

**Model Reduction of Wave Equations**  
**Theory and applications in Forward modeling and Imaging**

Zimmerling, Jörn

**DOI**

[10.4233/uuid:9fa0bdd9-29b4-489c-9799-b86e07e92813](https://doi.org/10.4233/uuid:9fa0bdd9-29b4-489c-9799-b86e07e92813)

**Publication date**

2018

**Document Version**

Final published version

**Citation (APA)**

Zimmerling, J. (2018). *Model Reduction of Wave Equations: Theory and applications in Forward modeling and Imaging*. [Dissertation (TU Delft), Delft University of Technology]. <https://doi.org/10.4233/uuid:9fa0bdd9-29b4-489c-9799-b86e07e92813>

**Important note**

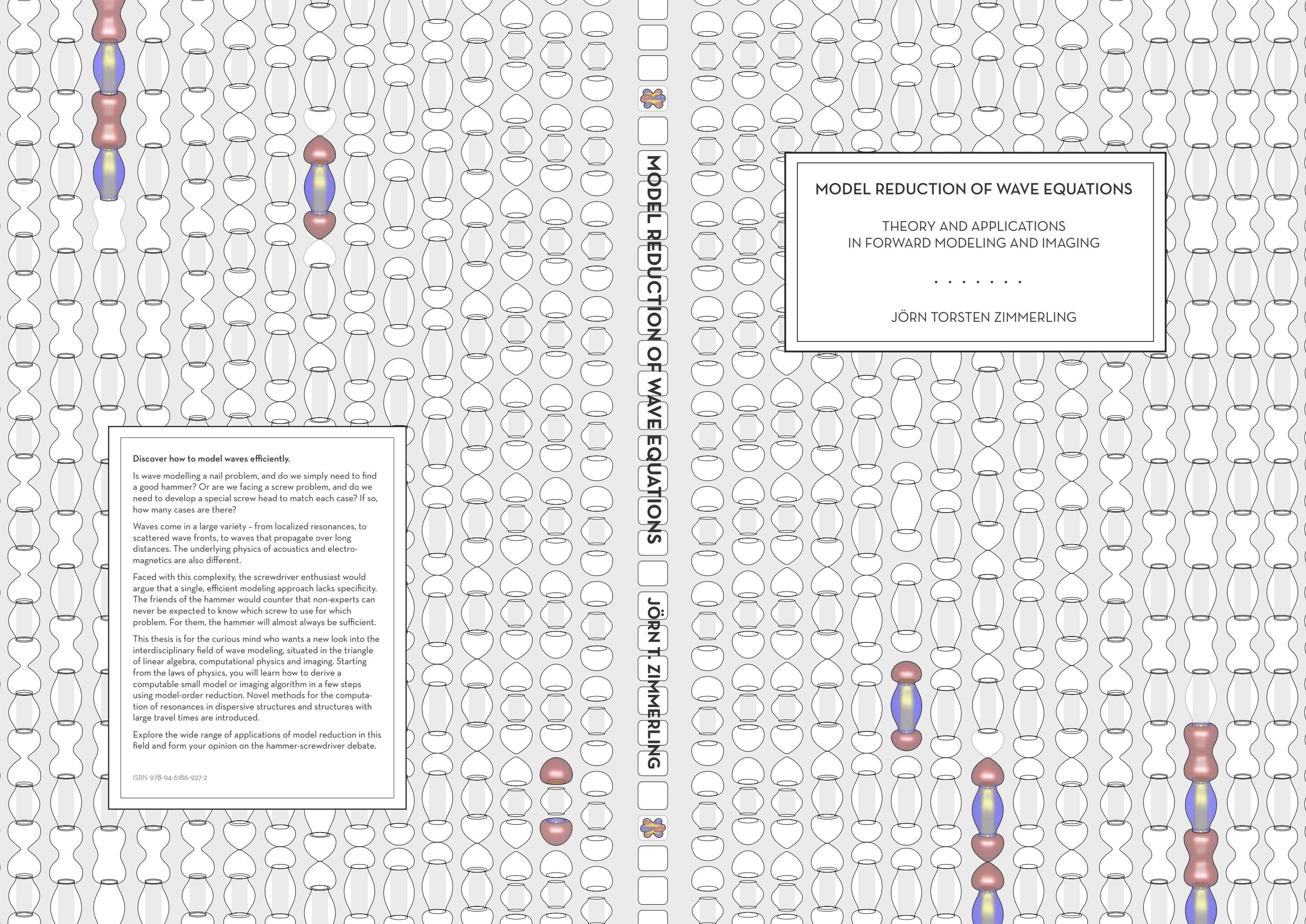
To cite this publication, please use the final published version (if applicable).  
Please check the document version above.

**Copyright**

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

**Takedown policy**

Please contact us and provide details if you believe this document breaches copyrights.  
We will remove access to the work immediately and investigate your claim.



MODEL REDUCTION OF WAVE EQUATIONS

JÖRN T. ZIMMERLING



# MODEL REDUCTION OF WAVE EQUATIONS

THEORY AND APPLICATIONS  
IN FORWARD MODELING AND IMAGING

.....

JÖRN TORSTEN ZIMMERLING

### Discover how to model waves efficiently.

Is wave modelling a nail problem, and do we simply need to find a good hammer? Or are we facing a screw problem, and do we need to develop a special screw head to match each case? If so, how many cases are there?

Waves come in a large variety - from localized resonances, to scattered wave fronts, to waves that propagate over long distances. The underlying physics of acoustics and electro-magnetics are also different.

Faced with this complexity, the screwdriver enthusiast would argue that a single, efficient modeling approach lacks specificity. The friends of the hammer would counter that non-experts can never be expected to know which screw to use for which problem. For them, the hammer will almost always be sufficient.

This thesis is for the curious mind who wants a new look into the interdisciplinary field of wave modeling, situated in the triangle of linear algebra, computational physics and imaging. Starting from the laws of physics, you will learn how to derive a computable small model or imaging algorithm in a few steps using model-order reduction. Novel methods for the computation of resonances in dispersive structures and structures with large travel times are introduced.

Explore the wide range of applications of model reduction in this field and form your opinion on the hammer-screwdriver debate.

ISBN 978-94-6186-927-2

# **MODEL REDUCTION OF WAVE EQUATIONS**

THEORY AND APPLICATIONS

IN

FORWARD MODELING AND IMAGING



# **MODEL REDUCTION OF WAVE EQUATIONS**

THEORY AND APPLICATIONS  
IN  
FORWARD MODELING AND IMAGING

## **PROEFSCHRIFT**

ter verkrijging van de graad van doctor  
aan de Technische Universiteit Delft,  
op gezag van de Rector Magnificus prof. dr. ir. T.H.J.J. van der Hagen  
voorzitter van het College voor Promoties,  
in het openbaar te verdedigen op  
maandag 2 juli 2018 om 12:30 uur

door

**Jörn Torsten ZIMMERLING**

Elektrotechnisch Ingenieur, Technische Universiteit Delft, Nederland  
geboren te Pinneberg, Duitsland

Dit proefschrift is goedgekeurd door de promotoren.

Samenstelling promotiecommissie bestaat uit:

Rector magnificus,	voorzitter
Prof. dr. H.P. Urbach	Technische Universiteit Delft, <i>promotor</i>
Dr. ir. R.F. Remis	Technische Universiteit Delft, <i>promotor</i>

*Onafhankelijke leden:*

Prof. dr. ir. C. Vuik	Technische Universiteit Delft
Prof. dr. W.M.J.M Coene	Technische Universiteit Delft
Prof. dr. L. Borcea	University of Michigan, U.S.A.
Dr. V.L. Druskin	Schlumberger Doll-Research, U.S.A.
Dr. M. Zaslavsky	Schlumberger Doll-Research, U.S.A.
Prof. dr. ir. A.-J. van der Veen	Technische Universiteit Delft, <i>reservelid</i>



This research was supported by the Dutch Technology Foundation STW (project number 14222 “GOOD VIBRATIONS” – FAST AND ROBUST WAVE FIELD COMPUTATIONS IN COMPLEX STRUCTURES USING KRYLOV RESONANCE EXPANSIONS), which is part of the Netherlands Organisation for Scientific Research (NWO), and which is partly funded by the Ministry of Economic Affairs.

*Keywords:* Maxwell Equations, Reduced-Order Modeling, Dispersive Media, Open Resonators, Krylov

*Printed by:* Ipskamp printing

*Front & Back:* Mathomorfose (SHB and JTZ)

Copyright © 2018 by J.T. Zimmerling

ISBN 978-94-6186-927-2

An electronic version of this dissertation is available at <http://repository.tudelft.nl/>.

*Alles was  
Ich sonst besitze, hat Natur und Glück  
Mir zugeteilt. Dies Eigentum allein  
Dank' ich der Tugend*

Gotthold Ephraim Lessing in *Nathan der Weise*

*Der Mensch ist Etwas, das überwunden werden muss:  
und darum sollst du deine Tugenden lieben denn du wirst an ihnen zugrunde gehn.*

Friedrich Nietzsche in *Also sprach Zarathustra*



# CONTENTS

<b>Summary</b>	<b>ix</b>
<b>Samenvatting</b>	<b>xi</b>
<b>Zusammenfassung</b>	<b>xiii</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Reduced-Order Modeling of Wave Propagation in Unbounded Domains . . .	3
1.2 Reduced-Order Modeling in Applications . . . . .	4
1.3 Reduced-Order Modeling for Imaging . . . . .	6
1.4 Thesis Contributions and Outline . . . . .	7
<b>2 From Physics to Discrete Systems</b>	<b>11</b>
2.1 Introduction . . . . .	12
2.2 Scalar Wave Equation on an Unbounded Domain . . . . .	12
2.3 Maxwell Equations on an Unbounded Domain . . . . .	13
2.4 Some General Wavefield Properties . . . . .	18
2.5 Spatial Discretization . . . . .	20
2.6 Introduction of a PML. . . . .	26
2.7 Solution of the System . . . . .	28
2.8 Conclusions. . . . .	32
<b>3 Model Order Reduction</b>	<b>33</b>
3.1 Introduction . . . . .	34
3.2 Projection-Based Model Order Reduction. . . . .	35
3.3 Computational Complexity of RKS, EKS and PKS . . . . .	41
3.4 Diffusion Approximation Versus Full Maxwell System. . . . .	42
3.5 Applications . . . . .	43
3.6 Conclusions. . . . .	61
<b>4 Phase-Preconditioning for Compression of Wavefields</b>	<b>65</b>
4.1 Introduction . . . . .	66
4.2 Motivation . . . . .	66
4.3 Structure-Preserving RKS . . . . .	69
4.4 Field Parametrization for SISO Problems . . . . .	70
4.5 Phase-Preconditioning for MIMO Systems . . . . .	76
4.6 Discrete Formulation . . . . .	79
4.7 Results . . . . .	81
4.8 Discussion on Parallel Implementation . . . . .	96
4.9 Conclusions. . . . .	100



<b>5</b>	<b>Inversion in the Reduced-Order Model Framework</b>	<b>103</b>
5.1	Introduction . . . . .	104
5.2	3D Anisotropic Resistivity . . . . .	104
5.3	Conclusions. . . . .	115
<b>6</b>	<b>Conclusions</b>	<b>117</b>
<b>A</b>	<b>Appendix – Phase-Preconditioning</b>	<b>121</b>
A.1	The Field Amplitudes in Two Dimensions. . . . .	121
<b>B</b>	<b>Appendix – Nuclear Sigma Imaging</b>	<b>125</b>
B.1	Fast Forward Model for Neutron Scattering . . . . .	126
B.2	Inversion Approach . . . . .	127
B.3	Synthetic Nine Layer Example (Inverse Crime) . . . . .	128
B.4	Inversion of Monte Carlo Data . . . . .	129
B.5	Conclusions. . . . .	133
	<b>References</b>	<b>135</b>
	<b>Glossary</b>	<b>143</b>
	<b>Notation</b>	<b>145</b>
	<b>Curriculum Vitae</b>	<b>147</b>

# SUMMARY

**H**OW do you look inside a box without opening it? How can we know whether or not a heart valve is functioning correctly without cutting a person open?

Imaging – the art of seeing the unseeable. A CT-scan at the doctor's office, crack detection in the wing of an airplane, and medical ultrasound are all examples of imaging techniques that allow us to inspect the interior of an object or person and enable us to observe features that are not directly visible to the naked eye. Science continuously improves upon existing imaging methods and occasionally invents new ones leading to improved image quality and faster image acquisition.

Many imaging applications rely on acoustic, electromagnetic, or elastodynamic waves for imaging. These methods use waves to illuminate a penetrable object, and then form an image of its interior from measurements of transmitted or scattered waves. In such imaging problems efficient computation of wavefields in complex geometries is key. New mathematical methods and algorithms are needed to keep up with the demands of the imaging industry – advancements in the computer industry alone cannot respond to the shift towards larger domains, higher resolution, and larger data sets.

This thesis is about reduced-order modeling of the equations that describe the dynamics of wave propagation. In reduced-order modeling, the aim is to systematically develop a small model that describes a complex system without losing information that is valuable for a specific application. Evaluating such a model is computationally much more efficient than direct evaluation of the unreduced system and in the context of imaging it can lighten the computational burden associated with imaging algorithms. The central question is, of course: How does one construct a model that describes the wave dynamics relevant to a particular application?

Wave equations are partial differential equations that interrelate the spatial and temporal variations of a particular physical wavefield quantity. When we discretize such equations in space, sparse systems of equations with hundreds of thousands or even millions of unknowns are obtained. Via projection onto a small subspace such a large-scale system can be reduced to a much smaller reduced system. The solution of this small system is called a reduced-order model. A properly constructed reduced-order model can be easily evaluated and gives an accurate wavefield description over a certain time or frequency interval or parameter range of interest.

In this thesis, we discuss different choices for the subspaces that are used for projection in model-order reduction. In particular, we show which types of subspaces are effective for wavefields that are localized and highly resonant and how to efficiently generate such subspaces by exploiting certain symmetry properties of the wave equations. We illustrate the effectiveness of the resulting reduced-order models by computing optical wavefield responses in three-dimensional metallic nano-resonators.



Not all wavefields are determined by a few resonances, of course. Waves can also travel over long distances without losing information; a property that is used by mobile phones every day. The reduction methods developed for resonating fields are not efficient for these types of propagation problems and require a different approach. In this thesis, we present a so-called phase-preconditioning reduction method, in which a specific subspace is generated that explicitly takes the large travel times of the waves into account. We demonstrate the effectiveness of this reduction approach using examples from geophysics, where waves with long travel times are frequently encountered or used to probe the subsurface of the Earth.

Finally, we show how reduced-order modeling techniques can be incorporated into advanced nonlinear imaging algorithms. Here, we focus on an imaging application in geophysics, where the goal is to retrieve the conductivity tensor of a bounded anomaly located in the subsurface of the Earth, based on measured electromagnetic field data collected on a borehole axis. We demonstrate that the use of reduced-order models in a nonlinear optimization framework does indeed lead to significant computational savings without sacrificing the quality of imaging results. To illustrate the wide applicability of model-order reduction techniques in imaging, an additional example from nuclear geophysical imaging is also presented.

## SAMENVATTING

**H**OE kijk je in een doos zonder de doos te openen? Hoe weten we of een hartklep nog goed functioneert, zonder een persoon open te snijden?

Imaging – de kunst van zien wat niet gezien kan worden. Een CT scan bij de dokter, het detecteren van scheurtjes in de vleugel van een vliegtuig of medische echografie, het zijn allemaal voorbeelden van afbeeldingstechnieken, welke ons in staat stellen om de binnenkant van een object of het inwendige van een persoon te inspecteren en kenmerken te kunnen waarnemen die met het blote oog niet waarneembaar zijn. Binnen de wetenschap worden er voortdurend nieuwe en verbeterde afbeeldingstechnieken ontwikkeld, waarmee op een steeds snellere wijze afbeeldingen met een hogere beeldkwaliteit kunnen worden verkregen.

Om een afbeelding te realiseren, maken een groot aantal technieken gebruik van elektromagnetische, akoestische of seismische golven. Door een doordringbaar object met deze golven te belichten, kan aan de hand van metingen van de doorgelaten en verstrooide golven een afbeelding van de inwendige structuur van het object worden gemaakt. Het efficiënt kunnen berekenen van golfvelden in complexe geometrieën is van essentieel belang bij dit soort golfveld-afbeeldingsproblemen en ontwikkelingen binnen imagingindustrie vragen om nieuwe wiskundige methoden en algoritmes die met grote datasets overweg kunnen om grote gebieden bij hoge resolutie af te kunnen beelden.

Dit proefschrift handelt over het opstellen van modellen van gereduceerde orde voor golfvergelijkingen. Het doel van dit reductieproces is om op een systematische wijze gereduceerde modellen op te stellen, waarmee het gedrag van een complex systeem kan worden beschreven zonder dat essentiële informatie verloren gaat. Het evalueren van zo'n gereduceerd model kost in vergelijking met een directe evaluatie van het niet-gereduceerde systeem aanzienlijk minder rekenkracht en -tijd en leidt bij toepassing tot efficiënte afbeeldingstechnieken. De centrale vraag is natuurlijk: Hoe construeren wij een model, waarmee de golfdynamica binnen een bepaalde toepassing nauwkeurig beschreven wordt?

Golfvergelijkingen zijn partiële differentiaalvergelijkingen en leggen een verband tussen de ruimtelijke en temporele variaties van een zekere golfveldgrootte. Discretizeren wij een golfvergelijking in de ruimte, dan verkrijgen wij een ijf stelsel van vergelijkingen met honderdduizenden of zelfs miljoenen onbekenden. Via projectie op een kleine deelruimte kan zo'n groot stelsel worden gereduceerd tot een veel kleiner stelsel. De oplossing van dit kleine stelsel wordt een model van gereduceerde orde genoemd. Een goed geconstrueerd model is snel te evalueren en geeft een nauwkeurige beschrijving van het golfveld over een zeker tijd- of frequentieinterval of over een parameterbereik, waarin wij geïnteresseerd zijn.

In dit proefschrift bespreken wij verschillende keuzes voor de deelruimtes die voor projectie binnen modelorderreductie worden gebruikt. We laten zien welke deel-



ruimtes tot een nauwkeurige beschrijving van gelokaliseerde en sterk resonante golven leiden en hoe deze deelruimtes op een efficiënte manier kunnen worden gegenereerd door gebruik te maken van zekere symmetrie-eigenschappen van de golfvergelijking. We illustreren de effectiviteit van de resulterende gereduceerde modellen door optische golfveldresponsies te berekenen in driedimensionale metallische nano-resonatoren.

Natuurlijk worden niet alle golfvelden door slechts een klein aantal resonanties beschreven. Golven kunnen ook over grote afstanden propageren zonder informatie te verliezen; een eigenschap, waarvan tijdens mobiel bellen dagelijks gebruik wordt gemaakt. De reductiemethoden voor resonante golfvelden zijn niet efficiënt voor dit soort propagatieproblemen en vereisen een andere aanpak. Wij introduceren daartoe een reductiemethode, waarin het golfveld wordt gepreconditioneerd en waarin de reistijden van de golven expliciet worden meegenomen in de constructie van de deelruimte die voor projectie wordt gebruikt. De effectiviteit van deze reductiemethode illustreren wij aan de hand van een aantal geofysische toepassingen, waarin golfvelden met lange reistijden vaak voorkomen.

Tot slot laten wij zien hoe modellen van gereduceerde orde kunnen worden opgenomen in geavanceerde niet-lineaire afbeeldingsalgoritmes. We concentreren ons hierbij op een afbeeldingstoepassing uit de geofysica, waarin het doel is om de geleidingstensor van een begrensde anomalie in de ondergrond te reconstrueren aan de hand van elektromagnetische velddata gemeten in een boorgat. We laten zien dat het gebruik van gereduceerde modellen in een niet-lineair optimalisatiekader inderdaad tot een aanzienlijke besparing aan rekentijd en benodigde rekenkracht oplevert zonder de kwaliteit van de afbeeldingen aan te tasten. Om de brede toepassing van modelordereductie te illustreren, bespreken wij als laatste ook nog een nucleaire afbeeldingstechniek welke wordt toegepast binnen de geofysica.

# ZUSAMMENFASSUNG

**W**IE sieht man in eine Schachtel oder Kiste, ohne sie zu öffnen? Wie bringt man in Erfahrung ob eine Herzklappe noch funktioniert, ohne einen Patienten aufzuschneiden?

Imaging (zu Deutsch: Abbilden) – Die Kunst das Unsichtbare sichtbar zu machen. Ein CT-Scan beim Arzt, das Entdecken von Rissen im Flügel eines Flugzeugs oder aus medizinischer Sicht eine Ultraschalluntersuchung, sind Vorbilder von Abbildungsverfahren, die es uns ermöglichen, das Innere eines Objektes oder einer Person zu inspizieren und Merkmale zu erkennen, die für das bloße Auge nicht sichtbar sind. Fortlaufend verbessert die Wissenschaft existierende bildgebende Verfahren und findet stets schnellere Möglichkeiten der Darstellung mit höherer Bildqualität.

Um eine Abbildung zu erstellen, verwenden viele Verfahren akustische, elektromagnetische oder elastische Wellen. Aus Messungen des gestreuten Feldes eines von Wellen bestrahlten und durchdringbarem Objektes kann eine Abbildung der inneren Struktur des Objektes erstellt werden. Das effiziente Berechnen dieser Wellen und Felder in komplexen Geometrien ist von äußerster Bedeutung für diese Abbildungsverfahren. Um mit den Entwicklungen in der Industrie für Abbildungsverfahren, wie zum Beispiel größere Abbildungsgebiete, höhere Auflösungen oder größere Datensätze, Schritt halten zu können, bedarf es neuer mathematischer Methoden und Algorithmen, da der Fortschritt in der Computerindustrie allein nicht ausreicht, um den Anforderungen der Industrie für bildgebende Verfahren gerecht zu werden.

In dieser Dissertation werden Modelle von reduzierter Ordnung entwickelt, die die dynamischen Eigenschaften von Wellen und Feldern beschreiben. Das Ziel beim Modellieren mit Modellen von reduzierter Ordnung ist es, ein komplexes System in einem kleinstmöglichen Modell zusammenzufassen, ohne dabei den Verlust anwendungsrelevanter Informationen zuzulassen. Das Auswerten eines solchen Modelles ist rechnerisch gesehen effizienter als das Berechnen eines nicht reduzierten Systems und im Zusammenhang mit Abbildungsverfahren können diese Modelle die benötigte Rechenkapazität senken. Die Schlüsselfrage für diese Methode ist: Wie erstellt man ein Modell, das die anwendungsspezifische Wellendynamik ausreichend genau beschreibt?

Wellengleichungen sind partielle Differentialgleichungen die räumliche und zeitliche Veränderungen einer bestimmten physikalischen Größe miteinander verbinden. Wenn eine solche Gleichung im Raum diskretisiert wird, so erhält man ein dünnbesetztes System von Gleichungen mit hunderttausenden oder sogar Millionen Unbekannten. Mittels Projektion auf einen kleinen Unterraum kann solch ein großes System reduziert werden zu einem kleineren System, ein Modell von reduzierter Ordnung. Ein angemessen reduziertes System kann schnell ausgewertet werden und gibt genaue Beschreibungen der Wellendynamik und Felder über ein gewisses Zeitintervall, Frequenzintervall oder Parameterintervall.



In dieser Doktorarbeit werden verschiedene Unterräume behandelt, die verwendet werden, um mittels Projektion Modelle von reduzierter Ordnung zu erhalten. Insbesondere wird gezeigt welche Arten von Unterräumen geeignet sind um lokalisierte, resonante Felder zu beschreiben, und wie solche Unterräume effizient, unter Berücksichtigung der Symmetrieeigenschaften von Wellengleichungen, berechnet werden können. Die Effizienz dieser Methode wird veranschaulicht am Beispiel der Berechnung von elektromagnetischen Resonanzen in dreidimensionalen, metallischen Nanoresonatoren.

Längst nicht alle natürlichen Wellen-Phänomene sind lokalisiert und resonant. Wellen können sich auch über große Strecken fortbewegen ohne Information zu verlieren, eine Eigenschaft die Mobiltelefone täglich nutzen. Die Verfahren, die entwickelt wurden, um resonante Felder zu reduzieren sind nicht sehr effizient für Felder, die sich über lange Strecken fortbewegen. In dieser Arbeit wird ein Verfahren, das wir "Phasen-vorkonditionierung" für Modelle reduzierter Ordnung nennen, eingeführt, das einen Unterraum generiert der explizit die Wellenlaufzeit in dem vorliegenden Material berücksichtigt. Mit Hilfe von Beispielen aus der Geophysik werden die Vorteile einer solchen Methode aufgezeigt, da große Wellenlaufzeiten in vielen Anwendungen in der Geophysik vorkommen. Abschließend wird in dieser Arbeit gezeigt, wie Modelle von reduzierter Ordnung in nichtlinearen

Abbildungsalgorithmen verwendet werden können. Dies geschieht hier am Vorbild eines Abbildungsverfahrens aus der Geophysik, bei dem ein Leitfähigkeitstensor einer begrenzten, unterirdischen Anomalie, durch elektromagnetische Messungen in einem Bohrloch, abgebildet wird. Es wird gezeigt, dass Modelle von reduzierter Ordnung, ohne die Abbildung zu beeinflussen, zu einer erheblichen Minderung der Berechnungskomplexität in nichtlinearen Optimierungsverfahren beitragen, die dieses Abbildungsproblem zu lösen versuchen. Um den breiten Anwendungsbereich von Modellen von reduzierter Ordnung zu demonstrieren wird als weiteres Beispiel ein nukleares Abbildungsverfahren aus der Geophysik behandelt.

# 1

## INTRODUCTION

*Mathematics is not a deductive science – that’s a cliché. When you try to prove a theorem, you don’t just list the hypotheses, and then start to reason. What you do is trial and error, experimentation, guesswork.*

Paul Halmos



WE experience the world primarily through our senses – we see an idyllic meadow in the woods, we hear the birds sing their songs, we feel the heat of a stone in the sun before we even touch it. These everyday sensual experiences rely on the laws of physics and the transport of energy over a distance. Over the centuries humans have mastered many of the domains of physics and have developed associated tools to prolong and enhance our quality of life. Electromagnetism, the physics that lets you see a meadow, is used to image bodies in medical applications like MRI- or CT-scans. Acoustical waves, like those responsible for transporting a bird's song from its hidden nest to your ear, are used in medical ultrasound to image a fetus.

Waves transporting energy, like the sun's rays warming a stone, also play an important role in our society; no mother would get a phone call on a Sunday if it weren't for the ability to send energy – and thus, information – over long distances. Waves are also used in many applications in imaging; we understand the interaction of waves with materials well enough to make claims about what a material looks like just by comparing the waves we send into a material to the waves that come out of the material after interacting with it.

In this thesis we are modeling wave propagation in complex materials in such a way that we can tell a computer to calculate the material interaction of the waves given a computer that is powerful enough. For many applications, computing the so-called wavefield in the whole space and to a high precision is unnecessary; the waves that are scattered during an ultrasound scan that don't reach a receiver, for example, are of little interest.

Standard techniques to make models to compute wave propagation can be very large and difficult to compute even on modern computers. For many applications this can be prohibitive: in medical imaging, you don't always have time to wait a week for an image to be computed; if an imaging device searches for land mines underground, fast computational methods are vital. Model order reduction techniques, which are extensively used in the area of system and control, can be used to lighten the computational burden associated with the computation of wave problems. Wave phenomena are described by partial differential equations (PDEs) and the materials that waves travel through manifest themselves as the coefficients in these PDEs. In this thesis we are interested in two types of wave problems:

1. **Forward problems:** The design of fast and efficient numerical solvers for PDEs with variable coefficients and in complex geometries. We develop reduced-order modeling techniques to lower the computational burden associated with wave simulations.
2. **Inverse problems:** Estimating the coefficients of a PDE from remote measurements of a wavefield. Specifically, utilizing reduced-order model approaches for fast and accurate imaging and inversion.

In this thesis we deal with the Maxwell equations, which describe electromagnetic wave propagation, and with the wave equation. We limit ourselves to these equations so as to not get lost in the notation and peculiarities of the different laws of physics; however, the

methods introduced in this work are also applicable to equations describing other types of wave phenomena (seismic waves for example).

The aim of this thesis is to assess reduced-order modeling techniques for wave equations, to develop efficient reduced-order modeling techniques for resonant as well as propagating wavefields, and to identify important application areas where these reduction techniques can be utilized.

## 1.1. REDUCED-ORDER MODELING OF WAVE PROPAGATION IN UNBOUNDED DOMAINS

Solutions of PDEs with variable coefficients in complex geometries are required in many applications, such as optimal design in optics and medical or geophysical imaging. Straightforward discretization of such equations is often computationally intractable. This, however, can be overcome by determining a computationally tractable surrogate, or reduced-order model. This model can then be evaluated in applications instead of solving the full-scale PDE.

There are two classes of reduced-order modeling techniques: (1) Techniques for finding approximate solutions of the PDE, so that construction and evaluation of the model is fast. (2) Techniques for building computationally “optimal” models. The construction of such models may be computationally expensive; however, their evaluation is fast and accurate. The first class of reduction techniques is referred to as “online” model reduction, while the second class is called “offline” model reduction. The goal of reduced-order modeling is generally a reduction in computational load [2].

The model order reduction approaches in this thesis are projection-based: a large wave operator is projected onto a small subspace. An approximate solution to the full-order system can be drawn from this projection at an insignificant computational cost compared to the full-order system.

Most applications encountered in wave propagation are on open domains, meaning that waves can leave the domain on which we model them. To model such open domains, we use a perfectly matched layer (PML) around the domain of interest that absorbs any outgoing wave without reflection. This is achieved by stretching the spatial coordinates inside this layer. Stretching depends nonlinearly on frequency and therefore the PML introduces a nonlinearity with respect to frequency into the modeling approach. The main computational cost of model reduction lies in the computation of a suitable projection space.

As subspaces we use various forms of Krylov subspaces, which have a proven track record in scientific computing [71, 72, 83, 17]. Many iterative solvers for matrix systems are based on polynomial Krylov subspaces (PKSs) [72, 82], originally introduced by Aleksey Krylov to study oscillations of ships [52]. In these methods, one searches for approximate solutions to the full-order system in a subspace spanned by polynomials of the full-order operator acting on a starting vector.

This polynomial space can be extended by adding negative powers of the full-order operator to arrive at so-called extended Krylov subspaces (EKS) [47, 22]. Polynomial and extended Krylov subspaces are particularly efficient for large, sparse, shifted systems as many matrix-vector products are required for their construction. A finite-



difference discretization of a wave equation is sparse; however, the system is only a shifted system in the absence of the nonlinearity introduced by the PML. Nonetheless, by using linearization techniques, PKS and EKS reduced-order models can still be obtained for wave equations on open domains [29, 28].

Rational Krylov subspaces (RKS), on the other hand, can deal with nonlinear systems directly. First introduced to compute eigenvalue problems [70], they are now used in many areas of scientific computing like control theory or parabolic PDEs [11, 50, 55].

Wave equations on open domains are challenging compared to diffusion equations due to the absorbing boundary conditions and due to the fact that information can travel long distances with waves. Therefore, the use of Krylov model order reduction techniques for wave problems has been limited so far [29, 68, 28].

## 1.2. REDUCED-ORDER MODELING IN APPLICATIONS

Several types of wave phenomena and special model reduction techniques for them are studied in this thesis. Localized wavefields in resonators are studied which have connections to classical eigenvalue problems. Waves that transport energy over many wavelengths are studied as these wave phenomena usually lead to a large computational grid. In addition, attention is paid to waves in lossy media since loss of information in such media allows for very small models.

### 1.2.1. REDUCED-ORDER MODELING IN GEOPHYSICS

In inverse problems in geophysics, several physical sensing mechanisms are used to reveal different properties of rock formations. For instance, a ground penetrating radar is used to estimate the electrical permittivity and conductivity of rock. Krylov-based electromagnetics solvers for the Maxwell system in the diffusion approximation have long been the workhorse of the geophysics community [11, 21, 51]. In such a solver, the discretized system of equations is projected onto a Krylov subspace in an iterative fashion. The projected system forms a reduced-order model (ROM) from which approximate solutions to the full-order system can be drawn.

Since rock formations in the Earth are typically conductive and thus lossy, the Maxwell equations do not form a hyperbolic system in this case. In this thesis we describe how model order reduction techniques can be applied to lossy wave equations and discuss the convergence of several Krylov based model order reduction techniques. As an application, the ground penetrating radar is considered. Further, model order reduction techniques for the lossy Maxwell equations are compared to the ones used for the diffusion approximation using a three-dimensional diffusion example.

Seismic exploration is another mechanism used to unveil the structure of the Earth, and is used extensively in geophysics to make accurate and large images of the subsurface for hydrocarbon exploration. Solving the resulting large models even for a single frequency is challenging due to the large size of the meshes required for such simulations. Model reduction techniques have been used to speed up such simulations via domain decomposition [26] or projection upon time-domain snapshots [63]. In the scope of this thesis, however, we do not deal with the elastodynamic wave equation, which describes the propagation of seismic waves.

### 1.2.2. RESONANCES IN OPEN DISPERSIVE SYSTEMS

Resonance is the basic principle behind most music instruments. The resonances of a violin, for instance, determine the timbre of a violin [44], and thus determine the response of it.

In model order reduction, a structure that only supports a few resonances can be approximated very efficiently once those resonances are known. In the context of this thesis we study resonances of dispersive materials – those for which the material parameters are dependent on the frequency of a wave. The efficient computation of time- and frequency-domain electromagnetic wavefields in linear dispersive media is extremely important in a wide variety of applications ranging from bioelectromagnetics [37] to nano-optics [58]. In the optical frequency range many noble metals have frequency-dependent material properties. Straightforward discretization of these dispersive Maxwell equations and solving the resulting nonlinear eigenvalue problem to obtain the resonances is not feasible due to the fine mesh that is needed to capture the complex sub-wavelength features. Common eigenvalue solvers for the resulting nonlinear eigenvalue problem cannot handle such large system of equations.

In this thesis, we develop a model order reduction method for dispersive systems, applicable to arbitrary three-dimensional configurations. The method is particularly effective for sub-wavelength resonating structures as encountered in nano-optics, for example, and both time- and frequency-domain fields can be computed simultaneously. In addition, the modes that dominate the electromagnetic field response at particular receiver locations can be determined directly at negligible additional computational costs and the method allows for so-called frequency sweeps as well, meaning that a *single* reduced-order model can be used for all frequencies within a certain frequency-interval of interest even in case of dispersive frequency-dependent dielectric materials.

### 1.2.3. COMPRESSION OF LARGE-SCALE WAVE PROPAGATION MODELS: PHASE-PRECONDITIONING

In exploration geophysics, waves are sent into the subsurface and the goal is to image the Earth from measured reflections. Earth models are large, and computations of wavefields in these applications are cumbersome. Moreover, waves travel over many wavelengths until they scatter back, leading to long travel times between sources and receivers. The reduced-order models based on rational Krylov subspaces that were developed for wave equations in unbounded domains are not performing well for these types of configurations, since they are based on data interpolation in the frequency-domain, where long travel times mean that the wavefields are highly oscillatory. The reduced-order models are limited by the Nyquist sampling rate in the frequency-domain, so their model order increases for oscillatory fields.

In general, preconditioning of Krylov subspaces for model reduction is a difficult and still open problem. However, we show that effective preconditioning is possible by incorporating asymptotic (high-frequency) wavefield solutions into a Krylov reduction framework.

Specifically, we use ideas from asymptotic Wentzel–Kramers–Brillouin analysis and Filon quadrature to factor out the highly oscillatory phase in wavefields. Then the wave operator can be projected onto a subspace that is frequency-dependent, to ob-



tain a reduced-order model that significantly outperforms the Nyquist limit. The subspace is spanned by functions defined by the product of frequency-independent smooth functions and frequency-dependent highly oscillatory phase terms that are computed from asymptotic analysis. The combination of asymptotic approximations gauged by a projection approach allows us to significantly compress the computed wavefields. The resulting “phase-preconditioned rational Krylov subspace method” (PPRKS) preserves the structure and properties of the PDE, although it is projected onto a subspace that is frequency-dependent.

### 1.3. REDUCED-ORDER MODELING FOR IMAGING

Reduced-order models have long been used for imaging in various applications. Examples range from diffusive optical tomography [48] to controlled source electromagnetics inversion in the oil and gas industry [10, 30]. An inverse problem is often formulated as a minimization problem of the form

$$\underset{m}{\operatorname{argmin}} (||d - \mathcal{F}\{m\}||), \quad (1.1)$$

with  $d$  the measured data and  $\mathcal{F}\{m\}$  the operator that maps the PDE coefficients  $m$  to the data. Assuming convexity of the functional in a neighborhood of the true coefficients and an initial guess in this neighborhood, one can use convex optimization to solve the imaging problem. This requires the evaluation of  $\mathcal{F}\{m\}$ , which means solving a PDE at every iteration of the chosen optimization algorithm for a parameter realization  $m$ . For real-time imaging applications like process control for optical lithography, solving full scale PDEs may be infeasible. Therefore, the operator  $\mathcal{F}\{m\}$  is replaced with a reduced-order operator as a surrogate for the full-scale operator to speed up the computations.

Two types of model reduction technique for inversion algorithms can be distinguished: reduction in the “online” and in the “offline” stage. Generally speaking, the online stage of an inversion algorithm is the stage in which one tries to minimize the mismatch between measured and modeled data. The offline stage consists of computations that can be carried out independent of the measurement data – for instance, computation of a parametric ROM, in case the range of PDE coefficients encountered in inversion is small and known *a priori*.

More recent approaches to inversion based on model order reduction are direct approaches rather than iterative [25]. Here, a ROM that interpolates the measurement data is constructed and the interpretation of such an interpolatory ROM as a projection of a full-scale finite-difference operator leads to direct imaging algorithms.

In process control for optical lithography, a known structure is fabricated on a silicon wafer and optical measurements are used to infer how well the structure was fabricated. In this specific application, the range of possible coefficients for the PDE is limited, and consequently one can construct a reduced-order model that is valid for the whole parameter range of interest [7, 13]. This can be done independent of a measurement or imaging algorithm (“offline”) and can take significant computation time as long as the constructed reduced-order model is easy to evaluate and accurate. The optimization problem (1.1) can then be solved using the reduced-order model as a surrogate for  $\mathcal{F}\{m\}$  to speed up the inverse problem.

In the case the parameter space for which the reduced-order model needs to be accurate is too large or unknown, one cannot construct a reduced-order model offline. This is the case for anisotropic three-dimensional resistivity imaging in boreholes. Since the displacement currents give a negligible contribution to the total induced currents in wet rock formations, the Maxwell equations can be approximated by a diffusion equation in this type of application. Since the response of wet rock formations is directionally dependent on the electric field, the formation is anisotropic and must be modeled with a rank two conductivity tensor.

Model order reduction techniques for these types of problems were developed in [31]. They are used to build a reduced-order model that interpolates the full-order model for all previously computed sets of PDE coefficients encountered during the optimization. Subsequently, the reduced-order model is used to solve the optimization problem stated in equation (1.1), only to run a full computation at the found minimum and to improve the reduced-order model. In this way, a reduced-order model is constructed and exploited “online.”

The described problem is a particularly difficult inverse problem as it seeks to reconstruct a symmetric tensor in three dimensions using diffusive measurements taken on a single line (the borehole) at multiple frequencies. Regularization of formulation (1.1) is therefore very important in this application. In this thesis it is shown that the full anisotropic conductivity tensor can be reconstructed from limited triaxial measurements. For an imaging application with inexact measurements, which is typically an underdetermined problem, this is no longer the case. However, it is shown that one does not obtain worse inversion results using the reduced-order method compared to solving the full PDE at every iteration. Further, the problem becomes easier to regularize in the model reduction framework.

## 1.4. THESIS CONTRIBUTIONS AND OUTLINE

The key contributions of this thesis can be summarized as follows:

1. Development of an efficient model reduction technique for the Maxwell equations with dispersive media [A].
2. Efficient computation of the spontaneous decay rate and resonance fields of arbitrarily shaped three-dimensional dispersive resonators [B].
3. Review and compare model order reduction techniques for the Maxwell equations with applications in geophysics [C].
4. Development of rational Krylov subspace techniques and phase-preconditioning to simulate and compress wave propagation on open domains. The developed techniques can handle localized resonant fields as well as fields propagating over many wavelength [D].
5. Application of “online” and “offline” reduced-order modeling to problems in imaging and inversion (see Chapter 5).



The thesis is structured into six chapters. In this chapter, we briefly touched upon the subjects relevant to the thesis, introduced the problems it deals with and outlined some solution approaches. In the next chapter, the wave equation and Maxwell equations are introduced, the properties of wavefields are discussed, and the dispersive Maxwell equations are used to show how to obtain a discrete system from continuous equations. Chapter 3 deals with model order reduction for forward problems. Several applications are presented to show how model order reduction reduces the computational requirements for problems in wave propagation. In Chapter 4, phase-preconditioning is introduced as a method for model order reduction of wave problems with long propagation delays. Chapter 5 shows the use of ROMs in imaging and inversion. Conclusions from this work are drawn in the final chapter.

## LIST PEER REVIEWED PUBLICATIONS

*Note:* Publications at mathematical conferences and in mathematical journals use alphabetical ordering of the authors, whereas publications in electrical engineering and optics use other guidelines. The references are in chronological order.

- [A] J. ZIMMERLING, LEI WEI, PAUL URBACH, AND ROB REMIS, *A Lanczos model-order reduction technique to efficiently simulate electromagnetic wave propagation in dispersive media*, *Journal of Computational Physics*, 315 (2016), pp. 348 – 362.
- [B] J. ZIMMERLING, LEI WEI, PAUL URBACH, AND ROB REMIS, *Efficient computation of the spontaneous decay rate of arbitrarily shaped 3D nanosized resonators: A Krylov model-order reduction approach*, *Applied Physics A*, 122 (2016), p. 158.
- [C] J. ZIMMERLING, V. DRUSKIN, R. REMIS, AND M. ZASLAVSKY, *Model-order reduction of electromagnetic fields in open domains*, *Geophysics*, 83 (2018), pp. WB61–WB70.
- [D] V. DRUSKIN, R. REMIS, M. ZASLAVSKY, AND J. ZIMMERLING, *Compressing large-scale wave propagation models via phase-preconditioned rational Krylov subspaces*, *SIAM Multiscale Modeling & Simulation*, (available on arxiv <https://arxiv.org/abs/1711.00942>), 2017, *submitted*.

## LIST OF CONFERENCE PROCEEDINGS AND TALKS

- [a] V. DRUSKIN, R. REMIS, M. ZASLAVSKY, AND J. ZIMMERLING, *Krylov model-order reduction of transient seismic wave propagation in unbounded domains*, in *Book of Abstracts, SIAM Conference on Mathematical and Computational Issues in the Geosciences (SIAM GS15)*, (Stanford, USA), June 2015, p. 105.

- [b] J. ZIMMERLING, L. WEI, H. URBACH, AND R. REMIS, *Efficient computation of the spontaneous decay rate of arbitrarily shaped 3D nanosized resonators – A Krylov model-order reduction approach*, in Conference proceedings, The 6th Conference on Metamaterials, Photonic Crystals, and Plasmonics (META 2015), (New York, USA), July 2015, pp. 657 – 662.
- [c] R. REMIS, V. DRUSKIN, A. MAMONOV, M. ZASLAVSKY, AND J. ZIMMERLING, *reduced-order models for large scale wave propagation*, in Book of abstracts, The 12th International Conference on Mathematical and Numerical Aspects of Wave Propagation (WAVES 2015), (Karlsruhe, Germany), July 2015, pp. 51 – 52.
- [d] V. DRUSKIN, R. REMIS, M. ZASLAVSKY, AND J. ZIMMERLING, *Efficient computation of electromagnetic wave fields on unbounded domains using stability-corrected wave functions and Krylov subspace projection methods*, in Conference proceedings (IEEE Xplore), International Conference on Electromagnetics in Advanced Applications (ICEAA15), (Turin, Italy), September 2015, pp. 19 – 22, DOI: 10.1109/ICEAA.2015.7297066, *invited talk*.
- [e] J. ZIMMERLING AND R. REMIS, *Krylov model-order reduction expansions for electromagnetic wave fields in strongly resonating structures*, in Conference proceedings (IEEE Xplore), International Conference on Electromagnetics in Advanced Applications (ICEAA15), (Turin, Italy), September 2015, pp. 23 – 26, DOI: 10.1109/ICEAA.2015.7297067, *invited talk*.
- [f] V. DRUSKIN, R. REMIS, M. ZASLAVSKY, AND J. ZIMMERLING, *Perfectly matched layers and rational Krylov subspaces with adaptive shifts for Maxwell systems*, in Book of Abstracts, SIAM Conference on Applied Linear Algebra (SIAM LA15), (Atlanta, USA), October 2015, p. 80, *invited talk*.
- [g] J. ZIMMERLING, V. DRUSKIN, R. REMIS AND M. ZASLAVSKY, *Efficient mode computations in open, dispersive, optical resonators*, in Conference Proceedings, Bi-Annual Meeting of the European Optical Society (EOSAM 2016), (Berlin, Germany), September 2016.
- [h] J. ZIMMERLING, V. DRUSKIN, R. REMIS, AND M. ZASLAVSKY, *Asymptotically corrected reduced-order modelling for wavefield computation with multiple sources*, in (78th EAGE Conference) - Workshop13, Methods and Challenges of Seismic Wave Modelling for Seismic Imaging - Seismic Modelling in the Frequency Domain, (Vienna, Austria), June 2016, p. C03.
- [i] R. REMIS, V. DRUSKIN, M. ZASLAVSKY, AND J. ZIMMERLING, *On rational Krylov subspace methods for large scale time and frequency-domain wavefield computations*, in Book of abstracts, SIAM Annual Meeting (SIAM AN16), (Boston MA, USA), July 2016, p. 100, *SIAG/LA minisymposium*.



- [j] [J. ZIMMERLING](#), V. DRUSKIN, R. REMIS, AND M. ZASLAVSKY, *Asymptotically corrected Krylov subspace model-order reduction of wavefields in travel-time dominated structures*, in Book of Abstracts, SIAM Annual Meeting (SIAM AN16), (Boston MA, USA), July 2016, p. 101, *SIAG/LA minisymposium*.
- [k] V. DRUSKIN, R. REMIS, M. ZASLAVSKY, AND [J. ZIMMERLING](#), *Phase-preconditioned rational Krylov subspaces for model reduction of large-scale wave propagation*, in Book of Abstracts, Numerical Linear Algebra and Applications (NL2A), (Luminy, France), October 2016, p. 49, *invited talk*.
- [l] [J. ZIMMERLING](#), V. DRUSKIN, R. REMIS, AND M. ZASLAVSKY, *Model order reduction of electromagnetic wave fields in open domains*, in Proceedings of the Sixth International Symposium in Three-Dimensional Electromagnetics, (Berkely, CA, USA), March 2017.
- [m] V. DRUSKIN, R. REMIS, M. ZASLAVSKY, AND [J. ZIMMERLING](#), *Stability-corrected wave functions and structure-preserving rational Krylov methods for large-scale wavefield simulations on open domains*, in Householder Symposium XX, Program and Abstracts (HHXX), (Blacksburg,VA, USA), June 2017.
- [n] [J. ZIMMERLING](#), V. DRUSKIN, R. REMIS, AND M. ZASLAVSKY, *Phase-preconditioned rational Krylov subspaces for wave simulation*, in Householder Symposium XX, Program and Abstracts (HHXX), (Blacksburg,VA, USA), June 2017.
- [o] V. L. DRUSKIN, R. F. REMIS, M. ZASLAVSKY, AND [J. ZIMMERLING](#), *Projection-based model-order reduction of large-scale Maxwell systems*, in 2017 International Conference on Electromagnetics in Advanced Applications (ICEAA), September 2017, pp. 385–388, *invited talk*.
- [p] R. REMIS, V. DRUSKIN, M. ZASLAVSKY, AND [J. ZIMMERLING](#), *Krylov subspaces for large scale wave field simulations*, in Icerm Workshop, Program and Abstracts, (Providence, RI, USA), November 2017, *invited talk*.
- [q] [J. ZIMMERLING](#), V. DRUSKIN, R. REMIS, AND M. ZASLAVSKY, *Phase-preconditioned rational Krylov subspaces for large-scale wave simulation*, in Icerm Workshop, Program and Abstracts, (Providence, RI, USA), November 2017, *invited talk*.
- [r] [J. ZIMMERLING](#), V. DRUSKIN, R. REMIS, AND M. ZASLAVSKY, *Rational Krylov subspaces for wavefield applications*, SIAM Conference on Applied Linear Algebra (SIAM ALA18), Program and Abstracts, (Hong Kong, China), May 2018, p. 103, *invited talk*.

# 2

2

## FROM PHYSICS TO DISCRETE SYSTEMS

*I didn't have time to write a short letter, so I wrote a long one instead.*

Mark Twain



## 2.1. INTRODUCTION

IN this chapter, we discuss how to obtain a discrete system from partial differential equations that describe the physics of electromagnetic fields or acoustic fields. Special attention is paid to wavefield properties such as symmetry and passivity and how they translate from analytical to the discrete domain. This chapter is divided into seven sections. First, the wave equation is discussed followed by the Maxwell equations. The properties of fields satisfying these equations are discussed as well. In order to find numerical approximations to the solutions of such equations the governing wave equations are spatially discretized. Since wave equations on open domains are considered in this thesis, absorbing boundary conditions are examined. Finally, formal solutions to the equations are given and the terminology used throughout the thesis is defined.

## 2.2. SCALAR WAVE EQUATION ON AN UNBOUNDED DOMAIN

We consider the scalar, isotropic, continuous wave equation on  $\mathbb{R}^k \times [0, \infty[$

$$\Delta u - \frac{1}{v^2} u_{tt} = -\frac{1}{v^2} \delta(t) \delta(x - x_S), \quad u|_{t=0} = 0, \quad u_t|_{t=0} = 0. \quad (2.1)$$

In this equation,  $\Delta$  denotes the  $k$ -dimensional Laplace operator and the position vector lies in a  $k$ -dimensional Euclidean space, i.e.  $x \in \mathbb{R}^k$  ( $1 \leq k \leq 3$ ). Furthermore,  $v(x) > 0$  is a wavespeed distribution in  $L^\infty[\mathbb{R}^k]$ , and  $u(x, t)$  is the wavefield with a compact support for all finite times.

After Laplace transformation, equation (2.1) turns into the Helmholtz equation

$$\Delta u - \frac{s^2}{v^2} u = -\frac{1}{v^2} \delta(x - x_S), \quad (2.2)$$

where  $s$  is the complex Laplace parameter with  $\Re(s) \geq 0$ . The Laplace domain wavefield  $u$  satisfies the limiting absorption principle, i.e.  $u$  vanishes at infinity for  $\Re(s) > 0$  and converges to the solution of Helmholtz's equation that satisfies the outgoing radiation condition as the Laplace parameter  $s$  approaches the imaginary axis via the right-half of the complex  $s$ -plane.

Let  $\Omega$  be a bounded subdomain of  $\mathbb{R}^k$  such that  $x_S \in \Omega$ . We now equivalently reduce the original problem on the unbounded domain to a problem on  $\Omega$  by considering equation (2.2) in the weak formulation and testing this equation with a testing function  $p$ . This gives

$$\int_{\Omega} \overline{p} \left( \Delta - \frac{s^2}{v^2} \right) u \, dx = -\frac{1}{v(x_S)^2} \overline{p}(x_S), \quad (2.3)$$

where the overbar denotes complex conjugation. After integration by parts, we obtain

$$-\int_{\Omega} (\nabla \overline{p}) \cdot (\nabla u) \, dx - \int_{\Omega} \overline{p} \frac{s^2}{v^2} u \, dx + \int_{\partial\Omega} \overline{p} \frac{\partial u}{\partial n} \, dx = -\frac{1}{v(x_S)^2} \overline{p}(x_S), \quad (2.4)$$

with  $\frac{\partial u}{\partial n}$  the derivative of  $u$  in the direction of the outward-pointing normal on  $\partial\Omega$ . Finally, introducing the Dirichlet-to-Neumann (DtN) map  $D(s)$  on  $\partial\Omega$  such that  $\frac{\partial u}{\partial n} = D(s)u$ ,

the above equation can be written as

$$-\int_{\Omega} (\nabla \bar{p}) \cdot (\nabla u) dx - \int_{\Omega} \bar{p} \frac{s^2}{v^2} u dx + \int_{\partial\Omega} \bar{p} D(s) u dx = -\frac{1}{v(x_S)^2} \bar{p}(x_S). \quad (2.5)$$

Without the boundary integral (third term on the left-hand side of the above equation) this equation is linear in  $s^2$ ; the DtN map, however, is a nonlinear function of frequency  $s$  [20].

#### NOTATION

To better draw similarities between continuous and discrete formulations, we will treat the complex-valued functions  $u$  and  $p$  as vectors from  $\mathbb{R}^{\infty}$  in the linear algebraic derivations and introduce the inner product

$$p^H u = \int_{\Omega} \bar{p} u dV. \quad (2.6)$$

We note, that  $u$  and  $p$  for  $k > 1$  have singularities at  $x_S$  that may make this inner product divergent. To avoid this, we assume by default that instead of  $\delta(x - x_S)$  we have some regular approximation of the delta function. After discretization,  $u$  and  $p$  become finite-dimensional vectors from  $\mathbb{R}^N$  and the issue of diverging integrals due to singularities disappears. In this notation, superscript  $H$  denotes the Hermitian transpose for vectors and an inner product with complex conjugation for functions. Analogously, the superscript  $T$  denotes the ordinary transpose for vectors and a bilinear form for functions, i.e. equation (2.6) without complex conjugation. Since the vectors are complex-valued such a bilinear form does not induce a norm. Operators in an integral form are printed with capital italic letters like  $Q$ , all other operators are printed in calligraphic font like  $\mathcal{A}$ . Physical vectors are printed bold, e.g., the electrical field strength is printed as  $\mathbf{E}$ . Laplace transformed field quantities carry a hat, e.g.  $\hat{\mathbf{E}}$ . Functions are printed in upright front in the time-domain and in italic in the Laplace domain. For linear combinations such as

$$q_m = \alpha_1 g^{[1]} + \alpha_2 g^{[2]} + \dots + \alpha_m g^{[m]},$$

with coefficients  $\alpha_i$  and expansion functions  $g^{[i]}$ , we write  $q_m = G_m z$  with  $z = [\alpha_1, \dots, \alpha_m]^T$  and the expansion functions are stored as columns in the function array  $G_m$ , i.e.,  $G_m \in \mathbb{R}^{\infty \times m}$  (sometimes referred to as a quasimatrix [81]). Finally, finite-dimensional matrices are printed using a capital sans serif font (like  $A$ ).

Using this notation, we now introduce the wave operator  $Q(s)$  to rewrite equation (2.5) as

$$p^H Q(s) u = -\frac{1}{v(x_S)^2} \bar{p}(x_S). \quad (2.7)$$

We note that the real and imaginary parts of  $Q(s)$  are self-adjoint.

## 2.3. MAXWELL EQUATIONS ON AN UNBOUNDED DOMAIN

We now consider the electromagnetic field equations since most applications in this thesis are of electromagnetic nature. The associated PDEs – the Maxwell equations – and



their variants are introduced in this section. The resulting field equations will be written in a unified form to introduce model order reduction concepts for all equations at once.

Electromagnetic fields are governed by the Maxwell equations

$$-\nabla \times \mathbf{H} + \mathbf{J}^{\text{cond}} + \partial_t \mathbf{D} = -\mathbf{J}^{\text{ext}} \quad (2.8)$$

and

$$\nabla \times \mathbf{E} + \partial_t \mathbf{B} = -\mathbf{K}^{\text{ext}}. \quad (2.9)$$

The fundamental unknowns in these equations are  $\mathbf{H}$  the magnetic field strength and  $\mathbf{E}$  the electrical field strength. These fields respond to the external sources on the right-hand side of equations (2.8) and (2.9). Here,  $\mathbf{K}^{\text{ext}}$  denotes an external magnetic current density and  $\mathbf{J}^{\text{ext}}$  an external electrical current density. Again, we limit our domain of interest to  $\Omega$  and assume that the sources  $\mathbf{K}^{\text{ext}}$  and  $\mathbf{J}^{\text{ext}}$  are compactly supported on this domain  $\Omega$ .

Further,  $\mathbf{J}^{\text{cond}}$  and  $\mathbf{D}$  are the conduction current density, and electrical flux density, respectively, and they are related to the electric field strength  $\mathbf{E}$  via constitutive relations. Finally,  $\mathbf{B}$  is the magnetic flux density which is related to the magnetic field strength  $\mathbf{H}$  via a constitutive relation. These constitutive relations describe the reaction of a material to the presence of a field. We consider four cases separately in this thesis:

1. **LOSSLESS ELECTROMAGNETICS** – For lossless, instantaneously reacting, isotropic media the constitutive relations are

$$\mathbf{B} = \mu \mathbf{H}, \quad \mathbf{J}^{\text{cond}} = \mathbf{0}, \quad \mathbf{D} = \varepsilon \mathbf{E}, \quad (2.10)$$

with  $\mu > 0$  the permeability, and  $\varepsilon > 0$  the permittivity. Both medium parameters are spatially dependent scalars.

2. **LOSSY ELECTROMAGNETICS** – For lossy, instantaneously reacting, isotropic media the conduction current no longer vanishes. The constitutive relations are

$$\mathbf{B} = \mu \mathbf{H}, \quad \mathbf{J}^{\text{cond}} = \sigma \mathbf{E}, \quad \mathbf{D} = \varepsilon \mathbf{E}, \quad (2.11)$$

with  $\sigma \geq 0$  the conductivity.

3. **SECOND-ORDER ELECTRIC DISPERSION** – Dispersive media are not instantaneously reacting – the constitutive relations generally include a (temporal) differential equation. In this thesis we are interested in materials that are electrically dispersive and whose constitutive relations are (at most) given by a second-order differential equation. The constitutive relation for such a material are

$$\mathbf{B} = \mu \mathbf{H}, \quad \mathbf{J}^{\text{cond}} = \sigma \mathbf{E}, \quad \mathbf{D} = \varepsilon \mathbf{E} + \mathbf{P}, \quad (2.12)$$

where the polarization density  $\mathbf{P}$  is related to  $\mathbf{E}$  via a second-order constitutive relation

$$\beta_3 \partial_t^2 \mathbf{P} + \beta_2 \partial_t \mathbf{P} + \beta_1 \mathbf{P} = \beta_0 \mathbf{E}, \quad (2.13)$$

with  $\beta_i$ ,  $i = 0, 1, 2, 3$ , the parameters describing the particular dispersive material.

4. **ANISOTROPIC DIFFUSION APPROXIMATION** – In the case the magnitude of the conduction current density  $\mathbf{J}^{\text{cond}}$  is much larger than the magnitude of the displacement current density  $\partial_t \mathbf{D}$ , the latter can be neglected and the Maxwell equations can be approximated by a diffusion equation. For these types of fields, the constitutive relations are

$$\mathbf{J}^{\text{cond}} = \sigma \mathbf{E}, \quad \text{and} \quad \mathbf{B} = \mu \mathbf{H}. \quad (2.14)$$

In the anisotropic case,  $\sigma$  is a positive-definite tensor of rank two.

### 2.3.1. LOSSLESS, INSTANTANEOUSLY REACTING, ISOTROPIC MEDIA

For lossless, instantaneously reacting, isotropic media the Maxwell equations simplify to

$$-\nabla \times \mathbf{H} + \varepsilon \partial_t \mathbf{E} = -\mathbf{J}^{\text{ext}} \quad (2.15)$$

and

$$\nabla \times \mathbf{E} + \mu \partial_t \mathbf{H} = -\mathbf{K}^{\text{ext}}, \quad (2.16)$$

where  $\varepsilon$  and  $\mu$  are spatially dependent scalars and one obtains a hyperbolic wave equation in a first-order formulation. For lossless electromagnetics the constitutive relation reads  $\mathbf{D} = \varepsilon \mathbf{E}$  and the conduction current vanishes. These equations can be written in matrix-operator form as

$$(\mathcal{D} + \mathcal{M} \partial_t) \mathcal{U} = \mathcal{B}', \quad (2.17)$$

with

$$\mathcal{D} = \begin{bmatrix} 0 & -\nabla \times \\ \nabla \times & 0 \end{bmatrix}, \quad \text{and} \quad \mathcal{M} = \begin{bmatrix} \varepsilon & 0 \\ 0 & \mu \end{bmatrix} \quad (2.18)$$

where  $\mathcal{D}$  contains the spatial derivatives,  $\mathcal{M}$  the medium parameters, the field quantities are combined in

$$\mathcal{U} = [\mathbf{E}_x, \mathbf{E}_y, \mathbf{E}_z, \mathbf{H}_x, \mathbf{H}_y, \mathbf{H}_z]^T$$

and the corresponding sources are collected in

$$\mathcal{B}' = -[\mathbf{J}_x^{\text{ext}}, \mathbf{J}_y^{\text{ext}}, \mathbf{J}_z^{\text{ext}}, \mathbf{K}_x^{\text{ext}}, \mathbf{K}_y^{\text{ext}}, \mathbf{K}_z^{\text{ext}}]^T.$$

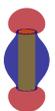
### 2.3.2. LOSSY, INSTANTANEOUSLY REACTING, ISOTROPIC MEDIA

In the case losses are present but the material is still instantaneously reacting this matrix operator form becomes

$$(\mathcal{D} + \mathcal{S} + \mathcal{M} \partial_t) \mathcal{U} = \mathcal{B}', \quad \text{with} \quad \mathcal{S} = \begin{bmatrix} \sigma & 0 \\ 0 & 0 \end{bmatrix}, \quad (2.19)$$

where  $\mathcal{S} \mathcal{U}$  contains the conduction current density  $\sigma \mathbf{E}$ . In the Laplace domain this equation can obviously also be written as a shifted system

$$(\mathcal{M}^{-1} \mathcal{D} + \mathcal{M}^{-1} \mathcal{S} + s \mathcal{I}) \hat{\mathcal{U}} = \mathcal{M}^{-1} \hat{\mathcal{B}}', \quad (2.20)$$



assuming vanishing initial conditions and with  $\mathcal{I}$  the identity operator. Testing equation (2.20) in the Laplace domain with a test field  $(\hat{\mathbf{E}}^t \quad \hat{\mathbf{H}}^t)$  in a  $\mathcal{W}_{\text{EM}} = \begin{bmatrix} \varepsilon & 0 \\ 0 & -\mu \end{bmatrix}$  bilinear form and integrating over  $\Omega$  we find

$$\int_{\Omega} -\hat{\mathbf{H}}^t \cdot \nabla \times \hat{\mathbf{E}} - \hat{\mathbf{E}}^t \cdot \nabla \times \hat{\mathbf{H}} + \hat{\mathbf{E}}^t \cdot (\sigma + s\varepsilon)\hat{\mathbf{E}} - \hat{\mathbf{H}}^t \cdot s\mu\hat{\mathbf{H}} dV = \int_{\Omega} -\hat{\mathbf{E}}^t \cdot \mathbf{j}^{\text{ext}} + \hat{\mathbf{H}}^t \cdot \hat{\mathbf{K}}^{\text{ext}} dV, \quad (2.21)$$

and using  $\nabla \cdot (\hat{\mathbf{E}}^t \times \hat{\mathbf{H}}) = \hat{\mathbf{H}} \cdot (\nabla \times \hat{\mathbf{E}}^t) - \hat{\mathbf{E}}^t \cdot (\nabla \times \hat{\mathbf{H}})$  this equation can be rewritten as

$$\begin{aligned} & \int_{\Omega} -(\nabla \times \hat{\mathbf{H}}^t) \cdot \hat{\mathbf{E}} - (\nabla \times \hat{\mathbf{E}}^t) \cdot \hat{\mathbf{H}} + \hat{\mathbf{E}}^t \cdot (\sigma + s\varepsilon)\hat{\mathbf{E}} - \hat{\mathbf{H}}^t \cdot s\mu\hat{\mathbf{H}} dV \\ & + \int_{\partial\Omega} (\hat{\mathbf{E}}^t \times \hat{\mathbf{H}} + \hat{\mathbf{H}}^t \times \hat{\mathbf{E}}) \cdot n dS = \int_{\Omega} -\hat{\mathbf{E}}^t \cdot \mathbf{j}^{\text{ext}} + \hat{\mathbf{H}}^t \cdot \hat{\mathbf{K}}^{\text{ext}} dV, \end{aligned} \quad (2.22)$$

Here  $n$  is the outward-pointing unit-normal on  $\partial\Omega$ . In our modeling approach we surround our domain of interest by an absorbing boundary layer that simulates the extension of our domain to infinity and absorbs all outgoing waves. The layer itself is terminated using the boundary conditions  $n \times \mathbf{E} = 0$  and  $n \times \mathbf{E}^t = 0$ , which need to be satisfied by the test and trial field. Under these conditions the boundary integral in the above equation vanishes and we obtain a symmetric operator. The introduction of the absorbing boundary layer is discussed in section 2.6.

To summarize, with boundary conditions included, the Maxwell equations form a symmetric operator in the  $\mathcal{W}_{\text{EM}}$  bilinear form from equation (2.21). Since the anti-symmetric boundary integral  $\int_{\partial\Omega} \dots dS$  vanishes for these boundary conditions we can swap the trial and test functions  $(\hat{\mathbf{E}}^t \quad \hat{\mathbf{H}}^t)$  and  $(\hat{\mathbf{E}} \quad \hat{\mathbf{H}})$  and obtain the same result. Finally, we rewrite the symmetric bilinear form (2.21) as

$$p^T Q(s) u = p^T b, \quad (2.23)$$

which is of the same form as the Helmholtz equation and  $Q(s)$  symmetric in the transpose bilinear form. (Or with  $Q(\bar{s})$  as adjoint in the Hermitian inner product). We do this in the understanding that  $u$  and  $p$  are 6-tuples with tuple elements from  $\mathbb{R}^{\infty}$  in the case of the instantaneously reacting, first-order electromagnetic equation.

### 2.3.3. DISPERSIVE MEDIA

In case the medium is electrically dispersive, the relation between the electric flux density and field strength is more involved. In this thesis we consider dispersive materials that follow a Debye, Drude, or Lorentz dispersion relation. To include this into the Maxwell equations, the polarization density  $\mathbf{P}$  is introduced and the electric flux density is given by  $\mathbf{D} = \varepsilon\mathbf{E} + \mathbf{P}$  with  $\varepsilon = \varepsilon_0\varepsilon_{\infty}$  and  $\varepsilon_{\infty}$  the instantaneous or high-frequency relative permittivity. The polarization vector  $\mathbf{P}$  is related to the electric field strength  $\mathbf{E}$  via the second-order constitutive relation

$$\beta_3 \partial_t^2 \mathbf{P} + \beta_2 \partial_t \mathbf{P} + \beta_1 \mathbf{P} = \beta_0 \mathbf{E}, \quad (2.24)$$

where the  $\beta_i$ ,  $i = 0, 1, 2, 3$ , are parameters describing the particular dispersive material of interest. For example, for a Drude material we have  $\beta_0 = \varepsilon_0 \omega_p^2$ ,  $\beta_1 = 0$ ,  $\beta_2 = \gamma_p$ , and

$\beta_3 = 1$ , where  $\omega_p$  is the volume plasma frequency and  $\gamma_p$  the collision frequency. The  $\beta$ -coefficients of other commonly used materials are summarized in Table 2.1. For a Lorentz medium we obtain the relation of a damped oscillator with damping term  $\beta_2 = 2\delta$ .

Table 2.1: Parameters to obtain common dispersion models with the general second-order dispersion model. Here,  $\tau$  is the characteristic relaxation time,  $\varepsilon_s$  is the static relative permittivity,  $\omega_p$  is the volume plasma frequency,  $\gamma_p$  is the collision frequency,  $\omega_0$  is the resonant plasma frequency, and  $\delta$  is the damping coefficient.

Medium	$\beta_0$	$\beta_1$	$\beta_2$	$\beta_3$
Lorentz	$\varepsilon_0(\varepsilon_0 s - \varepsilon_0 \infty)\omega_0^2$	$\omega_0^2$	$2\delta$	1
Drude	$\varepsilon_0\omega_p^2$	0	$\gamma_p$	1
Debye	$\varepsilon_0(\varepsilon_0 s - \varepsilon_0 \infty)$	1	$\tau$	0
Conductivity	$\sigma$	0	1	0

To try to find a self-adjoint formulation, we first rewrite the second-order constitutive relation in first-order form. To this end, we introduce the auxiliary field

$$\mathbf{U} = -\partial_t \mathbf{P} \quad (2.25)$$

and rewrite Eq. (2.24) as

$$\beta_3 \partial_t \mathbf{U} + \beta_2 \mathbf{U} - \beta_1 \mathbf{P} + \beta_0 \mathbf{E} = 0. \quad (2.26)$$

Combining these last two equations with Maxwell's equations, we arrive at the first-order system

$$-\nabla \times \mathbf{H} - \mathbf{U} + \varepsilon \partial_t \mathbf{E} = -\mathbf{J}^{\text{ext}}, \quad (2.27)$$

$$\mathbf{U} + \partial_t \mathbf{P} = 0, \quad (2.28)$$

$$\beta_2 \mathbf{U} - \beta_1 \mathbf{P} + \beta_0 \mathbf{E} + \beta_3 \partial_t \mathbf{U} = 0, \quad (2.29)$$

and

$$\nabla \times \mathbf{E} + \mu \partial_t \mathbf{H} = -\mathbf{K}^{\text{ext}}. \quad (2.30)$$

These equations can be written in matrix-operator form as

$$(\mathcal{D} + \mathcal{S} + \mathcal{M} \partial_t) \mathcal{U} = \mathcal{B}', \quad (2.31)$$

where  $\mathcal{U}$  and  $\mathcal{B}$  are the field and source vectors given by

$$\mathcal{U} = [E_x, E_y, E_z, P_x, P_y, P_z, U_x, U_y, U_z, H_x, H_y, H_z]^T \quad (2.32)$$

and

$$\mathcal{B}' = -[J_x^{\text{ext}}, J_y^{\text{ext}}, J_z^{\text{ext}}, 0, 0, 0, 0, 0, 0, K_x^{\text{ext}}, K_y^{\text{ext}}, K_z^{\text{ext}}]^T, \quad (2.33)$$



respectively. Furthermore,  $\mathcal{D}$  is a spatial differentiation matrix containing the curl operators from Maxwell's equations and  $\mathcal{S}$  and  $\mathcal{M}$  are medium matrices containing the medium parameters  $\beta_i$ ,  $\varepsilon$ , and  $\mu$ . Explicitly, the differentiation matrix is given by

$$\mathcal{D} = \begin{bmatrix} 0 & 0 & 0 & -\nabla \times \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \nabla \times & 0 & 0 & 0 \end{bmatrix} \quad (2.34)$$

and the medium matrices are

$$\mathcal{S} = \begin{bmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 \\ \beta_0 & -\beta_1 & \beta_2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad \mathcal{M} = \begin{bmatrix} \varepsilon & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \beta_3 & 0 \\ 0 & 0 & 0 & \mu \end{bmatrix}. \quad (2.35)$$

With these medium matrices in equation (2.20) we obtain a self-adjoint operator  $Q(s)$  in a bilinear form weighted with  $\mathcal{W}_{DP} = \text{diag}(\varepsilon, \beta_1/\beta_0, -\beta_3/\beta_0, -\mu)$  which again can be written in the symmetric integral form

$$p^T Q(s) u = p^T b. \quad (2.36)$$

### 2.3.4. ANISOTROPIC DIFFUSION APPROXIMATION

The electromagnetic diffusion approximation holds if the magnitude of the conduction current density  $\sigma \mathbf{E}$  is much larger than the magnitude of the displacement current density  $\partial_t \mathbf{D}$  in equation (2.8). This is the case for very lossy materials with a high conductivity or for very slowly varying fields; i.e. at low frequencies. Dropping the displacement current and eliminating  $\mathbf{H}$  in the Maxwell equations leads to the second-order PDE

$$\nabla \times \nabla \times \mathbf{E} + \mu \sigma \partial_t \mathbf{E} = -\nabla \times \mathbf{K}^{\text{ext}} - \partial_t \mu \mathbf{J}^{\text{ext}} \quad (2.37)$$

known as the diffusion approximation in the  $\mathbf{E}$ -field formulation. In case the conductivity is directionally dependent the medium is said to be anisotropic and the conductivity  $\sigma$  is a  $3 \times 3$  positive definite (rank 2) tensor as opposed to a positive scalar in the isotropic case. In the diffusive case we can define a Hermitian operator with a real spectrum via the weak form. Then using integration by parts we obtain a Hermitian operator which can again be written in the form  $p^H Q(s) u = p^H b$ .

## 2.4. SOME GENERAL WAVEFIELD PROPERTIES

In this section, the properties satisfied by the fields and operators presented in the previous section are discussed. These properties will be revisited after discretization in the next section and after model order reduction in the next chapter. The properties discussed are fundamental and need to be preserved after discretization and model order reduction.

In this thesis, we only consider passive materials – materials that do not produce any energy by themselves. This means that for isotropic materials  $\sigma \geq 0$  for wavefields

and  $\sigma > 0$  for diffusion needs to hold. Using passive materials leads to global energy conservation so that the energy in the region of interest  $\Omega$  is smaller or equal to the energy fed into the systems by initial conditions or sources. Therefore, all our operators  $Q(s)$  are passive as well. Passivity can be defined via the numerical range of the operator.

**DEFINITION 2.1.** — We define the nonlinear numerical range as

$$\mathcal{N}\{A(s)\} = \left\{ s \in \mathbb{C} : x^H A(s)x = 0 \quad \text{for some nontrivial } x \in \mathbb{C}^k \right\}. \quad (2.38)$$

This coincides with the ordinary numerical range definition for linear shifted systems  $A(s) = A' - sI$  (see [45], [32]). Passivity of a dynamic system is equivalent to the condition

$$\Re \mathcal{N}\{Q(s)\} \leq 0, \quad (2.39)$$

so that the region of convergence of the Laplace transform is the right half-plane. Passive systems do not produce any energy and strictly passive systems dissipate the input energy. Further, passive systems are always stable [78].

Time-domain quantities like the electric field strength are real functions of time. Consequently, the operator  $Q(s)$  and the field  $u(s)$  satisfy the Schwarz reflection principle

$$Q(\bar{s}) = \overline{Q(s)} \quad \text{and} \quad u(\bar{s}) = \overline{u(s)}, \quad (2.40)$$

and the spectrum of  $Q(s)$  is closed under conjugation.

Further, the symmetry bilinear forms of the various equations can be related to a field Lagrangian. More specifically, for the instantaneously reacting electromagnetic equations, an inner product with the symmetry operator  $\mathcal{W}_{EM}$  yields the free field Lagrangian as

$$\mathcal{L}_{\text{free}} = \frac{1}{2} u^H \mathcal{W}_{EM} u = \frac{1}{2} \int_{\Omega} \varepsilon |\hat{\mathbf{E}}|^2 - \mu |\hat{\mathbf{H}}|^2 dV. \quad (2.41)$$

Similarly, the symmetry operator found for the Maxwell equations with the second-order dispersion relation  $\mathcal{W}_{DP}$  yields

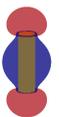
$$\mathcal{L} = \frac{1}{2} u^H \mathcal{W}_{DP} u = \mathcal{L}_{\text{free}} + \frac{1}{2} \int_{\Omega} \frac{\beta_1}{\beta_0} |\hat{\mathbf{P}}|^2 dV - \frac{1}{2} \int_{\Omega} \frac{\beta_3}{\beta_0} |\partial_t \hat{\mathbf{P}}|^2 dV. \quad (2.42)$$

As an illustration, if we substitute the values for a Lorentz medium we find

$$\mathcal{L} = \frac{1}{2} u^H \mathcal{W}_{DP} u = \mathcal{L}_{\text{free}} + \frac{1}{2} \int_{\Omega} \frac{1}{\varepsilon_0(\varepsilon_s - \varepsilon_{\infty})} |\hat{\mathbf{P}}|^2 dV - \frac{1}{2} \int_{\Omega} \frac{1}{\varepsilon_0(\varepsilon_s - \varepsilon_{\infty})\omega_0^2} |\partial_t \hat{\mathbf{P}}|^2 dV, \quad (2.43)$$

which can be recognized as the free field Lagrangian and the material Lagrangian of a material responding as an harmonic oscillator of frequency  $\omega_0$ , i.e. a Lorentz medium in absence of damping. (Compare with the Lagrangian of the harmonic oscillator  $\frac{1}{\omega_0^2} (\partial_t)^2 x + x = 0$ ).

The bilinear form is not only connected to the Lagrangian but also to reciprocity. To show this, assume that the test field  $(\hat{\mathbf{E}}^t \hat{\mathbf{H}}^t)$  in equation (2.22) satisfies Maxwell's equations and is causally related to the external sources  $(\hat{\mathbf{J}}^{\text{t:ext}} \hat{\mathbf{K}}^{\text{t:ext}})$ , which have a bounded



support in  $\Omega$ . Furthermore, at the boundary  $\partial\Omega$  of  $\Omega$  we assume that the tangential components of  $\mathbf{E}^t$  and  $\mathbf{E}$  vanish or, if an unbounded domain is of interest, we assume that in  $\Omega$ 's complement the medium is homogeneous. Now, substituting Maxwell's equations for the test field in equation (2.22) and taking either one of the above assumptions into account, we obtain the classical reciprocity relation

$$\int_{\Omega} -\hat{\mathbf{J}}^{t,\text{ext}} \cdot \hat{\mathbf{E}} + \hat{\mathbf{K}}^{t,\text{ext}} \cdot \mathbf{H} dV = \int_{\Omega} -\hat{\mathbf{E}}^t \cdot \hat{\mathbf{J}}^{\text{ext}} + \hat{\mathbf{H}}^t \cdot \hat{\mathbf{K}}^{\text{ext}} dV. \quad (2.44)$$

In inverse problems with multiple sources and receivers this reciprocity leads to invariance with respect to interchanging the role of sources and receivers.

### 2.4.1. RESONANCES

The field operator  $Q(s)$  can become singular leading to so-called scattering poles or resonances [49].

**DEFINITION 2.2.** — The resonance frequency and resonance mode of the operator  $Q(s)$  are given by the eigenpair  $(u_{\text{eig}}, s_{\text{eig}})$  that satisfies

$$Q(s_{\text{eig}})u_{\text{eig}} = 0. \quad (2.45)$$

In general, the eigenfrequency  $s_{\text{eig}}$  is complex, with an imaginary part that is related to the oscillation frequency of the mode and a real part that determines the temporal decay of the mode due to material losses or radiation losses as energy leaves the system through  $\partial\Omega$ .

## 2.5. SPATIAL DISCRETIZATION

In this section we show how to obtain a semi-discrete system from the continuous field equations. Furthermore, we introduce the discrete counterparts to the general wave-field properties discussed in the last section, to show that they are preserved after discretization. As an example we will derive the discrete system for the Maxwell equations with a second-order dispersion relation and then point out the peculiarities of the other cases considered. We start by considering instantaneously reacting materials and subsequently add dispersion.

We discretize the first-order field equations on a staggered Yee-grid [79] using primary and dual nodes in each Cartesian direction. For example, the primary and dual nodes in the  $y$ -direction are defined as

$$\Omega_y^{\text{p}} = \{y_q \in \mathbb{R}, q = 0, 1, \dots, N_y + 1, y_q > y_{q-1}\}, \quad (2.46)$$

and

$$\Omega_y^{\text{d}} = \{\hat{y}_q \in \mathbb{R}, q = 1, \dots, N_y + 1, \hat{y}_{q+1} > \hat{y}_q\}, \quad (2.47)$$

respectively, with corresponding step sizes given by

$$\delta_{y,q} = y_q - y_{q-1}, \quad q = 1, \dots, N_y + 1 \quad \text{and} \quad \hat{\delta}_{y,q} = \hat{y}_{q+1} - \hat{y}_q, \quad q = 1, \dots, N_y. \quad (2.48)$$

Dual step sizes carry a hat<sup>1</sup>, while primary step sizes do not. Grid nodes in the  $x$ - and  $z$ -direction are introduced in a similar manner with  $N_x$  dual step sizes in the  $x$ -direction and  $N_z$  dual step sizes in the  $z$ -direction.

On a Yee-grid, differentiation in each Cartesian direction can conveniently be described in terms of bidiagonal differentiation matrices. In particular, if we introduce the  $(N_y + 1)$ -by- $(N_y + 1)$  diagonal matrix of primary step sizes

$$W_y = \text{diag}(\delta_{y;1}, \delta_{y;2}, \dots, \delta_{y;N_y+1}) \quad (2.49)$$

and the  $N_y$ -by- $(N_y + 1)$  bidiagonal matrix  $\text{bidiag}_{N_y}(-1, 1)$  with  $-1$  on the diagonal and  $+1$  on the first upper diagonal, then differentiation of field quantities defined on primary grid nodes in the  $y$ -direction is carried out by the sparse differentiation matrix

$$Y = -W_y^{-1} \text{bidiag}_{N_y}(-1, 1)^T. \quad (2.50)$$

In a similar manner we can define a differentiation matrix that acts on field quantities defined on the dual nodes. Introducing the  $N_y$ -by- $N_y$  diagonal step size matrix

$$\hat{W}_y = \text{diag}(\hat{\delta}_{y;1}, \hat{\delta}_{y;2}, \dots, \hat{\delta}_{y;N_y}), \quad (2.51)$$

$$\hat{Y} = \hat{W}_y^{-1} \text{bidiag}_{N_y}(-1, 1) \quad (2.52)$$

computes two-point finite-differences of field quantities defined on dual nodes in the  $y$ -direction. Moreover, both differentiation matrices are related to each other via the obvious symmetry relation

$$\hat{Y}^T \hat{W}_y = -W_y Y. \quad (2.53)$$

Differentiation matrices  $X$ ,  $\hat{X}$ ,  $Z$ , and  $\hat{Z}$  in the  $x$ - and  $z$ -direction are defined in an analogous manner.

### 2.5.1. INSTANTANEOUSLY REACTING MEDIA

Discretizing the first-order Maxwell system on a standard Yee-grid and arranging the unknowns in lexicographical order, we arrive at the state-space representation

$$(D + S + M\partial_t)u = b'. \quad (2.54)$$

The order of this system is denoted by  $N$  and it is typically very large for real-world 3D problems (millions or even a billion of unknowns is not uncommon).

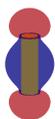
In the above representation, the sparse, spatial differentiation matrix is given by

$$D = \begin{bmatrix} 0 & D_h \\ D_e & 0 \end{bmatrix}, \quad (2.55)$$

with

$$D_h = \begin{bmatrix} 0 & \hat{Z} \otimes I_{N_y} \otimes I_{N_x+1} & -I_{N_z} \otimes \hat{Y} \otimes I_{N_x+1} \\ -\hat{Z} \otimes I_{N_y+1} \otimes I_{N_x} & 0 & I_{N_z} \otimes I_{N_y+1} \otimes \hat{X} \\ I_{N_z+1} \otimes \hat{Y} \otimes I_{N_x} & -I_{N_z+1} \otimes I_{N_y} \otimes \hat{X} & 0 \end{bmatrix} \quad (2.56)$$

<sup>1</sup>Not to be confused with Laplace domain quantities that also carry a hat.



and

$$D_e = \begin{bmatrix} 0 & -Z \otimes I_{N_y+1} \otimes I_{N_x} & I_{N_z+1} \otimes Y \otimes I_{N_x} \\ Z \otimes I_{N_y} \otimes I_{N_x+1} & 0 & -I_{N_z+1} \otimes I_{N_y} \otimes X \\ -I_{N_z} \otimes Y \otimes I_{N_x+1} & I_{N_z} \otimes I_{N_y+1} \otimes X & 0 \end{bmatrix}, \quad (2.57)$$

and  $\otimes$  is the Kronecker (tensor) product. Furthermore, the medium matrix  $S$  is given by

$$S = \begin{bmatrix} M_\sigma & 0 \\ 0 & 0 \end{bmatrix}, \quad (2.58)$$

where  $M_\sigma$  is a diagonal semi-positive definite matrix with (averaged) conductivity values on its diagonal. The medium matrix  $M$  is given by

$$M = \begin{bmatrix} M_\varepsilon & 0 \\ 0 & M_\mu \end{bmatrix}, \quad (2.59)$$

and both  $M_\varepsilon$  and  $M_\mu$  are diagonal and positive definite medium matrices with averaged permittivity and permeability values on their diagonal (in the case of isotropic media). The field vector is of the form

$$u = [e_x^T, e_y^T, e_z^T, h_x^T, h_y^T, h_z^T]^T, \quad (2.60)$$

where all field quantities are stored in lexicographical order in the corresponding field vectors  $e_i$  and  $h_i$ ,  $i = x, y, z$ . Finally, the finite-difference approximations of the external sources are stored in the source vector

$$b' = -[j_x^{\text{ext};T}, j_y^{\text{ext};T}, j_z^{\text{ext};T}, k_x^{\text{ext};T}, k_y^{\text{ext};T}, k_z^{\text{ext};T}]^T. \quad (2.61)$$

Premultiplying equation (2.54) by the inverse of the medium matrix  $M$ , we arrive at

$$(A + I\partial_t)u = M^{-1}b', \quad (2.62)$$

where we have introduced the system matrix as

$$A = M^{-1}(D + S). \quad (2.63)$$

This is the sparse system matrix for instantaneously reacting materials.

### SYMMETRY RELATIONS

To discuss the symmetry properties satisfied by the system matrix, we first introduce the diagonal step size matrices

$$W_e = \begin{bmatrix} \hat{W}_z \otimes \hat{W}_y \otimes W_x & 0 & 0 \\ 0 & \hat{W}_z \otimes W_y \otimes \hat{W}_x & 0 \\ 0 & 0 & W_z \otimes \hat{W}_y \otimes \hat{W}_x \end{bmatrix}, \quad (2.64)$$

and

$$W_h = \begin{bmatrix} W_z \otimes W_y \otimes \hat{W}_x & 0 & 0 \\ 0 & W_z \otimes \hat{W}_y \otimes W_x & 0 \\ 0 & 0 & \hat{W}_z \otimes W_y \otimes W_x \end{bmatrix}, \quad (2.65)$$

containing the voxel volumes of the Yee-cells on the diagonal. Using the symmetry relation of equation (2.53) (and the corresponding relations in the  $x$ - and  $z$ -directions), it is now easily verified that

$$D_h^T W_e = -W_h D_e \quad \text{and} \quad D_e^T W_h = -W_e D_h. \quad (2.66)$$

Furthermore, with

$$\tilde{W}_{EM} = \begin{pmatrix} W_e & 0 \\ 0 & -W_h \end{pmatrix} \quad (2.67)$$

we also have

$$D^T \tilde{W}_{EM} = \tilde{W}_{EM} D, \quad (2.68)$$

which leads to the symmetry property

$$A^T W_{EM} = W_{EM} A \quad \text{with} \quad W_{EM} = M \tilde{W}_{EM} = \tilde{W}_{EM} M = W_{EM}^T. \quad (2.69)$$

This symmetry property is the discrete analog of the previously discussed symmetry in an continuous setting. Since  $\tilde{W}$  contains the voxel volumes on the diagonal, a  $\tilde{W}$ -weighted inner-product is the finite-difference counterpart to spatial integration in the continuous domain. Analogues to the continuous case, the  $\frac{1}{2} u^H W u$  bilinear form with Hermitian transpose approximates the field Lagrangian. Finally, the symmetrized operator  $Q$  can be found as

$$(WA + W\partial_t)u = Q(\partial_t)u = WM^{-1}b' = b. \quad (2.70)$$

**REMARK 1.** — *Lossless media* – For lossless media the operator can be defined as  $A_{LL} = M^{-1}D$ . For this operator an anti-symmetric form can be found with weight matrix

$$W_{LL} = \begin{pmatrix} W_e & 0 \\ 0 & W_h \end{pmatrix} M \quad (2.71)$$

leading to the anti-symmetry relation

$$A_{LL}^T W_{LL} = -W_{LL} A_{LL}. \quad (2.72)$$

The  $W_{LL}$  weighted inner product  $\frac{1}{2} u^H W_{LL} u$  is the finite-difference approximation of the total energy of the field.

### 2.5.2. DISPERSIVE MEDIA

Loosely speaking, the main difference in setting up the semi-discrete Maxwell system for dispersive media is the presence of the polarization vectors  $\mathbf{P}$  and  $\mathbf{U} = -\partial_t \mathbf{P}$  in the field equations. These vectors are only active at points where a dispersive material is present. From a storage point of view, it is therefore advantageous to only keep the finite-difference approximations of  $\mathbf{P}$  and  $\mathbf{U}$  at these points in memory. Since the polarization is related to the electric field strength and electric field strength approximations are defined over the total computational domain, we need to introduce support<sup>2</sup>

<sup>2</sup>Sometimes called restriction matrix as  $I^{\text{sup}}u$  restricts  $u$  to the support domain.



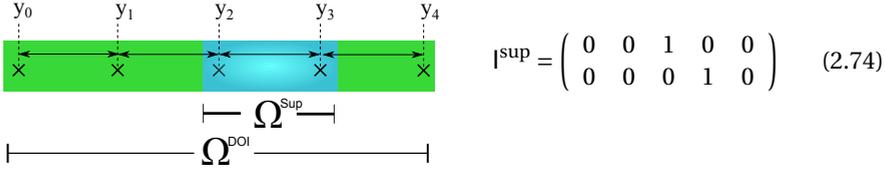


Figure 2.1: Construction of a support matrix. Grid points  $y_2$  and  $y_3$  are within the dispersive medium and thus  $y_2, y_3 \in \Omega^{\text{sup}}$ . The rows of  $I^{\text{sup}}$  are the basis vectors of  $\Omega^{\text{sup}}$  expressed in the basis vectors of  $\Omega^{\text{DOI}}$ .

matrices to implement the local dispersion relations. To this end, we define selection or logical projection matrices, which select the relevant electric field strength components from the total electric field vector. For example, if  $I_y^{\text{sup}}$  is the support matrix of a dispersive material in the  $y$ -direction and  $e_y$  contains all finite-difference approximations of the  $y$ -component of the electric field strength, then the vector  $I_y^{\text{sup}} e_y$  contains only those  $y$ -components of  $\mathbf{E}$  located within the dispersive material. An illustration of how the support matrix is constructed is shown in Figure 2.1.

Using this definition of the support matrices, the constitutive relation of equation (2.24) relating the electric and polarization fields to each other can be implemented in a straightforward manner. For example, for the  $y$ -component of equation (2.24) we have

$$B_{3,y} \partial_t u_y + B_{2,y} u_y - B_{1,y} p_y + B_{0,y} I_y^{\text{sup}} e_y = 0, \quad (2.73)$$

where the matrices  $B_{\{0,1,2,3\},y}$  are diagonal matrices with the (averaged) medium parameters  $\beta_{\{0,1,2,3\}}$  on their diagonal.

Using the Yee-grid introduced earlier, approximating the partial derivatives by two-point finite-difference formulas, and arranging the unknowns in lexicographical order, we now again arrive at the state-space representation

$$(D + S + M \partial_t) u = b'. \quad (2.75)$$

In this equation, the spatial differentiation matrix is given by

$$D = \begin{bmatrix} 0 & 0 & 0 & D_h \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ D_e & 0 & 0 & 0 \end{bmatrix}, \quad (2.76)$$

where  $D_h$  and  $D_e$  are given by equation (2.56) and (2.57), respectively. Furthermore, matrix  $S$  is given by

$$S = \begin{bmatrix} 0 & 0 & -I^{\text{sup};T} & 0 \\ 0 & 0 & I & 0 \\ B_0 I^{\text{sup}} & -B_1 & B_2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad (2.77)$$

where  $B_{\{0,1,2\}}$  are diagonal matrices only defined on the support of the dispersive media.

In addition,  $l^{\text{sup}}$  is the total support matrix and the medium matrix  $M$  is given by

$$M = \begin{bmatrix} M_\varepsilon & 0 & 0 & 0 \\ 0 & I & 0 & 0 \\ 0 & 0 & B_3 & 0 \\ 0 & 0 & 0 & M_\mu \end{bmatrix}, \quad (2.78)$$

where  $B_3$  is again a dispersion matrix and both  $M_\varepsilon$  and  $M_\mu$  are diagonal and positive definite medium matrices with averaged permittivity and permeability values on their diagonal. The field vector is now of the form

$$u = [e_x^T, e_y^T, e_z^T, p_x^T, p_y^T, p_z^T, u_x^T, u_y^T, u_z^T, h_x^T, h_y^T, h_z^T]^T, \quad (2.79)$$

where all field quantities are stored in lexicographical order in the corresponding field vectors  $e_i$ ,  $p_i$ ,  $u_i$ , and  $h_i$ ,  $i = x, y, z$ . Finally, the finite-difference approximations of the external sources are stored in the source vector

$$b' = -[j_x^{\text{ext};T}, j_y^{\text{ext};T}, j_z^{\text{ext};T}, 0, 0, 0, 0, 0, 0, k_x^{\text{ext};T}, k_y^{\text{ext};T}, k_z^{\text{ext};T}]^T. \quad (2.80)$$

Again, a symmetrized system can be obtained.

### SYMMETRY RELATIONS

The system matrix for media exhibiting relaxation is given by

$$A_{\text{DP}} = M^{-1}(D + S). \quad (2.81)$$

To discuss its symmetry properties, we introduce the matrix

$$\tilde{W}_{\text{DP}} = \begin{bmatrix} W_e & 0 & 0 & 0 \\ 0 & W_p & 0 & 0 \\ 0 & 0 & -W_u & 0 \\ 0 & 0 & 0 & -W_h \end{bmatrix} \quad (2.82)$$

and  $W_e$  and  $W_h$  as defined in equation (2.64) and (2.65). Furthermore,  $W_u$  and  $W_p$  are given by

$$W_u = B_0^{-1} l^{\text{sup}} W_e l^{\text{sup};T} \quad \text{and} \quad W_p = B_1 W_u = B_1 B_0^{-1} l^{\text{sup}} W_e l^{\text{sup};T}. \quad (2.83)$$

**PROPOSITION 2.1.** — The matrix  $A_{\text{DP}}$  is symmetric in the  $W_{\text{DP}}$  bilinear form as

$$A_{\text{DP}}^T W_{\text{DP}} = W_{\text{DP}} A_{\text{DP}} \quad \text{with} \quad W_{\text{DP}} = M \tilde{W}_{\text{DP}} = \tilde{W}_{\text{DP}} M = W_{\text{DP}}^T. \quad (2.84)$$

*Proof.* Given the above  $W_{\text{DP}}$  the symmetry property can be trivially verified via substitution.  $\square$

This symmetry property is similar in form to the symmetry relation for instantaneously reacting media. In our model order reduction method this symmetry plays a vital role. However, no such symmetry property can be found for general dispersion relations with an order higher than two; however, the proposition can be extended to



dispersion relations formed by sums of second-order dispersion relations. Finally, we mention that our approach relies on an analytic model for the dielectric constant. It can also handle measured dielectric data if the experimental data can be fitted by an arbitrary sum of multiple Drude and Lorentz dispersion curves. In that case the proposed method can still be used as the resulting system can be rewritten in a symmetric first-order form. Introducing auxiliary variables for every single medium again leads to a frequency-independent system. Thus, the final symmetrized discrete form of the dispersive Maxwell equations is

$$\mathbf{Q}(\partial_t) = \mathbf{W}_{\text{DP}}(\mathbf{A}_{\text{DP}} + \partial_t \mathbf{l})\mathbf{u} = \mathbf{b}, \quad (2.85)$$

with  $\mathbf{b} = \mathbf{W}_{\text{DP}}\mathbf{M}^{-1}\mathbf{b}'$ .

## 2.6. INTRODUCTION OF A PML

We confined the study of our wave equations to the domain  $\Omega$  and we assumed that all sources are supported inside this domain. Therefore, we impose a radiation condition at the boundary  $\partial\Omega$  which we denoted by  $D(s)$  – the Dirichlet to Neumann (DtN) map. During the discretization process in the last section we assumed Dirichlet boundary conditions which we now replace by radiation boundary conditions in this section to model an open domain.

The perfectly matched layers (PMLs) were introduced by Berenger in 1994 [8] for the Maxwell equations as an absorbing boundary condition, simulating the extension to an infinite space. The PML essentially reduces to complex-coordinate stretching [15], an approach introduced in the 1970s to study open quantum systems. In the discrete domain we introduce a few layers of complex frequency-dependent step sizes in the spatial coordinate normal to the boundary  $\partial\Omega$  of the form

$$h_i^c(s) = h_i + \frac{\hat{h}_i}{s}. \quad (2.86)$$

The real part of  $h^c$  absorbs evanescent waves, whereas the imaginary part absorbs propagating waves. Optimal step sizes  $h_i, \hat{h}_i$  can be found by rewriting the Dirichlet to Neumann map of a finite-difference grid as a continued fraction with the step sizes  $h_i, \hat{h}_i$  as coefficients. The coefficients of such a rational function that minimize the difference between the continuous and discrete DtN map can be found in closed form via Zolotarev rational approximation (see [20] for details).

This complex coordinate stretching introduces a domain  $\Omega_P$  surrounding the actual domain of interest. This can be reformulated to obtain an expression for the DtN that is only active on  $\partial\Omega$ . To this end, let  $\Omega_D$  be the interior of our domain of interest  $\Omega$  and let  $\partial\Omega$  denote its closed boundary surface. The configuration of these non-overlapping domains is illustrated in Figure 2.2 clearly showing that  $\Omega_D$  is simply connected,  $\partial\Omega$  and  $\Omega_P$  are not. Furthermore,  $\Omega_D$  and  $\Omega_P$  are  $k$ -dimensional domains, whereas  $\partial\Omega$  is of dimension  $k - 1$ .

Let the restriction of a vector or a matrix to a certain domain be denoted by subscripts, such that  $\mathbf{W}_D$  is the restriction of the matrix  $\mathbf{W}$  to  $\Omega_D$ ,  $\mathbf{W}_\partial$  to  $\partial\Omega$  and so on. These

restrictions can be defined via the previously introduced support matrices illustrated in Figure 2.1 as  $\mathbb{I}_D^{\text{sup}} u = u_D$  or  $W_D = \mathbb{I}_D^{\text{sup}} W \mathbb{I}_D^{\text{sup};T}$  for matrices.

Furthermore, let  $\hat{A}(s) - s^2 \mathbb{1}$  be the discrete operator on the whole domain including the complex, frequency-dependent step sizes inside the PML in a second-order form (The first-order form is entirely analogous). The Schur complement method is now used to permute the matrix  $\hat{A}(s) - s^2 \mathbb{1}$  into

$$\left( \begin{bmatrix} A_D & 0 & A_{D\partial} \\ 0 & A_P(s) & A_{P\partial}(s) \\ A_{\partial D} & A_{\partial P} & A_{\partial} \end{bmatrix} - s^2 \mathbb{1} \right) \begin{bmatrix} u_D \\ u_P \\ u_{\partial} \end{bmatrix} = \begin{bmatrix} W_D^{-1} b_D \\ 0 \\ 0 \end{bmatrix}, \quad (2.87)$$

with  $A_P(s)$  and  $A_{P\partial}(s)$  the only complex valued and frequency-dependent matrices. All matrices with double subscripts map values from the domain denoted by the second subscript into the domain denoted by the first subscript; e.g.,  $A_{D\partial}$  maps from  $\partial\Omega$  to  $\Omega_D$ . We can now rewrite this equation over the domain  $\Omega = \partial\Omega \cup \Omega_D$  eliminating the values of  $u_P$ , since we are only interested in the boundary contribution of  $u_P$  on  $u_{\partial}$ . We rewrite the equation described by the second row of equation (2.87) as  $u_P = (A_P(s) - s^2 \mathbb{1})^{-1} A_{P\partial}(s) u_{\partial}$  to arrive at

$$\left( \begin{bmatrix} A_D & A_{D\partial} \\ A_{\partial D} & A_{\partial} \end{bmatrix} - \begin{bmatrix} 0 & 0 \\ 0 & A_{\partial P}(W_P(s)A_P(s) - s^2 W_P(s))^{-1} W_P(s)A_{P\partial}(s) \end{bmatrix} - s^2 \mathbb{1} \right) \begin{bmatrix} u_D \\ u_{\partial} \end{bmatrix} = \begin{bmatrix} W_D^{-1} b_D \\ 0 \end{bmatrix}. \quad (2.88)$$

This exclusion of the PML eliminates certain eigenvalues whose eigenvectors are solely supported inside the PML. The DtN map  $A_{\partial P}(W_P(s)A_P(s) - s^2 W_P(s))^{-1} W_P(s)A_{P\partial}(s)$  is symmetric in the transpose bilinear form with  $W_{\partial}$  and  $(W_P(s)A_P(s) - s^2 W_P(s))^{-1}$  is symmetric as well. Thus, introduction of the PML transforms our shifted system into a non-linear system as the DtN map is a higher order rational function. However, the overall system is still symmetric in the frequency-independent bilinear form  $W$  such that the reciprocity results from the continuous case directly translate to the discretized domain. The Schwartz reflection principle still holds as the DtN map and the introduced complex step sizes follow the reflection principle. Moreover, stability and passivity is preserved as the numerical range of the resulting matrix still satisfies  $\Re \mathcal{N}\{Q(s)\} \leq 0$ . Further, the Hermitian bilinear form  $\frac{1}{2} u^H W u$  approximates the field Lagrangian, as the integral is replaced by summation with  $W$  containing the voxel volumes and medium parameters as grid weights.

**REMARK 2.** — *Diffusion* – In the case of diffusion no PML is needed as diffusive fields decay exponentially in space. Therefore, it suffices to terminate the domain with a Dirichlet boundary at a distance of a few diffusion length of the field. Further, to discretize anisotropic media a Lebedev-grid is used rather than a Yee-grid.

**REMARK 3.** — *Incorporation of source terms* – Until now little attention was paid to the fact that sources may have a temporal dependency, as we assumed a delta function as temporal dependence so far. For external sources described by an arbitrary time variation  $w(t)$ , we have

$$\mathbf{J}^{\text{ext}} = w(t) \mathbf{J}^{\text{sp}} \text{ or } \mathbf{K}^{\text{ext}} = w(t) \mathbf{K}^{\text{sp}}, \quad (2.89)$$



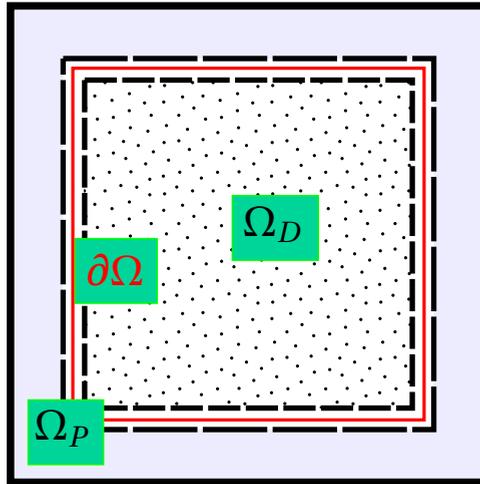


Figure 2.2: Used domain decomposition with the interior  $\Omega_D$  (dotted), the boundary  $\partial\Omega$  (red), and the exterior PML domain  $\Omega_P$  (blue) .

where the quantities with superscript “sp” are time independent, spatial densities and  $w(t)$  is the so-called wavelet that vanishes prior to the time instant  $t = 0$ .

In this thesis, our approach is to solve time- and frequency-domain wave problems using the Dirac distribution  $\delta(t)$  for excitation. The resulting Greens function responses are then used to obtain the solution due to a wavelet  $w(t)$  by convolution in the time-domain or simple multiplication in the frequency-domain.

**REMARK 4.** — *Integral and differential form* – We write our wave equation in two forms in this thesis, with a symmetric operator  $Q(s)$  defining an integral form and with an operator  $\mathcal{A}(s)$  provides us with a relation to shifted systems. For clarification we state both forms here. The shifted system and its discrete counterpart follows the equation

$$[\mathcal{A}(s) + s\mathcal{I}] u = \mathcal{W}^{-1} b \text{ and } [\hat{\mathcal{A}}(s) + s\mathbf{1}] \hat{u} = W(s)^{-1} b \quad (2.90)$$

and the symmetrized integral form follows as

$$Q(s)u = b \text{ and } Q(s)u = W(s) [\hat{\mathcal{A}}(s) + s\mathbf{1}] \hat{u} = b. \quad (2.91)$$

Since right-hand sides are supported inside  $\Omega$  (outside of the PML) only, they are frequency-independent. In the discrete case  $W$  has a dual purpose, making the equation symmetric and simultaneously approximating an integral by containing the voxel sizes of the Yee-cells on the diagonal.

## 2.7. SOLUTION OF THE SYSTEM

With the system formulation and the PML in place, the solution to the system can be introduced. Two approaches are considered: First we deal with the nonlinear problem and introduce the solution, next we linearize the nonlinearity introduced by the DtN map to

arrive at an unstable but shifted system. A stable solution to the introduced equations can nonetheless be derived via a stability correction procedure.

### 2.7.1. NONLINEAR PROBLEM

The nonlinear problem can formally be solved in the Laplace domain as

$$u = Q^{-1}(s)b \quad (2.92)$$

and quadrature rules for evaluating the Bromwich integral in the inverse Laplace transform can be used to obtain a time-domain solution (see [84]).

In the case of imaging or remote sensing applications we are typically not interested in the wavefield in the entirety of  $\Omega$  but at certain receiver locations only. With  $r$  denoting the receiver function, the transfer-function from sources to receivers can be defined as

$$f(r, b, s) = r^T Q^{-1}(s)b. \quad (2.93)$$

Since  $Q(s)$  is symmetric in the transpose bilinear form we find that  $f(r, b, s) = f(b, r, s)$ . Usually multiple sources and receivers are employed in an imaging problem so that next to a single-input single-output (SISO) transfer-function also a multiple-input multiple-output (MIMO) transfer-function can be defined. We introduce the MIMO problem via the wave equation in the time-domain governed by

$$\Delta u^{[l]} - \frac{1}{v^2} u_{tt}^{[l]} = -\frac{1}{v^2} \delta(t) \delta(x - x_S^l), \quad u|_{t=0} = 0, u_t|_{t=0} = 0, \quad (2.94)$$

on  $\mathbb{R}^k \times [0, \infty[$ , where the superscript  $l$  is the source index with  $l = 1, \dots, N_{\text{src}}$ . The weak formulation of the corresponding  $s$ -domain equations is (cf. equation (2.7))

$$p^H Q(s) u^{[l]} = -\bar{p}(x_S^l) \frac{1}{v(x_S^l)^2}, \quad \text{for } l = 1, 2, \dots, N_{\text{src}}. \quad (2.95)$$

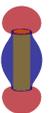
Assuming possibly coinciding source-receiver pairs, we collect all sources in the array  $B_s = [b^{[1]}, \dots, b^{[N_{\text{src}}]}]$ , with individual source contributions  $b^{[l]} = -\delta(x - x_S^l)/v^2(x_S^l)$  as columns and all receivers in the array  $B_r = [r^{[1]}, \dots, r^{[N_{\text{rec}}]}]$  with receiver contributions  $r^{[r]} = \delta(x - x_R^r)$ . Equivalently, we define  $U_s(s) = [u^{[1]}(s), \dots, u^{[N_{\text{src}}]}(s)]$  the array containing the fields to write the MIMO equation of (2.95) as

$$P^H Q(s) U_s = P^H B_s. \quad (2.96)$$

Finally, we define the MIMO transfer-function of size  $N_{\text{rec}} \times N_{\text{src}}$  as

$$F(s) = B_r^H U_s(s). \quad (2.97)$$

If sources and receivers coincide, the transfer-function  $F(s)$  is symmetric due to the symmetry of  $Q(s)$ .



### 2.7.2. PML LINEARIZATION - STABILITY CORRECTION

Computing  $Q^{-1}(s)b$  for multiple frequencies can be cumbersome or computationally not feasible for large 3D configurations with millions of unknowns. For certain model order reduction applications it is beneficial to linearize the PML, and to make use of the vast theory that was developed for shifted systems. Specifically, a fixed-frequency PML can be designed to be valid in a spectral range of interest to obtain a shifted operator on the domain of interest and in the domain where complex coordinate stretching has been applied ( $\Omega \cup \Omega_P$ )

$$\mathcal{W}(\mathcal{A}^{\text{fi}} + \partial_t \mathcal{I})u^{\text{fi}} = b, \quad (2.98)$$

where the superscript “fi” indicates that the PML frequency is fixed. We note that  $\mathcal{A}^{\text{fi}}$  is a complex operator due to the application of the complex-scaling method and the above system is therefore complex and unstable. Instability occurs since a fixed-frequency PML cannot handle the sign-change in the radiation condition that needs to occur when the imaginary part of the Laplace parameter  $s$  switches sign. Therefore, fixing the PML frequency approximates outgoing waves on one side of the imaginary axis and incoming waves on the other side of the imaginary axis. This violates stability and the Schwartz reflection principle.

However, real-valued and stable time-domain field responses can be obtained from the above system via a stability-correction procedure proposed in [28]. In particular, stable field approximations can be computed as (for details see [28])

$$\mathcal{U}^{\text{fi}}(t) = w(t) * 2\eta(t)\Re\left[\eta(\mathcal{A}^{\text{fi}})\exp(-\mathcal{A}^{\text{fi}}t)\mathcal{W}^{-1}b\right] \quad \text{for } t > 0, \quad (2.99)$$

where

$$\eta(z) = \begin{cases} 1 & \text{if } \Re(z) > 0, \\ \frac{1}{2} & \text{if } \Re(z) = 0, \\ 0 & \text{if } \Re(z) < 0 \end{cases}$$

is the complex Heaviside unit step function. Furthermore, by applying a one-sided Laplace transform to equation (2.103) we obtain the frequency-domain solution

$$\mathcal{U}^{\text{fi}}(s) = w(s) \left[ r(\mathcal{A}^{\text{fi}}, s) + r(\bar{\mathcal{A}}^{\text{fi}}, s) \right] \mathcal{W}^{-1}b, \quad (2.100)$$

where the overbar denotes complex conjugation and where we have introduced the filtered resolvent function

$$r(z, s) = \frac{\eta(z)}{z + s}. \quad (2.101)$$

In both equations the wavelet  $w(t)$  has been applied to the solution obtained from a delta source  $\delta(t)$ . The step function can be seen as a projection of the operator on the stable side of the imaginary axis and conjugation restores the Schwartz reflection principle.

The finite-difference solution can be obtained by substituting the finite-difference operators and field vectors into the expressions rather than the continuous expressions. The discrete analog of the solution given in equation (2.100) is

$$\hat{u}^{\text{fi}} = w(s) \left[ r(\mathcal{A}^{\text{fi}}, s) + r(\bar{\mathcal{A}}^{\text{fi}}, s) \right] \mathcal{W}^{-1}b, \quad (2.102)$$

and the discrete analog of equation (2.99) is given by

$$u^{\text{fi}}(t) = w(t) * 2\eta(t)\mathfrak{R}\epsilon \left[ \eta(A^{\text{fi}}) \exp(-A^{\text{fi}}t) W^{-1}b \right] \quad \text{for } t > 0. \quad (2.103)$$

Although equations (2.103) and (2.102) provide a explicit expression for the time- and frequency-domain electromagnetic field on our domain of interest, these expressions cannot be evaluated directly, since the order  $N$  of the discretized system matrix  $A$  is typically too large. For example,  $N$  may easily run into the millions for three-dimensional problems and direct evaluation of equation (2.103) or equation (2.102) is simply not feasible. However, equation (2.103) and (2.102) do serve as a starting point for the model order reduction method discussed in the next chapter.

**REMARK 5.** — *Modification of source vector* – The field solution in equation (2.102) can also be written as

$$\hat{u}^{\text{fi}} = \left[ r(A^{\text{fi}}, s)w(-A^{\text{fi}}) + r(\bar{A}^{\text{fi}}, s)w(-\bar{A}^{\text{fi}}) \right] W^{-1}b, \quad (2.104)$$

which becomes evident when evaluating the inverse Laplace transform via the residue theorem. The PML needs to be matched over the spectrum of the wavelet  $w(-A^{\text{fi}})$  for all angles of incidence. In addition, it can be seen that the parts of the source that actually excite stable waves are  $w(-A^{\text{fi}})\eta(A^{\text{fi}})W^{-1}b$  and  $w(-\bar{A}^{\text{fi}})\eta(\bar{A}^{\text{fi}})W^{-1}b$ .

**REMARK 6.** — *Second-order wave equation* – The second-order wave equation from equation (2.2) can be interpreted as a shifted system after linearization of the PML with shift  $s^2$  rather than  $-s$ . The solutions obtained for this second-order system will be in terms of  $\sqrt{A^{\text{fi}}}$  rather than  $A^{\text{fi}}$ . The stability correction is then incorporated by selecting the proper branch of the square root (see [28]).

### RESONANCES - FIXED-FREQUENCY PML

In the discrete setting, the notion of resonances is defined via the associated eigenvalue problem. In the continuous setting and in an infinite space, resonant modes are quasi-normal, e.g. normal in some weighted inner product space or bilinear form in which the associated differential operators are symmetric.

In the discrete setting with a linearized PML, the modes are also quasi-normal in the  $W$ -weighted bilinear form. Assume  $A^{\text{fi}}$  is the fixed-frequency PML, first-order operator. For two eigenpairs  $(y, \lambda)$  and  $(x, \mu)$  of this matrix, we find that

$$A^{\text{fi}}y = \lambda y \quad \text{and} \quad A^{\text{fi}}x = \mu x, \quad (2.105)$$

holds. Since the matrix  $A^{\text{fi}}$  is symmetric in the  $W$  bilinear form, the two following inner product are equal to one another as

$$x^T W A^{\text{fi}} y = \lambda x^T W y \quad \text{and} \quad y^T W A^{\text{fi}} x = \mu y^T W x. \quad (2.106)$$

By subtracting the second expression from the first we obtain

$$0 = (\lambda - \mu) y^T W x, \quad (2.107)$$



and for  $\lambda \neq \mu$  it follows that the eigenvectors are quasi-normal, that is, they satisfy  $y^T W x = 0$ .

For lossless media we can even find a stronger result since  $W_{LL} A_{LL}^{\text{fi}} = -A_{LL}^{\text{fi}} W_{LL}$  (see equation (2.72)). We find that

$$0 = (\lambda + \mu) y^T W_{LL} x. \quad (2.108)$$

Thus all modes have to be quasi-normal, except the corresponding unstable incoming and outgoing stable mode ( $\lambda = -\mu$ ). In the lossless electromagnetic case the unstable mode can be obtained by either flipping the sign of the H-field or E-field, thereby reversing the direction of Poynting's vector and thus the propagation direction. Specifically, let the lossless system for a stable eigenpair  $(e_1, h_1, s_1)$  with  $\Re s_1 > 0$  satisfy

$$\left( \begin{bmatrix} 0 & M_\varepsilon^{-1} D_h^{\text{fi}} \\ M_\varepsilon^{-1} D_e^{\text{fi}} & 0 \end{bmatrix} + s_1 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right) \begin{pmatrix} e_1 \\ h_1 \end{pmatrix} = 0. \quad (2.109)$$

Now, the unstable mode can be found by left-hand multiplication of  $\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$  and parti-

tion of unity in front of the field vectors as  $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ . We find

$$\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \left( \begin{bmatrix} 0 & M_\varepsilon^{-1} D_h^{\text{fi}} \\ M_\varepsilon^{-1} D_e^{\text{fi}} & 0 \end{bmatrix} + s_1 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right) \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{pmatrix} e_1 \\ h_1 \end{pmatrix} = 0, \quad (2.110)$$

which simplifies to

$$\left( \begin{bmatrix} 0 & M_\varepsilon^{-1} D_h^{\text{fi}} \\ M_\varepsilon^{-1} D_e^{\text{fi}} & 0 \end{bmatrix} - s_1 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right) \begin{pmatrix} e_1 \\ -h_1 \end{pmatrix} = 0. \quad (2.111)$$

Thus, the mode  $(e_1, -h_1, -s_1)$  satisfies the unstable field relations. If  $(e_1, h_1, s_1)$  is an outgoing mode, then  $(e_1, -h_1, -s_1)$  is an incoming mode since Poynting's vector flips sign as well. This analysis does not hold in case losses are involved.

## 2.8. CONCLUSIONS

In this chapter, the basic wave and diffusion equations considered in this thesis were discussed. Subsequently, the equations were discretized on a grid such that they can be solved numerically. The basic properties of wavefields were reviewed and it was shown that the discrete counterparts of these properties are preserved throughout discretization. All considered wave equations were written in a unified form with the symmetric operator  $Q(s)$  in the analytical domain and a sparse, symmetric matrix  $Q(s)$  in the discrete domain.

The domain of interest was surrounded by a PML in order to simulate open domains, which introduced a nonlinearity with respect to the Laplace frequency into the modeling approach. Finally, the solution to the wave equations was obtained for frequency-dependent and frequency-independent PMLs. The frequency-independent PML is used in the next chapter to leverage methods, developed for shifted systems and eigenvalue problems, for model order reduction. The formulation with the frequency-dependent PML is used to construct efficient reduced-order models and projection methods on subspaces that themselves depend on the Laplace parameter  $s$ .

# 3

3

## MODEL ORDER REDUCTION

*It is a curious historical fact that modern quantum mechanics began with two quite different mathematical formulations: the differential equation of Schroedinger and the matrix algebra of Heisenberg. The two apparently dissimilar approaches were proved to be mathematically equivalent.*

Richard P. Feynman



### 3.1. INTRODUCTION

**I**N the last chapter, the continuous wave equations and their discrete counterparts were introduced. In this chapter, we construct models that are considerably smaller than the straightforward discretizations introduced in the last chapter.

There are several reasons why one can reduce the model order of wave equations. First, finite-difference problems are typically over-discretized with respect to the Nyquist rate. For a second-order accurate finite-difference scheme in a lossless environment at least 10 point per wavelength should be used as the numerically introduced dispersion becomes too large otherwise. In addition, capturing complex sub-wavelength features or geometries typically leads to over-discretization; however, homogenization techniques can be used to avoid this. Second, for imaging applications we are typically only interested in the transfer-function for a few sources and receivers in a specific spectral range rather than the field in the whole simulated domain and for all frequencies. Therefore, a smaller model can be found that only contains the information relevant to the application. Furthermore, the transfer-function only needs to be approximated to the noise level of the measurements. Finally, losses present in the material or in the PML separate the spectrum of the operator from the imaginary axis, the area where we want to approximate the operator. Such a separation usually leads to the possibility of finding lower dimensional models. For example, low-order rational function approximations can usually be found if the domain where the function needs to be evaluated is well-separated from the domain containing the singularities of the function. This is also the reason why model reduction shows excellent results for diffusion problems as the spectrum of a diffusion operator is contained on the real semi-axis.

In this thesis, we consider projection-based model reduction where a large operator is projected onto a subspace of a smaller dimension. We consider rational Krylov subspaces (RKS), polynomial Krylov subspaces (PKS) and extended Krylov subspaces (EKS).

#### OUTLINE

First, projection-based reduced-order modeling is introduced in section 3.2 along with the reduced-order model solution. Subsequently, a rational Krylov subspace method is introduced to construct field approximations in the time- and frequency-domain. We show that this RKS approach is structure-preserving and that the transfer-function of reduced-order models based on this RKS is a Hermite interpolant of the transfer-function for a coinciding source-receiver pair. In this section we also show that polynomial and extended Krylov subspaces can be used for projection of the linearized wave operator. The symmetry and sparsity of the system matrix are exploited to obtain efficient model reduction algorithms. The approximation of resonances as Lanczos-Ritz pairs of the reduced-order modeling operator is discussed as well. The computational complexity of the three Krylov methods is analyzed in section 3.3. In the literature Krylov methods have extensively been used for diffusion equations and a comparison between these methods for wave and diffusion operators is given in section 3.4. The presented tech-

niques are then used in three example applications in section 3.5. First, the spontaneous decay rate of a quantum emitter in proximity of a dispersive nano-resonator is computed and the resonances are analyzed. Subsequently, examples from geophysics are discussed – the lossy Maxwell system is solved in a ground penetrating radar application and the Maxwell diffusion approximation is used for another exploration example. Finally, a lossless dielectric box is considered to show that RKS also performs well in the absence of material losses. Conclusions are given in section 3.6.

## 3.2. PROJECTION-BASED MODEL ORDER REDUCTION

In the last chapter, we arrived at the generic equation

$$Q(s)u = b, \quad (3.1)$$

for all of our PDEs. As a first step in our model order reduction framework, we restrict  $u$  to an  $m$ -dimensional subspace  $\mathcal{K}_m$ . Let the functions  $v^{[1]}, v^{[2]}, \dots, v^{[m]}$  form a complex basis  $V_m \in \mathbb{R}^{\infty \times m}$  of  $\mathcal{K}_m$  and let  $\hat{u}_m$  be the reduced-order solution in this space with an error  $e_m$  such that  $u = u_m + e_m$ . The reduced-order solution is now expanded as  $u_m = V_m z$  with expansion coefficients  $\alpha_i$  collected in the vector  $z = [\alpha_1, \dots, \alpha_m]^T$  so that

$$Q(s)V_m z + Q(s)e_m = b. \quad (3.2)$$

The residual is defined as  $r_m = Q(s)e_m$ , which in the context of wave propagation can be interpreted as the equivalent source causing the error. The reduced-order solution in equation (3.2) is not uniquely defined as we need a condition what to regard as error and what to regard as a solution. The Galerkin condition is such a condition and sets the residual orthogonal to the expansion space  $V_m^H r_m = 0$  or the error  $Q(s)$ -orthogonal to the expansion space. Under the Galerkin condition, the elements of  $z$  are uniquely defined through

$$V_m^H [Q(s)V_m z - b] = 0 \quad (3.3)$$

for all  $\Re\{s\} > 0$  provided that  $V_m$  is a basis and the passivity constraint  $\Re\{\mathcal{N}\{Q(s)\}\} \leq 0$  is satisfied<sup>1</sup>. In that case an invertible (for  $\Re\{s\} > 0$ )  $m \times m$  reduced-order operator  $R_m(s)$  can be defined as

$$R_m(s) = V_m^H Q(s) V_m \quad (3.4)$$

and the  $m$  expansion coefficients of the reduced-order model are found as

$$z(s) = R_m(s)^{-1} V_m^H b \quad (3.5)$$

defining the reduced-order solution  $u_m(s) = V_m z(s)$ . The reduced-order  $R_m$  is the projection of the operator  $Q(s)$  onto  $\mathcal{K}_m$ . The reduced-order MIMO transfer-function immediately follows as

$$F_m(s) = B_r^H U_m = B_r^H V_m R_m^{-1}(s) V_m^H B_s. \quad (3.6)$$

**PROPOSITION 3.1.** — The reduced-order operator  $R_m$  is passive, i.e.

$$\Re\{\mathcal{N}\{R_m\}\} \leq 0. \quad (3.7)$$

<sup>1</sup>Since we study propagating waves we can even assume that  $\Re\{\mathcal{N}\{Q(s)\}\} < 0$  holds and obtain a reduced-order model that is valid for all  $\Re\{s\} \geq 0$



*Proof.* It suffices to show that the numerical range of  $R_m$  is contained in the numerical range of  $Q(s)$ , that is  $\mathcal{N}\{R_m(s)\} \subseteq \mathcal{N}\{Q(s)\}$ . From the definition of the numerical range of the reduced-order model we find that

$$x_m^H R_m(s) x_m = (V_{m \times m})^H Q(s) (V_{m \times m}). \quad (3.8)$$

Now, every  $s$  that satisfies  $x_m^H R_m(s) x_m = 0$  also satisfies  $(V_{m \times m})^H Q(s) (V_{m \times m}) = 0$  and every point in the numerical range of  $R_m(s)$  is also included in the numerical range of  $Q(s)$ . Since  $Q(s)$  is passive, it follows that  $R_m$  is passive as well.  $\square$

Unfortunately, the operator  $Q(s)$  is not self-adjoint in the Hermitian inner product but only in the transpose bilinear form. Therefore,  $R_m(s)^H$  is equivalent to the reduced-order model of the adjoint  $Q(\bar{s})$ . However, for real subspaces  $V_{m; \mathfrak{R}\epsilon}$  the reduced-order model is structure preserving. (Take for instance  $\mathfrak{R}\epsilon \mathcal{K}_m \cup \mathfrak{I}\mathfrak{m} \mathcal{K}_m$  as a subspace).

**PROPOSITION 3.2.** — For a real basis  $V_{m; \mathfrak{R}\epsilon}$  the reduced-order model preserves the structure of the  $Q(s)$ ; i.e.  $R_{m; \mathfrak{R}\epsilon}(s)$  is symmetric and self-adjoint in the transpose bilinear form and follows the Schwartz reflection principle.

*Proof.* The symmetry of  $R_{m; \mathfrak{R}\epsilon}(s)$  follows from the symmetry of  $Q(s)$  since  $R_{m; \mathfrak{R}\epsilon}(s) = V_{m; \mathfrak{R}\epsilon}^H Q(s) V_{m; \mathfrak{R}\epsilon} = V_{m; \mathfrak{R}\epsilon}^T Q(s)^T \bar{V}_{m; \mathfrak{R}\epsilon} = R_{m; \mathfrak{R}\epsilon}(s)^T$ , since  $V_{m; \mathfrak{R}\epsilon}$  is a real. The same argument shows that the Schwartz reflection principle holds as

$$R_{m; \mathfrak{R}\epsilon}(\bar{s}) = V_{m; \mathfrak{R}\epsilon}^H Q(\bar{s}) V_{m; \mathfrak{R}\epsilon} = V_{m; \mathfrak{R}\epsilon}^H \bar{Q}(s) V_{m; \mathfrak{R}\epsilon} = \bar{R}_{m; \mathfrak{R}\epsilon}(s). \quad (3.9)$$

$\square$

Due to this proposition the spectrum of  $R_{m; \mathfrak{R}\epsilon}$  is closed under conjugation and  $\hat{u}_{m; \mathfrak{R}\epsilon}$  follows the Schwartz reflection principle. Thus, time-domain reduced order model solutions obtained after applying the inverse Laplace transform are real-valued.

In the following two sections we discuss the choice of  $\mathcal{K}_m$  and the consequences for the reduced-order model. First, we introduce rational Krylov subspaces and then polynomial and extended Krylov subspaces. For the latter two we project in a bilinear form like  $V_m^T Q(s) V_m$  with a complex basis  $V_m$ . Uniqueness, passivity and even stability of such a reduced-order solution cannot be guaranteed as the numerical range of the operator is not preserved.

### 3.2.1. RATIONAL KRYLOV SUBSPACES

As a first step towards an efficient rational Krylov methodology for multi-frequency wave-field problems, we construct field approximations or reduced-order models based on an interpolatory rational Krylov subspace containing single frequency solutions (snapshots) of the problem as trial and testing space. Specifically, our approach is to define an RKS of order  $m$  as

$$\mathcal{K}^m(\kappa) = \text{span}\{u(s_1), u(s_2), \dots, u(s_m)\} \quad (3.10)$$

with  $m$  distinct shifts  $\kappa = [s_1, \dots, s_m]$  and to use its real form, the RKS

$$\mathcal{K}_R^{2m}(\kappa) = \text{span}\{\mathfrak{R}\epsilon \mathcal{K}^m(\kappa), \mathfrak{I}\mathfrak{m} \mathcal{K}^m(\kappa)\}, \quad (3.11)$$

as a test and trial space. The real and imaginary parts of the snapshots  $u(s_i)$  spanning  $\mathcal{K}_R^{2m}(\kappa)$  are always linearly independent, since the eigenfunction expansion of the Dirac distribution appearing on the right-hand side of equation (2.2) has an infinite number of terms. Furthermore, from the symmetry  $u(\bar{s}) = \overline{u(s)}$  it follows that  $\mathcal{K}^m(\kappa) \subset \mathcal{K}_R^{2m}(\kappa)$  and  $\mathcal{K}^m(\bar{\kappa}) \subset \mathcal{K}_R^{2m}(\kappa)$  and a projection onto the subspace  $\mathcal{K}_R^{2m}(\kappa)$  will therefore preserve the Schwarz reflection principle leading to real-valued, time-domain wavefield approximations. The space  $\mathcal{K}_R^{2m}(\kappa)$  is equivalent to a rational Krylov subspace with shifts  $\bar{\kappa} \cup \kappa$ . The projection is structure preserving and leads to passive reduced-order models as shown in Propositions 3.1 and 3.2.

### INTERPOLATION PROPERTIES

The standard theory of Galerkin projection-based model reduction of passive, dynamic systems yields the following interpolation properties (e.g., see [3]).

**PROPOSITION 3.3.** — The projected RKS solution  $u_m(s)$  interpolates at the shifts, i.e.,

$$u_m(s) = u(s) \quad \forall s \in \kappa \cup \bar{\kappa} \quad (3.12)$$

and the SISO reduced-order transfer-function  $f_m(s)$  for a coinciding source and receiver is a Hermite interpolant of the SISO transfer-function  $f(s)$  at the shifts, that is,

$$f_m(s) = f(s) \quad \text{and} \quad \frac{d}{ds} f_m(s) = \frac{d}{ds} f(s) \quad \text{with } s \in \kappa \cup \bar{\kappa}. \quad (3.13)$$

*Proof.* Since  $\mathcal{K}^m(\kappa) \subset \mathcal{K}_R^{2m}(\kappa)$  and  $\mathcal{K}^m(\bar{\kappa}) \subset \mathcal{K}_R^{2m}(\kappa)$ , property (3.12) follows directly from the uniqueness of the Galerkin condition for passive problems. To prove (3.13), we first recall that the field error and residual are given by

$$e_m(s) = u(s) - u_m(s) \quad \text{and} \quad r_m(s) = b - Q(s)u_m(s),$$

respectively. From the Galerkin condition we obtain the relation

$$u_m^H(\bar{s}) r_m(s) = 0, \quad (3.14)$$

since  $u_m(\bar{s}) \in \mathcal{K}_R^{2m}$ . The error of the transfer-function can now be written as

$$f(s) - f_m(s) = b^H e_m(s) = u^H(\bar{s}) Q(s) e_m(s), \quad (3.15)$$

where we have used Schwarz's reflection principle. Since  $Q(s)e_m(s) = r_m(s)$  and the Galerkin condition of equation (3.14) holds, we can write

$$f(s) - f_m(s) = u^H(\bar{s}) r_m(s) = e_m^H(\bar{s}) r_m(s), \quad (3.16)$$

which has double zeros at  $s = \kappa \cup \bar{\kappa}$ , since the error and residual vanishes for these frequencies due to relation (3.12).  $\square$

For coinciding sources and receivers the subspace  $V_m$  contains the solutions and the adjoint solutions since we use a real basis, which leads to double interpolation. For



transfer-functions with non-coinciding sources and receivers, double interpolation can be achieved by using a subspace that contains  $Q(\bar{\kappa})^{-1}b_r$  and  $Q(\kappa)^{-1}b_s$ .

The outlined approach is most efficient if only a few singular Hankel values of the system contribute to the solution as is the case for resonating structures with a few excited and observable modes [7]. Then the frequency-domain response is well-described by a low-degree rational function and a rational Krylov technique will therefore quickly capture the desired wavefield response.

In the finite-difference framework, assume projection of the finite-difference operator from equation (2.88) with an excluded PML. Further assume the basis  $V_m$  to be  $W$  orthogonal as  $V_m^H W V_m = I_m$ . The reduced-order operator (for a second-order system) is then given by

$$R_m(s) = V_m^H W \left( \begin{bmatrix} A_D & A_{D\partial} \\ A_{\partial D} & A_{\partial} \end{bmatrix} - \begin{bmatrix} 0 & 0 \\ 0 & A_{\partial P}(W_P(s)A_P(s) - s^2 W_P(s))^{-1} W_P(s)A_{P\partial}(s) \end{bmatrix} - s^2 I \right) V_m, \quad (3.17)$$

and working out the inner-products we find

$$R_m(s) = \left( V_m^H W \begin{bmatrix} A_D & A_{D\partial} \\ A_{\partial D} & A_{\partial} \end{bmatrix} V_m - \begin{bmatrix} 0 & 0 \\ 0 & V_{m;\partial}^H W_{\partial} A_{\partial P}(W_P(s)A_P(s) - s^2 W_P(s))^{-1} W_P(s)A_{P\partial}(s) V_{m;\partial} \end{bmatrix} - s^2 I \right). \quad (3.18)$$

We observe that only the part of the operator concerned with the PML needs to be projected to obtain the reduced-order model for a new value of  $s$ .

The optimal Zolotarev rational approximants (see [20]) used for the PML construction make the size of the finite-difference problem (necessary for accurate approximation of the DtN map) in  $\mathbb{R}^k \setminus \Omega$  negligible compared to the grid in  $\Omega$ 's interior. In 2D for instance, the resulting equations that need to be solved in the PML can be solved efficiently with a block-cyclic solver [73], or with a band solver after sorting the PML system to a bandwidth of  $2k + 1$  (in 2D).

**REMARK 7.** — *Dispersive Media in RKS* – The rational Krylov subspaces can deal with operators that are nonlinear with respect to frequency. Therefore, an RKS can handle dispersive media without writing the constitutive relations and Maxwell equations as a first-order system, as long as these relations are passive. The snapshots of  $u(\kappa)$  are then computed for different wavespeeds at each frequency and upon evaluation the operator  $Q(s)$  is reprojected onto the basis. For the Maxwell equations in the diffusion approximation such a *parameter-dependent* rational Krylov subspace has been used to simulate dispersive media that follow a Cole-Cole model [33].

### 3.2.2. POLYNOMIAL AND EXTENDED KRYLOV SUBSPACES

In our model order reduction approach, our aim is to approximate the field solution  $\hat{u}(s) = [W\hat{A}(s) + sW]^{-1}\hat{b}$  by matrix functions that are easier to compute, for instance a polynomial or a low-order rational function. However, the introduction of coordinate

stretching to simulate open domains results in a discretized Maxwell system that is non-linear in frequency, which makes it computationally non-tractable to built high order, orthogonal polynomials in  $\hat{A}(s)$  as this requires recomputation of the polynomial for every value of  $s$ . Therefore, frequency-independent absorbing boundary conditions are desirable, as introduced in section 2.7.2. Starting point of the frequency-independent model reduction approach is equation (2.98), which for finite-dimensional vectors and in the Laplace domain, is given by

$$(A^{\hat{f}} + sI)\hat{u}^{\hat{f}} = W^{-1}b. \quad (3.19)$$

As projection spaces, polynomial Krylov subspaces

$$\mathcal{K}_{\text{PKS}}^m = \text{span} \left\{ W^{-1}b, A^{\hat{f}}W^{-1}b, \dots, \left( A^{\hat{f}} \right)^{m-1} W^{-1}b \right\}, \quad (3.20)$$

and extended Krylov subspaces

$$\mathcal{K}_{\text{EKS}}^{n_n, n_p} = \text{span} \left\{ \left( A^{\hat{f}} \right)^{-n_n} W^{-1}b, \dots, W^{-1}b, \dots, \left( A^{\hat{f}} \right)^{n_p-1} W^{-1}b \right\}, \quad (3.21)$$

are used with the starting vector  $W^{-1}b$  as equation (3.19) is a shifted system in this case.

Let  $V_m$  be a basis of the  $m$ -dimensional PKS that is  $W$  quasi-orthonormal in the transpose bilinear form. The PKS can be build iteratively following the condition

$$A^{\hat{f}}V_m = V_{m+1}T_{m+1} \quad (3.22)$$

with  $T_{m+1}$  a  $(m+1) \times m$  dimensional matrix. The  $i$ -th vector is a linear combination of the  $i-1$  preceding vectors and the vector  $Av^i$ . Therefore, it is clear that the leading  $m \times m$  block of  $T_{m+1}$  is upper Hessenberg. Multiplication by  $V_m^T W$  from the left leads to a symmetric matrix and reveals the leading  $m \times m$  block due to the quasi-orthonormality of  $V_m$  as

$$T_m = V_m^T W A^{\hat{f}} V_m = V_m^T W V_{m+1} T_{m+1}. \quad (3.23)$$

A symmetric, upper Hessenberg matrix is tridiagonal, so that the reduced-order model matrix  $T_m$  is tridiagonal. For completeness, we mention that the reduced system is given by  $R_m(s) = T_m + sI$  in this case. Due to the symmetry of the system matrix  $A^{\hat{f}}$  a reduced-order model can be computed via three-term recursion relations via the Lanczos algorithm [35]. In one iteration of the Lanczos algorithm only one sparse matrix-vector product and two vector-vector products are needed. For the extended Krylov subspace five-term recursion relations can be found leading to a symmetric pentadiagonal reduced-order model [47]. In order to use these efficient algorithms to construct the ROM we need to project the operator  $A^{\hat{f}}$  in the  $W$  weighted transpose bilinear form. Therefore, convergence of the algorithm cannot be proven as the algorithm can theoretically break down before convergence is reached. However, in applications such a breakdown has never been observed. For first-order, finite-difference operators it can even be proven that breakdown cannot occur in the first  $2d$  iterations, where  $d$  is the (smallest) number of grid cells between the support of  $b$  to the boundary of the domain.

As discussed in the last chapter, the linearization of the PML leads to unconditionally unstable matrices and stable field approximations can only be found via a stability-correction procedure. The stability-corrected solution is unconditionally stable and follows the Schwarz reflection principle as well.

The short recursion relations together with the fact that only the transfer-function

$$\hat{f}(r, \mathbf{b}, s) = r^T W \left[ r(\mathcal{A}^{\text{fi}}, s) + r(\overline{\mathcal{A}^{\text{fi}}}, s) \right] W^{-1} \mathbf{b} \quad (3.24)$$

is needed rather than the whole field approximation allows for the computation of the transfer-function without storing the basis. With  $V_m \in \mathbb{C}^{N \times m}$  as a basis matrix for the PKS or EKS, the reduced-order transfer-function can be written as

$$\hat{f}_m(r, \mathbf{b}, s) = r^T W \left[ V_m r(\mathbb{T}_m, s) V_m^T + \overline{V_m} r(\overline{\mathbb{T}_m}, s) \overline{V_m}^H \right] W^{-1} \mathbf{b} \quad (3.25)$$

using short term recurrence algorithms. Since the projection of the operator is computed with the short term recurrence algorithm, only the inner products  $r^T W V_m$  and  $V_m^T \mathbf{b}$  need to be kept in memory and the complete basis does not need to be stored. This makes these algorithms very memory efficient as only three vectors of the size of the computational domain need to fit inside the computational memory for PKS and five for EKS. In the time-domain the reduced-order solution can be written as

$$f_m(r, \mathbf{b}, t) = 2\eta(t) r^T W \Re \left[ V_m \eta(\mathbb{T}_m) \exp(-\mathbb{T}_m t) V_m^T \mathbf{b} \right] \quad \text{for } t > 0. \quad (3.26)$$

For SISO problems and with  $W^{-1} \mathbf{b}$  as starting vector for the EKS and PKS algorithm, the expressions further simplify as  $V_m^T W^{-1} \mathbf{b} = \sqrt{\mathbf{b}^T W^{-1} \mathbf{b}} e_1$  holds. The projected Maxwell operator  $\mathbb{T}_m = V_m^T W A^{\text{fi}} V_m$  is tridiagonal for PKS and pentadiagonal in the case of EKS, simplifying the evaluation of  $r(\mathbb{T}_m)$  or  $\exp(-\mathbb{T}_m t)$  even further.

**REMARK 8.** — *Numerical stability*—To enhance the numerical stability for ill-conditioned bilinear forms, it is possible to use basis vectors  $v^{[i]}$  that have unit Euclidean norm and are only  $W$ -orthogonal. The resulting reduced-order model is then symmetric with respect to the inverse of the diagonal matrix  $V_m^T W V_m$ , rather than being symmetric on its own.

Adding a vector to a PKS is computationally cheap and just requires one matrix-vector multiplication: it is equivalent in cost to one step of the Finite-Difference Time-Domain (FDTD) method. However, adding a negative power of matrix  $A^{\text{fi}}$  to an EKS is computationally more expensive, as one has to numerically solve a static (zero frequency) problem. Generally, EKS performs well if low frequencies need to be approximated, as it interpolates the original function and first  $n_n - 1$  derivatives at zero frequency. Finally, we mention that polynomial and extended Krylov field approximations of the stability-corrected transfer-function are stable by construction, since all eigenvalues of  $\eta(A^{\text{fi}})$  and  $\eta(\overline{A^{\text{fi}}})$  lie in the right half-plane.

The matrix  $A^{\text{fi}}$  is singular for wave propagation and diffusion and we therefore need to solve the system with the pseudo inverse of  $A^{\text{fi}}$  [66]. In the case of 2D wave propagation, this pseudo inverse corresponds to the law of Biot-Savart (the action of  $A^{\text{fi};\dagger} \mathbf{b}$ ). One has to essentially solve Poisson equations to compute the action of the pseudo inverse on a right-hand side [23]. This can be done with standard Poisson solvers or more generally with iterative solvers computing the minimum norm solution.

**REMARK 9.** — *Relation to shifted systems* – Finally, we mention that in our model reduction approach, we approximate shifted systems using polynomial and extended Krylov subspaces. The first-order Maxwell system leads to such a linear system if a fixed-frequency PML is used. The second-order Maxwell system leads to a quadratic equation

$$(D + M_{\varepsilon\mu} s^2 + M_{\mu\sigma} s)e = b'' \quad (3.27)$$

that can only be viewed as a shifted system in the lossless case, in the diffusion approximation or if  $\varepsilon$  is proportional to  $\sigma$  throughout the domain (see [65]). Thus for general lossy Maxwell systems, the first-order formulation is used in the proposed approach.

### 3.2.3. RESONANCES AS LANCZOS-RITZ PAIRS

Although the matrix  $A^{\text{fi}}$  has unstable eigenvalues, the stable eigenvalues for which the PML is matched are still valid finite-difference approximations of the resonances of the open system.

The Lanczos-Ritz pairs of the reduced-order model in turn approximate the mirrored<sup>2</sup> eigenvalues of the FD system and can thus be used to approximate the resonances of an open system. Let an eigenpair of  $T_m$  be given by  $(z_i, \lambda_i)$ , then the Lanczos-Ritz pairs are given by  $(V_m z_i, \lambda_i)$ ; i.e., the eigenvectors of the reduced-order model are the expansion coefficients of the basis. The modes are still quasi-normal in the  $W$  transpose inner product as the eigenvectors<sup>3</sup>  $z_i$  are quasi-normal in the transpose.

In an implementation it is usually not feasible to store the entire matrix  $V_m$  and the computational advantage of the short term recursion relations is memory efficiency as  $V_m$  is not needed to obtain the transfer-function. In the case a few resonance modes need to be approximated without storing  $V_m$ , two runs of the Lanczos algorithm are needed; in the first run the expansion coefficients  $z_i$  of the eigenvalues of interest are computed, while in the second run these coefficients are used during the Lanczos procedure to compute  $V_m z_i$ . As a measure to check if an eigenpair has converged, the residual

$$r_{\text{res}}^i = A^{\text{fi}} V_m z_i - \lambda_i V_m z_i = v^{[m+1]} \beta_{m+1}(z_i)_m \quad (3.28)$$

of that eigenpair is used, with  $\beta_{m+1} = (v^m)^T W v^m$  and  $(z_i)_m$  the last entry in the ROM eigenvector. This well-known relation shows that  $(V_m z_i, \lambda_i)$  is an approximate eigenpair of  $A^{\text{fi}}$  if  $\|r_{\text{res}}^i\|_2$  is small.

## 3.3. COMPUTATIONAL COMPLEXITY OF RKS, EKS AND PKS

In this section, we briefly discuss the computational complexity and memory requirements of the PKS, EKS, and RKS algorithms. With  $N$  denoting the order of the original unreduced Maxwell system and  $m$  the order of the reduced-order model, the performance of all three methods can be summarized as in Table 3.1. In this table, the order of the computational cost (in terms of  $N$ ) is  $N\psi(N)$  for solving a shifted system and  $N\phi(N)$

<sup>2</sup>Mirrored with respect to the origin as the eigenvalues of the matrix pencil  $(A^{\text{fi}}, -I)$  are the resonances.

<sup>3</sup>Existence of an eigenvalue decomposition cannot be guaranteed; however, orthogonality of eigenvectors still holds.



Table 3.1: Computational complexity, memory requirement and possible parallelization of the PKS, RKS and EKS method.

Method	Asymptotic complexity	Memory	Possible parallelization
RKS	$\psi(N)Nm$	$mN$	$m$
PKS	$Nm$	$3N$	–
EKS	$\phi(N)Nm$	$5N$	–

$N\psi(N)$  - Complexity of solving one shifted system

$N\phi(N)$  - Complexity of solving one static system ( $s = 0$ )

3

for a static system with zero shift. Thus  $\phi(N)$  and  $\psi(N)$  describe the cost of the particular solver used<sup>4</sup>.

From this table it is clear that there is a trade-off between memory requirements, parallelization and computational complexity. Specifically, one iteration of RKS is much more costly than one iteration of PKS or EKS, since it requires a solution of the unreduced Maxwell system at each iteration. On the other hand, RKS typically converges much faster than PKS or EKS as demonstrated in the numerical experiments at the end of this chapter. Another advantage of the RKS method is that the construction of the RKS basis is parallelizable, since multiple shifted systems can be solved in parallel. From a memory point of view, the PKS and EKS algorithms are more attractive due to the short term recurrence relations that allow for a memory efficient construction of these spaces. Only three vectors of the size of the computational domain need to be stored for PKS and five vectors for EKS. For the RKS method,  $m$  vectors need to be stored in memory. Finally, we note that in a sequential implementation of the RKS and PKS method, the computational cost of each iteration of RKS or PKS is approximately constant.

### 3.4. DIFFUSION APPROXIMATION VERSUS FULL MAXWELL SYSTEM

For the Maxwell equations in the diffusion approximation, polynomial or rational Krylov subspace methods clearly outperform time-stepping schemes [21, 51, 11]. In particular, to obtain the time-domain solution in the interval  $[0, T]$  to the diffusion equations, discretized with step size  $h$ , one needs  $m = \mathcal{O}(h^{-1}\sqrt{\log(h)T})$  PKS iterations as opposed to  $\mathcal{O}(h^{-2}T)$  time steps using an explicit second-order time-stepping scheme. Furthermore, the eigenvalues of the diffusion operator lie on the positive, real semi-axis and the optimal shifts for an RKS are known via a Zolotarev rational approximant [51]. In contrast to this, the CFL condition of the FDTD method for Maxwell's wave equations with lossless media yields an estimate of only  $\mathcal{O}(h^{-1}T)$  time steps. In fact, one can show that the PKS method for these types of problems is computationally equivalent to FDTD with an optimal time step in case of PEC boundary conditions and lossless media [67]. Material losses and the PML, however, shift the spectrum of the first-order wave operator away from the imaginary axis and into the complex plane. The separation of the spectrum

<sup>4</sup>If a sparse LU factorization is used to solve the 2D Helmholtz equations for instance,  $\psi(N) = \sqrt{N}$  [38].

from the area of approximation (the imaginary axis) enhances the convergence of PKS.

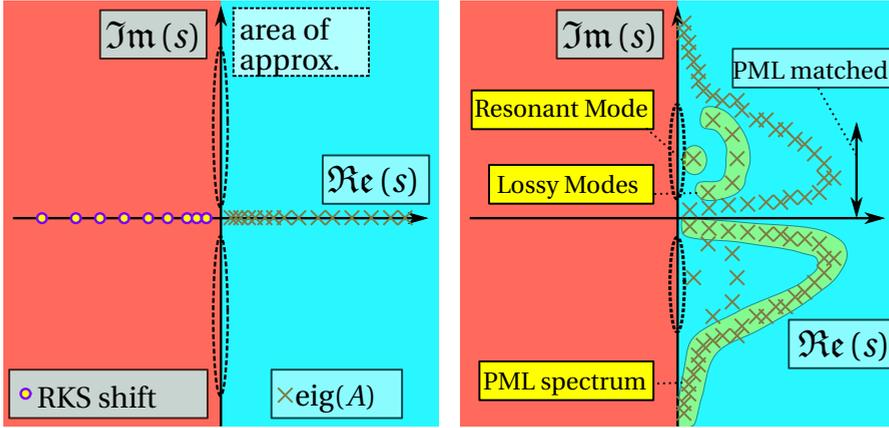
In Figure 3.1, we show the spectrum of a typical discretized first-order wave and second-order diffusion operator. All eigenvalues are shown as red crosses and are located in the right-half of the complex  $s$ -plane. The spectra of the diffusion and wave operator are clearly different – eigenvalues of the diffusion operator are located on the positive real line, while the wave operator has a (symmetric) spectrum that is completely located in the stable right half-plane  $\text{Re}(s) \geq 0$ . For frequency-domain problems, we want to obtain an accurate approximation in a frequency interval of interest that is usually located on the imaginary axis (encircled with a dashed black line). After discretization of the continuous wave operator, its spectrum becomes discrete and shifts into the complex plane due to material losses and the introduction of a PML. The fixed-frequency PML is only matched in a certain region of interest, corresponding to strongly damped eigenvalues with a large real part. Cavities or other resonant inclusions can introduce so-called resonance modes, which are characterized by eigenvalues with a small real part, as they lead to a slowly decaying exponent in the transfer-function of equation (3.26). The separation of the area of approximation and the spectrum also allows for efficient model reduction via rational Krylov subspaces. However, contrary to the diffusive case, the eigenvalues of the wave operator are not located on the real axis, but can be found in the complex plane (For more information about the spectrum of  $A(s)$  we refer to [49], for example). A priori selection of optimal shifts as in the diffusive case is therefore very difficult. In general, the frequency-domain response of a wave equation can be oscillatory which limits an interpolatory method like the RKS method. In conclusion, introducing losses into the Maxwell system allows us to find smaller reduced-order models. Losses lead to loss of information and allows for a low-order approximation that can capture the relevant features of the transfer-function. This will be illustrated with a numerical experiment in the application section.

### 3.5. APPLICATIONS

In this section, we show the performance of the introduced model reduction techniques in several applications. First, we consider the computation of the spontaneous decay rate of a quantum emitter in vicinity of a dispersive resonator. Next, applications from geophysics are considered. The effect of losses on the convergence behavior is studied for a ground penetrating radar. Further, for another example from geophysics, we consider the Maxwell equations in the diffusion approximation. The last example, a lossless dielectric box, shows that the developed methods don't rely on losses being present in the configuration. More specifically, the transfer-function of such a dielectric box can be well approximated by a low-degree rational function which allows for an approximation in a small rational Krylov subspace.

In this section, we show the wide applicability of model reduction for wave equations. Localized fields in resonators, fields in lossy environments and fields in configurations where the transfer-function can be described by a low-degree rational function are all well suited for model order reduction techniques.





(a) Illustration of the eigenvalues of the second-order diffusion operator.

(b) Illustration of the eigenvalues of the (linearized, stabilized) first-order wave operator  $A^{\text{fi}}$  including stability correction

Figure 3.1: Comparison of the spectra of a linearized wave operator and a diffusion operator. The unstable left half-plane ( $\text{Re}(s) < 0$ ) is depicted in red, while the stable right half-plane ( $\text{Re}(s) > 0$ ) is shown in blue.

### 3.5.1. SPONTANEOUS DECAY RATE AND DISPERSIVE RESONATORS

The spontaneous decay (SD) rate of a quantum emitter depends on its environment and can be modified by an electromagnetic resonance. This so-called Purcell effect [64] is a basic effect in quantum electrodynamics and it is well known that in the so-called weak-coupling regime, the SD rate can be computed classically and does not require a quantum mechanical treatment. Specifically, for electric-dipole transitions that take place at  $\mathbf{r} = \mathbf{r}_0$ , the SD rate  $\gamma$  normalized with respect to the decay rate  $\gamma_0$  in a reference medium can be computed as [58]

$$\gamma/\gamma_0 = P/P_0, \quad (3.29)$$

where

$$P = \frac{\omega}{2} \text{Im}[\bar{\mathbf{p}} \cdot \mathbf{E}(\mathbf{r}_0)] \quad (3.30)$$

is the power radiated by an electric dipole of the form  $\hat{\mathbf{j}}^{\text{ext}} = -i\omega\mathbf{p}\delta(\mathbf{r} - \mathbf{r}_0)$  with  $\mathbf{p}$  the dipole moment and  $P_0$  is the radiated power in the reference medium.

To determine the SD rate, the electric field at the location of the quantum emitter is required over a spectral interval of interest. For emitters located in the vicinity of dispersive nano-structures, we therefore have to solve the Maxwell equations for a coinciding source and receiver pair. The PKS algorithm is used for this application as the considered three-dimensional configurations with subwavelength geometries lead to very large finite-difference grids and only a few vectors fit into the computational memory. The PKS algorithm also allows us to efficiently compute and study the resonances in terms of the Lanczos-Ritz pairs.

**PROPOSITION 3.4.** — The normalized spontaneous decay rate of a quantum emitter can be computed in terms of the ROM as

$$\frac{P}{P_0} = \frac{\delta_v^3 6\pi c^3}{\varepsilon_r \omega^2} \Re \mathbf{e}_1^T \left[ \eta(\mathbb{T}_m)(\mathbb{T}_m + s\mathbb{I}_m)^{-1} + \eta(\bar{\mathbb{T}}_m)(\bar{\mathbb{T}}_m + s\mathbb{I}_m)^{-1} \right] \mathbf{e}_1. \quad (3.31)$$

*Proof.* The power radiated by a dipole inside a homogeneous medium with refractive index  $n$ , relative permittivity  $\varepsilon_r$ , operating at a frequency  $\omega$ , and having a dipole moment  $\mathbf{p}$  is given by [36]

$$P_0 = \frac{\omega^4 \|\mathbf{p}\|_2^2}{12\pi \varepsilon_0 c^3}. \quad (3.32)$$

As mentioned above, to compute the spontaneous decay rate, a PKS reduced-order model is used. In our finite-difference approximation we find that  $-\mathbf{b}$  is the finite-difference approximation of the delta pulse in the direction of  $\mathbf{p}$  scaled by the voxel volume of the corresponding source voxel  $\delta_v^3$  and with wavelet  $w = s\|\mathbf{p}\|$ . In addition, the receiver  $r = -\|\mathbf{p}\|\tilde{\mathbb{W}}^{-1}\mathbf{b}$  is an approximation of the delta function at the source location, since the source and receiver coincide for this application. The starting vector in the basis  $V_m$  is  $\mathbb{W}^{-1}\mathbf{b}$  and consequently  $V_m^T \mathbb{W} \mathbb{W}^{-1}\mathbf{b} = \mathbf{e}_1 \sqrt{\mathbf{b}^T \mathbb{W}^{-1}\mathbf{b}}$  is the scaled first unit vector. To compute the spontaneous decay rate, the inner product of the polarization vector with the electric field needs to be evaluated. With  $\tilde{\mathbb{W}}$  the matrix containing the Yee-voxel size this inner product is obtained in terms of the ROM as

$$\tilde{\mathbf{p}} \cdot \hat{\mathbf{E}}(\mathbf{x}_0) \approx w(s) r^T \tilde{\mathbb{W}} V_m \left[ \eta(\mathbb{T}_m)(\mathbb{T}_m + s\mathbb{I}_m)^{-1} + \eta(\bar{\mathbb{T}}_m)(\bar{\mathbb{T}}_m + s\mathbb{I}_m)^{-1} \right] V_m^T \mathbf{b}. \quad (3.33)$$

Substitution of the receiver vector  $r = -\|\mathbf{p}\|\tilde{\mathbb{W}}^{-1}\mathbf{b}$  in the above expression gives

$$\tilde{\mathbf{p}} \cdot \hat{\mathbf{E}}(\mathbf{x}_0) \approx -w(s) (\|\mathbf{p}\|\tilde{\mathbb{W}}^{-1}\mathbf{b})^T \tilde{\mathbb{W}} V_m \left[ \eta(\mathbb{T}_m)(\mathbb{T}_m + s\mathbb{I}_m)^{-1} + \eta(\bar{\mathbb{T}}_m)(\bar{\mathbb{T}}_m + s\mathbb{I}_m)^{-1} \right] V_m^T \mathbf{b},$$

and since the inner product of the source and receiver vector with the Lanczos vectors is known we find

$$\tilde{\mathbf{p}} \cdot \hat{\mathbf{E}}(\mathbf{x}_0) \approx -w(s) \|\mathbf{p}\| (\mathbf{b}^T \mathbb{W}^{-1}\mathbf{b}) \mathbf{e}_1^T \left[ \eta(\mathbb{T}_m)(\mathbb{T}_m + s\mathbb{I}_m)^{-1} + \eta(\bar{\mathbb{T}}_m)(\bar{\mathbb{T}}_m + s\mathbb{I}_m)^{-1} \right] \mathbf{e}_1.$$

Finally, using  $w(s) = s\|\mathbf{p}\|$  and  $s = -i\omega$  we obtain

$$\tilde{\mathbf{p}} \cdot \hat{\mathbf{E}}(\mathbf{x}_0) \approx i\omega \frac{\|\mathbf{p}\|^2 \delta_v^3}{\varepsilon_0 \varepsilon_r} \mathbf{e}_1^T \left[ r(\mathbb{T}_m) + r(\bar{\mathbb{T}}_m) \right] \mathbf{e}_1, \quad \text{with } r(z) = \frac{\eta(z)}{z + s}. \quad (3.34)$$

In this derivation we used that  $\mathbf{b}^T \mathbb{W}^{-1}\mathbf{b} = \frac{\delta_v^3}{\varepsilon_0 \varepsilon_r}$ , leading to a reduced-order approximation of the radiated power as

$$P = \frac{\delta_v^3 \omega^2 \|\mathbf{p}\|^2}{2\varepsilon_r \varepsilon_0} \Re \mathbf{e}_1^T \left[ \eta(\mathbb{T}_m)(\mathbb{T}_m + s\mathbb{I}_m)^{-1} + \eta(\bar{\mathbb{T}}_m)(\bar{\mathbb{T}}_m + s\mathbb{I}_m)^{-1} \right] \mathbf{e}_1. \quad (3.35)$$

Finally, the reduced-order approximation of the normalized spontaneous decay rate follows as

$$\frac{P}{P_0} = \frac{\delta_v^3 6\pi c^3}{\varepsilon_r \omega^2} \Re \mathbf{e}_1^T \left[ \eta(\mathbb{T}_m)(\mathbb{T}_m + s\mathbb{I}_m)^{-1} + \eta(\bar{\mathbb{T}}_m)(\bar{\mathbb{T}}_m + s\mathbb{I}_m)^{-1} \right] \mathbf{e}_1. \quad \square$$



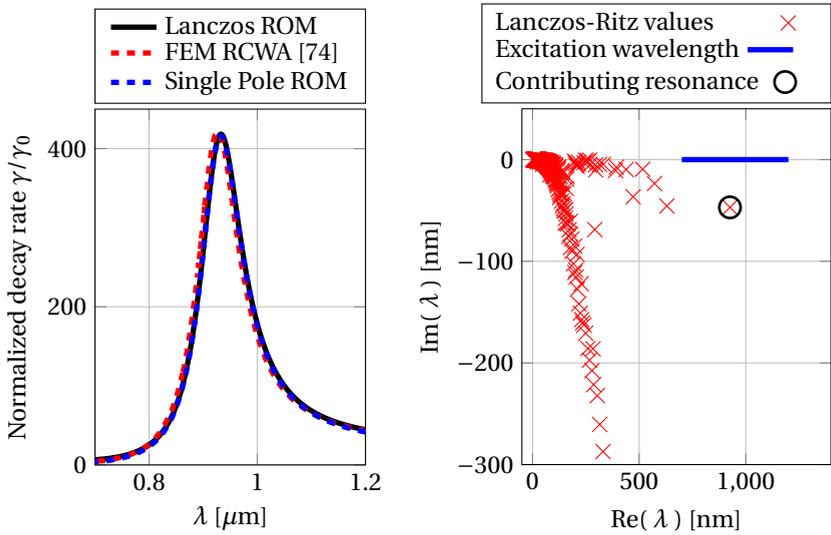
### GOLDEN NANOROD

As a first example, we consider a configuration similar to the one presented in [74], which consists of an electric dipole located in the vicinity of a golden nanorod (see Figure 3.4(a)). The rod has a diameter of 30 nm and a length of 100 nm and the dipole is located 10 nm above the upper surface of the rod. The background medium is homogeneous and is characterized by a refractive index of  $n = 1.5$ , while a Drude model with a plasma frequency  $\omega_p = 1.26 \cdot 10^{16} \text{ s}^{-1}$  and a collision frequency  $\gamma_p = 1.41 \cdot 10^{14} \text{ s}^{-1}$  is used as a constitutive relation for the golden nanorod. We are interested in the SD rate of the quantum emitter on a wavelength interval ranging from  $0.7 \mu\text{m}$  to  $1.2 \mu\text{m}$ . Discretizing the first-order Maxwell system of equations (2.27–2.30) such that the electromagnetic field and the geometry are well resolved for all wavelengths of interest, we obtain a semi-discrete Maxwell system as given by equation (2.75) with approximately 8.7 million unknowns. Given this large order, it is clear that direct evaluation of equation (2.102) is simply not feasible. We therefore construct the reduced-order model for the radiated power as given by equation (3.35) via the Lanczos reduction algorithm. For this particular example, it turns out that a model of order  $m = 4500$  is sufficient to accurately describe the SD rate of the quantum emitter over the entire wavelength interval of interest. Since the order of the original system is approximately 8.7 million, the order of the reduced-order model is about 1930 times smaller than the original system. Also note that the system matrix  $A^{\text{fi}}$  is sparse, and the Lanczos algorithm is therefore very efficient, since the algorithm is based on a three-term recurrence relation and the system matrix is only needed to compute a single matrix-vector multiplication at every iteration.

Now taking the homogeneous background medium as a reference medium to compute  $P_0$  (see equation (3.36)), we obtain the normalized decay rate curve shown in Figure 3.2(a) (black solid line). Also shown is the normalized SD rate as computed in [74] with 2D Rigorous Coupled-Wave Analysis (RCWA) that assumes rotational symmetry of the configuration (red dashed line). Clearly, both SD curves essentially overlap on the wavelength interval of interest.

In addition to the reduced-order models for the SD rate, we can also determine the dominant quasi-normal modes from the Lanczos decomposition. As an illustration, Figure 3.2(b) shows *all* eigenvalues of the reduced Lanczos matrix  $T_{4500}$  in the complex  $\lambda$ -plane. The reduced-order model takes the contribution of all these eigenvalues into account, but only one eigenvalue (encircled in Figure 3.2(b)) essentially contributes to the SD rate response. The mode that corresponds to this eigenvalue has converged and in Figure 3.3(a) we show the magnitude of the  $x$ -component of the electric field strength of this dominant quasi-normal mode for the quantum emitter configuration of Figure 3.4(a), while Figure 3.3(b) shows the  $y$ -component of the electric field strength. The solution can also be expanded in this one resonant mode as shown in Figure 3.2(a) (blue dashed line). Higher order modes that, do not contribute to the field excited by the dipole are shown in Figure 3.4(b) and Figure 3.4(c).

Finally, we remark that in exact arithmetic only those modes that are excited by the emitter can be captured by the Lanczos algorithm, since the (scaled) source vector  $W^{-1}\mathbf{b}$  is used as a starting vector. These are the modes of interest, of course, since possible other quasi-normal modes are simply not excited.



(a) Spontaneous decay rate as approximated by the ROM, shown alongside an RCWA comparison solution and the contribution of the main resonance. (b) Lanczos-Ritz values of the reduced-order model.

Figure 3.2: Spontaneous decay rate and the Lanczos-Ritz values of the simulated structure.

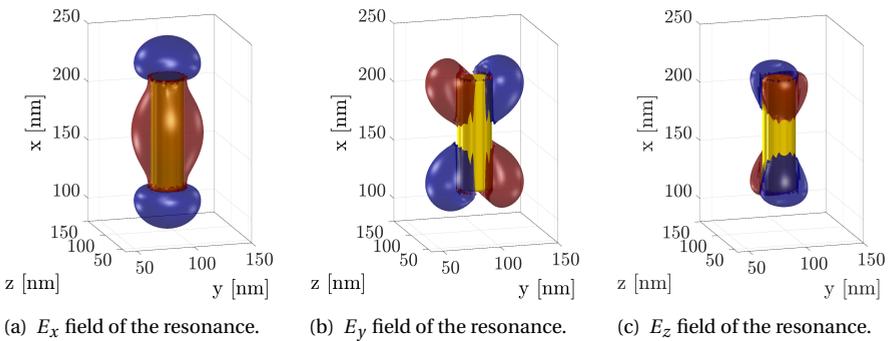


Figure 3.3: Isosurfaces of the electrical field strength of the fundamental resonance at  $\lambda = 926 - 47i$  nm.



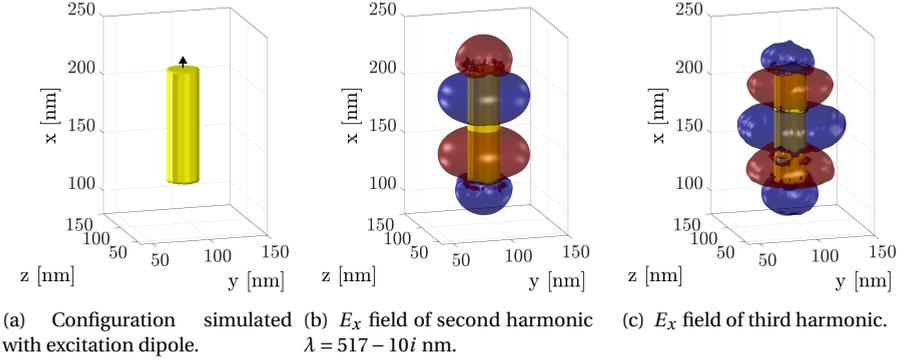


Figure 3.4: Second and third harmonic resonance fields of the nanorod configuration and the simulated configuration with the quantum emitter are shown. (a) Quantum emitter located above the golden nanorod, (b) the second and (c) the third harmonic resonance field.

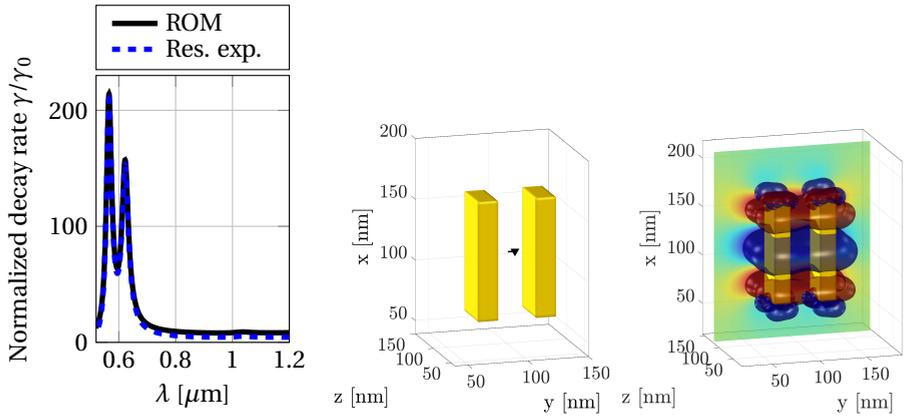
### PARALLEL PLATE OSCILLATOR

As a second example, we consider a quantum emitter located at the center of a gap between two rectangular golden nanorods (see Figure 3.5(b)).

The side lengths of each rectangle are given by  $102 \times 20 \times 40$  nm and the gap is 28 nm wide. We use the same Drude model as in the previous example as a constitutive relation for gold. The rods and emitter are embedded in a homogeneous background medium with a refractive index of  $n = 1.5$ .

Discretizing the first-order Maxwell system in space such that the electromagnetic field is well resolved for wavelengths running from  $0.52 \mu\text{m}$  to  $1.2 \mu\text{m}$ , a semi-discrete Maxwell system with approximately 7.2 million unknowns is obtained. Using the Lanczos algorithm, we construct reduced-order models of increasing order until convergence is reached. For this example, a reduced-order model of order  $m = 10000$  is sufficient to accurately describe the SD rate on the wavelength interval of interest (see Figure 3.5(a)). Again, the SD rate is computed by taking all eigenvalues of the reduced Lanczos matrix  $T_m$  into account. However, only a small number of eigenvalues actually contribute to the SD rate in this configuration, as shown by the dashed curve in Figure 3.5(a). The fields of one of the main contributing resonances for a dipole source oriented in the  $y$ -direction are shown in Figure 3.6.

Comparing the SD experiment for the nanorod with the experiment for the parallel-plate antenna, we observe that the two experiments differ not only in model complexity, but also in the wavelength-range of interest. Specifically, in the nanorod experiment only a single resonance essentially contributes to the SD rate on the wavelength interval of interest, while for the antenna two resonance modes contribute on a wavelength interval that is larger than in the first example (smaller wavelengths are considered in the second example compared with the first). Convergence therefore slows down, since the wavelength interval and the number of modes that need to be captured is larger. Moreover, smaller wavelengths put stronger constraints on the PML, which usually leads to a poorer conditioning of the system matrix.



(a) Normalized spontaneous decay rate as approximated by the ROM, shown alongside the contribution of the main resonances for a dipole oriented along the  $y$ -axis. (b) Simulated configuration of parallel plates. (c) Higher order resonance mode.

Figure 3.5: The normalized spontaneous decay rate for a quantum emitter between two parallel plates (a), the parallel plate configuration with the quantum emitter (b) and the isosurface of an high order anti-symmetric resonance mode (c).

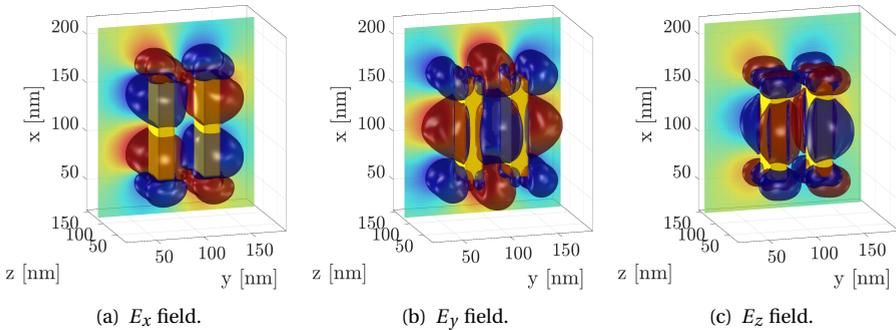


Figure 3.6: Isosurface of one of the two main contributing, anti-symmetric resonance mode of the parallel plate configuration with  $\lambda = 622 - 14i$  nm. Shown are the  $x$ -component (a), the  $y$ -component (b) and the  $z$ -component (c) of the electric field strength.



The construction of the reduced-order model in Matlab takes less than one hour for the first example and 2.5 hours for the second example on an Intel i5-3470 CPU 3.2 GHz under 64-bit Windows 7. As soon as the model is constructed, its evaluation on the complete wavelength interval of interest takes less than one second for one thousand uniformly sampled wavelength values. This “wavelength sweeping” feature is the main advantage of our reduced-order model approach.

Next to the spontaneous decay rate, the reduced-order modeling approach allows us to approximate the resonances of the system. In Figures 3.8 and 3.9 we show the magnitude of the  $x$ -,  $y$ - and  $z$ -components of the electric field strength of the quasi-normal modes corresponding to the fundamental resonances. A symmetric and an anti-symmetric mode can be found, where the symmetric mode has a higher frequency of oscillation.

The symmetric and anti-symmetric resonances and the difference in oscillation frequency can be analyzed with a simple circuit model of the parallel plate configuration. Imagine each of the parallel plates as a lossless  $LC$ -oscillator. Coupling two of these oscillators with coupling capacitors we end up with the circuit shown in Figure 3.7.

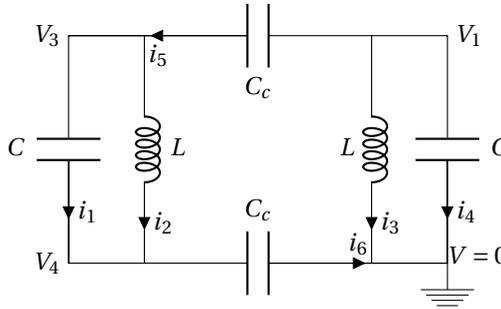


Figure 3.7: Two harmonic oscillators coupled with a coupling capacitor of size  $C_c$ .

To find the resonances of this circuit we analyze it in the Fourier domain and search for solutions that can sustain themselves in absence of sources. The sum of all currents in every node needs to be zero so that we find  $i_5 = i_6$ , which leads to  $V_1 = V_3 + V_4$ . Further, using the current-voltage relationships for each component in the Fourier domain we find that

$$(1 - \omega^2 CL)V_1 = -\omega^2 C_c V_4. \quad (3.36)$$

This equation has two solutions, namely, the symmetric and anti-symmetric oscillatory modes. The symmetric mode is equivalent to the resonance of the independent  $LC$ -circuits. In the symmetric oscillatory mode there is no voltage drop across  $C_c$  such that

$$\omega^{\text{sym}} = \frac{1}{\sqrt{LC}}, \quad \text{with } V_4 = 0, V_3 = V_1 \text{ and } i_5 = 0. \quad (3.37)$$

The two coupled oscillators are not interacting with each other in this case. In the anti-symmetric case, however, the oscillation frequency decreases to

$$\omega^{\text{anti}} = \frac{1}{\sqrt{L(C + C_c)}}, \quad \text{with } V_3 = 0, V_4 = V_1 \text{ and } i_5 \neq 0. \quad (3.38)$$

Clearly the anti-symmetric mode has a lower oscillation frequency than the symmetric one, which is equivalent to the coupled parallel plate oscillators discussed in this section. Obviously this model is an over simplification of the situation; however, it enables us to understand the difference in symmetric and anti-symmetric oscillation frequency. According to this model the resonance of the symmetric mode and a single plate would need to coincide. This is not the case for the parallel plate oscillator, as the single plate supports a fundamental resonance at a wavelength that is 50 nm larger than the one of the symmetric parallel plate configuration as shown in Figure 3.10.

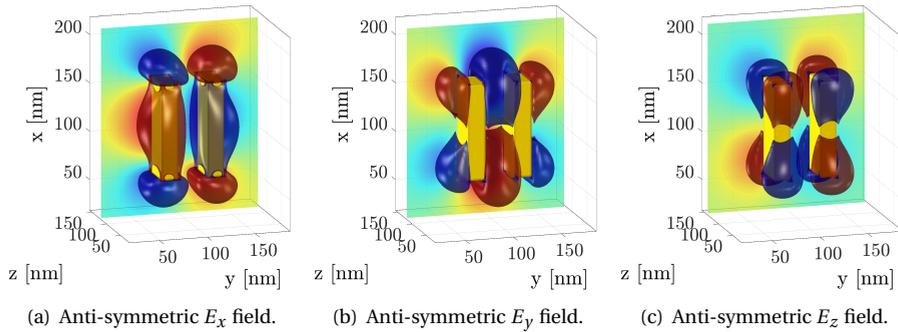


Figure 3.8: Isosurfaces of all three electric field strength components of the anti-symmetric resonance mode of the parallel plate configuration with  $\lambda = 1034 - 34i$  nm.

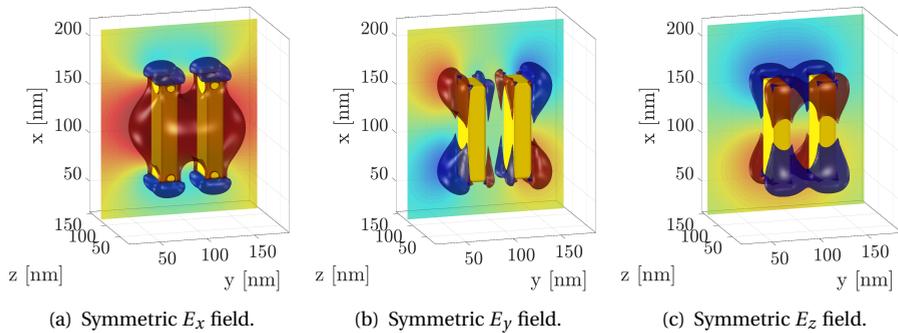


Figure 3.9: Isosurfaces of all three electric field strength components of the symmetric resonance mode of the parallel plate configuration with  $\lambda = 891 - 68i$  nm.



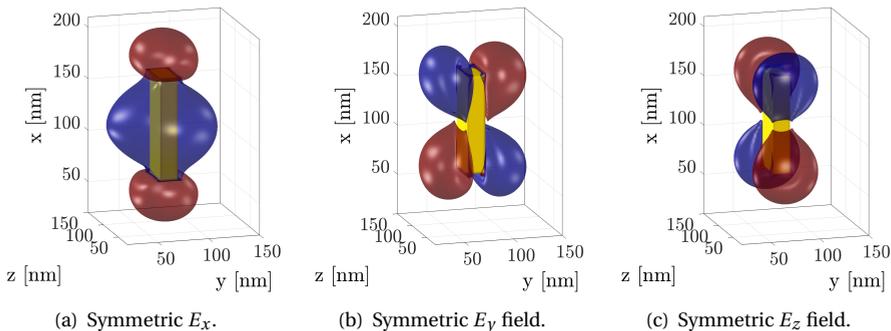


Figure 3.10: Isosurfaces of all three electric field strength components of the fundamental resonance mode of a single plate configuration with  $\lambda = 941 - 52i$  nm.

### 3.5.2. MODEL ORDER REDUCTION IN GEOPHYSICS

Krylov-based EM solvers for the Maxwell system in the diffusion approximation have long been the workhorse of the geophysics community (see [11, 21, 51]). In such a solver, the discretized system of equations is projected onto a Krylov subspace in an iterative fashion. The projected system forms a reduced-order model from which approximate solutions to the full order system can be drawn. In geoscience the inverse problem is typically of interest – imaging the conductivity of the Earth from a few measurements of a ground-penetrating electromagnetic field. Therefore, only the transfer-function from distinct sources to receivers are of interest. The domains encountered are typically open so that the computational domain is truncated by absorbing boundary conditions.

The Earth is a conductive medium with two electromagnetic regimes: the diffusion and wave regime. In the diffusion regime the displacement currents  $\epsilon \partial_t \mathbf{E}$  are much smaller than the conduction currents due to high conductivity profiles or low frequencies. In the wave regime low losses are present or high frequencies of operation or both and the displacement currents cannot be neglected.

There are several reasons to use model order reduction in geosciences. In the model order reduction framework the computation of Jacobians for the associated inverse problems is fast and very memory efficient. Further, the computational domains encountered are typically very large so that the computational demands can be lowered with model order reduction compared to traditional methods. The electromagnetic losses present in the Earth cause the spectrum of the Maxwell operator to shift into the complex plane away from the imaginary axis. Such a separation between the imaginary axis and the spectrum of the operator aids low-order approximations. From a physics stand point, dissipation typically leads to loss of information allowing for lower order approximations (see also page 56).

In this section we compare three types of projection-based Krylov model order reduction techniques for electromagnetic field problems as encountered in geophysics. We use several ground penetrating radar scenarios to illustrate the performance of the various Krylov reduction techniques. In addition, a 3D example for the diffusive Maxwell

equations is presented as well. Particular attention is paid to the effects of the conductivity in the lossy Maxwell system on the convergence speed of the ROMs.

### TWO-DIMENSIONAL GROUND PENETRATING RADAR

To show the approximation qualities of the introduced ROMs, we consider a configuration arising in exploration with a ground penetrating radar (GPR). We investigate a lossy subsurface with a box shaped anomaly and use a frequency band between 50 MHz and 3 GHz. The simulated configuration is shown in Figure 3.11(a), where the exact medium parameters are provided in the caption. An electrical dipole source oriented in the  $z$ -direction is used for excitation and the  $z$ -component of the electric field strength is measured at the ground-air interface. Finite-difference discretization with a second-order accurate operator at a maximum frequency of 3 GHz leads to a first-order Maxwell system with  $N = 62 \cdot 10^3$  unknowns. As comparison solution for our reduced-order models this full-order operator is used to compute the transfer-function for this configuration using a Finite-Difference Frequency-Domain (FDFD) method. This comparison solution is shown alongside the responses of the PKS, EKS and RKS reduced-order models in Figure 3.12. The order of the reduced-order models was increased until the reduced-order models and the comparison solution essentially overlap on the frequency interval of interest.

For the RKS method we choose equidistant shifts (interpolation points) on the imaginary axis between 50 MHz and 3 GHz. The RKS method leads to the smallest model with  $m = 55$ ; however, a single iteration is much more expensive than a PKS iteration. The PKS method needs  $m = 3400$  iterations until convergence, where one iteration is as expensive as an FDTD step. The EKS method needs  $n_n = 100$  inverse and  $n_p = 2200$  forward iterations. The only difference can be seen at very low frequencies where the frequency-independent and dependent PML deviate. Furthermore, the PKS method converges from large eigenvalues towards small eigenvalues meaning that high frequencies are approximated first. The EKS method with 100 inverse iterations interpolates at zero frequency and it converges from low and high frequencies towards the middle of the spectrum. The RKS method with uniform distributed shifts interpolates the response at all shifts and therefore converges on the whole interval of shifts simultaneously. The convergence of the different Krylov methods in the frequency-domain is shown in more detail in Figure 3.11(b), where the relative  $L_2$  error between the different Krylov ROMs and the FDFD comparison solution is shown as a function of the number of iterations. We measure the approximation by the relative  $L_2$  error in the time-domain given by

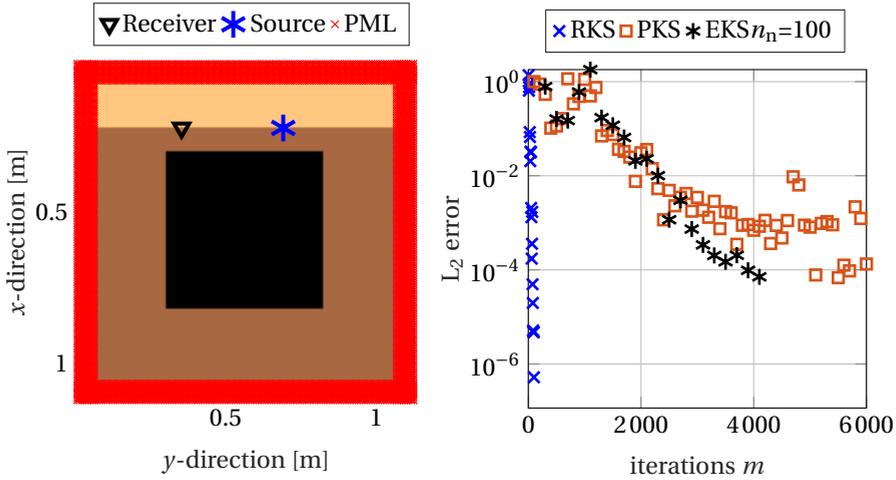
$$\text{err} = \frac{\|f_m(t) - f(t)\|_2}{\|f(t)\|_2}, \quad (3.39)$$

and use an FDFD comparison simulation to obtain  $f(t)$ .

The RKS-ROM converges up to machine precision to the FDFD comparison solution, since both solutions are computed using the frequency-dependent Maxwell operator. The error of the PKS and EKS ROMs stagnates around an error level of about  $10^{-4}$ , since these models use the frequency-independent PML formulation.

In Figure 3.13 the convergence of the PKS method in the time-domain is illustrated. Here, the time derivative of a Gaussian pulse, with a center frequency of 1.15 GHz





(a) Simulated GPR configuration, with an ( $\epsilon_r = 4, \sigma = 10^{-2} \text{ S/m}$ ) anomaly embedded in a ( $\epsilon_r = 2, \sigma = 5 \cdot 10^{-4} \text{ S/m}$ ) surface layer, with dry air ( $\epsilon_r = 1, \sigma = 0$ ) on top.. (b) Relative error of the PKS-, EKS- and RKS-ROMs. The PKS- and EKS-ROMs stagnate around a relative error of about  $\text{err} = 10^{-4}$  as the comparison solution was computed with an FDFD method with frequency-dependent PML.

Figure 3.11: Configuration of the ground penetrating radar application and convergence curves of the different model reduction techniques.

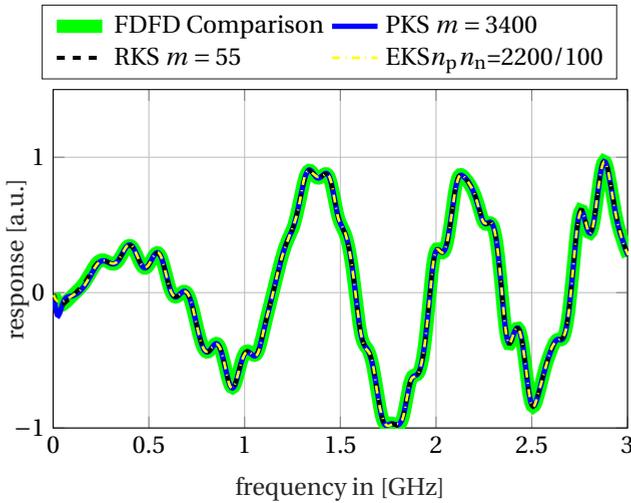


Figure 3.12: Imaginary part of the transfer-function over the frequency interval of interest computed with FDFD (green solid line), PKS after 3400 iterations (blue solid line), EKS after 2200 forward and 100 inverse iterations with  $A^{\text{fi}}$ , respectively, (yellow solid line), and RKS after 55 iterations (dashed line).

was used as a wavelet. The 1% cutoff frequency of this wavelet is 2.4 GHz and discretization is chosen such that we have about 10 points per smallest wavelength at this cutoff frequency. We observe that early times are approximated well after  $m = 1100$ ; however, late times containing the reflection of the anomaly did not converge yet. After  $m = 3400$  iterations the time-domain error is below 0.1% and the reduced-order approximation and FDFD comparison solution are indistinguishable. The direct arrival of the pulse and the reflections from the anomaly can clearly be distinguished in this experimental setup.

To study the effect of losses on the convergence rate of RKS-ROMs in the time-domain, we repeat the above experiment with different conductivity models. Specifically, we scale the conductivity profile by a factor of 0.5, 2, and 5. The relative time-domain errors (with respect to an FDTD comparison solution) for different iterations of RKS is shown in Figure 3.14, where we used the same wavelet as in the previous experiment. Higher losses clearly increase the convergence rate, as it shifts the spectrum of the operator further away from the imaginary axis and into the complex plane. For small dimensions of the Krylov subspace the errors in all experiments are similar due to the presence of a direct lossless path between the sources and receivers. Once this direct arrival is well approximated, the errors of the different conductivity models start to differ and higher losses positively influence the convergence rate. The convergence of RKS is not monotonic, however, since we use equidistant shifts for our experiments, meaning that we rebuild our subspace for each of the experiments.

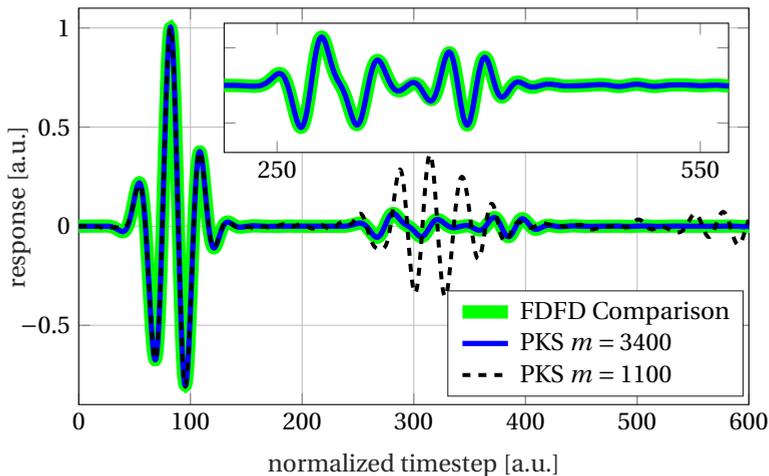


Figure 3.13: Time-domain trace as approximated by the PKS method after 1100 iterations and after 3400 iterations. For a PKS method the early times converge first. The reflections from the anomaly can be seen in the zoomed panel in the top right. The FDFD comparison response and ROM-PKS response for  $m = 3400$  are magnified for (normalized) times between 250 and 550.

### THREE-DIMENSIONAL GROUND PENETRATING RADAR

To test the algorithm for a three-dimensional ground penetrating radar application a similar box configuration as in the two-dimensional case is used. A cube shaped box with side length 40 cm and a relative permittivity of  $\epsilon_r = 4$  and conductivity of  $\sigma = 0.01$  S/m is



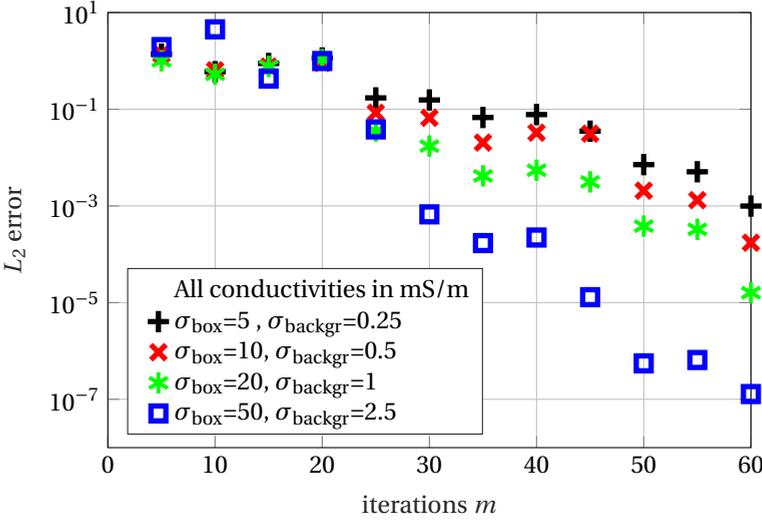


Figure 3.14: Time-domain convergence of the RKS method for four different media with increasing conductivity. All conductivities are given in mS/m and the red crosses correspond to the previously discussed experiment.

buried 60 cm deep in the subsurface which has a permittivity of  $\epsilon_r = 2$  and conductivity of  $\sigma = 0.5$  mS/m. A source and receiver are placed on the Earth-air interface, centered above the box and 23 cm separated. We are interested in simulating the frequency interval up to 2.4 GHz. Uniform discretization of a 3D volume of  $75 \text{ cm} \times 75 \text{ cm} \times 160 \text{ cm}$  with a Yee-grid leads to 10 million field unknowns and we use a PKS reduced-order model to compute the frequency-domain transfer-function.

The approximation of the transfer-function is shown in Figure 3.15 for reduced-order models of increasing dimensionality. After 3500 iterations the  $L_2$  error of the ROM is 1.5% and only small differences with the ROM of dimension 15000 can be seen. The PKS method clearly converges from high frequencies to the low frequencies similar to a power method for eigenvalue computation. I took around 30 minutes to compute the reduced-order model of order 3500 and only 0.5 s to evaluate 1000 frequency points using the model in a Matlab implementation.

### THREE-DIMENSIONAL ELECTROMAGNETIC DIFFUSION

In case the displacement currents  $\epsilon \partial_t \mathbf{E}$  are negligible compared to the conduction currents  $\sigma \mathbf{E}$ , we obtain the diffusion approximation of Maxwell's equations. In this computational example, we consider a three-dimensional configuration with a resistive, anisotropic anomaly. The geometrical configuration is a scaled version of the previous example and is shown in Figure 3.16(a), where the source and receiver are directed in the positive  $z$ -direction. The lower half space has a conductivity of  $\sigma = 1$  S/m and the resistive, anisotropic, cube anomaly has principal components  $\sigma_I = 0.01$  S/m,  $\sigma_{II} = 0.04$  S/m, and  $\sigma_{III} = 0.02$  S/m with a dip and azimuth angle of the anisotropy of 30 and 15 degrees. After discretization with finite-differences, the model has  $N = 1.5 \cdot 10^6$  field unknowns in

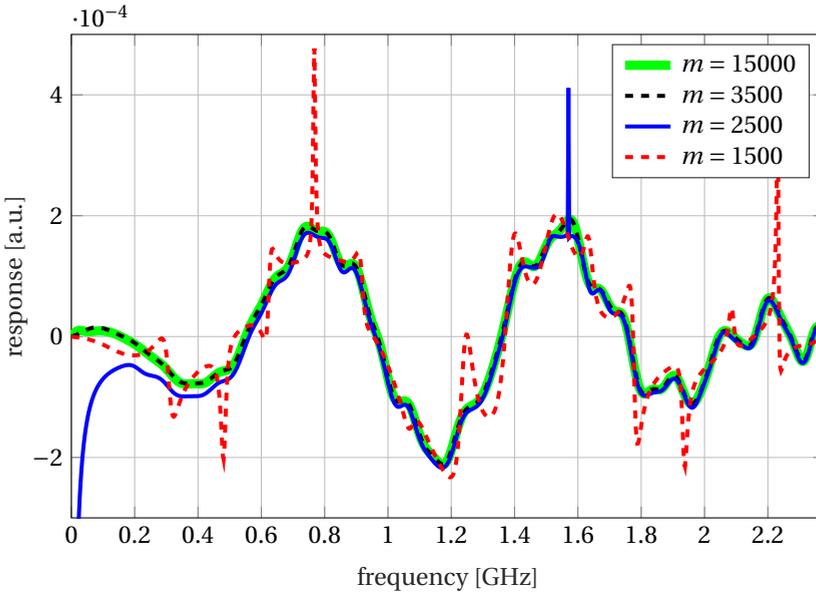


Figure 3.15: Imaginary part of the transfer-function over the frequency interval of interest computed with PKS after 1500, 2500, 3500 and 15000 iterations.

a second-order formulation.

The interpolation points of the RKS method are placed on the negative, real semi-axis and chosen asymptotically optimal assuming a uniform spectral distribution of the diffusion operator as described in [24]. Therefore, the shifts are known a priori and the dimension of the RKS can be increased by adding vectors one by one.

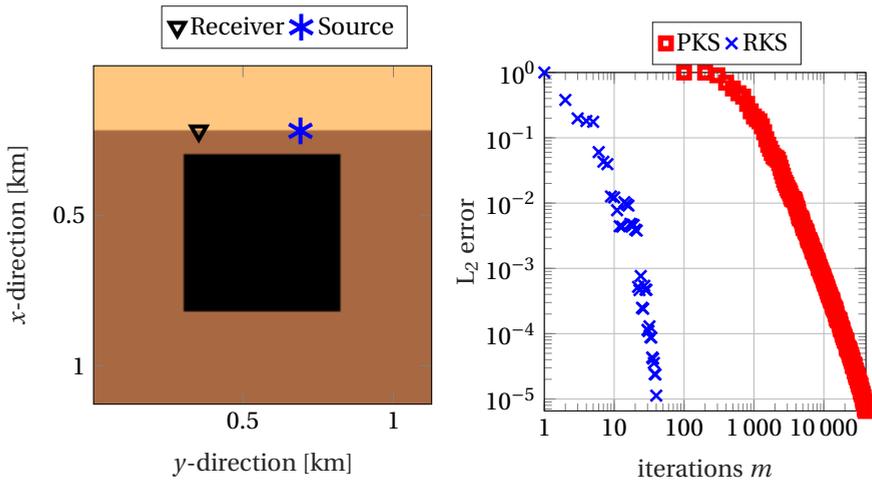
In Figure 3.16(b) the  $L_2$  error of the reduced-order model is shown as a function of the iteration number  $m$  for the PKS and RKS method. Both methods converge to a relative error of  $10^{-4}$  in  $m_{\text{PKS}} = 18700$  and  $m_{\text{RKS}} = 33$  iterations, respectively. To show the convergence within a single figure, double logarithmic axes are used. The RKS method clearly yields a smaller ROM.

The wall times for a sequential implementation are shown in Figure 3.16(b). The RKS method clearly outperforms the PKS method in terms of computation time and size of the ROM. However, the PKS uses less memory to construct the ROM. In section 3.3 the computational complexity of the reduced-order modeling methods was discussed in more detail. We choose to show the wall times in a semi-logarithmic and the  $L_2$  error as function of the iterations on a double-logarithmic scale as the asymptotic complexity of a single iteration of RKS is constant, and so is the complexity of PKS.

The time-domain transfer-function as approximated by the RKS and PKS method for the time interval of interest ( $t = 10^{-3} - 10^{-1}$  s) is shown in Figure 3.16(d). A slight difference at early times between these methods is caused by different averaging techniques employed in RKS and PKS for spatial discretization (see nodal homogenization and standard homogenization, respectively, from [56]). In this experiment, a Heaviside step function is used as a wavelet that is active for all negative times and zero for all posi-

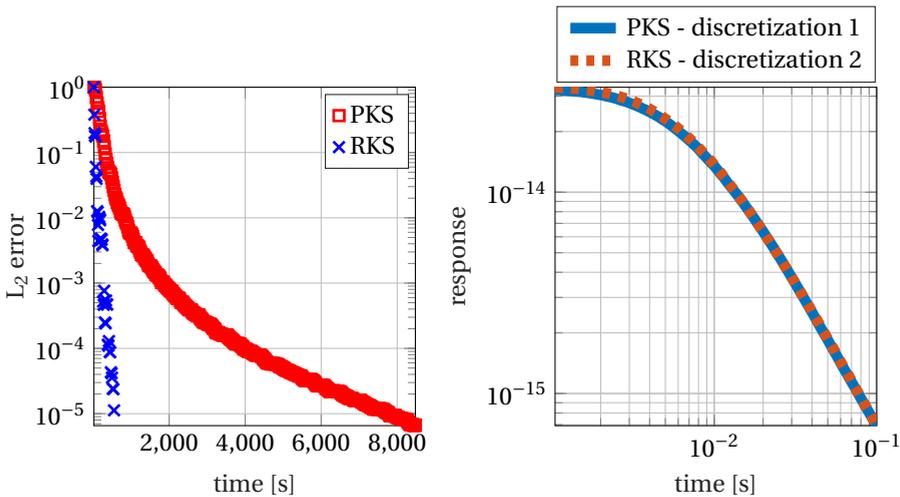


tive times (switch-off source). The observed  $L_2$  error between both methods for the time trace shown in Figure 3.16(d) is 3%. This 3D diffusive example shows that the convergence of PKS slows down as the error decreases, whereas the RKS shows linear convergence and RKS outperforms PKS in terms of computation time (see also Figure 3.16(b)).



(a) Cross section of the 3D model simulated with a lower half space of  $\sigma = 1$  S/m and a lossless upper half space. A resistive, anisotropic, cube anomaly with principal components  $\sigma_I = 0.01$  S/m,  $\sigma_{II} = 0.04$  S/m, and  $\sigma_{III} = 0.02$  S/m is present in the configuration. The dip and azimuth of the anisotropy are 30 and 15 degrees, respectively.

(b) Time-domain error of the RKS and PKS reduced-order models as a function of the iteration number  $m$  shown on a log-log scale.



(c) Time-domain error of the RKS and PKS reduced-order model for a given computation and time.

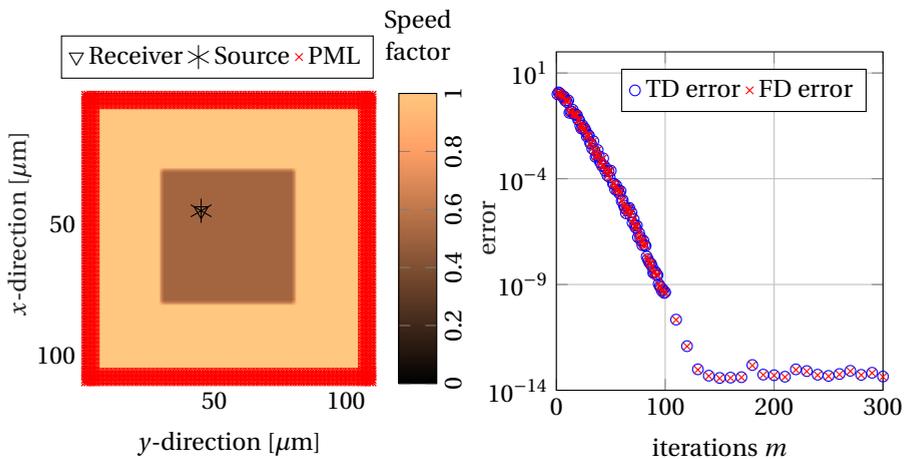
(d) Time-domain response for the converged RKS and PKS reduced-order models. To estimate the error introduced by finite-difference discretization the PKS and RKS method employ different homogenization methods.

Figure 3.16: Three-dimensional, anisotropic, electromagnetic diffusion example. The configuration under investigation (a), the convergence curve of an RKS and PKS method (b) is shown. The error is plotted against the computation times for the two methods (c) and the time-domain response of the system is depicted (d).



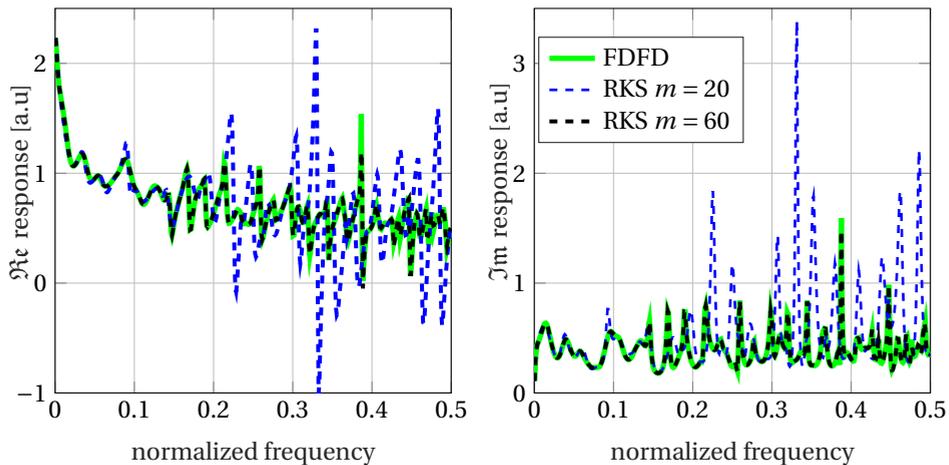
### 3.5.3. RESONANT BLOCK – STRUCTURE PRESERVING RKS

In this experiment we simulate the electrical field in a lossless two-dimensional dielectric box with permittivity  $\epsilon_r = 4$  excited by a line-source as depicted in Figure 3.17(a). Due to the two-dimensional and lossless nature of this example, the Helmholtz equation can be used to simulate this configuration in a second-order formulation. As discretization and normalization length we choose  $\Delta x = 1 \mu\text{m}$  and we excite the configuration using a Ricker wavelet, with a peak in the spectrum at  $\omega = 20 \cdot 10^{12} \text{rad/s}$  (normalized:  $\omega^n = 0.0667$ ). Since the source and receiver location coincide, we have Hermite interpolation of the transfer-function by the RKS method. To compute the rational Krylov subspace we use  $m$  equidistant shifts on the imaginary axis in the interval  $i \cdot s = [0, 0.5]$ , which results in  $m$  solves with the large system matrix of size  $N = 12 \cdot 10^3$ . The frequency-domain impulse response after  $m = 20$  and  $m = 60$  iterations is shown in Figure 3.18(a). The time- and frequency-domain (with wavelet) convergence is shown in Figure 3.17(b). For comparison, a Fourier method with 500 equidistant points in the interval  $i \cdot s = [0, 0.5]$  is used. Due to Parseval's identity, the time- and frequency-domain errors are equivalent. The method converges to machine precision after only  $m = 130$  iterations, meaning that the system of  $N = 12 \cdot 10^3$  unknowns can be reduced to  $2 \cdot 130$  without loss of accuracy. After  $m = 60$  the difference between the exact response and the reduced-order approximation is nearly indistinguishable by eye; only in the high frequency region the error is larger than the line-width. This convergence behavior is typical for resonant structures, where only a few eigenvalues and eigenvectors of the operator  $Q(s)$  need to be approximated to obtain a reasonable field description. The field response in the frequency-domain looks like a rational function due to the presence of resonances so that a rational Krylov subspace approximates the response well in a small subspace.



(a) Wavespeed factor in terms of  $c_0$  in the  $50 \mu\text{m} \times 50 \mu\text{m}$  dielectric box configuration. A dielectric box with wavespeed factor of 0.5 is embedded in an infinite space with wavespeed factor 1. (b) Time- and frequency-domain convergence of the RKS reduction method for the box configuration. A Ricker wavelet was used for excitation.

Figure 3.17: The box configuration and convergence in the time- and frequency-domain using the RKS method.



(a) Real part frequency-domain, reduced-order re- (b) Imaginary part frequency-domain, reduced-order  
 sponse with no wavelet applied after  $m=20$  and  $m=60$  response with no wavelet applied after  $m=20$  and  
 iterations.  $m=60$  iterations.

Figure 3.18: The real an imaginary part of the frequency-domain transfer-function of the box experiment.

### 3.6. CONCLUSIONS

In this chapter we have presented polynomial, extended, and rational Krylov reduced-order modeling techniques for wave equations on open domains. The model reduction techniques have been applied to dispersive, optical resonators to study the spontaneous decay rate of quantum emitters and ground penetrating radar configurations involving lossy media were considered to compare the convergence of all three model reduction methods. To illustrate the different convergence properties of the ROMs for diffusion and wavefield problems, an anisotropic, electromagnetic diffusion configuration was simulated. Finally, dielectric lossless box was considered to show that the RKS method can handle structures with many resonances and as a bridge to the next chapter, where a model order reduction technique is developed for configurations that are not dominated by resonances.

Polynomial and extended ROMs are based on a frequency-independent PML formulation in combination with a stability-correction procedure to obtain accurate transfer-function approximations between a given source and a given receiver. By exploiting the symmetry of the wave equations, we have shown that PKS- and EKS-ROMs can be constructed using short recurrence relations and it is not necessary to store a complete Krylov basis to evaluate transfer-function approximations for the fields.

The PKS reduced-order modeling techniques have been employed to compute the spontaneous decay rate of arbitrarily shaped 3D nano-sized resonators. By exploiting the symmetry of Maxwell equations in conjunction with a general second-order dispersion relation, reduced-order models for the spontaneous decay rate can be constructed very efficiently via a Lanczos-type reduction algorithm. The use of this algorithm allows



us to construct a single low-order model that is accurate on an entire spectral interval of interest and frequency sweeps can be performed at negligible cost given the low-order of the reduced-order models. Also the Lanczos-Ritz pairs can be used to study the quasi-normal modes of the resonators.

Large reduction factors can be achieved for general 3D resonators, since electromagnetic field responses in resonating structures are mainly dominated by a small number of quasi-normal modes. Moreover, discretizing 3D resonating nano-structures in space generally leads to heavily oversampled, semi-discrete Maxwell systems, since detailed subwavelength geometric features of the resonating structure need to be captured.

The small number of modes contributing to the electromagnetic response, suggests the possibility of solving this problem using rational Krylov subspaces. Rational Krylov subspaces generally show superior convergence for systems with only a few contributing eigenvalues. Therefore, future work should include the design of a rational Krylov subspace algorithm to compute the electromagnetic response of 3D resonators. In that case it is not necessary to work with a first-order form as the RKS method can handle nonlinearities with respect to frequency.

We point out that the developed approach can easily be extended to multipole dispersive media in order to simulate more complex dispersive materials. For instance a single resonator can be simulated with a dispersion relation given by the sum of a Drude and a Lorentz model. Furthermore, by storing the computed electric field strength values within the golden nanorod or by collecting the computed EM field at the boundary of the computational domain, we can determine the heat absorption or emission to the far field by invoking Poynting's theorem.

Constructing PKS- and EKS-ROMs can be realized at low computational cost, but the order of these models can be much larger than the order of RKS models. This makes PKS and EKS particularly well suited for online field computations. RKSs, on the other hand, generally yield the smallest models; however, for some applications and configurations the low cost of a single iteration of PKS or EKS outweighs this advantage. Furthermore, generating an RKS demands more memory as we have to save the basis. However, if the goal is to obtain a reduced-order model of the smallest order and this model can be computed offline, then RKS-ROMs may be preferred. Another advantage of an RKS approach over PKS- and EKS-ROM construction is that it does not require linearization of the PML and a stability-correction procedure is not necessary.

What all methods have in common is that the convergence rate of the models improves as the losses within the configuration increase. When solving wavefield problems on open domains, PMLs that simulate outward wave propagation, automatically introduce a loss mechanism into the system that moves the spectrum of the first-order Maxwell wave operator away from the imaginary axis, where the wavefield responses are approximated. If, in addition, material losses are present as well, the spectrum moves away even further into the complex frequency plane and this separation increases as the loss profiles of the different media increase. This separation has a positive effect on the convergence rate of the Krylov ROMs and lower order models are sufficient to accurately model the transfer-function responses between a source and a receiver. Physically, an increase in conductivity leads to a loss of information and smoother wave responses,

which can be more easily approximated, resulting in ROMs of a lower order. This was shown in this chapter using a GPR application where the convergence rate was studied as a function of the losses that are present in the configuration.

Finally, for the Maxwell equations in the diffusion approximation rational approximation of the transfer-function via RKS is clearly superior in computation time and resulting model order. The separation of the spectrum and the area of approximation are intrinsic to the problem in this case and optimal shifts for RKS can be found.

In summary, this chapter introduced model order reduction techniques for wave and diffusion problems and has identified several applications. It was demonstrated that reduced-order modeling significantly alleviates the computational burden in multiple applications ranging from geophysics to nano-photonics. The RKS method described in this chapter works well for wave equations if only a few singular Hankel values contribute to the system response as demonstrated by the last numerical experiment. It is now necessary to obtain an efficient model-reduction technique that can handle configurations in which this is no longer the case. The next chapter therefore incorporates asymptotic solutions into the model order reduction approach in order to handle configurations that are dominated by propagating waves.





## 4

# PHASE-PRECONDITIONING FOR COMPRESSION OF WAVEFIELDS

4

*Today's scientists have substituted mathematics for experiments, and they wander off through equation after equation, and eventually build a structure which has no relation to reality.*

Nikola Tesla

---

Parts of this chapter have been submitted for publication in the SIAM journal Multiscale Modeling & Simulation [D].

I thank Guillaume Leclerc for the help with the GPU benchmarking of the presented algorithm.



## 4.1. INTRODUCTION

**R**ATIONAL Krylov subspace (RKS) techniques are well-established and powerful tools for projection-based model reduction of time-invariant dynamic systems. For hyperbolic wavefield problems, such techniques perform well in configurations where only a few modes contribute to the field. RKS methods, however, are fundamentally limited by the Nyquist-Shannon sampling rate, making them unsuitable for the approximation of wavefields in configurations characterized by large travel times and propagation distances, since wavefield responses in such configurations are highly oscillatory in the frequency-domain. To overcome this limitation, we propose to precondition the RKSs by factoring out the rapidly varying frequency-domain field oscillations. The remaining amplitude functions are generally slowly varying functions of source position and spatial coordinate and allow for a significant compression of the approximation subspace. Our one-dimensional analysis together with numerical experiments for large-scale 2D acoustic models show superior approximation properties of preconditioned RKS compared with standard RKS model order reduction. The preconditioned RKS results in a reduction of the frequency sampling well below the Nyquist-Shannon rate, a weak dependence of the RKS size on the number of inputs and outputs for multiple-input/multiple-output (MIMO) problems, and, most importantly, in a significant coarsening of the finite-difference grid used to generate the RKS. A prototype implementation indicates that the preconditioned RKS algorithm is competitive in a modern high-performance computing environment.

## 4.2. MOTIVATION

The rational Krylov subspace approach is most efficient if only a few singular Hankel values of the system contribute to the solution as is the case for resonating structures with a few excited and observable modes [7]. A good example of such a structure is the dielectric box, the last example studied in the previous chapter in section 3.5.3. The frequency-domain response of such a configuration is well-described by a low-degree rational function and a rational Krylov technique will therefore quickly capture the desired wavefield response. For waves characterized by large travel times; however, this may no longer be the case, since such responses are highly oscillatory in the frequency-domain and sampling should at least take place at half the Nyquist-Shannon sampling rate. As an illustration, consider a source-receiver pair with an arrival at  $T^{\text{arr}}$  such that the source wavelet convolved with  $\delta(t - T^{\text{arr}})$  is measured. In the Laplace domain this translates to multiplication by  $\exp(-sT^{\text{arr}})$ , which means that according to the Nyquist sampling theorem the maximum frequency-domain sampling distance is  $\Delta s = \pi/T^{\text{arr}}$  on the imaginary axis. Clearly, the number of required frequency-domain samples increases as the travel time increases leading to prohibitory large rational Krylov subspaces. In this chapter we will incorporate travel time information to obtain basis functions that are less oscillatory to lower this sampling demand.

Numerical modeling of wave propagation in large domains is fundamental to many applications in design optimization and wavefield imaging. In the oil and gas industry, for instance, the solution of the Maxwell equations is required to invert electromagnetic measurements, while in seismic imaging the solution to the elastodynamic

wave equation is needed to ultimately image the subsurface of the Earth.

Finite-difference discretization of the governing wave equations leads to large-scale linear systems, whose solution is computationally intense. Imaging and optimization often use multiple frequencies, sources, and receivers, which leads to systems that need to be evaluated for multiple right-hand sides, time-steps or frequencies, depending on whether the problem is solved in the time- or frequency-domain. Therefore, these so-called multiple-input/multiple-output (MIMO) systems have a high demand on memory and computational power, causing long runtimes. To be more specific, let us consider a surface seismic imaging problem in a  $k$ -dimensional space ( $1 \leq k \leq 3$ ), with maximal propagation distance of  $N$  wavelengths. This would require the solution of a discretized system with  $O(N^k)$  state variables,  $O(N^{k-1})$  sources and receivers, and  $O(N)$  frequencies or time steps [57]. Model order reduction aims to reduce the complexity and computational burden of large-scale problems and here we target all three of these factors.

Recently, promising results were obtained in the time-domain via multiscale model reduction [16, 27]. The time-domain multiscale algorithms can be efficiently parallelized via domain-decomposition, but time stepping still needs to be carried out sequentially, while frequency-domain problems can be solved in parallel for different frequencies. Here we consider interpolatory projection-based model reduction in the frequency-domain, e.g., see [3]. The essence of this approach is the projection of the underlying system onto a rational Krylov subspace (originally introduced by Ruhe for eigenvalue computations [69]), which produces good quality, low-order approximations if the spectrum of the system is well separated from the frequency interval of interest, as in the case of diffusion PDEs, e.g., [5, 32, 50, 42]. In the context of wavefield modeling, such a separation of the clustered eigenvalues is introduced by losses present in the media or by the use of absorbing boundary conditions for the truncation of unbounded domains. Projection-based reduced-order models (ROMs) for wavefield problems may therefore exhibit fast convergence, especially for resonant configurations with few isolated resonant eigenmodes [9, 28, 29]. Some modifications of the RKS projection method can also be competitive for problems with smooth initial conditions leading to effective suppression of highly oscillatory eigenmodes [39].

Usually, the computational cost of projection-based reduced-order modeling is dominated by the generation of a suitable projection basis, e.g., see [7, 59]. In the interpolatory projection-based ROM the Helmholtz equation has to be solved at different frequencies (shifts) and the span of these solutions forms the RKS basis. The solution obtained from the Galerkin projection onto this subspace interpolates at the shifts, which are therefore also known as interpolation points. Moreover, for coinciding sources and receivers the transfer-function and its first derivative is interpolated at these points. A general drawback of an RKS approach is that the number of interpolation points can become large when wavefield solutions with large travel times or propagation distances are of interest. Such wavefields are highly oscillatory in the frequency-domain and the Nyquist-Shannon sampling theorem states that this oscillatory field should be sampled with at least one point per oscillation (two points per wavelength). Consequently, the number of interpolation points required to accurately represent the wavefield increases proportional to the propagation distance. Moreover, discretization grids in Helmholtz



solvers must also resolve wavefield oscillations. This requirement has an even more dramatic effect on the computational cost due to poor scalability of the available solvers. In favorable situations the best sampling rates approaching the Nyquist limit can be achieved with high-order spectral methods and their outgrowths. However, their cost per unknown can be significantly higher compared with less accurate low-order methods due to loss of sparsity. In this chapter, we show that the sampling demand can be significantly lowered by adding phase information to the model order reduction technique leading to phase-preconditioned RKSs (PPRKS). Preconditioning of Krylov subspaces for model reduction is a challenging and still open problem in general. However, to achieve it for particular applications one can try to incorporate the underlying physics and asymptotic analysis to arrive at PPRKS. Our approach is related to other known approaches in the field of oscillatory wave problem computation, such as preconditioners for Helmholtz solvers [34, 43], Filon quadrature [46], and a recent approach to data compression using phase-tracking [53].

4

In particular, we construct RKSs using polar decompositions of frequency-*dependent* basis functions. These decompositions consist of a product of *smooth* amplitude-functions and a known frequency-dependent *oscillatory* phase term. The phase term is determined from high-frequency asymptotic expansions such as the WKB approximation (Wentzel–Kramers–Brillouin) [6]. The amplitude-functions are computed by splitting the RKS into incoming and outgoing waves (by applying one-way wave operators) and factoring out the corresponding phase terms. Analogous to Filon quadrature, we handle the highly oscillatory phase-functions analytically and the smooth amplitude-function numerically. By developing a block version of phase-preconditioned RKS for MIMO problems, we are also able to factor out the main dependence of the RKS on the input (source) location. This feature, and the reduction of the number of interpolation points mentioned above, leads to a significant compression of the approximation space.

Finally, the resulting phase-preconditioned ROMs can also extrapolate to frequencies outside the interval of interpolation points, since the basis functions are frequency-dependent and the amplitude-functions are smooth for smoothly varying wavespeed profiles. This enables us to coarsen the second-order finite-difference grid used for the RKS generation.

In conclusion, with phase-preconditioned RKS we can effectively reduce all of the above mentioned factors contributing to the complexity of the MIMO wavefield problem. The overall goal is to approximate the transfer-functions from multiple sources to multiple receivers with a small reduced-order model that honors the physics of the underlying wave equation. The approach uses a coarse grid and low-frequency interpolation points to build an RKS and to obtain smooth amplitude-functions. Using high-frequency asymptotic expansions, this RKS is extrapolated to high frequencies and evaluated on a fine grid. The projection of a fine grid wave operator onto the extrapolated RKS gauges the ROM to the fine scale we intend to model. In this way, fine-scale wave scattering and large-scale wave propagation can be combined, which allows us to obtain a ROM valid for all time scales. RKS algorithms for wavefield problems are at a disadvantage compared with polynomial and extended Krylov subspace algorithms when it comes to computational memory consumption as the basis needs to be saved for RKS methods. The compression of the approximation space to a small number of amplitudes

and phases, however, leads to a reduction in the computational memory demand of the proposed method.

In section 4.3, we briefly recapitulate the RKS method for SISO problems. In section 4.4, we take the RKS approach as a starting point and introduce the phase-preconditioned RKS for one-dimensional SISO configurations. We show that phase preconditioning is structure-preserving and retains the interpolation properties of standard RKS Galerkin projection. The main result of this section is that for a piecewise constant wavespeed profile, the new method yields the exact solution with the number of interpolation points equal to the number of homogeneous layers, i.e., this number plays the same role as the problem dimensionality in a conventional RKS approach. Section 4.5 discusses the algorithm for higher spatial dimensions in a MIMO setting using a block version of phase-preconditioned RKS. Finite-difference implementation via a two-grid algorithm is discussed in section 4.6. In section 4.7 we illustrate the performance of the proposed RKS techniques through a number of two-dimensional numerical experiments. Section 4.8 discusses the implementation of the proposed method on parallel computation architectures and the conclusions can be found in section 4.9.

### 4.3. STRUCTURE-PRESERVING RKS

We start the development of the phase-preconditioning approach by briefly recapitulating the structure-preserving RKS method presented in section 3.2.1. The discussion of the phase-preconditioned RKS method is based on the Helmholtz equation in the weak form (see equation (2.5)),

$$p^H Q(s)u = -\frac{1}{v(x_S)^2} \bar{p}(x_S) = p^H b. \quad (4.1)$$

In a structure-preserving RKS approach, the weak solution of equation (4.1) is approximated by an element from the space  $\mathcal{K}_R^{2m}(\kappa)$ . The reduced-order solution is expanded as  $u_m = V_m z$  using the real basis  $V_m \in \mathbb{R}^{\infty \times 2m}$  of  $\mathcal{K}_R^{2m}(\kappa)$  with expansion coefficients  $\alpha_i$  collected in vector  $z = [\alpha_1, \dots, \alpha_{2m}]^T$ . These coefficients are obtained from a standard Galerkin procedure defined through the weak form of equation (2.5) leading to

$$z = [V_m^H Q(s) V_m]^{-1} V_m^H b \quad \text{or} \quad z = R_m^{-1}(s) V_m^H b \quad (4.2)$$

with  $b = -\delta(x - x_S)/v(x_S)^2$  and where  $R_m(s)$  is the  $2m \times 2m$  reduced-order operator given by  $R_m(s) = V_m^H Q(s) V_m$ . The reduced-order model is structure preserving as shown by the proposition in the previous chapter. Symmetry, the Schwarz reflection principle, and the numerical range are preserved as well.

We end this summary by introducing an alternative way of representing the reduced-order solution, which will be useful in the development of phase-preconditioned RKS methods. In expansion form the reduced-order solution can be written as

$$u_m = \sum_{i=1}^m \begin{bmatrix} d_i \\ \delta_i \end{bmatrix}^T \begin{bmatrix} u(s_i) \\ u(\bar{s}_i) \end{bmatrix}, \quad (4.3)$$

where the expansion coefficients  $d_i$  and  $\delta_i$  follow from the Galerkin condition and  $s_i$  are



the interpolation points collected in  $\kappa$ . Due to the mentioned linear independence<sup>1</sup> of the real and imaginary part of the snapshots, this representation is algebraically equivalent to  $u_m = V_m z$ , i.e. there exists a transform from the  $2m$  coefficients  $\alpha_i$  (collected in  $z$ ) to the coefficients  $d_i$  and  $\delta_i$  of equation (4.3).

#### 4.4. FIELD PARAMETRIZATION FOR SISO PROBLEMS

To enhance the convergence of an RKS approach for travel time dominated structures, we need to incorporate travel time information into the Krylov subspace, and thus into our basis functions. To this end, we assume that variations of the medium take place on a scale much larger than the wavelength at the considered frequencies, since this allows us to use a geometrical optics ansatz. Every basis vector belonging to the RKS is now split into an incoming and an outgoing wave and for each of these waves we factor out a strongly oscillating phase term  $\exp(\pm s T_{\text{eik}})$ , where  $T_{\text{eik}} = T_{\text{eik}}(x)$  is the eikonal time that solves the eikonal equation  $|\nabla T_{\text{eik}}(x)|^2 = \frac{1}{v(x)^2}$ . Splitting of the fields is realized using one-way wave equations. First we introduce this splitting for one-dimensional systems in section 4.4.1 and then generalize to higher dimensions in section 4.4.2.

##### 4.4.1. ONE-DIMENSIONAL FIELD PARAMETRIZATION

We decompose the field into an incoming and outgoing component by writing

$$u(s_j) = \exp(-s_j T_{\text{eik}}) c_{\text{out}}(s_j) + \exp(s_j T_{\text{eik}}) c_{\text{in}}(s_j). \quad (4.4)$$

For each component an oscillating phase term has been factored out and the amplitudes are determined from the single frequency snapshot solutions  $u(s_j)$  via one-way wave equations as

$$c_{\text{out}}(s_j) = \frac{v}{2s_j} \exp(s_j T_{\text{eik}}) \left( \frac{s_j}{v} u(s_j) - \frac{\partial}{\partial |x - x_S|} u(s_j) \right), \quad (4.5a)$$

and

$$c_{\text{in}}(s_j) = \frac{v}{2s_j} \exp(-s_j T_{\text{eik}}) \left( \frac{s_j}{v} u(s_j) + \frac{\partial}{\partial |x - x_S|} u(s_j) \right). \quad (4.5b)$$

In equation (4.5a), the incoming wave component of  $u(s_j)$  is filtered out leaving an outgoing component for which outgoing oscillations can be factored out. In equation (4.5b) the situation is reversed and the outgoing component of  $u(s_j)$  is filtered out. This decomposition is similar to the one chosen in [18] where Fourier integral operators are used for computation of wavefields. Finally, we note that using the above one-way wave equations for decomposition is equivalent to enforcing the condition

$$\exp(s_j T_{\text{eik}}) \frac{\partial}{\partial |x - x_S|} c_{\text{out}}(s_j) + \exp(-s_j T_{\text{eik}}) \frac{\partial}{\partial |x - x_S|} c_{\text{in}}(s_j) = 0, \quad (4.6)$$

<sup>1</sup>The real and imaginary parts of the snapshots  $u(s_i)$  spanning  $\mathcal{K}_R^{2m}(\kappa)$  are always linearly independent, since the eigenfunction expansion of the Dirac distribution appearing on the right-hand side of equation (2.2) has an infinite number of terms.

and the amplitudes  $c_{\text{out}}$  and  $c_{\text{in}}$  are spatially much smoother than the wavefield  $u$ , since the highly oscillatory phase term has been factored out.

Now to obtain a field approximation at frequency  $s$ , instead of projecting our operator onto single frequency solutions  $u(s_j)$ , we project it onto the phase-corrected basis functions  $\exp(-sT_{\text{eik}})c_{\text{out}}(s_j)$  and  $\exp(sT_{\text{eik}})c_{\text{in}}(s_j)$ . This is the central idea of our approach, which preserves the interpolation properties of the RKS. In particular, by introducing the phase-preconditioned subspace as

$$\mathcal{K}_{\text{EIK}}^{2m}(\kappa, s) = \text{span}\{\exp(-sT_{\text{eik}})c_{\text{out}}(s_1), \dots, \exp(-sT_{\text{eik}})c_{\text{out}}(s_m), \exp(sT_{\text{eik}})c_{\text{in}}(s_1), \dots, \exp(sT_{\text{eik}})c_{\text{in}}(s_m)\}, \quad (4.7)$$

and its symmetry-preserving real form

$$\mathcal{K}_{\text{EIK;R}}^{4m}(\kappa, s) = \text{span}\{\Re \mathcal{K}_{\text{EIK}}^{2m}(\kappa, s), \Im \mathcal{K}_{\text{EIK}}^{2m}(\kappa, s)\} \quad (4.8)$$

we can construct reduced-order models in the usual way, but now in terms of frequency-dependent basis functions. More precisely, let  $M \leq 4m$  be the dimension of  $\mathcal{K}_{\text{EIK;R}}^{4m}(\kappa, s)$  and let vectors  $v^{[1]}(s), v^{[2]}(s), \dots, v^{[M]}(s) \in \mathbb{R}^\infty$  form an orthonormal basis of  $\mathcal{K}_{\text{EIK;R}}^{4m}$ , then the field approximation drawn from this subspace can be written as

$$u_m(s) = \sum_{i=1}^M \alpha_i(s) v^{[i]}(s) \quad (4.9)$$

and the coefficients  $\alpha_i(s) \in \mathbb{C}$  can again be determined from the Galerkin condition. Note that  $m$  denotes the number of snapshots used to construct  $\mathcal{K}_{\text{EIK;R}}^{4m}$ , while  $M \leq 4m$  denotes the dimension of this subspace. The factor of 4 in the upper bound on  $M$  is due to splitting into incoming and outgoing fields, that can lead to a twice as large approximation subspace compared with unpreconditioned RKS with the same shifts. However, as we shall see in subsection 4.5.1, this increase can be circumvented in our implementation; the overall dimension of the preconditioned RKS is usually comparable to the dimension of a standard unpreconditioned RKS for the same accuracy, while using a smaller number of snapshots.

With  $V_{m;\text{EIK}}(s) \in \mathbb{R}^{\infty \times M}$  the real, orthonormal basis matrix of  $\mathcal{K}_{\text{EIK;R}}^{4m}(\kappa, s)$ , the reduced-order model that follows from the Galerkin condition can be written as a self-adjoint, time-invariant dynamic system

$$V_{m;\text{EIK}}(s)R_{m;\text{EIK}}(s)V_{m;\text{EIK}}^H(s)u_{m;\text{EIK}}(s) = b_m, \quad (4.10)$$

with

$$b_m = V_{m;\text{EIK}}(s)V_{m;\text{EIK}}^H(s)b \quad \text{and} \quad R_{m;\text{EIK}}(s) = V_{m;\text{EIK}}^H(s)Q(s)V_{m;\text{EIK}}(s).$$

The reduced-order model of equation (4.10) is the phase-corrected counterpart of the reduced-order model of equation (3.3). Furthermore, since  $\bar{c}_{\text{out/in}}(s_i) = c_{\text{out/in}}(\bar{s}_i)$  holds because the Schwarz reflection principle is satisfied, we can also express the reduced-order model of equation (4.9) in terms of the amplitude-functions  $c_{\text{in}}(s)$  and  $c_{\text{out}}(s)$  as (cf. equation (4.3))

$$u_m(s) = \sum_{i=1}^m \begin{bmatrix} a_i(s) \\ \alpha_i(s) \end{bmatrix}^T \begin{bmatrix} \exp(-sT_{\text{eik}})c_{\text{out}}(s_i) \\ \exp(-\bar{s}T_{\text{eik}})c_{\text{out}}(\bar{s}_i) \end{bmatrix} + \sum_{i=1}^m \begin{bmatrix} d_i(s) \\ \delta_i(s) \end{bmatrix}^T \begin{bmatrix} \exp(sT_{\text{eik}})c_{\text{in}}(s_i) \\ \exp(\bar{s}T_{\text{eik}})c_{\text{in}}(\bar{s}_i) \end{bmatrix} \quad (4.11)$$



with expansion coefficients  $a_i$ ,  $\alpha_i$ ,  $d_i$ , and  $\delta_i \in \mathbb{C}$  and where we have assumed that  $M = 4m$ . This formulation clearly shows that we use frequency-independent amplitudes, preconditioned by frequency-dependent phase-functions, employing conjugation to preserve the symmetry of the wave equation.

Finally, for field evaluations on the imaginary axis ( $s \in i\mathbb{R}$ ) the above expansion can be written more compactly as

$$u_m(s) = \sum_{i=1}^{2m} \begin{bmatrix} a_i(s) \\ \alpha_i(s) \end{bmatrix}^T \begin{bmatrix} \exp(-sT_{\text{eik}}) c(s_i) \\ \exp(sT_{\text{eik}}) c(-s_i) \end{bmatrix} \quad \text{with } s \in i\mathbb{R} \quad (4.12)$$

and where  $c(s_i) = c_{\text{out}}(s_i)$  for  $i = 1, 2, \dots, m$  and  $c(s_i) = \bar{c}_{\text{in}}(s_{i-m})$  for  $i = m+1, \dots, 2m$ .

The following results show that the PPRKS retains the structure-preserving, interpolatory-projection properties of standard RKS.

4

**LEMMA 4.1.** — The system of (4.10) is structure preserving, i.e.,  $W\{\mathcal{R}_{m;\text{EIK}}(s)\} \subseteq W\{Q(s)\}$  on the range (column space) of  $V_{m;\text{EIK}}(s)$ .

*Proof.* Let a nontrivial  $x_m$  be in the range of  $V_{m;\text{EIK}}(s)$ , that is  $x_m = V_{m;\text{EIK}}(s)y_m$ . Then

$$y_m^H \mathcal{R}_{m;\text{EIK}}(s)y_m = y_m^H V_{m;\text{EIK}}^H(s)Q(s)V_{m;\text{EIK}}(s)y_m = x_m^H Q(s)x_m.$$

□

Thus, phase-preconditioned reduced-order models can decrease the numerical range, as  $x_m$  lies in the range of  $V_{m;\text{EIK}}(s)$ ; however, the spectrum is always contained in the numerical range of  $Q(s)$  and does not increase.

**PROPOSITION 4.1.** — The SISO reduced-order transfer-function retains the interpolation properties of the unpreconditioned RKS with the same shifts stated in Proposition 3.3.

*Proof.* By construction,  $\mathcal{K}_{\text{EIK};R}^{4m}(\kappa, s) \supset \mathcal{K}_R^{2m}(\kappa)$  when  $s \in \kappa \cup \bar{\kappa}$ . According to Lemma 4.1, the ROM is passive given that the Galerkin problem has a unique solution. Therefore, the proof of Proposition 3.3 applies. □

One of the motivations to use this method is the expected fast convergence, when the parametrization of 4.4 is valid. In that case only a few phase-corrected, smooth amplitude-functions  $c_{\text{out/in}}$  are required to approximate the wavefield. Furthermore, in the RKS method discussed in the previous section, the number of required shifts or frequencies is dependent on the largest arrival time; however, in the phase-preconditioned RKS (PPRKS) discussed above, the arrival times are factored out and the number of shifts is dependent on the complexity of the wavespeed model  $v(x)$  rather than the largest arrival time. We make this explicit in the following proposition.

**PROPOSITION 4.2.** — Let a 1D problem have  $\ell$  homogeneous layers. Then there exist  $m \leq \ell + 1$  non-coinciding interpolation points, such that  $u_{m;\text{EIK}}(s) = u$ .

*Proof.* We start by noting that if the regions to the left and to the right of the source are considered as separate layers, then the solution to the one-dimensional wave equation consists of a superposition of left- and right-going waves of constant amplitude in each of the  $\ell + 1$  layers. For one-dimensional problems, the decomposition direction coincides with the travel direction of the wave; thus,  $c_{\text{out}}(\kappa)$  and  $c_{\text{in}}(\kappa)$  are piecewise constant. A piecewise-constant function with  $\ell + 1$  layers can be exactly represented by at most  $\ell + 1$  linear-independent piecewise-constant functions with the same jump locations. Let us prove from the opposite, i.e., assume, that there are no  $m \leq \ell + 1$  non-coinciding shifts  $\kappa_i$  such that  $c_{\text{out}}(\kappa_i)$  form a basis for all possible  $c_{\text{out}}(\kappa)$ . Then the number of shifts  $m$  yielding linear independent solutions should be less than  $\ell + 1$ . But by assumption there must be at least a single  $c_{\text{out}}(\kappa)$  not from the subspace. Then one can add this solution to the subspace, i.e., the true number of linearly-independent solutions is  $m + 1$ , which contradicts the assumption that this number is  $m$ . Analogously, we can prove the same statement for  $c_{\text{in}}(\kappa_i)$ .

In conclusion, there exist  $m \leq \ell + 1$  non-coinciding shifts  $\kappa_i$  such that  $c_{\text{out}}(\kappa_i)$  and  $c_{\text{in}}(\kappa_i)$  form respective bases for all possible  $c_{\text{out}}(\kappa)$  and  $c_{\text{in}}(\kappa)$ , i.e., the exact solution  $u$  will be in the projection subspace. Finally, due to Lemma 4.1 the exact solution  $u$  will be the unique solution of the Galerkin problem.  $\square$

The proposition can be extended to almost all arbitrary  $\ell + 1$  interpolation points, as the interpolation points that lead to  $\ell + 1$  linear-dependent functions have measure zero. Thus, phase-preconditioning allows us to obtain the exact solution with the number of interpolation points equal to the number of homogeneous layers, i.e., this number plays the same role as the problem dimensionality in a conventional RKS approach.

**REMARK 10.** — *Illustration of the Proposition 4.2*—We can illustrate Proposition 4.2 using a simple one-dimensional configuration with a slab inclusion. The domain is terminated with a PML on the right-hand side and with a reflecting Neumann boundary on the left-hand side. The slab inclusion has a wavespeed of 0.5 and sits between  $L_1$  and  $L_2$ . The medium, field and the field amplitudes are illustrated in Figure 4.1. There is no incoming field from the PML and the corresponding amplitude therefore vanishes between  $L_2$  and the PML. There now exist three shifts so that the incoming and outgoing amplitudes form a basis for functions that are piecewise constant in the corresponding layers, with the exception that the outgoing amplitude in the last layer will always be zero.

**REMARK 11.** — *Interpretation as Filon Quadrature*—Filon quadrature deals with the evaluation of oscillatory integrands. Using a standard quadrature to evaluate

$$y(s) = \int \exp(st) f(t) dt \quad (4.13)$$

with a smooth function  $f(t)$  that varies slowly compared to  $\exp(st)$  leads to

$$y(s) = \Delta t \sum_n a_n \exp(sn\Delta t) f(n\Delta t) \quad (4.14)$$

with quadrature weights  $a_n$  which requires  $s\Delta t < \pi$ . In Filon quadrature one makes the weights  $a_n$  dependent on  $s$  to arrive at quadrature formulas whose convergence depends



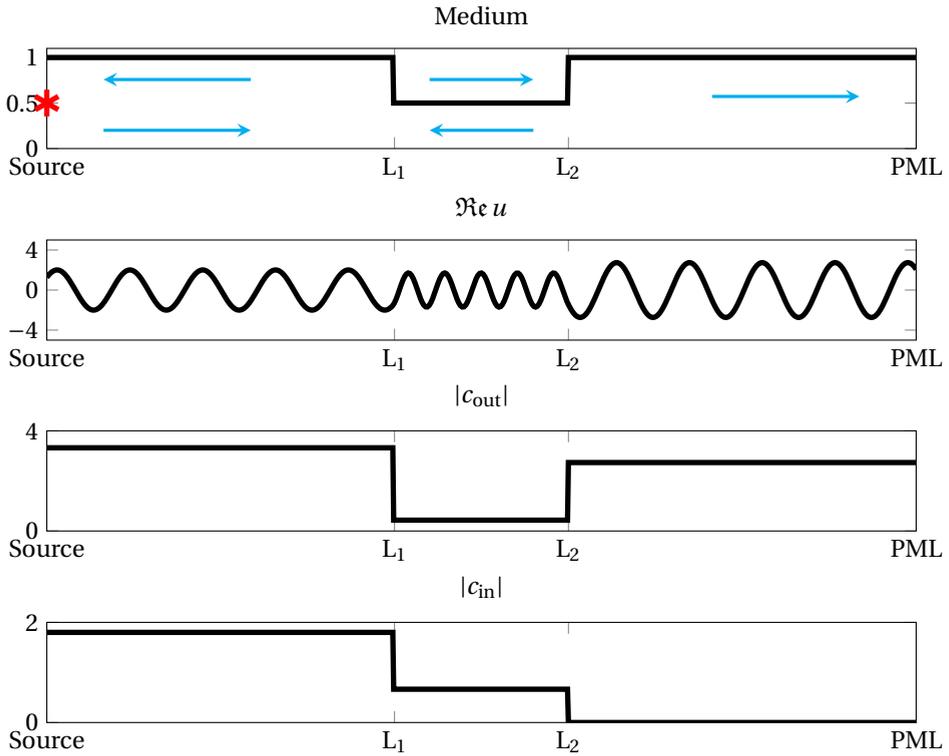


Figure 4.1: Decomposition of a wavefield in incoming and outgoing amplitudes. The medium consists of three homogenous layers.

merely on the properties of  $f(t)$ . Similarly, in phase-preconditioned rational Krylov we try to approximate an oscillatory frequency-domain transfer-function. By making our projection basis frequency-dependent we enhance the convergence of the reduced-order model.

#### 4.4.2. GENERALIZATION TO HIGHER DIMENSIONS

In higher spatial dimensions we again split the field in incoming and outgoing wave components and use a function  $g(z)$  to factor out a strongly varying phase. Specifically, for two- and three-dimensional problems we write

$$u(s_j) = g(s_j T_{\text{eik}}) c_{\text{out}}(s_j) + g(-s_j T_{\text{eik}}) c_{\text{in}}(s_j), \quad (4.15)$$

and project the problem onto the real and imaginary parts of the phase-preconditioned subspace

$$\begin{aligned} \mathcal{K}_{\text{EIK}}^{2m}(\kappa, s) = \text{span}\{ & g(s T_{\text{eik}}) c_{\text{out}}(s_1), \dots, g(s T_{\text{eik}}) c_{\text{out}}(s_m), \\ & g(-s T_{\text{eik}}) c_{\text{in}}(s_1), \dots, g(-s T_{\text{eik}}) c_{\text{in}}(s_m)\}, \end{aligned} \quad (4.16)$$

where  $g(z) = \exp(-z)/z$  in 3D, while  $g(z) = K_0(z)$  in 2D, with  $K_0$  the modified Bessel function of the second kind and order zero. The singular behavior of the field at the source location is factored out leading to a weaker dependence of the amplitude-functions  $c_{\text{out/in}}$  on the source location.

In higher spatial dimensions, the field amplitudes are again obtained via one-way wave equations, but this time along the eikonal rays leading to decomposition directions  $\pm \nabla T_{\text{eik}}$ . For 2D applications with  $\mathcal{K}_0(z)$  as incoming and  $\mathcal{K}_0(-z)$  as outgoing, we obtain the amplitude-functions

$$c_{\text{out}}(s_j) = \frac{s_j T}{\text{sign}(\Im(s_j)) i\pi} \left[ \mathcal{K}_1(-s_j T) u(s_j) - \mathcal{K}_0(-s_j T) \frac{v^2}{s_j} \nabla T \cdot \nabla u(s_j) \right] \quad (4.17)$$

and

$$c_{\text{in}}(s_j) = \frac{s_j T}{\text{sign}(\Im(s_j)) i\pi} \left[ \mathcal{K}_1(s_j T) u(s_j) + \mathcal{K}_0(s_j T) \frac{v^2}{s_j} \nabla T \cdot \nabla u(s_j) \right], \quad (4.18)$$

as shown in the appendix A.1. Analogous to the one-dimensional reduced-order solution of equation (4.11), we can write the reduced-order solution in higher spatial dimensions as

$$u_m = \sum_{i=1}^m \begin{bmatrix} a_i(s) \\ \alpha_i(s) \end{bmatrix}^T \left[ g(s T_{\text{eik}}) c_{\text{out}}(s_i) \right] + \sum_{i=1}^m \begin{bmatrix} d_i(s) \\ \delta_i(s) \end{bmatrix}^T \left[ g(-s T_{\text{eik}}) c_{\text{in}}(s_i) \right], \quad (4.19)$$

where the coefficients follow from the Galerkin condition.

Lemma 4.1 and Proposition 4.1 can be straightforwardly extended to the multi-dimensional case. However, Proposition 4.2 is not directly extendable to higher dimensions. As opposed to the one-dimensional case, a decomposition direction does not necessarily coincide with the travel direction of the wave and the field parametrization may be poor in such cases. This problem can be resolved, however, by considering MIMO



wavefield systems with multiple sources and receivers, since in this case we have a decomposition direction for each source and the span of these directions may properly capture the propagation direction of the waves. In the next section, we therefore focus on wavefield systems with multiple sources and multiple receivers. The problem may be additionally complicated by multivalued solutions of the eikonal equation. In most situations it is sufficient to use the rays corresponding to the minimal travel time; however, as we shall see for the case of internal resonant structures in section 4.7.3, it can be beneficial to split the subspace along multiple rays.

#### 4.5. PHASE-PRECONDITIONING FOR MIMO SYSTEMS

To introduce the reduced-order transfer-function, we define a block rational Krylov subspace

$$\mathcal{K}_B^{mN_{\text{src}}}(\kappa) = \text{span}\{U_s(s_1), U_s(s_2), \dots, U_s(s_m)\} \quad (4.20)$$

and its real counterpart containing  $\mathcal{K}_B^{mN_{\text{src}}}(\kappa)$  and  $\mathcal{K}_B^{mN_{\text{src}}}(\bar{\kappa})$  given by

$$\mathcal{K}_{B;R}^{2mN_{\text{src}}}(\kappa) = \text{span}\left\{\Re \mathcal{K}_B^{mN_{\text{src}}}(\kappa), \Im \mathcal{K}_B^{mN_{\text{src}}}(\kappa)\right\}. \quad (4.21)$$

The reduced-order model for the fields can now be constructed completely analogous to the SISO case. Specifically, with  $V_m$  a basis array that spans  $\mathcal{K}_{B;R}^{2mN_{\text{src}}}(\kappa)$ , we have

$$U_{s;m}(s) = V_m R_m(s)^{-1} V_m^H B_s \quad \text{with} \quad R_m(s) = V_m^H Q(s) V_m. \quad (4.22)$$

The reduced-order transfer-function now follows as

$$F_m(s) = B_s^H U_{s;m}(s), \quad (4.23)$$

and it is straightforward to show that the MIMO reduced-order transfer-function  $F_m(s)$  is a Hermite interpolant of the MIMO transfer-function  $F(s)$ . The proof of this statement is completely analogous to the proof of Proposition 3.3.

To formulate the phase-corrected extensions of the block-RKS method, we note that the block-RKS field approximation  $u_m^{[l]}(s)$  due to a source  $l$  can be written as

$$u_m^{[l]}(s) = \sum_{r=1}^{N_{\text{src}}} \sum_{i=1}^m \left[ \begin{array}{c} a_i^{[l]} \\ \alpha_i \end{array} \right]^T \left[ \begin{array}{c} u^{[r]}(s_i) \\ u^{[r]}(\bar{s}_i) \end{array} \right], \quad (4.24)$$

with  $s_i \in \kappa$ . In other words, the field approximation  $u_m^{[l]}(s)$  due to source  $l$  is a linear combination of single frequency solutions from *all* sources. A straightforward generalization of phase-preconditioning to MIMO systems is to use a field approximation  $u_m^{[l]}(s)$  that is a linear combination of phase-corrected incoming and outgoing fields from *all* sources. We write the field approximation as

$$u_m^{[l]}(s) = \sum_{r=1}^{N_{\text{src}}} \left( \sum_{j=1}^m \left[ \begin{array}{c} a_{rj}^{[l]} \\ \alpha_{rj}^{[l]} \end{array} \right]^T \left[ \begin{array}{c} g(sT_{\text{eik}}^{[r]})c_{\text{out}}^{[r]}(s_j) \\ g(\bar{s}T_{\text{eik}}^{[r]})c_{\text{out}}^{[r]}(\bar{s}_j) \end{array} \right] + \sum_{j=1}^m \left[ \begin{array}{c} a_{rj}^{[l]} \\ \delta_{rj}^{[l]} \end{array} \right]^T \left[ \begin{array}{c} g(-sT_{\text{eik}}^{[r]})c_{\text{in}}^{[r]}(s_j) \\ g(-\bar{s}T_{\text{eik}}^{[r]})c_{\text{in}}^{[r]}(\bar{s}_j) \end{array} \right] \right), \quad (4.25)$$

where  $T_{\text{eik}}^{[r]}$  is the eikonal solution corresponding to the  $r$ th source. The coefficients  $a_{rj}^{[l]}$ ,  $\alpha_{rj}^{[l]}$  and  $d_{rj}^{[l]}$ ,  $\delta_{rj}^{[l]}$  are found via the block-Galerkin condition. For  $N_{\text{src}} > 1$ , this approach accounts for multi-directional scattering by representing the field as a linear combination of phase-corrected functions with multiple directions  $\nabla T_{\text{eik}}^{[r]}$ .

The idea that we followed to justify the use of block Krylov methods is that the field caused by one source contains information about the field caused by a source with a different location (and frequency). In the context of phase-preconditioning, we can apply this idea a second time to obtain a block-preconditioned algorithm. This means that instead of using the phase-correction function  $g(T_{\text{eik}}^{[r]})$  to only correct  $c_{\text{out}}^{[r]}(s_j)$  for each source location  $r$  separately, we cross combine all phase-functions  $g(sT_{\text{eik}}^{[r_2]})$  with all amplitudes  $c_{\text{out}}^{[r_1]}(s_j)$  as indicated by the summations over  $r_1$  and  $r_2$ . This leads to a field approximation  $u_m^{[l]}(s)$  due to the  $l$ th source given by

$$u_m^{[l]}(s) = \sum_{r_2=1}^{N_{\text{src}}} \sum_{r_1=1}^{N_{\text{src}}} \left( \sum_{j=1}^m \begin{bmatrix} a_{r_1 r_2 j}^{[l]} \\ \alpha_{r_1 r_2 j}^{[l]} \end{bmatrix} \right)^T \left[ \begin{array}{c} g(sT_{\text{eik}}^{[r_2]})c_{\text{out}}^{[r_1]}(s_j) \\ g(\bar{s}T_{\text{eik}}^{[r_2]})c_{\text{out}}^{[r_1]}(\bar{s}j) \end{array} \right] + \sum_{j=1}^m \begin{bmatrix} d_{r_1 r_2 j}^{[l]} \\ \delta_{r_1 r_2 j}^{[l]} \end{bmatrix} \left[ \begin{array}{c} g(-sT_{\text{eik}}^{[r_2]})c_{\text{in}}^{[r_1]}(s_j) \\ g(-\bar{s}T_{\text{eik}}^{[r_2]})c_{\text{in}}^{[r_1]}(\bar{s}j) \end{array} \right]. \quad (4.26)$$

The expansion coefficients are found from the block-Galerkin condition. Basis vectors in this expression can be viewed as a tensor-product of the amplitudes  $c_{\text{out}}^{[r_1]}$  and the phase terms  $g(-\bar{s}T_{\text{eik}}^{[r_2]})$ , while the Hadamard product is used spatially.

Similarly to the multidimensional SISO case, Lemma 4.1 and Proposition 4.1 can be straightforwardly extended for the multidimensional MIMO case. Our experiments presented in section 4.7 also indicate that the number of interpolation points needed for multidimensional MIMO configurations is dependent on the complexity of the wave-speed model and not the largest travel time (as proven for the one-dimensional case in Proposition 4.2). For the phase-preconditioned expansions it is beneficial to use a block formulation and incorporate multiple sources, since this allows us to capture a large number of travel directions. Further, amplitudes originating from one source can be used to construct meaningful field approximations for another source. This stands in contrast to ordinary RKS, where using block methods may not necessarily lead to a drastic increase in convergence as fields solutions from different sources are usually close to orthogonal.

#### 4.5.1. SVD TRUNCATION OF THE EXPANSION AMPLITUDES

In the case of many sources and receivers it is possible to compress the field amplitudes  $c_{\text{out}}^{[r_1]}(s_j)$  and  $c_{\text{in}}^{[r_1]}(s_j)$  a posteriori using a thin singular value decomposition. Using this SVD we compress the reduced-order model and remove redundancy in the expansion of equation (4.26). In the phase-preconditioned approach redundancy of the basis occurs in two ways. First, smooth amplitude-functions corresponding to different frequencies can be close to linear dependent due to the frequency-dependent basis vectors in (4.26). Second, the amplitude-functions of multiple sources can be similar which leads to redundancy once we use cross combinations of amplitudes and phase-functions as in equation (4.26). The original block Krylov basis does not have these redundancies.

To realize SVD truncation, the amplitudes  $c_{\text{in}}^{[r_1]}(s_j)$  and  $c_{\text{out}}^{[r_1]}(s_j)$  are first normalized in pairs to have unit Euclidean norm. To be more specific,  $c_{\text{in}}^{[r_1]}(s_j)$  and  $c_{\text{out}}^{[r_1]}(s_j)$  are



normalized by

$$\sqrt{\|c_{\text{in}}^{[r_1]}(s_j)\|^2 + \|c_{\text{out}}^{[r_1]}(s_j)\|^2}$$

such that the sum of the squared singular values is  $2mN_{\text{src}}$ . In this way the ratio between the incoming and outgoing amplitude is preserved. The SVD of the  $2mN_{\text{src}}$  incoming and outgoing amplitudes is then computed separately and truncated after  $M_{\text{SVD}}^{\text{out}}$  and  $M_{\text{SVD}}^{\text{in}}$  left singular vectors to obtain the compressed amplitudes  $c_{\text{in};\text{SVD}}^j$  and  $c_{\text{out};\text{SVD}}^j$ . The original amplitudes  $c_{\text{in}}^{[r_1]}(s_j)$  and  $c_{\text{out}}^{[r_1]}(s_j)$  are associated with a specific frequency  $s_j$  and source  $r_1$ , whereas the compressed amplitudes  $c_{\text{in};\text{SVD}}^j$  and  $c_{\text{out};\text{SVD}}^j$  are associated with a singular value. The amplitudes are therefore no longer associated with a source or frequency and the corresponding subscripts are dropped and replaced by the singular value index  $j$ . The resulting reduced-order solution expressed in terms of these compressed amplitudes is given by

$$u_m^{[l]}(s) = \sum_{r=1}^{N_{\text{src}}} \left( \sum_{j=1}^{M_{\text{SVD}}^{\text{out}}} \begin{bmatrix} a_{rj}^{[l]} \\ \alpha_{rj}^{[l]} \end{bmatrix}^T \begin{bmatrix} g(sT_{\text{eik}}^{[r]})c_{\text{out};\text{SVD}}^j \\ g(\bar{s}T_{\text{eik}}^{[r]})\bar{c}_{\text{out};\text{SVD}}^j \end{bmatrix} + \sum_{j=1}^{M_{\text{SVD}}^{\text{in}}} \begin{bmatrix} d_{rj}^{[l]} \\ \delta_{rj}^{[l]} \end{bmatrix}^T \begin{bmatrix} g(-sT_{\text{eik}}^{[r]})c_{\text{in};\text{SVD}}^j \\ g(-\bar{s}T_{\text{eik}}^{[r]})\bar{c}_{\text{in};\text{SVD}}^j \end{bmatrix} \right), \quad (4.27)$$

where  $M_{\text{SVD}}^{\text{out/in}} \ll mN_{\text{src}}$ .

If we contract the outgoing amplitudes and the conjugate of the incoming amplitudes into one amplitude basis  $c_{\text{SVD}}^j$ , compute the SVD after pairwise normalization and evaluate on the imaginary line ( $s \in i\mathbb{R}$ ) we can expand the field as

$$u_m^{[l]}(s) = \sum_{r=1}^{N_{\text{src}}} \sum_{j=1}^{M_{\text{SVD}}} \begin{bmatrix} a_{rj}^{[l]} \\ \alpha_{rj}^{[l]} \end{bmatrix}^T \begin{bmatrix} g(sT_{\text{eik}}^{[r]})c_{\text{SVD}}^j \\ g(-sT_{\text{eik}}^{[r]})\bar{c}_{\text{SVD}}^j \end{bmatrix} \quad \text{with } s \in i\mathbb{R} \quad (4.28)$$

where  $M_{\text{SVD}} \ll 2mN_{\text{src}}$ . Here the  $c_{\text{SVD}}^j$  are first  $M_{\text{SVD}}$  left singular vectors of

$$[\bar{c}_{\text{in}}^{[r_1]} \ c_{\text{out}}^{[r_1]}] / \sqrt{\|c_{\text{in}}^{[r_1]}\|^2 + \|c_{\text{out}}^{[r_1]}\|^2}.$$

In our numerical experiments we show that the singular values of the contracted amplitudes  $c_{\text{SVD}}^j$  decay much faster than the ones of the block Krylov basis. In Figure 4.5 we plotted the decay of singular values of a matrix with pairwise normalized vectors  $c_{\text{SVD}}^j$  and normalized vectors  $u^{[r_1]}(s_j)$  as columns. Each SVD trace is normalized to the largest singular value to emphasize the decay. The singular values associated with  $c_{\text{SVD}}^j$  show a strong decay with a plateau at the level of the finite-difference error, whereas the singular values associated with the wavefield solutions  $u^{[r_1]}(s_j)$  barely show any decay before reaching the Nyquist sampling rate.

The compressibility of the amplitudes confirms that the chosen parametrization of the wavefield is valid. Phase-preconditioning enhances the convergence of RKS not by increasing the subspace but by preconditioning a small basis of problem specific amplitudes in dependence of the evaluation frequency using phase-functions. In our experiments we show that the number of contributing amplitudes is only weakly dependent on the number of sources.

The compression of the amplitudes offers several advantages. First, it significantly reduces the cost of evaluating the reduced-order model, since it reduces the number of inner products that need to be computed to obtain the reduced-order operator. Second, the cost associated with communicating and storing the reduced-order model is reduced as well. The compressed amplitude basis is only very weakly dependent on the source locations used to construct it. Therefore, we can reduce the number of sources (right-hand sides) for which the basis vectors need to be computed, since their response can be approximated from their eikonal travel time and the basis vectors computed from other sources. It is in line with our effort to reduce the computation at every stage of our algorithm.

## 4.6. DISCRETE FORMULATION

In this section we consider the discrete implementation of the introduced reduced-order modeling technique. Discretization and selection of the grid accuracy are addressed first, followed by a discussion on how we handle numerical dispersion.

### 4.6.1. FINITE-DIFFERENCE DISCRETIZATION

Our basic approach is to solve equation (2.5) (restated here as equation (4.29)) using phase-corrected single frequency solutions as expansion functions and to obtain a reduced-order solution for a complete spectral interval of interest via the Galerkin condition.

We consider a rectangular domain  $\Omega \in \mathbb{R}^k$  with constant  $v(x)$  on  $\mathbb{R}^k \setminus \Omega$  and discretize the equation

$$\int_{\Omega} \nabla \bar{p} \cdot \nabla u d\Omega - \int_{\Omega} \bar{p} \frac{s^2}{v^2} u d\Omega + \int_{\partial\Omega} \bar{p} D(s) u d\partial\Omega = -\frac{1}{v(x_S)^2} \bar{p}(x_S), \quad (4.29)$$

using finite-differences to obtain a linear shifted system with a matrix nonlinearly depending on  $s$ . Discretization of the first two terms in the above equation using a second-order accurate finite-difference scheme with constant step sizes is straightforward. To discretize the third term, we approximate the DtN map  $D(s)$  using nearly-optimal discrete perfectly-matched layers (PML) according to [20]. The optimal Zolotarev rational approximants used for the PML construction make the size of the finite-difference problem (necessary for accurate approximation of  $D(s)$ ) in  $\mathbb{R}^k \setminus \Omega$  negligible compared to the grid in  $\Omega$ 's interior. In 2D for instance, the resulting equations that need to be solved in the PML can be solved efficiently with a block-cyclic solver [73], or with a band solver after sorting the PML system to a bandwidth of  $2k + 1$  (in 2D). This discretization leads to the matrix equation

$$\mathbf{p}^H Q(s) \mathbf{u} = \mathbf{p}^H \mathbf{b}, \quad (4.30)$$

with  $\mathbf{b}$  the discrete approximation of the scaled delta function. The matrix  $Q(s)$  of order  $N$  inherits all properties of the continuous operator  $Q(s)$  and thus follows the Schwarz reflection principle, is symmetric in a bi-linear form ( $Q(\bar{s})$  is the adjoint of  $Q(s)$  in the Hermitian inner product) and has a nonlinear numerical range in the left-half of the complex  $s$ -plane. The single frequency solutions  $u(s_j)$  needed to build the rational Krylov subspace can be obtained using iterative solvers or Gaussian elimination. The eikonal



equation  $|\nabla T|^2 = \frac{1}{v^2}$  is solved on the same grid using a fast marching method [75]. The associated computational cost is negligible with respect to the computational cost of solving the Helmholtz equation.

#### 4.6.2. REALIZATION ON TWO GRIDS

For smooth media, the amplitudes  $c_{\text{out}}^{[r_1]}$  and  $c_{\text{in}}^{[r_1]}$  are smooth functions of the spatial coordinates, since the highly oscillatory part of the frequency-domain wavefield is factored out together with the source singularity. Therefore, linear combinations of these amplitudes can form reasonable approximations of amplitude distributions at higher frequencies. Consequently, the phase-corrected reduced-order models can extrapolate to frequencies outside the convex hull of the interpolation points. Building a reduced-order model that extrapolates to higher frequencies has the advantage that the amplitudes can be determined on grids that are much coarser than grids required by a direct method at these high frequencies. This significantly reduces the computational cost of solving the Helmholtz equation to obtain the amplitude-functions  $c_{\text{out/in}}$ , since the main cost of the algorithm is associated with solving shifted systems.

To be specific, let  $Q_{\text{fine}}(s)$  and  $Q_{\text{coarse}}(s)$  denote the matrix operators obtained by discretizing the wave operator  $Q(s)$  on a fine and coarse grid, respectively. Writing the transfer-function and field approximations obtained with these fine and coarse grid operators as  $F_m(s)$  and  $U_m(s)$  and  $F_{c;m}(s)$  and  $U_{c;m}(s)$ , respectively, we have

$$F_m(s) - F_{c;m}(s) = [U_m(s) - P_{\text{interp}}U_{c;m}(s)]^H Q_{\text{fine}}(s) [U_m(s) - P_{\text{interp}}U_{c;m}(s)], \quad (4.31)$$

which is essentially a block version of equation (3.16).  $U_m$  and  $U_{c;m}$  don't live on the same grid and the interpolation matrix  $P_{\text{interp}} \in \mathbb{R}^{N_{\text{fine}} \times N_{\text{coarse}}}$  makes the above equation consistent. In this case, however,  $U_m(s) - P_{\text{interp}}U_{c;m}(s)$  signifies the difference between the fine and coarse grid field solutions and the interpolation property as presented in Proposition 3.3 obviously does not hold here. In other words, using a coarse grid for construction and a fine grid for projection leads to a loss of the interpolation property of the reduced-order model. On the other hand, we do increase the accuracy of the coarse transfer-function as the errors introduced by interpolation and the coarse grid solution get squared at the interpolation points  $s \in \kappa \cup \bar{\kappa}$ .

The main drawback of using coarser grids is that the numerical dispersion error increases and the analytic phase term  $\exp(\pm s T_{\text{eik}})$  does not match the phase term of  $u(s)$  for large imaginary shifts. Fortunately, we can correct for this phase mismatch. To be precise, in the analytic case the phase term  $\exp(-s T_{\text{eik}})$  is used to cancel the high-frequency dominant term  $-s^2/v^2$  in the wave equation. To guarantee that this cancellation takes place in the discrete case and to match the discrete and analytic phase, we introduce the discrete, second-order, finite-difference gradient matrix  $D_{x_i}$  (see e.g. [14]) mapping from a primary to a dual grid in all spatial directions  $i = 1, \dots, k$  and adjust the wave speed model from  $v$  to  $v'$ , where  $v'$  follows from the requirement

$$\exp\left(2sT_{\text{eik;p}}^{[l]}\right) \sum_{i=1}^k [D_{x_i}^T \exp(-sT_{\text{eik;d}_{x_i}}^{[l]})]^2 = \frac{s^2}{v'^2}, \quad (4.32)$$

where  $T_{\text{eik;p}}$  is the eikonal time on the primary grid,  $T_{\text{eik;d},x_i}$  the eikonal time on the primary grid in all spatial directions except for  $x_i$ . This equation is thus an equation on the primary grid. This is the discrete counterpart of the (continuous) relation  $\exp(2sT_{\text{eik}}) [\nabla \exp(-sT_{\text{eik}})]^2 = \frac{s^2}{v^2}$ . This equation ensures that the high-frequency dominant term  $-s^2/v^2$  vanishes and the numerical dispersion error is minimized.

Finally, we note that this dispersion correction is only accurate in the dominant direction of  $\nabla T_{\text{eik}}^{[l]}$  and only works in the reduced-order modeling framework. In our block approach multiple directions are taken into account by incorporating multiple source locations. Therefore, errors occurring in the directions orthogonal to  $\nabla T_{\text{eik}}^{[l]}$  are corrected in the block projection framework by projection onto sources with different dominant directions.

## 4.7. RESULTS

In this section we illustrate the performance of the developed solution methods using three different two-dimensional numerical experiments. In our first set of experiments, we show the performance of the proposed preconditioning technique for wavefields in a smooth layered configuration. We simulate the same structure with and without grid coarsening to show the effects of both concepts. As a second example, we consider a non-smooth medium with jumps in the wavespeed profile to illustrate that the effectiveness of preconditioning decreases as the high-frequency geometrical optics argument is no longer valid. However, the method still exhibits excellent approximation properties even for non-smooth media. Finally, in the third experiment, a configuration with a resonant cavity present in a smooth geology is considered.

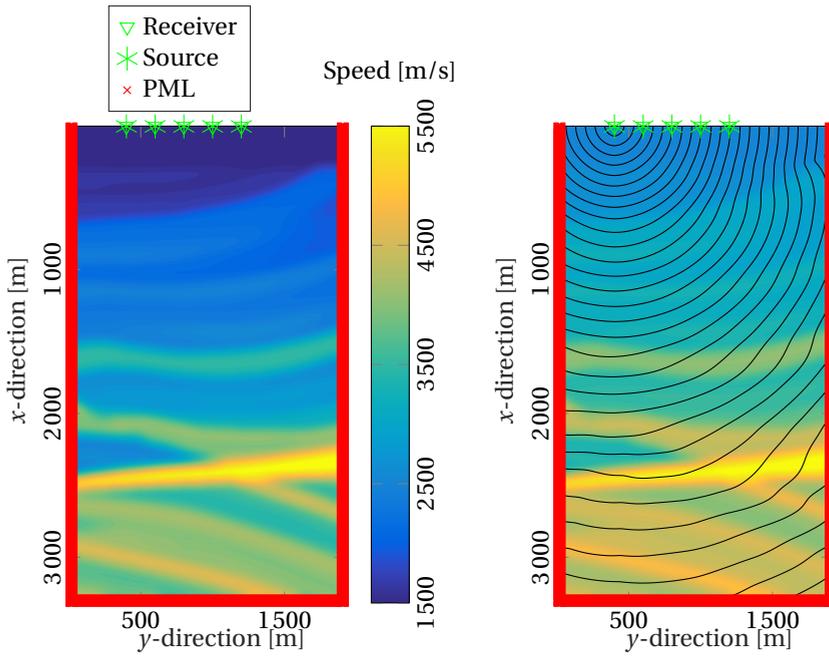
### 4.7.1. A GEOPHYSICAL STRUCTURE WITH A SMOOTH WAVESPEED PROFILE

To illustrate the effect of phase-preconditioning, we consider the smoothed geophysical structure illustrated in Figure 4.2(a). This model is obtained by smoothing a layered section of the acoustic Marmousi model [12] with a Hanning window of width  $h_{\text{Han}} = 200$  m leading to a discretized model of order  $N = 4 \cdot 10^5$ . Five coinciding source-receiver pairs are placed at the top boundary, where a perfectly reflecting boundary condition is imposed to model a water-air interface. A Ricker wavelet with a maximum in its spectrum at  $\omega^{\text{peak}} = 8$  Hz (13 ppw at 1% cut-off frequency) is used as a source signature and a near optimal eight-layer PML [20] is applied on the remaining outer boundaries to simulate outward wave propagation towards infinity. Finally, a fast marching method [75] is adopted to obtain the eikonal solution for this configuration.

The true solution  $u^{[4]}$  corresponding to the fourth source from the left at the frequency corresponding to 14.6 ppw is visualized in Figure 4.2(c). This solution shows “diving wave behavior” and a caustic can be seen at a depth of about 750 m in the left half of the configuration. The real part of the outgoing amplitude  $c_{\text{out}}^{[4]}$  is depicted in Figure 4.2(d), and is clearly spatially much smoother than the original wavefield. Reflections of the wavefield can easily be identified in this amplitude plot.

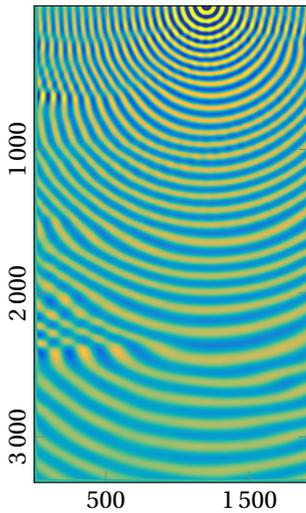
The overall time-domain errors of the block-RKS and preconditioned block-RKS reduced-order models without grid coarsening are shown in Figure 4.3(a), where we used a 500-point Fourier method to obtain a comparison solution. The overall time-domain



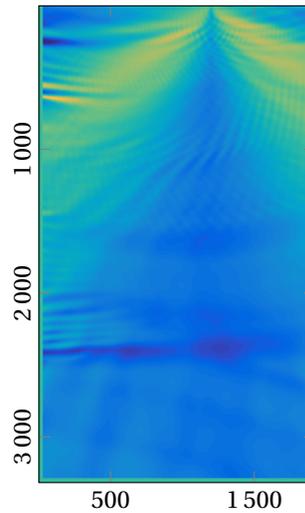


(a) Section of the wavespeed profile of the smoothed Marmousi model.

(b) Contour lines of the eikonal times for the leftmost source.



(c) Real part of the wavefield  $u^{[4]}$  excited by the fourth source from the left at a frequency corresponding to 14.6 ppw.

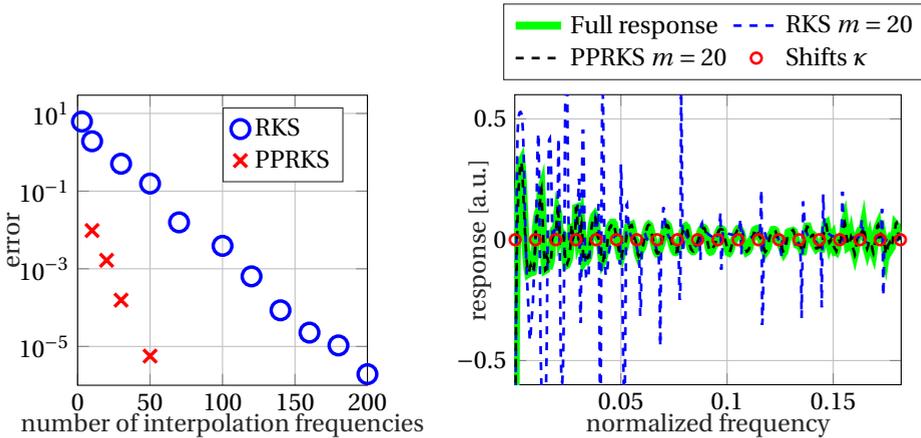


(d) Real part of the amplitude  $c_{out}^{[4]}$  excited by the fourth source from the left at a frequency corresponding to 14.6 ppw.

Figure 4.2: The smoothed Marmousi layers test configuration. Shown are the configuration (a), the contour lines of the eikonal times (b) and a solution vector prior (c) and after parametrization (d).

error is defined by the ratio of the r.m.s. error of all traces and the r.m.s. of the signals over *all* traces. Preconditioning the RKS method significantly decreases the number of interpolation points needed to reach a certain error level. To obtain an error of one percent in the time-domain, for example, the RKS algorithm needs about 80 interpolation points, while only 10 interpolation points are required in the phase-preconditioned algorithm. This fast convergence is due to the construction of the WKB-like field approximations at high frequencies in PPRKS, which already provide an accurate approximation of the Green's function at high frequencies and in smooth structures as considered in this example.

The real part of the transfer-function of the leftmost source to the rightmost receiver is shown in Figure 4.3(b) for structure-preserving and preconditioned block-RKS reduced-order models of order  $m = 20$ . The phase-preconditioned model coincides with the comparison solution on the complete frequency interval of interest. The main oscillations present in this Green's function response are due to the direct arrival of the wave and its first reflection from the salt layer located at a depth of about 2500 m. Typically, the PPRKS method provides a smooth approximation to the field response showing only small errors in the amplitudes or at highly oscillatory reflections. The structure-preserving RKS method, on the other hand, overshoots after every interpolation point causing spiking behavior as can be clearly seen in Figure 4.3(b).



(a) Time-domain convergence of RKS and PPRKS. The overall time-domain error is shown using a Ricker wavelet for excitation.

(b) Real part of the frequency-domain transfer-function from the leftmost source to the rightmost receiver after  $m = 20$  interpolation points.

Figure 4.3: Convergence curves of the RKS and PRKS (a) and the real part of the transfer-function for the leftmost source to the rightmost receiver (b).

#### GRID COARSENING AND SVD

The amplitude-functions  $c_{in}$  and  $c_{out}$  are spatially much smoother than the wavefield and therefore we expect that a coarser spatial grid can be employed. To investigate the effects of grid coarsening, we consider the same wavespeed profile as in the previous



example and place 12 coinciding source-receiver pairs at the top water-air interface instead of 5. For excitation, we use a modulated Gaussian pulse with a center frequency  $\omega^{\text{peak}}$  and its support essentially given by  $[0, 2\omega^{\text{peak}}]$ . The pulse is shifted in time such that it starts at  $t = 0$ . Spatial discretization is now chosen such that we have about 5.5 points per smallest wavelength, where the wavelength corresponds to the center frequency of the pulse and 2.7 ppw at the cut-off frequency of the pulse. With this choice, the step sizes of the grid are four times larger than in the previous example leading to a system that is sixteen times smaller with  $N = 2.5 \cdot 10^4$  unknowns. Using such a coarse grid to model wavefields without phase-preconditioning is obviously insufficient, but here we expect that the smoothness of the amplitude-functions  $c_{\text{in}}$  and  $c_{\text{out}}$  allows us to use a much coarser grid. During the evaluation of the reduced-order model we project an operator corresponding to a fine grid onto the phase-corrected RKS. For this example, we choose a fine operator using half the step size compared to the previous operator in order to show that the projection gauges the ROM to the operator used during projection.

For MIMO systems with grid coarsening, we define the error as the error averaged over all source-receiver combinations. We denote the elements of the finite-difference matrix transfer-function  $F_{\text{F}}(s)$  by  $f_{\text{F}}^{[ij]}(s) = b^{[i];H} u^{[j]}(s)$ , while the element of the ROM transfer-function  $F_m(s)$  are given by  $f_m^{[ij]}(s) = b^{[i];H} u_m^{[j]}(s)$ . Having introduced these elements, the average MIMO error as function of frequency is defined as

$$\text{err}_{\text{ROM}}^{\text{average}}(m, s) = \frac{\sqrt{\omega_{\text{max}}}}{N_{\text{src}}^2} \sum_{j=1}^{N_{\text{src}}} \sum_{i=1}^{N_{\text{src}}} \frac{|f_{\text{F}}^{[ij]}(s) - f_m^{[ij]}(s)|}{\left( \int_{\omega=0}^{\omega_{\text{max}}} |f_{\text{F}}^{[ij]}(i\omega)|^2 d\omega \right)^{1/2}}. \quad (4.33)$$

It is assumed that the comparison solution  $f_{\text{F}}^{[ij]}(t)$  is computed with a spatial discretization of sufficient accuracy. The averaged frequency-domain error definition of (4.33) gives a higher error, yet delivers more insight, than computing the overall error. The overall error is dominated by the monostatic elements  $f_{\text{F}}^{[ii]}$ , whose direct arrival contains most energy and is well approximated. Furthermore, the above error definition allows us to study the error as a function of frequency.

The phase-corrected RKS is build using  $m = 40$  equidistant shifts on the imaginary axis from  $\omega = 2.4 \cdot 10^{-3} \omega^{\text{peak}}$  (2383 points per smallest wavelength) to  $\omega = 1.1 \omega^{\text{peak}}$  (5 points per wavelength). In other words, the RKS interpolation frequencies uniformly cover the lower half of the support of the spectrum of the source wavelet. With  $m = 40$  interpolation points and  $N_{\text{src}} = 12$  source-receiver pairs, we have 480 amplitude-functions  $c_{\text{in}}$  and an additional 480 amplitude-functions  $c_{\text{out}}$ . Computing the SVD of the 960 amplitude-functions  $[c_{\text{out}} \bar{c}_{\text{in}}]$ , we observe that for this example, essentially only the first 100 singular functions contribute to the reduced-order model for the contracted amplitudes. We therefore use a truncated SVD that uses the first 100 SVD basis functions to represent the amplitudes.

The resulting time-domain trace from the leftmost source to the rightmost receiver is shown in Figure 4.4 compared to the trace obtained via a 500-point Fourier method using an operator with step sizes eight times smaller than the step sizes used in the coarse operator. Both responses clearly coincide on the considered time window

and the first arrival of the pulse, the complex interaction between the pulse and the upper layered medium, and the reflection of the pulse at the high contrast salt layer around  $t = 3000$  can be observed. The multiple reflection from source to salt layer, water/air interface, salt layer and back to the receiver can be seen around  $t = 6000$ .

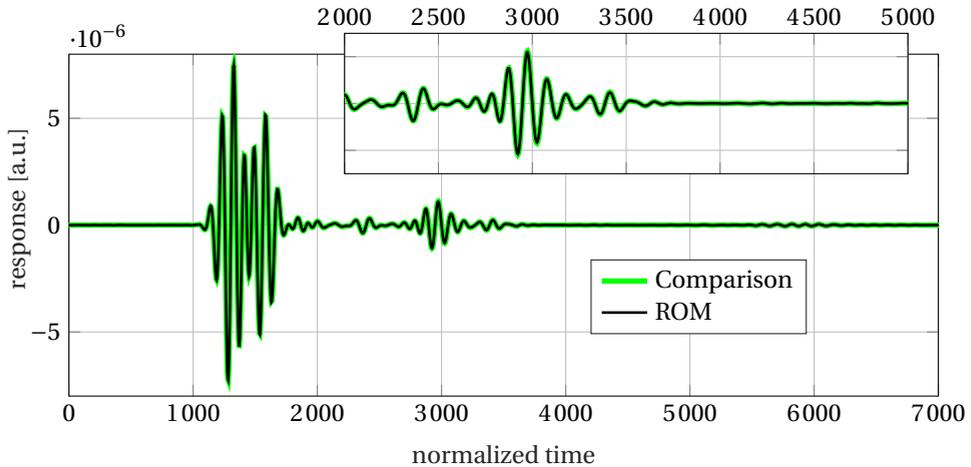
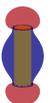


Figure 4.4: Time-domain trace from the leftmost source to the rightmost receiver after  $m = 40$  interpolation points.

In Figure 4.5 we plotted the decay of the singular values of a matrix with pairwise normalized vectors  $c_{\text{out}}^{[r_1]}(s_j)/\bar{c}_{\text{in}}^{[r_1]}(s_j)$  and normalized  $u^{[r_1]}(s_j)$  as columns. The decay of the singular values of the single frequency solutions that make up the RKS is shown in black. The singular values associated with the amplitude matrix are shown in blue. Prior to the SVD the vectors  $u^{[r_1]}$  were normalized, such that the sum of the squared singular values equals  $mN_s$ . The vectors  $c_{\text{out}}^{[r_1]}(s_j)$  and  $\bar{c}_{\text{in}}^{[r_1]}(s_j)$  were normalized in pairs together to reflect the ratio of the incoming and outgoing wave at each frequency. Their SVD is computed together such that the sum of the squares of all singular values adds up to  $mN_s$  as well. Finally, to show the decay in singular values we normalize the largest singular value to one for each of the curves. Figure 4.5 clearly shows that incoming and outgoing amplitudes are significantly compressible, whereas the RKS vectors are not. The singular values associated with RKS drop by less than a factor of 2 between the index of 50 and the index 400, indicating that the RKS can hardly be compressed. To show that after the compression the basis is essentially independent of the number of sources, we computed the SVD of the amplitudes  $[c_{\text{out}}^{[r_1]}(s_j) \bar{c}_{\text{in}}^{[r_1]}(s_j)]$  for  $N_{\text{src}} = 12, 24, 48$  and 96. The number of (normalized) singular values larger than 0.01 is shown in Table 4.1. It shows that the number of contributing singular vectors is basically independent of the number sources and so are the left singular vectors.

In Figure 4.6, the averaged error  $\text{err}_{\text{ROM}}^{\text{average}}$  over all  $12^2$  traces is shown along with the interpolation points used in the construction of the reduced-order model. The same Fourier method that was used to compute the comparison solution in Figure 4.4 is used here to compute the errors in the transfer-function. Furthermore, the figure shows the



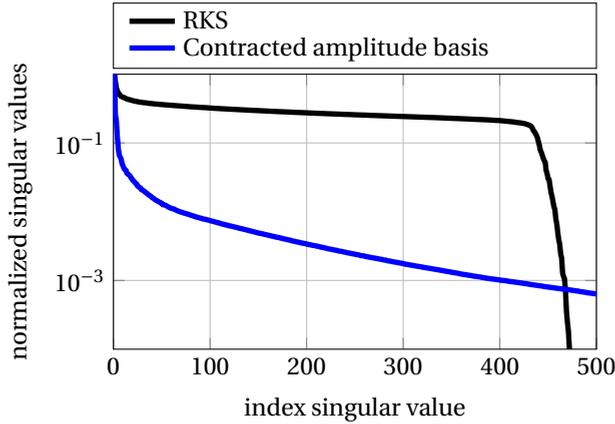


Figure 4.5: Decay of singular values of RKS vectors compared to the decay of  $[c_{\text{out}}, \bar{c}_{\text{in}}]$ .

Table 4.1: Number of (normalized) singular values larger than 0.01 in dependence of the number of sources

$N_{\text{src}}$	12	24	48	96
Number of (normalized) singular val. > 0.01 for $[c_{\text{out}}, \bar{c}_{\text{in}}]$	69	72	73	73
Number of (normalized) singular values > 0.01 for $u$	457	833	1369	1741
$m \cdot N_{\text{src}}$	480	960	1920	3840

error of an FDFD method with a normalized step size of 0.6, which is 20% larger than the step size used to compute the comparison solution and used for the operator that was projected onto the PPRKS. For higher frequencies such an operator has increasing dispersion, such that the solutions between the comparison FDFD method with normalized step size of 0.5 and the one of 0.6 don't match anymore. For low frequencies there is a small discrepancy due to the inability of both grids to approximate a delta source.

When introducing grid coarsening, the ROM no longer interpolates the transfer-function, but the error remains small and below 1% on the frequency interval covered by the interpolation points. In addition, we observe that the phase-preconditioned reduced-order models can extrapolate to higher frequencies to a certain extent, since the basis in PPRKS is frequency-dependent. The error only gradually grows outside the interpolation interval, which covers the lower half of the spectrum of the pulse, and at 2.4 points per smallest wavelength we end up with an error of about 5%.

Finally, in Figure 4.7 the averaged error in the transfer-function is shown as a function of the number of points per wavelength. Again, the PPRKS with 40 interpolation points is compared to the 500-point Fourier method, but this time the latter method uses the same coarse-grid operator that is used during construction of the PPRKS (instead of the operator that uses step sizes 0.6 as in the previous figure). Clearly, the RKS approach which uses a Galerkin condition to select optimal linear combinations with respect to a fine operator outperforms the direct Fourier method that uses the same operator to construct the field approximations. The RKS approach is gauged to the operator by using the Galerkin condition.

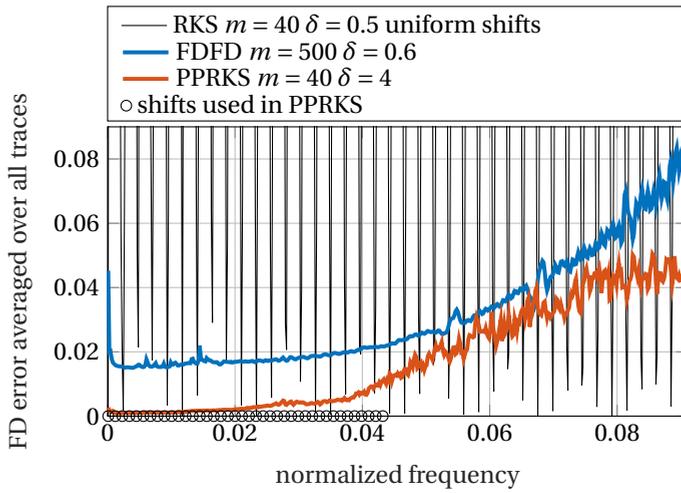


Figure 4.6: The relative error of eq. (4.33) of RKS, FDFD and PPRKS as function of frequency. The error of an FDFD method with a step size that is 20% larger than the one used to produce the comparison solution is shown in red.

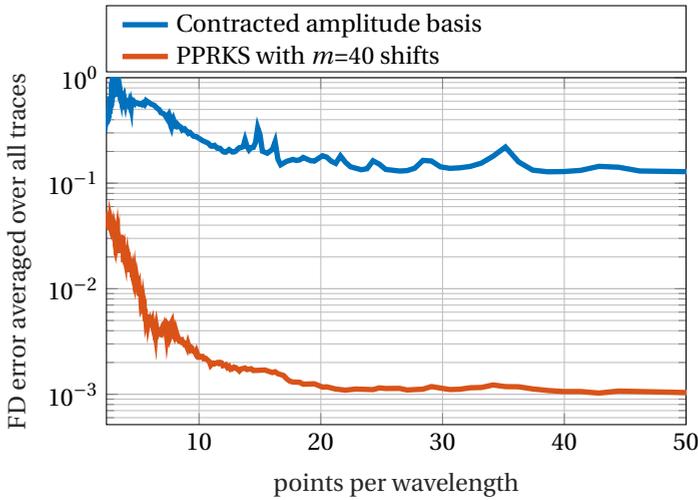


Figure 4.7: Comparison of the relative error of the reduced-order model  $err_{ROM;coarse}^{average}$  versus  $err_{FD;coarse}^{average}$  obtained by direct evaluation of the transfer-function on the coarse grid using  $Q_{coarse}(s)$  as a function of points per wavelength.



#### 4.7.2. A GEOPHYSICAL STRUCTURE WITH A NON-SMOOTH WAVESPEED PROFILE

The justification of the phase-preconditioned algorithm is based on a geometrical optics argument. This asymptotic argument is applicable for smooth media with spatial variations that take place on a scale much smaller than the wavelength. On the other hand, RKS reduced-order modeling is a valid approach independent of the medium considered and Proposition 4.2 shows that one-dimensional problems with piecewise constant wavespeeds need not be a problem for this approach to work. Therefore, let us turn to an unsmoothed variant of the layered geophysical structure from the Marmousi model considered earlier as depicted in Figure 4.8.

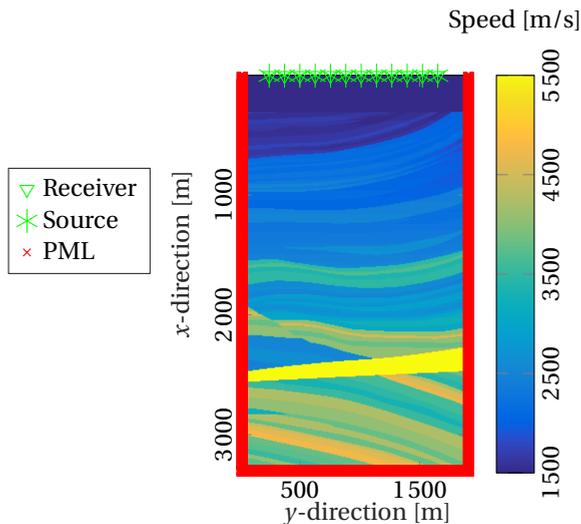


Figure 4.8: Section of the wavespeed profile from the Marmousi model without smoothing. Twelve coinciding source-receiver pairs are placed on the surface.

For this structure, we essentially follow the same procedure as before. Specifically, we again position 12 source-receiver pairs at the top air-water interface and use the same coarse grid operator as in the previous example to construct a phase-corrected RKS reduced-order model of order  $m = 40$  with interpolation points on the imaginary axis covering the lower half of the spectrum of the pulse such that we have 5 points per smallest wavelength for the highest interpolation frequency. The center frequency of the pulse is again chosen at 5.5 ppw. The only difference in model construction compared with the previous example, is that here we use a truncated SVD that takes 150 SVD basis functions into account, instead of the 100 basis functions in the previous example. Here, more basis functions are required, since the amplitude-functions are less smooth due to the non-smooth wavespeed profile of the present Marmousi model. Finally, the comparison solution is computed using a direct 500-point Fourier method using a spatially discretized operator with step sizes that are four times smaller than the step sizes used in the coarse operator. The coarse operator has a normalized step size of 4 and the

operator used to compute the comparison FDFD response has a normalized step size of 1.

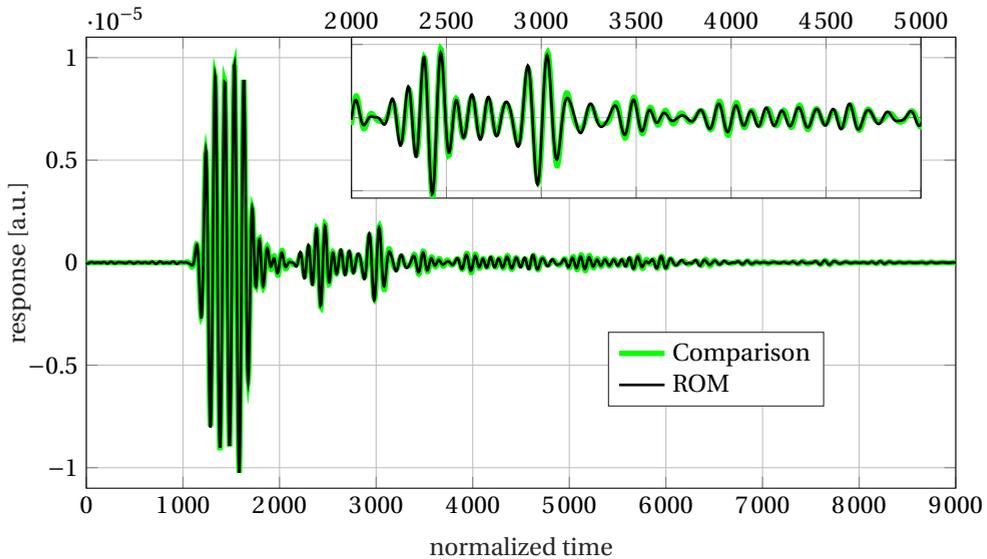
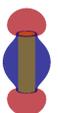


Figure 4.9: Time-domain trace from the leftmost source to the rightmost receiver after  $m = 40$  interpolation points and the comparison solution.

The resulting error is shown in Figure 4.10 along with the corresponding error curve for an FDFD method which used the coarse operator that constructed the PPRKS. In addition, the error of an ordinary RKS method is shown, which uses the fine operator for construction and evaluation. It uses  $m = 40$  shifts uniformly distributed on the whole spectral interval. An RKS method on the fine grid interpolates the FDFD response on the shifts, which leads to a strongly oscillatory error curve. This error curve clearly shows the advantage of phase-preconditioning with a dual grid approach – a lower error is achieved while solving considerably smaller shifted systems, and projecting on the same operator. The performance of the algorithm for smooth profiles is better than for non-smooth wavespeed profiles, especially for extrapolated frequencies. We also observe that the error decreases for lower frequencies, since lower frequencies have larger wavelengths and variations in the wavespeed profile take place on a scale smaller than these wavelengths of operation. Furthermore, compared to an FDFD method that uses a 20% coarser grid than the comparison solution, the PPRKS achieves lower errors across the whole spectral interval while the systems that need to be solved are much smaller. Especially in the area where the phase-preconditioned method has shifts it reproduces the comparison solution remarkably well. A similar error comparison is shown in Figure 4.11, where the error is plotted against points per wavelength.

To illustrate the effects of an increased error in the time-domain, we show the time trace for the most distant source-receiver pair in Figure 4.9 (for the same Gaussian pulse as used before in Figure 4.4) along with a comparison solution obtained with the 500-point Fourier method. We observe that the arrival times are approximated well; only



the amplitudes are slightly off. Throughout our numerical work, we have found that this result is typical for non-smooth problems. Furthermore, compared with the same trace computed for smooth media as shown in Figure 4.4, it is clear that a larger part of the pulse is scattered back to the receiver, as arrivals are visible on the complete time interval of observation for the non-smooth velocity profile considered here.

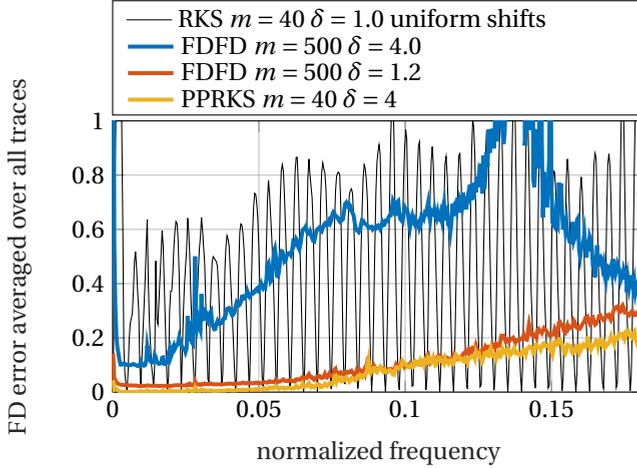


Figure 4.10: Comparison of the relative error  $\text{err}_{\text{ROM};\text{coarse}}^{\text{average}}$  and  $\text{err}_{\text{FD};\text{coarse}}^{\text{average}}$  for the non-smooth Marmousi layer configuration.

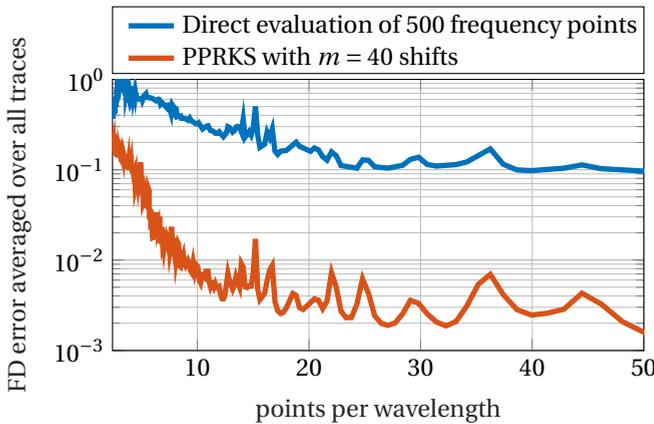


Figure 4.11: Comparison of relative error  $\text{err}_{\text{ROM};\text{coarse}}^{\text{average}}$  and  $\text{err}_{\text{FD};\text{coarse}}^{\text{average}}$  for the non-smooth Marmousi layer configuration.

### 4.7.3. A RESONANT CAVITY EMBEDDED IN A SMOOTH GEOLOGY

In this section we investigate the performance of our algorithm in a configuration with a resonant cavity. Figure 4.12(a) shows the wavespeed profile, which is inspired by borehole exploration. Coinciding source-receiver pairs are placed at the surface and inside a borehole of slow acoustical wavespeed.

The grid coarsening procedure and wavelet selection is equivalent to the previous examples. A coarse grid operator with the same accuracy as selected in the previous example is used to construct a phase-corrected RKS reduced-order model of order  $m = 40$  with interpolation points on the imaginary axis covering the lower half of the spectrum of the pulse such that 5 points per smallest wavelength are used for the highest interpolation frequency. The center frequency of the pulse is again chosen at 5.5 ppw.

To approximate the cavity-resonances with few interpolation points, we extend the approach discussed in this paper and factor out oscillations of resonance modes along the borehole. To do so, we take the fact that the eikonal time  $T_{\text{eik}}$  is multivalued into account. More specifically, each solution  $u(\kappa_i)$  is split using two different phase terms, a cavity-mode phase term and a propagation phase term. The eikonal phase term shown in Figure 4.12(b) shows a caustic inside the borehole, which has a low wavespeed compared to its surrounding. In this experiment we also factor out the cavity-mode phase term  $g(sT_{\text{eik;CM}})$ , where  $T_{\text{eik;CM}}$  follows the borehole as shown in Figure 4.12(c). The eikonal time of the cavity-mode  $T_{\text{eik;CM}}$  is not the second arrival, but it is chosen to correctly factor out resonances present in the borehole. At every interpolation point we split the field into four amplitudes as

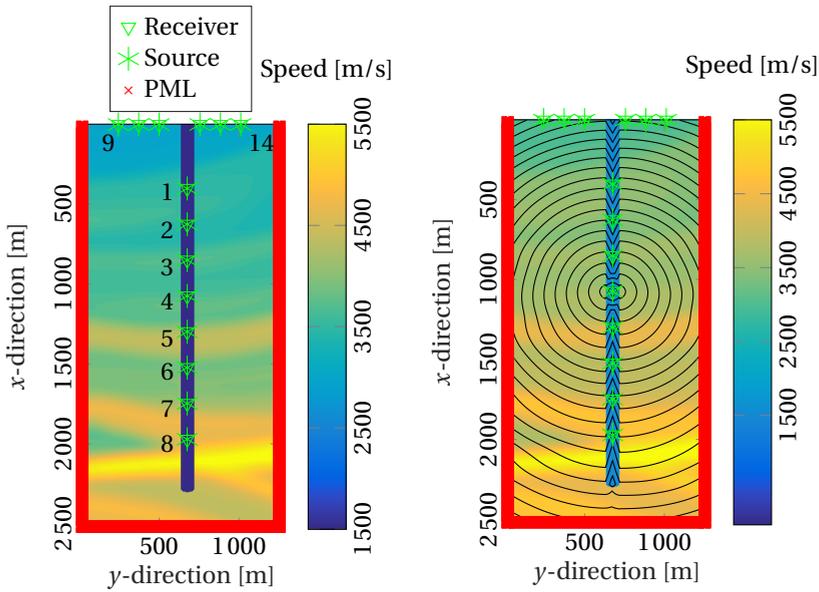
$$u^{[l]}(s_j) = g(s_j T_{\text{eik}}^{[l]}) c_{\text{out;eik}}^{[l]}(s_j) + g(-s_j T_{\text{eik}}^{[l]}) c_{\text{in;eik}}^{[l]}(s_j), \quad (4.34)$$

$$u^{[l]}(s_j) = g(s_j T_{\text{eik;CM}}^{[l]}) c_{\text{out;CM}}^{[l]}(s_j) + g(-s_j T_{\text{eik;CM}}^{[l]}) c_{\text{in;CM}}^{[l]}(s_j). \quad (4.35)$$

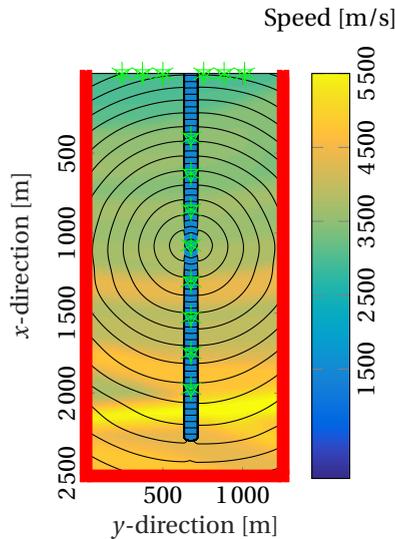
With 14 sources and 40 interpolation points we end up with 560 amplitudes for each of the four amplitudes  $c_{\text{in/out;eik/CM}}$ , which we compress to 30 each using an SVD. These compressed amplitudes are then used to construct the phase-preconditioned rational Krylov subspace on which the fine operator is projected.

For these types of configurations the time window of interest tends to be very long due to the resonant nature of the configuration. FDTD therefore requires very long runtimes, whereas the proposed algorithm just needs to evaluate the ROM on more frequencies to avoid aliasing. The time-domain trace of the top most source-receiver pair within the borehole is shown in Figure 4.13, where the emitted pulse bounces back and forth within the cavity. The reduced-order model captures this resonant behavior, showing that the resonant modes are well approximated. In Figure 4.14 a trace is shown from a source within the borehole to a surface receiver. In this trace it can be seen that the interaction of the pulse with the smooth geology is modeled correctly next to the repetitive trace shape caused by the resonant cavity. An ordinary RKS method with no grid coarsening would perform well on this problem, since it is mainly dominated by the resonant cavity; however, this would require solutions of the wave equation on a much finer grids than the proposed approach. Furthermore, contrary to the proposed approach, the RKS approximation deteriorates as the configuration size and thus the propagation distance from the cavity to the receiver increases.





(a) Simulated configuration with labeled sources. Eight source-receiver pairs are placed inside an elongated resonant cavity and six pairs are placed at the surface. (b) Contour lines of the eikonal time  $T_{\text{eik}}$  used to approximate the propagative part of the solution.



(c) Contour lines of the eikonal time  $T_{\text{eik;CM}}$  used to approximate the cavity-modes of the solution.

Figure 4.12: The source-receiver setup (a), and the eikonal times (b) and (c) used to decompose the fields into smooth amplitude functions are depicted

In this experiment we show that the developed algorithm shows potential for reduced-order modeling of resonant cavities within slowly varying media. The combination of an RKS method together with phase-preconditioning can approximate both resonant eigenmodes as well as propagative modes. We point out that this is just a first approach in order to include resonant structures into reduced-order models that are travel time dominated.

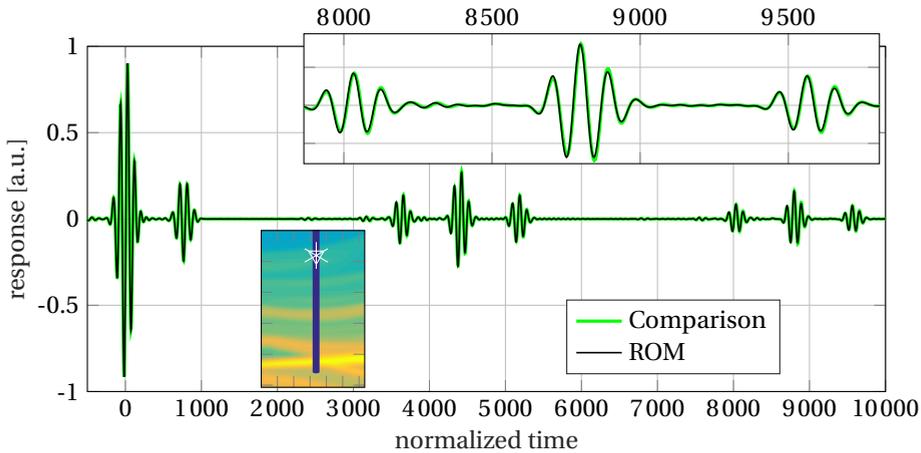


Figure 4.13: Time-domain trace of the coinciding source-receiver pair number 1 after  $m = 40$  interpolation points together with the comparison solution.

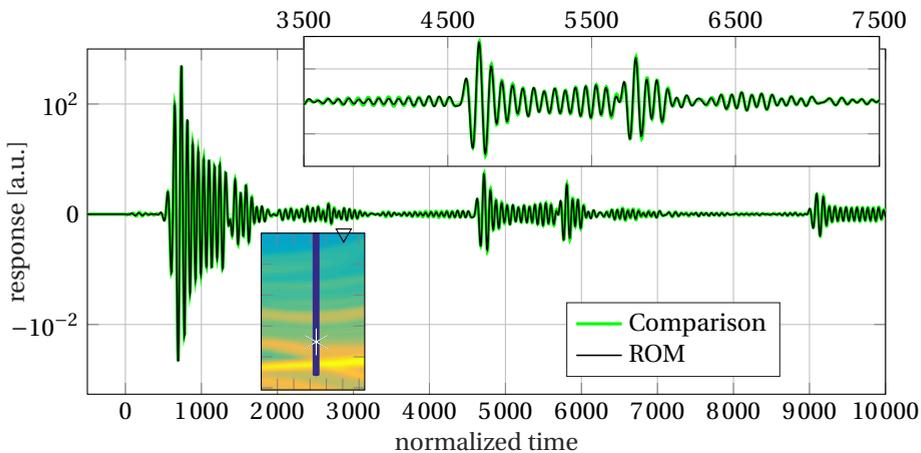
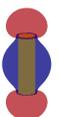


Figure 4.14: Time-domain responses of the resonant cavity test-case. Time-domain trace from source number 7 inside the borehole to the rightmost surface receiver number 14 after  $m = 40$  interpolation points.



#### 4.7.4. SMOOTHED COMPLEX GEOLOGY

In this last experiment we investigate the performance of the proposed algorithm in a complex geology with multiple caustics present in the eikonal solution. The simulated configurations is shown in Figure 4.15 along with the eikonal solution of the leftmost receiver is shown in Figure 4.16. In total 18 coinciding source-receiver pairs at the water-air interface are simulated. The contour lines of the eikonal time in Figure 4.16 clearly shows multiple caustics and a complex wavefront. For this example a coarse grid of  $\Delta x = 4$  m is used. The frequency-domain transfer-function for the most distant source-receiver pair is shown in Figure 4.17, after  $m = 40$  iterations and a truncation of the SVD of  $c_{in/out}$  after 80 vectors each. Until 4 ppw are reached, the ROM is in good agreement with the comparison solution consisting out of 500 frequency-domain solves of an operator with  $\Delta x = 4$  m. This example shows worse extrapolation then the previous examples as the frequency-domain response is not accurately reproduced until low points per wavelength. However in the area were RKS shifts are placed the signal is reproduced accurately. A Gaussian wavelet centered around 8 ppw with a cut-off at 3.5 ppw is used in this example to obtain the time-domain trace from the ninth source from the left to the rightmost receiver as shown in Figure 4.18. Comparison solution and ROM are indistinguishable for this example.

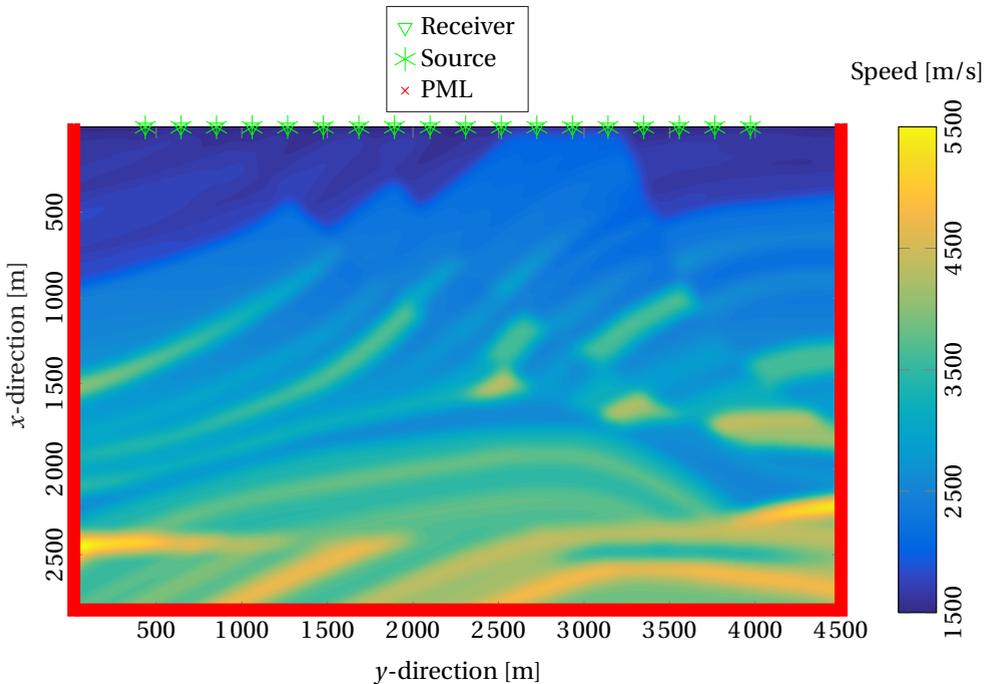


Figure 4.15: Smoothed section of the Marmousi model with multiple coinciding source-receiver pairs at the surface.

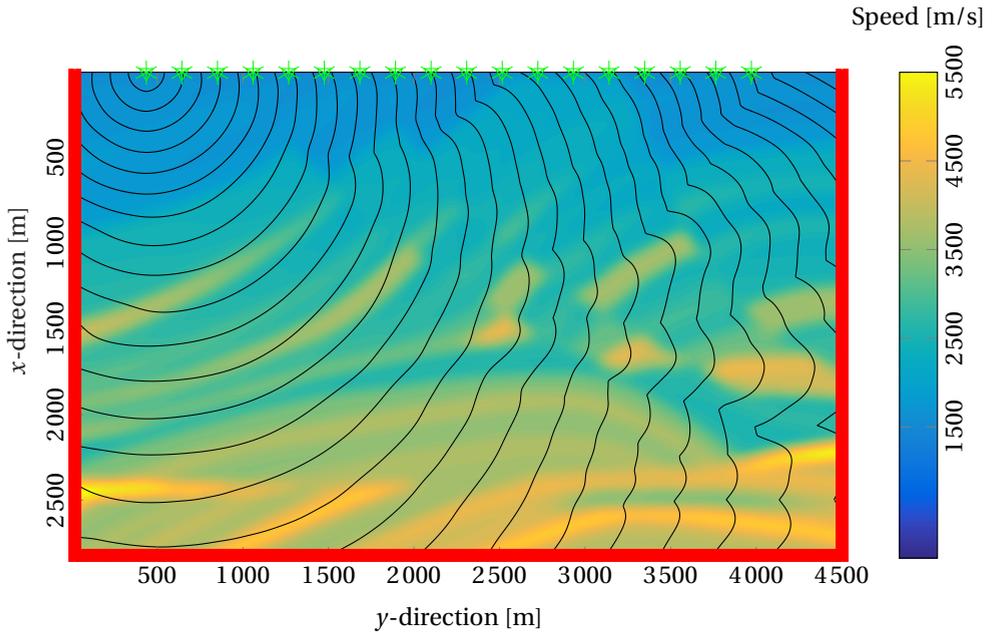


Figure 4.16: Contour lines of the eikonal time of the leftmost source. Multiple caustics are visible and the first arrival for the rightmost sources comes from a diving wave.

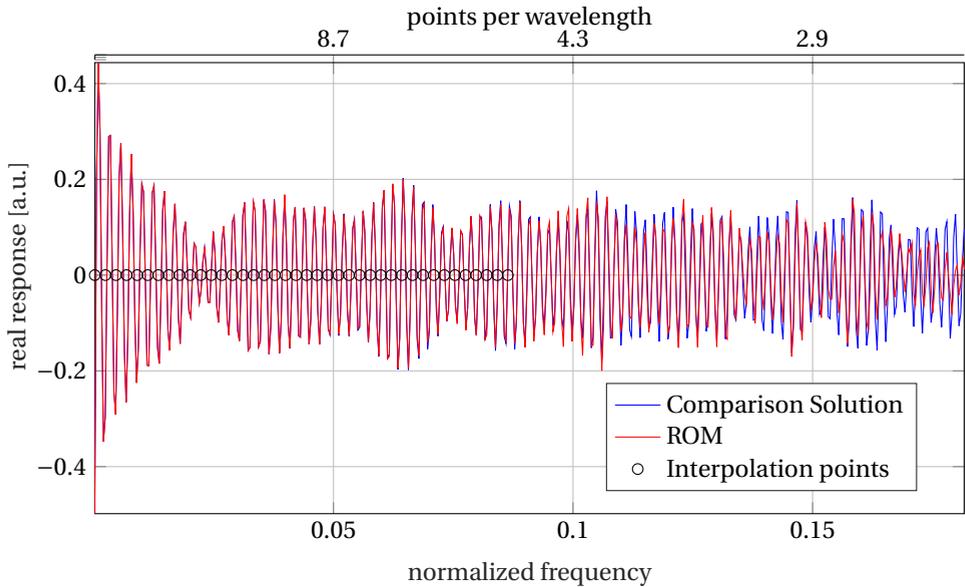


Figure 4.17: Frequency-domain transfer-function from the leftmost source to the rightmost receiver for the Marmousi test configuration with grid coarsening.



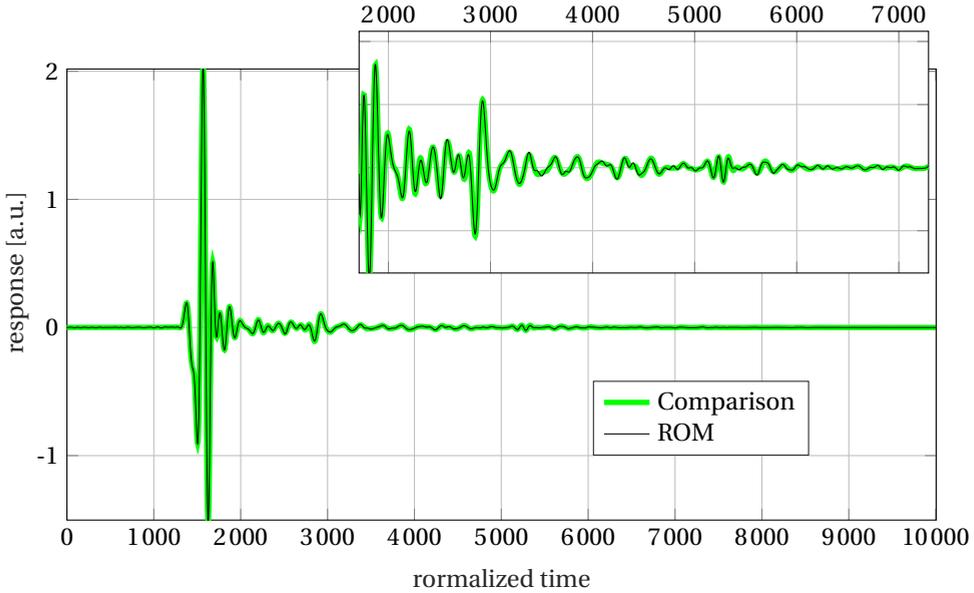


Figure 4.18: Time-domain trace from the middle source (9) to the rightmost receiver for a Gaussian wavelet with a center frequency at 8 ppw and a cut-off of 3.5ppw for the Marmousi test configuration with grid coarsening.

#### 4.8. DISCUSSION ON PARALLEL IMPLEMENTATION

The numerical experiments of the previous section (using a serial MATLAB prototype code) showed significant compression of large-scale wave propagation due to phase-preconditioning. To see how observed dimensionality reduction can be translated to computational cost reduction using modern high performance platforms, e.g., cloud computing, we consider the simplest parallel implementation, known in computer science literature as an “embarrassingly parallel workflow<sup>2</sup>.”

Like the majority of the projection-based model reduction methods, the PPRKS can be split into basis construction and ROM evaluation stages, as summarized in Figure 4.19. This figure is complemented by Table 4.2, where we compare computational cost estimates for PPRKS with standard RKS neglecting  $O(N_f)$  terms and considering only parallelism on the external level.

For both standard RKS and PPRKS the main cost of the first stage consists of the computation of the block-RKS and the rank-revealing subspace truncation via SVD. Phase-preconditioning adds the negligible cost of solving the eikonal equation and decomposing the waves into incoming/outgoing amplitudes via ((4.17) and (4.18)). In the table we assume that the block-RKS is computed by assigning solutions of Helmholtz problems for different frequencies and right-hand sides to separate workers, so that the PPRKS and RKS require  $N_{\text{src}} m_{\text{PPRKS}}$  and  $N_{\text{src}} m_{\text{RKS}}$  nodes, respectively. The PPRKS obviously reduces the number of Helmholtz solves; however, in the parallel implementation

<sup>2</sup>Term used for parallelization not requiring horizontal communication between nodes.

Table 4.2: Cost estimates for PPRKS and RKS.

Step	PPRKS	
	Computations per worker	# workers
Eikonal	$O(N_f \log N_f)$	$N_{\text{src}}$
Basis Comp	$O(N_c \psi(N_c))$	$N_{\text{src}} m_{\text{PPRKS}}$
SVD	$O(N_c [2N_{\text{src}} m_{\text{PPRKS}}]^2)$	1
Eval	$O(N_f N_{\text{src}}^2 M_{\text{SVD}}^2)$	$N_{\text{eval}}$
	RKS	
Basis Comp.	$O(N_f \psi(N_f))$	$N_{\text{src}} m_{\text{RKS}}$
SVD	$O(N_c [2N_{\text{src}} m_{\text{RKS}}]^2)$	1
Eval	$O(N_f N_{\text{src}}^2 m_{\text{RKS}}^2)$	1

## Legend

$N_{\text{src}}$	# sources
$N_c/f$	# coarse/fine grid notes
$m$	# ROM interpolation points
$N_{\text{eval}}$	# evaluation frequencies
$M_{\text{SVD}}$	Size SVD compressed vectors
$\psi(N_c/f)$	scaling function of Helmholtz solver

the most important cost reduction lies in a single solve. In our case, this cost is critical due to the high complexity and poor internal parallel scalability of available Helmholtz solvers. In the table,  $\psi(N)$  reflects this (usually faster than logarithmic) growth of the computational complexity of the Helmholtz solver.

Thus, the observed reduction in grid nodes from  $N_f$  to  $N_c$ , which lies between one and two orders of magnitude, can result in even stronger reductions of computation time. Subspace truncation is another poorly-scalable bottle-neck of the basis generation stage (e.g., see [59]) and the compound effect of the reduction of  $m_{\text{PPRKS}}$  and  $N_c$  compared to  $m_{\text{RKS}}$  and  $N_f$  is more than two orders.

The main cost of the second stage is the evaluation of the ROM frequency response at quadrature points in the frequency domain. In particular, the computation of the orthogonal basis and the Galerkin projection are the main bottlenecks with costs that grow linear with respect to the fine grid dimension. The dimension of the PPRKS approximation space is the product of the size of the compressed amplitude space and the number of sources, which is usually of the same order as the dimension needed for standard RKS<sup>3</sup>. Nonetheless, storage of the space is reduced by a factor of  $N_{\text{src}} N_f / N_c$  and the computation of the coarse grid amplitudes is obviously cheaper. However, the phase-preconditioned subspace is frequency-dependent, unlike standard RKS. Therefore, the Galerkin projection should be computed for every frequency for the entire operator  $Q(s)$ . This is not a significant disadvantage thanks to the possibility of an embarrassingly parallel implementation; for every evaluation frequency a separate worker can be assigned. Moreover, the compressed tensor-product representation (4.27) allows efficient lower level parallelization for the evaluation phase of PPRKS, i.e., column-wise, element-wise and via domain-decomposition of the inner products. Solving the Galerkin system as well as carrying out inverse Fourier transforms to the time-domain do not depend on the grid size and their costs can be neglected.

We choose to benchmark the prototype implementation of algorithm in two parts. The basis construction is benchmarked on a CPU and the ROM evaluation on a GPU, as our algorithm is intended for the modern high performance computing environment. Efficient Helmholtz solvers or solvers for large, sparse matrix systems are generally de-

<sup>3</sup>Recall that this is due to the tensor-product structure of the PPRKS approximation space given by (4.27).



4

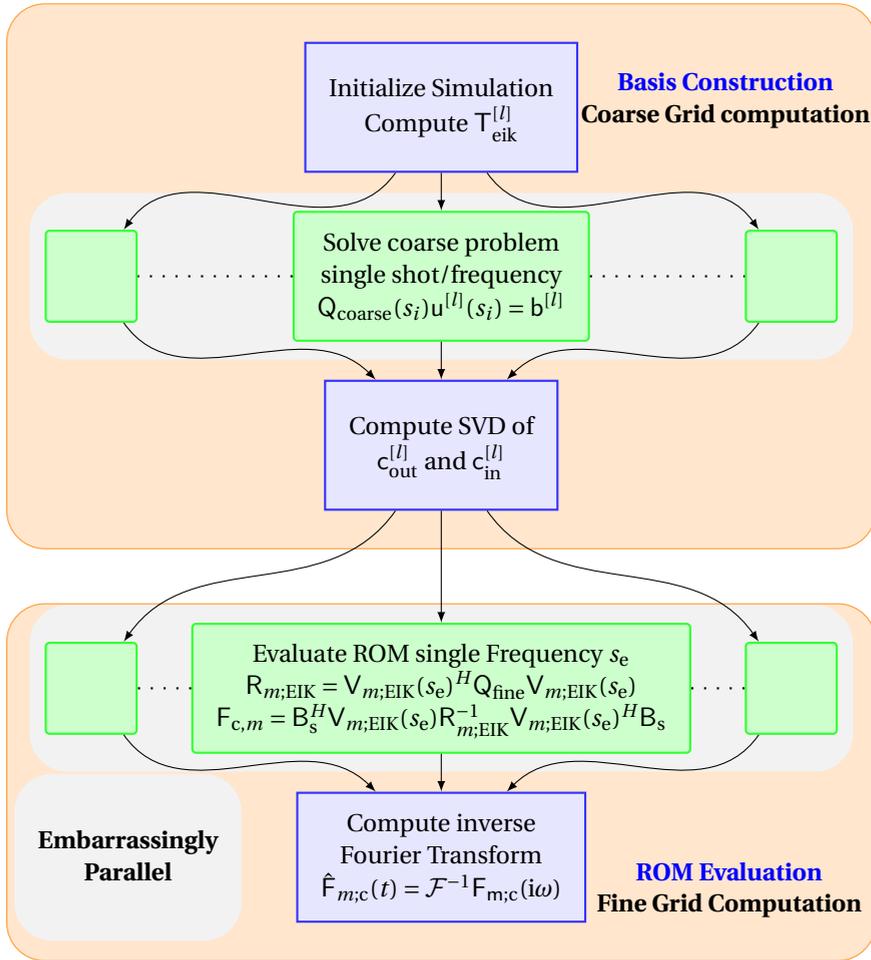


Figure 4.19: Overview of the proposed algorithm. External embarrassing parallelism is symbolized by parallel blocks. Internal parallelism within a block is also possible.

Table 4.3: Cost of the basis computation and evaluation of the reduced-order model.

Basis Computation comparison	Computation	Time	
Block solve fine grid	$\mathbf{Q}_{\text{fine}}(s_i)^{-1}\mathbf{B}$	10.3 s	
Single solve fine grid	$\mathbf{Q}_{\text{fine}}(s_i)^{-1}\mathbf{b}$	4.1 s	
Block solve coarse grid	$\mathbf{Q}_{\text{coarse}}(s_i)^{-1}\mathbf{B}$	0.6 s	
Single solve coarse grid	$\mathbf{Q}_{\text{coarse}}(s_i)^{-1}\mathbf{b}$	0.2 s	
Evaluation Step	Computation	Time	Scaling
Computing phase-functions	$\exp(sT_{\text{eik}})$	0.00546 s	$N_{\text{src}}N_f$
Hadamard Products	$\exp(sT_{\text{eik}})c_{\text{SVD}}$	0.01496 s	$M_{\text{SVD}}N_{\text{src}}N_f$
Galerkin inner product	$\mathbf{V}_{m;\text{EIK}}(s_e)^H \cdot \mathbf{Q}_{\text{fine}}\mathbf{V}_{m;\text{EIK}}(s_e)$	1.752 s	$N_f M_{\text{SVD}}^2 N_{\text{src}}^2$

veloped for CPUs. GPUs, however, are designed for fast, parallel computation of large inner products and therefore pre-eminently suitable for the evaluation stage of the proposed model order reduction technique.

For the smooth geophysical structure example given in this paper with  $N_f = 4 \cdot 10^5$ ,  $N_{\text{src}} = 12$  we compare the most important computation times in Table 4.3. The basis computation is performed on a CPU<sup>4</sup> and the ROM evaluation on a GPU<sup>5</sup>. In the proposed algorithm we solve the wave equation on a coarse grid only, leading to a much lower cost in basis construction than standard RKS where fine-scale systems need to be solved. This is especially important considering that for large 3D applications it can become infeasible to solve the wave equation on a fine grid as the scaling function  $\psi(\cdot)$  is much worse for 3D systems than for 2D systems. To show the cost of the evaluation of the reduced-order model we benchmarked the evaluation of a single frequency  $s_e$  on a GPU. The used model parameters are  $N_f = 4 \cdot 10^5$ ,  $M_{\text{SVD}} = 100$ ,  $N_{\text{src}} = 12$  and the results are given in Table 4.3. The computationally most involving part is the computation of the Galerkin inner product of left hand vectors  $\mathbf{V}_{m;\text{EIK}}(s_e)^H$  with the vectors  $\mathbf{Q}_{\text{fine}}\mathbf{V}_{m;\text{EIK}}(s_e)$ . Even for this relatively small example the computational cost of the solving a coarse system and projecting the ROM is smaller than evaluating the equation on a fine grid. We infer that especially for large-scale models these computation times become negligible with respect to basis construction, which scales worse. The phase-preconditioning approach drastically reduces the vertical communication of the algorithm as only coarse grid amplitudes and phases need to be transferred to all workers instead of fine grid RKS vectors. The ROM is essentially compressed and storage is drastically reduced.

In summary, the computational cost is shifted from the poorly scalable basis construction to the highly scalable evaluation stage where inner products can be computed in an embarrassingly parallel fashion on multiple GPUs. We should also mention significant storage reduction due to phase-preconditioning as the amplitude basis is smaller than the standard RKS basis for the same accuracy and is stored on the coarse grid only, which significantly reduces vertical communication.

<sup>4</sup>Solved using UMFPACK v 5.4.0 on a 4-Core Intel i5-4670 CPU@3.40 GHz with parallel BLAS level-3 routines.

<sup>5</sup>Double precision python implementation on an Nvidia GTX 1080 Ti.



## 4.9. CONCLUSIONS

In this chapter we have introduced phase-preconditioned rational Krylov subspaces for model order reduction and compression of wave propagation in unbounded domains targeting problems with large propagation distances. Preconditioning is achieved by splitting the RKS into incoming and outgoing waves and factoring out strongly-oscillating phase-terms using the WKB approximation. The remaining slowly-varying amplitude terms are SVD-compressed and then used in the construction of the preconditioned projection space via combinations of the singular vectors of the compressed space and the WKB phase terms computed for different inputs (sources). Finally, the ROM is evaluated via structure-preserving model reduction.

Phase-preconditioning has multiple objectives, namely, reduction of the number of required interpolation points, right-hand sides, and spatial discretization. The number of interpolation points needed for a non-preconditioned RKS method is fundamentally limited by the Nyquist frequency. However, phase-preconditioning weakens this dependence of the interpolation points on the Nyquist limit. We quantified this effect for one-dimensional SISO problems with piecewise constant coefficients, where the PPRKS solution is exact with the number of RKS shifts equal to the number of the homogeneous layers, i.e., this number plays the same role as the problem dimensionality in a conventional RKS approach. Thus in 1D the number of interpolation points needed is independent of the Nyquist rate. We do not have a rigorous estimate for the general case of multidimensional MIMO problems. However, numerical experiments show that the positive effects of preconditioning can increase due to simultaneous reduction of interpolation points and right-hand sides. In addition, factorization significantly relaxes requirements on the discretization grids for subspace computations, which is critical for large-scale problems due to the poor scalability of available Helmholtz solvers.

Furthermore, factorization reduces the computation cost and increases the model-reduction compression factor. More specifically, for a given approximation accuracy the SVD compressed amplitude space is much smaller than the RKS basis. Numerical experiments for sections of the 2D acoustic benchmark Marmousi problem show that the best cost reduction in subspace generation and compression is achieved for smooth wavespeed profiles; however, our approach is still competitive for discontinuous models and can even be adapted to include resonant substructures.

Finally, we point out that due to the tensor product-like structure of the MIMO preconditioned projection space, the dimension of this space is larger than the space of compressed amplitudes and can even be comparable to the conventional block-RKSs required for the same accuracy.

However, unlike subspace generation and compression, the projection is generally highly scalable and can be easily implemented in parallel on GPUs, leaving its computation-time insignificant.

In this chapter, we presented a prototype implementation of PPRKS for 2D problems using serial computation; however, our eventual target is high-performance computing of large-scale 3D seismic problems. In future work, we will also focus on optimal placement of the interpolation points. Specifically, we plan to investigate the approximation quality of the reduced-order models when we move the interpolation points away from the imaginary axis and into the complex plane. This can potentially improve both

the approximation properties of the preconditioned RKS for the case of bounded time intervals and the performance of Helmholtz iterative solvers used for RKS construction. As a natural extension of PPRKS, we will also focus on the modeling of wave propagation in dispersive media using PPRKS, since this will not introduce additional costs to the evaluation stage. Finally, we note that WKB-like asymptotic solutions are available for many discrete and continuous dynamical systems, which opens up a number of possibilities to extend phase-preconditioning to such problems and related matrix-function computations, in particular if the cost of the solution of the shifted systems is dominant in the RKS algorithm.





# 5

## INVERSION IN THE REDUCED-ORDER MODEL FRAMEWORK

5

*In metric, one milliliter of water occupies one cubic centimeter, weighs one gram, and requires one calorie of energy to heat up by one degree centigrade – which is 1 percent of the difference between its freezing point and its boiling point. An amount of hydrogen weighing the same amount has exactly one mole of atoms in it. Whereas in the American system, the answer to “How much energy does it take to boil a room-temperature gallon of water?” is “Go fuck yourself,” because you can’t directly relate any of those quantities.*

Josh Bazell in *Wild Thing*



## 5.1. INTRODUCTION

IN this chapter, we show the application of reduced-order modeling to inverse problems. The examples used highlight different model order reduction techniques for inversion and imaging. We distinguish two types of ROMs for this application: online and offline. In online reduced-order modeling a ROM is constructed and exploited during the imaging algorithm. On the other hand, if the construction is done prior to the imaging algorithm the reduced-model is constructed offline and exploited online.

In an imaging application one seeks to find the PDE coefficients  $m$  from measured data  $d$  using an operator  $\mathcal{F}\{m\}$  that maps  $m$  to the data. In general, a minimizer of the mismatch between simulated and measured data is desired which can be written as

$$m_{\text{inv}} = \underset{m}{\operatorname{arg\,min}} (||d - \mathcal{F}\{m\}||), \quad (5.1)$$

where  $||\cdot||$  is some suitably chosen norm. These types of imaging problems are typically ill-posed – they do not satisfy all of the Hadamard conditions for well-posedness. Thus, the solution may not exist or may not be unique or the solution does not depend continuously on the measured data. Especially in real measurement environments with noise, a different noise realization can lead to a different solution. To obtain a well-posed problem the above optimization problem needs to be regularized.

In this thesis we distinguish between inversion and imaging; inversion is generally an overdetermined problem and imaging is an underdetermined problem. As an example of online ROM three-dimensional anisotropic diffusion imaging in boreholes is considered. In the appendix B.5 we consider offline ROM inversion of nuclear neutron measurements in boreholes. The author also apologizes for the use of imperial units in the appendix section.

## 5.2. 3D ANISOTROPIC RESISTIVITY

If the parameter space for which the reduced-order model needs to be accurate is too big or unknown, one cannot construct a reduced-order model offline. However, parameter dependent ROMs can be build online in order to speed up the imaging of lossy rock formations. Due to the high electrical losses in wet rock formations, the Maxwell equations can be approximated by a diffusion equation in these types of applications, meaning that the displacements currents  $\epsilon \partial_t \mathbf{E}$  in equation (2.8) can be dropped to arrive at diffusion equation (2.37)). Furthermore, most rock formations are anisotropic in their response, and the conductivity  $\sigma$  in equation (2.37) becomes a symmetric, positive-definite  $3 \times 3$  tensor.

In this section we adapt the existing model order reduction approach (of [31]) and implement a weighted- $L_2$  regularization to preserve spatial jumps in the coefficients and solve underdetermined problems. Here, model reduction techniques are used to build a reduced-order model that interpolates the full-order model for all previously computed sets of PDE coefficients encountered during the optimization. Subsequently, the reduced-order model is used to solve the optimization problem stated in equation (5.1),

only to run a full-order computation at the found minimum and to improve the reduced-order model. This way a reduced-order model is constructed and exploited “online.”

The described problem is a particularly challenging inverse problem as it seeks to reconstruct a symmetric tensor in 3D using diffusive measurements taken on a single line (the borehole) at multiple frequencies. The problem is important as complex geologies like faults or thin shale layers are not imaged accurately within a 2D or 1D assumption, which is typically used in applications. The problem is intrinsically three-dimensional and anisotropic and only a 3D anisotropic imaging algorithm can therefore resolve these geologies.

In this section we also show that with enough measurement-data the full anisotropic tensor can be reconstructed exactly. For an imaging application, which is typically an underdetermined problem, with inexact measurements this is no longer the case. However, it is shown that one does not obtain worse inversion results using the reduced-order method compared to the solving the full PDE at every iteration. Furthermore, the problem becomes easier to regularize in the model reduction framework.

### 5.2.1. PARAMETER DEPENDENT ROM FOR IMAGING

We pose the imaging problem as an optimization problem, where we minimize a cost-function  $\Psi$  which is the sum of the squared data mismatch  $\chi^2$  and a regularization function  $\psi$  with regularization parameter  $\lambda$  as

$$\Psi[\sigma(\mathbf{x})] = \chi^2[\sigma(\mathbf{x})] + \lambda\phi[\sigma(\mathbf{x})], \quad (5.2)$$

with

$$\chi^2 = \frac{1}{2} \frac{\|W_{\text{weight}}(d - f[\sigma(\mathbf{x})])\|^2}{\|Wd\|^2}, \quad (5.3)$$

where  $W_{\text{weight}}$  is a diagonal weight-matrix with diagonal elements equal to the inverse of the standard deviation of the noise of each measurement and a frequency normalization to put all measurements on equal footing. Furthermore,  $\chi$  and  $\psi$  are scalar functions of the spatially dependent conductivity tensor  $\sigma(\mathbf{x})$ . In addition,  $d$  is a vector containing all measurements and  $f[\sigma(\mathbf{x})]$  is the measurement function that maps the conductivities to the measurements. Thus to compute  $f[\sigma(\mathbf{x})]$ , a PDE has to be solved for multiple source and receiver locations and multiple frequencies. Minimizing the cost-function is subject to the constraint that the conductivity tensor is positive-definite on the whole inversion domain.

We combine the optimization approach from [31] with the regularization presented in [1] to obtain a model order reduction algorithm for anisotropic imaging of the conductivity tensor. We take the second-order diffusion equation

$$\nabla \times \nabla \times \hat{\mathbf{E}} + \mu\sigma s\hat{\mathbf{E}} = -\nabla \times \hat{\mathbf{K}}^{\text{ext}} - s\mu\hat{\mathbf{j}}^{\text{ext}}, \quad (5.4)$$

as a starting point. We note that this diffusion equation is linear in the PDE coefficients  $\mu$  and  $\sigma$ . Discretizing the above equation on a Lebedev-grid leads to a discretized system of equations, which can be written as

$$A(C, s)u_l = b_l \quad (5.5)$$



where  $l$  denotes the source index. Here,  $A$  is a symmetric matrix in the transpose bilinear form that is linear in both  $C$  and  $s$  and has a real spectrum. In this form,  $C$  is the block-diagonal medium matrix with the finite-difference approximation of the conductivity tensor as blocks. We define the receiver matrix  $R = [r_1, \dots, r_{N_r}]$ , the source matrix  $B = [b_1, \dots, b_{N_s}]$ , and the transfer-function  $f(\sigma)$  as

$$f(\sigma) = \text{vec}([\text{vec}(R^T A(C, s_1)^{-1} B), \dots, \text{vec}(R^T A(C, s_{N_r})^{-1} B)]), \quad (5.6)$$

where we used the vectorization function  $\text{vec}(\cdot)$  which sorts a matrix lexicographically into a column vector. For a Gauss-Newton (GN) inversion approach we need to compute the Jacobian  $J = \frac{\partial f}{\partial C}$ , which has the sensitivities of the transfer-function to the medium parameters  $C$  as columns. Since our operator is self-adjoint in the transpose bilinear form, we can easily find the elements via the adjoint solutions

$$v_r = A(C, s)^{-1} r_r, \text{ for } r = 1, \dots, N_r. \quad (5.7)$$

The element of the Jacobian corresponding to the  $r$ -th receiver,  $l$ -th source and the  $j$ -th element of the medium matrix can be computed via adjoint solutions as

$$(J)_{rl,j} = -v_r^T \frac{\partial A}{\partial C_j} u_l = -v_r^T s M_\mu \frac{\partial C}{\partial C_j} u_l. \quad (5.8)$$

Thus, the computation of a Jacobian involves evaluation of the inner product of the solution and the adjoint solution over the voxels (indexed with  $j$ ) with which the imaging problem is discretized.

In our Gauss-Newton optimization we update the current  $i$ -th iterate  $C^i$  in the Gauss-Newton direction  $\Delta C_{GN}$  and perform a line search to find the step size  $\alpha$  that minimizes the cost-function in the Gauss-Newton search direction. Therefore, the update can be written as  $C^{i+1} = C^i + \alpha \Delta C_{GN}$ . During the line search, full-order PDEs need to be solved to evaluate the cost-function.

In the Model Reduction Gauss-Newton (MRGN) approach, the matrix  $A$  is projected onto the subspaces spanned by the solutions and adjoint solutions for all PDE coefficients  $C^0, \dots, C^i$  previously visited by the Gauss-Newton optimization routine. The reduced-order basis can be written as

$$W^i = [U^0, V^0, \dots, U^i, V^i] \text{ with } V^i = A(C^i, s)^{-1} R \text{ and } U^i = A(C^i, s)^{-1} B. \quad (5.9)$$

The reduced-order model follows as

$$H^i(C, s) = (W^i)^T A(C, s) W^i \quad (5.10)$$

or if the structure needs to be preserved it follows as

$$H_{sp}^i(C, s) = [\Re W^i \Im W^i]^H A(C, s) [\Re W^i \Im W^i]. \quad (5.11)$$

The ROM can now be updated efficiently by only reprojecting changes in the conductivity tensor  $C$ . Specifically, we have

$$H^i(C + \Delta C, s) = H^i(C, s) + (W^i)^T s M_\mu \Delta C W^i. \quad (5.12)$$

This reduced-order model defines a reduced-order transfer-function

$$\mathbf{f}_{\text{ROM}}^i(\sigma) = \text{vec}([\text{vec}(\mathbf{R}^T \mathbf{W}^i \mathbf{H}^i(\mathbf{C}, s_1)^{-1} (\mathbf{W}^i)^T \mathbf{B}), \dots, \mathbf{R}^T \mathbf{W}^i \mathbf{H}^i(\mathbf{C}, s_{N_f})^{-1} (\mathbf{W}^i)^T \mathbf{B}]). \quad (5.13)$$

The MRGN algorithm now uses two nested iterations. In the outer iteration the full PDE is solved and the ROM is updated. In the inner iterations a GN algorithm is employed which uses the ROM transfer-function to compute the system response and the Jacobian. Thus, no full-order PDE has to be solved inside the inner iterations. The full-order GN and MRGN algorithms are schematically shown in Figures 5.1 and 5.2, respectively.

Deflation of  $\mathbf{W}^i$  may become necessary as soon as its columns become linearly dependent or close to linearly dependent. Since our FD operator is self-adjoint, the reduced-order model transfer-function  $\mathbf{f}_{\text{ROM}}^i(\mathbf{C}, s)$  interpolates  $\mathbf{f}(\mathbf{C}, s)$  at all  $\mathbf{C}^n \forall n = 1, \dots, i$  at the operation frequencies.

**COROLLARY 5.1.** — The reduced-order transfer-function interpolates the transfer-function as

$$\mathbf{f}_{\text{ROM}}^i(\mathbf{C}, s) = \mathbf{f}^i(\mathbf{C}, s) \text{ with } \forall \{\mathbf{C}, s\} \in \{\mathbf{C}^n; s_h\} \forall \{n = 1, \dots, i; h = 1, \dots, N_f\}, \quad (5.14)$$

and

$$\frac{\partial \mathbf{f}_{\text{ROM}}^i(\mathbf{C}, s)}{\partial \mathbf{C}} = \frac{\partial \mathbf{f}^i(\mathbf{C}, s)}{\partial \mathbf{C}} \text{ with } \forall \{\mathbf{C}, s\} \in \{\mathbf{C}^n, s_h\} \forall \{n = 1, \dots, i; h = 1, \dots, N_f\}, \quad (5.15)$$

and

$$\frac{\partial \mathbf{f}_{\text{ROM}}^i(\mathbf{C}, s)}{\partial s} = \frac{\partial \mathbf{f}^i(\mathbf{C}, s)}{\partial s} \text{ with } \forall \{\mathbf{C}, s\} \in \{\mathbf{C}^n; s_h\} \forall \{n = 1, \dots, i; h = 1, \dots, N_f\}. \quad (5.16)$$

$$(5.17)$$

This is a corollary of Proposition 3.3.

The proof is analogous to the proof of Proposition 3.3 as solutions and adjoint-solutions are in the subspace and the operator is self-adjoint in the used transpose bilinear form. The reduced-order Jacobian also interpolates the full-order Jacobian as seen in equation (5.15).

### 5.2.2. REGULARIZATION FOR IMAGING

Conductivities in the subsurface are not smooth functions of the spatial coordinates as the conductivity can change abruptly when crossing layer beds or faults. Therefore, we choose to use a weighted  $L_2$  regularizer that preserves spatial jumps [1]. Let  $\sigma_{\text{ref}}$  be a reference medium or initial guess and let  $\sigma_i$  be the conductivity tensor at the  $i$ -th iterate of the Gauss-Newton optimization scheme. The regularization term of the  $i$ -th iterate is now given by

$$\phi_i(\sigma(\mathbf{x})) = \int_D b_i^{-2} \{[|\nabla \sigma(\mathbf{x}) - \sigma_{\text{ref}}|^2] + \delta_i^2\} dV \quad (5.18)$$



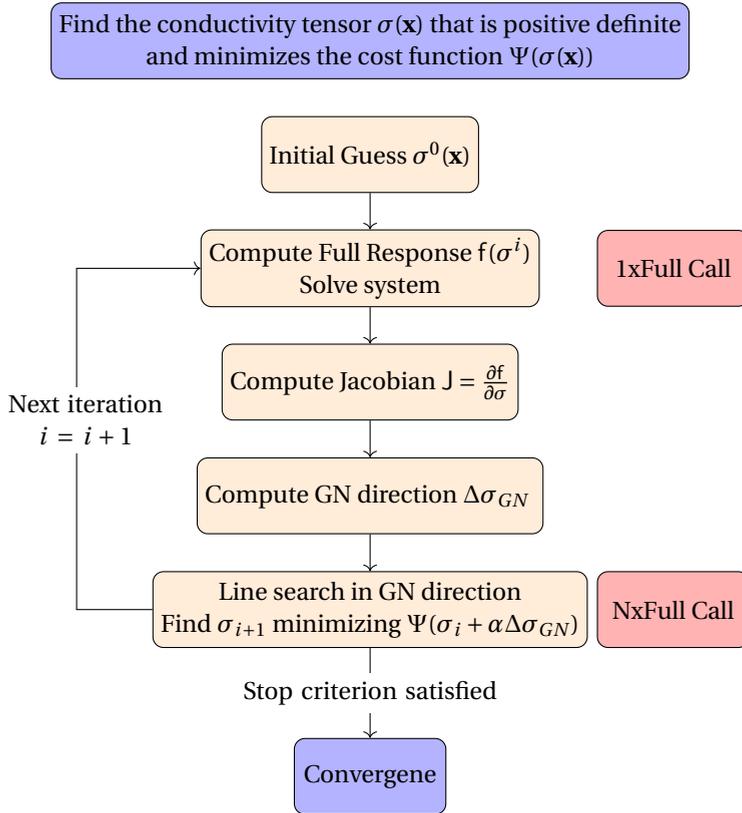


Figure 5.1: Ordinary Gauss-Newton inversion routine.

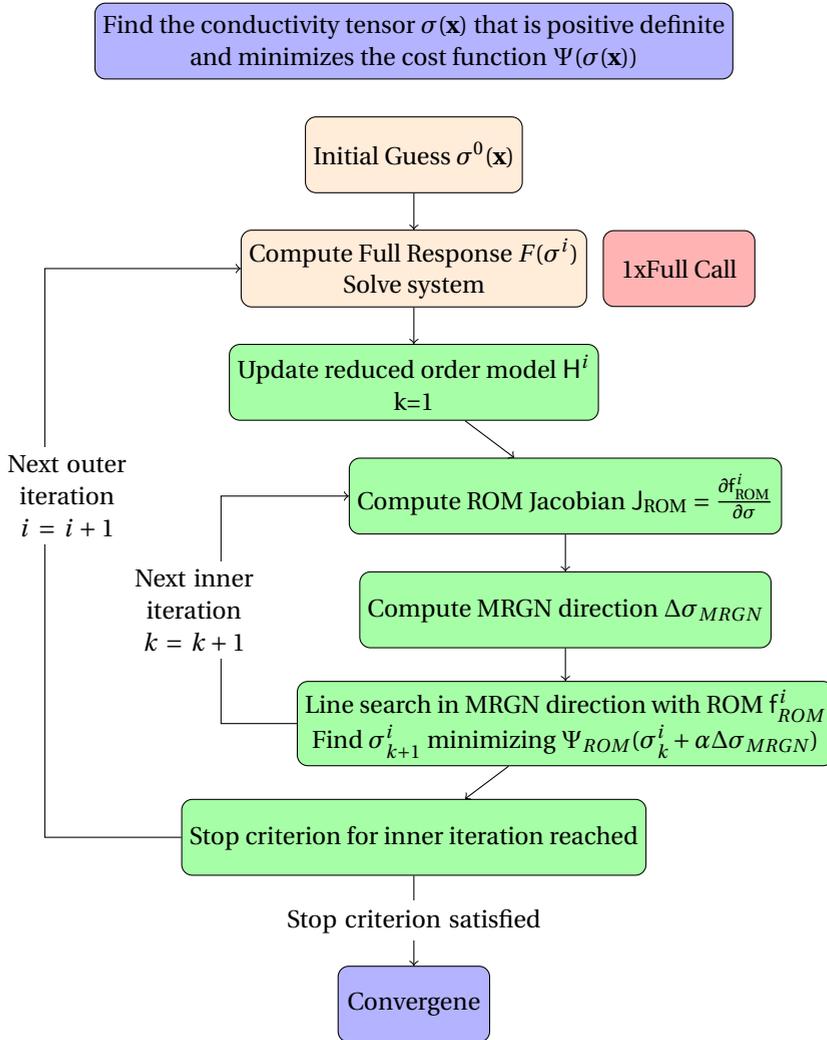


Figure 5.2: The MRGN inversion algorithm from [31].



with

$$b_i^2 = \int_D \{[|\nabla\sigma(\mathbf{x})_{i-1} - \sigma_{\text{ref}}|^2] + \delta_{i-1}^2\} dV \quad (5.19)$$

for unweighted  $L_2$  regularization with a normalization  $b_i$  that is not spatially varying. For weighted  $L_2$  regularization,  $b_i$  becomes spatially dependent as

$$b_i^2 = V[|\nabla\sigma(\mathbf{x})_{i-1} - \sigma^{\text{ref}}|^2] + \delta_{i-1}^2. \quad (5.20)$$

In these equations  $\delta_i^2 = \chi_i^2/dV$  is a small positive scalar that prevents division by zero and  $dV$  is a volume element used for discretization. We thus penalize introducing spatial variation of  $\sigma$  into the domain at every step of the optimization. In the weighted form, introducing variation in flat parts is penalized more than in areas with a large gradient in  $\sigma$ . We note that  $\phi_i(\sigma_{i-1}(\mathbf{x})) = 1$  holds for both regularizers. Furthermore, we choose a regularization parameter that reduces as the data mismatch reduces by setting  $\lambda_i = \chi_{i-1}^2$ . The MRGN algorithm adjusts the regularization term more frequently as a few inner iterations are used on one outer iteration. This makes the MRGN algorithm easier to regularize as an overdamping of the system only leads to more inner MRGN iterations which are computationally cheap.

5

### 5.2.3. INVERSION TEST

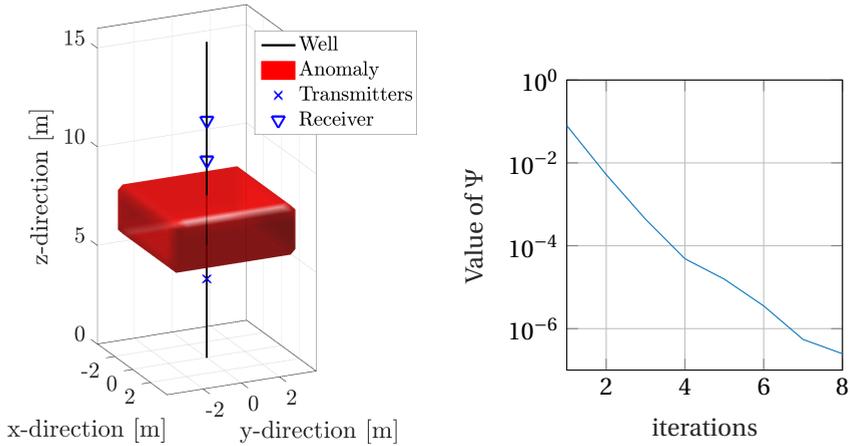
First, we test if all six components of the symmetric conductivity tensor can be reconstructed from the data, i.e. does the data hold enough information to find the exact conductivity values in a noiseless case. An anomaly of  $6 \text{ m} \times 6 \text{ m} \times 2 \text{ m}$  with a conductivity tensor of

$$\sigma = \begin{bmatrix} 3 & 0.3 & 0.1 \\ 0.3 & 2 & 0.2 \\ 0.1 & 0.2 & 1 \end{bmatrix} \text{ S/m}, \quad (5.21)$$

is embedded in a  $\sigma_{BG} = 1 \text{ S/m}$  background. Two triaxial receivers are placed above the anomaly and one triaxial source is placed below the anomaly. Three measurements are carried out at the frequencies  $f = 630, 315, 80 \text{ kHz}$ . The measurement setup is shown in Figure 5.3. As an initial guess  $2 \text{ S/m}$  is used for the diagonal elements and  $0.2 \text{ S/m}$  for the off diagonal elements.

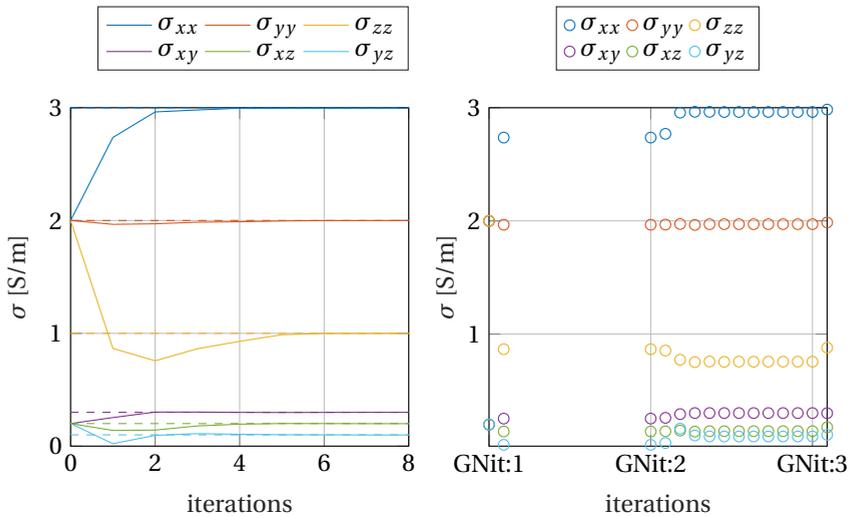
Within eight outer iterations of the Gauss-Newton algorithm, the cost-function was reduced by 4.5 orders of magnitude and all components of the conductivity tensor converged. The full-order PDE was solved 13 times in the optimization routine. To reach the same accuracy, the MRGN approach only needed seven evaluations of the full-order PDE. The reduced-order model transfer-function approximates the full-order transfer-function well enough after six iterations to reach the same accuracy. The cost-function for the MRGN approach is shown in Figure 5.3(b).

The path taken by the MRGN algorithm is visualized in Figure 5.4. The path of the outer iterations are visualized in the left figure and the path of the inner iterations is shown on the right. In the second Gauss-Newton iteration it can be seen that the Gauss-Newton direction of the  $\sigma_{zy}$  component changes sign within the inner iteration. This shows that with even two conductivity tensors as parameter-interpolation points the ROM already describes the response accurately.



(a) Measurement configuration used to invert the conductivity tensor of an anisotropic conductivity anomaly shown in red. (b) Value of the cost-function during the MRGN algorithm.

Figure 5.3: The measurement configuration with an anomaly of 6 m × 6 m × 2 m and embedded in a homogeneous isotropic background of 1 S/m (a). The reduction in cost function for the six parameter inversion test-case by the MRGN algorithm is depicted (b).

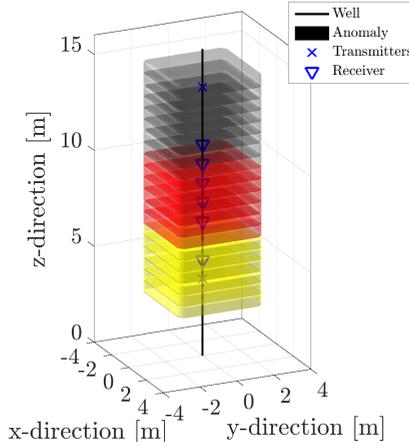


(a) Convergence path of the outer iterations. (b) Convergence path of the inner iterations.

Figure 5.4: Path of the MRGN algorithm for the six parameter inversion case.



### 5.2.4. 1D IMAGING



(a) Imaging setup for the 2D imaging test.

Figure 5.5: Source and receiver setup for the inversion test. The three layers are marked with colors black, red and yellow.

In this experiment, we consider a spatially varying anomaly to show the positive effect of MRGN on regularization. A  $4\text{m} \times 4\text{m} \times 10\text{m}$  anomaly with spatially varying conductivity tensor is placed along the borehole. The conductivity only varies as a function of the  $z$ -coordinate along the borehole. The anomaly has three sections as shown in Figure 5.5. The off-diagonal elements of the conductivity tensor are zero and the values of the diagonal elements in the three sections are

$$\sigma_{xx} = (0.05, 1, 0.1) \text{ S/m}, \quad (5.22)$$

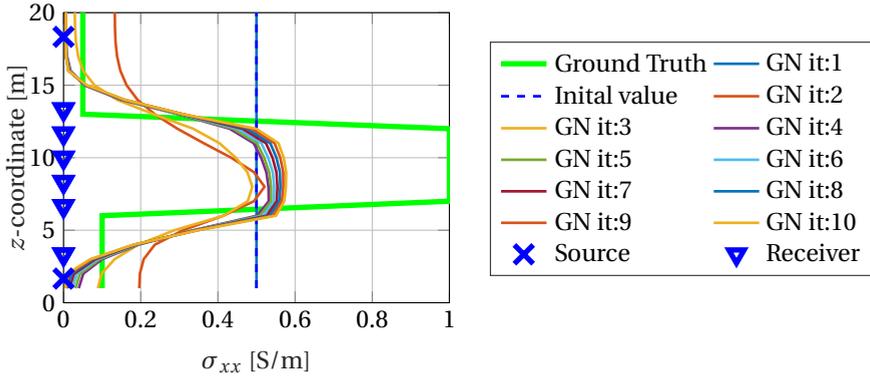
$$\sigma_{yy} = (0.1, 0.8, 0.2) \text{ S/m}, \quad (5.23)$$

$$\sigma_{zz} = (0.5, 0.5, 0.5) \text{ S/m}. \quad (5.24)$$

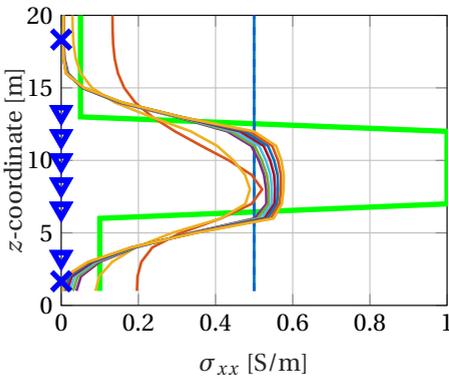
The background conductivity is taken to be  $1 \text{ S/m}$ . Six receivers in the  $z$ -direction and two sources in the  $x$ -direction are used, leading to a poor sensitivity with respect to the  $\sigma_{yy}$  component. The anomaly is discretized along the  $z$ -direction in equidistant steps of  $0.5 \text{ m}$ , where the first and last layers do not contain a source causing poor sensitivity, to see the effect of the chosen regularization.

The inversion results are shown in Figure 5.6. The cost-function was reduced by 2.5 orders of magnitude during seven outer iterations. The geometric average of the diagonal components of the inverted conductivity tensor are shown in Figure 5.6(a), the jump in conductivity along the borehole is reproduced. Since there is no receiver right across the second interface, the inverted conductivity has a slight shift in the boundary. In a real logging application the tool would move through the borehole and take multiple independent measurements which will result in a better spatial resolution.

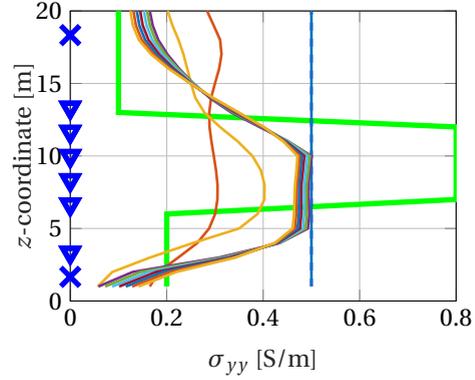
The conductivities during the outer Gauss-Newton iterations are shown in Figure 5.6(a) for  $\sigma_{xx}$  and in Figure 5.6(c) for  $\sigma_{yy}$ . The inversion results are better for the  $\sigma_{xx}$



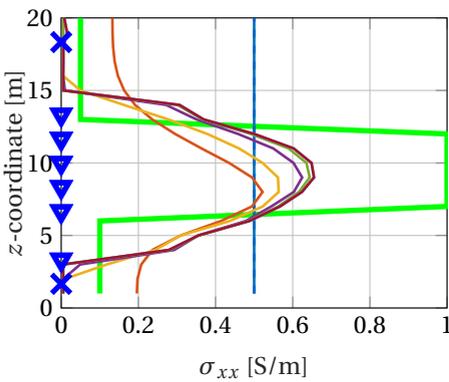
(a) Path of the optimization routine during inversion shown for the  $\sigma_{xx}$  component weighted- $L_2$  regularizer.



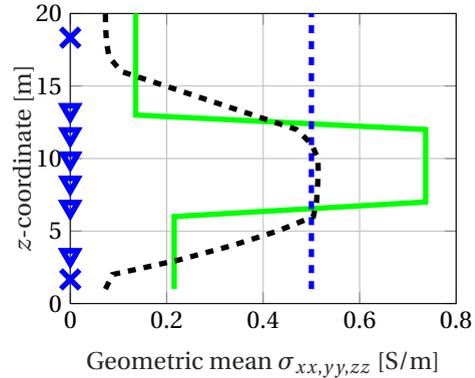
(b) Path of the optimization routine during inversion shown for the  $\sigma_{xx}$  component weighted- $L_2$  regularizer without the model order reduction approach.



(c) Path of the optimization routine during inversion shown for the  $\sigma_{yy}$  component with weighted- $L_2$  regularizer.



(d) Inversion result if the  $L_2$ -regularizer from equation (5.20) is used.



(e) Geometric average of the diagonal elements of the conductivity tensor. The final inversion result along with the ground truth.

Figure 5.6: Inversion results for the 1D imaging problem.



component as no transmitter or source in the  $y$ -direction is present, since we want to study the effect of regularization in an underdetermined imaging application.

We use a weighted  $L_2$  regularizer to preserve spatial jumps as the conductivity of a layered rock formation does not vary smoothly. The effect of the weighted and unweighted  $L_2$  regularizer can be understood when the inversion result from Figure 5.6(a) is compared to the one obtained with an unweighted  $L_2$  regularizer in Figure 5.6(b). The  $L_2$  regularizer penalizes an increase in spatial variation, whereas the weighted  $L_2$  regularizer penalizes an increase in spatial variation at positions that featured little spatial variation in the previous iteration. Therefore, the unweighted  $L_2$  regularizer smooths edges, while the weighted one preserves them. This is reproduced in our imaging results.

The MRGN approach needs six full-order PDE evaluation to reach the same accuracy as the full-order Gauss-Newton approach with 27 full-order PDE evaluations. Since the MRGN requires many internal iterations, the regularization term gets adjusted more frequently than in the full-order GN approach. This makes the MRGN algorithm easier to regularize and better behaved. The full-order GN approach takes very small steps towards the end of the algorithm as the regularization term prevents it from introducing more spatial variation as shown in Figure 5.6(b). The MRGN algorithm computed 36 Jacobians to reach the same accuracy as the full GN algorithm with 10 Jacobians.

5

### 5.2.5. 2D IMAGING

In the last example, we consider a diagonal conductivity tensor that is invariant in the  $z$ -direction along the borehole but varies in the  $xy$ -plane. In this test, we consider a diagonal conductivity tensor where all diagonal elements are equal  $\sigma_{xx} = \sigma_{yy} = \sigma_{zz}$ . However, we still invert for an anisotropic tensor and the imaging setup is less sensitive to  $\sigma_{zz}$  than to  $\sigma_{xx}$  and  $\sigma_{yy}$ . Close to the borehole there is a resistive region, whereas conductive regions are to the sides of the inversion domain. The whole domain is embedded in a conductive region with an isotropic conductivity of 1 S/m.

The inversion domain is  $15 \times 15 \times 15 \text{ m}^3$ ; however, we only discretize in the  $xy$ -plane with a uniform step size of one meter. This leads to 625 parameters that need to be estimated. We use one triaxial source and three triaxial receivers at three different frequencies  $f = 320, 127, 16 \text{ kHz}$ . The imaging configuration is shown in Figure 5.7 as a 3D illustration alongside a 2D cross section, where the borehole, indicated by the white circle, is drilled in the normal direction of the depicted plane.

In Figure 5.8 the imaging results of this configuration are shown. The reconstruction of all elements of the conductivity tensor are depicted in a plane through the three-dimensional inversion domain along with the ground truth. The reconstruction using the model reduction Gauss-Newton algorithm with a homogeneous initial guess is shown after five outer Gauss-Newton steps. The high conductivity region at the lower part of the image and the low conductivity region around the borehole are well imaged. The  $\sigma_{yy}$  component is very sensitive to spatial changes in the  $xz$ -plane and the  $\sigma_{xx}$  component to spatial changes in the  $zy$ -plane since a dipole source and receiver are mainly sensitive in the plane normal to their dipole orientation. The recovery of  $\sigma_{zz}$  is very poor as the setup is not very sensitive to this component of the conductivity tensor. The resistive region around the borehole is well imaged, but far away from the wellbore the results are less accurate. The geometrical mean of the inverted conductivities are shown

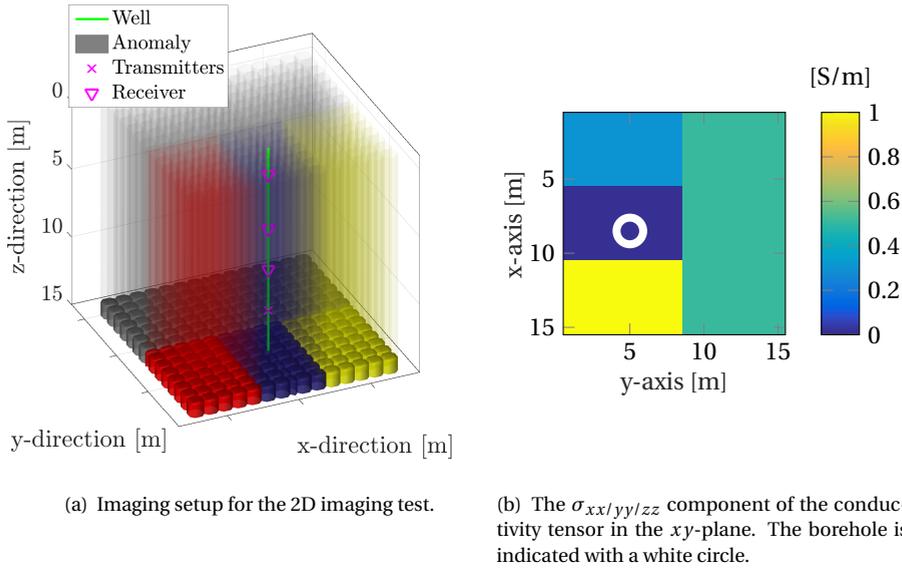


Figure 5.7: Source and receiver setup for the 2D imaging test. The measurement configuration is shown to the left and a cross section in the  $xy$ -plane is shown to the right. The medium only is invariant in the  $z$ -direction.

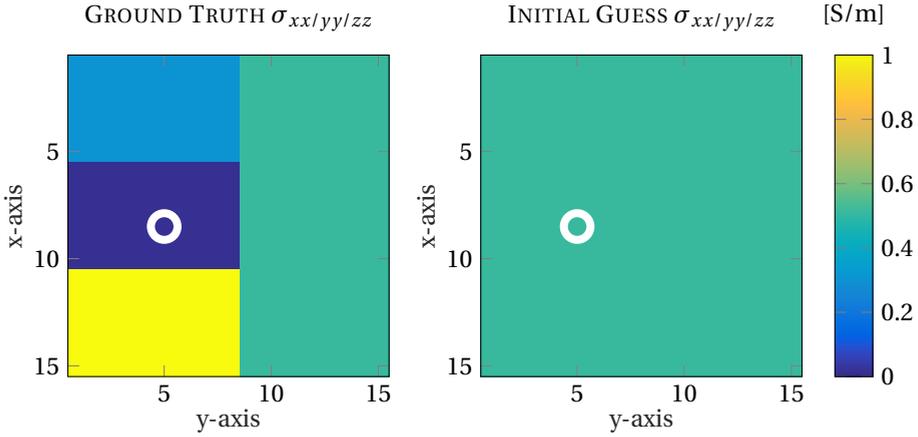
in Figure 5.8(c). The region is embedded in a very conductive background and the  $\sigma_{xx}$  component even recovers a conductive region close to the edge at  $y = 0$ .

The full GN algorithm computes 15 full-order PDE solution and the MRGN only needs 5 full-order computations to reach comparable results. Again, regularization of the MRGN method is easier as the cost-function was computed and adjusted more frequent in the MRGN algorithm.

### 5.3. CONCLUSIONS

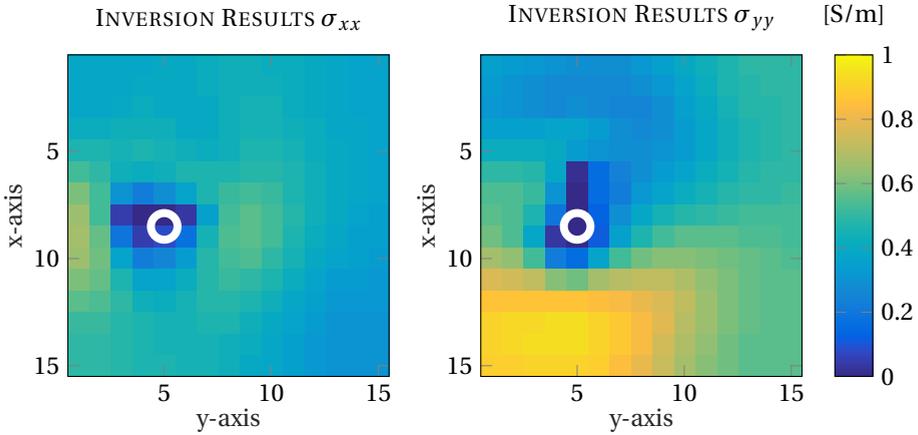
In this section, we showed how model order reduction methods can be used to speed up imaging applications by partially replacing full-order PDE solves with a fast surrogate reduced-order model. Significantly less PDE solves are needed to obtain imaging results of the same accuracy and regularization is simplified as the MRGN algorithm recomputes the cost-function more often, making it easier to introduce spatial variation of the PDE coefficients into the imaging domain. In the MRGN algorithm the work is basically shifted from solving large-scale PDE problems to computing large inner products needed for the computation of the reduced-order Jacobian. Due to recent advances in GPU architectures [4], entries of the Jacobian for all imaging parameters can be computed very fast in a parallel fashion. The conductivity tensor has to be positive definite such that a different parametrization into principal components and angles may be beneficial as the constraints on these are easier to implement than a positive definiteness constraint on a tensor.



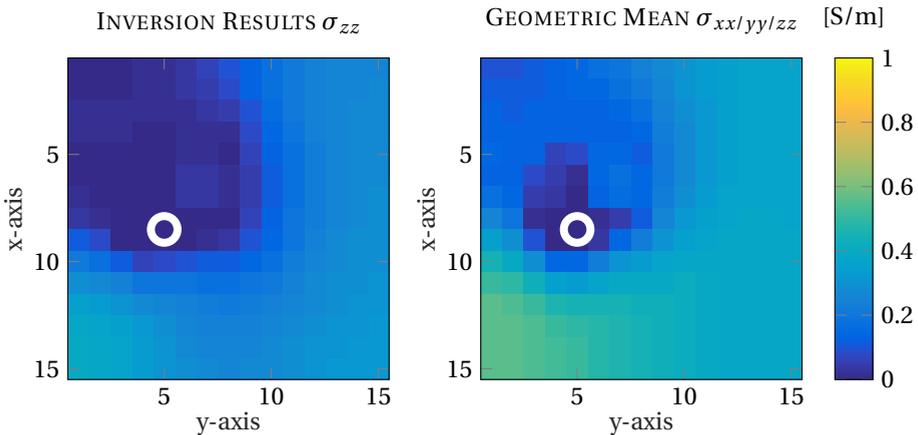


(a) Ground truth solution for  $\sigma_{xx/yy/zz}$  with an indication of the well (left). Initial guess and reference solution for all diagonal elements of the conductivity tensor (right).

5



(b) Inversion results for  $\sigma_{xx}$  and  $\sigma_{yy}$ .



(c) Inversion result for  $\sigma_{zz}$  shown alongside the inversion result of the geometric mean of all components.

Figure 5.8: Source and receiver setup for the inversion test along with the ground truth solution for all three diagonal components of the conductivity tensor.

# 6

## CONCLUSIONS

*Neque porro quisquam est, qui dolorem ipsum, quia dolor sit, amet, consectetur, adipisci velit [...].*

Cicero in *De finibus bonorum et malorum*

6

**R**EDUCED-ORDER MODELING of wave equations can lower the computational demand of existing forward and inverse wavefield simulations and allow us to solve wavefield problems, which are otherwise deemed too demanding. The aim of this thesis is to assess reduced-order modeling techniques for wave equations, to develop efficient reduced-order modeling techniques for resonant as well as propagating fields and to identify important application areas in which these techniques can have a significant impact. We have shown that with Krylov subspace model order reduction it is indeed possible to efficiently solve a wide class of time- and frequency-domain wavefield problems with applications ranging from high-frequency nano-optics to low-frequency geophysics. More specifically, the key contributions of this thesis are

1. the introduction of a novel model order reduction technique for three-dimensional electromagnetic wavefields in dispersive media,
2. an explicit reduced-order model expression for the spontaneous decay rate of a quantum emitter in three-dimensional and arbitrarily-shaped nano-resonators,
3. the development of a phase-preconditioning approach for rational Krylov model order reduction that allows us to very efficiently compute time- and frequency-domain wavefields in inhomogeneous media especially when these wavefields are characterized by large travel times, and
4. application of reduced-order modeling to a Gauss-Newton imaging technique for three-dimensional diffusive electromagnetic fields that is able to efficiently retrieve the elements of the conductivity tensor of a bounded anomaly based on limited measurement data collected on a borehole axis.



Since most wavefield problems are defined on open domains, wave propagation towards infinity needs to be taken into account when setting up the models that need to be solved on a computer. Simulating outward wave propagation can be realized by surrounding the domain of interest by a so-called Perfectly Matched Layer (PML) in which the spatial coordinates are stretched using frequency-dependent stretching functions. Since a stretching function depends in a nonlinear manner on frequency, the resulting wavefield models depend nonlinearly on frequency as well for spatial dimensions larger than one. Moreover, dispersive media introduce an additional nonlinearity to the problem, since the constitutive relations for dispersive media are generally nonlinear functions of frequency. In conclusion, the wavefield models that need to be discretized and solved on a computer depend in a nonlinear manner on frequency when simulating open wavefield problems or considering dispersive materials.

In this thesis, two approaches are followed to deal with this nonlinearity. In the first approach, we follow [28] and linearize the PML with respect to frequency. As shown in [28], an additional stability-correction procedure involving nonentire matrix functions is then necessary, but the main advantage of this approach is that it leads to models with system matrices that are frequency-independent for nondispersive media. Moreover, these matrices have certain symmetry properties due to reciprocity that may be exploited to efficiently construct reduced-order models based on standard polynomial or extended Laurent polynomial Krylov reduction techniques. In particular, polynomial reduction can be carried out using a three-term recurrence relation, while extended Krylov reduction requires a five-term recurrence relation. In this thesis we have shown that in electromagnetics this linearization approach can be extended to dispersive media. By following the linearization approach for the PML outlined above and by introducing an auxiliary field variable, the Maxwell equations and the constitutive relation can be written in a consistent first-order form with a system matrix that is again frequency-independent. This system matrix does not have the same symmetry property as the system matrix for nondispersive media, of course, but we have shown that a similar symmetry relation can be found for the system matrix in the dispersive case. In particular, we have shown that the latter system matrix is symmetric with respect to a particular Lagrangian bilinear form; it is this property that allows us to efficiently construct reduced-order models based on polynomial Krylov reduction in a similar way as for the nondispersive case. In addition, we have shown that the reduced-order models for the electromagnetic fields can be used to find an explicit reduced-order model expression for the spontaneous decay rate of a quantum emitter that is located in an arbitrarily-shaped three-dimensional metallic nano-resonator. This model allows for frequency or wavelength sweeps, meaning that a single model describes the decay rate over a complete frequency or wavelength interval of interest. Finally, we have shown that by running the reduction algorithm twice, only three basis vectors of the Krylov reduction space need to be kept in memory to explicitly determine the fields of the dominant modes that are present in the resonator and are excited by a given source. Due to this work, three-dimensional resonance fields of dispersive resonators with complex geometries can now be computed in a few hours on a laptop computer. Further, expansion of field responses in quasi-normal modes can be performed without a-priori selection of modes, contrary to other approaches.

In the second approach we deal with the nonlinear system directly and apply rational Krylov reduction to the wavefield models that describe wave propagation on open domains. No linearization and stability-correction is necessary in this case, but a rational Krylov method does require us to solve the large-scale models for particular frequencies to increase the dimension of the rational Krylov space. Since solving such a model is computationally expensive in general, we aim to keep the dimension of the space as small as possible. For problems where the wavefield response is determined by a small number of Hankel singular values (in resonant structures, for example), the subspace dimension can indeed be very small and rational Krylov reduction can be very efficient. However, for propagating waves travelling over large distances this is no longer the case, since waves with large travel times are highly oscillatory in the frequency-domain and these oscillations need to be captured according to the Nyquist-Shannon sampling theorem. Consequently, the sampling rate in the frequency-domain increases as the travel time of the wave increases. Since the frequency-domain solution at a sampling frequency is required to extend the rational Krylov subspace, this approach may become prohibitively expensive for wave propagation problems with large travel times.

In this thesis, however, we have proposed a preconditioning technique for rational Krylov subspace reduction that solves this problem. Specifically, we have first decomposed the field into incoming and outgoing waves and subsequently factored out the highly oscillatory part of these waves using high-frequency asymptotics for which only the eikonal equation has to be solved. In other words, we have written the frequency-domain wavefield as a product of a strongly oscillating phase-function, which is handled analytically, and a smoothly varying amplitude-function, which is handled numerically. Moreover, we have also demonstrated that if the real and imaginary parts of these field expansions are used as frequency-dependent basis vectors, the resulting rational Krylov reduction technique is structure-preserving, that is, essential wavefield properties of the unreduced system and fields (causality, Schwarz reflection principle, etc.) carry over to the reduced system and models. The models can also extrapolate in frequency, since the basis vectors are frequency dependent. In addition, we have demonstrated that significant additional wavefield compression is possible, since the amplitude-functions are smooth and only weakly dependent on the source location. Specifically, for configurations with multiple sources and receivers we have shown that redundancy in the set of amplitude-functions that is used to expand the fields can be removed via the singular value decomposition leading to models of an even smaller order. Finally, we have shown that smooth amplitude-functions allow us to carry out our field computations on spatial grids that are much coarser than the grids required if we solve for the frequency-domain fields directly. In conclusion, factoring out the strongly oscillating phase term has important consequences and enables us to reduce the model in multiple ways. Our numerical experiments have shown that the best reduction results are obtained for smoothly varying inhomogeneous media, which is to be expected since the high-frequency field expansions used to factor out the strongly oscillating phase term are more accurate for smoothly varying media and the resulting amplitude-functions are obviously smoother as well. For non-smooth media with jumps in the wavespeed profile, the high-frequency field expansions used for preconditioning are no longer valid, but our numerical results show that the proposed phase-preconditioned reduction approach still exhibits excel-



lent approximation properties even for such non-smooth wavespeed profiles.

To illustrate the effectiveness of model order reduction in imaging applications, we addressed the problem of finding the coefficients of the conductivity tensor of a bounded three-dimensional anomaly using diffusive electromagnetic fields and limited data sets collected on the axis of a borehole. We have shown how model order reduction methods can be used to speed up this large-scale and ill-posed imaging problem. In particular, we have shown that significant speed ups can be realized in a Gauss-Newton minimization framework by partially replacing full-order PDE solves with surrogate reduced-order models that can be evaluated very efficiently. Furthermore, compared to a standard Gauss-Newton approach, significantly fewer forward solutions are needed to obtain imaging results of the same accuracy with this method. Regularization is also simplified, since the Model Reduction Gauss-Newton (MRGN, [30]) algorithm evaluates the cost-function more often, making it easier to introduce spatial variation of the PDE coefficients into the imaging domain. Finally, by incorporating model order reduction into the Gauss-Newton algorithm, work is basically shifted from solving large-scale forward problems to computing large inner products needed for the computation of the reduced-order Jacobian. An important consequence of this feature is that the entries of the Jacobian can be computed very efficiently in parallel due to recent advances in GPU architectures [4]. In conclusion, the MRGN algorithm leads to substantial computational savings compared with a standard Gauss-Newton approach and a significant reduction in computation time as well, especially when implemented on a GPU architecture.

# A

## APPENDIX – PHASE-PRECONDITIONING

In this appendix we derive expressions for the wave amplitudes as given by equations (4.17) and (4.18) in Chapter 4.

### A.1. THE FIELD AMPLITUDES IN TWO DIMENSIONS

As is well known, the two-dimensional Laplace domain Greens function for a homogeneous medium with constant wavespeed  $c$  that describes causal wave propagation is given by

$$G(R, s) = \frac{1}{2\pi} \mathcal{K}_0(sR/c) = \frac{1}{2\pi} g(sR/c),$$

where  $R$  is the distance from the point where the two-dimensional Dirac distribution is active to the point of observation and  $g(z) = \mathcal{K}_0(z)$  is the modified Bessel function of the second kind and order zero. In Chapter 4, this Bessel function was used to realize the field splitting (see equation (4.15))

$$u = c_{\text{out}} g(sT) + c_{\text{in}} g(-sT),$$

where  $T$  is the eikonal that follows from the eikonal equation for the inhomogeneous medium of interest. The above splitting is obviously not unique and an additional relation is required to uniquely determine the wave amplitudes  $c_{\text{in}}$  and  $c_{\text{out}}$ . To this end, we take the spatial derivative of  $u$  in the direction of the rays  $\nabla T$  and we impose the condition

$$(\nabla T \cdot \nabla c_{\text{out}}) \mathcal{K}_0(sT) + (\nabla T \cdot \nabla c_{\text{in}}) \mathcal{K}_0(-sT) = 0.$$

This leads to the system of equations

$$\begin{bmatrix} \mathcal{K}_0(-sT) & \mathcal{K}_0(sT) \\ -\mathcal{K}_1(-sT) & \mathcal{K}_1(sT) \end{bmatrix} \begin{bmatrix} c_{\text{in}} \\ c_{\text{out}} \end{bmatrix} = \begin{bmatrix} u \\ -\frac{v^2}{s} \nabla T \cdot \nabla u \end{bmatrix}. \quad (\text{A.1})$$



A

Solving this system for the amplitudes  $c_{\text{in}}$  and  $c_{\text{out}}$ , we find

$$c_{\text{in}} = \frac{1}{W(sT)} \left[ \mathcal{K}_1(sT)u + \frac{v^2}{s} \mathcal{K}_0(sT) \nabla T \cdot \nabla u \right] \quad (\text{A.2})$$

and

$$c_{\text{out}} = \frac{1}{W(sT)} \left[ \mathcal{K}_1(-sT)u - \frac{v^2}{s} \mathcal{K}_0(-sT) \nabla T \cdot \nabla u \right], \quad (\text{A.3})$$

where the determinant is given by

$$W(z) = \mathcal{K}_0(-z)\mathcal{K}_1(z) + \mathcal{K}_0(z)\mathcal{K}_1(-z).$$

This expression for the determinant can be simplified using recurrence relations for modified Bessel functions of the second kind and the asymptotic behavior of these functions for large arguments. Specifically, if we compute the derivative of  $W$  with respect to  $z$  and use the recurrence relations 10.29.2 and 10.29.3 of [19], we find that the determinant satisfies

$$\frac{dW}{dz} = -\frac{W}{z}. \quad (\text{A.4})$$

Furthermore, for the modified Bessel functions we have

$$\mathcal{K}_\nu(z) \approx \sqrt{\pi/(2z)} \exp(-z)$$

as  $z \rightarrow \infty$  in  $|\arg(z)| < \frac{3}{2}\pi$  with a principle branch that corresponds to the principle value of the square root (see [19]). From this asymptotic behavior and equation (A.4) the determinant is now obtained as

$$W(z) = \begin{cases} \frac{i\pi}{z} & \text{if } \Im(z) > 0 \\ -\frac{i\pi}{z} & \text{if } \Im(z) < 0 \end{cases} \quad (\text{A.5})$$

and we arrive at the final expressions for the amplitudes as

$$c_{\text{in}} = \frac{sT}{\text{sign}[\Im(s)]i\pi} \left[ \mathcal{K}_1(sT)u + \frac{v^2}{s} \mathcal{K}_0(sT) \nabla T \cdot \nabla u \right] \quad (\text{A.6})$$

and

$$c_{\text{out}} = \frac{sT}{\text{sign}[\Im(s)]i\pi} \left[ \mathcal{K}_1(-sT)u - \frac{v^2}{s} \mathcal{K}_0(-sT) \nabla T \cdot \nabla u \right]. \quad (\text{A.7})$$

### A.1.1. BEHAVIOR OF THE AMPLITUDE FUNCTIONS AT $T(\mathbf{x}_S) = 0$

In order to obtain the correct field splitting at the source location, we need to analyze the behavior of the amplitude functions  $c_{\text{in/out}}$  near the source location. To carry out this analysis, we make use of the limiting forms (see 10.30.2 and 10.30.3 of [19])

$$\mathcal{K}_0(z) \approx -\ln z \quad \text{and} \quad \mathcal{K}_1(z) \approx 1/z \quad \text{as } z \rightarrow 0$$

to obtain

$$\lim_{sT \rightarrow 0} sT \mathcal{K}_0(\pm sT) = 0 \quad \text{and} \quad \lim_{sT \rightarrow 0} sT \mathcal{K}_1(\pm sT) = \pm 1.$$

With these results, the amplitude functions at the source location immediately follow as

$$c_{\text{in}}(\mathbf{x}_s) = \frac{u(\mathbf{x}_s)}{\text{sign}[\tilde{\mathcal{I}}\mathbf{m}(s)]i\pi} \quad \text{and} \quad c_{\text{out}}(\mathbf{x}_s) = -\frac{u(\mathbf{x}_s)}{\text{sign}[\tilde{\mathcal{I}}\mathbf{m}(s)]i\pi}. \quad (\text{A.8})$$

This result leads to interpolation of the original wavefield  $u$ , since the singularities of  $\mathcal{K}_0(sT)$  and  $\mathcal{K}_0(-sT)$  cancel out.





# B

## APPENDIX – NUCLEAR SIGMA IMAGING

In oil exploration, wells are drilled in places with high prospect for hydrocarbon deposits. For certain reservoirs wells are drilled at high angle or close to horizontal in order to maximize the reservoir contact and stay inside a thin target zone. To image the formations and allow for accurate well placement or evaluate physical properties of reservoirs, the formations are logged during the drilling – so-called Logging-While-Drilling (LWD)[40]. For the formation evaluation, multi-physics LWD platforms combining multiple nuclear and resistivity measurements are often used in high-angle and horizontal well [41], in order to determine formation dip and structure as well as porosity and water saturation.

For instance, as part of the measurement system the resistivity of the formation is derived from electromagnetic measurements (as shown in Chapter 5), the density and porosity from  $\gamma$ -ray measurements and the position of the logging tool in the borehole from a caliper measurement, derived from nuclear density. As part of the such an LWD system, high-energy neutrons are shot into the formation adjacent to the logging tool. The neutrons scatter in the formation, lose energy to become thermal neutrons and produce  $\gamma$ -rays. The thermal neutrons and  $\gamma$ -rays that return to the tool are measured in detectors with three different spacings from the neutron source. The various spacings allow one to obtain depth information about the scattering events. The detectors are called long-spaced  $\gamma$ -detector (LS), short-spaced  $\gamma$ -detector (SS) and the detector that is closest to the source is called TN for thermal neutrons. From these measurements, the equivalent cross section for the absorption of thermal neutrons, called sigma, can be derived and measured in capture units (c.u.), which is a relative measurement. Generally, it is used to get a measurement of water saturation, detect thin beds and shallow invasion of the borehole fluid into the formation[41].

After shooting neutrons into the formation, the count-rates in the detectors decay. These decays are used to derive an effective measure called “apparent sigma” which is used by well log analysts in saturation interpretation. Direct interpretation of sigma-logs in complex formations like high angle wells, where many thin and potentially in-



vaded beds are within the measurements depth of investigation, is a difficult task.

Inversion based interpretation of the measurements can help log analysts with this task. In order to deliver a consistent petrophysical interpretation of the formation all LWD-measurements should be integrated and processed together.

**B**

For instance, nuclear density images such as  $\gamma$ -ray density-measurements based on Compton scattering are one of the most commonly used LWD measurements. Inversion of these measurements deliver accurate estimates of the formation bed densities, thickness and their dip and azimuthal orientation in complex formations [76, 60, 85]. Common model parameterization can subsequently be used to perform a model based inversion of sigma measurements, where the formation structure is defined by the nuclear density inversion.

To make the inversion feasible, a model based approach is used where layer boundaries and the dip and azimuthal orientation of the formation is obtained from nuclear density measurements. Using this information and fast forward modeling of the formation neutron scattering one can invert the sigma of in the porous rock, as well as the depth the borehole fluid invaded into the formation. The parametrization used for model based inversion is thus the invasion radius, which is the depth of invasion of the borehole fluid into the rock and the sigma of the virgin fluid that fills the porous rock beyond the invaded area. In this approach we assume that the invasion is symmetric, though it is common that the gravity segregation contributes to the draped (“tear-drop”) invasion shape [61].

Scattering of neutrons in a formation is a stochastic process and modeled with Monte Carlo N-particle algorithm[80]. These methods are computationally intense and take days even on modern personal computers, which makes sigma inversion with these models infeasible. However, one can obtain an approximate fast forward model as explained in the next section to approximately model this measurement. This class of fast forward models does not fall into the class of projection-based reduced-order models as discussed in the rest of this thesis. However, it is an offline computed fast proxy model that allows inversion for sigma.

## B.1. FAST FORWARD MODEL FOR NEUTRON SCATTERING

The fast forward model is based on previously computed Monte Carlo(MC) simulations as described in [62, 80]. The model constructs an approximate detector count-rate in inhomogeneous media from previously simulated responses and sensitivities in homogeneous media. The precomputed responses from homogeneous background media are denoted  $n_B(\mathbf{x}_R, t, \Sigma_B, HI)$  and they are time-dependent count-rates for the different detector locations  $\mathbf{x}_R$ . The count-rates are dependent on the hydrogen index ( $HI$ ) of the material, and the sigma, is denoted by  $\Sigma$ . The spatial sensitivity of a sensor to changes in  $\Sigma$  is given by the spatial flux-sensitivity function, denoted as  $FSF(\mathbf{x}_R, \mathbf{x}, \Sigma_B)$ . The flux-sensitivity functions are also obtained from Monte Carlo simulations. They also have an azimuthal dependence which is embedded in the dependence of the  $FSF$  on  $\mathbf{x}$ .

The Fast Forward Model (FFM) uses a first-order perturbation approximation, which holds in case the simulated configuration is close to the one used to compute  $n_B(\mathbf{x}_R, t, \Sigma_B)$ , where the subscript  $R$  indicates that the function is evaluated at a receiver.

Formally this perturbation in  $\Sigma$  can be written as

$$n(\mathbf{x}_R, t, \Sigma, HI) = n_B(\mathbf{x}_R, t, \Sigma_B, HI) + n_B(\mathbf{x}_R, t, \Sigma_B, HI) \int_{\mathbb{R}^3} FSF(\mathbf{x}_R, \mathbf{x}, \Sigma_B) \frac{\Delta \Sigma(\mathbf{x})}{\Sigma_B(\mathbf{x})} d\mathbf{x}. \quad (\text{B.1})$$

Thus, the FSF is the measurement sensitivity to spatial changes in sigma for a given background formation sigma ( $\Sigma_B$ ). During a measurement, neutrons are first shot into the formation and the detectors count the events in predefined non-equidistant time-bins. Only time-bins of 40–405  $\mu\text{s}$  hold information relevant for the inversion of sigma, so that only these count-rates are modeled.

Monte Carlo simulations are used to build libraries of count-rates and FSF from homogeneous media for the FFM. The decay curves are computed for a range of homogeneous background sigmas and hydrogen indices. The 3D spatial FSF is also computed using the range of sigmas. The FSF and  $n_B$  are computed independently for each of the three detectors. It should be noted that in very inhomogeneous media the model is no longer valid as the perturbation assumption of a homogeneous formation is violated. The model interpolates the libraries to provide smooth and differentiable responses of the  $n_B$  and the FSF to be used by the inversion based on Gauss-Newton optimization.

## B.2. INVERSION APPROACH

The inversion for sigma is potentially part of a larger multi-physics inversion-based workflow [77, 86, 85].

Prior to the sigma inversion the layer boundaries, their orientation in space (dip and azimuth) and the porosity of each layer is determined using nuclear density inversion [76, 60]. In addition, the inversion determines the position of the tool inside the borehole, or distance of the tool to borehole wall (so-called standoff). From a mud log the properties of the used drilling fluid are known so that the sigma of the invading fluid is known. From spectroscopy processing, the sigma of the matrix formation without any fluids can be estimated, while the hydrogen index of the formation is also known from measurements. In every layer of the formation we are therefore interested in the invasion radius and the sigma of the virgin zone fluid. In other words, how deep did the drilling fluid penetrate into the formation layer and what is the sigma of the fluid present in the formation? In the model based approach these two quantities are used in parametrization every layer. The inversion is based on a Gauss-Newton algorithm, with a trust region approach and dog-leg steps (see sec. 3.3 [54]). We minimize a quadratic cost function and weigh the terms according to the noise present in the data.

The measurement window of the sigma measurement can extend across a few thin beds, such that the sigma of adjacent beds influences the decay rates. The sigma of beds that are far away do not influence this measurement. The inversion can therefore be parallelized by slicing the whole inversion domain into multiple overlapping windows, and simultaneously inverting for the medium parameters in all intervals. In the second step, the windows are “stitched” together to obtain a consistent inversion result, and if needed a second pass of simultaneous inversion can be performed to preserve the consistency between results in neighboring intervals. The measurement noise is mainly due to the discrete nature of counting particles. Therefore, the signal to noise ratio is



related to the count-rate as  $\sqrt{n_B}$  and a lower drilling speed (rate of penetration) leads to a better signal to noise ratio.

### B.3. SYNTHETIC NINE LAYER EXAMPLE (INVERSE CRIME)

First, the inversion is tested using a synthetic formation model where the FFM is used to generate the forward data (measurement responses). After that, noise is applied to the signal using the reference signal level that corresponds to a rate of penetration of 120 ft/h. The described validation procedure is an inverse crime as the same approximate FFM is used to generate the data and in the inversion. In this 9 layer example with a minimum bed thickness of 0.5 ft example we illustrate how the inversion domain and logs are segmented into five windows, processed in parallel. The well in the test example is deviated at  $85^\circ$ . First, the modeled density data are inverted in order to obtain the formation geometry for sigma inversion. A curtain section of the inverted and measured density data is displayed in Figure B.1. The figure also gives an indication of the simulated formation, the layers two to five are thin beds.

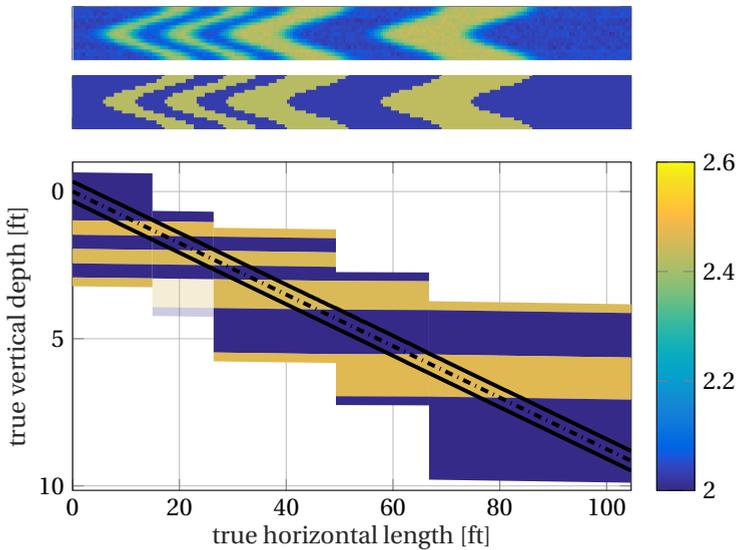


Figure B.1: Curtain section of the inverted density, showing measured and inverted density at the top. The horizontal axis in the top two images shows the azimuthal angle in the borehole.

The data shown in Figure B.2 is inverted using the standoff, boundary locations and dip obtained from density inversion and other measurements. Perfect knowledge of sigma in the invaded zone is assumed since the porosity of the invaded rock, sigma of the borehole fluid and sigma of the matrix rock are known from measurements and density inversion. However, the inversion is rather stable with respect to errors in these estimates.

The inversion was able to recover the formation properties as shown in Figure B.3. After five iterations of the Gauss-Newton algorithm the data mismatch and cost function

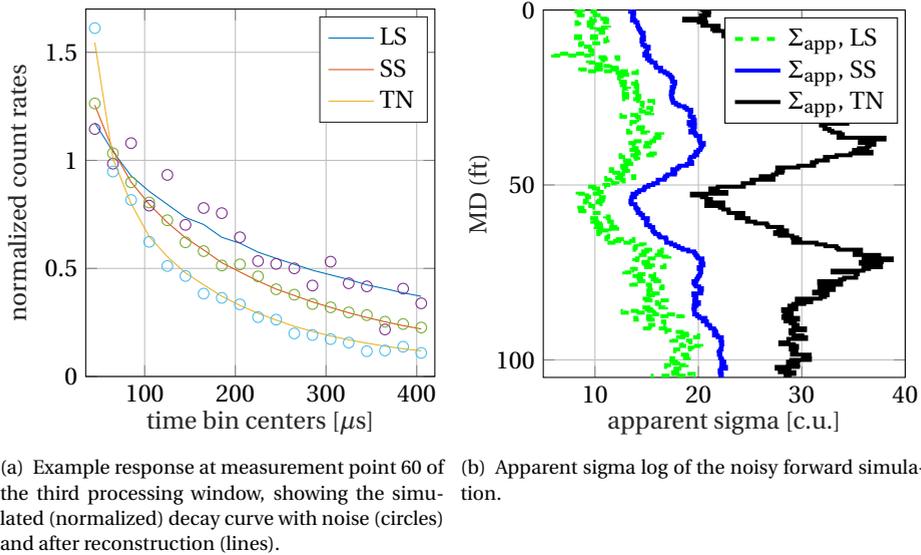


Figure B.2: Plots showing the inverted synthetic sigma data in the time-domain (a) and apparent sigma space (b).

are reduced by a factor of three in most processing windows. Small errors occur because of deep invasion, thin beds or errors in horizontal boundaries. The inversion result gives a much better understanding of the formation than the apparent sigma log from Figure B.2. After convergence of the algorithm we compute the Fischer information matrix and find that the invasion radius and the virgin sigma have a strong negative correlation. This means that the minimum found by the optimization algorithm is very shallow and recovering a larger invasion radius and smaller virgin sigma leads to only small increases in the cost function.

## B.4. INVERSION OF MONTE CARLO DATA

The nine layer inversion test was very optimistic, as we used the approximate forward model to generate noisy measurement. To really see the influence of the approximate FFM on the inversion result we use a Monte Carlo simulation to obtain more realistic logs in a high angle well. Especially interesting are the zones close to layer boundaries and in thin beds, where the linearity assumption of the FFM may break down. Furthermore, we again add the realistic measurement noise to the data. We assume perfect knowledge of boundaries and formation dip and azimuth as we do not have MC data for the density images.

A certain section of the formation simulated with the Monte Carlo solver is shown in Figure B.4. The well deviation is 80 degrees and measurements in four azimuthal sectors are taken every 0.5 ft along the well. This leads to 150 measurement points times four azimuthal directions and nineteen time-bins. The borehole has a diameter of 8.25 in



## B

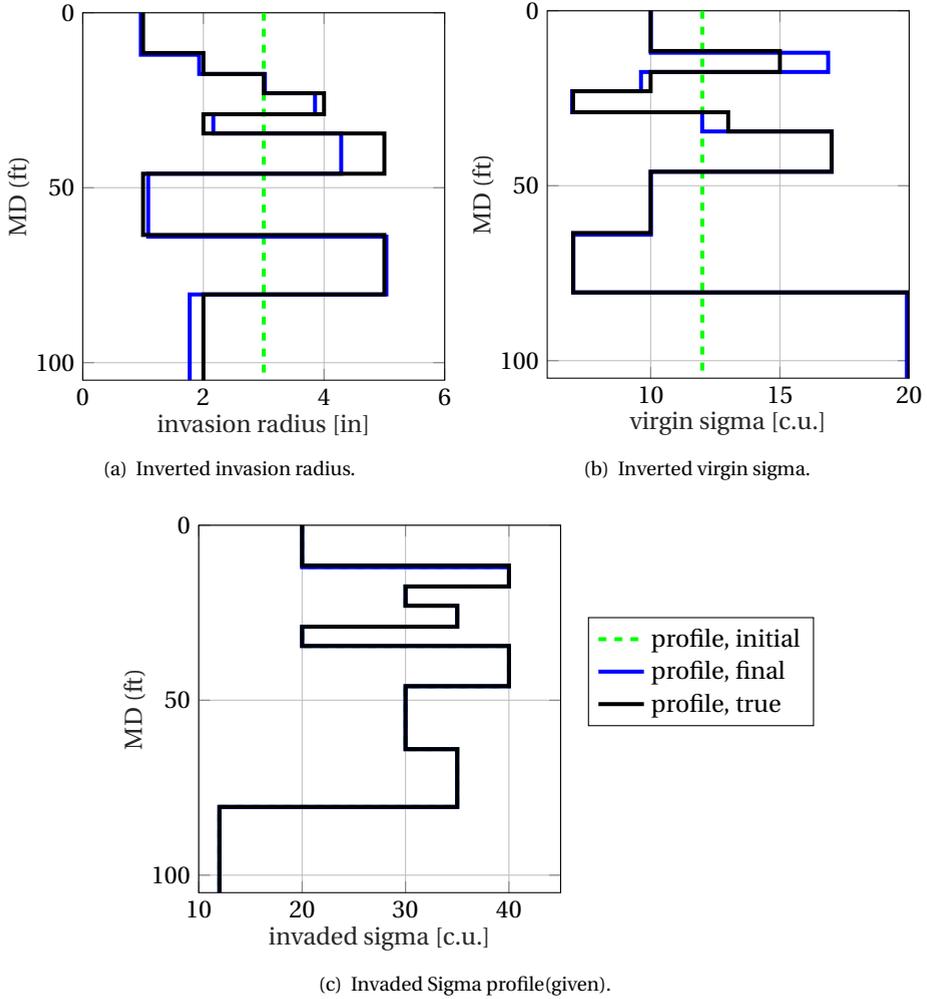


Figure B.3: Plots of the true and inverted formation parameters. The invaded sigma was not inverted but is given as reference to show the contrast in sigma.

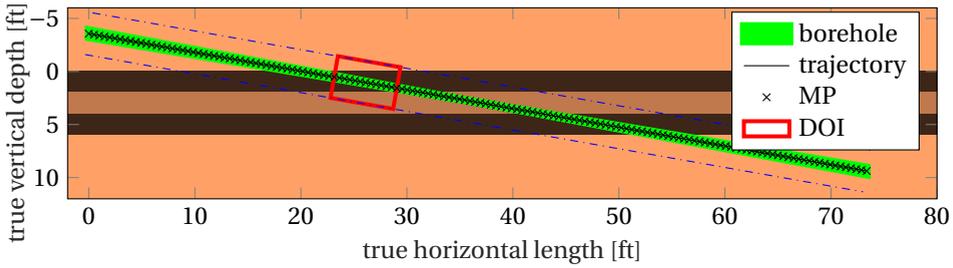


Figure B.4: Curtain view of formation used for the Monte Carlo test. Layer one and five are semi-infinite and layer two to four are 2ft wide. The indicated trajectories show the high angle well to the right and a vertical well trajectory for comparison on the left.

and we assume a centered tool. The formation consists of a thin beds filled with oil and a progressive invasion profile. The details of the formation are given in table B.1.

Table B.1: Formation details from the Monte Carlo inversion test.

Bore-hole	Layer	TST [ft]	Porosity [p.u.]	Virgin fluid	Matrix Sigma [c.u.]	Invasion radius [in]	Invaded Sigma [c.u.]	Virgin Sigma [c.u.]	Invaded H.I.	Virgin H.I.
200-ppk, Sigma 100.7, HI 0.916	0	$\infty$		pseudo-clay		0	n/a	24.3	n/a	0.2
	1	2	20	oil	7.1	6	25.9	10.0	0.184	0.192
	2	2	20	oil	7.1	4	25.9	10.0	0.184	0.192
	3	2	20	oil	7.1	2	25.9	10.0	0.184	0.192
	4	$\infty$		pseudo-clay		0	n/a	24.3	n/a	0.2

In Figure B.5 the inverted parameters and their correlation in the inversion are shown. As the initial guess for the inversion we use a sigma of 12 c.u. and an invasion radius of 3 in. The progressive invasion is well recovered; however, the inversion has problems recovering the virgin zone sigma of the deeply invaded layer. The cost function dropped over two and a half orders of magnitude within five iterations of the optimization algorithm. The sigma in the zones that were not invaded are recovered well. The error bars shown on the inversion result are derived from the model covariance matrix, linearizing the model near the solution, assuming that the noise and errors follow Gaussian distribution. They can, however, be used qualitatively. To obtain more accurate error bars and uncertainties one would need to perform a Bayesian inversion using more realistic noise models.

In Figure B.6 the apparent sigma logs of the Monte Carlo data and the FFM with the true formation data and the inverted formation parameters are shown. Clearly the mismatch between the FFM and the Monte Carlo simulation with the true formation parameters shows the shortcomings of the FFM. However, especially the short and long spaced detector are well matched by the inversion. The thermal neutron detector has the highest modeling mismatch. However, the inversion based interpretation of this log is advantageous over the apparent sigma log as the progressive invasion is clearly visible from the inversion result, contrary to the apparent sigma logs. The TN detector has the highest modeling error in the FFM and the inversion thus over-fits the count-rates from



B

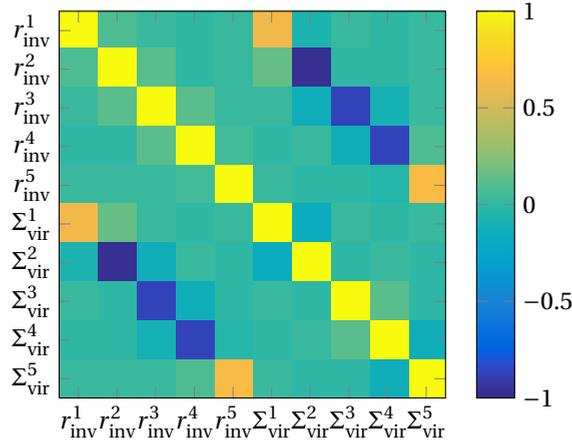
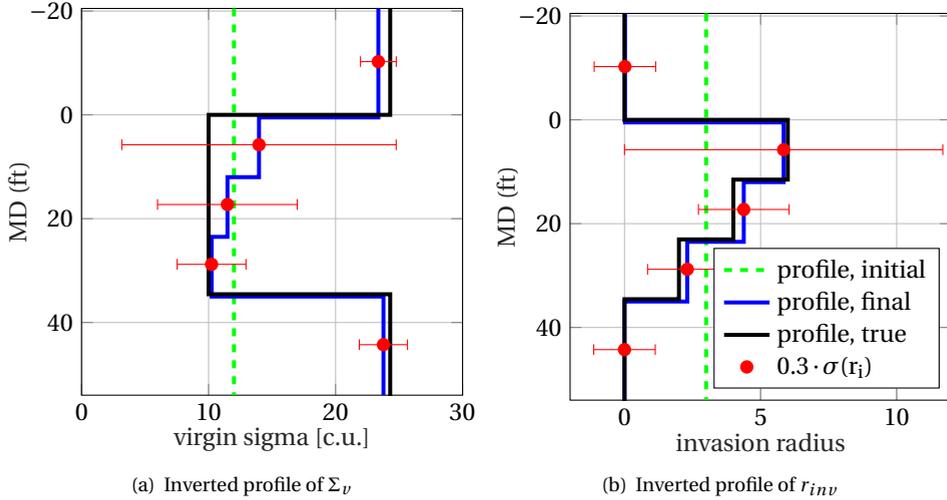


Figure B.5: Reconstruction of virgin sigma (a) and invasion radius (b) as well as the correlation of all inverted parameters in the inversion (c).

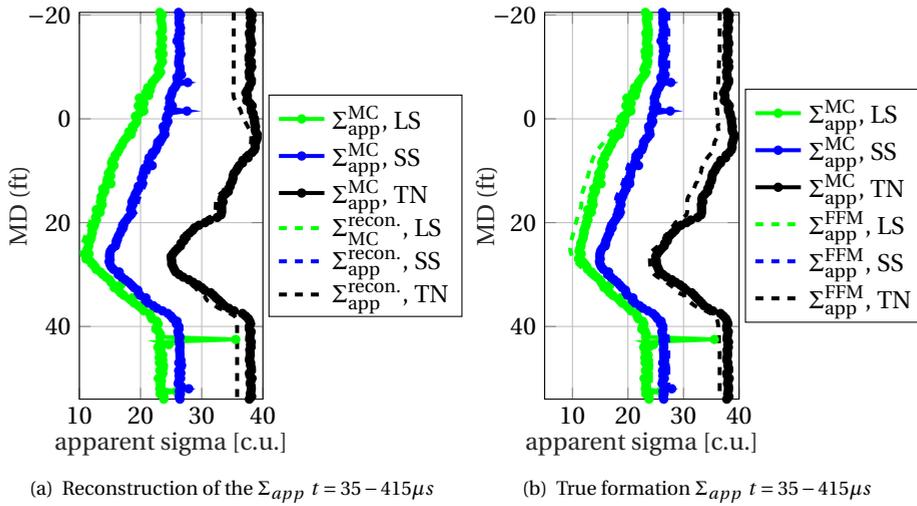


Figure B.6: Apparent sigma log of the measurement, true formation (a), and reconstructed formation (b).

this detector.

Since the well is highly deviated, the inversion parameters of neighboring beds are slightly correlated and due to the nature of the measurement, invasion radius and virgin sigma have a strong negative correlation. This experiment shows that model based inversion adds information to the log despite some shortcomings of the FFM modeling approach. The inversion results are much easier to interpret than the apparent sigma log.

## B.5. CONCLUSIONS

We developed and tested sigma inversion for multi-physics inversion in high-angle wells. It was integrated into a workflow that allows model based inversion to obtain a consisted interpretation of the measurement data. We validated the inversion using synthetic data, general-purpose Monte Carlo N-Particle data and field data. It was shown that time-domain inversion of sigma measurements allows recovery of virgin sigma in invaded beds.

Generally, deeply invaded beds and beds with small hydrogen index are causes for errors in the inversion, whereas thin beds are handled well. Within the inversion the invasion radius and the virgin sigma are negatively correlated as long as the sigma in the invasion region is larger than in the virgin zone.

The FFM can be seen as an offline computed reduced order model for the computation of neutron scattering. Performing this inverting with a Monte Carlo solver as forward model is not feasible due to the computational requirements for Monte Carlo simulations.





## REFERENCES

- [1] A. ABUBAKAR, T. M. HABASHY, V. L. DRUSKIN, L. KNIZHNERMAN, AND D. ALUM-BAUGH, *2.5D forward and inverse modeling for interpreting low-frequency electromagnetic measurements*, *Geophysics*, 73 (2008), pp. F165–F177, <https://doi.org/10.1190/1.2937466>.
- [2] A. ANTOULAS, *Approximation of Large-Scale Dynamical Systems*, Society for Industrial and Applied Mathematics, 2005, <https://doi.org/10.1137/1.9780898718713>.
- [3] A. ANTOULAS, C. BEATTIE, AND S. GUGERCIN, *Interpolatory Model Reduction of Large-Scale Dynamical Systems*, in *Efficient Modeling and Control of Large-Scale Systems*, J. Mohammadpour and K. Grigoriadis, eds., Springer US, Boston, MA, 2010, ch. 1, pp. 3–58.
- [4] H. S. AZAD, *Advances in GPU Research and Practice*, Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1st ed., 2016.
- [5] B. BECKERMANN AND L. REICHEL, *Error estimates and evaluation of matrix functions via the Faber transform*, *SIAM Journal on Numerical Analysis*, 47 (2009), pp. 3849–3883.
- [6] C. M. BENDER AND S. A. ORSZAG, *Advanced Mathematical Methods for Scientists and Engineers*, Springer-Verlag New York, 233 Spring Street, New York, NY 10013-1578, USA, 1 ed., 1999. Originally published by McGraw Hill, 1978.
- [7] P. BENNER, S. GUGERCIN, AND K. WILLCOX, *A Survey of Projection-Based Model Reduction Methods for Parametric Dynamical Systems*, *SIAM Review*, 57 (2015), pp. 483–531, <https://doi.org/10.1137/130932715>.
- [8] J.-P. BERENGER, *A perfectly matched layer for the absorption of electromagnetic waves*, *Journal of Computational Physics*, 114 (1994), pp. 185 – 200.
- [9] D. S. BINDEL, Z. BAI, AND J. W. DEMMEL, *Model Reduction for RF MEMS Simulation*, Springer Berlin Heidelberg, Berlin, Heidelberg, 2006, pp. 286–295.
- [10] L. BORCEA, V. DRUSKIN, A. V. MAMONOV, AND M. ZASLAVSKY, *A model reduction approach to numerical inversion for a parabolic partial differential equation*, *Inverse Problems*, 30 (2014), p. 125011.
- [11] R.-U. BÖRNER, O. G. ERNST, AND S. GÜTTEL, *Three-dimensional transient electromagnetic modelling using rational Krylov methods*, *Geophysical Journal International*, 202 (2015), pp. 2025–2043.



- [12] A. BOURGEOIS, M. BOURGET, P. LAILLY, M. POULET, P. RICARTE, AND R. VERSTEEG, *Marmousi, model and data*, The Marmousi Experience, European Association of Exploration Geophysicists, (1991), pp. 5–16.
- [13] N. V. BUDKO AND R. F. REMIS, *Electromagnetic inversion using a reduced-order three-dimensional homogeneous model*, *Inverse Problems*, 20 (2004), p. S17.
- [14] W. C. CHEW, *Electromagnetic theory on a lattice*, *Journal of Applied Physics*, 75 (1994), pp. 4843 – 4850, <https://doi.org/10.1063/1.355770>.
- [15] W. C. CHEW, J. M. JIN, AND E. MICHIELSSEN, *Complex coordinate stretching as a generalized absorbing boundary condition*, *Microwave and Optical Technology Letters*, 15 (1997), pp. 363–369.
- [16] E. T. CHUNG, Y. EFENDIEV, AND W. T. LEUNG, *Generalized Multiscale Finite Element Methods for Wave Propagation in Heterogeneous Media*, *Multiscale Modeling & Simulation*, 12 (2014), pp. 1691–1721, <https://doi.org/10.1137/130926675>.
- [17] B. N. DATTA, *Krylov subspace methods for large-scale matrix problems in control*, *Future Generation Computer Systems*, 19 (2003), pp. 1253 – 1263. Selected papers on Theoretical and Computational Aspects of Structural Dynamical Systems in Linear Algebra and Control.
- [18] L. DEMANET AND L. YING, *Fast wave computation via Fourier integral operators*, *Mathematics of computation*, 8 (2012), pp. 1455–1486.
- [19] *NIST Digital Library of Mathematical Functions*. <http://dlmf.nist.gov/>, Release 1.0.17 of 2017-12-22. F. W. J. Olver, A. B. Olde Daalhuis, D. W. Lozier, B. I. Schneider, R. F. Boisvert, C. W. Clark, B. R. Miller and B. V. Saunders, eds.
- [20] V. DRUSKIN, S. GÜTTEL, AND L. KNIZHNERMAN, *Near-Optimal Perfectly Matched Layers for Indefinite Helmholtz Problems*, *SIAM Review*, 58 (2016), pp. 90–116, <https://doi.org/10.1137/140966927>.
- [21] V. DRUSKIN AND L. KNIZHNERMAN, *Spectral approach to solving three-dimensional Maxwell's diffusion equations in the time and frequency domains*, *Radio Science*, 29 (1994), pp. 937–953, <https://doi.org/10.1029/94RS00747>.
- [22] V. DRUSKIN AND L. KNIZHNERMAN, *Extended Krylov Subspaces: Approximation of the Matrix Square Root and Related Functions*, *SIAM Journal on Matrix Analysis and Applications*, 19 (1998), pp. 755–771.
- [23] V. DRUSKIN AND L. KNIZHNERMAN, *Extended Krylov Subspaces: Approximation of the Matrix Square Root and Related Functions*, *SIAM Journal on Matrix Analysis and Applications*, 19 (1998), pp. 755–771, <https://doi.org/10.1137/S0895479895292400>.
- [24] V. DRUSKIN, L. KNIZHNERMAN, AND M. ZASLAVSKY, *Solution of Large Scale Evolutionary Problems Using Rational Krylov Subspaces with Optimized Shifts*, *Society of Industrial and Applied Mathematics Journal on Scientific Computing*, 31 (2009), pp. 3760–3780.

- [25] V. DRUSKIN, A. V. MAMONOV, A. E. THALER, AND M. ZASLAVSKY, *Direct, Nonlinear Inversion Algorithm for Hyperbolic Problems via Projection-Based Model Reduction*, SIAM Journal on Imaging Sciences, 9 (2016), pp. 684–747, <https://doi.org/10.1137/15M1039432>.
- [26] V. DRUSKIN, A. V. MAMONOV, AND M. ZASLAVSKY, *Multiscale S-Fraction Reduced-Order Models for Massive Wavefield Simulations*, Multiscale Modeling & Simulation, 15 (2017), pp. 445–475, <https://doi.org/10.1137/16M1072103>.
- [27] V. DRUSKIN, A. V. MAMONOV, AND M. ZASLAVSKY, *Multiscale S-Fraction Reduced-Order Models for Massive Wavefield Simulations*, Multiscale Modeling & Simulation, 15 (2017), pp. 445–475, <https://doi.org/10.1137/16M1072103>.
- [28] V. DRUSKIN AND R. REMIS, *A Krylov Stability-Corrected Coordinate-Stretching Method to Simulate Wave Propagation in Unbounded Domains*, SIAM Journal on Scientific Computing, 35 (2013), pp. B376–B400.
- [29] V. DRUSKIN, R. F. REMIS, AND M. ZASLAVSKY, *An extended Krylov subspace model-order reduction technique to simulate wave propagation in unbounded domains*, Journal of Computational Physics, 272 (2014), pp. 608 – 618.
- [30] V. DRUSKIN AND M. ZASLAVSKY, *On combining model reduction and Gauss-Newton algorithms for inverse partial differential equation problems*, Inverse Problems, 23 (2007), p. 1599.
- [31] V. DRUSKIN AND M. ZASLAVSKY, *On combining model reduction and Gauss-Newton algorithms for inverse partial differential equation problems*, Inverse Problems, 23 (2007), p. 1599.
- [32] V. DRUSKIN AND M. ZASLAVSKY, *On convergence of Krylov subspace approximations of time-invariant self-adjoint dynamical systems*, Linear Algebra and its Applications, 436 (2012), pp. 3883 – 3903. Special Issue dedicated to Heinrich Voss's 65th birthday.
- [33] V. DRUSKIN AND M. ZASLAVSKY, *On convergence of Krylov subspace approximations of time-invariant self-adjoint dynamical systems*, Linear Algebra and its Applications, 436 (2012), pp. 3883 – 3903. Special Issue dedicated to Heinrich Voss's 65th birthday.
- [34] B. ENGQUIST AND L. YING, *Sweeping preconditioner for the Helmholtz equation: moving perfectly matched layers*, Multiscale Modeling & Simulation, 9 (2011), pp. 686–710.
- [35] R. FREUND AND N. NACHTIGAL, *Software for simplified Lanczos and QMR algorithms*, Applied Numerical Mathematics, 19 (1995), pp. 319 – 341.
- [36] J. L. D. F.R.S., *Lxiii. on the theory of the magnetic influence on spectra; and on the radiation from moving ions*, The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science, 44 (1897), pp. 503–512, <https://doi.org/10.1080/14786449708621095>.



- [37] C. FURSE, D. CHRISTENSEN, AND C. DURNEY, *Basic Introduction to Bioelectromagnetics*, CRC Press, Boca Raton, 233 Spring Street, New York, NY 10013-1578, USA, 2 ed., 2009.
- [38] A. GEORGE AND E. NG, *On the complexity of sparse qr and lu factorization of finite-element matrices*, *SIAM Journal on Scientific and Statistical Computing*, 9 (1988), pp. 849–861, <https://doi.org/10.1137/0909057>.
- [39] T. GÖCKLER AND V. GRIMM, *Convergence Analysis of an Extended Krylov Subspace Method for the Approximation of Operator Functions in Exponential Integrators*, *SIAM Journal on Numerical Analysis*, 51 (2013), pp. 2189–2213.
- [40] R. GRIFFITHS, *Well Placement Fundamentals*, Schlumberger, 5599 San Felipe, Houston, Texas, United States, first ed., 2009.
- [41] R. GRIFFITHS, *EcoScope: User's Guide*, Schlumberger, 5599 San Felipe, Houston, Texas, United States, first ed., 2010.
- [42] S. GÜTTEL, *Rational Krylov methods for operator functions*, PhD thesis, Institut für Numerische Mathematik und Optimierung, Technische Universität Bergakademie Freiberg, 2010.
- [43] E. HABER AND S. MACLACHLAN, *A fast method for the solution of the Helmholtz equation*, *Journal of Computational Physics*, 230 (2011), pp. 4403 – 4418.
- [44] P. HUANG, S. SERAFIN, AND J. O. SMITH, *A 3-D waveguide mesh model of high-frequency violin body resonances*, in *Proc. Int. Computer Music Conf.*, Berlin, Germany, Aug. 2000, pp. 86–89.
- [45] I. GOHBERG AND P. LANCASTER AND L. RODMAN, *Matrix Polynomials*, SIAM, Philadelphia, PA 19104-2688 USA, 2009.
- [46] A. ISERLES, S. NØRSETT, AND S. OLVER, *Highly oscillatory quadrature: The story so far*, in *Numerical mathematics and advanced applications*, Springer, 2006, pp. 97–118.
- [47] C. JAGELS AND L. REICHEL, *Recursion relations for the extended Krylov subspace method*, *Linear Algebra and its Applications*, 434 (2011), pp. 1716 – 1732.
- [48] M. E. KILMER AND E. DE STURLER, *Recycling Subspace Information for Diffuse Optical Tomography*, *SIAM Journal on Scientific Computing*, 27 (2006), pp. 2140–2166, <https://doi.org/10.1137/040610271>.
- [49] S. KIM AND J. PASCIAK, *The computation of resonances in open systems using a perfectly matched layer*, *Mathematics of Computation*, 78 (2009), pp. 1375–1398.
- [50] L. KNIZHNERMAN, V. DRUSKIN, AND M. ZASLAVSKY, *On Optimal Convergence Rate of the Rational Krylov Subspace Reduction for Electromagnetic Problems in Unbounded Domains*, *SIAM Journal on Numerical Analysis*, 47 (2009), pp. 953–971, <https://doi.org/10.1137/080715159>.

- [51] L. KNIZHNERMAN, V. DRUSKIN, AND M. ZASLAVSKY, *On Optimal Convergence Rate of the Rational Krylov Subspace Reduction for Electromagnetic Problems in Unbounded Domains*, SIAM Journal on Numerical Analysis, 47 (2009), pp. 953–971, <https://doi.org/10.1137/080715159>.
- [52] A. N. KRYLOV, *On the numerical solution of the equation by which in technical questions frequencies of small oscillations of material systems are determined*, Izvestiya Akademii Nauk SSSR, Otdelenie Matematicheskikh i Estestvennykh Nauk 7:4, (1931), p. 491539.
- [53] Y. E. LI AND L. DEMANET, *Phase and amplitude tracking for seismic event separation*, Geophysics, 80 (2015), pp. WD59–WD72.
- [54] K. MADSEN, H. NIELSEN, AND O. TINGLEFF, *Method for non-linear least squares problems*, Lecture notes Technical University of Denmark, (2004).
- [55] C. MAGRUDER, C. BEATTIE, AND S. GUGERCIN, *Rational Krylov methods for optimal  $L_2$  model reduction*, in 49th IEEE Conference on Decision and Control (CDC), Dec 2010, pp. 6797–6802, <https://doi.org/10.1109/CDC.2010.5717454>.
- [56] S. MOSKOW, V. DRUSKIN, T. HABASHY, P. LEE, AND S. DAVYDYCHEVA, *A Finite Difference Scheme for Elliptic Equations with Rough Coefficients Using a Cartesian Grid Nonconforming to Interfaces*, SIAM Journal on Numerical Analysis, 36 (1999), pp. 442–464, <https://doi.org/10.1137/S0036142997318541>.
- [57] W. A. MULDER AND R.-E. PLESSIX, *Time- versus frequency-domain modelling of seismic wave propagation*, EAGE, 64th Annual Conference and Exhibition, Extended Abstract E015, (2002), p. 2730.
- [58] L. NOVOTNY AND B. HECHT, *Principles of Nano-Optics*, Cambridge University Press, 2 ed., 2012, <https://doi.org/10.1017/CB09780511794193>.
- [59] M. O’CONNELL, M. E. KILMER, E. DE STURLER, AND S. GUGERCIN, *Computing Reduced Order Models via Inner-Outer Krylov Recycling in Diffuse Optical Tomography*, SIAM Journal on Scientific Computing, 39 (2017), pp. B272–B297, <https://doi.org/10.1137/16M1062880>.
- [60] D. OMERAGIC, V. POLYAKOV, S. SHETTY, T. HABASHY, A. MAHESH, T. FRIEDEL, AND J.-M. DENICHO, *Integration Of Well Logs And Reservoir Geomodels For Formation Evaluation In High-Angle And Horizontal Wells*, SPWLA 53th Annual Logging Symposium, Colorado Springs, Colorado (2011).
- [61] D. OMERAGIC, L. SUN, V. POLYAKOV, Y.-H. CHEN, X. CAO, T. HABASHY, T. VIK, J. RASMUS, AND J.-M. DENICHO, *Characterizing Teardrop Invasion in Horizontal Wells in the Presence of Boundaries using LWD Directional Resistivity Measurements*, SPWLA 54th Annual Logging Symposium, New Orleans, Louisiana (2013).
- [62] E. ORTEGA, C. TORRES-VERDIN, AND W. E. PREGG, *Rapid forward modeling of multidetector logging-while-drilling Sigma measurements*, Geophysics, 79 (2014), pp. D253–D273.



- [63] V. PEREYRA, *Model order reduction with oblique projections for large scale wave propagation*, Journal of Computational and Applied Mathematics, 295 (2016), pp. 103 – 114. VIII Pan-American Workshop in Applied and Computational Mathematics.
- [64] E. M. PURCELL, H. C. TORREY, AND R. V. POUND, *Resonance Absorption by Nuclear Magnetic Moments in a Solid*, Physical Review, 69 (1946), pp. 37–38, <https://doi.org/10.1103/PhysRev.69.37>.
- [65] R. REMIS, *Reduced-Order Modeling of Transient Electromagnetic Fields*, PhD thesis, PhD thesis, Delft University of Technology, 1998.
- [66] R. F. REMIS, *Low-frequency model-order reduction of electromagnetic fields without matrix factorization*, IEEE Transactions on Microwave Theory and Techniques, 52 (2004), pp. 2298–2304, <https://doi.org/10.1109/TMTT.2004.834577>.
- [67] R. F. REMIS, *On the Relation Between FDTD and Fibonacci Polynomials*, J. Comput. Phys., 230 (2011), pp. 1382–1386, <https://doi.org/10.1016/j.jcp.2010.11.009>.
- [68] R. F. REMIS AND P. M. VAN DEN BERG, *Efficient computation of transient diffusive electromagnetic fields by a reduced modeling technique*, Radio Science, 33 (1998), pp. 191–204, <https://doi.org/10.1029/97RS03693>.
- [69] A. RUHE, *Rational Krylov sequence methods for eigenvalue computation*, Linear Algebra and its Applications, 58 (1984), pp. 391 – 405.
- [70] A. RUHE AND D. SKOOGH, *Rational Krylov Algorithms for Eigenvalue Computation and Model Reduction*, in Applied Parallel Computing. Large Scale Scientific and Industrial Problems., B. Kagström, J. Dongarra, E. Elmroth, and J. Waśniewski, eds., Lecture Notes in Computer Science, No. 1541, 1998, pp. 491–502.
- [71] Y. SAAD, *Iterative Methods for Sparse Linear Systems*, Society for Industrial and Applied Mathematics, second ed., 2003.
- [72] Y. SAAD AND M. H. SCHULTZ, *GMRES: A Generalized Minimal Residual Algorithm for Solving Nonsymmetric Linear Systems*, SIAM Journal on Scientific and Statistical Computing, 7 (1986), pp. 856–869, <https://doi.org/10.1137/0907058>.
- [73] A. A. SAMARSKIJ AND E. S. NIKOLAEV, *Numerical Methods for Grid Equations, Volume I Direct Methods*, Birkhäuser Basel, Allschwilerstrasse 10, 4055 Basel, Switzerland, 1 ed., 1989. (Translated from Russian).
- [74] C. SAUVAN, J. P. HUGONIN, I. S. MAKSYMOW, AND P. LALANNE, *Theory of the Spontaneous Optical Emission of Nanosize Photonic and Plasmon Resonators*, Physical Review Letters, 110 (2013), p. 237401.
- [75] J. A. SETHIAN, *A fast marching level set method for monotonically advancing fronts*, Proceedings of the National Academy of Sciences of the United States of America, 93 (1996), pp. 1591–159.

- [76] S. SHETTY, D. OMERAGIC, T. HABASHY, J. MILES, J. RASMUS, R. GRIFFITHS, AND C. MORRIS, *Inversion-Based Workflows For Interpretation Of Nuclear Density Images In High-Angle And Horizontal Wells*, SPWLA 53th Annual Logging Symposium, Cartagena, Colombia (2012).
- [77] S. SHETTY, D. OMERAGIC, T. HABASHY, J. MILES, J. RASMUS, R. GRIFFITHS, AND C. MORRIS, *3D Parametric Inversion for Interpretation of Logging-While-Drilling Density Images in High-Angle and Horizontal Wells*, PETROPHYSICS, 54 (2013), pp. 324–340.
- [78] D. C. SORENSEN, *Passivity preserving model reduction via interpolation of spectral zeros*, in 2003 European Control Conference (ECC), Sept 2003, pp. 974–978.
- [79] A. TAFLOVE AND S. HAGNESS, *Computational Electrodynamics: The Finite-Difference Time-Domain Method*, Artech House, 2nd ed., 2000.
- [80] X.-. M. C. TEAM, *MCNPA general Monte Carlo Nparticle transport code, Version 5 Volume I: Overview and Theory: LA-UR-03-1987*, Los Alamos National Laboratory, (1987).
- [81] A. TOWNSEND AND L. N. TREFETHEN, *Continuous analogues of matrix factorizations*, Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences, 471 (2014), <https://doi.org/10.1098/rspa.2014.0585>.
- [82] H. A. VAN DER VORST, *Bi-CGSTAB: A Fast and Smoothly Converging Variant of Bi-CG for the Solution of Nonsymmetric Linear Systems*, SIAM Journal on Scientific and Statistical Computing, 13 (1992), pp. 631–644, <https://doi.org/10.1137/0913035>.
- [83] D. WATKINS, *The Matrix Eigenvalue Problem*, Society for Industrial and Applied Mathematics, 2007, <https://doi.org/10.1137/1.9780898717808>.
- [84] J. A. C. WEIDEMAN AND L. N. TREFETHEN, *Parabolic and Hyperbolic Contours for Computing the Bromwich Integral*, Mathematics of Computation, 76 (2007), pp. 1341–1356.
- [85] H. XIE, C. MORRIS, J. RASMUS, , K. ITO, A. ABUBAKAR, AND E. J. STOCKHAUSEN, *Workflow for Determining Layer Properties from Density and Neutron Logs in High-Angle and Horizontal Wells*, SPE International Petroleum Exhibition & Conference, Abu Dhabi, UAE (2016).
- [86] H. XIE, D. OMERAGIC, S. SHETTY, T. HABASHY, J. MILES, J. RASMUS, R. GRIFFITHS, C. MORRIS, AND E. STOCKHAUSEN, *Improved consistency of inversion-based interpretation of LWD density images in complex horizontal well scenarios*, SPWLA 55th Annual Logging Symposium, Abu Dhabi (2014).
- [87] J. ZIMMERLING, V. DRUSKIN, M. ZASLAVSKY, AND R. F. REMIS, *Model-order reduction of electromagnetic fields in open domains*, Geophysics, 83 (2018), pp. WB61–WB70.



- [88] J. ZIMMERLING, L. WEI, P. URBACH, AND R. REMIS, *A Lanczos model-order reduction technique to efficiently simulate electromagnetic wave propagation in dispersive media*, Journal of Computational Physics, 315 (2016), pp. 348 – 362.
- [89] J. ZIMMERLING, L. WEI, P. URBACH, AND R. REMIS, *Efficient computation of the spontaneous decay rate of arbitrarily shaped 3D nanosized resonators: a Krylov model-order reduction approach*, Applied Physics A, 122 (2016), p. 158.

# GLOSSARY

A.D.	Anno Domini
A.u.c.	Ab Urbe Condita
cf.	confer (compare)
c.u.	capture units
CPU	Central Processing unit
DtN	Dirichlet to Neumann
EKS	Extended Krylov Subspace
e.g.	exempli gratia (for example)
FDFD	Finite-Difference Frequency-Domain
FDTD	Finite-Difference Time-Domain
FEM	Finite-Element Method
FFM	Fast forward model
GN	Gauss-Newton
GPU	Graphics Processing Unit
i.e.	id est (that is)
LWD	Logging-While-Drilling
MC	Monte Carlo
MRGN	Model Reduction Gauss-Newton
PDE	Partial Differential Equation
PKS	Polynomial Krylov Subspace
PPRKS	Phase-Preconditioned Rational Krylov Subspace
RCWA	Rigorous coupled-wave analysis
RKS	Rational Krylov Subspace
ROM	Reduced-Order Model
SD	Spontaneous Decay
STW	Stichting voor de Technische Wetenschappen
WKB	Wentzel–Kramers–Brillouin





# NOTATION

Example	Notation	Meaning
$A$	capital bold sans serif	matrix
$u$	bold sans serif	vector
$ \cdot $	single bars	absolute value
$\ \cdot\ $	double bars	2-norm
$\langle x y\rangle$	bra/ket	inner product
$\Re$	-	real part
$\Im$	-	imaginary part
$\hat{\cdot}$	hat	Laplace transformed quantity
$Q(s)$	capital italic	integral operator
$\mathcal{A}$	calligraphic	operator or collection of fields
$\mathbf{E}$	bold	field from physics
$s$	italic s	Laplace frequency
$\bar{\cdot}$	overbar	complex-conjugation
$\mathcal{N}\{\cdot\}$	calligraphic N	nonlinear numerical range
$k$	italic k	spatial dimensions
$n$	italic n	outward pointing normal





## CURRICULUM VITÆ

**JÖRN TORSTEN ZIMMERLING** was born at a very young age in Pinneberg, Germany on the 26<sup>th</sup> of April in 1991 (A. D.) or 2744 (A.u.c.). He obtained his Bachelor and Master's degree at the TU Delft *cum laude* in Electrical Engineering in 2012 and 2014, respectively. During his Bachelor Honors project he began to work with Rob Remis on ADE-FDTD algorithms. With his Master thesis "Modeling of wave propagation in open domains: A Krylov subspace approach" under the supervision of Rob Remis he became the Best Graduate of the TU Delft of 2014. This fruitful collaboration with Rob Remis was continued during a four-year PhD under the research grant "GOOD VIBRATIONS" supported by STW in the CAS group led by Alle-Jan van der Veen. During his PhD, Jörn Zimmerling collaborated with Vladimir Druskin, Mikhail Zaslavsky and Dzevat Omeragic at *Schlumberger-Doll Research*. For four summers, he worked on efficient wavefield solvers, imaging of the cross section of the absorption of thermal neutrons, and 3D anisotropic resistivity imaging. Another three months were spent at *ICERM* at Brown University during a semester program on "Mathematical and Computational Challenges in Radar and Seismic Reconstruction." His general scientific interest lies in forward and inverse problems for PDEs and in pedagogy; he taught the course "Wavefield Imaging" together with Rob Remis during his PhD. He was the president of the *SIAM Student Chapter* at the TU Delft in 2016 and Vice-president in 2017. In 2017 he won the *Delft Computational Science and Engineering* best poster prize. In 2010 Jörn rowed in the lightweight freshman eight of Proteus-Eretes. He is a fanatic cyclist, vivid cross country skier, enthusiast of classical music and supporter of European unity.

