INTENSITY-ONLY IMAGING WITH WAVES, RESTARTED INVERSE BORN SERIES, AND ANALYSIS OF COARSENING IN POLYCRYSTALLINE MATERIALS

by

Patrick Bardsley

A dissertation submitted to the faculty of
The University of Utah
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Mathematics

The University of Utah

August 2016

The University of Utah Graduate School

STATEMENT OF DISSERTATION APPROVAL

The dissertation of **Patrick Bardsley**

has been approved by the following supervisory committee members:

Fernando Guevara Vasquez ,	Chair(s)	**15 Apr 2016** Date Approved
Yekaterina Yuryevna Epshteyn ,	Chair(s)	**15 Apr 2016** Date Approved
Elena A. Cherkaev ,	Member	**15 Apr 2016** Date Approved
Graeme W. Milton ,	Member	**15 Apr 2016** Date Approved
Alexander V. Mamonov ,	Member	**15 Apr 2016** Date Approved

by **Peter Trapa** , Chair/Dean of

the Department/College/School of **Mathematics**

and by **David B. Kieda** , Dean of The Graduate School.

ABSTRACT

We first study the inverse problem of recovering a complex Schrödinger potential from a discrete set of measurements of the solution to the Schrödinger equation using different source terms. We solve this problem by generalizing the inverse Born series method to nonlinear mappings between Banach spaces. In this general setting, we show convergence and stability of inverse Born series follow from a single problem-specific bound. We show this bound for the inverse Schrödinger problem, and study numerically an application of this inverse problem to transient hydraulic tomography. Additionally, we develop a family of iterative methods based on truncated inverse Born series that are akin to iterative methods based on truncated Taylor series.

Next, we study the inverse problem of imaging scatterers in a homogeneous medium when only intensities of wavefields can be measured. Classic imaging methods, such as Kirchhoff migration, rely on phase information contained in full waveform data and thus cannot be used directly with intensity-only data. In situations where scattered wavefields are small compared to the incident wavefields, we can form and solve a linear least squares problem to recover a projection (on a known subspace) of full waveform data from intensity data. We show that for sufficiently high frequencies, this projection gives a Kirchhoff image asymptotically equivalent to the Kirchhoff image obtained from full waveform data. We also generalize this imaging method to using stochastic incident fields with autocorrelation measurements.

Finally, we study a mathematical model of grain growth in polycrystalline materials. We review a simplified 1D grain growth model and an entropy-based theory for the evolution of an important statistic harvested from this model, the GBCD. The theory suggests the GBCD evolves according to a Fokker-Planck equation, which we validate numerically. We derive methods to estimate times from the GBCD, thus fitting it to Fokker-Planck time scales. This allows for direct comparisons of the GBCD with the Fokker-Planck solution, where we find qualitative agreement. We also find an energy dissipation identity which Fokker-Planck solutions must satisfy. We verify the GBCD satisfies this identity both qualitatively and quantitatively, further validating the Fokker-Planck model of GBCD evolution.

For my family, my friends, and my loving wife Kristen.

CONTENTS

CHAPTERS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I am grateful to my committee members Professors Fernando Guevara Vasquez, Yekaterina Epshteyn, Elena Cherkaev, Alexander Mamonov and Graeme W. Milton. Their guidance throughout graduate school and the process of writing and defending a dissertation has been invaluable.

Professor Guevara Vasquez has been a truly outstanding mentor. I have always looked up to him, and I have learned more from him than I could have ever imagined. He has always pushed me to better myself and my work, yet he has always provided the help I needed to do so. His efforts and the training he has provided me with are very much appreciated. His mentorship and friendship remain important in my life, and I am very grateful for the time I have spent working with him. Professor Epshteyn has also been a wonderful mentor, always pushing me to learn new material and tools. She encouraged me to learn new programming languages and to pursue studies in mathematical statistics, where I ultimately obtained a master's degree I otherwise may not have. I would not be the mathematician I am today without her help. Professor Milton has been a great mentor, collaborator, and advocate. I am grateful for his generous support for several semesters and his collaborations with me, allowing me to stretch my understanding of the mathematical sciences. Professors Cherkaev and Mamonov have also provided terrific guidance. Their advice has greatly improved the quality and breadth of this thesis and my mathematical abilities in general.

I appreciate the mentorship of Professor Lajos Horváth. He has been an excellent influence on my studies and career, and he inspires confidence within me. I also wish to thank my peers Andy Thaler, Jason Albright, Greg Rice, Brent Kerby, Kyle Steffen, Michal Kordy, Anna Romanova, and Sean McAffe. I came to know these individuals as office-mates, fellow graduate students, and in many cases as both. The conversations I have had with them, from discussing mathematics to simple everyday

conversations, are some of my favorite times in graduate school.

I am fortunate to have too many people in my life whom I consider family and friends to acknowledge each of them individually. However, I wish to thank my parents Robert and Kathy Bardsley, my grandparents Quentin and LaRae Bardsley, my in-laws Mike and JoAnn Ethridge, Adam Ethridge, Auria Drozd, Steven Ethridge, Megan and Trent Hardy, and my "brothers and sisters" Jeremy Adams, Patty Adams, Nik Aksamit, David Brown, Tommy Cameron, Damon Call, Chris McDonald, and Tyler Phillips. I sincerely thank each of them for the love and support they have provided through this long journey.

The one person I am most grateful for is my wife and best friend Kristen Bardsley. She has supported me more than anyone during both highs and lows, and I attribute many of my successes to her. It has been a great experience going through this process with her by my side, and I am grateful for the love she has provided me through the years. I simply can't express the appreciation, love, and respect I have for her.

This dissertation is partially supported through the National Science Foundation grants DMS-0934664, DMS-1411577, DMS-1211359 and DMS-1112984.

CHAPTER 1

INTRODUCTION

Partial differential equations (PDEs) are common throughout applied mathematics. They are used to model physical phenomena by relating rates of change of certain physical quantities. The coefficients in such an equation carry physical information about the problem at hand, e.g., the speed of propagation of waves in a medium, the conductivity of a medium, or the temperature of a medium. Often, one is concerned with recovering information about these coefficients from measured data; this is the notion of an inverse problem. Other times, one is concerned with finding a PDE that describes the most significant traits of an observed physical phenomenon; this is the notion of mathematical modeling. These two facets of PDE study are the common themes throughout this dissertation. Here, we summarize our work on three specific problems: an inverse problem for an elliptic PDE (§1.1), an inverse problem for a hyperbolic PDE (§1.2), and the mathematical modeling of polycrystalline material dynamics (§1.3). Since these studies were developed independently, in this chapter (resp. thesis) we use notations that are specific to each section (resp. chapter).

1.1 An inverse Schrödinger problem

We consider the Schrödinger equation

$$\begin{cases} -\Delta u(\mathbf{x}) + q(\mathbf{x})u(\mathbf{x}) = \phi(\mathbf{x}), & \mathbf{x} \in \Omega, \\ u(\mathbf{x}) = 0, & \mathbf{x} \in \partial\Omega, \end{cases} \tag{1.1}$$

where $\Omega \subset \mathbb{R}^d$. Here $q(\mathbf{x})$ is the (possibly complex) Schrödinger potential and $\phi(\mathbf{x})$ is a source term supported in Ω. A solution $u \in H_0^1(\Omega)$ satisfying (1.1) in the weak sense exists if, e.g., $\phi \in L^2(\Omega)$ and $q \in L^\infty(\Omega)$ is such that the only solution to (1.1) with $\overline{q}$ instead of q and $\phi \equiv 0$ is $u \equiv 0$ (see, e.g., [23])[1]. Furthermore, u is unique if the only solution to (1.1) for $\phi \equiv 0$ is $u \equiv 0$. The inverse problem we consider here is to recover the potential q from knowledge of N source terms $\{\phi_i\}_{i=1}^N$, and the measurements

$$M_{i,j} = \int u_i(\mathbf{x})\phi_j(x)d\mathbf{x},\ i,j = 1,\ldots,N. \tag{1.2}$$

We call the $\{M_{i,j}\}_{i,j=1}^N$ “discrete internal measurements” because they provide information about u inside Ω. We do not expect a unique solution for this inverse problem since we are trying to recover a function q that generally lies in an infinite dimensional vector space, from finitely many measurements (1.2). However, in other situations, such as when sources and measurements are restricted to the boundary $\partial\Omega$ (see [40]) or when internal data of the form $q(\mathbf{x})u^2(\mathbf{x})$ is known for all $\mathbf{x} \in \Omega$ (see [2, 43]), the solution q is unique (under appropriate assumptions).

A motivating example for this inverse Schrödinger problem is transient hydraulic tomography, where the objective is to recover characteristics of an underground reservoir or aquifer from measurements of hydraulic pressure at a few wells resulting from fluid injection at one or more wells (see, e.g., [13]). The hydraulic pressure $v(\mathbf{x}, t)$ in an aquifer can be modeled by the parabolic PDE

$$S(\mathbf{x})\frac{\partial v}{\partial t}(\mathbf{x}, t) = \nabla \cdot (\sigma(\mathbf{x})\nabla v(\mathbf{x}, t)) - \psi(\mathbf{x}, t),$$

subject to initial and boundary conditions. Here ψ denotes a pressure source (the injection well) and S, σ, characterize physical properties of the aquifer that we are interested in knowing. Different injection wells can be thought of as having different source terms $\{\psi_i\}_{i=1}^N$. By a series of transformations, this imaging problem for S and σ can be recast as the inverse problem of finding the Schrödinger potential q in (1.1) from measurements $\{M_{i,j}\}_{i,j=1}^N$. The coefficients S, σ we were originally seeking can then be estimated from q.

---

[1] In §2.5.1 and in [3], we overlooked the condition involving $\overline{q}$. This was not fixed in Chapter 2 because of copyright reasons.

1.1.1 Inverse Born series

For the inverse Schrödinger problem, we have a nonlinear mapping f that maps a bounded Schrödinger potential $q \in L^\infty(\Omega)$ to the internal measurements $\{M_{i,j}\}_{i,j=1}^N \in (\mathbb{C}^{N\times N}, \|\cdot\|)$ where $\|\cdot\|$ denotes a suitable norm on $\mathbb{C}^{N\times N}$. Thus, $f : L^\infty(\Omega) \to (\mathbb{C}^{N\times N}, \|\cdot\|)$ is a nonlinear mapping from the Banach space $L^\infty(\Omega)$ to the Banach space $(\mathbb{C}^{N\times N}, \|\cdot\|)$ with $f(q) = \{M_{i,j}\}_{i,j=1}^N$. To recover q from $f(q)$, we can try to find an inverse mapping $g : (\mathbb{C}^{N\times N}, \|\cdot\|) \to L^\infty(\Omega)$ such that $g(f(q)) = q$. The inverse Born series method provides a systematic way of approximating such an inverse mapping g, close to a known reference coefficient q_{ref}. This method can be formulated for general nonlinear mappings between Banach spaces, as we now discuss.

The forward Born series represents a nonlinear mapping f from a Banach space X (which we refer to as a “coefficient space”) to a Banach space Y (a “measurement space”) as an infinite series involving multilinear mappings between the two spaces. Assuming a reference coefficient $x \in X$ is known, the forward Born series of f can be formally written as

$$d(h) = f(x+h) - f(x) = \sum_{n=1}^{\infty} a_n[x](h^{\otimes n}),$$

where $d(h) \in Y$ denotes measured data and $h \in X$ is a small perturbation about x. The $a_n[x]$ are linear operators on the tensor products $h^{\otimes n} = h \otimes h \otimes \cdots \otimes h$, and depend on the reference coefficient x. The inverse problem here is to recover the perturbation h from knowledge of $d(h)$, x, and $f(x)$. If we had a mapping $g : Y \to X$ that solves the inverse problem, i.e., $h = g(d(h))$, and if this mapping could be written as a series involving multilinear mappings between Y and X:

$$g(d) = \sum_{n=1}^{\infty} b_n[x](d^{\otimes n}),$$

then the operators $b_n[x]$ can be found formally by a recursion formula involving the $a_n[x]$ operators in a fashion similar to finding the Taylor series coefficients of the inverse f^{-1} of an analytic function f from the Taylor series coefficients of f. The series expansion of g is called “inverse Born series” and can be used to estimate a perturbation h about a known reference x from the data $d(h)$.

This concept was first introduced in the context of optical tomography and diffuse waves by Markel, O'Sullivan and Schotland [34]. Here, the authors were concerned with recovering a coefficient η that characterizes how light is absorbed in a body, from nonintrusive measurements d made only at the boundary of the body. The first convergence results of the inverse Born series were also shown in this physical context by Moskow and Schotland [35]. Here the authors showed the inverse Born series converges, provided certain bounds hold on the boundary data d. Moreover, they characterized the error in reconstructing $h = \eta - \eta_0$ for a known reference η_0, using the inverse Born series and showed it is a stable reconstruction method. The inverse Born series was later studied by Arridge et al. [1] in the context of electrical impedance tomography (EIT). In EIT, one is concerned with recovering the conductivity of a domain, using electrical measurements on the boundary of the domain. Arridge et al. [1] showed the same convergence and stability results hold for inverse Born series in this physical context.

1.1.2 Born series in a Banach setting

Motivated by the successful application of inverse Born series in these multiple contexts, we set out to apply the method to the inverse Schrödinger problem with discrete internal measurements. In doing so, we discovered the forward and inverse Born series can be generalized to nonlinear mappings between Banach spaces. This allows us to formulate the convergence and stability results of Moskow and Schotland [35] in a general setting, giving a common framework for the inverse Born series to be readily applied in many physical contexts. We show the previous studies of Moskow and Schotland [35], Arridge et al. [1], as well as the inverse Schrödinger problem with internal measurements all fit within this general framework. We also establish connections between Born series and the classical Taylor and Neumann series.

Given data $d \in Y$ and a reference coefficient $x_0 \in X$, we notice the inverse Born series can be written as an iterative method:

$$\begin{cases} x_1 = b_1[x_0](d), \\ x_{n+1} = x_n + b_{n+1}[x_0](d^{\otimes n+1}). \end{cases}$$

By truncating at step k and using x_k as a new reference coefficient to restart the inverse Born series, we develop a family of iterative methods we refer to as "restarted inverse Born series of order k" or RIBS(k). These iterative methods are closely related to well-established iterative methods based on Taylor expansions, such as the Gauss-Newton (see, e.g., [19]) and Chebyshev-Halley (see, e.g., [31]) methods. We perform a brief numerical study of convergence properties of the RIBS(k) method, but leave an in-depth investigation to future work.

In Chapter 2, we include a reprint of our publication [3] that details this study and its application to the inverse Schrödinger problem with internal measurements.

1.2 Imaging with waves

Inhomogeneities in an otherwise homogeneous medium can be imaged using waves. The basic principle is to probe a medium with a wave signal, wait for the signal to reflect off of inhomogeneities or scatterers in the medium, and record these backscattered echoes at one or more receivers. From these recordings, one can then image the scatterers' locations using classical imaging functionals such as Kirchhoff migration (travel-time migration) [11] or Multiple Signal Classification (MUSIC) [16]. The underlying PDE is the wave equation

$$\frac{1}{c^2}\frac{\partial^2 u}{\partial t^2} - \Delta u = f.$$

The goal of these imaging methods is to recover information of the wave speed $c(\vec{\boldsymbol{x}})$ of the medium from measurements of the wave field $u(\vec{\boldsymbol{x}}, t)$ and knowledge of the wave source $f(\vec{\boldsymbol{x}}, t)$. The Kirchhoff and MUSIC imaging functionals are well understood, but they both rely on the phase information contained in full waveform measurements of the scattered field. Here, full waveform measurements refer to measuring both the amplitude and phase of the complex field $\widehat{u}(\vec{\boldsymbol{x}}, \omega)$, where $\widehat{\cdot}$ denotes the Fourier transform of $u(x, t)$ with respect to the time variable. When this phase information is lost and only intensity (amplitude) measurements $|\widehat{u}(\vec{\boldsymbol{x}}, \omega)|^2$ can be made, these classical imaging methods cannot be used directly.

1.2.1 Intensity-only measurements

Intensity measurements arise in practice, e.g., when the response time of a receiver is much longer than a typical wave period. This is common in optical applications, such as optical coherence tomography [38, 39] and diffraction tomography [26, 27].

Another form of intensity measurement occurs when one measures autocorrelations of a signal, i.e., correlating a recorded signal with itself. Owing to the Wiener-Khinchin theorem (see, e.g., [32]), measuring the autocorrelation is equivalent to measuring the power spectrum of the signal, i.e., the intensity measurements for each frequency contained in a signal. Because intensity measurements are common in applications, imaging from such data is crucial to develop cheaper and more robust imaging techniques.

A common approach for imaging with intensities is known as phase retrieval. Here one tries to recover full waveform data (i.e., phase information) from intensity data, and then use this recovered field to form an image. Gerchberg and Saxton [28] developed an iterative method to fit an initial guess of full waveform data to the intensity measurements made at two different measurement planes. Gbur and Wolf [27] used a similar principle to essentially recover full waveform data on a single measurement plane, from the intensity measurements made on two different measurement planes. A method of Teague [42] uses a differential identity to relate measured intensity data on a plane with full waveform data measured on the plane's perimeter, ultimately allowing for full waveform data reconstruction. Still other methods use spatial constraints (e.g., [24]), or optimization techniques (e.g., [18]) to approximately recover phases.

Some intensity-only imaging methods bypass the phase retrieval step altogether. The approaches [12, 14, 44] each reformulate the problem of imaging a few point scatterers as a convex optimization problem involving low-rank matrices. These methods increase the dimensionality of the problem in that one tries to recover a matrix rather than a vector, but in doing so they also linearize the problem. By using specific illuminations (i.e., source signals), these methods can obtain exact recovery of a few point scatterers from single frequency intensity measurements. Other non-phase-retrieval methods fit assumed models of scatterers to measured intensity data (e.g., [41]), or simply treat intensity data as noisy measurements of full waveform data (e.g., [20]).

Recently, several imaging methods have been developed that do not recover phase information fully. Instead they recover sufficient phase information that one can then image with existing techniques (e.g., Kirchhoff migration or MUSIC). The method of Novikov, Moscoso, and Papanicolaou [36], uses linear combinations of single-source experiments and the polarization identity to measure inner products of single-source experiments from intensity data. They show these inner products determine full waveform data, up to a global phase that does not affect images when using, e.g., the MUSIC imaging functional. The work by Chen and Huang [15] uses intensity-only measurements scaled by the full waveform incident field as data to image with reverse time migration. Through a series of bounds, they show the image obtained using the scaled data is asymptotically equivalent to the image obtained from full waveform data. In [15], the asymptotic equivalence is in the far-field limit, meaning the distance between the sources/receivers and the scattering object tends to infinity.

1.2.2 Phase retrieval as a least squares problem

We develop an imaging method in [4] that also exploits the concept of an imaging functional being unaffected by imperfect knowledge of phases. Our method is built around a specific illumination strategy and a smallness assumption on the scattered field. We consider a physical setup consisting of multiple point sources located at $\{\vec{\boldsymbol{x}}_s\}_{s=1}^N$ and a single receiver located at $\vec{\boldsymbol{x}}_r$. Using source pairs (i.e., sending the *same* signal from a *pair* of locations $\vec{\boldsymbol{x}}_{s_1}, \vec{\boldsymbol{x}}_{s_2}$, for $s_1, s_2 \in \{1, \ldots, N\}$ simultaneously), we measure the intensity of the resulting wavefield at the receiver location $\vec{\boldsymbol{x}}_r$. Up to some approximations, we can view the problem of recovering the full waveform data from intensity data as a linear system. We show this linear system has a one-dimensional nullspace and thus does not admit a unique solution for full waveform data. However, we show its least squares solution gives sufficient data to image using Kirchhoff migration. Essentially, the Kirchhoff migration functional ignores the nullspace of this matrix for high frequencies. We also generalize our method to using stochastic source signals with autocorrelation measurements at the receiver, thus relaxing constraints on the knowledge of probing wave fields. A reprint of our publication that first proposes and details this imaging method and its specific illumination strategy is included in Chapter 3.

In Chapter 4, we generalize our imaging method to a simpler illumination strategy and framework. Our physical setup here consists of multiple receivers located at $\{\vec{\boldsymbol{x}}_r\}_{r=1}^N$ and a single point source located at $\vec{\boldsymbol{x}}_s$. We probe the medium with a source signal at $\vec{\boldsymbol{x}}_s$ and measure intensities of the resulting wavefield at each receiver location $\{\vec{\boldsymbol{x}}_r\}_{r=1}^N$. Again, up to some approximations we can view the problem of recovering the full waveform data from intensity data as a linear system. The linear system is underdetermined of size $N \times 2N$ since we are measuring intensity data at the receivers, i.e., we measure N *real* numbers, while the full waveform data consists of N *complex* or $2N$ real numbers. We show the nullspace of this system is $N-$dimensional, but more importantly, it again leaves Kirchhoff images unaffected for high frequencies. Moreover, the least squares solution of the system can be expressed simply as the scattered intensity data scaled by the incident field. Therefore, the least squares solution can be understood as a preprocessing step to use Kirchhoff migration with intensity data.

We recently realized that the recovered full waveform data and imaging functional we use in Chapter 4 is essentially the same as that in the preprint of Chen and Huang [15]. Although these methods are developed in different physical setups (we use a limited-aperture or array setup while Chen and Huang [15] use a full-aperture setup), they do share some similarities. Our method relies on a smallness assumption of the scattered field at the receivers, which is satisfied if scattering in the medium is sufficiently weak and the sources and receivers are close together. This smallness assumption is automatically satisfied in the setup considered by Chen and Huang [15] because the sources and receivers are located far from the scatterer. We use a stationary phase argument to show the asymptotic equivalence of Kirchhoff images for recovered data and full waveform data, in a high frequency limit. Chen and Huang [15] use bounds to show this asymptotic equivalence in the far field limit. Finally, in our approach we motivate the data we use for the Kirchhoff imaging functional as an incomplete phase retrieval resulting from the solution to a linear least squares problem. In [15] the imaging data is interpreted as preprocessed or “corrected” data.

1.3 Polycrystalline material dynamics

Polycrystalline materials are commonly found throughout science and engineering. These are materials composed of a multitude of small crystallites or grains, separated by interfaces known as grain boundaries. They are generally metastable and can undergo a rearrangement process known as grain growth or coarsening. Here, coarsening refers to the rearrangement of the polycrystalline material into an energetically preferable configuration by growing or shrinking its grain boundaries. Understanding how material properties change during coarsening is crucial to advance materials science technology. Thus, mathematical models of coarsening are needed.

1.3.1 Modeling grain growth

Many different mathematical models and numerical methods are used to study different aspects of coarsening on different length scales. For example, the mathematical models [17, 29, 37] are particle-based in the sense that grain boundaries are modeled as an alignment of individual particles with dynamics imposed on each particle. These models allow one to study local changes in the grain boundary network arising from intermolecular forces and interactions. However, due to computational cost these studies are generally restricted to a few individual grains. Therefore, other models have been proposed to study material dynamics on a large or macroscopic scale [6, 21, 22, 33]. These models are constructed using sets of coupled PDEs to impose curvature-driven growth on grain boundaries (known as the Mullins equation [25]), with boundary conditions imposed at locations where multiple grain boundaries meet (known as the Herring condition [30]). The models [21, 22], model grains as characteristic functions with diffuse boundaries, which is suitable to study the evolution of geometric material features, e.g., average grain sizes, grain lattice orientations, etc. Alternatively, the models [6, 33], employ sharp front tracking methods to accurately resolve grain boundary locations, which is useful when one is concerned with studying features of the grain boundary network itself.

1.3.2 The GBCD and a theory for its evolution

Of significant interest is understanding how a polycrystalline material rearranges itself according to the energetics of the grain boundary network. Energy in the network arises from the misalignment of the grain lattice of neighboring grains (termed misorientation) and/or a preferred direction of grain lattice orientation. Experimentally it is observed that grain boundaries with high energy tend to shrink while low-energy boundaries tend to grow. The recent mathematical models developed in [6, 33], thus impose grain boundary dynamics by using an interfacial energy density $\psi(\alpha, \theta)$ that gives the energy per unit length of grain boundary with misorientation α and normal direction $\mathbf{n} = (\cos(\theta), \sin(\theta))$. These advances in simulation eventually led to the discovery of the grain boundary character distribution (GBCD). The GBCD is the probability density $\rho(\alpha, \theta, t)$ that gives the probability of finding a grain boundary somewhere in the network, with a given α and θ, at time t. It is a stable statistic that can be collected easily in simulation and physical experiments, and is thus a leading candidate to characterize texture development in polycrystalline materials.

To better understand the behavior of the GBCD in two- and three-dimensional grain growth models, a one-dimensional coarsening model is proposed and analyzed in [5, 6, 7, 10]. The model captures crucial dynamics of grain growth in two and three dimensions, yet it simplifies the analysis of the grain boundary network. Dynamics are imposed in the model using an energy density $\psi(\alpha)$ that depends only on misorientation (i.e., no preferred growth direction). By defining a GBCD $\rho(\alpha, t)$ for the one-dimensional network, the authors [5, 6, 7, 10] derive an energy dissipation inequality that strongly suggests Fokker-Planck dynamics of ρ, i.e.,

$$\mu \frac{\partial \rho}{\partial t} = \frac{\partial}{\partial \alpha}\left(\sigma \frac{\partial \rho}{\partial \alpha} + \psi' \rho\right),$$

for some $\mu > 0, \sigma > 0$. The dissipation inequality they derive uses a crucial entropic assumption: the one-dimensional network evolves to maximize its entropy. Thus, the Fokker-Planck model for GBCD evolution is a theory that can be validated.

1.3.3 Validating the entropy-based theory

In Chapter 5, we review and extend the work of [5, 6, 7, 10] by providing numerical validations of their entropy-based theory. The first validation we perform is originally proposed and performed in [5, 6, 7, 8, 9, 10]. Essentially, by comparing the steady-state solution of the Fokker-Planck equation (a Boltzmann) with the GBCD harvested at the final time of simulation, one can estimate the coefficient σ and determine if the Boltzmann and final time GBCD agree qualitatively. We next extend this validation to intermediate times (rather than only at steady-state) by developing routines to assign Fokker-Planck times to the empirically harvested GBCD, allowing us to freely choose the coefficient μ. We refer to these routines as time-fitting procedures, and they are based on formulas for the evolution of expected values of given functions, where expectations are computed using the Fokker-Planck solution. Upon fitting the GBCD to Fokker-Planck time scales, we directly compare the GBCD with Fokker-Planck solutions at a few intermediate times, and determine they agree qualitatively. We provide an additional validation procedure in the form of an energy dissipation identity that is satisfied for solutions to the Fokker-Planck equation. By verifying the GBCD satisfies this identity both qualitatively and in a quantitative probabilistic sense, we find further evidence validating the Fokker-Planck model of the GBCD evolution.

1.4 References

[1] S. Arridge, S. Moskow, and J. C. Schotland, *Inverse Born series for the Calderon problem*, Inverse Problems, 28 (2012), pp. 035003, 16.

[2] G. Bal, K. Ren, G. Uhlmann, and T. Zhou, *Quantitative thermo-acoustics and related problems*, Inverse Problems, 27 (2011), p. 055007.

[3] P. Bardsley and F. Guevara Vasquez, *Restarted inverse Born series for the Schrödinger problem with discrete internal measurements*, Inverse Problems, 30 (2014), p. 045014.

[4] P. Bardsley and F. Guevara Vasquez, *Imaging with power controlled source pairs*, SIAM Journal on Imaging Sciences, 9 (2016), pp. 185–211.

[5] K. Barmak, E. Eggeling, M. Emelianenko, Y. Epshteyn, D. Kinderlehrer, R. Sharp, and S. Ta'asan, *Critical events, entropy, and the grain boundary character distribution*, Phys. Rev. B, 83 (2011), p. 134117.

[6] ——, *An entropy based theory of the grain boundary character distribution*, Discrete Contin. Dyn. Syst., 30 (2011), pp. 427–454.

[7] ——, *A theory and challenges for coarsening in microstructure*, in Analysis and numerics of partial differential equations, vol. 4 of Springer INdAM Ser., Springer, Milan, 2013, pp. 193–220.

[8] ——, *Materials microstructure: Entropy and curvature-driven coarsening*, RIMS, Research Institute for Mathematical Sciences, U. Kyoto, (2014), pp. 71–91.

[9] ——, *Recent developments in material microstructure: a theory of coarsening*, in Symposium NN - Mathematical and Computational Aspects of Materials Science, vol. 1753 of MRS Proceedings, 2015.

[10] K. Barmak, E. Eggeling, M. Emelianenko, Y. Epshteyn, D. Kinderlehrer, and S. Ta'asan, *Geometric growth and character development in large metastable networks*, Rend. Mat. Appl. (7), 29 (2009), pp. 65–81.

[11] N. Bleistein, J. K. Cohen, and J. W. Stockwell, Jr., *Mathematics of multidimensional seismic imaging, migration, and inversion*, vol. 13 of Interdisciplinary Applied Mathematics, Springer-Verlag, New York, 2001.

[12] E. J. Candès, T. Strohmer, and V. Voroninski, *PhaseLift: Exact and stable signal recovery from magnitude measurements via convex programming*, Comm. Pure Appl. Math., 66 (2013), pp. 1241–1274.

[13] M. Cardiff and W. Barrash, *3-D transient hydraulic tomography in unconfined aquifers with fast drainage response*, Water Resources Research, 47 (2011), p. W12518.

[14] A. Chai, M. Moscoso, and G. Papanicolaou, *Array imaging using intensity-only measurements*, Inverse Problems, 27 (2011), p. 015005.

[15] Z. Chen and G. Huang, *Phaseless imaging by reverse time migration: Acoustic waves*, arXiv, (2015).

[16] M. Cheney, *The linear sampling method and the MUSIC algorithm*, Inverse Problems, 17 (2001), pp. 591–595.

[17] F. Cleri, S. R. Phillpot, and D. Wolf, *Atomistic simulations of integranular fracture in symmetric-tilt grain boundaries*, Interface Science, 7 (1999), pp. 45–55.

[18] L. Crocco, M. D'Urso, and T. Isernia, *Inverse scattering from phaseless measurements of the total field on a closed curve*, J. Opt. Soc. Am. A, 21 (2004), pp. 622–631.

[19] P. Deuflhard, *Newton Methods For Nonlinear Problems: Affine Invariance and Adaptive Algorithms*, vol. 35 of Springer Series in Computational Mathematics, Springer, Heidelberg, 2011.

[20] A. J. Devaney, *Structure determination from intensity measurements in scattering experiments*, Phys. Rev. Lett., 62 (1989), pp. 2385–2388.

[21] M. Elsey, S. Esedoglu, and P. Smereka, *Diffusion generated motion for grain growth in two and three dimensions*, Journal of Computational Physics, 228 (2009), pp. 8015–8033.

[22] ——, *Large-scale simulation of normal grain growth via diffusion-generated motion*, Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences, 467 (2011), pp. 381–401.

[23] L. Evans, *Partial Differential Equations*, Graduate studies in mathematics, American Mathematical Society, 2010.

[24] J. R. Fienup, *Reconstruction of a complex-valued object from the modulus of its Fourier transform using a support constraint*, J. Opt. Soc. Am. A, 4 (1987), pp. 118–123.

[25] A. S. for Metals and U. M. Society, *Metal surfaces : structure, energetics and kinetics ; papers presented at a joint seminar of the American Society for Metals and the Metallurgical Society of AIME ; October 27 and 28, 1962*, American Soc. for Metals, 1963.

[26] G. Gbur and E. Wolf, *Diffraction tomography without phase information*, Opt. Lett., 27 (2002), pp. 1890–1892.

[27] G. Gbur and E. Wolf, *The information content of the scattered intensity in diffraction tomography*, Inform. Sci., 162 (2004), pp. 3–20.

[28] R. Gerchberg and W. Saxton, *A practical algorithm for the determination of phase from image and diffraction plane pictures*, Optik, 35 (1972).

[29] A. Haslam, S. Phillpot, D. Wolf, D. Moldovan, and H. Gleiter, *Mechanisms of grain growth in nanocrystalline fcc metals by molecular-dynamics simulation*, Materials Science and Engineering: A, 318 (2001), pp. 293–312.

[30] C. Herring, *Surface tension as a motivation for sintering*, in Fundamental Contributions to the Continuum Theory of Evolving Phase Interfaces in Solids, J. M. Ball, D. Kinderlehrer, P. Podio-Guidugli, and M. Slemrod, eds., Springer Berlin Heidelberg, 1999, pp. 33–69.

[31] F. Hettlich and W. Rundell, *A second degree method for nonlinear inverse problems*, SIAM J. Numer. Anal., 37 (2000), pp. 587–620.

[32] A. Ishimaru, *Wave propagation and scattering in random media*, IEEE/OUP Series on Electromagnetic Wave Theory, IEEE Press, New York, 1997.

[33] D. Kinderlehrer, I. Livshits, and S. Ta'asan, *A variational approach to modeling and simulation of grain growth*, SIAM J. Sci. Comput., 28 (2006), pp. 1694–1715 (electronic).

[34] V. A. Markel, J. A. O'Sullivan, and J. C. Schotland, *Inverse problem in optical diffusion tomography: IV: Nonlinear inversion formulas*, J. Opt. Soc. Am. A, 20 (2003), pp. 903–912.

[35] S. Moskow and J. C. Schotland, *Convergence and stability of the inverse scattering series for diffuse waves*, Inverse Problems, 24 (2008), p. 065005.

[36] A. Novikov, M. Moscoso, and G. Papanicolaou, *Illumination strategies for intensity-only imaging*, SIAM Journal on Imaging Sciences, 8 (2015), pp. 1547–1573.

[37] S. Phillpot, P. Keblinski, D. Wolf, and F. Cleri, *Synthesis and characterization of a polycrystalline ionic thin film by large-scale molecular-dynamics simulation*, Interface Science, 7 (1999), pp. 15–31.

[38] J. Schmitt, *Optical coherence tomography (OCT): A review*, IEEE Journal of Selected Topics in Quantum Electronics, 5 (1999), pp. 1205–1215.

[39] J. Schmitt, S. Lee, and K. Yung, *An optical coherence microscope with enhanced resolving power in thick tissue*, Optics Communications, 142 (1997), pp. 203 – 207.

[40] J. Sylvester and G. Uhlmann, *A global uniqueness theorem for an inverse boundary value problem*, Annals of mathematics, (1987), pp. 153–169.

[41] T. Takenaka, D. J. N. Wall, H. Harada, and M. Tanaka, *Reconstruction algorithm of the refractive index of a cylindrical object from the intensity measurements of the total field*, Microwave and Optical Technology Letters, 14 (1997), pp. 182–188.

[42] M. R. Teague, *Deterministic phase retrieval: a Green's function solution*, J. Opt. Soc. Am., 73 (1983), pp. 1434–1441.

[43] F. Triki, *Uniqueness and stability for the inverse medium problem with internal data*, Inverse Problems, 26 (2010), p. 095014.

[44] P. Yin and J. Xin, *PhaseLiftOff: An accurate and stable phase retrieval method based on difference of trace and Frobenius norms*, Commun. Math. Sci., 13 (2015), pp. 1033–1049.

CHAPTER 2

RESTARTED INVERSE BORN SERIES FOR THE SCHRÖDINGER PROBLEM WITH DISCRETE INTERNAL MEASUREMENTS

A joint work with Fernando Guevara Vasquez.

Originally published in the journal Inverse Problems. Bardsley, P., Guevara Vasquez, F. *Restarted inverse Born series for the Schrödinger problem with discrete internal measurements.* Volume 30, Issue 4, 045014. March 18, 2014. ©IOP Publishing. Reproduced with permission. All rights reserved. DOI: 10.1088/0266-5611/30/4/045014.

2.1 Introduction

We consider the problem of finding a Schrödinger potential $q(\mathbf{x})$ (which may be complex) from discrete internal measurements of the solution $u_i(\mathbf{x})$ to the Schrödinger equation

$$\begin{cases} -\Delta u_i + q u_i = \phi_i, & \text{for } \mathbf{x} \in \Omega, \\ u_i = 0, & \text{for } \mathbf{x} \in \partial\Omega, \end{cases} \tag{2.1}$$

in a closed bounded set $\Omega \subset \mathbb{R}^d$ for $d \geq 2$, and for different (known) source terms $\phi_i \in C^\infty(\Omega)$, $i = 1, \ldots, N$. We further assume $q \in L^\infty(\Omega)$ is known in $\Omega \backslash \widetilde{\Omega}$, where $\widetilde{\Omega}$ is a closed subset of Ω with a finite distance separating $\partial\widetilde{\Omega}$ and $\partial\Omega$.

The internal measurements we consider are of the form

$$D_{i,j} = \int_\Omega \phi_j(\mathbf{x}) u_i(\mathbf{x}) d\mathbf{x}, \text{for } i, j = 1, \ldots, N. \tag{2.2}$$

The measurement $D_{i,j}$ is a weighted average of the field u_i resulting from the $i-$th source term. Although it is not necessary for our method to work, we assume for simplicity the same source terms are used as weights for the averages.

A motivation for this inverse Schrödinger problem is transient hydraulic tomography (see, e.g., [4] for a review). The hydraulic pressure or head $v(\mathbf{x}, t)$ in an underground reservoir or aquifer Ω resulting from a source $\psi(\mathbf{x}, t)$ (the injection well) satisfies the initial value problem

$$\begin{cases} S v_t = \nabla \cdot (\sigma \nabla v) - \psi, & \text{for } \mathbf{x} \in \Omega,\ t > 0, \\ v(\mathbf{x}, t) = 0, & \text{for } \mathbf{x} \in \partial\Omega,\ t > 0, \\ v(\mathbf{x}, 0) = g(\mathbf{x}), & \text{for } \mathbf{x} \in \Omega. \end{cases} \tag{2.3}$$

Here $S(\mathbf{x})$ is the storage coefficient and $\sigma(\mathbf{x})$ the hydraulic conductivity of the aquifer. The inverse problem is to image both $S(\mathbf{x})$ and $\sigma(\mathbf{x})$ from a series of measurements made by fixing a source term at one well, and measuring the resulting pressure response at the other wells. We show in section 2.6 that the inverse problem of reconstructing $S(\mathbf{x})$ and $\sigma(\mathbf{x})$ from these sparse (and discrete) internal pressure measurements, can be recast as an inverse Schrödinger problem with discrete measurements as in (2.2).

The main tool we use here for solving the inverse Schrödinger problem is inverse Born series. Inverse Born series have been used to solve inverse problems in different contexts such as optical tomography [10, 11, 12, 13], the Calderón or electrical impedance tomography problem [1] and in inverse scattering for the wave equation [8].

In section 2.2 we generalize the inverse Born series convergence results of Moskow and Schotland [12] and Arridge et al. [1], to nonlinear mappings between Banach spaces. The convergence results of inverse Born series in this generalized setting are given in section 2.2.3 and proved in section 2.9, following the same pattern of the proofs in [1, 12]. This new framework is applied in section 2.3 to a few problems that have been solved before with inverse Born series. We also show that both forward and inverse Born series are closely related to Taylor series. Since the cost of calculating the $n-$th term in an inverse Born series grows exponentially with n, we restart it after having computed a few k terms (i.e., we truncate the series to k terms and iterate). We show in section 2.4 that restarting the inverse Born series gives a class of iterative methods that includes the Gauss-Newton and Chebyshev-Halley methods. For the discrete measurements Schrödinger problem, we prove that the necessary conditions for convergence of the inverse Born series are satisfied (section 2.5). Then in section 2.6, we explain how the transient hydraulic tomography problem can be transformed into a discrete measurement Schrödinger problem. Finally in section 4.5 we present numerical experiments comparing the performance of inverse Born series with other iterative methods and their effectiveness for reconstructing the Schrödinger potential in (2.1) and for solving the transient hydraulic tomography problem. We conclude in section 4.6 with a summary of our main results.

2.2 Forward and inverse Born series in Banach spaces

We start by extending the notion of Born series and inverse Born series [10, 12] to operators between Banach spaces, the idea being to give a common framework for the convergence proofs of the inverse Born series for diffuse waves [12], the Calderón problem [1] and the discrete internal measurements Schrödinger problem. This generalization also highlights that the inverse Born series are a systematic way of finding nonlinear approximate inverses for nonlinear mappings. The resulting approximate inverses are valid locally and have guaranteed error estimates.

In sections 2.2.1 and 2.2.2 we define forward and inverse Born series for a mapping f from a Banach space X (the parameter space) to another Banach space Y (the data space). Then in section 2.2.3 we state local convergence results for inverse Born series in Banach spaces that are valid under mild assumptions on the forward Born series. The proofs are included in section 2.9 as they are patterned after the proofs in [1, 12]. Examples of forward and inverse Born series are included in section 2.3.

2.2.1 Forward Born series

Let X and Y be Banach spaces and consider a mapping $f : X \to Y$. In inverse problems applications X is typically the parameter space and Y the data or measurements space. The forward problem is to find the measurements $y = f(x)$ from known parameters x. The inverse problem is to estimate the parameters x knowing the measurements y.

Born series involve operators in $\mathcal{L}(X^{\otimes n}, Y)$, i.e., bounded linear operators from $X^{\otimes n}$ to Y. Here we used the notation

$$X^{\otimes n} = \underbrace{X \otimes \cdots \otimes X}_{n \text{ times}}.$$

If $M(X^n)$ is the space of $n-$linear forms acting on X^n, the (elementary) tensor product $x_1 \otimes \ldots \otimes x_n \in X^{\otimes n}$, with $x_j \in X$, $j = 1, \ldots, n$, is a linear form acting on $M(X^n)$ such that $(x_1 \otimes \ldots \otimes x_n)(u) = u(x_1, \ldots, x_n)$, for $u \in M(X^n)$. The tensor product space $X^{\otimes n}$ is the subspace of the dual of $M(X^n)$ that is spanned by linear combinations of elementary tensor products, i.e., any $x \in X^{\otimes n}$ admits a (not necessarily unique) representation $x = \sum_{i=1}^k x_1^{(i)} \otimes \ldots \otimes x_n^{(i)}$. In general, $X^{\otimes n}$ is not a Banach space. In an abuse of notation we also denote by $X^{\otimes n}$ its completion under the *projective norm*:

$$\|x\|_{X^{\otimes n}} = \inf \left\{ \sum_{i=1}^k \|x_1^{(i)}\|_X \cdots \|x_n^{(i)}\|_X \ : \ x = \sum_{i=1}^k x_1^{(i)} \otimes \ldots \otimes x_n^{(i)} \right\}, \tag{2.4}$$

where the infimum is taken over all representations of x in terms of elementary tensors. Out of all the norms on a tensor product space we choose the projective norm because it has the properties (see, e.g., [14, Prop. 2.1, 2.3]):

1. For x_i in X, $i = 1, \ldots, n$,

$$\|x_1 \otimes \ldots \otimes x_n\|_{X^{\otimes n}} = \|x_1\|_X \ldots \|x_n\|_X.$$

2. If $a \in \mathcal{L}(X^{\otimes m}, Y)$ and $b \in \mathcal{L}(X^{\otimes n}, Y)$ then $a \otimes b \in \mathcal{L}(X^{\otimes(n+m)}, Y^{\otimes 2})$ is defined by $(a \otimes b)(u \otimes v) = a(u) \otimes b(v)$ for $u \in X^{\otimes m}$ and $v \in X^{\otimes n}$. Moreover, when the projective norm is used,

$$\|a \otimes b\|_{\mathcal{L}(X^{\otimes(n+m)}, Y^{\otimes 2})} = \|a\|_{\mathcal{L}(X^{\otimes m}, Y)} \|b\|_{\mathcal{L}(X^{\otimes n}, Y)}.$$

For the sake of clarity, and when there is no ambiguity, the norm subscripts are omitted.

Notice that a map $a \in \mathcal{L}(X^{\otimes n}, Y)$ can be identified to a bounded multilinear (or $n-$linear) map $\widetilde{a} : X^n \to Y$ defined by

$$\widetilde{a}(x_1, \ldots, x_n) = a(x_1 \otimes \cdots \otimes x_n),$$

and that $\|\widetilde{a}\| = \|a\|$, where

$$\|\widetilde{a}\| = \sup\{\|\widetilde{a}(x_1, \ldots, x_n)\|_Y \mid \|x_i\|_X \leq 1,\ i = 1, \ldots, n\}.$$

Remark 1 *The isometry $\|\widetilde{a}\| = \|a\|$ is only valid when the projective norm is used. It may be possible to extend the theory on forward and inverse Born series to other tensor product norms such as the* injective norm *(see, e.g., [14, §3]) or even to* reasonable crossnorms *(see, e.g., [14, §6]). However it is not clear to us if there is any advantage in doing so. Therefore we focus only on the projective norm because it gives an isometric isomorphism between bounded multilinear forms $X^n \to Y$ and $\mathcal{L}(X^{\otimes n}, Y)$ (see, e.g., [14, §2.2]).*

Forward Born series express the measurements for a parameter $x + h \in X$ near a known parameter $x \in X$, assuming knowledge of $y = f(x)$.

Definition 1 *A nonlinear map* $f : X \to Y$ *admits a Born series expansion at* $x \in X$ *if there are are bounded linear operators* $a_n \in \mathcal{L}(X^{\otimes n}, Y)$ *(possibly depending on* x*) such that*

$$d(h) = f(x+h) - f(x) = \sum_{n=1}^{\infty} a_n(h^{\otimes n}), \tag{2.5}$$

and the a_n *satisfy the bound*

$$\|a_n\| \leq \alpha\mu^n \textit{ for } n = 0, 1, \ldots. \tag{2.6}$$

It follows from the bounds on the operators a_n, that the Born series converges *locally*, i.e., when h is sufficiently small:

$$\|h\| < 1/\mu. \tag{2.7}$$

This restriction on the size of the perturbation h can be thought of as the radius of convergence of the expansion about the point x.

2.2.2 Inverse Born series

The purpose of inverse Born series is to recover h from knowing the difference in measurements $d(h) = f(x+h) - f(x)$ from a (known) reference combination of parameters x and measurements $y = f(x)$. The original idea in [10] is to write a power series of the data d,

$$g(d) = \sum_{n=1}^{\infty} b_n(d^{\otimes n}), \tag{2.8}$$

involving the operators $b_n \in \mathcal{L}(Y^{\otimes n}, X)$, which are obtained by requiring (formally) that g is the inverse of $d(h)$, i.e., $g(d(h)) = h$. By equating operators $\mathcal{L}(X^{\otimes n}, Y)$ with the same tensor power n, the operators b_n need to satisfy:

$$\begin{aligned}
I &= b_1(a_1) \\
0 &= b_1(a_2) + b_2(a_1 \otimes a_1) \\
0 &= b_1(a_3) + b_2(a_1 \otimes a_2) + b_2(a_2 \otimes a_1) + b_3(a_1 \otimes a_1 \otimes a_1) \\
&\vdots \\
0 &= \sum_{m=1}^{n} \sum_{s_1+\cdots+s_m=n} b_m(a_{s_1} \otimes \cdots \otimes a_{s_m})
\end{aligned} \tag{2.9}$$

where I is the identity in the parameter space X. The requirement that $b_1 a_1 = I$ is quite strong and may not be possible, for example when the measurement space Y is finite dimensional and X is infinite dimensional. Nevertheless if we assume that b_1 is both a right and left inverse of a_1 we can express the operators b_n in terms of the operators a_n and b_1:

$$\begin{aligned} b_2 &= -b_1 a_2 (b_1 \otimes b_1) \\ b_3 &= -(b_1 a_3 + b_2(a_1 \otimes a_2) + b_2(a_2 \otimes a_1))(b_1 \otimes b_1 \otimes b_1) \\ &\vdots \\ b_n &= -\left(\sum_{m=1}^{n-1} \sum_{s_1+\cdots+s_m=n} b_m (a_{s_1} \otimes \cdots \otimes a_{s_m}) \right) (b_1^{\otimes n}). \end{aligned} \tag{2.10}$$

Since an inverse of a_1 is not necessarily available, the key is to choose $b_1 \in \mathcal{L}(Y, X)$ as a regularized pseudoinverse of a_1 so that $b_1 a_1$ is close to the identity, at least in some subspace. This allows us to define the inverse Born series.

Definition 2 *Assume $f : X \to Y$ admits a Born series (Definition 1) and let $b_1 \in \mathcal{L}(Y, X)$. The* inverse Born series *for f using b_1 is the power series $g(d)$ given by* (2.8) *where the operators $b_n \in \mathcal{L}(Y^{\otimes n}, X)$ are defined for $n \geq 2$ by* (2.10)*. Here again we note the dependence of the operators b_n, $n \geq 2$, on the expansion point $x \in X$ and the operator b_1.*

We now state results that guarantee convergence of the inverse Born series, and give an error estimate between the limit of the inverse Born series and the true parameter perturbation h. The error estimate involves $\|(I - b_1 a_1)h\|$, that is how well the operator $b_1 a_1$ approximates the identity for h. These results require that both h and $d(h) = f(x + h) - f(x)$ are sufficiently small.

2.2.3 Inverse Born series local convergence

Convergence and stability for the forward and inverse Born series were established by Moskow and Schotland [12] for an inverse scattering problem for diffuse waves (see also section 2.3.3). Specifically they obtained bounds on the operators a_n in (2.27) similar to the bounds (2.6). With these bounds, it is possible to show convergence and stability of the inverse Born series and even give a reconstruction error bound [12].

The convergence and stability proofs in [12] for the diffuse wave problem carry out without major modifications to the general Banach space setting. We give in this section a summary of results analogous to those in [12]. The proofs are deferred to section 2.9, as they closely follow the proof pattern in [12].

The following Lemma shows that if the forward Born operators satisfy the bounds (2.6), the operators b_n are also bounded under a smallness condition on the linear operator b_1 that is used to prime the inverse Born series.

Lemma 1 *Assume $f : X \to Y$ admits a Born series and that*

$$\|b_1\| < \frac{1}{(1+\alpha)\mu}, \tag{2.11}$$

where α and μ are as in Definition (1). *Then the coefficients* (2.10) *of the inverse Born series satisfy the estimate*

$$\|b_n\| \leq \beta((1+\alpha)\mu\|b_1\|)^n, \ \textit{for } n \geq 2 \tag{2.12}$$

where

$$\beta = \|b_1\| \exp\left(\frac{1}{1-(1+\alpha)\mu\|b_1\|}\right). \tag{2.13}$$

Convergence of the inverse Born series follows from the bounds in Lemma 1 and a smallness condition on the data d.

Theorem 1 (Convergence of inverse Born series) *The inverse Born series* (2.8) *induced by b_1 and associated with the forward Born series* (2.5) *converges if*

$$\|b_1\| < \frac{1}{(1+\alpha)\mu} \tag{2.14}$$

and the data is sufficiently small

$$\|d\| < \frac{1}{(1+\alpha)\mu\|b_1\|}. \tag{2.15}$$

If h_* is the limit of the series, one can estimate the error due to truncating the series by

$$\left\| h_* - \sum_{n=1}^{N} b_n(d^{\otimes n}) \right\| \leq \beta \frac{((1+\alpha)\mu\|b_1\|\|d\|)^{N+1}}{1-(1+\alpha)\mu\|b_1\|\|d\|}.$$

Stability also follows using essentially the same proof as in [12].

Theorem 2 (Stability of inverse Born series) *Assume* $\|b_1\| < ((1+\alpha)\mu)^{-1}$ *and that we have two data* d_1 *and* d_2 *satisfying* $M = \max(\|d_1\|, \|d_2\|) < ((1+\alpha)\mu\|b_1\|)^{-1}$. *Let* $h_i = g(d_i)$ *for* $i = 1, 2$ *(i.e., the limit of the inverse Born series). Then the reconstructions are stable with respect to perturbations in the data in the sense that:*

$$\|h_1 - h_2\| < C\|d_1 - d_2\|, \tag{2.16}$$

where the constant C *depends on* M, α, μ, *and* $\|b_1\|$.

Theorem 1 guarantees convergence of the forward and inverse Born series:

$$d = \sum_{n=1}^{\infty} a_n(h^{\otimes n}) \quad \text{and} \quad h_* = \sum_{n=1}^{\infty} b_n(d^{\otimes n}). \tag{2.17}$$

The limit h_* of the inverse Born series is, in general, different from the true parameter perturbation h. The following theorem provides an estimate of the error $\|h - h_*\|$.

Theorem 3 (Error estimate) *Assuming that* $\|h\| \leq M$, $\|b_1 a_1 h\| \leq M$ *with*

$$M < \frac{1}{(1+\alpha)\mu},$$

and that the hypothesis of theorem 1 hold, i.e.,

$$\|b_1\| \leq \frac{1}{(1+\alpha)\mu} \text{ and } \|d\| \leq \frac{1}{(1+\alpha)\mu\|b_1\|},$$

we have the following error estimate for the reconstruction error of the inverse Born series:

$$\left\| h - \sum_{n=1}^{\infty} b_n(d^{\otimes n}) \right\| \leq C\|(I - b_1 a_1)h\|, \tag{2.18}$$

where the constant C *depends only on* M, α, β *and* μ *and* $\|b_1\|$.

The proofs of lemma 1, theorems 1, 2, and 3 can be found in section 2.9.

Remark 2 *To invoke theorems 1–3 for a specific mapping* f, *it is necessary to show the forward Born operators* a_n *satisfy certain bounds* (2.6). *By the bounded linear extension theorem (see, e.g., [9, §2.7]), it is sufficient to show the bound for elements of* $X^{\otimes n}$ *before completing the tensor product space with the projective norm. In other words, we only need to check that the bound* $\|a_n(x)\| \leq \alpha\mu^n\|x\|$ *holds for* x *that are* *finite linear combinations of elementary tensor products, i.e., for* $x = \sum_{i=1}^{k} x_1^{(i)} \otimes \cdots \otimes x_n^{(i)}$ *where* $x_j^{(i)} \in X$ *for all* $i = 1, \ldots, k$ *and* $j = 1, \ldots, n$*. Since we use the projective norm for tensor product spaces, another way of showing the bound* (2.6) *is to show it is satisfied by the associated multilinear operator* $\widetilde{a}_n : X^n \to Y$ *(see remark 1).*

2.3 Examples of forward and inverse Born series

We write examples of forward and inverse Born series in the framework of section 2.2. We start by showing in section 2.3.1 that forward and inverse Born series are intimately related to Taylor series. Another example is that of Neumann series (section 2.3.2). We also include the forward and inverse Born series from [1, 12], namely those for the diffuse waves for optical tomography (section 2.3.3) and the electrical impedance tomography problem (section 2.3.4). We finish the examples with the discrete internal measurements Schrödinger problem (section 2.3.5), which is the main application of inverse Born series that we are concerned with here.

2.3.1 Taylor series

- **Parameter space:** $X =$ Banach space
- **Measurement space:** $Y = X$ (for simplicity)
- **Forward map:** f analytic (see, e.g., [15])
- **Forward Born series coefficients:** About $x \in X$, the coefficients a_n can be any operators in $\mathcal{L}(X^{\otimes n}, X)$ agreeing with $f^{(n)}(x)/n!$ on the diagonal, i.e., for any $h \in X$,

  $$a_n(h^{\otimes n}) = \frac{1}{n!} f^{(n)}(x)(h^{\otimes n}).$$

  Here $f^{(n)}$ is the $n-$th Fréchet derivative of f, see, e.g., [16, §4.5] for a definition.

Here we use the theory of analytic functions between Banach spaces (see, e.g., [15]) which assumes that the function f is C^∞ and that the Taylor series of the function

$$f(x+h) = \sum_{n=0}^{\infty} \frac{1}{n!} f^{(n)}(x)(h^{\otimes n}) \tag{2.19}$$

converges absolutely and uniformly for h small enough. If in addition we assume that f admits a Born series expansion at x, then we have

$$d(h) = f(x+h) - f(x) = \sum_{n=1}^{\infty} \frac{1}{n!} f^{(n)}(x)(h^{\otimes n}) = \sum_{n=1}^{\infty} a_n(h^{\otimes n}).$$

That is the Taylor series and Born series coefficients, $f^{(n)}(x)/n!$ and a_n respectively, agree at the diagonal $h^{\otimes n}$.

Since f is C^∞, the Fréchet derivatives $f^{(n)}$ are symmetric in the sense that for any permutation π of $\{1, \ldots, n\}$ we have that

$$f^{(n)}(h_1 \otimes \cdots \otimes h_n) = f^{(n)}(h_{\pi(1)} \otimes \cdots \otimes h_{\pi(n)}).$$

The Born series coefficients a_n in general *do not* satisfy this property, however we can consider their symmetrization $\widetilde{a}_n : X^{\otimes n} \to Y$ defined by

$$\widetilde{a}_n(h_1 \otimes \cdots \otimes h_n) = \frac{1}{n!} \sum_{\pi} a_n(h_{\pi(1)} \otimes \cdots \otimes h_{\pi(n)}) \tag{2.20}$$

where the summation is taken over all permutations π of $\{1, \ldots, n\}$.

Clearly we have that

$$\widetilde{a}_n(h^{\otimes n}) = \frac{1}{n!} \sum_{\pi} a_n(h^{\otimes n}) = a_n(h^{\otimes n}),$$

and so we have the following equality:

$$d(h) = f(x+h) - f(x) = \sum_{n=1}^{\infty} \frac{1}{n!} f^{(n)}(x)(h^{\otimes n}) = \sum_{n=1}^{\infty} \widetilde{a}_n(h^{\otimes n}).$$

We then have two analytic functions that are equal for h sufficiently small, therefore the symmetric operators $\frac{1}{n!} f^{(n)}(x)$ and $\widetilde{a}_n$ must be identical (see [15]). Therefore the Born series and Taylor series coefficients are essentially the same, up to a symmetrization.

If $a_1 = f^{(1)}(x)$ is invertible (this is where the assumption $X = Y$ is used), we can apply the implicit function theorem (see, e.g., [15] or [16, §4.6]) to guarantee the existence of f^{-1} in a neighborhood of x. Moreover the inverse is analytic [15] in a neighborhood of $y = f(x)$ and admits a Taylor series near y

$$f^{-1}(y+d) = \sum_{n=0}^{\infty} \frac{1}{n!} (f^{-1})^{(n)}(y)(d^{\otimes n}). \tag{2.21}$$

On the other hand, if $b_1 = a_1^{-1}$ we can define an inverse Born series for f as in (2.8). By the error estimate for the inverse Born series (Theorem 3) we can guarantee that $h = g(d(h)) = g(f(x+h) - f(x))$ for h and $d(h)$ sufficiently small. Since f is invertible in a neighborhood of y we can also write g in terms f^{-1}

$$g(d) = f^{-1}(y+d) - f^{-1}(y) = f^{-1}(y+d) - x.$$

Using the Taylor series (2.21) for f^{-1} we can write

$$g(d) = \sum_{n=1}^{\infty} b_n(d^{\otimes n}) = \sum_{n=1}^{\infty} \frac{1}{n!}(f^{-1})^{(n)}(y)(d^{\otimes n}). \tag{2.22}$$

As is the case for the forward Born operators a_n, the inverse Born operators b_n are in general not symmetric. If we consider their symmetrization $\widetilde{b}_n$ (as in (2.20)), then we find that the symmetric operators $\widetilde{b}_n$ and $\frac{1}{n!}(f^{-1})^{(n)}(y)$ are the same. Therefore inverse Born series is a way of calculating (up to a symmetrization) the Taylor series for f^{-1} from the Taylor series for f.

2.3.2 Neumann series

- **Parameter space:** $X = \mathbb{R}^N$
- **Measurement space:** $Y = \mathbb{R}^{n \times n}$
- **Forward map:** $f(\mathbf{x}) = \mathbf{M}^T(\mathbf{L} - \operatorname{diag}(\mathbf{x}))^{-1}\mathbf{M}$, where $\mathbf{L} \in \mathbb{R}^{N\times N}$ is invertible and $\mathbf{M} \in \mathbb{R}^{N\times n}$.
- **Forward Born series coefficients:** About $\mathbf{0}$, the coefficients are $a_n(\mathbf{h}) = \mathbf{M}^T(\mathbf{L}^{-1}\operatorname{diag}(\mathbf{h}))^n\mathbf{L}^{-1}\mathbf{M}$.

The forward Born series in this is example comes from the Neumann series for the inverse of $\mathbf{L} - \operatorname{diag}(\mathbf{h})$, when it exists. Indeed if for some matrix induced norm $\|\mathbf{L}^{-1}\operatorname{diag}(\mathbf{h})\| < 1$, this inverse exists and is given by the Neumann series

$$(\mathbf{L} - \operatorname{diag}(\mathbf{h}))^{-1} = \left(\sum_{n=0}^{\infty}(\mathbf{L}^{-1}\operatorname{diag}(\mathbf{h}))^n\right)\mathbf{L}^{-1}. \tag{2.23}$$

The forward Born series is then

$$\begin{aligned} f(\mathbf{h}) - f(\mathbf{0}) &= \mathbf{M}^T(\mathbf{L} - \operatorname{diag}(\mathbf{h}))^{-1}\mathbf{M} - \mathbf{M}^T\mathbf{L}^{-1}\mathbf{M} \\ &= \sum_{n=1}^{\infty}\mathbf{M}^T(\mathbf{L}^{-1}\operatorname{diag}(\mathbf{h}))^n\mathbf{L}^{-1}\mathbf{M}. \end{aligned} \tag{2.24}$$

The inverse Born series can be defined by using as b_1 a regularized pseudoinverse of the linear map $a_1(\mathbf{h}) = \mathbf{M}^T\mathbf{L}^{-1}\operatorname{diag}(\mathbf{h})\mathbf{L}^{-1}\mathbf{M}$. By the convergence results of section 2.2.3, the inverse Born series converges under smallness conditions for $\mathbf{h}$, $f(\mathbf{h}) - f(\mathbf{0})$ and b_1.

This problem is motivated by a discretization of the Schrödinger equation $\Delta u - qu = \phi$ with finite differences. The matrix $\mathbf{L}$ is the finite difference discretization of the Laplacian and $\mathbf{h}$ is the Schrödinger potential at the discretization nodes. The matrix $\mathbf{M}$ corresponds to different source terms ϕ, which are also used to measure u (collocated sources and receiver setup as the one we use for the Schrödinger problem with discrete internal measurements in section 2.3.5). This example can be easily modified when the discretization of the qu term in the Schrödinger equation is not a diagonal matrix (as is often the case for finite elements). The collocated sources and receivers setup can be changed as well by using a matrix other than $\mathbf{M}^T$ in the definition of $f(\mathbf{x})$.

2.3.3 Optical tomography with diffuse waves model [12]

In the diffuse waves approximation for optical tomography (see, e.g., [2] for a review), the energy density $G_q(\mathbf{x}, \mathbf{y})$ resulting from a point source $\mathbf{y} \in \Omega$ satisfies a Schrödinger type equation:

$$\begin{cases} -\Delta_{\mathbf{x}} G_q(\mathbf{x}, \mathbf{y}) + q(\mathbf{x}) G_q(\mathbf{x}, \mathbf{y}) = -\delta(\mathbf{x} - \mathbf{y}), & \text{for } \mathbf{x} \in \Omega, \\ G_q(\mathbf{x}, \mathbf{y}) + \ell \mathbf{n}(\mathbf{x}) \cdot \nabla_{\mathbf{x}} G_q(\mathbf{x}, \mathbf{y}) = 0, & \text{for } \mathbf{x} \in \partial\Omega, \end{cases} \tag{2.25}$$

where the domain $\Omega \subset \mathbb{R}^d$, $d \geq 2$ has a smooth boundary $\partial\Omega$, and $q(\mathbf{x}) \geq 0$ is the absorption coefficient. The $\ell \geq 0$ in the Robin boundary condition is given and, as usual, $\mathbf{n}(\mathbf{x})$ denotes the unit outward pointing normal vector to $\partial\Omega$ at $\mathbf{x}$. The inverse problem here is to recover the absorption coefficient $q(\mathbf{x})$ from knowledge of $G_q(\mathbf{x}, \mathbf{y})$ on $\partial\Omega \times \partial\Omega$. This data amounts to taking measurements of the energy density at all $\mathbf{x} \in \partial\Omega$ for all source locations $\mathbf{y} \in \partial\Omega$ or to knowing the Robin-to-Dirichlet map for q. If the difference between the absorption coefficient $q(\mathbf{x})$ and a known reference coefficient $q_0(\mathbf{x})$ is supported in some $\widetilde{\Omega} \subset \Omega$ (with $\partial\Omega$ and $\partial\widetilde{\Omega}$ separated by a finite distance), then G_q satisfies the Lippmann-Schwinger type integral equation:

$$G_q(\mathbf{x}, \mathbf{y}) = G_{q_0}(\mathbf{x}, \mathbf{y}) + \int_{\widetilde{\Omega}} d\mathbf{z}\, G_{q_0}(\mathbf{x}, \mathbf{z})(q(\mathbf{z}) - q_0(\mathbf{z})) G_q(\mathbf{z}, \mathbf{y}). \tag{2.26}$$

Moskow and Schotland [12] show that the forward Born or scattering series for this problem can be defined as follows.

- **Parameter space:** $X = L^p(\widetilde{\Omega})$ for $2 \leq p \leq \infty$.
- **Measurement space:** $Y = L^p(\partial\Omega \times \partial\Omega)$
- **Forward map:** $f : q \to G_q(\mathbf{x}, \mathbf{y})|_{\partial\Omega\times\partial\Omega}$.
- **Forward Born series coefficients:** For $\eta_1, \ldots, \eta_n \in L^p(\widetilde{\Omega})$ and $\mathbf{x}_1, \mathbf{x}_2 \in \partial\Omega$, the coefficient for the Born series expansion about $q = q_0$ is

$$\begin{aligned}(a_n(\eta_1 \otimes \cdots \otimes \eta_n))(\mathbf{x}_1, \mathbf{x}_2) = \\ \int_{\widetilde{\Omega}^n} G_{q_0}(\mathbf{x}_1, \mathbf{y}_1) G_{q_0}(\mathbf{y}_1, \mathbf{y}_2) \ldots G_{q_0}(\mathbf{y}_{n-1}, \mathbf{y}_n) G_{q_0}(\mathbf{y}_n, \mathbf{x}_2) \\ \eta_1(\mathbf{y}_1) \ldots \eta_n(\mathbf{y}_n)\, d\mathbf{y}_1 \ldots d\mathbf{y}_n. \quad (2.27)\end{aligned}$$

In particular, the results of Moskow and Schotland [12] show that the operators a_n satisfy bounds similar to (2.6) assuming q_0 is constant and that q is sufficiently close to q_0. The authors formulate bounds on a_n in the context of multilinear operators $a_n : L^p(\widetilde{\Omega}^n) \to L^p(\partial\Omega \times \partial\Omega)$, but with minor modifications, the bounds also hold in the context of linear operators $a_n : (L^p(\widetilde{\Omega}))^{\otimes n} \to L^p(\partial\Omega \times \partial\Omega)$. Therefore one can define an inverse Born series through the procedure (2.10), and this series converges under appropriate conditions (see [12] and section 2.2.3).

2.3.4 The Calderón or electrical impedance tomography problem [1]

The electric potential inside a domain Ω with positive conductivity $\sigma(\mathbf{x}) \in L^\infty(\Omega)$ resulting from a point source located at $\mathbf{y} \in \Omega$ satisfies the equation

$$\begin{cases} \nabla_{\mathbf{x}} \cdot [\sigma(\mathbf{x}) \nabla_{\mathbf{x}} G_\sigma(\mathbf{x}, \mathbf{y})] = -\delta(\mathbf{x} - \mathbf{y}), & \text{for } \mathbf{x} \in \Omega \\ G_\sigma(\mathbf{x}, \mathbf{y}) + z\sigma \mathbf{n}(\mathbf{x}) \cdot \nabla_{\mathbf{x}} G_\sigma(\mathbf{x}, \mathbf{y}) = 0, & \text{for } \mathbf{x} \in \partial\Omega. \end{cases} \quad (2.28)$$

Here we assume the contact impedance $z \geq 0$ is known and that σ is constant on $\partial\Omega$. The domain Ω is also assumed to be in $\mathbb{R}^d$, $d \geq 2$ and with smooth boundary. The electric impedance tomography (EIT) problem consists in recovering the conductivity σ from the Robin-to-Dirichlet map, i.e., from knowledge of $G_\sigma(\mathbf{x}, \mathbf{y})$ on $\partial\Omega \times \partial\Omega$ (see, e.g., [3] for a review of EIT). If the difference between σ and a known reference conductivity σ_0 is supported in $\widetilde{\Omega} \subset \Omega$ (with $\partial\widetilde{\Omega}$ at a finite distance from $\partial\Omega$), G_σ satisfies the integral equation

$$G_\sigma(\mathbf{x},\mathbf{y}) = G_{\sigma_0}(\mathbf{x},\mathbf{y}) + \int_{\widetilde{\Omega}} d\mathbf{z}\; G_{\sigma_0}(\mathbf{x},\mathbf{z})\nabla_{\mathbf{z}} \cdot [(\sigma(\mathbf{z}) - \sigma_0(\mathbf{z}))\nabla_{\mathbf{z}} G_\sigma(\mathbf{z},\mathbf{y})]. \tag{2.29}$$

Integrating by parts and using that $\sigma = \sigma_0$ on $\partial\Omega$, G_σ obeys a Lippmann-Schwinger type equation:

$$G_\sigma(\mathbf{x},\mathbf{y}) = G_{\sigma_0}(\mathbf{x},\mathbf{y}) - \int_{\widetilde{\Omega}} d\mathbf{z}\; (\sigma(\mathbf{x}) - \sigma_0(\mathbf{x}))\nabla_{\mathbf{z}} G_{\sigma_0}(\mathbf{x},\mathbf{z}) \cdot \nabla_{\mathbf{z}} G_\sigma(\mathbf{z},\mathbf{y}). \tag{2.30}$$

As shown by Arridge et al. [1], one can then define a forward Born series that can be summarized as follows.

- **Parameter Space:** $X = L^\infty(\widetilde{\Omega})$.
- **Measurement space:** $Y = L^\infty(\partial\Omega \times \partial\Omega)$.
- **Forward map:** $f : \sigma \to G_\sigma(\mathbf{x},\mathbf{y})|_{\partial\Omega\times\partial\Omega}$.
- **Forward Born series coefficients:** For $\eta_1, \ldots, \eta_n \in L^\infty(\widetilde{\Omega})$ and $\mathbf{x}_1, \mathbf{x}_2 \in \partial\Omega$, the coefficient for the Born series expansion about $\sigma = \sigma_0$ is

$$\begin{aligned} &a_n(\eta_1 \otimes \cdots \otimes \eta_n)(\mathbf{x}_1,\mathbf{x}_2) = \\ &\quad (-1)^n \int_{\widetilde{\Omega}} d\mathbf{y}_1\; \eta_1(\mathbf{y}_1)\nabla_{\mathbf{y}_1} G_{\sigma_0}(\mathbf{y}_1,\mathbf{x}_1) \cdot \nabla_{\mathbf{y}_1} \int_{\widetilde{\Omega}} d\mathbf{y}_2\; \eta_2(\mathbf{y}_2)\nabla_{\mathbf{y}_2} G_{\sigma_0}(\mathbf{y}_2,\mathbf{y}_1)\cdot \\ &\qquad \cdots \nabla_{\mathbf{y}_{n-1}} \int_{\widetilde{\Omega}} d\mathbf{y}_n\; \eta_n(\mathbf{y}_n)\nabla_{\mathbf{y}_n} G_{\sigma_0}(\mathbf{y}_n,\mathbf{y}_{n-1}) \cdot \nabla_{\mathbf{y}_n} G_{\sigma_0}(\mathbf{y}_n,\mathbf{x}_2). \end{aligned} \tag{2.31}$$

Arridge et al. [1] show that for σ_0 constant, the operators a_n satisfy bounds similar to (2.6) and so an inverse Born series can be defined following the procedure (2.10). As in section 2.3.3, Arridge et al. [1] establish bounds on a_n as multilinear operators $a_n : L^\infty(\widetilde{\Omega}^n) \to L^\infty(\partial\Omega \times \partial\Omega)$, but with minor modifications, the bounds also hold for linear operators $a_n : (L^\infty(\widetilde{\Omega}))^{\otimes n} \to L^\infty(\partial\Omega \times \partial\Omega)$. The convergence of this series is established in [1] and can also be shown using the generalization in section 2.2.3.

2.3.5 The Schrödinger problem with discrete internal measurements

Instead of having infinitely many measurements as in the optical tomography inverse Schrödinger problem (outlined in section 2.3.3), we consider here the case where we only have access to *finitely many* internal measurements $D_{i,j}$ (see equation (2.2)) of the fields u_i, $i = 1, \dots, N$, satisfying (2.1). We also allow the Schrödinger potential in (2.1) to be complex (as discussed in section 2.6, this is useful when solving the transient hydraulic tomography problem).

The Green function $G_q(\mathbf{x}, \mathbf{y})$ for the problem (2.1) satisfies (2.25) with homogeneous Dirichlet boundary conditions (instead of homogeneous Robin boundary conditions). The fields u_i can be expressed in terms of the Green function G_q as

$$u_i(\mathbf{x}) = -\int_\Omega d\mathbf{y}\; G_q(\mathbf{x}, \mathbf{y})\phi_i(\mathbf{y}),\;\; i = 1, \dots, N. \tag{2.32}$$

If the difference between the Schrödinger potential $q(\mathbf{x})$ and known reference $q_0(\mathbf{x})$ is supported in $\widetilde{\Omega} \subset \Omega$ (with $\partial\widetilde{\Omega}$ and $\partial\Omega$ separated by a finite distance), G_q and G_{q_0} are still related by the Lippmann-Schwinger type equation (2.26). By a fixed point procedure we can define a forward Born series as follows.

- **Parameter Space:** $X = L^\infty(\widetilde{\Omega})$.
- **Measurement Space:** $Y = \mathbb{C}^{N\times N}$, with norm $\|\mathbf{A}\| = \max_{i,j=1,\dots N} |A_{i,j}|$.
- **Forward map:** Owing to (2.32), the data $\mathbf{D}$ in (2.2) becomes:

$$f : q \to \mathbf{D} = -\left[\int_{\Omega^2} d\mathbf{x}d\mathbf{y}\; \phi_i(\mathbf{y})\phi_j(\mathbf{x})G_q(\mathbf{x}, \mathbf{y})\right]_{i,j=1\dots N}.$$

- **Forward Born series coefficients:** For $\eta_1, \dots, \eta_n \in L^\infty(\widetilde{\Omega})$ the coefficient for the Born series expansion about q_0 is

$$\begin{aligned}&[a_n(\eta_1 \otimes \cdots \otimes \eta_n)]_{i,j} = \\ &\qquad (-1)^n \int_{\widetilde{\Omega}^{n+2}} G_{q_0}(\mathbf{x}, \mathbf{y}_1)G_{q_0}(\mathbf{y}_1, \mathbf{y}_2)\cdots G_{q_0}(\mathbf{y}_{n-1}, \mathbf{y}_n)G_{q_0}(\mathbf{y}_n, \mathbf{z})\cdot \\ &\qquad\qquad \eta_1(\mathbf{y}_1)\cdots\eta_n(\mathbf{y}_n)\phi_i(\mathbf{z})\phi_j(\mathbf{x})\; d\mathbf{z}d\mathbf{y}_1\cdots d\mathbf{y}_n d\mathbf{x},\end{aligned} \tag{2.33}$$

  for $i, j = 1, \dots, N$. Note that we have have assumed $\operatorname{supp}\phi_i \subset \widetilde{\Omega}$ so that instead of integrating over $\widetilde{\Omega}^n \times \Omega^2$ integrate over $\widetilde{\Omega}^{n+2}$.

We show in section 2.5 that the operators a_n satisfy the bounds (2.6) (with q_0 not necessarily constant), so it is possible to show convergence of the corresponding inverse Born series by the results of section 2.2.3.

2.4 Inverse Born series and iterative methods

The main goal of this section is to show that inverse Born series can be used to design superlinear[1] iterative methods converging to an approximation x_* of the true parameter x_{true} from knowing measurements $y_{\text{meas}} = f(x_{\text{true}})$ and the forward map $f : X \to Y$. The iterative methods we study here are of the form

$$\begin{cases} x_0 = \text{given}, \\ x_{n+1} = T_n(x_n), \text{ for } n \geq 0, \end{cases}$$

where $T_n : X \to X$. Of course, for such an iterative method to be useful, the iterates x_n need to converge to x_* as $n \to \infty$ (with an a priori rate of convergence) and one should be able to estimate the error $\|x_{\text{true}} - x_*\|$ between the desired parameter x_{true} and the limit x_*. Our results are in some sense a generalization of the result by Markel, O'Sullivan and Schotland [10] that shows that the limits of inverse Born series and the Newton-Kantorovich method are the same. The Newton-Kantorovich method is a "frozen" Gauss-Newton method, i.e., the Gauss-Newton method (which we recall in section 2.4.2), modified so that the pseudoinverse of the linearization of the forward map is found once and for all for the first iterate and used as is in subsequent iterates.

2.4.1 Inverse Born series as an iterative method

We start by reformulating the results of section 2.2.3 in the context of iterative methods. Let us assume that we have a good guess x_0 for x_{true}, and that we know the forward Born series about x_0, i.e., we know the coefficients $a_j[x_0] \in \mathcal{L}(X^{\otimes j}, Y)$ so that

$$f(x) - f(x_0) = \sum_{j=1}^{\infty} a_j[x_0](x - x_0)^{\otimes j}.$$

Theorem 1 means that for an appropriate choice of $b_1[x_0]$, if $\|x_0 - x_{\text{true}}\|$ and $\|f(x_0) - y_{\text{meas}}\|$ are sufficiently small then the inverse Born series

$$x_n - x_0 = \sum_{j=1}^{n} b_j[x_0](y_{\text{meas}} - f(x_0))^{\otimes j}, \tag{2.34}$$

---

[1] We recall that superlinear convergence of x_n to x_* means that $\|x_{n+1} - x_*\| \leq \epsilon_n \|x_n - x_*\|$, where $\epsilon_n \to 0$ as $n \to \infty$.

converges *linearly*[2] to some $x_* \in X$ as $n \to \infty$. Here we write explicitly the dependence of the inverse Born operators $b_n[x_0]$ (defined recursively as in (2.10)) on the reference parameter x_0. Notice that the inverse Born series (2.34) can be written as the iterative method,

$$\begin{cases} x_0 = \text{given}, \\ x_{n+1} = x_n + b_{n+1}[x_0](y_{\text{meas}} - f(x_0))^{\otimes(n+1)}, \text{ for } n \geq 0. \end{cases} \tag{2.35}$$

The error estimate of theorem 3 quantifies how close the limit x_* of the iterative method (2.35) is to the true parameter x_{true}, i.e., there is some $C > 0$ such that

$$\|x_* - x_{\text{true}}\| \leq C\|(I - b_1[x_0]a_1[x_0])(x_0 - x_{\text{true}})\|. \tag{2.36}$$

Unfortunately this is an expensive method to implement as the computational cost of each term $b_n[x_0]$ in the inverse Born series (see (2.10)) increases exponentially with n. Indeed if applying the forward Born operator $a_n[x_0]$ requires n forward problem solves (as is the case for the Schrödinger problem), an application of the inverse Born operator $b_n[x_0]$ involves $2^{n-1} - 1$ forward problem solves.

Remark 3 *We emphasize that the inverse Born series* (2.34) *and* (2.35) *does not require evaluating the forward map f at any other point than the initial iterate x_0. In inverse problems, this means the inverse Born series needs only solutions to the background problem, which may be less expensive to compute, perhaps because it corresponds to a homogeneous medium or a medium with other symmetries. In contrast, Gauss-Newton type methods and the restarted inverse Born series introduced in section 2.4.2 need to evaluate the forward map f (and its linearization) at every iterate x_n.*

2.4.2 Restarted inverse Born series (RIBS)

A natural idea to reduce the cost of inverse Born series is to use the $k-$th iterate of the inverse Born series (2.35) as the starting guess for a fresh run of inverse Born series. This gives rise to the following class of iterative methods:

---

[2] We recall that linear convergence rate of x_n to x_* means that there is some $0 < C < 1$ such that $\|x_{n+1} - x_*\| \leq C\|x_n - x^*\|$.

$$\begin{cases} x_0 = \text{given}, \\ x_{n+1} = x_n + \displaystyle\sum_{j=1}^{k} b_j[x_n](y_{\text{meas}} - f(x_n))^{\otimes j}, \text{ for } n \geq 0, \end{cases} \tag{2.37}$$

which we denote by RIBS(k).

If f is a differentiable mapping and we choose $b_1[x_n] = (f'(x_n))^\dagger$ (where the sign $\dagger$ stands for a regularized pseudoinverse of $f'(x_n)$), the RIBS(1) method is in fact the Gauss-Newton method:

$$\begin{cases} x_0 = \text{given}, \\ x_{n+1} = x_n + f'(x_n)^\dagger(y_{\text{meas}} - f(x_n)), \text{ for } n \geq 0, \end{cases} \tag{2.38}$$

and is quadratically convergent in a neighborhood of x_{true} under fairly mild conditions on f (for X and Y finite dimensional, see, e.g., [5]).

If in addition to choosing $b_1[x_n] = (f'(x_n))^\dagger$ we have $a_2[x_n] = f''(x_n)/2$, the RIBS(2) method can be written as

$$\begin{cases} x_0 = \text{given}, \\ x_{n+1} = x_n - f'(x_n)^\dagger \left[r_n - \dfrac{1}{2} f''(x_n)(f'(x_n)^\dagger r_n, f'(x_n)^\dagger r_n) \right], \text{ for } n \geq 0, \end{cases} \tag{2.39}$$

where $r_n \equiv y_{\text{meas}} - f(x_n)$. This is the so called Chebyshev-Halley method, which has been studied before by Hettlich and Rundell [7] in the context of inverse problems. This method is guaranteed to converge cubically when f'' is Lipschitz continuous [7].

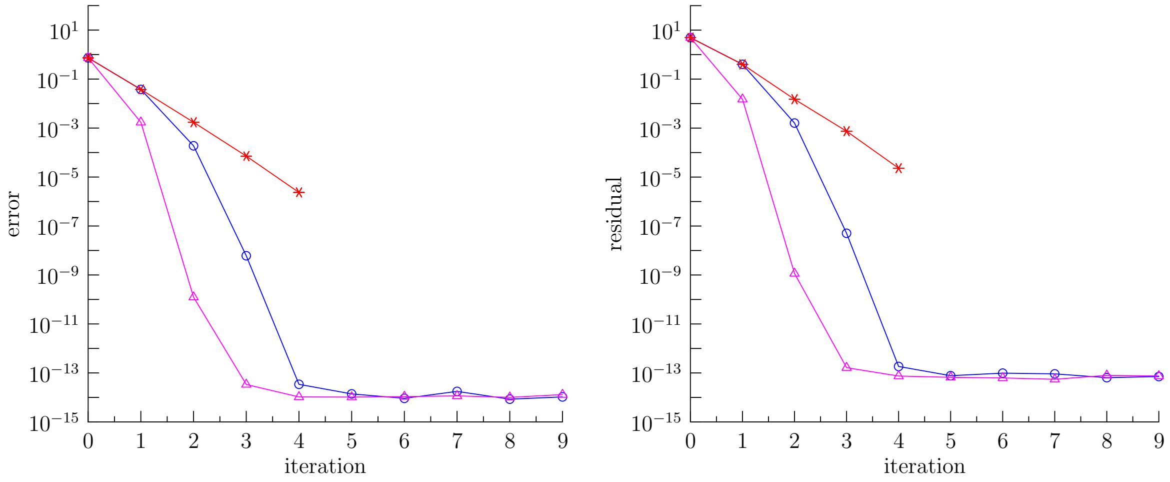
Remark 4 *Although the inverse Born series, and the Gauss-Newton and Chebyshev-Halley methods are guaranteed to converge (under appropriate assumptions), the limits may be different.*

2.4.3 Numerical experiments on a Neumann series toy problem

Here we compare the performance of inverse Born series, Gauss-Newton and Chebyshev-Halley on the Neumann series problem discussed in section 2.3.2. We used for discrete Laplacian $\mathbf{L}$ the matrix

$$\mathbf{L} = \begin{bmatrix} -3 & 1 & & & \\ 1 & -3 & 1 & & \\ & \cdots & & & \\ & & 1 & -3 & 1 \\ & & & 1 & -3 \end{bmatrix} \in \mathbb{R}^{256 \times 256}.$$

The true parameter is a vector with zero mean, independent, normal distributed entries and standard deviation 0.1. The measurement operator $\mathbf{M}$ is a 256×8 matrix with zero mean, independent, normal distributed entries and standard deviation 1. For the inverse Born series, b_1 is a pseudoinverse of the Jacobian of the forward problem, where the singular values smaller than 10^{-6} times the largest singular value (of the Jacobian) are treated as zeroes. The same pseudoinverse is applied to the Jacobian matrices involved in the Gauss-Newton and Chebyshev-Halley methods. The initial guess for all the methods is $\mathbf{x}_0 = \mathbf{0}$. For each method we display in Figure 2.1 (a) the quantity $\|\mathbf{x}_n - \mathbf{x}_*\|$. Since we do not have access to the limiting iterate, we simply took one more step of each method and used it instead of $\mathbf{x}_*$. The residual terms $\|f(\mathbf{x}_n) - f(\mathbf{x}_{\text{true}})\|$ are shown in Figure 2.1 (b). As expected, we see linear convergence for the iterates and the residuals from the truncated inverse Born series method. Also the first Gauss-Newton (resp. Chebyshev-Halley) iterate error and residual matches that of the first (resp. second) inverse Born series iterate. The Gauss-Newton method has the expected quadratic convergence of the error, while the Chebyshev-Halley exhibits super-quadratic convergence of the error.



(a) (b)

Figure 2.1. Convergence of (a) iterates $\|\mathbf{x}_n - \mathbf{x}_*\|$ and (b) residuals $\|f(\mathbf{x}_n) - f(\mathbf{x}_{\text{true}})\|$, for the inverse Born series $(*)$, Gauss-Newton $(\circ)$ and Chebyshev-Halley $(\triangle)$ methods. These methods are applied to the Neumann series problem of section 2.3.2.

2.5 Born series for the Schrödinger problem with discrete internal measurements

Recall from section 2.2.3 that local convergence of the forward and inverse Born series follows from showing that the forward Born operators a_n satisfy bounds of the type (2.6). We show in section 2.5.1 that bounds of the type (2.6) hold for the operators a_n for the Schrödinger problem with discrete internal measurements (defined in (2.33)). Then we report in section 2.5.2 a numerical approximation to the convergence radius of inverse Born series, in a setup related to the hydraulic tomography application of section 2.6.

2.5.1 Bounds on the forward Born operators

We recall from section 2.3.5 that the parameter space for this problem is $X = L^\infty(\widetilde{\Omega})$ where $\widetilde{\Omega} \subset \Omega$ and the distance between $\partial\Omega$ and $\partial\widetilde{\Omega}$ is positive. The difference between the unknown and the reference Schrödinger potentials is assumed to be supported in $\widetilde{\Omega}$. The measurements space is $Y = \mathbb{C}^{N\times N}$ where N is the number of sources used and the norm is the entry-wise ℓ_∞ norm of a matrix in $\mathbb{C}^{N\times N}$.

The proof of lemma 2 below follows a pattern similar to [12]. There are two main differences. The first is that we work with finitely many measurements. The second is that we allow the (possibly complex) reference Schrödinger potential q_0 to be in $L^\infty(\Omega)$, whereas in [12] the reference potential is assumed to be constant and real. The bound (2.6) immediately gives a smallness condition that is sufficient for convergence of the forward Born series. The smallness condition we obtain is identical to that in [12]. This is to be expected because the underlying equation is the same and only the measurements differ.

To prove lemma 2, we need that the reference Schrödinger potential $q_0(\mathbf{x}) \in L^\infty(\Omega)$ is such that the only solution to

$$\begin{cases} -\Delta u + q_0 u = 0, & \text{in } \Omega, \\ u = 0, & \text{on } \partial\Omega, \end{cases} \tag{2.40}$$

is $u = 0$. Such q_0 are sometimes called “nonresonant” and we assume that all the Schrödinger potentials that we deal with in what follows are nonresonant. We also need two properties for the Green function $G_{q_0}(\mathbf{x}, \mathbf{y})$ for the Schrödinger equation (as defined in section 2.3.5):

1. The function $\mathbf{x} \mapsto G_{q_0}(\mathbf{x}, \mathbf{y})$ is in $L^1(\Omega)$ for all $\mathbf{y} \in \Omega$,
2. The function $\mathbf{y} \mapsto \|G_{q_0}(\cdot, \mathbf{y})\|_{L^1(\Omega)}$ is in $L^\infty(\Omega)$.

These properties can be easily verified in both $\mathbb{R}^2$ and $\mathbb{R}^3$ for G_0 (i.e., when $q_0 \equiv 0$) and hold for general bounded q_0. Indeed, we have $(\Delta + q_0)(G_{q_0} - G_0) = -q_0 G_0$. Since the right hand side belongs to $L^2(\Omega)$, the difference $G_{q_0} - G_0$ must be in $H^2_{\text{loc}}(\Omega)$ by standard elliptic regularity estimates (see, e.g., [6]) and therefore continuous (by Sobolev embeddings). This argument shows that $(G_{q_0} - G_0)(\mathbf{x}, \mathbf{y})$ is continuous as function of $\mathbf{x}$ and for all $\mathbf{y}$. By reciprocity $G_{q_0} - G_0$ is continuous on $\Omega \times \Omega$. Therefore G_{q_0} satisfies the desired properties.

We can now show boundedness of the operators a_n for the Schrödinger equation with discrete measurements. The proof of the following lemma is similar to that in [12].

Lemma 2 *Let $q_0(\mathbf{x})$ be a (possibly complex) nonresonant Schrödinger potential. Then the operators a_n defined in* (2.33) *satisfy the bounds*

$$\|a_n\| \le \alpha \mu^n, \tag{2.41}$$

with $\alpha = \nu/\mu$, and where ν and μ are constants depending on Ω and q_0 only (see equations (2.43) *and* (2.44) *below for their definition). The norm on a_n is the operator norm in $\mathcal{L}(X^{\otimes n}, Y)$, with parameter space X and data space Y as in section 2.3.5.*

Proof. Following remark 2, we first establish the bound on the space of finite linear combinations of elementary tensor products of $L^\infty(\widetilde{\Omega})$. Let $\eta \in (L^\infty(\widetilde{\Omega}))^{\otimes n}$ with representation $\eta = \sum_{k=1}^N \eta_1^{(k)} \otimes \cdots \otimes \eta_n^{(k)}$ where $\eta_j^{(k)} \in L^\infty(\widetilde{\Omega})$, $j = 1, \ldots, n$, $k = 1, \ldots, N$, and observe

$$\begin{aligned}\|a_n(\eta)\| &= \sup_{i,j}\left|\left(\sum_{k=1}^{N} a_n(\eta_1^{(k)}\otimes\cdots\otimes\eta_n^{(k)})\right)_{i,j}\right| \\ &\le \sum_{k=1}^{N}\sup_{i,j}\int_{\widetilde{\Omega}^{n+2}}\big|G_{q_0}(\mathbf{x},\mathbf{y}_1)\cdots G_{q_0}(\mathbf{y}_{n-1},\mathbf{y}_n)\cdot \\ &\qquad \cdots G_{q_0}(\mathbf{y}_n,\mathbf{z})\eta_1^{(k)}(\mathbf{y}_1)\cdots\eta_n^{(k)}(\mathbf{y}_n)\phi_i(\mathbf{z})\phi_j(\mathbf{x})\big|d\mathbf{z}d\mathbf{y}_1\cdots d\mathbf{y}_n d\mathbf{x} \\ &\le \sum_{k=1}^{N}\|\eta_1^{(k)}\|_{L^\infty(\widetilde{\Omega})}\cdots\|\eta_n^{(k)}\|_{L^\infty(\widetilde{\Omega})}\sup_{i,j}\int_{\widetilde{\Omega}^{n+2}}\big|G_{q_0}(\mathbf{x},\mathbf{y}_1)\cdot \\ &\qquad \cdots G_{q_0}(\mathbf{y}_{n-1},\mathbf{y}_n)G_{q_0}(\mathbf{y}_n,\mathbf{z})\phi_i(\mathbf{z})\phi_j(\mathbf{x})\big|d\mathbf{z}d\mathbf{y}_1\cdots d\mathbf{y}_n d\mathbf{x}.\end{aligned} \tag{2.42}$$

Since this bound holds for all representations of η, it must hold for the infimum over all the representations of η, which gives the projective norm (2.4). Therefore the operator a_n is bounded on the space of finite linear combinations of elementary tensor products and

$$\begin{aligned}\|a_n(\eta)\| \le \|\eta\|_{(L^\infty(\widetilde{\Omega}))^{\otimes n}}\sup_{i,j}\int_{\widetilde{\Omega}^{n+2}}&\big|G_{q_0}(\mathbf{x},\mathbf{y}_1)\cdots G_{q_0}(\mathbf{y}_{n-1},\mathbf{y}_n)\cdot \\ &\cdots G_{q_0}(\mathbf{y}_n,\mathbf{z})\phi_i(\mathbf{z})\phi_j(\mathbf{x})\big|d\mathbf{z}d\mathbf{y}_1\cdots d\mathbf{y}_n d\mathbf{x}.\end{aligned}$$

By the bounded extension theorem (see, e.g., [9, §2.7]) this also gives an (identical) upper bound for the extension of a_n to the completion of $(L^\infty(\widetilde{\Omega}))^{\otimes n}$ under the projective norm.

Hence we can estimate the operator norm $\|a_1\|$ by

$$\begin{aligned}\|a_1\| &\le \sup_{i,j}\int_{\widetilde{\Omega}\times\widetilde{\Omega}\times\widetilde{\Omega}}\big|G_{q_0}(\mathbf{x},\mathbf{y}_1)G_{q_0}(\mathbf{y}_1,\mathbf{z})\phi_i(\mathbf{z})\phi_j(\mathbf{x})\big|d\mathbf{z}d\mathbf{y}_1 d\mathbf{x} \\ &\le \sup_{i,j}\int_{\widetilde{\Omega}}\int_{\widetilde{\Omega}}\big|G_{q_0}(\mathbf{y}_1,\mathbf{z})\phi_i(\mathbf{z})\big|d\mathbf{z}\int_{\widetilde{\Omega}}\big|G_{q_0}(\mathbf{x},\mathbf{y}_1)\phi_j(\mathbf{x})\big|d\mathbf{x}d\mathbf{y}_1 \\ &\le \sup_i\left(\sup_{\mathbf{x}\in\widetilde{\Omega}}\int_{\widetilde{\Omega}}\big|G_{q_0}(\mathbf{x},\mathbf{y})\phi_i(\mathbf{y})\big|d\mathbf{y}\right)^2|\Omega|.\end{aligned}$$

Since q_0 is assumed to be nonresonant and using that $\phi_i \in L^\infty(\Omega)$, the quantity

$$\nu = \left(\sup_i\sup_{\mathbf{x}\in\widetilde{\Omega}}\int_{\widetilde{\Omega}}\big|G_{q_0}(\mathbf{x},\mathbf{y})\phi_i(\mathbf{y})\big|d\mathbf{y}\right)^2|\Omega| \tag{2.43}$$

is bounded. We have established that $\|a_1\| \le \nu$.

For the remaining Born operators, we proceed recursively. Considering again (2.42) for $n \geq 2$, we have

$$\begin{aligned}\|a_n\| &\leq \sup_{i,j} \int_{\widetilde{\Omega}^{n+2}} \big|G_{q_0}(\mathbf{x},\mathbf{y}_1)G_{q_0}(\mathbf{y}_1,\mathbf{y}_2)\cdot \\ &\qquad \cdots G_{q_0}(\mathbf{y}_{n-1},\mathbf{y}_n)G_{q_0}(\mathbf{y}_n,\mathbf{z})\phi_i(\mathbf{z})\phi_j(\mathbf{x})\big| d\mathbf{z}d\mathbf{y}_1\cdots d\mathbf{y}_n d\mathbf{x} \\ &\leq \sup_{i,j}\left(\sup_{\mathbf{y}_1\in\widetilde{\Omega}}\int_{\widetilde{\Omega}}\big|G_{q_0}(\mathbf{x},\mathbf{y}_1)\phi_j(\mathbf{x})\big|d\mathbf{x}\right)\left(\sup_{\mathbf{y}_n\in\widetilde{\Omega}}\int_{\widetilde{\Omega}}\big|G_{q_0}(\mathbf{y}_n,\mathbf{z})\phi_i(\mathbf{z})\big|d\mathbf{z}\right) \\ &\cdot\int_{\widetilde{\Omega}^n}\big|G_{q_0}(\mathbf{y}_1,\mathbf{y}_2)\cdots G_{q_0}(\mathbf{y}_{n-1},\mathbf{y}_n)\big|d\mathbf{y}_1\cdots d\mathbf{y}_n \\ &\leq \left(\sup_i\sup_{\mathbf{x}\in\widetilde{\Omega}}\int_{\widetilde{\Omega}}\big|G_{q_0}(\mathbf{x},\mathbf{y})\phi_i(\mathbf{y})\big|d\mathbf{y}\right)^2 I_{n-1}\end{aligned}$$

where

$$I_{n-1} = \int_{\widetilde{\Omega}^n}\big|G_{q_0}(\mathbf{y}_1,\mathbf{y}_2)\cdots G_{q_0}(\mathbf{y}_{n-1},\mathbf{y}_n)\big|d\mathbf{y}_1\cdots d\mathbf{y}_n.$$

Estimating I_{n-1} we find that

$$\begin{aligned}I_{n-1} &\leq \sup_{\mathbf{y}_{n-1}\in\widetilde{\Omega}}\int_{\widetilde{\Omega}}\big|G_{q_0}(\mathbf{y}_{n-1},\mathbf{y}_n)\big|d\mathbf{y}_n\cdot\int_{\widetilde{\Omega}^{n-1}}\big|G_{q_0}(\mathbf{y}_1,\mathbf{y}_2)\cdots G_{q_0}(\mathbf{y}_{n-2},\mathbf{y}_{n-1})\big|d\mathbf{y}_1\cdots d\mathbf{y}_{n-1} \\ &\leq \mu I_{n-2},\end{aligned}$$

where the quantity

$$\mu = \sup_{\mathbf{x}\in\widetilde{\Omega}}\|G_{q_0}(\mathbf{x},\cdot)\|_{L^1(\widetilde{\Omega})} \tag{2.44}$$

is finite by the properties that G_{q_0} satisfies. Finally, noting that

$$\begin{aligned}I_1 &= \int_{\widetilde{\Omega}\times\widetilde{\Omega}}\big|G_{q_0}(\mathbf{y}_1,\mathbf{y}_2)\big|d\mathbf{y}_1 d\mathbf{y}_2 \\ &\leq \mu|\Omega|,\end{aligned}$$

it follows that

$$I_{n-1} \leq |\Omega|\mu^{n-1},$$

and thus

$$\|a_n\| \leq \left(\sup_i\sup_{\mathbf{x}\in\widetilde{\Omega}}\|G_{q_0}(\mathbf{x},\cdot)\|_{L^1(B_\rho(\mathbf{x}_i))}\right)^2|\Omega|\mu^{n-1} = \alpha\mu^n.$$

■

Remark 5 (L^p Bounds) *Bounds similar to those in lemma 2 can be proven when the parameter space is $X = L^2(\Omega)$ and the data space is $Y = \mathbb{C}^{N\times N}$, endowed with* *the Frobenius norm. Once we have bounds for the* ∞ *and* 2 *norms, it is possible to invoke the Riesz-Thorin theorem (as in [12]) to show bounds for* $2 \leq p \leq \infty$ *by interpolation. In this case the data space is* $X = L^p(\Omega)$ *and the parameter space is* $Y = \mathbb{C}^{N\times N}$*, endowed with the entry-wise* $p-$*norm (i.e., the* $p-$*norm of the* $\mathbb{C}^{N^2}$ *vector obtained by stacking the columns of a matrix in* $\mathbb{C}^{N\times N}$*).*

Having established norm bounds on the operators a_n for the discrete measurements Schrödinger problem, we can apply the results from section 2.2.3 to establish local convergence of the forward Born series, local convergence of the inverse Born series (provided the linear operator b_1 used to prime the series has sufficiently small norm, see theorem 1), stability of the inverse Born series (theorem 2) and even an error estimate (theorem 3). The actual choice of b_1 is discussed in section 4.5.

2.5.2 Numerical illustration

Applying theorem 1 to the Schrödinger problem with discrete measurements, we can expect the inverse Born series to converge when the difference d between the data for the unknown and reference Schrödinger potentials satisfies

$$\|d\| \leq \frac{1}{(1+\alpha)\mu\|b_1\|},$$

where the constants $\alpha = \nu/\mu$ and μ are constants defined by (2.43) and (2.44) and the norms are as in section 2.3.5.

In preparation for the application to hydraulic tomography, we consider the setup depicted in Figure 2.2 with computational domain $\Omega = [0,1]^2$. The distance between Ω and $\widetilde{\Omega}$ is $\epsilon \in [0,1/4]$ and the sources ϕ_i are supported in disks of radius 0.05 with centers $(0.2k, 0.2l)$, for $k,l = 1,\ldots,4$. The sources are $\phi_i(\mathbf{x}) = \phi(\mathbf{x} - \mathbf{x}_i)$ where $\mathbf{x}_i$ is the center of the disk support and ϕ is an infinitely smooth function with $0 \leq \phi(\mathbf{x}) \leq 1$. Although theorem 1 allows for the supports of the sources to overlap, we take them to be disjoint as this is the case in the hydraulic tomography application.

The constants μ and ν are approximated by solving appropriate (forward) Schrödinger problems with $q_0 = 0$. The grid we use for this purpose is uniform and consists of the nodes (kh, lh) for $k,l = 0,\ldots,400$ and $h = 1/400$. We display in Figure 2.3

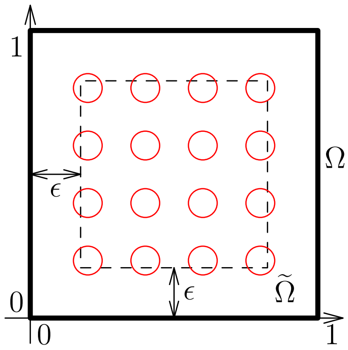


Figure 2.2. Setup for the numerical experiments with the Schrödinger problem with internal measurements. The domain Ω is the unit square. The domain $\widetilde{\Omega}$ where the Schrödinger potential is unknown is in dotted line and its boundary $\partial\widetilde{\Omega}$ is at a distance ϵ from $\partial\Omega$. The supports of the functions used as source terms/measurements are the red circle.

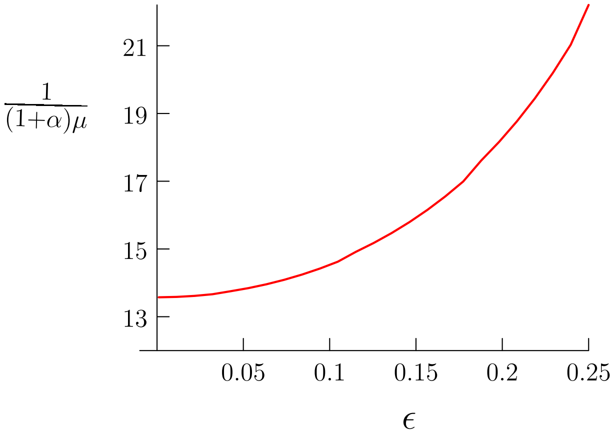


Figure 2.3. Numerical approximation of the radius of convergence for the inverse Born series for the Schrödinger problem with discrete internal measurements and assuming $\|b_1\| \geq 1$. The reference Schrödinger potential is $q_0 = 0$ and the setup is that given in Figure 2.2.

the radius of convergence of the inverse Born series predicted by theorem 1, assuming $\|b_1\| = 1$. We observe that the radius of convergence increases as ϵ increases, or in other words, the larger the region where we assume the Schrödinger potential is known, the larger the perturbations in the data the method can handle.

2.6 Application to transient hydraulic tomography

Consider an underground aquifer confined in a bounded domain Ω. The head or hydraulic pressure $u_i(\mathbf{x}, t)$ in the aquifer due to injecting water in the i−th well satisfies the equation

$$\begin{cases} S\dfrac{\partial u_i}{\partial t} = \nabla\cdot(\sigma\nabla u_i) - \phi_i, & \text{for } \mathbf{x}\in\Omega,\ t>0,\\ u_i(\mathbf{x},t) = 0, & \text{for } \mathbf{x}\in\partial\Omega,\ t>0,\\ u_i(\mathbf{x},0) = g(\mathbf{x}), & \text{for } \mathbf{x}\in\Omega. \end{cases} \tag{2.45}$$

where $i = 1,\dots,N$. Here we assume there are no sources or leaks of water in the aquifer, other than those prescribed at the wells. Hence the source term $\phi_i(\mathbf{x},t)$ is supported at the $i-$th well and represents the water injected at the $i-$th well. The physical properties of the aquifer are modeled by the storage coefficient $S(\mathbf{x})$ and the hydraulic conductivity $\sigma(\mathbf{x})$. The initial head (at $t = 0$) is given by $g(\mathbf{x})$.

The inverse problem of hydraulic tomography that we consider here, is to determine the coefficients σ and S from knowledge of the discrete internal measurements

$$M_{i,j}(t) = \int_\Omega \phi_j(\mathbf{x},t) * u_i(\mathbf{x},t)d\mathbf{x},\ i,j = 1,\dots,N, \tag{2.46}$$

where the convolution is in time. Physically these measurements correspond to time domain measurements at the $j-$th well of a spatial average of the hydraulic pressure u_i generated by injecting in the $i-$th well. Here for simplicity, we use for the impulse response (in time) of the $j-$th measurement well the function $\phi_j(\mathbf{x},t)$. In a more general setup, the injection and measurement "well functions" can be different.

2.6.1 Reformulation as a discrete internal measurements Schrödinger problem

The frequency domain version of problem (2.45) is

$$\begin{cases} \nabla\cdot(\sigma\nabla\widehat{u}_i) - \imath\omega S\widehat{u}_i = \widehat{\phi}_i, & \text{for } \mathbf{x}\in\Omega,\\ \widehat{u}_i = 0, & \text{for } \mathbf{x}\in\partial\Omega, \end{cases} \tag{2.47}$$

where the hat denotes Fourier transform in time, i.e.,

$$\widehat{u}_i(\mathbf{x},\omega) = \int_{\mathbb{R}} u_i(\mathbf{x},t)e^{-\imath\omega t}dt \quad \text{and} \quad \widehat{\phi}_i(\mathbf{x},\omega) = \int_{\mathbb{R}} \phi_i(\mathbf{x},t)e^{-\imath\omega t}dt.$$

The inverse problem is now to recover σ and S from the discrete internal measurements

$$\widehat{M_{i,j}}(\omega) = \int_\Omega \widehat{\phi}_j(\mathbf{x},\omega)\widehat{u}_i(\mathbf{x},\omega)d\mathbf{x}, \tag{2.48}$$

which is the Fourier transform in time of the discrete internal measurements for the time domain problem (2.46).

Next we use the Liouville transformation by defining $v_i = \sigma^{1/2}\widehat{u}_i$. If $\widehat{u}_i$ satisfies (2.47) then v_i must satisfy the Schrödinger equation

$$\begin{cases} \Delta v_i - \left(\dfrac{\Delta\sigma^{1/2}}{\sigma^{1/2}} + \dfrac{\imath\omega S}{\sigma}\right) v_i = \dfrac{\widehat{\phi}_i}{\sigma^{1/2}}, & \text{for } \mathbf{x} \in \Omega, \\ \qquad\qquad\qquad\qquad\qquad v_i = 0, & \text{for } \mathbf{x} \in \partial\Omega. \end{cases} \tag{2.49}$$

The internal measurements $\widehat{M_{i,j}}(\omega)$ can now be expressed in terms of v_i as

$$\widehat{M_{i,j}}(\omega) = \int_\Omega \widehat{\phi}_j(\mathbf{x},\omega)\widehat{u}_i(\mathbf{x},\omega)d\mathbf{x} = \int_\Omega \frac{\widehat{\phi}_j(\mathbf{x},\omega)}{\sigma^{1/2}(\mathbf{x})} v_i(\mathbf{x},\omega)d\mathbf{x}.$$

Hence the measurements $\widehat{M_{i,j}}(\omega)$ are of the form defined in (2.2) with test functions $\widehat{\phi}_i/\sigma^{1/2}$ (modeling both injection and measurement).

If we do have access to the inside of the wells (i.e., $\operatorname{supp}\widehat{\phi}_i$), it is reasonable to assume that σ is known in $\operatorname{supp}\widehat{\phi}_i$. Hence the test functions $\widehat{\phi}_i/\sigma^{1/2}$ are known and we can use any method for solving the inverse Schrödinger problem with discrete data to obtain an approximation to the complex Schrödinger potential

$$Q(\mathbf{x};\omega) = \frac{\Delta\sigma^{1/2}}{\sigma^{1/2}} + \frac{\imath\omega S}{\sigma}, \text{ for } \mathbf{x} \in \Omega. \tag{2.50}$$

Remark 6 *A limitation of transforming the hydraulic tomography problem into an inverse Schrödinger problem is that the conductivity σ appears as $\Delta\sigma^{1/2}/\sigma^{1/2}$ in the Schrödinger potential. Therefore any high (spatial) frequency components in $\sigma^{1/2}$ are magnified. The resulting Schrödinger potential can easily fall outside of the radius of convergence of the inverse Born series. It may be possible to overcome this limitation if we apply the inverse Born series to the hydraulic tomography problem directly (i.e., without doing the Liouville transform).*

2.6.2 Recovery of S and σ from one frequency

Once we have approximated $Q(\mathbf{x};\omega)$ for a single (known) frequency ω, the real part of $Q(\mathbf{x};\omega)$ can be used to estimate the hydraulic conductivity σ. This can be achieved by solving for $\sigma^{1/2}(\mathbf{x})$ in the equation

$$\Delta\sigma^{1/2} - \operatorname{Re}(Q(\mathbf{x};\omega))\sigma^{1/2} = 0,$$

on the aquifer without the wells, i.e.,

$$\Omega' \equiv \Omega \setminus \bigcup_{i=1}^{n} \operatorname{supp}\widehat{\phi}_i,$$

and with Dirichlet boundary conditions at $\partial\Omega'$ determined from the (assumed) knowledge of σ at the measurement wells and at $\partial\Omega$. An estimate of the storage coefficient S from $\mathrm{Im}(Q(\mathbf{x};\omega))$ and $\sigma(\mathbf{x})$ follows since

$$S(\mathbf{x}) = \sigma(\mathbf{x})\mathrm{Im}(Q(\mathbf{x};\omega))/\omega.$$

In principle, measurements $\widehat{M_{i,j}}(\omega)$ for one single frequency are enough to find both parameters $\sigma(\mathbf{x})$ and $S(\mathbf{x})$. Unfortunately, this procedure seems to be much more sensitive to changes in σ than to changes in S. This is due to $\Delta\sigma^{1/2}$ appearing in the expression of $Q(\mathbf{x};\omega)$ (see remark 6). We deal with this problem by using data for two frequencies as is explained below.

2.6.3 Recovery of S and σ from two frequencies

Here the data we have is $\widehat{M_{i,j}}(\omega_1)$ and $\widehat{M_{i,j}}(\omega_2)$ for two frequencies $\omega_1 \neq \omega_2$ and we use it to solve two discrete measurements Schrödinger problems for $Q(\mathbf{x};\omega_1)$ and $Q(\mathbf{x};\omega_2)$, for $\mathbf{x} \in \Omega$. A good rule of thumb is to choose the frequencies so that ω_1 is sufficiently low to make $\mathrm{Re}(Q(\mathbf{x};\omega_1))$ the largest term in $Q(\mathbf{x};\omega_1)$ and ω_2 is sufficiently large to make $\mathrm{Im}(Q(\mathbf{x};\omega_2))$ the largest term in $Q(\mathbf{x};\omega_2)$. For each point $\mathbf{x}$ in Ω' (the domain without the wells), we solve for $r_1(\mathbf{x})$ and $r_2(\mathbf{x})$ in the 2×2 system:

$$\begin{bmatrix} 1 & \imath\omega_1 \\ 1 & \imath\omega_2 \end{bmatrix} \begin{bmatrix} r_1(\mathbf{x}) \\ r_2(\mathbf{x}) \end{bmatrix} = \begin{bmatrix} Q(\mathbf{x};\omega_1) \\ Q(\mathbf{x};\omega_2) \end{bmatrix}. \tag{2.51}$$

Then to estimate the conductivity we solve for $\sigma^{1/2}$ in the equation:

$$\Delta\sigma^{1/2} - r_1(\mathbf{x})\sigma^{1/2} = 0, \text{ for } \mathbf{x} \in \Omega', \tag{2.52}$$

with Dirichlet boundary condition given by the knowledge of σ on $\partial\Omega'$. Once we know σ, the storage coefficient S can be easily obtained from r_2, indeed:

$$S(\mathbf{x}) = \sigma(\mathbf{x})r_2(\mathbf{x}). \tag{2.53}$$

2.7 Numerical Experiments

We now present numerical experiments comparing inverse Born series with the Gauss-Newton and Chebyshev-Halley methods for both the discrete internal measurements Schrödinger problem (section 2.7.1) and an application to transient hydraulic tomography (section 2.7.2).

2.7.1 Schrödinger potential reconstructions

As discussed in section 2.3.5, our objective is to recover an unknown Schrödinger potential q from the measurements $f(q) = \mathbf{D}$, where the entries $D_{i,j}$ of the $N \times N$ matrix $\mathbf{D}$ are given by (2.2).

We discretize the computational domain $\Omega = [0,1]^2$ with a uniform grid consisting of the nodes (kh, lh), for $k, l = 0, \ldots, 400$ and $h = 1/400$. We use a total of 16 measurement functions ϕ_j, which are smooth and satisfy: $\|\phi_j\|_{L^\infty(\Omega)} = 1$ for $j = 1, \ldots, 16$; ϕ_j is compactly supported on a circle of radius $\rho = 0.05$; and the centers of the wells are uniformly spaced in the domain at the points $(0.2m, 0.2n)$ for $m, n = 1, \ldots, 4$. The Laplacian in the Schrödinger equation is discretized with the usual five point finite differences stencil and the true Schrödinger potential is simply evaluated at the grid nodes. The measurements $D_{i,j} = \langle \phi_j, u_i \rangle_{L^2(\Omega)}$ involve integrals that are approximated by the trapezoidal rule on the grid. Measurements $f(q_0)$ for the reference potential q_0 are computed in the same grid. The data that we use for the reconstructions is $f(q) - f(q_0)$.

The reconstructions are performed on a different (coarser) grid consisting of the nodes (kh_c, lh_c) for $k, l = 0, \ldots, 80$ and $h_c = 1/80$. We compare the results obtained from a truncated inverse Born series of order 5, and 10 iterations of the Gauss-Newton and Chebyshev-Halley methods. These three reconstructions are applied to F, a coarse grid version of the map f. For instance, the reconstructions for the inverse Born series are

$$\sum_{n=1}^{k} B_n((f(q) - f(q_0))^{\otimes n}),$$

where the coefficients B_n are the inverse Born series coefficients for the coarse grid F (rather than those for the fine grid f, which would be an inverse crime). For the inverse Born series, the operator B_1 is a regularized pseudoinverse of A_1 (i.e., the linearization of the coarse grid forward map F) where the singular values of A_1 which are less than 0.01 times the largest singular value (of A_1) are treated as zero. The same regularization is used for the Jacobians involved in the Gauss-Newton and Chebyshev-Halley methods. We use $q_0 = 0$ as the reference potential for the inverse Born series as well as the initial guess for the iterative Gauss-Newton and

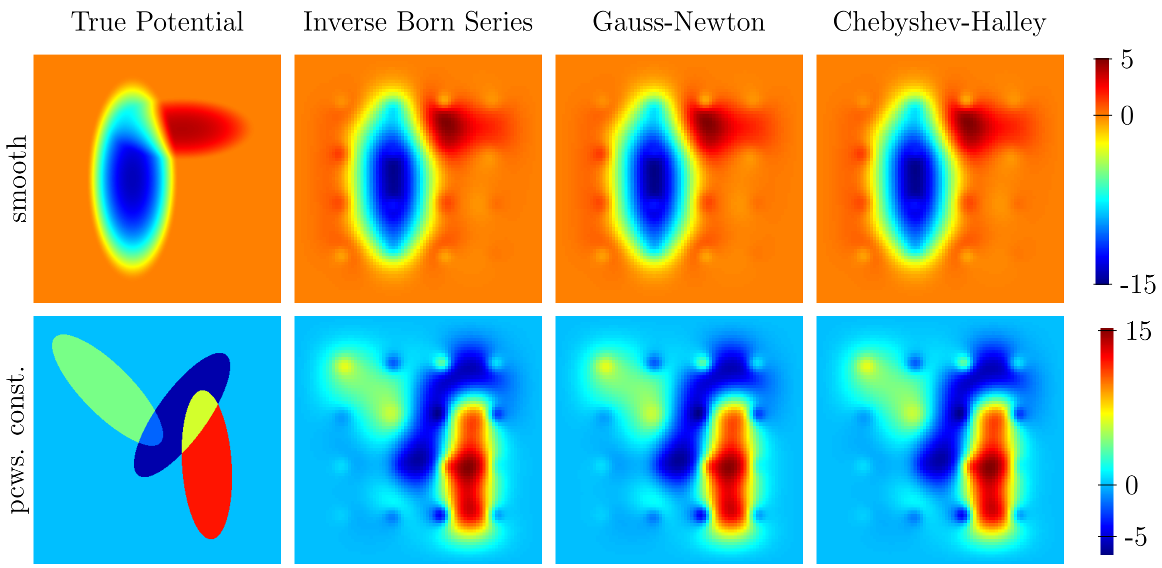


Figure 2.4. Comparison of reconstructions of a smooth (top) and piecewise constant (bottom) Schrödinger potential from discrete internal data at 16 locations and with no noise. The color scale is identical for all images in a row.

Chebyshev-Halley methods.

Figure 2.4 shows the reconstructions of a real smooth Schrödinger potential $-14 \leq q(\mathbf{x}) \leq 4$ and a real piecewise constant potential with $-6 \leq q(\mathbf{x}) \leq 12$. In both cases, the potential and the generated data are small enough to satisfy the hypotheses of theorem 3. Figure 2.5 displays the reconstructions of the same potentials from noisy data. The noisy data is obtained by first generating the true data $f(q) - f(q_0)$ as above, and then perturbing it with 1% zero mean additive Gaussian noise, i.e., with standard deviation $0.01 \max_{i,j} |(f(q) - f(q_0))_{i,j}|$. Similarly, Figure 2.6 displays the reconstructions with 5% additive Gaussian noise, i.e., with zero mean and standard deviation $0.05 \max_{i,j} |(f(q) - f(q_0))_{i,j}|$. In the experiments with noise present, the pseudoinverses of the Jacobians have been additionally regularized to compensate for the noise level (i.e. only singular values above 0.02 (resp. 0.06) times the largest singular value are retained for inversion for 1% (resp. 5%) noise).

2.7.2 Transient hydraulic tomography

In the frequency domain hydraulic tomography problem (see section 2.6), the objective is to estimate the hydraulic conductivity $\sigma(\mathbf{x})$ and the storage coefficient

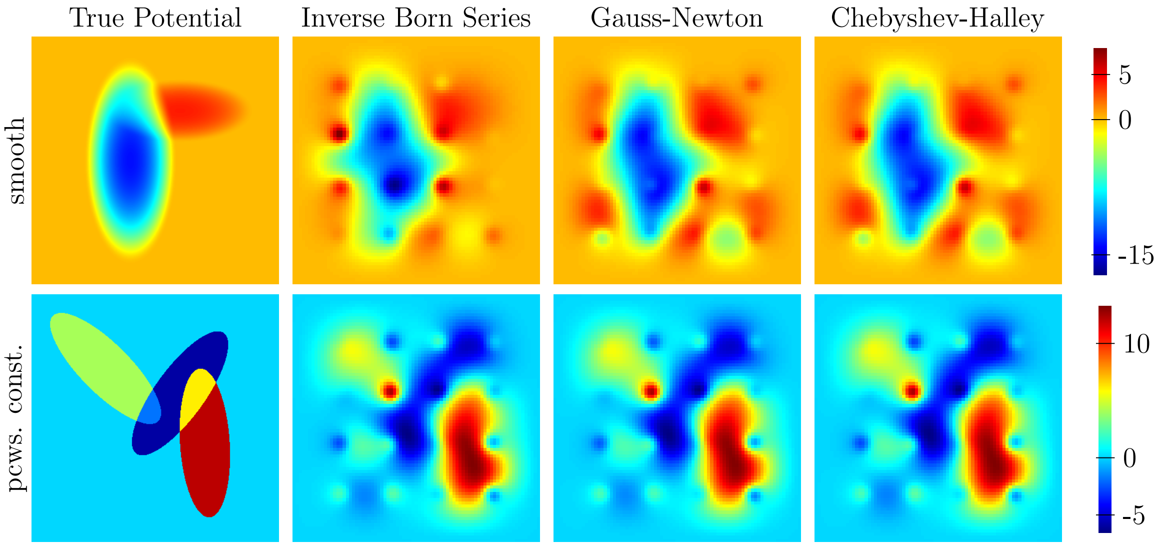


Figure 2.5. Comparison of reconstructions of a smooth (top) and piecewise constant (bottom) Schrödinger potential from discrete internal data at 16 locations and with 1% additive Gaussian noise. The color scale is identical for all images in a row.

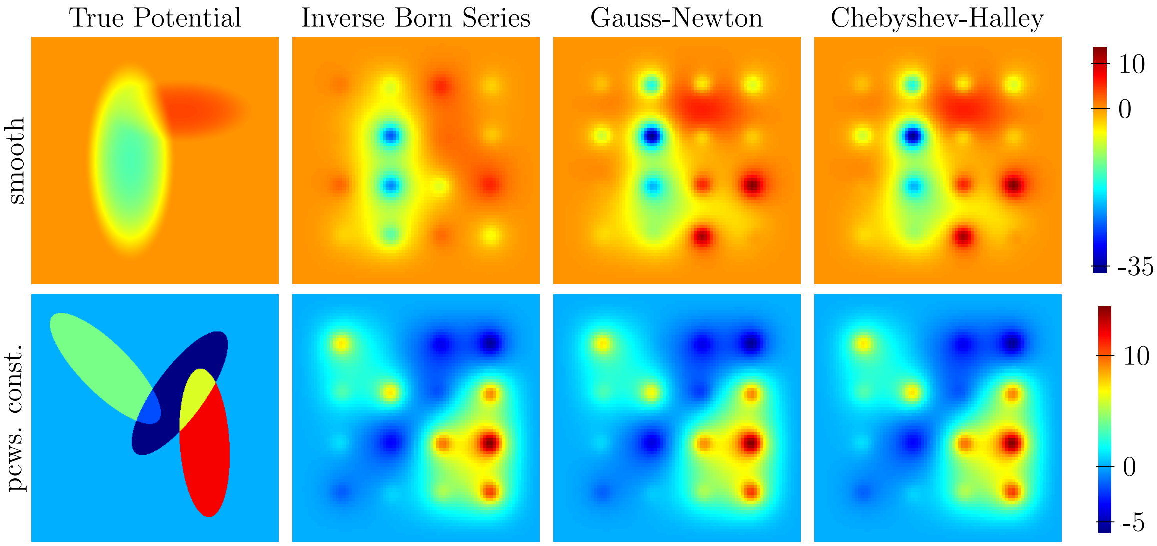


Figure 2.6. Comparison of reconstructions of a smooth (top) and piecewise constant (bottom) Schrödinger potential from discrete internal data at 16 locations and with 5% additive Gaussian noise. The color scale is identical for all images in a row.

$S(\mathbf{x})$ from the frequency-dependent measurements $\widehat{M}_{i,j}(\omega)$ defined in (2.48).

As before, the computational domain $\Omega = [0,1]^2$ is discretized with a uniform grid with nodes (kh, lh) for $k, l = 0, \ldots 400$ and $h = 1/400$. The true storage coefficient S is evaluated on this grid. The discretization of the term $\nabla \cdot [\sigma \nabla u]$ is done through the stencil

$$\begin{aligned}(\nabla \cdot [\sigma \nabla u])(kh, lh) \approx & \sigma_{k+1/2,l} \frac{u_{k+1,l} - u_{k,l}}{h^2} + \sigma_{k-1/2,l} \frac{u_{k-1,l} - u_{k,l}}{h^2} \\ & + \sigma_{k,l+1/2} \frac{u_{k,l+1} - u_{k,l}}{h^2} + \sigma_{k,l-1/2} \frac{u_{k,l-1} - u_{k,l}}{h^2},\end{aligned}$$

where $u_{k,l} \approx u(kh, lh)$ and similarly for σ. This means that the true conductivity is evaluated at the midpoints of the horizontal and vertical edges of the grid. The boundary points have a different stencil that takes into account the homogeneous Dirichlet boundary conditions, and that we do not include here for the sake of clarity.

The frequency domain measurement functions $\widehat{\phi}_i(\mathbf{x}, \omega)$ we use are, for simplicity, independent of the frequency ω and are given in $\mathbf{x}$ by the same 16 compactly supported smooth functions described in section 2.7.1. The measurements $\widehat{M}_{i,j}(\omega) = \langle \widehat{\phi}_j, \widehat{u}_i \rangle_{L^2(\Omega)}$ involve integrals over Ω that are evaluated by using the trapezoidal rule on the same grid that is used for the forward simulations. Recalling section 2.6.1, the measurements $\widehat{M}_{i,j}(\omega)$ can also be viewed as discrete internal measurements of a Schrödinger field v_i (see (2.49)) associated with the potential $Q(\mathbf{x}; \omega)$ defined in (2.50), i.e., $\widehat{\mathbf{M}}(\omega) = f(Q(\mathbf{x}; \omega))$ with well functions $\widehat{\phi}_i / \sigma^{1/2}$. We also compute measurements for the reference potential $Q_0 = 0$ on this grid using the well functions $\widehat{\phi}_i / \sigma^{1/2}$ (this corresponds to $S = 0$ and $\sigma = 1$). The measurements we use for reconstructions are $f(Q(\mathbf{x}; \omega)) - f(Q_0)$ (for two different frequencies).

Reconstructions are again performed on the coarse grid consisting of the nodes (kh_c, lh_c) for $k, l = 0, \ldots, 80$ and $h_c = 1/80$. For each method (inverse Born series order 5, Gauss-Newton, and Chebyshev-Halley), an approximation of the complex Schrödinger potential $Q(\mathbf{x}; \omega)$ is found from the frequency domain data $f(Q(\mathbf{x}; \omega)) - f(Q_0)$ for $\omega = 1, 10$. The parameters S and σ are then estimated with the procedure of section 2.6.3. The grid used for solving the problems (2.52) for the conductivity is the same coarse grid used for the reconstructions (to avoid an inverse crime). The boundary conditions for (2.52) are obtained from the true conductivity evaluated at appropriate points.

Figure 2.7 shows the reconstructions of the hydraulic conductivity σ and storage coefficient S when data has no noise. The conductivity σ is smooth and $|1-\sigma| < 0.8$. The storage coefficient S is also smooth and $-5 \leq S \leq 3$. We use the true conductivity σ inside the wells but the storage coefficient S inside the wells is computed, as in the rest of the domain, from (2.53). Reconstructions with 1% additive zero mean Gaussian noise are included in Figure 2.8. As before this means the noise has standard deviation $0.01 \max_{i,j} |[f(Q(\mathbf{x};\omega)) - f(Q_0)]_{i,j}|$, which is different for the two frequencies we use. Similarly, Figure 2.9 displays reconstructions with 5% additive zero mean Gaussian noise.

Remark 7 *In our experiments, the parameters σ and S are chosen so that the corresponding Schrödinger potential $Q(\mathbf{x};\omega)$ and the generated data are small enough to satisfy the hypotheses of theorem 3 (for $\omega = 1, 10$). This makes the contrasts in σ (especially) and S too small to represent a realistic problem (see, e.g., [4]). As noted before in remark 6, it may be possible to overcome this by using the inverse Born series on the hydraulic tomography problem directly.*

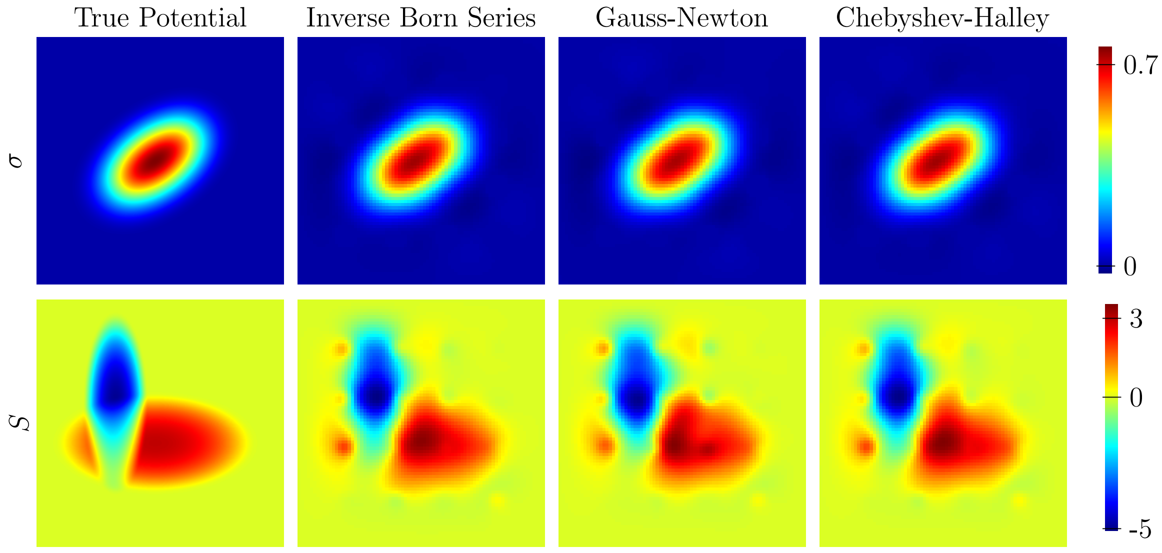


Figure 2.7. Hydraulic tomography reconstructions of the hydraulic conductivity $\sigma(\mathbf{x})$ (top) and the storage coefficient $S(\mathbf{x})$ (bottom) for noiseless data and different methods.

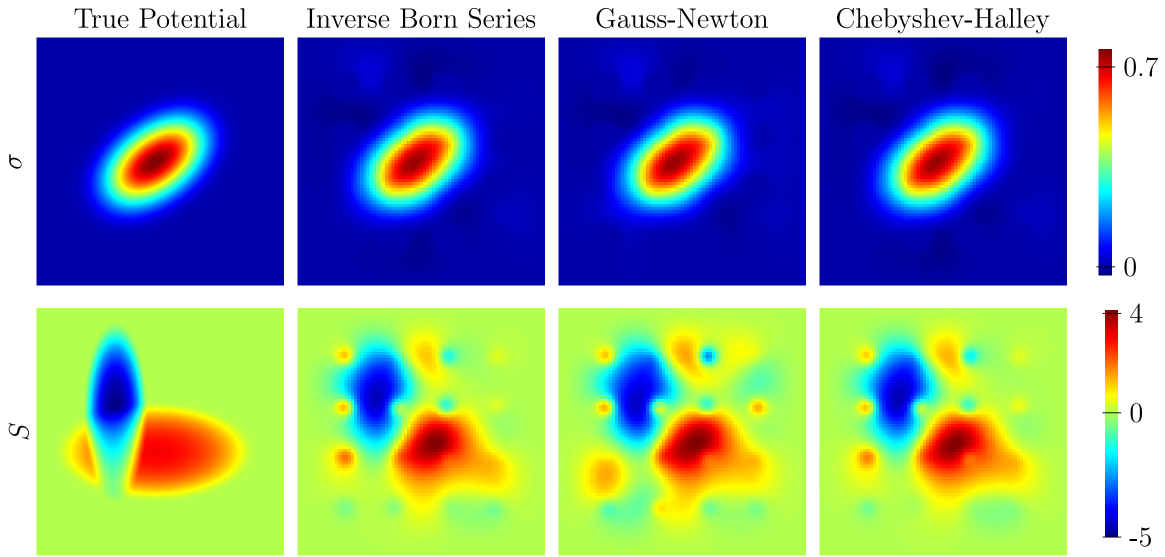


Figure 2.8. Hydraulic tomography reconstructions of the hydraulic conductivity $\sigma(\mathbf{x})$ (top) and the storage coefficient $S(\mathbf{x})$ (bottom) for data with 1% additive Gaussian noise and different methods.

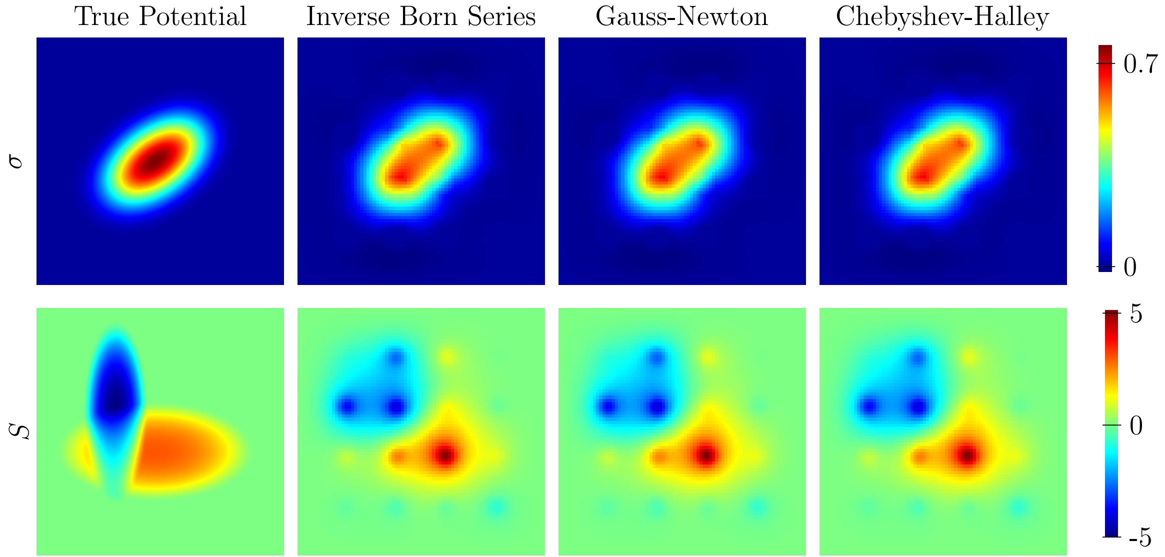


Figure 2.9. Hydraulic tomography reconstructions of the hydraulic conductivity $\sigma(\mathbf{x})$ (top) and the storage coefficient $S(\mathbf{x})$ (bottom) for data with 5% additive Gaussian noise and different methods.

2.8 Discussion

We show here that with little modification, the inverse Born series convergence results of Moskow and Schotland [12] can be generalized to mappings between Banach spaces. With this abstraction, we only need to show that the forward Born operators are bounded as in (2.6) to obtain convergence, stability and error estimates for the inverse Born series. Such results are then proven for the problem of finding the Schrödinger potential from discrete internal measurements. A nice byproduct of our approach is that we can relate forward and inverse Born series coefficients (up to a symmetrization) to the Taylor series coefficients of an analytic map and its inverse (provided it exists).

Since the cost of computing the $n-$th term of the inverse Born series increases exponentially in n, we also consider the iterative method obtained by restarting the inverse Born series after summing the first k terms. We obtain a class of methods that we call RIBS(k) and that includes the well-known Gauss-Newton and Chebyshev-Halley iterative methods. Our numerical results show these methods give reconstructions comparable to those obtained with the inverse Born series.

Among the future directions of this work would be to show the RIBS(k) method is convergent. We conjecture that the convergence rate of RIBS(k) is of order k. The RIBS(k) method is only locally convergent, meaning that we need to be already close to the solution for the method to converge. Globalization strategies that keep, when possible, this higher order convergence rate are needed.

The application we use to illustrate our method is a problem related to transient hydraulic tomography. Since we convert this problem to the problem of finding a Schrödinger potential and all the methods we use here are locally convergent, the contrasts that we can deal with are far from realistic ones. We believe that a proper globalization strategy will allow us to deal with higher contrasts. Another important question that we have not dealt with here is that of regularization. The only regularization that we consider here is the choice of the linear operator that primes the inverse Born series. By analogy with what can be done with the Gauss-Newton method, we believe it is possible to include specific a priori information about the true parameters by formulating the problem as minimizing the misfit plus a penalty term that takes into account the a priori information.

2.9 Inverse Born series convergence and stability proofs

The proofs in this section are an adaptation of the proofs in Moskow and Schotland [12] to inverse Born series in Banach spaces. The results are stated in section 2.2.3.

2.9.1 Proof of bounds for inverse Born series coefficients (lemma 1)

Proof. Since $\|a_n\| \leq \alpha\mu^n$, we can estimate for $n \geq 2$:

$$\begin{aligned} \|b_n\| &\leq \sum_{m=1}^{n-1} \sum_{s_1+\cdots+s_m=n} \|b_m\| \|a_{s_1}\| \cdots \|a_{s_m}\| \|b_1\|^n \\ &\leq \|b_1\|^n \sum_{m=1}^{n-1} \|b_m\| \sum_{s_1+\cdots+s_m=n} (\alpha\mu^{s_1}) \dots (\alpha\mu^{s_m}) \\ &= \|b_1\|^n \mu^n \sum_{m=1}^{n-1} \|b_m\| \alpha^m \sum_{s_1+\cdots+s_m=n} 1. \end{aligned} \tag{2.54}$$

The last sum is the number of partitions of the integer n into m ordered parts. Hence for $n \geq 2$, we get

$$\begin{aligned} \|b_n\| &\leq (\mu\|b_1\|)^n \sum_{m=1}^{n-1} \|b_m\| \alpha^m \binom{n-1}{m-1} \\ &\leq (\mu\|b_1\|)^n \left(\sum_{m=1}^{n-1} \|b_m\| \right) \left(\sum_{m=1}^{n-1} \alpha^m \binom{n-1}{m-1} \right) \\ &\leq (\mu\|b_1\|(\alpha+1))^n \sum_{m=1}^{n-1} \|b_m\|. \end{aligned} \tag{2.55}$$

To get the last inequality we used

$$\sum_{m=1}^{n-1} \alpha^m \binom{n-1}{m-1} = \sum_{m=0}^{n-2} \alpha^{m+1} \binom{n-1}{m} \leq \alpha \sum_{m=0}^{n-1} \alpha^{m+1} \binom{n-1}{m} = \alpha(1+\alpha)^{n-1} \leq (1+\alpha)^n.$$

Following [12] we can estimate the coefficients in the inverse Born series by

$$\|b_n\| \leq C_n (\mu\|b_1\|(\alpha+1))^n \|b_1\|, \text{ for } n \geq 2, \tag{2.56}$$

where the constants C_n are defined recursively by

$$C_2 = 1 \text{ and } C_{n+1} = 1 + ((\alpha+1)\mu\|b_1\|)^n \text{ for } n \geq 2. \tag{2.57}$$

The constants C_n are then

$$C_n = \prod_{m=2}^{n-1} (1 + ((\alpha+1)\mu\|b_1\|)^m) \leq \exp\left(\frac{1}{1-(\alpha+1)\mu\|b_1\|}\right). \tag{2.58}$$

where the bound for C_n can be derived as in [12] and is valid when $(\alpha+1)\mu\|b_1\| < 1$, which is one of the hypotheses. The result follows from the bounds (2.56) and (2.58). ■

2.9.2 Proof of local convergence of inverse Born series (theorem 1)

Proof. Using the estimate of lemma 1, we can dominate the term of the inverse Born series by a geometric series as follows

$$\|b_n(d^{\otimes n})\| \leq \beta((\alpha+1)\mu\|b_1\|\|d\|)^n. \tag{2.59}$$

Therefore the Born series is absolutely convergent when $(\alpha+1)\mu\|b_1\|\|d\| < 1$, which is one of the assumptions of this theorem. The tail of the series with terms the absolute values of the inverse Born series terms, can be estimated by noticing that

$$\sum_{N+1}^{\infty} \beta((\alpha+1)\mu\|b_1\|\|d\|)^n = \beta \frac{((\alpha+1)\mu\|b_1\|\|d\|)^{N+1}}{1-(\alpha+1)\mu\|b_1\|\|d\|}. \tag{2.60}$$

■

2.9.3 Proof of stability of inverse Born series (theorem 2)

Proof. We use an identity on tensor products to conclude that

$$\begin{aligned} \|h_1 - h_2\| &\leq \sum_{n=1}^{\infty} \|b_n(d_1^{\otimes n} - d_2^{\otimes n})\| \\ &= \sum_{n=1}^{\infty} \left\| b_n \left(\sum_{k=0}^{n-1} d_1^{\otimes k} \otimes (d_1 - d_2) \otimes d_2^{\otimes(n-k-1)} \right) \right\| \\ &\leq \sum_{n=1}^{\infty} n M^{n-1} \|b_n\| \|d_1 - d_2\|. \end{aligned} \tag{2.61}$$

The desired estimate follows from applying the estimate for the $\|b_n\|$ in lemma 1,

$$\begin{aligned} \|h_1 - h_2\| &\leq \|d_1 - d_2\| \sum_{n=1}^{\infty} n M^{n-1} \beta((\alpha+1)\mu\|b_1\|)^n \\ &\leq \|d_1 - d_2\| \frac{\beta}{M} \frac{1}{(1 - M(\alpha+1)\mu\|b_1\|)^2}, \end{aligned} \tag{2.62}$$

since we assumed that $M(\alpha+1)\mu\|b_1\| < 1$. Here we used the following inequality:

$$\beta \sum_{n=1}^{\infty} n M^{n-1} \delta^n = \frac{\beta}{M} \sum_{n=1}^{\infty} n (M\delta)^n \leq \frac{\beta}{M} \sum_{n=0}^{\infty} (n+1)(M\delta)^n = \frac{\beta}{M} \frac{1}{(1-M\delta)^2}$$

where $\delta \equiv (\alpha+1)\mu\|b_1\|$. ■

2.9.4 Proof of inverse Born series error estimate (theorem 3)

Proof. Taking the expression for d in (2.17) and replacing in the expression for h_* in (2.17) we get

$$h_* = \sum_{n=1}^{\infty} c_n(h^{\otimes n}), \tag{2.63}$$

where

$$\begin{aligned} c_1 &= b_1 a_1, \\ c_n &= \left(\sum_{m=1}^{n-1} b_m \left(\sum_{s_1+\cdots s_m = n} a_{s_1} \otimes \cdots \otimes a_{s_m} \right) \right) + b_n(a_1^{\otimes n}), \text{ for } n \geq 2. \end{aligned} \tag{2.64}$$

Using the expression (2.10) of b_n in terms of b_m, $1 \leq m \leq n-1$, we get for $n \geq 2$ that

$$c_n = \sum_{m=1}^{n-1} b_m \left(\sum_{s_1+\cdots s_m = n} a_{s_1} \otimes \cdots \otimes a_{s_m} \right) \left(I - (b_1 a_1)^{\otimes n} \right). \tag{2.65}$$

Hence the reconstruction error is

$$h - h_* = (h - b_1 a_1 h) - \sum_{n=2}^{\infty} \sum_{m=1}^{n-1} b_m \left(\sum_{s_1+\cdots s_m = n} a_{s_1} \otimes \cdots \otimes a_{s_m} \right) \left(h^{\otimes n} - (b_1 a_1 h)^{\otimes n} \right). \tag{2.66}$$

We now estimate the error:

$$\|h - h_*\| \leq \|h - b_1 a_1 h\| + \sum_{n=2}^{\infty} \sum_{m=1}^{n-1} \sum_{s_1+\cdots s_m = n} \|b_m\| \|a_{s_1}\| \cdots \|a_{s_m}\| \left\| h^{\otimes n} - (b_1 a_1 h)^{\otimes n} \right\|. \tag{2.67}$$

For $n \geq 1$ we can estimate:

$$\begin{aligned} \|h^{\otimes n} - (b_1 a_1 h)^{\otimes n}\| &= \left\| \sum_{k=0}^{n-1} h^{\otimes k} \otimes (h - b_1 a_1 h) \otimes (b_1 a_1 h)^{\otimes (n-k-1)} \right\| \\ &\leq n M^{n-1} \|h - b_1 a_1 h\|, \end{aligned} \tag{2.68}$$

where we used the hypothesis $\|h\| \leq M$, $\|b_1 a_1 h\| \leq M$. Since we assumed the Born series coefficients satisfy $\|a_n\| \leq \alpha\mu^n$ we get

$$\begin{aligned}\|h - h_*\| &\leq \|h - b_1 a_1 h\| \left(1 + \sum_{n=2}^{\infty} \sum_{m=1}^{n-1} \sum_{s_1 + \cdots s_m = n} \|b_m\| (\alpha\mu^{s_1}) \cdots (\alpha\mu^{s_m}) n M^{n-1}\right) \\ &= \|h - b_1 a_1 h\| \left(1 + \sum_{n=2}^{\infty} \sum_{m=1}^{n-1} \|b_m\| \alpha^m n \mu^n M^{n-1} \binom{n-1}{m-1}\right).\end{aligned} \tag{2.69}$$

Here we have used again the fact that the number of ordered partitions of n into m integers is

$$\sum_{s_1 + \cdots s_m = n} 1 = \binom{n-1}{m-1}.$$

Clearly we have that

$$\|h - h_*\| \leq \|h - b_1 a_1 h\| \left(1 + \sum_{n=2}^{\infty} n \mu^n M^{n-1} \left(\sum_{m=1}^{n-1} \|b_m\|\right) \left(\sum_{m=1}^{n-1} \alpha^m \binom{n-1}{m-1}\right)\right). \tag{2.70}$$

Now using the two facts:

$$\begin{aligned}\sum_{m=1}^{n-1} \|b_m\| &\leq \beta \sum_{m=1}^{n-1} ((\alpha+1)\mu\|b_1\|)^m \text{ (lemma 1)}, \\ \sum_{m=1}^{n-1} \alpha^m \binom{n-1}{m-1} &\leq (1+\alpha)^n \text{ (as in (2.55))},\end{aligned} \tag{2.71}$$

we get the inequality

$$\|h - h_*\| \leq \|h - b_1 a_1 h\| \left(1 + \sum_{n=2}^{\infty} \frac{n}{M} (\mu M(1+\alpha))^n \beta \sum_{m=1}^{n-1} ((\alpha+1)\mu\|b_1\|)^m\right). \tag{2.72}$$

Adding the $m = 0$ term to the geometric series over m and summing we get

$$\|h - h_*\| \leq \|h - b_1 a_1 h\| \left(1 + \frac{\beta}{M} \sum_{n=1}^{\infty} n (\mu M(1+\alpha))^n \frac{1 - ((\alpha+1)\mu\|b_1\|)^n}{1 - (\alpha+1)\mu\|b_1\|}\right). \tag{2.73}$$

The hypothesis $\mu M(\alpha+1) < 1$ and $\mu(\alpha+1)\|b_1\| < 1$ imply the quantity in parentheses is bounded and depends only on M, α, β and μ and $\|b_1\|$. ■

2.10 Acknowledgments

The authors would like to thank Liliana Borcea, Alexander V. Mamonov, Shari Moskow, John Schotland and Andy Thaler for insightful conversations on this subject.

FGV is grateful to Otmar Scherzer for pointing out reference [7]. The work of the authors was partially supported by the National Science Foundation grant DMS-0934664.

2.11 References

[1] S. Arridge, S. Moskow, and J. C. Schotland, *Inverse Born series for the Calderon problem*, Inverse Problems, 28 (2012), pp. 035003, 16.

[2] S. R. Arridge, *Optical tomography in medical imaging*, Inverse Problems, 15 (1999), p. R41.

[3] L. Borcea, *Electrical impedance tomography*, Inverse Problems, 18 (2002), pp. R99–R136.

[4] M. Cardiff and W. Barrash, *3-D transient hydraulic tomography in unconfined aquifers with fast drainage response*, Water Resources Research, 47 (2011), p. W12518.

[5] P. Deuflhard, *Newton Methods For Nonlinear Problems: Affine Invariance and Adaptive Algorithms*, vol. 35 of Springer Series in Computational Mathematics, Springer, Heidelberg, 2011.

[6] L. Evans, *Partial Differential Equations*, Graduate studies in mathematics, American Mathematical Society, 2010.

[7] F. Hettlich and W. Rundell, *A second degree method for nonlinear inverse problems*, SIAM J. Numer. Anal., 37 (2000), pp. 587–620.

[8] K. Kilgore, S. Moskow, and J. C. Schotland, *Inverse Born series for scalar waves*, J. Comput. Math., 30 (2012), pp. 601–614.

[9] E. Kreyszig, *Introductory functional analysis with applications*, Wiley Classics Library, John Wiley & Sons Inc., New York, 1989.

[10] V. A. Markel, J. A. O'Sullivan, and J. C. Schotland, *Inverse problem in optical diffusion tomography: IV: Nonlinear inversion formulas*, J. Opt. Soc. Am. A, 20 (2003), pp. 903–912.

[11] V. A. Markel and J. C. Schotland, *On the convergence of the Born series in optical tomography with diffuse light*, Inverse Problems, 23 (2007), pp. 1445–1465.

[12] S. Moskow and J. C. Schotland, *Convergence and stability of the inverse scattering series for diffuse waves*, Inverse Problems, 24 (2008), p. 065005.

[13] ——, *Numerical studies of the inverse Born series for diffuse waves*, Inverse Problems, 25 (2009), p. 095007.

[14] R. A. Ryan, *Introduction to tensor products of Banach spaces*, Springer Monographs in Mathematics, Springer-Verlag London Ltd., London, 2002.

[15] E. F. Whittlesey, *Analytic functions in Banach spaces*, Proc. Amer. Math. Soc., 16 (1965), pp. 1077–1083.

[16] E. Zeidler, *Nonlinear functional analysis and its applications: I: Fixed-point theorems*, Springer-Verlag, New York, 1986.

CHAPTER 3

ARRAY IMAGING WITH POWER-CONTROLLED SOURCE PAIRS

A joint work with Fernando Guevara Vasquez.

Originally published in the SIAM Journal on Imaging Sciences. Bardsley, P., Guevara Vasquez, F. *Imaging with power controlled source pairs.* Volume 9, Issue 1, Pg 185-211. February 09, 2016. DOI: 10.1137/15M1023191.

3.1 Introduction

Scatterers in a homogeneous medium can be imaged by probing the medium with a wave emanating from a point source, and recording the reflected waves at one or more receivers. An image of the scatterers can be generated by repeating this experiment while varying the position of the source and/or receiver and using classic methods such as Kirchhoff (travel time) migration (see, e.g., [2]) or MUSIC (see, e.g., [8]). We are concerned here with the case where only *intensity measurements* can be made at the receiver; destroying phase information that migration and MUSIC need to image. Intensity measurements occur, e.g., when the response time of the receiver is larger than the typical wave period or when it is more cost effective to measure intensities than the full waveform. This is typical in, e.g., optical coherence tomography [28, 29] and radar imaging [9]. Another situation is when the wave sources are stochastic and the measurements consist of correlations of the signal recorded at different points [12, 32]. In the special case of autocorrelations (i.e., correlating the signal with itself), the Wiener-Khinchin theorem guarantees we are measuring power spectra (see, e.g., [21]), another form of intensity measurements.

The setup we analyze consists of an array of sources and one single receiver that can only record power spectra, i.e., the intensity of the signal at certain frequency samples. A crucial assumption for our method is that we can use *source pairs*, meaning we can send *correlated signals* from *two different locations*. However we do allow for certain known delays or attenuations between the signals in a source pair. One way of achieving this would be to drive two transducers in an array with the *same* signal. This would be possible when the source phases are relatively easy to control, e.g., in acoustics.

At optical frequencies, controlling phases is challenging. For light, one could use a spatial light modulator or configurable mask (e.g., an LCD mask [20, 24]). By placing a light source behind the mask and letting light through only two small holes (e.g., pixels for an LCD mask), the holes behave as two point sources (Huygens-Fresnel principle). If the light source is a plane wave propagating in the direction perpendicular to the mask, the two point sources have identical phases (see Figure 3.1). The illuminations we obtain are similar to doing Young's double slit experiment with

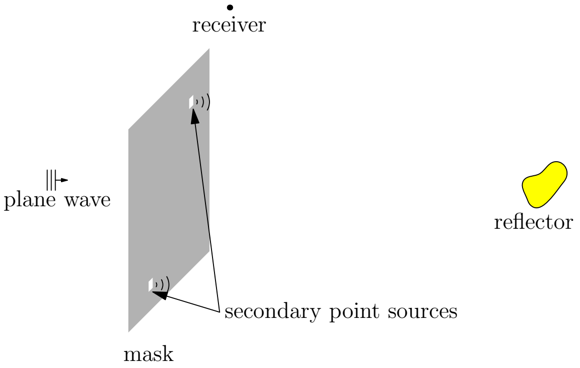


Figure 3.1. One way of obtaining two point sources with identical phases by using a plane wave, a configurable mask and the Huygens-Fresnel principle.

pinholes whose location can be adjusted. The same effect may be achieved with a Digital Micromirror Device (see, e.g., [27] for a review). Yet another way would be to use optical fibers (see, e.g., [1]) to divide a signal and pipe it along two paths of equal length to two different locations.

Our method can be used for imaging from both measurements of intensities (§3.1.1) and autocorrelations (§3.1.2).

3.1.1 Intensity only measurements

One way to deal with intensity measurements is *phase retrieval*, i.e., first recovering the phases from intensity measurements and using this reconstructed field to image. In diffraction tomography, intensity measurements at two different planes can be used to recover phases [17, 18, 34]. If additional information is known (e.g., the support of the scatterer), intensities at one single plane can be used [11, 22, 23]. Total or partial knowledge of the incident field can also be exploited to image from intensities at one single plane [10].

Chai, Moscoso, and Papanicolaou [6] recast the phase retrieval problem as the problem of recovering a rank-one matrix with convex programming techniques. With knowledge of the incident field, the location of a few point scatterers can then be resolved in both range and cross-range with monochromatic measurements. The same ideas can even be used to deal with multiple scattering [7]. Novikov, Moscoso, and Papanicolaou [25] use the polarization identity $4\mathrm{Re}(\boldsymbol{u}^*\boldsymbol{v}) = \|\boldsymbol{u}+\boldsymbol{v}\|^2 - \|\boldsymbol{u}-\boldsymbol{v}\|^2$, $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{C}^N$, and linear combinations of single-source experiments to recover dot products of two single-source experiments from intensity data. MUSIC can then be used to image with this quadratic functional of the full waveform data. Other convex programming techniques for the phase retrieval problem are [5, 36].

Here we do not completely recover phases of the total field, but we do recover the projection of the total field onto a known subspace and we show that this is enough to image with Kirchhoff migration. To achieve this, we assume the scattered field is small to reduce the problem of recovering the total field from intensity measurements to a linear system. The linear system has a one-dimensional nullspace that we can write explicitly in terms of the incident field. There is one (very sparse) linear system per frequency sample to solve, and the linear system has size comparable to twice the number of source positions. Intuitively we are recovering a field in $\mathbb{C}^N$ from $2N$ (or more) real measurements. Our main result (theorem 5) shows that vectors in the one-dimensional nullspace do not affect Kirchhoff migration. Hence we can use, without modification, Kirchhoff migration and its standard range and cross-range resolution estimates (see, e.g., [2]).

3.1.2 Correlation-based methods

In seismic imaging, correlations of traces (or recordings) at many receivers have been used to image the earth's subsurface, especially when the wave sources and their locations are not well known [30, 31, 32]. The idea is that correlations of the signals at two different locations contain information about the Green's function between the two locations, and this information can be exploited to image the medium and any scatterers. This principle can even be exploited to do opportunistic imaging with ambient noise [12, 13, 16]. Cross-correlations can also be used to image scatterers in a random medium [4, 14, 15]. In radar imaging, the measurements are in fact correlations [9], and so even stochastic processes can be used instead of deterministic signals [33, 35].

The method we present here can also be used to image scatterers using autocorrelations. We show it is possible to form an image by exploiting angular diversity in source pairs instead of cross-correlations among different receivers. Just as in the intensity measurements case, we are able to recover (up to a one-dimensional nullspace) full waveform array measurements. One advantage of using autocorrelations instead of cross-correlations is that the data acquisition at the (single) receiver is simpler. The drawback is that our illumination strategy requires us to illuminate with pairs of sources, but also with each of the sources in a pair on its own. To get the same full waveform data as an array with N sources, we need at least $3N$ different experiments. Another advantage of using autocorrelations is that the measurements are extremely robust to noise. As an example, our numerical experiments show that it is possible to image scatterers with an array that is sending noise from all possible source positions; the only assumption about the noise being that all the sources are independent stochastic processes except for two correlated sources whose positions we can control. Because of this robustness, it may be possible to use our imaging method in situations where the medium is to be probed in a nonintrusive way, i.e., active imaging with waves that are of the same magnitude as the ambient noise.

3.1.3 Contents

The particular physical setup we consider is described in §3.2. The illumination strategy with source pairs is explained in §3.3, which leads to a phase retrieval problem that can be formulated as a linear system (§3.4). The least-squares solution to the linear system is then used as data for imaging with Kirchhoff migration, and we show that this gives essentially the same images as full waveform data (§3.5). The extension to stochastic source pairs is given in §3.6. Then we show that our method is robust to additive noise when using autocorrelations (§3.7). Numerical experiments illustrating our method are given in §4.5 and we conclude with a discussion in §4.6.

3.2 Array imaging for full waveform measurements

Here we introduce the experimental setup we consider (§3.2.1) and recall the classic Kirchhoff migration imaging method (§3.2.2).

3.2.1 Experimental setup

The physical setup is illustrated in Figure 3.2. We probe a homogeneous medium with waves originating from N point sources with locations $\vec{\boldsymbol{x}}_s \in \mathcal{A}$, $s = 1, 2, \ldots, N$. For simplicity we consider a linear array in 2D or a square array in 3D, i.e., $\mathcal{A} =$

Figure 3.2. Physical setup for array imaging with an array $\mathcal{A}$ of sources $\vec{\boldsymbol{x}}_s$ and a single receiver $\vec{\boldsymbol{x}}_r$. The scatterer is represented by a compactly supported reflectivity function $\rho(\vec{\boldsymbol{x}})$.

$[-a/2, a/2]^{d-1} \times \{0\}$, where $d = 2$ or 3 is the dimension. Our imaging strategy imposes only mild restrictions on the source positions, so other array shapes may be considered. Waves are recorded at a *single* known receiver location $\vec{\boldsymbol{x}}_r$.

The total field generated by the array (or incident field) can be written as

$$\widehat{u}_{\text{inc}}(\vec{\boldsymbol{x}}, \omega) = \boldsymbol{g}_0(\vec{\boldsymbol{x}}, \omega)^{\mathsf{T}} \boldsymbol{f}(\omega), \tag{3.1}$$

where

$$\boldsymbol{g}_0(\vec{\boldsymbol{x}}, \omega) = \left[\widehat{G}_0(\vec{\boldsymbol{x}}, \vec{\boldsymbol{x}}_1, \omega), \widehat{G}_0(\vec{\boldsymbol{x}}, \vec{\boldsymbol{x}}_2, \omega), \ldots, \widehat{G}_0(\vec{\boldsymbol{x}}, \vec{\boldsymbol{x}}_N, \omega)\right]^{\mathsf{T}} \in \mathbb{C}^N, \tag{3.2}$$

and the source driving signals are $\boldsymbol{f}(\omega) = [\widehat{f}_1(\omega), \widehat{f}_2(\omega), \ldots, \widehat{f}_N(\omega)]^{\mathsf{T}}$. Since we assume waves propagate through a homogeneous medium, we used the outgoing free space Green function,

$$\widehat{G}_0(\vec{\boldsymbol{x}}, \vec{\boldsymbol{y}}, \omega) = \begin{cases} \dfrac{\imath}{4} H_0^{(1)}(k|\vec{\boldsymbol{x}} - \vec{\boldsymbol{y}}|), & d = 2, \\ \dfrac{\exp(\imath k|\vec{\boldsymbol{x}} - \vec{\boldsymbol{y}}|)}{4\pi|\vec{\boldsymbol{x}} - \vec{\boldsymbol{y}}|}, & d = 3. \end{cases} \tag{3.3}$$

Here $H_0^{(1)}$ is the zeroth order Hankel function of the first kind, $k = \omega/c_0$ is the wavenumber, ω is the angular frequency and c_0 is a known constant background wave speed. For functions of time, the Fourier transform convention we use is

$$\widehat{f}(\omega) = \int f(t) e^{\imath\omega t} dt, \text{ and } f(t) = \frac{1}{2\pi} \int \widehat{f}(\omega) e^{-\imath\omega t} d\omega, \text{ where } f \in L^2(\mathbb{R}). \tag{3.4}$$

The scatterers we want to image are represented by a reflectivity function (or scattering potential) $\rho(\vec{\boldsymbol{x}})$ with compact support $\mathcal{P}$. We assume the scatterers are sufficiently weak (e.g., $\rho \ll 1$) so that the total field at the receiver can be expressed using the Born approximation:

$$\widehat{u}(\vec{\boldsymbol{x}}_r, \omega) = (\boldsymbol{g}_0 + \boldsymbol{p})^{\mathsf{T}} \boldsymbol{f}, \tag{3.5}$$

where the array response vector is

$$\boldsymbol{p}(\vec{\boldsymbol{x}}, \omega) = k^2 \int_{\mathcal{P}} d\vec{\boldsymbol{y}} \rho(\vec{\boldsymbol{y}}) \widehat{G}_0(\vec{\boldsymbol{x}}, \vec{\boldsymbol{y}}, \omega) \boldsymbol{g}_0(\vec{\boldsymbol{y}}, \omega). \tag{3.6}$$

We emphasize here that we neglect multiple scattering events, which leaves out situations where the sources, scatterer, and receiver are not in direct line of sight of each other.

3.2.2 Kirchhoff migration

By, e.g., using illuminations $\boldsymbol{f}(\omega) = \boldsymbol{e}_i$, $i = 1, \ldots, N$ corresponding to the canonical basis vectors, it is possible to obtain the array response vector $\boldsymbol{p}(\vec{\boldsymbol{x}}_r, \omega)$ from the measurements (3.5). The scatterers can then be imaged using the Kirchhoff migration functional (see, e.g., [2]), which for a single frequency ω is

$$\Gamma_{\text{KM}}[\boldsymbol{p}, \omega](\vec{\boldsymbol{y}}) = \overline{\widehat{G}_0}(\vec{\boldsymbol{y}}, \vec{\boldsymbol{x}}_r, \omega) \boldsymbol{g}_0(\vec{\boldsymbol{y}}, \omega)^* \boldsymbol{p}(\vec{\boldsymbol{x}}_r, \omega), \tag{3.7}$$

where $\vec{\boldsymbol{y}}$ represents a point in the image. This image has a Rayleigh or cross-range (i.e., in the direction parallel to the array) resolution of $\lambda L/a$, where L is the distance from the array to the scatterer (see, e.g., [2]). To get range (i.e., in the direction perpendicular to the array) resolution we need to integrate $\Gamma_{\text{KM}}[\boldsymbol{p}, \omega](\vec{\boldsymbol{y}})$ for frequencies ω in some frequency band $\mathcal{B} = [-\omega_{max}, -\omega_{min}] \cup [\omega_{min}, \omega_{max}]$, the same frequency band of the signals $\boldsymbol{f}(\omega)$. The range resolution is then $c_0/(\omega_{max} - \omega_{min})$ (see, e.g., [2]). We discuss this imaging functional further in section 3.5.

3.3 Intensity only measurements

We begin this section by making a smallness assumption on the scattered field $\boldsymbol{p}$ with respect to the incident field when evaluated *at the receiver* $\vec{\boldsymbol{x}}_r$ (assumption 1). This allows us to neglect higher order terms in $\boldsymbol{p}$ and formulate the recovery of $\boldsymbol{p}$ as a linear system (§3.3.2).

Assumption 1 *We assume the reflectivity* ρ *with* $\operatorname{supp}(\rho) = \mathcal{P}$ *is such that*

$$|\boldsymbol{p}(\vec{\boldsymbol{x}}_r, \omega)| \ll |\boldsymbol{g}_0(\vec{\boldsymbol{x}}_r, \omega)| \tag{3.8}$$

where $\boldsymbol{g}_0$ *and* $\boldsymbol{p}$ *are given by* (3.2) *and* (3.6)*, respectively.*

One physical setup where (3.8) is satisfied (for fixed k) is when the reflectivity $\|\rho\|_\infty \ll 1$ and the receiver location $\vec{\boldsymbol{x}}_r$ is near the source array $\mathcal{A}$, as in Figure 3.2. The condition breaks down if the receiver location $\vec{\boldsymbol{x}}_r$ (resp. the source array $\mathcal{A}$) is too close to the scatterer support $\mathcal{P}$ since the Green functions in (3.6) become large. Condition (3.8) also breaks down (for fixed $\mathcal{A}$ and $\vec{\boldsymbol{x}}_r$) if $k^2\|\rho\|_\infty$ is large.

Remark 8 *We consider for simplicity real reflectivities* ρ*. Complex conductivities that are independent of the frequency* ω *may also be considered with minor modifications.*

3.3.1 Illumination strategy

The data we use come from probing the medium with N_p source pairs that are sending signals with known power and phase difference. Since the number of distinct source pairs out of an array with N sources is $N(N-1)/2$ we must have $N_p \le N(N-1)/2$. We assume the power and phase differences remain the same for all N_p illuminations. The case where these quantities depend on the source pair is left for future studies. To be more precise, the illumination corresponding to the $m-$th source pair $(i(m), j(m)) \in \{1, \dots, N\}^2$ is

$$\boldsymbol{f}_m(\omega) = \mathsf{F}_m \begin{bmatrix} \alpha(\omega) \\ \beta(\omega) \end{bmatrix}, \text{ where } \mathsf{F}_m = [\boldsymbol{e}_{i(m)}, \boldsymbol{e}_{j(m)}] \in \mathbb{R}^{N\times 2}. \tag{3.9}$$

We emphasize that only $|\alpha|^2$, $|\beta|^2$ and the phase difference $\phi(\omega) \equiv \arg(\overline{\alpha}\beta)$ is assumed to be known for the signals α and β. A particular case is when the same signal is sent from the source pair, i.e., $\beta = \alpha$ and $\phi(\omega) = 0$.

The intensity of the field u_m arising from the source pair illumination $\boldsymbol{f}_m$ is

$$|\widehat{u}_m(\vec{\boldsymbol{x}}_r, \omega)|^2 = \overline{\boldsymbol{g}^\mathsf{T} \boldsymbol{f}_m} \boldsymbol{f}_m^\mathsf{T} \boldsymbol{g} = \boldsymbol{g}^* \mathsf{F}_m \begin{bmatrix} |\alpha|^2 & \overline{\alpha}\beta \\ \overline{\beta}\alpha & |\beta|^2 \end{bmatrix} \mathsf{F}_m^\mathsf{T} \boldsymbol{g}, \tag{3.10}$$

where we used $\boldsymbol{g} = \boldsymbol{g}_0 + \boldsymbol{p}$. Note that since $\overline{\alpha}\beta = |\alpha||\beta|e^{\imath\phi}$, the inner 2×2 Hermitian matrix is uniquely determined by the magnitudes of α and β and their phase difference ϕ. By using the single-source reference illumination $\boldsymbol{e}_i$ we additionally measure

$$|\widehat{u}_i^0(\vec{\boldsymbol{x}}_r, \omega)|^2 = \boldsymbol{g}^* \boldsymbol{e}_i \boldsymbol{e}_i^{\mathsf{T}} \boldsymbol{g}, \qquad \text{for } i = 1, \ldots, N. \tag{3.11}$$

The data we exploit to recover $\boldsymbol{p}$ is obtained by subtracting the appropriate reference illuminations (3.11) from (3.10), that is

$$\begin{aligned} d_m(\vec{\boldsymbol{x}}_r, \omega) &= |\widehat{u}_m|^2 - |\alpha|^2 |\widehat{u}_{i(m)}^0|^2 - |\beta|^2 |\widehat{u}_{j(m)}^0|^2 \\ &= \boldsymbol{g}^* \mathsf{F}_m \begin{bmatrix} 0 & \overline{\alpha}\beta \\ \overline{\beta}\alpha & 0 \end{bmatrix} \mathsf{F}_m^{\mathsf{T}} \boldsymbol{g}. \end{aligned}$$

3.3.2 A linear system to recover the total field

By recalling that $\boldsymbol{g} = \boldsymbol{g}_0 + \boldsymbol{p}$, the measurements d_m are

$$d_m(\vec{\boldsymbol{x}}_r, \omega) = (\boldsymbol{g}_0 + \boldsymbol{p})^* \mathsf{F}_m \begin{bmatrix} 0 & \overline{\alpha}\beta \\ \overline{\beta}\alpha & 0 \end{bmatrix} \mathsf{F}_m^{\mathsf{T}} (\boldsymbol{g}_0 + \boldsymbol{p}).$$

To make the following expressions concise, we denote by D the Hermitian matrix

$$\mathsf{D} = \begin{bmatrix} 0 & \overline{\alpha}\beta \\ \overline{\beta}\alpha & 0 \end{bmatrix}. \tag{3.12}$$

By the smallness assumption 1, we may neglect the quadratic terms in $\boldsymbol{p}$ and collect all measurements for $m = 1, \ldots, N_p$ as a single vector $\boldsymbol{d} \in \mathbb{R}^{N_p}$:

$$\begin{aligned} \begin{bmatrix} d_1(\vec{\boldsymbol{x}}_r, \omega) \\ d_2(\vec{\boldsymbol{x}}_r, \omega) \\ \vdots \\ d_{N_p}(\vec{\boldsymbol{x}}_r, \omega) \end{bmatrix} \approx \boldsymbol{d}(\vec{\boldsymbol{x}}_r, \omega) &= \operatorname{Re}\left(\begin{bmatrix} \boldsymbol{g}_0^* \mathsf{F}_1 \mathsf{D} \mathsf{F}_1^{\mathsf{T}} \\ \boldsymbol{g}_0^* \mathsf{F}_2 \mathsf{D} \mathsf{F}_2^{\mathsf{T}} \\ \vdots \\ \boldsymbol{g}_0^* \mathsf{F}_{N_p} \mathsf{D} \mathsf{F}_{N_p}^{\mathsf{T}} \end{bmatrix} (\boldsymbol{g}_0 + 2\boldsymbol{p}) \right) \\ &= \mathsf{M}(\vec{\boldsymbol{x}}_r, \omega) \begin{bmatrix} \operatorname{Re}(\boldsymbol{g}_0 + 2\boldsymbol{p}) \\ \operatorname{Im}(\boldsymbol{g}_0 + 2\boldsymbol{p}) \end{bmatrix}, \end{aligned} \tag{3.13}$$

where the $N_p \times 2N$ real matrix M is given by

$$\mathsf{M}(\vec{\boldsymbol{x}}_r, \omega) = \begin{bmatrix} \operatorname{Re}(\boldsymbol{g}_0^* \mathsf{F}_1 \mathsf{D} \mathsf{F}_1^{\mathsf{T}}) & -\operatorname{Im}(\boldsymbol{g}_0^* \mathsf{F}_1 \mathsf{D} \mathsf{F}_1^{\mathsf{T}}) \\ \operatorname{Re}(\boldsymbol{g}_0^* \mathsf{F}_2 \mathsf{D} \mathsf{F}_2^{\mathsf{T}}) & -\operatorname{Im}(\boldsymbol{g}_0^* \mathsf{F}_2 \mathsf{D} \mathsf{F}_2^{\mathsf{T}}) \\ \vdots & \vdots \\ \operatorname{Re}(\boldsymbol{g}_0^* \mathsf{F}_{N_p} \mathsf{D} \mathsf{F}_{N_p}^{\mathsf{T}}) & -\operatorname{Im}(\boldsymbol{g}_0^* \mathsf{F}_{N_p} \mathsf{D} \mathsf{F}_{N_p}^{\mathsf{T}}) \end{bmatrix}. \tag{3.14}$$

Note that by construction, the matrix M has at most 4 nonzero elements per row, and is thus a very sparse matrix for N large.

3.4 Analysis of the linear system for the total field

We now address the question of whether there is enough information in the measurements $\boldsymbol{d} \in \mathbb{R}^{N_p}$ to recover the array response vector $\boldsymbol{p} \in \mathbb{C}^N$. The main result of this section is Theorem 4, where we show that with appropriately chosen pairs of sources, $\mathsf{M}^\dagger \boldsymbol{d}$ (i.e., the Moore-Penrose pseudoinverse of M times $\boldsymbol{d}$) gives $\boldsymbol{p}$ up to a complex scalar multiple of the vector $\boldsymbol{g}_0$, which is known a priori.

Let us first consider the case where we take measurements using all possible source pairs, i.e., that $N_p = N(N-1)/2$. Clearly, we need $N \geq 5$ to guarantee that $N_p \geq 2N$, i.e., that the matrix M has more rows than columns and the system $\boldsymbol{d} = \mathsf{M}[\mathrm{Re}(\boldsymbol{g}_0 + 2\boldsymbol{p})^\mathsf{T}, \mathrm{Im}(\boldsymbol{g}_0 + 2\boldsymbol{p})^\mathsf{T}]^\mathsf{T}$ is overdetermined.

Instead of using all possible source pairs, we use the following strategy which for $N \geq 5$, guarantees $N_p = 2N$.

Strategy to choose source pairs:

1. All 10 distinct source pairs between the source positions $\{1, \ldots, 5\}$.
2. For source position $s > 5$, choose any two different source pairs of the form (s, i) and (s, j) where $i, j \in \{1, \ldots, 5\}$.

This strategy is illustrated in Figure 3.3. More source pairs can be added without affecting the recoverability of $\boldsymbol{p}$ (Theorem 4). We now make the following assumption on the first 5 source positions.

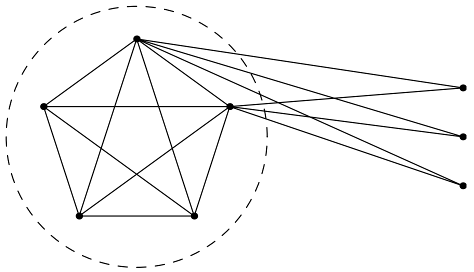


Figure 3.3. An example illustrating the strategy to choose the source pairs for $N = 8$ source positions. Each source position is represented by a node in the graph, and source pairs are represented by edges. The first 5 source positions are in the circle.

Assumption 2 *We assume the receiver is located at a position* $\vec{\boldsymbol{x}}_r$ *such that for* $i, j = 1, \ldots, 5$, *the vector* $\boldsymbol{g}_0 \equiv \boldsymbol{g}_0(\vec{\boldsymbol{x}}_r, \omega)$ *satisfies*

$$Re\big(\boldsymbol{g}_0\big)_i \neq 0,\ \ Im\big(\boldsymbol{g}_0\big)_i \neq 0,\ \ \text{and } Re\big(\boldsymbol{g}_0\big)_i Im\big(\boldsymbol{g}_0\big)_j \neq Re\big(\boldsymbol{g}_0\big)_j Im\big(\boldsymbol{g}_0\big)_i. \tag{3.15}$$

Additionally for one pair $i, j \in [1, \ldots, 5]$ *we assume*

$$\begin{aligned} \cos(\phi) \left(Re\big(\boldsymbol{g}_0\big)_i Im\big(\boldsymbol{g}_0\big)_j - Re\big(\boldsymbol{g}_0\big)_j Im\big(\boldsymbol{g}_0\big)_i \right) \neq & \\ - \sin(\phi) \left(Re(\boldsymbol{g}_0)_i Re\big(\boldsymbol{g}_0\big)_j + Im\big(\boldsymbol{g}_0\big)_i Im\big(\boldsymbol{g}_0\big)_j \right). & \end{aligned} \tag{3.16}$$

This assumption is by no means necessary for the end result (Theorem 4) to hold, but it is sufficient. If $d = 3$, condition (3.15) is equivalent to the geometric condition

$$|\vec{\boldsymbol{x}}_i - \vec{\boldsymbol{x}}_r| \notin \frac{\lambda}{4}\mathbb{Z} \text{ and } |\vec{\boldsymbol{x}}_i - \vec{\boldsymbol{x}}_r| - |\vec{\boldsymbol{x}}_j - \vec{\boldsymbol{x}}_r| \notin \frac{\lambda}{2}\mathbb{Z} \text{ for all } i, j = 1, \ldots, 5, \tag{3.17}$$

while condition (3.16) implies for one pair $i, j \in [1, \ldots, 5]$ that

$$|\vec{\boldsymbol{x}}_r - \vec{\boldsymbol{x}}_i| - |\vec{\boldsymbol{x}}_r - \vec{\boldsymbol{x}}_j| \notin \frac{\lambda}{2}\mathbb{Z} - \frac{\lambda}{2\pi}\phi. \tag{3.18}$$

Here the set $(\lambda/2)\mathbb{Z}$ is the set of all integer multiples of $\lambda/2$, where $\lambda = 2\pi c_0/\omega$ is the wavelength. In $d = 2$, conditions similar to (3.17) and (3.18) are sufficient for large enough frequencies ω because of the Hankel function asymptotic

$$H_0^{(1)}(t) = \sqrt{\frac{2}{\pi t}} \exp[\imath(t - (\pi/4))](1 + \mathcal{O}(1/t)),\ \text{ as } t \to \infty.$$

Lemma 3 *Provided* $\alpha \neq 0$, $\beta \neq 0$, $Re(\overline{\alpha}\beta) \neq 0$, *the source pairs are chosen with the above strategy and assumption 2 holds, the matrix* $\mathsf{M} \equiv \mathsf{M}(\vec{\boldsymbol{x}}_r, \omega)$ *satisfies*

$$\operatorname{null} \mathsf{M} = \operatorname{span} \left\{ \begin{bmatrix} -Im\big(\boldsymbol{g}_0(\vec{\boldsymbol{x}}_r, \omega)\big) \\ Re\big(\boldsymbol{g}_0(\vec{\boldsymbol{x}}_r, \omega)\big) \end{bmatrix} \right\}. \tag{3.19}$$

Proof. For clarity of exposition, we adopt the notation

$$a_i = \operatorname{Re}\big(\boldsymbol{g}_0\big)_i,\ \ b_i = \operatorname{Im}\big(\boldsymbol{g}_0\big)_i,$$

for $i = 1, \ldots, N$ and with $\boldsymbol{g}_0 \equiv \boldsymbol{g}_0(\vec{\boldsymbol{x}}_r, \omega)$. The proposed vector spanning the nullspace is $[\boldsymbol{v}^\mathsf{T}, \boldsymbol{w}^\mathsf{T}]^\mathsf{T} = [-\operatorname{Im}(\boldsymbol{g}_0)^\mathsf{T}, \operatorname{Re}(\boldsymbol{g}_0)^\mathsf{T}]^\mathsf{T}$ and has components $v_i = -b_i$ and $w_i = a_i$ for $i = 1, \ldots, N$.

The proof is by induction on the number of sources N. For the purpose of the induction argument, we denote by $\mathsf{M}^{(N)}$ the measurement matrix $\mathsf{M}(\vec{\boldsymbol{x}}_r, \omega)$ corresponding to N sources, which if we use the strategy explained above, must be a $2N \times 2N$ real matrix. For the base case $N = 5$ of the induction, $\mathsf{M}^{(5)}$ can be written as

$$\mathsf{M}^{(5)} = \begin{bmatrix} A_2^- & A_1^+ & 0 & 0 & 0 & B_2^+ & B_1^- & 0 & 0 & 0 \\ A_3^- & 0 & A_1^+ & 0 & 0 & B_3^+ & 0 & B_1^- & 0 & 0 \\ A_4^- & 0 & 0 & A_1^+ & 0 & B_4^+ & 0 & 0 & B_1^- & 0 \\ A_5^- & 0 & 0 & 0 & A_1^+ & B_5^+ & 0 & 0 & 0 & B_1^- \\ 0 & A_3^- & A_2^+ & 0 & 0 & 0 & B_3^+ & B_2^- & 0 & 0 \\ 0 & A_4^- & 0 & A_2^+ & 0 & 0 & B_4^+ & 0 & B_2^- & 0 \\ 0 & A_5^- & 0 & 0 & A_2^+ & 0 & B_5^+ & 0 & 0 & B_2^- \\ 0 & 0 & A_4^- & A_3^+ & 0 & 0 & 0 & B_4^+ & B_3^- & 0 \\ 0 & 0 & A_5^- & 0 & A_3^+ & 0 & 0 & B_5^+ & 0 & B_3^- \\ 0 & 0 & 0 & A_5^- & A_4^+ & 0 & 0 & 0 & B_5^+ & B_4^- \end{bmatrix},$$

where we have used

$$A_i^{\pm} = |\alpha||\beta|(\cos(\phi)a_i \pm \sin(\phi)b_i), \qquad B_i^{\pm} = |\alpha||\beta|(\cos(\phi)b_i \pm \sin(\phi)a_i). \tag{3.20}$$

Using the expressions (3.20), the leading principal 9×9 minor of $\mathsf{M}^{(5)}$ is

$$\begin{aligned} |\mathsf{M}^{(5)}_{1:9,1:9}| = -4|\alpha|^9|\beta|^9\cos^2(\phi)\left(\cos(\phi)(b_3a_1 - b_1a_3) + \sin(\phi)(b_3b_1 + a_3a_1)\right) \times \\ a_5(b_3a_2 - b_2a_3)(b_2a_1 - a_2b_1)(b_5a_4 - a_5b_4). \end{aligned}$$

Therefore if assumption 2 holds and $\cos\phi \neq 0$ (which we get from $\mathrm{Re}(\overline{\alpha}\beta) \neq 0$), we must have $\operatorname{rank}\mathsf{M}^{(5)} \geq 9$. By direct calculations, we have that

$$\operatorname{null}\mathsf{M}^{(5)} = \operatorname{span}\left\{[-b_1, \ldots, -b_5, a_1, \ldots, a_5]^{\mathsf{T}}\right\}.$$

Thus the base case $N = 5$ holds and $\operatorname{rank}\mathsf{M}^{(5)} = 9$.

For the induction hypothesis we assume that $N \geq 5$ and that

$$\operatorname{null}\mathsf{M}^{(N)} = \operatorname{span}\left\{[-\boldsymbol{b}^{\mathsf{T}}, \boldsymbol{a}^{\mathsf{T}}]^{\mathsf{T}}\right\},$$

where $\boldsymbol{a} = [a_1, \ldots, a_N]^{\mathsf{T}}$ and $\boldsymbol{b} = [b_1, \ldots, b_N]^{\mathsf{T}}$. If the first $2N$ source pairs to construct $\mathsf{M}^{(N+1)}$ are chosen in exactly the same way as the source pairs to construct $\mathsf{M}^{(N)}$, and the last two source pairs are, e.g., $(N+1,1)$ and $(N+1,2)$, we must have for any $\boldsymbol{v}, \boldsymbol{w} \in \mathbb{R}^N$ and $v_{N+1}, w_{N+1} \in \mathbb{R}$ that

$$\mathsf{M}^{(N+1)} \begin{bmatrix} \boldsymbol{v} \\ v_{N+1} \\ \boldsymbol{w} \\ w_{N+1} \end{bmatrix} = \begin{bmatrix} \mathsf{M}^{(N)} \begin{bmatrix} \boldsymbol{v} \\ \boldsymbol{w} \end{bmatrix} \\ A^-_{N+1} v_1 + A^+_1 v_{N+1} + B^+_{N+1} w_1 + B^-_1 w_{N+1} \\ A^-_{N+1} v_2 + A^+_2 v_{N+1} + B^+_{N+1} w_2 + B^-_2 w_{N+1} \end{bmatrix}. \tag{3.21}$$

Hence if $[\boldsymbol{v}^\mathsf{T}, v_{N+1}, \boldsymbol{w}^\mathsf{T}, w_{N+1}]^\mathsf{T} \in \operatorname{null} \mathsf{M}^{(N+1)}$, then we must have $[\boldsymbol{v}^\mathsf{T}, \boldsymbol{w}^\mathsf{T}]^\mathsf{T} \in \operatorname{null} \mathsf{M}^{(N)}$, i.e., there is some real $k \neq 0$ such that $\boldsymbol{v} = -k\boldsymbol{b}$ and $\boldsymbol{w} = k\boldsymbol{a}$. Equating the last two components of (3.21) to zero and using that $v_i = -kb_i$ and $w_i = ka_i$ for $i = 1, 2$, one gets the linear system

$$\begin{bmatrix} A^+_1 & B^-_1 \\ A^+_2 & B^-_2 \end{bmatrix} \begin{bmatrix} v_{N+1} \\ w_{N+1} \end{bmatrix} = \begin{bmatrix} kA^-_{N+1} b_1 - kB^+_{N+1} a_1 \\ kA^-_{N+1} b_2 - kB^+_{N+1} a_2 \end{bmatrix}.$$

Since $A^+_1 B^-_2 - A^+_2 B^-_1 = |\alpha|^2 |\beta|^2 (a_1 b_2 - a_2 b_1) \neq 0$, the unique solution to this system is $v_{N+1} = -kb_{N+1}$ and $w_{N+1} = ka_{N+1}$. Thus the desired result holds for any $N \geq 5$. ■

In Figure 3.4, we show the condition number of $\mathsf{M}(\vec{\boldsymbol{x}}_r, \omega)$ (i.e., σ_1/σ_{2N-1} the ratio of the largest singular value to the smallest nonzero singular value) over a frequency band. The experimental setup is that given in §4.5 and corresponds to sending exactly the same signal from both locations in a source pair (i.e., $\alpha = \beta$ and $\phi = 0$). Figure 3.4a shows the condition number of M with $\vec{\boldsymbol{x}}_r$ chosen so that assumption 2 is satisfied, while Figure 3.4b shows the condition number of M with $\vec{\boldsymbol{x}}_r$ chosen so that assumption 2 is violated for some frequencies. In both cases, we see improved conditioning by using more than $2N$ source pair experiments.

We now tie $\mathsf{M}^\dagger \boldsymbol{d}$ to the array response vector $\boldsymbol{p}$.

Theorem 4 *Under the assumptions of lemma 3 it is possible to recover* $\boldsymbol{p} \equiv \boldsymbol{p}(\vec{\boldsymbol{x}}_r, \omega)$ *from the intensity data* $\boldsymbol{d}$ *up to a complex scalar multiple of* $\boldsymbol{g}_0 \equiv \boldsymbol{g}_0(\vec{\boldsymbol{x}}_r, \omega)$*, more precisely,* $\mathsf{M}^\dagger \boldsymbol{d}$ *determines the vector* $\boldsymbol{p} + \zeta \boldsymbol{g}_0$ *where*

$$\zeta \equiv \zeta(\vec{\boldsymbol{x}}_r, \omega) = \frac{1}{2} - \imath \frac{Im(\boldsymbol{g}_0^* \boldsymbol{p})}{\boldsymbol{g}_0^* \boldsymbol{g}_0}. \tag{3.22}$$

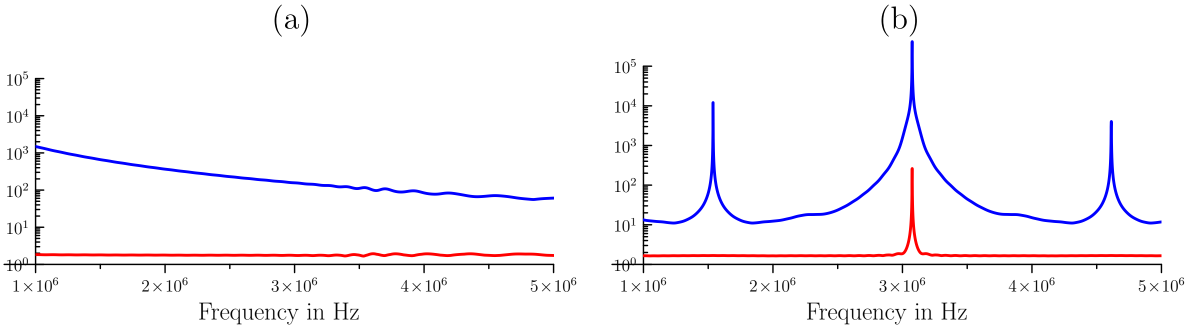


Figure 3.4. Condition number of $\mathsf{M}(\vec{\boldsymbol{x}}_r, \omega)$ with receiver location $\vec{\boldsymbol{x}}_r$ chosen so that (a) assumption 2 is satisfied, (b) assumption 2 is violated for some frequencies. The number of source pair experiments used is $N_p = N(N-1)/2$ (in red) and $N_p = 2N$ (in blue).

Proof. Recalling the form of our data we have

$$\boldsymbol{d} = \mathsf{M} \begin{bmatrix} \mathrm{Re}\big(\boldsymbol{g}_0 + 2\boldsymbol{p}\big) \\ \mathrm{Im}\big(\boldsymbol{g}_0 + 2\boldsymbol{p}\big) \end{bmatrix}.$$

By lemma 3, the matrix M has a one dimensional nullspace therefore

$$\mathsf{M}^\dagger \boldsymbol{d} = \begin{bmatrix} \mathrm{Re}(\boldsymbol{g}_0 + 2\boldsymbol{p}) \\ \mathrm{Im}(\boldsymbol{g}_0 + 2\boldsymbol{p}) \end{bmatrix} - \widetilde{\zeta} \begin{bmatrix} -\mathrm{Im}(\boldsymbol{g}_0) \\ \mathrm{Re}(\boldsymbol{g}_0) \end{bmatrix},$$

where $\widetilde{\zeta} \in \mathbb{R}$ is found by enforcing orthogonality with $[-\mathrm{Im}(\boldsymbol{g}_0)^\mathsf{T}, \mathrm{Re}(\boldsymbol{g}_0)^\mathsf{T}]^\mathsf{T}$, i.e.,

$$\widetilde{\zeta} = \frac{1}{\boldsymbol{g}_0^* \boldsymbol{g}_0} [-\mathrm{Re}(\boldsymbol{g}_0 + 2\boldsymbol{p})^\mathsf{T} \mathrm{Im}(\boldsymbol{g}_0) + \mathrm{Im}(\boldsymbol{g}_0 + 2\boldsymbol{p})^\mathsf{T} \mathrm{Re}(\boldsymbol{g}_0)] = \frac{2\mathrm{Im}(\boldsymbol{g}_0^* \boldsymbol{p})}{\boldsymbol{g}_0^* \boldsymbol{g}_0}.$$

Thus from $\mathsf{M}^\dagger \boldsymbol{d}$ we can get the $\mathbb{C}^N$ vector

$$\frac{1}{2}[\mathrm{Re}(\boldsymbol{g}_0 + 2\boldsymbol{p}) + \widetilde{\zeta}\mathrm{Im}(\boldsymbol{g}_0)] + \frac{\imath}{2}[\mathrm{Im}(\boldsymbol{g}_0 + 2\boldsymbol{p}) - \widetilde{\zeta}\mathrm{Re}(\boldsymbol{g}_0)] = \frac{1}{2}\boldsymbol{g}_0 + \boldsymbol{p} - \frac{\imath}{2}\widetilde{\zeta}\boldsymbol{g}_0 = \boldsymbol{p} + \zeta \boldsymbol{g}_0,$$

where the scalar $\zeta \equiv \zeta(\vec{\boldsymbol{x}}_r, \omega) \in \mathbb{C}$ is given by (3.22). ■

3.5 Kirchhoff migration imaging

We now show that we can image with the reconstructed field $\boldsymbol{p} + \zeta \boldsymbol{g}_0$ instead of $\boldsymbol{p}$ by using Kirchhoff migration. This is because the Kirchhoff migration image of $\zeta \boldsymbol{g}_0$ is negligible compared to the image of $\boldsymbol{p}$ for high frequencies. In order to show that this nullspace vector does not affect the imaging, we need to make sure the receiver satisfies the following condition.

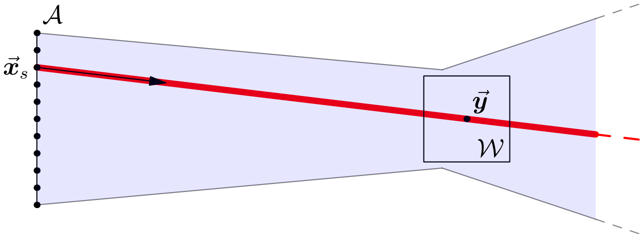


Figure 3.5. Given an array $\mathcal{A}$ and a region $\mathcal{W}$ containing the scatterers to image, assumption 3 ensures the receiver location $\vec{\boldsymbol{x}}_r$ is outside of the blue shaded region. This guarantees the Kirchhoff images using data $\boldsymbol{p}$ and the recovered $\boldsymbol{p} + \zeta \boldsymbol{g}_0$ are essentially the same. A ray indicating the direction $\vec{\boldsymbol{x}}_s - \vec{\boldsymbol{y}}$ for particular $\vec{\boldsymbol{x}}_s \in \mathcal{A}$ and $\vec{\boldsymbol{y}} \in \mathcal{W}$ is shown in red. If $\vec{\boldsymbol{x}}_r$ is outside the blue shaded region, we have $(\vec{\boldsymbol{x}}_s - \vec{\boldsymbol{x}}_r)/|\vec{\boldsymbol{x}}_s - \vec{\boldsymbol{x}}_r| \neq (\vec{\boldsymbol{x}}_s - \vec{\boldsymbol{y}})/|\vec{\boldsymbol{x}}_s - \vec{\boldsymbol{y}}|$ for all $\vec{\boldsymbol{x}}_s \in \mathcal{A}$ and all $\vec{\boldsymbol{y}} \in \mathcal{W}$.

Assumption 3 (Geometric imaging conditions) *For a scattering potential with support $\mathcal{P}$ contained inside an image window $\mathcal{W}$, we assume $\vec{\boldsymbol{x}}_r$ satisfies*

$$\frac{\vec{\boldsymbol{x}}_s - \vec{\boldsymbol{x}}_r}{|\vec{\boldsymbol{x}}_s - \vec{\boldsymbol{x}}_r|} \neq \frac{\vec{\boldsymbol{x}}_s - \vec{\boldsymbol{y}}}{|\vec{\boldsymbol{x}}_s - \vec{\boldsymbol{y}}|}, \tag{3.23}$$

for $s = 1, \ldots, N$ and $\vec{\boldsymbol{y}} \in \mathcal{W}$.

One way to guarantee assumption 3 holds is to place the receiver at location $\vec{\boldsymbol{x}}_r$ outside of the shaded region in Figure 3.5.

Theorem 5 *Provided assumption 3 holds and for large frequencies ω, the image of the reconstructed array response vector is*

$$\Gamma_{KM}[\boldsymbol{p} + \zeta \boldsymbol{g}_0, \omega](\vec{\boldsymbol{y}}) \approx \Gamma_{KM}[\boldsymbol{p}, \omega](\vec{\boldsymbol{y}}).$$

Proof. First we approximate the Kirchhoff imaging functional (3.7) by an integral over the array $\mathcal{A}$, i.e.,

$$\begin{aligned}\Gamma_{\mathrm{KM}}[\zeta \boldsymbol{g}_0, \omega](\vec{\boldsymbol{y}}) &= \overline{\widehat{G}}(\vec{\boldsymbol{x}}_r, \vec{\boldsymbol{y}}, \omega) \boldsymbol{g}_0(\vec{\boldsymbol{y}}, \omega)^* \zeta(\vec{\boldsymbol{x}}_r, \omega) \boldsymbol{g}_0(\vec{\boldsymbol{x}}_r, \omega) \\ &\sim \zeta(\vec{\boldsymbol{x}}_r, \omega) \int_{\mathcal{A}} d\boldsymbol{x}_s C(\boldsymbol{x}_s) \exp\left(\imath \omega c_0^{-1} \left(|\vec{\boldsymbol{x}}_s - \vec{\boldsymbol{x}}_r| - |\vec{\boldsymbol{x}}_s - \vec{\boldsymbol{y}}| - |\vec{\boldsymbol{y}} - \vec{\boldsymbol{x}}_r|\right)\right),\end{aligned} \tag{3.24}$$

where the symbol $\sim$ means equal up to a constant and $C(\boldsymbol{x}_s)$ collects smooth geometric spreading terms.

Let us first use the stationary phase method (see, e.g., [2]) on the integral over $\mathcal{A}$. In the high frequency limit $\omega \to \infty$, the dominant contribution comes from stationary points of the phase, i.e., the points $\vec{\boldsymbol{x}}_s$ for which

$$\nabla_{\vec{\boldsymbol{x}}_s} \Big(|\vec{\boldsymbol{x}}_s - \vec{\boldsymbol{x}}_r| - |\vec{\boldsymbol{x}}_s - \vec{\boldsymbol{y}}| - |\vec{\boldsymbol{y}} - \vec{\boldsymbol{x}}_r| \Big) = 0.$$

The stationary points must then satisfy

$$\frac{\vec{\boldsymbol{x}}_s - \vec{\boldsymbol{x}}_r}{|\vec{\boldsymbol{x}}_s - \vec{\boldsymbol{x}}_r|} = \frac{\vec{\boldsymbol{x}}_s - \vec{\boldsymbol{y}}}{|\vec{\boldsymbol{x}}_s - \vec{\boldsymbol{y}}|}.$$

Thus by assumption 3, there are no stationary points in the phase of the integral over the array $\mathcal{A}$ appearing in (3.24). Neglecting boundary effects, this integral goes to zero faster than any polynomial in ω (see, e.g., [3]).

We now show that in the high frequency limit $\omega \to \infty$, we have $\zeta(\vec{\boldsymbol{x}}_r, \omega) \to 1/2$. Recalling (3.22), we have

$$\begin{aligned} \zeta(\vec{\boldsymbol{x}}_r, \omega) &= \frac{1}{2} + \frac{\boldsymbol{g}_0(\vec{\boldsymbol{x}}_r, \omega)^* \boldsymbol{p}(\vec{\boldsymbol{x}}_r, \omega) - \boldsymbol{p}(\vec{\boldsymbol{x}}_r, \omega)^* \boldsymbol{g}_0(\vec{\boldsymbol{x}}_r, \omega)}{\boldsymbol{g}_0(\vec{\boldsymbol{x}}_r, \omega)^* \boldsymbol{g}_0(\vec{\boldsymbol{x}}_r, \omega)} \\ &\sim \frac{1}{2} + \frac{\omega^2}{c_0^2} \int_{\mathcal{P}} d\vec{\boldsymbol{z}} \int_{\mathcal{A}} d\boldsymbol{x}_s C(\boldsymbol{x}_s) \exp\left(\imath \omega c_0^{-1} \big(|\vec{\boldsymbol{x}}_s - \vec{\boldsymbol{z}}| + |\vec{\boldsymbol{z}} - \vec{\boldsymbol{x}}_r| - |\vec{\boldsymbol{x}}_s - \vec{\boldsymbol{x}}_r| \big) \right) \\ &\quad - \frac{\omega^2}{c_0^2} \int_{\mathcal{P}} d\vec{\boldsymbol{z}} \int_{\mathcal{A}} d\boldsymbol{x}_s C(\boldsymbol{x}_s) \exp\left(\imath \omega c_0^{-1} \big(|\vec{\boldsymbol{x}}_s - \vec{\boldsymbol{x}}_r| - |\vec{\boldsymbol{x}}_s - \vec{\boldsymbol{z}}| - |\vec{\boldsymbol{z}} - \vec{\boldsymbol{x}}_r| \big) \right), \end{aligned} \tag{3.25}$$

where $C(\vec{\boldsymbol{x}}_s)$ collects geometric spreading terms and $|\boldsymbol{g}_0(\vec{\boldsymbol{x}}_r, \omega)|^{-2}$, which is actually independent of the frequency ω. By assumption 3, the integrals over $\mathcal{A}$ in (3.25) do not have any stationary points. Thus if we neglect boundary terms, these integrals must go to zero faster than any polynomial in ω (see, e.g., [3]), meaning that $\zeta(\vec{\boldsymbol{x}}_r, \omega) \to 1/2$ as $\omega \to \infty$. Thus $\Gamma_{\mathrm{KM}}[\zeta \boldsymbol{g}_0, \omega](\vec{\boldsymbol{y}}) \to 0$ as $\omega \to \infty$. ■

Remark 9 *Theorem 5 is a high frequency asymptotic result. At a given frequency, the image $\Gamma_{KM}[\boldsymbol{g}_0, \omega](\vec{\boldsymbol{y}})$ may not be negligible with respect to $\Gamma_{KM}[\boldsymbol{p}, \omega](\vec{\boldsymbol{y}})$, especially since we assumed the scattered field $\boldsymbol{p}$ is much smaller than $\boldsymbol{g}_0$ (assumption 1). This problem did not appear in the numerical experiments reported in §4.5. However it does appear if one makes $\boldsymbol{p}$ smaller by, e.g., choosing a smaller reflectivity. One can remedy this by recalling the expression for ζ in (3.22) and imaging with data $\boldsymbol{p} + (\zeta - 1/2)\boldsymbol{g}_0$ instead of $\boldsymbol{p} + \zeta \boldsymbol{g}_0$. This makes the error in the data comparable to $\boldsymbol{p}$*

(since $\|(\zeta - 1/2)\boldsymbol{g}_0\| \leq \|\boldsymbol{p}\|$*) instead of being comparable to* $\|\boldsymbol{g}_0\|$*, while also vanishing as the frequency increases.*

3.6 Autocorrelation measurements

Up to this point we have assumed deterministic control over the source illuminations. In this section we relax this control by driving the array with stochastic signals. We start in section 3.6.1 by recalling an ergodicity result of Garnier and Papanicolaou [12] which guarantees that if Gaussian stochastic processes are used to drive the sources, the realization average of the total field can be well approximated by time averages of the total field. Then in section 3.6.2 we adapt the source pair illumination strategy to pairs of sources driven by two correlated Gaussian processes, with (known) correlation identical for different pairs. From these pairwise illuminations we measure empirical autocorrelations to obtain intensity measurements that are essentially (up to ergodic averaging) the same as those using the deterministic strategy of section 3.3.1.

3.6.1 Stochastic array illuminations

We consider array illuminations $\boldsymbol{f}(t) \in \mathbb{C}^N$ given by a stationary Gaussian process with mean zero and with correlation the $N \times N$ matrix function

$$\mathsf{R}(\tau) = \langle \overline{\boldsymbol{f}}(t) \boldsymbol{f}^{\mathsf{T}}(t+\tau) \rangle. \tag{3.26}$$

Here $\langle \cdot \rangle$ denotes the expectation with respect to realizations of $\boldsymbol{f}$, and in an abuse of notation we have denoted by $\boldsymbol{f}(t)$ the time domain vector of signals driving the array. Since $R_{s,s'}(\tau) = \langle \overline{f}_s(t) f_{s'}(t+\tau) \rangle = \overline{\langle \overline{f}_{s'}(t+\tau) f_s(t) \rangle} = \overline{R_{s',s}(-\tau)}$ for $s, s' = 1, \ldots, N$, we have $\mathsf{R}(\tau) = \mathsf{R}^*(-\tau)$ and so $\widehat{\mathsf{R}}(\omega)$ is a Hermitian $N \times N$ matrix.

The total field u at the receiver arising from the array illumination $\boldsymbol{f}$ is, in the time domain,

$$u(\vec{\boldsymbol{x}}_r, t) = \sum_{s=1}^{N} \int dt' G(\vec{\boldsymbol{x}}_r, \vec{\boldsymbol{x}}_s, t - t') f_s(t'), \tag{3.27}$$

where G is the Born approximation of the inhomogeneous Green function, i.e.,

$$G(\vec{\boldsymbol{x}}_r, \vec{\boldsymbol{x}}_s, t) = \frac{1}{2\pi} \int d\omega e^{-\imath\omega t} \Big[\widehat{G}_0(\vec{\boldsymbol{x}}_r, \vec{\boldsymbol{x}}_s, \omega) + k^2 \int_{\mathcal{P}} d\vec{\boldsymbol{z}} \rho(\vec{\boldsymbol{z}}) \widehat{G}_0(\vec{\boldsymbol{x}}_r, \vec{\boldsymbol{z}}, \omega) \widehat{G}_0(\vec{\boldsymbol{z}}, \vec{\boldsymbol{x}}_s, \omega) \Big].$$

The empirical autocorrelation of u is

$$\psi(\vec{\boldsymbol{x}}_r,\tau)=\frac{1}{2T}\int_{-T}^{T}\overline{u}(\vec{\boldsymbol{x}}_r,t)u(\vec{\boldsymbol{x}}_r,t+\tau)dt, \tag{3.28}$$

where T is a known measurement time. Following Garnier and Papanicolaou [12], we formulate proposition 1 regarding the statistical stability and ergodicity of (4.18). This proposition is essentially the same as [12, Proposition 4.1], but we make small modifications to allow for complex fields and more general correlations in space. We include it here for the sake of completeness and the proof can be found in §3.10.

Proposition 1 *Assume* $\boldsymbol{f}$ *satisfies* (3.26)*. The expectation (w.r.t. realizations of* $\boldsymbol{f}$*) of the empirical autocorrelation* (4.18) *is independent of measurement time* T*:*

$$\langle\psi(\vec{\boldsymbol{x}}_r,\tau)\rangle=\Psi(\vec{\boldsymbol{x}}_r,\tau), \tag{3.29}$$

where

$$\begin{aligned}\Psi(\vec{\boldsymbol{x}}_r,\tau)&=\sum_{s,s'=1}^{N}\int dt'\int dt''\overline{G}(\vec{\boldsymbol{x}}_r,\vec{\boldsymbol{x}}_r,-t')G(\vec{\boldsymbol{x}}_r,\vec{\boldsymbol{x}}_{s'},\tau-t'')R_{s,s'}(t''-t')\\&=\frac{1}{2\pi}\int d\omega e^{-\imath\omega\tau}\boldsymbol{g}(\vec{\boldsymbol{x}}_r,\omega)^*\widehat{\mathsf{R}}(\omega)\boldsymbol{g}(\vec{\boldsymbol{x}}_r,\omega).\end{aligned} \tag{3.30}$$

Furthermore, (4.18) *is ergodic, i.e.,*

$$\psi(\vec{\boldsymbol{x}}_r,\tau)\xrightarrow{T\to\infty}\Psi(\vec{\boldsymbol{x}}_r,\tau). \tag{3.31}$$

3.6.2 Pairwise stochastic illuminations

We make N_p illuminations each corresponding to using only two distinct sources $(i(m),j(m))\in\{1,\ldots,N\}^2$, $m=1,\ldots,N_p$. The correlation matrix for the $m-$th experiment has the form

$$\widehat{\mathsf{R}}_m(\omega)=\mathsf{F}_m\mathsf{C}(\omega)\mathsf{F}_m^{\mathsf{T}}, \tag{3.32}$$

where $\mathsf{F}_m=[\boldsymbol{e}_{i(m)},\boldsymbol{e}_{j(m)}]\in\mathbb{R}^{N\times 2}$ and $\mathsf{C}(\omega)$ is a known 2×2 Hermitian positive semidefinite matrix that represents the correlation between the two sources and is assumed to be the same for all experiments. For instance, if we send the same signal with power spectrum $F(\omega)$ from both sources in a pair, this correlation matrix is

$$\mathsf{C}(\omega) = F(\omega) \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}.$$

By the ergodicity (3.31) of proposition 1, when we measure the empirical autocorrelation ψ_m of u_m at the receiver $\vec{\boldsymbol{x}}_r$ for long enough time T, the empirical autocorrelation is close to an intensity measurement, i.e.,

$$\widehat{\Psi}_m(\vec{\boldsymbol{x}}_r, \omega) = \boldsymbol{g}(\vec{\boldsymbol{x}}_r, \omega)^* \mathsf{F}_m \mathsf{C}(\omega) \mathsf{F}_m^{\mathsf{T}} \boldsymbol{g}(\vec{\boldsymbol{x}}_r, \omega). \tag{3.33}$$

By using appropriate single-source illuminations driven by a signal with known correlation, it is possible to measure

$$\widehat{\Psi}_i^0(\vec{\boldsymbol{x}}_r, \omega) = \boldsymbol{g}^*(\vec{\boldsymbol{x}}_r, \omega) \boldsymbol{e}_i \boldsymbol{e}_i^{\mathsf{T}} \boldsymbol{g}(\vec{\boldsymbol{x}}_r, \omega). \text{ for } i = 1, \ldots, N. \tag{3.34}$$

From (3.33) and (3.34) we obtain the $m-$th measurement

$$\begin{aligned} d_m(\vec{\boldsymbol{x}}_r, \omega) &= \widehat{\Psi}_m(\vec{\boldsymbol{x}}_r, \omega) - C_{11}(\omega) \widehat{\Psi}^0_{i(m)}(\vec{\boldsymbol{x}}_r, \omega) - C_{22}(\omega) \widehat{\Psi}^0_{j(m)}(\vec{\boldsymbol{x}}_r, \omega) \\ &= \boldsymbol{g}(\vec{\boldsymbol{x}}_r, \omega)^* \mathsf{F}_m \mathsf{D}(\omega) \mathsf{F}_m^{\mathsf{T}} \boldsymbol{g}(\vec{\boldsymbol{x}}_r, \omega), \end{aligned} \tag{3.35}$$

where the matrix D is 2×2, Hermitian with zero diagonal, i.e., precisely of the same form as the matrix D we encountered in the intensity measurements case (3.12).

Proceeding analogously as in section 3.3.1 and recalling that $\boldsymbol{g} = \boldsymbol{g}_0 + \boldsymbol{p}$ we have

$$d_m(\vec{\boldsymbol{x}}_r, \omega) = \big(\boldsymbol{g}_0 + \boldsymbol{p}\big)^* \mathsf{F}_m \mathsf{D}(\omega) \mathsf{F}_m^{\mathsf{T}} \big(\boldsymbol{g}_0 + \boldsymbol{p}\big).$$

Collecting the measurements for $m = 1, \ldots, N_p$ and neglecting the quadratic term in $\boldsymbol{p}$ we have the approximate data

$$\begin{bmatrix} d_1(\vec{\boldsymbol{x}}_r, \omega) \\ d_2(\vec{\boldsymbol{x}}_r, \omega) \\ \vdots \\ d_{N_p}(\vec{\boldsymbol{x}}_r, \omega) \end{bmatrix} \approx \boldsymbol{d}(\vec{\boldsymbol{x}}_r, \omega) = \mathsf{M}(\vec{\boldsymbol{x}}_r, \omega) \begin{bmatrix} \mathrm{Re}\big(\boldsymbol{g}_0 + 2\boldsymbol{p}\big) \\ \mathrm{Im}\big(\boldsymbol{g}_0 + 2\boldsymbol{p}\big) \end{bmatrix}, \tag{3.36}$$

where the matrix $\mathsf{M} \in \mathbb{R}^{N_p \times 2N}$ is again given by (3.14). Thus, the data (3.36) obtained by measuring the empirical autocorrelation (4.18) and using correlated pair illuminations, are essentially the same as the data obtained using deterministic source pairs (3.13). Hence the analysis of the matrix M of §3.4 holds and we can use Kirchhoff migration as we did in §3.5 for the intensity measurements case.

Remark 10 (Uncorrelated background illumination) *The proposed illumination strategy is robust with respect to noise and even allows us to send the* same *Gaussian signal from the* $m-th$ *source pair* $(i(m), j(m))$ *and* independent *Gaussian signals from all remaining sources on the array. If the independent signals have the same spectral density* $F(\omega)$ *as the source pair signal, the correlation matrix for the* $m-th$ *experiment is*

$$\widehat{\mathsf{R}}_m(\omega) = F(\omega)\left(\mathsf{I} + \mathsf{F}_m \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \mathsf{F}_m^{\mathsf{T}}\right), \tag{3.37}$$

where I *is* $N \times N$ *identity matrix. By subtracting from the autocorrelation for the* $m-th$ *experiment, the autocorrelation for a reference illumination that sends* independent *Gaussian signals with correlation matrix* $F(\omega)\mathsf{I}$*, it is possible to obtain* $m-th$ *measurement* (3.35) *with*

$$\mathsf{D}(\omega) = F(\omega) \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

3.7 Additive noise

Here we discuss the effects of additive instrumental noise in autocorrelated measurements of the total field. The total field at $\vec{\boldsymbol{x}}_r$ resulting from illuminating with the $m-$th pair and tainted with additive noise is $u_m(\vec{\boldsymbol{x}}_r, t) + \xi(t)$. We assume the noise ξ is a stationary Gaussian process with mean zero and spectral density

$$\widehat{\Xi}(\omega) = \exp\left(\frac{-l_c^2(\omega - \omega_0)^2}{4\pi}\right). \tag{3.38}$$

Here l_c represents the correlation time of the noise (i.e., $\Xi(\tau) = \langle \overline{\xi}(t)\xi(t+\tau)\rangle \approx 0$ for $\tau \gg l_c$) and ω_0 is the central angular frequency of the noise. If the noise ξ is independent of the signals used to drive the source pairs, it can be shown using the techniques of §3.10 that

$$\frac{1}{2T}\int_{-T}^{T} dt\big(\overline{u}_m(\vec{\boldsymbol{x}}_r, t) + \overline{\xi}(t)\big)\big(u_m(\vec{\boldsymbol{x}}_r, t+\tau) + \xi(t+\tau)\big) \xrightarrow{T\to\infty} \Psi_m(\vec{\boldsymbol{x}}_r, \tau) + \Xi(\tau),$$

where Ψ_m is given by (3.30).

Assuming the same form of instrumental noise in the single-source reference measurements, the $m-$th measurement $d_m(\vec{\boldsymbol{x}}_r, \omega)$ is

$$d_m(\vec{\boldsymbol{x}}_r, \omega) = \big(\boldsymbol{g}_0 + \boldsymbol{p}\big)^* \mathsf{F}_m \mathsf{D}(\omega) \mathsf{F}_m^{\mathsf{T}} \big(\boldsymbol{g}_0 + \boldsymbol{p}\big) + C\widehat{\Xi}(\omega).$$

for some $C \in \mathbb{R}$. Neglecting the terms which are quadratic in $\boldsymbol{p}$ and going back to the time domain we have

$$\begin{aligned} d_m(\vec{\boldsymbol{x}}_r,\tau) \approx \frac{1}{2\pi}\int d\omega e^{-\imath\omega\tau}\Big[&\boldsymbol{g}_0(\vec{\boldsymbol{x}}_r,\omega)^*\mathsf{F}_m\mathsf{D}(\omega)\mathsf{F}_m^{\mathsf{T}}\boldsymbol{g}_0(\vec{\boldsymbol{x}}_r,\omega)\\ &+\boldsymbol{g}_0(\vec{\boldsymbol{x}}_r,\omega)^*\mathsf{F}_m\mathsf{D}(\omega)\mathsf{F}_m^{\mathsf{T}}\boldsymbol{p}(\vec{\boldsymbol{x}}_r,\omega)\\ &+\boldsymbol{p}(\vec{\boldsymbol{x}}_r,\omega)^*\mathsf{F}_m\mathsf{D}(\omega)\mathsf{F}_m^{\mathsf{T}}\boldsymbol{g}_0(\vec{\boldsymbol{x}}_r,\omega)\Big] + C\Xi(\tau), \end{aligned}$$

with the slight abuse of notation of using d_m for both time and frequency domain quantities. The second and third terms in the integrand are incident-scattered field correlations and contain the available information about the scattering potential $\rho(\vec{\boldsymbol{y}})$.

For simplicity, we now focus on the case where the source pair signals have correlation matrix

$$\mathsf{D}(\omega) = F(\omega)\begin{bmatrix} 0 & e^{\imath\omega\phi}\\ e^{-\imath\omega\phi} & 0\end{bmatrix}.$$

Such correlation corresponds to sending a signal from one of the sources in a pair and a copy of the same signal delayed by ϕ from the other source. For a point scatterer at $\vec{\boldsymbol{y}}$, the incident-scattered terms have peaks at delay times $\tau(\vec{\boldsymbol{y}})$ corresponding to differences between travel times of a reflected path and direct path, i.e., for the $m-$th experiment the peaks occur at the four possible delays

$$\tau(\vec{\boldsymbol{y}}) = \begin{cases} \pm((|\vec{\boldsymbol{x}}_{j(m)} - \vec{\boldsymbol{y}}| + |\vec{\boldsymbol{y}} - \vec{\boldsymbol{x}}_r| - |\vec{\boldsymbol{x}}_{i(m)} - \vec{\boldsymbol{x}}_r|)/c_0 + \phi),\\ \pm((|\vec{\boldsymbol{x}}_{i(m)} - \vec{\boldsymbol{y}}| + |\vec{\boldsymbol{y}} - \vec{\boldsymbol{x}}_r| - |\vec{\boldsymbol{x}}_{j(m)} - \vec{\boldsymbol{x}}_r|)/c_0 - \phi). \end{cases}$$

Consider then the minimal delay time $\tau_{\min}(\vec{\boldsymbol{y}})$ given by

$$\tau_{\min}(\vec{\boldsymbol{y}}) = \min_{\vec{\boldsymbol{x}}_s,\vec{\boldsymbol{x}}_{s'}\in\mathcal{A}}\left|\frac{|\vec{\boldsymbol{x}}_s - \vec{\boldsymbol{y}}| + |\vec{\boldsymbol{y}} - \vec{\boldsymbol{x}}_r| - |\vec{\boldsymbol{x}}_{s'} - \vec{\boldsymbol{x}}_r|}{c_0} \pm \phi\right|, \tag{3.39}$$

that is the minimal delay time we expect the incident-scattered correlations to peak. If we assume the additive noise decorrelates much faster than the first incident-scattered arrival from $\vec{\boldsymbol{y}}$ (i.e., $l_c \ll \tau_{\min}(\vec{\boldsymbol{y}})$), then the information of the scatterer $\rho(\vec{\boldsymbol{y}})$ contained in $d_m(\vec{\boldsymbol{x}}_r,\tau)$ is essentially unchanged (up to ergodic averaging). Hence we can stably image using the proposed method at $\vec{\boldsymbol{y}}$ provided $\tau_{\min}(\vec{\boldsymbol{y}}) \gg l_c$.

3.8 Numerical experiments

Here we include 2D numerical experiments of our proposed imaging routine for scalings corresponding to acoustics (§3.8.1) and optics (§3.8.2). We demonstrate the stochastic source pair illumination strategy for the acoustic regime, i.e., we compute the autocorrelations for time domain data. In the optic regime this is an expensive calculation, so we use instead power spectra (i.e., deterministic illuminations). We also provide experiments for physical setups where our method is not expected to work (§3.8.3).

3.8.1 Acoustic regime

For imaging in an acoustic regime, our choice of physical parameters corresponds to ultrasound in water. We choose the background wave velocity to be $c_0 = 1500$ m/s. The central frequency for all signals (sources and additive noise) is 3 MHz, which gives a central wavelength of $\lambda_0 = 0.5$ mm. We center a source array $\mathcal{A}$ at the origin consisting of 41 point sources at coordinates $\vec{\boldsymbol{x}}_s = (0, -10\lambda_0 + (s-1)\lambda_0/2)$ for $s = 1, \ldots, 41$. A single receiver is located at the coordinate $\vec{\boldsymbol{x}}_r = (-20\lambda_0, -20\lambda_0)$ (see Figure 3.2).

We generate a stationary Gaussian time signal $f(t)$ with mean zero and correlation function

$$F(\tau) = \exp\left(-\pi \frac{\tau^2}{t_c^2}\right),$$

using the Wiener-Khinchin theorem. The correlation time $t_c \approx 1.25$ μs which gives the signal an effective frequency band $[1, 5]$ MHz. We generate time signals of length $2T$ for $T \approx 260$ μs with 8001 uniformly spaced samples. This sampling is enough to resolve the frequencies in the angular frequency band $\mathcal{B}$, while T is sufficient to observe ergodic averaging (see §3.6). By placing the *same* realization of this signal $\widehat{f}(\omega)$ at the locations $\vec{\boldsymbol{x}}_{i(m)}$ and $\vec{\boldsymbol{x}}_{j(m)}$ we generate the pair illumination $\boldsymbol{f}_m(\omega) = \widehat{f}(\omega)(\boldsymbol{e}_{i(m)} + \boldsymbol{e}_{j(m)})$. Similarly, by placing an independent realization of $\widehat{f}(\omega)$ at location $\vec{\boldsymbol{x}}_i$ we generate the single-source reference illumination $\boldsymbol{f}_i^0(\omega) = \widehat{f}(\omega)\boldsymbol{e}_i$.

For all experiments, synthetic data are generated in the frequency domain using the Born approximation. We assume 3D wave propagation for simplicity so that G_0 is given by (3.3) for $d = 3$. The $m-$th measurement is obtained through the formula

$$d_m(\vec{\boldsymbol{x}}_r, \omega) = \widehat{\Psi}_m(\vec{\boldsymbol{x}}_r, \omega) - \widehat{\Psi}^0_{i(m)}(\vec{\boldsymbol{x}}_r, \omega) - \widehat{\Psi}^0_{j(m)}(\vec{\boldsymbol{x}}_r, \omega),$$

where

$$\begin{aligned}\widehat{\Psi}_m(\vec{\boldsymbol{x}}_r,\omega) &= \left|\big(\boldsymbol{g}_0(\vec{\boldsymbol{x}}_r,\omega)+\boldsymbol{p}(\vec{\boldsymbol{x}}_r,\omega)\big)^{\mathsf{T}}\boldsymbol{f}_m(\omega)\right|^2,\\ \widehat{\Psi}_i^0(\vec{\boldsymbol{x}}_r,\omega) &= \left|\big(\boldsymbol{g}_0(\vec{\boldsymbol{x}}_r,\omega)+\boldsymbol{p}(\vec{\boldsymbol{x}}_r,\omega)\big)^{\mathsf{T}}\boldsymbol{f}_i^0(\omega)\right|^2,\end{aligned}$$

with $\boldsymbol{g}_0$ and $\boldsymbol{p}$ defined by (3.2) and (3.6), respectively. Thus, the quadratic term $\boldsymbol{p}^{\mathsf{T}}\mathsf{H}_m\overline{\boldsymbol{p}}$ is present in each measurement d_m.

For these simulations we use the full set of pair illuminations, which for $N = 41$ source locations, generates a measurement matrix $\mathsf{M}(\vec{\boldsymbol{x}}_r,\omega) \in \mathbb{R}^{820\times 82}$. We use the Moore-Penrose pseudoinverse $\mathsf{M}^{\dagger}$ to recover $\boldsymbol{p}+\zeta\boldsymbol{g}_0$ for each $\omega \in \mathcal{B}$. When the number of sources N and thus the dimension of M is large (recall $\mathsf{M} \in \mathbb{R}^{N(N-1)/2\times N}$), the pseudoinverse could become computationally expensive. However, the system is sparse as it contains only 4 nonzero elements per row, so linear least-squares solvers that exploit sparsity (e.g., CGLS [19]) may be more efficient than our approach. Furthermore, as discussed in §3.4 we can reduce the size of M to $2N \times 2N$ while keeping the nullspace of M one-dimensional by using an appropriate subset of source pairs.

We form an image at $\vec{\boldsymbol{y}} \in \mathcal{W} = \{(100\lambda_0+i\lambda_0/2.5, j\lambda_0/2.5),\ \text{for } i,j=-25,\ldots,25\}$ using the Kirchhoff migration functional (§3.2.2), summed over the bandwidth band $\mathcal{B}$,

$$\Gamma_{\mathrm{KM}}[\boldsymbol{p}+\zeta\boldsymbol{g}_0](\vec{\boldsymbol{y}}) = \int_{\mathcal{B}} d\omega\Gamma_{\mathrm{KM}}[\boldsymbol{p}+\zeta\boldsymbol{g}_0,\omega](\vec{\boldsymbol{y}}).$$

For our first experiment, we place a single point reflector at the location $\vec{\boldsymbol{y}} = (100\lambda_0, 0)$, with reflectivity $\rho(\vec{\boldsymbol{y}}) = 1\times 10^{-8}$ (roughly equivalent to a reflector of area λ_0^2 with reflectivity 0.04). The migrated image (Figure 3.6a) indeed exhibits the cross-range (Rayleigh) resolution estimate $\lambda_0 L/a \approx 5\lambda_0$ and range resolution estimate $c_0/|\mathcal{B}| \approx 1\lambda_0$. Note that there is a trade-off in the choice of the reflectivity. We need ρ to be sufficiently small so that assumption 1 is valid and quadratic terms of $\boldsymbol{p}$ in (3.13) are negligible. However the smaller ρ is, the longer the acquisition time T has to be in order to better observe the reflected-incident correlations in the data.

In our second experiment (Figure 3.6b), we consider two oblique reflectors located at $\vec{\boldsymbol{y}}_1 = (99\lambda_0, -2\lambda_0)$ and $\vec{\boldsymbol{y}}_2 = (103\lambda_0, 4\lambda_0)$ each with $\rho(\vec{\boldsymbol{y}}_i) = 1\times 10^{-8}$. We include a reconstruction of an extended scatterer (line segment) in Figure 3.7. Here the line

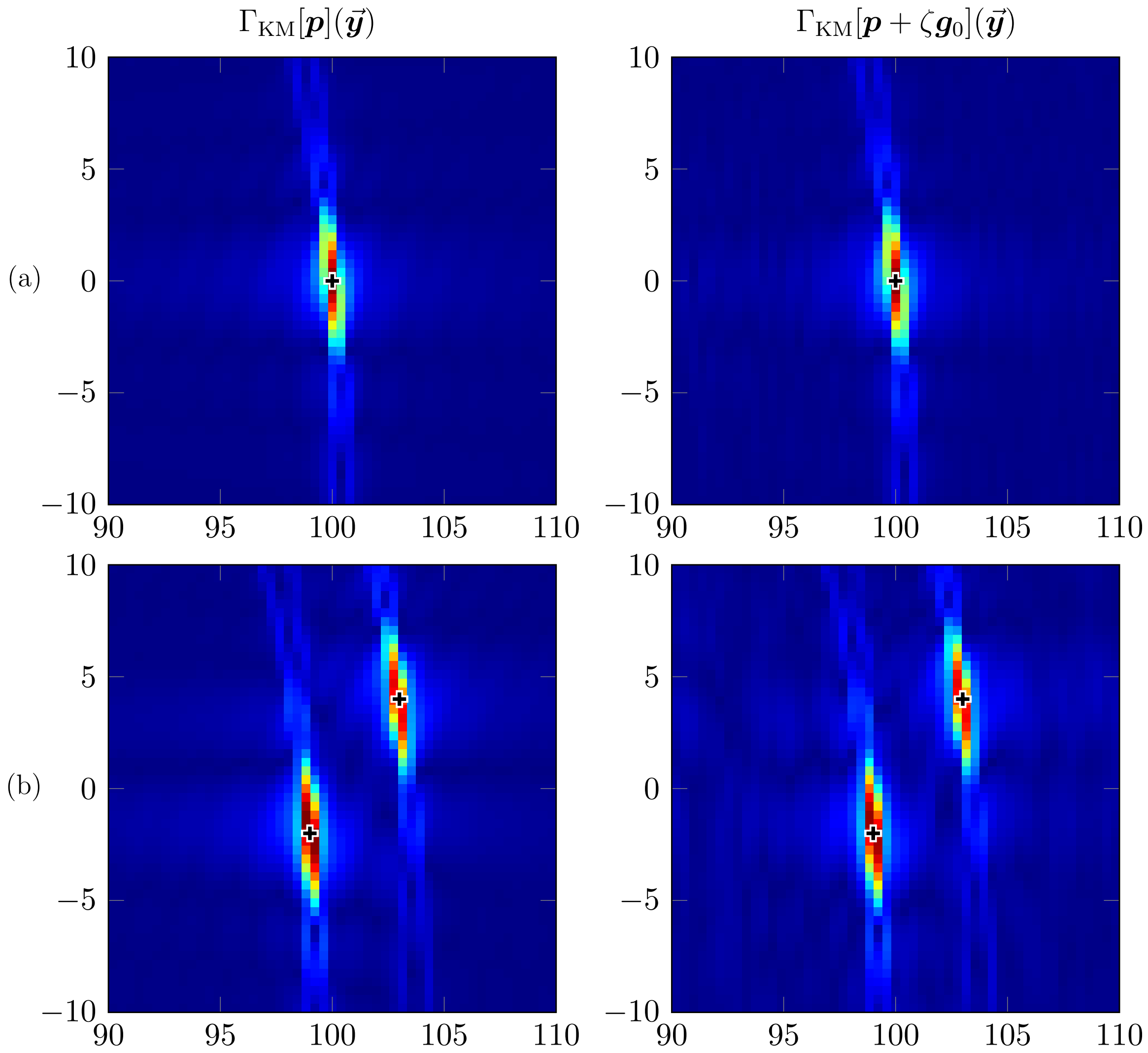


Figure 3.6. Kirchhoff images of (a) one point and (b) two point reflectors, whose true positions are indicated with crosses. The left column uses the full waveform data $\boldsymbol{p}$, while the right column use the recovered data $\boldsymbol{p}+\zeta\boldsymbol{g}_0$. The horizontal and vertical axes display the range and cross-range respectively, with scales in central wavelengths λ_0.

segment is generated as a set of point reflectors each with $\rho(\vec{\boldsymbol{y}}_i) = 1 \times 10^{-9}$ uniformly spaced by $\lambda_0/8$.

We now demonstrate the robustness of the proposed method with respect to additive noise (see section 3.7). Here we have taken a realization of the data for a single point reflector (Figure 3.6a) and perturbed each measurement with additive noise as follows. The $m-$th signal $\widehat{u}_m(\vec{\boldsymbol{x}}_r, \omega)$ has total power $p_m = \int |\widehat{u}_m(\vec{\boldsymbol{x}}_r, \omega)|^2 d\omega$.

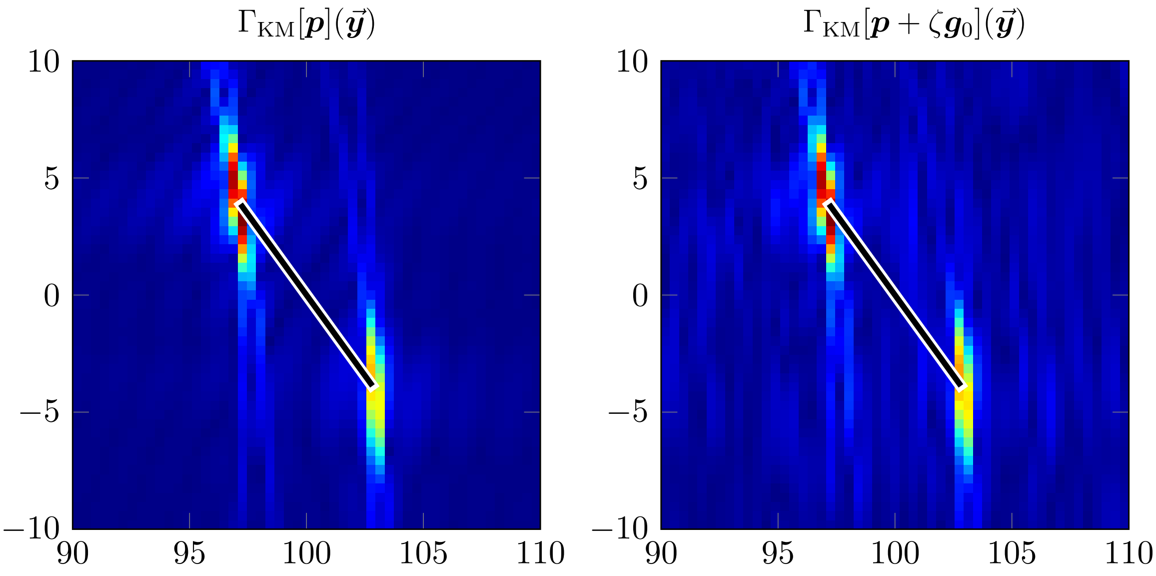


Figure 3.7. Kirchhoff images of an extended reflector. The left image uses the full waveform data $\boldsymbol{p}$, while the right image uses the recovered data $\boldsymbol{p}+\zeta\boldsymbol{g}_0$. The horizontal and vertical axes display the range and cross-range respectively, with scales in central wavelengths λ_0.

We construct a Gaussian signal $\xi_m(t)$ with mean zero, spectral density (3.38), $l_c \approx 1.25$ μs and total power 1. This allows us to obtain the perturbed total field $\widehat{u}_m(\vec{\boldsymbol{x}}_r,\omega)+\sqrt{\nu p_m}\widehat{\xi}_m(\omega)$ for some $\nu > 0$. The $m-$th measurement with additive noise is thus $d_m(\vec{\boldsymbol{x}}_r,\omega) = |\widehat{u}_m(\vec{\boldsymbol{x}}_r,\omega)|^2+\nu p_m|\widehat{\xi}(\omega)|^2$. Thus the ratio of the signal power to the noise power is $1/\nu$. The signal-to-noise ratio (SNR) is then

$$\text{SNR}_m = -10\log_{10}(\nu)\text{dB}.$$

Figure 3.8 shows the reconstruction from data with $\text{SNR}_m = 0$ dB for each m, meaning that the signal and the noise have the same power.

Lastly we perform an experiment that sends as the $m-$th illumination the usual correlated pair illumination $\boldsymbol{f}_m$, and uncorrelated noise from the remaining sources on the array $\mathcal{A}$ (see remark 10). To generate this illumination we place the *same* realization of the signal $\widehat{f}(\omega)$ at the locations $\boldsymbol{x}_{i(m)}$ and $\boldsymbol{x}_{j(m)}$, and *independent* realizations of $\widehat{f}(\omega)$ at the remaining source locations. Similarly, a reference illumination is generated by placing *independent* realizations of $\widehat{f}(\omega)$ at *all* locations on the array $\mathcal{A}$. By measuring the autocorrelation of the resulting fields we obtain data that are

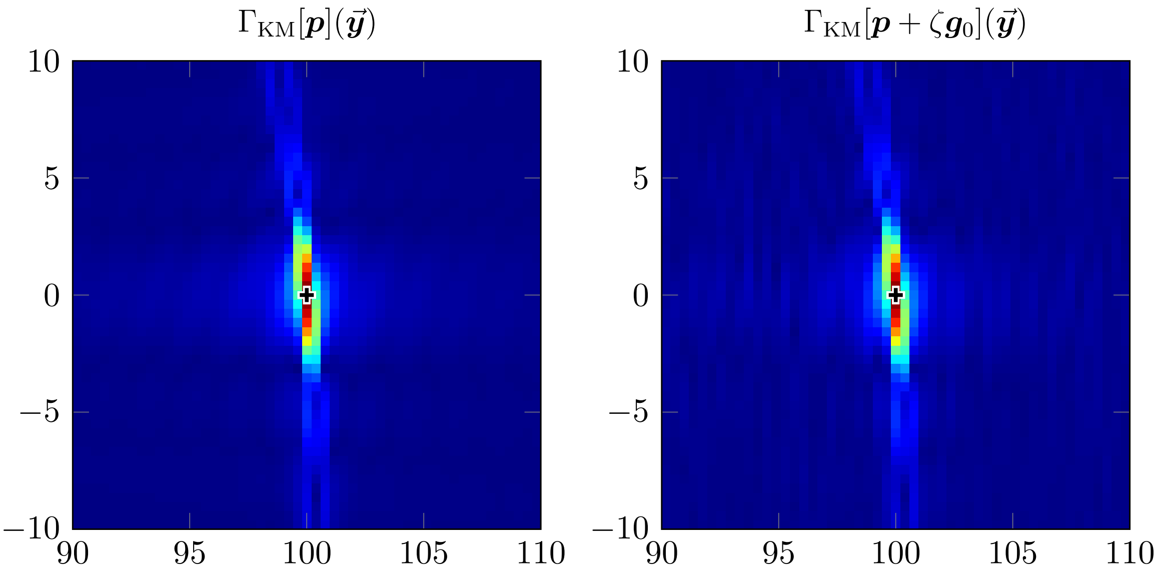


Figure 3.8. Additive noise: (left) array response vector migration $\Gamma_{\mathrm{KM}}[\boldsymbol{p}](\vec{\boldsymbol{y}})$, (right) recovered array response vector migration $\Gamma_{\mathrm{KM}}[\boldsymbol{p}+\zeta\boldsymbol{g}_0](\vec{\boldsymbol{y}})$ for $\mathrm{SNR}_m = 0\mathrm{dB}$. The horizontal and vertical axes display the range and cross-range respectively measured in central wavelengths λ_0.

essentially the same form as $d_m(\vec{\boldsymbol{x}}_r,\omega)$. Figure 3.9 shows this experiment with the single point reflector located at $\vec{\boldsymbol{y}} = (100\lambda_0, 0)$ and reflectivity $\rho(\vec{\boldsymbol{y}}) = 1\times 10^{-8}$.

3.8.2 Optic regime

For imaging in an optic regime, we use the background wave velocity $c_0 = 3\times 10^8$ m/s and central frequency $\approx$ 589 THz which gives a central wavelength $\lambda_0 \approx 509$ nm. Our source array $\mathcal{A}$ is again centered at the origin, but now consists of 501 point sources located at coordinates $\vec{\boldsymbol{x}}_s = \big(0, -0.005 + (s-1)(0.01/500)\big)$ for $s = 1, \ldots, 501$, and we set $\vec{\boldsymbol{x}}_r = (-0.005, -0.005)$. This corresponds to centering a 1cm source array at the origin with sources spaced approximately 20μm apart and a receiver placed slightly behind. We use this spacing for sources since it is physically attainable by using a Digital Micromirror Device (e.g., the Texas Instruments DLP4500 has a micromirror pitch of 7.6μm). Our numerical experiments assume the sources are points, which is not the case if we were using such a device (since each pixel is several wavelengths across). The effect of this approximation is not studied here.

We generate intensity data $\boldsymbol{d}(\vec{\boldsymbol{x}}_r,\omega)$ as

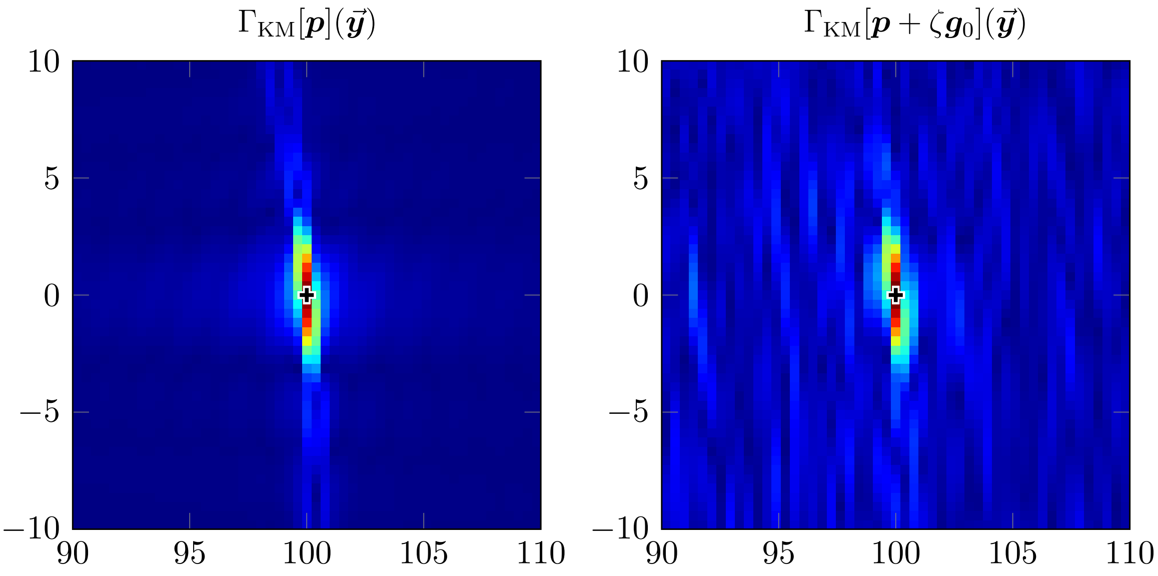


Figure 3.9. Uncorrelated background illumination: (left) array response vector migration $\Gamma_{\mathrm{KM}}[\boldsymbol{p}](\vec{\boldsymbol{y}})$, (right) recovered array response vector migration $\Gamma_{\mathrm{KM}}[\boldsymbol{p}+\zeta\boldsymbol{g}_0](\vec{\boldsymbol{y}})$ for $\mathrm{SNR}_m = 0\mathrm{dB}$. The horizontal and vertical axes display the range and cross-range respectively measured in central wavelengths λ_0.

$$d_m(\boldsymbol{x}_r,\omega) = \left|\left(\boldsymbol{g}_0+\boldsymbol{p}\right)^T(\boldsymbol{e}_{i(m)}+\boldsymbol{e}_{j(m)})\right|^2 - \left|\left(\boldsymbol{g}_0+\boldsymbol{p}\right)^T\boldsymbol{e}_{i(m)}\right|^2 - \left|\left(\boldsymbol{g}_0+\boldsymbol{p}\right)^T\boldsymbol{e}_{j(m)}\right|^2,$$

for 1000 (angular) frequencies ω uniformly spaced in the frequency band $[429, 750]$ THz. This corresponds to performing the source pair experiments (source pair illuminations and single-source reference illuminations) for 1000 different monochromatic visible light sources with wavelengths $\lambda \in [400, 700]$ nm, equally spaced in frequency. Since there are a large number of sources in this setup ($N = 501$), we implement the strategy discussed in §3.4 to reduce the number of source pair experiments from $N_p = N(N-1)/2$ to $N_p = 2N$.

As before, we use the pseudoinverse $\mathsf{M}^\dagger$ to recover $\boldsymbol{p}+\zeta\boldsymbol{g}_0$ for each frequency $\omega \in \mathcal{B}$, and then use the Kirchhoff migration functional (§3.2.2) to form an image. Here we use the image window $\mathcal{W} = \{(0.05 + i\lambda_0/2.5, j\lambda_0/2.5),\ \text{for}\ i,j = -25,\ldots,25\}$. In Figure 3.10 we demonstrate the migrated image for two point reflectors placed at $\vec{\boldsymbol{y}}_1 = (0.05 - 2\lambda_0, 2\lambda_0)$ and $\vec{\boldsymbol{y}}_2 = (0.05 + 3\lambda_0, -4\lambda_0)$ each with reflectivity $\rho(\vec{\boldsymbol{y}}_i) = 1\times 10^{-15}$. Although we are significantly undersampling the data in frequency and on the array (the source spacing is of about $40\lambda_0$, much larger than $\lambda_0/2$), the spot sizes

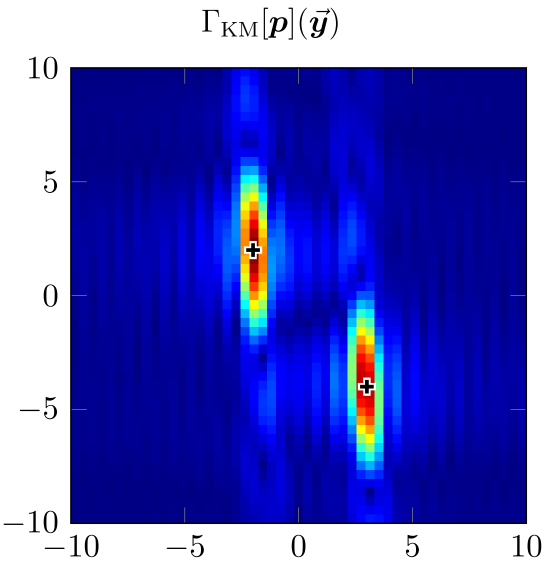


Figure 3.10. Optic regime: (left) array response vector migration $\Gamma_{\mathrm{KM}}[\boldsymbol{p}](\vec{\boldsymbol{y}})$, (right) recovered array response vector migration $\Gamma_{\mathrm{KM}}[\boldsymbol{p}+\zeta\boldsymbol{g}_0](\vec{\boldsymbol{y}})$. The horizontal and vertical axes display the range and cross-range respectively measured in central wavelengths λ_0 from the center $(0.05, 0)$.

still exhibit the Kirchhoff migration resolution estimates (§3.2.2) of $\lambda_0 L/a \approx 5\lambda_0$ in cross-range and $c_0/|\mathcal{B}| \approx 2\lambda_0$ in range.

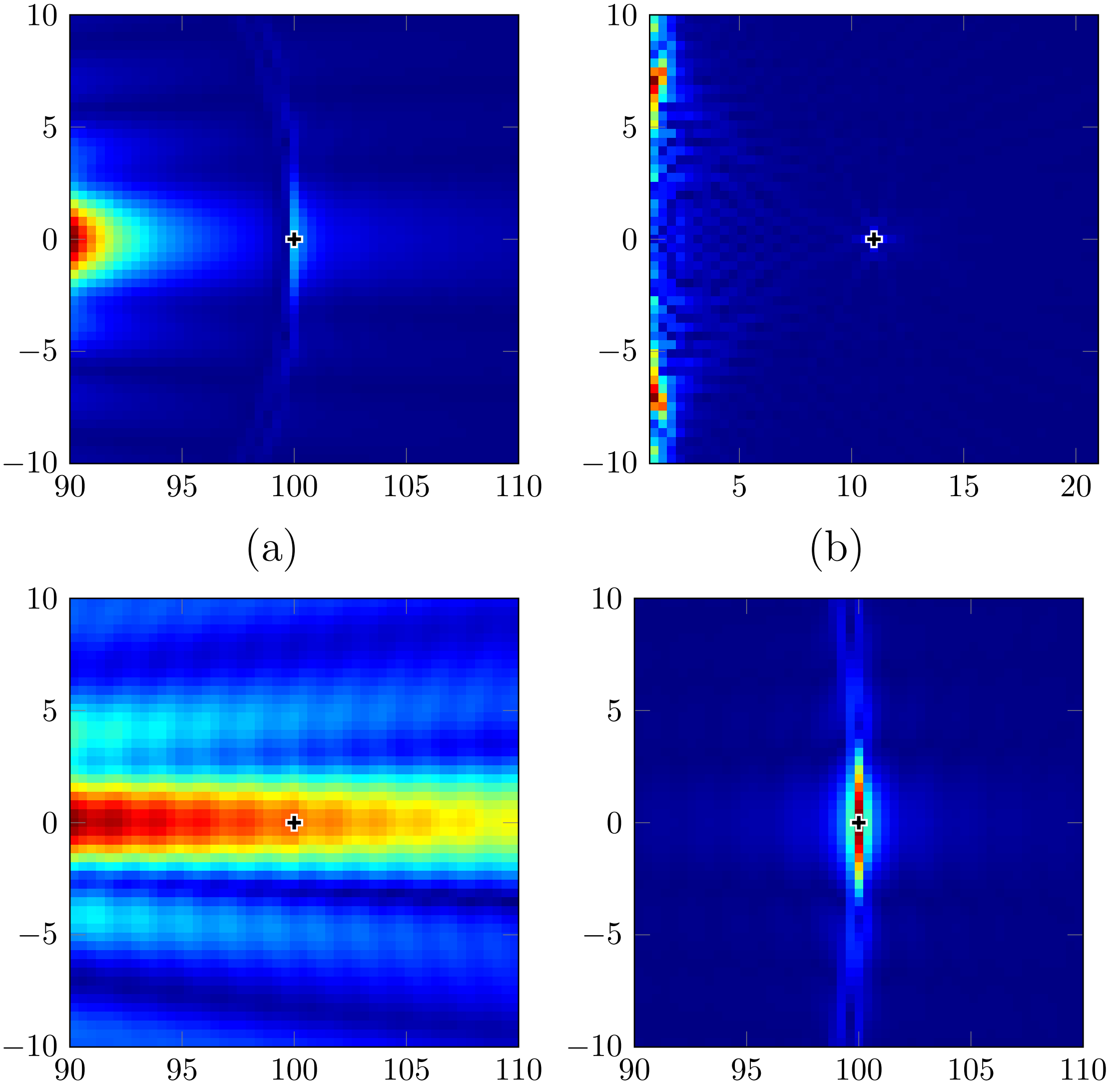
3.8.3 Breakdown of the imaging method

Here we consider physical setups in the acoustic regime ($c_0 = 1500$ m/s, $(2\pi)^{-1}\mathcal{B} = [1, 5]$ MHz, $\lambda_0 = 0.5$ mm) for which assumptions 1 and/or 3 are violated. For these experiments we fix the source positions $\vec{\boldsymbol{x}}_s = (0, -10\lambda_0 + (s-1)\lambda_0/2)$ for $s = 1, \dots, 41$ and vary a point scatterer location $\vec{\boldsymbol{y}}$, reflectivity size $\rho(\vec{\boldsymbol{y}})$ and receiver location $\vec{\boldsymbol{x}}_r$. Deterministic intensity data are collected as

$$d_m(\omega) = \left|(\boldsymbol{g}_0 + \boldsymbol{p})(\boldsymbol{e}_{i(m)} + \boldsymbol{e}_{j(m)})\right|^2 - \left|(\boldsymbol{g}_0 + \boldsymbol{p})\boldsymbol{e}_{i(m)}\right|^2 - \left|(\boldsymbol{g}_0 + \boldsymbol{p})\boldsymbol{e}_{j(m)}\right|^2,$$

for each $m = 1, \dots, N(N-1)/2$ and 1000 equally spaced frequencies $\omega \in \mathcal{B}$.

In Figure 3.11 we show the Kirchhoff migrated images of the recovered array response vector $\Gamma_{\mathrm{KM}}[\boldsymbol{p} + \zeta\boldsymbol{g}_0]$ for the situations indicated in Table 3.1. From Figures 3.11a, 3.11b and 3.11c, our method appears most sensitive to the smallness assumption 1. In these situations, the quadratic term $\boldsymbol{p}^*\mathsf{H}_m\boldsymbol{p}$ is non-negligible and



(a) (b)

(c) (d)

Figure 3.11. Breakdown of imaging method: migrated images of the recovered array response vector $\Gamma_{\mathrm{KM}}[\boldsymbol{p}+\zeta\boldsymbol{g}_0]$ for setups violating assumptions 1 and/or 3. The physical setups are given above and correspond to (a) $\vec{\boldsymbol{x}}_r$ placed too close to $\vec{\boldsymbol{y}}$, (b) $\mathcal{A}$ placed too close to $\vec{\boldsymbol{y}}$, (c) large reflectivity ρ and (d) $\vec{\boldsymbol{x}}_r$ placed in front of the array $\mathcal{A}$. The true position of the point scatterer is indicated with a cross.

Table 3.1. Breakdown of imaging method: physical setups and violated assumptions for each of the numerical experiments shown in Figure 3.11.

	Assumptions violated	$\vec{\boldsymbol{y}}$	$\rho(\vec{\boldsymbol{y}})$	$\vec{\boldsymbol{x}}_r$
(a) Receiver near scatterer	1 and 3	$(100\lambda_0, 0)$	1×10^{-8}	$(85\lambda_0, 0)$
(b) Sources near scatterer	1	$(11\lambda_0, 0)$	1×10^{-8}	$(-200\lambda_0, 0)$
(c) Large reflectivity	1	$(100\lambda_0, 0)$	1×10^{-3}	$(-10\lambda_0, 0)$
(d) No geometric condition	3	$(100\lambda_0, 0)$	1×10^{-8}	$(10\lambda_0, 0)$

the linearization we use in §3.3.2 is invalid, which leads to artifacts in the images. As seen in Figure 3.11d where only assumption 3 is violated, the method appears more robust than expected.

By the ergodicity result (proposition 1) we are guaranteed these experiments also indicate the performance of our method when using stochastic source pairs, at least in the limit $T \to \infty$. One additional situation where we expect our method to breakdown, is when using stochastic source illuminations with large correlation times t_c. If t_c is sufficiently large, incident-incident correlations may dominate the data, preventing us from resolving the reflected arrivals necessary for imaging. This follows from a discussion similar to that in §3.7.

3.9 Discussion

By sending correlated signals from different pairs of locations we have shown that the intensity data can be approximated by a linear system on the condition that smallness assumption 1 holds. This linear system has a known one-dimensional nullspace provided the sources and receiver satisfy the distance conditions given by assumption 2, which allows for the recovery of $\boldsymbol{p} + \zeta \boldsymbol{g}_0$. We show this quantity is enough to use standard migration techniques (e.g., Kirchhoff migration Γ_{KM}) provided the sources and receiver satisfy the additional geometric conditions of assumption 3. Thus we obtain full waveform resolution estimates for an image formed from intensity-only data.

Our method relies on assumption 1 to neglect quadratic terms in $\boldsymbol{p}$. As expected our method breaks down if such terms are large, e.g., when the sources or the receiver are near the scatterers or for large scattering potentials.

Our method relies only on knowledge of paired source locations and the correlation of the signals being sent. This allows us to relax illumination control by using paired stochastic signals. By measuring autocorrelations of the resulting fields, we obtain essentially the same intensity data as with using deterministic source pairs. These stochastic illuminations can be created, e.g., by using a configurable mask that is parallel to the wave fronts of an incoherent plane wave.

The linear system we solve has size $2N \times 2N$ and is very sparse (up to 4 nonzero entries per row). In our simulations we used $\mathsf{M}^\dagger$, however sparse solvers such as CGLS (see, e.g., [19]) could be used. To form the system we need at least $3N$ different illuminations, $2N$ pair illuminations plus N reference illuminations. However, in our illumination strategy, the phase of the source signals does not need to be known. We replace the direct phase control by the natural phase modulation that comes from the different positions of the signals.

We use the geometric imaging conditions (assumption 3) to show the nullspace of M does not affect imaging via Γ_{KM}. This assumption imposes some restrictions on the juxtaposition of the sources and receiver and in turn on the forms of illuminations we can consider. For example, using a stationary phase argument, it can be shown the autocorrelation of the total field is negligible if spatially continuous array illuminations (rather than paired point sources) are used. In future work, we would like to address this more thoroughly to determine if more general illuminations can be used. It may also be interesting to see if the source pair strategy we propose works for other imaging setups.

Lastly we recall our method assumes line-of-sight between sources, scatterers and receiver and a homogeneous linear response from the medium. We leave the study of more complicated setups and media for future work.

3.10 Proof of Proposition 1

In this section we prove proposition 1 which details the statistical stability of the measured autocorrelation (4.18) with respect to realizations of the illumination $\boldsymbol{f}$. The theorem and proof are patterned after the result by Garnier and Papanicolaou [12, Proposition 4.1], only we make small modifications to allow for complex fields and the form (3.26) of the correlation function $\mathsf{R}(\tau)$.

Proof. Since we are assuming $\boldsymbol{f}$ is a stationary process in t, the resulting total field u is also a stationary random process in t. So we have

$$\langle \overline{u}(\vec{\boldsymbol{x}}_r, t) u(\vec{\boldsymbol{x}}_r, t+\tau)\rangle = \langle \overline{u}(\vec{\boldsymbol{x}}_r, 0) u(\vec{\boldsymbol{x}}_r, \tau)\rangle,$$

which allows us to compute

$$\begin{aligned}\langle \psi(\vec{\boldsymbol{x}}_r, \tau)\rangle &= \frac{1}{2T}\int_{-T}^{T} dt \langle \overline{u}(\vec{\boldsymbol{x}}_r, t) u(\vec{\boldsymbol{x}}_r, t+\tau)\rangle \\ &= \frac{1}{2T}\int_{-T}^{T} dt \langle \overline{u}(\vec{\boldsymbol{x}}_r, 0) u(\vec{\boldsymbol{x}}_r, \tau)\rangle = \langle \overline{u}(\vec{\boldsymbol{x}}_r, 0) u(\vec{\boldsymbol{x}}_r, \tau)\rangle.\end{aligned}$$

So (4.18) is independent of T. By expressing the quantity $\langle \overline{u}(\vec{\boldsymbol{x}}_r, 0) u(\vec{\boldsymbol{x}}_r, \tau)\rangle$ through the Green's function G we verify (3.29):

$$\begin{aligned}\langle \psi(\vec{\boldsymbol{x}}_r, \tau)\rangle &= \sum_{p,p'=1}^{N} \int dt' \int dt'' \overline{G}(\vec{\boldsymbol{x}}_r, \vec{\boldsymbol{x}}_p, -t') G(\vec{\boldsymbol{x}}_r, \vec{\boldsymbol{x}}_{p'}, \tau - t'') \langle \overline{f}_p(t') f_{p'}(t'')\rangle \\ &= \sum_{p,p'=1}^{N} \int dt' \int dt'' \overline{G}(\vec{\boldsymbol{x}}_r, \vec{\boldsymbol{x}}_p, -t') G(\vec{\boldsymbol{x}}_r, \vec{\boldsymbol{x}}_{p'}, \tau - t'') R_{p,p'}(t'' - t') \\ &= \frac{1}{2\pi}\int d\omega e^{-\imath\omega\tau} \boldsymbol{g}(\vec{\boldsymbol{x}}_r, \omega)^{*} \widehat{\mathsf{R}}(\omega) \boldsymbol{g}(\vec{\boldsymbol{x}}_r, \omega).\end{aligned}$$

To show the ergodicity (3.31), we need to compute the variance of ψ. We first compute the covariance as

$$\begin{aligned}\mathrm{Cov}\big(\psi(\vec{\boldsymbol{x}}_r, \tau), \psi(\vec{\boldsymbol{x}}_r, \tau+\Delta\tau)\big) &= \sum_{p,p',q,q'=1}^{N} \frac{1}{(4T)^2}\int_{-T}^{T}\int_{-T}^{T} dt dt' \int ds ds' du du' \\ &\times G(\vec{\boldsymbol{x}}_r, \vec{\boldsymbol{x}}_p, s) \overline{G}(\vec{\boldsymbol{x}}_r, \vec{\boldsymbol{x}}_{p'}, u-\tau) \overline{G}(\vec{\boldsymbol{x}}_r, \vec{\boldsymbol{x}}_q, s') G(\vec{\boldsymbol{x}}_r, \vec{\boldsymbol{x}}_{q'}, u' - \tau - \Delta\tau) \\ &\times \Big(\langle f_p(t-s) \overline{f}_{p'}(t-u) \overline{f}_q(t'-s') f_{q'}(t'-u')\rangle \\ &- \langle f_p(t-s)\overline{f}_{p'}(t-u)\rangle \langle \overline{f}_q(t'-s') f_{q'}(t'-u')\rangle\Big).\end{aligned} \tag{3.40}$$

The product of the second order moments is

$$\langle f_p(t-s)\overline{f}_{p'}(t-u)\rangle \langle \overline{f}_q(t'-s') f_{q'}(t'-u')\rangle = R_{p',p}(u-s) R_{q,q'}(s'-u').$$

Since $\boldsymbol{f}(t)$ is Gaussian (in time), the fourth order moment is given by the complex Gaussian moment theorem (see, e.g., [26]) as

$$\begin{aligned}\langle f_p(t-s) \overline{f}_{p'}(t-u) \overline{f}_q(t'-s') f_{q'}(t'-u')\rangle &= R_{p',p}(u-s) R_{q,q'}(s'-u') \\ &+ R_{q,p}(t-t'-s+s') R_{p',q'}(t'-t-u'+u).\end{aligned}$$

We now integrate over the t, t' variables to obtain

$$\begin{aligned}
&\frac{1}{4T^2}\int_{-T}^{T} dt \int_{-T}^{T} dt' \Big(\langle f_p(t-s)\overline{f}_{p'}(t-u)\overline{f}_q(t'-s')f_{q'}(t'-u')\rangle \\
&\qquad\qquad - \langle f_p(t-s)\overline{f}_{p'}(t-u)\rangle\langle \overline{f}_q(t'-s')f_{q'}(t'-u')\rangle \Big) \\
&= \frac{1}{4T^2}\int_{-T}^{T} dt \int_{-T}^{T} dt' R_{q,p}(t-t'-s+s')R_{p',q'}(t'-t-u'+u) \\
&= \int d\omega \int d\omega' \operatorname{sinc}^2\big((\omega-\omega')T\big) e^{\imath\omega'(s-s')} e^{-\imath\omega(u-u')} \widehat{R}_{q,p}(\omega)\widehat{R}_{p',q'}(\omega').
\end{aligned}$$

Plugging this into (3.40) we obtain

$$\begin{aligned}
&\operatorname{Cov}\big(\psi(\vec{\boldsymbol{x}}_r,\tau),\psi(\vec{\boldsymbol{x}}_r,\tau+\Delta\tau)\big) = \sum_{p,p',q,q'=1}^{N} \int d\omega \int d\omega' \operatorname{sinc}^2\big((\omega-\omega')T\big) \\
&\qquad \times \widehat{G}(\vec{\boldsymbol{x}}_r,\vec{\boldsymbol{x}}_p,\omega')\overline{\widehat{G}}(\vec{\boldsymbol{x}}_r,\vec{\boldsymbol{x}}_{p'},\omega)\overline{\widehat{G}}(\vec{\boldsymbol{x}}_r,\vec{\boldsymbol{x}}_q,\omega)\widehat{G}(\vec{\boldsymbol{x}}_r,\vec{\boldsymbol{x}}_{q'},\omega)\widehat{R}_{q,p}(\omega)\widehat{R}_{p',q'}(\omega')e^{\imath\omega\Delta\tau} \\
&= \int d\omega \int d\omega' \operatorname{sinc}^2\big((\omega-\omega')T\big)\big(\boldsymbol{g}(\vec{\boldsymbol{x}}_r,\omega)^*\widehat{\mathsf{R}}(\omega)\boldsymbol{g}(\vec{\boldsymbol{x}}_r,\omega')\big) \\
&\qquad\qquad\qquad \times \big(\boldsymbol{g}(\vec{\boldsymbol{x}}_r,\omega)^*\widehat{\mathsf{R}}(\omega')\boldsymbol{g}(\vec{\boldsymbol{x}}_r,\omega')\big)e^{\imath\omega\Delta\tau},
\end{aligned}$$

where $\boldsymbol{g} = \boldsymbol{g}_0 + \boldsymbol{p}$ is given by (3.2) and (3.6), and $\big(\widehat{\mathsf{R}}(\omega)\big)_{i,j} = \widehat{R}_{i,j}(\omega)$ is a $\mathbb{C}^{N\times N}$ Hermitian matrix for each ω. Then taking $T \to \infty$ we compute the variance as

$$T \operatorname{Var}\big(\psi_m(\vec{\boldsymbol{x}}_r,\tau)\big) \xrightarrow{T\to\infty} \int d\omega \Big|\boldsymbol{g}(\vec{\boldsymbol{x}}_r,\omega)^*\widehat{\mathsf{R}}(\omega)\boldsymbol{g}(\vec{\boldsymbol{x}}_r,\omega)\Big|^2,$$

and so the variance is $\mathcal{O}(1/T)$ as $T \to \infty$. This establishes (3.31). ■

3.11 Acknowledgements

The authors would like to thank Alexander Mamonov, Andy Thaler, and Greg Rice for insightful discussions related to this project. PB is especially grateful to Graeme W. Milton for his generous support.

3.12 References

[1] G. Agrawal, *Fiber-Optic Communication Systems*, Wiley Series in Microwave and Optical Engineering, Wiley, 2012.

[2] N. Bleistein, J. K. Cohen, and J. W. Stockwell, Jr., *Mathematics of multidimensional seismic imaging, migration, and inversion*, vol. 13 of Interdisciplinary Applied Mathematics, Springer-Verlag, New York, 2001.

[3] N. Bleistein and R. A. Handelsman, *Asymptotic expansions of integrals*, Dover Publications, Inc., New York, second ed., 1986.

[4] L. Borcea, G. Papanicolaou, and C. Tsogka, *Interferometric array imaging in clutter*, Inverse Problems, 21 (2005), pp. 1419–1460.

[5] E. J. Candès, T. Strohmer, and V. Voroninski, *PhaseLift: Exact and stable signal recovery from magnitude measurements via convex programming*, Comm. Pure Appl. Math., 66 (2013), pp. 1241–1274.

[6] A. Chai, M. Moscoso, and G. Papanicolaou, *Array imaging using intensity-only measurements*, Inverse Problems, 27 (2011), p. 015005.

[7] ——, *Imaging strong localized scatterers with sparsity promoting optimization*, SIAM J. Imaging Sci., 7 (2014), pp. 1358–1387.

[8] M. Cheney, *The linear sampling method and the MUSIC algorithm*, Inverse Problems, 17 (2001), pp. 591–595.

[9] M. Cheney and B. Borden, *Fundamentals of radar imaging*, vol. 79 of CBMS-NSF Regional Conference Series in Applied Mathematics, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2009.

[10] L. Crocco, M. D'Urso, and T. Isernia, *Faithful non-linear imaging from only-amplitude measurements of incident and total fields*, Opt. Express, 15 (2007), pp. 3804–3815.

[11] J. R. Fienup, *Reconstruction of a complex-valued object from the modulus of its Fourier transform using a support constraint*, J. Opt. Soc. Am. A, 4 (1987), pp. 118–123.

[12] J. Garnier and G. Papanicolaou, *Passive sensor imaging using cross correlations of noisy signals in a scattering medium*, SIAM J. Imaging Sci., 2 (2009), pp. 396–437.

[13] ——, *Resolution analysis for imaging with noise*, Inverse Problems, 26 (2010), pp. 074001, 22.

[14] ——, *Correlation-based virtual source imaging in strongly scattering random media*, Inverse Problems, 28 (2012), p. 075002.

[15] ——, *Role of scattering in virtual source array imaging*, SIAM J. Imaging Sci., 7 (2014), pp. 1210–1236.

[16] J. Garnier, G. Papanicolaou, A. Semin, and C. Tsogka, *Signal to noise ratio analysis in virtual source array imaging*, SIAM J. Imaging Sci., 8 (2015), pp. 248–279.

[17] G. Gbur and E. Wolf, *The information content of the scattered intensity in diffraction tomography*, Inform. Sci., 162 (2004), pp. 3–20.

[18] R. Gerchberg and W. Saxton, *A practical algorithm for the determination of phase from image and diffraction plane pictures*, Optik, 35 (1972).

[19] P. C. Hansen, *Rank-deficient and discrete ill-posed problems: Numerical aspects of linear inversion*, SIAM Monographs on Mathematical Modeling and Computation, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1998.

[20] D. Huang, H. Timmers, A. Roberts, N. Shivaram, and A. S. Sandhu, *A low-cost spatial light modulator for use in undergraduate and graduate optics labs*, American Journal of Physics, 80 (2012), pp. 211–215.

[21] A. Ishimaru, *Wave propagation and scattering in random media*, IEEE/OUP Series on Electromagnetic Wave Theory, IEEE Press, New York, 1997.

[22] M. H. Maleki and A. J. Devaney, *Phase-retrieval and intensity-only reconstruction algorithms for optical diffraction tomography*, J. Opt. Soc. Am. A, 10 (1993), pp. 1086–1092.

[23] M. H. Maleki, A. J. Devaney, and A. Schatzberg, *Tomographic reconstruction from optical scattered intensities*, J. Opt. Soc. Am. A, 9 (1992), pp. 1356–1363.

[24] J. Neff, R. Athale, and S. Lee, *Two-dimensional spatial light modulators: A tutorial*, Proceedings of the IEEE, 78 (1990), pp. 826–855.

[25] A. Novikov, M. Moscoso, and G. Papanicolaou, *Illumination strategies for intensity-only imaging*, SIAM Journal on Imaging Sciences, 8 (2015), pp. 1547–1573.

[26] I. Reed, *On a moment theorem for complex Gaussian processes*, Information Theory, IRE Transactions on, 8 (1962), pp. 194–195.

[27] Y.-X. Ren, R.-D. Lu, and L. Gong, *Tailoring light with a digital micromirror device*, Annalen der Physik, 527 (2015), pp. 447–470.

[28] J. Schmitt, *Optical coherence tomography (OCT): A review*, IEEE Journal of Selected Topics in Quantum Electronics, 5 (1999), pp. 1205–1215.

[29] J. Schmitt, S. Lee, and K. Yung, *An optical coherence microscope with enhanced resolving power in thick tissue*, Optics Communications, 142 (1997), pp. 203 – 207.

[30] G. T. Schuster, *Resolution limits for crosswell migration and traveltime tomography*, Geophysical Journal International, 127 (1996), pp. 427–440.

[31] ——, *Seismic Interferometry*, Cambridge University Press, 2009.

[32] G. T. Schuster, J. Yu, J. Sheng, and J. Rickett, *Interferometric/daylight seismic imaging*, Geophysical Journal International, 157 (2004), pp. 838–852.

[33] D. Tarchi, K. Lukin, J. Fortuny-Guasch, A. Mogyla, P. Vyplavin, and A. Sieber, *SAR imaging with noise radar*, IEEE Transactions on Aerospace and Electronic Systems, 46 (2010), pp. 1214–1225.

[34] M. R. Teague, *Deterministic phase retrieval: a Green's function solution*, J. Opt. Soc. Am., 73 (1983), pp. 1434–1441.

[35] R. Vela, R. Narayanan, K. Gallagher, and M. Rangaswamy, *Noise radar tomography*, in Radar Conference (RADAR), IEEE, 2012, pp. 0720–0724.

[36] P. Yin and J. Xin, *PhaseLiftOff: An accurate and stable phase retrieval method based on difference of trace and Frobenius norms*, Commun. Math. Sci., 13 (2015), pp. 1033–1049.

CHAPTER 4

KIRCHHOFF MIGRATION WITHOUT PHASES

A joint work with Fernando Guevara Vasquez.

4.1 Introduction

In Chapter 3 we developed a method to image scatterers in a homogeneous medium from intensity-only measurements. Our method used a specific illumination strategy of source pairs, meaning we sent correlated signals from two different locations on a *source array* and measured the intensity of the resulting wavefield at a *single receiver* location. We showed a projection of full waveform data could be recovered from a collection of these intensity measurements (i.e., intensity measurements from multiple source pair experiments) by solving a linear least-squares problem. Moreover, we showed this projection was sufficient to image with Kirchhoff migration. In this chapter, we generalize and in fact simplify this imaging method by considering a reciprocal physical setup consisting of a *single source* location and a *receiver array* where intensities are recorded. Since we have only one source, we can no longer use source pairs and thus we can only record intensities for a single experiment. This means we work with less data than in Chapter 3, and our illumination strategy is greatly simplified. However, even in this simplified setup we can still exploit the principle that Kirchhoff migration does not require complete phase information to form an image.

The setup we now analyze consists of one source and N receivers, all of known location. The receivers can only record intensities and only the source intensity is known. If the scattered field is small compared to the probing field (at the receivers) then the scattered field projected onto a known subspace can be found from the intensity data by solving an underdetermined, real, least-squares problem of size $N \times 2N$ (per frequency sample). This system is underdetermined because we are trying to use the intensity data, i.e., N real measurements, to recover the scattered field, i.e., N complex or $2N$ real numbers. Fortunately, a stationary phase argument similar to that in §3.5, shows that the error made by projecting the scattered field does not affect Kirchhoff imaging (for high frequencies). Moreover the least-squares problem is typically well conditioned and its solution is embarrassingly simple: it merely costs about $2N$ complex operations (additions or multiplications). Hence our method is comparable in computational cost to Kirchhoff migration. Well-known resolution studies for Kirchhoff migration, discussed in §3.2.2, can also be used for this method.

4.1.1 Contents

The physical setup and notations we use for this chapter are described in §4.2.1, where we also briefly review Kirchhoff migration in this setting. Using the Born approximation, we formulate the problem of recovering the full wave scattered field at the array from intensity-only measurements as a linear least-squares problem §4.2.2. In §4.3 we analyze and solve the least-squares problem and show that its solution can be used with Kirchhoff migration. We extend this imaging method to stochastic illuminations and autocorrelation measurements in §4.4. Numerical experiments for an optic regime are provided in §4.5, and we conclude with a discussion in §4.6.

4.2 Wave propagation and intensity-only measurements

Here we introduce the setup we work with and briefly recall Kirchhoff migration. As in Chapter 3, we use the Fourier transform convention for functions of time t:

$$\widehat{f}(\omega) = \int_{-\infty}^{\infty} dt f(t) e^{\imath\omega t}, \ f(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} d\omega \widehat{f}(\omega) e^{-\imath\omega t}, \ \text{ for } \ f(t), \ \widehat{f}(\omega) \in L^2(\mathbb{R}). \tag{4.1}$$

4.2.1 Wave propagation and imaging in a homogeneous medium

The physical setup we consider is depicted in Figure 4.1. We probe the medium with a point source located at $\vec{\boldsymbol{x}}_s$. Waves are recorded on an array of receivers $\vec{\boldsymbol{x}}_r = (\boldsymbol{x}_r, 0) \in \mathcal{A}$ for $r = 1, \ldots, N$, where $\mathcal{A} \subset \mathbb{R}^d \times \{0\}$ and $d = 2, 3$ is the dimension. We use the notation $\boldsymbol{x}$ for the first $d-1$ components of a vector $\vec{\boldsymbol{x}} \in \mathbb{R}^d$. For simplicity we consider a linear array in 2D or a square array in 3D, i.e., $\mathcal{A} = [-a/2, a/2]^{d-1} \times \{0\}$, however other shapes may be considered. We impose only mild conditions on the positions of the source and receivers, in particular that the source is not in the array. We assume the medium contains scatterers with reflectivity $\rho(\vec{\boldsymbol{y}})$ with $\text{supp}(\rho) = \mathcal{P}$, and background wave velocity c_0.

The total field arriving at the receiver location $\vec{\boldsymbol{x}}_r$ from frequency modulation $\widehat{f}(\omega)$ at the source location $\vec{\boldsymbol{x}}_s$ is

$$\widehat{u}(\vec{\boldsymbol{x}}_r, \vec{\boldsymbol{x}}_s, \omega) = \widehat{G}(\vec{\boldsymbol{x}}_r, \vec{\boldsymbol{x}}_s, \omega)\widehat{f}(\omega), \tag{4.2}$$

where $\widehat{G}$ is the Green's function for the (inhomogeneous) medium in the frequency

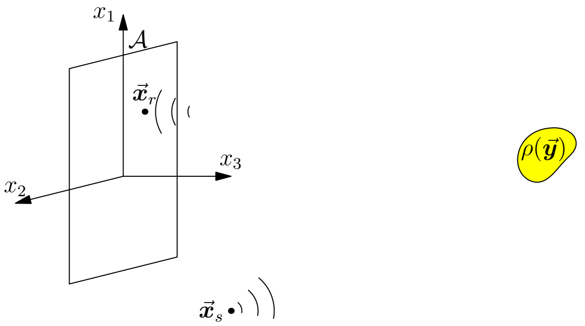


Figure 4.1. Physical setup for an array $\mathcal{A}$ of receivers with a wave source located at $\vec{\boldsymbol{x}}_s$. The scatterer is characterized by the compactly supported function $\rho(\vec{\boldsymbol{y}})$ (in yellow).

domain. We assume the scatterers are weak so that multiple scatterings may be neglected and by the Born approximation

$$\widehat{G}(\vec{\boldsymbol{x}}_r,\vec{\boldsymbol{x}}_s,\omega) \approx \widehat{G}_0(\vec{\boldsymbol{x}}_r,\vec{\boldsymbol{x}}_s,\omega) + k^2 \int_{\mathcal{P}} d\vec{\boldsymbol{y}}\rho(\vec{\boldsymbol{y}})\widehat{G}_0(\vec{\boldsymbol{x}}_r,\vec{\boldsymbol{y}},\omega)\widehat{G}_0(\vec{\boldsymbol{y}},\vec{\boldsymbol{x}}_s,\omega), \tag{4.3}$$

where $\widehat{G}_0$ is the Green's function for the Helmhotz equation given by (3.3).

We express the total fields received on the array with linear algebra notation as

$$u(\vec{\boldsymbol{x}}_r,\vec{\boldsymbol{x}}_s,\omega) = \boldsymbol{e}_r^{\mathsf{T}}\big(\boldsymbol{g}_0(\vec{\boldsymbol{x}}_s,\omega) + \boldsymbol{p}(\vec{\boldsymbol{x}}_s,\omega)\big)\widehat{f}(\omega), \quad \text{for } r = 1,\dots,N,$$

where the vector $\boldsymbol{g}_0$ is the vector of direct arrivals (or incident field) at the array

$$\boldsymbol{g}_0(\vec{\boldsymbol{x}}_s,\omega) = \Big[\widehat{G}_0(\vec{\boldsymbol{x}}_1,\vec{\boldsymbol{x}}_s,\omega), \widehat{G}_0(\vec{\boldsymbol{x}}_2,\vec{\boldsymbol{x}}_s,\omega), \cdots, \widehat{G}_0(\vec{\boldsymbol{x}}_N,\vec{\boldsymbol{x}}_s,\omega)\Big]^{\mathsf{T}}, \tag{4.4}$$

and the array response vector (or scattered field at the array) is

$$\boldsymbol{p}(\vec{\boldsymbol{x}}_s,\omega) = k^2 \int_{\mathcal{P}} d\vec{\boldsymbol{z}}\boldsymbol{g}_0(\vec{\boldsymbol{z}},\omega)\widehat{G}_0(\vec{\boldsymbol{x}}_s,\vec{\boldsymbol{z}},\omega)\rho(\vec{\boldsymbol{z}}). \tag{4.5}$$

Note, the vectors $\boldsymbol{g}_0$ and $\boldsymbol{p}$ now represent fields arriving at different receiver locations $\{\vec{\boldsymbol{x}}_r\}_{r=1}^N$ from a single illumination at $\vec{\boldsymbol{x}}_s$, instead of fields arriving at a single receiver location $\vec{\boldsymbol{x}}_r$ from illuminations at different source locations $\{\vec{\boldsymbol{x}}_s\}_{s=1}^N$ as was the case in Chapter 3. In this sense, the vectors $\boldsymbol{g}_0$ and $\boldsymbol{p}$ defined in (4.4) and (4.5) are reciprocal of those defined in (3.2) and (3.6), respectively.

When full waveform measurements are available, i.e., $u(\vec{\boldsymbol{x}}_r,\vec{\boldsymbol{x}}_s,\omega)$ is known for $r = 1,\dots,N$, the scattered field $\boldsymbol{p}$ can be obtained from the total field at the array,

$\boldsymbol{g}_0$ and $\widehat{f}(\omega)$. The scatterers in the medium can be imaged with Kirchhoff migration applied to $\boldsymbol{p}$, which, for a single frequency ω and the physical setup we consider here has the form:

$$\Gamma_{\mathrm{KM}}\big[\boldsymbol{p},\omega\big](\vec{\boldsymbol{y}}) = \overline{\widehat{G}}_0(\vec{\boldsymbol{x}}_s,\vec{\boldsymbol{y}},\omega)\boldsymbol{g}_0(\vec{\boldsymbol{y}},\omega)^*\boldsymbol{p}(\vec{\boldsymbol{x}}_s,\omega), \tag{4.6}$$

where $\vec{\boldsymbol{y}}$ is a point in the image. The Kirchhoff migration functional has been studied extensively (see, e.g., Bleistein et al. [2] for a review). As discussed in §3.2.2, we can expect a cross-range resolution of $\lambda L/a$, and integrating $\Gamma_{\mathrm{KM}}[\boldsymbol{p},\omega]$ over a frequency band $\mathcal{B} = [-\omega_{\max},-\omega_{\min}] \cup [\omega_{\min},\omega_{\max}]$, a range resolution of $c_0/(\omega_{\max}-\omega_{\min})$.

4.2.2 Intensity-only measurements

Using the illumination $\widehat{f}(\omega)$ at the source location $\vec{\boldsymbol{x}}_s$, the *intensity-only* measurement of the wave field at $\vec{\boldsymbol{x}}_r \in \mathcal{A}$ is

$$|u(\vec{\boldsymbol{x}}_r,\vec{\boldsymbol{x}}_s,\omega)|^2 = |\widehat{f}(\omega)|^2\boldsymbol{e}_r^{\mathsf{T}}\left[\overline{(\boldsymbol{g}_0+\boldsymbol{p})}\odot(\boldsymbol{g}_0+\boldsymbol{p})\right], \tag{4.7}$$

where the operator $\odot$ denotes the componentwise or Hadamard product of two vectors and $\{\boldsymbol{e}_r\}_{r=1}^N$ is the standard orthonormal basis of $\mathbb{R}^N$. Our objective is to find as much as we can about $\boldsymbol{p}$ from the vector of measurements $[|u(\vec{\boldsymbol{x}}_r,\vec{\boldsymbol{x}}_s,\omega)|^2]_{r=1,\ldots,N}$. This is done by linearization, so we need to assume that the scattered field $\boldsymbol{p}$ is small compared to the direct arrival $\boldsymbol{g}_0$ at the array.

Assumption 4 *The position of the receivers, the source, and the reflectivity are such that* $|\boldsymbol{e}_r^{\mathsf{T}}\boldsymbol{p}| \ll |\boldsymbol{e}_r^{\mathsf{T}}\boldsymbol{g}_0|$, $r=1,\ldots,N$.

This assumption is satisfied, e.g., if the reflectivity is sufficiently small and the source $\vec{\boldsymbol{x}}_s$ is near the receiver array (as is shown in Figure 4.1). With assumption 4 we can neglect quadratic terms in $\boldsymbol{p}$ to approximate the intensity measurements (4.7) by a vector $\boldsymbol{d}(\vec{\boldsymbol{x}}_s,\omega)$ defined by

$$|u(\vec{\boldsymbol{x}}_r,\vec{\boldsymbol{x}}_s,\omega)|^2 \approx \boldsymbol{e}_r^{\mathsf{T}}\boldsymbol{d}(\vec{\boldsymbol{x}}_s,\omega) \equiv |\widehat{f}(\omega)|^2\boldsymbol{e}_r^{\mathsf{T}}\mathrm{Re}\left[\overline{\boldsymbol{g}}_0\odot(\boldsymbol{g}_0+2\boldsymbol{p})\right].$$

This is not, strictly speaking, a linear system for $\boldsymbol{p}\in\mathbb{C}^N$ since $\boldsymbol{u}\to\mathrm{Re}(\boldsymbol{u})$ is not a linear mapping from $\mathbb{C}^N$ to $\mathbb{C}^N$. However we can write an underdetermined linear system for the real and imaginary parts of $\boldsymbol{g}_0+\boldsymbol{p}$ as follows

$$|\widehat{f}(\omega)|^2 \mathsf{M}(\vec{\boldsymbol{x}}_s,\omega) \begin{bmatrix} \mathrm{Re}(\boldsymbol{g}_0 + 2\boldsymbol{p}) \\ \mathrm{Im}(\boldsymbol{g}_0 + 2\boldsymbol{p}) \end{bmatrix} = \boldsymbol{d}(\vec{\boldsymbol{x}}_s,\omega), \tag{4.8}$$

where the matrix $\mathsf{M}(\vec{\boldsymbol{x}}_s,\omega) \in \mathbb{R}^{N\times 2N}$ is given by

$$\mathsf{M}(\vec{\boldsymbol{x}}_s,\omega) = \left[\mathrm{diag}\left(\mathrm{Re}(\boldsymbol{g}_0)\right) \quad \mathrm{diag}\left(\mathrm{Im}(\boldsymbol{g}_0)\right)\right]. \tag{4.9}$$

We give in the next section an explicit solution to the least-squares problem (4.8).

4.3 Migrating a least-squares estimate of the scattered field

The first step in our imaging method consists of a cheap least-squares preprocessing step that gives an approximation to the array response vector (§4.3.1). The second step is to migrate this approximation with standard Kirchhoff migration §4.3.2. Crucially we show in theorem 6 that the mistake we make by using this approximation of the array response vector does not affect the Kirchhoff images.

4.3.1 Recovering a projection of the array response vector

We start by finding a simple and explicit expression to the pseudoinverse of the matrix M that we obtained from linearizing the problem of finding the real and imaginary parts of the array response vector $\boldsymbol{p}$. This can be used to recover from the data $\boldsymbol{d}$ the orthogonal projection of $[\mathrm{Re}(\boldsymbol{p})^\mathsf{T}, \mathrm{Im}(\boldsymbol{p})^\mathsf{T}]^\mathsf{T}$ onto a known N dimensional subspace (that depends only on $\boldsymbol{g}_0$). Moreover the process is well conditioned.

First notice that the matrix M is full-rank. Indeed a simple calculation gives that $\mathsf{M}\mathsf{M}^\mathsf{T} = \mathrm{diag}(\overline{\boldsymbol{g}}_0 \odot \boldsymbol{g}_0)$. This matrix is clearly invertible because it is a diagonal matrix with the moduli of 2D or 3D Green functions on the diagonal. Hence the Moore-Penrose pseudoinverse $\mathsf{M}^\dagger$ can be written explicitly

$$\mathsf{M}^\dagger = \mathsf{M}^\mathsf{T}(\mathsf{M}\mathsf{M}^\mathsf{T})^{-1} = \begin{bmatrix} \mathrm{diag}\left(\mathrm{Re}(\boldsymbol{g}_0)\right) \\ \mathrm{diag}\left(\mathrm{Im}(\boldsymbol{g}_0)\right) \end{bmatrix} \mathrm{diag}(\overline{\boldsymbol{g}}_0 \odot \boldsymbol{g}_0)^{-1}. \tag{4.10}$$

We can use $\mathsf{M}^\dagger$ to see what information about $\boldsymbol{p}$ we can recover from the right hand side $\boldsymbol{d}$ in the least-squares problem (4.8). Since M has an N dimensional nullspace, we can only expect to recover the orthogonal projection of $[\mathrm{Re}(\boldsymbol{p})^\mathsf{T}, \mathrm{Im}(\boldsymbol{p})^\mathsf{T}]^\mathsf{T}$ onto $\mathrm{range}(\mathsf{M}^\mathsf{T}) = (\mathrm{null}(\mathsf{M}))^\perp$. This projection has a simple form when we write it in $\mathbb{C}^N$, as can be seen in the next proposition.

Proposition 2 *Provided* $|\widehat{f}(\omega)|^2 \neq 0$*, the intensity measurements* $\boldsymbol{d}$ *determine*

$$\widetilde{\boldsymbol{p}} \equiv \boldsymbol{p} + (\overline{\boldsymbol{g}}_0)^{-1} \odot \boldsymbol{g}_0 \odot \overline{\boldsymbol{p}}, \tag{4.11}$$

where the inverse of a vector is understood componentwise. Moreover $\widetilde{\boldsymbol{p}}$ *can be obtained in about* $2N$ *complex operations from* $\boldsymbol{d}$ *with*

$$\widetilde{\boldsymbol{p}} = |\widehat{f}(\omega)|^{-2}(\overline{\boldsymbol{g}}_0)^{-1} \odot \boldsymbol{d} - \boldsymbol{g}_0. \tag{4.12}$$

Proof. Since we use the first (resp. last) N rows of $\mathsf{M}^\dagger$ to recover the real (resp. imaginary) part of a vector in $\mathbb{C}^N$, it is convenient to consider the matrix

$$\begin{bmatrix}\mathsf{I} & i\mathsf{I}\end{bmatrix} \mathsf{M}^\dagger = \operatorname{diag}(\boldsymbol{g}_0) \operatorname{diag}(\overline{\boldsymbol{g}}_0 \odot \boldsymbol{g}_0)^{-1} = \operatorname{diag}(\overline{\boldsymbol{g}}_0)^{-1},$$

where I is the $N \times N$ identity matrix. To see what information about $\boldsymbol{p}$ we can recover from the right hand side $\boldsymbol{d}$ in the least-squares system (4.8) we can evaluate:

$$\begin{aligned}
|\widehat{f}(\omega)|^{-2} \begin{bmatrix}\mathsf{I} & i\mathsf{I}\end{bmatrix} \mathsf{M}^\dagger \boldsymbol{d} &= \operatorname{diag}(\overline{\boldsymbol{g}}_0)^{-1} \mathsf{M} \begin{bmatrix}\operatorname{Re}(\boldsymbol{g}_0 + 2\boldsymbol{p}) \\ \operatorname{Im}(\boldsymbol{g}_0 + 2\boldsymbol{p})\end{bmatrix} \\
&= \operatorname{diag}(\overline{\boldsymbol{g}}_0)^{-1}[\operatorname{Re}(\boldsymbol{g}_0)^2 + \operatorname{Im}(\boldsymbol{g}_0)^2 + 2\operatorname{Re}(\boldsymbol{g}_0)\operatorname{Re}(\boldsymbol{p}) + 2\operatorname{Im}(\boldsymbol{g}_0)\operatorname{Im}(\boldsymbol{p})] \\
&= \boldsymbol{g}_0 + \operatorname{diag}(\overline{\boldsymbol{g}}_0)^{-1}(\boldsymbol{g}_0 \odot \overline{\boldsymbol{p}} + \overline{\boldsymbol{g}}_0 \odot \boldsymbol{p}) \\
&= \boldsymbol{g}_0 + \boldsymbol{p} + \boldsymbol{g}_0 \odot (\overline{\boldsymbol{g}}_0)^{-1} \odot \overline{\boldsymbol{p}} = \boldsymbol{g}_0 + \widetilde{\boldsymbol{p}}.
\end{aligned}$$

Hence we can get $\widetilde{\boldsymbol{p}}$ from the intensity data $\boldsymbol{d}$ with essentially N complex multiplications and N complex additions. ■

A natural question to ask is whether we can obtain $\widetilde{\boldsymbol{p}}$ in a stable manner from $\boldsymbol{d}$. This can be answered by looking at the conditioning of M, i.e., the ratio of the largest singular value σ_1 of M to σ_N, the smallest one. These are easily obtained from the square roots of the eigenvalues of the diagonal matrix $\mathsf{M}\mathsf{M}^T = \operatorname{diag}(\overline{\boldsymbol{g}}_0 \odot \boldsymbol{g}_0)$. Hence the conditioning of M is the ratio of the largest to the smallest moduli of the entries of $\boldsymbol{g}_0$:

$$\operatorname{cond}(\mathsf{M}) = \begin{cases} \dfrac{\max_r \left|H_0^{(1)}(k|\vec{\boldsymbol{x}}_r - \vec{\boldsymbol{x}}_s|)\right|}{\min_r \left|H_0^{(1)}(k|\vec{\boldsymbol{x}}_r - \vec{\boldsymbol{x}}_s|)\right|}, & \text{for } d = 2, \\ \dfrac{\max_r \left|\vec{\boldsymbol{x}}_r - \vec{\boldsymbol{x}}_s\right|}{\min_r \left|\vec{\boldsymbol{x}}_r - \vec{\boldsymbol{x}}_s\right|}, & \text{for } d = 3. \end{cases} \tag{4.13}$$

In Figure 4.2 we show the condition number of $\mathsf{M}(\vec{\boldsymbol{x}}_s, \omega)$ plotted over an optical frequency band. The experimental setup is that given in §4.5. The condition number

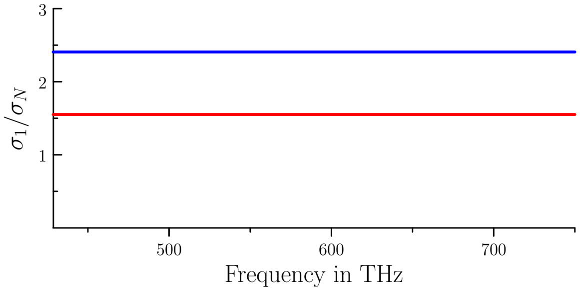


Figure 4.2. Condition number of $\mathsf{M}(\vec{\boldsymbol{x}}_s, \omega)$ for $d = 2$ (red) and $d = 3$ (blue) for the setup given in §4.5

(4.13) is clearly independent of frequency for $d = 3$, and for $d = 2$ we have the approximation for high frequencies:

$$\operatorname{cond} \mathsf{M}(\vec{\boldsymbol{x}}_s, \omega) = \frac{\max_r |\vec{\boldsymbol{x}}_r - \vec{\boldsymbol{x}}_s|^{1/2}}{\min_r |\vec{\boldsymbol{x}}_r - \vec{\boldsymbol{x}}_s|^{1/2}}(1 + \mathcal{O}(1/\omega)), \quad \text{as } \omega \to \infty.$$

This approximation follows from the Hankel function asymptotic (see, e.g., [7])

$$H_0^{(1)}(t) = \sqrt{\frac{2}{\pi t}} \exp[\imath(t - \pi/4)](1 + \mathcal{O}(1/t)), \quad \text{as } t \to \infty.$$

Thus the conditioning of M is determined by the ratio of largest to smallest source-to-receiver distances.

4.3.2 Kirchhoff migration

We now show that migrating the recovered data $\widetilde{\boldsymbol{p}}$ (4.11) using Γ_{KM} gives essentially the same image as migrating the true data $\boldsymbol{p}$. We establish this result by means of a stationary phase argument but in order to do this, we need the following assumption on the location of the source $\vec{\boldsymbol{x}}_s$.

Assumption 5 (Geometric imaging conditions) *For a scattering potential with support contained inside an image window $\mathcal{W}$, we assume $\vec{\boldsymbol{x}}_s$ satisfies*

$$\frac{\vec{\boldsymbol{x}}_r - \vec{\boldsymbol{x}}_s}{|\vec{\boldsymbol{x}}_r - \vec{\boldsymbol{x}}_s|} \neq \frac{\vec{\boldsymbol{x}}_r - \vec{\boldsymbol{y}}}{|\vec{\boldsymbol{x}}_r - \vec{\boldsymbol{y}}|},$$

for $r = 1, \ldots, N$, and $\vec{\boldsymbol{y}} \in \mathcal{W}$.

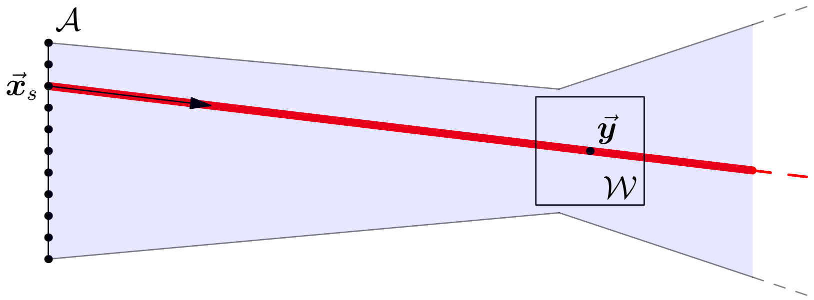


Figure 4.3. Illustration of assumption 5. If $\vec{\boldsymbol{x}}_s$ is outside of the light blue region then $(\vec{\boldsymbol{x}}_r - \vec{\boldsymbol{x}}_s)/|\vec{\boldsymbol{x}}_r - \vec{\boldsymbol{x}}_s| \neq (\vec{\boldsymbol{x}}_r - \vec{\boldsymbol{y}})/|\vec{\boldsymbol{x}}_r - \vec{\boldsymbol{y}}|$ for all $\vec{\boldsymbol{x}}_r \in \mathcal{A}$ and $\vec{\boldsymbol{y}} \in \mathcal{W}$.

We interpret this assumption as a restriction on the placement of our source location $\vec{\boldsymbol{x}}_s$ as follows. Fix a receiver position $\vec{\boldsymbol{x}}_r$ and consider the cone

$$\mathcal{K}(\vec{\boldsymbol{x}}_r) = \left\{ \alpha \frac{\vec{\boldsymbol{y}} - \vec{\boldsymbol{x}}_r}{|\vec{\boldsymbol{y}} - \vec{\boldsymbol{x}}_r|} : \alpha > 0, \vec{\boldsymbol{y}} \in \mathcal{W} \right\}.$$

As long as $\vec{\boldsymbol{x}}_s \notin \mathcal{K}(\vec{\boldsymbol{x}}_r)$, then we have that $(\vec{\boldsymbol{x}}_s - \vec{\boldsymbol{x}}_r)/|\vec{\boldsymbol{x}}_s - \vec{\boldsymbol{x}}_r| \neq (\vec{\boldsymbol{y}} - \vec{\boldsymbol{x}}_r)/|\vec{\boldsymbol{y}} - \vec{\boldsymbol{x}}_r|$ for any $\vec{\boldsymbol{y}} \in \mathcal{W}$, i.e., assumption 5 holds for $\vec{\boldsymbol{x}}_r$. Ensuring this is satisfied for all receiver locations $\vec{\boldsymbol{x}}_r$ for $r = 1, \ldots, N$, we require $\vec{\boldsymbol{x}}_s \notin \cup_{r=1}^{N} \mathcal{K}(\vec{\boldsymbol{x}}_r)$. In Figure 4.3 we illustrate this assumption. Here, the dark blue region depicts the cone $\mathcal{K}(\vec{\boldsymbol{x}}_r)$ while the union of cones $\cup_{r=1}^{N} \mathcal{K}(\vec{\boldsymbol{x}}_r)$ is depicted by the light blue region. Assumption 5 simply requires $\vec{\boldsymbol{x}}_s$ to be outside the light blue region.

Theorem 6 *Provided assumption 5 holds, the image of the reconstructed array response vector is*

$$\Gamma_{KM}\big[\boldsymbol{p} + (\overline{\boldsymbol{g}}_0)^{-1} \odot \boldsymbol{g}_0 \odot \overline{\boldsymbol{p}}, \omega\big](\vec{\boldsymbol{y}}) \approx \Gamma_{KM}\big[\boldsymbol{p}, \omega\big](\vec{\boldsymbol{y}}).$$

Proof. We begin by approximating the Kirchhoff migration functional (4.6) by an integral over the array $\mathcal{A}$:

$$\begin{aligned}
&\Gamma_{\mathrm{KM}}\left[(\overline{\boldsymbol{g}}_0)^{-1} \odot \boldsymbol{g}_0 \odot \overline{\boldsymbol{p}}\right](\vec{\boldsymbol{y}}) \\
&\quad = \overline{\widehat{G}}_0(\vec{\boldsymbol{x}}_s, \vec{\boldsymbol{y}}, \omega)\boldsymbol{g}_0(\vec{\boldsymbol{y}}, \omega)^* \left[(\overline{\boldsymbol{g}}_0(\vec{\boldsymbol{x}}_s, \omega))^{-1} \odot \boldsymbol{g}_0(\vec{\boldsymbol{x}}_s, \omega) \odot \overline{\boldsymbol{p}}(\vec{\boldsymbol{x}}_s, \omega)\right] \\
&\quad \sim k^2 \int_{\mathcal{P}} d\vec{\boldsymbol{z}} \int_{\mathcal{A}} d\boldsymbol{x}_r C(\vec{\boldsymbol{x}}_s, \vec{\boldsymbol{x}}_r, \vec{\boldsymbol{z}}, \vec{\boldsymbol{y}}) \times \\
&\qquad \exp\left(\imath k(2|\vec{\boldsymbol{x}}_r - \vec{\boldsymbol{x}}_s| - |\vec{\boldsymbol{x}}_r - \vec{\boldsymbol{z}}| - |\vec{\boldsymbol{z}} - \vec{\boldsymbol{x}}_s| - |\vec{\boldsymbol{x}}_r - \vec{\boldsymbol{y}}| - |\vec{\boldsymbol{y}} - \vec{\boldsymbol{x}}_s|)\right).
\end{aligned} \tag{4.14}$$

Here $\sim$ denotes equal up to a constant and $C(\vec{\boldsymbol{x}}_s, \vec{\boldsymbol{x}}_r, \vec{\boldsymbol{z}}, \vec{\boldsymbol{y}})$ is a smooth real valued function that collects the various $\widehat{G}_0$ geometric spreading terms.

Now we apply the method of stationary phase (see, e.g., [2]) to the integral over $\mathcal{A}$. In the large wavenumber limit $k \to \infty$, dominant contributions to the integral come from stationary points of the phase, i.e., points $\vec{\boldsymbol{x}}_r$ that satisfy

$$\nabla_{\vec{\boldsymbol{x}}_r} \left(2|\vec{\boldsymbol{x}}_r - \vec{\boldsymbol{x}}_s| - |\vec{\boldsymbol{x}}_r - \vec{\boldsymbol{z}}| - |\vec{\boldsymbol{z}} - \vec{\boldsymbol{x}}_s| - |\vec{\boldsymbol{x}}_r - \vec{\boldsymbol{y}}| - |\vec{\boldsymbol{y}} - \vec{\boldsymbol{x}}_s|\right) = 0.$$

This expression is equivalent to

$$\frac{\vec{\boldsymbol{x}}_r - \vec{\boldsymbol{x}}_s}{|\vec{\boldsymbol{x}}_r - \vec{\boldsymbol{x}}_s|} = \frac{1}{2}\left(\frac{\vec{\boldsymbol{x}}_r - \vec{\boldsymbol{y}}}{|\vec{\boldsymbol{x}}_r - \vec{\boldsymbol{y}}|} + \frac{\vec{\boldsymbol{x}}_r - \vec{\boldsymbol{z}}}{|\vec{\boldsymbol{x}}_r - \vec{\boldsymbol{z}}|}\right). \tag{4.15}$$

If (4.15) holds then we must have

$$\begin{aligned} \left|\frac{\vec{\boldsymbol{x}}_r - \vec{\boldsymbol{x}}_s}{|\vec{\boldsymbol{x}}_r - \vec{\boldsymbol{x}}_s|}\right|^2 &= \frac{1}{4}\left|\frac{\vec{\boldsymbol{x}}_r - \vec{\boldsymbol{y}}}{|\vec{\boldsymbol{x}}_r - \vec{\boldsymbol{y}}|} + \frac{\vec{\boldsymbol{x}}_r - \vec{\boldsymbol{z}}}{|\vec{\boldsymbol{x}}_r - \vec{\boldsymbol{z}}|}\right|^2 \\ \iff 1 &= \frac{\vec{\boldsymbol{x}}_r - \vec{\boldsymbol{y}}}{|\vec{\boldsymbol{x}}_r - \vec{\boldsymbol{y}}|} \cdot \frac{\vec{\boldsymbol{x}}_r - \vec{\boldsymbol{z}}}{|\vec{\boldsymbol{x}}_r - \vec{\boldsymbol{z}}|}. \end{aligned} \tag{4.16}$$

Since $(\vec{\boldsymbol{x}}_r - \vec{\boldsymbol{y}})/|\vec{\boldsymbol{x}}_r - \vec{\boldsymbol{y}}|$ and $(\vec{\boldsymbol{x}}_r - \vec{\boldsymbol{z}})/|\vec{\boldsymbol{x}}_r - \vec{\boldsymbol{z}}|$ are both unit vectors, it follows from the Cauchy-Schwarz *equality* that (4.16) holds only if $\vec{\boldsymbol{z}} = \vec{\boldsymbol{y}}$. Thus stationary points must satisfy

$$\frac{\vec{\boldsymbol{x}}_r - \vec{\boldsymbol{x}}_s}{|\vec{\boldsymbol{x}}_r - \vec{\boldsymbol{x}}_s|} = \frac{\vec{\boldsymbol{x}}_r - \vec{\boldsymbol{y}}}{|\vec{\boldsymbol{x}}_r - \vec{\boldsymbol{y}}|},$$

where $\vec{\boldsymbol{y}} \in \mathcal{W}$. By assumption 5 there are no such stationary points and therefore, neglecting boundary effects, this integral vanishes faster than any polynomial power of ω (see, e.g., [3]). ■

Remark 11 *We used a similar idea in [1] to show that with multiple sources, a single receiver, and a specific pairwise illumination scheme it is possible to image with sole knowledge of the intensities of the wave fields at the receiver and of the probing fields. We approached the problem by estimating the array response vector (a vector in $\mathbb{C}^N$) with $2N$ (or more) real measurements, which are essentially the measured intensities for $2N$ or more different pairs of sources. The results of §4.3.1 and §4.3.2 can be modified by reciprocity to apply to the setup we considered in [1]. Hence images similar to those in [1] can be obtained without the pairwise illumination scheme and the number of required illuminations is reduced from $3N$ to N.*

4.4 Stochastic illuminations and autocorrelations

Our imaging method can also be used when the source is driven by a stationary stochastic process (for which we only assume knowledge of the autocorrelation or power spectra), and only empirical autocorrelations are measured at the receiver locations.

To be more precise, the source at $\vec{\boldsymbol{x}}_s$ is driven by $f(t)$, a stationary mean zero Gaussian process with autocorrelation function

$$\langle \overline{f}(t) f(t+\tau) \rangle = F(\tau). \tag{4.17}$$

Here $\langle \cdot \rangle$ denotes expectation with respect to realizations of f and we recall that $F(\tau) = \overline{F}(-\tau)$. In the time domain, the field recorded at $\vec{\boldsymbol{x}}_r$ is

$$u(\vec{\boldsymbol{x}}_r, \vec{\boldsymbol{x}}_s, t) = \frac{1}{2\pi} \int d\omega e^{-\imath\omega t} \widehat{G}(\vec{\boldsymbol{x}}_r, \vec{\boldsymbol{x}}_s, \omega) \widehat{f}(\omega), \qquad \text{for } r = 1, \ldots, N,$$

where we assume $\widehat{G}$ is given by the Born approximation (4.3).

The measurements at the receiver locations $\vec{\boldsymbol{x}}_r$ are the empirical autocorrelations:

$$\psi(\vec{\boldsymbol{x}}_r, \vec{\boldsymbol{x}}_s, \tau) = \frac{1}{2T} \int_{-T}^{T} dt \overline{u}(\vec{\boldsymbol{x}}_r, \vec{\boldsymbol{x}}_s, t) u(\vec{\boldsymbol{x}}_r, \vec{\boldsymbol{x}}_s, t+\tau) \qquad \text{for } r = 1, \ldots, N, \tag{4.18}$$

where T is a fixed acquisition time. As shown by Garnier and Papanicolaou [4], these measurements are independent of the acquisition time T and ergodic as we summarize in the following proposition.

Proposition 3 *Assume $f(t)$ is a stationary mean zero Gaussian process satisfying* (4.17). *The expectation of the empirical autocorrelations* (4.18) *is independent of the acquisition time T:*

$$\langle \psi(\vec{\boldsymbol{x}}_r, \vec{\boldsymbol{x}}_s, \tau) \rangle = \Psi(\vec{\boldsymbol{x}}_r, \vec{\boldsymbol{x}}_s, \tau),$$

where

$$\Psi(\vec{\boldsymbol{x}}_r, \vec{\boldsymbol{x}}_s, \tau) = \frac{1}{2\pi} \int d\omega e^{-\imath\omega\tau} \widehat{F}(\omega) \boldsymbol{e}_r^{\mathsf{T}} \left[\overline{\boldsymbol{g}}(\vec{\boldsymbol{x}}_s, \omega) \odot \boldsymbol{g}(\vec{\boldsymbol{x}}_s, \omega) \right], \tag{4.19}$$

with $\boldsymbol{g} \equiv \boldsymbol{g}_0 + \boldsymbol{p}$. Furthermore, (4.18) *is ergodic, i.e.,*

$$\psi(\vec{\boldsymbol{x}}_r, \vec{\boldsymbol{x}}_s, \tau) \xrightarrow{T\to\infty} \Psi(\vec{\boldsymbol{x}}_r, \vec{\boldsymbol{x}}_s, \tau). \tag{4.20}$$

Proof. The proof is a straightforward application of [4, Proposition 4.1]. ■

The ergodicity (4.20) of this proposition guarantees that for sufficiently large acquisition time T, the autocorrelation $\psi(\vec{\boldsymbol{x}}_r, \vec{\boldsymbol{x}}_s, \tau)$ is close to an intensity measurement, i.e.,

$$\widehat{\psi}(\vec{\boldsymbol{x}}_r, \vec{\boldsymbol{x}}_s, \omega) \xrightarrow{T\to\infty} \widehat{\Psi}(\vec{\boldsymbol{x}}_r, \vec{\boldsymbol{x}}_s, \omega) = \widehat{F}(\omega)\boldsymbol{e}_r^{\mathsf{T}} \left[\overline{(\boldsymbol{g}_0 + \boldsymbol{p})} \odot (\boldsymbol{g}_0 + \boldsymbol{p})\right].$$

Proceeding analogously as in §4.2.2, we neglect the quadratic term in $\boldsymbol{p}$:

$$\widehat{\Psi}(\vec{\boldsymbol{x}}_r, \vec{\boldsymbol{x}}_s, \omega) \approx \widehat{F}(\omega)\boldsymbol{e}_r^{\mathsf{T}} \mathrm{Re}\left(\overline{\boldsymbol{g}}_0 \odot (\boldsymbol{g}_0 + 2\boldsymbol{p})\right).$$

The collection of autocorrelations for $r = 1, \ldots, N$ can be expressed, approximately, as

$$\left[\widehat{\Psi}(\vec{\boldsymbol{x}}_r, \vec{\boldsymbol{x}}_s, \omega)\right]_{r=1,\ldots,N} \approx \boldsymbol{d}(\vec{\boldsymbol{x}}_s, \omega) \equiv \widehat{F}(\omega)\mathsf{M}(\vec{\boldsymbol{x}}_s, \omega) \begin{bmatrix}\mathrm{Re}(\boldsymbol{g}_0 + 2\boldsymbol{p}) \\ \mathrm{Im}(\boldsymbol{g}_0 + 2\boldsymbol{p})\end{bmatrix},$$

where $\mathsf{M}(\vec{\boldsymbol{x}}_s, \omega) \in \mathbb{R}^{N\times 2N}$ is given by (4.9). Therefore, the techniques developed in §4.3 can be applied to image from the autocorrelation measurements (4.18).

4.5 Numerical experiments

We now provide 2D numerical experiments of our proposed imaging method. The physical scalings we use correspond to an optic regime. We use the background wave velocity of $c_0 = 3 \times 10^8$ m/s and central frequency of about 590 THz, which gives a central wavelength λ_0 of about 509 nm. Our receiver array $\mathcal{A}$ is a linear array centered at the origin and consists of 501 receivers located at coordinates $\vec{\boldsymbol{x}}_r = (0, -5 + (r-1)(10/500))$mm for $r = 1, \ldots, 501$. This corresponds to using a 1 cm linear array of receivers spaced approximately 20μm apart. We place the wave source at coordinate $\vec{\boldsymbol{x}}_s = (5, -7.5)$mm to guarantee assumption 5 is satisfied. We begin with experiments in the deterministic setting (§4.5.1) followed by an experiment in the stochastic setting (§4.5.2). Lastly, we investigate situations where assumptions 4 and/or 5 are violated and our method is not expected to work (§4.5.3). For all experiments, we assume 3D wave propagation for simplicity so that $\widehat{G}_0$ is given by (3.3) for $d = 3$.

4.5.1 Deterministic illuminations

Using the illumination $\widehat{f}(\omega) \equiv 1$, we generate the intensity data $\boldsymbol{d}(\vec{\boldsymbol{x}}_s, \omega)$ using the Born approximation:

$$\boldsymbol{d}(\vec{\boldsymbol{x}}_s, \omega) = \overline{(\boldsymbol{g}_0 + \boldsymbol{p})} \odot (\boldsymbol{g}_0 + \boldsymbol{p}),$$

with $\boldsymbol{g}_0$ and $\boldsymbol{p}$ defined by (4.4) and (4.5), respectively, for 100 uniformly spaced frequencies in the frequency band $[430, 750]$ THz. This corresponds to obtaining intensity data for 100 different monochromatic illuminations with wavelengths $\lambda \in [400, 700]$ nm, equally spaced in the frequency band. Note that the quadratic term $\overline{\boldsymbol{p}} \odot \boldsymbol{p}$ is present in our data, but using assumption 4 we proceed assuming $\boldsymbol{d}$ is well approximated by the linear system (4.8).

We recover the approximate array response vector $\widetilde{\boldsymbol{p}} = (\overline{\boldsymbol{g}}_0)^{-1} \odot \boldsymbol{d} - \boldsymbol{g}_0$ for each frequency ω in the angular frequency band $\mathcal{B}$, where $(2\pi)^{-1}\mathcal{B} = [430, 750]$ THz. An image is then formed using the Kirchhoff migration functional integrated over $\mathcal{B}$:

$$\Gamma_{\text{KM}}[\widetilde{\boldsymbol{p}}](\vec{\boldsymbol{y}}) = \int_{\mathcal{B}} d\omega \Gamma_{\text{KM}}[\widetilde{\boldsymbol{p}}, \omega](\vec{\boldsymbol{y}}),$$

where Γ_{KM} is defined in (4.6). Here we consider image points $\vec{\boldsymbol{y}} \in \mathcal{W} = \{(50\text{mm} + i\lambda_0/2.5, j\lambda_0/2.5), \text{ for } i, j = -25, \ldots, 25\}$.

For our first experiment, we consider a point reflector located at coordinate $\vec{\boldsymbol{y}} = (50, 0)$mm with refractive index perturbation $\rho(\vec{\boldsymbol{y}}) = 1 \times 10^{-15}$ (roughly equivalent to a reflector of area $(\lambda_0)^2$ and reflectivity 5978). The migrated images of the true array response vector $\boldsymbol{p}$ and the recovered array response vector $\widetilde{\boldsymbol{p}}$ are shown in Figure 4.4a. Although we are significantly undersampling both in frequency and on the array (recall the spacing between receivers is approx. $20\mu\text{m} \gg \lambda_0/2$), the images still exhibit the cross-range (Rayleigh) resolution estimate $\lambda_0 L/a \approx 5\lambda_0$ and range resolution estimate $c_0/|\mathcal{B}| \approx 1\lambda_0$. Our second experiment (Figure 4.4b) uses two point reflectors located at coordinates $\vec{\boldsymbol{y}}_1 = (50\text{mm} - 3\lambda_0, -\lambda_0)$ and $\vec{\boldsymbol{y}}_2 = (50\text{mm} + 6\lambda_0, 5\lambda_0)$, each with $\rho(\vec{\boldsymbol{y}}_i) = 1 \times 10^{-15}$. We show an extended scatterer (a disk) in Figure 4.5. The disk is generated as a set of point reflectors, each with $\rho(\vec{\boldsymbol{y}}_i) = 1 \times 10^{-15}$ separated by $\lambda_0/4$.

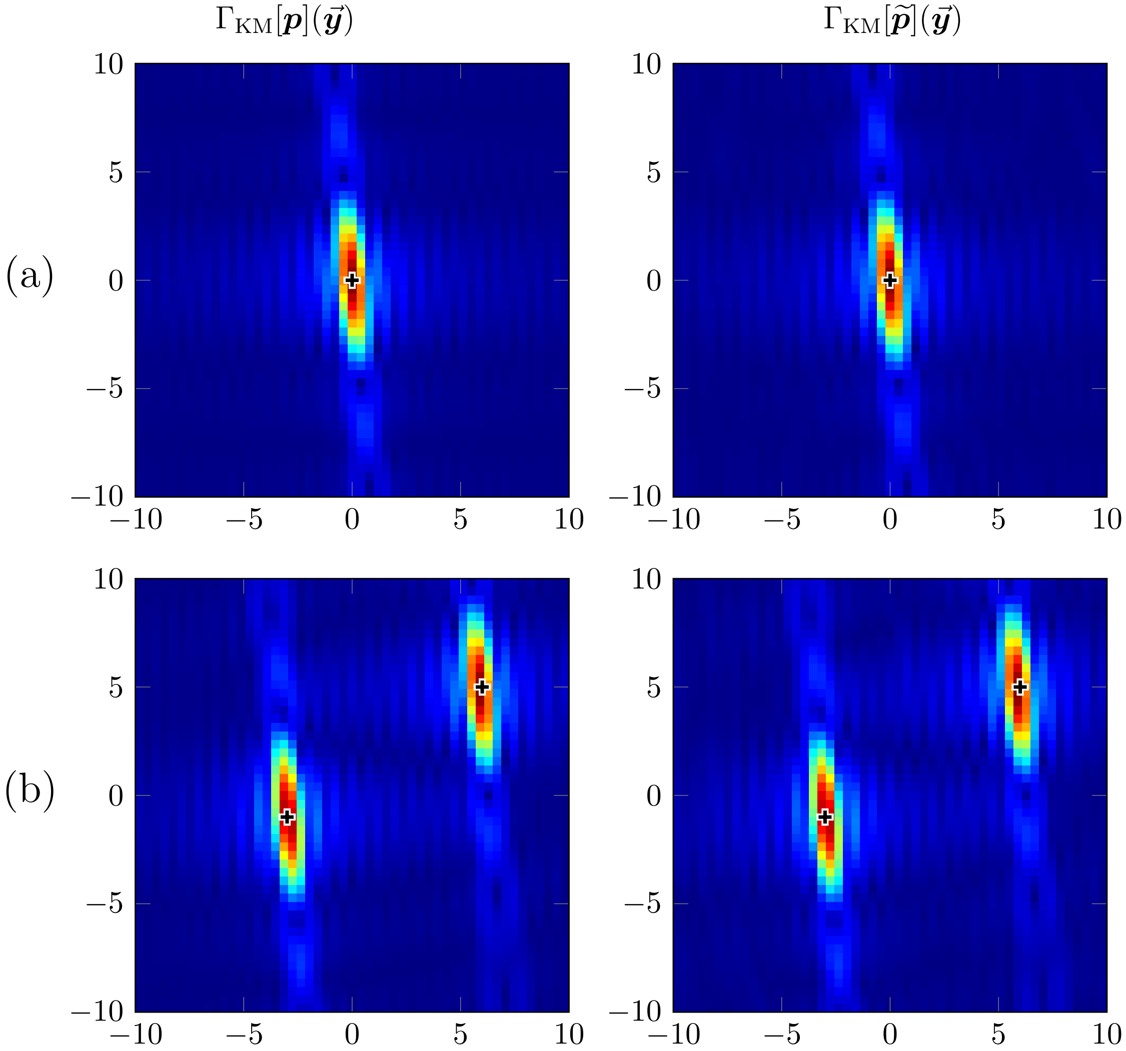


Figure 4.4. Kirchhoff images of (a) one and (b) two point scatterers whose true locations are indicated by crosses. The left column uses the full array response vector $\boldsymbol{p}$ while the right column uses the array response vector $\widetilde{\boldsymbol{p}}$ recovered from intensity data. The horizontal and vertical axes display the range and cross-range respectively, measured in central wavelengths λ_0 from $(50, 0)$mm.

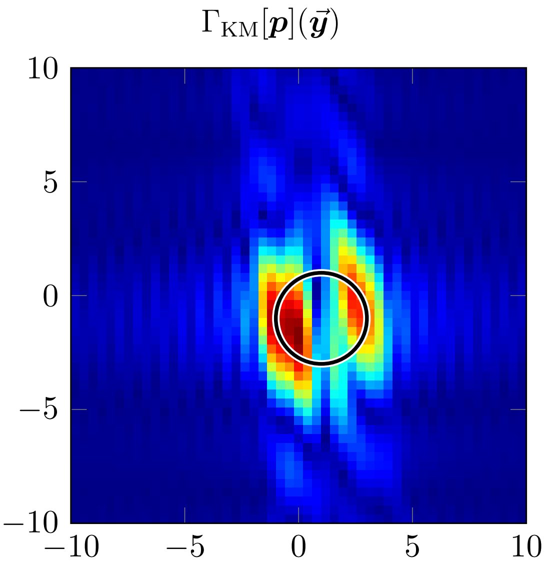


Figure 4.5. Kirchhoff images of an extended scatterer (disk). The boundary of the disk is indicated by the black and white circle. The left image uses the true array response vector $\boldsymbol{p}$ while the right image uses the array response vector $\widetilde{\boldsymbol{p}}$ recovered from intensity measurements. The horizontal and vertical axes display the range and cross-range respectively, measured in central wavelengths λ_0 from $(50, 0)$mm.

4.5.2 Stochastic illumination

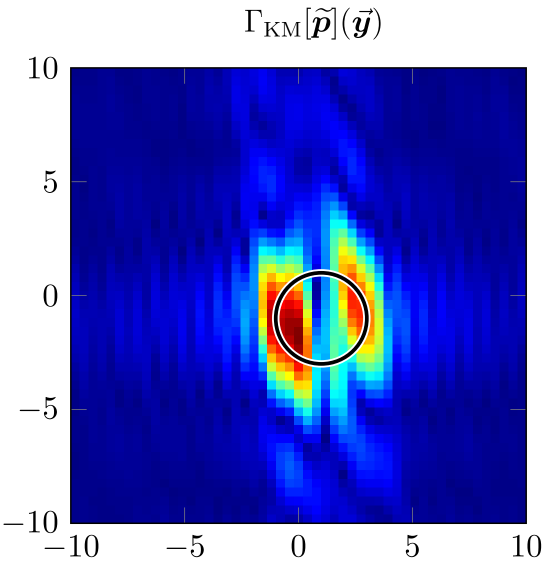
Here we image with power spectrum data $\boldsymbol{d}$ generated from a stochastic illumination as in §4.4. Since we work in an optic regime, adequately sampling signals in the time domain and performing the autocorrelations (4.18) is an expensive calculation. We instead use the Wiener-Khinchin theorem [5] to simulate power spectrum measurements directly.

We assume the wave source at $\vec{\boldsymbol{x}}_s$ is driven by a stationary mean zero Gaussian process $f(t)$ with correlation function $\langle \overline{f(t)} f(t+\tau)\rangle = F(\tau)$. By the Wiener-Khinchin theorem, $\widehat{f}(\omega)$ is a mean zero Gaussian process with correlation function

$$\langle \overline{\widehat{f}(\omega)} \widehat{f}(\omega')\rangle = 2\pi\delta(\omega-\omega')\widehat{F}(\omega). \tag{4.21}$$

Thus frequency samples of $\widehat{f}(\omega)$ are *independent* normal random variables with variance proportional to $\widehat{F}(\omega)$. Here we use

$$\widehat{F}(\omega) = t_c \exp\left(\frac{-(\omega-\omega_0)^2}{4\pi/t_c^2}\right), \tag{4.22}$$



where $(2\pi)^{-1}\omega_0 = 590$ THz is the central frequency and $t_c = 150 \times 10^{-12}$ sec is the correlation time of $f(t)$ (i.e., $F(\tau) \approx 0$ for $\tau \gg t_c$). This choice of t_c gives the signal an effective frequency band of $(2\pi)^{-1}\mathcal{B} = [430, 750]$ THz (i.e., $\widehat{F}(\omega) \approx 0$ for $\omega \notin \mathcal{B}$). Using (4.21) and (4.22) we generate frequency samples $\widehat{f}(\omega_i)$ for 100 frequencies ω_i equally spaced in $\mathcal{B}$.

For a large enough acquisition time T, the empirical autocorrelations (4.18) give frequency domain measurements proportional to

$$\widehat{\psi}(\vec{\boldsymbol{x}}_r, \omega) = \left|\boldsymbol{e}_r^{\mathsf{T}}(\boldsymbol{g}_0 + \boldsymbol{p})\widehat{f}(\omega)\right|^2 \quad \text{for } r = 1, \ldots, N.$$

Thus for each frequency $\omega_i \in \mathcal{B}$ we generate the power spectrum data

$$\boldsymbol{d}(\vec{\boldsymbol{x}}_s, \omega_i) = \overline{\Big((\boldsymbol{g}_0 + \boldsymbol{p})\widehat{f}(\omega_i)\Big)} \odot \Big((\boldsymbol{g}_0 + \boldsymbol{p})\widehat{f}(\omega_i)\Big). \tag{4.23}$$

Because correlations are robust with respect to additive noise, we also consider autocorrelations with additive noise:

$$\widehat{\psi}(\vec{\boldsymbol{x}}_r, \omega) = \left|\boldsymbol{e}_r^{\mathsf{T}}(\boldsymbol{g}_0 + \boldsymbol{p})\widehat{f}(\omega) + \widehat{\eta}_r(\omega)\right|^2 \quad \text{for } r = 1, \ldots, N,$$

where the noise $\widehat{\eta}_r(\omega)$ is an *independent* mean zero Gaussian process with correlation function given by (4.21) and (4.22) for each $r = 1, \ldots, N$. Here we set the noise power equal to 10% of the signal power at each receiver, i.e.,

$$\int d\omega |\widehat{\eta}_r(\omega)|^2 = \frac{1}{10}\int d\omega |\boldsymbol{e}_r^{\mathsf{T}}(\boldsymbol{g}_0 + \boldsymbol{p})\widehat{f}(\omega)|^2, \quad \text{for } r = 1, \ldots, N.$$

Noisy data for each frequency $\omega_i \in \mathcal{B}$ is then generated as

$$\boldsymbol{d}(\vec{\boldsymbol{x}}_s, \omega_i) = \overline{\Big((\boldsymbol{g}_0 + \boldsymbol{p})\widehat{f}(\omega_i) + \widehat{\boldsymbol{\eta}}(\omega_i)\Big)} \odot \Big((\boldsymbol{g}_0 + \boldsymbol{p})\widehat{f}(\omega_i) + \widehat{\boldsymbol{\eta}}(\omega_i)\Big), \tag{4.24}$$

where $\widehat{\boldsymbol{\eta}}(\omega) = \left[\widehat{\eta}_1(\omega), \cdots, \widehat{\eta}_N(\omega)\right]^{\mathsf{T}}$. We can indeed consider noise with much larger power (e.g., noise power equal to 100% signal power), but to compensate we then need additional frequency samples to maintain sufficient averaging in migration images. In Figure 4.6 we show the migrated images $\Gamma_{\mathrm{KM}}[\widetilde{\boldsymbol{p}}]$ for $\widetilde{\boldsymbol{p}} = (2\pi\widehat{F}\overline{\boldsymbol{g}}_0)^{-1} \odot \boldsymbol{d} - \boldsymbol{g}_0$ recovered from clean data (4.23) and from noisy data (4.24).

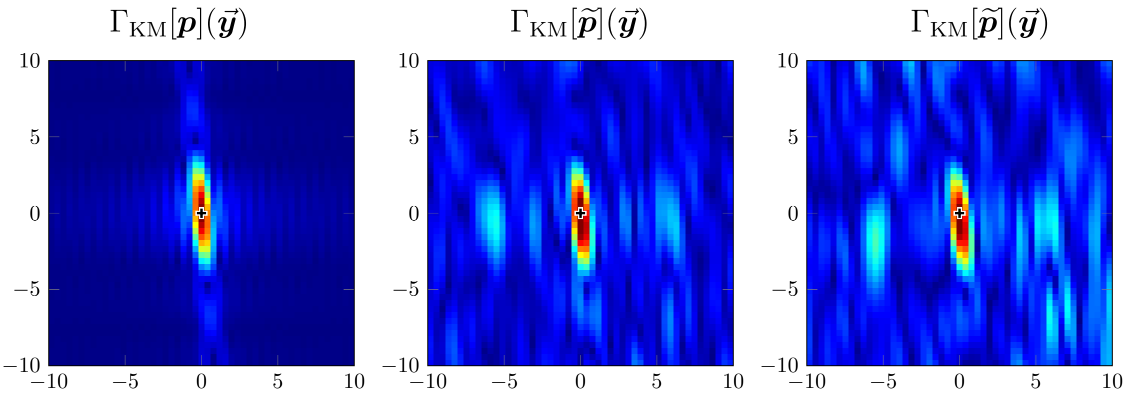


Figure 4.6. Kirchhoff images of a point scatterer using a stochastic illumination and autocorrelation measurements. The images are generated using (left) the true array response vector $\boldsymbol{p}$, (center) $\widetilde{\boldsymbol{p}}$ recovered from clean power spectrum data (4.23) and (right) $\widetilde{\boldsymbol{p}}$ recovered from noisy power spectrum data (4.24). The horizontal and vertical axes display the range and cross-range respectively, measured in central wavelengths λ_0 from $(50, 0)$mm.

4.5.3 Breakdown of the method

We now investigate situations where assumptions 4 and/or 5 are violated. For these experiments, we fix the receiver array $\mathcal{A}$ (again consisting of 501 receivers with locations $\vec{\boldsymbol{x}}_r$ given above) while varying the source position $\vec{\boldsymbol{x}}_s$, the reflector location $\vec{\boldsymbol{y}}$ and the reflectivity $\rho(\vec{\boldsymbol{y}})$.

In Table 4.1, we indicate a few situations where our imaging method breaks down. The corresponding migrated images of the recovered array response vector $\Gamma_{\mathrm{KM}}[\widetilde{\boldsymbol{p}}]$ are shown in Figure 4.7. From Figures 4.7a, 4.7b and 4.7c we see the imaging method is most sensitive to breaking assumption 4. In these situations the quadratic term $\overline{\boldsymbol{p}} \odot \boldsymbol{p}$ cannot be neglected in the intensity data (4.7) and thus the linear system we

Table 4.1. Breakdown of imaging method: physical setups and violated assumptions for each of the numerical experiments shown in Figure 4.7.

	Assumptions violated	$\vec{\boldsymbol{y}}$	$\rho(\vec{\boldsymbol{y}})$	$\vec{\boldsymbol{x}}_s$
(a) Source near scatterer	4 and 5	$(50\text{mm}, 0)$	10^{-15}	$(50\text{mm} - 10\lambda_0, 0)$
(b) Receivers near scatterer	4	$(11\lambda_0, 0)$	10^{-15}	$(-50\text{mm}, 0)$
(c) Large reflectivity	4	$(50\text{mm}, 0)$	10^{-10}	$(5, -75)$mm
(d) No geometric condition	5	$(50\text{mm}, 0)$	10^{-15}	$(5\text{mm}, 0)$

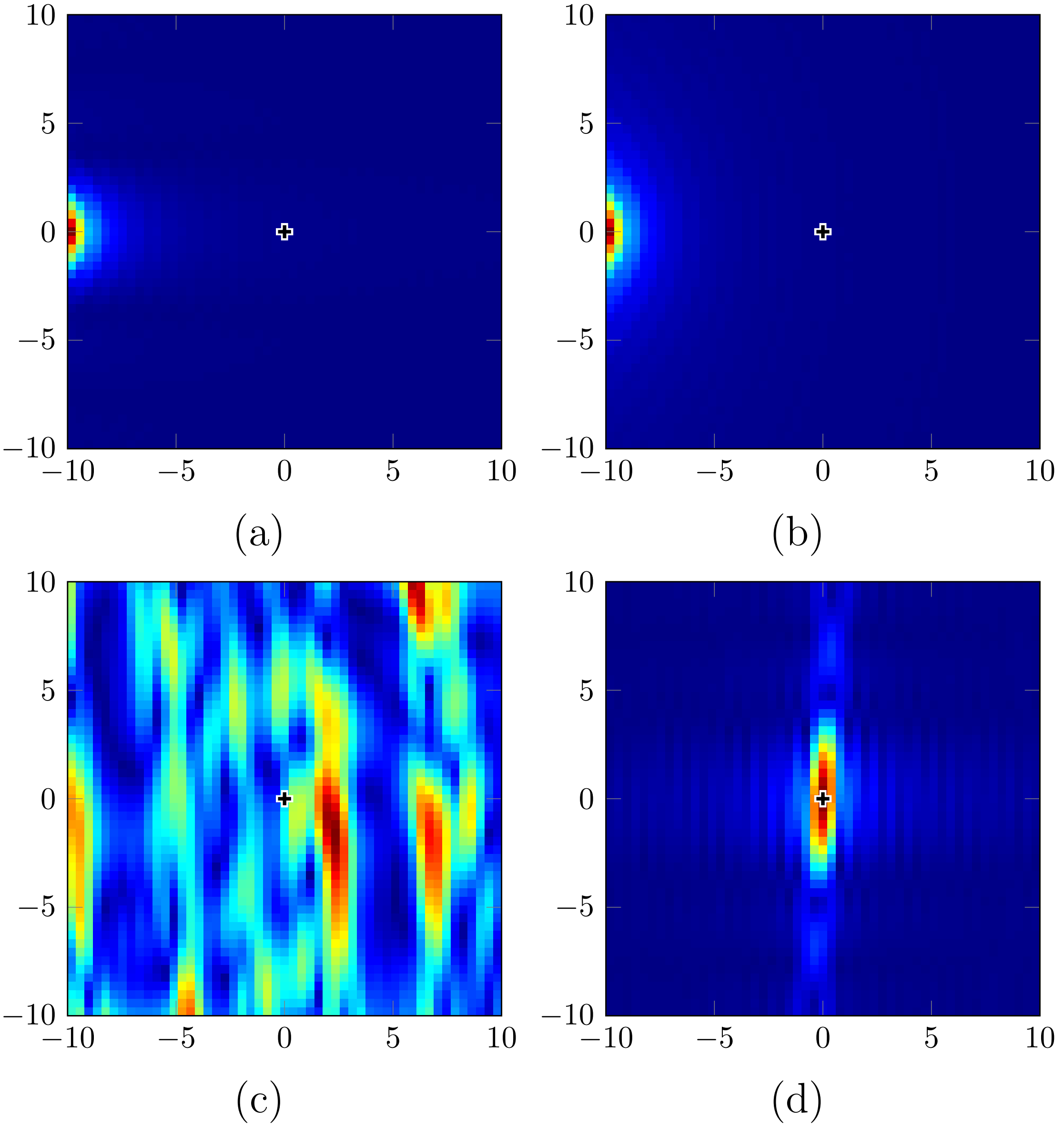


Figure 4.7. Breakdown of imaging method: migrated images $\Gamma_{\mathrm{KM}}[\widetilde{\boldsymbol{p}}]$ for setups violating assumptions 4 and/or 5. Details of each setup are listed above and correspond to (a) $\vec{\boldsymbol{x}}_s$ placed to close to $\vec{\boldsymbol{y}}$, (b) $\mathcal{A}$ placed too close to $\vec{\boldsymbol{y}}$, (c) large reflectivity ρ and (d) $\vec{\boldsymbol{x}}_s$ placed in front of the array $\mathcal{A}$. The axes are measured in central wavelengths λ_0 from the scatterer's true location $\vec{\boldsymbol{y}}$ which is indicated by a cross.

consider in (4.8) is no longer a good approximation. This leads to artifacts in the images. Figure 4.7d demonstrates the imaging method is more robust than expected with respect to assumption 5 and the position of $\vec{\boldsymbol{x}}_s$.

4.6 Discussion and future work

We have shown that when the scattered field is small compared to the incident field (assumption 4), one can consider the problem of recovering full-waveform data from intensity measurements as a linear least-squares problem. For N receivers, the corresponding real matrix is $N \times 2N$, so all we can expect to recover is the projection of the real and imaginary parts of the full-waveform data onto an N dimensional subspace. This turns out to be sufficient to image with Kirchhoff migration (theorem 6). Crucially, we do not need to manipulate the fields at the receiver end, e.g., to introduce phases. The least-squares problems we obtain are usually well conditioned and the computational cost of solving them ($\mathcal{O}(N)$ complex operations, each) is negligible compared to the cost of Kirchhoff migration. Since we make no assumptions on the source phases, our method adapts well to situations where the source is driven by a Gaussian process and the measurements are autocorrelations at the receiver locations.

The fundamental principle we have used here is that the imaging method (in this case Kirchhoff migration) does not require all the data (in this case the full-waveform scattered field) to form an image. For Kirchhoff migration this is exploited, e.g., by undersampling in frequency and/or using only a few sources or receivers to image. We have shown that there is another way in which one can use incomplete data, as projections of the array response vector on certain subspaces leave the Kirchhoff images unaffected. A similar principle is what is exploited by Novikov et al. [6] to image with intensities, since they show that knowing inner products of single-source experiments is enough to image with MUSIC. It would be interesting to carry this idea further and see whether the same preprocessing we use here works for MUSIC and also whether it is possible to image scatterers with even fewer data.

4.7 Acknowledgements

The work of P. Bardsley and F. Guevara Vasquez was partially supported by the National Science Foundation grant DMS-1411577.

4.8 References

[1] P. Bardsley and F. Guevara Vasquez, *Imaging with power controlled source pairs*, SIAM Journal on Imaging Sciences, 9 (2016), pp. 185–211.

[2] N. Bleistein, J. K. Cohen, and J. W. Stockwell, Jr., *Mathematics of multidimensional seismic imaging, migration, and inversion*, vol. 13 of Interdisciplinary Applied Mathematics, Springer-Verlag, New York, 2001.

[3] N. Bleistein and R. A. Handelsman, *Asymptotic expansions of integrals*, Dover Publications, Inc., New York, second ed., 1986.

[4] J. Garnier and G. Papanicolaou, *Passive sensor imaging using cross correlations of noisy signals in a scattering medium*, SIAM J. Imaging Sci., 2 (2009), pp. 396–437.

[5] A. Ishimaru, *Wave propagation and scattering in random media*, IEEE/OUP Series on Electromagnetic Wave Theory, IEEE Press, New York, 1997.

[6] A. Novikov, M. Moscoso, and G. Papanicolaou, *Illumination strategies for intensity-only imaging*, SIAM Journal on Imaging Sciences, 8 (2015), pp. 1547–1573.

[7] F. W. J. Olver, D. W. Lozier, R. F. Boisvert, and C. W. Clark, eds., *NIST Handbook of Mathematical Functions*, Cambridge University Press, New York, NY, 2010.

CHAPTER 5

DISSIPATION MECHANISM: EVOLUTION OF MATERIAL MICROSTRUCTURE AND THE GBCD.

A joint work with Yekaterina Epshteyn and David Kinderlehrer.

5.1 Introduction

Cellular networks are commonly found in the natural world. An example of such a network occurs in polycrystalline materials, which find many useful applications throughout science and engineering. These are materials composed of a multitude of small crystallites or grains, separated by interfaces known as grain boundaries (see, e.g., Figure 5.1). These materials are generally metastable and can undergo a rearrangement and growth process under specific conditions, e.g., exposing the material to a heat source such as in annealing. This rearrangement and growth process is often referred to as coarsening or grain growth. As a material undergoes coarsening, two competing processes take place. A total energy is minimized by reducing the total number of grain boundaries in the material, while simultaneously, available space in the material is filled. As the number of grain boundaries decreases, the average size of the grains or cells increases, and still more grain boundaries are eliminated in order to maintain the space-filling constraint. Understanding the mathematical relationship between the grain boundary network and the total energy is the primary theme of this chapter.

There are many models and numerical routines to simulate coarsening in poly-

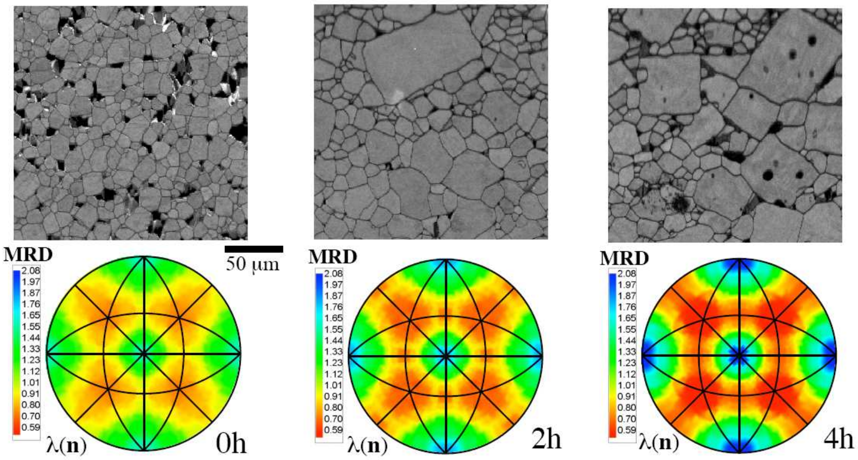


Figure 5.1. Coarsening in polycrystalline materials: (top) polycrystalline microstructures in a real material, (bottom) stereograms of the distribution of interfacial orientations. These images are compiled with permission from the journal "Zeitschrift für Metallkunde", IJMR 2005, pp. 207-210, by E.P. Gorzkowski, T. Sano, C.-S. Kim, G.S. Rohrer, H.M. Chan, and M.P. Harmer, ©Carl Hanser Verlag GmbH & Co. KG, Muenchen, [21].

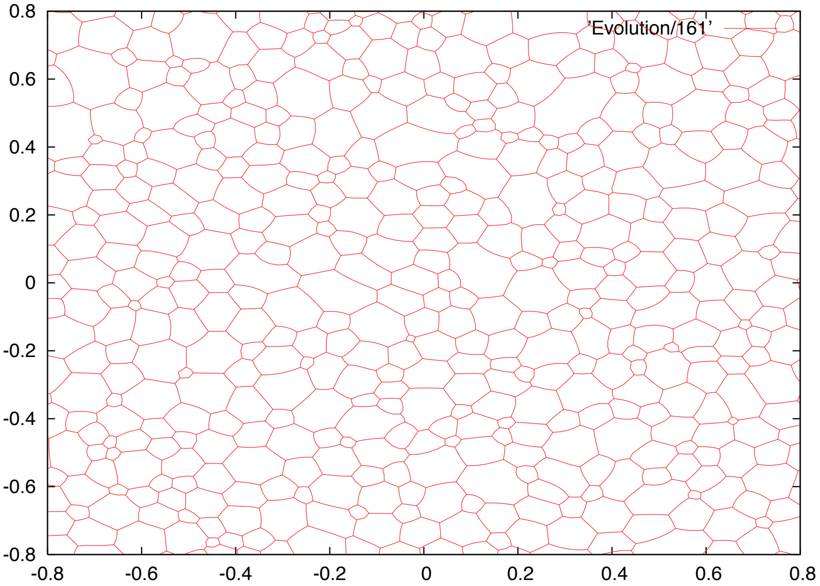


Figure 5.2. An instance of the grain boundary network during the 2D coarsening simulation in [4, 5].

crystalline materials, see, e.g., [5, 12, 18, 19, 22, 26, 30], and each one of them is well suited for studying a particular aspect of grain growth. Molecular dynamics (particle-based) models [12, 22, 30] are best used to study local topological changes happening during coarsening, while implicit methods that treat grains as diffuse characteristic functions [18, 19] are useful to study geometric features such as average grain size or average number of sides of a grain. In recent years, explicit front tracking methods [26, 5] have been developed, which are crucial to study grain boundary dynamics and intergranular energies on a macroscopic scale (see, e.g., Figure 5.2). This is the approach employed by K. Barmak, E. Eggeling, M. Emelianenko, Y. Epshteyn, D. Kinderlehrer, R. Sharp, and S. Ta'asan [4, 5, 6, 9] to study of the evolution of large grain boundary networks, and is the method of study we pursue here.

To characterize material coarsening on a macroscopic scale, one can collect a statistic from the material microstructure and develop a theory for its evolution. Until recently, research has traditionally focused on geometric features of the polycrystalline system such as average grain size or preferred direction of grain lattice orientation. However, owing to the recent advances in experimentation and simulation [1, 25, 26, 32, 33], a new statistic has been discovered, the grain boundary character distribution

(GBCD). The GBCD is an empirical distribution of the relative length (in 2D) or area (in 3D) of interface with a given lattice misorientation and grain boundary normal [1, 4, 5, 6, 9, 25, 26, 32, 33]. It is a statistic that can be harvested from grain boundary networks both *in vivo* and *in silico* and is a leading candidate to characterize texture development [1, 4, 5, 6, 9, 25, 26, 32, 33].

To better understand the dynamics of GBCDs collected from physical materials, in [4, 5, 6, 9] the authors employ a simplified 1D coarsening model of a grain boundary network. This 1D model captures important dynamics of coarsening in polycrystalline materials. Grain boundaries are modeled by 1D subintervals as they grow/shrink to minimize an associated energy in the system, and critical events (the elimination of small shrinking grain boundaries) occur which mimic space-filling constraints. The authors [4, 5, 6, 9] define an associated GBCD and find that its behavior agrees qualitatively with GBCDs harvested both from large-scale 2D simulations as well as physical experiments. Using an assumption that the grain boundary network evolves to maximize configurational entropy, they develop a theory for the GBCD evolution. Their theory suggests the GBCD of the 1D coarsening model evolves as the solution of a Fokker-Planck equation. We review this 1D model and entropy-based theory in §5.2 and §5.3, respectively.

The next step is to validate the Fokker-Planck theory of GBCD evolution. There are some challenges here as there are two a priori unknown coefficients in the Fokker-Planck PDE that must be specified before comparing traits of the GBCD with traits of the Fokker-Planck solution. More precisely, one must first specify an unknown diffusion coefficient and an unknown time scale coefficient. These unknown coefficients arise from two physical uncertainties: the grain boundary system generates entropy at an unknown rate or “temperature”, and the coarsening simulation time scale needs to be compared with the time scale of the Fokker-Planck solution [4, 5, 6, 7, 8, 9]

The unknown diffusion coefficient can be found, as shown in [4, 5, 6, 7, 8, 9], using convergence properties of the Fokker-Planck solution. It is known that the Fokker-Planck solution converges in relative entropy to a Boltzmann distribution with an associated “temperature” (see, e.g., [20]). The authors [4, 5, 6, 7, 8, 9] use this concept to find the unique “temperature” whose corresponding Boltzmann distribution best matches the GBCD harvested at the final time of simulation, i.e., at steady state. Here, the final time for simulations is typically the simulation time corresponding to 80% removal of the initial grain boundaries. A direct comparison of the final time GBCD and this optimal Boltzmann distribution show very good agreement, providing a first and important validation of their theory. In §5.4.1 we recall this validation in more detail, and demonstrate its use on an empirically harvested GBCD in §5.4.4.

In a similar spirit, in §5.4.2 we overcome the time scaling difficulties by developing time-matching procedures based on established theory of the Fokker-Planck PDE. These time-matching procedures fit the GBCD to the Fokker-Planck time scale and allow us to freely choose the unknown time scale coefficient. The procedures are developed by first deriving a formula for the time-varying expected value of a given function, where expectations are computed using the Fokker-Planck solution. Inverting this formula (or approximately inverting it), we obtain a formula to recover time from the function's expected value. This allows us to estimate the "Fokker-Planck time" of the GBCD by computing expectations with the GBCD instead of the Fokker-Planck solution. After the time scale is set in this manner, in §5.4.4, we compare the GBCD with the corresponding Fokker-Planck solution at a few intermediate times and indeed find qualitative agreement.

In §5.5, we use an energy dissipation principle to validate the Fokker-Planck model of GBCD evolution [4, 5, 6, 9]. The Fokker-Planck equation is equivalently expressed as the Wasserstein-2 (W2) gradient flow of a free energy (see [23] for the original proof of this equivalence). By collecting facts of W2 gradient flows and mass transport theory (see, e.g., [2, 23, 35]), we find an energy dissipation identity that must be satisfied for the Fokker-Planck solution. This identity tells us that if the system exhibits Fokker-Planck dynamics, the system must dissipate its energy according to the W2 metric derivative (see §5.5.3 or [2, 23] for more detailed analyses). In §5.5.4, we apply this principle to the GBCD harvested from the 1D coarsening model, and determine that this dissipation identity is approximately satisfied, another validation that the GBCD evolves according to the Fokker-Planck equation.

In §5.6, we quantitatively validate the Fokker-Planck theory of [4, 5, 6, 9] by taking a probabilistic approach to the energy dissipation identity. Because the 1D grain boundary network is initialized randomly, the evolution of the GBCD is random. Thus, the energy dissipated by the system between two times is a random variable that depends on the particular realization of the grain boundary network. With this in mind, we ask if the GBCD dissipation samples converge in some sense to W2 metric derivative samples (Fokker-Planck dissipation samples). We first estimate the distribution of the GBCD energy dissipation samples and the distribution of the W2 metric derivative samples using histograms and multiple realizations of the 1D grain boundary system. Then, using a pseudometric known as an $f-$divergence, we compute the "distance" between the two distributions. This gives us a probabilistic and quantitative understanding of how well the Fokker-Planck equation models GBCD evolution. Moreover, we find numerical evidence of convergence of these energy samples, in distribution, as the number of initial grain boundaries is increased. This suggests that the Fokker-Planck equation models GBCD evolution more accurately in a continuum or many particles limit, i.e., an infinite number of initial grain boundaries.

We conclude our study in §5.7 by summarizing our findings, discussing ongoing work, and commenting on ideas for future research directions.

5.2 A simplified 1D coarsening model [4, 5, 6, 9]

Here we discuss a simplified coarsening model of grain boundaries, that was first employed in studies of the GBCD evolution by K. Barmak, E. Eggeling, M. Emelianenko, Y. Epshteyn, D. Kinderlehrer, R. Sharp, and S. Ta'asan [4, 5, 6, 9]. This simplified 1D critical event model is driven by boundary conditions and has similar dissipative properties as that of 2D and 3D grain growth systems. Important dynamics of grain growth processes are captured in that grain boundaries grow and shrink to minimize an internal energy and it accurately models critical events, i.e., the disappearance of shrinking grain boundaries. Yet, it also simplifies analysis and allows one to find a macroscopic description of the grain boundary dynamics by characterizing the evolution of its GBCD.

As it was presented in [4, 5, 6, 9], an interval $[0, L]$ is considered, which has been randomly subdivided into n smaller subintervals. The endpoints of the subintervals

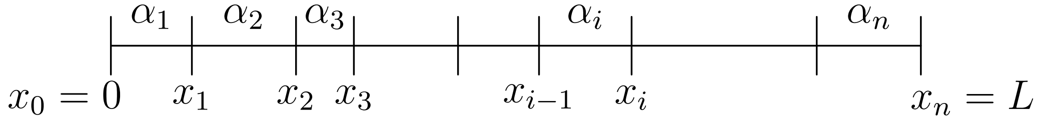


Figure 5.3. Configuration of a typical 1D grain boundary system. The subinterval $[x_{i-1}, x_i]$ models a grain boundary with misorientation parameter α_i while the endpoints $\{x_i\}_{i=0}^{n}$ are 1D analogues of triple junctions. Here misorientation parameters denote the difference in angle of the crystalline lattice orientation of two neighboring grains.

are denoted as x_i for $i = 0, \ldots, n$ where $x_0 = 0$ and $x_n = L$, and for simplicity, the interval endpoints $x_0 \equiv x_n$ are periodically identified. A random misorientation $\alpha_i \in \Omega = [-\pi/4, \pi/4]$ is assigned to each of the subintervals $[x_{i-1}, x_i]$ for $i = 1, \ldots, n$. In many simulations (including ours), the misorientations are sampled as $\alpha_i \sim \text{Unif}(-\pi/4, \pi/4)$ i.i.d., although other distributions may also be used. The $i-$th subinterval $[x_{i-1}, x_i]$ then models a grain boundary (not a grain itself) with grain lattice misorientation α_i. The endpoints $\{x_i\}_{i=0}^{n}$ where two grain boundaries meet are then analogues in 1D of triple junctions (i.e., nodal points where three grain boundaries meet) in 2D and 3D simulations. A typical grain boundary configuration is depicted in Figure 5.3.

Each grain boundary is assigned a grain boundary energy given by a smooth energy density function ψ. By assuming ψ depends only on the misorientation of neighboring grains, one assumes there are no body forces (e.g., gravity) acting on the grain boundary system. The quantity $\psi(\alpha)$ then gives the internal energy per unit length of grain boundary with misorientation α. The total internal energy in the grain boundary system at a given time t is then defined as (see [4, 5, 6, 9])

$$E(t) = \sum_{i=1}^{N} \psi(\alpha_i)\ell_i(t), \tag{5.1}$$

where $\ell_i(t) = x_i(t) - x_{i-1}(t)$ denotes the length of the $i-$th grain boundary. Next, gradient flow dynamics are imposed on the endpoints $\{x_i(t)\}_{i=0}^{n}$ with respect to the internal energy E:

$$\dot{x}_i(t) = -\frac{\partial}{\partial x_i} E(t) = \psi(\alpha_{i+1}) - \psi(\alpha_i) \quad \text{for } i = 0, \ldots, n. \tag{5.2}$$

The velocity of the $i-$th grain boundary (i.e., how fast the $i-$th subinterval is growing or shrinking) is then given by

$$v_i(t) = \dot{x}_i(t) - \dot{x}_{i-1}(t) = \psi(\alpha_{i+1}) - 2\psi(\alpha_i) + \psi(\alpha_{i-1}) \quad \text{for } i = 0, \ldots, n. \tag{5.3}$$

In 1D, the grain growth simulation is simply a sorting problem. Beginning at time $t_0 = 0$, the grain boundary system is initialized as indicated above. Lists of the endpoint locations $\{x_i(t_j)\}_{i=0}^n$, lengths of grain boundaries $\{\ell_i(t_j)\}_{i=1}^n$, and misorientation parameters $\{\alpha_i\}_{i=1}^n$ are then created. The velocities of the grain boundaries $\{v_i(t_j)\}_{i=1}^n$ are computed according to (5.3), which allows one to compute the admissible time step

$$\tau_j = \min_{\substack{i=1,\ldots,n \\ v_i<0}} - \left(\frac{\ell_i(t_j)}{v_i(t_j)} \right).$$

The time step τ_j is the time elapsed until one of the grain boundaries has length zero and thus disappears from the simulation. This grain boundary, its misorientation parameter, and one of its endpoints are then removed from the system. The remaining endpoint locations are updated according to (5.2), and the simulation time is updated as $t_{j+1} = t_j + \tau_j$. This removal and update process is repeated for the set of times $\{t_j\}_{j=1}^N$ (termed critical events or removal events) until a stopping criterion is satisfied. In most simulations (including ours), this criterion is chosen to be 80% removal of the initial grain boundaries. After this time in the simulation, it has been observed that finite system size effects strongly influence and in fact stagnate the evolution of the GBCD.

At any time t of the simulation, the GBCD is defined using a histogram of the relative length of grain boundary with a given misorientation parameter [4, 5, 6, 9]. The state space $\Omega = [-\pi/4, \pi/4]$ is divided into m uniform bins of size $h = \pi/(2m)$, and the GBCD is defined as the probability density

$$\rho(\alpha, t) := \sum_{i=1}^{n} \frac{\ell_i(t)}{Lh} \mathcal{I}\big(\alpha_i \in ((k-1)h, kh]\big) \quad \text{for } \alpha \in ((k-1)h, kh], \tag{5.4}$$

where $\mathcal{I}$ is the indicator function satisfying

$$\mathcal{I}\big(\alpha_i \in ((k-1)h, kh]\big) = \begin{cases} 1 & \text{if } \alpha_i \in ((k-1)h, kh], \\ 0 & \text{otherwise.} \end{cases} \tag{5.5}$$

The $1/h$ scaling in (5.4) is to ensure ρ is a probability density on Ω, i.e., $\int_\Omega \rho d\alpha = 1$. Thus, $\rho(\alpha, t)$ essentially gives the probability of finding a grain boundary with misorientation parameter α at time t.

5.3 Fokker-Planck dynamics: a review of the entropy-based theory [4, 5, 6, 9]

Here we review the entropy-based theory developed in [4, 5, 6, 9] which characterizes the evolution of the GBCD (5.4) as the solution of a Fokker-Planck equation. In §5.3.1, we recall the derivation of a dissipation inequality for a free energy associated with the grain boundary system. In this derivation, the authors [4, 5, 6, 9] made a crucial modeling assumption that the grain boundary system evolves to maximize configurational entropy. In §5.3.2, we review a refined version of this dissipation inequality. This refinement uses concepts from mass transport theory (see, e.g., [35]) and suggests the grain boundary network evolves to minimize viscous dissipation. This in turn suggests the GBCD evolves according to a Fokker-Planck PDE.

5.3.1 Dissipation of the free energy

It is straightforward to show the coarsening system is dissipative for the internal energy $E(t)$ defined in (5.1):

$$\frac{d}{dt}E(t) = \sum_{i=1}^{n} \frac{\partial E}{\partial x_i}\dot{x}_i(t) = -\sum_{i=1}^{n}(\dot{x}_i(t))^2 \leq 0. \tag{5.6}$$

Now, for a small enough $\tau > 0$ such that the time interval $(t, t+\tau)$ *does not contain any critical events*, by the Fundamental Theorem of Calculus one has

$$\int_t^{t+\tau} \sum_{i=1}^{n} \big(\dot{x}_i(s)\big)^2 ds = -\int_t^{t+\tau} \frac{d}{ds}E(s)ds = E(t) - E(t+\tau). \tag{5.7}$$

Using the triangle inequality and Young's inequality[1] one obtains

---

[1] $2ab \leq a^2 + b^2$

$$\begin{aligned}\sum_{i=1}^{n} v_i^2 = \sum_{i=1}^{n} |v_i^2| &= \sum_{i=1}^{n} \left|(\dot{x}_i - \dot{x}_{i-1})^2\right| \\ &= \sum_{i=1}^{n} \left|\dot{x}_i^2 + \dot{x}_{i-1}^2 - 2\dot{x}_i\dot{x}_{i-1}\right| \\ &\leq \sum_{i=1}^{n} \dot{x}_i^2 + \dot{x}_{i-1}^2 + |2\dot{x}_i\dot{x}_{i-1}| \\ &\leq \sum_{i=1}^{n} \dot{x}_i^2 + \dot{x}_{i-1}^2 + \left|\dot{x}_i^2 + \dot{x}_{i-1}^2\right| \leq 4\sum_{i=1}^{n} \dot{x}_i^2.\end{aligned} \tag{5.8}$$

Combining (5.7) and (5.8), the dissipation inequality can be shown:

$$\int_t^{t+\tau} \frac{1}{4}\sum_{i=1}^{n} v_i^2(s)ds + E(t+\tau) \leq E(t).$$

This dissipation inequality is valid under the assumption that there are no critical events in the system, and resembles a similar dissipation inequality for 2D grain growth systems [24, 26]. Using the definition of the GBCD (5.4) and internal energy (5.1), it was shown in [4, 5, 6, 9] this can be further expressed as

$$\mu_0 \int_t^{t+\tau} \int_\Omega |\dot{\rho}(\alpha, s)|^2 \, d\alpha ds + \int_\Omega \rho(\alpha, t+\tau)\psi(\alpha)d\alpha \leq \int_\Omega \rho(\alpha, t)\psi(\alpha)d\alpha, \tag{5.9}$$

where μ_0 is some positive constant and for notational brevity

$$\dot{\rho}(\alpha, t) = \frac{\partial}{\partial t}\rho(\alpha, t).$$

In [4, 5, 6, 9] a crucial modeling assumption was made to account for time intervals that *do contain critical events*. Namely, it was assumed the grain boundary system evolves to maximize configurational entropy. Mathematically, this assumes the grain boundary system evolves so that

$$\mu_0 \int_t^{t+\tau} \int_\Omega |\dot{\rho}|^2 d\alpha ds + \int_\Omega (\rho\psi + \sigma\rho\log\rho)\, d\alpha|_{t+\tau} \leq \int_\Omega (\rho\psi + \sigma\rho\log\rho)\, d\alpha|_t, \tag{5.10}$$

for some $\sigma > 0$ will be determined later. This assumption is consistent with the lack of reversibility of the system upon a grain boundary removal event. The system lacks knowledge of which boundary was removed, and as such the uncertainty of its initial state increases (i.e., its entropy increases). Also note that for the case $\psi \equiv 0$, by minimizing the left hand side of (5.10) the GBCD ρ tends to the uniform distribution on Ω (i.e., all grain boundary misorientations are equally likely).

Note that a free energy of the grain boundary system at time t can be defined as (see [4, 5, 6, 9]):

$$F_\sigma(\rho(\cdot,t)) = \int_\Omega \rho(\alpha,t)\psi(\alpha) + \sigma\rho(\alpha,t)\log\rho(\alpha,t)d\alpha. \tag{5.11}$$

Hence, as shown in [4, 5, 6, 9], inequality (5.10) can be reformulated as

$$\mu_0 \int_t^{t+\tau}\int_\Omega |\dot{\rho}(\alpha,s)|^2\, d\alpha ds + F_\sigma(\rho(\cdot,t+\tau)) \le F_\sigma(\rho(\cdot,t)). \tag{5.12}$$

5.3.2 Mass transport paradigm and the W2 implicit scheme

Inequality (5.12) expresses that the free energy at time $t+\tau$ is bounded above by the free energy at time t. As observed in [4, 5, 6, 9], the term

$$\int_t^{t+\tau}\int_\Omega |\dot{\rho}(\alpha,s)|^2\, d\alpha ds$$

fails as a proper dissipation mechanism as it does not represent energy lost due to frictional or viscous forces. To refine this term, in [4, 5, 6, 9] the authors derived the additional bound

$$\frac{C}{\min_\Omega \rho(\alpha,t)}\int_t^{t+\tau}\int_\Omega v^2(\alpha,s)\rho(\alpha,s)d\alpha ds \le \int_t^{t+\tau}\int_\Omega |\dot{\rho}(\alpha,s)|^2\, d\alpha ds, \tag{5.13}$$

for some $C > 0$, where ρ and v are related through the continuity equation

$$\frac{\partial\rho}{\partial t}(\alpha,t) + \frac{\partial}{\partial\alpha}(v(\alpha,t)\rho(\alpha,t)) = 0.$$

The term on the left hand side of (5.13) is a proper dissipation term as it represents loss of energy due to viscous forces. Owing to a formulation of the W2 metric by Benamou and Brenier [10], the left hand side of (5.13) in turn bounds the W2 distance $d^2(\cdot,\cdot)$ between $\rho(\alpha,t)$ and $\rho(\alpha,t+\tau)$:

$$\frac{1}{\tau}d^2(\rho(\cdot,t),\rho(\cdot,t+\tau)) = \inf_{f,v}\int_t^{t+\tau}\int_\Omega v(\alpha,s)^2 f(\alpha,s)d\alpha ds \tag{5.14}$$

where the infimum is taken over all f, v satisfying

$$\begin{cases} \dfrac{\partial f}{\partial t}(\alpha,t) + \dfrac{\partial}{\partial\alpha}(v(\alpha,s)f(\alpha,t)) = 0, \\ f(\alpha,t) = \rho(\alpha,t), f(\alpha,t+\tau) = \rho(\alpha,t+\tau). \end{cases}$$

Assuming $\rho(\alpha) > \delta > 0$ for all $\alpha \in \Omega$, by combining (5.12), (5.13) and (5.14), it was shown in [4, 5, 6, 9] that

$$\frac{\mu}{2\tau} d^2(\rho(\cdot,t), \rho(\cdot,t+\tau)) + F_\sigma(\rho(\cdot,t+\tau)) \le F_\sigma(\rho(\cdot,t)), \tag{5.15}$$

where $\mu > 0$ is some constant.

The dissipation inequality (5.15) is highly suggestive of a mass transport implicit scheme (see, e.g., §5.5.3, [23]). That is, assuming the grain boundary network evolves in a manner that minimizes viscous energy dissipation (i.e., its viscous energy dissipation is given by the W2 metric as in (5.14)), one obtains the GBCD at time $t + \tau$ from knowledge of the GBCD at time t by solving the minimization problem

$$\rho(\alpha, t+\tau) = \arg\ \min_{\rho^*} \frac{\mu}{2\tau} d^2(\rho(\cdot,t), \rho^*(\cdot)) + F_\sigma(\rho^*(\cdot)). \tag{5.16}$$

As shown by R. Jordan, D. Kinderlehrer, and F. Otto [23], such a minimization problem is a time-discrete implicit scheme with respect to the W2 metric for the solution of a Fokker-Planck equation. More precisely, as $\tau \to 0$ the solution $\rho(\alpha, t+\tau)$ of (5.16) converges to the solution of the PDE

$$\mu \frac{\partial \rho}{\partial t} = \frac{\partial}{\partial \alpha}\Big(\sigma \frac{\partial \rho}{\partial \alpha} + \psi' \rho\Big), \tag{5.17}$$

with periodic boundary conditions and initial data $\rho(\alpha, t)$.

We summarize the above theory. In [4, 5, 6, 9], two crucial assumptions are made to develop an evolution equation for the GBCD ρ. Namely, it was assumed that the grain boundary network evolves to simultaneously maximize configurational entropy (5.10) while minimizing viscous dissipation (5.14). With these assumptions in place, one finds the GBCD evolves according to a time-discrete form of the Fokker-Planck equation. The remainder of this chapter is focused on validating this theory numerically.

5.4 Determining the parameters σ and μ

In order to validate the Fokker-Planck equation (5.17) as a model for GBCD evolution, the unknown parameters σ and μ need to be specified. The diffusive parameter σ can be estimated using a relative-entropy procedure originally developed in [4, 5, 6, 7, 8, 9]. We recall this method in §5.4.1. Then in §5.4.2, we develop two novel time-matching procedures to estimate Fokker-Planck times from the GBCD ρ. These procedures abandon the simulation times $\{t_j\}$ in favor of times recovered from theoretical formulas, allowing us to freely choose the parameter μ. In §5.4.3 we show μ only scales the time of the Fokker-Planck solution, and so an obvious choice for simplicity is choosing $\mu \equiv 1$. With σ and μ determined, we can directly compare the GBCD with the solution of the corresponding Fokker-Planck equation (5.17). Numerical demonstrations of these routines and comparisons are provided in §5.4.4.

5.4.1 Relative entropy test

Here we review the relative entropy validation procedure developed in [4, 5, 6, 7, 8, 9]. To estimate the parameter σ we look to the stationary solution of the Fokker-Planck PDE (5.17), which is given by the Boltzmann distribution [20]

$$\rho_\sigma(\alpha) = \frac{1}{Z_\sigma} e^{-\frac{\psi(\alpha)}{\sigma}},$$

with partition function (i.e., normalization factor)

$$Z_\sigma = \int_\Omega e^{-\frac{\psi(\alpha)}{\sigma}} d\alpha.$$

Solutions of the Fokker-Planck PDE (5.17) have the property that they converge exponentially fast to ρ_σ in Kullback-Leibler (KL) relative entropy. KL relative entropy, or KL divergence, is a particular $f-$divergence, i.e., it is a particular function that measures the difference between two probability densities (see, e.g., [15, 16]). We denote the KL relative entropy Φ between the GBCD $\rho(\alpha, t)$ and the Boltzmann distribution $\rho_\sigma(\alpha)$ as

$$\begin{aligned}\Phi(\rho(\cdot,t),\rho_\sigma(\cdot)) &= \int_\Omega \rho(\alpha,t)\log\frac{\rho(\alpha,t)}{\rho_\sigma(\alpha)}d\alpha \\ &= \frac{1}{\sigma}F_\sigma(\rho(\cdot,t)) + \log(Z_\sigma) \\ &= \frac{F_\sigma(\rho(\cdot,t)) - F_\sigma(\rho_\sigma(\cdot))}{\sigma}.\end{aligned}$$

If the GBCD $\rho(\alpha, t)$ indeed evolves according to (5.17), it must converge exponentially fast to $\rho_\sigma(\alpha)$ in KL relative entropy as $t \to \infty$. In other words, the free energy of the grain boundary system must decrease (exponentially fast) to the free energy of the Boltzmann distribution $\rho_\sigma(\alpha)$. This concept was introduced in [4, 5, 6, 7, 8, 9] to estimate σ as described below.

Consider the final time of the harvested GBCD $\rho(\alpha, t_N)$. As stated in §5.2, typically t_N is the simulation time when 80% of the initial grain boundaries have been removed. Assuming $\rho(\alpha, t_N)$ as the steady-state of the GBCD, the parameter σ can be estimated via

$$\sigma = \arg\ \min_{\lambda>0} \Phi(\rho(\cdot, t_N), \rho_\lambda(\cdot)). \tag{5.18}$$

This finds the unique value of λ for which the free energies $F_\lambda(\rho(\cdot, t_N))$ and $F_\lambda(\rho_\lambda(\cdot))$ best agree, or in other words, the value of λ for which the Boltzmann distribution $\rho_\lambda(\alpha)$ best matches $\rho(\alpha, t_N)$. We demonstrate this procedure numerically in §5.4.4 and defer to [4, 5, 6, 7, 8, 9] for the original formulation of this validation procedure.

5.4.2 Time-matching formulae

Here we provide two time-matching procedures which can be used to assign Fokker-Planck times to empirically harvested GBCDs. The main principle we use here is that expected values evolve in a predictable way, assuming the GBCD indeed solves the Fokker-Planck equation (5.17). Our first method (§5.4.2.1) inverts a formula for the time evolution of the expectation of certain eigenfunctions. These eigenfunctions are observed to be heavily weighted near the boundary $\partial\Omega$. We thus provide another method (§5.4.2.2) that uses expectations of functions which are heavily weighted on the interior of Ω, and is based on inverting numerical quadrature formulas.

5.4.2.1 Time-matching via eigenfunctions

Denote ρ^{fp} as the solution of the Fokker-Planck equation (5.17). We can express ρ^{fp} using an eigenfunction expansion (see, e.g., [20, Sec. 5.2.5] or [31, Sec. 5.4]) as

$$\rho^{\text{fp}}(\alpha, t) = \sum_{i=1}^{\infty} a_i p_i(\alpha) e^{-\frac{\nu_i}{\mu} t}, \tag{5.19}$$

where p_i are eigenfunctions (with eigenvalue $-\nu_i$) of the operator

$$\mathcal{L}(u) = \frac{\partial}{\partial\alpha}\Big(\sigma\frac{\partial u}{\partial\alpha} + \psi' u\Big). \tag{5.20}$$

The coefficients a_i in (5.19) are obtained from a projection of the initial data:

$$a_i = \int_\Omega \rho^{\text{fp}}(\alpha, 0) q_i(\alpha) d\alpha,$$

where q_i are eigenfunctions (again with eigenvalue $-\nu_i$) of the adjoint operator

$$\mathcal{L}^\dagger(u) = \sigma \frac{\partial^2 u}{\partial \alpha^2} - \psi' \frac{\partial u}{\partial \alpha}. \tag{5.21}$$

Together, the eigenfunctions q_i and p_i form a biorthogonal system on Ω, i.e.,

$$\int_\Omega p_i(\alpha) q_j(\alpha) d\alpha = \delta_{ij},$$

where δ_{ij} is the usual Kronecker delta.

Now, consider observing the expected value of q_i over time. Using representation (5.19) and the orthogonality of p_i, q_j for $i \neq j$, we obtain

$$\begin{aligned} f(t) = \int_\Omega q_i(\alpha) \rho^{\text{fp}}(\alpha, t) d\alpha &= \sum_{j=0}^{\infty} a_j e^{-\frac{\nu_j}{\mu} t} \int_\Omega p_j(\alpha) q_i(\alpha) d\alpha \\ &= a_i e^{-\frac{\nu_i}{\mu} t} \\ &= \left(\int_\Omega q_i(\alpha) \rho^{\text{fp}}(\alpha, 0) d\alpha \right) e^{-\frac{\nu_i}{\mu} t} \\ &= f(0) e^{-\frac{\nu_i}{\mu} t}. \end{aligned}$$

Thus, we can recover the time t from ρ^{fp} as

$$\begin{aligned} t &= -\frac{\mu}{\nu_i} \log \left(\frac{f(t)}{f(0)} \right) \\ &= -\frac{\mu}{\nu_i} \left[\log \left(\int_\Omega \rho^{\text{fp}}(\alpha, t) q_i(\alpha) \right) - \log \left(\int_\Omega \rho^{\text{fp}}(\alpha, 0) q_i(\alpha) \right) \right]. \end{aligned} \tag{5.22}$$

For a given value of μ and eigenfunction q_i, we can use (5.22) as a formula to estimate Fokker-Planck times for the GBCD ρ:

$$t_j^{\text{ef}} := -\frac{\mu}{\nu_i} \left[\log \left(\int_\Omega \rho(\alpha, t_j) q_i(\alpha) \right) - \log \left(\int_\Omega \rho(\alpha, t_0) q_i(\alpha) \right) \right] \tag{5.23}$$

for $j = 0, \ldots, N$. Here the superscript "ef" is to denote "eigen-fit". We define the eigen-fit GBCD as

$$\rho^{\text{ef}}(\alpha, t_j^{\text{ef}}) := \rho(\alpha, t_j) \quad \text{for } j = 0, \ldots, N. \tag{5.24}$$

Thus, by using the GBCD ρ to compute expected values instead of the Fokker-Planck solution ρ^{fp}, we match the GBCD with its most representative Fokker-Planck time.

We remark on a few points here. First, the eigenfunctions q_i do not necessarily have an explicit representation for a given energy density ψ. Thus in practice, we will estimate the eigenfunctions and eigenvalues numerically by discretizing the operator $\mathcal{L}^\dagger$ and computing the spectrum of the discrete operator. Furthermore, it can be shown (see, e.g., [31]) the eigenvalues $-\nu_i \leq 0$ for all i, i.e., the operators $\mathcal{L}, \mathcal{L}^\dagger$, are negative semidefinite. Theoretically there should be no difference in computing eigen-fit times using different eigenpairs $(q_i, -\nu_i)$. However, this procedure is observed to work better when using eigenpairs with larger nonzero eigenvalues $-\nu_i$ since higher modes decay more rapidly. We include numerical simulations of the eigen-fit time-matching procedure in §5.4.4.

5.4.2.2 Time-matching via numerical quadrature

Consider a function $w(\alpha) \in C_0^\infty(\Omega)$, i.e., a smooth function with compact support on Ω such that $w(\alpha) \geq 0$ for all $\alpha \in \Omega$. Its expected value at time t is given by

$$f(t) = \int_\Omega \rho^{\text{fp}}(\alpha, t) w(\alpha) d\alpha,$$

where ρ^{fp} denotes the solution of the Fokker-Planck equation (5.17). We derive an evolution equation for $f(t)$ by formally computing $f'(t)$:

$$\begin{aligned} f'(t) = \int_\Omega \frac{\partial \rho^{\text{fp}}}{\partial t} w d\alpha &= \int_\Omega \frac{1}{\mu} \frac{\partial}{\partial \alpha} \left(\sigma \frac{\partial \rho^{\text{fp}}}{\partial \alpha} + \psi' \rho^{\text{fp}} \right) w d\alpha \\ &\overset{1}{=} - \int_\Omega \frac{1}{\mu} \left(\sigma \frac{\partial \rho^{\text{fp}}}{\partial \alpha} + \psi' \rho^{\text{fp}} \right) w' d\alpha \\ &\overset{2}{=} \int_\Omega \frac{1}{\mu} \left(\sigma \rho^{\text{fp}} w'' - \psi' \rho^{\text{fp}} w' \right) d\alpha = \int_\Omega \rho^{\text{fp}} \frac{1}{\mu} \left(\sigma w'' - \psi' w' \right) d\alpha, \end{aligned}$$

where equalities 1 and 2 indicate integration by parts and the boundary terms vanish from the compact support of w. Denoting

$$g(t) = \int_\Omega \rho^{\text{fp}}(\alpha, t) \frac{1}{\mu} \left(\sigma w''(\alpha) - \psi'(\alpha) w'(\alpha) \right) d\alpha = \int_\Omega \rho^{\text{fp}}(\alpha, t) \mathcal{L}^\dagger w(\alpha) d\alpha$$

where $\mathcal{L}^\dagger$ is given by (5.21), we have the ODE

$$f'(t) = g(t).$$

If $f(t_j)$ and $g(t_j)$ are known for some time t_j, the ODE $f'(t) = g(t)$ can be solved approximately for $f(t_{j+1})$ using, e.g., the trapezoidal rule

$$f(t_{j+1}) = f(t_j) + (t_{j+1} - t_j) \frac{g(t_{j+1}) + g(t_j)}{2}. \tag{5.25}$$

Now, if we instead assume $t_j, f(t_j), f(t_{j+1}), g(t_j)$ and $g(t_{j+1})$ are known, we can invert (5.25) to solve for t_{j+1}:

$$\begin{aligned} t_{j+1} &= t_j + 2\frac{f(t_{j+1}) - f(t_j)}{g(t_{j+1}) + g(t_j)} \\ &= t_j + 2\mu\frac{\int_\Omega \rho^{\text{fp}}(\alpha, t_j)w(\alpha)d\alpha - \int_\Omega \rho^{\text{fp}}(\alpha, t_{j-1})w(\alpha)d\alpha}{\int_\Omega \rho^{\text{fp}}(\alpha, t_j)\mathcal{L}^\dagger w(\alpha)d\alpha + \int_\Omega \rho^{\text{fp}}(\alpha, t_{j-1})\mathcal{L}^\dagger w(\alpha)d\alpha}. \end{aligned} \tag{5.26}$$

Note there is no restriction to using the trapezoidal rule above, we simply choose it for its second-order behavior and its simplicity.

For a given value of μ and function $w(\alpha)$, we can use (5.26) as a formula to estimate Fokker-Planck times for the GBCD ρ:

$$t_j^{\text{qf}} := \begin{cases} 0, & j = 0, \\ t_j^{\text{qf}} + 2\mu\dfrac{\int_\Omega \rho(\alpha, t_j)w(\alpha)d\alpha - \int_\Omega \rho(\alpha, t_{j-1})w(\alpha)d\alpha}{\int_\Omega \rho(\alpha, t_j)\mathcal{L}^\dagger w(\alpha)d\alpha + \int_\Omega \rho(\alpha, t_{j-1})\mathcal{L}^\dagger w(\alpha)d\alpha}, & j > 0. \end{cases} \tag{5.27}$$

Here the superscript "qf" is to denote "quadrature-fit". As before we define the associated time-matched GBCD as

$$\rho^{\text{qf}}(\alpha, t_j^{\text{qf}}) := \rho(\alpha, t_j) \quad \text{for } j = 0, \dots, N. \tag{5.28}$$

We remark this procedure is a slight generalization of the eigen-fit method. The eigen-fit method uses $w(\alpha) \equiv q_i(\alpha)$, i.e., an eigenfunction of $\mathcal{L}^\dagger$. In this case, the ODE $f'(t) = g(t)$ can be easily solved and inverted exactly to recover formula (5.22). However, the eigenfunctions q_i are observed (numerically) to be heavily weighted near the boundary $\partial\Omega$. Thus, eigen-fit times computed from expectations of q_i mostly use information of the GBCD near the boundary $\partial\Omega$. This boundary information is known to be less accurate than GBCD information in the interior of Ω, due to finite size effects of the grain boundary system. Using the quadrature-fit method, we have control over the function $w(\alpha)$, so we can choose one heavily weighted in the interior of Ω and hopefully obtain more accurate time estimates. Numerical simulations of the quadrature-fit time-matching procedure are included in §5.4.4.

5.4.3 Time scaling effect of μ

The effect of μ in the Fokker-Planck equation (5.17) is simply a global time scaling. This can be seen as follows. Assume $\rho^{\text{fp}}(\alpha, t)$ satisfies (5.17) and define a new function

$\widetilde{\rho}$ as

$$\widetilde{\rho}(\alpha, t) := \rho^{\mathrm{fp}}(\alpha, \mu t).$$

Thus, $\widetilde{\rho}$ is a time-dilated version of the Fokker-Planck solution ρ^{fp}. By direct computation we find $\widetilde{\rho}$ satisfies

$$\begin{aligned}
\frac{\partial}{\partial t}\widetilde{\rho}(\alpha, t) = \frac{\partial}{\partial t}\rho^{\mathrm{fp}}(\alpha, \mu t) &= \left(\mu \frac{\partial}{\partial t'}\rho^{\mathrm{fp}}(\alpha, t')\right)\Bigg|_{t'=\mu t} \\
&= \left(\frac{\partial}{\partial \alpha}\left(\sigma \frac{\partial \rho^{\mathrm{fp}}}{\partial \alpha}(\alpha, t') + \psi'(\alpha)\rho^{\mathrm{fp}}(\alpha, t')\right)\right)\Bigg|_{t'=\mu t} \\
&= \frac{\partial}{\partial \alpha}\left(\sigma \frac{\partial \rho^{\mathrm{fp}}}{\partial \alpha}(\alpha, \mu t) + \psi'(\alpha)\rho^{\mathrm{fp}}(\alpha, \mu t)\right) \\
&= \frac{\partial}{\partial \alpha}\left(\sigma \frac{\partial \widetilde{\rho}}{\partial \alpha}(\alpha, t) + \psi'(\alpha)\widetilde{\rho}(\alpha, t)\right).
\end{aligned}$$

Hence, $\widetilde{\rho}$ satisfies the Fokker-Planck equation (5.17) with $\mu \equiv 1$.

From this observation we note that by choosing a value of μ to use one of the time-matching formulae (5.23) or (5.27), we are simply choosing a *global* time scaling for the GBCD. For simplicity then, we will use $\mu \equiv 1$ when computing GBCD times $\{t_j^{\mathrm{ef}}\}$ or $\{t_j^{\mathrm{qf}}\}$ as this will not affect qualitative behaviors, merely the global time scale on which they occur.

5.4.4 Numerical validation and parameter estimation

We begin our numerical demonstrations by illustrating the relative entropy validation procedure reviewed in §5.4.1. We consider a GBCD ρ harvested from a 1D coarsening simulation (§5.2) with energy density $\psi(\alpha) = 1 + 2\alpha^2$ and 2^{15} initial grain boundaries. We compute the KL relative entropy $\Phi\big(\rho(\cdot, t_j), \rho_\lambda(\cdot)\big)$ for all of the collection times $\{t_j\}_{j=0}^{N}$ and for 30 uniformly spaced values of $\lambda \in [0.02, 0.04]$. The resulting relative entropy curves are shown in Figure 5.4(a). By computing the minimum over all λ of $\Phi(\rho(\cdot, t_N), \rho_\lambda(\cdot))$ at the final time t_N, i.e., using (5.18), we estimate $\sigma \approx 0.03069$. Its corresponding relative entropy curve is depicted in red. We also plot the corresponding Boltzmann distribution $\rho_\sigma(\alpha)$ (red) against the final time GBCD $\rho(\alpha, t_N)$ (blue) in Figure 5.4(b). The agreement between the two densities is a first validation of the Fokker-Planck model of the GBCD evolution as discussed in [4, 5, 6, 7, 8, 9].

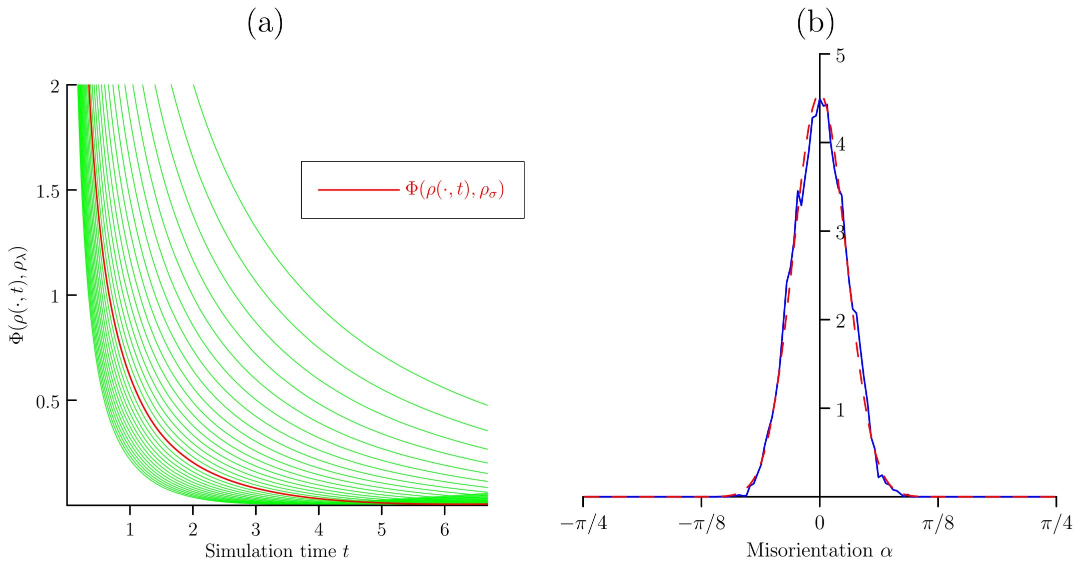


Figure 5.4. Relative entropy: (a) KL relative entropy curves $\Phi(\rho, \rho_\lambda)$ for 30 trial values of $\lambda \in [0.01, 0.05]$. The red curve depicts the optimal curve for $\sigma = 0.03069$, (b) comparison of the final time GBCD $\rho(\alpha, t_N)$ (blue) with the optimal Boltzmann distribution $\rho_\sigma(\alpha)$ (red dashed).

Next, we numerically investigate the time-fitting procedures of §5.4.2. As a benchmark test, we apply the time-fitting formulae (5.23) and (5.27) to the solution ρ^{fp} of the Fokker-Planck equation (5.17) to see if the correct times are recovered, at least approximately. To this end, we first generate a numerical solution ρ^{fp} of (5.17) with $\mu \equiv 1$, $\sigma = 0.03069$, $\psi = 1 + 2\alpha^2$, using an initial GBCD as initial data $\rho^{\mathrm{fp}}(\alpha, 0)$. The numerical solution is obtained by discretizing the operator $\mathcal{L}(u) = \frac{\partial}{\partial\alpha}(\sigma\frac{\partial u}{\partial\alpha} + \psi' u)$ with second-order finite differences with periodic boundary conditions, and using the backward Euler method to solve for the time steps $t_j^{\mathrm{fp}} = (j-1)/999$ for $j = 1, \ldots, 1000$.

To test the eigen-fit procedure, we discretize the operator $\mathcal{L}^\dagger(u) = \sigma\frac{\partial^2 u}{\partial\alpha^2} - \psi'\frac{\partial u}{\partial\alpha}$ with second-order finite differences with periodic boundary conditions, and compute the discrete eigenpairs $(q_i, -\nu_i)$. Using formula (5.22) with $\mu \equiv 1$ and the eigenpair $(q_i, -\nu_i)$ with largest nonzero eigenvalue $-\nu_i (\approx -4)$, we recover the times $\{t_j^{\mathrm{ef}}\}$. We plot the recovered times against the true Fokker-Planck times (backward Euler times) $\{t_j^{\mathrm{fp}}\}$ in Figure 5.5(a). The linear relation here shows the eigen-fit method indeed recovers the true times approximately. The slight break from linearity at later times shows the method has difficulties recovering times as the solution ρ^{fp} nears its steady-

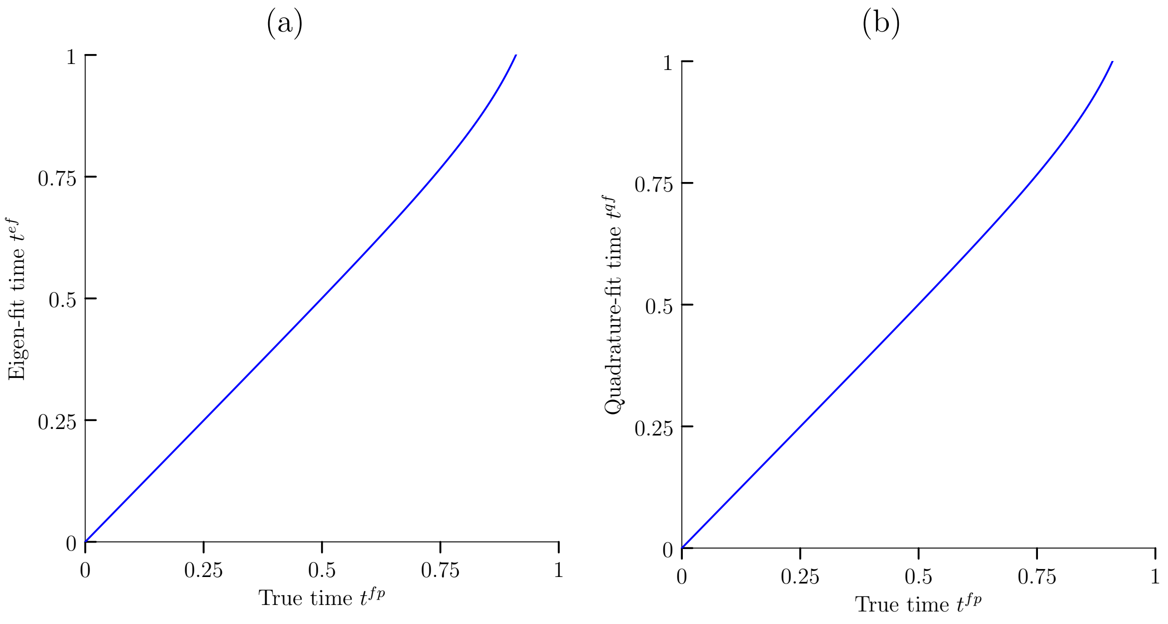


Figure 5.5. A comparison of the true Fokker-Planck times $\{t_j^{\mathrm{fp}}\}$ with the (a) the eigen-fit recovered times $\{t_j^{\mathrm{ef}}\}$ and (b) the quadrature-fit recovered times $\{t_j^{\mathrm{qf}}\}$. The linear relation shows the recovery formulas approximately recover the true Fokker-Planck times.

state.

To test the quadrature-fit procedure, we use formula (5.27) with $\mu \equiv 1$ and function $w(\alpha) = \cos(2\alpha)$ to recover the times $\{t_j^{\mathrm{qf}}\}$. We again use the second-order finite difference approximation of $\mathcal{L}^\dagger$ in formula (5.27). We plot these recovered times against the true Fokker-Planck times $\{t_j^{\mathrm{fp}}\}$ in Figure 5.5(b). Here we again see a linear relation indicating the quadrature-fit method approximately recovers the true times. Similar to the eigen-fit method, there appears to be a slight break in the time recovery as the solution ρ^{fp} approaches its steady-state.

We now apply the time-fitting procedures to the GBCD ρ. For the eigen-fit method, we again discretize the operator $\mathcal{L}^\dagger$ and compute its eigenpairs $(q_i, -\nu_i)$. We use (5.23) with $\mu \equiv 1$ and the eigenpair $(q_i, -\nu_i)$ having largest nonzero eigenvalue $-\nu_i(\approx -4)$ to compute the Fokker-Planck times $\{t_j^{\mathrm{ef}}\}$, and define the eigen-fit GBCD ρ^{ef} according to (5.24). We compare the eigen-fit GBCD ρ^{ef} with the Fokker-Planck solution ρ^{fp} (obtained as above) at a few times. The two densities are plotted in Figure 5.6 at eigen-fit times $t_j^{\mathrm{ef}} = 0.057, 0.136, 0.263, 1$. These times correspond

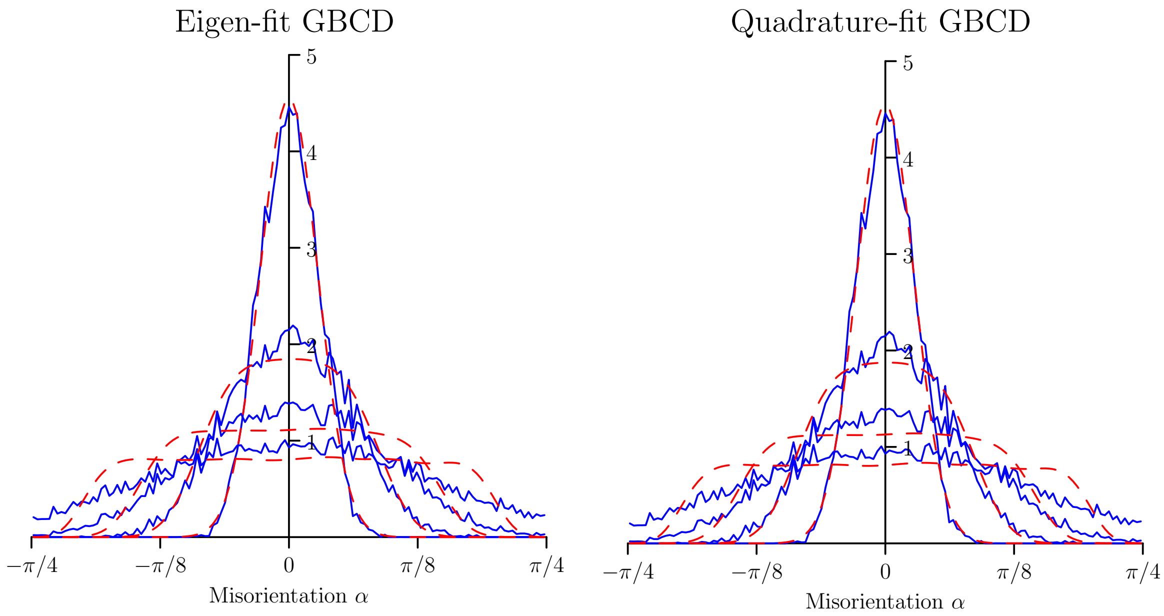


Figure 5.6. Comparison of the time-fitted GBCD (blue) with the Fokker–Planck solution ρ^{fp} (red dashed). Here ρ^{fp} is obtained using the backward Euler method. The left image compares the eigen-fit GBCD ρ^{ef} with ρ^{fp} at times $t_j^{\text{ef}} = 0.057, 0.136, 0.263, 1$. The right image compares the quadrature-fit GBCD ρ^{qf} with ρ^{fp} at times $t_j^{\text{qf}} = 0.059, 0.14, 0.267, 1$.

to approximately $20\%, 40\%, 60\%, 80\%$ initial grain boundary removal, respectively. The agreement between the time-matched GBCD ρ^{ef} (blue) and the Fokker-Planck solution ρ^{fp} (red) further validates the Fokker-Planck model of the GBCD evolution.

For the quadrature-fit method, we use (5.27) with $\mu \equiv 1$ and $w(\alpha) = \cos(2\alpha)$ to compute the Fokker-Planck times $\{t_j^{\text{qf}}\}$. We then define the quadrature-fit GBCD ρ^{qf} according to (5.28). In Figure 5.6, we compare ρ^{qf} with the Fokker-Planck solution ρ^{fp} (obtained as above) at quadrature-fit times $t_j^{\text{qf}} = 0.059, 0.14, 0.267, 1$, corresponding to approximately $20\%, 40\%, 60\%, 80\%$ initial grain boundary removal, respectively. Again we observe validation of the Fokker-Planck model of GBCD evolution as we see good agreement between the densities ρ^{qf} (blue) and ρ^{fp} (red).

Lastly, we compare the time-matched GBCDs ρ^{ef} and ρ^{qf} with the Fokker-Planck solution ρ^{fp} obtained from the eigenfunction expansion (5.19). To compute ρ^{fp} we discretize the operators $\mathcal{L}$ and $\mathcal{L}^\dagger$, compute the eigenpairs of the discrete operators, compute the coefficients a_i using an initial GBCD as the initial data, and finally form

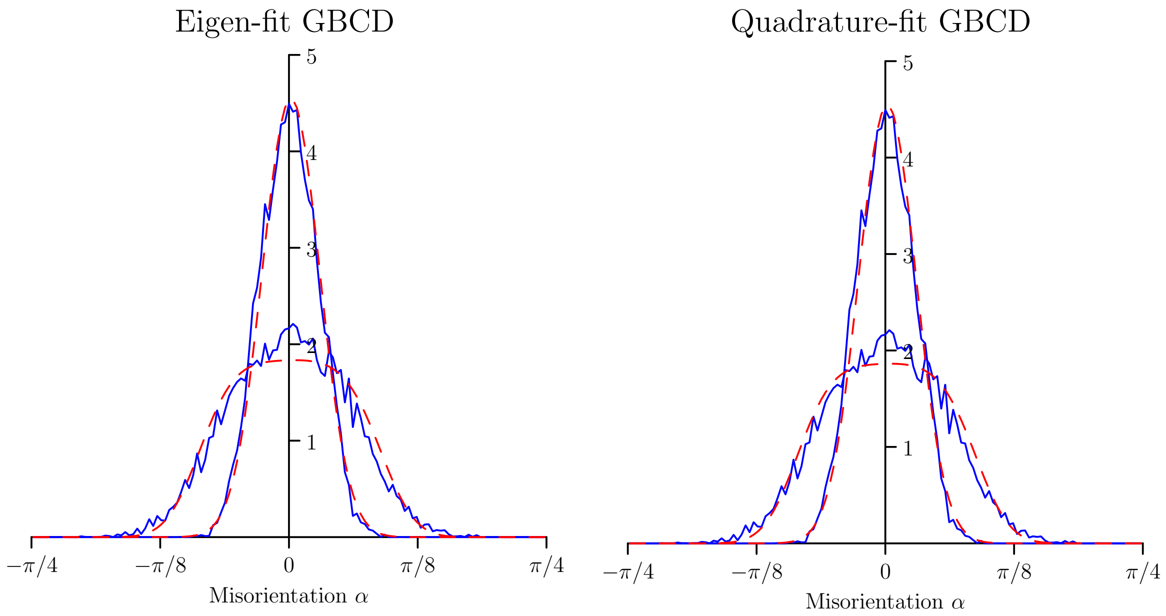


Figure 5.7. Density comparisons the time-fitted GBCDs (blue) with the Fokker–Planck solution ρ^{fp} (red dashed). Here ρ^{fp} is obtained using the eigenfunction expansion (5.19). The left image compares the eigen-fit GBCD ρ^{ef} with ρ^{fp} at times $t_j^{\mathrm{ef}} = 0.263, 1$. The right image compares the quadrature-fit GBCD ρ^{qf} with ρ^{fp} at times $t_j^{\mathrm{qf}} = 0.267, 1$.

the expansion (5.19) with $\mu \equiv 1$. In Figure 5.7 we compare ρ^{ef} and ρ^{qf} with ρ^{fp} at Fokker-Planck times corresponding to approximately 60% and 80% initial grain boundary removal. We note the qualitative agreement is similar to that of Figure 5.6, an additional validation of the Fokker-Planck model of GBCD evolution.

5.5 Gradient flows and dissipation mechanisms

In this section we study gradient flows in several contexts. It was first shown in [23] that the Fokker-Planck equation (5.17) can be regarded as the W2 gradient flow of the free energy F_σ. Thus to validate the Fokker-Planck equation (5.17) as a description of the GBCD evolution, we seek an energy dissipation principle the GBCD must satisfy if it is a W2 gradient flow of F_σ. We begin in §5.5.1 by recalling a classical gradient flow in $\mathbb{R}^d$, and reformulating it as variational problem. This variational formulation can be generalized and in §5.5.2 we consider an example L^2 gradient flow in the space $H^1(\Omega)$, namely, the diffusion equation. From a simple quadratic inequality, we derive an energy dissipation identity that holds only for solutions to the diffusion equation. These examples are patterned after a series of lecture notes by M. Peletier [28, 29]. They serve as analogies of our main result in §5.5.3, where we derive an energy dissipation identity that holds if and only if a function is a W2 gradient flow of the free energy F_σ, i.e., the solution of the Fokker-Planck equation (5.17). In §5.5.4, we determine that the GBCD approximately satisfies this identity which validates the Fokker-Planck model of GBCD evolution.

5.5.1 Gradient flows in Euclidean space

Consider the gradient flow for $x(t) \in \mathbb{R}^d$ of an energy function $\mathcal{F} : \mathbb{R}^d \to \mathbb{R}$ given by the system

$$\frac{d}{dt}x(t) = -\nabla\mathcal{F}(x(t)) \iff \dot{x}(t) + \nabla\mathcal{F}(x(t)) = 0. \tag{5.29}$$

If initial conditions $x(0)$ are known, the backward Euler approximation of this system is given by

$$\frac{x(\tau) - x(0)}{\tau} + \nabla\mathcal{F}(x(\tau)) = 0. \tag{5.30}$$

Notice, $x(\tau)$ then satisfies the stationary condition implied by the minimization problem

$$\min_{x(\tau)\in\mathbb{R}^d} \frac{\|x(\tau) - x(0)\|^2}{2\tau} + \mathcal{F}(x(\tau)), \tag{5.31}$$

where $\|\cdot\|$ denotes the standard Euclidean norm in $\mathbb{R}^d$. Thus, we may approximate the solution of the gradient flow system (5.29) by finding the minimizer of (5.31). Moreover, as $\tau \to 0$, the approximation (5.30) becomes exact. These are motivating observations to generalize gradient flows to infinite dimensional vector spaces.

5.5.2 An L^2 gradient flow

Consider an analogous minimization problem for functions in $H^1(\Omega)$, where we recall $\Omega = [-\pi/4, \pi/4] \subset \mathbb{R}$:

$$\min_{\substack{u(\cdot,\tau)\in H^1(\Omega)\\ \frac{\partial u}{\partial x}|_{\partial\Omega}=0}} \frac{\|u(\cdot,\tau) - u(\cdot,0)\|^2_{L^2(\Omega)}}{2\tau} + \frac{1}{2}\int_\Omega \left|\frac{\partial u}{\partial x}(x,\tau)\right|^2 dx, \tag{5.32}$$

where $u(x,0)$ is assumed to be known and $\|\cdot\|_{L^2(\Omega)}$ denotes the L^2 norm of functions defined on Ω. Setting the first variation of this expression to zero, we obtain the stationary condition

$$\int_\Omega \frac{u(x,\tau)-u(x,0)}{\tau} v(x)dx = -\int_\Omega \frac{\partial u}{\partial x}(x,\tau)\frac{\partial v}{\partial x}(x)dx, \tag{5.33}$$

for all functions $v \in H^1(\Omega)$. Recalling the backward Euler time derivative approximation

$$\frac{u(x,\tau)-u(x,0)}{\tau} \approx \frac{\partial u}{\partial t}(x,\tau),$$

we recognize (5.33) as a weak approximation of the diffusion equation

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}, \tag{5.34}$$

with Neumann boundary conditions $\frac{\partial u}{\partial x}\big|_{\partial\Omega} = 0$ and $u(x,0)$ as initial data. As $\tau \to 0$, the minimizer of (5.32) converges to the solution of (5.34) with initial data $u(x,0)$. Thus, (5.32) is a variational formulation of the diffusion equation (as $\tau \to 0$). Noting the explicit dependence of (5.32) on the L^2 metric and energy $\mathcal{F}(u) = \frac{1}{2}\int \left|\frac{\partial u}{\partial x}\right|^2 dx$ in (5.32), the diffusion equation is considered the L^2 gradient flow of the energy functional $\mathcal{F}(u)$.

We now derive an energy dissipation identity that only solutions of (5.34) satisfy. For any sufficiently regular function $u(x,t)$ satisfying $\frac{\partial u}{\partial x}|_{\partial\Omega} = 0$, we observe the rate of change of the energy $\mathcal{F}(u)$:

$$-\frac{d}{dt}\mathcal{F}(u) = -\int_\Omega \frac{\partial u}{\partial x}\frac{\partial^2 u}{\partial x \partial t}dx = \int_\Omega \frac{\partial^2 u}{\partial x^2}\frac{\partial u}{\partial t}dx, \tag{5.35}$$

where we have integrated by parts in the last equality and the boundary terms vanish because $\frac{\partial u}{\partial x}\big|_{\partial\Omega} = 0$. We also have

$$0 \le \int_\Omega \left(\frac{\partial u}{\partial t} - \frac{\partial^2 u}{\partial x^2}\right)^2 dx = \int_\Omega \left(\left|\frac{\partial u}{\partial t}\right|^2 + \left|\frac{\partial^2 u}{\partial x^2}\right|^2 - 2\frac{\partial u}{\partial t}\frac{\partial^2 u}{\partial x^2}\right) dx, \tag{5.36}$$

where equality with zero holds if and only if $\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}$ almost everywhere in Ω. Rearranging (5.36), we have

$$\int_\Omega \frac{\partial u}{\partial t}\frac{\partial^2 u}{\partial x^2}dx \le \frac{1}{2}\int_\Omega \left(\left|\frac{\partial u}{\partial t}\right|^2 + \left|\frac{\partial^2 u}{\partial x^2}\right|^2\right) dx. \tag{5.37}$$

Finally, we combine (5.35) and (5.37) to obtain

$$-\frac{d}{dt}\mathcal{F}(u) = \int_\Omega \frac{\partial^2 u}{\partial x^2}\frac{\partial u}{\partial t}dx \le \frac{1}{2}\int_\Omega \left(\left|\frac{\partial u}{\partial t}\right|^2 + \left|\frac{\partial^2 u}{\partial x^2}\right|^2\right) dx, \tag{5.38}$$

where equality holds if and only if $\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}$ almost everywhere in Ω. Thus, (5.38) is satisfied with equality if and only if $u(x,t)$ is a solution of the diffusion equation (5.34), i.e., the L^2 gradient flow of the energy $\mathcal{F}$.

5.5.2.1 Estimating the dissipation of L^2 gradient flows

Assuming u solves the diffusion equation (5.34), we compute the rate of dissipation, which by (5.38) is

$$-\frac{d}{dt}\mathcal{F}(u(\cdot,t)) = \int_\Omega \left(\frac{\partial u}{\partial t}(x,t)\right)^2 dx.$$

Expressing the time derivatives as limits we have

$$\begin{aligned}\lim_{\tau\to 0}\frac{\mathcal{F}(u(\cdot,t) - \mathcal{F}(u(\cdot,t+\tau))}{\tau} &= \lim_{\tau\to 0}\frac{1}{\tau^2}\int_\Omega (u(x,t+\tau)-u(x,t))^2\,dx \\ &= \lim_{\tau\to 0}\frac{1}{\tau^2}\|u(\cdot,t+\tau)-u(\cdot,t)\|^2_{L^2(\Omega)},\end{aligned}$$

where we have assumed sufficient smoothness of u to pass the limit outside of the integral. By dropping the limits and multiplying by τ, we can estimate the energy dissipation of an L^2 gradient flow using the L^2 metric:

$$\mathcal{F}(u(\cdot,t) - \mathcal{F}(u(\cdot,t+\tau)) \approx \frac{1}{\tau}\|u(\cdot,t+\tau)-u(\cdot,t)\|^2_{L^2(\Omega)}. \tag{5.39}$$

In §5.5.3, we will find an analogous estimate for the energy dissipation of a W2 gradient flow.

5.5.3 A Wasserstein-2 gradient flow

For a W2 gradient flow, we consider the state space

$$\mathcal{P}(\Omega) = \{\rho : \Omega \to [0,\infty) \big| \int_\Omega \rho(\alpha)d\alpha = 1, \int_\Omega \alpha^2\rho(\alpha)d\alpha < \infty\}, \tag{5.40}$$

where $\Omega = [-\pi/4, \pi/4] \subset \mathbb{R}$. Thus, $\mathcal{P}(\Omega)$ is the space of probability densities on Ω with finite second moment. For densities $\rho(\cdot,\tau) \in \mathcal{P}(\Omega)$ we can write the variational form

$$\min_{\rho(\cdot,\tau)\in\mathcal{P}(\Omega)} \frac{d^2(\rho(\cdot,\tau),\rho(\cdot,0))}{2\tau} + F_\sigma(\rho(\cdot,\tau)), \tag{5.41}$$

where $\rho(\cdot,0) \in \mathcal{P}(\Omega)$ is a known probability density, $d(\cdot,\cdot)$ denotes the W2 distance (5.14), and $F_\sigma(\cdot)$ is the free energy defined in (5.11). Clearly, (5.41) is analogous to

(5.32) with the W2 metric $d(\cdot,\cdot)$ in place of the L^2 metric and with energy functional $F_\sigma : \mathcal{P}(\Omega) \to \mathbb{R}$, i.e., the free energy (5.11). R. Jordan, D. Kinderlehrer and F. Otto introduced this scheme in [23] and showed as $\tau \to 0$, the minimizer of (5.41) converges to the solution of the Fokker-Planck equation (5.17) with $\mu \equiv 1$ and $\rho(\alpha, 0)$ as initial data. Thus, we interpret the dynamics implied by the Fokker-Planck equation (5.17) as the W2 gradient flow of the free energy F_σ.

Analogous to §5.5.2, we now derive a dissipation identity that is satisfied only for the solution of the Fokker-Planck equation (5.17) with $\mu \equiv 1$, i.e., the W2 gradient flow of F_σ. The following theorem is essentially a collection of a few facts of W2 gradient flows and W2 spaces. We assume sufficient smoothness of the underlying functions for clarity and precision. However, similar statements can be made in a much more general setting. We defer to [2, 3] for these generalized statements.

Theorem 7 (W2 dissipation identity) *For* $\Omega = [-\pi/4, \pi/4] \subset \mathbb{R}$, *define* $\mathcal{P}(\Omega)$ *as in* (5.40). *We assume* $\rho(\cdot, t') \in \mathcal{P}(\Omega)$ *for all* $t' \in (t, t+\tau)$ *and satisfies the following:*

$$\rho \in C^2(\Omega \times (t, t+\tau)), \tag{5.42}$$

$$\rho(\alpha, t') > \delta > 0, \quad \forall t' \in (t, t+\tau), \tag{5.43}$$

$$\rho(-\pi/4, t') = \rho(\pi/4, t'), \quad \forall t' \in (t, t+\tau) \tag{5.44}$$

$$d(\rho(\cdot, t_1), \rho(\cdot, t_2)) \le \int_{t_1}^{t_2} f(s)ds \quad \textit{for } t < t_1 \le t_2 < t+\tau, \tag{5.45}$$

for some $f \in L^1((t, t+\tau))$ *where* $d(\cdot,\cdot)$ *denotes the W2 metric* (5.14). *The energy dissipation identity*

$$-\frac{d}{dt}F_\sigma(\rho) = \frac{1}{2}\int_\Omega \left(\left(\sigma \frac{1}{\rho}\frac{\partial \rho}{\partial \alpha} + \psi' \right)^2 + v^2 \right) \rho d\alpha$$

holds if and only if $v = -(\sigma \frac{1}{\rho}\frac{\partial \rho}{\partial \alpha} + \psi')$, *where* v *and* ρ *satisfy continuity equation*

$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial \alpha}(v\rho) = 0. \tag{5.46}$$

Proof. As discussed in [2, Definition 2.28], assumption 5.45 implies $\rho(\cdot, t) : t \mapsto \rho(\alpha, t)$ is an absolutely continuous curve in $\mathcal{P}(\Omega)$. By [2, Theorem 2.29], if $\rho(\cdot, t)$ is an absolutely continuous curve then there must exist a velocity field $v(\alpha, t)$ such that

$$\int_\Omega v^2(\alpha, t')\rho(\alpha, t')d\alpha < \infty \quad \text{for } t' \in (t, t+\tau),$$

and the continuity equation (5.46) holds in the sense of distributions. By our assumed smoothness (5.42) of ρ and the Riesz Representation Theorem, (5.46) must hold in the classic sense for some v. Furthermore, since ρ is a probability density we have $\int_\Omega \rho(\alpha,t')d\alpha = 1$ for all $t' \in (t, t+\tau)$, and so

$$0 = \frac{d}{dt}\int_\Omega \rho d\alpha = \int_\Omega \frac{\partial \rho}{\partial t} dt = -\int_\Omega \frac{\partial}{\partial \alpha}(v\rho)d\alpha = (v\rho)\Big|_{-\pi/4}^{\pi/4}.$$

Then, using the periodic boundary conditions (5.44) and positivity (5.43) of ρ, we have

$$\big(v(\pi/4,t') - v(-\pi/4,t')\big)\rho(\pi/4,t') = 0 \implies v(-\pi/4,t') = v(\pi/4,t'). \tag{5.47}$$

Analogous to our derivation of the L^2 dissipation inequality (5.38), we now compute the dissipation rate of the free energy F_σ:

$$\begin{aligned}
-\frac{d}{dt}F_\sigma(\rho) = -\int_\Omega \frac{\partial}{\partial t}\left(\sigma\rho\log\rho + \psi\rho\right)d\alpha &= -\int_\Omega \left(\sigma\frac{\partial \rho}{\partial t}\log\rho + \sigma\frac{\partial \rho}{\partial t} + \psi\frac{\partial \rho}{\partial t}\right)d\alpha \\
&= -\int_\Omega (\sigma\log\rho + \sigma + \psi)\frac{\partial \rho}{\partial t}d\alpha \\
&= \int_\Omega (\sigma\log\rho + \sigma + \psi)\frac{\partial}{\partial \alpha}(v\rho)d\alpha \\
&= -\int_\Omega \left(\sigma\frac{1}{\rho}\frac{\partial \rho}{\partial \alpha} + \psi'\right)v\rho d\alpha,
\end{aligned} \tag{5.48}$$

where we have integrated by parts in the last line and the boundary terms have vanished owing to the periodic boundary conditions (5.44) and (5.47). We also have

$$\begin{aligned}
0 &\le \int_\Omega \left(\left(\sigma\frac{1}{\rho}\frac{\partial \rho}{\partial \alpha} + \psi'\right) + v\right)^2 \rho d\alpha \\
&= \int_\Omega \left(\left(\sigma\frac{1}{\rho}\frac{\partial \rho}{\partial \alpha} + \psi'\right)^2 + v^2 + 2\left(\sigma\frac{1}{\rho}\frac{\partial \rho}{\partial \alpha} + \psi'\right)v\right)\rho d\alpha,
\end{aligned} \tag{5.49}$$

where equality with zero holds if and only if $v = -\left(\sigma\frac{1}{\rho}\frac{\partial\rho}{\partial\alpha} + \psi'\right)$, which follows by the assumed positivity (5.43) and continuity (5.42) of ρ. Rearranging (5.49), we have

$$-\int_\Omega \left(\sigma\frac{1}{\rho}\frac{\partial \rho}{\partial \alpha} + \psi'\right)v\rho \le \frac{1}{2}\int_\Omega \left(\left(\sigma\frac{1}{\rho}\frac{\partial \rho}{\partial \alpha} + \psi'\right)^2 + v^2\right)\rho d\alpha. \tag{5.50}$$

Finally, we combine (5.48) and (5.50) to obtain

$$-\frac{d}{dt}F_\sigma(\rho) = -\int_\Omega \left(\sigma\frac{1}{\rho}\frac{\partial \rho}{\partial \alpha} + \psi'\right)v\rho d\alpha \le \frac{1}{2}\int_\Omega \left(\left(\sigma\frac{1}{\rho}\frac{\partial \rho}{\partial \alpha} + \psi'\right)^2 + v^2\right)\rho d\alpha, \tag{5.51}$$

where equality holds if and only if

$$v = -\left(\sigma \frac{1}{\rho}\frac{\partial \rho}{\partial \alpha} + \psi'\right).$$

■

We remark at the similarity between (5.51) and (5.38). In (5.38), we formed a dissipation inequality that is satisfied with equality only for a solution of the diffusion equation, i.e., the L^2 gradient flow of an energy $\mathcal{F}$. Similarly, inequality (5.51) is satisfied with equality only for a density ρ such that

$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial \alpha}\left(-\left(\sigma \frac{\partial \rho}{\partial \alpha}\frac{1}{\rho} + \psi'\right)\rho\right) = 0 \iff \frac{\partial \rho}{\partial t} = \frac{\partial}{\partial \alpha}\left(\sigma \frac{\partial \rho}{\partial \alpha} + \psi'\rho\right).$$

Thus, (5.51) is satisfied with equality if and only if $\rho(\alpha, t)$ is the solution of the Fokker-Planck equation (5.17) with $\mu \equiv 1$, i.e., the W2 gradient flow of the free energy F_σ (owing to the equivalence of (5.17) and (5.41) established in [23]). Finally, we note that if $v = +\left(\sigma \frac{1}{\rho}\frac{\partial \rho}{\partial \alpha} + \psi'\right)$, then the system described by ρ is nondissipative since we have

$$\frac{d}{dt}F_\sigma(\rho) = \int_\Omega \left(\sigma \frac{1}{\rho}\frac{\partial \rho}{\partial \alpha} + \psi'\right)^2 \rho d\alpha > 0.$$

5.5.3.1 Estimating the dissipation of W2 gradient flows

From (5.51) we find the solution ρ^{fp} of the Fokker-Planck equation (5.17) with $\mu \equiv 1$ dissipates the free energy F_σ as

$$-\frac{d}{dt}F_\sigma(\rho^{\text{fp}}) = \int_\Omega v^2 \rho^{\text{fp}} d\alpha, \tag{5.52}$$

where $v = -\left(\sigma \frac{1}{\rho^{\text{fp}}}\frac{\partial \rho^{\text{fp}}}{\partial \alpha} + \psi'\right)$. We can estimate (5.52) from below as

$$\begin{aligned} -\frac{d}{dt}F_\sigma(\rho^{\text{fp}}) = \int_\Omega v^2 \rho^{\text{fp}} d\alpha &\overset{1}{=} \lim_{\tau \to 0} \frac{1}{\tau} \int_t^{t+\tau} \int_\Omega v^2 \rho^{\text{fp}} d\alpha ds \\ &\geq \lim_{\tau \to 0} \frac{1}{\tau} \inf_{\widetilde{\rho}, \widetilde{v}} \int_t^{t+\tau} \int_\Omega \widetilde{v}^2 \widetilde{\rho} d\alpha ds, \end{aligned}$$

where equality 1 is the mean value theorem for integrals. Here the inf is taken over all $\widetilde{\rho}, \widetilde{v}$ satisfying

$$\begin{cases} \dfrac{\partial}{\partial t}\widetilde{\rho} + \dfrac{\partial}{\partial \alpha}(\widetilde{v}\widetilde{\rho}) = 0, \\ \widetilde{\rho}(\alpha, t) = \rho^{\text{fp}}(\alpha, t), \quad \widetilde{\rho}(\alpha, t+\tau) = \rho^{\text{fp}}(\alpha, t+\tau). \end{cases}$$

Recalling the Benamou-Brenier formulation of the W2 metric (5.14) and writing the time derivative of F_σ as a limit we have

$$\lim_{\tau\to 0}\frac{F_\sigma(\rho^{\mathrm{fp}}(\cdot,t)) - F_\sigma(\rho^{\mathrm{fp}}(\cdot,t+\tau))}{\tau} \geq \lim_{\tau\to 0}\frac{1}{\tau^2}d^2(\rho^{\mathrm{fp}}(\cdot,t),\rho^{\mathrm{fp}}(\cdot,t+\tau)). \tag{5.53}$$

Remark 12 (W2 metric derivative) *Numerical observations and a theoretical result found in [2, Proposition 3.22], show that ρ^{fp} in fact satisfies* (5.53) *with* equality[2]*:*

$$\lim_{\tau\to 0}\frac{F_\sigma(\rho^{fp}(\cdot,t)) - F_\sigma(\rho^{fp}(\cdot,t+\tau))}{\tau} = \lim_{\tau\to 0}\frac{1}{\tau^2}d^2(\rho^{fp}(\cdot,t),\rho^{fp}(\cdot,t+\tau)).$$

Thus analogous to §5.5.2.1, by dropping the limits and multiplying by τ, we estimate the dissipation of a W2 gradient flow using the W2 metric:

$$F_\sigma(\rho^{fp}(\cdot,t)) - F_\sigma(\rho^{fp}(\cdot,t+\tau)) \approx \frac{1}{\tau}d^2(\rho^{fp}(\cdot,t),\rho^{fp}(\cdot,t+\tau)). \tag{5.54}$$

5.5.4 Numerical validation of W2 dissipation

We now use the dissipation estimate (5.54) to numerically validate the Fokker-Planck model of GBCD evolution. The principle we use here is that if a density evolves according to the Fokker-Planck equation (5.17), it must dissipate energy according to (5.54). As a benchmark test, we first generate a numerical solution of the Fokker-Planck equation (5.17) and determine the dissipation estimate (5.54) indeed holds. Next, we consider a time-matched GBCD harvested from the 1D critical event model, and find the estimate (5.54) is approximately satisfied for the GBCD, away from initial times. Thus, the GBCD exhibits W2 dissipation of the free energy F_σ, validating it as an approximate solution of the Fokker-Planck equation (5.17).

We begin with the benchmark test. We construct a numerical solution of (5.17) with $\mu \equiv 1$, $\sigma = 0.03069$, $\psi = 1+2\alpha^2$, using an initial GBCD as initial data. To solve (5.17) numerically, we use second-order finite differences for spatial derivatives and solve for 1000 time steps using the backward Euler method to estimate $\rho^{\mathrm{fp}}(\alpha,t_j)$ where $t_j = (j-1)/999$ for $j = 1,\ldots,1000$. We compute the free energy dissipation $F_\sigma(\rho^{\mathrm{fp}}_j) - F_\sigma(\rho^{\mathrm{fp}}_{j+1})$ and the W2 distances $d^2(\rho^{\mathrm{fp}}_j,\rho^{\mathrm{fp}}_{j+1})/\tau_j$ where we have abbreviated $\rho^{\mathrm{fp}}_j =$

---

[2] A similar dissipation identity holds for gradient flows for general metrics and sufficiently convex energy functionals, i.e., not only for the W2 gradient flow of F_σ in one-dimension. See [2] for details.

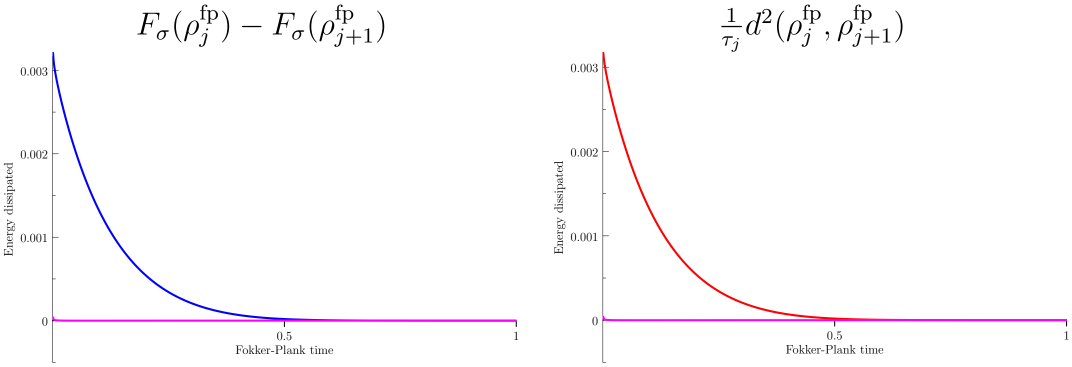


Figure 5.8. Fokker-Planck dissipation: the energy dissipation $F_\sigma(\rho_j^{\mathrm{fp}}) - F_\sigma(\rho_{j+1}^{\mathrm{fp}})$ (left) and the W2 distances $d^2(\rho_j^{\mathrm{fp}}, \rho_{j+1}^{\mathrm{fp}})/\tau_j$ (right). The difference between the two curves is depicted in magenta (approx. 0). Here the density ρ^{fp} is a numerical solution of the Fokker-Planck equation (5.17). The agreement between the curves indicates (5.54) is satisfied and thus ρ^{fp} is indeed an approximate solution of (5.17).

$\rho^{\mathrm{fp}}(\alpha, t_j)$ and $\tau_j = t_{j+1} - t_j$ for $j = 1, \ldots, 999$. Here the W2 distances are computed using a numerical algorithm developed by D. Kinderlehrer and N. Walkington [27]. In Figure 5.8 we plot the energy dissipation and the W2 distances. The difference between the two curves, depicted in magenta, indicates the free energy dissipation agrees with the W2 distances, and thus (5.54) is satisfied. As expected, we validate the numerical solution ρ^{fp} satisfies the Fokker-Planck equation (5.17).

Now we apply the dissipation estimate (5.54) to time-matched GBCDs. We first apply the test to a GBCD harvested from the 1D simulation (§5.2) with $n = 2^{15}$ initial grain boundaries. We use the energy density $\psi = 1 + 2\alpha^2$ and harvest the GBCD at the set of collection times $\{t_j\}_{j=1}^N$. The parameter $\sigma \approx 0.03069$ is estimated using the relative entropy test of §5.4.1. Next, we implement the quadrature-fit method of §5.4.2 to fit the GBCD to Fokker-Planck times:

$$\rho^{\mathrm{qf}}(\alpha, t_j^{\mathrm{qf}}) = \rho(\alpha, t_j) \quad \text{for } j = 0, \ldots, N,$$

where t_j^{qf} is given by (5.27) (with $\mu \equiv 1$). Finally, we compute the energy dissipation $F_\sigma(\rho_j^{\mathrm{qf}}) - F_\sigma(\rho_{j+1}^{\mathrm{qf}})$ and W2 distances $d^2(\rho_j^{\mathrm{qf}}, \rho_{j+1}^{\mathrm{qf}})/\tau_j^{\mathrm{qf}}$, where we have abbreviated $\rho_j^{\mathrm{qf}} = \rho^{\mathrm{qf}}(\alpha, t_j^{\mathrm{qf}})$ and $\tau_j^{\mathrm{qf}} = t_{j+1}^{\mathrm{qf}} - t_j^{\mathrm{qf}}$ for $j = 0, \ldots, N-1$. We compare the energy dissipation and the W2 distances in Figure 5.9, where we see qualitative agreement between the curves, away from initial times. Because the solution of (5.17) must satisfy the dissipation estimate (5.54), we have further validation that the Fokker-Planck equation characterizes the GBCD evolution. We believe the disagreement at these earlier times can be understood as transient startup effects of the 1D coarsening simulation. That is, at the beginning of the 1D simulation high-energy grain boundaries (i.e., grain boundaries with misorientation parameters α near the boundary of $\Omega = [-\pi/4, \pi/4]$) are not eliminated as rapidly as in later times of the simulation (cf. Figure 5.6). We focus more on the later times of the simulation since the 1D grain boundary system at these times is more representative of the theory derived in [4, 5, 6, 9].

In Figure 5.9, we also show the dissipation estimate for GBCDs harvested from simulations having $n = 2^{17}$ and $n = 2^{19}$ initial grain boundaries. We observe as the number of initial grain boundaries increases, the dissipation estimate (5.54) improves. This improvement suggests the Fokker-Planck equation is a more accurate description of GBCD evolution as the number of grain boundaries $n \to \infty$, i.e., in a continuum or many particles limit.

5.6 Statistical validation of Fokker-Planck dynamics

We now quantify the agreement between the GBCD evolution and Fokker-Planck dynamics by asking how well (5.54) is satisfied. Since the grain boundary system is initialized randomly, the GBCD itself is random and thus the energy dissipation $F_\sigma(\rho(\cdot, t_j)) - F_\sigma(\rho(\cdot, t_{j+1}))$ and W2 distances $d^2(\rho(\cdot, t_j), \rho(\cdot, t_{j+1}))$ between any two times are themselves random variables. Therefore, we consider generating many realizations of the grain boundary system and associated statistics to determine if the energy dissipation samples and W2 distance samples agree in a statistical sense. In §5.6.1 we develop a procedure to quantify the agreement between the distributions of the energy dissipation samples and the W2 distance samples. We implement this procedure in §5.6.2 and observe the energy samples appear to converge, in distribution, as the number of initial grain boundaries n increases, i.e., in a continuum or many particles limit. To summarize this observation we formulate Conjecture 1, and characterize the implications of such convergence for the grain boundary system dynamics.

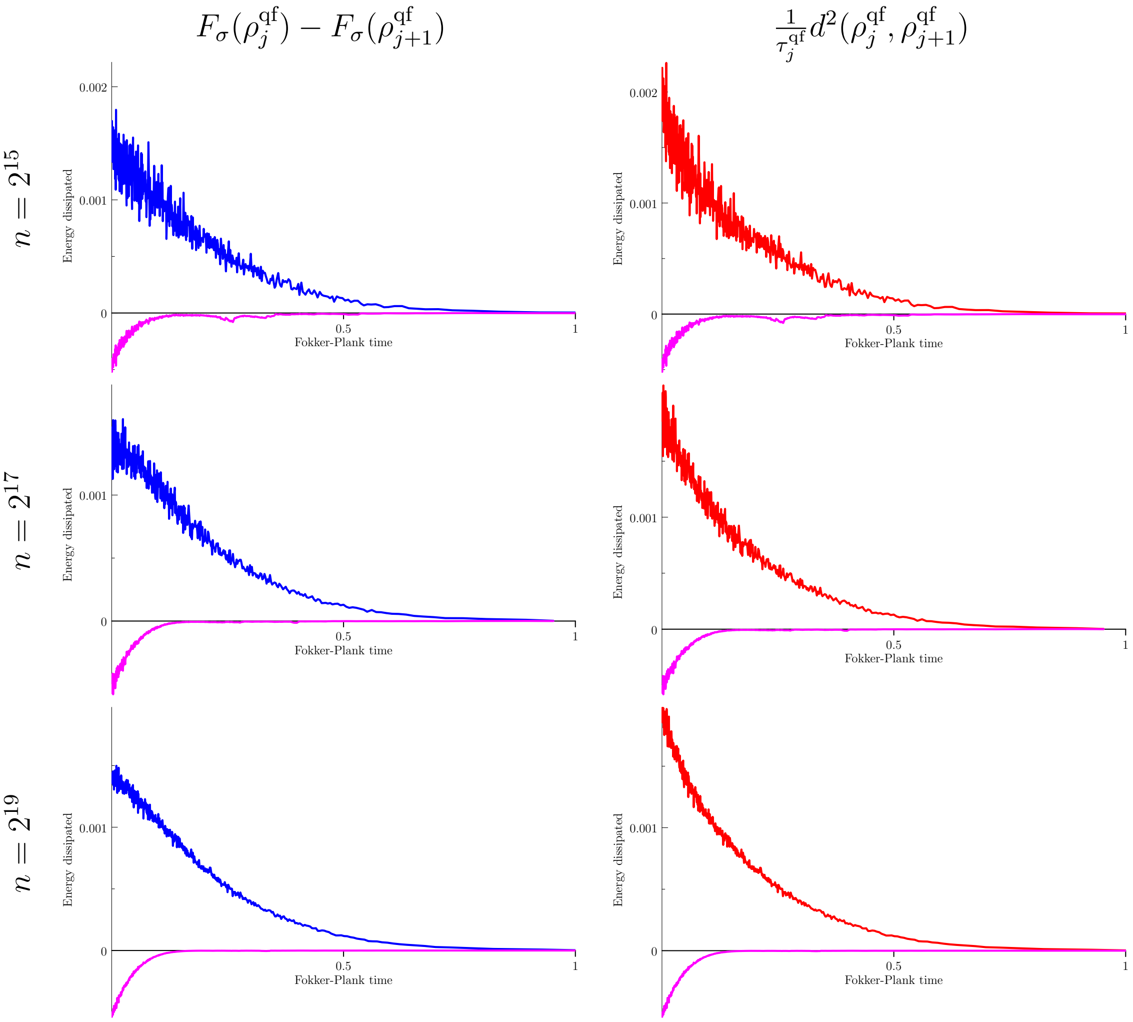


Figure 5.9. GBCD dissipation: the energy dissipation $F_\sigma(\rho_j^{\mathrm{qf}}) - F_\sigma(\rho_{j+1}^{\mathrm{qf}})$ (left) and the W2 distances $d^2(\rho_j^{\mathrm{qf}}, \rho_{j+1}^{\mathrm{qf}})/\tau_j^{\mathrm{qf}}$ (right) for GBCDs harvested from 1D simulations for different numbers of initial grain boundaries n (indicated). Here the time scales have been set using the quadrature-fit method of §5.4.2.2 with $\mu \equiv 1$. The difference between the two curves is depicted in magenta and indicates (5.54) is satisfied, thus validating the Fokker-Planck model of GBCD evolution.

5.6.1 Energy samples as random variables

Consider the 1D coarsening simulation of §5.2. The grain boundary system is initialized randomly both by the random subdivision of the interval $[0, L]$ into subintervals (grain boundaries), and by the initial prescription of random misorientation parameters α_i. Thus, statistics harvested from the grain boundary system are inherently random. For a given realization of the grain boundary system, we collect the statistics

$$\{\rho(\alpha, t_j)\}_{j=0}^{N}, \quad \sigma, \quad \{t_j^{\mathrm{ef}}\}_{j=0}^{N}, \quad \{t_j^{\mathrm{qf}}\}_{0=1}^{N},$$

$$\{F_\sigma(\rho(\cdot, t_j))\}_{j=0}^{N}, \quad d^2(\rho(\cdot, t_j), \rho(\cdot, t_{j+1}))\}_{j=0}^{N-1},$$

and more importantly, the energetic statistics (as they are energetic quantities)

$$\begin{aligned} \{\Delta F_j\}_{j=0}^{N-1} &= \{F_\sigma(\rho(\cdot, t_j)) - F_\sigma(\rho(\cdot, t_{j+1}))\}_{j=0}^{N-1}, \\ \{W_j\}_{j=0}^{N-1} &= \left\{ \frac{d^2(\rho(\cdot, t_j), \rho(\cdot, t_{j+1}))}{t_{j+1}^{\mathrm{qf}} - t_j^{\mathrm{qf}}} \right\}_{j=0}^{N-1}. \end{aligned} \tag{5.55}$$

In this context, in §5.5.4 we verified that $\Delta F_j \approx W_j$ for each $j = 0, \ldots, N-1$ and for a single realization of the 1D grain boundary system. This in turn validated the Fokker-Planck model of GBCD evolution qualitatively for the given realization. We now validate the model in a quantitative and probabilistic sense by considering multiple realizations of the grain boundary system.

Denote by $f_{\Delta F}(e, t)$ and $f_W(e, t)$, the time-dependent probability densities for the energetic random variables ΔF and W, respectively. That is, $f_{\Delta F}(e, t)$ gives the probability of observing an energy dissipation sample e at time t, and likewise for $f_W(e, t)$. The state space for these probability densities is $\mathbb{R}^+ = [0, \infty)$ since energy is a non-negative quantity. For the time being, we assume $f_{\Delta F}(e, t)$ and $f_W(e, t)$ are known, but we will estimate them empirically in §5.6.2.

To quantify the difference between the probability distributions of ΔF and W, we use an $f-$divergence. An $f-$divergence is a real-valued non-negative function, that measures the difference between two probability densities (see, e.g., [13, 14]). Crucially, an $f-$divergence is zero if and only if the densities, and therefore the associated distributions, are equal. Of the many choices of $f-$divergence forms, we focus on the KL divergence of $f_{\Delta F}$ from f_W at time t which is given by

$$\Phi(f_{\Delta F}(\cdot,t), f_W(\cdot,t)) = \int_0^\infty f_{\Delta F}(e,t)\log\left(\frac{f_{\Delta F}(e,t)}{f_W(e,t)}\right)de. \tag{5.56}$$

Our choice of KL divergence is because it also provides an upper bound on the total variation distance between the two distributions, i.e., the L^1 norm between the densities:

$$0 \le \frac{1}{2}\|f_{\Delta F}(\cdot,t) - f_W(\cdot,t)\|^2_{L^1(\mathbb{R}^+)} \le \Phi(f_{\Delta F}(\cdot,t), f_W(\cdot,t)), \tag{5.57}$$

which is known as the Csiszár-Kullback-Pinsker inequality (see, e.g., [11]). Thus, convergence of $f_{\Delta F}(e,t)$ to $f_W(e,t)$ in KL divergence, implies convergence in L^1. These two quantities give a measurement of the difference between the probability densities $f_{\Delta F}(e,t)$ and $f_W(e,t)$ at a given time t, with smaller values indicating statistical agreement.

For global (in time) quantification, we then integrate the time-dependent curves $\Phi(f_{\Delta F}(\cdot,t), f_W(\cdot,t))$ and $\|f_{\Delta F}(\cdot,t) - f_W(\cdot,t)\|^2_{L^1(\mathbb{R}^+)}$:

$$\begin{aligned}\int_0^\infty \Phi(f_{\Delta F}(\cdot,t), f_W(\cdot,t))dt &= \int_0^\infty\int_0^\infty f_{\Delta F}(e,t)\log\left(\frac{f_{\Delta F}(e,t)}{f_W(e,t)}\right)dedt,\\ \int_0^\infty \|f_{\Delta F}(\cdot,t) - f_W(\cdot,t)\|^2_{L^1(\mathbb{R}^+)}dt &= \int_0^\infty\left(\int_0^\infty |f_{\Delta F}(e,t) - f_W(e,t)|de\right)^2 dt.\end{aligned} \tag{5.58}$$

Small values of these global quantities indicate the densities $f_{\Delta F}(e,t)$ and $f_W(e,t)$ agree at all times t, i.e., the energy dissipated as the GBCD evolves is statistically equivalent to the energy dissipated as the Fokker-Planck solution evolves. Thus, small values of the quantities (5.58) will statistically validate the Fokker-Planck model of GBCD evolution.

5.6.2 Convergence of the energetic statistics in a continuum limit

To estimate $f_{\Delta F}(e,t)$ and $f_W(e,t)$ we use empirical histograms of two-dimensional energy×time samples as follows. Multiple (independent) realizations $r = 1, \ldots, R$ of the 1D coarsening simulation are generated, each realization using the same number of initial grain boundaries n. For each realization r, we estimate $\sigma^{(r)}$ using the relative entropy procedure of §5.4.1, and then collect the statistics $\{(t_j^{\mathrm{qf}})^{(r)}\}_{j=0}^{N}$, $\{\Delta F_j^{(r)}\}_{j=0}^{N-1}$ and $\{W_j^{(r)}\}_{j=0}^{N-1}$. We uniformly discretize a suitable energy×time window $E \times T \subset \mathbb{R}^2$ (see Remark 13) with mesh given by $[E_i, T_k]$ for $i = 1, \ldots, I$ and $k = 1, \ldots, K$. Finally, we form the density estimators $\widehat{f}_{\Delta F}(e,t)$ and $\widehat{f}_W(e,t)$ as normalized histograms of the two-dimensional energy×time samples $(\Delta F_j^{(r)}, (t_j^{\mathrm{qf}})^{(r)})$ and $(W_j^{(r)}, (t_j^{\mathrm{qf}})^{(r)})$, respectively. More precisely, for $(e,t) \in [E_i, E_{i+1}] \times [T_k, T_{k+1}]$ we define

$$\begin{aligned} \widehat{f}_{\Delta F}(e,t) &:= \frac{1}{H(t)} \sum_{r=1}^{R} \sum_{j=0}^{N-1} \mathcal{I}\big(\Delta F_j^{(r)} \in [E_i, E_{i+1}]\big) \mathcal{I}\big((t_j^{\mathrm{qf}})^{(r)} \in [T_k, T_{k+1}]\big), \\ \widehat{f}_W(e,t) &:= \frac{1}{H(t)} \sum_{r=1}^{R} \sum_{j=0}^{N-1} \mathcal{I}\big(W_j^{(r)} \in [E_i, E_{i+1}]\big) \mathcal{I}\big((t_j^{\mathrm{qf}})^{(r)} \in [T_k, T_{k+1}]\big), \end{aligned} \tag{5.59}$$

where $\mathcal{I}$ is the indicator function (5.5) and the normalization is

$$H(t) = \sum_{j=0}^{N-1} \sum_{r=1}^{R} \mathcal{I}\big((t_j^{\mathrm{qf}})^{(r)} \in [T_k, T_{k+1}]\big) \quad \text{for } t \in [T_k, T_{k+1}].$$

This normalization ensures that for any $t \in T$ we have

$$\int_E \widehat{f}_{\Delta F}(e,t)de = \int_E \widehat{f}_W(e,t)de = 1.$$

Remark 13 (The energy×time window) *We use the energy×time window* $E \times T = [0, E_{max}] \times [T_{min}, T_{max}]$. *For our simulations, we use a nonzero initial time* $T_{min} = 0.1$ *as we wish to ignore transient startup effects of the 1D coarsening simulation (see our discussion in §5.5.4). This time corresponds to approximately* 20% *initial grain boundary removal in all realizations* $r = 1, \ldots, R$. *Similarly, we use* $T_{max} = 1$ *as we wish to ignore any effects of using a finite-sized grain boundary system, and these are most prominent at later times. The maximal energy* $E_{max} = 2.5 \times 10^{-3}$ *is simply chosen larger than any of the observations* $\Delta F_j^{(r)}$ *and* $W_j^{(r)}$, *i.e.,*

$$E_{max} > \max_{j,r} \Delta F_j^{(r)} \quad \text{and } E_{max} > \max_{j,r} \Delta F_j^{(r)}.$$

In Figure 5.10 we display the density estimators $\widehat{f}_{\Delta F}(e,t)$ and $\widehat{f}_W(e,t)$ computed for different numbers of initial grain boundaries $n = 2^{15}, 2^{16}, 2^{17}, 2^{18}$. Each estimator is generated from $R = 1000$ independent realizations of the 1D coarsening simulation. At least to the eye, the densities appear similar. For each n, we compute the KL

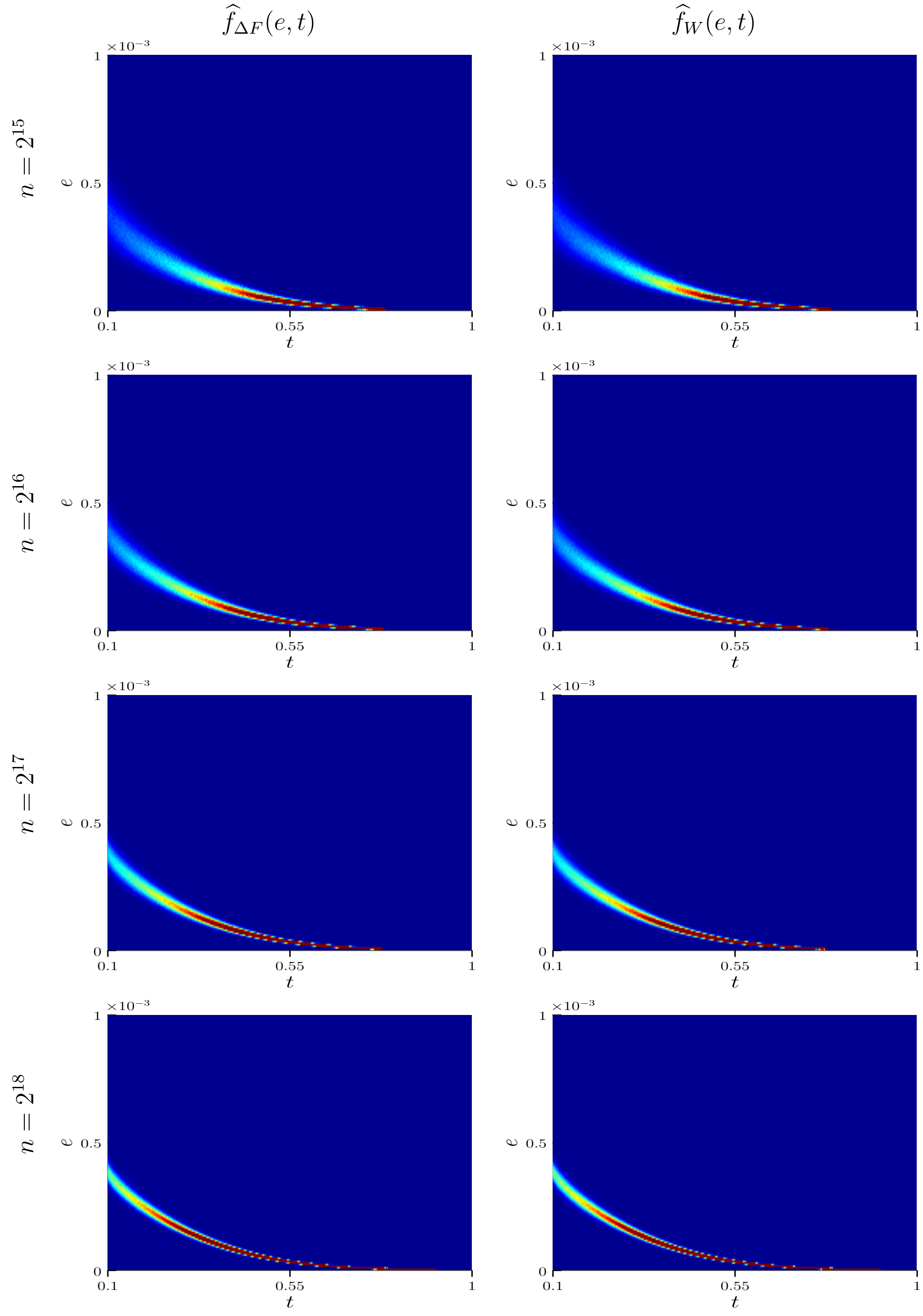


Figure 5.10. Density estimators (left) $\widehat{f}_{\Delta F}(e,t)$ and (right) $\widehat{f}_{W}(e,t)$ computed using (5.59), $n = 2^x$ initial grain boundaries and $R = 1000$ realizations.

Table 5.1. The global quantifier values (5.58) for density estimators $\widehat{f}_{\Delta F}(e,t)$ and $\widehat{f}_W(e,t)$ for grain boundary systems with n initial grain boundaries. Here we observe convergence as the number of initial grain boundaries n is increased, suggesting a better statistical agreement between the energy dissipation samples and W2 distance samples as $n \to \infty$.

n	$\int_T \Phi(\widehat{f}_{\Delta F}, \widehat{f}_W)$	$\int_T \Vert\widehat{f}_{\Delta F} - \widehat{f}_W\Vert^2_{L^1(E)}$
2^{15}	0.045322	0.037916
2^{16}	0.023644	0.018025
2^{17}	0.014968	0.010740
2^{18}	0.009261	0.007013

divergence curve $\Phi(\widehat{f}_{\Delta F}(\cdot,t), \widehat{f}_W(\cdot,t))$ and the L^1 error curve $\Vert\widehat{f}_{\Delta F}(\cdot,t) - \widehat{f}_W(\cdot,t)\Vert^2_{L^1(E)}$. Finally, we integrate the KL divergence and L^1 error curves over the time interval T to compute the global quantification values $\int_T \Phi(\widehat{f}_{\Delta F}, \widehat{f}_W)dt$ and $\int_T \Vert\widehat{f}_{\Delta F} - \widehat{f}_W\Vert^2_{L^1}dt$. The results are displayed in Table 5.1. Here we notice a striking phenomenon: the global values $\int_T \Phi(\widehat{f}_{\Delta F}, \widehat{f}_W)dt$ and $\int_T \Vert\widehat{f}_{\Delta F} - \widehat{f}_W\Vert^2_{L^1(E)}dt$ exhibit convergence toward zero as the number of initial grain boundaries n increases. Moreover, this convergence is approximately linear in n. We summarize these findings in the following conjecture.

Conjecture 1 (W2 dissipation in a continuum limit) *Recall n as the number of initial grain boundaries and the energy×time window $E \times T = [0, E_{max}] \times [T_{min}, T_{max}]$ defined in Remark 13. Table 5.1 provides evidence of global convergence of $f_{\Delta F}(e,t)$ to $f_W(e,t)$:*

$$0 \leq \frac{1}{2}\int_T \Vert f_{\Delta F}(\cdot,t) - f_W(\cdot,t)\Vert^2_{L^1(E)}dt \leq \int_T \Phi(f_{\Delta F}(\cdot,t), f_W(\cdot,t))dt \xrightarrow{n\to\infty} 0.$$

Note that in Conjecture 1 we have used the true densities $f_{\Delta F}(e,t)$ and $f_W(e,t)$ since the estimators $\widehat{f}_{\Delta F}(e,t)$ and $\widehat{f}_W(e,t)$ can be made arbitrarily close to the true densities by using sufficiently many realizations R while refining the energy×time grid discretization (see, e.g., [15]).

We interpret Conjecture 1 and the observed convergence in KL divergence Φ as follows. Assume the energy dissipated by the grain boundary system at time t is distributed according to a density $f(e,t;\theta_0)$ where θ_0 is an unknown parameter, but the form of the density is otherwise known. Minimizing the KL relative entropy

$$\min_{\theta} \Phi(f(\cdot,t;\theta_0), f(\cdot,t;\theta)) = \min_{\theta} \int_E f(e,t;\theta_0) \log\left(\frac{f(e,t;\theta_0)}{f(e,t;\theta)}\right) de$$

with respect to the parameter θ, is equivalent to finding the maximum likelihood estimator $\widehat{\theta}$ for θ_0. That is, the parameter $\widehat{\theta}$ that most accurately predicts the true density based on observed data (see, e.g., [13]). If we think of θ as parameterizing various dissipation mechanisms and their associated dynamics, a minimum of the relative entropy characterizes the most likely dissipation mechanism and its associated dynamics. Since zero is the unique global minimum of the relative entropy $\Phi(f_{\Delta F}(\cdot,t), f_W(\cdot,t))$, if $\Phi(f_{\Delta F}(\cdot,t), f_W(\cdot,t)) \to 0$ as $n \to \infty$, the most likely dissipation mechanism for the GBCD will be W2 dissipation. Therefore, in a continuum limit, Fokker-Planck dynamics would be the most likely description of the GBCD evolution.

Another interpretation of Conjecture 1 follows from the observed convergence in L^1 of the probability densities $f_{\Delta F}(e,t)$ and $f_W(e,t)$. Owing to Scheffe [34], convergence of densities in L^1 implies convergence in distribution of the associated random variables. We have observed evidence that the distribution of the energy dissipation samples ΔF converges to the distribution of W2 distance samples W, in a continuum limit $n \to \infty$. This convergence would imply the energy dissipation of the GBCD to be statistically equivalent to energy dissipation of the Fokker-Planck solution (5.17), in the limit $n \to \infty$.

5.7 Conclusion

The empirical GBCD (5.4) harvested from the 1D coarsening simulation described in §5.2 possesses many traits of the solution of the Fokker-Planck equation (5.17). In this chapter, we have described and performed validation tests that support this claim. We first estimated the diffusive coefficient σ and set time scales for the GBCD by appealing to the theory of Fokker-Planck solutions. This allowed us to directly compare the GBCD with the corresponding Fokker-Planck solution where we found qualitative agreement. We then studied W2 gradient flows and found a dissipation identity the GBCD must satisfy if it indeed evolves according to the Fokker-Planck PDE. We determined the GBCD approximately satisfies this identity both qualitatively and quantitatively. Moreover, we found as the number of initial grain boundaries n increases, the GBCD satisfies this identity more accurately. We interpreted this as indicating that the Fokker-Planck equation (5.17) is a more accurate model of the GBCD evolution in a continuum or many particles limit.

Current work on this project has shifted toward analyzing the dissipative behavior of GBCDs harvested from 2D simulations. The 2D simulation is outlined in [5, 17], and in our work we use a modified code originally developed by E. Eggeling and S. Ta'asan [17]. Although the simulation is two-dimensional, the harvested GBCD is still a one-dimensional probability density (assuming there is no directional dependence in growth), and so all of the analysis provided in this chapter can be applied. The simulations are computationally expensive, and at this time only relatively small simulations are available ($\sim 10,000$ initial grains). However, even GBCDs harvested from these small-scale simulations exhibit similar qualitative behaviors as those harvested from the 1D simulations. In Figure 5.11 we show results of the relative entropy procedure (see §5.4.1) applied to GBCDs harvested from grain boundary systems having $10,000$ and $20,000$ initial grains. Here we see qualitative agreement between the final time GBCDs and the optimal Boltzmann distributions. We can also use the energy dissipation estimate (5.54) for these GBCDs and we find it indeed holds (approximately). We depict this estimate in Figure 5.12 for the two grain boundary systems of $10,000$ and $20,000$ initial grains. These preliminary results again suggest the Fokker-Planck equation may be a good approximation of the evolution of GBCDs harvested from 2D simulations. However, the system sizes here are too small to truly validate the Fokker-Planck model; thus, we hope to soon apply these validation techniques to GBCDs harvested from much larger simulations.

In the near future, we hope to address time scalings (see §5.4.2) using other methods, e.g., with dimensional analysis. We are also interested in studying the case when grains grow in a preferred direction and the interfacial energy density ψ depends on grain boundary orientation, i.e., $\psi = \psi(\alpha, \theta)$ where $\mathbf{n} = (\cos(\theta), \sin(\theta))$ is the grain boundary normal direction. Extending the present theory and techniques to this case is an important step since these energy densities are more representative of the dynamics of true polycrystalline materials.

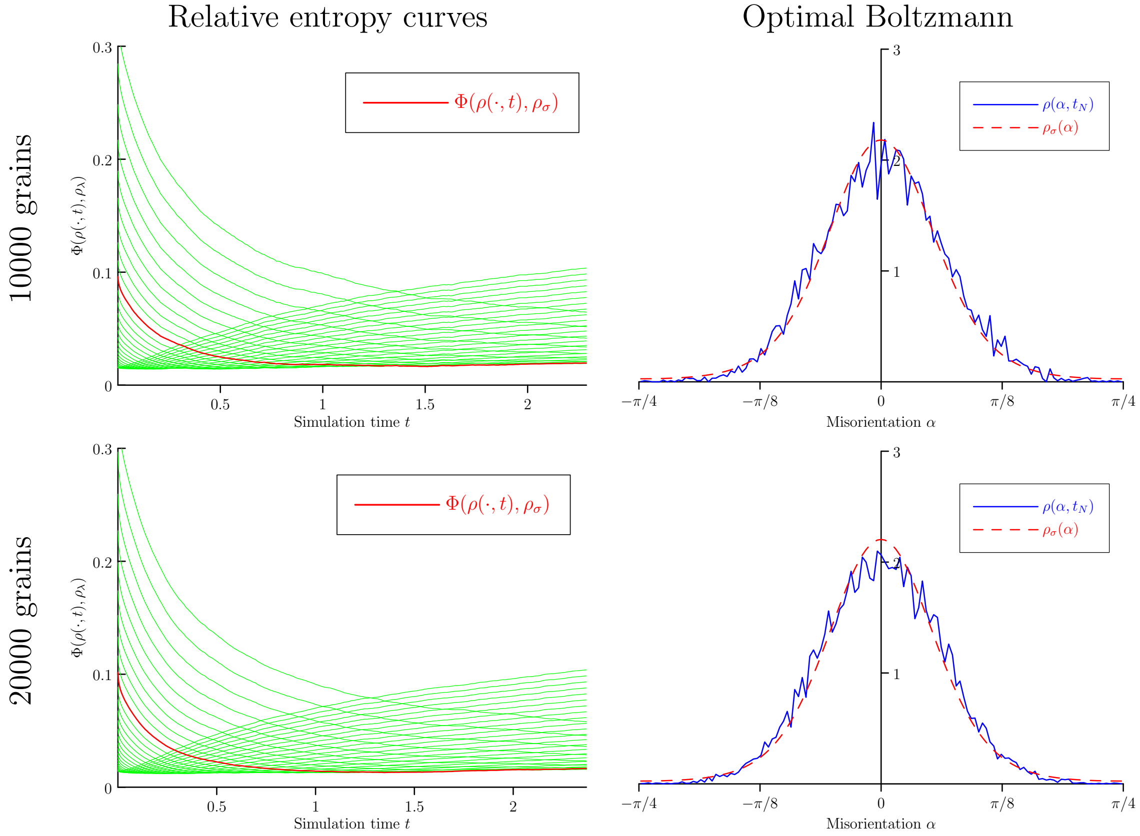


Figure 5.11. Relative entropy results for a GBCD harvested from a 2D simulation with (top) $10,000$ and (bottom) $20,000$ initial grains. The left column shows KL relative entropy curves $\Phi(\rho,\rho_\lambda)$ for 30 trial values of $\lambda \in [0.08, 0.2]$. The right column compares the final time GBCD $\rho(\alpha,t_N)$ (blue) with the optimal Boltzmann distribution $\rho_\sigma(\alpha)$ (red dashed). The optimal value is estimated as $\sigma = 0.113$ for both $n = 10,000$ (top) and $n = 20,000$ (bottom).

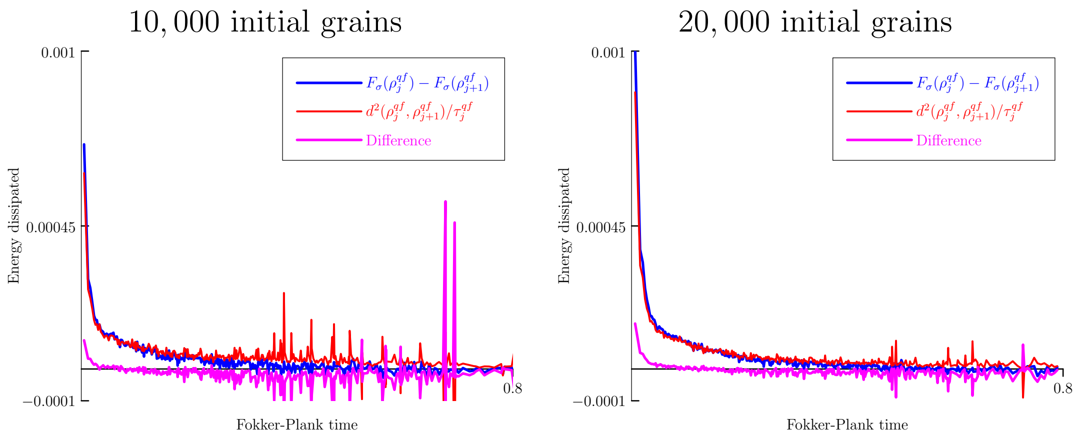


Figure 5.12. Dissipation validation of GBCDs harvested from a 2D simulation with (left) 10, 000 and (right) 20, 000 initial grains. The energy dissipation $F_\sigma(\rho_j^{\mathrm{qf}}) - F_\sigma(\rho_{j+1}^{\mathrm{qf}})$ (blue) is compared with the W2 distances $d^2(\rho_j^{\mathrm{qf}}, \rho_{j+1}^{\mathrm{qf}})/\tau_j^{\mathrm{qf}}$ (red). Their difference is depicted in magenta.

5.8 References

[1] B. Adams, S. Ta'Asan, D. Kinderlehrer, I. Livshits, D. Mason, C.-T. Wu, W. Mullins, G. Rohrer, A. Rollett, and D. Saylor, *Extracting grain boundary and surface energy from measurement of triple junction geometry*, Interface Science, 7 (1999), pp. 321–337.

[2] L. Ambrosio and N. Gigli, *Modelling and Optimisation of Flows on Networks: Cetraro, Italy 2009, Editors: Benedetto Piccoli, Michel Rascle*, Springer Berlin Heidelberg, Berlin, Heidelberg, 2013, ch. A User's Guide to Optimal Transport, pp. 1–155.

[3] L. Ambrosio, N. Gigli, and G. Savaré, *Gradient Flows: in Metric Spaces and in the Space of Probability Measures*, Lectures in math, Birkhäuser, 2005.

[4] K. Barmak, E. Eggeling, M. Emelianenko, Y. Epshteyn, D. Kinderlehrer, R. Sharp, and S. Ta'asan, *Critical events, entropy, and the grain boundary character distribution*, Phys. Rev. B, 83 (2011), p. 134117.

[5] ——, *An entropy based theory of the grain boundary character distribution*, Discrete Contin. Dyn. Syst., 30 (2011), pp. 427–454.

[6] ——, *A theory and challenges for coarsening in microstructure*, in Analysis and numerics of partial differential equations, vol. 4 of Springer INdAM Ser., Springer, Milan, 2013, pp. 193–220.

[7] ——, *Materials microstructure: Entropy and curvature-driven coarsening*, RIMS, Research Institute for Mathematical Sciences, U. Kyoto, (2014), pp. 71–91.

[8] ——, *Recent developments in material microstructure: a theory of coarsening*, in Symposium NN - Mathematical and Computational Aspects of Materials Science, vol. 1753 of MRS Proceedings, 2015.

[9] K. Barmak, E. Eggeling, M. Emelianenko, Y. Epshteyn, D. Kinderlehrer, and S. Ta'asan, *Geometric growth and character development in large metastable networks*, Rend. Mat. Appl. (7), 29 (2009), pp. 65–81.

[10] J.-D. Benamou and Y. Brenier, *A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem*, Numerische Mathematik, 84 (2000), pp. 375–393.

[11] F. Bolley and C. Villani, *Weighted Csiszár-Kullback-Pinsker inequalities and applications to transportation inequalities*, in Annales de la Faculté des sciences de Toulouse: Mathématiques, vol. 14(3), 2005, pp. 331–352.

[12] F. Cleri, S. R. Phillpot, and D. Wolf, *Atomistic simulations of integranular fracture in symmetric-tilt grain boundaries*, Interface Science, 7 (1999), pp. 45–55.

[13] T. Cover and J. Thomas, *Elements of Information Theory*, A Wiley-Interscience Publication, Wiley, 2006.

[14] I. Csiszár and P. Shields, *Information theory and statistics: a tutorial*, Foundations and Trends in Communications and Information Theory, 1 (2004), pp. 417–528.

[15] A. DasGupta, *Asymptotic Theory of Statistics and Probability*, Springer Texts in Statistics, Springer New York, 2008.

[16] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications*, Stochastic Modelling and Applied Probability, Springer, 2009.

[17] E. Eggeling and S. Ta'asan, *2D curvature driven grain growth.* C++ Code, 2008.

[18] M. Elsey, S. Esedoglu, and P. Smereka, *Diffusion generated motion for grain growth in two and three dimensions*, Journal of Computational Physics, 228 (2009), pp. 8015–8033.

[19] ——, *Large-scale simulation of normal grain growth via diffusion-generated motion*, Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences, 467 (2011), pp. 381–401.

[20] C. W. Gardiner, *Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences*, vol. 13 of Springer Series in Synergetics, Springer-Verlag, Berlin, third ed., 2004.

[21] E. P. Gorzkowski, T. Sano, C.-S. Kim, G. S. Rohrer, H. M. Chan, and M. P. Harmer, *Changes in the distribution of interfaces in PMN-35 mol% PT as a function of time*, Zeitschrift für Metallkunde, 96 (2005), pp. 207–210.

[22] A. Haslam, S. Phillpot, D. Wolf, D. Moldovan, and H. Gleiter, *Mechanisms of grain growth in nanocrystalline fcc metals by molecular-dynamics simulation*, Materials Science and Engineering: A, 318 (2001), pp. 293–312.

[23] R. Jordan, D. Kinderlehrer, and F. Otto, *The variational formulation of the Fokker-Planck equation*, SIAM J. Math. Anal, 29 (1998), pp. 1–17.

[24] D. Kinderlehrer and C. Liu, *Evolution of grain boundaries*, Mathematical Models and Methods in Applied Sciences, 11 (2001), pp. 713–729.

[25] D. Kinderlehrer, I. Livshits, G. S. Rohrer, S. Ta'asan, and P. Yu, *Mesoscale simulation of the evolution of the grain boundary character distribution*, in Materials Science Forum, vol. 467, Trans Tech Publ, 2004, pp. 1063–1068.

[26] D. Kinderlehrer, I. Livshits, and S. Ta'asan, *A variational approach to modeling and simulation of grain growth*, SIAM J. Sci. Comput., 28 (2006), pp. 1694–1715 (electronic).

[27] D. Kinderlehrer and N. J. Walkington, *Approximation of parabolic equations using the Wasserstein metric*, ESAIM: Mathematical Modelling and Numerical Analysis, 33 (2010), pp. 837–852.

[28] M. Peletier, *Energies, gradient flows, and large deviations: a modelling point of view.* Lecture notes: Mathematical Principles for and Advances in Continuum Mechanics, 2012.

[29] ——, *Variational modelling: Energies, gradient flows, and large deviations*, Preprint, (2014).

[30] S. Phillpot, P. Keblinski, D. Wolf, and F. Cleri, *Synthesis and characterization of a polycrystalline ionic thin film by large-scale molecular-dynamics simulation*, Interface Science, 7 (1999), pp. 15–31.

[31] H. Risken, *The Fokker-Planck Equation: Methods of Solution and Applications*, vol. 18 of Springer Series in Synergetics, Springer-Verlag, Berlin, second ed., 1989.

[32] G. S. Rohrer, *Influence of interface anisotropy on grain growth and coarsening*, Annual Review of Materials Research, 35 (2005), pp. 99–126.

[33] A. D. Rollett, S.-B. Lee, R. Campman, and G. Rohrer, *Three-dimensional characterization of microstructure by electron back-scatter diffraction*, Annual Review of Materials Research, 37 (2007), pp. 627–658.

[34] H. Scheffe, *A useful convergence theorem for probability distributions*, Ann. Math. Statist., 18 (1947), pp. 434–438.

[35] C. Villani, *Topics in Optimal Transportation*, Graduate studies in mathematics, American Mathematical Society, 2003.

CHAPTER 6

CONCLUSION

6.1 Inverse Born series

In Chapter 2, we generalized the inverse Born series convergence and stability proofs of Moskow and Schotland [10] to nonlinear mappings between Banach spaces. This generalization allows for the inverse Born series machinery to be used in a given physical setting by showing a single problem-specific bound. Moreover, this generalization highlights the connection between inverse Born series and classical Taylor and Neumann series. Indeed we find that for analytic functions, Born series expansions and Taylor series expansions are equivalent up to a symmetrization. We also discovered a new class of iterative methods, which we refer to as restarted inverse Born series RIBS(k) since they are based on truncating and restarting inverse Born series. These methods are closely related to well-known iterative methods based on truncated Taylor expansions, e.g., the Gauss-Newton and Chebyshev-Halley methods.

In the future, we plan to study the RIBS(k) methods in more detail. At this point, we have only proposed the methods and performed a simple numerical convergence study. A rigorous convergence analysis needs to be carried out for the methods to be used in practice. Furthermore, we currently only expect these methods to be locally convergent owing to the local convergence of inverse Born series themselves. Thus, we hope to pursue globalization strategies to extend the convergence regions of these methods to arbitrary starting guesses.

Another interest we have is to characterize improvements that are made to a linear model (i.e., Born series of order 1) by considering a k-linear model (i.e., Born series of order k). In the linear case, i.e., when solving for h from $d = a_1(h)$ where a_1 is a linear mapping, it is well-known the SVD of a_1 provides the most efficient basis to represent h with. The best (linear) inversion we can hope for is the projection of h onto the singular vectors with nonzero singular values (or above measurement noise thresholds). When considering a k-linear model, i.e., $\sum_{j=1}^{k} a_j(h^{\otimes j})$ where a_j are linear mappings of the tensor products $h^{\otimes j}$, we may find a different or improved basis to represent h with. This improved basis might be characterized by lifting the problem to a higher dimension and considering bases of higher-order tensors.

6.2 Intensity-only imaging

In Chapter 3, we developed a novel method to image scatterers in a homogeneous medium from intensity-only measurements. The method uses N point sources, a single receiver and a specific illumination strategy of source pairs (i.e., sending the same signal from two different locations simultaneously). With a smallness assumption on scattered fields, we showed that a projection of full waveform data onto a known subspace can be recovered by solving, in the least squares sense, a linear system involving the intensity data measured at the receiver for multiple source pair experiments. Using a stationary phase argument, we showed this imperfect knowledge of the full waveform data does not affect Kirchhoff migration images in a high frequency limit. Thus, the least squares solution of the linear system can be used as input data for Kirchhoff migration.

In Chapter 4, we developed a similar imaging method for a setup consisting of N receivers and a single point source. With a smallness assumption on the scattered field, we again showed a projection of full waveform data can be recovered by solving, in the least squares sense, a linear system that involves the intensity data measured at the N receivers. In this case, the linear system is underdetermined of size $N \times 2N$, and has an N-dimensional nullspace that can be characterized in terms of the incident wavefield. We showed this nullspace does not affect Kirchhoff images for high frequencies, and thus, as before, we can use the least squares solution of the linear system as input data to image with Kirchhoff migration. This generalization of our previous work thus relaxes the illumination strategy from using source pairs to using a single source. Moreover, the least squares solution has a simple expression in terms of the incident field, and can be interpreted as a preprocessing step to use Kirchhoff migration with intensity data (see also [7]).

We generalized each of our imaging methods to using stochastic illuminations with autocorrelations measured at the receiver locations (instead of intensities). Since correlations are robust with respect to measurement noise, we expect these methods to perform well in situations having a low signal-to-noise ratio, e.g., nonintrusive imaging when probing wavefields have the same magnitude as background ambient noise.

In future work, we would like to apply similar imaging principles to different physical problems. For example, optical stellar interferometry (see, e.g., [9, 11]) is an inverse source problem (i.e., one wishes to determine the location and/or magnitude of an unknown wave source), where measured data consists of correlations of intensity measurements. In the frequency domain, these measurements are quartic in the wavefield rather than quadratic as in the intensity-only case. Thus, we may be able to image with a classic imaging functional (e.g., matched field) by also preprocessing the intensity correlation data, but new machinery and analysis will likely need to be developed.

6.3 Evolution of the GBCD

In Chapter 5, we provided numerical validations of an entropy-based theory for the evolution of the GBCD first proposed and studied in [1, 2, 3, 6]. This theory suggests the GBCD evolves according to a Fokker-Planck equation. Our first numerical validation was a repeat of the relative entropy procedure developed in [1, 2, 3, 4, 5, 6]. Here, we compared the steady-state Fokker-Planck solution (a Boltzmann distribution) with the GBCD harvested from the final time of simulation. We found, as did the original authors [1, 2, 3, 4, 5, 6], a qualitative agreement between the two densities, giving a first validation of their entropy-based theory. Next, we developed novel time-matching procedures to assign Fokker-Planck times to the empirical GBCD. We then compared the Fokker-Planck solution with the time-fitted GBCD at intermediate times. Here we again found a qualitative agreement between the densities, a further validation of the Fokker-Planck model of GBCD evolution.

Our final validation tool was derived by exploring the energy dissipation of W2 gradient flows. Here we found a free energy dissipation identity that must be satisfied if the GBCD evolves as a Fokker-Planck solution. We determined the GBCD approximately satisfies this identity, again validating the Fokker-Planck model of GBCD evolution. Moreover, by taking a statistical approach and sampling many realizations of the GBCD, we found numerical evidence that the Fokker-Planck model of GBCD evolution improves as the number of grain boundaries tends to infinity (i.e., in a continuum limit).

Among our ideas for future work, we are most interested in applying our same validation principles to GBCDs harvested from 2D simulations. In simulations where the energy density ψ depends only on misorientations α and not on normal directions θ, the harvested GBCDs are still one-dimensional and our validation principles can still be applied. This could determine if the Fokker-Planck model accurately describes GBCD evolution in this setting, and perhaps aid in development of a theory for the evolution of such GBCDs. We also plan to investigate GBCDs harvested from simulations having an energy density ψ that does depend on θ, since these are more exemplary of those encountered in real materials. In these situations the harvested GBCDs are two-dimensional, further complicating the analysis. Moreover, numerically computing W2 distances between multivariate densities is currently not possible except for special cases (see [8, Section 5]). Since our energy dissipation principle relies on such computations, perhaps an alternate energetic validation principle should be considered for 2D GBCDs.

6.4 References

[1] K. Barmak, E. Eggeling, M. Emelianenko, Y. Epshteyn, D. Kinderlehrer, R. Sharp, and S. Ta'asan, *Critical events, entropy, and the grain boundary character distribution*, Phys. Rev. B, 83 (2011), p. 134117.

[2] ——, *An entropy based theory of the grain boundary character distribution*, Discrete Contin. Dyn. Syst., 30 (2011), pp. 427–454.

[3] ——, *A theory and challenges for coarsening in microstructure*, in Analysis and numerics of partial differential equations, vol. 4 of Springer INdAM Ser., Springer, Milan, 2013, pp. 193–220.

[4] ——, *Materials microstructure: Entropy and curvature-driven coarsening*, RIMS, Research Institute for Mathematical Sciences, U. Kyoto, (2014), pp. 71–91.

[5] ——, *Recent developments in material microstructure: a theory of coarsening*, in Symposium NN - Mathematical and Computational Aspects of Materials Science, vol. 1753 of MRS Proceedings, 2015.

[6] K. Barmak, E. Eggeling, M. Emelianenko, Y. Epshteyn, D. Kinderlehrer, and S. Ta'asan, *Geometric growth and character development in large metastable networks*, Rend. Mat. Appl. (7), 29 (2009), pp. 65–81.

[7] Z. Chen and G. Huang, *Phaseless imaging by reverse time migration: Acoustic waves*, arXiv, (2015).

[8] D. Kinderlehrer and N. J. Walkington, *Approximation of parabolic equations using the Wasserstein metric*, ESAIM: Mathematical Modelling and Numerical Analysis, 33 (2010), pp. 837–852.

[9] A. Labeyrie, S. Lipson, and P. Nisenson, *An Introduction to Optical Stellar Interferometry*, Cambridge University Press, 2006.

[10] S. Moskow and J. C. Schotland, *Convergence and stability of the inverse scattering series for diffuse waves*, Inverse Problems, 24 (2008), p. 065005.

[11] P. D. Nuñez, *Towards optical intensity interferometry for high angular resolution stellar astrophysics*, PhD thesis, University of Utah, Salt Lake City, UT, August 2012.