analyzing the convergence of (1.11) is thus equivalent to studying the location of the eigenvalues of $\mathbb{A}(k,\tau)$. Given $k \geq 1$, we shall prove in Section 3.1 that for sufficiently small $\tau > 0$, all the eigenvalues of $\mathbb{A}(k,\tau)$ lie inside the unit circle in the complex plane, which yields the convergence of (1.9). To do that, we first give a suitable form for the eigenvalue equation of $\mathbb{A}(k,\tau)$, see (3.13) for the expression of this equation. We remark that (3.13) is not in polynomial form and includes a scalar product in $\mathbb{C}^{n_\sigma}$ with the presence of the matrices B, M, H and the descent step τ. The eigenvalue equation (3.13) for $k \geq 1$, which is a more general and complicated version of (2.9) for $k = 1$, can be viewed as the sum of several components and controlling their signs is still the key idea of our proof. In case an eigenvalue λ of $\mathbb{A}(k,\tau)$ is real, (2.9) contains some semi positive-definite operators. However, in case λ is complex but not real, we lose the semi positive-definite property in general. We shall tackle this by combining the real and the imaginary parts of the eigenvalue equation (3.13) so that we obtain some new semi positive-definite operators. The study is then completed thanks to several technical lemmas. What we obtain is a sufficient condition on τ with explicit bound that ensures the convergence of the semi-implicit k-step one-shot method ($k \geq 1$). In particular, for $\|B\| < 1$, the bound of τ depends only on $\|B\|, \|M\|, \|H\|$, the number of inner iterations k and the regularization parameter α; this bound does not depend on the dimensions of σ, u, p. The main difference between the two techniques respectively presented in Chapter 2 and Chapter 3 lies on the choice of a technical parameter θ_0 in Lemma 3.A.4: in Chapter 2 we can take $0 < \theta_0 \leq \frac{\pi}{4}$ but in Chapter 3 we have to take $0 < \theta_0 < \frac{\pi}{4}$ since our new estimates contain a division by $\cos 2\theta_0$. After finishing the convergence proof for the multi-step one-shot methods, we perform several numerical experiments on a toy 2D Helmholtz inverse problem and present their results in Section 3.2. First, we consider in Section 3.2.1 the case of noise-free data, without regularization ($\alpha = 0$). Two series of experiments are performed to study the dependence of the multi-step one-shot methods on the descent step τ and on the number of inner iterations k, for fixed descent step τ. Next, we consider in Section 3.2.2 the case where the measurements are affected by different levels of noise: $\varepsilon = 1\%, 3\%$ and 5%. The regularization parameter α is then chosen by trial and error. In particular, we observe that

the semi-implicit k-step one-shot methods, with some small values of k, require a similar number of outer iterations as the semi-implicit gradient descent algorithm to achieve the same accuracy. This proves the potential of these methods since only a few inner iterations are used. Moreover, we study in Section 3.2.3 the robustness with respect to the size of the discretized problem. Our experiments indeed show that the number of outer iterations is not much affected by the size of the discrete system. Finally, we consider in Section 3.2.4 the dependence of the number of outer iterations on the norm of B. Although we did not analyze the convergence rate, we observe from the numerical experiments that the smaller the norm of B, the faster the convergence.
The content of this chapter is extracted from the article:

> [4] M. Bonazzoli, H. Haddar, and T.A. Vu (2023, preprint). On the convergence analysis of one-shot inversion methods. hal-04151014

Convergence analysis in some particular cases. In Chapter 4, we analyze the convergence of the semi-implicit multi-step one-shot methods (1.9) in two particular cases. The first particular one that we shall present in Section 4.1 is an extension to inverse problems where the inverse parameter is real but the state and adjoint state are complex, let us call them *complex inverse problems*. The aim is to show that the complex inverse problem can be translated into the previous framework of the *real inverse problem* in Section 1.2. Indeed, we study the same state equation $u = Bu + M\sigma + F$ as (1.1) where σ is still in $\mathbb{R}^{n_\sigma}$ but u is now in $\mathbb{C}^{n_u}$ instead of $\mathbb{R}^{n_u}$. The measurement $Hu(\sigma)$ where $H \in \mathbb{C}^{n_g \times n_u}$ is also now in a complex space. The two assumptions for the complex inverse problem are set as in (1.4): $\rho(B) < 1$ and $H(I-B)^{-1}M$ is injective. With the Lagrangian technique for the complex equation, we define the adjoint state p of $u(\sigma)$. Then, by doubling the sizes of u and p (actually by taking their real and imaginary parts), we obtain a real inverse problem that has the same structure as in Section 1.2. Moreover, we show that this newly obtained problem satisfies the assumptions (1.4), that is to say, the case of complex inverse problems is also covered by our theory. The rest of the chapter is dedicated to the so-called *scalar case* (i.e. $n_u = n_\sigma = n_g = 1$) without regularization parameter (i.e. $\alpha = 0$), which is the simplest configuration. We shall study not only the usual gradient descent and multi-step one-shot algorithms presented before but also the shifted algorithms named *shifted gradient descent* and *shifted multi-step one-shot*. Each system of errors of any of these algorithms is indeed in the form of a fixed point iteration governed by a 3×3 matrix. Therefore, the eigenvalue equation associated with each algorithm is just a third-order polynomial. Thanks to Jury-Marden Criterion (that provides a necessary and sufficient condition on the coefficients of a real-coefficient polynomial so that all its zeros lie strictly inside the unit circle in the complex plane), we establish sufficient and even necessary conditions on the descent step τ that ensure the convergence of these algorithms. The ranges of admissible τ are plotted so that we can better understand the performances of one-shot algorithms when the number of inner iterations tends to infinity.
The content of this chapter is extracted from the report:

> [3] M. Bonazzoli, H. Haddar, and T.A. Vu (2022). Convergence analysis of multi-step one-shot methods for linear inverse problems. Research Report RR-9477, Inria Saclay, ENSTA ParisTech. hal-03727759

Combination of multi-step one-shot and domain decomposition methods for linear inverse problems. The multi-step one-shot inversion methods introduced in Chapter

1 can be applied to the case where the iterative solver for the forward problem is a domain decomposition method. The goal of Chapter 5 is to study the combination of the multi-step one-shot and domain decomposition methods for the linearized inverse conductivity problem. The non-linear case will be studied later in Chapter 6. To begin Chapter 5, we recall in Section 5.1 the linearized inverse conductivity problem with Neumann forward problem, which has been introduced in Section 1.4.2, and, using the classical Lagrangian technique, we define the corresponding adjoint state. Next, we apply a domain decomposition method, or more specifically, the nonoverlapping OSM to solve the forward and adjoint problems. For the sake of simplification, we consider the case where the domain is splitted into two subdomains on which we will apply the OSM, and the unknown conductivity contrast is located in the inner subdomain. When applying the OSM to solve the forward and adjoint problems, we find an algorithm that simultaneously calculates the forward and adjoint states with respect to a given conductivity contrast. Moreover, like in Algorithm 1.4, this algorithm can be compactly rewritten in terms of the interface impedance variables coming from the forward and adjoint problems. We then analyze the convergence of this algorithm under some reasonable assumptions, and show that, in particular, these are satisfied in the case of circular domain presented in Appendix 5.A. By combining this OSM algorithm for the simultaneous calculation of the forward and adjoint states with the gradient descent parameter update, we obtain a domain decomposition one-shot algorithm that solves the linearized inverse conductivity problem. Section 5.2 is dedicated to the discretized versions of the algorithms studied in the previous section, obtained by discretizing the forward and adjoint problems using the finite element method. We observe that the discretized version of the combined algorithm looks close to but does not exactly match the abstract framework introduced in Section 1.2. We therefore propose a new scheme as follows. We construct an alternative discrete inverse problem in terms of the impedance variable of the state by working directly on the discretized version of nonoverlapping OSM for the forward problem. This way allows us to derive an alternative scheme where the adjoint state is replaced by the numerical adjoint for the reformulation of the problem in terms of impedance variables. This second scheme indeed respects the abstract framework in Section 1.2, and is presented in Section 5.2.3.

Combination of multi-step one-shot methods and domain decomposition for non-linear inverse problems. Chapter 6, the last chapter of the thesis, is dedicated to the application of multi-step one-shot inversion methods and domain decomposition methods to the non-linear inverse conductivity problem. This chapter follows the structure of the previous one. Also, we still treat the case of Neumann forward problem, and consider the case where the domain is splitted into two subdomains on which we will apply the domain decomposition method and the unknown conductivity is located in the inner subdomain. To begin Chapter 6, we recall in Section 6.1 the non-linear inverse conductivity problem with Neumann forward problem, which has been introduced in Section 1.4.1, and using the classical Lagrangian technique, we define the corresponding adjoint state Next, we apply the nonoverlapping OSMs to solve the forward and adjoint problems. We find an algorithm that simultaneously calculates the forward and adjoint states with respect to a given conductivity, and furthermore, this algorithm can be compactly rewritten in terms of the impedance variables coming from the forward and adjoint problems. We then analyze the algorithm convergence under some reasonable assumptions, and show that, in particular, these are satisfied in the case of circular domain presented in Appendix 6.A. By combining this OSM algorithm for the simultaneous calculation of the forward and adjoint states with the gradient descent parameter update, we obtain a domain decomposition one-

shot algorithm that solves the non-linear inverse conductivity problem. Section 6.2 is then dedicated to the discretized versions of the algorithms studied in the previous section, obtained by discretizing the forward and adjoint problems using the finite element methods. Finally, we provide in Section 6.3 several numerical experiments to compare the performance of the classical gradient descent algorithm (with direct solvers for the forward and adjoint problems) with the domain decomposition multi-step one-shot algorithm. Like in Section 3.2, we observe that very few inner iterations are enough to give good convergence of the combined algorithm.

To end this manuscript, we present in Chapter 7 some conclusions and some research directions to extend our methods.

The contributions of this thesis led to the following publications and presentations.

Publication

[2022]

- [3] M. Bonazzoli, H. Haddar, and T.A. Vu (2022). Convergence analysis of multi-step one-shot methods for linear inverse problems. Research Report RR-9477, Inria Saclay, ENSTA ParisTech. hal-03727759

[2024]

- M. Bonazzoli, H. Haddar, and T.A. Vu (2024). Convergence analysis of semi-implicit multi-step one-shot methods for regularized linear inverse problems. In *Proceedings of The 16th International Conference on Mathematical and Numerical Aspects of Wave Propagation*. https://doi.org/10.17617/3.MBE4AA

Preprint

[2023]

- [4] M. Bonazzoli, H. Haddar, and T.A. Vu (2023, preprint). On the convergence analysis of one-shot inversion methods. hal-04151014

Conferences

[2024]

- (Talk) 16th International Conference on Mathematical and Numerical Aspects of Wave Propagation. *Convergence analysis of semi-implicit multi-step one-shot methods for regularized linear inverse problems.* Berlin, Germany, 30/06-05/07/2024.
- (Talk) 22nd Workshop on Optimization and Scientific Computing. Ba Vi, Vietnam, 25-27/04/2024.

[2023]

- (Talk) Mini-symposium at the European Conference on Numerical Mathematics and Advanced Applications (ENUMATH 2023). *Convergence analysis of semi-implicit multi- step one-shot methods for regularized linear inverse problems.* Lisbon, Portugal, 04-08/09/2023.

- (Poster) Research School on Iterative Methods for Partial Differential Equations (IMPDE 2023). Sorbonne Université, Paris, France, 15-16/05/2023.

[2022]

- (Talk) 3ème Rencontre Jeunes Chercheuses Jeunes Chercheurs (JCJC) Ondes. Inria Université Côte d'Azur (Sophia-Antipolis), France, 28-30/11/2022.
- (Talk) Mini-symposium at PICOF 2022: Inversion methods for problems with singular parameters. Université de Caen Normandie, France, 25-27/10/2022.
- (Talk) Congrès Jeunes Chercheurs en Mathématiques et leurs Applications (CJC-MA). Université du Littoral Côte d'Opale, Calais, France, 21-23/09/2022.
- (Poster) Research School on Domain Decomposition for Optimal Control Problems. CIRM, Marseille, France, 05-09/09/2022.
- (Talk) Seminar of PhD Students at CMAP and CMLS at École polytechnique. *Convergence analysis for multi-step one-shot methods.* École polytechnique, France, 08/06/2022.

Appendix 1.A An application of the unique continuation principle

When studying the uniqueness of the solutions to the subproblems in the OSM applied to the Dirichlet Helmholtz problem (1.22), we see that the Robin boundary condition is assigned only on a part of the subdomain boundary, see (1.26). Hence we cannot directly conclude the uniqueness of the subproblems using the maximum principle or the energy method, since they require information on the full boundary. One way to tackle this issue is using the unique continuation principle (see [9, Section 8.3]), saying that any H^2-solution u of a second order equation must be identically zero if it satisfies the two following conditions: u vanishes in a non-empty open subset of the domain and $|\triangle u|$ has a suitable estimation in the whole domain.

Lemma 1.A.1 (Unique continuation principle)**.** *Let Ω be a bounded and simply connected domain in $\mathbb{R}^n$ and let $u_i, 1 \leq i \leq m$, be real-valued functions in $C^2(\Omega)$ satisfying*

$$|\triangle u_i| \leq c \sum_{j=1}^{m} (|u_j| + |\nabla u_j|) \quad \text{in } \Omega$$

for $1 \leq i \leq m$ and some constant c. Assume that u_i vanishes in a neighborhood of some point $x_0 \in \Omega$ for $1 \leq i \leq m$. Then u_i is identically zero in Ω for $1 \leq i \leq m$.

The proof of Lemma 1.A.1 can be found in [9, pages 312–315] in case $n = 3$. Now we are ready to prove the following theorem.

Theorem 1.A.2 (The uniqueness theorem for the Helmholtz equation with both Dirichlet and Neumann boundary conditions)**.** *Let $\Omega \subset \mathbb{R}^n$ be a bounded and simply connected domain with Lipschitz boundary $\partial\Omega$ and let $\Gamma \subset \partial\Omega$ be a part of the boundary with non-empty interior in $\partial\Omega$, i.e. Γ is the intersection between $\partial\Omega$ and some open set in $\mathbb{R}^n$. Assume that $u \in H^2(\Omega)$ solves*

$$\begin{cases} (-\triangle - \omega^2)u = 0 & \text{in } \Omega, \\ u = 0 \text{ and } \frac{\partial u}{\partial \nu} = 0 & \text{on } \Gamma \subset \partial\Omega \end{cases}$$

where $\omega \in \mathbb{R}, \omega > 0$. Then $u = 0$ in Ω.

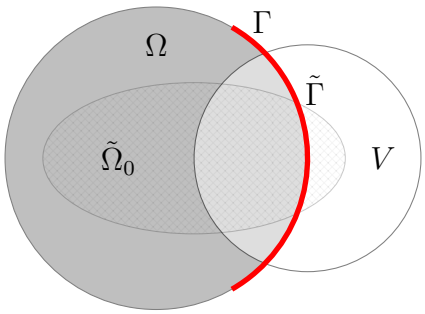


Figure 1.3: Domain Ω extended by V.

Proof. Since the interior of Γ in $\partial\Omega$ is non-empty, there exists a subset $\tilde{\Gamma}$ of Γ with non-empty interior in $\partial\Omega$ and an open neighborhood V of $\tilde{\Gamma}$ such that

$$V \backslash \overline{\Omega} \neq \varnothing \text{ and } V \cap \partial\Omega = \tilde{\Gamma} \subset \Gamma.$$

Let $\tilde{\Omega} := \Omega \cup V$ then

$$\partial\Omega \cap \tilde{\Omega} = \partial\Omega \cap (\Omega \cup V) = \partial\Omega \cap V \subset \Gamma. \tag{1.40}$$

We define

$$\tilde{u}(x) := u(x)\mathbb{1}_\Omega(x),\ \forall x \in \tilde{\Omega}.$$

We show that $\triangle\tilde{u}$ exists and equals $(\triangle u)\mathbb{1}_\Omega$. Indeed, let $\phi \in C_c^\infty(\tilde{\Omega})$ be a test function with $\mathrm{supp}(\phi) \subset \tilde{\Omega}_0$ where $\tilde{\Omega}_0 \subset \tilde{\Omega}$ such that $\overline{\tilde{\Omega}_0}$ is a compact subset in $\tilde{\Omega}$. Before calculating some integrals, we first prove that

$$\partial(\tilde{\Omega}_0 \cap \Omega) \subset \partial\tilde{\Omega}_0 \cup \Gamma. \tag{1.41}$$

Indeed, given any $x \in \partial(\tilde{\Omega}_0 \cap \Omega)$, we have $x \in \partial(\tilde{\Omega}_0 \cap \Omega) \subset \partial\tilde{\Omega}_0 \cup \partial\Omega$. There are two cases.
Case 1. $x \in \partial\tilde{\Omega}_0$. Obviously, $x \in \partial\tilde{\Omega}_0 \cup \Gamma$.
Case 2. $x \notin \partial\tilde{\Omega}_0$. In this case we must have $x \in \partial\Omega$. Also, since $\tilde{\Omega}_0 \cap \Omega \subset \tilde{\Omega}_0 \subset \tilde{\Omega}$, we have $\overline{\tilde{\Omega}_0 \cap \Omega} \subset \overline{\tilde{\Omega}_0} \subset \tilde{\Omega}$ and in particular, $x \in \partial(\tilde{\Omega}_0 \cap \Omega) \subset \tilde{\Omega}$. Hence $x \in \partial\Omega \cap \tilde{\Omega}$. Using (1.40), we deduce that $x \in \Gamma$.
By combining the results of the two above cases, we obtain (1.41). Let $\nu = (\nu_i)_{i=1}^n$ be the coordinate (in the canonical basis of $\mathbb{R}^n$) of the outer normal vector ν directed to the exterior of $\tilde{\Omega}_0 \cap \Omega$. For $1 \le i \le n$ we have

$$\int_{\tilde{\Omega}} \left(\frac{\partial u}{\partial x_i}\mathbb{1}_\Omega\right)\phi\,\mathrm{dx} = \int_{\tilde{\Omega}_0\cap\Omega} \frac{\partial u}{\partial x_i}\phi\,\mathrm{dx} = \int_{\partial(\tilde{\Omega}_0\cap\Omega)} \phi u\nu_i\,\mathrm{ds} - \int_{\tilde{\Omega}_0\cap\Omega} u\frac{\partial\phi}{\partial x_i}\,\mathrm{dx} = -\int_{\tilde{\Omega}} \tilde{u}\frac{\partial\phi}{\partial x_i}\,\mathrm{dx}.$$

Here we remark that for every $x \in \partial(\tilde{\Omega}_0 \cap \Omega)$ we always have $\phi(x)u(x)\nu_i(x) = 0$ thanks to (1.41) and the fact that $\phi(x) = 0$ if $x \in \partial\tilde{\Omega}_0$ and $u(x) = 0$ if $x \in \Gamma$. Therefore $\frac{\partial\tilde{u}}{\partial x_i}$ exists and equals $\frac{\partial u}{\partial x_i}\mathbb{1}_\Omega$. Next, for $1 \le i \le n$ we have

$$\begin{aligned}
\int_{\tilde{\Omega}}((\triangle u)\mathbb{1}_\Omega)\phi\,\mathrm{dx} &= \sum_{i=1}^n \int_{\tilde{\Omega}} \left(\frac{\partial^2 u}{\partial x_i^2}\mathbb{1}_\Omega\right)\phi\,\mathrm{dx} \\
&= \sum_{i=1}^n \int_{\tilde{\Omega}_0\cap\Omega} \frac{\partial^2 u}{\partial x_i^2}\phi\,\mathrm{dx} \\
&= \sum_{i=1}^n \left(\int_{\partial(\tilde{\Omega}_0\cap\Omega)} \phi\frac{\partial u}{\partial x_i}\nu_i\,\mathrm{ds} - \int_{\tilde{\Omega}_0\cap\Omega} \frac{\partial u}{\partial x_i}\frac{\partial\phi}{\partial x_i}\,\mathrm{dx}\right) \\
&= \sum_{i=1}^n \left(\int_{\partial(\tilde{\Omega}_0\cap\Omega)} \phi\frac{\partial u}{\partial x_i}\nu_i\,\mathrm{ds} - \int_{\tilde{\Omega}} \frac{\partial\tilde{u}}{\partial x_i}\frac{\partial\phi}{\partial x_i}\,\mathrm{dx}\right) \\
&= \int_{\partial(\tilde{\Omega}_0\cap\Omega)} \phi\frac{\partial u}{\partial\nu}\,\mathrm{ds} - \int_{\tilde{\Omega}} \nabla\tilde{u}\cdot\nabla\phi\,\mathrm{dx} \\
&= -\int_{\tilde{\Omega}} \nabla\tilde{u}\cdot\nabla\phi\,\mathrm{dx}.
\end{aligned}$$

Here we remark that for every $x \in \partial(\tilde{\Omega}_0 \cap \Omega)$ we always have $\phi(x)\frac{\partial u}{\partial\nu}(x) = 0$ thanks to (1.41) and the fact that $\phi(x) = 0$ if $x \in \partial\tilde{\Omega}_0$ and $\frac{\partial u}{\partial\nu}(x) = 0$ if $x \in \Gamma$. Hence, $\triangle\tilde{u}$ exists and equals $(\triangle u)\mathbb{1}_\Omega$. We deduce that

$$-\triangle\tilde{u} - \omega^2\tilde{u} = (-\triangle - \omega^2)u\mathbb{1}_\Omega = 0 \quad \text{in } \tilde{\Omega},$$

which implies

$$|\triangle\tilde{u}| = |\omega^2\tilde{u}| \le \omega^2|\tilde{u}| \quad \text{in } K$$

for every $K \subset \tilde{\Omega}$ such that $K \neq \varnothing$, $\overline{K}$ is a compact subset in $\tilde{\Omega}$ and $(V\backslash\overline{\Omega}) \cap K \neq \varnothing$. Moreover, $\tilde{u} = 0$ in the non-empty open set $(V\backslash\overline{\Omega}) \cap K \subset V$. Therefore, by the unique continuation principle (Lemma 1.A.1), we have $\tilde{u} = 0$ in K. Since K is arbitrary, we conclude that $\tilde{u} = 0$ in $\tilde{\Omega}$ and in particular, $u = 0$ in Ω. □

Chapter 2

One-step one-shot methods for linear inverse problems

Contents

In this chapter, we analyze the convergence of the semi-implicit one-step one-shot methods (1.11). The error system of this algorithm is a fixed point iteration governed by a 3×3 *block iteration matrix*. Analyzing the convergence of (1.11) is thus equivalent to studying the spectral radius of the associated iteration matrix: the algorithm converges if and only if the spectral radius of the iteration matrix is less than 1. To study the location of the iteration matrix eigenvalues, we first find an expression for the eigenvalue equation of the block matrix, which is not in polynomial form. The eigenvalue equation can be viewed as the sum of some components and controlling their signs is the key idea of our proof. The case of real eigenvalues is simple thanks to the presence of semi positive-definite operators. However, we lose this property when dealing with complex eigenvalues in general. We shall tackle this difficulty by separating the real and imaginary parts of the eigenvalue equation, then suitably combining them so that we retrieve some semi positive-definite operators.

This chapter is organized as follows. The aim of Section 2.1 is to write the eigenvalue equation for the convergence study of one-step one-shot methods. Next, in Section 2.2, we investigate the location of the eigenvalues in the complex plane. Our aim is to find sufficiently small descent steps τ such that the eigenvalue equation does not admit any solutions outside and on the unit circle $|z| = 1$. The case of real eigenvalues is directly proved using a simple estimation. The case of complex eigenvalues is more complicated and hence divided into four subcases after we succeed in separating the real and imaginary parts of the eigenvalue equation. The results on the descent step τ in these subcases are obtained with the help of several technical lemmas. In Section 2.3, we synthesize all our attained results in a convergence theorem. In particular, the condition of method convergence does not depend on the dimensions of the state and adjoint unknowns, nor on the dimension of the inverse problem.

The content of this chapter is extracted from the article:

> [4] M. Bonazzoli, H. Haddar, and T.A. Vu (2023, preprint). On the convergence analysis of one-shot inversion methods. hal-04151014

2.1 Block iteration matrix and eigenvalue equation

To analyze the convergence of the semi-implicit one-step one-shot method (1.11)

$$\begin{cases} \sigma^{n+1} = \sigma^n - \tau M^* p^n - \tau\alpha\sigma^{n+1}, \\ u^{n+1} = Bu^n + M\sigma^{n+1} + F, \\ p^{n+1} = B^* p^n + H^*(Hu^n - g), \end{cases} \qquad \text{(1.11 recalled)}$$

we first express $(\sigma^{n+1}, u^{n+1}, p^{n+1})$ in terms of (σ^n, u^n, p^n), by inserting the expression for σ^{n+1} into the iteration for u^{n+1} in (1.11), so that system (1.11) is rewritten as

$$\begin{cases} \sigma^{n+1} = \frac{1}{1+\tau\alpha}\sigma^n - \frac{\tau}{1+\tau\alpha}M^* p^n, \\ u^{n+1} = Bu^n + \frac{1}{1+\tau\alpha}M\sigma^n - \frac{\tau}{1+\tau\alpha}MM^* p^n + F, \\ p^{n+1} = B^* p^n + H^* Hu^n - H^* g. \end{cases} \tag{2.1}$$

Now, we consider the errors $(\sigma^n - \sigma_\alpha^{\mathrm{ex}}, u^n - u(\sigma_\alpha^{\mathrm{ex}}), p^n - p(\sigma_\alpha^{\mathrm{ex}}))$ with respect to the regularized solution at the n-th iteration, and, by abuse of notation, we designate them by (σ^n, u^n, p^n). We obtain that these errors satisfy

$$\begin{cases} \sigma^{n+1} = \frac{1}{1+\tau\alpha}\sigma^n - \frac{\tau}{1+\tau\alpha}M^* p^n, \\ u^{n+1} = Bu^n + \frac{1}{1+\tau\alpha}M\sigma^n - \frac{\tau}{1+\tau\alpha}MM^* p^n, \\ p^{n+1} = B^* p^n + H^* Hu^n, \end{cases} \tag{2.2}$$

or equivalently, by putting in evidence the block iteration matrix

$$\begin{bmatrix} \sigma^{n+1} \\ u^{n+1} \\ p^{n+1} \end{bmatrix} = \begin{bmatrix} \frac{1}{1+\tau\alpha}I & 0 & -\frac{\tau}{1+\tau\alpha}M^* \\ \frac{1}{1+\tau\alpha}M & B & -\frac{\tau}{1+\tau\alpha}MM^* \\ 0 & H^*H & B^* \end{bmatrix} \begin{bmatrix} \sigma^n \\ u^n \\ p^n \end{bmatrix}. \tag{2.3}$$

Now recall that a fixed point iteration converges if and only if the spectral radius of its iteration matrix is less than 1. Therefore in the following proposition we establish an eigenvalue equation for the iteration matrix of the semi-implicit one-step one-shot method.

Proposition 2.1.1. *Assume that $\lambda \in \mathbb{C}$ is an eigenvalue of the iteration matrix in (2.3). If $\lambda \notin \mathrm{Spec}(B)$ then $\exists\, y \in \mathbb{C}^{n_\sigma}, \|y\| = 1$ such that:*

$$(1+\tau\alpha)\lambda - 1 + \tau\lambda\langle M^*(\lambda I - B^*)^{-1}H^*H(\lambda I - B)^{-1}My, y\rangle = 0. \tag{2.4}$$

In particular, $\lambda = 1$ is not an eigenvalue of the iteration matrix.

Proof. Since $\lambda \in \mathbb{C}$ is an eigenvalue of the iteration matrix, there exists a non-zero vector $(\tilde{p}, \tilde{u}, y) \in \mathbb{C}^{n_u + n_u + n_\sigma}$ such that

$$\begin{cases} \lambda y = \frac{1}{1+\tau\alpha}y - \frac{\tau}{1+\tau\alpha}M^*\tilde{p}, \\ \lambda\tilde{u} = B\tilde{u} + \frac{1}{1+\tau\alpha}My - \frac{\tau}{1+\tau\alpha}MM^*\tilde{p}, \\ \lambda\tilde{p} = B^*\tilde{p} + H^*H\tilde{u}. \end{cases}$$

By inserting the first equation into the second equation, we simplify this system of equations as

$$\begin{cases} \lambda y = \frac{1}{1+\tau\alpha}y - \frac{\tau}{1+\tau\alpha}M^*\tilde{p}, \\ \lambda\tilde{u} = B\tilde{u} + \lambda My, \\ \lambda\tilde{p} = B^*\tilde{p} + H^*H\tilde{u}. \end{cases} \tag{2.5}$$

The second equation in (2.5) gives us directly $\tilde{u}$ in terms of y:

$$\tilde{u} = \lambda(\lambda I - B)^{-1} M y. \tag{2.6}$$

The third equation in (2.5) at first gives us $\tilde{p}$ in terms of $\tilde{u}$:

$$\lambda \tilde{p} = (\lambda I - B^*)^{-1} H^* H \tilde{u},$$

then by combing with equation (2.6), we obtain $\tilde{p}$ in terms of y:

$$\tilde{p} = \lambda(\lambda I - B^*)^{-1} H^* H (\lambda I - B)^{-1} M y. \tag{2.7}$$

We also see that $y \neq 0$; indeed if $y = 0$ then equations (2.6) and (2.7) yield $u = 0$ and $p = 0$, that is a contradiction. Next, inserting the expression of $\tilde{p}$ in equation (2.7) back into the first equation in (2.5), we get

$$\lambda y = \frac{1}{1+\tau\alpha} y - \frac{\tau}{1+\tau\alpha} \lambda M^* (\lambda I - B^*)^{-1} H^* H (\lambda I - B)^{-1} M y,$$

which leads to

$$\left[(1+\alpha\tau)\lambda - 1\right] y + \tau\lambda M^* (\lambda I - B^*)^{-1} H^* H (\lambda I - B)^{-1} M y = 0. \tag{2.8}$$

Finally, by taking the scalar product of (2.8) with y, then dividing by $\|y\|^2$, we obtain (2.4). Now assume that $\lambda = 1$ is an eigenvalue of the iteration matrix, then (2.4) yields

$$\alpha + \|H(I - B)^{-1} M y\|^2 = 0,$$

which cannot be true due to the injectivity of $H(I - B)^{-1} M$. □

In the following sections we will show that, for sufficiently small τ, equation (2.4) cannot hold if $|\lambda| \geq 1$, thus algorithm (1.11) converges. It is convenient to rewrite (2.4) as

$$(1+\tau\alpha)\lambda^2 - \lambda + \tau \langle M^* \left(I - B^*/\lambda\right)^{-1} H^* H \left(I - B/\lambda\right)^{-1} M y, y \rangle = 0. \tag{2.9}$$

For the analysis we use auxiliary technical results proved in Appendix 3.A.

2.2 Location of the eigenvalues in the complex plane

We now turn our attention to the eigenvalues λ for which (2.9) holds. We would like to derive conditions on the descent step τ such that all the eigenvalues lie inside the unit circle which would ensure the convergence for the scheme (1.11). We start with the simple case of real eigenvalues.

Proposition 2.2.1 (Real eigenvalues)**.** *Equation* (2.9) *admits no solution* $\lambda \in \mathbb{R}, \lambda \neq 1, |\lambda| \geq 1$ *for all* $\tau > 0$.

Proof. When $\lambda \in \mathbb{R}\backslash\{0\}$ equation (2.9) becomes

$$(1+\tau\alpha)\lambda^2 - \lambda + \tau \|H(I - B/\lambda)^{-1} M y\|^2 = 0.$$

If $\lambda \in \mathbb{R}, \lambda \neq 1, |\lambda| \geq 1$ then $(1+\tau\alpha)\lambda^2 - \lambda \geq \lambda^2 - \lambda > 0$, thus the left-hand side of the above equation is positive for any $\tau > 0$. □

For the general case of complex eigenvalues, the study is much more complicated and technical. First, we study separately the very particular and simple case where $B = 0$.

Proposition 2.2.2. *When $B = 0$, the eigenvalue equation (2.9) cannot hold for $\lambda \in \mathbb{C}, |\lambda| \geq 1$ if $\tau > 0$ and*

$$(\|H\|^2\|M\|^2 - \alpha)\tau < 1.$$

Proof. When $B = 0$, HM is injective by (1.4), and equation (2.9) becomes $(1 + \tau\alpha)\lambda^2 - \lambda + \tau\|HMy\|^2 = 0$. One can prove the proposition by explicitly computing the roots of the second order polynomial with respect to λ. One can also apply the following elementary Lemma 2.2.3 below, which can be deduced from Marden's work [38] or Appendix 4.A. Indeed, the previous polynomial can be written as $P(\lambda) = a_0 + a_1\lambda + \lambda^2$ with $a_0 = \frac{\tau\|HMy\|^2}{1+\tau\alpha}$ and $a_1 = \frac{-1}{1+\tau\alpha}$. We have $a_0 - a_1 + 1 = \frac{\tau(\|HMy\|^2+\alpha)+2}{1+\tau\alpha} > 0$ and $a_0 + a_1 + 1 = \frac{\tau(\|HMy\|^2+\alpha)}{1+\tau\alpha} > 0$ since $\|HMy\|^2 > 0$ thanks to the injectivity of HM. Therefore, according to Lemma 2.2.3, the roots of P stay strictly inside the unit circle of the complex plane if and only if $|a_0| = \frac{\tau\|HMy\|^2}{1+\tau\alpha} < 1$, which is equivalent to $(\|HMy\|^2 - \alpha)\tau < 1$. Since $\|y\| = 1$, the later inequality is verified if $(\|H\|^2\|M\|^2 - \alpha)\tau < 1$, which is the statement of the proposition. □

Lemma 2.2.3. *Let $a_0, a_1 \in \mathbb{R}$, then all roots of $\mathcal{P}(z) = a_0 + a_1 z + z^2$ stay (strictly) inside the unit circle of the complex plane if and only if*

$$|a_0| < 1 \quad \text{and} \quad (a_0 - a_1 + 1)(a_0 + a_1 + 1) > 0.$$

Now we consider the more complicated case where $0 \neq \rho(B) < 1$. The following proposition summarizes the results we obtained.

Proposition 2.2.4 (Complex eigenvalues)**.** *If $0 \neq \rho(B) < 1$, there exists $\tau > 0$ sufficiently small such that equation (2.9) admits no solution $\lambda \in \mathbb{C}\backslash\mathbb{R}, |\lambda| \geq 1$. In particular, if $\|B\| < 1$, given any $\delta_0 > 0$ and $0 < \theta_0 \leq \frac{\pi}{4}$, one can choose*

$$\tau < \min_{1\leq i\leq 3} \left(\frac{\|H\|^2\|M\|^2}{(1 - \|B\|)^4}\varphi_i(\|B\|) + C_i\alpha \right)^{-1},$$

where

$$\varphi_1(b) := 4b^2, \quad \varphi_2(b) := \frac{1}{2\sin\frac{\theta_0}{2}}(1 + b)^2(1 - b)^2, \quad \varphi_3(b) := \frac{2c}{\delta_0}b^2,$$

$$C_1 := \sqrt{2} - 1, \quad C_2 := \sqrt{2} + \frac{1}{2\sin\frac{\theta_0}{2}} - 1, \quad C_3 := \frac{\sqrt{c}}{\delta_0} - 1 \quad \text{and} \quad c := \frac{1 + 2\delta_0\sin\frac{3\theta_0}{2} + \delta_0^2}{\cos^2\frac{3\theta_0}{2}}.$$

Proof. **Step 1. Rewrite equation (2.9) by separating real and imaginary parts.**

Let $\lambda = R(\cos\theta + \mathrm{i}\sin\theta)$ in polar form where $R = |\lambda| \geq 1$ and $\theta \in (-\pi, \pi)$, $\theta \neq 0$. Write $1/\lambda = r(\cos\phi + \mathrm{i}\sin\phi)$ in polar form where $r = 1/|\lambda| = 1/R \leq 1$ and $\phi = -\theta \in (-\pi, \pi)$. By Lemma 3.A.3, we have

$$\left(I - \frac{B}{\lambda}\right)^{-1} = P(\lambda) + \mathrm{i}Q(\lambda), \quad \left(I - \frac{B^*}{\lambda}\right)^{-1} = P(\lambda)^* + \mathrm{i}Q(\lambda)^*$$

where $P(\lambda)$ and $Q(\lambda)$ are $\mathbb{C}^{n_u\times n_u}$ matrices that satisfy the following bounds in the case $\|B\| < 1$ for all $|\lambda| \geq 1$:

$$\|P(\lambda)\| \le p := \frac{1}{1-\|B\|}, \tag{2.10}$$

$$\|Q(\lambda)\| \le q_1 := \frac{\|B\|}{1-\|B\|} \quad \text{and} \quad \|Q(\lambda)\| \le |\sin\theta| q_2 \quad \text{with} \quad q_2 := \frac{\|B\|}{(1-\|B\|)^2}. \tag{2.11}$$

These bounds still hold in the case $0 \neq \rho(B) < 1$ with

$$p = (1+\|B\|)s(B)^2, \quad q_1 = \|B\|s(B)^2 \quad \text{and} \quad q_2 = \|B\|s(B)^2. \tag{2.12}$$

To simplify the notation, we will not explicitly write the dependence of P and Q on λ. Now we rewrite (2.9) as

$$(1+\tau\alpha)(\lambda^2-\lambda) + \tau\alpha(\lambda-1) + \tau\alpha + \tau G(P^* + \mathrm{i}Q^*, P + \mathrm{i}Q) = 0 \tag{2.13}$$

where

$$G(X,Y) := \langle M^* X H^* H Y M y, y\rangle \in \mathbb{C}, \quad X, Y \in \mathbb{C}^{n_u \times n_u}.$$

Notice that G is a bilinear form and $G(X,Y) = G(Y^*, X^*)^*$ so that $G(X,Y) + G(Y^*, X^*)$ is real. With these properties of G, we expand (2.13) and take its real and imaginary parts, which yields

$$(1+\tau\alpha)\,\mathrm{Re}(\lambda^2-\lambda) + \tau\alpha\,\mathrm{Re}(\lambda-1) + \tau\alpha + \tau[G(P^*,P) - G(Q^*,Q)] = 0, \tag{2.14}$$

and

$$(1+\tau\alpha)\,\mathrm{Im}(\lambda^2-\lambda) + \tau\alpha\,\mathrm{Im}(\lambda-1) + \tau[G(P^*,Q) + G(Q^*,P)] = 0. \tag{2.15}$$

Step 2. Use a suitable combination of equations (2.14) and (2.15).

Let $\gamma \in \mathbb{R}$. Multiplying equation (2.15) with γ then summing it with equation (2.14), we obtain:

$$\begin{aligned}(1+\tau\alpha)[\mathrm{Re}(\lambda^2-\lambda) + \gamma\,\mathrm{Im}(\lambda^2-\lambda)] + \tau\alpha[\mathrm{Re}(\lambda-1) + \gamma\,\mathrm{Im}(\lambda-1)] \\ + \tau\alpha + \tau[G(P^*,P) - G(Q^*,Q) + \gamma G(P^*,Q) + \gamma G(Q^*,P)] = 0,\end{aligned}$$

or equivalently,

$$\begin{aligned}(1+\tau\alpha)[\mathrm{Re}(\lambda^2-\lambda) + \gamma\,\mathrm{Im}(\lambda^2-\lambda)] + \tau\alpha[\mathrm{Re}(\lambda-1) + \gamma\,\mathrm{Im}(\lambda-1)] \\ + \tau\alpha + \tau G(P^* + \gamma Q^*, P + \gamma Q) - (1+\gamma^2)\tau G(Q^*,Q) = 0.\end{aligned} \tag{2.16}$$

Now we consider four cases for λ as in Lemma 3.A.4 (see Figure 2.1):

- *Case 1.* $\mathrm{Re}(\lambda^2-\lambda) \ge 0$;
- *Case 2.* $\mathrm{Re}(\lambda^2-\lambda) < 0$ and $\theta \in [\theta_0, \pi-\theta_0] \cup [-\pi+\theta_0, -\theta_0]$ for fixed $0 < \theta_0 \le \frac{\pi}{4}$;
- *Case 3.* $\mathrm{Re}(\lambda^2-\lambda) < 0$ and $\theta \in (-\theta_0, \theta_0)$ for fixed $0 < \theta_0 \le \frac{\pi}{4}$;
- *Case 4.* $\mathrm{Re}(\lambda^2-\lambda) < 0$ and $\theta \in (\pi-\theta_0, \pi) \cup (-\pi, -\pi+\theta_0)$ for fixed $0 < \theta_0 \le \frac{\pi}{4}$.

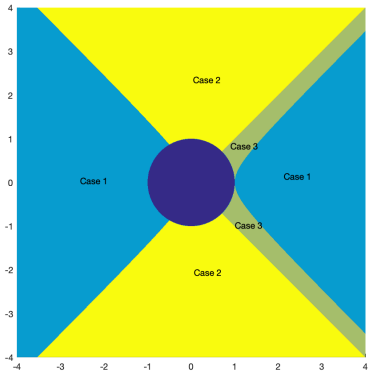


Figure 2.1: Illustration of the regions in the complex plane associated with the cases 1, 2 and 3 indicated in the proof of Proposition 2.2.4 for $\theta_0 = \frac{\pi}{4}$. The center circle is the unit circle.

Cases 1, 2 and 3 are respectively treated in Lemmas 2.2.5, 2.2.6 and 2.2.7 below. Notice that case 4 corresponds to an empty set, according to Lemma 3.A.4 *(iv)*. The statement of the proposition easily follows from the combination of Lemmas 2.2.5, 2.2.6 and 2.2.7. In particular, the parameter θ_0 is mainly used for case 2 to obtain a lower bound for $|\lambda - 1|$. This is why we require $\theta_0 > 0$. We need $\theta_0 \leq \frac{\pi}{4}$ to design a lower bound for a suitable combination of $\mathrm{Re}(\lambda^2 - \lambda)$ and $\mathrm{Im}(\lambda^2 - \lambda)$ in case 3. □

In the lemmas below we shall make use of the following obvious property where $\|y\| = 1$:

$$0 \leq G(X^*, X) = \|HXMy\|^2 \leq (\|H\|\|M\|\|X\|)^2, \quad \forall X \in \mathbb{C}^{n_u \times n_u}. \tag{2.17}$$

Lemma 2.2.5 (Case 1)**.** *Let $0 \neq \rho(B) < 1$ and let $|\lambda| \geq 1$ with $\mathrm{Re}(\lambda^2 - \lambda) \geq 0$. Then equation (2.9) cannot hold if one chooses*

$$\tau < \left(\|H\|^2\|M\|^2 s(B)^4 \cdot 4\|B\|^2 + (\sqrt{2} - 1)\alpha\right)^{-1}.$$

Moreover, if $\|B\| < 1$, the result is also true if

$$\tau < \left(\frac{\|H\|^2\|M\|^2}{(1 - \|B\|)^4} \cdot 4\|B\|^2 + (\sqrt{2} - 1)\alpha\right)^{-1}.$$

Proof. Define

$$\gamma_1 = \gamma_1(\lambda) := \begin{cases} 1 & \text{if } \mathrm{Im}(\lambda^2 - \lambda) \geq 0, \\ -1 & \text{if } \mathrm{Im}(\lambda^2 - \lambda) < 0 \end{cases}$$

as in Lemma 3.A.4 *(i)*. Writing (2.16) for $\gamma = \gamma_1$ and using $\gamma_1^2 = 1$ we obtain

$$\begin{aligned}(1 + \tau\alpha)[\mathrm{Re}(\lambda^2 - \lambda) + \gamma_1 \mathrm{Im}(\lambda^2 - \lambda)] + \tau\alpha[\mathrm{Re}(\lambda - 1) + \gamma_1 \mathrm{Im}(\lambda - 1)] \\ + \tau\alpha + \tau G(P^* + \gamma_1 Q^*, P + \gamma_1 Q) - 2\tau G(Q^*, Q) = 0.\end{aligned} \tag{2.18}$$

Since $G(P^* + \gamma_1 Q^*, P + \gamma_1 Q) \geq 0$ by (2.17) and $\tau\alpha \geq 0$, then the left-hand side of (2.18) is positive if τ satisfies

$$(1 + \tau\alpha)[\mathrm{Re}(\lambda^2 - \lambda) + \gamma_1 \mathrm{Im}(\lambda^2 - \lambda)] - \tau\alpha|\,\mathrm{Re}(\lambda - 1) + \gamma_1 \mathrm{Im}(\lambda - 1)| - 2\tau G(Q^*, Q) > 0,$$

or equivalently,

$$1 + \tau\alpha - \tau\alpha \frac{|\,\mathrm{Re}(\lambda - 1) + \gamma_1 \mathrm{Im}(\lambda - 1)|}{\mathrm{Re}(\lambda^2 - \lambda) + \gamma_1 \mathrm{Im}(\lambda^2 - \lambda)} - 2\tau \frac{G(Q^*, Q)}{\mathrm{Re}(\lambda^2 - \lambda) + \gamma_1 \mathrm{Im}(\lambda^2 - \lambda)} > 0. \tag{2.19}$$

Notice that the choice of γ_1 ensures $\operatorname{Re}(\lambda^2-\lambda)+\gamma_1\operatorname{Im}(\lambda^2-\lambda)>0$. By Lemma 3.A.4 *(i)* we have

$$\frac{|\operatorname{Re}(\lambda-1)+\gamma_1\operatorname{Im}(\lambda-1)|}{\operatorname{Re}(\lambda^2-\lambda)+\gamma_1\operatorname{Im}(\lambda^2-\lambda)}\leq\frac{\sqrt{1+\gamma_1^2}|\lambda-1|}{|\lambda(\lambda-1)|}=\frac{\sqrt{2}}{|\lambda|}\leq\sqrt{2}.$$

Using again Lemma 3.A.4 *(i)* and (2.17), we have

$$\begin{aligned}\frac{G(Q^*,Q)}{\operatorname{Re}(\lambda^2-\lambda)+\gamma_1\operatorname{Im}(\lambda^2-\lambda)}\leq\frac{(\|H\|\|M\|\sin\theta|q_2)^2}{2|\sin(\theta/2)|}&=2\left|\sin\frac{\theta}{2}\right|\cos^2\frac{\theta}{2}\|H\|^2\|M\|^2q_2^2\\&\leq 2\|H\|^2\|M\|^2q_2^2.\end{aligned}$$

Inserting the two previous inequalities in (2.19) gives the desired results using definitions (2.11) and (2.12) of q_2.

□

Lemma 2.2.6 (Case 2)**.** *Let $0\neq\rho(B)<1$ and let $|\lambda|\geq 1$, $\operatorname{Re}(\lambda^2-\lambda)<0$, $\theta\in[\theta_0,\pi-\theta_0]\cup[-\pi+\theta_0,-\theta_0]$ for given $0<\theta_0\leq\frac{\pi}{4}$. Then equation (2.9) cannot hold if one chooses*

$$\tau<\left(\|H\|^2\|M\|^2s(B)^4\cdot\frac{(1+2\|B\|)^2}{2\sin\frac{\theta_0}{2}}+\left(\sqrt{2}+\frac{1}{2\sin\frac{\theta_0}{2}}-1\right)\alpha\right)^{-1}.$$

Moreover, if $\|B\|<1$, the result is also true if

$$\tau<\left(\frac{\|H\|^2\|M\|^2}{(1-\|B\|)^2}\cdot\frac{(1+\|B\|)^2}{2\sin\frac{\theta_0}{2}}+\left(\sqrt{2}+\frac{1}{2\sin\frac{\theta_0}{2}}-1\right)\alpha\right)^{-1}.$$

Proof. Define

$$\gamma_2=\gamma_2(\lambda):=\begin{cases}-1 & \text{if } \operatorname{Im}(\lambda^2-\lambda)\geq 0,\\ 1 & \text{if } \operatorname{Im}(\lambda^2-\lambda)<0\end{cases}$$

as in Lemma 3.A.4 *(ii)*. Writing (2.16) for $\gamma=\gamma_2$ and using $\gamma_2^2=1$ we obtain

$$\begin{aligned}(1+\tau\alpha)[\operatorname{Re}(\lambda^2-\lambda)+\gamma_2\operatorname{Im}(\lambda^2-\lambda)]+\tau\alpha[\operatorname{Re}(\lambda-1)+\gamma_2\operatorname{Im}(\lambda-1)]\\+\tau\alpha+\tau G(P^*+\gamma_2Q^*,P+\gamma_2Q)-2\tau G(Q^*,Q)=0.\end{aligned}\tag{2.20}$$

Since $G(Q^*,Q)\geq 0$ by (2.17), the left-hand side of (2.20) is negative if τ satisfies

$$\begin{aligned}(1+\tau\alpha)[\operatorname{Re}(\lambda^2-\lambda)+\gamma_2\operatorname{Im}(\lambda^2-\lambda)]+\tau\alpha|\operatorname{Re}(\lambda-1)+\gamma_2\operatorname{Im}(\lambda-1)|\\+\tau\alpha+\tau G(P^*+\gamma_2Q^*,P+\gamma_2Q)<0,\end{aligned}$$

or equivalently,

$$\begin{aligned}-1-\tau\alpha+\tau\alpha\frac{|\operatorname{Re}(\lambda-1)+\gamma_2\operatorname{Im}(\lambda-1)|}{|\operatorname{Re}(\lambda^2-\lambda)+\gamma_2\operatorname{Im}(\lambda^2-\lambda)|}+\frac{\tau\alpha}{|\operatorname{Re}(\lambda^2-\lambda)+\gamma_2\operatorname{Im}(\lambda^2-\lambda)|}\\+\tau\frac{G(P^*+\gamma_2Q^*,P+\gamma_2Q)}{|\operatorname{Re}(\lambda^2-\lambda)+\gamma_2\operatorname{Im}(\lambda^2-\lambda)|}<0.\end{aligned}\tag{2.21}$$

Notice that the choice of γ_2 ensures $\operatorname{Re}(\lambda^2-\lambda)+\gamma_2\operatorname{Im}(\lambda^2-\lambda)<0$. In the following we derive upper bounds independent from λ for the terms appearing with the negative sign in (2.21). By Lemma 3.A.4 *(ii)* we have

$$\frac{|\operatorname{Re}(\lambda-1)+\gamma_2\operatorname{Im}(\lambda-1)|}{|\operatorname{Re}(\lambda^2-\lambda)+\gamma_2\operatorname{Im}(\lambda^2-\lambda)|}\leq\frac{\sqrt{1+\gamma_2^2}|\lambda-1|}{|\lambda(\lambda-1)|}=\frac{\sqrt{2}}{|\lambda|}\leq\sqrt{2}$$

and

$$\frac{1}{|\operatorname{Re}(\lambda^2-\lambda)+\gamma_2\operatorname{Im}(\lambda^2-\lambda)|}\leq\frac{1}{2\sin\frac{\theta_0}{2}}.$$

Using again Lemma 3.A.4 *(ii)* and (2.17), we have

$$\frac{G(P^*+\gamma_2Q^*,P+\gamma_2Q)}{|\operatorname{Re}(\lambda^2-\lambda)+\gamma_2\operatorname{Im}(\lambda^2-\lambda)|}\leq\frac{\|H\|^2\|M\|^2(p+q_1)^2}{2\sin\frac{\theta_0}{2}}.$$

Inserting these previous inequalities in (2.21) gives the desired results using definitions (2.10), (2.11) and (2.12) of p and q_1. □

Lemma 2.2.7 (Case 3). *Let $0\neq\rho(B)<1$ and let $|\lambda|\geq 1$, $\operatorname{Re}(\lambda^2-\lambda)<0$, $\theta\in(-\theta_0,\theta_0)$ for given $0<\theta_0\leq\frac{\pi}{4}$. For any $\delta_0>0$, equation (2.9) cannot hold if one chooses*

$$\tau<\left(\|H\|^2\|M\|^2s(B)^4\cdot\frac{2c}{\delta_0}\|B\|^2+\left(\frac{\sqrt{c}}{\delta_0}-1\right)\alpha\right)^{-1}$$

where $c=c(\theta_0,\delta_0):=\left(1+2\delta_0\sin\frac{3\theta_0}{2}+\delta_0^2\right)/\cos^2\frac{3\theta_0}{2}$. Moreover, if $0<\|B\|<1$, the result is also true if

$$\tau<\left(\frac{\|H\|^2\|M\|^2}{(1-\|B\|)^{-4}}\cdot\frac{2c}{\delta_0}\|B\|^2+\left(\frac{\sqrt{c}}{\delta_0}-1\right)\alpha\right)^{-1}.$$

Proof. Define

$$\gamma_3=\gamma_3(\operatorname{sign}(\theta)):=\begin{cases}\left(\delta_0+\sin\frac{3\theta_0}{2}\right)/\cos\frac{3\theta_0}{2} & \text{if } \theta>0,\\ -\left(\delta_0+\sin\frac{3\theta_0}{2}\right)/\cos\frac{3\theta_0}{2} & \text{if } \theta<0\end{cases}$$

as in Lemma 3.A.4 *(iii)*. Writing (2.16) for $\gamma=\gamma_3$ we obtain

$$\begin{aligned}(1+\tau\alpha)[\operatorname{Re}(\lambda^2-\lambda)+\gamma_3\operatorname{Im}(\lambda^2-\lambda)]+\tau\alpha[\operatorname{Re}(\lambda-1)+\gamma_3\operatorname{Im}(\lambda-1)]\\ +\tau\alpha+\tau G(P^*+\gamma_3Q^*,P+\gamma_3Q)-\tau(1+\gamma_3^2)G(Q^*,Q)=0.\end{aligned}\tag{2.22}$$

Since $G(P^*+\gamma_3Q^*,P+\gamma_3Q)\geq 0$ by (2.17) and $\tau\alpha\geq 0$, the left-hand side of (2.22) is positive if τ satisfies

$$(1+\tau\alpha)[\operatorname{Re}(\lambda^2-\lambda)+\gamma_3\operatorname{Im}(\lambda^2-\lambda)]-\tau\alpha|\operatorname{Re}(\lambda-1)+\gamma_3\operatorname{Im}(\lambda-1)|-\tau(1+\gamma_3^2)G(Q^*,Q)>0,$$

or equivalently,

$$1+\tau\alpha-\tau\alpha\frac{|\operatorname{Re}(\lambda-1)+\gamma_3\operatorname{Im}(\lambda-1)|}{\operatorname{Re}(\lambda^2-\lambda)+\gamma_3\operatorname{Im}(\lambda^2-\lambda)}-\tau(1+\gamma_3^2)\frac{G(Q^*,Q)}{\operatorname{Re}(\lambda^2-\lambda)+\gamma_3\operatorname{Im}(\lambda^2-\lambda)}>0.\tag{2.23}$$

Notice that the choice of γ_3 ensures $\operatorname{Re}(\lambda^2-\lambda)+\gamma_3\operatorname{Im}(\lambda^2-\lambda)>0$, also

$$1+\gamma_3^2=1+\frac{\left(\delta_0+\sin\frac{3\theta_0}{2}\right)^2}{\cos^2\frac{3\theta_0}{2}}=\frac{1+2\delta_0\sin\frac{3\theta_0}{2}+\delta_0^2}{\cos^2\frac{3\theta_0}{2}}=:c$$

is a constant greater than δ_0^2. By Lemma 3.A.4 *(iii)* we have

$$\frac{|\operatorname{Re}(\lambda-1)+\gamma_3\operatorname{Im}(\lambda-1)|}{\operatorname{Re}(\lambda^2-\lambda)+\gamma_3\operatorname{Im}(\lambda^2-\lambda)}\leq\frac{\sqrt{1+\gamma_3^2}}{\delta_0}=\frac{\sqrt{c}}{\delta_0}.$$

Using again Lemma 3.A.4 *(iii)* and (2.17), we have

$$\frac{G(Q^*, Q)}{\mathrm{Re}(\lambda^2 - \lambda) + \gamma_1 \,\mathrm{Im}(\lambda^2 - \lambda)} \leq \frac{(\|H\|\|M\|| \sin\theta| q_2)^2}{2\delta_0 |\sin(\theta/2)|} = \frac{2}{\delta_0} \left|\sin\frac{\theta}{2}\right| \cos^2\frac{\theta}{2} \|H\|^2 \|M\|^2 q_2^2$$
$$\leq \frac{2}{\delta_0} \|H\|^2 \|M\|^2 q_2^2.$$

Inserting the two previous inequalities in (2.23) gives the desired results using definitions (2.11) and (2.12) of q_2. □

2.3 Final result ($k = 1$)

Considering Proposition 2.2.2 (for $B = 0$), Proposition 2.2.1 (for real eigenvalues and $B \neq 0$) and taking the bound in Proposition 2.2.4 (for complex eigenvalues and $B \neq 0$), we obtain a sufficient condition on the descent step τ to ensure convergence of the semi-implicit one-step one-shot method.

Theorem 2.3.1 (Convergence of semi-implicit one-step one-shot)**.** *Under assumption* (1.4)*, the one-step one-shot method* (1.11) *converges for sufficiently small τ. In particular, for $\|B\| < 1$, there exist an explicit piecewise (at most a 4th order) polynomial function $\mathcal{P}_1$ and a pure constant $C > 0$ such that $\mathcal{P}_1 > 0$ on $[0, 1)$ and it is enough to take*

$$\tau < \left(\frac{\|H\|^2 \|M\|^2}{(1 - \|B\|)^4} \mathcal{P}_1(\|B\|) + C\alpha \right)^{-1}.$$

We emphasize that the bound on τ in Theorem 2.3.1 depends only on $\|B\|, \|M\|, \|H\|$ and the regularization parameter α. This bound does not depend on the dimensions of σ, u and g.

Chapter 3

Multi-step one-shot methods for linear inverse problems

Contents

By developing the technique presented in the previous chapter for the semi-implicit one-step one-shot method, we shall analyze in this chapter the convergence of the semi-implicit k-step one-shot method (1.9). This slightly-modified technique is also applicable to the case $k = 1$, that explains why it is included in this chapter even though it was studied in the previous chapter. The main difficulty in the case $k \geq 1$ is that the eigenvalue equation now depends on k and is much more complicated. We then perform some numerical experiments on a toy conductivity inverse problem given by the Helmholtz equation. In these tests, we observe that very few inner iterations may still guarantee good convergence of the inversion algorithm, even in the presence of noisy data.

This chapter is organized as follows. The aim of Section 3.1 is to give a sufficient condition on the descent step τ that ensures the convergence of the multi-step one-shot methods (1.9). We first write in Section 3.1.1 the eigenvalue equation for the convergence study of the multi-step one-shot methods, then investigate in Section 3.1.2 the location of the eigenvalues in the complex plane. This is similar to what we did in Section 3.1.1 and 2.2. Our aim is to find sufficiently small descent steps τ such that the eigenvalue equation does not admit any solutions outside and on the unit circle $|z| = 1$. The final convergence theorem is stated in Section 3.1.3. Next, in Section 3.2, we do several numerical tests on the performance of the different algorithms on a toy 2D Helmholtz inverse problem. In the first experiment presented in Section 3.2.1, the measurements are noise-free and the

parameter that we desire to reconstruct is discretized in a low-dimensional space, while in the second experiment presented in Section 3.2.2, the measurements are affected by noise and the reconstructed parameter is discretized in a higher dimensional space. Moreover, using the same configuration for experiments as in Section 3.2.1 but a change mainly on the setting of the unknown conductivity contrast, we study in Section 3.2.3 the robustness with respect to the size of the discretized problem and in Section 3.2.4 the dependence of the number of outer iterations on the norm of B.

The content of this chapter is extracted from the article:

> [4] M. Bonazzoli, H. Haddar, and T.A. Vu (2023, preprint). On the convergence analysis of one-shot inversion methods. hal-04151014

3.1 Convergence of the multi-step one-shot method ($k \geq 1$)

We now tackle the general case of semi-implicit multi-step one-shot methods, that is algorithm (1.9)

$$\begin{cases} \sigma^{n+1} = \sigma^n - \tau M^* p^n - \tau\alpha\sigma^{n+1}, \\ u_0^{n+1} = u^n, p_0^{n+1} = p^n, \\ \text{for } \ell = 0, 1, \dots, k-1 : \\ \quad \left| \begin{array}{l} u_{\ell+1}^{n+1} = B u_\ell^{n+1} + M\sigma^{n+1} + F, \\ p_{\ell+1}^{n+1} = B^* p_\ell^{n+1} + H^*(H u_\ell^{n+1} - g), \end{array} \right. \\ u^{n+1} := u_k^{n+1}, p^{n+1} := p_k^{n+1} \end{cases} \tag{1.9 recalled}$$

with $k \geq 1$. The procedure is quite similar to the case $k = 1$ in Chapter 2 but with more involved technicalities.

3.1.1 Block iteration matrix and eigenvalue equation

Let $k \geq 1$ be the number of inner iterations for u and p. First we express $(\sigma^{n+1}, u^{n+1}, p^{n+1})$ in terms of (σ^n, u^n, p^n) in a matrix form as for the case $k = 1$. More precisely, the system (1.9) can be equivalently written as

$$\begin{cases} \sigma^{n+1} = \sigma^n - \tau M^* p^n, \\ u^{n+1} = B^k u^n + T_k M \sigma^n - \tau T_k M M^* p^n + T_k F, \\ p^{n+1} = [(B^*)^k - \tau X_k M M^*] p^n + U_k u^n + X_k M \sigma^n + X_k F - T_k^* H^* g \end{cases} \tag{3.1}$$

where

$$T_k := I + B + \dots + B^{k-1} = (I - B)^{-1}(I - B^k), \tag{3.2}$$

$$U_k := (B^*)^{k-1} H^* H + (B^*)^{k-2} H^* H B + \dots + H^* H B^{k-1}, \tag{3.3}$$

$$X_k := \begin{cases} (B^*)^{k-2} H^* H T_1 + (B^*)^{k-3} H^* H T_2 + \dots + H^* H T_{k-1} & \text{if } k \geq 2, \\ 0 & \text{if } k = 1. \end{cases} \tag{3.4}$$

Before analyzing recursion (3.1), we gather in the following lemma some useful properties of T_k, U_k and X_k.

Lemma 3.1.1. *We have the following properties.*

(i) The matrices U_k and X_k can be respectively rewritten as

$$U_k = \sum_{i+j=k-1} (B^*)^i H^* H B^j, \quad \forall k \geq 1,$$

and

$$X_k = \sum_{l=0}^{k-2} \sum_{i+j=l} (B^*)^i H^* H B^j = \sum_{l=1}^{k-1} U_l, \quad \forall k \geq 2.$$

(ii) The matrices U_k and X_k are self-adjoint: $U_k^* = U_k$, $X_k^* = X_k$.

(iii) We have the relation

$$U_k T_k - X_k B^k + X_k = T_k^* H^* H T_k, \quad \forall k \geq 1. \tag{3.5}$$

Proof. Property *(i)* is easy to check from the definitions (3.2), (3.3) and (3.4). Property *(ii)* straightforwardly follows from *(i)*.

Now we prove *(iii)*. For $k = 1$, we have $U_1 = H^*H$, $T_1 = I$ and $X_1 = 0$, hence the identity is verified. For $k \geq 2$, we remark that $X_{k+1} = B^* X_k + H^* H T_k$ by definition (3.4). Thanks to *(ii)* we know that X_{k+1} is self-adjoint, hence $X_{k+1} = X_{k+1}^* = X_k B + T_k^* H^* H$. On the other hand, from *(i)* we get that $X_{k+1} = X_k + U_k$. Thus,

$$X_k + U_k = X_k B + T_k^* H^* H, \quad \text{or equivalently,} \quad U_k = X_k (B - I) + T_k^* H^* H.$$

Finally,

$$U_k T_k = X_k (B - I) T_k + T_k^* H^* H T_k = X_k (B^k - I) + T_k^* H^* H T_k.$$

□

Now, we consider the errors $(\sigma^n - \sigma_\alpha^{\text{ex}}, u^n - u(\sigma_\alpha^{\text{ex}}), p^n - p(\sigma_\alpha^{\text{ex}}))$ with respect to the regularized solution at the n-th iteration, and, by abuse of notation, we denote them by (σ^n, u^n, p^n). We obtain that the errors satisfy

$$\begin{cases} \sigma^{n+1} = \frac{1}{1+\tau\alpha}\sigma^n - \frac{\tau}{1+\tau\alpha} M^* p^n, \\ u^{n+1} = B^k u^n + \frac{1}{1+\tau\alpha} T_k M \sigma^n - \frac{\tau}{1+\tau\alpha} T_k M M^* p^n, \\ p^{n+1} = \left[(B^*)^k - \frac{\tau}{1+\tau\alpha} X_k M M^* \right] p^n + U_k u^n + \frac{1}{1+\tau\alpha} X_k M \sigma^n, \end{cases} \tag{3.6}$$

or equivalently, by putting in evidence the block iteration matrix

$$\begin{bmatrix} \sigma^{n+1} \\ u^{n+1} \\ p^{n+1} \end{bmatrix} = \begin{bmatrix} \frac{1}{1+\tau\alpha} I & 0 & -\frac{\tau}{1+\tau\alpha} M^* \\ \frac{1}{1+\tau\alpha} T_k M & B^k & -\frac{\tau}{1+\tau\alpha} T_k M M^* \\ \frac{1}{1+\tau\alpha} X_k M & U_k & (B^*)^k - \frac{\tau}{1+\tau\alpha} X_k M M^* \end{bmatrix} \begin{bmatrix} \sigma^n \\ u^n \\ p^n \end{bmatrix}. \tag{3.7}$$

Proposition 3.1.2. *Assume that $\lambda \in \mathbb{C}$ is an eigenvalue of the iteration matrix in (3.7). If $\lambda \notin \text{Spec}(B)$ then $\exists\, y \in \mathbb{C}^{n_\sigma}, \|y\| = 1$ such that:*

$$(1+\tau\alpha)\lambda - 1 + \tau\lambda \langle M^* [\lambda I - (B^*)^k]^{-1} [(\lambda - 1) X_k + T_k^* H^* H T_k] (\lambda I - B^k)^{-1} M y, y \rangle = 0. \tag{3.8}$$

In particular, $\lambda = 1$ is not an eigenvalue of the iteration matrix.

The proof is similar to the proof of Proposition 2.1.1. The slight difference is that in the calculation we use (3.5) to simplify some terms and exploit the fact that T_k and $(\lambda I - B^k)^{-1}$ commute.

Proof. Since $\lambda \in \mathbb{C}$ is an eigenvalue of the iteration matrix, there exists a non-zero vector $(\tilde{p}, \tilde{u}, y) \in \mathbb{C}^{n_u+n_u+n_\sigma}$ such that

$$\begin{cases} \lambda y = \frac{1}{1+\tau\alpha} y - \frac{\tau}{1+\tau\alpha} M^* \tilde{p}, \\ \lambda \tilde{u} = B^k \tilde{u} + \frac{1}{1+\tau\alpha} T_k M y - \frac{\tau}{1+\tau\alpha} T_k M M^* \tilde{p}, \\ \lambda \tilde{p} = \left[(B^*)^k - \frac{\tau}{1+\tau\alpha} X_k M M^* \right] \tilde{p} + U_k \tilde{u} + \frac{1}{1+\tau\alpha} X_k M y. \end{cases}$$

By inserting the first equation into the second and the third equations, we simplify this system of equations as

$$\begin{cases} \lambda y = \frac{1}{1+\tau\alpha} y - \frac{\tau}{1+\tau\alpha} M^* \tilde{p}, \\ \lambda \tilde{u} = B^k \tilde{u} + \lambda T_k M y, \\ \lambda \tilde{p} = (B^*)^k \tilde{p} + U_k \tilde{u} + \lambda X_k M y. \end{cases} \tag{3.9}$$

The second equation in (3.9) gives us directly $\tilde{u}$ in terms of y:

$$\tilde{u} = \lambda (\lambda I - B^k)^{-1} T_k M y = \lambda T_k (\lambda I - B^k)^{-1} M y. \tag{3.10}$$

The third equation in (3.9) gives us $\tilde{p}$ in terms of $\tilde{u}$ and y:

$$\tilde{p} = [\lambda I - (B^*)^k]^{-1} U_k \tilde{u} + \lambda [\lambda I - (B^*)^k]^{-1} X_k M y,$$

then by combing with equation (3.10), we obtain $\tilde{p}$ in terms of y:

$$\tilde{p} = \lambda (\lambda I - (B^*)^k)^{-1} V (\lambda I - B^k)^{-1} M y, \tag{3.11}$$

where

$$V \coloneqq U_k T_k + X_k (\lambda I - B^k) = (\lambda - 1) X_k + T_k^* H^* H T_k$$

thanks to (3.5). We also see that $y \neq 0$; indeed if $y = 0$ then equations (3.10) and (3.11) give $u = 0$ and $p = 0$, that is a contradiction. Next, inserting the expression of $\tilde{p}$ in equation (3.11) back into the first equation in (3.9), we get

$$\lambda y = \frac{1}{1+\tau\alpha} y - \frac{\tau}{1+\tau\alpha} \lambda M^* (\lambda I - (B^*)^k)^{-1} V (\lambda I - B^k)^{-1} M y,$$

which leads to

$$\left[(1+\alpha\tau)\lambda - 1 \right] y + \tau \lambda M^* [\lambda I - (B^*)^k]^{-1} V (\lambda I - B^k)^{-1} M y = 0. \tag{3.12}$$

Finally, by taking the scalar product of (3.12) with y, then dividing by $\|y\|^2$, we obtain (3.8). Now assume that $\lambda = 1$ is an eigenvalue of the iteration matrix, then (3.8) yields

$$\alpha + \|H(I - B)^{-1} M y\|^2 = 0,$$

which cannot be true due to the injectivity of $H(I - B)^{-1} M$. □

In the following sections we will show that, for sufficiently small τ, equation (3.8) admits no solution $|\lambda| \geq 1$, thus algorithm (1.9) converges. When $\lambda \neq 0$, it is convenient to rewrite (3.8) as

$$(1+\tau\alpha)\lambda^2 - \lambda + \tau\langle M^* \left[I - (B^*)^k/\lambda\right]^{-1} [(\lambda-1)X_k + T_k^* H^* H T_k] \left(I - B^k/\lambda\right)^{-1} My, y\rangle = 0. \tag{3.13}$$

Remark 3.1.3. The simple scalar case where $n_u, n_\sigma, n_g = 1$ and $\alpha = 0$ is analyzed in Section 4.2, for which necessary and sufficient conditions on τ are derived.

Remark 3.1.4. Note that when $B = 0$ and $k \geq 2$, the semi-implicit k-step one-shot (1.9) is equivalent to the semi-implicit gradient descent method (1.7), which converges if and only if $(\rho(A^*A) - \alpha)\tau < 2$.

For the analysis we use some auxiliary results proved in Appendix 3.A, and the following bounds for $s(B^k), T_k, X_k$.

Lemma 3.1.5. *If $\|B\| < 1$ then for every $k \geq 1$:*

$$s(B^k) = s((B^*)^k) \leq \frac{1}{1 - \|B\|^k}, \quad \|T_k\| \leq \frac{1 - \|B\|^k}{1 - \|B\|}$$

and

$$\|X_k\| \leq \frac{\|H\|^2(1 - k\|B\|^{k-1} + (k-1)\|B\|^k)}{(1 - \|B\|)^2}.$$

Proof. The bound for $s(B^k)$ is proved using Lemma 3.A.2 and $\|B^k\| \leq \|B\|^k$. Next, from (3.2) we have

$$\|T_k\| \leq 1 + \|B\| + \ldots + \|B\|^{k-1} = \frac{1 - \|B\|^k}{1 - \|B\|}.$$

From (3.4), if $k \geq 2$ we have

$$\begin{aligned}\|X_k\| &\leq \|H\|^2\Big(\|B\|^{k-2} + \|B\|^{k-3}(1 + \|B\|) + \ldots + (1 + \|B\| + \ldots + \|B\|^{k-2})\Big) \\ &= \|H\|^2(1 + 2\|B\| + \ldots + (k-1)\|B\|^{k-2}) = \frac{\|H\|^2(1 - k\|B\|^{k-1} + (k-1)\|B\|^k)}{(1 - \|B\|)^2}.\end{aligned}$$

□

3.1.2 Location of eigenvalues in the complex plane

We first establish conditions on the descent step $\tau > 0$ such that the real eigenvalues stay inside the unit disk. Recall that we have already proved that $\lambda = 1$ is not an eigenvalue for any k.

Proposition 3.1.6 (Real eigenvalues)**.** *Let $0 \neq \rho(B) < 1$ and $\lambda \in \mathbb{R}, \lambda \neq 1, |\lambda| \geq 1$. Then equation (3.13) cannot hold if $\tau > 0$ and*

$$\left(\|M\|^2\|X_k\|s(B^k)^2 - \frac{1}{2}\alpha\right)\tau < 1.$$

Moreover, if $\|B\| < 1$, the result is also true if $\tau > 0$ and

$$\left(\frac{\|H\|^2\|M\|^2}{(1 - \|B\|)^2(1 - \|B\|^k)^2}\left(1 - k\|B\|^{k-1} + (k-1)\|B\|^k\right) - \frac{1}{2}\alpha\right)\tau < 1.$$

Proof. When $\lambda \in \mathbb{R}$ equation (3.13) can be rewritten as

$$(1+\tau\alpha)\lambda^2 - \lambda + \tau\|HT_k\left(I - \frac{B^k}{\lambda}\right)^{-1} My\|^2 + \tau(\lambda-1)\langle M^*\left[I - \frac{(B^*)^k}{\lambda}\right]^{-1} X_k \left(I - \frac{B^k}{\lambda}\right)^{-1} My, y\rangle = 0. \quad (3.14)$$

We show that we can choose τ so that the left-hand side of the above equation is positive. First, we note that

$$\left|\langle M^*\left[I - \frac{(B^*)^k}{\lambda}\right]^{-1} X_k \left(I - \frac{B^k}{\lambda}\right)^{-1} My, y\rangle\right| \leq \|M\|^2\|X_k\|s(B^k)^2.$$

If $\lambda > 1$, we rewrite equation (3.14) again as

$$(1+\tau\alpha)\lambda(\lambda-1) + \tau\alpha + \tau\|HT_k\left(I - \frac{B^k}{\lambda}\right)^{-1} My\|^2 + \tau(\lambda-1)\langle M^*\left[I - \frac{(B^*)^k}{\lambda}\right]^{-1} X_k \left(I - \frac{B^k}{\lambda}\right)^{-1} My, y\rangle = 0.$$

Since $\lambda(\lambda-1) \geq \lambda - 1$, $\|HT_k\left(I - \frac{B^k}{\lambda}\right)^{-1} My\|^2 \geq 0$ and $\tau\alpha \geq 0$, we choose τ such that

$$(1+\tau\alpha)(\lambda-1) - \tau(\lambda-1)\|M\|^2\|X_k\|s(B^k)^2 > 0,$$

or equivalently,

$$1 + \tau\alpha - \tau\|M\|^2\|X_k\|s(B^k) > 0.$$

If $\lambda \leq -1$, we consider equation (3.14). Since $\|HT_k\left(I - \frac{B^k}{\lambda}\right)^{-1} My\|^2 \geq 0$, we choose τ such that

$$(1+\tau\alpha)\lambda^2 - \lambda - \tau(1-\lambda)\|M\|^2\|X_k\|s(B^k)^2 > 0,$$

or equivalently,

$$\frac{(1+\tau\alpha)\lambda^2 - \lambda}{1-\lambda} - \tau\|M\|^2\|X_k\|s(B^k)^2 > 0,$$

Since the function $\lambda \mapsto \frac{(1+\tau\alpha)\lambda^2-\lambda}{\lambda-1}$ is decreasing on $(-\infty, -1]$, it suffices to choose τ such that

$$1 + \frac{\alpha}{2}\tau - \tau\|M\|^2\|X_k\|s(B^k)^2 > 0,$$

which proves the first statement of the proposition. Finally, the case $\|B\| < 1$ can be deduced using Lemma 3.1.5. □

For the general case of complex eigenvalues, the study is much more complicated and technical. The following proposition summarizes the results obtained.

Proposition 3.1.7 (Complex eigenvalues)**.** *If $0 \neq \rho(B) < 1$, there exists $\tau > 0$ sufficiently small such that equation (3.13) admits no solution $\lambda \in \mathbb{C}\backslash\mathbb{R}$, $|\lambda| \geq 1$. In particular, if $\|B\| < 1$, given any $\delta_0 > 0$ and $0 < \theta_0 < \frac{\pi}{4}$, one can choose*

$$\tau < \min_{1\leq i\leq 3}\left(\frac{\|H\|^2\|M\|^2}{(1-\|B\|)^2(1-\|B\|^k)^2}\psi_i(k, \|B\|) + C_i\alpha\right)^{-1},$$

where

$$\psi_1(k,b) := 4b^{2k} + \sqrt{2}[1 - kb^{k-1} + (k-1)b^k](1+b^k),$$
$$\psi_2(k,b) := \left(\frac{1}{2\sin\frac{\theta_0}{2}}(1-b^k)^2 + \sqrt{2}(1 - kb^{k-1} + (k-1)b^k)\right)(1+b^k)^2,$$
$$\psi_3(k,b) := \frac{2c\sin\frac{\theta_0}{2}}{\delta_0}b^{2k} + \frac{\sqrt{c}}{\delta_0}[1 - kb^{k-1} + (k-1)b^k](1+b^{2k}) + 2\max\left(\frac{\sqrt{c}}{\delta_0}, \frac{\sqrt{c}}{\cos 2\theta_0}\right)[1 - kb^{k-1} + (k-1)b^k]b^k,$$

$$C_1 := \sqrt{2} - 1, \quad C_2 := \sqrt{2} + \frac{1}{2\sin\frac{\theta_0}{2}} - 1, \quad C_3 := \frac{\sqrt{c}}{\delta_0} - 1 \quad \textit{and} \quad c := \frac{1 + 2\delta_0\sin\frac{3\theta_0}{2} + \delta_0^2}{\cos^2\frac{3\theta_0}{2}}.$$

Proof. **Step 1. Rewrite equation (3.13) by separating real and imaginary parts.**

Let $\lambda = R(\cos\theta + \mathrm{i}\sin\theta)$ in polar form where $R = |\lambda| \geq 1$ and $\theta \in (-\pi, \pi)$, $\theta \neq 0$. Write $1/\lambda = r(\cos\phi + \mathrm{i}\sin\phi)$ in polar form where $r = 1/|\lambda| = 1/R \leq 1$ and $\phi = -\theta \in (-\pi, \pi)$. By Lemma 3.A.3 applied to $T = B^k$, we have

$$\left(I - \frac{B^k}{\lambda}\right)^{-1} = P_k(\lambda) + \mathrm{i}Q_k(\lambda), \quad \left(I - \frac{(B^*)^k}{\lambda}\right)^{-1} = P_k(\lambda)^* + \mathrm{i}Q_k(\lambda)^*$$

where $P_k(\lambda)$ and $Q_k(\lambda)$ are $\mathbb{C}^{n_u \times n_u}$ matrices that satisfy the following bounds in the case $\|B\| < 1$ for all $|\lambda| \geq 1$:

$$\|P_k(\lambda)\| \leq p := \frac{1}{1 - \|B\|^k}, \tag{3.15}$$

$$\|Q_k(\lambda)\| \leq q_1 := \frac{\|B\|^k}{1 - \|B\|^k} \quad \text{and} \quad \|Q_k(\lambda)\| \leq q_2|\sin\theta| \quad \text{with} \quad q_2 := \frac{\|B\|^k}{(1 - \|B\|^k)^2} \tag{3.16}$$

These bounds still hold in the case $0 \neq \rho(B) < 1$ with

$$p := (1 + \|B^k\|)s(B^k)^2, \quad q_1 := \|B^k\|s(B^k)^2, \quad \text{and} \quad q_2 := \frac{\|B^k\|}{1 - \|B^k\|}. \tag{3.17}$$

To simplify the notation, we will not explicitly write the dependence of P_k and Q_k on λ. Now we rewrite (3.13) as

$$(1+\tau\alpha)(\lambda^2 - \lambda) + \tau\alpha(\lambda - 1) + \tau\alpha + \tau G_k(P_k^* + \mathrm{i}Q_k^*, P_k + \mathrm{i}Q_k) + \tau(\lambda - 1)L_k(P_k^* + \mathrm{i}Q_k^*, P_k + \mathrm{i}Q_k) = 0 \tag{3.18}$$

where

$$G_k(X,Y) = \langle M^*XT_k^*H^*HT_kYMy, y\rangle, \quad L_k(X,Y) = \langle M^*XX_kYMy, y\rangle, \quad X, Y \in \mathbb{C}^{n_u \times n_u}.$$

Notice that G_k is a bilinear form and $G_k(X,Y) = G_k(Y^*, X^*)^*$ so that $G_k(X,Y) + G_k(Y^*, X^*)$ is real. Similarly, L_k has the same properties as G_k (note that $X_k^* = X_k$

by Lemma 3.1.1). With these properties of G_k and L_k, we expand (3.1.2) and take its real and imaginary parts, which yields

$$(1+\tau\alpha)\,\mathrm{Re}(\lambda^2-\lambda)+\tau\alpha\,\mathrm{Re}(\lambda-1)+\tau\alpha+\tau G_{1,k}+\tau[\mathrm{Re}(\lambda-1)L_{1,k}-\mathrm{Im}(\lambda-1)L_{2,k}]=0 \tag{3.19}$$

and

$$(1+\tau\alpha)\,\mathrm{Im}(\lambda^2-\lambda)+\tau\alpha\,\mathrm{Im}(\lambda-1)+\tau G_{2,k}+\tau[\mathrm{Im}(\lambda-1)L_{1,k}+\mathrm{Re}(\lambda-1)L_{2,k}]=0 \tag{3.20}$$

where

$$G_{1,k}\coloneqq G_k(P_k^*,P_k)-G_k(Q_k^*,Q_k),\quad G_{2,k}\coloneqq G_k(P_k^*,Q_k)+G_k(Q_k^*,P_k),$$
$$L_{1,k}\coloneqq L_k(P_k^*,P_k)-L_k(Q_k^*,Q_k)\quad\text{and}\quad L_{2,k}\coloneqq L_k(P_k^*,Q_k)+L_k(Q_k^*,P_k).$$

Step 2. Use a suitable combination of equations (3.19) and (3.20).
Let $\gamma\in\mathbb{R}$. Multiplying equation (3.20) with γ then summing it with equation (3.19), we obtain:

$$\begin{aligned}(1+\tau\alpha)[\mathrm{Re}(\lambda^2-\lambda)+\gamma\,\mathrm{Im}(\lambda^2-\lambda)]+\tau\alpha[\mathrm{Re}(\lambda-1)+\gamma\,\mathrm{Im}(\lambda-1)]+\tau\alpha\\ +\tau G_k(P_k^*+\gamma Q_k^*,P_k+\gamma Q_k)-\tau(1+\gamma^2)G_k(Q_k^*,Q_k)\\ +\tau[\mathrm{Re}(\lambda-1)+\gamma\,\mathrm{Im}(\lambda-1)]L_{1,k}+\tau[\gamma\,\mathrm{Re}(\lambda-1)-\mathrm{Im}(\lambda-1)]L_{2,k}=0.\end{aligned} \tag{3.21}$$

Now we consider four cases of λ as in the proof for $k=1$ namely:

- *Case 1.* $\mathrm{Re}(\lambda^2-\lambda)\geq 0$;
- *Case 2.* $\mathrm{Re}(\lambda^2-\lambda)<0$ and $\theta\in[\theta_0,\pi-\theta_0]\cup[-\pi+\theta_0,-\theta_0]$ for fixed $0<\theta_0<\frac{\pi}{4}$;
- *Case 3.* $\mathrm{Re}(\lambda^2-\lambda)<0$ and $\theta\in(-\theta_0,\theta_0)$ for fixed $0<\theta_0<\frac{\pi}{4}$;
- *Case 4.* $\mathrm{Re}(\lambda^2-\lambda)<0$ and $\theta\in(\pi-\theta_0,\pi)\cup(-\pi,-\pi+\theta_0)$ for fixed $0<\theta_0<\frac{\pi}{4}$.

Cases 1, 2 and 3 are respectively treated in Lemmas 3.1.9, 3.1.10 and 3.1.11 below. Notice that case 4 corresponds to an empty set, according to the Lemma 3.A.4 *(iv)*. The statement of the proposition easily follows from the combination of Lemmas 3.1.9, 3.1.10 and 3.1.11. □

Remark 3.1.8. The main difference of the two techniques respectively presented in Chapter 2 and Chapter 3 lies in the choice of the technical parameter θ_0 in Lemma 3.A.4: in Chapter 2 we can take $0<\theta_0\leq\frac{\pi}{4}$ but in Chapter 3 we have to take $0<\theta_0<\frac{\pi}{4}$ since the new estimates contains a division by $\cos 2\theta_0$, see Lemma 3.1.11 the below.

In the lemmas below we shall make use of the following obvious properties where $\|y\|=1$:

$$0\leq G_k(X^*,X)=\|HT_kXMy\|^2\leq(\|H\|\|M\|\|T_k\|\|X\|)^2,\quad\forall X\in\mathbb{C}^{n_u\times n_u}. \tag{3.22}$$

$$\begin{aligned}|L_{1,k}|=|L_k(P_k^*,P_k)-L_k(Q_k^*,Q_k)|&\leq|L_k(P_k^*,P_k)|+|L_k(Q_k^*,Q_k)|\\ &\leq\|X_k\|\|M\|^2(\|P_k\|^2+\|Q_k\|^2)\leq\|X_k\|\|M\|^2(p^2+q_1^2)\end{aligned} \tag{3.23}$$

and

$$\begin{aligned}|L_{2,k}|=|L_k(P_k^*,Q_k)+L_k(Q_k^*,P_k)|&\leq|L_k(P_k^*,Q_k)|+|L_k(Q_k^*,P_k)|\\ &\leq 2\|X_k\|\|M\|^2\|P_k\|\|Q_k\|\leq 2\|X_k\|\|M\|^2pq_1.\end{aligned} \tag{3.24}$$

Lemma 3.1.9 (Case 1). *Let $\rho(B) < 1$ and let $|\lambda| \geq 1$ with $\mathrm{Re}(\lambda^2 - \lambda) \geq 0$. Then equation (3.13) cannot hold if $\tau > 0$ and*

$$\left(4\|H\|^2\|M\|^2\|T_k\|^2\|B^k\|^2 s(B^k)^4 + \sqrt{2}\|M\|^2\|X_k\|(1+2\|B^k\|)^2 s(B^k)^4 + (\sqrt{2}-1)\alpha\right)\tau < 1.$$

Moreover, if $\|B\| < 1$, the result is also true if $\tau > 0$ and

$$\left(\frac{\|H\|^2\|M\|^2}{(1-\|B\|)^2(1-\|B\|^k)^2}\psi_1(k,\|B\|) + (\sqrt{2}-1)\alpha\right)\tau < 1.$$

where $\psi_1(k,b) := 4b^{2k} + \sqrt{2}(1 - kb^{k-1} + (k-1)b^k)(1+b^k)$.

Proof. Define

$$\gamma_1 = \gamma_1(\lambda) := \begin{cases} 1 & \text{if } \mathrm{Im}(\lambda^2 - \lambda) \geq 0, \\ -1 & \text{if } \mathrm{Im}(\lambda^2 - \lambda) < 0 \end{cases}$$

as in Lemma 3.A.4 *(i)*. Writing (3.21) for $\gamma = \gamma_1$ as in Lemma 3.A.4 and using $\gamma_1^2 = 1$ we obtain

$$\begin{aligned}(1+\tau\alpha)[\mathrm{Re}(\lambda^2-\lambda) + \gamma_1 \mathrm{Im}(\lambda^2-\lambda)] + \tau\alpha[\mathrm{Re}(\lambda-1) + \gamma_1 \mathrm{Im}(\lambda-1)] + \tau\alpha \\ + \tau G_k(P_k^* + \gamma_1 Q_k^*, P_k + \gamma_1 Q_k) - 2\tau G_k(Q_k^*, Q_k) \\ + \tau[\mathrm{Re}(\lambda-1) + \gamma_1 \mathrm{Im}(\lambda-1)]L_{1,k} + \tau[\gamma_1 \mathrm{Re}(\lambda-1) - \mathrm{Im}(\lambda-1)]L_{2,k} = 0. \quad (3.25)\end{aligned}$$

Since $G_k(P_k^* + \gamma_1 Q_k^*, P_k + \gamma_1 Q_k) \geq 0$ by (3.22) and $\tau\alpha \geq 0$, the left-hand side of (3.25) is positive if τ satisfies

$$\begin{aligned}(1+\tau\alpha)[\mathrm{Re}(\lambda^2-\lambda) + \gamma_1 \mathrm{Im}(\lambda^2-\lambda)] - \tau\alpha[\mathrm{Re}(\lambda-1) + \gamma_1 \mathrm{Im}(\lambda-1)] - 2\tau G_k(Q_k^*, Q_k) \\ - \tau[\mathrm{Re}(\lambda-1) + \gamma_1 \mathrm{Im}(\lambda-1)]|L_{1,k}| - \tau[\gamma_1 \mathrm{Re}(\lambda-1) - \mathrm{Im}(\lambda-1)]|L_{2,k}| > 0,\end{aligned}$$

or equivalently,

$$\begin{aligned}1 + \tau\alpha - \tau\alpha\frac{|\mathrm{Re}(\lambda-1) + \gamma_1 \mathrm{Im}(\lambda-1)|}{\mathrm{Re}(\lambda^2-\lambda) + \gamma_1 \mathrm{Im}(\lambda^2-\lambda)} - 2\tau\frac{G_k(Q_k^*, Q_k)}{\mathrm{Re}(\lambda^2-\lambda) + \gamma_1 \mathrm{Im}(\lambda^2-\lambda)} \\ - \tau\frac{|\mathrm{Re}(\lambda-1) + \gamma_1 \mathrm{Im}(\lambda-1)|}{\mathrm{Re}(\lambda^2-\lambda) + \gamma_1 \mathrm{Im}(\lambda^2-\lambda)}|L_{1,k}| - \tau\frac{|\gamma_1 \mathrm{Re}(\lambda-1) - \mathrm{Im}(\lambda-1)|}{\mathrm{Re}(\lambda^2-\lambda) + \gamma_1 \mathrm{Im}(\lambda^2-\lambda)}|L_{2,k}| > 0. \quad (3.26)\end{aligned}$$

Notice that the choice of γ_1 ensures $\mathrm{Re}(\lambda^2-\lambda) + \gamma_1 \mathrm{Im}(\lambda^2-\lambda) > 0$. In the following we derive upper bounds independent from λ for the terms appearing with the negative sign in (3.26). By Lemma 3.A.4 *(i)* we have

$$\frac{|\mathrm{Re}(\lambda-1) + \gamma_1 \mathrm{Im}(\lambda-1)|}{\mathrm{Re}(\lambda^2-\lambda) + \gamma_1 \mathrm{Im}(\lambda^2-\lambda)} \leq \frac{\sqrt{1+\gamma_1^2}|\lambda-1|}{|\lambda(\lambda-1)|} = \frac{\sqrt{2}}{|\lambda|} \leq \sqrt{2}$$

and

$$\frac{|\gamma_1 \mathrm{Re}(\lambda-1) - \mathrm{Im}(\lambda-1)|}{\mathrm{Re}(\lambda^2-\lambda) + \gamma_1 \mathrm{Im}(\lambda^2-\lambda)} \leq \frac{\sqrt{1+\gamma_1^2}|\lambda-1|}{|\lambda(\lambda-1)|} = \frac{\sqrt{2}}{|\lambda|} \leq \sqrt{2}.$$

Using again Lemma 3.A.4 *(i)* and (3.24), we have

$$\begin{aligned}\frac{G_k(Q_k^*, Q_k)}{\mathrm{Re}(\lambda^2-\lambda) + \gamma_1 \mathrm{Im}(\lambda^2-\lambda)} &\leq \frac{(\|H\|\|M\|\|T_k\|\sin\theta|q_2)^2}{2|\sin(\theta/2)|} \\ &= 2\left|\sin\frac{\theta}{2}\right|\cos^2\frac{\theta}{2}\|H\|^2\|M\|^2\|T_k\|^2 q_2^2 \leq 2\|H\|^2\|M\|^2\|T_k\|^2 q_2^2.\end{aligned}$$

Recall from (3.23) and (3.24) that

$$|L_{1,k}| \leq \|X_k\|\|M\|^2(p^2+q_1^2), \quad |L_{2,k}| \leq 2\|X_k\|\|M\|^2 pq_1.$$

Inserting these previous inequalities in (3.26) gives the desired result thanks to expressions (3.15), (3.16) and (3.17) of respectively p, q_1 and q_2, and thanks to Lemma 3.1.5. □

Lemma 3.1.10 (Case 2)**.** *Let $\rho(B) < 1$ and let $|\lambda| \geq 1$, $\mathrm{Re}(\lambda^2 - \lambda) < 0$, $\theta \in [\theta_0, \pi - \theta_0] \cup [-\pi + \theta_0, -\theta_0]$ for given $0 < \theta_0 < \frac{\pi}{4}$. Then equation (3.13) cannot hold if $\tau > 0$ and*

$$\left(\left(\frac{1}{2\sin\frac{\theta_0}{2}}\|H\|^2\|M\|^2\|T_k\|^2 + \sqrt{2}\|M\|^2\|X_k\|\right)(1+2\|B^k\|)^2 s(B^k)^4 + (\sqrt{2}-1)\alpha\right)\tau < 1.$$

Moreover, if $\|B\| < 1$, the result is also true if

$$\tau < \left(\frac{\|H\|^2\|M\|^2}{(1-\|B\|)^2(1-\|B\|^k)^2}\psi_2(k,\|B\|) + (\sqrt{2}-1)\alpha\right)^{-1}$$

where $\psi_2(k,b) = \left(\frac{1}{2\sin\frac{\theta_0}{2}}(1-b^k)^2 + \sqrt{2}(1-kb^{k-1}+(k-1)b^k)\right)(1+b^k)^2$.

Proof. Define

$$\gamma_2 = \gamma_2(\lambda) := \begin{cases} -1 & \text{if } \mathrm{Im}(\lambda^2-\lambda) \geq 0, \\ 1 & \text{if } \mathrm{Im}(\lambda^2-\lambda) < 0 \end{cases}$$

as in Lemma 3.A.4 *(ii)*. Writing (3.21) for $\gamma = \gamma_2$ as in Lemma 3.A.4 *(ii)* and using $\gamma_2^2 = 1$, we obtain

$$\begin{aligned}(1+\tau\alpha)[\mathrm{Re}(\lambda^2-\lambda) + \gamma_2\,\mathrm{Im}(\lambda^2-\lambda)] + \tau\alpha[\mathrm{Re}(\lambda-1) + \gamma_2\,\mathrm{Im}(\lambda-1)] + \tau\alpha \\ + \tau G_k(P_k^* + \gamma_2 Q_k^*, P_k + \gamma_2 Q_k) - 2\tau G_k(Q_k^*, Q_k) \\ + \tau[\mathrm{Re}(\lambda-1) + \gamma_2\,\mathrm{Im}(\lambda-1)]L_{1,k} + \tau[\gamma_2\,\mathrm{Re}(\lambda-1) - \mathrm{Im}(\lambda-1)]L_{2,k} = 0. \quad (3.27)\end{aligned}$$

Since $G_k(Q_k^*, Q_k) \geq 0$ by (3.22), the left-hand side of (3.27) is negative if τ satisfies

$$\begin{aligned}(1+\tau\alpha)[\mathrm{Re}(\lambda^2-\lambda) + \gamma_2\,\mathrm{Im}(\lambda^2-\lambda)] + \tau\alpha|\,\mathrm{Re}(\lambda-1) + \gamma_2\,\mathrm{Im}(\lambda-1)| + \tau\alpha \\ + \tau G_k(P_k^* + \gamma_2 Q_k^*, P_k + \gamma_2 Q_k) \\ + \tau|\,\mathrm{Re}(\lambda-1) + \gamma_1\,\mathrm{Im}(\lambda-1)||L_{1,k}| + \tau|\gamma_1\,\mathrm{Re}(\lambda-1) - \mathrm{Im}(\lambda-1)||L_{2,k}| < 0.\end{aligned}$$

or equivalently,

$$\begin{aligned}-1 - \tau\alpha + \tau\alpha\frac{|\,\mathrm{Re}(\lambda-1) + \gamma_2\,\mathrm{Im}(\lambda-1)|}{|\,\mathrm{Re}(\lambda^2-\lambda) + \gamma_2\,\mathrm{Im}(\lambda^2-\lambda)|} \\ + \frac{\tau\alpha}{|\,\mathrm{Re}(\lambda^2-\lambda) + \gamma_2\,\mathrm{Im}(\lambda^2-\lambda)|} + \tau\frac{G_k(P_k^* + \gamma_2 Q_k^*, P_k + \gamma_2 Q_k)}{|\,\mathrm{Re}(\lambda^2-\lambda) + \gamma_2\,\mathrm{Im}(\lambda^2-\lambda)|} \\ + \tau\frac{|\,\mathrm{Re}(\lambda-1) + \gamma_2\,\mathrm{Im}(\lambda-1)|}{|\,\mathrm{Re}(\lambda^2-\lambda) + \gamma_2\,\mathrm{Im}(\lambda^2-\lambda)|}|L_{1,k}| + \tau\frac{|\gamma_2\,\mathrm{Re}(\lambda-1) - \mathrm{Im}(\lambda-1)|}{|\,\mathrm{Re}(\lambda^2-\lambda) + \gamma_2\,\mathrm{Im}(\lambda^2-\lambda)|}|L_{2,k}| < 0.\end{aligned} \quad (3.28)$$

Notice that the choice of γ_2 ensures $\mathrm{Re}(\lambda^2-\lambda) + \gamma_2\,\mathrm{Im}(\lambda^2-\lambda) < 0$. In the following we derive upper bounds independent from λ for the terms appearing with the positive sign in (3.28). By Lemma 3.A.4 *(ii)* we have

$$\frac{|\,\mathrm{Re}(\lambda-1) + \gamma_2\,\mathrm{Im}(\lambda-1)|}{|\,\mathrm{Re}(\lambda^2-\lambda) + \gamma_2\,\mathrm{Im}(\lambda^2-\lambda)|} \leq \frac{\sqrt{1+\gamma_2^2}|\lambda-1|}{|\lambda(\lambda-1)|} = \frac{\sqrt{2}}{|\lambda|} \leq \sqrt{2},$$

$$\frac{|\gamma_2 \operatorname{Re}(\lambda-1) - \operatorname{Im}(\lambda-1)|}{|\operatorname{Re}(\lambda^2-\lambda) + \gamma_2 \operatorname{Im}(\lambda^2-\lambda)|} \leq \frac{\sqrt{1+\gamma_2^2}|\lambda-1|}{|\lambda(\lambda-1)|} = \frac{\sqrt{2}}{|\lambda|} \leq \sqrt{2},$$

and

$$\frac{1}{|\operatorname{Re}(\lambda^2-\lambda) + \gamma_2 \operatorname{Im}(\lambda^2-\lambda)|} \leq \frac{1}{2\sin\frac{\theta_0}{2}}.$$

Using again Lemma 3.A.4 *(ii)* and (3.22), we have

$$\frac{G_k(P_k^* + \gamma_2 Q_k^*, P_k + \gamma_2 Q_k)}{|\operatorname{Re}(\lambda^2-\lambda) + \gamma_2 \operatorname{Im}(\lambda^2-\lambda)|} \leq \frac{\|H\|^2\|M\|^2\|T_k\|^2(p+q_1)^2}{2\sin\frac{\theta_0}{2}}.$$

Recall from (3.23) and (3.24) that

$$|L_{1,k}| \leq \|X_k\|\|M\|^2(p^2+q_1^2), \quad |L_{2,k}| \leq 2\|X_k\|\|M\|^2 p q_1.$$

Inserting these previous inequalities in (3.28) gives the desired result thanks to expressions (3.15), (3.16) and (3.17) of respectively p and q_1, and thanks to Lemma 3.1.5. □

Lemma 3.1.11 (Case 3)**.** *Let $\rho(B) < 1$ and let $|\lambda| \geq 1$, $\operatorname{Re}(\lambda^2-\lambda) < 0$, $\theta \in (-\theta_0, \theta_0)$ for given $0 < \theta_0 < \frac{\pi}{4}$. For any $\delta_0 > 0$, equation (3.13) cannot hold if $\tau > 0$ and*

$$\left(\left[\frac{2c\sin\frac{\theta_0}{2}}{\delta_0}\|H\|^2\|M\|^2\|T_k\|^2\|B^k\|^2 + \frac{\sqrt{c}}{\delta_0}\|M\|^2\|X_k\|(1+2\|B^k\|+2\|B^k\|^2)\right.\right.$$
$$\left.\left.+2\max\left(\frac{\sqrt{c}}{\delta_0}, \frac{\sqrt{c}}{\cos 2\theta_0}\right)\|M\|^2\|X_k\|(\|B^k\|+\|B^k\|^2)\right] s(B^k)^2 + (\sqrt{2}-1)\alpha\right)\tau < 1$$

where $c = c(\theta_0, \delta_0) := \left(1 + 2\delta_0 \sin\frac{3\theta_0}{2} + \delta_0^2\right) / \cos^2\frac{3\theta_0}{2}$. Moreover, if $\|B\| < 1$, the result is also true if $\tau > 0$ and

$$\left(\frac{\|H\|^2\|M\|^2}{(1-\|B\|)^2(1-\|B\|^k)^2}\psi_3(k, \|B\|) + \left(\frac{\sqrt{c}}{\delta_0} - 1\right)\alpha\right)\tau < 1$$

where

$$\psi_3(k,b) := \frac{2c\sin\frac{\theta_0}{2}}{\delta_0}b^{2k} + \left(\frac{\sqrt{c}}{\delta_0}(1+b^{2k}) + 2\max\left(\frac{\sqrt{c}}{\delta_0}, \frac{\sqrt{c}}{\cos 2\theta_0}\right)b^k\right)(1 - kb^{k-1} + (k-1)b^k)b^k.$$

Proof. Define

$$\gamma_3 = \gamma_3(\operatorname{sign}(\theta)) := \begin{cases} \left(\delta_0 + \sin\frac{3\theta_0}{2}\right) / \cos\frac{3\theta_0}{2} & \text{if } \theta > 0, \\ -\left(\delta_0 + \sin\frac{3\theta_0}{2}\right) / \cos\frac{3\theta_0}{2} & \text{if } \theta < 0 \end{cases}$$

as in Lemma 3.A.4 *(iii)*. Writing (3.21) for $\gamma = \gamma_3$ we obtain

$$(1+\tau\alpha)[\operatorname{Re}(\lambda^2-\lambda) + \gamma_3 \operatorname{Im}(\lambda^2-\lambda)] + \tau\alpha[\operatorname{Re}(\lambda-1) + \gamma_3 \operatorname{Im}(\lambda-1)] + \tau\alpha$$
$$+ \tau G_k(P_k^* + \gamma_3 Q_k^*, P_k + \gamma_3 Q_k) - (1+\gamma_3^2)\tau G_k(Q_k^*, Q_k)$$
$$+ \tau\left([\operatorname{Re}(\lambda-1) + \gamma_3 \operatorname{Im}(\lambda-1)]L_{1,k} + [\gamma_3 \operatorname{Re}(\lambda-1) - \operatorname{Im}(\lambda-1)]L_{2,k}\right) = 0. \quad (3.29)$$

Since $G(P^* + \gamma_3 Q^*, P + \gamma_3 Q) \geq 0$ and $\tau\alpha \geq 0$, the left-hand side of (3.29) is positive if τ satisfies

$$(1+\tau\alpha)[\mathrm{Re}(\lambda^2-\lambda)+\gamma_3\,\mathrm{Im}(\lambda^2-\lambda)]-\tau\alpha|\,\mathrm{Re}(\lambda-1)+\gamma_3\,\mathrm{Im}(\lambda-1)|-(1+\gamma_3^2)\tau G_k(Q_k^*,Q_k)$$
$$-\tau|\,\mathrm{Re}(\lambda-1)+\gamma_3\,\mathrm{Im}(\lambda-1)||L_{1,k}|+|\gamma_3\,\mathrm{Re}(\lambda-1)-\mathrm{Im}(\lambda-1)||L_{2,k}|>0,$$

or equivalently,

$$1+\tau\alpha-\tau\alpha\frac{|\,\mathrm{Re}(\lambda-1)+\gamma_3\,\mathrm{Im}(\lambda-1)|}{\mathrm{Re}(\lambda^2-\lambda)+\gamma_3\,\mathrm{Im}(\lambda^2-\lambda)}-\tau(1+\gamma_3^2)\frac{G_k(Q_k^*,Q_k)}{\mathrm{Re}(\lambda^2-\lambda)+\gamma_3\,\mathrm{Im}(\lambda^2-\lambda)}$$
$$-\tau\frac{|\,\mathrm{Re}(\lambda-1)+\gamma_3\,\mathrm{Im}(\lambda-1)|}{\mathrm{Re}(\lambda^2-\lambda)+\gamma_3\,\mathrm{Im}(\lambda^2-\lambda)}|L_{1,k}|-\tau\frac{|\gamma_3\,\mathrm{Re}(\lambda-1)-\mathrm{Im}(\lambda-1)|}{\mathrm{Re}(\lambda^2-\lambda)+\gamma_3\,\mathrm{Im}(\lambda^2-\lambda)}|L_{2,k}|>0. \quad (3.30)$$

Notice that the choice of γ_3 ensures $\mathrm{Re}(\lambda^2-\lambda)+\gamma_3\,\mathrm{Im}(\lambda^2-\lambda)>0$, also

$$1+\gamma_3^2=1+\frac{\left(\delta_0+\sin\frac{3\theta_0}{2}\right)^2}{\cos^2\frac{3\theta_0}{2}}=\frac{1+2\delta_0\sin\frac{3\theta_0}{2}+\delta_0^2}{\cos^2\frac{3\theta_0}{2}}=:c$$

is a constant greater than δ_0^2. In the following we derive upper bounds independent from λ for the terms appearing with the negative sign in (3.30). By Lemma 3.A.4 *(iii)* we have

$$\frac{|\,\mathrm{Re}(\lambda-1)+\gamma_3\,\mathrm{Im}(\lambda-1)|}{\mathrm{Re}(\lambda^2-\lambda)+\gamma_3\,\mathrm{Im}(\lambda^2-\lambda)}\leq\frac{\sqrt{1+\gamma_3^2}}{\delta_0}=\frac{\sqrt{c}}{\delta_0}$$

and

$$\frac{|\gamma_3\,\mathrm{Re}(\lambda-1)-\mathrm{Im}(\lambda-1)|}{\mathrm{Re}(\lambda^2-\lambda)+\gamma_3\,\mathrm{Im}(\lambda^2-\lambda)}\leq\max\left(\frac{\sqrt{1+\gamma_3^2}}{\delta_0},\frac{\sqrt{1+\gamma_3^2}}{\cos 2\theta_0}\right)=\max\left(\frac{\sqrt{c}}{\delta_0},\frac{\sqrt{c}}{\cos 2\theta_0}\right).$$

Using again Lemma 3.A.4 *(iii)* and (3.22), we have

$$\frac{G_k(Q_k^*,Q_k)}{\mathrm{Re}(\lambda^2-\lambda)+\gamma_3\,\mathrm{Im}(\lambda^2-\lambda)}\leq\frac{(\|H\|\|M\|\|T_k\|\sin\theta|q_2)^2}{2\delta_0|\sin(\theta/2)|}$$
$$=\frac{2}{\delta_0}\left|\sin\frac{\theta}{2}\right|\cos^2\frac{\theta}{2}\|H\|^2\|M\|^2\|T_k\|^2q_2^2\leq\frac{2}{\delta_0}\|H\|^2\|M\|^2\|T_k\|^2q_2^2.$$

Recall from (3.23) and (3.24) that

$$|L_{1,k}|\leq\|X_k\|\|M\|^2(p^2+q_1^2),\quad |L_{2,k}|\leq 2\|X_k\|\|M\|^2pq_1.$$

Inserting these previous inequalities in (3.30) gives the desired result thanks to expressions (3.15), (3.16) and (3.17) of respectively p, q_1 and q_2, and thanks to Lemma 3.1.5. □

3.1.3 Final result ($k \geq 1$)

Considering Proposition 3.1.6 (for real eigenvalues) and taking the bound in Proposition 3.1.7 (for complex eigenvalues), we finally obtain a sufficient condition on the descent step τ to ensure convergence of the multi-step one-shot method.

Theorem 3.1.12 (Convergence of semi-implicit k-step one-shot, $k \geq 1$)**.** *Under assumption (1.4), the k-step one-shot method (1.9), $k \geq 1$, converges for sufficiently small τ. In particular, for $\|B\| < 1$, there exist an explicit piecewise (at most a $(4k)$th order) polynomial function $\mathcal{P}_k$ and a pure constant $C > 0$ independent of k such that $\mathcal{P}_k > 0$ on $[0, 1)$ and it is enough to take $\tau > 0$ and*

$$\left(\frac{\|H\|^2 \|M\|^2}{(1 - \|B\|)^2 (1 - \|B\|^k)^2} \mathcal{P}_k(\|B\|) + C\alpha \right) \tau < 1.$$

We emphasize that the bound of τ in Theorem 3.1.12 depends only on $\|B\|, \|M\|, \|H\|$, the number of inner iterations k and the regularization parameter α. This bound in fact does not depend on the dimensions of σ, u and g. Also, Theorem 3.1.12 includes the case $B = 0$, but the bound for τ in this case is quite far from optimal. Note that the optimal bound for τ in the case $B = 0$ can be found in Proposition 2.2.2 (for $k = 1$) and Remark 3.1.4 (for $k \geq 2$).

3.2 Numerical experiments on a toy problem

In order to compare the performance of classical gradient descent algorithm (1.7) with the k-step one-shot algorithms (1.9), we propose the following toy model related to the inverse conductivity problem in a cavity. Given $\Omega \subset \mathbb{R}^2$ an open bounded regular domain, we consider a system of M direct problems, each of which is the Helmholtz equation for the linearized scattered field $u \in H^1(\Omega)$ given by

$$\begin{cases} \operatorname{div}(\sigma_0 \nabla u) + \omega^2 u = \operatorname{div}(\sigma \nabla u_0), & \text{in } \Omega, \\ u = 0, & \text{on } \partial\Omega, \end{cases} \tag{3.31}$$

where the incident field $u_0 \in H^1(\Omega)$ satisfies

$$\begin{cases} \operatorname{div}(\sigma_0 \nabla u_0) + \omega^2 u_0 = 0, & \text{in } \Omega, \\ u_0 = f, & \text{on } \partial\Omega, \end{cases} \tag{3.32}$$

with the boundary data $f = f_i \in H^{1/2}(\partial\Omega)$, $1 \leq i \leq M$. Here σ and $\sigma_0 \in L^\infty(\Omega)$ are supposed to be positive functions such that the support $\overline{\Omega}_0$ of σ is strictly included inside Ω. Equation (3.31) is obtained by linearizing the scattered field problem using the Born approximation, and σ is the *conductivity contrast* with respect to σ_0 (see also Section 1.4.2). The variational formulation of (3.31) is: find $u \in H_0^1(\Omega)$ such that

$$\int_\Omega \sigma_0 \nabla u \cdot \nabla v \, \mathrm{d}x - \int_\Omega \omega^2 u v \, \mathrm{d}x = \int_{\Omega_0} \sigma \nabla u_0 \cdot \nabla v \, \mathrm{d}x, \quad \forall v \in H_0^1(\Omega). \tag{3.33}$$

We consider the inverse problem of retrieving σ from measurements $g := \sigma_0 \frac{\partial u}{\partial \nu}\Big|_{\partial\Omega}$ (we indeed collect M different measurements).
By discretizing u using $\mathbb{P}^1$-Lagrange finite elements on a mesh $\mathcal{T}_h(\Omega)$ of Ω, and σ by $\mathbb{P}^0$-Lagrange finite elements on a coarser mesh $\widetilde{\mathcal{T}}_{h'}(\Omega_0)$ of Ω_0, the discretization of (3.33) leads to a linear system of the form

$$A_1 \vec{u} = A_2 \vec{\sigma}, \tag{3.34}$$

where $\vec{u} \in \mathbb{R}^{n_u}$, $\vec{\sigma} \in \mathbb{R}^{n_\sigma}$ with n_u denoting the number of inner nodes of $\mathcal{T}_h(\Omega)$ and n_σ denoting the number of triangles in $\widetilde{\mathcal{T}}_{h'}(\Omega_0)$. In order to rewrite the linear system in the

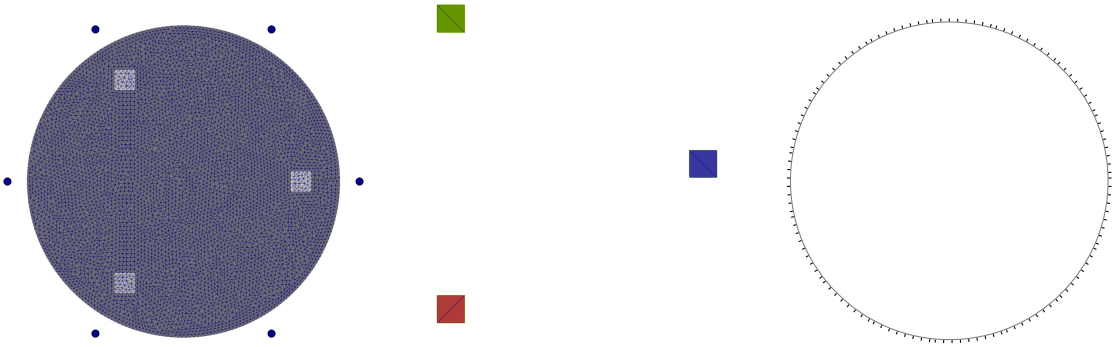
form (1.1) with a controllable norm for the matrix B, we choose to parameterize σ_0 as $\sigma_0 = \tilde{\sigma}_0 + \delta\sigma_{\rm r}$ where $\delta > 0$ is a small parameter and $\sigma_{\rm r} \leq 1$ is a random function. With this choice of σ_0, we can write the matrix A_1 as $A_1 = A_{11} + \delta A_{12}$ where A_{11} corresponds to the discretization of the bilinear form $\int_\Omega (\tilde{\sigma}_0 \nabla u \cdot \nabla v - \omega^2 uv)\,\mathrm{d}x$ on $H_0^1(\Omega) \times H_0^1(\Omega)$ and A_{12} corresponds to the discretization of the bilinear form $\int_\Omega \sigma_{\rm r} \nabla u \cdot \nabla v\,\mathrm{d}x$ on $H_0^1(\Omega) \times H_0^1(\Omega)$. For ω^2 not a Dirichlet eigenvalue of $-\operatorname{div}(\tilde{\sigma}_0 \nabla u)$ in Ω, the matrix A_{11} is invertible and we can equivalently write (3.34) as

$$\vec{u} = A_{11}^{-1}(-\delta A_{12}\vec{u} + A_2\vec{\sigma})$$

which is in the form (1.1) with $B = -\delta A_{11}^{-1}A_{12}$, $M = A_{11}^{-1}A_2$ and $F = 0$. Indeed $\|B\| < 1$ for sufficiently small δ. We employed the finite element library FreeFEM [26] to generate the matrices A_{11}, A_{12}, A_2 and the measurement operator $H \in \mathbb{R}^{n_g \times n_u}$, which is the discretization of the normal trace operator $\sigma_0 \frac{\partial u}{\partial \nu}\big|_{\partial\Omega}$, where n_g denotes the number of nodes on $\partial\Omega$.

For the numerical tests below, we set $\omega = 2\pi$, $\tilde{\sigma}_0 = 1$, $\delta = 0.01$ and the mesh size $h = \lambda/20 = 0.05$ where $\lambda = \sqrt{\tilde{\sigma}_0}2\pi/\omega = 1$. The domain Ω is the disk of radius $R = 2\lambda$ and Ω_0 is formed by three squares as in Figures 3.1 and 3.4. To generate measurements, we use $M = 6$ resulting in six different incident fields u_0 corresponding to imposing boundary data $f = f_i$, $1 \leq i \leq 6$, where $f_i(x) = Y_0(\omega|x - y_i|)$ with y_i located (outside Ω) on the circle of radius $R + 0.25\lambda$ (see Figure 3.1). The function Y_0 is the Bessel function of the second kind of order zero. The cost functional is then the sum of the cost functionals associated with each of the incident fields. We take as exact solution $\sigma^{\rm ex} = 10$ in each square, and as initial guess $\sigma^0 = 12$ in each square.

3.2.1 The case of noise-free data

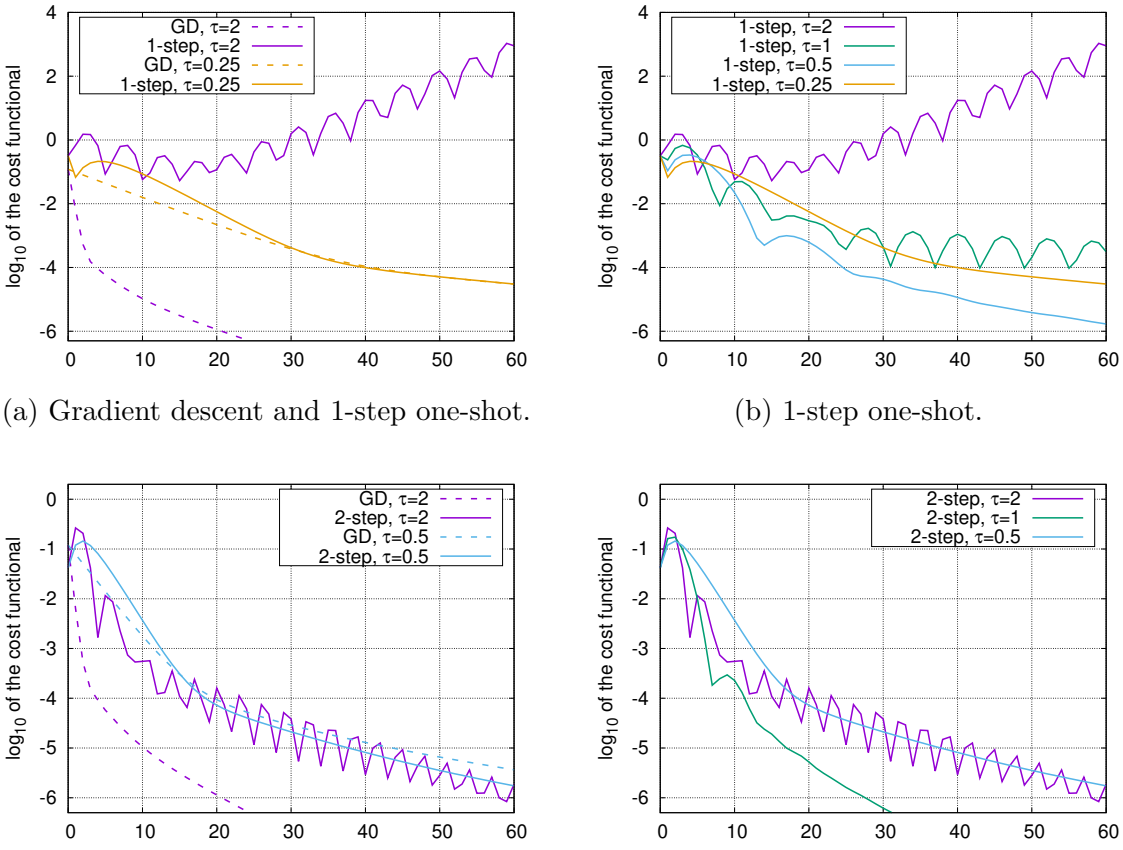


(a) Mesh for u ($n_u = 5849$). (b) Mesh for σ ($n_\sigma = 6$). (c) Mesh for g ($n_g = 125$).

Figure 3.1: The configuration for the experiment in Section 3.2.1. The small circles outside the mesh for u indicate the source locations y_i, $1 \leq i \leq 6$.

We first consider the case of noise-free data, without regularization ($\alpha = 0$). Here, the domain Ω_0 is formed by three squares of size $\lambda/4$ distributed as shown in Figure 3.1. We set the mesh size $h' = \lambda/4$ so that each square is divided into two triangles (see Figure 3.1b), which gives $n_\sigma = 6$. The mesh used for Ω (see Figure 3.1a) leads to $n_u = 5849$. The boundary mesh used for generating the data g (see Figure 3.1c) is two times coarser than the mesh for u and this gives $n_g = 125$. We compare the performances of k-step one-shot

methods (1.9) (which coincide with (1.8) in the present case $\alpha = 0$). Recall that k is the number of inner iterations on the direct and adjoint problems. We consider two series of experiments.



(a) Gradient descent and 1-step one-shot.

(b) 1-step one-shot.

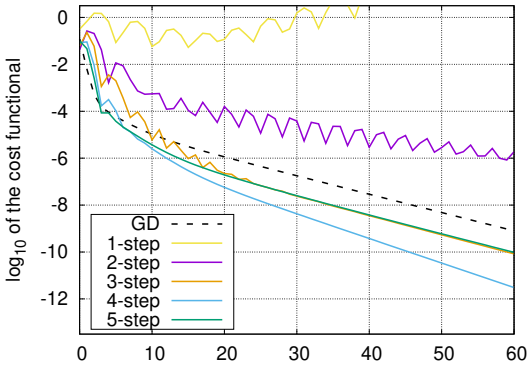
(c) Gradient descent and 2-step one-shot.

(d) 2-step one-shot.

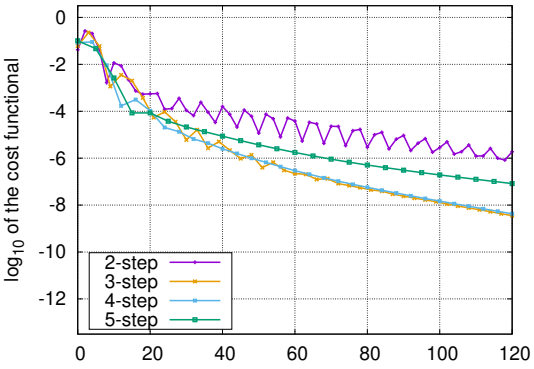
Figure 3.2: Convergence curves of gradient descent and k-step one-shot: dependence on the descent step τ.

In the first one, we study the dependence on the descent step τ. In Figure 3.2a–3.2b and 3.2c–3.2d we respectively fix $k = 1$ and $k = 2$ and compare k-step one-shot methods with the gradient descent method. We plot in semi-log scale the value of the cost functional in terms of the (outer) iteration number n in (1.6) and (1.9). We can verify that for sufficiently small τ, the k-step one-shot methods converge. This is not always the case for larger value of τ. In particular, for $\tau = 2$, while gradient descent and 2-step one-shot converge, 1-step one-shot diverges. Oscillations may appear on the convergence curve for certain values of τ, but they gradually disappear when τ gets smaller. For sufficiently small τ, the convergence curves of one-shot methods are comparable to the one of gradient descent.

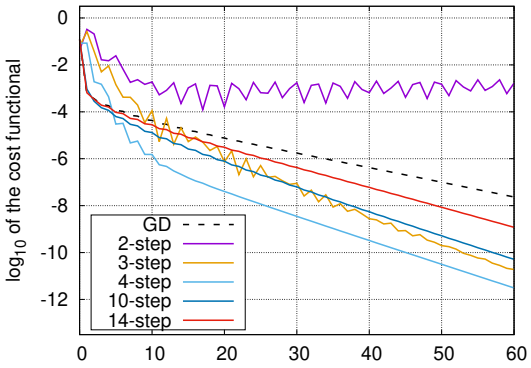
In the second series of experiments, we study the dependence on the number of inner iterations k, for fixed τ. First (Figures 3.3a and 3.3c), we investigate for which k the convergence curve of k-step one-shot is comparable with the one of gradient descent, and, as in the previous figures, on the horizontal axis we indicate the (outer) iteration number n. For $\tau = 2$ (see Figure 3.3a), we observe that for $k = 3, 4$ the convergence curves of k-step one-shot are close to the one of gradient descent. Note that with 3 inner iterations the L^2 error between u^n and the exact solution to the forward problem ranges between $59 \cdot 10^{-9}$ and 0.48334 for different n in (1.9); in fact, this error is rather significant at



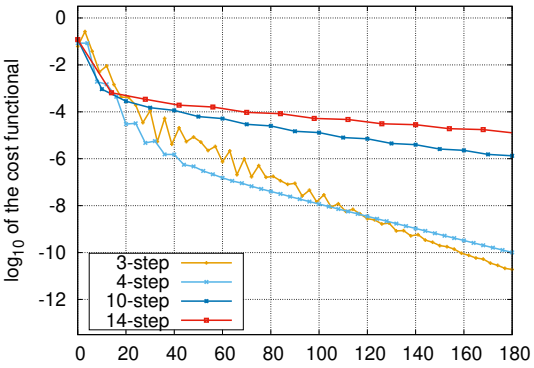
(a) Gradient descent and k-step one-shot with $\tau = 2$.



(b) k-step one-shot with $\tau = 2$ in terms of the accumulated inner iteration number.



(c) Gradient descent and k-step one-shot with $\tau = 2.5$.

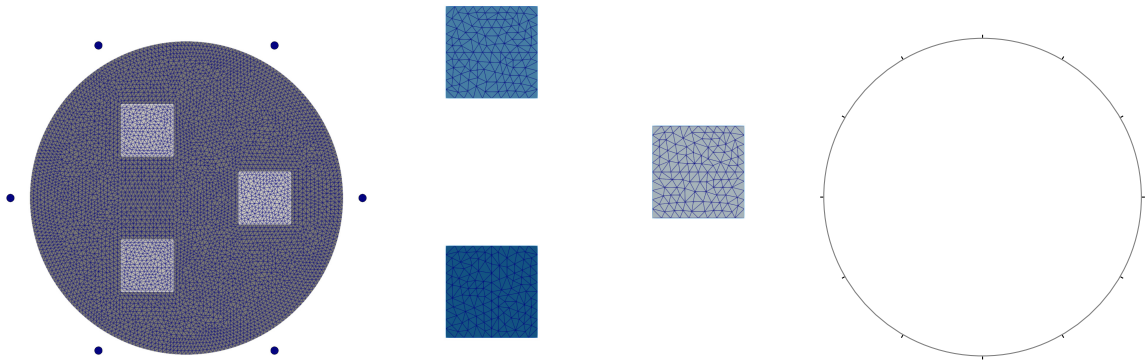


(d) k-step one-shot with $\tau = 2.5$ in terms of the accumulated inner iteration number.

Figure 3.3: Convergence curves of gradient descent and k-step one-shot: dependence on the number of inner iterations k.

the beginning of the iteration, then it reduces as we get closer to the convergence for the parameter σ. Therefore, incomplete inner iterations on the forward problem are enough to have good precision on the solution of the inverse problem. In the particular case $\tau = 2.5$ (see Figure 3.3c), we observe an interesting phenomenon: when $k = 3, 4, 10$, with k-step one-shot the cost functional decreases even faster than with gradient descent. For larger values of k, for example $k = 14$, the convergence curve of one-shot method gets closer to the one of gradient descent as one may expect. Next, in Figures 3.3b and 3.3d, we display the results of the same experiment as in Figures 3.3a and 3.3c, but this time on the horizontal axis we indicate the accumulated inner iteration number, which is equal to kn where n is the number of outer iterations. This would allow us to compare the overall speed of convergence among k-step one-shot methods. For $\tau = 2$ (respectively for $\tau = 2.5$), $k = 3$ and $k = 4$ (respectively $k = 3$) appear to provide the fastest rate of convergence. This confirms the potential interest that this type of method would have in solving large-scale inverse problem since few inner iterations are capable of providing fast convergence.

3.2.2 The case of noisy data



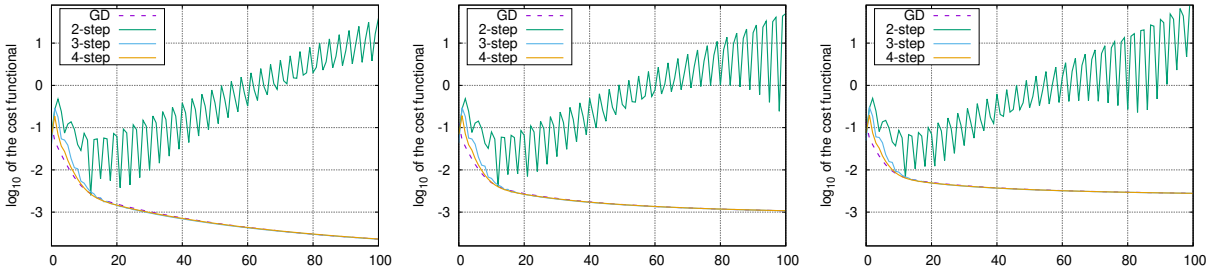
(a) Mesh for u ($n_u = 5582$). (b) Mesh for σ ($n_\sigma = 698$). (c) Mesh for g^ε ($n_g = 12$).

Figure 3.4: The configuration for the experiment in Section 3.2.2. The small circles outside the mesh for u indicate the source locations y_i, $1 \le i \le 6$.

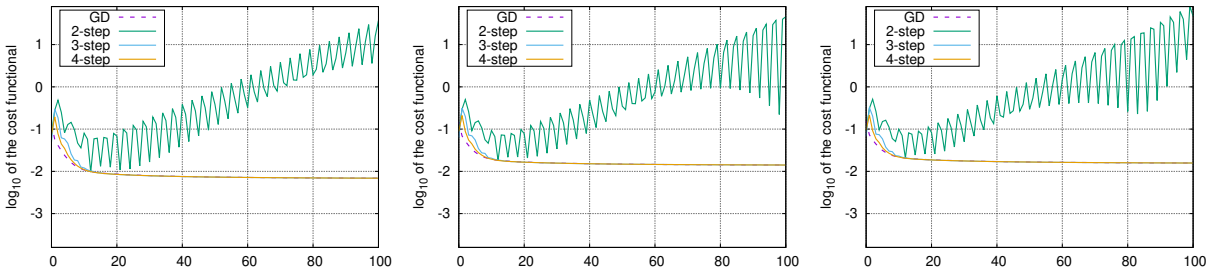
We consider now the case where the measurements g are corrupted with noise. More specifically, we replace the vector $g = Hu(\sigma^{\text{ex}})$ in the cost functional by the vector $g^\varepsilon \in \mathbb{R}^{n_g}$ with $g_i^\varepsilon := g_i + \varepsilon_i g_i$, where the ε_i are random numbers uniformly distributed between $-\varepsilon$ and ε for a noise level ε. In order to hit the ill-posedness of the inverse problem, we artificially increase the size of the discretization space for the parameter σ to $n_\sigma = 698$ by enlarging the size of the squares (now the size of their edges equals $2\lambda/3$ and the distance of their center from the boundary equals λ, see Figure 3.4b). We also show in Figure 3.4a the mesh for u, which is about 20 times finer than the boundary mesh used for generating the 6 data g^ε (see Figure 3.4c). In this new configuration, $n_u = 5582$ and $n_g = 12$. Notice also that the meshes of the squares in Figure 3.4a and 3.4b do not coincide.

We perform several numerical tests with different noise levels: $\varepsilon = 1\%, 3\%$ and 5%. We choose the regularization parameter α depending on the noise level ε in order to optimize the accuracy of the reconstruction. This choice, which is made by trial and error, does not affect much the convergence of the algorithms (see Figures 3.5 and 3.6) and mainly reduces the size of the oscillations for the reconstructed σ. The convergence curves displayed in Figure 3.5 for different noise levels show that the semi-implicit k-step one-shot methods ($k = 3, 4$) require a similar number of outer iterations as the semi-implicit gradient descent algorithm to achieve the same precision. We also see from Figure 3.5 that the convergence curves with $\alpha \neq 0$ look less steep than those with $\alpha = 0$. Since in the case of noisy data the convergence for the cost functional does not imply in general the accuracy of the reconstructed parameter σ, we also plot convergence curves for the relative error on σ in Figure 3.6 to check the quality of the reconstruction. Also in these plots we see that the semi-implicit k-step one-shot methods ($k = 3, 4$) require a similar number of outer iterations as the semi-implicit gradient descent algorithm to achieve the same accuracy. Indeed, this proves the potential of these methods since only a few inner iterations are used. Moreover, the chosen values of $\alpha \neq 0$ adapted to the noise level give fairly better relative error curves. Finally, to better show the effect of the regularization, Figures 3.7a–3.7f display the final reconstructions for σ by semi-implicit 3-step one-shot, whose relative error curves were presented in Figures 3.6a–3.6f. On the left of the color scale we also indicate the actual maximum and minimum values attained for each reconstruction; note that this range gets

wider for higher noise level. These plots confirm the benefit of regularization for treating noisy data.



(a) $\varepsilon = 1\%$, $\alpha = 0$. (b) $\varepsilon = 3\%$, $\alpha = 0$. (c) $\varepsilon = 5\%$, $\alpha = 0$.

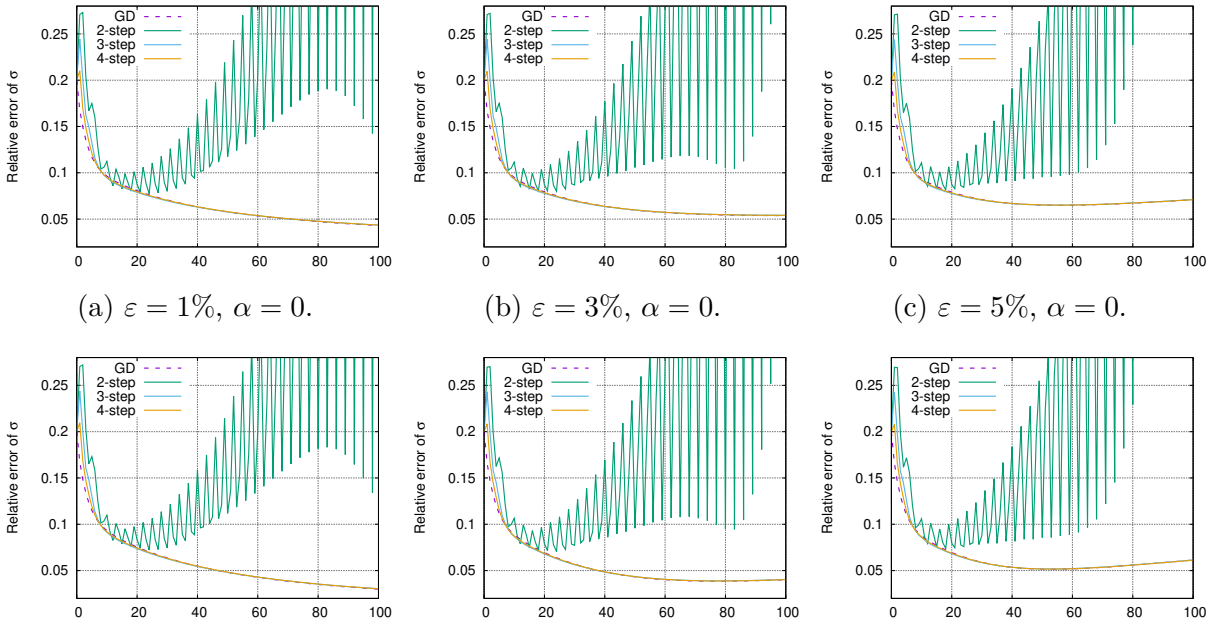


(d) $\varepsilon = 1\%$, $\alpha = 10^{-4}$. (e) $\varepsilon = 3\%$, $\alpha = 2 \cdot 10^{-4}$. (f) $\varepsilon = 5\%$, $\alpha = 2 \cdot 10^{-4}$.

Figure 3.5: Convergence curves of semi-implicit gradient descent and k-step one-shot with different noise levels ε and regularization parameters α. The descent step is $\tau = 4.7$.

3.2.3 Robustness with respect to the size of the discretized problem

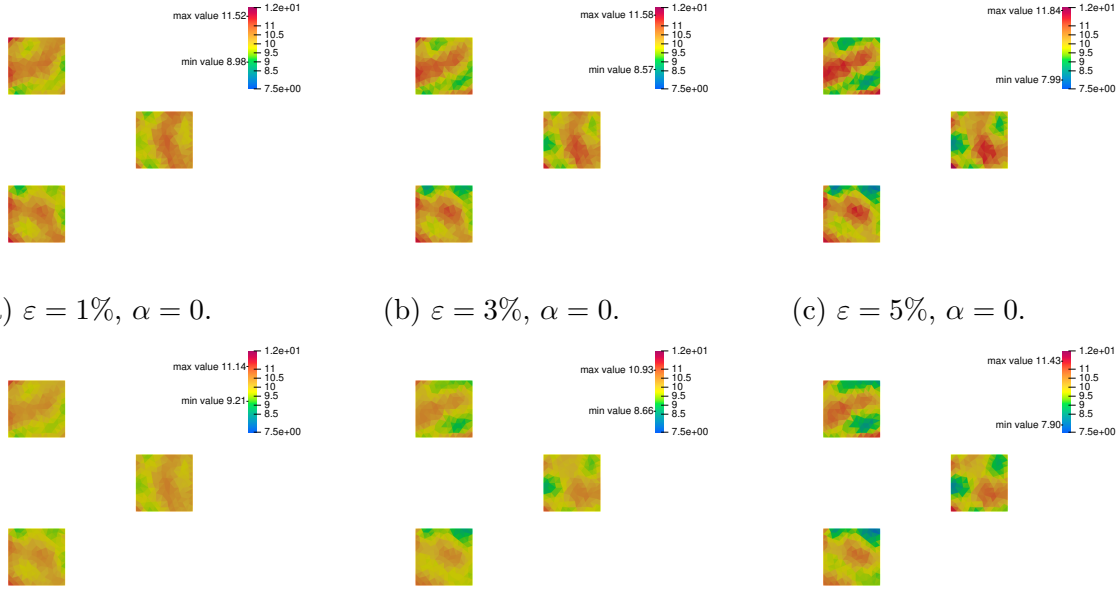
In the following experiment, we would like to confirm that the convergence rate of the k-step one-shot methods is asymptotically independent of the size of the discretized problem similarly to what is known for the gradient descent algorithms. In order to modify the number of discretization points without modifying the structure of the background media, we replace the random contribution σ_{r} with the constant function $\sigma_{\mathrm{r}} = 1$. This means that we choose $\sigma_0 = \tilde{\sigma}_0 + \delta$. (we fix $\delta = 0.01$). We keep the same configuration as in the experiment presented in Figure 3.4 in terms of number of sources, number of measurements and values of the exact conductivity and of the initial guess (i.e. $\sigma^{\mathrm{ex}} = 10$ and $\sigma^0 = 12$ in each square), but we change the configuration of the support of the unknown conductivity to investigate another test case. The exact data g_i, $1 \leq i \leq 6$ are generated using a fine mesh corresponding to a mesh size $h = \lambda/60$ and a number of degrees of freedom 157844. The data are then corrupted with $\varepsilon = 3\%$ relative random noise. For the inversion, we use two different meshes respectively corresponding to mesh sizes $h = \lambda/40$ ($n_u = 70093$) and $h = \lambda/55$ ($n_u = 128490$). This experiment is illustrated by Figure 3.8, where the mesh with $h = \lambda/40$ is displayed. We also show the mesh for the unknown σ, that results in $n_\sigma = 10$ unknowns. For the inversion, we choose to show the result for the semi-implicit 2-step one-shot method with fixed descent step $\tau = 2$ and regularization parameter $\alpha = 2 \cdot 10^{-4}$. As shown in Figure 3.9, the number of outer iterations is not much affected by the size of the discrete system. We observe the same behavior for other semi-implicit k-step one-shot methods with $k \geq 3$.



(a) $\varepsilon = 1\%$, $\alpha = 0$. (b) $\varepsilon = 3\%$, $\alpha = 0$. (c) $\varepsilon = 5\%$, $\alpha = 0$.

(d) $\varepsilon = 1\%$, $\alpha = 10^{-4}$. (e) $\varepsilon = 3\%$, $\alpha = 2 \cdot 10^{-4}$. (f) $\varepsilon = 5\%$, $\alpha = 2 \cdot 10^{-4}$.

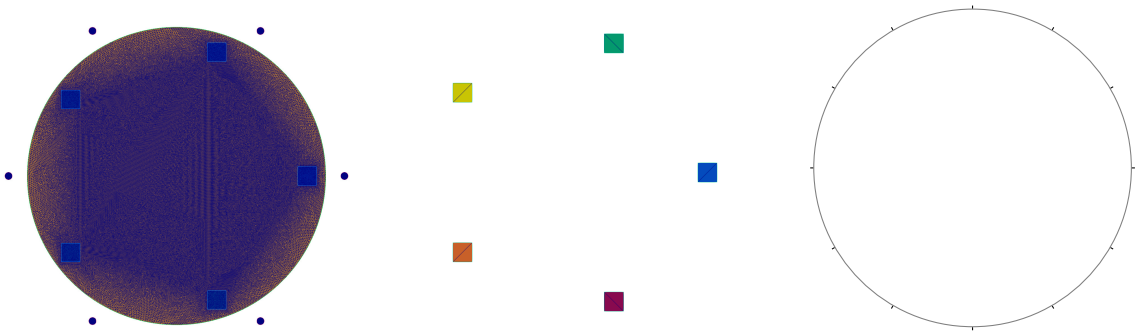
Figure 3.6: Convergence curves for the relative error on the parameter σ of semi-implicit gradient descent and k-step one-shot with different noise levels ε and regularization parameters α. The descent step is $\tau = 4.7$.



(a) $\varepsilon = 1\%$, $\alpha = 0$. (b) $\varepsilon = 3\%$, $\alpha = 0$. (c) $\varepsilon = 5\%$, $\alpha = 0$.

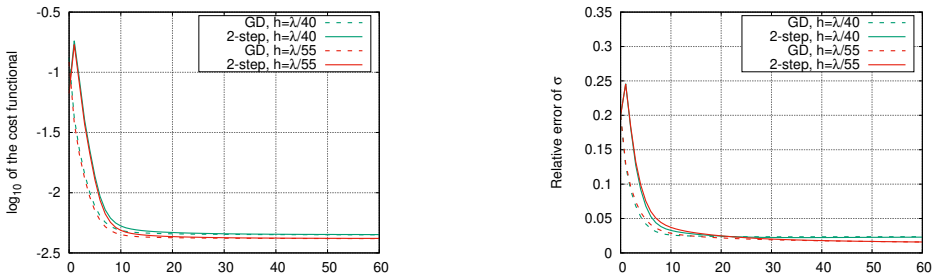
(d) $\varepsilon = 1\%$, $\alpha = 10^{-4}$. (e) $\varepsilon = 3\%$, $\alpha = 2 \cdot 10^{-4}$. (f) $\varepsilon = 5\%$, $\alpha = 2 \cdot 10^{-4}$.

Figure 3.7: Reconstructed σ by semi-implicit 3-step one-shot with different noise levels ε.



(a) Mesh for u ($h = \lambda/40$). (b) Mesh for σ ($n_\sigma = 10$). (c) Mesh for g^ε ($n_g = 12$).

Figure 3.8: The configuration for the experiment of Sections 3.2.3 and 3.2.4. The values of the exact conductivity is $\sigma^{\text{ex}} = 10$ in each square. The small circles outside the mesh for u indicate the source locations y_i, $1 \leq i \leq 6$.

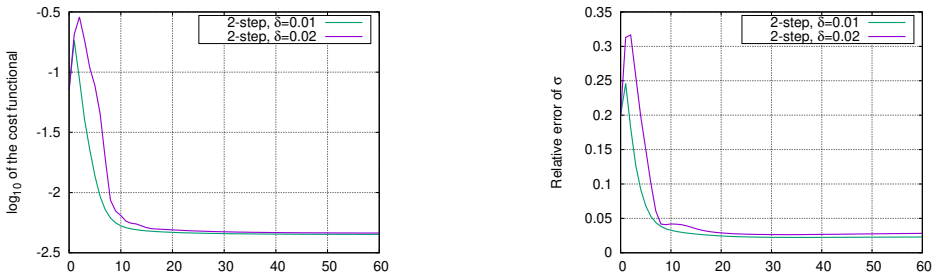


(a) $\log_{10}$ of the cost functional. (b) Relative error on σ.

Figure 3.9: Result for the experiment in Section 3.2.3 displaying the convergence rate for semi-implicit 2-step one-shot and gradient descent with different mesh sizes for the state equations. The descent step is $\tau = 2$, the noise level is $\varepsilon = 3\%$ and the regularization parameter is $\alpha = 2 \cdot 10^{-4}$.

3.2.4 Dependence of the number of outer iterations on the norm of B

The purpose of this experiment is to illustrate how the convergence speed would depend on the norm of B. Although we did not analyze the convergence rate, one indeed expects that the smaller the norm of B, the faster the convergence. This is what is observed in Figure 3.10 where we repeat the same experiment as in Figure 3.9 for $h = \lambda/40$, but we vary the values of δ ($\delta = 0.01$ and $\delta = 0.02$). For a given accuracy, the number of outer iterations for $\delta = 0.01$ is less than the one for $\delta = 0.02$.



(a) $\log_{10}$ of the cost functional. (b) Relative error on σ.

Figure 3.10: Result for the experiment in Section 3.2.4 displaying the dependence of the convergence rate for semi-implicit 2-step one-shot with respect to δ. The descent step is $\tau = 2$, the noise level is $\varepsilon = 3\%$ and the regularization parameter is $\alpha = 2 \cdot 10^{-4}$.

Appendix 3.A Some useful lemmas for the convergence analysis

We state auxiliary results about matrices like those appearing in the eigenvalue equations (2.9) and (3.13).

Lemma 3.A.1. *Let $(\mathbb{C}^{n\times n}, \|\cdot\|)$ be a normed space and $T \in \mathbb{C}^{n\times n}$. If $\rho(T) < 1$, then*

$$\sum_{k=0}^{\infty} T^k \text{ converges and } \sum_{k=0}^{\infty} T^k = (I-T)^{-1}.$$

Moreover, if $\|T\| < 1$, $\|(I-T)^{-1}\| \leq \frac{1}{1-\|T\|}$.

Lemma 3.A.2. *Let $T \in \mathbb{C}^{n\times n}$ such that $\rho(T) < 1$. Set*

$$s(T) \coloneqq \sup_{z\in\mathbb{C}, |z|\geq 1} \|(I - T/z)^{-1}\| \tag{3.35}$$

then $0 < \|(I-T)^{-1}\| \leq s(T) = s(T^*) < +\infty$. Moreover, if $\|T\| < 1$, $0 < s(T) \leq \dfrac{1}{1-\|T\|}$.

Proof. The existence of $s(T)$ (and also $s(T^*)$) is deduced from the fact that the functional $z \mapsto \|(I - T/z)^{-1}\|$, with $z \in \mathbb{C}, |z| \geq 1$, is well-defined and continuous. For every $z \in \mathbb{C}, |z| \geq 1$ we have

$$\|(I - T/z)^{-1}\| = \|\left((I - T/z)^{-1}\right)^*\| = \|(I - T^*/z^*)^{-1}\| \leq s(T^*)$$

and

$$\|(I - T^*/z)^{-1}\| = \|\left((I - T^*/z)^{-1}\right)^*\| = \|(I - T/z^*)^{-1}\| \leq s(T),$$

thus $s(T) = s(T^*)$. The second part of conclusion is obtained by Lemma 3.A.1. □

The following lemma says that, for $T \in \mathbb{C}^{n\times n}$ and $\lambda \in \mathbb{C}, |\lambda| \geq 1$, we can decompose $\left(I - \frac{T}{\lambda}\right)^{-1} = P(\lambda) + \mathrm{i}Q(\lambda)$ and $\left(I - \frac{T^*}{\lambda}\right)^{-1} = P(\lambda)^* + \mathrm{i}Q(\lambda)^*$, and gives bounds for $P(\lambda)$ and $Q(\lambda)$.

Lemma 3.A.3. *Let $T \in \mathbb{C}^{n\times n}$ such that $\rho(T) < 1$ and $\lambda \in \mathbb{C}, |\lambda| \geq 1$. Write $\frac{1}{\lambda} = r(\cos\phi + \mathrm{i}\sin\phi)$ in polar form, where $0 < r \leq 1$ and $\phi \in [-\pi, \pi]$. Then*

$$\left(I - \frac{T}{\lambda}\right)^{-1} = P(\lambda) + \mathrm{i}Q(\lambda) \quad \text{and} \quad \left(I - \frac{T^*}{\lambda}\right)^{-1} = P(\lambda)^* + \mathrm{i}Q(\lambda)^*$$

where

$$P(\lambda) = (I - r\cos\phi\, T)(I - 2r\cos\phi\, T + r^2T^2)^{-1}$$

and

$$Q(\lambda) = r\sin\phi\, T(I - 2r\cos\phi\, T + r^2T^2)^{-1}$$

are $\mathbb{C}^{n\times n}$-valued functions. We also have the following properties:

(i) $\|P(\lambda)\| \leq (1 + \|T\|)\, s(T)^2$ *and* $\|Q(\lambda)\| \leq |\sin\phi|\|T\|s(T)^2 \leq \|T\|s(T)^2$.

(ii) Moreover if $\|T\| < 1$ *then*

$$\|P(\lambda)\| \le \frac{1}{1-\|T\|} \quad \textit{and} \quad \|Q(\lambda)\| \le \frac{\|T\|}{1-\|T\|}.$$

Proof. The first part of the lemma is verified by direct computation, using

$$\begin{aligned} (I - T/\lambda)^{-1} &= (I - T/\lambda^*)\left[(I - T/\lambda)(I - T/\lambda^*)\right]^{-1}, \\ (I - T^*/\lambda)^{-1} &= \left[(I - T^*/\lambda^*)(I - T^*/\lambda)\right]^{-1}(I - T^*/\lambda^*) \\ \text{and } (I - T/\lambda)(I - T/\lambda^*) &= I - 2r\cos\phi\, T + r^2T^2. \end{aligned}$$

After that, with the help of Lemma 3.A.2, it is not difficult to show the inequalities in *(i)*. To prove *(ii)*, first observe that the two series

$$\sum_{k=0}^{\infty} r^k \cos(k\phi)T^k \quad \text{and} \quad \sum_{k=1}^{\infty} r^k \sin(k\phi)T^k$$

converge. Then, by expanding and simplifying the left-hand sides, we can show that

$$\Big(\sum_{k=0}^{\infty} r^k \cos(k\phi)T^k\Big)(I - 2r\cos\phi\, T + r^2T^2) = I - r\cos\phi\, T$$

and

$$\Big(\sum_{k=1}^{\infty} r^k \sin(k\phi)T^k\Big)(I - 2r\cos\phi\, T + r^2T^2) = r\sin\phi\, T$$

so $P(\lambda)$ and $Q(\lambda)$ can be expressed as the series above, and the inequalities in *(ii)* follow. □

In Sections 2.2 and 3.1.2 we identify different cases of $\lambda \in \mathbb{C}$ and we need corresponding estimations, given in the following lemma.

Lemma 3.A.4. *For* $\lambda \in \mathbb{C} \setminus \mathbb{R}, |\lambda| \ge 1$ *we write* $\lambda = R(\cos\theta + \mathrm{i}\sin\theta)$ *in polar form where* $R \ge 1$, $\theta \in (-\pi, \pi)$, $\theta \ne 0$.

(i) For λ *satisfying* $\mathrm{Re}(\lambda^2 - \lambda) \ge 0$, *let* $\gamma_1 = \gamma_1(\lambda) := \begin{cases} 1 & \text{if } \mathrm{Im}(\lambda^2 - \lambda) \ge 0, \\ -1 & \text{if } \mathrm{Im}(\lambda^2 - \lambda) < 0 \end{cases}$ *then*

$$\mathrm{Re}(\lambda^2 - \lambda) + \gamma_1 \mathrm{Im}(\lambda^2 - \lambda) \ge |\lambda(\lambda - 1)| \ge 2|\sin(\theta/2)|.$$

(ii) Let $0 < \theta_0 \le \frac{\pi}{4}$. *For* λ *satisfying* $\mathrm{Re}(\lambda^2 - \lambda) < 0$ *and* $\theta \in [\theta_0, \pi - \theta_0] \cup [-\pi + \theta_0, -\theta_0]$, *let* $\gamma_2 = \gamma_2(\lambda) := \begin{cases} -1 & \text{if } \mathrm{Im}(\lambda^2 - \lambda) \ge 0, \\ 1 & \text{if } \mathrm{Im}(\lambda^2 - \lambda) < 0 \end{cases}$ *then*

$$|\mathrm{Re}(\lambda^2 - \lambda) + \gamma_2 \mathrm{Im}(\lambda^2 - \lambda)| \ge |\lambda(\lambda - 1)| \ge 2\sin(\theta_0/2).$$

(iii) Let $0 < \theta_0 \le \frac{\pi}{4}$ *and* $\delta_0 > 0$. *For* λ *satisfying* $\mathrm{Re}(\lambda^2 - \lambda) < 0$ *and* $\theta \in (-\theta_0, \theta_0)\backslash\{0\}$, *let* $\gamma_3 = \gamma_3(\mathrm{sign}(\theta)) := \begin{cases} \left(\delta_0 + \sin\frac{3\theta_0}{2}\right) / \cos\frac{3\theta_0}{2} & \text{if } \theta > 0, \\ -\left(\delta_0 + \sin\frac{3\theta_0}{2}\right) / \cos\frac{3\theta_0}{2} & \text{if } \theta < 0 \end{cases}$ *then*

$$\mathrm{Re}(\lambda^2 - \lambda) + \gamma_3 \mathrm{Im}(\lambda^2 - \lambda) \ge 2\delta_0|\sin(\theta/2)|$$

and

$$\frac{|\operatorname{Re}(\lambda-1)+\gamma_3\operatorname{Im}(\lambda-1)|}{\operatorname{Re}(\lambda^2-\lambda)+\gamma_3\operatorname{Im}(\lambda^2-\lambda)} \le \frac{\sqrt{1+\gamma_3^2}}{\delta_0}.$$

Moreover, if $0<\theta_0<\frac{\pi}{4}$ *then*

$$\frac{|\gamma_3\operatorname{Re}(\lambda-1)-\operatorname{Im}(\lambda-1)|}{\operatorname{Re}(\lambda^2-\lambda)+\gamma_3\operatorname{Im}(\lambda^2-\lambda)} \le \max\left(\frac{\sqrt{1+\gamma_3^2}}{\delta_0}, \frac{\sqrt{1+\gamma_3^2}}{\cos 2\theta_0}\right).$$

(iv) Let $0<\theta_0\le\frac{\pi}{4}$. *There exists no* λ *satisfying* $\operatorname{Re}(\lambda^2-\lambda)<0$ *and* $\theta\in(\pi-\theta_0,\pi)\cup(-\pi,-\pi+\theta_0)$.

Proof. (i) Notice that $\gamma_1^2=1$, $\gamma_1\operatorname{Im}(\lambda^2-\lambda)\ge 0$. We have

$$\begin{aligned}\Big[\operatorname{Re}(\lambda^2-\lambda)+\gamma_1\operatorname{Im}(\lambda^2-\lambda)\Big]^2 &= \Big[\operatorname{Re}(\lambda^2-\lambda)\Big]^2+\Big[\operatorname{Im}(\lambda^2-\lambda)\Big]^2+2\gamma_1\operatorname{Re}(\lambda^2-\lambda)\operatorname{Im}(\lambda^2-\lambda)\\ &\ge \Big[\operatorname{Re}(\lambda^2-\lambda)\Big]^2+\Big[\operatorname{Im}(\lambda^2-\lambda)\Big]^2=|\lambda^2-\lambda|^2,\end{aligned}$$

which yields $\operatorname{Re}(\lambda^2-\lambda)+\gamma_1\operatorname{Im}(\lambda^2-\lambda)\ge|\lambda(\lambda-1)|$. Finally,

$$|\lambda-1|=|R\cos\theta-1+\mathrm{i}R\sin\theta|=\sqrt{R^2+1-2R\cos\theta}\ge\sqrt{2-2\cos\theta}=2\left|\sin\frac{\theta}{2}\right|$$

since the function $R\mapsto R^2+1-2R\cos\theta$, for $R\ge1$, is increasing.
(ii) In this case we have $\frac{\theta}{2}\in\left[\frac{\theta_0}{2},\frac{\pi}{2}-\frac{\theta_0}{2}\right]\cup\left[-\frac{\pi}{2}+\frac{\theta_0}{2},-\frac{\theta_0}{2}\right]$ so $\left|\sin\frac{\theta}{2}\right|\ge\sin\frac{\theta_0}{2}$. We also notice that $\gamma_2^2=1$ and $\gamma_2\operatorname{Im}(\lambda^2-\lambda)\le0$. Similar to *(i)*, we have $|\operatorname{Re}(\lambda^2-\lambda)+\gamma_2\operatorname{Im}(\lambda^2-\lambda)|=-\operatorname{Re}(\lambda^2-\lambda)-\gamma_2\operatorname{Im}(\lambda^2-\lambda)\ge|\lambda(\lambda-1)|\ge2|\sin(\theta/2)|$, that implies the conclusion.
(iii) Note that $\cos2\theta>0, -\frac{\pi}{2}<2\theta<\frac{\pi}{2}$, and $\sin2\theta$ has the same sign as θ and γ_3, so we have

$$\begin{aligned}\operatorname{Re}(\lambda^2-\lambda)+\gamma_3\operatorname{Im}(\lambda^2-\lambda) &= R(R\cos2\theta-\cos\theta+\gamma_3R\sin2\theta-\gamma_3\sin\theta)\\ &\ge\cos2\theta-\cos\theta+\gamma_3\sin2\theta-\gamma_3\sin\theta\\ &=-2\sin\tfrac{3\theta}{2}\sin\tfrac{\theta}{2}+2\gamma_3\cos\tfrac{3\theta}{2}\sin\tfrac{\theta}{2}\\ &=2\sin\tfrac{\theta}{2}\left(\gamma_3\cos\tfrac{3\theta}{2}-\sin\tfrac{3\theta}{2}\right).\end{aligned}$$

Then we consider two cases: if $0<\theta<\theta_0$ then $\gamma_3>0$, $\left|\sin\frac{\theta}{2}\right|=\sin\frac{\theta}{2}>0$, $0<\frac{3\theta}{2}<\frac{3\theta_0}{2}<\frac{\pi}{2}$ and $\gamma_3\cos\frac{3\theta}{2}-\sin\frac{3\theta}{2}>\gamma_3\cos\frac{3\theta_0}{2}-\sin\frac{3\theta_0}{2}=\delta_0$; if $-\theta_0<\theta<0$ then $-\gamma_3>0$, $\left|\sin\frac{\theta}{2}\right|=-\sin\frac{\theta}{2}>0$, $-\frac{\pi}{2}<-\frac{3\theta_0}{2}<\frac{3\theta}{2}<0$ and $-\gamma_3\cos\frac{3\theta}{2}+\sin\frac{3\theta}{2}>-\gamma_3\cos\frac{3\theta_0}{2}-\sin\frac{3\theta_0}{2}=\delta_0$.

Next, we will show that $\frac{|\operatorname{Re}(\lambda-1)+\gamma_3\operatorname{Im}(\lambda-1)|}{\operatorname{Re}(\lambda^2-\lambda)+\gamma_3\operatorname{Im}(\lambda^2-\lambda)}$ is bounded. First,

$$\begin{aligned}\frac{|\operatorname{Re}(\lambda-1)+\gamma_3\operatorname{Im}(\lambda-1)|}{\operatorname{Re}(\lambda^2-\lambda)+\gamma_3\operatorname{Im}(\lambda^2-\lambda)} &= \frac{|(\cos\theta+\gamma_3\sin\theta)R-1|}{R[(\cos2\theta+\gamma_3\sin2\theta)R-(\cos\theta+\gamma_3\sin\theta)]}\\ &\le \frac{|(\cos\theta+\gamma_3\sin\theta)R-1|}{(\cos2\theta+\gamma_3\sin2\theta)R-(\cos\theta+\gamma_3\sin\theta)}.\end{aligned}$$

Since γ_3 does not depend on R, let us study $f_1(R)\coloneqq\left(\frac{aR-1}{bR-a}\right)^2$ where $a\coloneqq\cos\theta+\gamma_3\sin\theta$ and $b\coloneqq\cos2\theta+\gamma_3\sin2\theta$. We observe that:

- $a > 0$ and $b > 0$. Indeed, $\cos\theta > 0$, $\cos 2\theta > 0$, and θ and γ_3 have the same sign.
- $bR - a > 0$ since $\mathrm{Re}(\lambda^2 - \lambda) + \gamma_3 \mathrm{Im}(\lambda^2 - \lambda) > 0$, thus $R > \frac{a}{b}$.
- $a^2 > b$ (equivalently $\frac{a}{b} > \frac{1}{a}$), since $a^2 = \cos^2\theta + \gamma_3^2 \sin^2\theta + \gamma_3 \sin 2\theta > \cos^2\theta - \sin^2\theta + \gamma_3 \sin 2\theta = b$.

Now, $f_1'(R) = 2 \cdot \frac{aR-1}{bR-a} \cdot \frac{b-a^2}{(bR-a)^2} < 0$ for $R > \frac{a}{b} > \frac{1}{a}$ and we would like to have $\frac{a}{b} < 1$ so that $f_1(R) \leq f_1(1), \forall R \geq 1$. Indeed $\frac{a}{b} < 1$ is equivalent to

$$\cos\theta + \gamma_3 \sin\theta < \cos 2\theta + \gamma_3 \sin 2\theta \Leftrightarrow |\gamma_3| > \frac{\left|\sin\frac{3\theta}{2}\right|}{\cos\frac{3\theta}{2}},$$

which is true since

$$|\gamma_3| = \frac{\delta_0 + \sin\frac{3\theta_0}{2}}{\cos\frac{3\theta_0}{2}} > \frac{\left|\sin\frac{3\theta}{2}\right|}{\cos\frac{3\theta}{2}} + \varepsilon_0 \quad \text{where} \quad \varepsilon_0 = \frac{\delta_0}{\cos\frac{3\theta_0}{2}}.$$

Then we study

$$f_1(1) = \left(\frac{\cos\theta - 1 + \gamma_3 \sin\theta}{\cos 2\theta - \cos\theta + \gamma_3(\sin 2\theta - \sin\theta)}\right)^2 = \left(\frac{-\sin\frac{\theta}{2} + \gamma_3 \cos\frac{\theta}{2}}{-\gamma_3 \sin\frac{3\theta}{2} + \gamma_3^2 \cos\frac{3\theta}{2}}\right)^2 \gamma_3^2.$$

We have:

- $(-\sin\frac{\theta}{2} + \gamma_3 \cos\frac{\theta}{2})^2 \leq 1 + \gamma_3^2$ by Cauchy-Schwartz inequality;
- $\gamma_3^2 = |\gamma_3|^2 > \frac{\gamma_3 \sin\frac{3\theta}{2}}{\cos\frac{3\theta}{2}} + \varepsilon_0|\gamma_3|$ that leads to $-\gamma_3 \sin\frac{3\theta}{2} + \gamma_3^2 \cos\frac{3\theta}{2} > \varepsilon_0 \cos\frac{3\theta}{2}|\gamma_3| = \delta_0|\gamma_3|$;

hence $f_1(1) \leq \frac{1+\gamma_3^2}{\delta_0^2}$ and this yields $\frac{|\mathrm{Re}(\lambda-1)+\gamma_3 \mathrm{Im}(\lambda-1)|}{\mathrm{Re}(\lambda^2-\lambda)+\gamma_3 \mathrm{Im}(\lambda^2-\lambda)} \leq \frac{\sqrt{1+\gamma_3^2}}{\delta_0}$.

Finally, we show that if $0 < \theta_0 < \frac{\pi}{4}$ then $\frac{|\gamma_3 \mathrm{Re}(\lambda-1) - \mathrm{Im}(\lambda-1)|}{\mathrm{Re}(\lambda^2-\lambda)+\gamma_3 \mathrm{Im}(\lambda^2-\lambda)}$ is bounded. Indeed, we have

$$\begin{aligned}\frac{|\gamma_3 \mathrm{Re}(\lambda-1) - \mathrm{Im}(\lambda-1)|}{\mathrm{Re}(\lambda^2-\lambda)+\gamma_3 \mathrm{Im}(\lambda^2-\lambda)} &= \frac{|(\gamma_3\cos\theta - \sin\theta)R - \gamma_3|}{R[(\cos 2\theta + \gamma_3 \sin 2\theta)R - (\cos\theta + \gamma_3 \sin\theta)]} \\ &\leq \frac{|(\gamma_3\cos\theta - \sin\theta)R - \gamma_3|}{(\cos 2\theta + \gamma_3 \sin 2\theta)R - (\cos\theta + \gamma_3 \sin\theta)}.\end{aligned}$$

Since γ_3 does not depend on R, let us study $f_2(R) \coloneqq \left(\frac{cR-\gamma_3}{bR-a}\right)^2$ where $c \coloneqq \gamma_3\cos\theta - \sin\theta$ and a, b as above. We observe that:

- $\gamma_3 b - ca$ and θ have the same sign. Indeed, $\gamma_3 b - ca = (\gamma_3^2+1)\sin\theta\cos\theta$. Consequently, we always have $(\gamma_3 b - ca)\gamma_3 > 0$.
- We always have $\frac{\gamma_3}{c} > 1$. Indeed, if $\theta > 0$ then $c > 0$ since $\gamma_3 = \frac{\delta_0 + \sin\frac{3\theta_0}{2}}{\cos\frac{3\theta_0}{2}} > \frac{\sin\theta}{\cos\theta}$, also $\frac{\gamma_3}{c} = \frac{\gamma_3}{\gamma_3\cos\theta - \sin\theta} > 1$; if $\theta < 0$ then $c < 0$ since $-\gamma_3 = \frac{\delta_0 + \sin\frac{3\theta_0}{2}}{\cos\frac{3\theta_0}{2}} > -\frac{\sin\theta}{\cos\theta}$, also $\frac{\gamma_3}{c} = \frac{-\gamma_3}{-\gamma_3\cos\theta + \sin\theta} > 1$.

Now, $f_2'(R) = 2 \cdot \frac{\frac{c}{\gamma_3}R-1}{bR-a} \cdot \frac{(\gamma_3 b - ca)\gamma_3}{(bR-a)^2}$, so, thanks to the above results, $f_2(R)$ decreases for $1 \le R < \frac{\gamma_3}{c}$ and increases for $R > \frac{\gamma_3}{c}$. Moreover, like for $f_1(1)$, we can estimate

$$f_2(1) = \left(\frac{-\cos\frac{\theta}{2} - \gamma_3 \sin\frac{\theta}{2}}{-\gamma_3 \sin\frac{3\theta}{2} + \gamma_3^2 \cos\frac{3\theta}{2}} \right)^2 \gamma_3^2 \le \frac{1+\gamma_3^2}{\delta_0^2}$$

and $\lim_{R\to+\infty} f_2(R) = \left(\frac{\gamma_3\cos\theta - \sin\theta}{\cos 2\theta + \gamma_3 \sin 2\theta}\right)^2 \le \frac{1+\gamma_3^2}{\cos 2\theta_0}$. Therefore

$$\frac{|\gamma_3 \operatorname{Re}(\lambda-1) - \operatorname{Im}(\lambda-1)|}{\operatorname{Re}(\lambda^2-\lambda) + \gamma_3 \operatorname{Im}(\lambda^2-\lambda)} \le \max\left(\frac{\sqrt{1+\gamma_3^2}}{\delta_0}, \frac{\sqrt{1+\gamma_3^2}}{\cos 2\theta_0} \right).$$

(iv) For $\theta \in (\pi-\theta_0, \pi) \cup (-\pi, -\pi+\theta_0)$, we have $\cos 2\theta > 0$ since $2\theta \in \left(\frac{3\pi}{2}, 2\pi\right) \cup \left(-2\pi, -\frac{3\pi}{2}\right)$, while $\cos\theta < 0$. Hence $\operatorname{Re}(\lambda^2-\lambda) = R(R\cos 2\theta - \cos\theta) > 0$. □

Chapter 4

Convergence analysis in some particular cases

Contents

In this chapter, we analyze the convergence of the semi-implicit multi-step one-shot methods (1.9) in two particular cases. The first one is an extension to inverse problems where the inverse parameter is real but the state and adjoint state are complex, let us call them *complex inverse problems*. The aim of this section is to show that the complex inverse problem can be rewritten so that it falls into the framework of the *real inverse problem* in Section 1.2. In other words, we ensure that our framework stated with real variables is enough for both real and complex inverse problems. The rest of the chapter is dedicated to the so-called *scalar case* (i.e. $n_u = n_\sigma = n_g = 1$) without regularization (i.e. $\alpha = 0$), which is the simplest configuration. As before, the algorithm convergence is deduced by studying the spectral radius of the associated iteration matrix. As the eigenvalue equation of a matrix is in polynomial form and not too complicated in the scalar case, we can apply Jury-Marden Criterion, which provides a necessary and sufficient condition on the coefficients of a real-coefficient polynomial so that all its zeros lie strictly inside the unit circle in the complex plane. That is how we establish the sufficient and even necessary conditions on the descent step τ that ensure the algorithm convergence. We shall study not only the algorithms introduced in Section 1.2 but also their so-called *shifted* variants. Although in this case we only deal with 3×3 matrices and third order polynomials, the obtained results about the admissible ranges for τ are not trivial. We shall plot these admissible ranges to visualize their behaviors when the number of inner iterations increases.

This chapter is structured as follows. The extension to complex inverse problems is studied in Section 4.1. First, using the Lagrangian technique for the complex equation, we define the adjoint state for the complex forward problem. By doubling the sizes of the forward and adjoint states, we obtain a transformed inverse problem with real variables.

We then show that the assumptions for the original complex inverse problem can be reformulated for the newly obtained problem to prove that the case of complex inverse problems is also covered by our theory. Next, Section 4.2 is dedicated to the convergence analysis of the scalar case without regularization parameter. We shall study four algorithms, two of them were introduced in Section 1.2 and the others are their shifted variants. After adapting notations to the scalar case, we derive the 3×3 iteration matrix associated with each algorithm. As the eigenvalue equation of a 3×3 matrix is in the form of a third order polynomial, we can apply the Jury-Marden Criterion to obtain necessary and sufficient conditions on the descent step τ that ensure the convergence of the algorithms. The admissible ranges of τ for these four algorithms are then plotted to illustrate their dependence with respect to the number of inner iterations.

The content of this chapter is extracted from the report:

> [3] M. Bonazzoli, H. Haddar, and T.A. Vu (2022). Convergence analysis of multi-step one-shot methods for linear inverse problems. Research Report RR-9477, Inria Saclay, ENSTA ParisTech. hal-03727759

4.1 Inverse problem with complex forward problem and real parameter

In this section we show that a linear inverse problem with associated complex forward problem and real parameter can be transformed into a linear inverse problem which matches with the real model at the beginning of Section 1.2, so that the previous theory applies. More precisely, here we study the state equation

$$u = Bu + M\sigma + F$$

where $u \in \mathbb{C}^{n_u}$, $\sigma \in \mathbb{R}^{n_\sigma}$, $B \in \mathbb{C}^{n_u \times n_u}, M \in \mathbb{C}^{n_u \times n_\sigma}$ and $F \in \mathbb{C}^{n_u}$. Measuring $g = Hu(\sigma)$, where $H \in \mathbb{C}^{n_g \times n_u}$, we consider the linear inverse problem of retrieving σ from the knowledge of g. To solve the inverse problem we write its regularized least squares formulation: given σ^{ex} the exact solution of the inverse problem and $g := Hu(\sigma^{\text{ex}})$ (g can also be a noisy version of $Hu(\sigma^{\text{ex}})$), we look for the regularized solution $\sigma_\alpha^{\text{ex}}$ defined by

$$\sigma_\alpha^{\text{ex}} = \text{argmin}_{\sigma \in \mathbb{R}^{n_\sigma}} J(\sigma) \quad \text{where } J(\sigma) := \frac{1}{2}\|Hu(\sigma) - g\|^2 + \frac{\alpha}{2}\|\sigma\|^2,\ \alpha \geq 0.$$

Using the Lagrangian technique with

$$\mathcal{L}(u, v, \sigma) = \frac{1}{2}\|Hu - g\|^2 + \frac{\alpha}{2}\|\sigma\|^2 + \text{Re}\langle Bu + m\sigma + F - u, v\rangle,$$

we can define the adjoint state $p = p(\sigma)$ such that

$$p = B^* p + H^*(Hu(\sigma) - g),$$

which allows us to compute

$$\nabla J(\sigma) = \text{Re}(M^* p) + \alpha\sigma.$$

By separating the real and imaginary parts of all vectors and matrices $u = u_1 + \mathrm{i}u_2$, $p = p_1 + \mathrm{i}p_2$, $B = B_1 + \mathrm{i}B_2$, $M = M_1 + \mathrm{i}M_2$, $F = F_1 + \mathrm{i}F_2$, $H = H_1 + \mathrm{i}H_2$, $g = g_1 + \mathrm{i}g_2$, we

can transform this inverse problem with complex forward problem into the inverse problem with real forward problem introduced at the beginning of Section 1.2. Indeed, note that

$$B^* = B_1^* - \mathrm{i}B_2^*, M^* = M_1^* - \mathrm{i}M_2^* \text{ and } H^* = H_1^* - \mathrm{i}H_2^*,$$

so we have

$$\begin{cases} u_1 + \mathrm{i}u_2 = (B_1 + \mathrm{i}B_2)(u_1 + \mathrm{i}u_2) + (M_1 + \mathrm{i}M_2)\sigma + (F_1 + \mathrm{i}F_2), \\ p_1 + \mathrm{i}p_2 = (B_1^* - \mathrm{i}B_2^*)(p_1 + \mathrm{i}p_2) + (H_1^* - \mathrm{i}H_2^*)[(H_1 + \mathrm{i}H_2)(u_1 + \mathrm{i}u_2) - (g_1 + \mathrm{i}g_2)], \\ \nabla J(\sigma) = \mathrm{Re}[(M_1^* - \mathrm{i}M_2^*)(p_1 + \mathrm{i}p_2)] + \alpha\sigma, \end{cases}$$

which implies

$$\begin{cases} u_1 = B_1u_1 - B_2u_2 + M_1\sigma + F_1, \\ u_2 = B_2u_1 + B_1u_2 + M_2\sigma + F_2, \\ p_1 = B_1^*p_1 + B_2^*p_2 + (H_1^*H_1 + H_2^*H_2)u_1 - (H_2^*H_1 - H_1^*H_2)u_2 - (H_1^*g_1 + H_2^*g_2), \\ p_2 = -B_2^*p_1 + B_1^*p_2 + (H_2^*H_1 - H_1^*H_2)u_1 + (H_1^*H_1 + H_2^*H_2)u_2 - (-H_2^*g_1 + H_1^*g_2), \\ \nabla J(\sigma) = M_1^*p_1 + M_2^*p_2 + \alpha\sigma. \end{cases}$$

By setting

$$\tilde{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}, \tilde{p} = \begin{bmatrix} p_1 \\ p_2 \end{bmatrix}, \tilde{B} = \begin{bmatrix} B_1 & -B_2 \\ B_2 & B_1 \end{bmatrix}, \tilde{M} = \begin{bmatrix} M_1 \\ M_2 \end{bmatrix}, \tilde{F} = \begin{bmatrix} F_1 \\ F_2 \end{bmatrix}, \tilde{H} = \begin{bmatrix} H_1 & -H_2 \\ H_2 & H_1 \end{bmatrix}, \tilde{g} = \begin{bmatrix} g_1 \\ g_2 \end{bmatrix},$$

we obtain

$$\begin{cases} \tilde{u} = \tilde{B}\tilde{u} + \tilde{M}\sigma + \tilde{F}, \\ \tilde{p} = \tilde{B}^*\tilde{p} + \tilde{H}^*(\tilde{H}\tilde{u} - \tilde{g}), \\ \nabla J(\sigma) = \tilde{M}^*\tilde{p} + \alpha\sigma, \end{cases}$$

which has the same structure as the inverse problem in Section 1.2. We finish this section by two lemmas that match the assumptions of the inverse problem with complex state variable with the assumptions of the transformed inverse problem with real state variable.

Lemma 4.1.1. $\mathrm{Spec}(\tilde{B}) = \mathrm{Spec}(B) \cup \{\overline{z} : z \in \mathrm{Spec}(B)\}$.

Proof. Let I and $\tilde{I}$ be respectively the identity matrices of dimensions n_u and $2n_u$. By writing

$$\tilde{B} = \begin{bmatrix} B_1 & -B_2 \\ B_2 & B_1 \end{bmatrix} = \underbrace{\begin{bmatrix} I & I \\ \mathrm{i}I & -\mathrm{i}I \end{bmatrix}}_{C^{-1}} \begin{bmatrix} \overline{B} & 0 \\ 0 & B \end{bmatrix} \underbrace{\begin{bmatrix} \frac{1}{2}I & -\frac{\mathrm{i}}{2}I \\ \frac{1}{2}I & \frac{i}{2}I \end{bmatrix}}_{C} \tag{4.1}$$

where

$$\overline{B} \coloneqq B_1 - \mathrm{i}B_2,$$

we find that $\det\left(\tilde{B} - \lambda\tilde{I}\right) = \det\left(\overline{B} - \lambda I\right)\det(B - \lambda I)$. The conclusion is then deduced thanks to the fact that $\mathrm{Spec}(\overline{B}) = \{\overline{z} : z \in \mathrm{Spec}(B)\}$. □

Lemma 4.1.2. *Assume that $\rho(B) < 1$, and $H(I - B)^{-1}M$ is injective. Then $\rho(\tilde{B}) < 1$, and $\tilde{H}(\tilde{I} - \tilde{B})^{-1}\tilde{M}$ is injective, where $\tilde{I} \in \mathbb{R}^{2n_u \times 2n_u}$ is the identity matrix.*

Proof. The previous lemma says that $\rho(\tilde{B}) = \rho(B) < 1$. Therefore $(\tilde{I} - \tilde{B})^{-1}$ is well-defined and thanks to (4.1),

$$\begin{aligned}(\tilde{I} - \tilde{B})^{-1} &= \underbrace{\begin{bmatrix} I & I \\ \mathrm{i}I & -\mathrm{i}I \end{bmatrix}}_{C^{-1}} \begin{bmatrix} (I-\overline{B})^{-1} & 0 \\ 0 & (I-B)^{-1} \end{bmatrix} \underbrace{\begin{bmatrix} \frac{1}{2}I & -\frac{\mathrm{i}}{2}I \\ \frac{1}{2}I & \frac{i}{2}I \end{bmatrix}}_{C} \\ &= \frac{1}{2}\begin{bmatrix} (I-\overline{B})^{-1} + (I-B)^{-1} & -\mathrm{i}(I-\overline{B})^{-1} + \mathrm{i}(I-B)^{-1} \\ \mathrm{i}(I-\overline{B})^{-1} - \mathrm{i}(I-B)^{-1} & (I-\overline{B})^{-1} + (I-B)^{-1} \end{bmatrix}.\end{aligned}$$

We thus have

$$\begin{aligned}\tilde{H}(\tilde{I} - \tilde{B})^{-1}\tilde{M} &= \frac{1}{2}\begin{bmatrix} H_1 & -H_2 \\ H_2 & H_1 \end{bmatrix} \begin{bmatrix} (I-\overline{B})^{-1} + (I-B)^{-1} & -\mathrm{i}(I-\overline{B})^{-1} + \mathrm{i}(I-B)^{-1} \\ \mathrm{i}(I-\overline{B})^{-1} - \mathrm{i}(I-B)^{-1} & (I-\overline{B})^{-1} + (I-B)^{-1} \end{bmatrix} \begin{bmatrix} M_1 \\ M_2 \end{bmatrix} \\ &= \frac{1}{2}\begin{bmatrix} \overline{H}(I-\overline{B})^{-1} + H(I-B)^{-1} & -\mathrm{i}\overline{H}(I-\overline{B})^{-1} + \mathrm{i}H(I-B)^{-1} \\ \mathrm{i}\overline{H}(I-\overline{B})^{-1} - \mathrm{i}H(I-B)^{-1} & \overline{H}(I-\overline{B})^{-1} + H(I-B)^{-1} \end{bmatrix} \begin{bmatrix} M_1 \\ M_2 \end{bmatrix} \\ &= \frac{1}{2}\begin{bmatrix} \overline{H}(I-\overline{B})^{-1}\overline{M} + H(I-B)^{-1}M \\ \mathrm{i}\overline{H}(I-\overline{B})^{-1}\overline{M} - \mathrm{i}H(I-B)^{-1}M \end{bmatrix}\end{aligned}$$

where

$$\overline{H} \coloneqq H_1 - \mathrm{i}H_2 \text{ and } \overline{M} \coloneqq M_1 - \mathrm{i}M_2.$$

Now assume that there exists $x \in \mathbb{C}^{n_\sigma}$ such that $\tilde{H}(\tilde{I} - \tilde{B})^{-1}\tilde{M}x = 0$, then

$$\begin{cases} [\overline{H}(I-\overline{B})^{-1}\overline{M} + H(I-B)^{-1}M]x = 0, \\ [\mathrm{i}\overline{H}(I-\overline{B})^{-1}\overline{M} - \mathrm{i}H(I-B)^{-1}M]x = 0, \end{cases}$$

or equivalently,

$$\begin{cases} [H(I-\overline{B})^{-1}\overline{M} + H(I-B)^{-1}M]x = 0, \\ [-\overline{H}(I-\overline{B})^{-1}\overline{M} + H(I-B)^{-1}M]x = 0. \end{cases}$$

By summing up these two equations, we deduce that $H(I-B)^{-1}Mx = 0$, then $x = 0$ thanks to the injectivity of $H(I-B)^{-1}M$. □

4.2 Convergence study for the scalar case without regularization

In this section, we shall study the convergence analysis of several algorithms in the scalar case ($n_u = n_\sigma = n_g = 1$) and without regularization (i.e. $\alpha = 0$). The considered algorithms are presented in the following.

- Usual gradient descent (usual GD):

$$\begin{cases} \sigma^{n+1} = \sigma^n - \tau M^* p^n, \\ u^n = Bu^n + M\sigma^n + F, \\ p^n = B^* p^n + H^*(Hu^n - g); \end{cases} \tag{4.2}$$

- Shifted gradient descent (shifted GD):

$$\begin{cases} \sigma^{n+1} = \sigma^n - \tau M^* p^n, \\ u^{n+1} = Bu^{n+1} + M\sigma^n + F, \\ p^{n+1} = B^* p^{n+1} + H^*(Hu^{n+1} - g); \end{cases} \tag{4.3}$$

- k-step one-shot:

$$\begin{cases} \sigma^{n+1} = \sigma^n - \tau M^* p^n, \\ u_0^{n+1} = u^n, p_0^{n+1} = p^n : \\ \text{for } \ell = 0, 1, \dots, k-1, \\ \quad \left| \begin{array}{l} u_{\ell+1}^{n+1} = B u_\ell^{n+1} + M\sigma^{n+1} + F, \\ p_{\ell+1}^{n+1} = B^* p_\ell^{n+1} + H^*(H u_\ell^{n+1} - g), \end{array} \right. \\ u^{n+1} := u_k^{n+1}, p^{n+1} := p_k^{n+1}; \end{cases} \tag{4.4}$$

- Shifted k-step one-shot:

$$\begin{cases} \sigma^{n+1} = \sigma^n - \tau M^* p^n, \\ u_0^{n+1} = u^n, p_0^{n+1} = p^n, \\ \text{for } \ell = 0, 1, \dots, k-1 : \\ \quad \left| \begin{array}{l} u_{\ell+1}^{n+1} = B u_\ell^{n+1} + M\sigma^{n} + F, \\ p_{\ell+1}^{n+1} = B^* p_\ell^{n+1} + H^*(H u_\ell^{n+1} - g), \end{array} \right. \\ u^{n+1} := u_k^{n+1}, p^{n+1} := p_k^{n+1}. \end{cases} \tag{4.5}$$

Link with the algorithms in Section 1.2. Algorithm (4.2) is algorithm (1.6) when the regularization parameter $\alpha = 0$. Also, as we mentioned in Section 1.2, the algorithm k-step one-shot (1.8) coincides with the semi-implicit k-step one-shot algorithm (1.9) when $\alpha = 0$, that is, the algorithm given by (4.4).

Two shifted algorithms (4.3) and (4.5). The shifted gradient descent (4.3) and the shifted k-step one-shot (4.5) are analyzed in our report [3]. Historically, the one-step version (i.e. $k = 1$) of (4.5), that is

$$\textbf{shifted one-step one-shot:} \quad \begin{cases} \sigma^{n+1} = \sigma^n - \tau M^* p^n, \\ u^{n+1} = B u^n + M\sigma^n + F, \\ p^{n+1} = B^* p^n + H^*(H u^n - g), \end{cases}$$

was introduced at first since the state u^{n+1} is directly updated using σ^n from the previous outer iteration (see e.g. [17]). Instead, the one-step one-shot, given by

$$\textbf{one-step one-shot:} \quad \begin{cases} \sigma^{n+1} = \sigma^n - \tau M^* p^n, \\ u^{n+1} = B u^n + M\sigma^{n+1} + F, \\ p^{n+1} = B^* p^n + H^*(H u^n - g), \end{cases}$$

has a delay in updating the state u since it employs σ^{n+1} instead of σ^n. This is not an issue in practice since updating σ is usually cheap. We also remark that when k tends to infinity, the k-step one-shot (4.4) and shifted k-step one-shot (4.5) formally converge to the usual gradient descent (4.2) and the shifted gradient descent (4.3), respectively. In practice, the usual gradient descent and k-step one-shot present better performance than their shifted versions. However, here we analyze also algorithms (4.3) and (4.5) for completeness.

4.2.1 Notations and preliminary calculation

In the scalar case, that is when $n_u = n_\sigma = n_g = 1$, we change the notation from capital to lower case letters:

$$B \leftarrow b \in \mathbb{R}, b < 1, \quad M \leftarrow m \in \mathbb{R}, m \neq 0, \quad H \leftarrow h \in \mathbb{R}, h \neq 0,$$

$$T_k \leftarrow t_k = 1 + b + ... + b^{k-1} = \frac{1-b^k}{1-b}, \quad U_k \leftarrow u_k = kh^2b^{k-1} \tag{4.6}$$

$$X_k \leftarrow x_k = \begin{cases} 0, & k = 1, \\ h^2[1 + 2b + 3b^2 + ... + (k-1)b^{k-2}], & k \geq 2. \end{cases}$$

The identity $1 + 2x + 3x^2 + ... + nx^{n-1} = \left(\frac{1-x^{n+1}}{1-x}\right)' = \frac{1-(n+1)x^n+nx^{n+1}}{(1-x)^2}$ says that

$$x_k = h^2\frac{1 - kb^{k-1} + (k-1)b^k}{(1-b)^2}, \quad k \geq 1, \tag{4.7}$$

where we set $b^{k-1} = 1$ when $k = 1$ and $b = 0$. Now for each of algorithms (4.2), (4.3), (4.4), (4.5), we write the iterations for the errors in the scalar case and the corresponding iteration matrix $\mathcal{M}$ such that $[\sigma^{n+1}, u^{n+1}, p^{n+1}]^\intercal = \mathcal{M}[\sigma^n, u^n, p^n]^\intercal$.

- Usual gradient descent:

$$\begin{cases} \sigma^{n+1} = \sigma^n - \tau m p^n, \\ u^n = bu^n + m\sigma^n, \\ p^n = bp^n + h^2u^n, \end{cases} \qquad \mathcal{M} = \begin{bmatrix} 1 & 0 & -m\tau \\ m(1-b)^{-1} & 0 & -m^2(1-b)^{-1}\tau \\ h^2m(1-b)^{-2} & 0 & -h^2m^2(1-b)^{-2}\tau \end{bmatrix}; \tag{4.8}$$

- Shifted gradient descent:

$$\begin{cases} \sigma^{n+1} = \sigma^n - \tau m p^n, \\ u^{n+1} = bu^{n+1} + m\sigma^n, \\ p^{n+1} = bp^{n+1} + h^2u^{n+1}, \end{cases} \qquad \mathcal{M} = \begin{bmatrix} 1 & 0 & -m\tau \\ m(1-b)^{-1} & 0 & 0 \\ h^2m(1-b)^{-2} & 0 & 0 \end{bmatrix}; \tag{4.9}$$

- k-step one-shot:

$$\begin{cases} \sigma^{n+1} = \sigma^n - \tau m p^n, \\ u^{n+1} = b^k u^n + mt_k\sigma^n - \tau m^2 t_k p^n, \\ p^{n+1} = (b^k - \tau m^2 x_k)p^n + u_k u^n + mx_k\sigma^n, \end{cases} \qquad \mathcal{M} = \begin{bmatrix} 1 & 0 & -m\tau \\ mt_k & b^k & -m^2t_k\tau \\ mx_k & u_k & b^k - m^2x_k\tau \end{bmatrix}; \tag{4.10}$$

- Shifted k-step one-shot:

$$\begin{cases} \sigma^{n+1} = \sigma^n - \tau m p^n, \\ u^{n+1} = b^k u^n + mt_k\sigma^n, \\ p^{n+1} = b^k p^n + u_k u^n + mx_k\sigma^n, \end{cases} \qquad \mathcal{M} = \begin{bmatrix} 1 & 0 & -m\tau \\ mt_k & b^k & 0 \\ mx_k & u_k & b^k \end{bmatrix}. \tag{4.11}$$

4.2.2 Necessary and sufficient conditions for convergence

In this simpler scalar case, we will be able to prove sufficient and also necessary conditions on the descent step τ for convergence. Our strategy to study the spectral radius $\rho(\mathcal{M})$ is as follows:

1. Compute $\det(\mathcal{M} - \lambda I)$ to write the eigenvalue equation $\mathcal{P}(\lambda) = 0$. For the considered methods, $\mathcal{P}$ turns out to be a polynomial of degree 3, $\mathcal{P}(\lambda) = a_0 + a_1\lambda + a_2\lambda^2 + \lambda^3$, where $a_0, a_1, a_2 \in \mathbb{R}$ depend on h, m, b, τ. For the computations, the identity $u_k t_k - b^k x_k + x_k = h^2 t_k^2$, which is the scalar version of (3.5), can be helpful.

2. Apply to $\mathcal{P}$ Lemma 4.2.1, which states a necessary and sufficient condition for a real coefficient polynomial of degree 3 to have all roots inside the unit circle of the complex plane. Then deduce conditions on τ.

Lemma 4.2.1. *Let $a_0, a_1, a_2 \in \mathbb{R}$, then all roots of $\mathcal{P}(z) = a_0 + a_1 z + a_2 z^2 + z^3$ stay (strictly) inside the unit circle of the complex plane if and only if*

$$(a_0 - 1)(a_0 + 1) < 0, \tag{4.12}$$
$$(a_0^2 - a_2 a_0 + a_1 - 1)(a_0^2 + a_2 a_0 - a_1 - 1) > 0, \tag{4.13}$$
$$(a_0 + a_2 - a_1 - 1)(a_0 + a_2 + a_1 + 1) < 0. \tag{4.14}$$

The proof of Lemma 4.2.1 is in Appendix 4.A and is mainly based on Marden's works [38].

Descent step for the usual gradient descent

Here, the coefficients of $\mathcal{P}$ are

$$a_0 = 0, \quad a_1 = 0, \quad a_2 = h^2 m^2 (1 - b)^{-2}\tau - 1.$$

Conditions (4.12) and (4.13) of Lemma 4.2.1 are automatically satisfied. Condition (4.14) gives

$$0 < \tau < \frac{2(1 - b)^2}{h^2 m^2}.$$

Descent step for the shifted gradient descent

Here, the coefficients of $\mathcal{P}$ are

$$a_0 = 0, \quad a_1 = h^2 m^2 (1 - b)^{-2}\tau, \quad a_2 = -1.$$

Condition (4.12) of Lemma 4.2.1 is automatically satisfied, condition (4.14) is automatically satisfied for $\tau > 0$, and condition (4.13) gives us

$$\tau < \frac{(1 - b)^2}{h^2 m^2}.$$

Descent step for k-step one-shot

Here, the coefficients of $\mathcal{P}$ are

$$a_0 = -s^2, \quad a_1 = m^2(h^2 t_k^2 - x_k)\tau + (s^2 + 2s), \quad a_2 = m^2 x_k \tau - (2s + 1)$$

where $s = b^k$. Condition (4.12) of Lemma 4.2.1 is obviously satisfied since $|b| < 1$. Next we deal with condition (4.13). The computation shows that

$$a_0^2 - a_2 a_0 + a_1 - 1 = m^2(h^2 t_k^2 - x_k + x_k s^2)\tau + \underbrace{(s-1)^3(s+1)}_{<0}, \tag{4.15}$$

$$a_0^2 + a_2 a_0 - a_1 - 1 = -m^2(h^2 t_k^2 - x_k + x_k s^2)\tau + \underbrace{(s-1)(s+1)^3}_{<0} \tag{4.16}$$

and

$$h^2 t_k^2 - x_k + x_k s^2 = \frac{h^2 b^{k-1}(1-b^k)[k-(k+1)b+kb^k-(k-1)b^{k+1}]}{(1-b)^2}. \tag{4.17}$$

Lemma 4.2.2. *$k-(k+1)b+kb^k-(k-1)b^{k+1} > 0$, $\forall |b| < 1$, $\forall k \geq 1$.*

Proof. We write $k-(k+1)b+kb^k-(k-1)b^{k+1} = (1-b)A$ where $A = k+1-\frac{1-b^k}{1-b}+(k-1)b^k$. It suffices to show $A > 0$. If $k = 1$ then $A = 1 > 0$. If either k is even, or $k \geq 3$ is odd and $0 \leq b < 1$, then $(k-1)b^k \geq 0$ and $\left|\frac{1-b^k}{1-b}\right| = |b^{k-1}+b^{k-2}+...+b+1| \leq |b^{k-1}|+|b^{k-2}|+...+|b|+1 < k$ give us the conclusion. If $k \geq 3$ is odd and $-1 < b < 0$ then $(k-1)(1+b^k+1) > 0$ and $\frac{1-b^k}{1-b} < 1$ therefore $A = 1 + \left(1 - \frac{1-b^k}{1-b}\right) + (k-1)(1+b^k) > 0$. □

Then, condition (4.13) imposes

- $\tau < \frac{(1-b)^2(1+b^k)(1-b^k)^2}{h^2 m^2 b^{k-1}[k-(k+1)b+kb^k-(k-1)b^{k+1}]}$ if $b^{k-1} > 0$;
- $\tau < \frac{(1-b)^2(1+b^k)^3}{h^2 m^2 b^{k-1}[-k+(k+1)b-kb^k+(k-1)b^{k+1}]}$ if $b^{k-1} < 0$;
- no condition on τ if $k \geq 2$ and $b = 0$.

Finally we check condition (4.14). We have $a_0 + a_2 + a_1 + 1 = h^2 m^2 t_k^2 \tau > 0$ and

$$a_0 + a_2 - a_1 - 1 = \frac{h^2 m^2 (1 - 2kb^{k-1} + 2kb^k - b^{2k})}{(1-b)^2}\tau - 2(1+s)^2,$$

therefore, condition (4.14) gives

- $\tau < \frac{2(1-b)^2(1+b^k)^2}{h^2 m^2 (1-2kb^{k-1}+2kb^k-b^{2k})}$ if $1 - 2kb^{k-1} + 2kb^k - b^{2k} > 0$;
- no condition on τ if $1 - 2kb^{k-1} + 2kb^k - b^{2k} \leq 0$.

In the following lemma we study the quantity $1 - 2kb^{k-1} + 2kb^k - b^{2k}$ that appears above.

Lemma 4.2.3. *Let $f_k(b) = 1 - 2kb^{k-1} + 2kb^k - b^{2k}$ for $k \in \mathbb{N}^*$ and $-1 \leq b \leq 1$.*

(i) $f_1(b) = -(1-b)^2 < 0, \forall -1 < b < 1$.

(ii) $f_2(b) = 1 - 4b + 4b^2 - b^4$ *has a unique solution* $b = -1+\sqrt{2}$ *in* $(-1,1)$*; and* $f_2(b) > 0$ *if* $-1 < b < -1+\sqrt{2}$, $f_2(b) < 0$ *if* $-1+\sqrt{2} < b < 1$.

(iii) If $k \geq 3$ is odd then $f_k(b)$ has exactly two solutions $b_1(k) < b_2(k)$ in $(-1,1)$; if $k \geq 2$ is even then $f_k(b)$ has a unique solution $b_3(k)$ in $(-1,1)$. Moreover, for every odd $k \geq 3$:

- $-1 < b_1(k) < 0 < b_2(k) < 1$;
- $f_k(b) > 0 \Leftrightarrow b_1(k) < b < b_2(k)$;
- $f_k(b) < 0 \Leftrightarrow -1 < b < b_1(k) \vee b_2(k) < b < 1$.

and for every even $k \geq 2$:

- $0 < b_3(k) < 1$;
- $f_k(b) > 0 \Leftrightarrow -1 < b < b_3(k)$;
- $f_k(b) < 0 \Leftrightarrow b_3(k) < b < 1$.

(iv) $\lim\limits_{\substack{k \text{ odd} \\ k\to\infty}} b_1(k) = -1$ *and* $\lim\limits_{\substack{k \text{ odd} \\ k\to\infty}} b_2(k) = 1 = \lim\limits_{\substack{k \text{ even} \\ k\to\infty}} b_3(k) = 1$.

Proof. (i) and *(ii)* are easy to verify. *(iii)* It remains to consider $k \geq 3$. We have

$$f_k'(b) = b^{k-2}\left[-2k(k-1) + 2k^2 b - 2kb^{k+1}\right], -1 < b < 1.$$

Set

$$g_k(b) = -2k(k-1) + 2k^2 b - 2kb^{k+1}, \quad -1 \leq b \leq 1, k \geq 3.$$

Case 1. [$k \geq 3$ is odd] By studying the sign of $g_k'(b)$, we find that

- g_k has a unique solution $v_1(k)$ in $(-1, 1)$ and $0 < v_1(k) < \sqrt[k]{\frac{k}{k+1}} < 1$;
- $g_k(b) > 0 \Leftrightarrow v_1(k) < b < 1$;
- $g_k(b) < 0 \Leftrightarrow -1 < b < v_1(k)$.

Next, by studying the sign of $f_k'(b)$, we find that

- $f_k(b)$ has exactly two solutions $b_1(k) < b_2(k)$ in $(-1, 1)$ and $-1 < b_1(k) < 0 < b_2(k) < 1$;
- $f_k(b) > 0 \Leftrightarrow b_1(k) < b < b_2(k)$;
- $f_k(b) < 0 \Leftrightarrow -1 < b < b_1(k) \vee b_2(k) < b < 1$.

Case 2. [$k \geq 4$ is even] By studying the sign of $g_k'(b)$, we find that

- g_k has a unique solution $v_2(k)$ in $(-1, 1)$ and $0 < v_2(k) < \sqrt[k]{\frac{k}{k+1}} < 1$;
- $g_k(b) > 0 \Leftrightarrow v_2(k) < b < 1$;
- $g_k(b) < 0 \Leftrightarrow 0 < b < v_2(k)$.

Next, by studying the sign of $f_k'(b)$, we find that

- $f_k(b)$ has a unique solution $b_3(k)$ in $(-1, 1)$ and $0 < b_3(k) < 1$;
- $f_k(b) > 0 \Leftrightarrow -1 < b < b_3(k)$;
- $f_k(b) < 0 \Leftrightarrow b_3(k) < b < 1$.

(iv) We have

$$f_k\left(\frac{1}{2}\right) = 1 - \frac{k}{2^{k-1}} - \frac{1}{2k}, \quad \forall k \geq 3 \quad \text{and} \quad f_k\left(-\frac{1}{2}\right) = 1 - \frac{3k}{2^{k-1}} - \frac{1}{2k}, \quad \forall \text{ odd } k \geq 3,$$

hence for sufficiently large k we have $f_k\left(\frac{1}{2}\right) > 0$ and for sufficiently large odd k we have $f_k\left(-\frac{1}{2}\right) > 0$. By the table of signs of f_k, we conclude that $b_1(k) < -\frac{1}{2}$ for large odd k, $b_2(k) > \frac{1}{2}$ for large odd k and $b_3(k) > \frac{1}{2}$ for large even k.
Case 1. [$k \geq 3$ is odd and sufficiently large] First we work with $b_1(k)$. We have

$$1 - 2kb_1(k)^{k-1} + 2kb_1(k)^k - b_1(k)^{2k} = 0$$

and $b_1(k) < -\frac{1}{2}$ so

$$-b_1(k)^{2k} + 2kb_1(k)^k + 1 = 2kb_1(k)^{k-1} = \underbrace{[-2kb_1(k)^k]}_{>0} \cdot \frac{1}{-b_1(k)} < [-2kb_1(k)^k].2 = -4kb_1(k)^k,$$

which leads to

$$b_1(k)^{2k} - 6kb_1(k)^k - 1 > 0 \Leftrightarrow [b_1(k)^k - 3k]^2 > 1 + 9k^2.$$

Since $-1 < b_1(k) < 0$ and k is odd, this tells us that

$$-1 < b_1(k) < -(-3k + \sqrt{1+9k^2})^{1/k} = \frac{-1}{(3k + \sqrt{1+9k^2})^{1/k}} < \frac{-1}{(7k)^{1/k}},$$

which yields $\lim\limits_{\substack{k \text{ odd} \\ k \to \infty}} b_1(k) = -1$. Next, we have

$$1 - 2kb_2(k)^{k-1} + 2kb_2(k)^k - b_2(k)^{2k} = 0$$

and $b_2(k) > \frac{1}{2}$ so

$$-b_2(k)^{2k} + 2kb_2(k)^k + 1 = 2kb_2(k)^{k-1} = 2kb_2(k)^k \cdot \frac{1}{b_2(k)} < 4kb_2(k)^k,$$

which leads to

$$b_2(k)^{2k} + 2kb_2(k)^k - 1 > 0 \Leftrightarrow [b_2(k)^k + k]^2 > 1 + k^2.$$

Since $0 < b_2(k) < 1$, this tells us that

$$1 > b_2(k) > (-k + \sqrt{1+k^2})^{1/k} = \frac{1}{(k + \sqrt{1+k^2})^{1/k}} > \frac{1}{(3k)^{1/k}},$$

which yields $\lim\limits_{\substack{k \text{ even} \\ k \to \infty}} b_2(k) = 1$.
Case 2. [$k \geq 4$ is even and sufficiently large] We repeat the same arguments as $b_2(k)$ for $b_3(k)$. □

In summary, we have the following proposition.

Proposition 4.2.4 (Convergence of k-step one-shot). *Let* $\eta_1(k,b) \coloneqq +\infty$ *and*

$$\eta_{21}(k,b) \coloneqq \frac{(1-b)^2(1+b^k)(1-b^k)^2}{b^{k-1}[k-(k+1)b+kb^k-(k-1)b^{k+1}]};$$

$$\eta_{22}(k,b) \coloneqq \frac{-(1-b)^2(1+b^k)^3}{b^{k-1}[k-(k+1)b+kb^k-(k-1)b^{k+1}]};$$

$$\eta_3(k,b) \coloneqq \frac{2(1-b)^2(1+b^k)^2}{1-2kb^{k-1}+2kb^k-b^{2k}}$$

then the necessary and sufficient condition for the convergence of k-step one-shot in the scalar case is of the form $\tau < \frac{\eta(k,b)}{h^2m^2}$ where $\eta(k,b)$ is defined as follows:

(i) $\eta(1,b) = \eta_{21}(1,b) = (1-b)^3(1+b), -1 < b < 1$;

(ii) for odd $k \geq 3$,

$$\eta(k,b) = \begin{cases} \eta_{21}(k,b), & -1 < b \leq b_1(k) \vee b_2(k) \leq b < 1, \\ \min\{\eta_{21}(k,b), \eta_3(k,b)\}, & b_1(k) < b < b_2(k) \wedge b \neq 0, \\ 2, & b = 0 \end{cases}$$

where $-1 < b_1(k) < 0 < b_2(k) < 1$ *are the two solutions of*

$$1-2kb^{k-1}+2kb^k-b^{2k} = 0, \quad -1 < b < 1;$$

(iii) for even $k \geq 2$,

$$\eta(k,b) = \begin{cases} \eta_{21}(k,b), & b_3(k) \leq b < 1, \\ \min\{\eta_{21}(k,b), \eta_3(k,b)\}, & 0 < b < b_3(k), \\ 2, & b = 0, \\ \min\{\eta_{22}(k,b), \eta_3(k,b)\}, & -1 < b < 0 \end{cases}$$

where $0 < b_3(k) < 1$ *is the unique solution of*

$$1-2kb^{k-1}+2kb^k-b^{2k} = 0, \quad -1 < b < 1.$$

Note that $\lim\limits_{\substack{k \text{ odd} \\ k\to\infty}} b_1(k) = -1$ and $\lim\limits_{\substack{k \text{ odd} \\ k\to\infty}} b_2(k) = 1 = \lim\limits_{\substack{k \text{ even} \\ k\to\infty}} b_3(k)$, so the behavior of τ when $k \to \infty$ is consistent with the result $\tau < \frac{2(1-b)^2}{h^2m^2}, -1 < b < 1$ for the usual gradient descent. For illustrations of the function $\eta(k,b)$ for different k see section 4.2.3.

Descent step for shifted k-step one-shot

Here, the coefficients of the polynomial $\mathcal{P}$ of the eigenvalue equation are

$$a_0 = h^2m^2v_k\tau - s^2, \quad a_1 = h^2m^2y_k\tau + s^2 + 2s, \quad a_2 = -2s - 1, \tag{4.18}$$

where $s = b^k$, $y_k = \frac{x_k}{h^2} = \frac{1-kb^{k-1}+(k-1)b^k}{(1-b)^2}$ and $v_k = t_k^2 - y_k = \frac{b^{k-1}[k-(k+1)b+b^{k+1}]}{(1-b)^2}$. Note that v_k and b^{k-1} have the same sign, also $v_k = 0$ if and only if $k \geq 2$ and $b = 0$, since it is easy to show that $k-(k+1)b+b^{k+1} > 0$, $\forall |b| < 1, \forall k \geq 1$. Then, condition (4.12) of Lemma 4.2.1 imposes

- $\tau < \frac{1+s^2}{h^2m^2v_k} = \frac{(1-b)^2(1+b^{2k})}{h^2m^2b^{k-1}[k-(k+1)b+b^{k+1}]}$ if $b^{k-1} > 0$;
- $\tau < \frac{-1+s^2}{h^2m^2v_k} = \frac{(1-b)^2(-1+b^{2k})}{h^2m^2b^{k-1}[k-(k+1)b+b^{k+1}]}$ if $b^{k-1} < 0$;
- no condition on τ if $k \geq 2$ and $b = 0$.

Next we study condition (4.13). We have

$$a_0^2 - a_2a_0 + a_1 - 1 = v_k^2(h^2m^2\tau)^2 + [(-2s^2+2s+1)v_k + y_k]h^2m^2\tau + \underbrace{(s-1)^3(s+1)}_{<0}$$

and

$$a_0^2 + a_2a_0 - a_1 - 1 = v_k^2(h^2m^2\tau)^2 - [(2s^2+2s+1)v_k + y_k]h^2m^2\tau + \underbrace{(s-1)(s+1)^3}_{<0},$$

each of which, considered as a second order polynomial of $h^2m^2\tau$ if $v_k \neq 0$, has exactly two roots of opposite signs. Therefore if $v_k \neq 0$, condition (4.13) is equivalent to $(h^2m^2\tau - r_1)(h^2m^2\tau - r_2) > 0$ where

$$r_1 := \frac{(2s^2-2s-1)v_k - y_k + \sqrt{(-4s+5)v_k^2 + y_k^2 + 2(-2s^2+2s+1)v_ky_k}}{2v_k^2} > 0$$

and

$$r_2 := \frac{(2s^2+2s+1)v_k + y_k + \sqrt{(8s^2+12s+5)v_k^2 + y_k^2 + 2(2s^2+2s+1)v_ky_k}}{2v_k^2} > 0.$$

Lemma 4.2.5. *r_1 and r_2 cannot be both strictly less than $\frac{1+s^2}{v_k}$. r_1 and r_2 cannot be both strictly less than $\frac{-1+s^2}{v_k}$.*

Proof. Either $r_1 < \frac{1+s^2}{v_k}$ or $r_1 < \frac{-1+s^2}{v_k}$ implies $(s^2+4s+1)v_k^2 + (s^2+1)v_ky_k > 0$. Either $r_2 < \frac{1+s^2}{v_k}$ or $r_2 < \frac{-1+s^2}{v_k}$ implies $(s^2+4s+1)v_k^2 + (s^2+1)v_ky_k < 0$. □

Thanks to this lemma we see that condition (4.13), in combination with condition (4.12), gives

- $\tau < \frac{1}{h^2m^2}\min\{r_1, r_2\}$ if $b^{k-1} \neq 0$;
- $\tau < \frac{1}{h^2m^2}$ if $k \geq 2$ and $b = 0$.

Finally, we have $a_0 + a_2 + a_1 + 1 = h^2m^2t_k^2\tau > 0$ and

$$a_0 + a_2 - a_1 - 1 = \frac{h^2m^2}{(1-b)^2}[-1 + 2kb^{k-1} - 2kb^k + b^{2k}]\tau - 2(1-b^k)^2,$$

thus condition (4.14) is equivalent to

- $\tau < \frac{2(1-b)^2(1-b^k)^2}{h^2m^2(-1+2kb^{k-1}-2kb^k+b^{2k})}$ if $1 - 2kb^{k-1} + 2kb^k - b^{2k} < 0$;
- no condition on τ if $1 - 2kb^{k-1} + 2kb^k - b^{2k} \geq 0$.

One can look again at Lemma 4.2.3 for the analysis of $1-2kb^{k-1}+2kb^k-b^{2k}$. In summary, we have the following proposition.

Proposition 4.2.6 (Convergence of shifted k-step one-shot)**.** *Let*

$$\kappa_{11}(k,b) := \frac{(1-b)^2(1+b^{2k})}{b^{k-1}[k-(k+1)b+b^{k+1}]};$$

$$\kappa_{12}(k,b) := \frac{(1-b)^2(-1+b^{2k})}{b^{k-1}[k-(k+1)b+b^{k+1}]};$$

$$t_k := \frac{1-b^k}{1-b}, \quad y_k := \frac{1-kb^{k-1}+(k-1)b^k}{(1-b)^2}, \quad s := b^k, \quad v_k := t_k^2 - y_k,$$

$$\kappa_{21}(k,b) := \frac{(2s^2-2s-1)v_k - y_k + \sqrt{(-4s+5)v_k^2+y_k^2+2(-2s^2+2s+1)v_k y_k}}{2v_k^2},$$

$$\kappa_{22}(k,b) := \frac{(2s^2+2s+1)v_k + y_k + \sqrt{(8s^2+12s+5)v_k^2+y_k^2+2(2s^2+2s+1)v_k y_k}}{2v_k^2},$$

$$\kappa_2(k,b) := \min\{\kappa_{21}(k,b), \kappa_{22}(k,b)\};$$

$$\kappa_3(k,b) := \frac{2(1-b)^2(1-b^k)^2}{-1+2kb^{k-1}-2kb^k+b^{2k}}$$

then the necessary and sufficient condition for the convergence of shifted k-step one-shot in the scalar case is of the form $\tau < \frac{\kappa(k,b)}{h^2m^2}$ where $\kappa(k,b)$ is defined as follows:

(i) $\kappa(1,b) = \min\{\kappa_{11}(1,b), \kappa_2(1,b), \kappa_3(1,b)\}$*, also note that*

$$\kappa_{11}(1,b) = 1+b^2, \quad \kappa_{21}(1,b) = \frac{2b^2-2b-1+\sqrt{-4b+5}}{2},$$

$$\kappa_{22}(1,b) = \frac{2b^2+2b+1+\sqrt{8b^2+12b+5}}{2}, \quad \kappa_3(1,b) = 2(1-b)^2;$$

(ii) for odd $k \geq 3$*,*

$$\kappa(k,b) = \begin{cases} \min\{\kappa_{11}(k,b), \kappa_2(k,b), \kappa_3(k,b)\}, & -1 < b < b_1(k) \vee b_2(k) < b < 1, \\ \min\{\kappa_{11}(k,b), \kappa_2(k,b)\}, & b_1(k) \leq b \leq b_2(k) \wedge b \neq 0, \\ 1, & b = 0 \end{cases}$$

where $-1 < b_1(k) < 0 < b_2(k) < 1$ *are the two solutions of*

$$1-2kb^{k-1}+2kb^k-b^{2k} = 0, \quad -1 < b < 1;$$

(iii) for even $k \geq 2$*,*

$$\kappa(k,b) = \begin{cases} \min\{\kappa_{11}(k,b), \kappa_2(k,b), \kappa_3(k,b)\}, & b_3(k) < b < 1, \\ \min\{\kappa_{11}(k,b), \kappa_2(k,b)\}, & 0 < b \leq b_3(k), \\ 1, & b = 0, \\ \min\{\kappa_{12}(k,b), \kappa_2(k,b)\}, & -1 < b < 0 \end{cases}$$

where $0 < b_3(k) < 1$ *is the unique solution of*

$$1-2kb^{k-1}+2kb^k-b^{2k} = 0, \quad -1 < b < 1.$$

Remark 4.2.7. In implementation, we rewrite $\kappa_{21}(k,b)$ as

$$\frac{b(1-b)^2(b^k-1)}{k-(k+1)b+b^{k+1}}+\frac{2\cdot\left[-b^k+1+\frac{b(1-b)^2(1-b^k)y_k}{k-(k+1)b+b^{k+1}}\right]}{y_k+v_k+\sqrt{(-4s+5)v_k^2+y_k^2+2(-2s^2+2s+1)v_ky_k}}$$

to avoid numerical errors. Also in this formula, we see that $\kappa_{21}(k,b)\overset{k\to\infty}{\longrightarrow}(1-b)^2$ (note that $y_k=\frac{1-kb^{k-1}+(k-1)b^k}{(1-b)^2}\overset{k\to\infty}{\longrightarrow}\frac{1}{(1-b)^2}$ and $v_k=t_k^2-y_k\overset{k\to\infty}{\longrightarrow}0$).

For illustrations of the function $\kappa(k,b)$ for different k see section 4.2.3.

4.2.3 Comparison of the bounds for the descent step

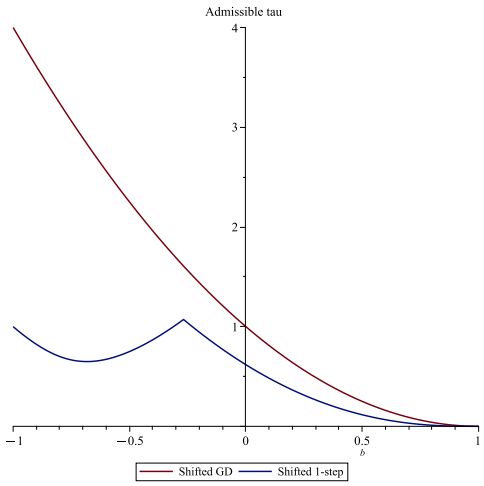
In summary, in the scalar case, the necessary and sufficient convergence conditions on the descent step $\tau>0$ are:

$$\tau<\frac{2(1-b)^2}{h^2m^2},\qquad \tau<\frac{(1-b)^2}{h^2m^2},\qquad \tau<\frac{\eta(k,b)}{h^2m^2},\qquad \tau<\frac{\kappa(k,b)}{h^2m^2},$$

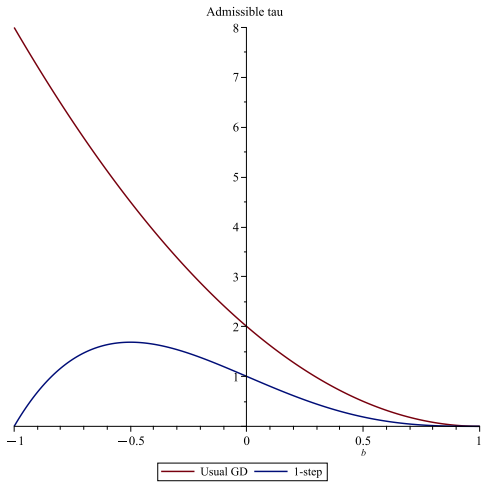
respectively for usual GD, shifted GD, k-step one-shot (with $\eta(k,b)$ given in Proposition 4.2.4), shifted k-step one-shot (with $\kappa(k,b)$ given in Proposition 4.2.6). By taking $m=h=1$, in Figure 4.1 we plot for different k the functions: $b\mapsto 2(1-b)^2$ (usual GD), $b\mapsto(1-b)^2$ (shifted GD), $b\mapsto\eta(k,b)$ (k-step one-shot) and $b\mapsto\kappa(k,b)$ (shifted k-step one-shot).

From these plots we can draw two important conclusions. First, when k increases the visualized curves for k-step one-shot and shifted k-step one-shot tend to the corresponding curves for usual and shifted gradient descent, as expected. Second, even in this scalar case, it appears difficult to establish a simplified expression for $\eta(k,b)$ in Proposition 4.2.4 and $\kappa(k,b)$ in Proposition 4.2.6 to find a practical upper bound for the descent step τ.

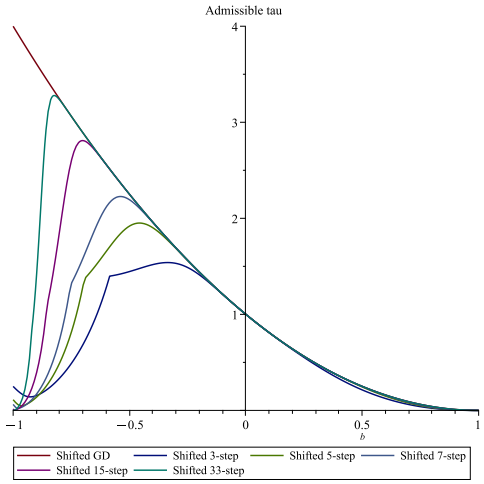
Remark 4.2.8. For $k\geq 2$, we observe that for some b the admissible range of τ of k-step one-shot is larger than the one of usual GD, that is not intuitive. This is indeed verified numerically using FreeFEM: when $b=0.2$ and $\tau=2.08$, 2-step one-shot converges while the usual GD does not.



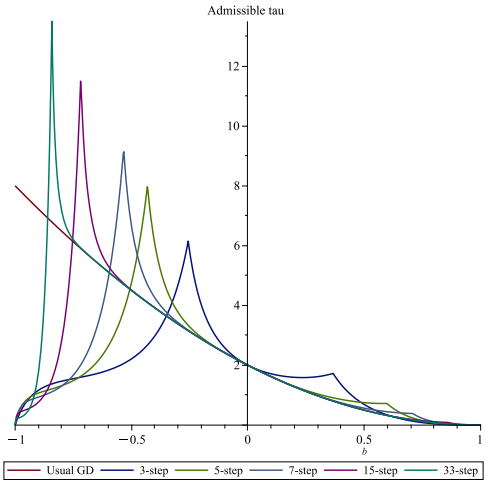
(a) Shifted 1-step one-shot



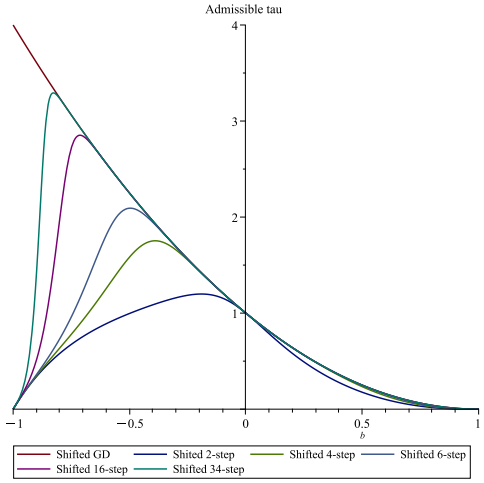
(b) 1-step one-shot



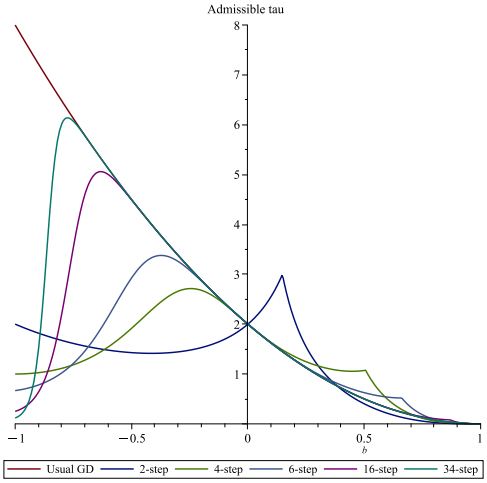
(c) Shifted k-step one-shot, odd $k \geq 3$



(d) k-step one-shot, odd $k \geq 3$



(e) Shifted k-step one-shot, even $k \geq 2$



(f) k-step one-shot, even $k \geq 2$

Figure 4.1: Admissible τ in the scalar case as a function of b.

Appendix 4.A A proof of Lemma 4.2.1 based on Marden's works

Definition 4.A.1. We say that a complex coefficient polynomial has property $\mathscr{P}$ if all its zeros lie (strictly) inside the unit circle $|z| = 1$.

We recall some definitions from Marden's works [38].

Definition 4.A.2. Let $P(z) = a_0 + a_1 z + ... + a_n z^n$ where $a_k \in \mathbb{R}, k = 0, ..., n$ (we do not require $a_n \neq 0$ here). We define

$$\tilde{P}(z) \coloneqq a_n + a_{n-1} z + ... + a_0 z^n$$

and call it the reverse polynomial of P. One can also see that $\tilde{P}(z) = z^n P(1/z)$.

Definition 4.A.3. Let $P(z) = a_0 + a_1 z + ... + a_n z^n$ where $a_k \in \mathbb{R}, k = 0, ..., n$. We define a polynomial sequence $\{P_k\}_{0 \leq k \leq n}$ where

$$P_k(z) = a_0^{(k)} + a_1^{(k)} z + ... + a_{n-k}^{(k)} z^{n-k}$$

as follows:

- $P_0 = P$;
- $P_{k+1} = a_0^{(k)} P_k - a_{n-k}^{(k)} \tilde{P}_k$ for $0 \leq k \leq n-1$.

Then we define

$$m_k(P) = a_0^{(1)} a_0^{(2)} \cdots a_0^{(k)}, \quad 1 \leq k \leq n.$$

The coefficients of these polynomials can be gathered in the following table, that we call Marden's table:

	1	z	z^2	...	z^{n-1}	z^n
P_0	a_0	a_1	a_2	...	a_{n-1}	a_n
$\tilde{P}_0$	a_n	a_{n-1}	a_{n-2}	...	a_1	a_0
P_1	$a_0^{(1)}$	$a_1^{(1)}$	$a_2^{(1)}$	...	$a_{n-1}^{(1)}$	
$\tilde{P}_1$	$a_{n-1}^{(1)}$	$a_{n-2}^{(1)}$	$a_{n-3}^{(1)}$	...	$a_0^{(1)}$	
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
P_{n-1}	$a_0^{(n-1)}$	$a_1^{(n-1)}$				
$\tilde{P}_{n-1}$	$a_1^{(n-1)}$	$a_0^{(n-1)}$				
P_n	$a_0^{(n)}$					

We have a nice and simple criterion mainly based on the works of Marden [37, 38] and Jury [29, 30], known as Jury-Marden Criterion:

Theorem 4.A.4 (Jury-Marden Criterion)**.** *The polynomial P has property $\mathscr{P}$ if and only if*

$$a_0^{(1)} < 0; \quad a_0^{(k)} > 0, \forall\, 2 \leq k \leq n.$$

This necessary and sufficient condition is mentioned several times in the literature (see e.g. [2, Theorem 3.10]), but it is not easy to find an explicit proof, so we provide a proof for the reader's convenience. Before proving this result, we apply Jury-Marden Criterion to a polynomial of degree 3 and obtain precisely Lemma 4.2.1, that is the following proposition.

Proposition 4.A.5. *Let $P(z) = a_0 + a_1 z + a_2 z^2 + z^3, z \in \mathbb{C}$ where $a_0, a_1, a_2 \in \mathbb{R}$. Then P has property $\mathscr{P}$ if and only if*

$$\begin{cases} (a_0 - 1)(a_0 + 1) < 0, \\ (a_0^2 - a_2 a_0 + a_1 - 1)(a_0^2 + a_2 a_0 - a_1 - 1) > 0, \\ (a_0 + a_2 - a_1 - 1)(a_0 + a_2 + a_1 + 1) < 0. \end{cases}$$

Proof. By directly applying Jury-Marden Criterion to P, we obtain Marden's table as follows:

	1	z	z^2	z^3
$P_0 = P$	a_0	a_1	a_2	1
$\tilde{P}_0$	1	a_2	a_1	a_0
P_1	$a_0^2 - 1$	$a_1 a_0 - a_2$	$a_2 a_0 - a_1$	
$\tilde{P}_1$	$a_2 a_0 - a_1$	$a_1 a_0 - a_2$	$a_0^2 - 1$	
P_2	$(a_0^2 - 1)^2 - (a_2 a_0 - a_1)^2$	$(a_1 a_0 - a_2)(a_0^2 - a_2 a_0 + a_1 - 1)$		
$\tilde{P}_2$	$(a_1 a_0 - a_2)(a_0^2 - a_2 a_0 + a_1 - 1)$	$(a_0^2 - 1)^2 - (a_2 a_0 - a_1)^2$		

and

$$P_3(x) = \left[(a_0^2 - 1)^2 - (a_2 a_0 - a_1)^2\right]^2 - (a_1 a_0 - a_2)^2 (a_0^2 - a_2 a_0 + a_1 - 1)^2.$$

Hence

$$\begin{aligned} a_0^{(1)} &= a_0^2 - 1 = (a_0 - 1)(a_0 + 1), \\ a_0^{(2)} &= (a_0^2 - 1)^2 - (a_2 a_0 - a_1)^2 = (a_0^2 - a_2 a_0 + a_1 - 1)(a_0^2 + a_2 a_0 - a_1 - 1), \\ a_0^{(3)} &= \left[(a_0^2 - 1)^2 - (a_2 a_0 - a_1)^2\right]^2 - (a_1 a_0 - a_2)^2 (a_0^2 - a_2 a_0 + a_1 - 1)^2 \\ &= \left[(a_0^2 + a_2 a_0 - a_1 - 1)^2 - (a_1 a_0 - a_2)^2\right] (a_0^2 - a_2 a_0 + a_1 - 1)^2 \\ &= [a_0^2 + (a_2 - a_1) a_0 + a_2 - a_1 - 1][a_0^2 + (a_2 + a_1) a_0 - a_2 - a_1 - 1] \\ &\quad (a_0^2 - 1 - a_2 a_0 + a_1)^2 \\ &= (a_0 + 1)(a_0 + a_2 - a_1 - 1)(a_0 - 1)(a_0 + a_2 + a_1 + 1)(a_0^2 - 1 - a_2 a_0 + a_1)^2. \end{aligned}$$

Then the condition $a_0^{(1)} < 0, a_0^{(2)} > 0, a_0^{(3)} > 0$, after being simplified, is equivalent to three inequalities of the statement. □

Now, to prove Jury-Marden Criterion, we need the following two results.

Theorem 4.A.6 (Marden, [38], Theorem 42.1)**.** *Let P be a real coefficient polynomial of n-th degree. If the sequence*

$$m_1(P), m_2(P), ..., m_n(P)$$

has exactly p negative elements and $n - p$ positive elements (hence no null elements), then P has p complex roots (including multiplicities) inside the unit circle $|z| = 1$, no roots on this circle and $n - p$ complex roots (including multiplicities) outside this circle.

Lemma 4.A.7 (Schur, [43])**.** *Let $P(z) = a_0 + a_1 z + ... + a_n z^n$ where $a_k \in \mathbb{R}, \forall 1 \leq k \leq n$. Assume that $|a_0| < |a_n|$. Then $\deg \tilde{P}_1 = n - 1$, and P has property $\mathscr{P}$ if and only if $\tilde{P}_1$ has property $\mathscr{P}$.*

Proof of Jury-Marden Criterion 4.A.4. The sufficient condition for P having property $\mathscr{P}$ is a direct consequence of Marden's Theorem 4.A.6. It remains to prove the necessary one.

For that, we will prove the following statement $M(n)$ by induction: "For every real-coefficient polynomial P of n-th degree having property $\mathscr{P}$, the sequence $a_0^{(1)}, ..., a_0^{(n)}$ obtained by Marden's algorithm must satisfy

$$a_0^{(1)} < 0, \quad a_0^{(k)} > 0, \forall\, 2 \leq k \leq n."$$

To check $M(1)$, let $P(z) = a_0 + a_1 z$ where $a_0, a_1 \in \mathbb{R}, a_1 \neq 0$. Then $P(z) = 0 \Leftrightarrow z = -a_0/a_1$ and $|-a_0/a_1| < 1 \Leftrightarrow |a_0| < |a_1| \Leftrightarrow a_0^{(1)} = a_0^2 - a_1^2 < 0$.

Now supposing that $M(n-1)$ is true for some $n \in \mathbb{N}, n \geq 2$, we show that $M(n)$ is true. Let $P(z) = a_0 + a_1 z + ... + a_n z^n$ where $a_k \in \mathbb{R}, k = 0, ..., n$ and $a_n \neq 0$. Assume that P has property $\mathscr{P}$. First, $a_0^{(1)} = a_0^2 - a_n^2 < 0$. Indeed, let $z_1, z_2, ..., z_n$ be the n zeros including multiplicities of P, then by Viète's formulas $z_1 z_2 \cdots z_n = (-1)^n (a_0/a_n)$. Taking the module of both sides of this identity and noting that P has property $\mathscr{P}$, we have $|a_0/a_n| < 1$, thus $a_0^{(1)} = a_0^2 - a_n^2 < 0$. Next, by Lemma 4.A.7, $\tilde{P}_1$ is of $(n-1)$-th degree and it also has property $\mathscr{P}$. Marden's table for $\tilde{P}_1$ can be easily found:

	1	z	z^2	...	z^{n-3}	z^{n-2}	z^{n-1}
$\tilde{P}_1$	$a_{n-1}^{(1)}$	$a_{n-2}^{(1)}$	$a_{n-3}^{(1)}$	...	$a_2^{(1)}$	$a_1^{(1)}$	$a_0^{(1)}$
P_1	$a_0^{(1)}$	$a_1^{(1)}$	$a_2^{(1)}$	...	$a_{n-3}^{(1)}$	$a_{n-2}^{(1)}$	$a_{n-1}^{(1)}$
$-P_2$	$-a_0^{(2)}$	$-a_1^{(2)}$	$-a_2^{(2)}$	...	$-a_{n-3}^{(2)}$	$-a_{n-2}^{(2)}$	
$-\tilde{P}_2$	$-a_{n-2}^{(2)}$	$-a_{n-3}^{(2)}$	$a_{n-4}^{(1)}$	...	$-a_1^{(2)}$	$-a_0^{(2)}$	
P_3	$a_0^{(3)}$	$a_1^{(3)}$	$a_2^{(3)}$	...	$a_{n-3}^{(3)}$		
$\tilde{P}_3$	$a_{n-3}^{(3)}$	$a_{n-4}^{(3)}$	$a_{n-5}^{(3)}$	...	$a_0^{(3)}$		
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
P_{n-1}	$a_0^{(n-1)}$	$a_1^{(n-1)}$					
$\tilde{P}_{n-1}$	$a_1^{(n-1)}$	$a_0^{(n-1)}$					
P_n	$a_0^{(n)}$						

By $M(n-1)$, we must then have $-a_0^{(2)} < 0,\ a_0^{(k)} > 0,\ \forall 3 \leq k \leq n$. □

Chapter 5

Combination of multi-step one-shot and domain decomposition methods for linear inverse problems

Contents

The goal of this chapter is to apply the formalism of multi-step one-shot inversion methods introduced in Chapter 1 to the specific case where the iterative method to solve the direct problem is based on domain decomposition. We shall specifically consider the case of nonoverlapping OSM. We consider the model problem of the inverse conductivity problem that has been outlined in the introduction (see Section 1.4). This chapter is dedicated to the linearized problem as a preparatory work for the more general non-linear case considered in the last chapter. We shall treat the case of Neumann forward problems. The case of Dirichlet problems can be treated in a similar way. To simplify the presentation of the method, we consider the case where the domain is splitted into two subdomains on which we will apply the domain decomposition method and the unknown conductivity is located in the inner subdomain. The most natural application of multi-step one-shot methods to this specific setting shows that we cannot phrase the resulting algorithm exactly as in the abstract framework in Chapter 1. In fact, there are multiple possibilities to write the iterative schemes depending on applying the one-shot methods on the continuous or discretized setting. We can also have some variants of the algorithms depending on the choice of variables of the problems: either the interface values of the impedance parameters or the state variables inside the domain. By considering the problem at the discretized level, we propose a new scheme that respects the abstract structure in Chapter 1.

This chapter is organized as follows. Section 5.1 is dedicated to the case of the linearized conductivity inverse problem. After stating the forward and inverse problems, we construct the adjoint problem to obtain a convenient expression of the derivative of the least squares cost functional. Next, we apply the nonoverlapping OSM to the forward and adjoint problems in the continuous setting. Then, we give the combined algorithms after rephrasing the OSM as an iterative method for the impedance interface values. In Section 5.2, we write the algorithms for the discretized versions of the forward and adjoint problems using finite element methods. The goal is to see whether the algorithms fit into the abstract framework of the previous chapter. In fact, we prove that this is not the case and propose an alternative scheme working directly on the discretized version. The second scheme that we shall derive this way is presented in Section 5.2.3.

5.1 The linearized forward and inverse problems

5.1.1 The forward problem

We here recall the linearized conductivity inverse problem (1.39) where the data is collected from the input of some currents on the boundary associated with the Neumann boundary conditions. Given a positive-definite $\sigma_0 \in L^\infty(\Omega)$, given $\sigma \in V_0(\Omega)$ where

$$V_0(\Omega) \coloneqq \{\sigma \in L^\infty(\Omega) : \sigma = 0 \text{ in a neighborhood of } \partial\Omega\},$$

and given some parameter $\eta > 0$, the forward problem can be stated as: Seek $u = u(\sigma) \in H^1(\Omega)$ such that

$$\begin{cases} -\operatorname{div}(\sigma_0 \nabla u) + \eta u = -\operatorname{div}(\sigma \nabla u^{\text{inc}}) & \text{in } \Omega, \\ \sigma_0 \dfrac{\partial u}{\partial \nu} = 0 & \text{on } \partial\Omega, \end{cases} \tag{5.1}$$

where ν denotes the outer normal vector on $\partial\Omega$ and the incident field $u^{\text{inc}} \in H^1(\Omega)$ satisfies

$$\begin{cases} -\operatorname{div}(\sigma_0 \nabla u^{\text{inc}}) + \eta u^{\text{inc}} = 0 & \text{in } \Omega, \\ \sigma_0 \dfrac{\partial u^{\text{inc}}}{\partial \nu} = f & \text{on } \partial\Omega \end{cases} \tag{5.2}$$

for a given $f \in H^{-\frac{1}{2}}(\partial\Omega)$. For later use, we recall that the variational formulation of problem (5.1) can be written as: $u \in H^1(\Omega)$ such that

$$\int_\Omega \sigma_0 \nabla u \cdot \nabla v \,\mathrm{dx} + \int_\Omega \eta u v \,\mathrm{dx} = \int_\Omega \sigma \nabla u^{\text{inc}} \cdot \nabla v \,\mathrm{dx}, \quad \forall v \in H^1(\Omega). \tag{5.3}$$

5.1.2 The inverse problem and classical gradient descent

For the inverse problem, we would like to retrieve some $\sigma^{\text{ex}} \in V_0(\Omega)$ from measuring the trace $g \coloneqq u(\sigma^{\text{ex}})|_{\partial\Omega}$. To do that, we introduce the least squares cost functional

$$J(\sigma) \coloneqq \frac{1}{2} \int_{\partial\Omega} (u(\sigma) - g)^2 \,\mathrm{ds}. \tag{5.4}$$

Proposition 5.1.1. *The functional J defined in (5.4) is differentiable on $V_0(\Omega)$ and its derivative can be expressed for all $\sigma \in V_0(\Omega)$ as*

$$\langle J'(\sigma), h \rangle = \int_\Omega h \nabla u^{\text{inc}} \cdot \nabla p(\sigma) \,\mathrm{dx}, \ \forall h \in V_0(\Omega) \tag{5.5}$$

where the adjoint state $p = p(\sigma) \in H^1(\Omega)$ is determined by

$$\begin{cases} -\operatorname{div}(\sigma_0 \nabla p) + \eta p = 0 & \text{in } \Omega, \\ \sigma_0 \dfrac{\partial p}{\partial \nu} = u(\sigma) - g & \text{on } \partial\Omega. \end{cases} \tag{5.6}$$

Proof. It is quite easy to prove that the mapping $\sigma \in V_0(\Omega) \mapsto u(\sigma) \in H^1(\Omega)$ is differentiable and that $w_h := \langle u'(\sigma), h\rangle$ is in $H^1(\Omega)$ and verifies

$$\begin{cases} -\operatorname{div}(\sigma_0 \nabla w_h) + \eta w_h = -\operatorname{div}(h \nabla u^{\text{inc}}) & \text{in } \Omega, \\ \sigma_0 \dfrac{\partial w_h}{\partial \nu} = 0 & \text{on } \partial\Omega. \end{cases} \tag{5.7}$$

This implies in particular that J is differentiable on $V_0(\Omega)$ and

$$\langle J'(\sigma), h\rangle = \int_{\partial\Omega} w_h (u(\sigma) - g)\, \mathrm{ds}. \tag{5.8}$$

The variational formulation of (5.7) can be written as

$$\int_\Omega \sigma_0 \nabla w_h \cdot \nabla v \,\mathrm{dx} + \int_\Omega \eta w_h v \,\mathrm{dx} = \int_\Omega h \nabla u^{\text{inc}} \cdot \nabla v \,\mathrm{dx}, \ \forall v \in H^1(\Omega) \tag{5.9}$$

and the variational formulation for $p(\sigma)$ defined in (5.6) can be written as

$$\int_\Omega \sigma_0 \nabla p(\sigma) \cdot \nabla \phi \,\mathrm{dx} + \int_\Omega \eta p(\sigma) \phi \,\mathrm{dx} = \int_{\partial\Omega} (u(\sigma) - g)\phi \,\mathrm{ds}, \ \forall \phi \in H^1(\Omega). \tag{5.10}$$

Taking $v = p(\sigma)$ in (5.9) and $\phi = w_h$ in (5.10) shows that

$$\int_\Omega h \nabla u^{\text{inc}} \cdot \nabla p(\sigma) \,\mathrm{dx} = \int_{\partial\Omega} w_h (u(\sigma) - g) \,\mathrm{ds} \overset{(5.8)}{=} \langle J'(\sigma), h\rangle,$$

which is (5.5).

□

Proposition 5.1.1 shows that the Riesz representative of $J'(\sigma)$ in $L^2(\Omega)$ formally equals $\nabla u^{\text{inc}} \cdot \nabla p(\sigma)$. Therefore, the gradient descent algorithm for σ using this representation for the gradient can be written as

$$\sigma^{n+1} = \sigma^n - \tau \nabla u^{\text{inc}} \cdot \nabla p(\sigma^n). \tag{5.11}$$

In practice, this scheme may be unstable since the $\nabla u^{\text{inc}} \cdot \nabla p(\sigma)$ can be sensitive to errors in the data g. There are several ways to stabilize the scheme. One possibility is to smooth $J'(\sigma)$, e.g. by taking the Riesz presentation in a higher Sobolev norm. The classical option is to consider the weighted H^1 representative $d(\sigma) \in H^1_0(\Omega)$ such that

$$-\delta \triangle d(\sigma) + d(\sigma) = \nabla u^{\text{inc}} \cdot \nabla p(\sigma) \ \text{ in } \Omega,$$

where $\delta > 0$ is a so-called regularization parameter. Another option that we should adopt in our later numerical simulations is to reduce the number of unknowns for σ by discretizing σ on a coarse subdivision of Ω. More specifically, we split Ω as the disjoint union of some subdomains $\Omega_i, i = 0, 1, 2..., M$ where Ω_0 is the neighborhood of Ω where $\sigma = 0$. Then we seek σ as

$$\sigma = \sum_{i=1}^{M} \sigma_i \mathbb{1}_{\Omega_i} \text{ where } \sigma_i \in \mathbb{R}, i = 1, 2, ..., M.$$

We can then view J as a function of $(\sigma_i)_{i=1}^M$ and it is quite straightforward to conclude from Proposition 5.1.1 that J is differentiable with respect to each σ_i and

$$\frac{\partial J}{\partial \sigma_i}(\sigma) = \int_{\Omega_i} \nabla u^{\text{inc}} \cdot \nabla p(\sigma)\, \mathrm{dx}.$$

In this case, the gradient descent algorithm can be written as

$$\sigma_i^{n+1} = \sigma_i^n - \tau \int_{\Omega_i} \nabla u^{\text{inc}} \cdot \nabla p(\sigma^n)\, \mathrm{dx},\ i = 1, ..., M \text{ where } \sigma^n = \sum_{i=1}^M \sigma_i^n \mathbb{1}_{\Omega_i}.$$

5.1.3 The application of OSM to the forward and adjoint problems

OSM for the forward problem

We here outline the application of the nonoverlapping OSM to the linearized forward problem (5.1). In order to ease the presentation, we consider the splitting of the domain Ω into two disjoint subdomains Ω_0 and Ω_1 where Ω_0 is the neighborhood of $\Gamma := \partial\Omega$. We further assume that the unknown conductivity contrast σ equals 0 in Ω_0 and equals σ_1 in Ω_1 (here σ_1 can be a function), i.e. $\sigma = \sigma_1 \mathbb{1}_{\Omega_1}$. We remark that the following schemes can be easily extended to the cases where the subdomain Ω_1 is also splitted into several subdomains.

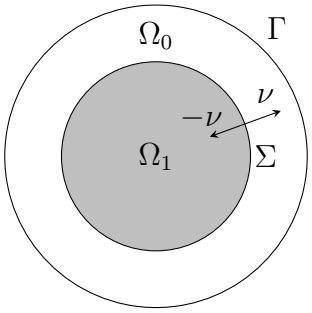


Figure 5.1: Illustration for the domain Ω and its decomposition.

We first rewrite the forward problem (5.1) as a transmission problem between the subdomains Ω_0 and Ω_1. Let $\beta \in \mathbb{C}$ be a given parameter. The choice of β is already discussed in Section 1.4: if $\eta > 0$ (the coercive case), we often take $\beta \in \mathbb{R}$, $\beta > 0$ and if $\eta < 0$ (the case of the Helmholtz equation), we often take $\beta = \mathrm{i}\tilde{\beta} \in \mathbb{C}$ where $\tilde{\beta} \in \mathbb{R}$, $\tilde{\beta} > 0$. We denote by Σ the interface between Ω_0 and Ω_1, which is also the boundary of Ω_1. We also denote by ν the normal unit vector on Γ directed toward the exterior of Ω and use the same notation for the normal vector on Σ directed toward the exterior of Ω_1 (see Figure (5.1)). For $i = 0, 1$, let us denote by $u_i \in H^1(\Omega_i)$ the restriction of the solution u to the domain Ω_i. Then problem (5.1) can be equivalently written as follows:

$$\begin{cases} -\operatorname{div}(\sigma_0 \nabla u_0) + \eta u_0 = 0 & \text{in } \Omega_0, \\ \sigma_0 \dfrac{\partial u_0}{\partial \nu} = 0 & \text{on } \Gamma, \\ \sigma_0 \dfrac{\partial u_0}{\partial \nu} - \beta u_0 = \sigma_0 \dfrac{\partial u_1}{\partial \nu} - \sigma_1 \dfrac{\partial u^{\text{inc}}}{\partial \nu} - \beta u_1 & \text{on } \Sigma, \end{cases} \tag{5.12}$$

and

$$\begin{cases} -\operatorname{div}(\sigma_0\nabla u_1)+\eta u_1 = -\operatorname{div}(\sigma_1\nabla u^{\mathrm{inc}}) & \text{in } \Omega_1,\\ \sigma_0\dfrac{\partial u_1}{\partial \nu}-\sigma_1\dfrac{\partial u^{\mathrm{inc}}}{\partial \nu}+\beta u_1 = \sigma_0\dfrac{\partial u_0}{\partial \nu}+\beta u_0 & \text{on } \Sigma. \end{cases} \tag{5.13}$$

Then, the nonoverlapping OSM can be seen as a fixed-point iteration applied to (5.12) and (5.13). More precisely, it consists in the following induction for the sequences $(u_{0;\ell})_{\ell\geq 0}$ and $(u_{1;\ell})_{\ell\geq 0}$ with a given initial guess $(u_{0;0}, u_{1;0})$:

$$\begin{cases} -\operatorname{div}(\sigma_0\nabla u_{0;\ell+1})+\eta u_{0;\ell+1} = 0 & \text{in } \Omega_0,\\ \sigma_0\dfrac{\partial u_{0;\ell+1}}{\partial \nu} = 0 & \text{on } \partial\Omega,\\ \sigma_0\dfrac{\partial u_{0;\ell+1}}{\partial \nu}-\beta u_{0;\ell+1} = \sigma_0\dfrac{\partial u_{1;\ell}}{\partial \nu}-\sigma_1\dfrac{\partial u^{\mathrm{inc}}}{\partial \nu}-\beta u_{1;\ell} & \text{on } \Sigma, \end{cases} \tag{5.14}$$

and

$$\begin{cases} -\operatorname{div}(\sigma_0\nabla u_{1;\ell+1})+\eta u_{1;\ell+1} = -\operatorname{div}(\sigma_1\nabla u^{\mathrm{inc}}) & \text{in } \Omega_1,\\ \sigma_0\dfrac{\partial u_{1;\ell+1}}{\partial \nu}-\sigma_1\dfrac{\partial u^{\mathrm{inc}}}{\partial \nu}+\beta u_{1;\ell+1} = \sigma_0\dfrac{\partial u_{0;\ell}}{\partial \nu}+\beta u_{0;\ell} & \text{on } \Sigma. \end{cases} \tag{5.15}$$

As already mentioned in Section 1.3, the numerical implementation of the transmission conditions on Σ may suffer from inaccurate evaluation of the normal derivatives. This is why it is more convenient to rewrite an induction for the impedance values

$$\lambda_{0;\ell+1} \coloneqq \sigma_0\frac{\partial u_{1;\ell}}{\partial \nu}-\sigma_1\frac{\partial u^{\mathrm{inc}}}{\partial \nu}-\beta u_{1;\ell} \quad\text{and}\quad \lambda_{1;\ell+1} \coloneqq \sigma_0\frac{\partial u_{0;\ell}}{\partial \nu}+\beta u_{0;\ell}. \tag{5.16}$$

Using these impedance variables, (5.14) and (5.15) respectively lead to

$$\begin{cases} -\operatorname{div}(\sigma_0\nabla u_{0;\ell})+\eta u_{0;\ell} = 0 & \text{in } \Omega_0,\\ \sigma_0\dfrac{\partial u_{0;\ell}}{\partial \nu} = 0 & \text{on } \partial\Omega,\\ \sigma_0\dfrac{\partial u_{0;\ell}}{\partial \nu}-\beta u_{0;\ell} = \lambda_{0;\ell} & \text{on } \Sigma, \end{cases} \tag{5.17}$$

and

$$\begin{cases} -\operatorname{div}(\sigma_0\nabla u_{1;\ell})+\eta u_{1;\ell} = -\operatorname{div}(\sigma_1\nabla u^{\mathrm{inc}}) & \text{in } \Omega_1,\\ \sigma_0\dfrac{\partial u_{1;\ell}}{\partial \nu}-\sigma_1\dfrac{\partial u^{\mathrm{inc}}}{\partial \nu}+\beta u_{1;\ell} = \lambda_{1;\ell} & \text{on } \Sigma. \end{cases} \tag{5.18}$$

One indeed has

$$\lambda_{0;\ell+1} \overset{(5.16)}{=} \left(\sigma_0\frac{\partial u_{1;\ell}}{\partial \nu}-\sigma_1\frac{\partial u^{\mathrm{inc}}}{\partial \nu}+\beta u_{1;\ell}\right)-2\beta u_{1;\ell} \overset{(5.18)}{=} \lambda_{1;\ell}-2\beta u_{1;\ell} \text{ on } \Sigma$$

and

$$\lambda_{1;\ell+1} \overset{(5.16)}{=} \left(\sigma_0\frac{\partial u_{0;\ell}}{\partial \nu}-\beta u_{0;\ell}\right)+2\beta u_{0;\ell} \overset{(5.17)}{=} \lambda_{0;\ell}+2\beta u_{0;\ell} \text{ on } \Sigma.$$

Therefore,

$$\begin{cases} \lambda_{0;\ell+1} = \lambda_{1;\ell}+2\beta u_{1;\ell} \text{ on } \Sigma,\\ \lambda_{1;\ell+1} = \lambda_{0;\ell}+2\beta u_{0;\ell} \text{ on } \Sigma. \end{cases} \tag{5.19}$$

For the sake of compacting the notation, let us introduce the following two operators:

$$\mathcal{A}_0 : \lambda \in L^2(\Sigma) \mapsto u \in H^1(\Omega_0)$$

such that

$$\begin{cases} -\operatorname{div}(\sigma_0 \nabla u) + \eta u = 0 & \text{in } \Omega_0, \\ \sigma_0 \dfrac{\partial u}{\partial \nu} = 0 & \text{on } \Gamma, \\ \sigma_0 \dfrac{\partial u}{\partial \nu} - \beta u = \lambda & \text{on } \Sigma, \end{cases} \tag{5.20}$$

and

$$\mathcal{A}_1 : (\lambda, \sigma) \in L^2(\Sigma) \times L^\infty(\Omega_1) \mapsto u \in H^1(\Omega_1)$$

such that

$$\begin{cases} -\operatorname{div}(\sigma_0 \nabla u) + \eta u = -\operatorname{div}(\sigma \nabla u^{\text{inc}}) & \text{in } \Omega_1, \\ \sigma_0 \dfrac{\partial u}{\partial \nu} - \sigma \dfrac{\partial u^{\text{inc}}}{\partial \nu} + \beta u = \lambda & \text{on } \Sigma. \end{cases} \tag{5.21}$$

In addition, for $i = 0, 1$, we introduce

$$\gamma_{\Omega_i,\Sigma} : H^1(\Omega_i) \to L^2(\Sigma) \tag{5.22}$$

the trace operator on Σ. Then, thanks to (5.19), the nonoverlapping OSM algorithm for the forward problem can be synthetically written as the following induction with a given initial guess $(\lambda_{0;0}, \lambda_{1;0})$:

$$\begin{cases} \lambda_{0;\ell+1} = \lambda_{1;\ell} - 2\beta \gamma_{\Omega_1,\Sigma} \mathcal{A}_1(\lambda_{1;\ell}, \sigma_1), \\ \lambda_{1;\ell+1} = \lambda_{0;\ell} + 2\beta \gamma_{\Omega_0,\Sigma} \mathcal{A}_0(\lambda_{0;\ell}). \end{cases} \tag{5.23}$$

We can even write (5.23) into a more compact form by setting: $\lambda_{[\ell]} = (\lambda_{0;\ell}, \lambda_{1;\ell}) \in L^2(\Sigma)^2$ as

$$\lambda_{[\ell+1]} = \mathcal{A}(\lambda_{[\ell]}, \sigma_1) \tag{5.24}$$

where $\mathcal{A} : L^2(\Sigma)^2 \times L^\infty(\Omega_1) \to L^2(\Sigma)^2$ is defined by

$$\mathcal{A}(\boldsymbol{\lambda}, \sigma) := (\boldsymbol{\lambda_1} - 2\beta \gamma_{\Omega_1,\Sigma} \mathcal{A}_1(\boldsymbol{\lambda_1}, \sigma), \boldsymbol{\lambda_0} + 2\beta \gamma_{\Omega_0,\Sigma} \mathcal{A}_0(\boldsymbol{\lambda_0})), \tag{5.25}$$

with $\boldsymbol{\lambda} = (\boldsymbol{\lambda_0}, \boldsymbol{\lambda_1}) \in L^2(\Sigma)^2$ and $\sigma \in L^\infty(\Omega_1)$.

OSM for the adjoint equation

With the same domain setting as in the previous section 5.1.3, we now consider the application of the nonoverlapping DDM to the adjoint problem (5.6) associated with the forward problem (5.1). Similarly, we shall write the OSM algorithm for the adjoint state p step by step as follows.

For $i = 0, 1$, let us denote by $p_i \in H^1(\Omega_i)$ the restriction of the solution p to the domain Ω_i. By using the same parameter β appearing in (5.12) and (5.13), we can equivalently write problem (5.6) as follows:

$$\begin{cases} -\operatorname{div}(\sigma_0 \nabla p_0) + \eta p_0 = 0 & \text{in } \Omega_0, \\ \sigma_0 \dfrac{\partial p_0}{\partial \nu} = u(\sigma_1) - g & \text{on } \Gamma, \\ \sigma_0 \dfrac{\partial p_0}{\partial \nu} - \beta p_0 = \sigma_0 \dfrac{\partial p_1}{\partial \nu} - \beta p_1 & \text{on } \Sigma \end{cases} \tag{5.26}$$

and

$$\begin{cases} -\operatorname{div}(\sigma_0 \nabla p_1) + \eta p_1 = 0 & \text{in } \Omega_1, \\ \sigma_0 \dfrac{\partial p_1}{\partial \nu} + \beta p_1 = \sigma_0 \dfrac{\partial p_0}{\partial \nu} + \beta p_0 & \text{on } \Sigma. \end{cases} \tag{5.27}$$

The nonoverlapping OSM algorithm applied to (5.26) and (5.27) is described by the following induction for the sequences $(p_{0;\ell})_{\ell\geq 0}$ and $(p_{1;\ell})_{\ell\geq 0}$ with a given initial guess $(p_{0;0}, p_{1;0})$:

$$\begin{cases} -\operatorname{div}(\sigma_0\nabla p_{0;\ell+1}) + \eta p_{0;\ell+1} = 0 & \text{in } \Omega_0, \\ \sigma_0\dfrac{\partial p_{0;\ell+1}}{\partial \nu} = u(\sigma_1) - g & \text{on } \partial\Omega, \\ \sigma_0\dfrac{\partial p_{0;\ell+1}}{\partial \nu} - \beta p_{0;\ell+1} = \sigma_0\dfrac{\partial p_{1;\ell}}{\partial \nu} - \beta p_{1;\ell} & \text{on } \Sigma, \end{cases} \tag{5.28}$$

and

$$\begin{cases} -\operatorname{div}(\sigma_0\nabla p_{1;\ell+1}) + \eta p_{1;\ell+1} = 0 & \text{in } \Omega_1, \\ \sigma_0\dfrac{\partial p_{1;\ell+1}}{\partial \nu} + \beta p_{1;\ell+1} = \sigma_0\dfrac{\partial p_{0;\ell}}{\partial \nu} + \beta p_{0;\ell} & \text{on } \Sigma. \end{cases} \tag{5.29}$$

By introducing the impedance values

$$\hat{\lambda}_{0;\ell+1} \coloneqq \sigma_0\frac{\partial p_{1;\ell}}{\partial \nu} - \beta p_{1;\ell}, \quad \text{and} \quad \hat{\lambda}_{1;\ell+1} \coloneqq \sigma_0\frac{\partial p_{0;\ell}}{\partial \nu} + \beta p_{0;\ell}, \tag{5.30}$$

we can equivalently rewrite the induction for the sequences $(p_{0;\ell})_{\ell\geq 0}$ and $(p_{1;\ell})_{\ell\geq 0}$, defined by (5.28) and (5.29), as

$$\begin{cases} -\operatorname{div}(\sigma_0\nabla p_{0;\ell}) + \eta p_{0;\ell} = 0 & \text{in } \Omega_0, \\ \sigma_0\dfrac{\partial p_{0;\ell}}{\partial \nu} = u(\sigma_1) - g & \text{on } \Gamma, \\ \sigma_0\dfrac{\partial p_{0;\ell}}{\partial \nu} - \beta p_{0;\ell} = \hat{\lambda}_{0;\ell} & \text{on } \Sigma \end{cases} \tag{5.31}$$

and

$$\begin{cases} -\operatorname{div}(\sigma_0\nabla p_{1;\ell}) + \eta p_{1;\ell} = 0 & \text{in } \Omega_1, \\ \sigma_0\dfrac{\partial p_{1;\ell}}{\partial \nu} + \beta p_{1;\ell} = \hat{\lambda}_{1;\ell} & \text{on } \Sigma. \end{cases} \tag{5.32}$$

One also has the following induction on the impedance values

$$\begin{cases} \hat{\lambda}_{0;\ell+1} = \hat{\lambda}_{1;\ell} - 2\beta p_{1;\ell} \text{ on } \Sigma, \\ \hat{\lambda}_{1;\ell+1} = \hat{\lambda}_{0;\ell} + 2\beta p_{0;\ell} \text{ on } \Sigma. \end{cases} \tag{5.33}$$

Let us introduce the following two operators:

$$\hat{\mathcal{A}}_0 : (\lambda, \psi) \in L^2(\Sigma) \times H^{-\frac{1}{2}}(\Gamma) \mapsto u \in H^1(\Omega_0)$$

such that

$$\begin{cases} -\operatorname{div}(\sigma_0\nabla u) + \eta u = 0 & \text{in } \Omega_0, \\ \sigma_0\dfrac{\partial u}{\partial \nu} = \psi & \text{on } \Gamma, \\ \sigma_0\dfrac{\partial u}{\partial \nu} - \beta u = \lambda & \text{on } \Sigma \end{cases} \tag{5.34}$$

and

$$\hat{\mathcal{A}}_1 : \lambda \in L^2(\Sigma) \mapsto u \in H^1(\Omega_1)$$

such that

$$\begin{cases} -\operatorname{div}(\sigma_0\nabla u) + \eta u = 0 & \text{in } \Omega_1, \\ \sigma_0\dfrac{\partial u}{\partial \nu} + \beta u = \lambda & \text{on } \Sigma. \end{cases} \tag{5.35}$$

We also introduce

$$\gamma_\Gamma : H^1(\Omega_0) \to L^2(\Gamma) \tag{5.36}$$

the trace operator on Γ. Then, thanks to (5.33), the nonoverlapping OSM algorithm for the adjoint problem (5.6) can be synthetically written as the following induction with a given initial guess $(\hat{\lambda}_{0;0}, \hat{\lambda}_{1;0})$:

$$\begin{cases} \hat{\lambda}_{0;\ell+1} = \hat{\lambda}_{1;\ell} - 2\beta\gamma_{\Omega_1,\Sigma}\hat{\mathcal{A}}_1(\hat{\lambda}_{1;\ell}), \\ \hat{\lambda}_{1;\ell+1} = \hat{\lambda}_{0;\ell} + 2\beta\gamma_{\Omega_0,\Sigma}\hat{\mathcal{A}}_0(\hat{\lambda}_{0;\ell}, \gamma_\Gamma u(\sigma_1) - g). \end{cases} \tag{5.37}$$

The trace operators $\gamma_{\Omega_i,\Sigma}$, $i = 0, 1$ are already defined in (5.22). We can even write (5.37) into a more compact form by setting: $\hat{\lambda}_{[\ell]} = (\hat{\lambda}_{0;\ell}, \hat{\lambda}_{1;\ell}) \in L^2(\Sigma)^2$ as

$$\hat{\lambda}_{[\ell+1]} = \hat{\mathcal{A}}(\hat{\lambda}_{[\ell]}, \gamma_\Gamma u(\sigma_1) - g) \tag{5.38}$$

where $\hat{\mathcal{A}} : L^2(\Sigma)^2 \times H^{-\frac{1}{2}}(\Gamma) \to L^2(\Sigma)^2$ is defined by

$$\hat{\mathcal{A}}(\boldsymbol{\lambda}, \psi) := (\boldsymbol{\lambda_1} - 2\beta\gamma_{\Omega_1,\Sigma}\hat{\mathcal{A}}_1(\boldsymbol{\lambda_1}), \boldsymbol{\lambda_0} + 2\beta\gamma_{\Omega_0,\Sigma}\hat{\mathcal{A}}_0(\boldsymbol{\lambda_0}, \psi)), \tag{5.39}$$

with $\boldsymbol{\lambda} = (\boldsymbol{\lambda_0}, \boldsymbol{\lambda_1}) \in L^2(\Sigma)^2$ and $\psi \in H^{-\frac{1}{2}}(\Gamma)$.

OSM for the simultaneous calculation of the forward and adjoint states

In the nonoverlapping OSM algorithm applied to the adjoint equation (5.37), one assumes that the state $u(\sigma_1)$ is already known. In practice, it is indeed more natural to perform the iterations in (5.37) in parallel to the OSM iterations for the forward problem (5.23). The combination of (5.23) and (5.37) leads to the following scheme that simultaneously computes the forward and adjoint states: given $\sigma_1 \in L^\infty(\Omega_1)$ and initial guess $(\lambda_{0;0}, \lambda_{1;0}, \hat{\lambda}_{0;0}, \hat{\lambda}_{1;0}) \in L^2(\Sigma)^4$,

$$\begin{cases} \lambda_{0;\ell+1} = \lambda_{1;\ell} - 2\beta\gamma_{\Omega_1,\Sigma}\mathcal{A}_1(\lambda_{1;\ell}, \sigma_1), \\ \lambda_{1;\ell+1} = \lambda_{0;\ell} + 2\beta\gamma_{\Omega_0,\Sigma}\mathcal{A}_0(\lambda_{0;\ell}), \\ \hat{\lambda}_{0;\ell+1} = \hat{\lambda}_{1;\ell} - 2\beta\gamma_{\Omega_1,\Sigma}\hat{\mathcal{A}}_1(\hat{\lambda}_{1;\ell}), \\ \hat{\lambda}_{1;\ell+1} = \hat{\lambda}_{0;\ell} + 2\beta\gamma_{\Omega_0,\Sigma}\hat{\mathcal{A}}_0(\hat{\lambda}_{0;\ell}, \gamma_\Gamma\mathcal{A}_0(\lambda_{0;\ell}) - g). \end{cases} \tag{5.40}$$

Using the notations $\mathcal{A}$ and $\hat{\mathcal{A}}$ respectively introduced in (5.25) and (5.39), we can also rewrite (5.40) more compactly as: given $\sigma_1 \in L^\infty(\Omega_1)$ and initial guess $(\lambda_{[0]}, \hat{\lambda}_{[0]}) \in \{L^2(\Sigma)^2 \times L^2(\Sigma)^2\}$,

$$\begin{cases} \lambda_{[\ell+1]} = \mathcal{A}(\lambda_{[\ell]}, \sigma_1), \\ \hat{\lambda}_{[\ell+1]} = \hat{\mathcal{A}}(\hat{\lambda}_{[\ell]}, \gamma_\Gamma\mathcal{A}_0(\lambda_{0;\ell}) - g), \end{cases} \tag{5.41}$$

where $\lambda_{[\ell]} = (\lambda_{0;\ell}, \lambda_{1;\ell}) \in L^2(\Sigma)^2$ and $\hat{\lambda}_{[\ell]} = (\hat{\lambda}_{0;\ell}, \hat{\lambda}_{1;\ell}) \in L^2(\Sigma)^2$. The convergence of this coupled scheme may be established using the following abstract theorem.

Theorem 5.1.2. *Consider the sequence $(\lambda_{[\ell]}, \hat{\lambda}_{[\ell]})$ defined by (5.41) for a given $\sigma_1 \in L^\infty(\Omega_1)$ and given initial guess $(\lambda_{0;0}, \lambda_{1;0}, \hat{\lambda}_{0;0}, \hat{\lambda}_{1;0}) \in L^2(\Sigma)^4$. Let $\|.\|$ denote the $L^2(\Sigma)^2$ norm. We make the following assumptions:*

(i) there exists a subspace U of $H^{-\frac{1}{2}}(\Gamma)$, a constant $\hat{C} \in (0, 1)$ and a constant $C > 0$ such that

$$\|\hat{\mathcal{A}}(\boldsymbol{\lambda}, \psi)\| \leq \hat{C}\|\boldsymbol{\lambda}\| + C\|\psi\|, \ \forall \boldsymbol{\lambda} \in L^2(\Sigma)^2, \forall \psi \in U,$$

(ii) $\gamma_\Gamma \mathcal{A}_0(\lambda_{0;\ell})$ *stays in* U *for all* ℓ.

Then the sequence $(\lambda_{[\ell]}, \hat{\lambda}_{[\ell]})$, $\ell \geq 0$ *is convergent.*

Proof. By the definitions of $\mathcal{A}_0$, $\mathcal{A}_1$, $\mathcal{A}$, $\hat{\mathcal{A}}_0$, $\hat{\mathcal{A}}_1$ and $\hat{\mathcal{A}}$ respectively in (5.20), (5.21), (5.25), (5.34), (5.35) and (5.39), we observe that $\mathcal{A}(\boldsymbol{\lambda}, \sigma)$ is linear with respect to $(\boldsymbol{\lambda}, \sigma) \in L^2(\Sigma)^2 \times L^\infty(\Omega_1)$, $\hat{\mathcal{A}}(\boldsymbol{\lambda}, \psi)$ is linear with respect to $(\boldsymbol{\lambda}, \psi) \in L^2(\Sigma)^2 \times H^{-\frac{1}{2}}(\Gamma)$ and $\mathcal{A}(\boldsymbol{\lambda}, 0) = \hat{\mathcal{A}}(\boldsymbol{\lambda}, 0)$, $\forall \boldsymbol{\lambda} \in L^2(\Sigma)^2$. Therefore,

$$\lambda_{[\ell+1]} - \lambda_{[\ell]} = \mathcal{A}(\lambda_{[\ell]}, \sigma_1) - \mathcal{A}(\lambda_{[\ell-1]}, \sigma_1) = \mathcal{A}(\lambda_{[\ell]} - \lambda_{[\ell-1]}, 0) = \hat{\mathcal{A}}(\lambda_{[\ell]} - \lambda_{[\ell-1]}, 0), \forall \ell \geq 1.$$

Since $0 \in U$, by assumption *(i)* we have

$$\|\lambda_{[\ell+1]} - \lambda_{[\ell]}\| \leq \hat{C}\|\lambda_{[\ell]} - \lambda_{[\ell-1]}\|, \forall \ell \geq 1.$$

This implies

$$\|\lambda_{[k+\ell+1]} - \lambda_{[k+\ell]}\| \leq \hat{C}^\ell \|\lambda_{[k+1]} - \lambda_{[k]}\|, \forall k, \ell \geq 0. \tag{5.42}$$

In particular, we have

$$\|\lambda_{[\ell+1]} - \lambda_{[\ell]}\| \leq \hat{C}^\ell \|\lambda_{[1]} - \lambda_{[0]}\|, \forall \ell \geq 0$$

so that the sequence $\|\lambda_{[\ell+1]} - \lambda_{[\ell]}\|$, $\ell \geq 0$ converges to 0 as $\ell \to \infty$. Now given $\varepsilon > 0$ arbitrarily, there exists $k_0 \geq 1$ such that

$$\|\lambda_{[k+1]} - \lambda_{[k]}\| \leq \varepsilon, \ \forall k \geq k_0.$$

Then for all $k \geq k_0$ and $m \geq 1$, we have

$$\begin{aligned}\|\lambda_{[k+m]} - \lambda_{[k]}\| \leq \sum_{\ell=0}^{m-1} \|\lambda_{[k+\ell+1]} - \lambda_{[k+\ell]}\| &\overset{(5.42)}{\leq} \sum_{\ell=0}^{m-1} \hat{C}^\ell \|\lambda_{[k+1]} - \lambda_{[k]}\| \\ &= \frac{1-\hat{C}^m}{1-\hat{C}} \|\lambda_{[k+1]} - \lambda_{[k]}\| \leq \frac{\varepsilon}{1-\hat{C}}.\end{aligned}$$

We conclude that $(\lambda_{[\ell]})_{\ell \geq 0}$ is a Cauchy sequence and thus convergent.

Next, for $\ell \geq 1$ we have

$$\begin{aligned}\hat{\lambda}_{[\ell+1]} - \hat{\lambda}_{[\ell]} &= \hat{\mathcal{A}}(\hat{\lambda}_{[\ell]}, \gamma_\Gamma \mathcal{A}_0(\lambda_{0;\ell}) - g) - \hat{\mathcal{A}}(\hat{\lambda}_{[\ell-1]}, \gamma_\Gamma \mathcal{A}_0(\lambda_{0;\ell-1}) - g) \\ &= \hat{\mathcal{A}}(\hat{\lambda}_{[\ell]} - \hat{\lambda}_{[\ell-1]}, \gamma_\Gamma \mathcal{A}_0(\lambda_{0;\ell}) - \gamma_\Gamma \mathcal{A}_0(\lambda_{0;\ell-1})).\end{aligned}$$

Assumption *(ii)* ensures that $\gamma_\Gamma \mathcal{A}_0(\lambda_{0;\ell}) - \gamma_\Gamma \mathcal{A}_0(\lambda_{0;\ell-1}) \in U$. Then assumption *(i)* gives us

$$\begin{aligned}\|\hat{\lambda}_{[\ell+1]} - \hat{\lambda}_{[\ell]}\| &\leq \hat{C}\|\hat{\lambda}_{[\ell]} - \hat{\lambda}_{[\ell-1]}\| + \underbrace{C\|\gamma_\Gamma\|\|\mathcal{A}_0\|}_{M \geq 0}\|\lambda_{0;\ell} - \lambda_{0;\ell-1}\| \\ &\leq \hat{C}\|\hat{\lambda}_{[\ell]} - \hat{\lambda}_{[\ell-1]}\| + M\|\lambda_{[\ell]} - \lambda_{[\ell-1]}\|, \ \forall \ell \geq 1.\end{aligned}$$

By induction, it is easy to show that for all $k \geq 0$ and $\ell \geq 1$,

$$\begin{aligned}\|\hat{\lambda}_{[k+\ell+1]} - \hat{\lambda}_{[k+\ell]}\| &\leq \hat{C}^\ell \|\hat{\lambda}_{[k+1]} - \hat{\lambda}_{[k]}\| + M \sum_{j=0}^{\ell-1} \hat{C}^{\ell-j-1} \|\lambda_{[k+j+1]} - \lambda_{[k+j]}\| \\ &\overset{(5.42)}{\leq} \hat{C}^\ell \|\hat{\lambda}_{[k+1]} - \hat{\lambda}_{[k]}\| + M \sum_{j=0}^{\ell-1} \hat{C}^{\ell-j-1} \hat{C}^j \|\lambda_{[k+1]} - \lambda_{[k]}\| \\ &= \hat{C}^\ell \|\hat{\lambda}_{[k+1]} - \hat{\lambda}_{[k]}\| + \ell \hat{C}^{\ell-1} M \|\lambda_{[k+1]} - \lambda_{[k]}\|. \end{aligned} \tag{5.43}$$

In particular, we have

$$\|\hat{\lambda}_{[\ell+1]} - \hat{\lambda}_{[\ell]}\| \leq \hat{C}^{\ell}\|\hat{\lambda}_{[1]} - \hat{\lambda}_{[0]}\| + \ell\hat{C}^{\ell-1}M\|\lambda_{[1]} - \lambda_{[0]}\|, \tag{5.44}$$

so the sequence $\|\hat{\Lambda}_{[\ell+1]} - \hat{\Lambda}_{[\ell]}\|$, $\ell \geq 0$ converges to 0 as $\ell \to \infty$. Now given $\varepsilon > 0$, there exists $k_1 \geq 1$ such that

$$\|\lambda_{[k+1]} - \lambda_{[k]}\| \leq \varepsilon \text{ and } \|\hat{\lambda}_{[k+1]} - \hat{\lambda}_{[k]}\| \leq \varepsilon, \ \forall k \geq k_1.$$

Then for all $k \geq k_1$ and $m \geq 2$, we have

$$\begin{aligned}
\|\hat{\lambda}_{[k+m]} - \hat{\lambda}_{[k]}\| &\leq \|\hat{\lambda}_{[k+1]} - \hat{\lambda}_{[k]}\| + \sum_{\ell=1}^{m-1}\|\hat{\lambda}_{[k+\ell+1]} - \hat{\lambda}_{[k+\ell]}\| \\
&\overset{(5.43)}{\leq} \|\hat{\lambda}_{[k+1]} - \hat{\lambda}_{[k]}\| + \sum_{\ell=1}^{m-1}(\hat{C}^{\ell}\|\hat{\lambda}_{[k+1]} - \hat{\lambda}_{[k]}\| + \ell\hat{C}^{\ell-1}M\|\lambda_{[k+1]} - \lambda_{[k]}\|) \\
&= \sum_{\ell=0}^{m-1}\hat{C}^{\ell}\|\hat{\lambda}_{[k+1]} - \hat{\lambda}_{[k]}\| + \sum_{\ell=1}^{m-1}\ell\hat{C}^{\ell-1}M\|\lambda_{[k+1]} - \lambda_{[k]}\| \\
&\leq \sum_{\ell=0}^{m-1}\hat{C}^{\ell}\varepsilon + \sum_{\ell=1}^{m-1}\ell\hat{C}^{\ell-1}M\varepsilon.
\end{aligned}$$

Therefore

$$\|\hat{\lambda}_{[k+m]} - \hat{\lambda}_{[k]}\| \leq \underbrace{\left(\frac{1-\hat{C}^m}{1-\hat{C}} + \frac{1 - m\hat{C}^{m-1} + (m-1)\hat{C}^m}{(1-\hat{C})^2}M\right)}_{\text{bounded}}\varepsilon.$$

We conclude that $(\hat{\lambda}_{[\ell]})_{\ell\geq 0}$ is also a Cauchy sequence and thus convergent. □

This theorem can be for example applied to the case of circular domains, for which the assumptions *(i)* and *(ii)* are proved in Appendix 5.A.

An algorithm combining gradient descent with domain decomposition

We here use the same notation as in Section 5.1.3, i.e. the domain Ω is splitted into two disjoint subdomains Ω_0 and Ω_1 where Ω_0 is the neighborhood of $\Gamma := \partial\Omega$. We assume that the unknown conductivity contrast σ equals 0 in Ω_0 and equals σ_1 in Ω_1, i.e. $\sigma = \sigma_1 \mathbb{1}_{\Omega_1}$. Then the 1-step one-shot inversion algorithm combining the nonoverlapping OSM (5.41) and the gradient descent (5.11) can be written as follows: given some initial guess $(\sigma_1^0, \lambda^0, \hat{\lambda}^0)$, we iterate using the induction

$$\begin{cases}
\sigma_1^{n+1} = \sigma_1^n - \tau\nabla u^{\text{inc}} \cdot \nabla\hat{\mathcal{A}}_1(\hat{\lambda}_1^n), \\
\lambda^{n+1} = \mathcal{A}(\lambda^n, \sigma_1^{n+1}), \\
\hat{\lambda}^{n+1} = \hat{\mathcal{A}}(\hat{\lambda}^n, \gamma_\Gamma\mathcal{A}_0(\lambda_0^n) - g),
\end{cases} \tag{5.45}$$

where $\lambda^n = (\lambda_0^n, \lambda_1^n)$, $\hat{\lambda}^n = (\hat{\lambda}_0^n, \hat{\lambda}_1^n)$; $\mathcal{A}$ and $\hat{\mathcal{A}}$ are respectively defined in (5.25) and (5.39). The k-step version of the algorithm (5.45) can be also written as: given some initial guess $(\sigma_1^0, \lambda^0, \hat{\lambda}^0)$, we iterate using the induction

$$\begin{cases}
\sigma_1^{n+1} = \sigma_1^n - \tau\nabla u^{\text{inc}} \cdot \nabla\hat{\mathcal{A}}_1(\hat{\lambda}_1^n), \\
\lambda_{[0]}^{n+1} = \lambda^n, \hat{\lambda}_{[0]}^{n+1} = \hat{\lambda}^n, \\
\text{for } \ell = 0, 1, ..., k-1 : \\
\quad \left| \begin{array}{l} \lambda_{[\ell+1]}^{n+1} = \mathcal{A}(\lambda_{[\ell]}^n, \sigma_1^{n+1}), \\ \hat{\lambda}_{[\ell+1]}^{n+1} = \hat{\mathcal{A}}(\hat{\lambda}_{[\ell]}^n, \gamma_\Gamma\mathcal{A}_0(\lambda_{[\ell]}^n) - g), \end{array} \right. \\
\lambda^{n+1} = \lambda_{[k]}^{n+1}, \hat{\lambda}^{n+1} = \hat{\lambda}_{[k]}^{n+1},
\end{cases} \tag{5.46}$$

where $\lambda^n = (\lambda_0^n, \lambda_1^n)$, $\hat{\lambda}^n = (\hat{\lambda}_0^n, \hat{\lambda}_1^n)$, $\lambda_{[\ell]}^n = (\lambda_{0;\ell}^n, \lambda_{1;\ell}^n)$, $\hat{\lambda}_{[\ell]}^n = (\hat{\lambda}_{0;\ell}^n, \hat{\lambda}_{1;\ell}^n)$; $\mathcal{A}$ and $\hat{\mathcal{A}}$ are respectively defined in (5.25) and (5.39). As we shall explain in the next section, the scheme (5.46) does not exactly fall into the abstract framework for the multi-step one-shot (1.8) for the linearized problem studied in Section 1.2. The main reason behind this is that $\hat{\lambda}$ does not exactly present the numerical adjoint of λ. An alternative scheme that fits into the framework of Section 1.2 will be proposed in Section 5.2.3.

In the case where Ω_1 is the disjoint union of the subdomains $\Omega_{1,i}, i = 1, 2..., n_{\sigma_1}$ and σ_1 is given by

$$\sigma_1 = \sum_{i=1}^{n_{\sigma_1}} \sigma_{1,i} \mathbb{1}_{\Omega_{1,i}} \text{ where } \sigma_{1,i} \in \mathbb{R}, i = 1, 2, ..., n_{\sigma_1},$$

the first step in scheme (5.46) should be replaced with

$$\sigma_{1,i}^{n+1} = \sigma_{1,i}^n - \tau \int_{\Omega_{1,i}} \nabla u^{\text{inc}} \cdot \nabla \hat{\mathcal{A}}_1(\hat{\lambda}_1^n) \, \mathrm{dx}, \quad i = 1, 2, ..., n_{\sigma_1}, \tag{5.47}$$

where $\sigma_1^n = \sum_{i=1}^{n_{\sigma_1}} \sigma_{1,i}^n \mathbb{1}_{\Omega_{1,i}}$, $n \geq 0$.

5.2 Discretized versions of the algorithms studied in 5.1.3

The main goal of this section is to give the discretized version of (5.46). In order to do that, we first need to provide the discretized versions for the operators $\mathcal{A}$ and $\hat{\mathcal{A}}$, then step by step we provide the discretized version of the nonoverlapping OSM (5.41). Let us give some details in the case where these discretized versions are obtained by discretizing the associated problems using finite element methods. For $i = 0, 1$, we assume that Ω_i is discretized using a regular triangulation $\mathcal{T}(\Omega_i)$. We further assume that Σ and Γ are respectively discretized using edge elements $\mathcal{S}(\Sigma)$ and $\mathcal{S}(\Gamma)$. We recall the definition of the following $\mathbb{P}^1$ finite element spaces:

$$\begin{aligned}
\mathcal{V}(\Omega_i) &:= \left\{ u \in C^0(\Omega_i) : u|_{T \in \mathcal{T}(\Omega_i)} \in \mathbb{P}^1 \right\} = \text{Span}\{\phi_i^j, 1 \leq j \leq n_i\}, \ i = 0, 1, \\
\mathcal{V}(\Sigma) &:= \left\{ \lambda \in C^0(\Sigma) : \lambda|_{S \in \mathcal{S}(\Sigma)} \in \mathbb{P}^1 \right\} = \text{Span}\{\phi_\Sigma^j, 1 \leq j \leq n_\Sigma\}, \\
\text{and } \mathcal{V}(\Gamma) &:= \left\{ g \in C^0(\Gamma) : g|_{S \in \mathcal{S}(\Gamma)} \in \mathbb{P}^1 \right\} = \text{Span}\{\phi_\Gamma^j, 1 \leq j \leq n_\Gamma\},
\end{aligned}$$

where n_i, $i = 0, 1$ is the number of degrees of freedom for the solution in Ω_i, n_Γ and n_Σ are the number of degrees of freedom respectively for the measurement on Γ and for the impedance on Σ. A discretized function in any of these spaces is determined by the corresponding column vectors with respect to the chosen basis.

We also consider the following case where σ_1 is of the form

$$\sigma_1 = \sum_{i=1}^{n_{\sigma_1}} \sigma_{1,i} \mathbb{1}_{\Omega_{1,i}} \text{ where } \sigma_{1,i} \in \mathbb{R}, i = 1, 2, ..., n_{\sigma_1},$$

and we set

$$S_1 := \text{Col}(\sigma_{1;i})_{i=1}^{n_{\sigma_1}}.$$

5.2.1 Discretized versions of the operators $\mathcal{A}$ and $\hat{\mathcal{A}}$

Let us prepare some necessary matrices for writing the discretized version of the operators $\mathcal{A}$ and $\hat{\mathcal{A}}$. We respectively denote by $\tilde{M}_\Sigma \in \mathbb{R}^{n_\Sigma \times n_\Sigma}$, and $\tilde{M}_\Gamma \in \mathbb{R}^{n_\Gamma \times n_\Gamma}$ the Σ-interface and Γ-interface mass matrices. More precisely,

$$\tilde{M}_{\Sigma;p,q} \coloneqq \int_\Sigma \phi_\Sigma^q \phi_\Sigma^p \,\mathrm{ds},\ 1 \le p, q \le n_\Sigma, \tag{5.48}$$

and

$$\tilde{M}_{\Gamma;p,q} = \int_\Gamma \phi_\Gamma^p \phi_\Gamma^q \,\mathrm{ds},\ 1 \le p, q \le n_\Gamma. \tag{5.49}$$

Also, for $i = 0, 1$, we introduce $\tilde{R}_i \in \mathbb{R}^{n_\Sigma \times n_i}$ the interpolation matrix from $\mathcal{V}(\Omega_i)$ to $\mathcal{V}(\Sigma)$. Then

$$\tilde{R}_i = \tilde{M}_\Sigma^{-1} \tilde{M}_{i;\Sigma},\ \text{or equivalently, } \tilde{M}_{i;\Sigma} = \tilde{M}_\Sigma \tilde{R}_i, \tag{5.50}$$

where $\tilde{M}_{i;\Sigma} \in \mathbb{R}^{n_\Sigma \times n_i}$ is defined by

$$\tilde{M}_{i;\Sigma;p,q} \coloneqq \int_\Sigma \phi_i^q \phi_\Sigma^p \,\mathrm{ds},\ 1 \le p \le n_\Sigma, 1 \le q \le n_i. \tag{5.51}$$

Now let $\tilde{A}_i \in \mathbb{R}^{n_i \times n_i}, i = 0, 1$ be the matrices arising from the discretization of $-\triangle + \eta$ along with the boundary condition $\partial_\nu \mp \beta$ on Σ. More precisely, $\tilde{A}_i$ is defined by

$$\tilde{A}_i = \tilde{K}_i + \beta \tilde{R}_i^\intercal \tilde{M}_\Sigma \tilde{R}_i,\ i = 0, 1 \tag{5.52}$$

where the local matrices of the problem $\tilde{K}_i \in \mathbb{R}^{n_i \times n_i}$, $i = 0, 1$, are defined by

$$\tilde{K}_{i;p,q} \coloneqq \int_{\Omega_i} \sigma_0 \nabla \phi_i^q \cdot \nabla \phi_i^p \,\mathrm{dx} + \int_{\Omega_i} \eta \phi_i^q \phi_i^p \,\mathrm{dx},\ 1 \le p, q \le n_i, i = 0, 1. \tag{5.53}$$

In addition, we introduce $\tilde{H} \in \mathbb{R}^{n_\Sigma \times n_i}$ the interpolation matrix from $\mathcal{V}(\Omega_0)$ into $\mathcal{V}(\Gamma)$. Then

$$\tilde{H} = \tilde{M}_\Gamma^{-1} \tilde{M}_{0;\Gamma} \tag{5.54}$$

where $\tilde{M}_{0;\Gamma} \in \mathbb{R}^{n_\Gamma \times n_0}$ is defined by

$$\tilde{M}_{0;\Gamma;p,q} \coloneqq \int_\Gamma \phi_0^q \phi_\Gamma^p \,\mathrm{ds},\ 1 \le p \le n_\Gamma, 1 \le q \le n_0. \tag{5.55}$$

Finally, we consider the matrix $\tilde{Q}_1 \in \mathbb{R}^{n_1 \times n_{\sigma_i}}$ associated with the gradient of the cost functional, which is given by

$$\tilde{Q}_{1;p,q} \coloneqq \int_{\Omega_{1,q}} \sigma_{1,q} \nabla u^{\mathrm{inc}} \cdot \nabla \phi_1^p \,\mathrm{dx},\ 1 \le p \le n_1, 1 \le q \le n_{\sigma_1}. \tag{5.56}$$

Now we are ready for discretizing the operators $\mathcal{A}_i$, $i = 0, 1$. The variational formulation for $u = \mathcal{A}_0(\lambda)$ in (5.20) is given by

$$\int_{\Omega_0} \sigma_0 \nabla u \cdot \nabla v \,\mathrm{dx} + \int_{\Omega_0} \eta u v \,\mathrm{dx} + \int_\Sigma \beta u v \,\mathrm{ds} = -\int_\Sigma \lambda v \,\mathrm{ds},\ \forall v \in H^1(\Omega_0).$$

Therefore, the discrete equivalent of the operator $\mathcal{A}_0$ is the application $\Lambda \in \mathbb{R}^{n_\Sigma} \mapsto U \in \mathbb{R}^{n_0}$ such that

$$\tilde{A}_0 U = -\tilde{R}_0^\intercal \tilde{M}_\Sigma \Lambda = -\tilde{M}_{0;\Sigma}^\intercal \Lambda.$$

With an abuse of notation, we shall set in the following

$$\boxed{\mathcal{A}_0(\Lambda) \coloneqq -\tilde{A}_0^{-1}\tilde{M}_{0;\Sigma}^{\intercal}\Lambda,\ \forall \Lambda \in \mathbb{R}^{n_\Sigma}.} \tag{5.57}$$

The variational formulation for $u = \mathcal{A}_1(\lambda, \sigma_1)$ in (5.21) is given by

$$\int_{\Omega_1} \sigma_0 \nabla u \cdot \nabla v \,\mathrm{d}x + \int_{\Omega_1} \eta u v \,\mathrm{d}x + \int_{\Sigma} \beta u v \,\mathrm{d}s = \int_{\Omega_1} \sigma_1 \nabla u^{\mathrm{inc}} \cdot \nabla v \,\mathrm{d}x + \int_{\Sigma} \lambda v \,\mathrm{d}s,\ \forall v \in H^1(\Omega_1).$$

Therefore, the discrete equivalent of the operator $\mathcal{A}_1$ is the application $(\Lambda, S_1) \in \mathbb{R}^{n_\Sigma} \times \mathbb{R}^{n_{\sigma_1}} \mapsto U \in \mathbb{R}^{n_0}$ such that

$$\tilde{A}_1 U = \tilde{Q}_1 S_1 + \tilde{R}_1^{\intercal}\tilde{M}_\Sigma \Lambda = \tilde{Q}_1 S_1 + \tilde{M}_{1;\Sigma}^{\intercal}\Lambda.$$

With an abuse of notation, we shall set in the following

$$\boxed{\mathcal{A}_1(\Lambda, S_1) \coloneqq \tilde{A}_1^{-1}\tilde{M}_{1;\Sigma}^{\intercal}\Lambda + \tilde{A}_1^{-1}\tilde{Q}_1 S_1,\ \forall (\Lambda, S_1) \in \mathbb{R}^{n_\Sigma} \times \mathbb{R}^{n_{\sigma_1}}.} \tag{5.58}$$

The variational formulation for $u = \hat{\mathcal{A}}_0(\lambda, \psi)$ in (5.34) is given by

$$\int_{\Omega_0} \sigma_0 \nabla u \cdot \nabla v \,\mathrm{d}x + \int_{\Omega_0} \eta u v \,\mathrm{d}x + \int_{\Sigma} \beta u v \,\mathrm{d}s = \int_{\Gamma} \psi v \,\mathrm{d}s - \int_{\Sigma} \lambda v \,\mathrm{d}s,\ \forall v \in H^1(\Omega_0).$$

Therefore, the discrete equivalent of the operator $\hat{\mathcal{A}}_0$ is the application $(\Lambda, \Psi) \in \mathbb{R}^{n_\Sigma} \times \mathbb{R}^{n_\Gamma} \mapsto U \in \mathbb{R}^{n_0}$ such that

$$\tilde{A}_0 U = \tilde{H}^{\intercal}\tilde{M}_\Gamma \Psi - \tilde{R}_0^{\intercal}\tilde{M}_\Sigma \Lambda = \tilde{H}^{\intercal}\tilde{M}_\Gamma \Psi - \tilde{M}_{0;\Sigma}^{\intercal}\Lambda.$$

With an abuse of notation, we shall set in the following

$$\boxed{\hat{\mathcal{A}}_0(\Lambda, \Psi) \coloneqq -\tilde{A}_0^{-1}\tilde{M}_{0;\Sigma}^{\intercal}\Lambda + \tilde{A}_0^{-1}\tilde{H}^{\intercal}\tilde{M}_\Gamma \Psi,\ \forall (\Lambda, \Psi) \in \mathbb{R}^{n_\Sigma} \times \mathbb{R}^{n_\Gamma}.} \tag{5.59}$$

Finally, the variational formulation for $u = \hat{\mathcal{A}}_1(\Lambda)$ in (5.35) is given by

$$\int_{\Omega_1} \sigma_0 \nabla u \cdot \nabla v \,\mathrm{d}x + \int_{\Omega_1} \eta u v \,\mathrm{d}x + \int_{\Sigma} \beta u v \,\mathrm{d}s = \int_{\Sigma} \lambda v \,\mathrm{d}s,\ \forall v \in H^1(\Omega_1).$$

Therefore, the discrete equivalent of the operator $\hat{\mathcal{A}}_1$ is the application $\Lambda \in \mathbb{R}^{n_\Sigma} \mapsto U \in \mathbb{R}^{n_0}$ such that

$$\tilde{A}_1 U = \tilde{R}_1^{\intercal}\tilde{M}_\Gamma \Lambda = \tilde{M}_{1;\Sigma}^{\intercal}\Lambda.$$

With an abuse of notation, we shall set in the following

$$\boxed{\hat{\mathcal{A}}_1(\Lambda) \coloneqq \tilde{A}_1^{-1}\tilde{M}_{1;\Sigma}^{\intercal}\Lambda,\ \forall \Lambda \in \mathbb{R}^{n_\Sigma}.} \tag{5.60}$$

Then, similarly to (5.25), with an abuse of notation, we define the operator $\mathcal{A} : \{\mathbb{R}^{n_\Sigma} \times \mathbb{R}^{n_\Sigma}\} \times \mathbb{R}^{n_{\sigma_1}} \to \{\mathbb{R}^{n_\Sigma} \times \mathbb{R}^{n_\Sigma}\}$ by

$$\mathcal{A}(\boldsymbol{\Lambda}, S_1) \coloneqq (\boldsymbol{\Lambda_1} - 2\beta\tilde{R}_1\mathcal{A}_1(\boldsymbol{\Lambda_1}, S_1), \boldsymbol{\Lambda_0} + 2\beta\tilde{R}_0\mathcal{A}_0(\boldsymbol{\Lambda_0})) \tag{5.61}$$

where $\boldsymbol{\Lambda} = (\boldsymbol{\Lambda_0}, \boldsymbol{\Lambda_1}) \in \{\mathbb{R}^{n_\Sigma} \times \mathbb{R}^{n_\Sigma}\}$ and $S_1 \in \mathbb{R}^{n_{\sigma_1}}$. This definition can be written in a more explicit way as

$$\boxed{\mathcal{A}(\mathbf{\Lambda}, S_1) = \mathbb{B}\mathbf{\Lambda} + \mathbb{M}S_1} \tag{5.62}$$

where

$$\mathbb{B} := \begin{bmatrix} 0 & I - 2\beta\tilde{R}_1\tilde{A}_1^{-1}\tilde{M}_{1;\Sigma}^{\intercal} \\ I - 2\beta\tilde{R}_0\tilde{A}_0^{-1}\tilde{M}_{0;\Sigma}^{\intercal} & 0 \end{bmatrix} \in \mathbb{R}^{2n_\Sigma \times 2n_\Sigma} \tag{5.63}$$

and

$$\mathbb{M} := \begin{bmatrix} -2\beta\tilde{R}_1\tilde{A}_1^{-1}\tilde{Q}_1 \\ 0 \end{bmatrix} \in \mathbb{R}^{2n_\Sigma \times n_{\sigma_1}}. \tag{5.64}$$

Also, similarly to (5.39), with an abuse of notation, we define the operator $\hat{\mathcal{A}} : \{\mathbb{R}^{n_\Sigma} \times \mathbb{R}^{n_\Sigma}\} \times \mathbb{R}^{n_\Gamma} \to \{\mathbb{R}^{n_\Sigma} \times \mathbb{R}^{n_\Sigma}\}$ by

$$\hat{\mathcal{A}}(\mathbf{\Lambda}, \Psi) := (\mathbf{\Lambda_1} - 2\beta\tilde{R}_1\hat{\mathcal{A}}_1(\mathbf{\Lambda_1}), \mathbf{\Lambda_0} + 2\beta\tilde{R}_0\hat{\mathcal{A}}_0(\mathbf{\Lambda_0}, \Psi)) \tag{5.65}$$

where $\mathbf{\Lambda} = (\mathbf{\Lambda_0}, \mathbf{\Lambda}_1) \in \{\mathbb{R}^{n_\Sigma} \times \mathbb{R}^{n_\Sigma}\}$ and $\Psi \in \mathbb{R}^{n_\Gamma}$. This definition can be written in a more explicit way as

$$\boxed{\hat{\mathcal{A}}(\mathbf{\Lambda}, \Psi) = \mathbb{B}\mathbf{\Lambda} + \mathbb{E}\Psi} \tag{5.66}$$

where $\mathbb{B}$ is defined in (5.63) and

$$\mathbb{E} := \begin{bmatrix} 0 \\ 2\beta\tilde{R}_0\tilde{A}_0^{-1}\tilde{H}^{\intercal}\tilde{M}_\Gamma \end{bmatrix} \in \mathbb{R}^{2n_\Sigma \times n_\Gamma}. \tag{5.67}$$

Using the discretized versions of the operators $\mathcal{A}_0$, $\mathcal{A}$, $\hat{\mathcal{A}}_1$ and $\hat{\mathcal{A}}$ respectively defined by (5.57), (5.62), (5.60) and (5.66), we will give in the following section the discretized versions of algorithms (5.45) and (5.46). We end this section by a lemma that links the spectral radius of $\mathbb{B}$ with the choice of the Robin parameter β.

Lemma 5.2.1. *The matrix $\mathbb{B}$ defined by (5.63) satisfies $\|\mathbb{B}\|_2 < 1$ if the Robin parameter $\beta > 0$ is chosen sufficiently small.*

Proof. By (5.50) and (5.52) we have

$$I - 2\beta\tilde{R}_i\tilde{A}_i^{-1}\tilde{M}_{i;\Sigma}^{\intercal} = I - 2\beta\tilde{M}_\Sigma^{-1}\tilde{M}_{i;\Sigma}(\tilde{K}_i + \beta\tilde{R}_i^{\intercal}\tilde{M}_\Sigma\tilde{R}_i)^{-1}\tilde{M}_{i;\Sigma}^{\intercal} = I - V_i(\beta),\ i = 1, 2,$$

where

$$V_i(\beta) := 2\beta\tilde{M}_\Sigma^{-1}\tilde{M}_{i;\Sigma}(\tilde{K}_i + \beta\tilde{R}_i^{\intercal}\tilde{M}_\Sigma\tilde{R}_i)^{-1}\tilde{M}_{i;\Sigma}^{\intercal},\ \beta > 0, i = 1, 2.$$

For $i = 0, 1$, it is easy to verify that $V_i(\beta)$ is symmetric and positive definite for every $\beta > 0$, also $V_i(\beta)$ is bounded since $\lim_{\beta \to 0^+} V_i(\beta) = 0$ and $\lim_{\beta \to +\infty} V_i(\beta) = 2\tilde{M}_\Sigma^{-1}\tilde{M}_{i;\Sigma}(\tilde{R}_i^{\intercal}\tilde{M}_\Sigma\tilde{R}_i)^{-1}\tilde{M}_{i;\Sigma}^{\intercal}$. By (5.63),

$$\mathbb{B} = \begin{bmatrix} 0 & I - V_1(\beta) \\ I - V_0(\beta) & 0 \end{bmatrix},$$

therefore $\mathbb{B}^{\intercal}\mathbb{B}$ is of the form

$$\mathbb{B}^{\intercal}\mathbb{B} = \begin{bmatrix} (I - V_1(\beta))^2 & 0 \\ 0 & (I - V_0(\beta))^2 \end{bmatrix}$$

where $V_i(\beta)$, $i = 1, 2$, are symmetric, positive definite and bounded. We deduce that

$$\|B\|_2 = \rho(\mathbb{B}^{\intercal}\mathbb{B}) = \Big(\max_{i=0,1} \rho(I - V_i(\beta))\Big)^2,$$

which leads to the statement of the lemma. □

5.2.2 Discretized versions of algorithms (5.45) and (5.46)

Let $\Lambda_i^n \in \mathbb{R}^{n_\Sigma}$, $\hat{\Lambda}_i^n \in \mathbb{R}^{n_\Sigma}$, $i = 0, 1$ and $G \in \mathbb{R}^{n_\Gamma}$ be the column vectors associated respectively with λ_i^n, $\hat{\lambda}_i^n$, $i = 0, 1$ and $g \in \mathcal{V}(\Gamma)$ as in (5.45). Let us set

$$\Lambda^n = (\Lambda_0^n, \Lambda_1^n) \text{ and } \hat{\Lambda}^n = (\hat{\Lambda}_0^n, \hat{\Lambda}_1^n).$$

for $n \geq 0$. Then the 1-step one-shot algorithm (5.45) with the gradient defined in (5.5) can be synthetically written in the discrete form as

$$\begin{cases} S_1^{n+1} = S_1^n - \tau \tilde{Q}_1^\intercal \hat{\mathcal{A}}_1(\hat{\Lambda}_1^n), \\ \Lambda^{n+1} = \mathcal{A}(\Lambda^n, S_1^{n+1}), \\ \hat{\Lambda}^{n+1} = \hat{\mathcal{A}}(\hat{\Lambda}^n, \tilde{H}\mathcal{A}_0(\Lambda_0^n) - G), \end{cases} \tag{5.68}$$

where the operators $\mathcal{A}_0$, $\mathcal{A}$, $\hat{\mathcal{A}}_1$ and $\hat{\mathcal{A}}$ are respectively defined by (5.57), (5.62), (5.60) and (5.66). Indeed, using the expressions of $\mathcal{A}_0$ in (5.57) and of $\hat{\mathcal{A}}$ in (5.66), we have

$$\hat{\mathcal{A}}(\hat{\Lambda}^n, \tilde{H}\mathcal{A}_0(\Lambda_0^n) - G) = \mathbb{B}\hat{\Lambda}^n + \mathbb{E}(-\tilde{H}\tilde{A}_0^{-1}\tilde{M}_{0;\Sigma}^\intercal \Lambda_0^n - G) = \mathbb{B}\hat{\Lambda}^n + \mathbb{D}\Lambda^n - \mathbb{E}G$$

where

$$\begin{aligned} \mathbb{D} := \mathbb{E}\underbrace{\begin{bmatrix} -\tilde{H}\tilde{A}_0^{-1}\tilde{M}_{0;\Sigma}^\intercal & 0 \end{bmatrix}}_{\in \mathbb{R}^{n_\Gamma \times 2n_\Sigma}} &= \begin{bmatrix} 0 \\ 2\beta \tilde{R}_0 \tilde{A}_0^{-1} \tilde{H}^\intercal \tilde{M}_\Gamma \end{bmatrix} \begin{bmatrix} -\tilde{H}\tilde{A}_0^{-1}\tilde{M}_{0;\Sigma}^\intercal & 0 \end{bmatrix} \\ &= \begin{bmatrix} 0 & 0 \\ -2\beta \tilde{R}_0 \tilde{A}_0^{-1} \tilde{H}^\intercal \tilde{M}_\Gamma \tilde{H} \tilde{A}_0^{-1} \tilde{M}_{0;\Sigma}^\intercal & 0 \end{bmatrix} \in \mathbb{R}^{2n_\Sigma \times 2n_\Sigma}. \end{aligned} \tag{5.69}$$

Also, using the definition of $\hat{\mathcal{A}}_1$ in (5.60), we can rewrite

$$\tilde{Q}_1^\intercal \hat{\mathcal{A}}_1(\hat{\Lambda}_1^n) = \mathbb{L}\hat{\Lambda}^n$$

where

$$\mathbb{L} := \begin{bmatrix} 0 & \tilde{Q}_1^\intercal \tilde{A}_1^{-1} \tilde{M}_{1;\Sigma}^\intercal \end{bmatrix} \in \mathbb{R}^{n_{\sigma_1} \times 2n_\Sigma}. \tag{5.70}$$

With these expressions, scheme (5.68) can be more explicitly written as

$$\boxed{\begin{cases} S_1^{n+1} = S_1^n - \tau \mathbb{L}\hat{\Lambda}^n \\ \Lambda^{n+1} = \mathbb{B}\Lambda^n + \mathbb{M}S_1^{n+1}, \\ \hat{\Lambda}^{n+1} = \mathbb{B}\hat{\Lambda}^n + \mathbb{D}\Lambda^n - \mathbb{E}G, \end{cases}} \tag{5.71}$$

where $\mathbb{B}$, $\mathbb{M}$, $\mathbb{D}$, $\mathbb{E}$ and $\mathbb{L}$ are respectively defined by (5.63), (5.64), (5.69), (5.67) and (5.70).

The structure of algorithm (5.71) is slightly different from the framework introduced in Section 1.2. We shall explain later how one can write a scheme that falls into the framework of Section 1.2 and how one can equivalently rewrite (5.71) so that it resembles that scheme. We end this section by indicating the k-step version of algorithm (5.46).

Let $\Lambda_i^n \in \mathbb{R}^{n_\Sigma}$, $\hat{\Lambda}_i^n \in \mathbb{R}^{n_\Sigma}$, $i = 0, 1$ and $G \in \mathbb{R}^{n_\Gamma}$ be the column vectors associated respectively with λ_i^n, $\hat{\lambda}_i^n$, $i = 0, 1$ and $g \in \mathcal{V}(\Gamma)$ as in (5.45). Let us set

$$\Lambda^n = (\Lambda_0^n, \Lambda_1^n), \Lambda_{[\ell]}^n = (\Lambda_{0;\ell}^n, \Lambda_{1;\ell}^n), \hat{\Lambda}^n = (\hat{\Lambda}_0^n, \hat{\Lambda}_1^n) \text{ and } \hat{\Lambda}_{[\ell]}^n = (\hat{\Lambda}_{0;\ell}^n, \hat{\Lambda}_{1;\ell}^n).$$

for $n \geq 0$ and $\ell \geq 0$. Then, the k-step one-shot algorithm (5.46) with the gradient defined in (5.5) can be synthetically written in the discrete form as

$$\begin{cases} S_1^{n+1} = S_1^n - \tau \tilde{Q}_1^\intercal \hat{\mathcal{A}}_1(\hat{\Lambda}_1^n), \\ \Lambda_{[0]}^{n+1} = \Lambda^n, \hat{\Lambda}_{[0]}^{n+1} = \hat{\Lambda}^n, \\ \text{for } \ell = 0, 1, ..., k-1 : \\ \quad \left| \begin{array}{l} \Lambda_{[\ell+1]}^{n+1} = \mathcal{A}(\Lambda_{[\ell]}^n, S_1^{n+1}), \\ \hat{\Lambda}_{[\ell+1]}^{n+1} = \hat{\mathcal{A}}(\hat{\Lambda}_{[\ell]}^n, \tilde{H}\mathcal{A}_0(\Lambda_{[\ell]}^n) - G), \end{array} \right. \\ \Lambda^{n+1} = \Lambda_{[k]}^{n+1}, \hat{\Lambda}^{n+1} = \hat{\Lambda}_{[k]}^{n+1}, \end{cases} \tag{5.72}$$

where the operators $\mathcal{A}_0$, $\mathcal{A}$, $\hat{\mathcal{A}}_1$ and $\hat{\mathcal{A}}$ are respectively defined by (5.57), (5.62), (5.60) and (5.66). With these expression, the scheme can be more explicitly written as:

$$\boxed{\begin{cases} S_1^{n+1} = S_1^n - \tau \mathbb{L}\hat{\Lambda}^n, \\ \Lambda_0^{n+1} = \Lambda^n, \hat{\Lambda}_0^{n+1} = \hat{\Lambda}^n, \\ \text{for } \ell = 0, 1, \ldots, k-1 : \\ \quad \left| \begin{array}{l} \Lambda_{[\ell+1]}^{n+1} = \mathbb{B}\Lambda_{[\ell]}^{n+1} + \mathbb{M}S_1^{n+1}, \\ \hat{\Lambda}_{[\ell+1]} = \mathbb{B}\hat{\Lambda}_{[\ell]}^{n+1} + \mathbb{D}\Lambda_{[\ell]}^{n+1} - \mathbb{E}G, \end{array} \right. \\ \Lambda^{n+1} := \Lambda_{[k]}^{n+1}, \hat{\Lambda}^{n+1} := \hat{\Lambda}_{[k]}^{n+1} \end{cases}} \tag{5.73}$$

where $\mathbb{B}$, $\mathbb{M}$, $\mathbb{D}$, $\mathbb{E}$ and $\mathbb{L}$ are respectively defined by (5.63), (5.64), (5.69), (5.67) and (5.70).

5.2.3 An alternative algorithm based on the numerical adjoint

The goal of this section is to build an algorithm inspired by the abstract framework in Section 1.2, by working directly on the discretized version of nonoverlapping OSM for the direct problem and the numerical equivalent of the cost functional.

Using the notations for the discretization already introduced in Section 5.2.1, one can easily observe that solving the forward problem using (5.12) and (5.13) is equivalent to solve the coupled system

$$\begin{cases} \tilde{A}_0 U_0 = -\tilde{M}_{0;\Sigma}^\intercal \Lambda_0, \\ \Lambda_0 = \Lambda_1 - 2\beta \tilde{R}_1^\intercal U_1, \\ \tilde{A}_1 U_1 = \tilde{Q}_1 S_1 + \tilde{M}_{1;\Sigma}^\intercal \Lambda_1, \\ \Lambda_1 = \Lambda_0 + 2\beta \tilde{R}_0^\intercal U_0, \end{cases} \tag{5.74}$$

where the matrices $\tilde{A}_i, \tilde{M}_{i;\Sigma}, \tilde{R}_i$, $i = 0, 1$ and $\tilde{Q}_1$ are respectively defined in (5.52), (5.51), (5.50) and (5.56). Eliminating U_0 and U_1 in this system using the expressions

$$U_0 = -\tilde{A}_0^{-1} \tilde{M}_{0;\Sigma}^\intercal \Lambda_0 \text{ and } U_1 = \tilde{A}_1^{-1}(\tilde{Q}_1 S_1 + \tilde{M}_{1;\Sigma}^\intercal \Lambda_1), \tag{5.75}$$

one ends up with the following equivalent writing of the forward problem in terms of $\Lambda = (\Lambda_0, \Lambda_1)$ as

$$\boxed{\Lambda = \mathbb{B}\Lambda + \mathbb{M}S_1} \tag{5.76}$$

where $\mathbb{B}$ and $\mathbb{M}$ are respectively defined by (5.63) and (5.64). This is compatible with the discretized version of the iterative equation (5.24) for Λ. Indeed, using the notations introduced in Section 5.2.1, the discretized version of (5.24) is written as

$$\Lambda_{[\ell+1]} = \mathcal{A}(\Lambda_{[\ell]}, S_1) \overset{(5.62)}{=} \mathbb{B}\Lambda_{[\ell]} + \mathbb{M}S_1, \quad \ell = 0, 1, \ldots$$

Taking $\ell \to \infty$ in this fixed point iteration also yields (5.76).

We now need to introduce the numerical cost functional associated with the inverse conductivity problem. The measurement $u_0(\sigma_1)|_\Gamma$ are represented at the discrete level by $\tilde{H}U_0$ where we recall that $\tilde{H}$ is the interpolation matrix from $\mathcal{V}(\Omega_0)$ to $\mathcal{V}(\Gamma)$ defined by (5.54) and U_0 is the solution of (5.74). Then, by the expression of U_0 in term of Λ_0 in (5.75), we have

$$\tilde{H}U_0 = -\tilde{H}\tilde{A}_0^{-1}\tilde{M}_{0;\Sigma}^\intercal \Lambda_0 = \mathbb{H}\Lambda$$

where

$$\mathbb{H} := \begin{bmatrix} -\tilde{H}\tilde{A}_0^{-1}\tilde{M}_{0;\Sigma}^\intercal & 0 \end{bmatrix} \in \mathbb{R}^{n_\Gamma \times 2n_\Sigma}. \tag{5.77}$$

Let us denote by S_1^{exact} the exact conductivity we want to retrieve and G the corresponding measurements. The discrete least square cost functional can then be defined as

$$J(S_1) := \frac{1}{2}(\mathbb{H}\Lambda - G)^\intercal \tilde{M}_\Gamma (\mathbb{H}\Lambda - G) \tag{5.78}$$

where Λ is the solution of (5.76). We thus obtain the discretized version of the conductivity inverse problem:

$$\begin{array}{ll} \text{direct problem:} & \Lambda(S_1) = \mathbb{B}\Lambda(S_1) + \mathbb{M}S_1, \\ \text{inverse problem:} & \text{retrieve } S_1 \text{ from } G, \\ \text{cost functional:} & J(S_1) = \frac{1}{2}(\mathbb{H}\Lambda(S_1) - G)^\intercal \tilde{M}_\Gamma (\mathbb{H}\Lambda(S_1) - G). \end{array} \tag{5.79}$$

The numerical adjoint for (5.79)

Since (5.79) falls into our abstract framework in Section 1.2, we can define the numerical adjoint $\tilde{\Lambda} = \tilde{\Lambda}(S_1)$ of Λ as

$$\tilde{\Lambda} = \mathbb{B}^\intercal \tilde{\Lambda} + \mathbb{H}^\intercal \tilde{M}_\Gamma (\mathbb{H}\Lambda - G). \tag{5.80}$$

Then, the gradient of the cost functional J in (5.78) is given by

$$\nabla J(\Lambda) = \mathbb{M}^\intercal \tilde{\Lambda}. \tag{5.81}$$

The k-step one-shot alternative algorithm then reads

$$\begin{cases} S_1^{n+1} = S_1^n - \tau \mathbb{M}^\intercal \tilde{\Lambda}^n, \\ \Lambda_0^{n+1} = \Lambda^n, \tilde{\Lambda}_0^{n+1} = \tilde{\Lambda}^n, \\ \text{for } \ell = 0, 1, \dots, k-1 : \\ \quad \left| \begin{array}{l} \Lambda_{[\ell+1]}^{n+1} = \mathbb{B}\Lambda_{[\ell]}^{n+1} + \mathbb{M}S_1^{n+1}, \\ \tilde{\Lambda}_{[\ell+1]} = \mathbb{B}^\intercal \tilde{\Lambda}_{[\ell]}^{n+1} + \mathbb{H}\tilde{M}_\Gamma(\mathbb{H}\Lambda - G), \end{array} \right. \\ \Lambda^{n+1} := \Lambda_{[k]}^{n+1}, \hat{\Lambda}^{n+1} := \hat{\Lambda}_{[k]}^{n+1}. \end{cases} \tag{5.82}$$

Numerical interpretation of $\tilde{\Lambda}$ in comparison with $\hat{\Lambda}$

Direct implementation of (5.80) is not practical. One should relate $\tilde{\Lambda}$ to some solutions of local problems posed on Ω_0 and Ω_1. Thanks to the definition of $\mathbb{B}$ in (5.63) and of $\mathbb{H}$ in (5.77), equation (5.80) can be explicitly written as

$$\begin{aligned} &\begin{bmatrix} I & -(I - 2\beta\tilde{R}_0\tilde{A}_0^{-1}\tilde{M}_{0;\Sigma}^\intercal) \\ -(I - 2\beta\tilde{R}_1\tilde{A}_1^{-1}\tilde{M}_{1;\Sigma}^\intercal) & I \end{bmatrix} \tilde{\Lambda} \\ &= \begin{bmatrix} \tilde{M}_{0;\Sigma}\tilde{A}_0^{-1}\tilde{H}^\intercal \tilde{M}_\Gamma \tilde{H}\tilde{A}_0^{-1}\tilde{M}_{0;\Sigma}^\intercal & 0 \\ 0 & 0 \end{bmatrix} \Lambda + \begin{bmatrix} \tilde{M}_{0;\Sigma}\tilde{A}_0^{-1}\tilde{H}^\intercal \tilde{M}_\Gamma G \\ 0 \end{bmatrix}, \end{aligned}$$

which yields

$$\tilde{\Lambda}_0 - (I - 2\beta\tilde{R}_0\tilde{A}_0^{-1}\tilde{M}_{0;\Sigma}^{\intercal})\tilde{\Lambda}_1 = \tilde{M}_{0;\Sigma}\tilde{A}_0^{-1}\tilde{H}^{\intercal}\tilde{M}_{\Gamma}\tilde{H}\tilde{A}_0^{-1}\tilde{M}_{0;\Sigma}^{\intercal}\Lambda_0 + \tilde{M}_{0;\Sigma}\tilde{A}_0^{-1}\tilde{H}^{\intercal}\tilde{M}_{\Gamma}G \quad (5.83)$$

and

$$-(I - 2\beta\tilde{R}_1\tilde{A}_1^{-1}\tilde{M}_{1;\Sigma}^{\intercal})\tilde{\Lambda}_0 + \tilde{\Lambda}_1 = 0. \quad (5.84)$$

We first work with the simpler equation (5.84). Using the definition of $\hat{\mathcal{A}}_1$ in (5.60), we can rewrite (5.84) as

$$\tilde{\Lambda}_1 = \tilde{\Lambda}_0 - 2\beta\tilde{R}_1\tilde{A}_1^{-1}\tilde{M}_{1;\Sigma}^{\intercal}\tilde{\Lambda}_0 = \tilde{\Lambda}_0 - 2\beta\tilde{R}_1\hat{\mathcal{A}}_1(\tilde{\Lambda}_0).$$

Then, using the definition of $\hat{\mathcal{A}}_0$ in (5.59) and of $\hat{\mathcal{A}}_0$ in (5.59), we can rewrite (5.83) as

$$\begin{aligned}\tilde{\Lambda}_0 &= \tilde{\Lambda}_1 - 2\beta\tilde{R}_0\tilde{A}_0^{-1}\tilde{M}_{0;\Sigma}^{\intercal}\tilde{\Lambda}_1 - \tilde{M}_{0;\Sigma}\tilde{A}_0^{-1}\tilde{H}^{\intercal}\tilde{M}_{\Gamma}(-\tilde{H}\tilde{A}_0^{-1}\tilde{M}_{0;\Sigma}^{\intercal}\Lambda_0 - G)\\ &= \tilde{\Lambda}_1 + 2\beta\tilde{R}_0\mathcal{A}_0(\tilde{\Lambda}_1) - \tilde{M}_{0;\Sigma}\tilde{A}_0^{-1}\tilde{H}^{\intercal}\tilde{M}_{\Gamma}(\tilde{H}\mathcal{A}_0(\Lambda_0) - G)\\ &= \tilde{\Lambda}_1 + 2\beta\tilde{R}_0\mathcal{A}_0(\tilde{\Lambda}_1) - \tilde{M}_{0;\Sigma}\hat{\mathcal{A}}_0(0, \tilde{H}\mathcal{A}_0(\Lambda_0) - G).\end{aligned}$$

In summary, we obtain

$$\begin{cases}\tilde{\Lambda}_1 = \tilde{\Lambda}_0 - 2\beta\tilde{R}_1\hat{\mathcal{A}}_1(\tilde{\Lambda}_0),\\ \tilde{\Lambda}_0 = \tilde{\Lambda}_1 + 2\beta\tilde{R}_0\mathcal{A}_0(\tilde{\Lambda}_1) - \tilde{M}_{0;\Sigma}\hat{\mathcal{A}}_0(0, \tilde{H}\mathcal{A}_0(\Lambda_0) - G).\end{cases} \quad (5.85)$$

In the following we shall derive the equations (5.87) for $\hat{\Lambda} = (\hat{\Lambda}_0, \hat{\Lambda}_1)$, the impedance values associated with the adjoint problem (5.6), and compare these equations with (5.85). Indeed, using the notations introduced in Section 5.2.1, the discretized version of the iterative scheme (5.41) is written as

$$\begin{cases}\Lambda_{[\ell+1]} = \mathcal{A}(\Lambda_{[\ell]}, S_1),\\ \hat{\Lambda}_{[\ell+1]} = \hat{\mathcal{A}}(\hat{\Lambda}_{[\ell]}, \tilde{H}\mathcal{A}_0(\Lambda_{0;\ell}) - G)\end{cases}$$

where $\mathcal{A}_0, \mathcal{A}, \hat{\mathcal{A}}$ are respectively defined in (5.57), (5.62) and (5.66). Taking $\ell \to \infty$ in this system yields

$$\begin{cases}\Lambda = \mathcal{A}(\Lambda, S_1),\\ \hat{\Lambda} = \hat{\mathcal{A}}(\hat{\Lambda}, \tilde{H}\mathcal{A}_0(\Lambda_0) - G)\end{cases}$$

where $\Lambda = (\Lambda_0, \Lambda_1)$ and $\hat{\Lambda} = (\hat{\Lambda}_0, \hat{\Lambda}_1)$ are respectively the impedance values associated with the forward and adjoint problems (5.1), (5.6). In particular, the second equation in (5.2.3), that is,

$$\hat{\Lambda} = \hat{\mathcal{A}}(\hat{\Lambda}, \tilde{H}\mathcal{A}_0(\Lambda_0) - G) \quad (5.86)$$

can be more explicitly written as

$$\begin{cases}\hat{\Lambda}_0 = \hat{\Lambda}_1 - 2\beta\tilde{R}_1\hat{\mathcal{A}}_1(\hat{\Lambda}_1),\\ \hat{\Lambda}_1 = \hat{\Lambda}_0 + 2\beta\tilde{R}_0\hat{\mathcal{A}}_0(\hat{\Lambda}_0, \tilde{H}\mathcal{A}_0(\Lambda_0) - G).\end{cases} \quad (5.87)$$

We clearly observe that the role of $\tilde{\Lambda}_0$ in (5.85) is played by $\hat{\Lambda}_1$ in (5.87) and the role of $\tilde{\Lambda}_1$ in (5.85) is played by $\hat{\Lambda}_0$ in (5.87). This means that one would obtain a rewriting of the scheme (5.85) in a form similar to the abstract framework in Section 1.2 if we phrase

it in terms of variables (Λ_0, Λ_1) and $(\tilde{\Lambda}_1, \tilde{\Lambda}_0)$. Let us set $\bar{\Lambda} \coloneqq (\tilde{\Lambda}_1, \tilde{\Lambda}_0)$ then scheme (5.85) is equivalent to

$$\bar{\Lambda} = \bar{\mathcal{A}}(\bar{\Lambda}, \tilde{H}\mathcal{A}_0(\Lambda_0) - G) \tag{5.88}$$

where $\bar{\mathcal{A}} : \{\mathbb{R}^{n_\Sigma} \times \mathbb{R}^{n_\Sigma}\} \times \mathbb{R}^{n_\Gamma} \to \{\mathbb{R}^{n_\Sigma} \times \mathbb{R}^{n_\Sigma}\}$ is defined by

$$\bar{\mathcal{A}}(\mathbf{\Lambda}, \Psi) \coloneqq (\mathbf{\Lambda_1} - 2\beta\tilde{R}_1\hat{\mathcal{A}}_1(\mathbf{\Lambda_1}), \mathbf{\Lambda_0} + 2\beta\tilde{R}_0\mathcal{A}_0(\mathbf{\Lambda_0}) - \tilde{M}_{0;\Sigma}\hat{\mathcal{A}}_0(0, \Psi)), \tag{5.89}$$

with $\mathbf{\Lambda} = (\mathbf{\Lambda_0}, \mathbf{\Lambda}_1) \in \{\mathbb{R}^{n_\Sigma} \times \mathbb{R}^{n_\Sigma}\}$ and $\Psi \in \mathbb{R}^{n_\Gamma}$. This is to compare with

$$\hat{\mathcal{A}}(\mathbf{\Lambda}, \Psi) \coloneqq (\mathbf{\Lambda_1} - 2\beta\tilde{R}_1\hat{\mathcal{A}}_1(\mathbf{\Lambda_1}), \mathbf{\Lambda_0} + 2\beta\tilde{R}_0\hat{\mathcal{A}}_0(\mathbf{\Lambda_0}, \Psi)). \tag{5.65 recalled}$$

We can see that although the two equations (5.88) and (5.86) are similar, the operator $\bar{\mathcal{A}}$ defined in (5.89) does not exactly match the operator $\hat{\mathcal{A}}$ defined in (5.65).

Appendix 5.A Convergence analysis of (5.41) in the case of a circular domain

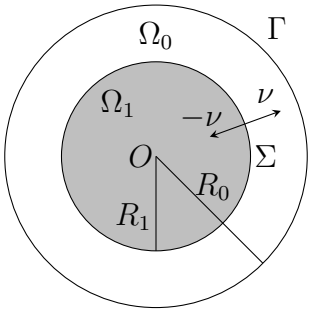


Figure 5.2: Illustration for the domain Ω and its decomposition.

We investigate in the following the convergence of (5.41) in the case of a circular domain. Let (r, θ) be the polar coordinates in $\mathbb{R}^2$. Given $m \in \mathbb{N}^*$, we define

$$V_m(R) := \{v : \partial B(0,R) \to \mathbb{C} \text{ such that } v(\theta) = c_v \mathrm{e}^{\mathrm{i}m\theta}, \forall \theta \text{ for some constant } c_v\}, \ \forall R > 0.$$

We set the following assumptions.

Assumption 5.A.1. Ω and Ω_1 are respectively the circles of radii R_0 and R_1, and both are centered at the origin (see Figure 5.2). $\Omega_0 = \Omega \backslash \overline{\Omega}_1$ is the corresponding annulus.

Assumption 5.A.2. $\sigma_0 > 0$ and $\sigma = \sigma_1 \mathbb{1}_{\Omega_1}$ where $\sigma_1 \in \mathbb{R}, \sigma_1 > 0$.

Assumption 5.A.3. The boundary datum f in the equation of the incident field (5.2) belongs to $V_m(R_0)$ and the initial guesses $\lambda_{0;0}, \lambda_{1;0}, \hat{\lambda}_{0;0}, \hat{\lambda}_{1;0}$ of algorithm (5.41) also belong to $V_m(R_1)$, for some $m \in \mathbb{N}^*$.

Let J_α and Y_α respectively denote the α-order Bessel functions of the first and second kinds. By Assumption 5.A.3, given two constants $\sigma_0 > 0$ and $\eta \in \mathbb{R}$, we are interested in the solutions of

$$-\operatorname{div}(\sigma_0 \nabla u) + \eta u = 0$$

in the form

$$(\mathcal{C}_1 A_m(r) + \mathcal{C}_2 B_m(r)) \mathrm{e}^{\mathrm{i}m\theta} \quad (\mathcal{C}_1 \text{ and } \mathcal{C}_2 \text{ are constants}), \tag{5.90}$$

where the functions A_m and B_m are defined as follows: let $\tilde{k} := \sqrt{\frac{|\eta|}{\sigma_0}}$ then

- if $\eta < 0$ (the case of the Helmholtz equation),

$$A_m(r) = J_m(\tilde{k}r) \text{ and } B_m(r) = Y_m(\tilde{k}r),$$

- if $\eta = 0$,

$$A_m(r) = r^m \text{ and } B_m(r) = r^{-m},$$

- if $\eta > 0$ (the coercive case),

$$A_m(r) = J_m(\mathrm{i}\tilde{k}r) \text{ and } B_m(r) = Y_m(\mathrm{i}\tilde{k}r).$$

We note that B_m is not regular at 0 in all the cases. We shall use the solution form (5.90) to give the expressions of the operators $\mathcal{A}$ and $\hat{\mathcal{A}}$ that play an important role in algorithm (5.41).

First, we give the expression of the incident field. We make Assumptions 5.A.1–5.A.3 and let $f = c_f \mathrm{e}^{\mathrm{i}m\theta} \in V_m(R_0)$. It is easy to verify that the incident field u^{inc} defined by

$$\begin{cases} -\operatorname{div}(\sigma_0 \nabla u^{\mathrm{inc}}) + \eta u^{\mathrm{inc}} = 0 & \text{in } \Omega, \\ \sigma_0 \dfrac{\partial u^{\mathrm{inc}}}{\partial \nu} = f & \text{on } \partial\Omega \end{cases} \qquad \text{(5.2 recalled)}$$

is expressed as

$$\boxed{u^{\mathrm{inc}}(r,\theta) = \frac{A_m(r)}{\sigma_0 A_m'(R_0)} c_f \mathrm{e}^{\mathrm{i}m\theta}.} \tag{5.91}$$

We have the following lemma.

Lemma 5.A.4. *We make Assumptions 5.A.1–5.A.3. Given $\eta \in \mathbb{R}$, we consider the equation*

$$-\operatorname{div}(\sigma_0 \nabla u) + \eta u = -\operatorname{div}(\sigma \nabla u^{\mathrm{inc}}) \text{ in } \Omega \tag{5.92}$$

where u^{inc} is given by (5.91). We define the radial function $\tilde{w}_m$ as follows:

$$\tilde{w}_m(r) = -\frac{\mathrm{i}\pi \tilde{k} c_f}{4\sigma_0^2 J_m'(\tilde{k}R_0)} \left(H_m^{(1)}(\tilde{k}r) \int_0^r J_m(\tilde{k}t)^2 t \,\mathrm{d}t + J_n(\tilde{k}r) \int_r^{R_1} H_n^{(1)}(\tilde{k}t) J_m(\tilde{k}t) t \,\mathrm{d}t \right)$$

if $r < R_1$, and

$$\tilde{w}_m(r) = -\frac{\mathrm{i}\pi \tilde{k} c_f}{4\sigma_0^2 J_m'(\tilde{k}R_0)} H_m^{(1)}(\tilde{k}r) \int_0^{R_1} J_m(\tilde{k}t)^2 t \,\mathrm{d}t$$

if $R_1 < r < R_0$. Then a particular solution $u^{\mathrm{part}} = u^{\mathrm{part}}(r,\theta;\sigma_1)$ of (5.92) is given by

$$u^{\mathrm{part}}(r,\theta;\sigma_1) = \sigma_1 w_m(r) \mathrm{e}^{\mathrm{i}m\theta} \tag{5.93}$$

where

$$w_m(r) \coloneqq \begin{cases} \tilde{w}_m(r) & \text{if } \eta < 0, \\ 0 & \text{if } \eta = 0, \\ \tilde{w}_m(\mathrm{i}r) & \text{if } \eta > 0. \end{cases} \tag{5.94}$$

Proof. We first study the case $\eta = 0$. Indeed, for $\eta = 0$, by definition (5.2) the incident field u^{inc} verifies

$$-\operatorname{div}(\sigma_0 \nabla u^{\mathrm{inc}}) = 0,$$

which yields $\triangle u^{\mathrm{inc}} = 0$. It is easy to check that equation (5.92) then becomes

$$-\operatorname{div}(\sigma_0 \nabla u) = 0 \text{ in } \Omega,$$

hence we can take $u^{\mathrm{part}} = 0$ as a particular solution.
We now study the case $\eta < 0$ (the case of the Helmholtz equation). Let us begin from a slightly more general context. We consider the Helmholtz equation

$$\triangle \tilde{u} + \tilde{k}^2 \tilde{u} = \tilde{f} \chi_{\Omega_1} \text{ in } \Omega \tag{5.95}$$

where $\tilde{k}$ is the wavenumber and $\tilde{f} = \tilde{f}(r,\theta) = CJ_m(\tilde{k}r)\mathrm{e}^{\mathrm{i}m\theta}$ for some constant C. We denote by $H_\alpha^{(1)}$ the α-order Hankel function of the first kind, then a particular solution of (5.95) is given by

$$\tilde{u}(x) = \int_{\Omega_1} \tilde{f}(y)\phi(x,y)\,\mathrm{d}y \tag{5.96}$$

where

$$\phi(x,y) \coloneqq -\frac{\mathrm{i}}{4}H_0^{(1)}(\tilde{k}|x-y|),\ (x,y) \in \mathbb{R}^2 \times \mathbb{R}^2$$

is the Green function. Moreover, let (r_x,θ_x) and (r_y,θ_y) respectively denote the polar coordinates of $x \in \mathbb{R}^2$ and $y \in \mathbb{R}^2$ then we can expand ϕ as

$$\phi(x,y) = -\frac{\mathrm{i}}{4}\left(H_0^{(1)}(\tilde{k}r_x)J_0(\tilde{k}r_y) + \frac{1}{2}\sum_{\substack{n\in\mathbb{Z}\\ n\neq 0}} H_n^{(1)}(\tilde{k}r_x)J_n(\tilde{k}r_y)\mathrm{e}^{\mathrm{i}n\theta_x}\mathrm{e}^{-\mathrm{i}n\theta_y}\right)$$

if $r_x > r_y$, and

$$\phi(x,y) = -\frac{\mathrm{i}}{4}\left(J_0(\tilde{k}r_x)H_0^{(1)}(\tilde{k}r_y) + \frac{1}{2}\sum_{\substack{n\in\mathbb{Z}\\ n\neq 0}} J_n(\tilde{k}r_x)H_n^{(1)}(\tilde{k}r_y)\mathrm{e}^{-\mathrm{i}n\theta_x}\mathrm{e}^{\mathrm{i}n\theta_y}\right)$$

if $r_x < r_y$. Before giving a more explicit expression for $\tilde{u}$ in (5.96), we recall that $\Omega_1 = B(0,R_1)$ and $\tilde{f}(r,\theta) = CJ_m(\tilde{k}r)\mathrm{e}^{\mathrm{i}m\theta}$. Now given x such that $r_x < R_1$. Since ϕ is not defined when $y = x$, we rewrite (5.96) as

$$\tilde{u}(x) = \int_{r_y<r_x} \tilde{f}(y)\phi(x,y)\,\mathrm{d}y + \int_{r_x<r_y<R_1} \tilde{f}(y)\phi(x,y)\,\mathrm{d}y\,,$$

which can be expanded as

$$\begin{aligned}
\tilde{u}(x) = \ & -\frac{\mathrm{i}C}{4}\Bigg(H_0^{(1)}(\tilde{k}r_x)\int_0^{r_x} J_0(\tilde{k}r_y)J_m(\tilde{k}r_y)r_y\,\mathrm{d}r_y\int_0^{2\pi}\mathrm{e}^{\mathrm{i}m\theta_y}\,\mathrm{d}\theta_y \\
& + \frac{1}{2}\sum_{\substack{n\in\mathbb{Z}\\ n\neq 0}} H_n^{(1)}(\tilde{k}r_x)\int_0^{r_x} J_n(\tilde{k}r_y)J_m(\tilde{k}r_y)r_y\,\mathrm{d}r_y\,\mathrm{e}^{\mathrm{i}n\theta_x}\int_0^{2\pi}\mathrm{e}^{\mathrm{i}(m-n)\theta_y}\,\mathrm{d}\theta_y\Bigg) \\
& -\frac{\mathrm{i}C}{4}\Bigg(J_0(\tilde{k}r_x)\int_{r_x}^{R_1} H_0^{(1)}(\tilde{k}r_y)J_m(\tilde{k}r_y)r_y\,\mathrm{d}r_y\int_0^{2\pi}\mathrm{e}^{\mathrm{i}m\theta_y}\,\mathrm{d}\theta_y \\
& + \frac{1}{2}\sum_{\substack{n\in\mathbb{Z}\\ n\neq 0}} J_n(\tilde{k}r_x)\int_{r_x}^{R_1} H_n^{(1)}(\tilde{k}r_y)J_m(\tilde{k}r_y)r_y\,\mathrm{d}r_y\,\mathrm{e}^{-\mathrm{i}n\theta_x}\int_0^{2\pi}\mathrm{e}^{\mathrm{i}(n+m)\theta_y}\,\mathrm{d}\theta_y\Bigg)
\end{aligned}$$

Since $\int_0^{2\pi}\mathrm{e}^{\mathrm{i}n\theta_y}\,\mathrm{d}\theta_y = 0$, $\forall n \in \mathbb{Z}\backslash\{0\}$, we obtain

$$\begin{aligned}
\tilde{u}(x) = \ & -\frac{\mathrm{i}C}{8}H_m^{(1)}(\tilde{k}r_x)\int_0^{r_x} J_m(\tilde{k}r_y)^2 r_y\,\mathrm{d}r_y\cdot 2\pi\mathrm{e}^{\mathrm{i}m\theta_x} \\
& -\frac{\mathrm{i}C}{8}J_n(\tilde{k}r_x)\int_{r_x}^{R_1} H_n^{(1)}(\tilde{k}r_y)J_m(\tilde{k}r_y)r_y\,\mathrm{d}r_y\cdot 2\pi\mathrm{e}^{\mathrm{i}m\theta_x} \\
= \ & -\frac{\mathrm{i}\pi C}{4}\left(H_m^{(1)}(\tilde{k}r_x)\int_0^{r_x} J_m(\tilde{k}r_y)^2 r_y\,\mathrm{d}r_y + J_n(\tilde{k}r_x)\int_{r_x}^{R_1} H_n^{(1)}(\tilde{k}r_y)J_m(\tilde{k}r_y)r_y\,\mathrm{d}r_y\right)\mathrm{e}^{\mathrm{i}m\theta_x}.
\end{aligned}$$

Similarly, if $R_1 < r_x < R_0$, we have

$$\tilde{u}(x) = -\frac{\mathrm{i}\pi C}{4}H_m^{(1)}(\tilde{k}r_x)\int_0^{R_1} J_m(\tilde{k}r_y)^2 r_y\,\mathrm{d}r_y\,\mathrm{e}^{\mathrm{i}m\theta_x}.$$

Therefore, we conclude that

$$\tilde{u}(r,\theta) = \frac{\mathrm{i}\pi C}{4}\left(H_m^{(1)}(\tilde{k}r)\int_0^r J_m(\tilde{k}t)^2 t\,\mathrm{d}t + J_n(\tilde{k}r)\int_r^{R_1} H_n^{(1)}(\tilde{k}t)J_m(\tilde{k}t)t\,\mathrm{d}t\right)\mathrm{e}^{\mathrm{i}m\theta} \tag{5.97}$$

if $r < R_1$, and

$$\tilde{u}(r,\theta) = -\frac{\mathrm{i}\pi C}{4}H_m^{(1)}(\tilde{k}t)\int_0^{R_1} J_m(\tilde{k}t)^2 t\,\mathrm{d}t\,\mathrm{e}^{\mathrm{i}m\theta} \tag{5.98}$$

if $R_1 < r < R_0$, is a particular solution to (5.95).
Now we go back to equation (5.92). For $\eta < 0$, we have $A_m(r) = J_m(\tilde{k}r)$. By (5.91), the incident field u^{inc} in this case is given by

$$u^{\mathrm{inc}}(r,\theta) = \frac{c_f}{\sigma_0\tilde{k}J_m'(\tilde{k}R_0)}J_m(\tilde{k}r)\mathrm{e}^{\mathrm{i}m\theta} \text{ with } \tilde{k} = \sqrt{\frac{|\eta|}{\sigma_0}}, \tag{5.99}$$

and verifies

$$-\operatorname{div}(\nabla u^{\mathrm{inc}}) = \tilde{k}^2 u^{\mathrm{inc}} \text{ in } \Omega.$$

Thus, we can rewrite (5.92) as

$$-\operatorname{div}(\nabla u) - \frac{\omega^2}{\sigma_0}u = \begin{cases} 0 & \text{in } \Omega_0, \\ -\dfrac{\sigma_1}{\sigma_0}\operatorname{div}(\nabla u^{\mathrm{inc}}) = \sigma_1\dfrac{\tilde{k}^2}{\sigma_0}u^{\mathrm{inc}} & \text{in } \Omega_1, \end{cases}$$

or equivalently,

$$\triangle u^{\mathrm{inc}} + \tilde{k}^2 u^{\mathrm{inc}} = -\sigma_1\frac{\tilde{k}^2}{\sigma_0}u^{\mathrm{inc}}\chi_{\Omega_1} \text{ in } \Omega.$$

We therefore apply (5.97) and (5.98) for

$$\tilde{f} = -\sigma_1\frac{\tilde{k}^2}{\sigma_0}u^{\mathrm{inc}} \overset{(5.99)}{=} CJ_m(\tilde{k}r)\mathrm{e}^{\mathrm{i}m\theta} \text{ with } C = -\sigma_1\frac{\tilde{k}c_f}{\sigma_0^2 J_m'(\tilde{k}R_0)}$$

to obtain the conclusion of the lemma.
The case $\eta > 0$ can be deduced from the case $\eta < 0$ by a change of variables $x_{\mathrm{new}} = \mathrm{i}x$, $x \in \mathbb{R}^2$. □

Now we are ready for giving the expressions of $\mathcal{A}_i$.

Lemma 5.A.5. *We make Assumptions 5.A.1–5.A.3 and given $\eta \in \mathbb{R}$.*

(i) The operator $\hat{\mathcal{A}}_0$ defined in (5.34) can be expressed as

$$\hat{\mathcal{A}}_0(\lambda,\psi)(r,\theta) = (C_1(r;\beta)c_\lambda + C_2(r;\beta)c_\psi)\mathrm{e}^{\mathrm{i}m\theta} \tag{5.100}$$

for $\lambda = c_\lambda\mathrm{e}^{\mathrm{i}m\theta} \in V_m(R_1)$ and $\psi = c_\psi\mathrm{e}^{\mathrm{i}m\theta} \in V_m(R_0)$, where

$$C_1(r;\beta) := \frac{A_m(r)}{(\sigma_0 B_m'(R_1) - \beta B_m'(R_1))D(\beta)} - \frac{A_m'(R_0)B_m(r)}{B_m'(R_0)(\sigma_0 B_m'(R_1) - \beta B_m'(R_1))D(\beta)}, \tag{5.101}$$

$$C_2(r;\beta) := \frac{-A_m(r)}{\sigma_0 B_m'(R_0)D(\beta)} + \left(\frac{1}{\sigma_0 B_m'(R_0)} + \frac{A_m'(R_0)}{\sigma_0 B_m'(R_0)^2 D(\beta)}\right)B_m(r)$$

and

$$D(\beta) \coloneqq \frac{\sigma_0 A'_m(R_1) - \beta A_m(R_1)}{\sigma_0 B'_m(R_1) - \beta B_m(R_1)} - \frac{A'_m(R_0)}{B'_m(R_0)}. \tag{5.102}$$

In particular, the operator $\mathcal{A}_0$ *defined in* (5.20) *can be expressed as*

$$\mathcal{A}_0(\lambda)(r,\theta) = \hat{\mathcal{A}}_0(\lambda,0)(r,\theta) = C_1(r;\beta)c_\lambda \mathrm{e}^{\mathrm{i}m\theta} \tag{5.103}$$

for $\lambda = c_\lambda \mathrm{e}^{\mathrm{i}m\theta} \in V_m(R_1)$.

(ii) The operator $\mathcal{A}_1$ *defined in* (5.21) *can be expressed as*

$$\mathcal{A}_1(\lambda,\sigma_1)(r,\theta) = (C_3(r;\beta)c_\lambda + C_4(r;\beta)\sigma_1)\mathrm{e}^{\mathrm{i}m\theta} \tag{5.104}$$

for $\lambda = c_\lambda \mathrm{e}^{\mathrm{i}m\theta} \in V_m(R_1)$, *where*

$$C_3(r;\beta) \coloneqq \frac{A_m(r)}{\sigma_0 A'_m(R_0) + \beta A_m(R_0)} \tag{5.105}$$

and

$$C_4(r;\beta) \coloneqq \frac{\sigma_0 w'_m(R_0) + \beta w_m(R_0) - \frac{c_f}{\sigma_0}}{\sigma_0 A'_m(R_0) + \beta A_m(R_0)} A_m(r) + w_m(r),$$

with w_m *defined in* (5.94).

In particular, the operator $\hat{\mathcal{A}}_1$ *defined in* (5.35) *can be expressed as*

$$\hat{\mathcal{A}}_1(\lambda)(r,\theta) = \mathcal{A}_1(\lambda,0)(r,\theta) = C_3(r;\beta)c_\lambda \mathrm{e}^{\mathrm{i}m\theta} \tag{5.106}$$

for $\lambda = c_\lambda \mathrm{e}^{\mathrm{i}m\theta} \in V_m(R_1)$.

(iii) The operator $\mathcal{A}$ *defined in* (5.25) *can be expressed as*

$$\mathcal{A}(\boldsymbol{\lambda},\sigma_1) = \mathcal{B}(\beta)\boldsymbol{\lambda} + \mathcal{M}(\beta)\sigma_1 \mathrm{e}^{\mathrm{i}m\theta} \tag{5.107}$$

for $\boldsymbol{\lambda} \in V_m(R_1)^2$, *where*

$$\mathcal{B}(\beta) \coloneqq \begin{bmatrix} 0 & 1 - 2\beta C_3(R_1;\beta) \\ 1 + 2\beta C_1(R_1;\beta) & 0 \end{bmatrix} \tag{5.108}$$

and

$$\mathcal{M}(\beta) \coloneqq \begin{bmatrix} 0 \\ -2\beta C_4(R_1;\beta) \end{bmatrix}.$$

(iv) The operator $\hat{\mathcal{A}}$ *defined in* (5.39) *can be expressed as*

$$\mathcal{A}(\boldsymbol{\lambda},\psi) = \mathcal{B}(\beta)\boldsymbol{\lambda} + \mathcal{E}(\beta)c_\psi \mathrm{e}^{\mathrm{i}m\theta}, \tag{5.109}$$

for $\boldsymbol{\lambda} \in V_m(R_1)^2$ *and* $\psi = c_\psi \mathrm{e}^{\mathrm{i}m\theta} \in V_m(R_0)$, *where* $\mathcal{B}(\beta)$ *is defined in* (5.108) *and*

$$\mathcal{E}(\beta) \coloneqq \begin{bmatrix} 0 \\ 2\beta C_2(R_1;\beta) \end{bmatrix}.$$

Moreover, the sequences $(\lambda_{[\ell]})_{\ell\geq 0}$ *and* $(\hat{\lambda}_{[\ell]})_{\ell\geq 0}$ *defined by algorithm* (5.41) *are in* $V_m(R_1)^2$, *also* $\gamma_\Gamma \mathcal{A}_0(\lambda_{0;\ell}) \in V_m(R_0), \forall \ell \geq 0$.

Proof. *(i)* We recall that

$$\hat{\mathcal{A}}_0 : (\lambda, \psi) \in L^2(\Sigma) \times H^{-\frac{1}{2}}(\Gamma) \mapsto u \in H^1(\Omega_0)$$

such that

$$\begin{cases} -\operatorname{div}(\sigma_0 \nabla u) + \eta u = 0 & \text{in } \Omega_0, \\ \sigma_0 \dfrac{\partial u}{\partial \nu} = \psi & \text{on } \Gamma, \\ \sigma_0 \dfrac{\partial u}{\partial \nu} - \beta u = \lambda & \text{on } \Sigma. \end{cases} \qquad \text{(5.34 recalled)}$$

We seek the solution to (5.34) of the form

$$u = (c_1 A_m(r) + c_2 B_m(r)) \mathrm{e}^{\mathrm{i}m\theta}. \tag{5.110}$$

The boundary conditions on Γ and Σ in (5.34) give

$$\begin{cases} \sigma_0(c_1 A'_m(R_0) + c_2 B'_m(R_0)) = c_\psi, \\ \sigma_0(c_1 A'_m(R_1) + c_2 B'_m(R_1)) - \beta(c_1 A_m(R_1) + c_2 B_m(R_1)) = c_\lambda. \end{cases}$$

The first line of this system yields

$$c_2 = -\frac{A'_m(R_0)}{B'_m(R_0)} c_1 + \frac{c_\psi}{\sigma_0 B'_m(R_0)}, \tag{5.111}$$

then inserting this result into the second line and dividing by $\sigma_0 B'_m(R_1) - \beta B_m(R_1) \neq 0$ gives

$$D(\beta) c_1 + \frac{c_\psi}{\sigma_0 B'_m(R_0)} = \frac{c_\lambda}{\sigma_0 B'_m(R_1) - \beta B_m(R_1)},$$

where

$$D(\beta) := \frac{\sigma_0 A'_m(R_1) - \beta A_m(R_1)}{\sigma_0 B'_m(R_1) - \beta B_m(R_1)} - \frac{A'_m(R_0)}{B'_m(R_0)}.$$

We thus obtain

$$c_1 = \frac{c_\lambda}{(\sigma_0 B'_m(R_1) - \beta B'_m(R_1)) D(\beta)} - \frac{c_\psi}{\sigma_0 B'_m(R_0) D(\beta)}.$$

Inserting this expression of c_2 into (5.111) leads to

$$\begin{aligned} c_2 = & -\frac{A'_m(R_0)}{B'_m(R_0)} c_1 + \frac{c_\psi}{\sigma_0 B'_m(R_0)} \\ = & -\frac{A'_m(R_0)}{B'_m(R_0)(\sigma_0 B'_m(R_1) - \beta B'_m(R_1)) D(\beta)} c_\lambda + \left(\frac{1}{\sigma_0 B'_m(R_0)} + \frac{A'_m(R_0)}{\sigma_0 B'_m(R_0)^2 D(\beta)} \right) c_\psi. \end{aligned}$$

Plugging the above results of c_1 and c_2 into (5.110), we obtain the result of the lemma.
(ii) For $\sigma = \sigma_1$ constant in Ω_1, we recall that

$$\mathcal{A}_1 : (\lambda, \sigma_1) \in L^2(\Sigma) \times L^\infty(\Omega_1) \mapsto u \in H^1(\Omega_1)$$

such that

$$\begin{cases} -\operatorname{div}(\sigma_0 \nabla u) + \eta u = -\operatorname{div}(\sigma_1 \nabla u^{\mathrm{inc}}) & \text{in } \Omega_1, \\ \sigma_0 \dfrac{\partial u}{\partial \nu} - \sigma_1 \dfrac{\partial u^{\mathrm{inc}}}{\partial \nu} + \beta u = \lambda & \text{on } \Sigma. \end{cases} \qquad \text{(5.21 recalled)}$$

Let $v \coloneqq u - u^{\text{part}}$ where u^{part} is defined in (5.93), then v verifies

$$\begin{cases} -\operatorname{div}(\sigma_0 \nabla v) + \eta v = 0 & \text{in } \Omega_1, \\ \sigma_0 \dfrac{\partial v}{\partial \nu} + \beta v + \sigma_0 \dfrac{\partial u^{\text{part}}}{\partial \nu} + \beta u^{\text{part}} - \sigma_1 \dfrac{\partial u^{\text{inc}}}{\partial \nu} = \lambda & \text{on } \Sigma. \end{cases}$$

Since $\lambda = c_\lambda \mathrm{e}^{\mathrm{i}m\theta} \in V_m(R_1)$, we seek v of the form $v = cA_m(r)\mathrm{e}^{\mathrm{i}m\theta}$. Using the expressions of u^{part} in (5.93) and u^{inc} in (5.91), the boundary condition on Σ in (5.A) gives

$$\sigma_0 c A_m'(R_0) + \beta c A_m(R_0) + \sigma_0\sigma_1 w_m'(R_0) + \beta\sigma_1 w_m(R_0) - \sigma_1 \frac{c_f}{\sigma_0} = c_\lambda.$$

Hence,

$$c = \frac{1}{\sigma_0 A_m'(R_0) + \beta A_m(R_0)} \left(\left(\sigma_0 w_m'(R_0) + \beta w_m(R_0) - \frac{c_f}{\sigma_0} \right) \sigma_1 + c_\lambda \right).$$

Computing $u = v + u^{\text{part}} = (cA_m(r) + \sigma_1 w_m(R))\mathrm{e}^{\mathrm{i}m\theta}$ then gives the result of the lemma.
(iii) For $\sigma = \sigma_1$ constant in Ω_1, we recall that $\mathcal{A} : L^2(\Sigma)^2 \times L^\infty(\Omega_1) \to L^2(\Sigma)^2$ is defined by

$$\mathcal{A}(\boldsymbol{\lambda}, \sigma_1) \coloneqq (\boldsymbol{\lambda_1} - 2\beta\gamma_{\Omega_1,\Sigma}\mathcal{A}_1(\boldsymbol{\lambda_1}, \sigma_1), \boldsymbol{\lambda_0} + 2\beta\gamma_{\Omega_0,\Sigma}\mathcal{A}_0(\boldsymbol{\lambda_0})), \qquad \text{(5.25 recalled)}$$

where $\boldsymbol{\lambda} = (\boldsymbol{\lambda_0}, \boldsymbol{\lambda_1}) \in L^2(\Sigma)^2$. Assume that $\boldsymbol{\lambda_i} = c_{\boldsymbol{\lambda_i}}\mathrm{e}^{\mathrm{i}m\theta} \in V_m(R_1)$, $i = 0, 1$. By the expressions of $\mathcal{A}_1$ and $\mathcal{A}_0$ respectively in (5.104) and (5.103), we have

$$\boldsymbol{\lambda_1} - 2\beta\gamma_{\Omega_1,\Sigma}\mathcal{A}_1(\boldsymbol{\lambda_1}, \sigma_1) = (1 - 2\beta C_3(R_1;\beta))\boldsymbol{\lambda_1} - 2\beta C_4(R_1;\beta)\sigma_1 \mathrm{e}^{\mathrm{i}m\theta}$$

and

$$\boldsymbol{\lambda_0} + 2\beta\gamma_{\Omega_0,\Sigma}\mathcal{A}_0(\boldsymbol{\lambda_0}) = (1 + 2\beta C_1(R_1;\beta))\boldsymbol{\lambda_0},$$

which provides the result of the lemma.
(iv) We recall that $\hat{\mathcal{A}} : L^2(\Sigma)^2 \times H^{-\frac{1}{2}}(\Gamma) \to L^2(\Sigma)^2$ is defined by

$$\hat{\mathcal{A}}(\boldsymbol{\lambda}, \psi) \coloneqq (\boldsymbol{\lambda_1} - 2\beta\gamma_{\Omega_1,\Sigma}\hat{\mathcal{A}}_1(\boldsymbol{\lambda_1}), \boldsymbol{\lambda_0} + 2\beta\gamma_{\Omega_0,\Sigma}\hat{\mathcal{A}}_0(\boldsymbol{\lambda_0}, \psi)), \qquad \text{(5.39 recalled)}$$

where $\boldsymbol{\lambda} = (\boldsymbol{\lambda_0}, \boldsymbol{\lambda_1}) \in L^2(\Sigma)^2$ and $\psi \in H^{-\frac{1}{2}}(\Gamma)$. Assume that $\boldsymbol{\lambda_i} = c_{\boldsymbol{\lambda_i}}\mathrm{e}^{\mathrm{i}m\theta} \in V_m(R_1)$, $i = 0, 1$ and recall that $\psi = c_\psi \mathrm{e}^{\mathrm{i}m\theta} \in V_m(R_0)$. By the expressions of $\hat{\mathcal{A}}_1$ and $\hat{\mathcal{A}}_0$ respectively in (5.106) and (5.100), we have

$$\boldsymbol{\lambda_1} - 2\beta\gamma_{\Omega_1,\Sigma}\hat{\mathcal{A}}_1(\boldsymbol{\lambda_1}) = (1 - 2\beta C_3(R_1;\beta))\boldsymbol{\lambda_1}$$

and

$$\boldsymbol{\lambda_0} + 2\beta\gamma_{\Omega_0,\Sigma}\hat{\mathcal{A}}_0(\boldsymbol{\lambda_0}, \psi) = (1 + 2\beta C_1(R_1;\beta))\boldsymbol{\lambda_0} + 2\beta C_2(R_1;\beta)c_\psi \mathrm{e}^{\mathrm{i}m\theta},$$

which provides the result of the lemma.
The last conclusion of the lemma is immediate thanks to the expressions of $\mathcal{A}$, $\hat{\mathcal{A}}$ and $\mathcal{A}_0$ respectively in (5.107), (5.109) and (5.103). □

In the case $\eta = 0$, we have the following result about the spectral radius of $\mathcal{B}(\beta)$.

Lemma 5.A.6. *In the case $\eta = 0$, the matrix $\mathcal{B}(\beta)$ defined in (5.108) has spectral radius less than 1 if $\beta > 0$ is sufficiently small.*

Proof. In the case $\eta = 0$, we recall that $A_m(r) = r^m$ and $B_m(r) = r^{-m}$. Let

$$c_3 \coloneqq \lim_{\beta \to 0^+} C_3(R_1; \beta), d \coloneqq \lim_{\beta \to 0^+} D(\beta) \text{ and } c_1 \coloneqq \lim_{\beta \to 0^+} C_1(R_1; \beta).$$

By (5.105), (5.102) and (5.101), we respectively have

$$c_3 = \frac{A_m(R_1)}{\sigma_0 A'_m(R_0)} = \frac{R_1^m}{\sigma_0 m R_0^{m-1}} > 0,$$

$$d = \frac{A'_m(R_1)}{B'_m(R_1)} - \frac{A'_m(R_0)}{B'_m(R_0)} = \frac{m R_1^{m-1}}{-m R_1^{-m-1}} - \frac{m R_0^{m-1}}{-m R_0^{-m-1}} = R_0^{2m} - R_1^{2m} > 0$$

and

$$\begin{aligned} c_1 = & \frac{A_m(R_1)}{\sigma_0 B'_m(R_1) d} - \frac{A'_m(R_0) B_m(R_1)}{B'_m(R_1) \sigma_0 B'_m(R_0) d} \\ = & \frac{R_1^m}{-\sigma_0 m R_1^{-m-1} d} - \frac{m R_0^{m-1} R_1^{-m}}{(-m) R_0^{-m-1} \sigma_0 (-m) R_1^{-m-1} d} \\ = & -\frac{R_1^{2m+1} + R_0^{2m} R_1}{\sigma_0 m d} < 0. \end{aligned}$$

Therefore, the matrix $\mathcal{B}(\beta)$ defined in (5.108) can be written as

$$\mathcal{B}(\beta) = \begin{bmatrix} 0 & 1 - 2\beta c_3 + O(\beta^2) \\ 1 + 2\beta c_1 + O(\beta^2) & 0 \end{bmatrix}$$

where $c_3 > 0$ and $c_1 < 0$, which yields the conclusion of the lemma. □

We conclude that in the case $\eta = 0$, there exists a sufficiently small $\beta > 0$ such that the two assumptions in Theorem 5.1.2 are both satisfied. Indeed, assumptions *(i)* and *(ii)* are respectively verified by Lemmas 5.A.5 and 5.A.6. Therefore, algorithm (5.41) is convergent.

Chapter 6

Combination of multi-step one-shot and domain decomposition methods for non-linear inverse problems

Contents

In the previous chapter, we already presented the application of multi-step one-shot inversion methods and domain decomposition methods to the linearized inverse conductivity problem. In the same spirit, we shall study in this chapter the non-linear inverse conductivity problem. We still treat the case of Neumann forward problems, and consider the case where the domain is splitted into two subdomains on which we will apply the domain decomposition method and the unknown conductivity is located in the inner subdomain. Similarly as in Chapter 5, we can have some variants of the algorithms depending on the choice of variables of the problems: either the interface values of the impedance parameters or the state variables inside the domain. We furthermore consider the non-linear problem at the discretized level and investigate these algorithms in this setting. The material presented in this part follows the one in Chapter 5. We mainly repeat the same arguments and choose to repeat the details to make the chapter (in most parts) self contained and easier to read. Indeed the important difference is that the iterative operator now depends on the inverse parameter in a non-linear way. Finally, we conduct a numerical validation of the derived algorithms. In particular, we still observe that very few inner iterations are enough to give good convergence of the algorithm combining multi-step one-shot and OSM.

This chapter is organized as follows. Section 6.1 is dedicated to the case of the non-linear inverse conductivity problem. After stating the forward and inverse problems, we construct the adjoint problem to obtain a convenient expression of the derivative of the least-square cost functional. Next, we apply the nonoverlapping OSM to the forward and adjoint problems in the continuous setting. Then, we give the combined algorithms after rephrasing OSM as an iterative method for the impedance boundary values. In Section 6.2, we write the algorithms for the discretized versions of the forward and adjoint problems using finite element methods. We close this chapter by Section 6.3, where we provide several numerical experiments in two different configurations.

6.1 The forward and inverse problems

6.1.1 The forward problem

We here recall the forward problem where the data is collected from the input of some currents on the boundary associated the Neumann boundary conditions. Given some constants $\underline{\sigma}, \overline{\sigma} > 0$, we define

$$V(\Omega) \coloneqq \{\sigma \in L^\infty(\Omega) : \underline{\sigma} < \sigma < \overline{\sigma} \text{ a.e. in } \Omega\}.$$

Given $\sigma \in V(\Omega)$ and given some parameter $\eta > 0$, the forward problem can be stated as: Seek $u \equiv u(\sigma) \in H^1(\Omega)$ such that

$$\begin{cases} -\operatorname{div}(\sigma \nabla u) + \eta u = 0 & \text{in } \Omega, \\ \sigma \dfrac{\partial u}{\partial \nu} = f & \text{on } \partial\Omega, \end{cases} \tag{6.1}$$

where ν denotes the outer normal vector on $\partial\Omega$ and $f \in H^{-\frac{1}{2}}(\partial\Omega)$. For later use, we recall that the variational formulation of problem (6.1) can be written as: $u \equiv u(\sigma) \in H^1(\Omega)$ such that

$$\int_\Omega \sigma \nabla u \cdot \nabla v \,\mathrm{dx} + \int_\Omega \eta u v \,\mathrm{dx} = \int_{\partial\Omega} f v \,\mathrm{ds}, \quad \forall v \in H^1(\Omega). \tag{6.2}$$

6.1.2 The inverse problem and classical gradient descent

For the inverse problem, we would like to retrieve some $\sigma^{\mathrm{ex}} \in V(\Omega)$ from measuring the trace $g \coloneqq u(\sigma^{\mathrm{ex}})|_{\partial\Omega}$. To do so, we introduce the least-square cost functional

$$J(\sigma) \coloneqq \frac{1}{2} \int_{\partial\Omega} (u(\sigma) - g)^2 \,\mathrm{ds}. \tag{6.3}$$

Proposition 6.1.1. *The functional J defined in (6.3) is differentiable on $V(\Omega)$ and its derivative can be expressed for all $\sigma \in V(\Omega)$ as*

$$\langle J'(\sigma), h \rangle = \int_\Omega h \nabla u(\sigma) \cdot \nabla p(\sigma) \,\mathrm{dx} \tag{6.4}$$

where $h \in L^\infty(\Omega)$ such that $\sigma + h \in V(\Omega)$ and the adjoint state $p \equiv p(\sigma) \in H^1(\Omega)$ is determined by

$$\begin{cases} -\operatorname{div}(\sigma \nabla p) + \eta p = 0 & \text{in } \Omega, \\ \sigma \dfrac{\partial p}{\partial \nu} = g - u(\sigma) & \text{on } \partial\Omega. \end{cases} \tag{6.5}$$

Proof. Since $\sigma \in V(\Omega)$, $\underline{\sigma} < \sigma < \overline{\sigma}$ a.e. in Ω. Taking $v = u(\sigma)$ in (6.2) yields

$$\underline{\sigma}\int_\Omega |\nabla u(\sigma)|^2 + \eta\int_\Omega u(\sigma)^2\,\mathrm{d}x\,\mathrm{d}x \le \int_\Omega \sigma|\nabla u(\sigma)|^2\,\mathrm{d}x + \int_\Omega \eta u(\sigma)^2\,\mathrm{d}x = \int_{\partial\Omega} f u(\sigma)\,\mathrm{d}s \le \|f\|_{H^{-\frac{1}{2}}(\partial\Omega)}\|u\|_{H^{\frac{1}{2}}(\partial\Omega)}.$$

We then deduce that

$$\min(\underline{\sigma},\eta)\|u\|^2_{H^1(\Omega)} \le \|f\|_{H^{-\frac{1}{2}}(\partial\Omega)}\|u\|_{H^{\frac{1}{2}}(\partial\Omega)} \le \|\gamma_0\|\|f\|_{H^{-\frac{1}{2}}(\partial\Omega)}\|u\|_{H^1(\Omega)}$$

where $\gamma_0 : H^1(\Omega) \to H^{\frac{1}{2}}(\partial\Omega)$ is the trace operator on $\partial\Omega$. Hence,

$$\|u(\sigma)\|_{H^1(\Omega)} \le \frac{\|\gamma_0\|}{\min(\underline{\sigma},\eta)}\|f\|_{H^{-\frac{1}{2}}(\partial\Omega)}. \tag{6.6}$$

Let $h \in L^\infty(\Omega)$ such that $\sigma + h \in V(\Omega)$. By definition, $u(\sigma+h)$ satisfies

$$\begin{cases} -\operatorname{div}(\sigma+h)(\nabla u(\sigma+h)) + \eta u(\sigma+h) = 0 & \text{in } \Omega, \\ \sigma\dfrac{\partial u(\sigma+h)}{\partial \nu} = 0 & \text{on } \partial\Omega. \end{cases} \tag{6.7}$$

and similarly to (6.6), we have

$$\|u_{\sigma+h}\|_{H^1(\Omega)} \le \frac{\|\gamma_0\|}{\min(\underline{\sigma},\eta)}\|f\|_{H^{-\frac{1}{2}}(\partial\Omega)}. \tag{6.8}$$

Now let $w_{\sigma,h} \coloneqq u(\sigma+h) - u(\sigma)$. By taking the difference of (6.7) and (6.1), one can verify that

$$-\operatorname{div}(\sigma\nabla w_{\sigma,h}) + \eta w_{\sigma,h} = \operatorname{div}(h\nabla u(\sigma+h)) \text{ in } \Omega. \tag{6.9}$$

The variational formulation of (6.9) can be written as

$$\int_\Omega \sigma\nabla w_{\sigma,h}\cdot\nabla v\,\mathrm{d}x + \int_\Omega \eta w_{\sigma,h} v\,\mathrm{d}x = -\int_\Omega h\nabla u(\sigma+h)\cdot\nabla v\,\mathrm{d}x,\ \forall v \in H^1(\Omega). \tag{6.10}$$

With the same technique as in (6.6), taking $v = w_{\sigma,h}$ in (6.10) yields

$$\|w_{\sigma,h}\|_{H^1(\Omega)} \le \frac{1}{\min(\underline{\sigma},\eta)}\|u(\sigma+h)\|_{H^1(\Omega)}\|h\|_{L^\infty(\Omega)} \overset{(6.8)}{\le} C\|f\|_{H^{-\frac{1}{2}}(\partial\Omega)}\|h\|_{L^\infty(\Omega)} \tag{6.11}$$

where $C \coloneqq \frac{\|\gamma_0\|}{(\min(\underline{\sigma},\eta))^2}$. Next, thanks to the Lax-Milgram Theorem, we can define $v_{\sigma,h} \in H^1(\Omega)$ the solution of

$$\int_\Omega \sigma\nabla v_{\sigma,h}\cdot\nabla v\,\mathrm{d}x + \int_\Omega \eta v_{\sigma,h} v\,\mathrm{d}x = -\int_\Omega h\nabla u(\sigma)\cdot\nabla v\,\mathrm{d}x,\ \forall v \in H^1(\Omega). \tag{6.12}$$

By taking the difference of (6.10) and (6.12), the error $e_{\sigma,h} \coloneqq w_{\sigma,h} - v_{\sigma,h} = u(\sigma+h) - u(\sigma) - v_{\sigma,h}$ verifies

$$\int_\Omega \sigma\nabla e_{\sigma,h}\cdot\nabla v\,\mathrm{d}x + \int_\Omega \eta e_{\sigma,h} v\,\mathrm{d}x = -\int_\Omega h\nabla w_{\sigma,h}\cdot\nabla v\,\mathrm{d}x,\ \forall v \in H^1(\Omega).$$

Taking $v = e_{\sigma,h}$ in this equation gives

$$\|e_{\sigma,h}\|_{H^1(\Omega)} \le \|h\|_{L^\infty(\Omega)}\|w_{\sigma,h}\|_{H^1(\Omega)} \overset{(6.11)}{\le} C\|f\|_{H^{-\frac{1}{2}}(\partial\Omega)}\|h\|^2_{L^\infty(\Omega)},$$

thus $\frac{\|e_{\sigma,h}\|_{\bar{H}(\Omega)}}{\|h\|_{L^\infty(\Omega)}} \to 0$ as $\|h\|_{L^\infty(\Omega)} \to 0$. We therefore conclude that

$$v_{\sigma,h} = \langle u'(\sigma), h\rangle$$

and in particular, J is differentiable on $V(\Omega)$ and

$$\langle J'(\sigma), h\rangle = \int_{\partial\Omega} v_{\sigma,h}(u(\sigma) - g)\,\mathrm{ds}. \tag{6.13}$$

The variational formulation for $p(\sigma)$ defined in (6.5) can be written as

$$\int_\Omega \sigma\nabla p(\sigma)\cdot\nabla\phi\,\mathrm{dx} + \int_\Omega \eta p(\sigma)\phi\,\mathrm{dx} = \int_{\partial\Omega}(g - u(\sigma))\phi\,\mathrm{ds},\ \forall\phi\in H^1(\Omega). \tag{6.14}$$

Taking $v = -p(\sigma)$ in (6.12) and $\phi = -v_{\sigma,h}$ in (6.14) shows that

$$\int_\Omega h\nabla u(\sigma)\cdot\nabla p(\sigma)\,\mathrm{dx} = \int_{\partial\Omega} v_{\sigma,h}(u(\sigma)-g)\,\mathrm{ds} \overset{(6.13)}{=} \langle J'(\sigma), h\rangle,$$

which is (6.4). □

Proposition 6.1.1 shows that the Riesz representative of $J'(\sigma)$ in $L^2(\Omega)$ formally equals $\nabla u(\sigma)\cdot\nabla p(\sigma)$. Therefore, the gradient descent algorithm for σ using this representation for the gradient can be written as

$$\sigma^{n+1} = \sigma^n - \tau\nabla u(\sigma^n)\cdot\nabla p(\sigma^n). \tag{6.15}$$

We refer to Section 5.1.2 for the discussion on how to regularize the descent direction. The one we adopt in our later numerical simulations is to reduce the number of unknowns for σ by discretizing σ on a coarse subdivision of Ω. More specifically, we split Ω as the disjoint union of some subdomains $\Omega_i, i = 0, 1, 2..., M$ where Ω_0 is the neighborhood of Ω where $\sigma = \sigma_0$ is known. Then we seek σ as

$$\sigma = \sigma_0\mathbb{1}_{\Omega_0} + \sum_{i=1}^M \sigma_i\mathbb{1}_{\Omega_i} \text{ where } \sigma_i\in\mathbb{R}, i = 1, ..., M \text{ are unknown.} \tag{6.16}$$

We can then view J as a function of $(\sigma_i)_{i=1}^M$ and it is quite straightforward to conclude from Proposition 6.1.1 that J is differentiable with respect to each σ_i and

$$\frac{\partial J}{\partial\sigma_i}(\sigma) = \int_{\Omega_i}\nabla u(\sigma)\cdot\nabla p(\sigma)\,\mathrm{dx}.$$

In this case, the gradient descent algorithm can be written as

$$\sigma_i^{n+1} = \sigma_i^n - \tau\int_{\Omega_i}\nabla u(\sigma^n)\cdot\nabla p(\sigma^n)\,\mathrm{dx},\ i = 1, ..., M \text{ where } \sigma^n = \sigma_0\mathbb{1}_{\Omega_0} + \sum_{i=1}^M \sigma_i^n\mathbb{1}_{\Omega_i}. \tag{6.17}$$

6.1.3 The application of OSM to the forward and adjoint problems

OSM for the forward problem

We here outline the application of the nonoverlapping OSM to the forward problem (6.1). In order to ease the presentation, we consider the splitting of the domain Ω into two

disjoint subdomains Ω_0 and Ω_1 where Ω_0 is the neighborhood of $\Gamma := \partial\Omega$. We assume that the conductivity σ equals σ_0 in Ω_0 and equals σ_1 in Ω_1 (here σ_1 can be a function), i.e. $\sigma = \sigma_0 \mathbb{1}_{\Omega_0} + \sigma_1 \mathbb{1}_{\Omega_1}$. Furthermore, we assume that σ_0 is known. We remark that the following schemes can be easily extended to the cases where the subdomain Ω_1 is also splitted into several subdomains.

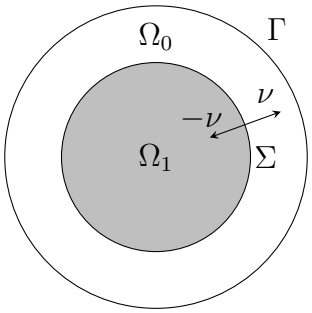


Figure 6.1: Illustration for the domain Ω and its decomposition.

We first rewrite the forward problem (6.1) as a transmission problem between the subdomains Ω_0 and Ω_1. We denote by Σ the interface between Ω_0 and Ω_1, which is also the boundary of Ω_1. We also denote by ν the normal unit vector on Γ directed toward the exterior of Ω and use the same notion for the normal vector on Σ directed toward the exterior of Ω_1 (see Figure (6.1)). For $i = 0, 1$, let us denote by $u_i \in H^1(\Omega_i)$ the restriction of the solution u in the domain Ω_i. Then problem (6.1) can be equivalently written as follows:

$$\begin{cases} -\operatorname{div}(\sigma_0 \nabla u_0) + \eta u_0 = 0 & \text{in } \Omega_0, \\ \sigma_0 \dfrac{\partial u_0}{\partial \nu} = f & \text{on } \Gamma, \\ \sigma_0 \dfrac{\partial u_0}{\partial \nu} - \beta u_0 = \sigma_1 \dfrac{\partial u_1}{\partial \nu} - \beta u_1 & \text{on } \Sigma. \end{cases} \tag{6.18}$$

and

$$\begin{cases} -\operatorname{div}(\sigma_1 \nabla u_1) + \eta u_1 = 0 & \text{in } \Omega_1, \\ \sigma_1 \dfrac{\partial u_1}{\partial \nu} + \beta u_1 = \sigma_0 \dfrac{\partial u_0}{\partial \nu} + \beta u_0 & \text{on } \Sigma, \end{cases} \tag{6.19}$$

where $\beta \neq 0$ is a given parameter (see Section 1.4). The nonoverlapping OSM can be seen as a fixed-point iteration applied to (6.18) and (6.19). More precisely, it consists in the following induction for the sequences $(u_{0;\ell})_{\ell \geq 0}$ and $(u_{1;\ell})_{\ell \geq 0}$ with a given initial guess $(u_{0;0}, u_{1;0})$:

$$\begin{cases} -\operatorname{div}(\sigma_0 \nabla u_{0;\ell+1}) + \eta u_{0;\ell+1} = 0 & \text{in } \Omega_0, \\ \sigma_0 \dfrac{\partial u_{0;\ell+1}}{\partial \nu} = f & \text{on } \partial\Omega, \\ \sigma_0 \dfrac{\partial u_{0;\ell+1}}{\partial \nu} - \beta u_{0;\ell+1} = \sigma_1 \dfrac{\partial u_{1;\ell}}{\partial \nu} - \beta u_{1;\ell} & \text{on } \Sigma \end{cases} \tag{6.20}$$

and

$$\begin{cases} -\operatorname{div}(\sigma_1 \nabla u_{1;\ell+1}) + \eta u_{1;\ell+1} = 0 & \text{in } \Omega_1, \\ \sigma_1 \dfrac{\partial u_{1;\ell+1}}{\partial \nu} + \beta u_{1;\ell+1} = \sigma_0 \dfrac{\partial u_{0;\ell}}{\partial \nu} + \beta u_{0;\ell} & \text{on } \Sigma. \end{cases} \tag{6.21}$$

As we already mentioned in Section 1.3, it is more convenient to rewrite an induction for the impedance values

$$\lambda_{0;\ell+1} := \sigma_1 \frac{\partial u_{1;\ell}}{\partial \nu} - \beta u_{1;\ell} \quad \text{and} \quad \lambda_{1;\ell+1} := \sigma_0 \frac{\partial u_{0;\ell}}{\partial \nu} + \beta u_{0;\ell}. \tag{6.22}$$

Using these impedance variables, (6.20) and (6.21) respectively lead to

$$\begin{cases} -\operatorname{div}(\sigma_0 \nabla u_{0;\ell}) + \eta u_{0;\ell} = 0 & \text{in } \Omega_0, \\ \sigma_0 \dfrac{\partial u_{0;\ell}}{\partial \nu} = f & \text{on } \partial\Omega, \\ \sigma_0 \dfrac{\partial u_{0;\ell}}{\partial \nu} - \beta u_{0;\ell} = \lambda_{0;\ell} & \text{on } \Sigma \end{cases} \tag{6.23}$$

and

$$\begin{cases} -\operatorname{div}(\sigma_1 \nabla u_{1;\ell}) + \eta u_{1;\ell} = 0 & \text{in } \Omega_1, \\ \sigma_1 \dfrac{\partial u_{1;\ell}}{\partial \nu} + \beta u_{1;\ell} = \lambda_{1;\ell} & \text{on } \Sigma. \end{cases} \tag{6.24}$$

One indeed has

$$\lambda_{0;\ell+1} \overset{(6.22)}{=} \left(\sigma_1 \frac{\partial u_{1;\ell}}{\partial \nu} + \beta u_{1;\ell}\right) - 2\beta u_{1;\ell} \overset{(6.24)}{=} \lambda_{1;\ell} - 2\beta u_{1;\ell} \text{ on } \Sigma$$

and

$$\lambda_{1;\ell+1} \overset{(6.22)}{=} \left(\sigma_0 \frac{\partial u_{0;\ell}}{\partial \nu} - \beta u_{0;\ell}\right) + 2\beta u_{0;\ell} \overset{(6.23)}{=} \lambda_{0;\ell} + 2\beta u_{0;\ell} \text{ on } \Sigma.$$

Therefore,

$$\begin{cases} \lambda_{0;\ell+1} = \lambda_{1;\ell} - 2\beta u_{1;\ell} \text{ on } \Sigma, \\ \lambda_{1;\ell+1} = \lambda_{0;\ell} + 2\beta u_{0;\ell} \text{ on } \Sigma. \end{cases} \tag{6.25}$$

For the sake of compacting the notation, let us introduce the following two operators:

$$\mathcal{A}_0 : \lambda \in L^2(\Sigma) \mapsto u \in H^1(\Omega_0)$$

such that

$$\begin{cases} -\operatorname{div}(\sigma_0 \nabla u) + \eta u = 0 & \text{in } \Omega_0, \\ \sigma_0 \dfrac{\partial u}{\partial \nu} = f & \text{on } \Gamma, \\ \sigma_0 \dfrac{\partial u}{\partial \nu} - \beta u = \lambda & \text{on } \Sigma \end{cases} \tag{6.26}$$

and

$$\mathcal{A}_1 : (\lambda, \sigma) \in L^2(\Sigma) \times L^\infty(\Omega_1) \mapsto u \in H^1(\Omega_1)$$

such that

$$\begin{cases} -\operatorname{div}(\sigma \nabla u) + \eta u = 0 & \text{in } \Omega_1, \\ \sigma \dfrac{\partial u}{\partial \nu} + \beta u = \lambda & \text{on } \Sigma. \end{cases} \tag{6.27}$$

In addition, for $i = 0, 1$, we introduce

$$\gamma_{\Omega_i,\Sigma} : H^1(\Omega_i) \to L^2(\Sigma) \tag{6.28}$$

the trace operator on Σ. Then, thanks to (6.25), the nonoverlapping OSM algorithm for the forward problem can be synthetically written as the following induction with a given initial guess $(\lambda_{0;0}, \lambda_{1;0})$:

$$\begin{cases} \lambda_{0;\ell+1} = \lambda_{1;\ell} - 2\beta \gamma_{\Omega_1,\Sigma} \mathcal{A}_1(\lambda_{1;\ell}, \sigma_1), \\ \lambda_{1;\ell+1} = \lambda_{0;\ell} + 2\beta \gamma_{\Omega_0,\Sigma} \mathcal{A}_0(\lambda_{0;\ell}). \end{cases} \tag{6.29}$$

We can even write (6.29) into a more compact form by setting: $\lambda_{[\ell]} = (\lambda_{1;\ell}, \lambda_{2;\ell}) \in L^2(\Sigma)^2$ as

$$\lambda_{[\ell+1]} = \mathcal{A}(\lambda_{[\ell]}, \sigma_1) \tag{6.30}$$

where $\mathcal{A} : L^2(\Sigma)^2 \times L^\infty(\Omega_1) \to L^2(\Sigma)^2$ is defined by

$$\mathcal{A}(\boldsymbol{\lambda}, \sigma) \coloneqq (\boldsymbol{\lambda_1} - 2\beta\gamma_{\Omega_1,\Sigma}\mathcal{A}_1(\boldsymbol{\lambda_1}, \sigma), \boldsymbol{\lambda_0} + 2\beta\gamma_{\Omega_0,\Sigma}\mathcal{A}_0(\boldsymbol{\lambda_0})), \tag{6.31}$$

here $\boldsymbol{\lambda} = (\boldsymbol{\lambda_0}, \boldsymbol{\lambda_1}) \in L^2(\Sigma)^2$.

OSM for the adjoint equation

With the same domain setting as in the previous section 6.1.3, we now consider the application of the nonoverlapping DDM to the adjoint problem (6.5) associated with the forward problem (6.1). Similarly, we shall write the OSM algorithm for the adjoint state p step by step as follows.

For $i = 0, 1$, let us denote by $p_i \in H^1(\Omega_i)$ the restriction of the solution p to the domain Ω_i. By using the same parameter β appearing in (6.18) and (6.19), we can equivalently write problem (6.5) as follows:

$$\begin{cases} -\operatorname{div}(\sigma_0 \nabla p_0) + \eta p_0 = 0 & \text{in } \Omega_2, \\ \sigma_0 \dfrac{\partial p_0}{\partial \nu} = g - u(\sigma) & \text{on } \Gamma, \\ \sigma_0 \dfrac{\partial p_0}{\partial \nu} - \beta p_0 = \sigma_1 \dfrac{\partial p_1}{\partial \nu} - \beta p_1 & \text{on } \Sigma \end{cases} \tag{6.32}$$

and

$$\begin{cases} -\operatorname{div}(\sigma_1 \nabla p_1) + \eta p_1 = 0 & \text{in } \Omega_1, \\ \sigma_1 \dfrac{\partial p_1}{\partial \nu} + \beta p_1 = \sigma_0 \dfrac{\partial p_0}{\partial \nu} + \beta p_0 & \text{on } \Sigma. \end{cases} \tag{6.33}$$

The nonoverlapping OSM algorithm applied to (6.32) and (6.33) is described by the following induction for the sequences $(p_{0;\ell})_{\ell \geq 0}$ and $(p_{1;\ell})_{\ell \geq 0}$ with a given initial guess $(p_{0;0}, p_{1;0})$:

$$\begin{cases} -\operatorname{div}(\sigma_0 \nabla p_{0;\ell+1}) + \eta p_{0;\ell+1} = 0 & \text{in } \Omega_0, \\ \sigma_0 \dfrac{\partial p_{0;\ell+1}}{\partial \nu} = g - u(\sigma) & \text{on } \partial\Omega, \\ \sigma_0 \dfrac{\partial p_{0;\ell+1}}{\partial \nu} - \beta p_{0;\ell+1} = \sigma_1 \dfrac{\partial p_{1;\ell}}{\partial \nu} - \beta p_{1;\ell} & \text{on } \Sigma \end{cases} \tag{6.34}$$

and

$$\begin{cases} -\operatorname{div}(\sigma_1 \nabla p_{1;\ell+1}) + \eta p_{1;\ell+1} = 0 & \text{in } \Omega_1, \\ \sigma_1 \dfrac{\partial p_{1;\ell+1}}{\partial \nu} + \beta p_{1;\ell+1} = \sigma_0 \dfrac{\partial p_{0;\ell}}{\partial \nu} + \beta p_{0;\ell} & \text{on } \Sigma. \end{cases} \tag{6.35}$$

By introducing the impedance values

$$\hat{\lambda}_{0;\ell+1} \coloneqq \sigma_1 \frac{\partial p_{1;\ell}}{\partial \nu} - \beta p_{1;\ell}, \quad \text{and} \quad \hat{\lambda}_{1;\ell+1} \coloneqq \sigma_2 \frac{\partial p_{0;\ell}}{\partial \nu} + \beta p_{0;\ell}, \tag{6.36}$$

we can equivalently rewrite the induction for the sequences $(p_{0;\ell})_{\ell \geq 0}$ and $(p_{1;\ell})_{\ell \geq 0}$, defined by (6.34) and (6.35), as

$$\begin{cases} -\operatorname{div}(\sigma_0 \nabla p_{0;\ell}) + \eta p_{0;\ell} = 0 & \text{in } \Omega_0, \\ \sigma_0 \dfrac{\partial p_{0;\ell}}{\partial \nu} = g - u(\sigma) & \text{on } \Gamma, \\ \sigma_0 \dfrac{\partial p_{0;\ell}}{\partial \nu} - \beta p_{0;\ell} = \hat{\lambda}_{0;\ell} & \text{on } \Sigma. \end{cases} \tag{6.37}$$

and

$$\begin{cases} -\operatorname{div}(\sigma_1 \nabla p_{1;\ell}) + \eta p_{1;\ell} = 0 & \text{in } \Omega_1, \\ \sigma_1 \dfrac{\partial p_{1;\ell}}{\partial \nu} + \beta p_{1;\ell} = \hat{\lambda}_{1;\ell} & \text{on } \Sigma. \end{cases} \tag{6.38}$$

One also has the following induction on the impedance values

$$\begin{cases} \hat{\lambda}_{0;\ell+1} = \hat{\lambda}_{1;\ell} - 2\beta p_{1;\ell} \text{ on } \Sigma, \\ \hat{\lambda}_{1;\ell+1} = \hat{\lambda}_{0;\ell} + 2\beta p_{0;\ell} \text{ on } \Sigma. \end{cases} \tag{6.39}$$

Let us introduce the operator

$$\hat{\mathcal{A}}_0 : (\lambda, \psi) \in L^2(\Sigma) \times \times H^{-\frac{1}{2}}(\Gamma) \mapsto u \in H^1(\Omega_0)$$

such that

$$\begin{cases} -\operatorname{div}(\sigma_0 \nabla u) + \eta u = 0 & \text{in } \Omega_0, \\ \sigma_0 \dfrac{\partial u}{\partial \nu} = \psi & \text{on } \Gamma, \\ \sigma_0 \dfrac{\partial u}{\partial \nu} - \beta u = \lambda & \text{on } \Sigma. \end{cases} \tag{6.40}$$

We also introduce

$$\gamma_\Gamma : H^1(\Omega_2) \to L^2(\Gamma) \tag{6.41}$$

the trace operator on Γ. Then, thanks to (6.39), the nonoverlapping OSM algorithm for the adjoint problem (6.5) can be synthetically written as the following induction with a given initial guess $(\hat{\lambda}_{0;0}, \hat{\lambda}_{1;0})$:

$$\begin{cases} \hat{\lambda}_{0;\ell+1} = \hat{\lambda}_{1;\ell} - 2\beta \gamma_{\Omega_1,\Sigma} \mathcal{A}_1(\hat{\lambda}_{1;\ell}, \sigma_1), \\ \hat{\lambda}_{1;\ell+1} = \hat{\lambda}_{0;\ell} + 2\beta \gamma_{\Omega_0,\Sigma} \hat{\mathcal{A}}_0(\hat{\lambda}_{0;\ell}, g - \gamma_\Gamma u(\sigma)). \end{cases} \tag{6.42}$$

Here we recall that the trace operators $\gamma_{\Omega_i,\Sigma}$, $i = 0, 1$ and the operator $\mathcal{A}_1$ are already defined respectively in (6.28) and (6.27). We can even write (6.42) into a more compact form by setting: $\hat{\lambda}_{[\ell]} = (\hat{\lambda}_{0;\ell}, \hat{\lambda}_{1;\ell})$ as

$$\hat{\lambda}_{[\ell+1]} = \hat{\mathcal{A}}(\hat{\lambda}_{[\ell]}, \sigma_1, g - \gamma_\Gamma u(\sigma)) \tag{6.43}$$

where $\hat{\mathcal{A}} : L^2(\Sigma)^2 \times L^\infty(\Omega_1) \times H^{-\frac{1}{2}}(\Gamma) \to L^2(\Sigma)^2$ is defined by

$$\hat{\mathcal{A}}(\boldsymbol{\lambda}, \sigma, \psi) := (\boldsymbol{\lambda_1} - 2\beta \gamma_{\Omega_1,\Sigma} \mathcal{A}_1(\boldsymbol{\lambda_1}, \sigma), \boldsymbol{\lambda_0} + 2\beta \gamma_{\Omega_0,\Sigma} \hat{\mathcal{A}}_0(\boldsymbol{\lambda_0}, \psi)), \tag{6.44}$$

here $\boldsymbol{\lambda} = (\boldsymbol{\lambda_0}, \boldsymbol{\lambda_1}) \in L^2(\Sigma)^2$.

OSM for the simultaneous calculation of the forward and adjoint states

In the nonoverlapping OSM algorithm applied to the adjoint equation (6.42), one assumes that the state $u(\sigma)$ is already known. In practice, it is indeed more natural to perform the iterations in (6.42) in parallel to the OSM iterations for the forward problem (6.29). The combination of (6.29) and (6.42) leads to the following scheme that simultaneously computes the forward and adjoint states: given some initial guess $(\lambda_{0;0}, \lambda_{1;0}, \hat{\lambda}_{0;0}, \hat{\lambda}_{1;0})$,

$$\begin{cases} \lambda_{0;\ell+1} = \lambda_{1;\ell} - 2\beta \gamma_{\Omega_1,\Sigma} \mathcal{A}_1(\lambda_{1;\ell}, \sigma_1), \\ \lambda_{1;\ell+1} = \lambda_{0;\ell} + 2\beta \gamma_{\Omega_0,\Sigma} \mathcal{A}_0(\lambda_{0;\ell}), \\ \hat{\lambda}_{0;\ell+1} = \hat{\lambda}_{1;\ell} - 2\beta \gamma_{\Omega_1,\Sigma} \mathcal{A}_1(\hat{\lambda}_{1;\ell}, \sigma_1), \\ \hat{\lambda}_{1;\ell+1} = \hat{\lambda}_{0;\ell} + 2\beta \gamma_{\Omega_0,\Sigma} \hat{\mathcal{A}}_0(\hat{\lambda}_{0;\ell}, g - \gamma_\Gamma \mathcal{A}_0(\lambda_{0;\ell})). \end{cases} \tag{6.45}$$

Using the notations $\mathcal{A}$ and $\hat{\mathcal{A}}$ respectively introduced in (6.31) and (6.44), we can also rewrite (6.45) more compactly as: given some initial guess $(\lambda_{[0]}, \hat{\lambda}_{[0]})$,

$$\begin{cases} \lambda_{[\ell+1]} = \mathcal{A}(\lambda_{[\ell]}, \sigma_1), \\ \hat{\lambda}_{[\ell+1]} = \hat{\mathcal{A}}(\hat{\lambda}_{[\ell]}, \sigma_1, g - \gamma_\Gamma \mathcal{A}_0(\lambda_{0;\ell})). \end{cases} \tag{6.46}$$

where $\lambda_{[\ell]} = (\lambda_{0;\ell}, \lambda_{1;\ell}) \in L^2(\Sigma)^2$ and $\hat{\lambda}_{[\ell]} = (\hat{\lambda}_{0;\ell}, \hat{\lambda}_{1;\ell}) \in L^2(\Sigma)^2$. The convergence of this scheme in some simplified cases may be done using the following abstract result.

Theorem 6.1.2. *Consider the sequence $(\lambda_{[\ell]}, \hat{\lambda}_{[\ell]})$ defined by (6.46) for a given $\sigma_1 \in L^\infty(\Omega_1)$ and given initial guess $(\lambda_{0;0}, \lambda_{1;0}, \hat{\lambda}_{0;0}, \hat{\lambda}_{1;0})$. Let $\|.\|$ be the in $L^2(\Sigma)^2$ norm. We make the following assumptions:*

(i) There exists a subspace U of $H^{-\frac{1}{2}}(\Gamma)$, a constant $\tilde{C} \in (0,1)$ and a constant $C > 0$ such that

$$\|\hat{\mathcal{A}}(\boldsymbol{\lambda}, \sigma_1, \psi)\| \leq \tilde{C}\|\boldsymbol{\lambda}\| + C\|\psi\|, \ \forall (\boldsymbol{\lambda}, \psi) \in L^2(\Sigma)^2 \times U,$$

(ii) $\gamma_\Gamma \mathcal{A}_0(\lambda_{0;\ell})$ stays in U for all ℓ.

Then the sequence $(\lambda_{[\ell]}, \hat{\lambda}_{[\ell]})$, $\ell \geq 0$ is convergent.

Proof. By the definitions of $\mathcal{A}_0$, $\mathcal{A}_1$, $\mathcal{A}$, $\hat{\mathcal{A}}_0$ and $\hat{\mathcal{A}}$ respectively in (6.26), (6.27), (6.31), (6.40) and (6.44), we observe that $\mathcal{A}(\boldsymbol{\lambda}, \sigma) = \hat{\mathcal{A}}(\boldsymbol{\lambda}, \sigma, f)$, $\forall (\boldsymbol{\lambda}, \sigma, f) \in L^2(\Sigma)^2 \times L^\infty(\Omega_1) \times H^{-\frac{1}{2}}(\Gamma)$ and $\hat{\mathcal{A}}(\boldsymbol{\lambda}, \sigma, \psi)$ is linear with respect to $(\boldsymbol{\lambda}, \psi) \in L^2(\Sigma)^2 \times H^{-\frac{1}{2}}(\Gamma)$. Therefore,

$$\lambda_{[\ell+1]} - \lambda_{[\ell]} = \mathcal{A}(\lambda_{[\ell]}, \sigma_1) - \mathcal{A}(\lambda_{[\ell-1]}, \sigma_1) = \hat{\mathcal{A}}(\lambda_{[\ell]} - \lambda_{[\ell-1]}, \sigma_1, 0), \forall \ell \geq 1.$$

Since $0 \in U$, by assumption *(i)* we have

$$\|\lambda_{[\ell+1]} - \lambda_{[\ell]}\| \leq \hat{C}\|\lambda_{[\ell]} - \lambda_{[\ell-1]}\|, \forall \ell \geq 1.$$

This implies

$$\|\lambda_{[k+\ell+1]} - \lambda_{[k+\ell]}\| \leq \hat{C}^\ell \|\lambda_{[k+1]} - \lambda_{[k]}\|, \forall k, \ell \geq 0.$$

Particularly, we have

$$\|\lambda_{[\ell+1]} - \lambda_{[\ell]}\| \leq \hat{C}^\ell \|\lambda_{[1]} - \lambda_{[0]}\|, \forall \ell \geq 0.$$

We conclude that $(\lambda_{[\ell]})_{\ell \geq 0}$ is a Cauchy sequence and thus convergent using the same arguments as in the proof of Theorem 5.1.2.

Next, for $\ell \geq 1$ we have

$$\begin{aligned} \hat{\lambda}_{[\ell+1]} - \hat{\lambda}_{[\ell]} &= \hat{\mathcal{A}}(\hat{\lambda}_{[\ell]}, \sigma_1, g - \gamma_\Gamma \mathcal{A}_0(\lambda_{0;\ell})) - \hat{\mathcal{A}}(\hat{\lambda}_{[\ell-1]}, \sigma_1, g - \gamma_\Gamma \mathcal{A}_0(\lambda_{0;\ell-1})) \\ &= \hat{\mathcal{A}}(\hat{\lambda}_{[\ell]} - \hat{\lambda}_{[\ell-1]}, \sigma_1, \gamma_\Gamma \mathcal{A}_0(\lambda_{0;\ell-1}) - \gamma_\Gamma \mathcal{A}_0(\lambda_{0;\ell})) \\ &= \hat{\mathcal{A}}(\hat{\lambda}_{[\ell]} - \hat{\lambda}_{[\ell-1]}, \sigma_1, \gamma_\Gamma \hat{\mathcal{A}}_0(\lambda_{0;\ell-1} - \lambda_{0;\ell}, 0)). \end{aligned}$$

Here we note that $\mathcal{A}_0(\lambda) = \hat{\mathcal{A}}_0(\lambda, f)$, $\forall (\lambda, f) \in L^2(\Sigma) \times H^{-\frac{1}{2}}(\Gamma)$ and $\hat{\mathcal{A}}_0(\lambda, \psi)$ is linear with respect to $(\lambda, \psi) \in L^2(\Sigma) \times H^{-\frac{1}{2}}(\Gamma)$. Also, assumption *(ii)* ensures that

$$\gamma_\Gamma \hat{\mathcal{A}}_0(\lambda_{0;\ell-1} - \lambda_{0;\ell}, 0) = \gamma_\Gamma \mathcal{A}_0(\lambda_{0;\ell-1}) - \gamma_\Gamma \mathcal{A}_0(\lambda_{0;\ell}) \in U.$$

Hence, assumption *(i)* gives us

$$\|\hat{\lambda}_{[\ell+1]} - \hat{\lambda}_{[\ell]}\| \leq \hat{C}\|\hat{\lambda}_{[\ell]} - \hat{\lambda}_{[\ell-1]}\| + \underbrace{C\|\gamma_\Gamma\|\|\mathcal{A}_0(\cdot,0)\|}_{M\geq 0}\|\lambda_{0;\ell} - \lambda_{0;\ell-1}\|$$
$$\leq \hat{C}\|\hat{\lambda}_{[\ell]} - \hat{\lambda}_{[\ell-1]}\| + M\|\lambda_{[\ell]} - \lambda_{[\ell-1]}\|, \ \forall \ell \geq 1.$$

Here we remark that $\mathcal{A}_0(\cdot,0)$, derived from (6.26), is the linear operator defined by

$$\hat{\mathcal{A}}_0(\cdot,0) : \lambda \in L^2(\Sigma) \mapsto u \in H^1(\Omega_0)$$

such that

$$\begin{cases} -\operatorname{div}(\sigma_0 \nabla u) + \eta u = 0 & \text{in } \Omega_0, \\ \sigma_0 \dfrac{\partial u}{\partial \nu} = 0 & \text{on } \Gamma, \\ \sigma_0 \dfrac{\partial u}{\partial \nu} - \beta u = \lambda & \text{on } \Sigma. \end{cases}$$

Using the same arguments as in the proof of Theorem 5.1.2, we conclude that $(\hat{\lambda}_{[\ell]})_{\ell\geq 0}$ is also a Cauchy sequence and thus convergent. □

An algorithm combining gradient descent with domain decomposition

We here use the same notation as in Section 6.1.3, i.e. the domain Ω is splitted into two disjoint subdomains Ω_0 and Ω_1. We assume that the conductivity σ equals σ_0 in Ω_0 and equals σ_1 in Ω_1, i.e. $\sigma = \sigma_0 \mathbb{1}_{\Omega_0} + \sigma_1 \mathbb{1}_{\Omega_1}$. Furthermore, we assume that σ_0 is known and σ_1 is unknown. Then the 1-step one-shot inversion algorithm combining the nonoverlapping OSM (6.46) and the gradient descent (6.15) can be written as follows: given some initial guess $(\sigma_1^0, \lambda^0, \hat{\lambda}^0)$, we iterate using the induction

$$\begin{cases} \sigma_1^{n+1} = \sigma_1^n - \tau \nabla \mathcal{A}_1(\lambda_1^n, \sigma_1^n) \cdot \nabla \hat{\mathcal{A}}_1(\hat{\lambda}_1^n, \sigma_1^n), \\ \lambda^{n+1} = \mathcal{A}(\lambda^n, \sigma^n), \\ \hat{\lambda}^{n+1} = \hat{\mathcal{A}}(\hat{\lambda}^n, \sigma^n, g - \gamma_\Gamma \mathcal{A}_2(\lambda_2^n, \sigma_2^n)), \end{cases} \tag{6.47}$$

where $\lambda^n = (\lambda_0^n, \lambda_1^n)$, $\hat{\lambda}^n = (\hat{\lambda}_0^n, \hat{\lambda}_1^n)$; $\mathcal{A}$ and $\hat{\mathcal{A}}$ are respectively defined in (6.31) and (6.44). The k-step version of algorithm (6.47) can be also written as: given some initial guess $(\sigma_1^0, \lambda^0, \hat{\lambda}^0)$, we iterate using the induction

$$\begin{cases} \sigma_1^{n+1} = \sigma_1^n - \tau \nabla \mathcal{A}_1(\lambda_1^n, \sigma_1^n) \cdot \nabla \hat{\mathcal{A}}_1(\hat{\lambda}_1^n, \sigma_1^n), \\ \lambda_{[0]}^{n+1} = \lambda^n, \hat{\lambda}_{[0]}^{n+1} = \hat{\lambda}^n, \\ \text{for } \ell = 0, 1, ..., k-1 : \\ \quad \left| \begin{array}{l} \lambda_{[\ell+1]}^{n+1} = \mathcal{A}(\lambda_{[\ell]}^n, \sigma^{n+1}), \\ \hat{\lambda}_{[\ell+1]}^{n+1} = \hat{\mathcal{A}}(\hat{\lambda}_{[\ell]}^n, \gamma_\Gamma \mathcal{A}_0(\lambda_{[\ell]}^n) - g), \end{array} \right. \\ \lambda^{n+1} = \lambda_{[k]}^{n+1}, \hat{\lambda}^{n+1} = \hat{\lambda}_{[k]}^{n+1}, \end{cases} \tag{6.48}$$

where $\lambda^n = (\lambda_0^n, \lambda_1^n)$, $\hat{\lambda}^n = (\hat{\lambda}_0^n, \hat{\lambda}_1^n)$, $\lambda_{[\ell]}^n = (\lambda_{0;\ell}^n, \lambda_{1;\ell}^n)$, $\hat{\lambda}_{[\ell]}^n = (\hat{\lambda}_{0;\ell}^n, \hat{\lambda}_{1;\ell}^n)$; $\mathcal{A}$ and $\hat{\mathcal{A}}$ are respectively defined in (6.31) and (6.44). The discretized version of scheme (6.48) will be studied in the next section, in the same way as we did in Section 5.2.

In the case where Ω_1 is the disjoint union of the subdomains $\Omega_{1,i}, i = 1, 2..., n_{\sigma_1}$ and σ_1 is given by

$$\sigma_1 = \sum_{i=1}^{n_{\sigma_1}} \sigma_{1,i} \mathbb{1}_{\Omega_{1,i}} \text{ where } \sigma_{1,i} \in \mathbb{R}, i = 1, 2, ..., n_{\sigma_1}, \tag{6.49}$$

the first step in scheme (6.48) should be replaced with

$$\sigma_{1,i}^{n+1} = \sigma_{1,i}^{n} - \tau \int_{\Omega_{1,i}} \nabla \mathcal{A}_1(\lambda_1^n, \sigma_1^n) \cdot \nabla \hat{\mathcal{A}}_1(\hat{\lambda}_1^n, \sigma_1^n)\, \mathrm{dx}, \quad i = 1, 2, ..., n_{\sigma_1}, \tag{6.50}$$

where $\sigma_1^n = \sum_{i=1}^{n_{\sigma_1}} \sigma_{1,i}^n \mathbb{1}_{\Omega_{1,i}}$, $n \geq 0$.

6.2 Discretized versions of the algorithms studied in 6.1.3

The main goal of this section is to give the discretized version of (6.48). We follow the same steps as in Section 5.2 by discretizing the problems using finite element methods. For $i = 0, 1$, we assume that Ω_i is discretized using a regular triangulation $\mathcal{T}(\Omega_i)$. We further assume that Σ and Γ are respectively discretized using edge elements $\mathcal{S}(\Sigma)$ and $\mathcal{S}(\Gamma)$. We define the following $\mathbb{P}^1$ finite element spaces:

$$\begin{aligned} \mathcal{V}(\Omega_i) &\coloneqq \left\{u \in C^0(\Omega_i) : u|_{T\in\mathcal{T}(\Omega_i)} \in \mathbb{P}^1\right\} = \mathrm{Span}\{\phi_i^j, 1 \leq j \leq n_i\},\ i = 0, 1, \\ \mathcal{V}(\Sigma) &\coloneqq \left\{\lambda \in C^0(\Sigma) : \lambda|_{S\in\mathcal{S}(\Sigma)} \in \mathbb{P}^1\right\} = \mathrm{Span}\{\phi_\Sigma^j, 1 \leq j \leq n_\Sigma\}, \\ \text{and } \mathcal{V}(\Gamma) &\coloneqq \left\{g \in C^0(\Gamma) : g|_{S\in\mathcal{S}(\Gamma)} \in \mathbb{P}^1\right\} = \mathrm{Span}\{\phi_\Gamma^j, 1 \leq j \leq n_\Gamma\}, \end{aligned}$$

where n_i, $i = 0, 1$ is the number of degrees of freedom for the solution in Ω_i, n_Γ and n_Σ are the number of degrees of freedom respectively for the measurement on Γ and for the impedance on Σ. A discretized function in any of these spaces is determined by the corresponding column vectors with respect to the chosen basis.

We also consider the following case where σ_1 is of the form

$$\sigma_1 = \sum_{i=1}^{n_{\sigma_1}} \sigma_{1,i} \mathbb{1}_{\Omega_{1,i}} \text{ where } \sigma_{1,i} \in \mathbb{R}, i = 1, 2, ..., n_{\sigma_1},$$

and we set

$$S_1 \coloneqq \mathrm{Col}(\sigma_{1;i})_{i=1}^{n_{\sigma_1}}.$$

6.2.1 Discretized version of the operators $\mathcal{A}$ and $\hat{\mathcal{A}}$

Let us prepare some necessary matrices for writing the discretized version of the operators $\mathcal{A}$ and $\hat{\mathcal{A}}$. We respectively denote by $\tilde{M}_\Sigma \in \mathbb{R}^{n_\Sigma \times n_\Sigma}$, and $\tilde{M}_\Gamma \in \mathbb{R}^{n_\Gamma \times n_\Gamma}$ the Σ-interface and Γ-interface mass matrices. More precisely,

$$\tilde{M}_{\Sigma;p,q} \coloneqq \int_\Sigma \phi_\Sigma^q \phi_\Sigma^p \,\mathrm{ds},\ 1 \leq p, q \leq n_\Sigma, \tag{6.51}$$

and

$$\tilde{M}_{\Gamma;p,q} = \int_\Gamma \phi_\Gamma^p \phi_\Gamma^q \,\mathrm{ds},\ 1 \leq p, q \leq n_\Gamma. \tag{6.52}$$

Also, for $i = 0, 1$, we introduce $\tilde{R}_i \in \mathbb{R}^{n_\Sigma \times n_i}$ the interpolation matrix from $\mathcal{V}(\Omega_i)$ to $\mathcal{V}(\Sigma)$. Then

$$\tilde{R}_i = \tilde{M}_\Sigma^{-1} \tilde{M}_{i;\Sigma} \tag{6.53}$$

where $\tilde{M}_{i;\Sigma} \in \mathbb{R}^{n_\Sigma \times n_i}$ is defined by

$$\tilde{M}_{i;\Sigma;p,q} \coloneqq \int_\Sigma \phi_i^q \phi_\Sigma^p \,\mathrm{ds},\ 1 \leq p \leq n_\Sigma, 1 \leq q \leq n_i. \tag{6.54}$$

Now let $\tilde{A}_i \in \mathbb{R}^{n_i \times n_i}, i = 0, 1$ be the matrices arising from the discretization of $-\triangle + \eta$ along with the boundary condition $\partial_\nu \mp \beta$ on Σ. More precisely, let $\tilde{M}_i$, $i = 0, 1$ be the mass matrices in Ω_i defined by

$$\tilde{M}_{i;p,q} \coloneqq \int_{\Omega_i} \phi_\Sigma^q \phi_\Sigma^p \, \mathrm{dx}, \ 1 \leq p, q \leq n_i, i = 0, 1 \tag{6.55}$$

and let $\tilde{K}_i \in \mathbb{R}^{n_i \times n_i}$, $i = 0, 1$ be the local stiffness matrices of the problem defined by

$$\tilde{K}_{i;p,q} \coloneqq \int_{\Omega_i} \sigma_i \nabla \phi_i^q \cdot \nabla \phi_i^p \, \mathrm{dx}, \ 1 \leq p, q \leq n_i, \ i = 0, 1 \tag{6.56}$$

then we define $\tilde{A}_0$ as

$$\tilde{A}_0 = \tilde{K}_0 + \eta \tilde{M}_0 + \beta \tilde{R}_0^\intercal \tilde{M}_\Sigma \tilde{R}_0 \tag{6.57}$$

and $\tilde{A}_1 = \tilde{A}_1(S_1)$ as

$$\tilde{A}_1(S_1) = \tilde{K}_1(S_1) + \eta \tilde{M}_1 + \beta \tilde{R}_1^\intercal \tilde{M}_\Sigma \tilde{R}_1. \tag{6.58}$$

In addition, we introduce $\tilde{H} \in \mathbb{R}^{n_\Sigma \times n_i}$ the interpolation matrix from $\mathcal{V}(\Omega_0)$ into $\mathcal{V}(\Gamma)$. Then

$$\tilde{H} = \tilde{M}_\Gamma^{-1} \tilde{M}_{0;\Gamma} \tag{6.59}$$

where $\tilde{M}_{0;\Gamma} \in \mathbb{R}^{n_\Gamma \times n_0}$ is defined by

$$\tilde{M}_{0;\Gamma;p,q} \coloneqq \int_\Gamma \phi_0^q \phi_\Gamma^p \, \mathrm{ds}, \ 1 \leq p \leq n_\Gamma, 1 \leq q \leq n_0. \tag{6.60}$$

We also introduce $b \in \mathbb{R}^{n_0}$ defined by

$$b_p \coloneqq \int_\Gamma f \phi_0^p \, \mathrm{ds}, \ 1 \leq p \leq n_0, \tag{6.61}$$

or equivalently,

$$b = \tilde{H}^\intercal \tilde{M}_\Gamma f.$$

Finally, we consider the matrix $\tilde{Q}_i \in \mathbb{R}^{n_1 \times n_1}$, $i = 1, .., n_{\sigma_1}$ associated with the gradient of the cost functional, which is given by

$$\tilde{Q}_{i;p,q} \coloneqq \int_{\Omega_{1,i}} \nabla \phi_1^q \cdot \nabla \phi_1^p \, \mathrm{dx}, \ 1 \leq p \leq n_1, 1 \leq q \leq n_{\sigma_1}.$$

Now we are ready for discretizing the operators $\mathcal{A}_i$, $i = 0, 1$. The variational formulation for $u = \mathcal{A}_0(\lambda)$ in (6.26) is given by

$$\int_{\Omega_0} \sigma_0 \nabla u \cdot \nabla v \, \mathrm{dx} + \int_{\Omega_0} \eta u v \, \mathrm{dx} + \int_\Sigma \beta u v \, \mathrm{ds} = \int_\Gamma f v \, \mathrm{ds} - \int_\Sigma \lambda v \, \mathrm{ds}, \ \forall v \in H^1(\Omega_0).$$

Therefore, the discrete equivalent of the operator $\mathcal{A}_0$ is the application $\Lambda \in \mathbb{R}^{n_\Sigma} \mapsto U \in \mathbb{R}^{n_0}$ such that

$$\tilde{A}_0 U = \tilde{H}^\intercal \tilde{M}_\Gamma f - \tilde{R}_0^\intercal \tilde{M}_\Sigma \Lambda = -\tilde{M}_{0;\Sigma}^\intercal \Lambda + \tilde{H}^\intercal \tilde{M}_\Gamma f.$$

With an abuse of notation, we shall set in the following

$$\boxed{\mathcal{A}_0(\Lambda) \coloneqq -\tilde{A}_0^{-1} \tilde{M}_{0;\Sigma}^\intercal \Lambda + \tilde{A}_0^{-1} \tilde{H}^\intercal \tilde{M}_\Gamma f, \ \forall \Lambda \in \mathbb{R}^{n_\Sigma}.} \tag{6.62}$$

The variational formulation for $u = \mathcal{A}_1(\lambda, \sigma_1)$ in (6.27) is given by

$$\int_{\Omega_1} \sigma_1 \nabla u \cdot \nabla v \, \mathrm{dx} + \int_{\Omega_1} \eta u v \, \mathrm{dx} + \int_\Sigma \beta u v \, \mathrm{ds} = \int_\Sigma \lambda v \, \mathrm{ds}, \ \forall v \in H^1(\Omega_1).$$

Therefore, the discrete equivalent of the operator $\mathcal{A}_1$ is the application $(\Lambda, S_1) \in \mathbb{R}^{n_\Sigma} \times \mathbb{R}^{n_{\sigma_1}} \mapsto U \in \mathbb{R}^{n_0}$ such that

$$\tilde{A}_1(S_1)U = \tilde{R}_1^\intercal \tilde{M}_\Sigma \Lambda = \tilde{M}_{1;\Sigma}^\intercal \Lambda.$$

With an abuse of notation, we shall set in the following

$$\boxed{\mathcal{A}_1(\Lambda, S_1) \coloneqq \tilde{A}_1(S_1)^{-1} \tilde{M}_{1;\Sigma}^\intercal \Lambda, \ \forall (\Lambda, S_1) \in \mathbb{R}^{n_\Sigma} \times \mathbb{R}^{n_{\sigma_1}}.} \tag{6.63}$$

Finally, the variational formulation for $u = \hat{\mathcal{A}}_0(\lambda, \psi)$ in (6.40) is given by

$$\int_{\Omega_0} \sigma_0 \nabla u \cdot \nabla v \, \mathrm{d}x + \int_{\Omega_0} \eta u v \, \mathrm{d}x + \int_\Sigma \beta u v \, \mathrm{d}s = \int_\Gamma \psi v \, \mathrm{d}s - \int_\Sigma \lambda v \, \mathrm{d}s, \ \forall v \in H^1(\Omega_0).$$

Therefore, the discrete equivalent of the operator $\hat{\mathcal{A}}_0$ is the application $(\Lambda, \Psi) \in \mathbb{R}^{n_\Sigma} \times \mathbb{R}^{n_\Gamma} \mapsto U \in \mathbb{R}^{n_0}$ such that

$$\tilde{A}_0 U = \tilde{H}^\intercal \tilde{M}_\Gamma \Psi - \tilde{R}_0^\intercal \tilde{M}_\Sigma \Lambda = \tilde{H}^\intercal \tilde{M}_\Gamma \Psi - \tilde{M}_{0;\Sigma}^\intercal \Lambda.$$

With an abuse of notation, we shall set in the following

$$\boxed{\hat{\mathcal{A}}_0(\Lambda, \Psi) \coloneqq -\tilde{A}_0^{-1} \tilde{M}_{0;\Sigma}^\intercal \Lambda + \tilde{A}_0^{-1} \tilde{H}^\intercal \tilde{M}_\Gamma \Psi, \ \forall (\Lambda, \Psi) \in \mathbb{R}^{n_\Sigma} \times \mathbb{R}^{n_\Gamma}.} \tag{6.64}$$

Then, similarly to (6.31), with an abuse of notation, we define the operator $\mathcal{A} : \{\mathbb{R}^{n_\Sigma} \times \mathbb{R}^{n_\Sigma}\} \times \mathbb{R}^{n_{\sigma_1}} \to \{\mathbb{R}^{n_\Sigma} \times \mathbb{R}^{n_\Sigma}\}$ by

$$\mathcal{A}(\boldsymbol{\Lambda}, S_1) \coloneqq (\boldsymbol{\Lambda_1} - 2\beta \tilde{R}_1 \mathcal{A}_1(\boldsymbol{\Lambda_1}, S_1), \boldsymbol{\Lambda_0} + 2\beta \tilde{R}_0 \mathcal{A}_0(\boldsymbol{\Lambda_0})) \tag{6.65}$$

where $\boldsymbol{\Lambda} = (\boldsymbol{\Lambda_0}, \boldsymbol{\Lambda_1}) \in \{\mathbb{R}^{n_\Sigma} \times \mathbb{R}^{n_\Sigma}\}$ and $S_1 \in \mathbb{R}^{n_{\sigma_1}}$. This definition can be written in a more explicit way as

$$\boxed{\mathcal{A}(\boldsymbol{\Lambda}, S_1) = \mathbb{B}(S_1)\boldsymbol{\Lambda} + \mathbb{E}f} \tag{6.66}$$

where

$$\mathbb{B}(S_1) \coloneqq \begin{bmatrix} 0 & I - 2\beta \tilde{R}_1 \tilde{A}_1(S_1)^{-1} \tilde{M}_{1;\Sigma}^\intercal \\ I - 2\beta \tilde{R}_0 \tilde{A}_0^{-1} \tilde{M}_{0;\Sigma}^\intercal & 0 \end{bmatrix} \in \mathbb{R}^{2n_\Sigma \times 2n_\Sigma} \tag{6.67}$$

and

$$\mathbb{E} \coloneqq \begin{bmatrix} 0 \\ 2\beta \tilde{R}_0 \tilde{A}_0^{-1} \tilde{H}^\intercal \tilde{M}_\Gamma \end{bmatrix} \in \mathbb{R}^{2n_\Sigma \times n_\Gamma}. \tag{6.68}$$

Also, similarly to (6.44), with an abuse of notation, we define the operator $\hat{\mathcal{A}} : \{\mathbb{R}^{n_\Sigma} \times \mathbb{R}^{n_\Sigma}\} \times \mathbb{R}^{n_{\sigma_1}} \times \mathbb{R}^{n_\Gamma} \to \{\mathbb{R}^{n_\Sigma} \times \mathbb{R}^{n_\Sigma}\}$ by

$$\hat{\mathcal{A}}(\boldsymbol{\Lambda}, S_1, \Psi) \coloneqq (\boldsymbol{\Lambda_1} - 2\beta \tilde{R}_1 \mathcal{A}_1(\boldsymbol{\Lambda_1}, S_1), \boldsymbol{\Lambda_0} + 2\beta \tilde{R}_0 \hat{\mathcal{A}}_0(\boldsymbol{\Lambda_0}, \Psi)) \tag{6.69}$$

where $\boldsymbol{\Lambda} = (\boldsymbol{\Lambda_0}, \boldsymbol{\Lambda_1}) \in \{\mathbb{R}^{n_\Sigma} \times \mathbb{R}^{n_\Sigma}\}$. This definition can be written in a more explicit way as

$$\boxed{\hat{\mathcal{A}}(\boldsymbol{\Lambda}, S_1, \Psi) = \mathbb{B}(S_1)\boldsymbol{\Lambda} + \mathbb{E}\Psi} \tag{6.70}$$

where $\mathbb{B}$ and $\mathbb{E}$ are respectively defined in (6.67) and (6.68).

6.2.2 Discretized version of algorithm (6.48)

Let $\Lambda_i^n \in \mathbb{R}^{n_\Sigma}$, $\hat{\Lambda}_i^n \in \mathbb{R}^{n_\Sigma}$, $i = 0, 1$ and $G \in \mathbb{R}^{n_\Gamma}$ be the column vectors associated respectively with λ_i^n, $\hat{\lambda}_i^n$, $i = 0, 1$ and $g \in \mathcal{V}(\Gamma)$ as in (6.47). Let us set

$$\Lambda^n = (\Lambda_0^n, \Lambda_1^n) \text{ and } \hat{\Lambda}^n = (\hat{\Lambda}_0^n, \hat{\Lambda}_1^n).$$

for $n \geq 0$. Then the 1-step one-shot algorithm (6.47) with the gradient defined in (6.4) can be synthetically written in the discrete form as

$$\begin{cases} S_{1,i}^{n+1} = S_{1,i}^n - \tau \mathcal{A}_1(\hat{\Lambda}_1^n, S_1^n)^\intercal Q_i \mathcal{A}_1(\Lambda^n, S_1^n), \ 1 \leq i \leq n_{\sigma_1}, \\ \Lambda^{n+1} = \mathcal{A}(\Lambda^n, S_1^{n+1}), \\ \hat{\Lambda}^{n+1} = \hat{\mathcal{A}}(\hat{\Lambda}^n, S_1^{n+1}, G - \tilde{H}\mathcal{A}_0(\Lambda_0^n)), \end{cases} \tag{6.71}$$

where the operators $\mathcal{A}_0$, $\mathcal{A}_1$, $\mathcal{A}$ and $\hat{\mathcal{A}}$ are respectively defined by (6.62), (6.63), (6.66) and (6.70). Indeed, using the expression of $\mathcal{A}$ in (6.66), we have

$$\mathcal{A}(\Lambda^n, S_1^{n+1}) = \mathbb{B}(S_1^{n+1})\Lambda^n + \mathbb{E}f.$$

Next, using the expressions of $\mathcal{A}_0$ in (6.62) and of $\hat{\mathcal{A}}$ in (6.70), we have

$$\begin{aligned} \hat{\mathcal{A}}(\hat{\Lambda}^n, S_1^{n+1}, G - \tilde{H}\mathcal{A}_0(\Lambda_0^n)) = &\ \mathbb{B}(S_1^{n+1})\hat{\Lambda}^n + \mathbb{E}(G + \tilde{H}\tilde{A}_0^{-1}\tilde{M}_{0;\Sigma}^\intercal \Lambda_0^n - \tilde{H}\tilde{A}_0^{-1}\tilde{H}^\intercal \tilde{M}_\Gamma f) \\ = &\ \mathbb{B}\hat{\Lambda}^n + \mathbb{D}\Lambda^n + \mathbb{E}G - \mathbb{F}f \end{aligned}$$

where

$$\begin{aligned} \mathbb{D} := \mathbb{E}\underbrace{\begin{bmatrix} \tilde{H}\tilde{A}_0^{-1}\tilde{M}_{0;\Sigma}^\intercal & 0 \end{bmatrix}}_{\in \mathbb{R}^{n_\Gamma \times 2n_\Sigma}} &= \begin{bmatrix} 0 \\ 2\beta \tilde{R}_0 \tilde{A}_0^{-1}\tilde{H}^\intercal \tilde{M}_\Gamma \end{bmatrix} \begin{bmatrix} \tilde{H}\tilde{A}_0^{-1}\tilde{M}_{0;\Sigma}^\intercal & 0 \end{bmatrix} \\ &= \begin{bmatrix} 0 & 0 \\ 2\beta \tilde{R}_0 \tilde{A}_0^{-1}\tilde{H}^\intercal \tilde{M}_\Gamma \tilde{H}\tilde{A}_0^{-1}\tilde{M}_{0;\Sigma}^\intercal & 0 \end{bmatrix} \in \mathbb{R}^{2n_\Sigma \times 2n_\Sigma} \end{aligned} \tag{6.72}$$

and

$$\begin{aligned} \mathbb{F} := \mathbb{E}\underbrace{\tilde{H}\tilde{A}_0^{-1}\tilde{H}^\intercal \tilde{M}_\Gamma}_{\in \mathbb{R}^{n_\Gamma \times n_\Gamma}} &= \begin{bmatrix} 0 \\ 2\beta \tilde{R}_0 \tilde{A}_0^{-1}\tilde{H}^\intercal \tilde{M}_\Gamma \end{bmatrix} \tilde{H}\tilde{A}_0^{-1}\tilde{H}^\intercal \tilde{M}_\Gamma \\ &= \begin{bmatrix} 0 \\ -2\beta \tilde{R}_0 \tilde{A}_0^{-1}\tilde{H}^\intercal \tilde{M}_\Gamma \tilde{H}\tilde{A}_0^{-1}\tilde{H}^\intercal \tilde{M}_\Gamma \end{bmatrix} \in \mathbb{R}^{2n_\Sigma \times n_\Gamma}. \end{aligned} \tag{6.73}$$

Also, using the definition of $\mathcal{A}_1$ in (6.63), we can rewrite

$$\mathcal{A}_1(\hat{\Lambda}_1^n, S_1^n)^\intercal Q_i \mathcal{A}_1(\Lambda^n, S_1^n) = (\hat{\Lambda}_1^n)^\intercal \mathbb{L}_i(S_1^n)\Lambda^n, \ 1 \leq i \leq n_{\sigma_1}$$

where

$$\mathbb{L}_i := (\tilde{A}_1(S_1)^{-1}\tilde{M}_{1;\Sigma}^\intercal)^\intercal Q_i \tilde{A}_1(S_1)^{-1}\tilde{M}_{1;\Sigma}^\intercal \in \mathbb{R}^{2n_\Sigma \times 2n_\Sigma}, \ 1 \leq i \leq n_{\sigma_1}. \tag{6.74}$$

With these expressions, scheme (6.71) can be more explicitly written as

$$\boxed{\begin{cases} S_{1,i}^{n+1} = S_{1,i}^n - \tau(\hat{\Lambda}_1^n)^\intercal \mathbb{L}_i(S_1^n)\Lambda^n, \ 1 \leq i \leq n_{\sigma_1}, \\ \Lambda^{n+1} = \mathbb{B}(S_1^{n+1})\Lambda^n + \mathbb{E}f, \\ \hat{\Lambda}^{n+1} = \mathbb{B}(S_1^{n+1})\hat{\Lambda}^n + \mathbb{D}\Lambda^n + \mathbb{E}G - \mathbb{F}f, \end{cases}} \tag{6.75}$$

where $\mathbb{B}$, $\mathbb{D}$, $\mathbb{E}$, $\mathbb{F}$ and $\mathbb{L}_i$ are respectively defined by (6.67), (6.72), (6.68), (6.73) and (5.70). We remark that the structure of algorithm (6.75) is different from the framework introduced in Section 1.2. This is obvious due to the fact that we are working with a non-linear problem.

We end this section by indicating the k-step version of algorithm (6.48). Let $\Lambda_i^n \in \mathbb{R}^{n_\Sigma}$, $\hat{\Lambda}_i^n \in \mathbb{R}^{n_\Sigma}$, $i = 0, 1$ and $G \in \mathbb{R}^{n_\Gamma}$ be the column vectors associated respectively with λ_i^n, $\hat{\lambda}_i^n$, $i = 0, 1$ and $g \in \mathcal{V}(\Gamma)$ as in (6.47). Let us set

$$\Lambda^n = (\Lambda_0^n, \Lambda_1^n), \Lambda_{[\ell]}^n = (\Lambda_{0;\ell}^n, \Lambda_{1;\ell}^n), \hat{\Lambda}^n = (\hat{\Lambda}_0^n, \hat{\Lambda}_1^n) \text{ and } \hat{\Lambda}_{[\ell]}^n = (\hat{\Lambda}_{0;\ell}^n, \hat{\Lambda}_{1;\ell}^n).$$

for $n \geq 0$ and $\ell \geq 0$. Then, the k-step one-shot algorithm (6.48) with the gradient defined in (6.4) can be synthetically written in the discrete form as

$$\begin{cases} S_{1,i}^{n+1} = S_{1,i}^n - \tau \hat{\mathcal{A}}_1(\hat{\Lambda}_1^n, S_1^n)^\intercal Q_i \mathcal{A}_1(\Lambda^n, S_1^n), \ 1 \leq i \leq n_{\sigma_1}, \\ \Lambda_{[0]}^{n+1} = \Lambda^n, \hat{\Lambda}_{[0]}^{n+1} = \hat{\Lambda}^n, \\ \text{for } \ell = 0, 1, ..., k-1 : \\ \quad \left| \begin{array}{l} \Lambda_{[\ell+1]}^{n+1} = \mathcal{A}(\Lambda_{[\ell]}^n, S_1^{n+1}), \\ \hat{\Lambda}_{[\ell+1]}^{n+1} = \hat{\mathcal{A}}(\hat{\Lambda}_{[\ell]}^n, S_1^{n+1}, G - \tilde{H}\mathcal{A}_0(\Lambda_{[\ell]}^n)), \end{array} \right. \\ \Lambda^{n+1} = \Lambda_{[k]}^{n+1}, \hat{\Lambda}^{n+1} = \hat{\Lambda}_{[k]}^{n+1}, \end{cases} \tag{6.76}$$

where the operators $\mathcal{A}_0$, $\mathcal{A}_1$, $\mathcal{A}$ and $\hat{\mathcal{A}}$ are respectively defined by (6.62), (6.63), (6.66) and (6.70). With these expression, the scheme can be more explicitly written as:

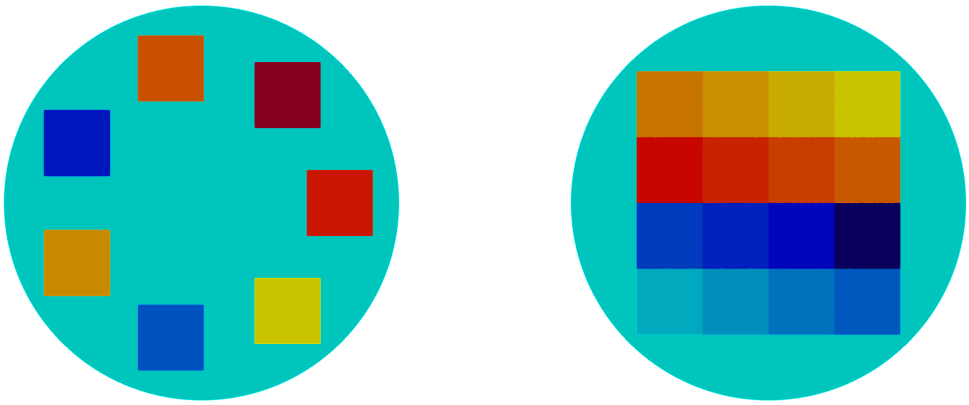
$$\begin{cases} S_{1,i}^{n+1} = S_{1,i}^n - \tau(\hat{\Lambda}_1^n)^\intercal \mathbb{L}_i(S_1^n)\Lambda^n, \ 1 \leq i \leq n_{\sigma_1}, \\ \Lambda_0^{n+1} = \Lambda^n, \hat{\Lambda}_0^{n+1} = \hat{\Lambda}^n, \\ \text{for } \ell = 0, 1, \ldots, k-1 : \\ \quad \left| \begin{array}{l} \Lambda_{[\ell+1]}^{n+1} = \mathbb{B}(S_1^{n+1})\Lambda_{[\ell]}^{n+1} + \mathbb{E}f, \\ \hat{\Lambda}_{[\ell+1]} = \mathbb{B}(S_1^{n+1})\hat{\Lambda}_{[\ell]}^{n+1} + \mathbb{D}\Lambda_{[\ell]}^{n+1} + \mathbb{E}G - \mathbb{F}f, \end{array} \right. \\ \Lambda^{n+1} := \Lambda_{[k]}^{n+1}, \hat{\Lambda}^{n+1} := \hat{\Lambda}_{[k]}^{n+1} \end{cases} \tag{6.77}$$

where $\mathbb{B}(.)$, $\mathbb{D}$, $\mathbb{E}$, $\mathbb{F}$ and $\mathbb{L}_i$ are respectively defined by (6.67), (6.72), (6.68), (6.73) and (6.74).

6.3 Numerical experiments

We compare the performance of the classical gradient descent algorithm (with direct solvers for the forward and adjoint problems) with the domain decomposition k-step one-shot algorithm (6.48) to solve the Neumann non-linear inverse conductivity problem with Dirichlet measurements (1.38) (see also Section 6.1.2). We focus on the case where the unknown conductivity σ is defined by (6.16) in the global domain Ω and similarly by (6.49) in the inner subdomain Ω_1, that is, the inverse problem has $M = n_{\sigma_1}$ scalar unknowns, while the background value σ_0 is known. These unknowns are iteratively updated as in (6.17) and (6.50).

The computational domain Ω is a disk, and we consider two configurations for the conductivity: in the first one the unknowns $\sigma_1, \ldots, \sigma_M$ are the values on $M = 7$ disjoint squares (see Figure 6.2, left), while in the second one the unknowns $\sigma_1, \ldots, \sigma_M$ are the values on $M = 16$ touching squares (see Figure 6.2, right). In the domain decomposition



(a) (b)

Figure 6.2: Two configurations for the scalar unknowns defining the conductivity σ on the computational domain.

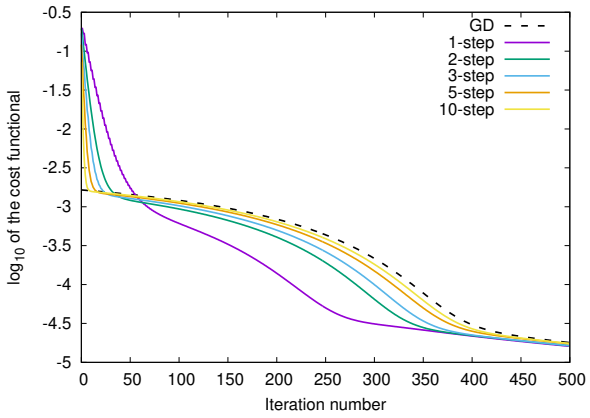
iterations, described in Section 6.1.3, the inner subdomain Ω_1 is then the union of these $n_{\sigma_1} = M$ squares. The background subdomain is $\Omega_0 = \Omega \setminus \overline{\Omega}_1$.

In all the experiments, we take $\sigma_0 = 1$, $\eta = 10^{-6}$, and N_{m} Neumann data $f_m(x, y) = m \sin(m \operatorname{atan2}(y, x))$, with $m = 1, \ldots, N_{\mathrm{m}}$, in the forward problem (6.1). The corresponding N_{m} synthetic Dirichlet measurements g_m are built on a coarser mesh with respect to the one where the forward and adjoint problems are solved, in order to avoid an inverse crime. The different values of N_{m} and mesh sizes will be specified below. Both the state and adjoint state are discretized with $\mathbb{P}^1$ Lagrange finite elements. In the domain decomposition (inner) iterations, we take $\beta = 10$ for the impedance parameter in the transmission conditions on the interface Σ between the two subdomains Ω_0 and Ω_1 (which are non-overlapping). The substructured iteration variables λ and $\hat{\lambda}$ are discretized with $\mathbb{P}^1$ Lagrange finite elements on a mesh of the interface Σ. The (outer) iteration on the inverse problem unknowns is stopped after 500 iterations.

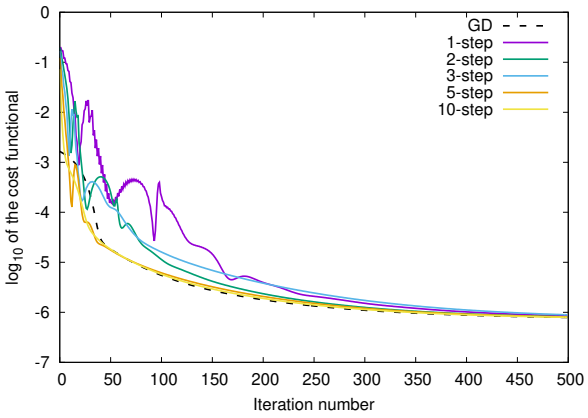
In the next two paragraphs we present the results obtained respectively for the two configurations in Figure 6.2. In the convergence plots, we use solid lines for Domain Decomposition (DD) k-step one-shot methods and dashed lines for the classical Gradient Descent (GD) algorithm. We plot both the cost functional in $\log_{10}$ scale (top rows in the figures below) and the relative error on the conductivity (bottow rows), with respect to the outer iteration number.

First configuration (disjoint squares) Here the computational domain Ω is a disk of radius 2. The unknown values of σ are supported on 7 disjoint squares of side 2/3, whose centers have distance 0.6 from the disk boundary (see Figure 6.2, left). The mesh size of the forward and adjoint problems mesh is $h = 1/20$, which gives $n_u = 5625$ degrees of freedom, while the mesh size of the synthetic measurements mesh is $h_{\mathrm{m}} = 2h$, which gives $n_g = 125$ (surface) degrees of freedom for each measurement. The exact conductivity that we wish to reconstruct is given by $\sigma_1 = 10$, $\sigma_2 = 9$, $\sigma_3 = 12$, $\sigma_4 = 7$, $\sigma_5 = 14$, $\sigma_6 = 5$, $\sigma_7 = 16$, and as initial guess we choose $\sigma_i^0 = 20$ for all $i = 1, \ldots, 7$.

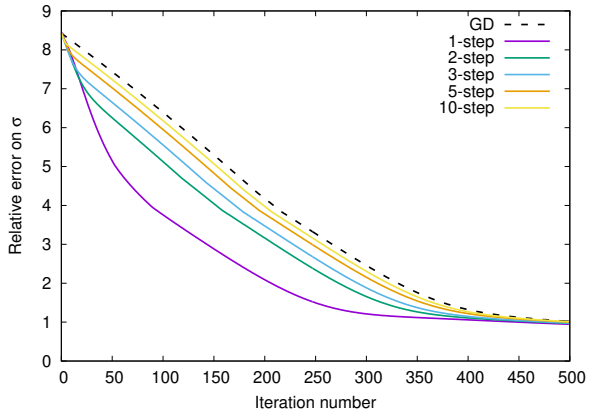
We first study in Figure 6.3 the influence on the convergence of the number k of domain decomposition iterations, taking just one measurement ($N_{\mathrm{m}} = 1$) for now. As expected, as k increases the convergence curves of DD k-step one-shot methods tend to the one of classical GD. In this experiment, for 'small' descent steps such as $\tau = 100$ (Figure 6.3,



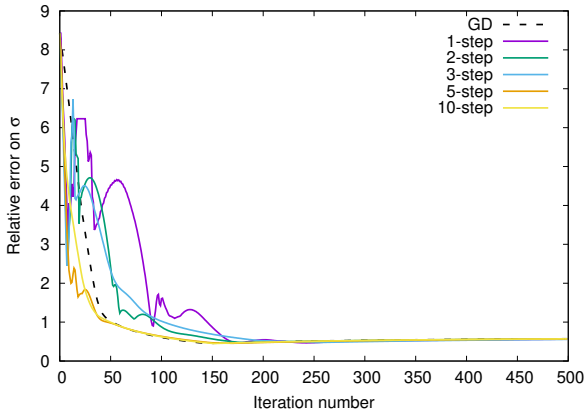
(a) Cost functional, $N_{\mathrm{m}} = 1$, $\tau = 100$.



(b) Cost functional, $N_{\mathrm{m}} = 1$, $\tau = 1000$.



(c) Relative error on σ, $N_{\mathrm{m}} = 1$, $\tau = 100$.



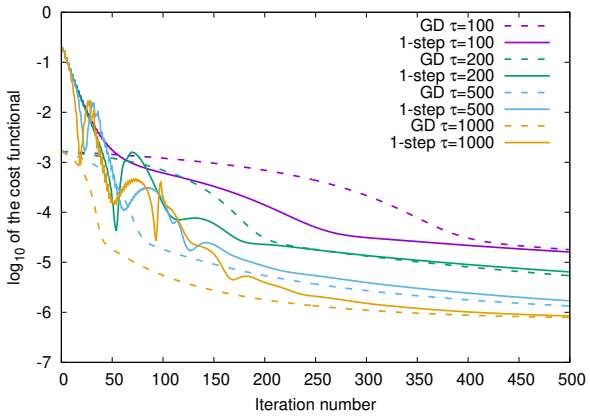
(d) Relative error on σ, $N_{\mathrm{m}} = 1$, $\tau = 1000$.

Figure 6.3: For the first configuration, convergence curves (cost functional, top, and relative error on σ, bottom) of gradient descent (dashed line) and domain decomposition k-step one-shot (solid lines): dependence on the number of inner iterations k.

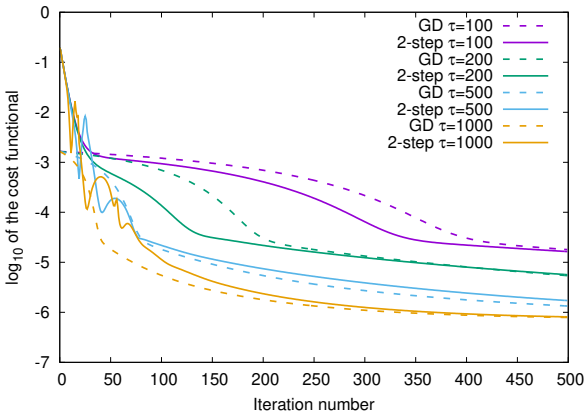
left), the DD k-step one-shot methods give even better relative errors than GD. For larger descent steps such as $\tau = 1000$ (Figure 6.3, right), oscillations appear for small k in the first part of the convergence history. In all the considered cases, the DD k-step one-shot methods are convergent, even with just one domain decomposition iteration and large descent step $\tau = 1000$.

In Figure 6.4 we fix the number of domain decomposition iterations to $k = 1$ (left) and $k = 2$ (right) and vary the descent step τ to study intermediate values with respect to those chosen in the previous Figure 6.3. We observe that, when τ increases, the DD 1-step and 2-step one-shot methods are affected by oscillations, which appear sooner for the 1-step method: for instance, $\tau = 200$ gives oscillations for the 1-step method, but not for the 2-step method. However, in all the cases, the asymptotic behavior of DD k-step one-shot methods after several outer iterations is like the one of the corresponding classical GD method.

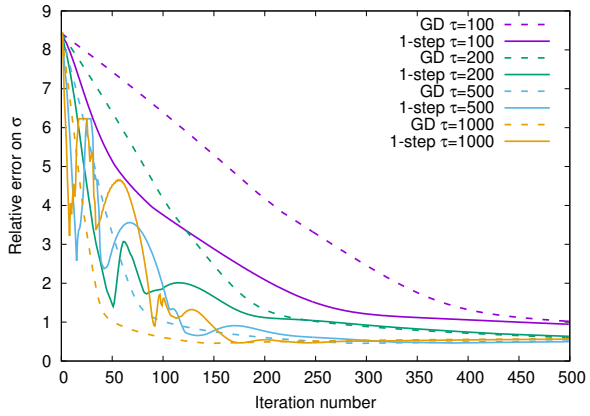
We now test the performance of the DD k-step one-shot methods for more than one measurement: in Figure 6.5 we vary $N_{\mathrm{m}} = 1, 2, 3$, and take just one domain decomposition iteration ($k = 1$). As expected, increasing the number of measurements improve the relative error curves, which with higher N_{m} become steeper in the first part of the convergence history. It is worth noticing that in the implementation of the DD k-step one-shot algorithm (6.48) (with (6.50) replacing the first line) in the case of several measurements, for



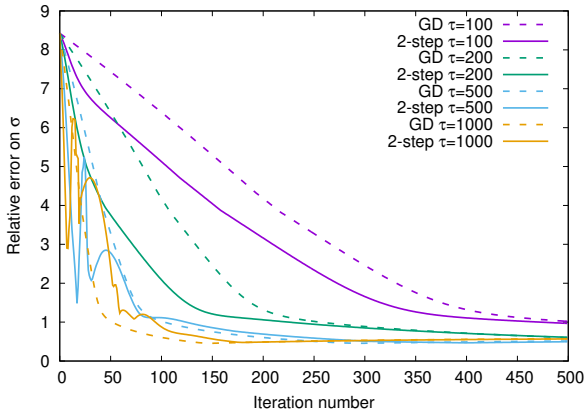
(a) Cost functional, $N_{\mathrm{m}} = 1$, $k = 1$.



(b) Cost functional, $N_{\mathrm{m}} = 1$, $k = 2$.



(c) Relative error on σ, $N_{\mathrm{m}} = 1$, $k = 1$.



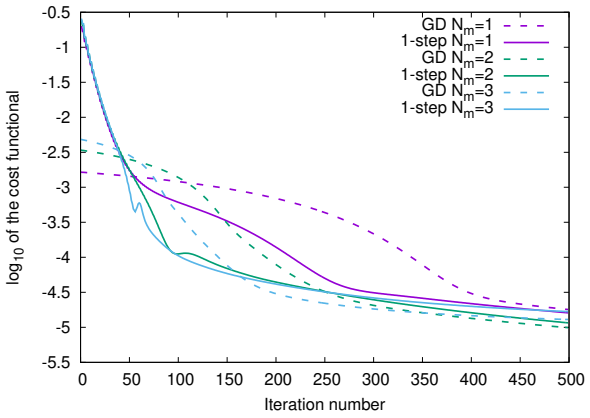
(d) Relative error on σ, $N_{\mathrm{m}} = 1$, $k = 2$.

Figure 6.4: For the first configuration, convergence curves (cost functional, top, and relative error on σ, bottom) of gradient descent (dashed lines) and domain decomposition k-step one-shot (solid lines): dependence on the descent step τ.

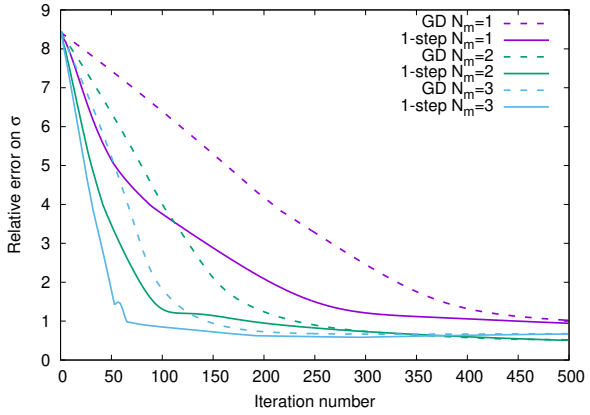
each measurement we need to store $(\lambda_{[k]}^{n+1}, \hat{\lambda}_{[k]}^{n+1})$ so that, after updating the conductivity, they can be taken as initial guesses for the next domain decomposition iterations for the corresponding measurement.

Second configuration (touching squares) Here the computational domain Ω is a disk of radius $3/2$. The unknown values of σ are supported on 16 touching squares of side $1/2$ (see Figure 6.2, right). The mesh size of the forward and adjoint problems mesh is $h = 1/20$, which gives $n_u = 3285$ degrees of freedom, while the mesh size of the synthetic measurements mesh is $h_{\mathrm{m}} = 1.5\,h$, which gives $n_g = 125$ (surface) degrees of freedom for each measurement.

Figure 6.6 shows the exact conductivity to be reconstructed: indeed, in this second configuration we wish to test whether the algorithm is able to distinguish in Ω_1 the regions where the exact conductivity is equal to the background value $\sigma_0 = 1$, from the regions where it is different from σ_0 and equal to 5. As initial guess we choose $\sigma_i^0 = \sigma_0$ (the background value) for all $i = 1, \ldots, 16$. We represent in Figure 6.7 the reconstructed conductivity after 500 iterations of the classical GD method (left) and of the DD 1-step one-shot method (right), with descent step $\tau = 4$ and $N_{\mathrm{m}} = 3$ measurements. We can observe that the reconstructions obtained with the two methods are similar, with slightly more dispersed values in the case of the DD 1-step one-shot method.



(a) Cost functional, $k = 1$, $\tau = 100$.



(b) Relative error on σ, $k = 1$, $\tau = 100$.

Figure 6.5: For the first configuration, convergence curves (cost functional, top, and relative error on σ, bottom) of gradient descent (dashed lines) and domain decomposition k-step one-shot (solid lines): dependence on the number of measurements N_{m}.

We perform also in this second configuration the experiments presented for the first configuration in the previous paragraph (we only adapt the range for the choice of the descent step). First, in Figure 6.8 we vary the number k of domain decomposition iterations, at fixed descent steps ($\tau = 0.5$, left, and $\tau = 4$, right) and with one measurement ($N_{\mathrm{m}} = 1$). We can see that as k increases the relative error curves of DD k-step one-shot methods tend to the one of classical GD as expected, while this is not the case for the cost functional curves for which there is a large gap with respect to the one of classical GD. In all the considered cases, even if some oscillations appear initially, the DD k-step one-shot methods are convergent, even with just one domain decomposition iteration and large descent step $\tau = 4$.

In Figure 6.9 we study the influence on the convergence of the descent step τ, and fix the number of domain decomposition iterations to $k = 1$ (left) and $k = 2$ (right). We observe that, when τ increases, the DD 1-step and 2-step one-shot methods are affected by more oscillations, which have a wider amplitude for the 1-step method compared to the 2-step method. However, in all the cases, at least for the relative error plots, the asymptotic behavior of DD k-step one-shot methods after several outer iterations is like the one of the corresponding classical GD method.

Finally, in Figure 6.10 we test the performance of the DD k-step one-shot methods for more than one measurement, by varying $N_{\mathrm{m}} = 1, 2, 3$. We take just one domain

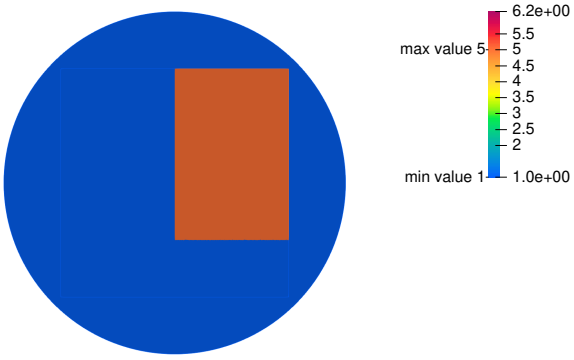
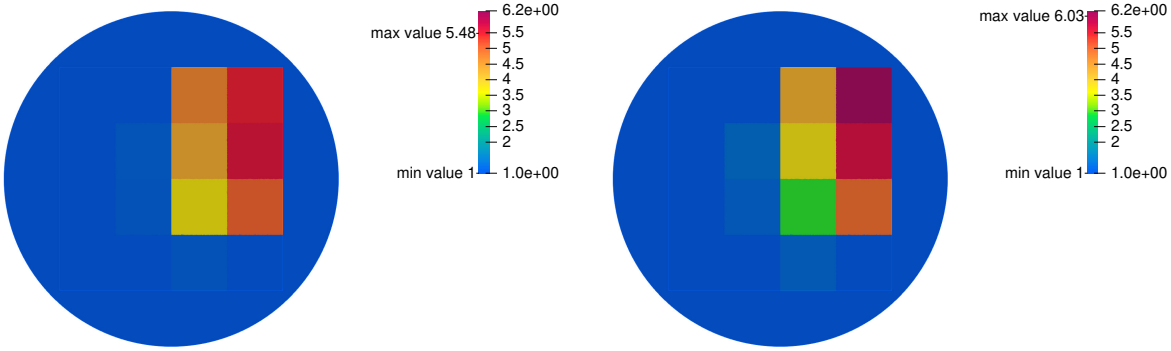


Figure 6.6: The exact conductivity in the second configuration.

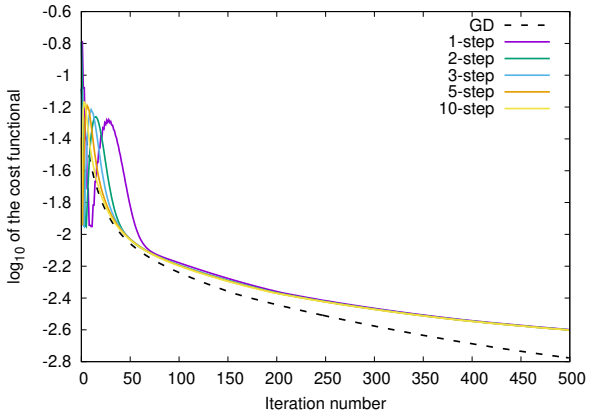


(a) Gradient descent. (b) DD 1-step one-shot.

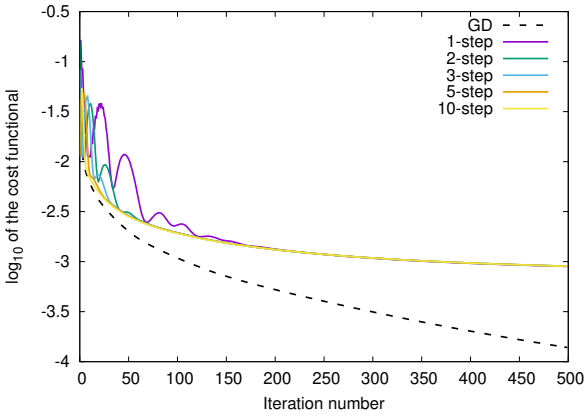
Figure 6.7: Reconstructed conductivity σ in the second configuration, with $\tau = 4$, $N_{\mathrm{m}} = 3$.

decomposition iteration ($k = 1$), and fix $\tau = 0.5$ (Figure 6.10, left) and $\tau = 4$ (Figure 6.10, right). As expected, increasing the number of measurements improve the relative error curves, although, especially for the largest descent step $\tau = 4$, the gain of passing from 2 to 3 measurements is more moderate than the gain of passing from 1 to 2 measurements.

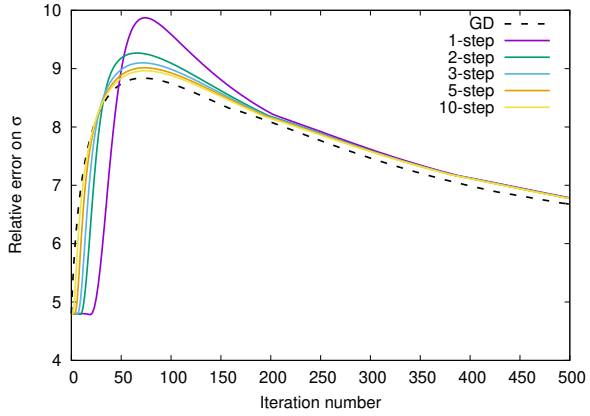
Overall, these promising numerical results for small 2D problems show the potential of domain decomposition one-shot methods in solving large-scale inverse problems, for which it would not be possible to use the classical gradient descent method with a direct solver because of its high memory cost.



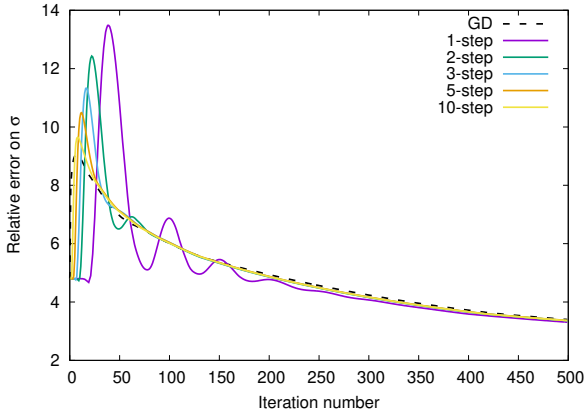
(a) Cost functional, $N_{\mathrm{m}} = 1$, $\tau = 0.5$.



(b) Cost functional, $N_{\mathrm{m}} = 1$, $\tau = 4$.

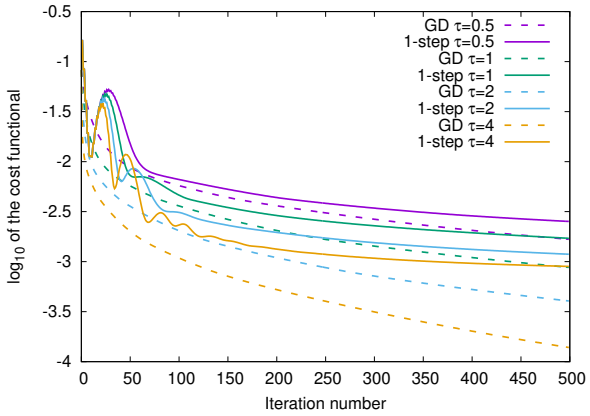


(c) Relative error on σ, $N_{\mathrm{m}} = 1$, $\tau = 0.5$.

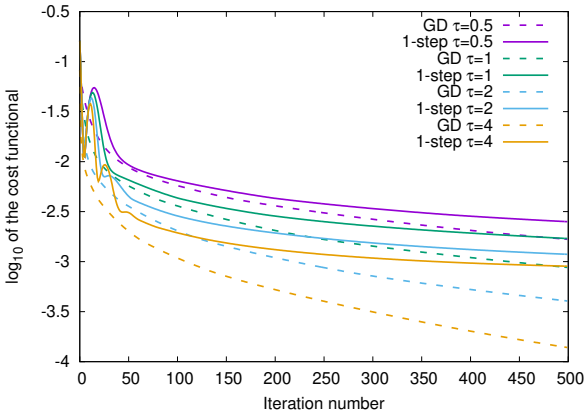


(d) Relative error on σ, $N_{\mathrm{m}} = 1$, $\tau = 4$.

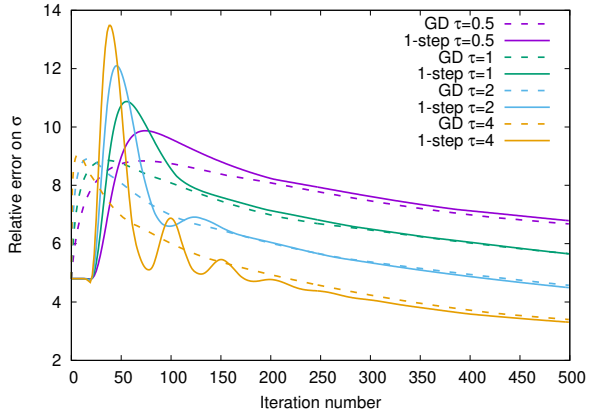
Figure 6.8: For the second configuration, convergence curves (cost functional, top, and relative error on σ, bottom) of gradient descent (dashed line) and domain decomposition k-step one-shot (solid lines): dependence on the number of inner iterations k.



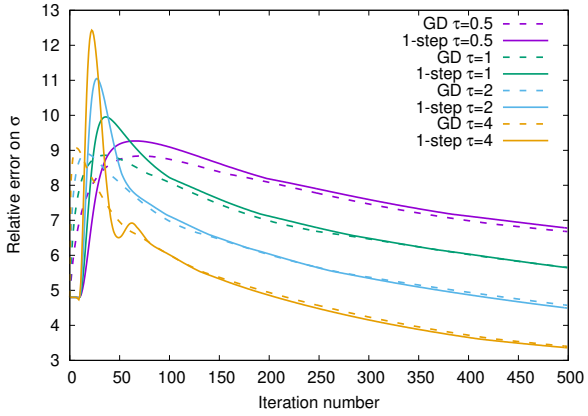
(a) Cost functional, $N_{\mathrm{m}} = 1$, $k = 1$.



(b) Cost functional, $N_{\mathrm{m}} = 1$, $k = 2$.

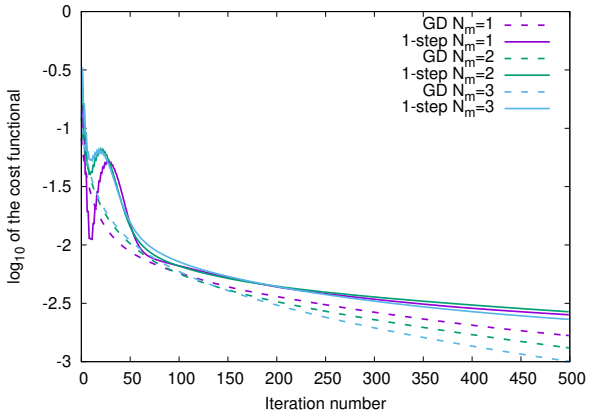


(c) Relative error on σ, $N_{\mathrm{m}} = 1$, $k = 1$.

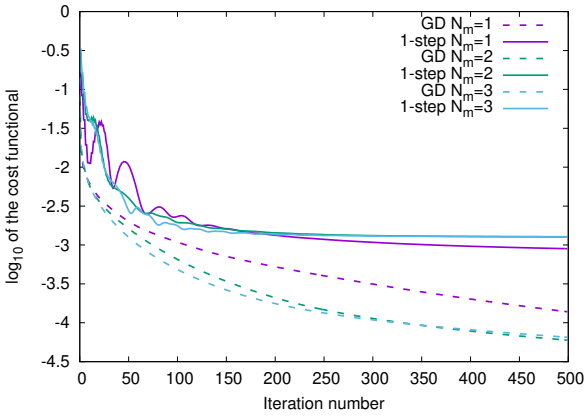


(d) Relative error on σ, $N_{\mathrm{m}} = 1$, $k = 2$.

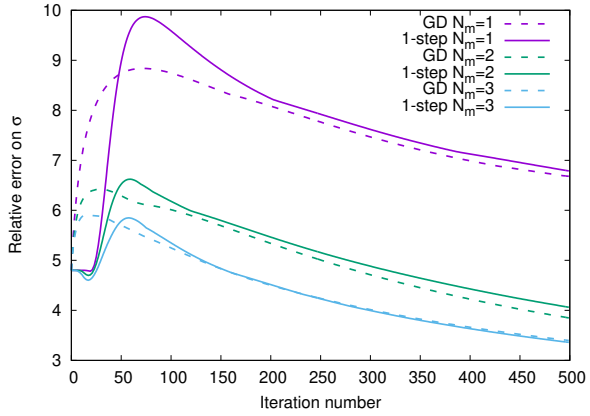
Figure 6.9: For the second configuration, convergence curves (cost functional, top, and relative error on σ, bottom) of gradient descent (dashed lines) and domain decomposition k-step one-shot (solid lines): dependence on the descent step τ.



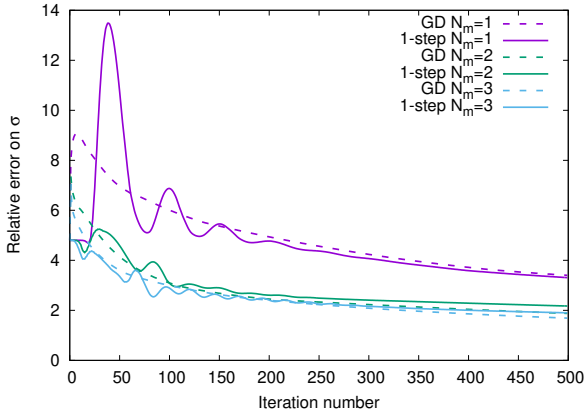
(a) Cost functional, $k = 1$, $\tau = 0.5$.



(b) Cost functional, $k = 1$, $\tau = 4$.



(c) Relative error on σ, $k = 1$, $\tau = 0.5$.



(d) Relative error on σ, $k = 1$, $\tau = 4$.

Figure 6.10: For the second configuration, convergence curves (cost functional, top, and relative error on σ, bottom) of gradient descent (dashed lines) and domain decomposition k-step one-shot (solid lines): dependence on the number of measurements N_{m}.

Appendix 6.A Convergence analysis of (6.46) in the case of a circular domain

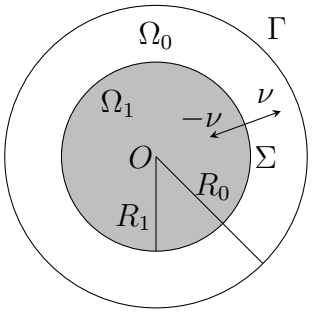


Figure 6.11: Illustration for the domain Ω and its decomposition.

We investigate in the following the convergence of (6.46) in the case of a circular domain. Let (r, θ) be the polar coordinates in $\mathbb{R}^2$. Given $m \in \mathbb{N}^*$, we define

$$V_m(R) \coloneqq \{v : \partial B(0,R) \to \mathbb{C} \text{ such that } v(\theta) = c_v \mathrm{e}^{\mathrm{i}m\theta}, \forall \theta \text{ for some constant } c_v\}, \ \forall R > 0.$$

We set the following assumptions.

Assumption 6.A.1. Ω and Ω_1 are respectively the circles of radii R_0 and R_1, and both are centered at the origin (see Figure 6.11). $\Omega_0 = \Omega \backslash \overline{\Omega}_1$ is the corresponding annulus.

Assumption 6.A.2. $\sigma_0 \in \mathbb{R}, \sigma > 0$ is known and $\sigma = \sigma_0 \mathbb{1}_{\Omega_0} + \sigma_1 \mathbb{1}_{\Omega_1}$ where $\sigma_1 \in \mathbb{R}, \sigma_1 > 0$.

Assumption 6.A.3. The boundary datum f in equation (6.1) belongs to $V_m(R_0)$ and the initial guesses $\lambda_{0;0}, \lambda_{1;0}, \hat{\lambda}_{0;0}, \hat{\lambda}_{1;0}$ of algorithm (6.46) also belong to $V_m(R_1)$, for some $m \in \mathbb{N}^*$.

Let J_α and Y_α respectively denote the α-order Bessel functions of the first and second kinds. By Assumption 6.A.3, given two constants $\sigma_0 > 0$ and $\eta \in \mathbb{R}$, we are interested in the solutions of

$$-\operatorname{div}(\sigma \nabla u) + \eta u = 0$$

in the form

$$(\mathcal{C}_1 A_{m,\sigma}(r) + \mathcal{C}_2 B_{m,\sigma}(r)) \mathrm{e}^{\mathrm{i}m\theta} \quad (\mathcal{C}_1 \text{ and } \mathcal{C}_2 \text{ are constants}) \tag{6.78}$$

where the functions $A_{m,\sigma}$ and $B_{m,\sigma}$ are defined as

$$A_{m,\sigma}(r) \coloneqq \begin{cases} J_m\left(\sqrt{\frac{-\eta}{\sigma}} r\right) & \text{if } \eta < 0, \\ r^m & \text{if } \eta = 0, \\ J_m\left(\mathrm{i}\sqrt{\frac{\eta}{\sigma}} r\right) & \text{if } \eta > 0 \end{cases} \quad \text{and} \quad B_{m,\sigma}(r) \coloneqq \begin{cases} Y_m\left(\sqrt{\frac{-\eta}{\sigma}} r\right) & \text{if } \eta < 0, \\ r^{-m} & \text{if } \eta = 0, \\ Y_m\left(\mathrm{i}\sqrt{\frac{\eta}{\sigma}} r\right) & \text{if } \eta > 0. \end{cases}$$

We note that B_m is not regular at 0. We shall use the solution form (6.78) to give the expressions of the operators $\mathcal{A}$ and $\hat{\mathcal{A}}$ that play an important role in algorithm (6.46).

Lemma 6.A.4. *We make Assumptions 6.A.1–6.A.3 and given $\eta \in \mathbb{R}$. The operators $\hat{\mathcal{A}}_0$, $\mathcal{A}_1$, $\mathcal{A}$ and $\hat{\mathcal{A}}$ respectively defined in (6.40), (6.27),(6.31) and (6.27) can be expressed in some particular cases as follows.*

(i) The operator $\hat{\mathcal{A}}_0$ defined in (6.40) can be expressed as

$$\hat{\mathcal{A}}_0(\lambda,\psi)(r,\theta) = (C_1(r;\beta)c_\lambda + C_2(r;\beta)c_\psi)\mathrm{e}^{\mathrm{i}m\theta} \tag{6.79}$$

for $\lambda = c_\lambda \mathrm{e}^{\mathrm{i}m\theta} \in V_m(R_1)$ and $\psi = c_\psi \mathrm{e}^{\mathrm{i}m\theta} \in V_m(R_0)$, where

$$C_1(r;\beta) \coloneqq \frac{A_{m,\sigma_0}(r)}{(\sigma_0 B'_{m,\sigma_0}(R_1) - \beta B'_{m,\sigma_0}(R_1))D(\beta)} - \frac{A'_{m,\sigma_0}(R_0)B_{m,\sigma_0}(r)}{B'_{m,\sigma_0}(R_0)(\sigma_0 B'_{m,\sigma_0}(R_1) - \beta B'_{m,\sigma_0}(R_1))D(\beta)}, \tag{6.80}$$

$$C_2(r;\beta) \coloneqq \frac{-A_{m,\sigma_0}(r)}{\sigma_0 B'_{m,\sigma_0}(R_0)D(\beta)} + \left(\frac{1}{\sigma_0 B'_{m,\sigma_0}(R_0)} + \frac{A'_{m,\sigma_0}(R_0)}{\sigma_0 B'_{m,\sigma_0}(R_0)^2 D(\beta)}\right) B_{m,\sigma_0}(r)$$

and

$$D(\beta) \coloneqq \frac{\sigma_0 A'_{m,\sigma_0}(R_1) - \beta A_{m,\sigma_0}(R_1)}{\sigma_0 B'_{m,\sigma_0}(R_1) - \beta B_{m,\sigma_0}(R_1)} - \frac{A'_{m,\sigma_0}(R_0)}{B'_{m,\sigma_0}(R_0)}. \tag{6.81}$$

In particular, the operator $\mathcal{A}_0$ defined in (5.20) can be expressed as

$$\mathcal{A}_0(\lambda)(r,\theta) = \hat{\mathcal{A}}_0(\lambda,f)(r,\theta) = (C_1(r;\beta)c_\lambda + C_2(r;\beta)c_f)\mathrm{e}^{\mathrm{i}m\theta} \tag{6.82}$$

for $\lambda = c_\lambda \mathrm{e}^{\mathrm{i}m\theta} \in V_m(R_1)$ and $f = c_f \mathrm{e}^{\mathrm{i}m\theta} \in V_m(R_0)$,

(ii) The operator $\hat{\mathcal{A}}_1$ defined in (6.27) can be expressed as

$$\mathcal{A}_1(\lambda,\sigma_1)(r,\theta) = C_3(r,\sigma_1;\beta)c_\lambda \mathrm{e}^{\mathrm{i}m\theta} \tag{6.83}$$

for $\lambda = c_\lambda \mathrm{e}^{\mathrm{i}m\theta} \in V_m(R_1)$, where

$$C_3(r,\sigma_1;\beta) \coloneqq \frac{A_{m,\sigma_1}(r)}{\sigma_1 A'_{m,\sigma_1}(R_1) + \beta A_{m,\sigma_1}(R_1)}. \tag{6.84}$$

(iii) The operator $\hat{\mathcal{A}}$ defined in (6.44) can be expressed as

$$\hat{\mathcal{A}}(\boldsymbol{\lambda},\sigma_1,\psi) = \mathcal{B}(\sigma_1;\beta)\boldsymbol{\lambda} + \mathcal{E}(\beta)c_\psi \mathrm{e}^{\mathrm{i}m\theta} \tag{6.85}$$

for $\boldsymbol{\lambda} \in V_m(R_1)^2$ and $\psi = c_\psi \mathrm{e}^{\mathrm{i}m\theta} \in V_m(R_0)$, where

$$\mathcal{B}(\sigma_1;\beta) \coloneqq \begin{bmatrix} 0 & 1 - 2\beta C_3(R_1,\sigma_1;\beta) \\ 1 + 2\beta C_1(R_1;\beta) & 0 \end{bmatrix} \tag{6.86}$$

and

$$\mathcal{E}(\beta) \coloneqq \begin{bmatrix} 0 \\ 2\beta C_2(R_1;\beta) \end{bmatrix}.$$

In particular, the operator $\mathcal{A}$ defined in (6.31) can be expressed as

$$\mathcal{A}(\boldsymbol{\lambda},\sigma_1) = \hat{\mathcal{A}}(\boldsymbol{\lambda},\sigma_1,f) = \mathcal{B}(\sigma_1;\beta)\boldsymbol{\lambda} + \mathcal{E}(\beta)c_f \mathrm{e}^{\mathrm{i}m\theta} \tag{6.87}$$

for $\boldsymbol{\lambda} \in V_m(R_1)^2$ and $f = c_f \mathrm{e}^{\mathrm{i}m\theta} \in V_m(R_0)$.

Moreover, the sequences $(\lambda_{[\ell]})_{\ell\geq 0}$ and $(\hat{\lambda}_{[\ell]})_{\ell\geq 0}$ defined by algorithm (6.46) are in $V_m(R_1)^2$, also $\gamma_\Gamma \mathcal{A}_0(\lambda_{0;\ell}) \in V_m(R_0), \forall \ell \geq 0$.

Proof. *(i)* We recall that

$$\hat{\mathcal{A}}_0 : (\lambda, \psi) \in L^2(\Sigma) \times H^{-\frac{1}{2}}(\Gamma) \mapsto u \in H^1(\Omega_0)$$

such that

$$\begin{cases} -\operatorname{div}(\sigma_0 \nabla u) + \eta u = 0 & \text{in } \Omega_0, \\ \sigma_0 \dfrac{\partial u}{\partial \nu} = \psi & \text{on } \Gamma, \\ \sigma_0 \dfrac{\partial u}{\partial \nu} - \beta u = \lambda & \text{on } \Sigma. \end{cases} \qquad \text{(6.40 recalled)}$$

We seek the solution to (6.40) of the form

$$u = (c_1 A_{m,\sigma_0}(r) + c_2 B_{m,\sigma_0}(r)) \mathrm{e}^{\mathrm{i}m\theta}. \tag{6.88}$$

The boundary conditions on Γ and Σ in (6.40) give

$$\begin{cases} \sigma_0(c_1 A'_{m,\sigma_0}(R_0) + c_2 B'_{m,\sigma_0}(R_0)) = c_\psi, \\ \sigma_0(c_1 A'_{m,\sigma_0}(R_1) + c_2 B'_{m,\sigma_0}(R_1)) - \beta(c_1 A_{m,\sigma_0}(R_1) + c_2 B_{m,\sigma_0}(R_1)) = c_\lambda. \end{cases}$$

The first line of this system yields

$$c_2 = -\frac{A'_{m,\sigma_0}(R_0)}{B'_{m,\sigma_0}(R_0)} c_1 + \frac{c_\psi}{\sigma_0 B'_{m,\sigma_0}(R_0)}, \tag{6.89}$$

then inserting this result into the second line and dividing by $\sigma_0 B'_{m,\sigma_0}(R_1) - \beta B_{m,\sigma_0}(R_1) \neq 0$ gives

$$D(\beta) c_1 + \frac{c_\psi}{\sigma_0 B'_{m,\sigma_0}(R_0)} = \frac{c_\lambda}{\sigma_0 B'_{m,\sigma_0}(R_1) - \beta B_{m,\sigma_0}(R_1)},$$

where

$$D(\beta) := \frac{\sigma_0 A'_{m,\sigma_0}(R_1) - \beta A_{m,\sigma_0}(R_1)}{\sigma_0 B'_{m,\sigma_0}(R_1) - \beta B_{m,\sigma_0}(R_1)} - \frac{A'_{m,\sigma_0}(R_0)}{B'_{m,\sigma_0}(R_0)}.$$

We thus obtain

$$c_1 = \frac{c_\lambda}{(\sigma_0 B'_{m,\sigma_0}(R_1) - \beta B'_{m,\sigma_0}(R_1)) D(\beta)} - \frac{c_\psi}{\sigma_0 B'_{m,\sigma_0}(R_0) D(\beta)},$$

Inserting this expression of c_2 into (6.89) leads to

$$\begin{aligned} c_2 = & -\frac{A'_{m,\sigma_0}(R_0)}{B'_{m,\sigma_0}(R_0)} c_1 + \frac{c_\psi}{\sigma_0 B'_{m,\sigma_0}(R_0)} \\ = & -\frac{A'_{m,\sigma_0}(R_0)}{B'_{m,\sigma_0}(R_0)(\sigma_0 B'_{m,\sigma_0}(R_1) - \beta B'_{m,\sigma_0}(R_1)) D(\beta)} c_\lambda \\ & + \left(\frac{1}{\sigma_0 B'_{m,\sigma_0}(R_0)} + \frac{A'_{m,\sigma_0}(R_0)}{\sigma_0 B'_{m,\sigma_0}(R_0)^2 D(\beta)} \right) c_\psi. \end{aligned}$$

Plugging the above results of c_1 and c_2 into (6.88), we obtain the result of the lemma.
(ii) For $\sigma = \sigma_1$ constant,

$$\mathcal{A}_1 : (\lambda, \sigma_1) \in L^2(\Sigma) \times L^\infty(\Omega_1) \mapsto u \in H^1(\Omega_1)$$

such that

$$\begin{cases} -\operatorname{div}(\sigma_1 \nabla u) + \eta u = 0 & \text{in } \Omega_1, \\ \sigma_1 \dfrac{\partial u}{\partial \nu} + \beta u = \lambda & \text{on } \Sigma. \end{cases} \qquad \text{(6.27 recalled with } \sigma = \sigma_1)$$

Since $\lambda = c_\lambda \mathrm{e}^{\mathrm{i}m\theta} \in V_m(R_1)$, we seek the solution to (6.27) of the form $u = c_3 A_{m,\sigma_1}(r)\mathrm{e}^{\mathrm{i}m\theta}$. The boundary condition on Σ in (6.27) gives

$$\sigma_1 c_3 A'_{m,\sigma_1}(R_1) + \beta c_3 A_{m,\sigma_1}(R_1) = c_\lambda.$$

Hence

$$c_3 = \frac{c_\lambda}{\sigma_1 A'_{m,\sigma_1}(R_1) + \beta A_{m,\sigma_1}(R_1)},$$

which provides the result of the lemma.
(iii) We recall that $\hat{\mathcal{A}} : L^2(\Sigma)^2 \times L^\infty(\Omega_1) \times H^{-\frac{1}{2}}(\Gamma) \to L^2(\Sigma)^2$ is defined by

$$\hat{\mathcal{A}}(\boldsymbol{\lambda}, \sigma, \psi) \coloneqq (\boldsymbol{\lambda_1} - 2\beta\gamma_{\Omega_1,\Sigma}\mathcal{A}_1(\boldsymbol{\lambda_1}, \sigma), \boldsymbol{\lambda_0} + 2\beta\gamma_{\Omega_0,\Sigma}\hat{\mathcal{A}}_0(\boldsymbol{\lambda_0}, \psi)), \qquad \text{(6.44 recalled)}$$

where $\boldsymbol{\lambda} = (\boldsymbol{\lambda_0}, \boldsymbol{\lambda_1}) \in L^2(\Sigma)^2$, $\boldsymbol{\lambda} = (\boldsymbol{\lambda_0}, \boldsymbol{\lambda_1}) \in L^2(\Sigma)^2$. Assume that $\boldsymbol{\lambda_i} = c_{\boldsymbol{\lambda_i}}\mathrm{e}^{\mathrm{i}m\theta} \in V_m(R_1)$, $i = 0, 1$ and recall that $\psi = c_\psi \mathrm{e}^{\mathrm{i}m\theta} \in V_m(R_0)$. By the expressions of $\mathcal{A}_1$ and $\hat{\mathcal{A}}_0$ respectively in (6.83) and (6.79), we have

$$\boldsymbol{\lambda_1} - 2\beta\gamma_{\Omega_1,\Sigma}\mathcal{A}_1(\boldsymbol{\lambda_1}, \sigma_1) = (1 - 2\beta C_3(R_1, \sigma_1; \beta))\boldsymbol{\lambda_1}$$

and

$$\boldsymbol{\lambda_0} + 2\beta\gamma_{\Omega_0,\Sigma}\hat{\mathcal{A}}_0(\boldsymbol{\lambda_0}, \psi) = (1 + 2\beta C_1(R_1; \beta))\boldsymbol{\lambda_0} + 2\beta C_2(R_1; \beta)c_\psi \mathrm{e}^{\mathrm{i}m\theta},$$

which provides the result of the lemma.
The last conclusion of the lemma is immediate thanks to the expressions of $\mathcal{A}$, $\hat{\mathcal{A}}$ and $\mathcal{A}_0$ respectively in (6.87), (6.85) and (6.82). □

In the case $\eta = 0$, we have the following result about the spectral radius of $\mathcal{B}(\sigma_1; \beta)$.

Lemma 6.A.5. *In the case $\eta = 0$, the matrix $\mathcal{B}(\sigma_1; \beta)$ defined in (6.86) has spectral radius less than 1 if $\beta > 0$ is sufficiently small.*

Proof. In the case $\eta = 0$, we recall that $A_{m,\sigma}(r) = r^m$ and $B_{m,\sigma}(r) = r^{-m}$. Let

$$c_3(\sigma_1) \coloneqq \lim_{\beta \to 0^+} C_3(R_1, \sigma_1; \beta), d \coloneqq \lim_{\beta \to 0^+} D(\beta) \text{ and } c_1 \coloneqq \lim_{\beta \to 0^+} C_1(R_1; \beta).$$

By (6.84), (6.81) and (6.80), we respectively have

$$c_3(\sigma_1) = \frac{A_{m,\sigma_1}(R_1)}{\sigma_1 A'_{m,\sigma_1}(R_0)} = \frac{R_1^m}{\sigma_1 m R_0^{m-1}} > 0,$$

$$d = \frac{A'_{m,\sigma_0}(R_1)}{B'_{m,\sigma_0}(R_1)} - \frac{A'_{m,\sigma_0}(R_0)}{B'_{m,\sigma_0}(R_0)} = \frac{mR_1^{m-1}}{-mR_1^{-m-1}} - \frac{mR_0^{m-1}}{-mR_0^{-m-1}} = R_0^{2m} - R_1^{2m} > 0$$

and

$$\begin{aligned} c_1 &= \frac{A_{m,\sigma_0}(R_1)}{\sigma_0 B'_{m,\sigma_0}(R_1)d} - \frac{A'_{m,\sigma_0}(R_0)B_{m,\sigma_0}(R_1)}{B'_{m,\sigma_0}(R_1)\sigma_0 B'_{m,\sigma_0}(R_0)d} \\ &= \frac{R_1^m}{-\sigma_0 m R_1^{-m-1}d} - \frac{mR_0^{m-1}R_1^{-m}}{(-m)R_0^{-m-1}\sigma_0(-m)R_1^{-m-1}d} \\ &= -\frac{R_1^{2m+1} + R_0^{2m}R_1}{\sigma_0 m d} < 0. \end{aligned}$$

Therefore, the matrix $\mathcal{B}(\sigma_1;\beta)$ defined in (6.86) can be written as

$$\mathcal{B}(\sigma_1;\beta) = \begin{bmatrix} 0 & 1 - 2\beta c_3(\sigma_1) + O(\beta^2) \\ 1 + 2\beta c_1 + O(\beta^2) & 0 \end{bmatrix}$$

where $c_3(\sigma_1) > 0$ and $c_1 < 0$, which yields the conclusion of the lemma. □

We conclude that in the case $\eta = 0$, there exists a sufficiently small $\beta > 0$ such that the two assumptions in Theorem 6.1.2 are both satisfied. Indeed, assumptions *(i)* and *(ii)* are respectively verified by Lemmas 6.A.4 and 6.A.5. Therefore, algorithm (6.46) is convergent.

Chapter 7

Conclusion and outlook

For linear inverse problems, we have proved sufficient conditions on the descent step for the convergence of semi-implicit multi-step one-shot methods, with a regularization parameter $\alpha \geq 0$. This complements the results obtained in our research report [3] for explicit schemes with no regularization. Although these bounds on the descent step are not optimal, to our knowledge no other bounds, explicit in the number of inner iterations, are available in the literature for multi-step one-shot methods. Furthermore, we have shown in the numerical experiments that very few inner iterations on the forward and adjoint problems may be enough to guarantee results similar to the classical gradient descent algorithm. Next, we have considered the case where the iterative solvers are based on domain decomposition methods instead of generic fixed point iterations. This has been done for both linearized and non-linear inverse conductivity problems, and, more specifically, we have applied the nonoverlapping OSM in a configuration of two disjoint subdomains. After rephrasing the OSM as an iterative method in terms of the impedance boundary variables, we have studied an algorithm that simultaneously calculates the forward and adjoint states, and proved that this algorithm converges under some reasonable assumptions. Moreover, we have derived an algorithm that combines multi-step one-shot and OSM, then studied all algorithms not only in the continuous setting but also in the discretized setting. In particular, in the case of linearized inverse conductivity problems, we have proposed at the discretized level a new scheme that fits our abstract framework in Chapter 1. Finally, our numerical tests have confirmed the potential of the combined algorithms in solving inverse problems with a few number of inner iterations.

A numerical validation for large-scale problems arising from realistic applications needs to be carried on. In particular, the inner fixed point iterations could be replaced by more efficient Krylov subspace methods, such as conjugate gradient or GMRES, and one could use L-BFGS instead of gradient descent as optimization algorithm. Another interesting issue is how to adapt the number of inner iterations in the course of the outer iterations. Moreover, based on this linear inverse problem study, we plan to tackle the convergence analysis of multi-step one-shot methods in the challenging case of non-linear inverse problems, as studied in the seminal paper [24]. Finally, although we have given a preliminary result about the convergence of the algorithm that simultaneously calculates the forward and adjoint states, the convergence analysis of the derived domain decomposition one-shot methods (which do not fall into our abstract framework in Chapter 1) still remains to be tackled.

Bibliography

[1] L. Audibert, H. Girardon, H. Haddar, and P. Jolivet. Inversion of eddy-current signals using a level-set method and block Krylov solvers. *SIAM Journal on Scientific Computing*, 45(3):B366–B389, 2023.

[2] S. Barnett. *Polynomials and linear control systems*, volume 77 of *Pure Appl. Math.* Marcel Dekker, Inc., New York, NY, 1983.

[3] M. Bonazzoli, H. Haddar, and T.A. Vu. Convergence analysis of multi-step one-shot methods for linear inverse problems. Research Report RR-9477, Inria Saclay; ENSTA ParisTech, July 2022.

[4] M. Bonazzoli, H. Haddar, and T.A. Vu. On the convergence analysis of one-shot inversion methods. Submitted in 2023.

[5] M. Burger and W. Mühlhuber. Iterative regularization of parameter identification problems by sequential quadratic programming methods. *Inverse Problems*, 18:943–969, 2002.

[6] S Chaabane, B Charfi, and H Haddar. Reconstruction of discontinuous parameters in a second order impedance boundary operator. *Inverse Problems*, 32(10):105004, 2016.

[7] Slim Chaabane, Houssem Haddar, and Rahma Jerbi. A combination of Kohn-Vogelius and DDM methods for a geometrical inverse problem. *Inverse Problems*, 39(9):095001, 2023.

[8] Xavier Claeys and Emile Parolin. Robust treatment of cross-points in optimized Schwarz methods. *Numerische Mathematik*, 151(2):405–442, 2022.

[9] David Colton and Rainer Kress. *Inverse Acoustic and Electromagnetic Scattering Theory*. Number volume 93 in Applied Mathematical Sciences. Springer Nature, fourth edition edition, 2019.

[10] B. Després, P. Joly, and J. E. Roberts. A domain decomposition method for the harmonic Maxwell equations. In *Iterative Methods in Linear Algebra (Brussels, 1991)*, pages 475–484. North-Holland, 1992.

[11] Bruno Després. Domain decomposition method and the Helmholtz problem. In *Second International Conference on Mathematical and Numerical Aspects of Wave Propagation (Newark, DE, 1993)*, pages 197–206. SIAM, 1993.

[12] V. Dolean, P. Jolivet, and F. Nataf. *An Introduction to Domain Decomposition Methods: Algorithms, Theory, and Parallel Implementation*. Society for Industrial and Applied Mathematics, Philadelphia, PA, 2015.

[13] Martin J. Gander and Hui Zhang. Schwarz methods by domain truncation. *Acta Numerica*, 31:1–134, 2022.

[14] N. Gauger, A. Griewank, A. Hamdi, C. Kratzenstein, E.Özkaya, and T. Slawig. Automated extension of fixed point PDE solvers for optimal design with bounded retardation. In *Constrained Optimization and Optimal Control for Partial Differential Equations, International Series of Numerical Mathematics*, pages 99–122. Springer Basel, 2012.

[15] Shihua Gong, Martin J. Gander, Ivan G. Graham, David Lafontaine, and Euan A. Spence. Convergence of parallel overlapping domain decomposition methods for the Helmholtz equation. *Numerische Mathematik*, 152(2):259–306, 2022.

[16] A. Greenbaum. *Iterative Methods for Solving Linear Systems*. Number 17 in Frontiers in Applied Mathematics. Soc. for Industrial and Applied Math, Philadelphia, 1997.

[17] A. Griewank. Projected Hessians for Preconditioning in One-Step One-Shot Design Optimization. In *Large-Scale Nonlinear Optimization*, volume 83, pages 151–171. Springer US, Boston, MA, 2006. Series Title: Nonconvex Optimization and Its Applications.

[18] S. Günther, N. R. Gauger, and Q. Wang. Simultaneous single-step one-shot optimization with unsteady PDEs. *Journal of Computational and Applied Mathematics*, 294:12–22, 2016.

[19] E. Haber and U. M. Ascher. Preconditioned all-at-once methods for large, sparse parameter estimation problems. *Inverse Problems*, 17(6):1847–1864, 2001.

[20] Houssem Haddar and Rainer Kress. Conformal mappings and inverse boundary value problems. *Inverse Problems*, 21(3):935–953, 2005.

[21] Houssem Haddar and Rainer Kress. Conformal mapping and impedance tomography. *Inverse Problems*, 26(7):074002, 2010.

[22] A. Hamdi and A. Griewank. Reduced quasi-Newton method for simultaneous design and optimization. *Computational Optimization and Applications*, 49(3):521–548, 2009.

[23] A. Hamdi and A. Griewank. Properties of an augmented Lagrangian for design optimization. *Optimization Methods and Software*, 25(4):645–664, 2010.

[24] Martin Hanke, Andreas Neubauer, and Otmar Scherzer. A convergence analysis of the Landweber iteration for nonlinear ill-posed problems. *Numerische Mathematik*, 72(1):21–37, 1995.

[25] S.B. Hazra, V. Schulz, J. Brezillon, and N.R. Gauger. Aerodynamic shape optimization using simultaneous pseudo-timestepping. *Journal of Computational Physics*, 204(1):46–64, 2005.

[26] F. Hecht. New development in FreeFem++. *J. Numer. Math.*, 20(3-4):251–265, 2012.

[27] Victor Isakov. *Inverse Problems for Partial Differential Equations*. Number volume 127 in Applied Mathematical Sciences. Springer, third edition, 2017.

[28] Rahma Jerbi. *Doctoral thesis: Combined inversion methods for inverse conductivity problems*. University of Sfax, Tunisia, 2023.

[29] E.I. Jury. On the roots of a real polynomial inside the unit circle and a stability criterion for linear discrete systems. *IFAC Proceedings Volumes*, 1(2):142–153, 1963. 2nd International IFAC Congress on Automatic and Remote Control: Theory, Basle, Switzerland, 1963.

[30] E.I. Jury. *Theory and Applications of the Z-Transform Method*. New York, 1964.

[31] B. Kaltenbacher. Regularization based on all-at-once formulations for inverse problems. *SIAM Journal on Numerical Analysis*, 54(4):2594–2618, 2016.

[32] B Kaltenbacher. All-at-once versus reduced iterative methods for time dependent inverse problems. *Inverse Problems*, 33(6):064002, 2017.

[33] B. Kaltenbacher, A. Kirchner, and B. Vexler. Goal oriented adaptivity in the IRGNM for parameter identification in PDEs II: all-at-once formulations. *Inverse Problems*, 30:045002, 2014.

[34] P.-L. Lions. On the Schwarz alternating method. II. In *Domain Decomposition Methods, Tony Chan, Roland Glowinski, Jacques Périaux, and Olof Widlund, Editors*, pages 47–70. SIAM, 1989.

[35] P.-L. Lions. On the Schwarz alternating method. In *First International Symposium*

on Domain Decomposition Methods for Partial Differential Equations, Tony F. Chan, Roland Glowinski, Jacques Périaux, and Olof Widlund, Editors. SIAM, 1990.

[36] P.-L. Lions. On the Schwarz alternating method. III. In *Third International Symposium on Domain Decomposition Methods for Partial Differential Equations, Tony F. Chan, Roland Glowinski, Jacques Périaux, and Olof Widlund, Editors*. SIAM, 1990.

[37] M. Marden. *The geometry of the zeros of a polynomial in a complex variable*, volume 3 of *Math. Surv.* American Mathematical Society (AMS), Providence, RI, 1949.

[38] M. Marden. *Geometry of Polynomials*. Number 3 in Mathematical Surveys and Monographs. American Math. Soc, Providence, RI, 2nd edition, 1966.

[39] Tram Thi Ngoc Nguyen. Landweber–Kaczmarz for parameter identification in time-dependent inverse problems: All-at-once versus reduced version. *Inverse Problems*, 35(3):035009, 2019.

[40] Tram Thi Ngoc Nguyen. Bi-level iterative regularization for inverse problems in nonlinear PDEs. *Inverse Problems*, 40(4):045020, 2024.

[41] E. Özkaya and N. R. Gauger. Single-step One-shot Aerodynamic Shape Optimization. In *Optimal Control of Coupled Systems of Partial Differential Equations*, volume 158, pages 191–204. Birkhäuser Basel, Basel, 2009. Series Title: International Series of Numerical Mathematics.

[42] V. Schulz and I. Gherman. One-Shot Methods for Aerodynamic Shape Optimization. In *MEGADESIGN and MegaOpt - German Initiatives for Aerodynamic Simulation and Optimization in Aircraft Design*, volume 107, pages 207–220. Springer Berlin Heidelberg, Berlin, Heidelberg, 2009. Series Title: Notes on Numerical Fluid Mechanics and Multidisciplinary Design.

[43] I. Schur. Über Potenzreihen, die im Innern des Einheitskreises beschränkt sind. *Journal für die reine und angewandte Mathematik (Crelles Journal)*, 1917(147):205–232, 1917.

[44] H.A. Schwarz. Über einen Grenzübergang durch alternierendes Verfahren. *Vierteljahrss- chrift der Naturforschenden Gesellschaft in Zürich*, pages 15:272–286, 1870.

[45] A. Shenoy, M. Heinkenschloss, and E. M. Cliff. Airfoil design by an all-at-once method. *International Journal of Computational Fluid Dynamics*, 11(1-2):3–25, 1998.

[46] S. Ta'asan. "One Shot" Methods for Optimal Control of Distributed Parameter Systems I: Finite Dimensional Control. Technical Report 91-2, ICASE, Hampton, 1991.

[47] S. Ta'asan, G. Kuruvila, and M. Salas. Aerodynamic design and optimization in one shot. In *30th Aerospace Sciences Meeting and Exhibit*, Reno, NV, U.S.A., 1992. American Institute of Aeronautics and Astronautics.

[48] A. Tarantola and B. Valette. Generalized nonlinear inverse problems solved using the least squares criterion. *Reviews of Geophysics*, 20(2):219–232, 1982.

[49] P.-H. Tournier, I. Aliferis, M. Bonazzoli, M. de Buhan, M. Darbas, V. Dolean, F. Hecht, P. Jolivet, I. El Kanfoud, C. Migliaccio, F. Nataf, C. Pichot, and S. Semenov. Microwave tomographic imaging of cerebrovascular accidents by using high-performance computing. *Parallel Computing*, 85:88–97, 2019.

[50] T. van Leeuwen and F. J. Herrmann. Mitigating local minima in full-waveform inversion by expanding the search space. *Geophysical Journal International*, 195(1):661–667, 2013.

[51] T. van Leeuwen and F. J. Herrmann. A penalty method for PDE-constrained optimization in inverse problems. *Inverse Problems*, 32(1):015007, 2015.

ECOLE
DOCTORALE
DE MATHEMATIQUES
HADAMARD

Title: One-shot inversion methods and domain decomposition

Key words: inverse problems, one-shot methods, domain decomposition methods, convergence analysis, parameter identification

Abstract:
Our main goal is to analyze the convergence of a gradient-based optimization method, to solve inverse problems for parameter identification, in which the corresponding forward and adjoint problems are solved by an iterative solver. Coupling the iterations for the three unknowns (the inverse problem parameter, the forward problem solution and the adjoint problem solution) yields the so-called one-shot inversion methods. Many numerical experiments showed that using very few inner iterations for the forward and adjoint problems may still lead to a good convergence for the inverse problem. This motivates us to develop a rigorous convergence theory for one-shot methods using a fixed small number of inner iterations, with a semi-implicit scheme for the parameter update and a regularized cost functional. Our theory covers a general class of linear inverse problems in the finite-dimensional discrete setting, for which the forward and adjoint problems are solved by generic fixed point iteration methods. By studying the spectral radius of the block iteration matrix of the coupled iterations, we prove that for sufficiently small descent steps the (semi-implicit) one-shot methods converge. In particular, in the scalar case, where the unknowns belong to one-dimensional spaces, we establish not only sufficient but even necessary convergence conditions on the descent step. Next, we apply one-shot methods to (linearized and then non-linear) inverse conductivity problems, and solve the forward and adjoint problems by domain decomposition methods, more specifically nonoverlapping optimized Schwarz methods. We analyze a domain decomposition algorithm that simultaneously calculates the forward and adjoint solutions for a given conductivity. By combining this algorithm with the gradient descent parameter update, we obtain a domain decomposition one-shot method that solves the inverse problem. We propose two discretized versions of the coupled algorithm, the second of which (in the case of the linearized inverse conductivity problem) falls into the abstract framework of our convergence theory. Finally, several numerical experiments are provided to illustrate the performance of the one-shot methods, in comparison with the classical gradient descent in which the forward and adjoint problems are solved using direct solvers. In particular, we observe that, even in the case of noisy data, very few inner iterations may still guarantee good convergence of the one-shot methods.

Titre : Méthodes d'inversion de type one-shot et décomposition de domaine

Mots clés : problèmes inverses, méthodes de type one-shot, méthodes de décomposition de domaine, analyse de convergence, identification de paramètres

Résumé :
Notre objectif principal est d'analyser la convergence d'une méthode d'optimisation basée sur le gradient, pour résoudre des problèmes inverses d'identification de paramètres, dans laquelle les problèmes directs et adjoints correspondants sont résolus par un solveur itératif. Le couplage des itérations pour les trois inconnues (le paramètre du problème inverse, la solution du problème direct et la solution du problème adjoint) donne ce que l'on appelle les méthodes d'inversion de type one-shot. De nombreux tests numériques ont montré que l'utilisation de très peu d'itérations internes pour les problèmes directs et adjoints peut néanmoins conduire à une bonne convergence pour le problème inverse. Cela nous motive à développer une théorie de convergence rigoureuse pour les méthodes de type one-shot en utilisant un petit nombre fixe d'itérations internes, avec un schéma semi-implicite pour la mise à jour du paramètre et une fonction de coût régularisée. Notre théorie couvre une classe générale de problèmes inverses linéaires dans le cadre discret de dimension finie, pour lesquels les problèmes directs et adjoints sont résolus par des méthodes génériques d'itération de point fixe. En étudiant le rayon spectral de la matrice par blocs des itérations couplées, nous prouvons que pour des pas de descente suffisamment petits, les méthodes de type one-shot (semi-implicites) convergent. En particulier, dans le cas scalaire, où les inconnues appartiennent à des espaces à une dimension, nous établissons des conditions de convergence suffisantes et même nécessaires sur le pas de descente. Ensuite, nous appliquons des méthodes de type one-shot aux problèmes inverses de conductivité (linéarisés et puis non linéaires), et résolvons les problèmes directs et adjoints par des méthodes de décomposition de domaines, plus spécifiquement des méthodes de Schwarz optimisées sans recouvrement. Nous analysons un algorithme de décomposition de domaine qui calcule simultanément les solutions directe et adjointe pour une conductivité donnée. En combinant cet algorithme avec la mise à jour du paramètre par descente de gradient, nous obtenons une méthode one-shot de décomposition de domaine qui résout le problème inverse. Nous proposons deux versions discrétisées de l'algorithme couplé, dont la seconde (dans le cas du problème inverse de conductivité linéarisé) s'inscrit dans le cadre abstrait de notre théorie de convergence. Enfin, plusieurs expériences numériques sont fournies pour illustrer les performances des méthodes de type one-shot, en comparaison avec la méthode de descente de gradient classique dans laquelle les problèmes directs et adjoints sont résolus par des solveurs directs. En particulier, nous observons que, même dans le cas de données bruitées, très peu d'itérations internes peuvent toujours garantir une bonne convergence des méthodes de type one-shot.

Institut Polytechnique de Paris
91120 Palaiseau, France