# Simultaneous-source seismic data acquisition and processing with compressive sensing

by

Haneet Wason

B.Sc., University of Calgary, 2010

A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF

**Doctor of Philosophy**

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES
(Geophysics)

The University of British Columbia
(Vancouver)

August 2017

# Abstract

The work in this thesis adapts ideas from the field of compressive sensing (CS) that lead to new insights into acquiring and processing seismic data, where we can fundamentally rethink how we design seismic acquisition surveys and process acquired data to minimize acquisition- and processing-related costs. Current efforts towards dense source/receiver sampling and full azimuthal coverage to produce high-resolution images of the subsurface have led to the deployment of multiple sources across survey areas. A step ahead from multisource acquisition is simultaneous-source acquisition, where multiple sources fire shots at near-simultaneous/random times resulting in overlapping shot records, in comparison to no overlaps during conventional sequential-source acquisition. Adoption of simultaneous-source techniques has helped to improve survey efficiency and data density. The engine that drives simultaneous-source technology is simultaneous-source separation — a methodology that aims to recover conventional sequential-source data from simultaneous-source data. This is essential because many seismic processing techniques rely on dense and periodic (or regular) source/receiver sampling. We address the challenge of source separation through a combination of tailored simultaneous-source acquisition design and sparsity-promoting recovery via convex optimization using $\ell_1$ objectives. We use CS metrics to investigate the relationship between marine simultaneous-source acquisition design and data reconstruction fidelity, and consequently assert the importance of randomness in the acquisition system in combination with an appropriate choice for a sparsifying transform (i.e., curvelet transform) in the reconstruction algorithm. We also address the challenge of minimizing the cost of expensive, dense, periodically-sampled and replicated time-lapse surveying and data processing by adapting ideas from distributed compressive sensing. We show that compressive randomized time-lapse surveys need not be replicated to attain acceptable levels of data repeatability, as long as we know the shot positions (post acquisition) to a sufficient degree of accuracy. We conclude by comparing sparsity-promoting and rank-minimization recovery techniques for marine simultaneous-source separation, and demonstrate that recoveries are comparable; however, the latter approach readily scales to large-scale seismic data and is computationally faster.

# Lay Summary

Adapting ideas from the field of compressive sensing, we design economic compressive (or subsampled) randomized simultaneous-source acquisitions and develop processing techniques to address the challenge of source separation. We recover dense periodic (or regular) conventional sequential-source data from subsampled randomized simultaneous-source data via structure promotion, i.e., curvelet-domain sparsity or low rank. Adapting ideas from distributed compressive sensing, we show that compressive randomized time-lapse surveys need not be replicated to attain acceptable levels of data repeatability, and since irregular spatial sampling is inevitable in the real world, it would be better to focus on knowing what the shot positions were (post acquisition) to a sufficient degree of accuracy, than aiming to replicate them. In a nutshell, compressive randomized simultaneous-source acquisitions and subsequent processing techniques provide flexibility in acquisition geometries for better survey-area coverage, and speedup acquisition — effectively minimizing acquisition- and processing-related costs.

# Preface

The work presented in this thesis has been carried out at the Department of Earth, Ocean and Atmospheric Sciences at The University of British Columbia, Vancouver, Canada, under the supervision of Professor Felix J. Herrmann as part of the Seismic Laboratory for Imaging and Modelling (SLIM).

A version of Chapter 2 has been published in a geophysical magazine (Herrmann, F. J., H. Wason, and T. T. Y. Lin, 2011, Compressive sensing in seismic exploration: an outlook on a new paradigm: CSEG Recorder, 36. Part 1 [April Edition], 19–33. Part 2 [June Edition], 34–39). Felix wrote the theoretical sections and parts of the experimental section. I conducted experiments for land and marine simultaneous-source academic case study and wrote the corresponding section. Tim was involved in the setup of these experiments.

A version of Chapter 3 has been published in a geophysical journal (Mansour, H., H. Wason, T. T. Y. Lin, and F. J. Herrmann, 2012, Randomized marine acquisition with compressive sampling matrices: Geophysical Prospecting, 60, 648–662). The manuscript was written jointly. Hassan led the theoretical sections and corresponding experiments, and I led the seismic experimental sections. Tim and Felix were involved in editing the manuscript.

A version of Chapter 4 has been published as an expanded abstract (Wason, H., and F. J. Herrmann, 2013, Time-jittered ocean bottom seismic acquisition: SEG Technical Program Expanded Abstracts, 32, 1–6). I was the lead investigator and manuscript composer.

A version of Chapter 5 has been published in a geophysical journal (Oghenekohwo, F., H. Wason, E. Esser, and F. J. Herrmann, 2017, Low-cost time-lapse seismic with distributed compressive sensing-Part 1: exploiting common information among the vintages, Geophysics, 82, P1–P13). The manuscript was written jointly. I conducted experiments for the synthetic seismic case study and wrote the corresponding section. F. Oghenekohwo provided the relevant literature of the subject, conducted the stylized experiments and wrote the corresponding section. F. Herrmann wrote most of the theoretical sections with significant inputs from F. Oghenekohwo and I. E. Esser was involved in the algorithm evaluation. Given the equivalent contributions made by F. Oghenekohwo and I, we mutually agreed to include this chapter verbatim in both our dissertations. This chapter appears as Chapter 2 in F. Oghenekohwo's dissertation. F. Oghenekohwo has granted permission for this chapter to also appear in my dissertation.

A version of Chapter 6 has been published in a geophysical journal (Wason, H., F. Oghenekohwo, and F. J. Herrmann, 2017, Low-cost time-lapse seismic with distributed compressive sensing-Part 2: impact on repeatability, 82, P15–P30). I was the lead investigator and manuscript composer. F. Oghenekohwo simulated the monitor model for the SEAM Phase 1 model. F. Herrmann was involved in editing the manuscript. Since this chapter is a necessary extension of Chapter 5, this chapter also appears verbatim as Chapter 3 in F. Oghenekohwo's dissertation. I have granted permission for this chapter to also appear in F. Oghenekohwo's dissertation.

A version of Chapter 7 has been published in a geophysical journal (Kumar, R., H. Wason, and F. J. Herrmann, 2015, Source separation for simultaneous towed-streamer marine acquisition — A compressed sensing approach, Geophysics, 80, WD73–WD88). I encountered the source-separation problem for towed-streamer acquisitions during an internship. Subsequently, the problem was investigated at SLIM. The manuscript was written jointly with equal contribution from Rajiv and I. I led the theoretical section on sparsity-promoting technique and conducted the corresponding experiments. Rajiv led the theoretical section on rank-minimization technique and conducted the corresponding experiments. Felix was involved in editing the manuscript. Given the equivalent contributions made by R. Kumar and I, we mutually agreed to include this chapter verbatim in both our dissertations. This chapter appears as Chapter 4 in R. Kumar's dissertation. R. Kumar has granted permission for this chapter to also appear in my dissertation.

MATLAB and its parallel computing toolbox has been used to prepare all the examples in this thesis. The SPG$\ell_1$ solver is a public toolbox provided by Ewout van den Berg and Michael P. Friedlander. The SPG-LR solver is coded by Aleksandr Y. Aravkin. The Curvelet toolbox is provided by Emmanuel J. Candès, Laurent Demanet, David L. Donoho and Lexing Ying.

# Table of Contents

# List of Tables

# List of Figures

# Glossary

BPDN - Basis pursuit denoise
CS - Compressive sensing
DCS - Distributed compressive sensing
FDCT - Fast discrete curvelet transform
HSS - Hierarchical semi-separable
IRS - Independent recovery strategy
JRM - Joint recovery model (or method)
LASSO - Least absolute shrinkage and selection operator
NFDCT - Nonequispaced fast discrete curvelet transform
NRMS - Normalized root-mean-square
OBC - Ocean-bottom cable
OBN - Ocean-bottom node
RIP - Restricted isometry property
SLO - Simultaneous long offset
S/N - Signal-to-noise ratio
SPG - Spectral projected gradient
SVD - Singular value decomposition

# Acknowledgments

I would first and foremost like to thank my advisor, Professor Dr. Felix J. Herrmann, for giving me a position in the SLIM group, where I have tackled some of the most interesting problems and learnt a great deal from the diverse nature of these problems. I am grateful to him for his support throughout my time in graduate school. I am very fortunate to have been a part of the SLIM group. This experience has most definitely made me a better student.

I would also like to thank Professor Dr. Eldad Haber and Professor Dr. Michael P. Friedlander for serving on my advisory committee and for generously offering their time, support and advice.

My journey through graduate school would not have been memorable if it were not for the people I have worked with. I would like to thank each and every member of the SLIM group, students and postdoctorate fellows, for being a part of my journey that has been nothing short of a roller-coaster ride. The exciting and stimulating interactions with my colleagues has increased my knowledge of scientific research tremendously. I wish them all the very best for their careers.

The SLIM group would not be able to function smoothly without the support of Miranda Joyce and Henryk Modzelewski. A very special thanks to them for always being so concerned, generous and helpful. I truly value their friendship.

I am very grateful to PGS (Weybridge, UK) for the hospitality during my internship. I would like to thank Sverre Brandsberg-Dahl and Andreas Klaedtke for giving me this opportunity. Special thanks to Andreas Klaedtke and Rolf Baardman for great technical discussions. This experience improved my understanding of my research topic and its relevance to the industry immensely.

I would like to thank Dr. Eric Verschuur for providing the Gulf of Suez dataset, which I use in Chapters 2, 3, 4 and 7, and the BG group for providing the synthetic 3D COMPASS velocity model that I use in Chapters 5 and 6. Many thanks to the authors of SPG$\ell_1$, CurveLab, Madagascar, IWAVE, SPG-LR, the Marmousi velocity model and the BP salt model, which I use throughout this work.

This work was financially supported in part by the Natural Sciences and Engineering Research Council of Canada Collaborative Research and Development Grant DNOISE II (CDRP J 375142-08). This research was carried out as part of the SINBAD II project with the support of the member organizations of the SINBAD Consortium. I also wish to acknowledge the SENAI CIMATEC Supercomputing Center for Industrial Innovation, with support from BG Brasil and the Brazilian Authority for Oil, Gas and Biofuels (ANP), for the provision and operation of computational facilities and the commitment to invest in Research and Development.

# Chapter 1

# Introduction

## 1.1 Seismic exploration

Geophysical surveys determine characteristics of the earth's subsurface by measuring the physical differences between rock types or physical discontinuities without seeing them directly by digging, drilling or tunnelling. Geophysical surveys are classified as seismic and non-seismic surveys that include magnetic and electromagnetic surveys, gravitational surveys, seismic surveys, etc. Seismic is perhaps the most commonly used geophysical technique to locate potential oil and natural gas deposits in the geologic structures within the earth. Seismic exploration techniques involve the collection of massive data volumes, where regularly sampled wavefields exhibit up to a 5-dimensional structure (1D for the time dimension × 2D for the receiver positions × 2D for the source positions), and their exploitation during processing.

### 1.1.1 Static vs. dynamic geometries

Seismic surveys can be conducted on onshore (land) and offshore (marine). Land and marine surveys operate on the same basic principles but differ operationally. Both land and marine surveys include a source and a receiver but may differ in the geometry of the receiver system, the density of measurements made over a given area, and the type of source and receiver (or sensor) used (Caldwell and Walker, 2011).

#### Land

Land seismic data acquisition uses primarily two types of seismic sources — vibroseis vehicles (vibrators mounted on trucks) or a low-impact explosive charge — that generate acoustic waves, which propagate deep into the earth. Each time an acoustic wave encounters a change in the rock formation, part of the wave is reflected back to the surface where an array of sensors records the returning sound waves. The receivers are typically geophones, which are like small microphones pushed into the soil to measure the ground motion (Caldwell and Walker, 2011). Since the receivers are fixed on the earth's surface, we refer to this acquisition geometry as the "static" geometry. Figure 1.1 illustrates a land seismic survey.

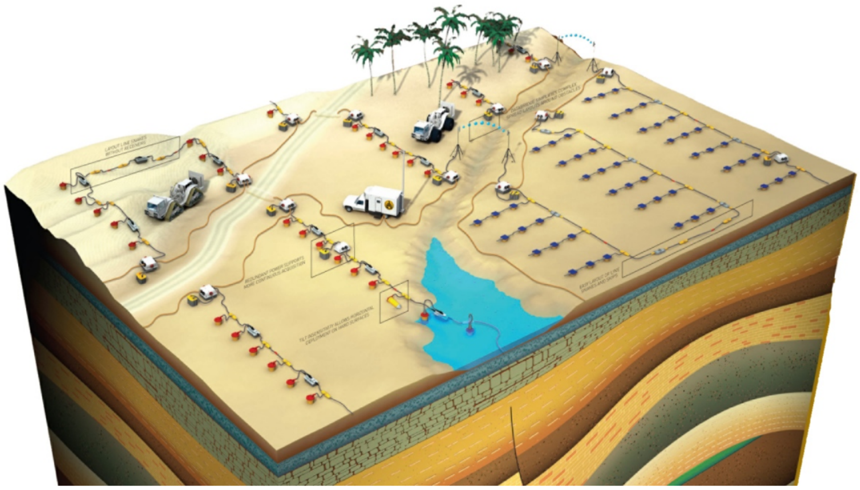


**Figure 1.1:** Schematic of land seismic survey. Image courtesy ION (www.iongeo.com).

### Marine

In water, the energy source is typically an array of air guns, i.e., guns with different sized air-chambers filled with compressed air. The source is towed behind a seismic survey vessel and releases bursts of high pressure energy (acoustic pulses) into the water. The returning sound waves are detected and recorded by sensors that are either hydrophones (measure wave pressure) spaced out along a series of cables, i.e., streamers towed behind a survey vessel, or arrays placed on the seafloor, i.e., ocean-bottom seismic sensors comprising of hydrophones and/or geophones. Figure 1.2 illustrates the different receiver geometries used in marine seismic surveying (Caldwell and Walker, 2011). Similar to land acquisition, we refer to the ocean-bottom surveys as "static" marine surveys, while the towed-streamer surveys are referred to as "dynamic" marine surveys. Ocean-bottom surveys can further be classified as ocean-bottom cable (OBC) surveys, wherein data are recorded using cables laid on the seabed (acquisition geometry 2 in Figure 1.2), or ocean-bottom node (OBN) surveys, wherein sensor nodes (which may or may not be connected by cables) are placed on the seafloor (Figure 1.3).

Figure 1.4 provides a list of the different types of marine surveys (Caldwell and Walker, 2011). 2D surveys comprise of a single source vessel towing a single streamer along a single line, called a sail line, over a survey area. In 3D surveying, groups of sail lines are acquired, i.e., 3D acquisition is the acquisition of many 2D lines closely spaced over the area. Surveys that are acquired repeatedly over the same area, particularly on established (or producing) fields to monitor changes in the reservoir over time due to production, are known as 4D or time-lapse surveys. The duration between surveys can be on the order of months or years. Figure 1.5 illustrates a 2D and 3D survey geometry.

Seismic data volumes are a collection of seismic traces. A seismic trace represents the response of the elastic wavefield to velocity and density contrasts across interfaces of layers of rock or sediments as energy travels from a source through the subsurface to a receiver or receiver array [Schlumberger Oilfield Glossary]. The convention adopted by the Society of Exploration Geophysicists (SEG) for display of (zero-phase) seismic data is as follows: If the signal arises from a reflection that indicates

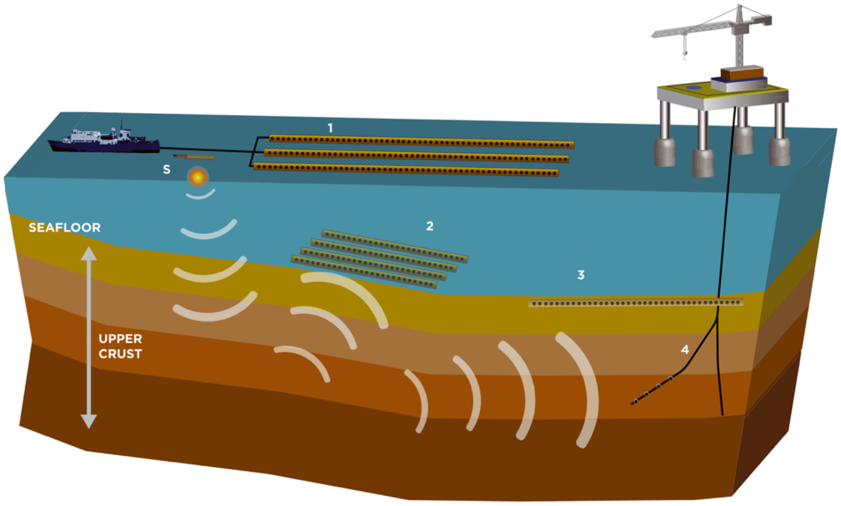


**Figure 1.2:** Schematic of different marine seismic surveys. “1” illustrates a towed-streamer geometry, “2” an ocean-bottom geometry, “3” a buried seafloor array, and “4” a VSP (vertical seismic profile) geometry, where the receivers are positioned in a well. [Source: Caldwell and Walker]

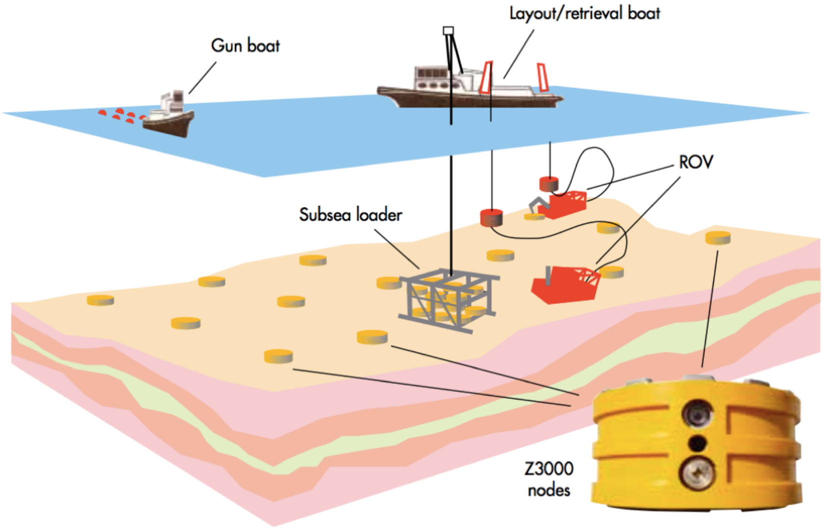


**Figure 1.3:** Schematic of ocean-bottom node survey. Remotely operated vehicles (ROVs) are used to deploy and recover sensor nodes. [Source: Caldwell and Walker]

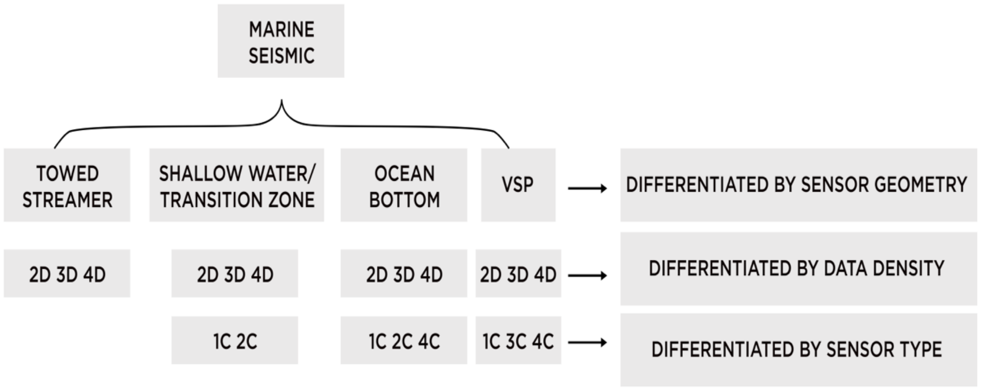


**Figure 1.4:** Summary of the majority of different types of marine seismic surveys. The letter "D" represents dimension and the letter "C" represents component ($Z$, $X$, $Y$). [Source: Caldwell and Walker]

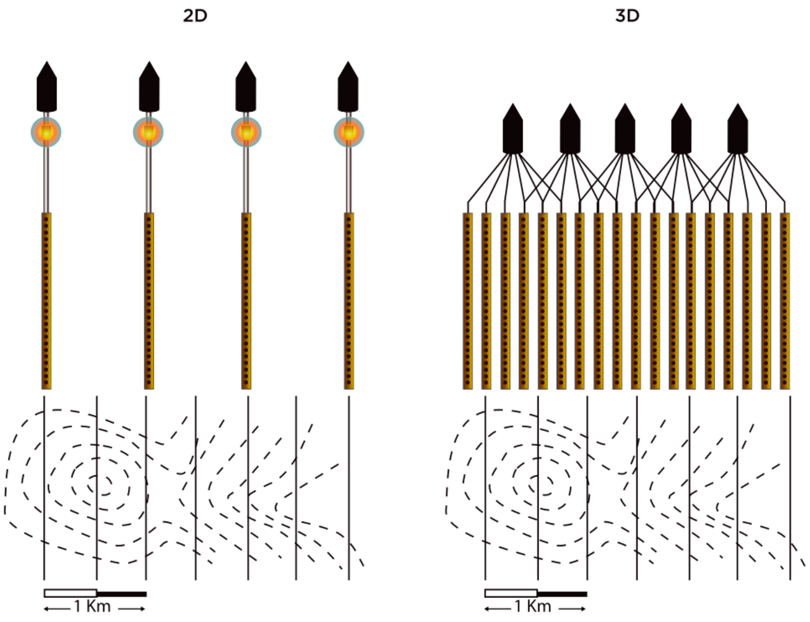


**Figure 1.5:** Illustration of difference between 2D and 3D survey geometry for same survey area. The dashed lines suggest subsurface structure contour lines. [Source: Caldwell and Walker]

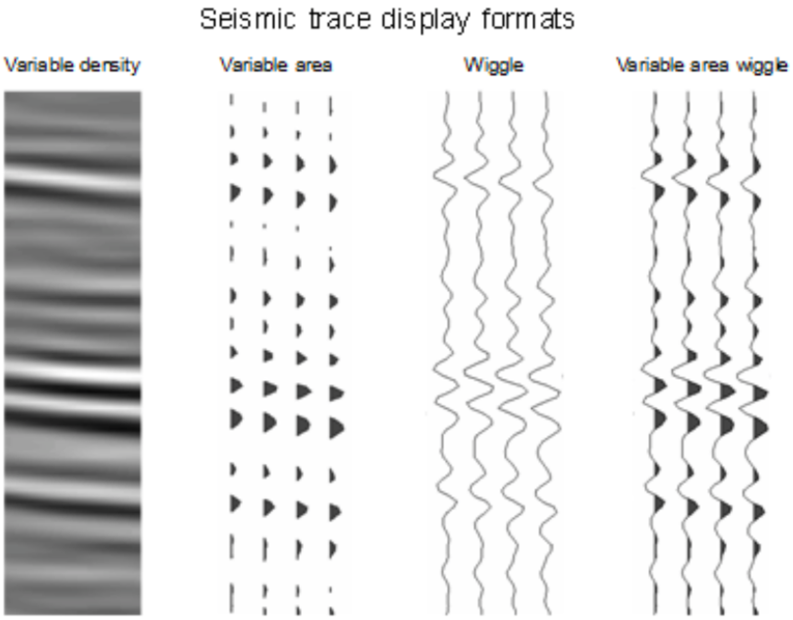


**Figure 1.6:** Four most common seismic trace display formats. [Source: SEG Wiki]

an increase in acoustic impedance (product of density and seismic velocity), the polarity is positive and is displayed as a peak. If the signal arises from a reflection that indicates a decrease in acoustic impedance, the polarity is negative and is displayed as a trough [Schlumberger Oilfield Glossary]. Figure 1.6 illustrates the four most common trace display formats [SEG wiki].

## 1.2 Simultaneous-source acquisition

Most of the commonly used processing algorithms, e.g., amplitude-versus-offset (AVO) analysis, surface-related multiple elimination (SRME) (Verschuur et al., 1992), estimation of primaries by sparse inversion (EPSI) (van Groenestijn and Verschuur, 2009; Lin and Herrmann, 2013), wave-equation based inversion techniques such as reverse-time migration (RTM) and full-waveform inversion (FWI) need dense and periodic (or regular) coverage of the survey area to produce high-resolution images of the subsurface. The need for dense sampling and full azimuthal coverage have led to the use of multiple source vessels, and simultaneous-source acquisition techniques. In seismic-acquisition literature, the term “azimuth” is defined as the angle at the source location between the sail line and the direction to a given receiver. The “simultaneous-source acquisition” methodology is also referred to as “blended acquisition” (Beasley et al., 1998; de Kok and Gillespie, 2002; Beasley, 2008; Berkhout, 2008; Moldoveanu, 2010; Abma et al., 2013; Mosher et al., 2014).

Long restricted to land acquisition, simultaneous-source methodology has now been proven in a marine environment (see references below). In principle, simultaneous-source methodology involves firing multiple sources at near-simultaneous/random-dithered times, hence, resulting in overlaps between shot records, as opposed to no overlaps during conventional (periodically-sampled) acquisition. In land seismic acquisition, simultaneous sources have revolutionized the way surveys are acquired, providing much greater sampling and consequently better imaging (Bagaini, 2010, and the references therein; Krohn and Neelamani, 2008; Neelamani et al., 2008). Marine simultaneous-source acquisition was invented in the year 1998-99 by the researchers at WesternGeco, however, its implementation had to overcome some fairly significant hurdles. The invention came just before the big downturn in the seismic industry that lasted until 2004, resulting in little or no interest for this new technology. Cost of additional vessels was another big hurdle to surmount. Advent of wide-azimuth (WAZ) surveys, fortunately, made multivessel operations more common, and as more companies saw its benefits, the simultaneous source concept began to gain momentum (Duey, 2012).

Since 2010-11, marine simultaneous-source acquisition has been an emerging technology stimulating both geophysical research and commercial efforts. The benefits of this methodology are substantial since seismic data is acquired in an economic and environmentally more sustainable way, i.e., data is acquired in less time, as compared to conventional acquisition, by firing multiple sources at near simultaneous/random times, or more data is acquired within the same time or a combination of both. Seismic acquisition literature contains a whole slew of works that have explored the concept of simultaneous- or blended-source activation (Beasley et al., 1998; de Kok and Gillespie, 2002; Beasley, 2008; Berkhout, 2008; Hampson et al., 2008; Moldoveanu, 2010; Berkhout, 2012; Abma et al., 2013; Mosher et al., 2014). However, there are challenges associated with simultaneous-source acquisition. Since many subsurface attribute inversion schemes (e.g., AVO analysis, SRME, EPSI, FWI, RTM, etc.) still rely on single-source prestack data volumes, one of the main challenges of simultaneous-source acquisition is to recover conventional sequential (or periodic) data from simultaneous data, i.e., estimate interference-free shot (and receiver) gathers, and particularly recover subtle late reflections of low amplitudes that can be overlaid by interfering seismic responses from other shots. This is known as *source separation*, also referred to as “deblending”.

Stefani et al. (2007), Moore et al. (2008) and Akerberg et al. (2008) have observed that as long as the sources are fired at suitably randomly-dithered times, the resulting interferences (in simultaneous data) will appear noise-like in specific gather domains such as common-offset and common-receiver, turning the separation into a typical (random) noise removal procedure. Application to land acquisition is reported in Bagaini and Ji (2010). Subsequent source-separation (or deblending) techniques, which aim to remove noise-like source crosstalk, vary from vector-median filters (Huo et al., 2009) to inversion-type algorithms (Moore, 2010; Abma et al., 2010; Mahdad et al., 2011; Doulgeris et al., 2012) to a combination of both (Baardman and van Borselen, 2013). The former are mostly “processing” techniques where the interfering energy (i.e., source crosstalk) is removed and not mapped back to coherent energy, at least not in a single step alone, while the latter are designed to take advantage of sparse representations of coherent seismic signals, which is advantageous because they exploit inherent structure in seismic data. Maraschini et al. (2012), Cheng and Sacchi (2013) and Kumar et al. (2015b) use matrix rank-reduction scheme for source separation. Recent success of simultaneous-source field trials, its implementation and source-separation techniques, have increased the industry’s confidence in this technology (Beasley et al., 2012; Abma et al., 2012, 2013; Mosher et al., 2014), where it has been observed that simultaneous-source surveys can be acquired faster and at lower costs.

Theoretical results from Compressive Sensing (CS, Donoho (2006); Candès et al. (2006c)) suggest that there is a direct relationship between the acquisition design and the expected fidelity of the achieved structure-promoting recovery, however, the aforementioned works did not investigate this link between the specific properties of the acquisition system and the sparsity-based recovery, especially in marine acquisition. Compressive sensing is a novel nonlinear sampling paradigm effective for signals that have a sparse representation in some transform domain. It is a novel paradigm in the sense that it provides theory for the link between the acquisition design, sparsity-promoting signal recovery and signal reconstruction quality. For land-based acquisition, the CS-related schemes (suitable for forward-modelling in the computer) presented by Neelamani et al. (2008) and Neelamani et al. (2010) suggest the use of noise-like signals as sweeps, Lin and Herrmann (2009a) use randomly phase-encoded vibroseis sweeps, and Herrmann et al. (2009) use impulsive sources that require modulation of each source by a randomly determined scaling in amplitude.

Analysis of pragmatic marine acquisition schemes in terms of CS arguments remains challenging because, while the success of CS hinges on an incoherent (or random) sampling technique, adapting this approach to real-life problems in exploration seismology is subject to physical constraints on the placement, type, and number of (simultaneous) sources, and number of receivers. Therefore, one of the objectives of this thesis is to propose a pragmatic sampling technique for marine simultaneous-source acquisition that adapts ideas from CS and no longer relies on the Nyquist sampling criteria. This technique, termed *time-jittered marine*, aims to achieve increased source sampling density and shorter acquisition times, and thus mitigate acquisition-related costs. Pioneered by Hennenfent and Herrmann (2008) and Herrmann (2010), this thesis presents a pragmatic CS simultaneous-source acquisition technique, specifically for marine, and addresses the challenge of source separation by sparsity-promoting recovery via convex optimization using $\ell_1$ objectives. Over the course of this research, Mosher et al. (2014) reported successful field application of randomized CS surveys and also showed the advantage of recovery via structure promotion in contrast to simply “processing” the acquired data. This encouraged further research on CS acquisition designs and processing, and their implication on time-lapse seismic.

### 1.2.1 Static vs. dynamic simultaneous-source acquisition

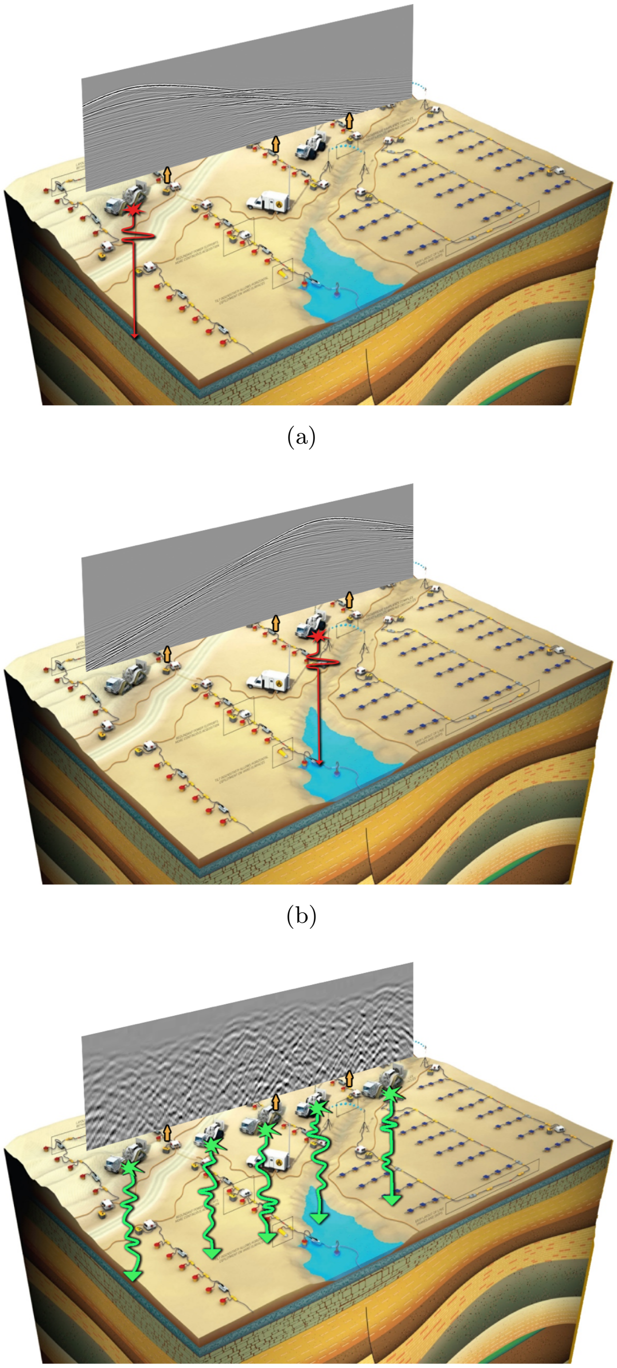
Based on the acquisition geometry, i.e., static or dynamic, there are different ways of acquiring simultaneous data.

#### Land simultaneous-source acquisition

An instance of land simultaneous-source acquisition is shown in Figure 1.7, wherein multiple vibroseis vehicles fire shots simultaneously. Specifically, sequential impulsive sources are replaced by impulsive simultaneous ‘phase-encoded’ sources. Chapter 2 presents this scenario in more details. After simultaneous data is acquired, the aim is then to recover individual sequential shot records as acquired during conventional acquisition.

#### Marine simultaneous-source acquisition

As mentioned above, and also explained in this thesis, to render possible artifacts induced by simultaneous-source firing incoherent, the key for simultaneous-source acquisition is the inclusion of randomization in the acquisition design, e.g., randomizing shot-firing times, randomizing source/receiver positions, randomizing distance between sail lines, etc. Randomization of shot-firing times depends on whether data is acquired with dynamic towed streamers or static receivers (OBCs or OBNs). Figure 1.8 illustrates the variability (or randomness) in shot-firing times for static and dynamic marine acquisition geometries. For dynamic towed streamers, randomness in shot-firing time



(a)

(b)

(c)

**Figure 1.7:** Schematic of land simultaneous-source acquisition. (a) and (b) Individual shots acquired in a conventional survey. (c) Simultaneous shot acquired in a simultaneous-source survey. Images courtesy ION (www.iongeo.com).

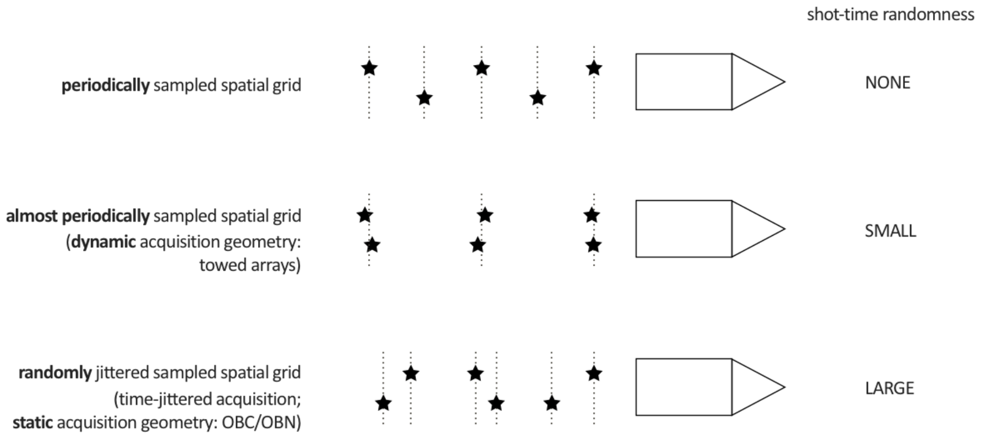


**Figure 1.8:** Shot-time randomness (or variability) for static and dynamic marine simultaneous-source acquisitions.

is small, on the order of a few seconds, i.e., sources fire within 1 or 2 second(s) of each other. This is because receivers are in motion, and therefore the returning sounds waves from the subsurface need to be captured within the time-frame of one shot record, which is typically 10 seconds. Moreover, moving arrays can only be compensated for a couple of meters. In contrast, static geometries enjoy large degrees of randomness, on the order of tens of seconds, in shot-firing times since the receivers are fixed.

Two instances of dynamic simultaneous-source acquisitions are over/under (or multilevel) source acquisition (Hill et al., 2006; Moldoveanu et al., 2007; Lansley et al., 2007; Long, 2009; Hegna and Parkes, 2012; Hoy et al., 2013) and Simultaneous-Long Offset (SLO) acquisition (Long et al., 2013, and the references therein). Over/under source acquisition extends the recorded bandwidth at the low and high ends of the spectrum because the depths of the sources produce complementary ghost functions, avoiding deep notches in the spectrum, while SLO acquisition provides a longer coverage in offsets without the need to tow very long streamer cables that can be problematic to deal with in the field (Chapter 7). Both acquisitions generate simultaneous data volumes that need to be separated in to corresponding individual data volumes for further processing (Figures 1.9 and 1.10).

The proposed CS simultaneous-source acquisition scheme, i.e., time-jittered marine acquisition is an instance of static simultaneous-source acquisition, wherein a single (and/or multiple) source vessel(s) sail(s) across an ocean-bottom array firing air guns at randomly jittered-time instances, which translate to (sub-Nyquist or subsampled) jittered shot positions for a given (fixed) speed of the source vessel. The basic idea of jittered subsampling is to regularly decimate the interpolation grid and subsequently perturb the coarse-grid sample points on the fine grid while controlling the maximum gap size between adjacent sample points (i.e., shot locations). Figure 1.11 illustrates this acquisition scheme for a single source vessel and two air gun arrays with receivers (OBC) recording continuously, resulting in a continuous subsampled simultaneous time-jittered data volume. The proposed acquisition scheme leads to improved spatial sampling of recovered (or separated) data, and speedup in acquisition compared to conventional periodically-sampled acquisition (Chapters 4-6). The improvement in spatial sampling is a result of separation and interpolation of simultaneous data, which is acquired on (relatively) coarsely-sampled spatial grids, to finely-sampled spatial

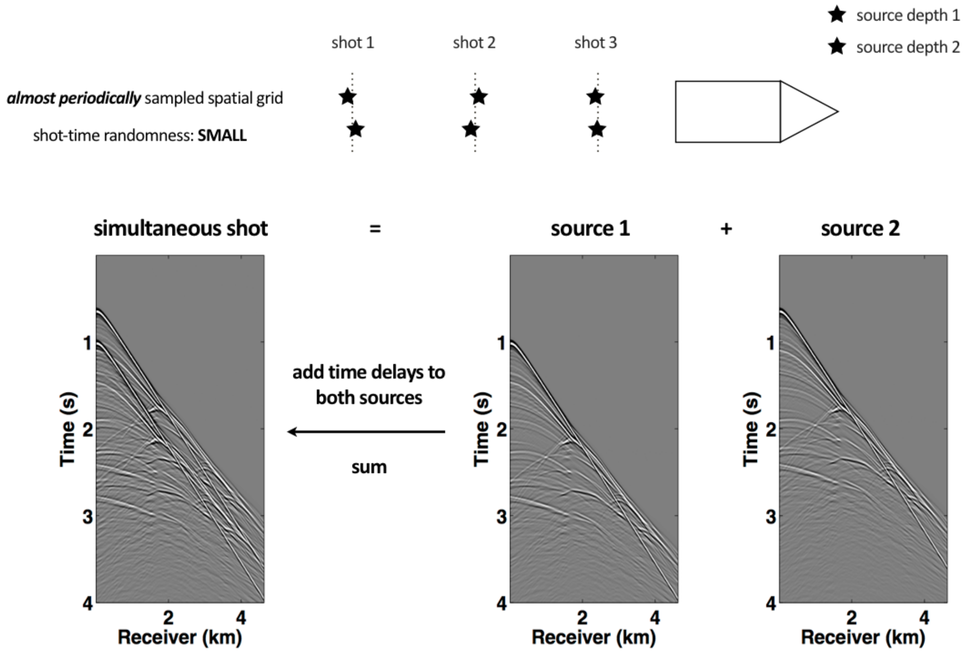


**Figure 1.9:** Schematic of dynamic over/under marine simultaneous-source acquisition. Simultaneous data acquired in the field is separated in to individual source components using source-separation techniques.

grids. Since static geometries provide better control over receiver positioning compared to dynamic towed streamers, the latter is a relatively more challenging scenario for source separation (Chapters 4-7). Additional challenges include processing of massive (especially 3D) simultaneous data volumes in computationally efficient ways, i.e., to reduce computational time at each step of the recovery algorithm, and efficient ways to store recovered data volumes in memory. Chapters 4 and 7 provide details on how we address these challenges.

## 1.3 Time-lapse seismic

Time-lapse (or 4D) seismic techniques involve acquisition, processing and interpretation of multiple 2D or 3D seismic surveys over a producing field with the aim of understanding the changes in the reservoir over time, particularly its behaviour during production (Lumley, 2001; Fanchi, 2001). The need for high degrees of data repeatability in time-lapse seismic has lead to the need to replicate dense surveys that are mostly OBC/OBN surveys, since these surveys provide better control over receiver positioning compared to towed streamers. Densely sampled and replicated surveys are expensive, and generate dense time-lapse data volumes whose processing is also computationally expensive. Therefore, the challenge is to minimize the cost of time-lapse surveying and data processing without impacting data repeatability. Owing to the positive impact of simultaneous-source acquisition on the industry, i.e., improved survey efficiency and data density, two key questions arise: “What are the implications of randomization on the attainable repeatability of time-lapse seismic?”, and “Should randomized time-lapse surveys be replicated?” These questions are of great importance because the incorporation of simultaneous-source acquisition in time-lapse seismic can

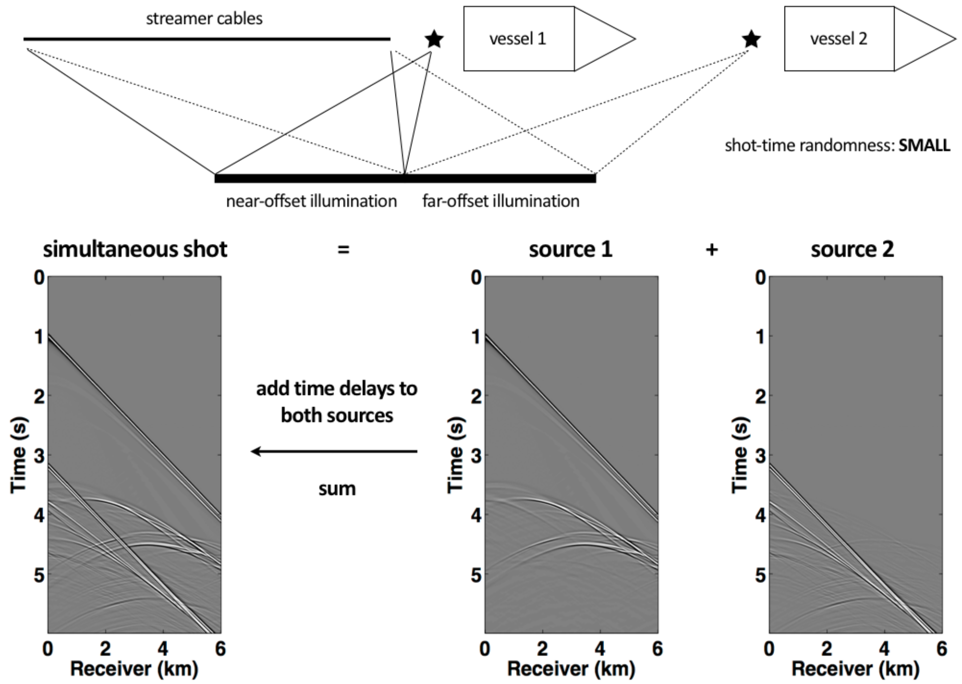


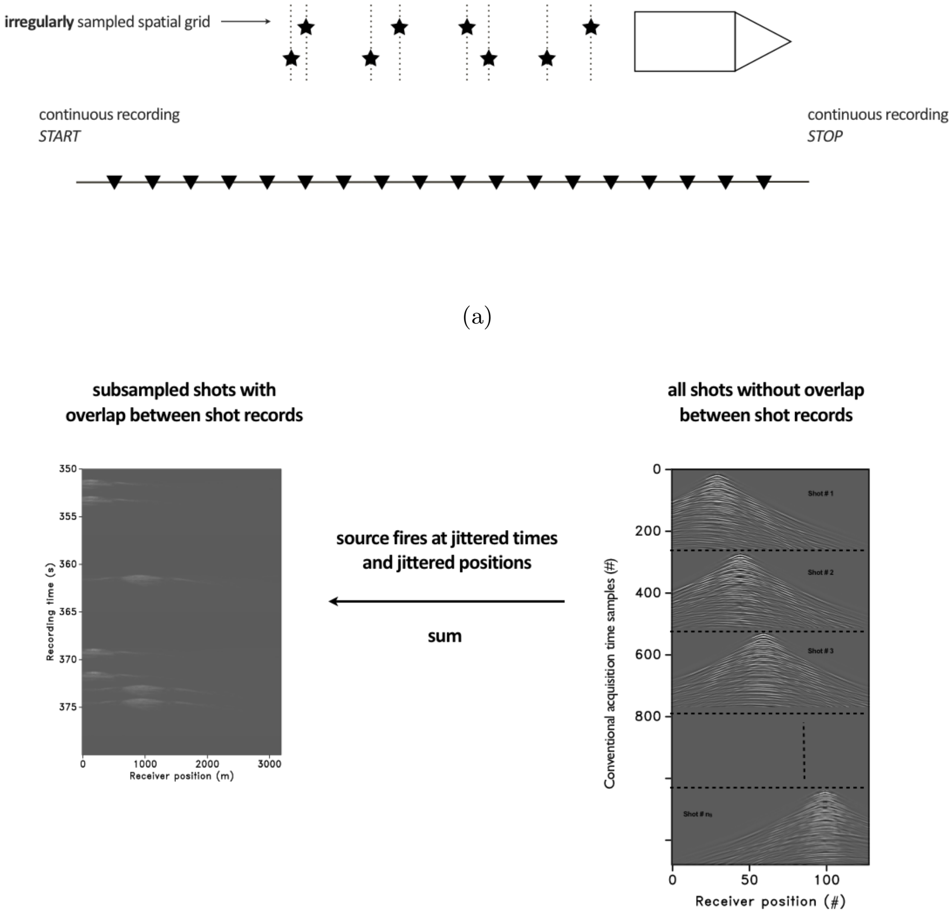
**Figure 1.10:** Schematic of dynamic marine simultaneous-long acquisition. Simultaneous data acquired in the field is separated in to individual source components using source-separation techniques. Note that the streamer length is 6 km and the second source vessel is deployed one spread-length (6 km) ahead of the main seismic vessel.

significantly change the current paradigm of time-lapse seismic that relies on expensive dense periodic sampling and replication of the baseline and monitor surveys (Lumley and Behrens, 1998). Therefore, another objective of this thesis is to analyze the effects of simultaneous-source (or randomized) surveys in time-lapse seismic by comparing repeatability of data recovered from fully replicated randomly-subsampled surveys and nonreplicated randomly-subsampled surveys. To this end, we present a new approach that explicitly exploits common information shared by the different time-lapse vintages. Note that we refer to the baseline and monitor data as the time-lapse vintages. The presented joint-recovery method (JRM), which is derived from distributed compressive sensing (DCS, Baron et al. (2009)), inverts for the common component and innovations with respect to this common component.

## 1.4 Objectives

The main purpose of this thesis is to develop practical compressive randomized marine simultaneous-source acquisitions and source-separation techniques by adapting ideas from compressive sensing. The objectives can be summarized as follows:

1. Adapt ideas from CS to design pragmatic marine simultaneous-source acquisition that acquires data economically with a reduced environmental imprint, whereby cost of surveys depends on certain inherent structure in seismic data rather than on the Nyquist sampling criteria. Investigate the relationship between the acquisition design and recovery quality.



(a)

(b)

**Figure 1.11:** Schematic of static marine simultaneous-source acquisition. This also illustrates the design of our proposed *time-jittered marine* acquisition. Source separation aims to recover densely sampled interference-free data by unravelling overlapping shot records and interpolation to a fine source grid (Chapters 4–6).

2. Address the challenge of source separation through a combination of tailored simultaneous-source acquisition design and sparsity-promoting recovery via convex optimization using $\ell_1$ objectives.

3. Compare sparsity-promoting and rank-minimization recovery techniques for static and dynamic marine simultaneous-source acquisitions. Develop a thorough understanding of the performance and limitations of the two techniques.

4. Adapt ideas from distributed compressive sensing to analyze the implications of simultaneous-source (or randomized) surveys on the attainable repeatability of time-lapse seismic.

## 1.5 Contributions

To the best of our knowledge, this work represents a first step towards a comprehensive study of designing marine simultaneous-source acquisitions by adapting ideas from compressive sensing. We use CS metrics, such as mutual coherence and restricted isometry property, to investigate the relationship between the acquisition design and data reconstruction fidelity. Consequently, we are able to assert the importance of randomness in the acquisition system in combination with an appropriate choice for a sparsifying transform in the reconstruction algorithm. This leads to new insights into acquiring and processing seismic data where we can fundamentally rethink on how we design acquisition surveys. Compressive (or subsampled) randomized (or jittered) simultaneous-source acquisitions and subsequent recovery techniques lead to improved wavefield reconstruction by increasing source-sampling density, and also speedup acquisition.

This work also presents a first instance of adapting ideas from CS and DCS to assess the effects (or risks) of random or irregular spatial sampling (i.e., samples that do not lie on a regular or periodic grid) in the field on time-lapse data, and demonstrate that high-quality data recoveries are the norm and not the exception. The main finding that compressive randomized time-lapse surveys need not be replicated to attain similar/acceptable levels of repeatability is significant since it can potentially change the current paradigm of time-lapse seismic that relies on expensive dense periodic sampling and replication of time-lapse surveys. Using a joint-recovery model (JRM) to process compressive randomized time-lapse data, we observe that recovery of the vintages improves when the time-lapse surveys are not replicated, since independent surveys give additional structural information. Moreover, since irregular spatial sampling is inevitable in the real world, it would be better to focus on knowing what the shot positions were (post acquisition) to a sufficient degree of accuracy, than aiming to replicate them. Recent successes of randomized surveys in the field (see, e.g., Mosher et al. (2014)) show that this can be achieved in practice.

## 1.6 Outline

The thesis comprises of eight chapters, including this introduction. We begin by presenting a synopsis of compressive sensing in Chapter 2, where we outline its three key principles: *(i)* Find representations that reveal structure of data, e.g, sparse (few nonzero entries) or compressible (can be well-approximated by a sparse signal). Examples of such representations are sparsifying transform domains, e.g., Fourier, curvelets, etc. *(ii)* Design a randomized subsampling scheme, which turns subsampling related artifacts into incoherent noise that is not sparse or compressible. *(iii)* Recover artifact-free fully sampled data by promoting structure, i.e., sparse recovery via one-norm minimization. Motivated by the challenges of the Nyquist's sampling criterion and "curse of dimensionality" (exponential increase in volume when adding extra dimensions to data collection) in exploration seismology, we discuss how we adapt ideas from CS to land and (static) marine simultaneous-source seismic acquisitions. We demonstrate the benefits of random subsampling over periodic subsampling by conducting of 1D synthetic experiments using random realizations of harmonic signals. We also empirically demonstrate that curvelets lead to compressible (real-world signals are not strictly sparse) representation of seismic data compared to wavelets, wave atoms, etc. Hence, we use curvelets for recovery via sparsity promotion.

Since the primary objective of this thesis to develop practical compressive randomized *marine* simultaneous-source acquisitions, the focus henceforth is on marine simultaneous-source acquisitions only. In Chapter 3, we identify (static) marine simultaneous-source acquisition as a linear subsampling system, which we subsequently analyze using metrics from compressive sensing, such as mutual coherence and restricted isometry property. Importance of sparsity-based recovery for simultaneous-source acquisition has been noticed by many authors. However, few have thoroughly investigated the underlying interaction between acquisition design and reconstruction fidelity, especially in the marine setting. According to CS, a sparsifying transform that is incoherent with the CS matrix can significantly impact the reconstruction quality. We demonstrate that the CS matrix resulting from our proposed sampling scheme is incoherent with the curvelet transform. We quantitatively verify the importance of randomness in the acquisition system and more compressible transforms by comparing recoveries from synthetic seismic experiments (using data from a real seismic line) for three different simultaneous acquisition schemes, namely "ideal" simultaneous acquisition, random time-dithering and periodic time-dithering acquisition, with different subsamplings for each.

While the observations made from Chapters 2 and 3 on adapting CS ideas to simultaneous-source acquisitions are very encouraging, it is important to not forget that compressive marine simultaneous-source acquisitions are beneficial to the seismic industry only if they are physically realizable in the field. We learnt this lesson while presenting the work at a conference meeting, where the nonrealistic nature of our acquisition scheme was revealed to us by communications with industry experts. This lead to a detailed investigation of the proposed compressive acquisition scheme to render it practical, and the findings of which are reported in Chapter 4. We develop a pragmatic compressive marine simultaneous-source acquisition scheme, termed *time-jittered marine*, wherein a single (and/or multiple) source vessel(s) sails across an ocean-bottom array firing air guns at jittered-time instances, which translate to jittered shot positions for a given (fixed) speed of the source vessel. The simultaneous data are time compressed, and are therefore acquired economically with a small environmental imprint. We demonstrate that all the observations made in Chapters 2 and 3 hold true for the pragmatic time-jittered marine acquisition. We conduct synthetic seismic experiments to recover interference-free densely sampled data from compressive (or subsampled) simultaneous data via sparsity promotion using 2D and 3D curvelet transforms (Candès et al., 2006a; Ying et al., 2005).

Chapter 5 delves in to the implications of compressive randomized acquisitions on the attainable repeatability of time-lapse seismic. Adapting ideas from CS and DCS, we aim to address the challenges of the current time-lapse seismic paradigm, such as, reliance on expensive dense periodic and replicated time-lapse surveys. We demonstrate that under certain assumptions, high-quality prestack data can be obtained from randomized subsampled measurements that are observed from nonreplicated surveys. We present a joint-recovery method (JRM) that exploits common information among the vintages leading to significant improvements in recovery quality of the time-lapse vintages when the same *on-the-grid* shot locations are not revisited. We compare joint recovery of time-lapse data with independent recovery, where each vintage is recovered independently. The acquisition is low cost because the measurements are subsampled. We conduct numerous 1D synthetic stylized experiments to test the JRM. We also confirm that high degrees of repeatability are achievable from the proposed time-jittered marine acquisition scheme. We assume measurements are taken *on-the-grid* (i.e., a discrete grid where the measurements lie "exactly" on the grid) and ignore errors related to taking measurements off the grid, a more realistic scenario that is dealt with in Chapter 6.

Since irregular or off-the-grid spatial sampling of sources and receivers is inevitable in field seismic acquisitions, it is important to analyze the implications of randomization on time-lapse seismic in this realistic setting. In Chapter 6, we extend our time-jittered marine acquisition to time-lapse surveys by designing acquisition on irregular spatial grids that render simultaneous, subsampled, and irregular measurements. We adapt the JRM to incorporate a regularization operator that maps traces from an irregular spatial grid to a regular periodic grid. The recovery method is therefore a combined operation of regularization, interpolation (estimating missing fine-grid traces from subsampled coarse-grid data), and source separation (unraveling overlapping shot records). We introduce the nonequispaced fast discrete curvelet transform (NFDCT, Hennenfent et al. (2010)) and its application to recover periodic densely sampled seismic lines from simultaneous and irregular measurements via sparsity-promoting inversion. We conduct a series of synthetic seismic experiments with different random realizations of the time-jittered marine acquisition with irregular sampling.

Chapter 7 addresses the source-separation problem for the more challenging dynamic towed-streamer acquisitions. We formulate the problem as a CS problem, which we subsequently solve by promoting two types of structure in seismic data, i.e., sparse and low rank. We simulate two simultaneous towed-streamer acquisitions, namely over/under and SLO. For recovery via rank minimization, we adopt the hierarchical semiseparable (HSS) matrix representation method proposed by Chandrasekaran et al. (2006) to exploit low-rank structure at high frequencies. We also combine the singular-value-decomposition-free matrix factorization approach recently developed by Lee et al. (2010) with the Pareto curve approach proposed by Berg and Friedlander (2008) that renders this framework suitable for large-scale seismic data because it avoids expensive singular value decompositions (SVDs), a necessary step in traditional rank-minimization-based methods. We compare recovery via sparsity promotion and rank minimization in terms of separation quality, computational time, and memory usage. We also make comparisons with the NMO-based median-filtering-type technique proposed by Chen et al. (2014).

In Chapter 8, we summarize the work done in this thesis, discuss certain associated limitations and propose ideas to address them by means of future research directions.

# Chapter 2

# Compressive sensing in seismic exploration: an outlook on a new paradigm

## 2.1 Summary

Many seismic exploration techniques rely on the collection of massive data volumes that are subsequently mined for information during processing. While this approach has been extremely successful in the past, current efforts toward higher resolution images in increasingly complicated regions of the Earth continue to reveal fundamental shortcomings in our workflows. Chiefly amongst these is the so-called "curse of dimensionality" exemplified by Nyquist's sampling criterion, which disproportionately strains current acquisition and processing systems as the size and desired resolution of our survey areas continues to increase. We offer an alternative sampling method leveraging recent insights from compressive sensing towards seismic acquisition and processing for data that, from a traditional point of view, are considered to be undersampled. The main outcome of this approach is a new technology where acquisition and processing related costs are decoupled the stringent Nyquist sampling criterion.

At the heart of our approach lies randomized incoherent sampling that breaks subsampling-related interferences by turning them into harmless noise, which we subsequently remove by promoting sparsity in a transform-domain. Acquisition schemes designed to fit into this regime no longer grow significantly in cost with increasing resolution and dimensionality of the survey area, but instead its cost ideally only depends on transform-domain sparsity of the expected data. Our contribution is twofold. First, we demonstrate by means of carefully designed numerical experiments that ideas from compressive sensing can be adapted to seismic acquisition. Second, we leverage the property that seismic data volumes are well approximated by a small percentage of curvelet coefficients. Thus curvelet-domain sparsity allows us to recover conventionally-sampled seismic data volumes from compressively-sampled data volumes whose size exceeds this percentage by only a small factor. Because compressive sensing combines transformation and encoding by a single linear encoding step, this technology is directly applicable to seismic acquisition and therefore

---

A version of this chapter has been published in CSEG Recorder, 2011, vol. 36, Part 1 [April Edition]: pp. 19–33, Part 2 [June Edition]: pp. 34–39.

constitutes a new paradigm where acquisitions costs scale with transform-domain sparsity instead of with the gridsize. We illustrate this principle by showcasing recovery of a real seismic line from simulated compressively sampled acquisitions.

## 2.2 Inspiration

### 2.2.1 Nyquist sampling and the curse of dimensionality

The livelihood of exploration seismology depends on our ability to collect, process, and image extremely large seismic data volumes. The recent push towards full-waveform approaches only exacerbates this reliance, and we, much like researchers in many other fields in science and engineering, are constantly faced with the challenge to come up with new and innovative ways to mine this overwhelming barrage of data for information. This challenge is especially daunting in exploration seismology because our data volumes sample wavefields that exhibit structure in up to five dimensions (two coordinates for the sources, two for the receivers, and one for time). When acquiring and processing this high-dimensional structure, we are not only confronted with Nyquist's sampling criterion but we also face the so-called "curse of dimensionality", which refers to the exponential increase in volume when adding extra dimensions to our data collection.

These two challenges are amongst the largest impediments to progress in the application of more sophisticated seismic methods to oil and gas exploration. In this chapter, we introduce a new methodology adapted from the field of "compressive sensing" or "compressive sampling" (CS in short throughout the article, Candès et al., 2006c; Donoho, 2006; Mallat, 2009), which is aimed at removing these impediments via dimensionality reduction techniques based on randomized subsampling. With this dimensionality reduction, we arrive at a sampling framework where the sampling rates are no longer scaling directly with the gridsize, but by transform-domain compression; more compressible data requires less sampling.

### 2.2.2 Dimensionality reduction by compressive sensing

Current nonlinear data-compression techniques are based on high-resolution linear sampling (e.g., sampling by a CCD chip in a digital camera) followed by a nonlinear encoding technique that consists of transforming the samples to some transformed domain, where the signal's energy is encoded by a relatively small number of significant transform-domain coefficients (Mallat, 2009). Compression is accomplished by keeping only the largest transform-domain coefficients. Because this compression is lossy, there is an error after decompression. A compression ratio expresses the compressed-signal size as a fraction of the size of the original signal. The better the transform captures the energy in the sampled data, the larger the attainable compression ratio for a fixed loss.

Even though this technique underlies the digital revolution of many consumer devices, including digital cameras, music, movies, etc., it does not seem possible for exploration seismology to scale in a similar fashion because of two major hurdles. First, high-resolution data has to be collected during the linear sampling step, which is already prohibitively expensive for exploration seismology. Second, the encoding phase is nonlinear. This means that if we select a compression ratio that is too high, the decompressed signal may have an unacceptable error, in the worst case making it necessary to repeat collection of the high-resolution samples.

By replacing the combination of high-resolution sampling and nonlinear compression by a single randomized subsampling technique that combines sampling and encoding in one single linear step, CS addresses many of the above shortcomings. First of all, randomized subsampling has the distinct advantage that the encoding is linear and does not require access to high-resolution data during encoding. This opens possibilities to sample incrementally and to process data in the compressed domain. Second, encoding through randomized sampling suppresses subsampling related artifacts. Coherent subsampling related artifacts—whether these are caused by periodic missing traces or by cross-talk between coherent simultaneous-sources—are turned into relatively harmless incoherent Gaussian noise by randomized subsampling (see e.g. Herrmann and Hennenfent, 2008; Hennenfent and Herrmann, 2008; Herrmann et al., 2009, for seismic applications of this idea).

By solving a sparsity-promoting problem (Candès et al., 2006c; Donoho, 2006; Herrmann et al., 2007; Mallat, 2009), we reconstruct high-resolution data volumes from the randomized samples at the moderate cost of a minor oversampling factor compared to data volumes obtained after conventional compression (see e.g. Donoho et al., 1999a, for wavelet-based compression). With sufficient sampling, this nonlinear recovery outputs a set of largest transform-domain coefficients that produces a reconstruction with a recovery error comparable with the error incurred during conventional compression. As in conventional compression this error is controllable, but in case of CS this recovery error depends on the sampling ratio—i.e., the ratio between the number of samples taken and the number of samples of the high-resolution data. Because compressively sampled data volumes are much smaller than high-resolution data volumes, we reduce the dimensionality and hence the costs of acquisition, storage, and possibly of data-driven processing.

We mainly consider recovery methods that derive from compressive sampling. Therefore our method differs from interpolation methods based on pattern recognition (Spitz, 1999), plane-wave destruction (Fomel et al., 2002) and data mapping (Bleistein et al., 2001), including parabolic, apex-shifted Radon and DMO-NMO/AMO (Trad, 2003; Trad et al., 2003; Harlan et al., 1984; Hale, 1995; Canning and Gardner, 1996; Bleistein et al., 2001; Fomel, 2003; Malcolm et al., 2005). To benefit fully from this new sampling paradigm, we will translate and adapt its ideas to exploration seismology while evaluating their performance. Here lies our main contribution. Before we embark on this mission we first share some basic insights from compressive sensing in the context of a well-known problem in geophysics: recovery of time-harmonic signals, which is relevant for missing-trace interpolation.

Compressive sensing is based on three key elements: randomized sampling, sparsifying transforms, and sparsity-promotion recovery by convex optimization. By themselves, these elements are not new to geophysics. Spiky deconvolution and high-resolution transforms are all based on sparsity-promotion (Taylor et al., 1979; Oldenburg et al., 1981; Ulrych and Walker, 1982; Levy et al., 1988; Sacchi et al., 1994) and analyzed by mathematicians (Santosa and Symes, 1986; Donoho and Logan, 1992); wavelet transforms are used for seismic data compression (Donoho et al., 1999a); randomized samples have been shown to benefit Fourier-based recovery from missing traces (Trad et al., 2003; Xu et al., 2005; Abma and Kabir, 2006; Zwartjes and Sacchi, 2007). The novelty of CS lies in the combination of these concepts into a comprehensive theoretical framework that provides design principles and performance guarantees.

### 2.2.3 Examples

#### Periodic versus uniformly-random subsampling

Because Nyquist's sampling criterion guarantees perfect reconstruction of arbitrary bandwidth-limited signals, it has been the leading design principle for seismic data acquisition and processing. This explains why acquisition crews go at length to place sources and receivers as finely and as regularly as possible. Although this approach spearheaded progress in our field, CS proves that periodic sampling at Nyquist rates may be far from optimal when the signal of interest exhibits some sort of structure, such as when the signal permits a transform-domain representation with few significant and many zero or insignificant coefficients. For this class of signals (which includes nearly all real-world signals) it suffices to sample randomly with fewer samples than that determined by Nyquist.

Take any arbitrary time-harmonic signal. According to compressive sensing, we can guarantee its recovery from a very small number of samples drawn at random times. In the seismic situation, this corresponds to using seismic arrays with fewer geophones selected uniformly-randomly from an underlying regular sampling grid with spacings defined by Nyquist (meaning it does not violate the Nyquist sampling theorem). By taking these samples randomly instead of periodically, the majority of artifacts directly due to incomplete sampling will behave like Gaussian white noise (Hennenfent and Herrmann, 2008; Donoho et al., 2009) as illustrated in Figure 2.1. We observe that for the same number of samples the subsampling artifacts can behave very differently.

In the geophysical community, subsampling-related artifacts are commonly known as "spectral leakage" (Xu et al., 2005), where energy from each frequency is leaked to other frequencies. Understandably, the amount of spectral leakage depends on the degree of subsampling: the higher this degree the more leakage. However, the characteristics of the artifacts themselves depend on the irregularity of the sampling. The more uniformly-random our sampling is, the more the leakage behaves as zero-centered Gaussian noise spread over the entire frequency spectrum.

Compressive sensing schemes aim to design acquisition that specifically create Gaussian-noise like subsampling artifacts (Donoho et al., 2009). As opposed to coherent subsampling related artifacts (Figure 2.1(f)), these noise-like artifacts (Figure 2.1(d)) can subsequently be removed by a sparse recovery procedure, during which the artifacts are separated from the signal and amplitudes are restored. Of course, the success of this method also hinges on the degree of subsampling, which determines the noise level, and the sparsity level of the signal.

By carrying out a random ensemble of experiments, where random realizations of harmonic signals are recovered from randomized samplings with decreasing sampling ratios, we confirm this behavior empirically. Our findings are summarized in Figure 2.2. The estimated spectra are obtained by solving a sparsifying program with the Spectral Projected Gradient for $\ell_1$ solver (SPGL1 - Berg and Friedlander, 2008) for signals with $k$ non-zero entries in the Fourier domain. We define these spectra by randomly selecting $k$ entries from vectors of length 600 and populating these with values drawn from a Gaussian distribution with unit standard deviation. As we will show below, the solution of each of these problems corresponds to the inversion of a matrix whose aspect ratio (the ratio of the number of columns over the number of rows) increases as the number of samples decreases.

To get reasonable estimates, each experiment is repeated 100 times for the different subsampling schemes and for varying sampling ratios ranging from 1/2 to 1/6. The reconstruction error is the number of vector entries at which the estimated spectrum and the true spectrum disagree by more than $10^{-4}$. This error counts both false positives and false negatives. The averaged results for the different experiments are summarized in Figures 2.2(a) and 2.2(b), which correspond to regular and random subsampling, respectively. The horizontal axes in these plots represent the relative underdeterminedness of the system, i.e., the ratio of the number $k$ of nonzero entries in the spectrum to the number of acquired data points $n$. The vertical axes denote the percentage of erroneous entries. The different curves represents the different subsampling factors. In each plot, the curves from top to bottom correspond to sampling ratios of 1/2 to 1/6.

Figure 2.2(a) shows that, regardless of the subsampling factor, there is no range of relative underdeterminedness for which the spectrum, and hence the signal, can accurately be recovered from regular subsamplings. Sparsity is not enough to discriminate the signal components from the spectral leakage. The situation is completely different in Figure 2.2(b) for the random sampling. In this case, one can observe that for a subsampling ratio of 1/2 exact recovery is possible for $0 < k/n \lesssim 1/4$. The main purpose of these plots is to qualitatively show the transition from successful to failed recovery. The quantitative interpretation for these diagrams of the transition is less well understood but also observed in phase diagrams in the literature (Donoho and Tanner, 2009; Donoho et al., 2009). A possible explanation for the observed behavior of the error lies in the nonlinear behavior of the solvers and on an error not measured in the $\ell_2$ sense.

### 2.2.4 Main contributions

We propose and analyze randomized sampling schemes, termed compressive seismic acquisition. Under specific conditions, these schemes create favourable recovery conditions for seismic wavefield reconstructions that impose transform-domain sparsity in Fourier or Fourier-related domains (see e.g. Sacchi et al., 1998; Xu et al., 2005; Zwartjes and Sacchi, 2007; Herrmann et al., 2007; Hennenfent and Herrmann, 2008; Tang et al., 2009). Our contribution is twofold. First, we demonstrate by means of carefully designed numerical experiments on synthetic and real data that compressive sensing can successfully be adapted to seismic acquisition, leading to a new generation of randomized acquisition and processing methodologies where high-resolution wavefields can be sampled and reconstructed with a controllable error. We introduce a number of performance measures that allow us to compare wavefield recoveries based on different sampling schemes and sparsifying transforms. Second, we show that accurate recovery can be accomplished for compressively sampled data volumes sizes that exceed the size of conventional transform-domain compressed data volumes by a small factor. Because compressive sensing combines transformation and encoding by a single linear encoding step, this technology is directly applicable to seismic acquisition and to dimensionality reduction during processing. We verify this claim by a series of experiments on real data. We also show that the linearity of CS allows us to extend this technology to seismic data processing. In either case, sampling, storage, and processing costs scale with transform-domain sparsity.

### 2.2.5 Outline

First, we briefly present the key principles of CS, followed by a discussion on how to adapt these principles to the seismic situation. For this purpose, we introduce measures that quantify reconstruction and recovery errors and expresses the overhead that CS imposes. We use these measures to compare the performance of different transform domains and sampling strategies during reconstruction. We then use this information to evaluate and apply this new sampling technology towards acquisition and processing of a 2D seismic line.

## 2.3 Basics of compressive sensing

In this section, we give a brief overview of CS and concise recovery criteria. CS relies on specific properties of the compressive-sensing matrix and the sparsity of the to-be-recovered signal.

### 2.3.1 Recovery by sparsity-promoting inversion

Consider the following linear forward model for sampling

$$\mathbf{b} = \mathbf{A}\mathbf{x}_0, \tag{2.1}$$

where $\mathbf{b} \in \mathbb{R}^n$ represents the compressively sampled data consisting of $n$ measurements. Suppose that a high-resolution data $\mathbf{f}_0 \in \mathbb{R}^N$, with $N$ the ambient dimension, has a sparse representation $\mathbf{x}_0 \in \mathbb{R}^N$ in some known transform domain. For now, we assume that this representation is the identity basis—i.e., $\mathbf{f}_0 = \mathbf{x}_0$. We will also assume that the data is noise free. According to this model, measurements are defined as inner products between rows of $\mathbf{A}$ and high-resolution data.

The sparse recovery problem involves the reconstruction of the vector $\mathbf{x}_0 \in \mathbb{R}^N$ given incomplete measurements $\mathbf{b} \in \mathbb{R}^n$ with $n \ll N$. This involves the inversion of an underdetermined system of equations defined by the matrix $\mathbf{A} \in \mathbb{R}^{n\times N}$, which represents the sampling operator that collects the acquired samples from the model, $\mathbf{f}_0$.

The main contribution of CS is to come up with conditions on the compressive-sampling matrix $\mathbf{A}$ and the sparse representation $\mathbf{x}_0$ that guarantee recovery by solving a convex sparsity-promoting optimization problem. This sparsity-promoting program leverages sparsity of $\mathbf{x}_0$ and hence overcomes the singular nature of $\mathbf{A}$ when estimating $\mathbf{x}_0$ from $\mathbf{b}$. After sparsity-promoting inversion, the recovered representation for the signal is given by

$$\tilde{\mathbf{x}} = \arg\min_{\mathbf{x}} ||\mathbf{x}||_1 \quad \text{subject to} \quad \mathbf{b} = \mathbf{A}\mathbf{x}. \tag{2.2}$$

In this expression, the symbol $\tilde{}$ represents estimated quantities and the $\ell_1$ norm $\|\mathbf{x}\|_1$ is defined as $\|\mathbf{x}\|_1 \overset{\text{def}}{=} \sum_{i=1}^N |x[i]|$, where $x[i]$ is the $i^{th}$ entry of the vector $\mathbf{x}$.

Minimizing the $\ell_1$ norm in equation 2.2 promotes sparsity in $\mathbf{x}$ and the equality constraint ensures that the solution honors the acquired data. Among all possible solutions of the (severely) underdetermined system of linear equations ($n \ll N$) in equation 2.1, the optimization problem in equation 2.2 finds a sparse or, under certain conditions, the sparsest (i.e., smallest $\ell_0$ norm (Donoho and Huo, 2001)) possible solution that exactly explains the data.

### 2.3.2 Recovery conditions

The basic idea behind CS (see e.g. Candès et al., 2006c; Mallat, 2009) is that recovery is possible and stable as long as any subset $S$ of $k$ columns of the $n \times N$ matrix $\mathbf{A}$—with $k \leq N$ the number of nonzeros in $\mathbf{x}$—behave approximately as an orthogonal basis. In that case, we can find a constant $\hat{\delta}_k$ for which we can bound the energy of the signal from above and below —i.e.,

$$(1 - \hat{\delta}_k)\|\mathbf{x}_S\|_{\ell_2}^2 \leq \|\mathbf{A}_S\mathbf{x}_S\|_{\ell_2}^2 \leq (1 + \hat{\delta}_k)\|\mathbf{x}_S\|_{\ell_2}^2, \tag{2.3}$$

where $S$ runs over sets of all possible combinations of columns with the number of columns $|S| \leq k$ (with $|S|$ the cardinality of $S$). The smaller $\hat{\delta}_k$, the more energy is captured and the more stable the inversion of $\mathbf{A}$ becomes for signals $\mathbf{x}$ with maximally $k$ nonzero entries.

The key factor that bounds the restricted-isometry constants $\hat{\delta}_k > 0$ from above is the mutual coherence amongst the columns of $\mathbf{A}$—i.e.,

$$\hat{\delta}_k \leq (k-1)\mu \tag{2.4}$$

with

$$\mu = \max_{1 \leq i \neq j \leq N} |\mathbf{a}_i^H \mathbf{a}_j|, \tag{2.5}$$

where $\mathbf{a}_i$ is the $i^{\text{th}}$ column of $\mathbf{A}$ and $^H$ denotes the Hermitian transpose.

Matrices for which $\hat{\delta}_k$ is small contain subsets of $k$ columns that are incoherent. Random matrices, with Gaussian *i.i.d.* entries with variance $n^{-1}$ have this property, whereas deterministic constructions almost always have structure.

For these random Gaussian matrices (there are other possibilities such as Bernouilli or restricted Fourier matrices that accomplish approximately the same behavior, see e.g. Candès et al., 2006c; Mallat, 2009), the mutual coherence is small. For this type of CS matrices, it can be proven that Equation 2.3 holds and Equation 2.2 recovers $\mathbf{x}_0$'s exactly with high probability as long as this vector is maximally $k$ sparse with

$$k \leq C \cdot \frac{n}{\log_2(N/n)}, \tag{2.6}$$

where $C$ is a moderately sized constant. This result proves that for large $N$, recovery of $k$ nonzeros only requires an oversampling ratio of $n/k \approx C \cdot \log_2 N$, as opposed to taking all $N$ measurements.

The above result is profound because it entails an oversampling with a factor $C \cdot \log_2 N$ compared to the number of nonzeros $k$. Hence, the number of measurements that are required to recover these nonzeros is much smaller than the ambient dimension ($n \ll N$ for large $N$) of high-resolution data. Similar results hold for compressible instead of strictly sparse signals while measurements can be noisy (Candès et al., 2006c; Mallat, 2009). In that case, the recovery error depends on the noise level and on the transform-domain compression rate—i.e., the decay of the magnitude-sorted coefficients.

In summary, according to CS (Candès et al., 2006b; Donoho, 2006), the solution $\tilde{\mathbf{x}}$ of equation 2.2 and $\mathbf{x}_0$ coincide when two conditions are met, namely 1) $\mathbf{x}_0$ is sufficiently sparse, i.e., $\mathbf{x}_0$ has few nonzero entries, and 2) the subsampling artifacts are incoherent, which is a direct consequence of measurements with a matrix whose action mimics that of a Gaussian matrix.

Unfortunately, most rigorous results from CS, except for work by Rauhut et al. (2008), are valid for orthonormal measurement and sparsity bases only and the computation of the recovery conditions for realistically sized seismic problems remains computational prohibitive. To overcome these important shortcomings, we will in the next section introduce a number of practical and computable performance measures that allow us to design and compare different compressive-seismic acquisition strategies.

## 2.4 Compressive-sensing design

As we have seen, the machinery that supports sparse recovery from incomplete data depends on specific properties of the compressive-sensing matrix. It is important to note that CS is not meant to be blindly applied to arbitrary linear inversion problems. To the contrary, the success of a sampling scheme operating in the CS framework hinges on the design of new acquisition strategies that are both practically feasible and lead to favourable conditions for sparse recovery. Mathematically speaking, the resulting CS sampling matrix needs to both be realizable and behave as a Gaussian matrix. To this end, the following key components need to be in place:

1. **a sparsifying signal representation** that exploits the signal's structure by mapping the energy into a small number of significant transform-domain coefficients. The smaller the number of significant coefficients, the better the recovery;
2. **sparse recovery by transform-domain one-norm minimization** that is able to handle large system sizes. The fewer the number of matrix-vector evaluations, the faster and more practically feasible the wavefield reconstruction;
3. **randomized seismic acquisition** that breaks coherent interferences induced by deterministic subsampling schemes. Randomization renders subsampling related artifacts—including aliases and simultaneous source crosstalk—harmless by turning these artifacts into incoherent Gaussian noise;

Given the complexity of seismic data in high dimensions and field practicalities of seismic acquisition, the mathematical formulation of CS outlined in the previous section does not readily apply to seismic exploration. Therefore, we will focus specifically on the design of source subsampling schemes that favor recovery and on the selection of the appropriate sparsifying transform. Because theoretical results are mostly lacking, we will guide ourselves by numerical experiments that are designed to measure recovery performance.

During seismic data acquisition, data volumes are collected that represent discretizations of analog finite-energy wavefields in two or more dimensions including time. We recover the discretized wavefield $\mathbf{f}$ by inverting the compressive-sampling matrix

$$\mathbf{A} := \overbrace{\mathbf{R}}^{\text{restriction}} \quad \underbrace{\mathbf{M}}_{\text{measurement}} \quad \overbrace{\mathbf{S}^H}^{\text{synthesis}} \tag{2.7}$$

with the sparsity-promoting program:

$$\widetilde{\mathbf{f}} = \mathbf{S}^H \widetilde{\mathbf{x}} \quad \text{with} \quad \widetilde{\mathbf{x}} = \arg\min_{\mathbf{x}} \|\mathbf{x}\|_1 := \sum_{p=0}^{P-1} |x[i]| \quad \text{subject to} \quad \mathbf{A}\mathbf{x} = \mathbf{b}. \tag{2.8}$$

This formulation differs from standard compressive sensing because we allow for a wavefield representation that is redundant—i.e., $\mathbf{S} \in \mathbb{C}^{P \times N}$ with $P \geq N$. Aside from results reported by Rauhut et al. (2008), which show that recovery with redundant frames is determined by the RIP constant $\hat{\delta}$ of the restricted sampling and sparsifying matrices that is least favorable, there is no practical algorithm to compute these constants. Therefore, our hope is that the above sparsity-promoting optimization program, which finds amongst all possible transform-domain vectors the vector $\widetilde{\mathbf{x}} \in \mathbb{R}^P$ that has the smallest $\ell_1$-norm, recovers high-resolution data $\widetilde{\mathbf{f}} \in \mathbb{R}^N$.

### 2.4.1 Seismic wavefield representations

One of the key ideas of CS is leveraging structure within signals to reduce sampling. Typically, structure translates into transform-domains that concentrate the signal's energy in as few as possible significant coefficients. The size of seismic data volumes, along with the complexity of its high-dimensional and highly directional wavefront-like features, makes it difficult to find a transform that accomplishes this task.

To meet this challenge, we only consider transforms that are fast (at the most $N \log N$ with $N$ the number of samples), multiscale (splitting the Fourier spectrum into dyadic frequency bands), and multidirectional (splitting Fourier spectrum into second dyadic angular wedges). For reference, we also include separable 2D wavelets in our study. We define this wavelet transform as the Kronecker product (denoted by the symbol $\otimes$) of two 1D wavelet transforms: $\mathbf{W} = \mathbf{W}_1 \otimes \mathbf{W}_1$ with $\mathbf{W}_1$ the 1D wavelet-transform matrix.

### Separable versus non-separable transforms

There exists numerous signal representations that decompose a multi-dimensional signal with respect to directional and localized elements. For the appropriate representation of seismic wavefields, we limit our search to non-separable curvelets (Candès et al., 2006a) and wave atoms (Demanet and Ying, 2007). The elements of these transforms behave approximately as high-frequency asymptotic eigenfunctions of wave equations (see e.g. Smith, 1998; Candès and Demanet, 2005; Candès et al., 2006a; Herrmann et al., 2008), which makes these two representations particularly well suited for our task of representing seismic data parsimoniously.

Unlike wavelets, which compose curved wavefronts into a superposition of multiscale “fat dots” with limited directionality, curvelets and wave atoms compose wavefields as a superposition of highly anisotropic localized and multiscale waveforms, which obey a so-called parabolic scaling principle. For curvelets in the physical domain, this principle translates into a support with its length proportional to the square of the width. At the fine scales, this scaling leads to curvelets that become increasingly anisotropic, i.e., needle-like. Each dyadic frequency band is split into a number of overlapping angular wedges that double in every other dyadic scale. This partitioning results in increased directionality at the fine scales. This construction makes curvelets well adapted to data with impulsive wavefront-like features. Figure 2.3(a) shows the multiscale and multidirectional 2D curvelets in the time-space domain and the frequency-wavenumber domain. Curvelets approximate curved singularities, i.e., wavefronts, in a nonadaptive manner with very few significant curvelet coefficients (Figure 2.3(b)). Wave atoms, on the other hand, are anisotropic because it is their wavelength, not the physical length of the individual wave atoms, that depends quadratically on their width. By construction, wave atoms are more appropriate for data with oscillatory patterns. Because seismic data sits somewhere between these two extremes, we include both transforms in our study.

### Approximation error

For an appropriately chosen representation magnitude-sorted transform-domain coefficients often decay rapidly–i.e., the magnitude of the $j^{th}$ largest coefficient is $\mathcal{O}(j^{-s})$ with $s \geq 1/2$. For orthonormal bases, this decay rate is directly linked to the decay of the nonlinear approximation error (see e.g. Mallat, 2009). This error is expressed by

$$\sigma(k) = \|\mathbf{f} - \mathbf{f}_k\| = \mathcal{O}(k^{1/2-s}), \tag{2.9}$$

with $\mathbf{f}_k$ the reconstruction from the largest $k$ - coefficients. Notice that this error does not account for discretization errors (cf. Equation 2.16), which we ignore.

Unfortunately, this relationship between the decay rates of the magnitude-sorted coefficients and the decay rate of the nonlinear approximation error does not hold for redundant transforms. Also, there are many coefficient sequences that explain the data $\mathbf{f}$ making them less sparse–i.e., expansions with respect to this type of signal representations are not unique. For instance, analysis by the curvelet transform of a single curvelet does not produce a single non-zero entry in the curvelet coefficient vector.

To address this issue, we use an alternative definition for the nonlinear approximation error, which is based on the solution of a sparsity-promoting program. With this definition, the $k$-term nonlinear-approximation error is computed by taking the $k-$largest coefficients from the vector that solves

$$\min_{\mathbf{x}} \|\mathbf{x}\|_1 \quad \text{subject to} \quad \mathbf{S}^H\mathbf{x} = \mathbf{f}. \tag{2.10}$$

Because this vector is obtained by inverting the synthesis operator $\mathbf{S}^H$ with a sparsity-promoting program, this vector is always sparser than the vector obtained by applying the analysis operator $\mathbf{S}$ directly.

To account for different redundancies in the transforms, we study signal-to-noise ratios (S/Ns) as a function of the sparsity ratio $\rho = k/P$ (with $P = N$ for orthonormal bases) defined as

$$\mathrm{S/N}(\rho) = -20 \log \frac{\|\mathbf{f} - \mathbf{f}_\rho\|}{\|\mathbf{f}\|}. \tag{2.11}$$

The smaller this ratio, the more coefficients we ignore, the sparser the transform-coefficient vector becomes, which in turn leads to a smaller S/N. In our study, we include $\mathbf{f}_\rho$ that are derived from either the analysis coefficients or from the synthesis coefficients. The latter coefficients are solutions of the above sparsity-promoting program (Equation 2.10).

### Empirical approximation errors

The above definition gives us a metric to compare recovery S/Ns of seismic data for wavelets, curvelets, and wave atoms. We make this comparison on a common-receiver gather (Figure 2.4) extracted from a Gulf of Suez data set. Because the current implementations of wave atoms (Demanet and Ying, 2007) only support data that is square, we padded the 178 traces with zeros to 1024 traces. The temporal and spatial sampling interval of the high-resolution data are 0.004s and 25m, respectively. Because this zero-padding biases the $\rho$, we apply a correction.

Our results are summarized in Figure 2.5 and they clearly show that curvelets lead to rapid improvements in S/N as the sparsity ratio increases. This effect is most pronounced for synthesis coefficients, benefiting remarkably from sparsity promotion. By comparison, wave atoms benefit not as much, and wavelet even less. This behavior is consistent with the overcompleteness of these transforms, the curvelet transform matrix has the largest redundancy (a factor of about eight in 2D) and is therefore the tallest. Wave atoms only have a redundancy of two and wavelets are orthogonal. Since our method is based on sparse recovery, this experiment suggests that sparse recovery from subsampling would potentially benefit most from curvelets. However, this is not the only factor that determines the performance of our compressive-sampling scheme.

### 2.4.2 Subsampling of shots

Aside from obtaining good reconstructions from small compression ratios, breaking the periodicity of coherent sampling is paramount to the success of sparse recovery—whether this involves selection of subsets of sources or the design of incoherent simultaneous-source experiments. To underline the importance of maximizing incoherence in seismic acquisition, we conduct two experiments where common-source gathers are recovered from subsets of sequential and simultaneous-source experiments. To make useful comparisons, we keep for each survey the number of source experiments, and hence the size of the collected data volumes, the same.

#### Coherent versus incoherent sampling

Mathematically, sequential and simultaneous acquisition only differ in the definition of the measurement basis. For sequential-source acquisition, this sampling matrix is given by the Kronecker product of two identity bases—i.e., $\mathbf{I} := \mathbf{I}_{N_s} \otimes \mathbf{I}_{N_t}$, which is a $N \times N$ identity matrix with $N = N_t \times N_s$, the product of the number of time samples $N_t$ and the number of shots $N_s$. For simultaneous acquisition, where all sources fire simultaneously, this matrix is given by $\mathbf{M} := \mathbf{G}_{N_s} \otimes \mathbf{I}_{N_t}$ with $\mathbf{G}_{N_s}$ a $N_s \times N_s$ Gaussian matrix with *i.i.d.* entries. In both cases, we use a restriction operator $\mathbf{R} := \mathbf{R}_{n_s} \otimes \mathbf{I}_{N_t}$ to model the collection of incomplete data by reducing the number of shots to $n_s \ll N_s$. This restriction acts on the source coordinate only.

Roughly speaking, CS predicts superior recovery for compressive-sampling matrices with smaller coherence. According to Equation 2.5, this coherence depends on the interplay between the restriction, measurement, and synthesis matrices. To make a fair comparison, we keep the restriction matrix the same and study the effect of having measurement matrices that are either given by the identity or by a random Gaussian matrix. Physically, the first CS experiment corresponds to surveys with sequential shots missing. The second CS experiment corresponds to simultaneous-source experiments with simultaneous source experiments missing. Examples of both measurements for the real common-receiver gather of Figure 2.4 are plotted in Figure 2.6. Both data sets have 50% of the original size. Remember that the horizontal axes in the simultaneous experiment no longer has a physical meaning. Notice also that there is no observable coherent crosstalk amongst the simultaneous sources.

Multiplication of orthonormal sparsifying bases by random measurement matrices turns into random matrices with a small mutual coherence amongst the columns. This property also holds (but only approximately) for redundant signal representations with synthesis matrices that are wide and have columns that are linearly dependent. This suggests improved performance using random incoherent measurement matrices. To verify this statement empirically, we compare sparse recoveries with Equation 2.8 from data plotted in Figure 2.6(a).

Despite the fact that simultaneous acquisition with all sources firing simultaneously may not be easily implementable in practice[1], this approach has been applied successfully to reduce simulation and imaging costs (Herrmann et al., 2009; Herrmann, 2009; Lin and Herrmann, 2009a,b). In the "eyeball norm", the recovery from the simultaneous data is as expected clearly superior (cf. Figures 2.6(b)). The difference plots (cf. Figures 2.6(c)) confirm this observation and show very little coherent signal loss for the recovery from simultaneous data. This is consistent with CS, which predicts improved performance for sampling schemes that are more incoherent. Because this quali-

---

[1] Although one can easily imagine a procedure in the field where a "supershot" is created by some stacking procedure.

tative statement depends on the interplay between the sampling and the sparsifying transform, we conduct an extensive series of experiments to get a better idea on the performance of these two different sampling schemes for different sparsifying transforms. We postpone our analysis of the quantitative behavior of the recovery S/Ns to after that discussion.

### Sparse recovery errors

The examples of the previous section clearly illustrate that randomized sampling is important, and that randomized simultaneous acquisition leads to better recovery compared to randomized subsampling of sequential sources. To establish this observation more rigorously, we calculate estimates for the recovery error as a function of the sampling ratio $\delta = n/N$ by conducting a series of 25 controlled recovery experiments. For each $\delta \in [0.2, 0.8]$, we generate 25 realizations of the randomized compressive-sampling matrix. Applying these matrices to our common-receiver gather (Figure 2.4) produces 25 different data sets that are subsequently used as input to sparse recovery with wavelets, curvelets, and wave atoms. For each realization, we calculate the $\mathrm{S/N}(\delta)$ with

$$\mathrm{S/N}(\delta) = -20 \log \frac{\|\mathbf{f} - \widetilde{\mathbf{f}}_\delta\|}{\|\mathbf{f}\|}, \tag{2.12}$$

where

$$\widetilde{\mathbf{f}}_\delta = \mathbf{S}^H \widetilde{\mathbf{x}}_\delta \quad \text{and} \quad \widetilde{\mathbf{x}}_\delta = \arg\min_{\mathbf{x}} \|\mathbf{x}\|_1 \quad \text{subject to} \quad \mathbf{A}_\delta \mathbf{x} = \mathbf{b}.$$

For each experiment, the recovery of $\widetilde{\mathbf{f}}_\delta$ is calculated by solving this optimization problem for 25 different realizations of $\mathbf{A}_\delta$ with $\mathbf{A}_\delta := \mathbf{R}_\delta \mathbf{M}_\delta \mathbf{S}^H$, where $\mathbf{R}_\delta := \mathbf{R}_{n_s} \otimes \mathbf{I}_{N_t}$ with $\delta = n_s/N_s$. For each simultaneous experiment, we also generate different realizations of the measurement matrix $\mathbf{M} := \mathbf{G}_{N_s} \otimes \mathbf{I}_{N_t}$.

From these randomly selected experiments, we calculate the average S/Ns for the recovery error, $\overline{\mathrm{S/N}}(\delta)$, including its standard deviation. By selecting $\delta$ evenly on the interval $\delta \in [0.2, 0.8]$, we obtain reasonable reliable estimates with error bars. Results of this exercise are summarized in Figure 2.7. From these plots it becomes immediately clear that simultaneous acquisition greatly improves recovery for all three transforms. Not only are the S/Ns better, but the spread in S/Ns amongst the different reconstructions is also much less, which is important for quality assurance. The plots validate CS, which predicts improved recovery for increased sampling ratios. Although somewhat less pronounced as for the approximation S/Ns in Figure 2.5, our results again show superior performance for curvelets compared to wave atoms and wavelets. This observation is consistent with our earlier empirical findings.

### Empirical oversampling ratios

The key factor that establishes CS is the sparsity ratio $\rho$ that is required to recover wavefields with errors that do not exceed a predetermined nonlinear approximation error (cf. Equation 2.11). The latter sets the fraction of largest coefficients that needs to be recovered to meet a preset minimal S/N for reconstruction.

Motivated by Mallat (2009), we introduce the oversampling ratio $\delta/\rho \geq 1$. For a given $\delta$, we obtain a target S/N from $\overline{\mathrm{S/N}}(\delta)$. Then, we find the smallest $\rho$ for which the nonlinear recovery S/N is greater or equal to $\overline{\mathrm{S/N}}(\delta)$. Thus, the oversampling ratio $\delta/\rho \geq 1$ expresses the sampling overhead required by compressive sensing. This measure helps us to determine the performance of our CS scheme numerically. The smaller this ratio, the smaller the overhead and the more economically favorable this technology becomes compared to conventional sampling schemes.

We calculate for each $\delta \in [0.2,\, 0.8]$

$$\delta/\rho \quad \text{with} \quad \rho = \inf\{\tilde{\rho} : \quad \overline{\mathrm{S/N}}(\delta) \leq \mathrm{S/N}(\tilde{\rho})\}. \tag{2.13}$$

When the sampling ratio approaches one from below ($\delta \to 1$), the data becomes more and more sampled leading to smaller and smaller recovery errors. To match this decreasing error, the sparsity ratio $\rho \to 1$ and consequently we can expect this oversampling ratio to go to one, $\delta/\rho \to 1$.

Remember that in the CS paradigm, acquisition costs grow with the permissible recovery S/N that determines the sparsity ratio. Conversely, the costs of conventional sampling grow with the size of the sampling grid irrespective of the transform-domain compressibility of the wavefield, which in higher dimensions proves to be a major difficulty.

The numerical results of our experiments are summarized in Figure 2.8. Our calculations use empirical S/Ns for both the approximation errors of the synthesis coefficients as a function of $\rho$ and the recovery errors as a function of $\delta$. The estimated curves lead to the following observations. First, as the sampling ratio increases the oversampling ratio decreases, which can be understood because the recovery becomes easier and more accurate. Second, recoveries from simultaneous data have significantly less overhead and curvelets outperform wave atoms, which in turn perform significantly better than wavelets. All curves converge to the lower limit (depicted by the dashed line) as $\delta \to 1$. Because of the large errorbars in the recovery S/Ns (cf. Figure 2.7), the results for the recovery from missing sequential sources are less clear. However, general trends predicted by CS are also observable for this type of acquisition, but the performance is significantly worse than for recovery with simultaneous sources. Finally, the observed oversampling ratios are reasonable for both curvelet and wave atoms.

## 2.5 An academic case study

Now that we established that high S/N's are achievable with modest oversampling ratios, we study the performance of our recovery algorithm on a seismic line from the Gulf of Suez by comparing two simultaneous-source scenarios with coincident source-receiver positions:

- **'Land' acquisition with random amplitude encoding:** Here, sequential impulsive sources are replaced by impulsive simultaneous 'phase-encoded' sources. Mathematically, simultaneous measurements are obtained by replacing the sampling matrix for the sources—normally given by identity matrix—by a measurement matrix obtained by phase encoding along the source coordinate. Following Romberg (2008) and Herrmann et al. (2009), we define the measurement matrix by the following Kronecker product

  $$\mathbf{M} := \left[ \mathbf{I} \otimes \overbrace{\mathrm{diag}\left(\boldsymbol{\eta}\right) \mathcal{F}_s^{*} \mathrm{diag}\left(e^{\hat{i}\boldsymbol{\theta}}\right) \mathcal{F}_s}^{\text{Gaussian matrix}} \otimes \mathbf{I} \right]. \tag{2.14}$$

  In this expression, conventional sampling, which corresponds to the action of the identity matrix $\mathbf{I}$, is replaces by a 'random phase encoding' consisting of applying a Fourier transform along the source coordinate ($\mathcal{F}_s$), followed by uniformly drawn random phase rotations $\theta \in [0, \pi]$, an inverse Fourier transform ($\mathcal{F}_s^*$), and a multiplication by a a random sign vector (i.e., multiplication by $\operatorname{diag}(\boldsymbol{\eta})$ with $(\eta) \in N(0,1)$). As shown by Romberg (2008), the combined action of these operations corresponds to the action of a Gaussian matrix at reduced computational costs (see also Herrmann et al., 2009). Application of this matrix to a conventionally-sampled seismic line turns sequential impulsive source into a simultaneous 'supershot' where all sources fire simultaneously with weights drawn from a single Gaussian distribution. As before, the restriction operator selects a subset of $n_s'$ 'supershots' generated by different randomly-weighted simultaneous sources. After restriction along the source coordinate, the sampling matrix has an aspect(or undersampling) ratio of $\delta = n_s'/n_s$. An example of this type of sampling, resulting in a seismic line consisting of $n_s' \ll n_s$ supershots, is included in Figure 2.9. In this Figure, $n_s$ single impulsive-source experiments (2.9(a) left-hand-side plot) become $n_s'$ simultaneous-source experiments (juxtapose Figure 2.9(a) left-hand-side plot and 2.9(b) left-hand-side plot). While this sort of sampling is perhaps physically unrealizable—i.e., we typically do not have large numbers of vibroseis trucks available—it gives us the most favorable recovery conditions from the compressive-sensing perspective. Therefore, our 'Land' acquisition will serve as a benchmark with which we can compare alternative and physically more realistic acquisition scenarios.

- **'Marine' acquisition with random-time dithering:** Here, sequential acquisition with a single air gun is replaced by continuous acquisition with multiple air guns that fire continuously at random times and at random locations. In this scenario, a seismic line is mapped into a single long 'supershot'. Mathematically, this type of acquisition is represented by the following sampling operator

  $$\mathbf{RM} := [\mathbf{I} \otimes \mathbf{T}]. \tag{2.15}$$

  In this expression, the linear operator $\mathbf{T}$ turns sequential recordings (Figure 2.9(a) right-hand-side plot) with synchronized impulsive shots (Figure 2.9(a) left-hand-side plot) into continuous recordings with $n_s^*$ impulsive sources firing at random positions (Figure 2.9(c) left-hand-side plot), selected uniformly-random from $[1 \cdots n_s]$ discrete source indices and from discrete random time indices, selected uniformly from $(0 \cdots (n_s^* - 1) \times n_t)]$ time indices. Note that $\mathbf{T}$ acts both on the shot and the time coordinate. The resulting data is one long supershot' that contains a superposition of $n_s^*$ impulsive shots. For plotting reasons, we reshaped in Figure 2.9(c) (right-hand-side plot) this long record into multiple shorter records. Notice that this type of 'Marine' acquisition is physically realizable as long as the number of simultaneous sources involved is limited.

Aside from mathematical factors, such as the mutual coherence (cf. Equation 2.5) that determines the recovery quality, there are also economical factors to consider. For this purpose, Berkhout (2008) proposed two performance indicators, which quantify the cost savings associated with simultaneous and continuous acquisition. The first measure compares the number of sources involved in conventional and simultaneous acquisition and is expressed in terms of the source-density ratio

$$\text{SDR} = \frac{\text{number of sources in the simultaneous survey}}{\text{number of sources in the conventional survey}}. \tag{2.16}$$

For 'Land data' acquisition, this quantity equals $\text{SDR}_{Land} = (n_s \times n_s')/n_s = n_s'$ and for 'Marine data' $\text{SDR}_{\text{Marine}} = n_s^*/n_s$. Remember that the number of sources refers the number of sources firing and not the number of source experiments. Clearly, 'Land' acquisition has a significant higher SDR.

Aside from the number of sources, the cost of acquisition is also determined by survey-time ratio

$$\mathrm{STR} = \frac{\text{time of the conventional sequential survey}}{\text{time of the continuous and simultaneous recording}}. \tag{2.17}$$

Ignoring overhead in sequential shooting, this quantity equals $\mathrm{STR}_{\mathrm{Land}} = n_s/n'_s$ in the first and $\mathrm{STR}_{\mathrm{Marine}} = n_s \times T_0/T$ with $T_0$ the time of a single sequential experiment and $T$ the total survey time of the continuous recording. The overall economic performance is measured by the product of these two ratios. For 'Land' acquisition this product is proportional to $n_s$ and for 'Marine' acquisition proportional to $n_s^* \times T_0/T$.

As we have seen from our discussion on compressive sensing, recovery depends on the mutual coherence of the sampling matrix. So, the challenge really lies in the design of acquisition scenarios that obtain the lowest mutual coherence while maximizing the above two economic performance indicators. To get a better insight in how these factors determine the quality of recovered data, we conduct a series of experiments by simulating possible alternative acquisition strategies on a perviously traditionally recorded real seismic line.

First, we simulate 'Land' data for $\delta = 0.5$ (64 simultaneous source experiments with all sources firing) and study the recovery based on 2D and 3D curvelets. The former is based on a 2D discrete curvelet transform along the source and receiver coordinates, and the discrete wavelet transform along the remaining time coordinate:

$$\mathbf{S} := \mathbf{C}_2 \otimes \mathbf{W}. \tag{2.18}$$

We conduct a similar experiment for the 'Marine case'. In this case, we randomly select 128 shots from the total survey time $T = \delta \times (n_s - 1) \times T_0$, yielding the same aspect ratio for the sampling matrix.

Figures 2.10 and 2.11 summarize the results for 'Land' and 'Marine' acquisition using recoveries based on the 2D and 3D curvelet transform. The following observations can be made. First, it is clear that accurate recovery is possible by solving an $\ell_1$ optimization problem using SPG$\ell_1$ (Berg and Friedlander, 2008) while limiting the number of iterations for the 2D case to 500 and the 3D case to 200. Second, the recovery results for 3D recovery of 'Land' data show and improvement 1.3 dB by exploiting 3D structure of the wavefronts. Similarly, we find an improvement of 3.9 dB for the 'Marine' case. Both observations can be explained by the fact that the 3D curvelet transforms attains higher sparsity because it explores continuity of the wavefield along all three coordinate axes. Second, 'Land' acquisition clearly favors recovery by curvelet-domain sparsity promotion compared to 'Marine' acquisition. This is true despite the fact that the subsampling ratio, i.e., the aspect ratio of the sampling matrices, are the same. Clearly this difference lies in the mutual coherence of the sampling matrix. The columns of the sampling matrix for 'Land' acquisition are more incoherent and hence more independent and this favors recovery. These observations are confirmed by the S/Ns, which for 'Land' acquisition equal 10.3 dB and 11.6dB, for the 2D/3D recovery, respectively, and 7.2 dB and 11.1 dB, for 'Marine' acquisition.

Unfortunately, recovery quality is not the only consideration. The economics expressed by the SDR and STR also play a role. In the above setting, the 'Land' acquisition has a SDR = 64 and STR = 2 while the 'Marine' acquisition has SDR = 1 and STR = 2. Clearly, the SDR for land acquisition may not be realistic.

## 2.6 Discussion

The presented results illustrate that we are at the cusp of exciting new developments where acquisition workflows are no longer impeded by subsampling related artifacts. Instead, we arrive at acquisition schemes that control these artifacts. We accomplish by applying the following design principles: (i) **randomize**—break coherent aliases by introducing randomness, e.g. by designing randomly perturbed acquisition grids, or by designing randomized simultaneous sources; and (ii) **sparsify**—utilize sparsifying transforms in conjunction with sparsity-promoting programs that separate signal and subsampling artifacts and that restore amplitudes. The implications of randomized incoherent sampling go far beyond the examples presented here. For instance, our approach is applicable to land acquisition for physically realizable sources (Krohn and Neelamani, 2008; Romberg, 2008) and can be used to compute solutions to wavefield simulations (Herrmann et al., 2009) and to compute full waveform inversion (Herrmann et al., 2009) faster. Because randomized sampling is linear (Bobin et al., 2008), wavefield reconstructions and processing can be carried out incrementally as more compressive data becomes available.

Indeed, compressive sensing offers enticing perspectives towards the design of future Land and Marine acquisition systems. In order for this technology to become successful the following issues need to be addressed, namely the performance of recovery

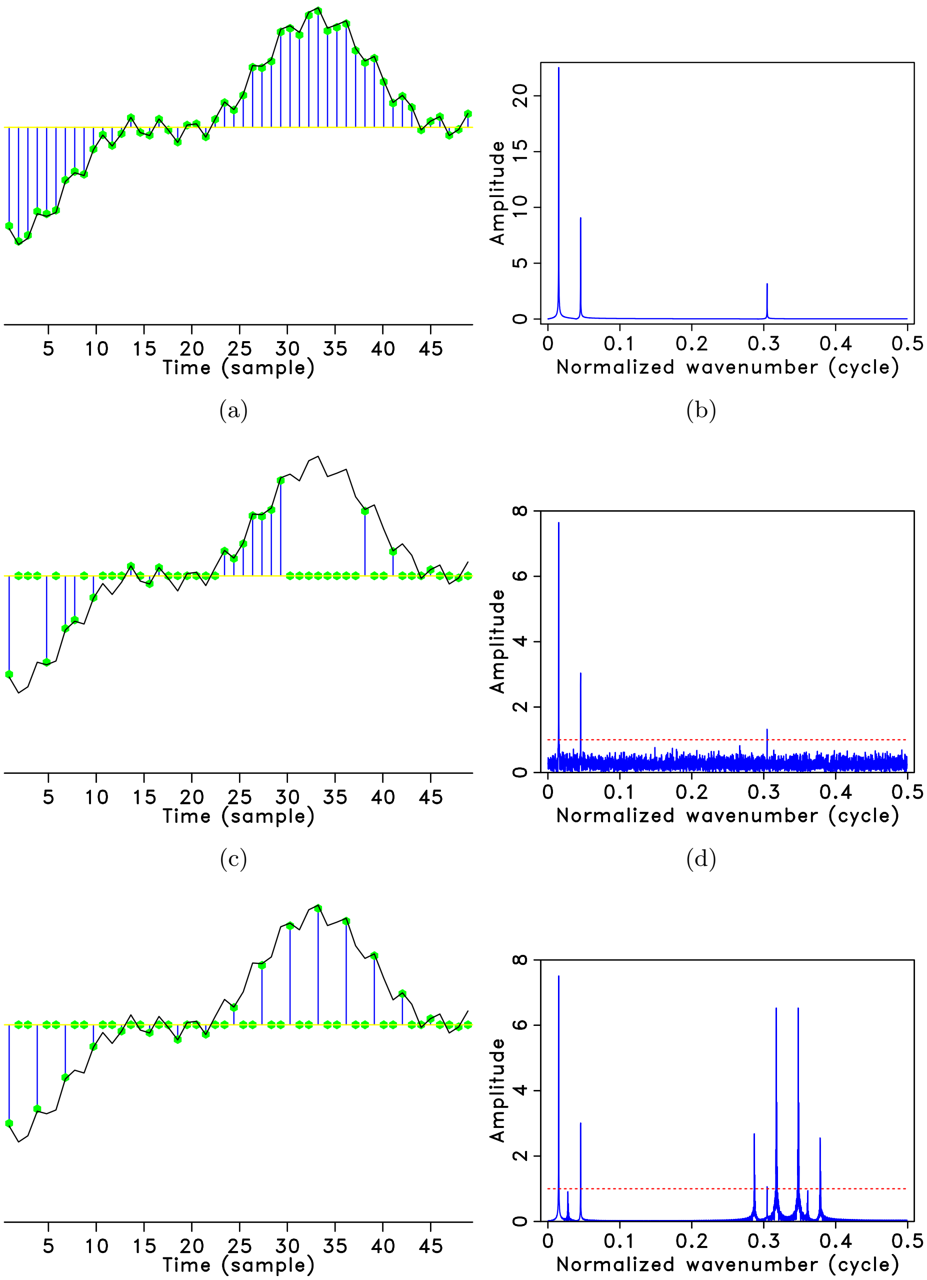
- from field data including all its idiosyncrasies. This will require an concerted effort from practitioners in the field and theoreticians. For Marine acquisition, recent work by Moldoveanu (2010) has shown early indications that randomized jittered sampling leads to improved imaging.
- from discrete data with quantization errors. Addressing this issue calls for integration of digital-to-analog conversion into compressive and recent progress has been made in this area (see e.g. Güntürk et al., 2010);
- from Land data that has the imprint of statics. Addressing this issue will be essential because severe static effects may adversely affect transform-domain sparsity on which recovery from compressive-sampled data relies.

## 2.7 Conclusions

Following ideas from compressive sensing, we made the case that seismic wavefields can be reconstructed with a controllable error from randomized subsamplings. By means of carefully designed numerical experiments on synthetic and real data, we established that compressive sensing can indeed successfully be adapted to seismic data acquisition, leading to a new generation of randomized acquisition and processing methodologies.

With carefully designed experiments and the introduction of performance measures for nonlinear approximation and recovery errors, we established that curvelets perform best in recovery, closely followed by wave atoms, and with wavelets coming in as a distant third, which is consistent with the directional nature of seismic wavefronts. This finding is remarkable for the following reasons: (i) it underlines the importance of sparsity promotion, which offsets the "costs" of redundancy and (ii) it shows that the relative sparsity ratio effectively determines the recovery performance rather than the absolute number of significant coefficients. Our observation of significantly improved recovery for simultaneous-source acquisition also confirms predictions of compressive sensing. Finally, our analysis showed that accurate recoveries are possible from compressively sampled data volumes that exceed the size of conventionally compressed data volumes by only a small factor.

The fact that compressive sensing combines sampling and compression in a single linear encoding step has profound implications for exploration seismology that include: a new randomized sampling paradigm, where the cost of acquisition are no longer dominated by resolution and size of the acquisition area, but by the desired reconstruction error and transform domain sparsity of the wavefield, and a new paradigm for randomized processing and inversion, where dimensionality reductions will allow us to mine high-dimensional data volumes for information in ways, which previously, would have been computationally infeasible.

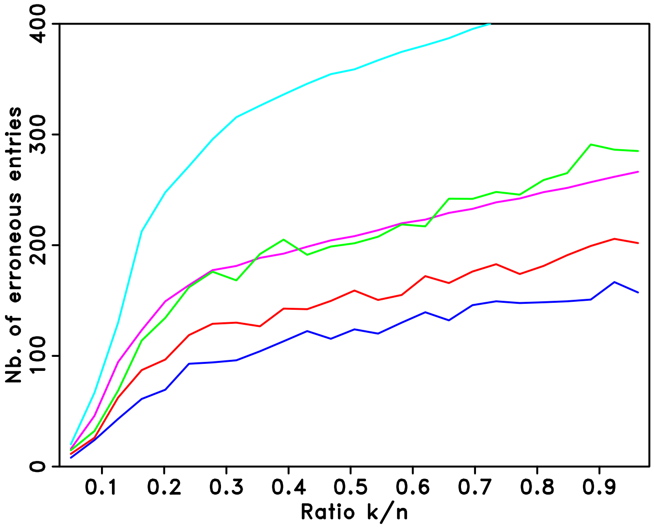


(a) (b)

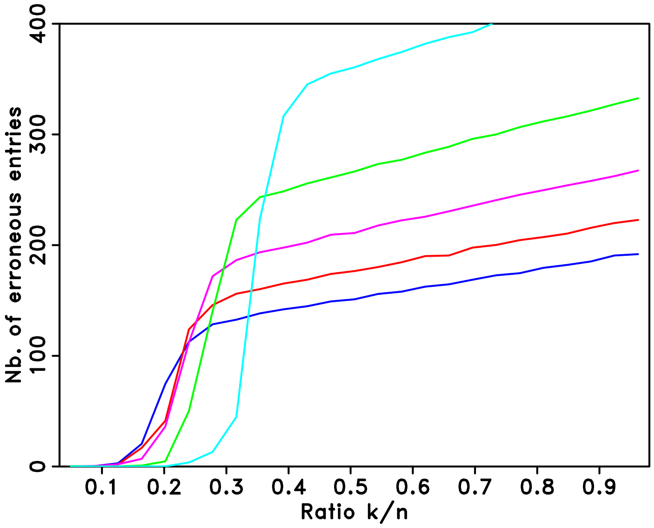
(c) (d)

(e) (f)

**Figure 2.1:** Different (sub)sampling schemes and their imprint in the Fourier domain for a signal that is the superposition of three cosine functions. Signal **(a)** regularly sampled above Nyquist rate, **(c)** randomly three-fold undersampled according to a discrete uniform distribution, and **(e)** regularly three-fold undersampled. The respective amplitude spectra are plotted in **(b)**, **(d)** and **(f)**. Unlike aliases, the subsampling artifacts due to random subsampling can easily be removed using a standard denoising technique, e.g., nonlinear thresholding (dashed line), effectively recovering the original signal (adapted from (Hennenfent and Herrmann, 2008)).

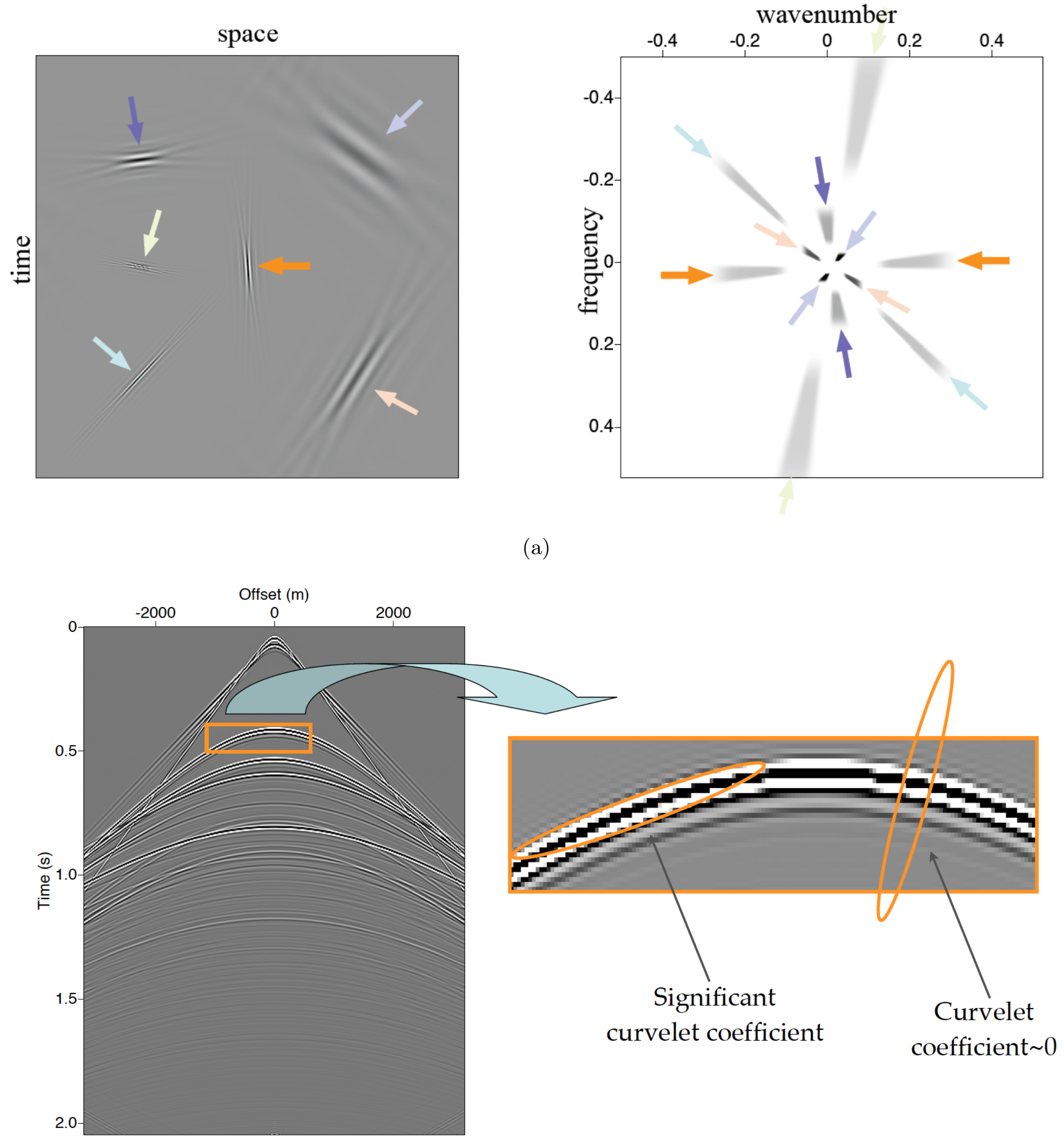


(a)



(b)

**Figure 2.2:** Averaged recovery error percentages for a $k$-sparse Fourier vector reconstructed from $n$ time samples taken **(a)** regularly and **(b)** uniformly-randomly. In each plot, the curves from top to bottom correspond to a subsampling factor ranging from two to six (adapted from Hennenfent and Herrmann (2008)).



(a)

(b)

**Figure 2.3:** Curvelets and seismic data. **(a)** 2D curvelets in the time-space domain and the frequency-wavenumber domain. **(b)** Curvelets approximate curved singularities, i.e., wavefronts, with very few significant curvelet coefficients.

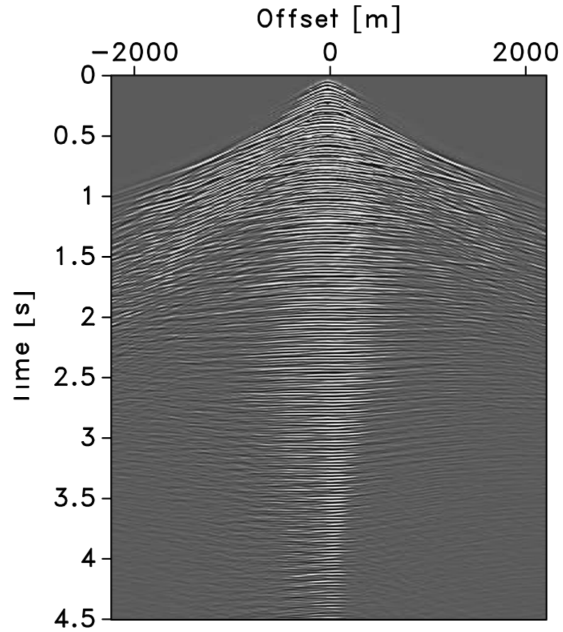


**Figure 2.4:** Real common-receiver gather from Gulf of Suez data set.

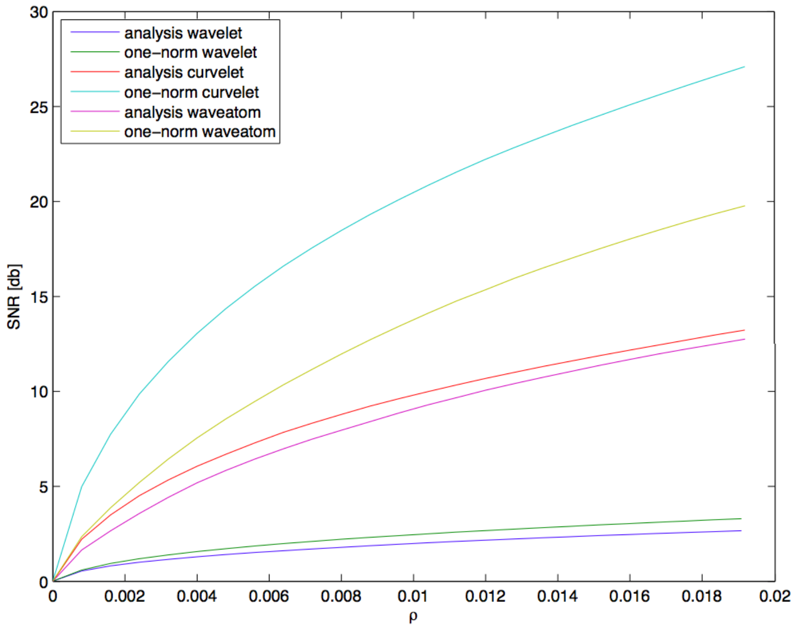
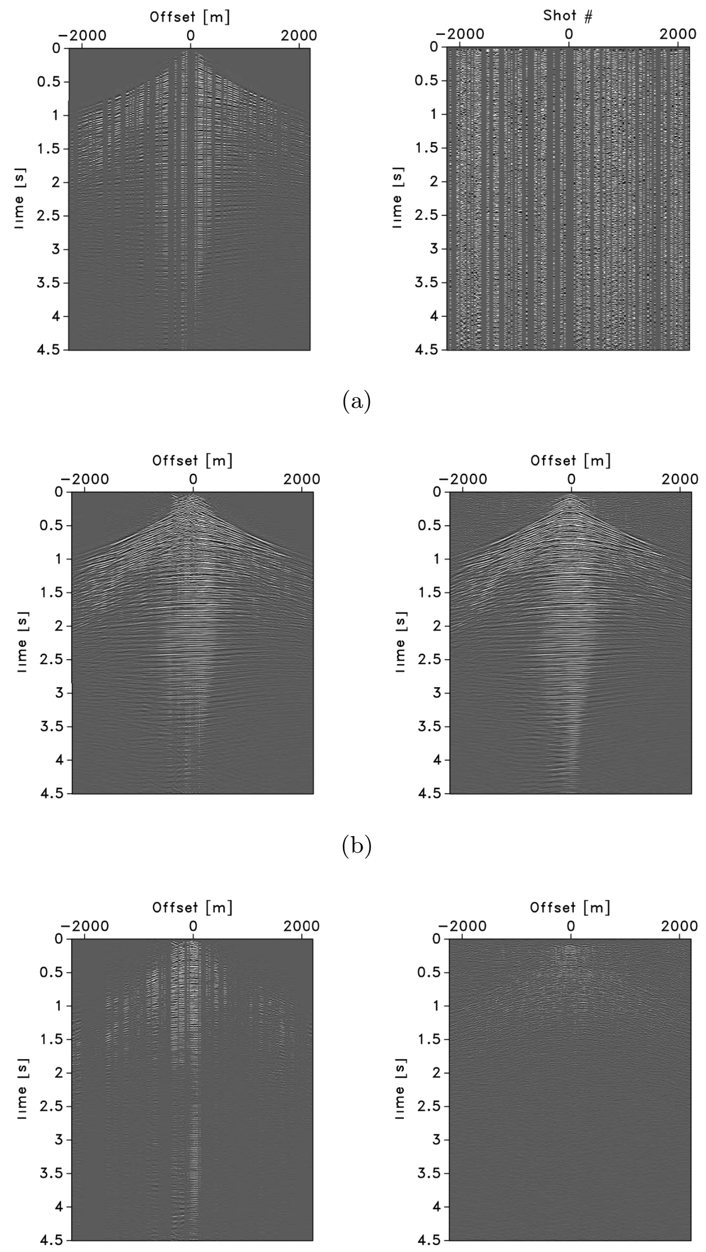


**Figure 2.5:** Signal-to-noise ratios (S/Ns) for the nonlinear approximation errors of the common-receiver gather plotted in Figure 2.4. The S/Ns are plotted as a function of the sparsity ratio $\rho \in (0, 0.02]$. The plots include curves for the errors obtained from the analysis and one-norm minimized synthesis coefficients. Notice the significant improvement in S/Ns for the synthesis coefficients obtained by solving Equation 2.10.



(a)

(b)

(c)

**Figure 2.6:** Recovery from a compressively-sampled common-receiver gather with 50% ($\delta = 0.5$) of the sources missing. **(a)** Left: Receiver gather with sequential shots selected uniformly at random. **(a)** Right: The same but for random simultaneous shots. **(b)** Left: Recovery from incomplete data in **(a)** left-hand-side plot. **(b)** Right: The same but now for the data in **(a)** right-hand-side plot. **(c)** Left: Difference plot between the data in Figure 2.4 and the recovery in **(b)** left-hand-side plot. **(c)** Right: The same but now for recovery from simultaneous data in **(a)** right-hand-side plot. Notice the remarkable improvement in the recovery from simultaneous data.

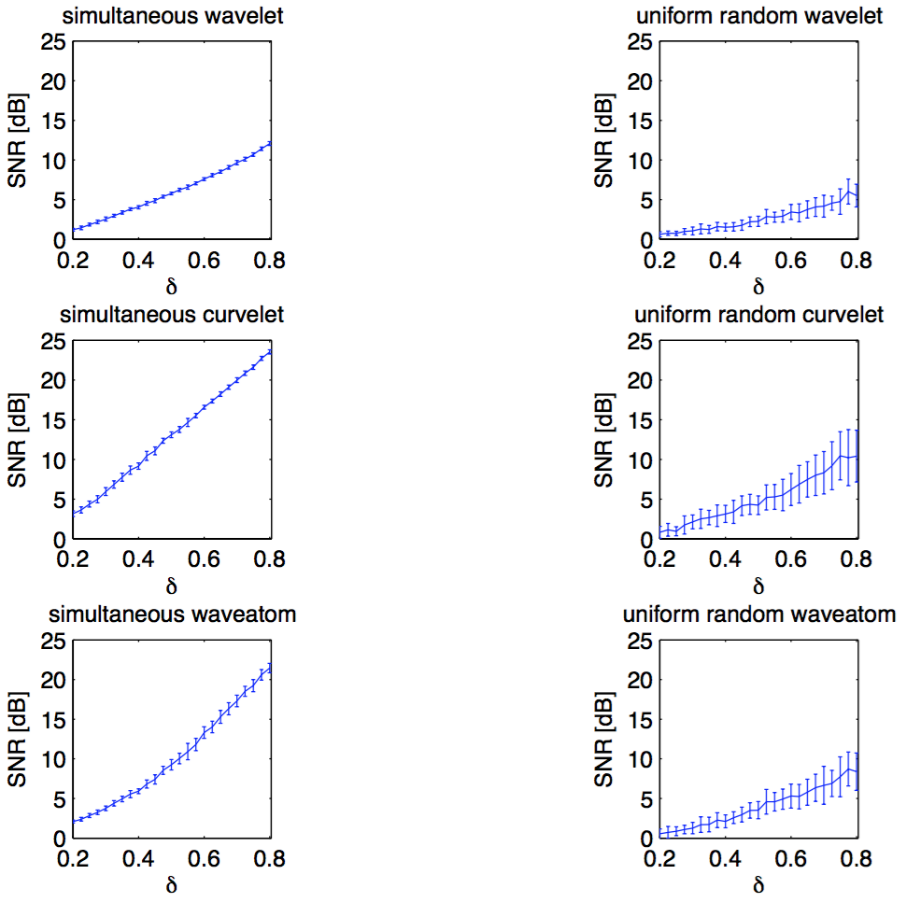


**Figure 2.7:** S/Ns (cf. Equation 2.12) for nonlinear sparsity-promoting recovery from compressively sampled data with 20% − 80% of the sources missing ($\delta \in [0.2,\ 0.8]$). The results summarize 25 experiments for 25 different values of $\delta \in [0.2,\ 0.8]$. The plots include estimates for the standard deviations. From these results, it is clear that simultaneous acquisition (results in the left column) is more conducive to sparsity-promoting recovery. Curvelet-based recovery seems to work best, especially towards high percentages of data missing.

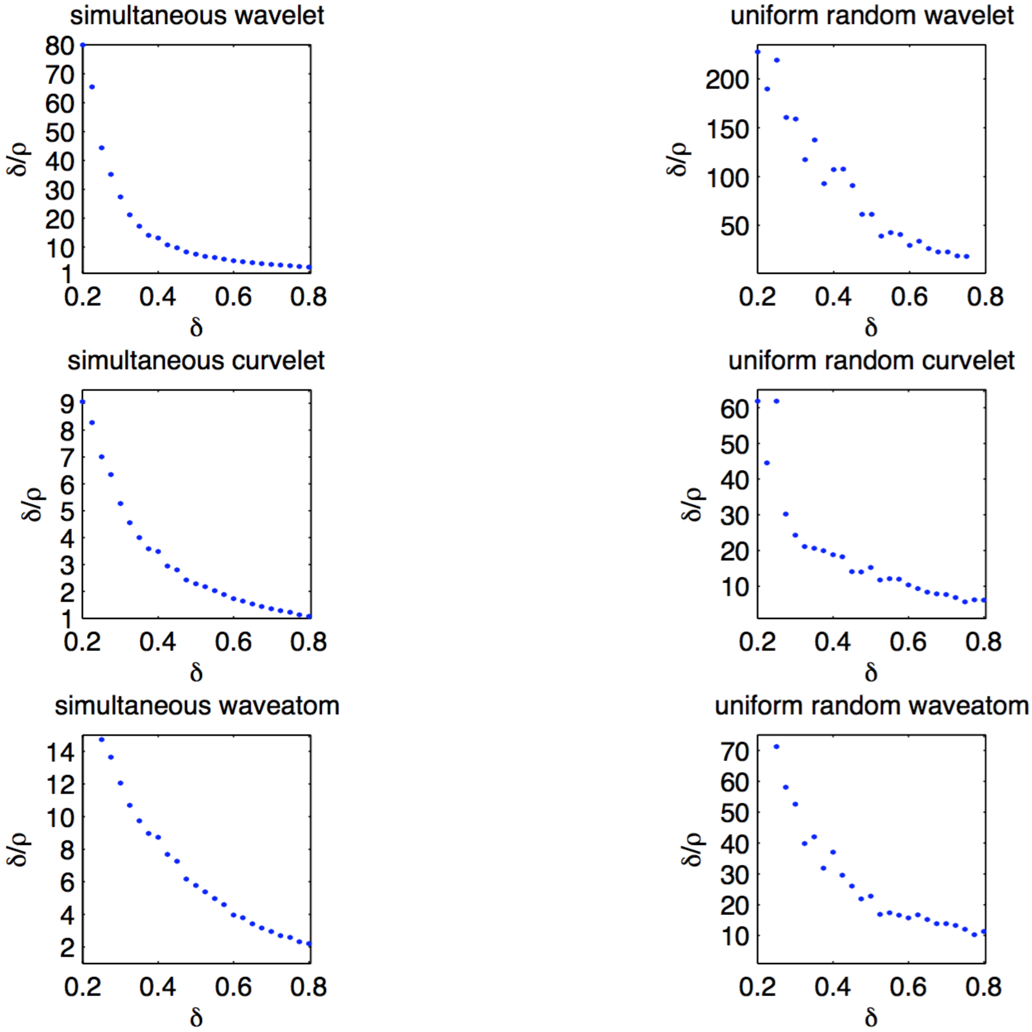
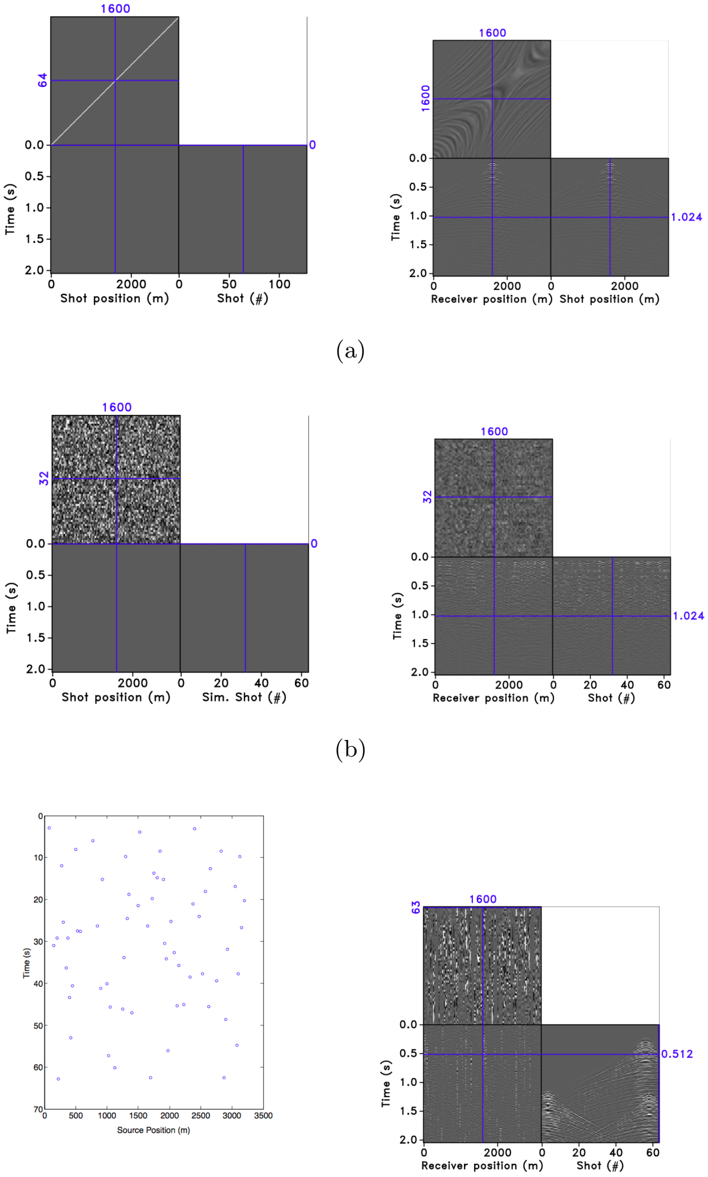


**Figure 2.8:** Oversampling ratio $\delta/\rho$ as a function of the sampling ratio $\delta$ (cf. Equation 2.13) for sequential- and simultaneous-source experiments. As expected, the overhead is smallest for simultaneous acquisition and curvelet-based recovery.

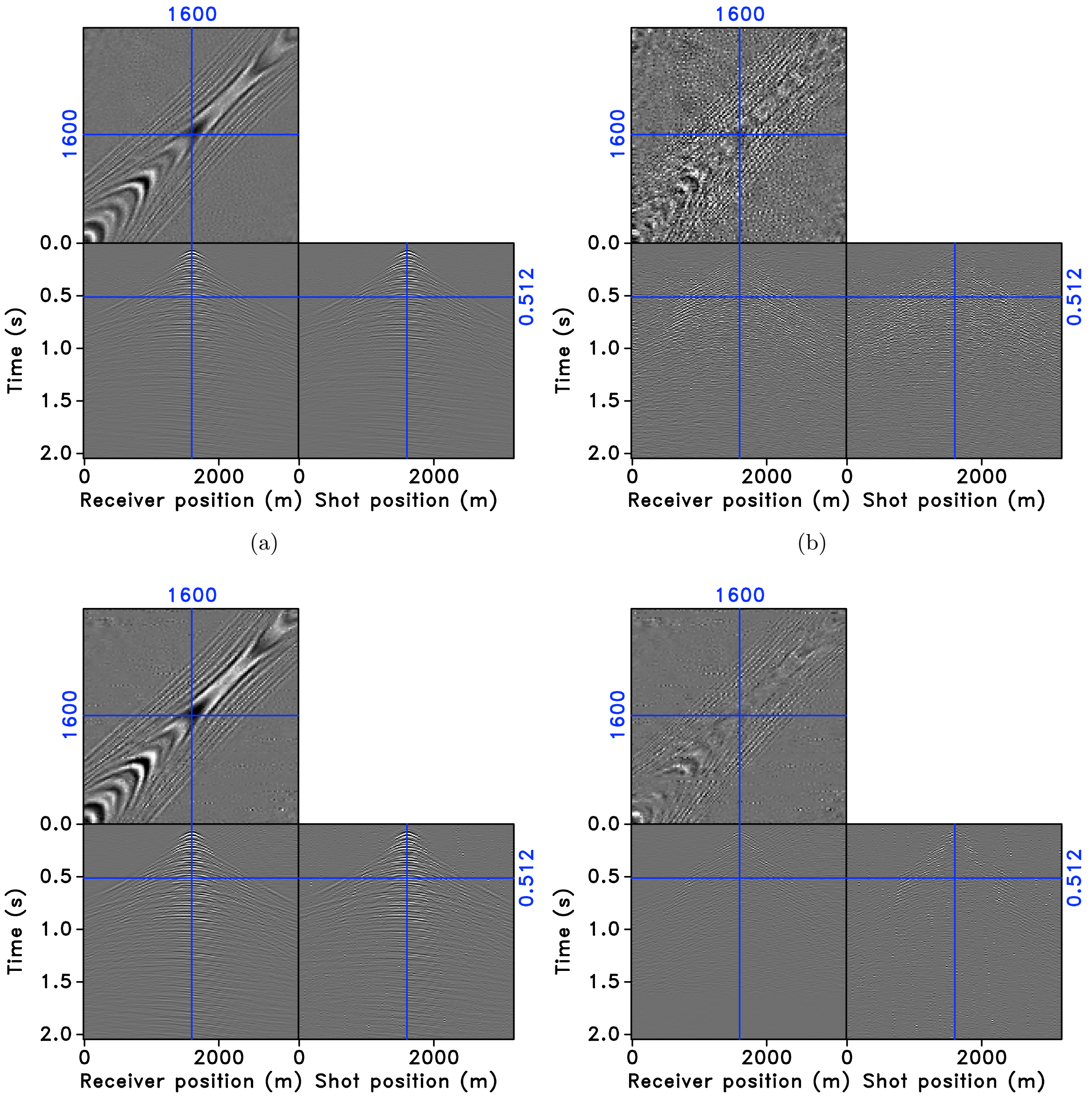


(a)

(b)

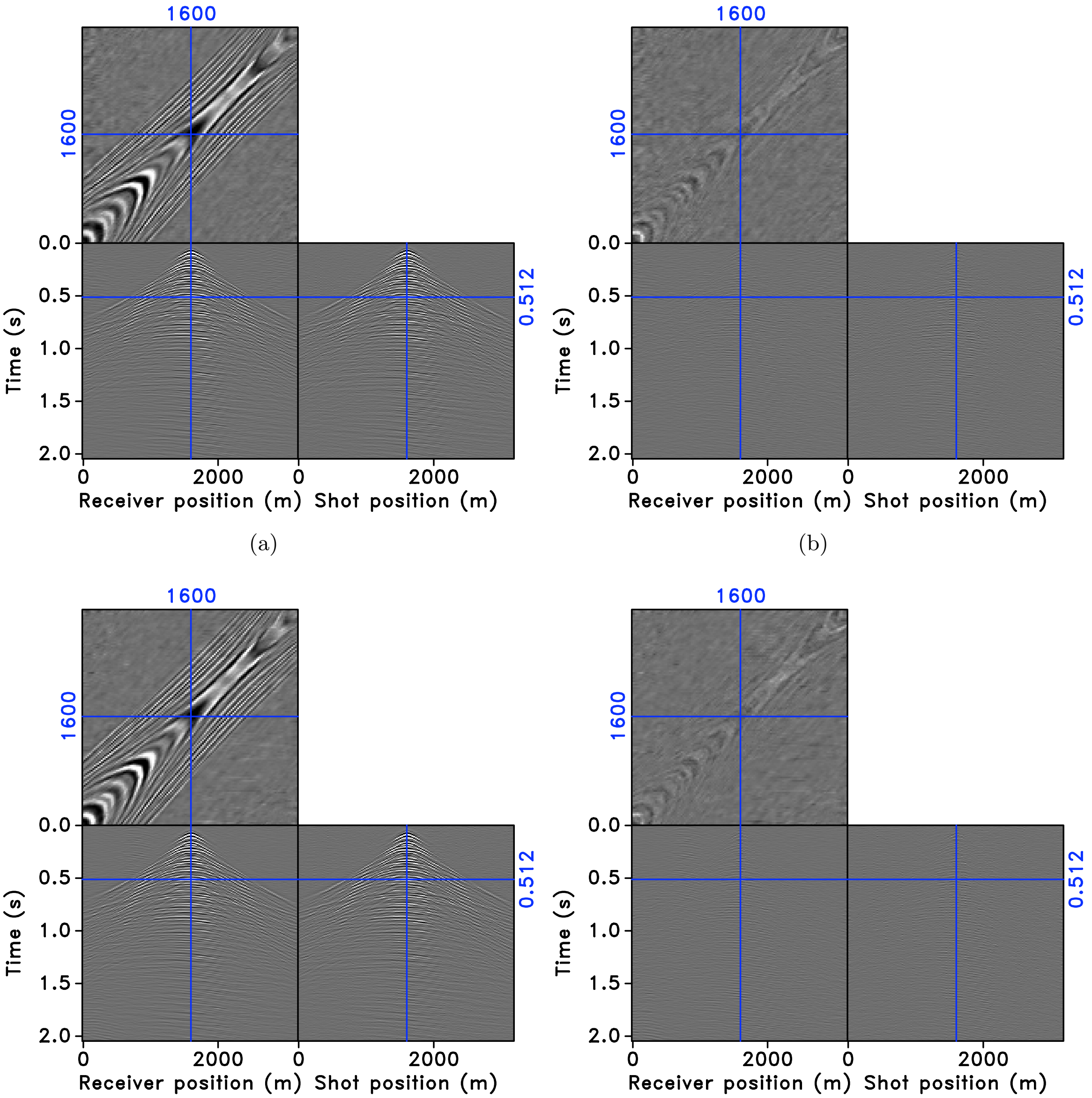
(c)

**Figure 2.9:** Different acquisition scenarios. **(a)** Left: Impulsive sources for conventional sequential acquisition, yielding 128 shot records for 128 receivers and 512 time samples. **(a)** Right: Corresponding fully sampled sequential data. **(b)** Left: Simultaneous sources for 'Land' acquisition with 64 simultaneous-source experiments. Notice that all shots fire simultaneously in this case. **(b)** Right: Corresponding compressively sampled land data. **(c)** Left: Simultaneous sources for 'Marine' acquisition with 128 sources firing at random times and locations during a continuous total 'survey' time of $T = 262$s. **(c)** Right: Corresponding 'Marine' data plotted as a conventional seismic line.



(a) (b) (c) (d)

**Figure 2.10:** Sparsity-promoting recovery with $\delta = 0.5$ with the 2D curvelet transforms. **(a)** 2D curvelet-based recovery from 'Land' data (10.3 dB). **(b)** The corresponding difference plot. **(c)** 2D curvelet-based recovery from 'Marine' data (7.2 dB). **(d)** Corresponding difference plot. Notice the improvement in recovery from 'Land' data.



(a) (b)

(c) (d)

**Figure 2.11:** Sparsity-promoting recovery with $\delta = 0.5$ with the 3D curvelet transforms. **(a)** 3D curvelet-based recovery from 'Land' data (11.6 dB). **(b)** The corresponding difference plot. **(c)** 3D curvelet-based recovery from 'Marine' data (11.1 dB). **(d)** Corresponding difference plot. Notice the improvement in recovery compared to 2D curvelet based recovery.

# Chapter 3

# Randomized marine acquisition with compressive sampling matrices

## 3.1 Summary

Seismic data acquisition in marine environments is a costly process that calls for the adoption of simultaneous-source or randomized acquisition - an emerging technology that is stimulating both geophysical research and commercial efforts. Simultaneous marine acquisition calls for the development of a new set of design principles and post-processing tools. In this chapter, we discuss the properties of a specific class of randomized simultaneous acquisition matrices and demonstrate that sparsity-promoting recovery improves the quality of the reconstructed seismic data volumes. We propose a practical randomized marine acquisition scheme where the sequential sources fire air guns at only randomly time-dithered instances. We demonstrate that the recovery using sparse approximation from random time-dithering with a single source approaches the recovery from simultaneous-source acquisition with multiple sources. Established findings from the field of compressive sensing indicate that the choice of the sparsifying transform that is incoherent with the compressive sampling matrix can significantly impact the reconstruction quality. Leveraging these findings, we then demonstrate that the compressive sampling matrix resulting from our proposed sampling scheme is incoherent with the curvelet transform. The combined measurement matrix exhibits better isometry properties than other transform bases such as a non-localized multidimensional Fourier transform. We illustrate our results with simulations of "ideal" simultaneous-source marine acquisition, which dithers both in time and space, compared with periodic and randomized time-dithering.

## 3.2 Introduction

Data acquisition in seismic exploration forms one of the bottlenecks in seismic imaging and inversion. It involves the collection and processing of massive data volumes, which can be up to 5-dimensional in nature (2D for the source positions × 2D for the receiver positions × 1D for the time dimension). Constrained by the Nyquist sampling rate, the increasing sizes of these data volumes pose a fundamental shortcoming in the traditional sampling paradigm as the size and desired

---

A version of this chapter has been published in Geophysical Prospecting, 2012, vol. 60, pp. 648–662.

resolution of the survey areas continue to grow.

Conventional marine acquisition is carried out as single-source experiments of the subsurface response. This means that the size of the field recording is a product of the number of source locations, the number of receiver locations active per source experiment, and the number of discretized time samples proportional to the length of the reflection series. To reduce the threat of aliasing, the source and receiver locations are preferably bound to a grid spacing of less than 50 meters, therefore the number of receivers and sources are in turn direct functions of the size of the survey area.

Geological considerations, such as the presence of salt bodies in the Gulf of Mexico, create a need for wide azimuth coverage (i.e., receivers at large offsets), which pushes survey areas to cover thousands of square kilometres, with transverse lengths of 50 kilometres or more. However, adhering to a conventional single-source recording scheme requires recording vessels to move at a pace no more than 10 kilometres per hour to maintain the desired source spacing. This represents a direct conflict to the productivity of surveys, and makes large area acquisition particularly expensive. On the other hand, it is entirely possible and even desirable in terms of streamer stability for survey vessels to move faster (up to operational limits such as streamer drag), but doing so will not usually allow sufficient time for the single-source seismic responses to completely decay before the vessels have reached the next source position.

Several works in the seismic imaging literature have explored the concept of simultaneous or blended source activation to account for this situation Beasley et al. (1998); de Kok and Gillespie (2002); Beasley (2008); Berkhout (2008); Hampson et al. (2008). When sources are fired simultaneously, the main issue is the resulting interference between the responses of the different sources that makes it difficult to estimate interference-free shot gathers. Therefore, the challenge is to recover subtle late reflections that can be overlaid by interfering seismic response from other shots. We will show that this challenge can be effectively addressed through recovery by sparsity promotion.

Recently, "compressed sensing" (Donoho, 2006; Candès and Tao, 2006) has emerged as an alternate sampling paradigm in which randomized sub-Nyquist sampling is used to capture the structure of the data with the assumption that it is sparse or compressible in some transform domain. A signal is said to admit a sparse (or compressible) representation in a transform domain if only a small number $k$ of the transform coefficients are nonzero (or if the signal can be well approximated by the $k$ largest-in-magnitude transform coefficients). In seismic exploration, data consists of wavefronts that exhibit structure in multiple dimensions. With the appropriate data transformation, we capture this structure by a small number of significant transform coefficients resulting in a sparse representation of data.

We rely on the compressed sensing literature to analyze a physically realizable simultaneous-source marine acquisition technology where acquisition related costs are no longer determined by the Nyquist sampling criteria. We also propose a random time-dithered acquisition scheme whose performance using a single source approaches that of simultaneous-source acquisition. Under this paradigm, data are no longer collected as separate shot records with single-source experiments. Instead, we continuously record over the whole acquisition process, collecting a single long "super-shot" record that is acquired over a time interval shorter than the cumulative time of conventional marine acquisition (excluding downtime and overhead such as vessel turning). We then recover the canonical sequential single-source shot record by solving a sparsity promoting problem. The contributions of this work can be summarized as follows:

- We develop the relation between simultaneous-source sampling that is physically realizable in a marine setting and curvelet-based sparse recovery.

- We propose a random time-dithering marine acquisition scheme which can deliver even in the case of a single source a sparse recovery performance that approaches that of simultaneous-source acquisition with multiple sources with random time and space dithering.
- Through Monte-Carlo simulations, we estimate quantitative measures introduced in Compressive Sensing that predict the performance of sparse recovery for a particular acquisition design.
- Through simulated experiments, we demonstrate the performance of the proposed sampling matrices in recovering real prestack seismic data.

## 3.3 Related work

The earliest works on simultaneous acquisition were formulated with land acquisition in mind, where the prevalent use of vibroseis sweeps have allowed the freedom of employing sophisticated codes in source signatures as a way to differentiate between the responses due to different simultaneous sources (Allen et al., 1998; Beasley et al., 1998; Bagaini, 2006; Lin and Herrmann, 2009a). Marine acquisition, on the other hand, rarely employs the marine vibrators that are the analog of the vibroseis due to poor signal-to-noise ratio and low-frequency content. Consequently, the marine case of simultaneous acquisition was less well-explored, as the more commonly used impulsive air-gun sources are considerably more rigid when it comes to manipulating its signature. Literature directly discussing the marine case topic did not appear until Beasley et al. (1998), where the source "encoding" are limited to dithered activation times and locations. Near-simultaneous marine cases were also discussed in de Kok and Gillespie (2002) and Hampson et al. (2008). The main messages in these works seem to be that interferences due to simultaneous firing are often ignorable after stacking and simple filtering, and thus do not seriously impact the imaging step.

However, many subsurface attribute inversion schemes still rely on single-source prestack data. Recovering these volumes from simultaneous marine recordings did not truly become feasible until the recognition that, as long as the shot timings are suitably randomly delayed, the resulting interferences will appear noise-like in specific gather domains such as common-offset and common-receiver. This property differentiates these events from responses due to single-source experiments that remain coherent in these gather domains. This observation was reported in Stefani et al. (2007); Moore et al. (2008) and Akerberg et al. (2008) with application to land acquisition in Bagaini and Ji (2010). Subsequent processing techniques, which aim to remove noise-like source crosstalk, vary from simple filters (Huo et al., 2009) to more sophisticated inversion-type algorithms (Moore, 2010; Abma et al., 2010; Mahdad et al., 2011). The latter are designed to take advantage of sparse representations for coherent seismic signals.

The aforementioned works did not investigate the link between sparsity-based recovery and the specific properties of the acquisition system, but theoretical results from compressive sensing *do* suggest a direct relationship between acquisition design and the expected fidelity of the achievable recovery. An analysis for practical acquisition systems exist in terms of incoherency arguments (Blacquière et al., 2009, which interestingly also considers receiver-side blending), but analysis in terms of compressive sensing arguments remain challenging, as most existing mathematical results in compressive sensing deal with rather abstract acquisition systems. Evidently, existing works relating to seismic acquisition and compressive sensing only seem to suggest schemes suitable for forward-modelling in the computer. Neelamani et al. (2008, 2010) suggested an acquisition that uses noise-like signals as sweeps for land-based acquisition. Herrmann et al. (2009) on the other hand uses impulsive sources but requires modulation of each source by a randomly-determined scaling in amplitude. These papers also suggested recovery using sparse inversion of the curvelet representation of the data. We will keep the sparse inversion technique but consider more practical acquisition systems suitable for field marine acquisition.

## 3.4 Compressed sensing overview

Compressive sensing (abbreviated as CS throughout the chapter) is a process of acquiring random linear measurements of a signal and then reconstructing it by utilizing the prior knowledge that the signal is sparse or compressible in some transform domain. One of the main advantages of CS is that it combines sampling and compression in a single linear step, thus reducing the cost of traditional Nyquist sampling followed by dimensionality reduction through data encoding. A direct application which can benefit from this feature of CS is seismic acquisition where the acquisition costs are now quantified by the transform-domain sparsity of seismic data instead of by the grid size.

### 3.4.1 The sparse recovery problem

Suppose that $\mathbf{x}_0$ is an $P$ dimensional vector with at most $k \ll P$ nonzero entries. The sparse recovery problem involves solving an underdetermined system of equations

$$\mathbf{b} = \mathbf{A}\mathbf{x_0}, \tag{3.1}$$

where $\mathbf{b} \in \mathbb{C}^n$, $n < P$ represents the compressively sampled data of $n$ measurements, and $\mathbf{A} \in \mathbb{C}^{n \times P}$ represents the measurement matrix. When $\mathbf{x_0}$ is sparse—i.e., when there are only $k < n$ nonzero entries in $\mathbf{x_0}$— sparsity-promoting recovery can be achieved by solving the $\ell_0$ minimization problem

$$\tilde{\mathbf{x}} = \arg\min_{\mathbf{x} \in \mathbb{C}^P} \|\mathbf{x}\|_0 \quad \text{subject to } \mathbf{b} = \mathbf{A}\mathbf{x}, \tag{3.2}$$

where $\tilde{\mathbf{x}}$ represents the sparse approximation of $\mathbf{x_0}$, and the $\ell_0$ norm $\|\mathbf{x}\|_0$ is the number of non-zero entries in $\mathbf{x_0}$. Note that if the $\ell_0$ minimization problem were solvable in practice and every $n \times n$ submatrix of $\mathbf{A}$ is invertible, then $\tilde{\mathbf{x}} = \mathbf{x_0}$ when $k < n/2$ (Donoho and Elad, 2003).

However, $\ell_0$ minimization is a combinatorial problem and quickly becomes intractable as the dimensions increase. Instead, the basis pursuit (BP) convex optimization problem shown below can be used to recover an estimate $\tilde{\mathbf{x}}$ at the cost of decreasing the level of recoverable sparsity $k$ — e.g. $k \lesssim n/\log(N/n) < n/2$ when $\mathbf{A}$ is a Gaussian matrix with independent identically distributed (i.i.d.) entries (Candès et al., 2006c; Donoho, 2006). The BP problem is given by

$$\tilde{\mathbf{x}} = \arg\min_{\mathbf{x} \in \mathbb{C}^P} \|\mathbf{x}\|_1 \quad \text{subject to } \mathbf{b} = \mathbf{A}\mathbf{x}, \tag{3.3}$$

where $\tilde{\mathbf{x}}$ represents the sparse (or compressible) approximation of $\mathbf{x_0}$, and the $\ell_1$ norm $\|\mathbf{x}\|_1$ is the sum of absolute values of the elements of a vector $\mathbf{x}$. The BP problem typically finds a sparse or (under some conditions) the sparsest solution that explains the data exactly.

Finally, we note that $\mathbf{x_0}$ can be the sparse expansion of a physical domain signal $\mathbf{f}_0 \in \mathbb{C}^N$ in some transform domain characterized by the operator $\mathbf{S} \in \mathbb{C}^{P \times N}$ with $P \geq N$. In this case, $\mathbf{A}$ can be composed of the product of a sampling operator $\mathbf{RM}$ with the sparsifying operator $\mathbf{S}$ such that $\mathbf{A} = \mathbf{RMS}^H$, where $^H$ denotes the Hermitian transpose. Consequently, the acquired measurements $\mathbf{b}$ are given by

$$\mathbf{b} = \mathbf{A}\mathbf{x_0} = \mathbf{RMf_0}.$$

We will elaborate more on this concept of sparse recovery in the randomized marine acquisition section.

### 3.4.2 Recovery conditions

Next, we discuss some conditions that make unique recovery possible despite the fact that the linear system we are solving is underdetermined, meaning that we have fewer equations than unknowns. We present two sets of conditions, those that guarantee recovery of any arbitrary signal $\mathbf{x_0}$, and those that are specialized for a particular class of signals $\mathbf{x_0}$.

Suppose that the vector $\mathbf{x_0}$ is an arbitrary signal that can be well approximated by the vector $\mathbf{x_k}$ which contains only the largest $k$ coefficients of $\mathbf{x_0}$, i.e., the largest $k \ll P$ nonzero entries of $\mathbf{x_k}$ contain most of the energy of $\mathbf{x_0}$. Let $a$ be a number larger than 1. As long as any subset $\Lambda$ of $ak$ columns of the $n \times P$ matrix $\mathbf{A}$ are linearly independent and constitute a submatrix $\mathbf{A}_\Lambda$ which is invertible and has a small condition number (close to 1), there exists some algorithm that recovers $\mathbf{x_0}$ exactly.

To quantify this property of $\mathbf{A}$, Candès and Tao (2005) define the *restricted isometry property* (RIP) which states that there exists a constant $0 < \delta_{ak} < 1$ for which

$$(1 - \delta_{ak})\|\mathbf{u}\|_2^2 \leq \|\mathbf{A}_\Lambda \mathbf{u}\|_2^2 \leq (1 + \delta_{ak})\|\mathbf{u}\|_2^2, \tag{3.4}$$

where $\Lambda$ is any subset of $\{1 \dots P\}$ of size $|\Lambda| \leq ak$, $\mathbf{A}_\Lambda$ is the submatrix of $\mathbf{A}$ whose columns are indexed by $\Lambda$, and $\mathbf{u}$ is an arbitrary $k$-dimensional vector. The definition above indicates that if every submatrix $\mathbf{A}_\Lambda$ has an RIP constant that is close to zero, then its condition number approaches 1. More precisely, let $\sigma_{min}$ and $\sigma_{max}$ be the smallest and largest singular values of $\mathbf{A}_\Lambda$, respectively, the RIP constant

$$\delta_{ak} = \sup_{\Lambda \in \{1, \dots P\}} \max\{1 - \sigma_{\min}, \sigma_{\max} - 1\}. \tag{3.5}$$

That is, the RIP constant is the smallest upper bound on the maximum of $\{1 - \sigma_{\min}, \sigma_{\max} - 1\}$ for all subsets $\Lambda \in \{1, \dots P\}$ of size $ak$.

The RIP constant is difficult to compute since it requires evaluating $\delta_{ak}$ for every subset of $ak$ columns of $\mathbf{A}$ and there are $\binom{P}{ak}$ of such subsets. However, it is possible to find theoretical upper bounds on the RIP constant for matrices whose entries are drawn i.i.d from sub-Gaussian distributions. Otherwise, Monte Carlo simulations are used to approximate the value of $\delta_{ak}$. It was shown in (Candès et al., 2006c) that if $\delta_{(a+1)k} < \frac{a-1}{a+1}$ (e.g. $\delta_{3k} < \frac{1}{3}$), then the BP problem can recover an approximation $\tilde{\mathbf{x}}$ to $\mathbf{x_0}$ with an error bounded by

$$\|\tilde{\mathbf{x}} - \mathbf{x_0}\|_2 \leq \frac{C(\delta_{ak})}{\sqrt{k}}\|\mathbf{x_k} - \mathbf{x_0}\|_1, \tag{3.6}$$

where $C(\delta_{ak})$ is a well-behaved constant. This error bound indicates that if the matrix $\mathbf{A}$ has the RIP for a specific sparsity level $k$, then the recovery error is bounded by the best $k$-term approximation error of the signal. Finally, we note that the condition on the RIP constant was later improved by Candès (2008) to $\delta_{2k} < \sqrt{2} - 1$.

An easier to compute but less informative characterization of $\mathbf{A}$ is the *mutual coherence* $\mu(\mathbf{A})$. The mutual coherence, which measures correlations between the columns of $\mathbf{A}$, provides an upper bound on $\delta_k < (k-1)\mu(\mathbf{A})$ and is given by

$$\mu(\mathbf{A}) = \max_{1 \leq i \neq j \leq P} \frac{|\mathbf{a}_i^H \mathbf{a}_j|}{(\|\mathbf{a}_i\|_2 \cdot \|\mathbf{a}_j\|_2)}, \tag{3.7}$$

where $\mathbf{a}_i$ is the $i^{th}$ column of $\mathbf{A}$. It is evident from equation (3.7) that the mutual coherence is much easier to compute than the RIP constant. If we normalize the columns of $\mathbf{A}$ and form the Gram matrix $\mathbf{G} = \mathbf{A}^H\mathbf{A}$, the mutual coherence is then simply equal to the maximum off-diagonal element of $\mathbf{G}$. Therefore, the mutual coherence is the largest absolute normalized inner product between different columns of $\mathbf{A}$ (Bruckstein et al., 2009). Because near orthogonal matrices have small correlations amongst their columns, matrices with small mutual coherence favor recovery.

## 3.5 Compressed sensing and randomized marine acquisition

Our focus in this section is on the design of source subsamplingschemes that favor recovery in combination with the selection of the appropriate sparsifying transform. To illustrate the importance of transform-domain sparsity and mutual coherence, we include sparse recovery by the curvelet transform and the 3D Fourier transform in our simulations. Note that this Fourier transform is a simple 3D transform and should not be confused with windowed Fourier transforms that are typically used in seismic data processing.

### 3.5.1 Randomized marine acquisition as a CS problem

Consider marine data organized in a seismic line with $N_s$ sources, $N_r$ receivers, and $N_t$ time samples. For simplicity, we assume that all sources see the same receivers, which makes our method applicable to marine acquisition with ocean-bottom cables. The seismic line can be reshaped into an $N$ dimensional vector $\mathbf{f}$, where $N = N_s N_r N_t$. It is well known that seismic data admit sparse representations by curvelets that capture “wavefront sets” efficiently (see e.g. Smith, 1998; Candès and Demanet, 2005; Candès et al., 2006a; Herrmann et al., 2008, and the references therein). Therefore, we wish to recover a sparse approximation $\tilde{\mathbf{f}}$ of the discretized wavefield $\mathbf{f}$ from measurements $\mathbf{b} = \mathbf{RMf}$.

Let $\mathbf{S}$ be a sparsifying operator that characterizes a transform domain of $\mathbf{f}$, such that $\mathbf{S} \in \mathbb{C}^{P\times N}$ with $P \geq N$. When $\mathbf{S}$ is an orthonormal basis, i.e. $P = N$ and $\mathbf{SS}^H = \mathbf{S}^H\mathbf{S} = \mathbf{I}$ where $\mathbf{I}$ is the identity matrix, the signal $\mathbf{f}$ admits a unique transform domain representation $\mathbf{x} = \mathbf{Sf}$. On the other hand, if $\mathbf{S}$ is a tight frame with $P \geq N$ and $\mathbf{S}^H\mathbf{S} = \mathbf{I}$, as in the case of the redundant curvelet transform (Candès et al., 2006a), then the expansion of $\mathbf{f}$ in $\mathbf{S}$ is not unique. The sparse approximation $\tilde{\mathbf{f}}$ is obtained by solving the inverse problem

$$\mathbf{A} := \mathbf{RMS}^H \tag{3.8}$$

with the basis pursuit sparsity-promoting program

$$\tilde{\mathbf{x}} = \arg\min_{\mathbf{x}} \|\mathbf{x}\|_1 \quad \text{subject to} \quad \mathbf{Ax} = \mathbf{b}, \tag{3.9}$$

yielding $\tilde{\mathbf{f}} = \mathbf{S}^H\tilde{\mathbf{x}}$. To solve this one-norm optimization problem, we use the SPG$\ell_1$ solver (Berg and Friedlander, 2008).

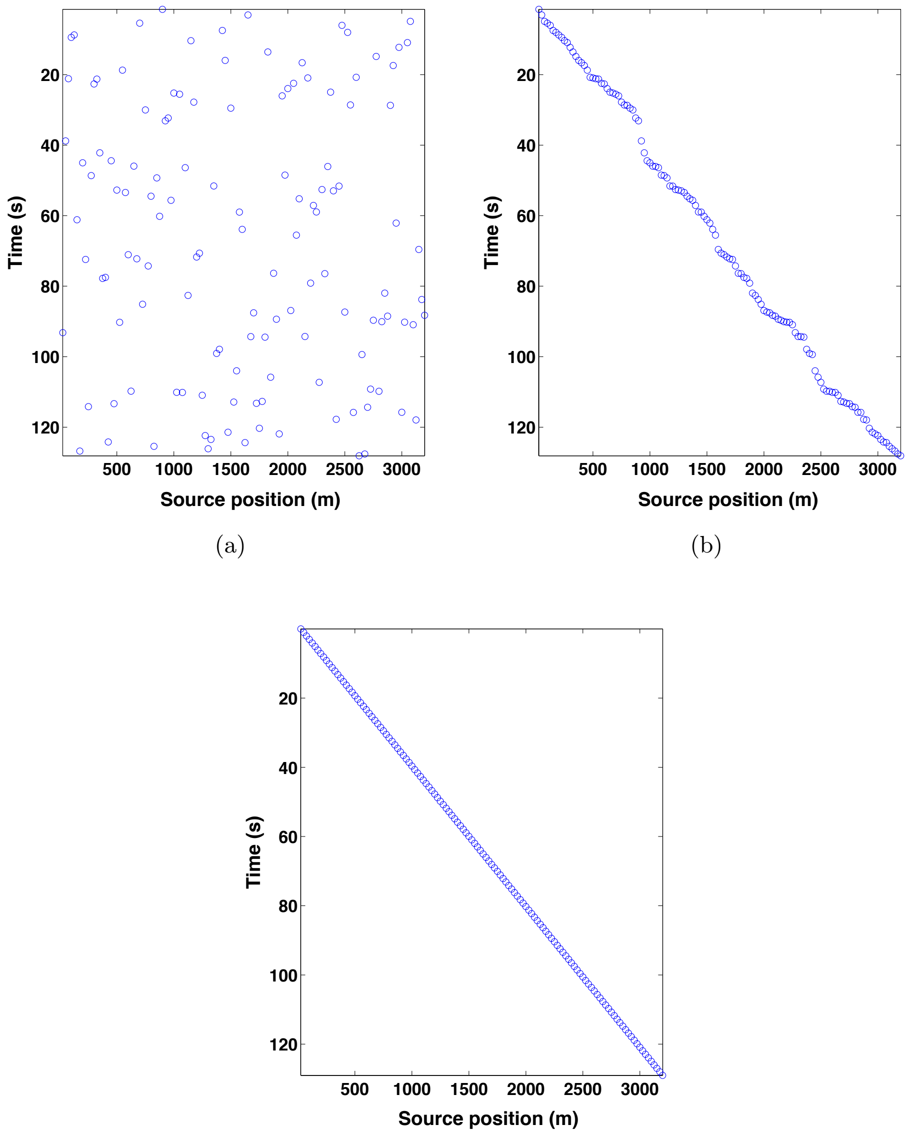
By solving a sparsity-promoting problem (Candès and Tao, 2006; Donoho, 2006; Herrmann et al., 2007; Mallat, 2009), it is possible to reconstruct high-resolution data volumes from the randomized samples at moderate sampling overhead compared to data volumes obtained after conventional compression (see e.g., Donoho et al. (1999b) for wavelet-based compression, and Herrmann (2010) for empirical oversampling rates for seismic problems). As in conventional compression, the recovery error is controllable, but in the case of CS this recovery error depends on the sampling ratio $\gamma = \frac{n}{N}$. This ratio is given by the number of compressive samples and the number of conventionally acquired samples. From a simultaneous-source marine seismic perspective, this is the ratio between the size of the continuous and simultaneous recordings and the size of the conventional sequential data. From the perspective of the proposed single-source randomly time-dithered marine acquisition scheme, this is the ratio between the size of the randomly overlapping sequential source recordings and the size of the conventional non-overlapping sequential data. Consequently, in both the simultaneous-source and the randomized acquisition scenarios, the survey time is reduced for a fixed number of shots. Conversely, the number of shots recorded can also be increased given the same amount of recording time as a conventional survey, which is useful for projects afflicted with poor shot coverage.

### 3.5.2 Designing the randomized operator

The design of the linear sampling operator $\mathbf{RM}$ is critical to the success of the recovery algorithm. $\mathbf{RM}$ may in some cases be separable and composed of an $n \times N$ restriction matrix $\mathbf{R}$ multiplying an $N \times N$ mixing matrix $\mathbf{M}$. This is not true in the case of simultaneous-source or the single source random time-dithered marine acquisition where, as we will illustrate, the sampling operator $\mathbf{RM}$ is nonseparable. In the simultaneous marine acquisition scenario, the classic sequential acquisition with a single air gun is replaced with continuous acquisition with multiple air guns firing at random times and at random locations that span the entire survey area. This "ideal" simultaneous-source sampling scheme is illustrated in Figure 3.1(a) where the circles indicate the firing times and locations of the multiple sources. Such simultaneous acquisition schemes require an air gun to be located at each source location throughout the survey, which is infeasible. Alternatively, it may be possible yet costly to send out several vessels with air guns that swarm over an ocean-bottom array.

We present a new alternative which requires a very small number of vessels (possibly one) that map the entire survey area while firing sequential shots at randomly time-dithered instances. In the random time-dithered acquisition scheme a single air gun or multiple air guns are fired sequentially with random lag intervals between shots. This random time-dithered marine acquisition scheme is illustrated in Figure 3.1(b) where, similar to the simultaneous source scheme, the firing times are still random but the source positions are sorted with respect to survey time. Remember that the ordered acquisition is still random by virtue of the random time shifts as opposed to the case of a periodic time-dithering scheme where we simply decrease the intershot time delays as depicted in Figure 3.1(c). In the remainder of this chapter, we use the CS criteria of the previous section to analyze the efficacy of this random time-dithered scheme combined with the appropriate transform $\mathbf{S}$ and sparse approximation algorithm in recovering the discretized wavefield $\mathbf{f}$ from measurements $\mathbf{b} = \mathbf{RMf}$.

First we develop the structure of the sampling operator $\mathbf{RM}$. Suppose as before we have $N_s$



(a)

(b)

(c)

**Figure 3.1:** Examples of (a) random dithering in source location and trigger times, (b) sequential locations and random time-triggers, and (c) periodic source firing triggers.

shots, $N_r$ receivers, and every shot decays after $N_t$ time samples. We first map the seismic line into a series of sequential shots $\mathbf{f}$ of total length $N = N_s \times N_t \times N_r$ and apply the sampling operator $\mathbf{RM}$ to reduce $\mathbf{f}$ to a single long "supershot" of length $n \ll N$ that consists of a superposition of $N_s$ impulsive shots. Again, to make the analysis more tractable, we ignore varying detector coverage by assuming a fixed receiver spread.

Since the subsampling/mixing is performed in the source-time domain, the resulting sampling operator is defined as follows

$$\mathbf{RM} := [\mathbf{I} \otimes \mathbf{T}], \tag{3.10}$$

where $\otimes$ is the Kronecker product, $\mathbf{I}$ is an $N_r \times N_r$ identity matrix, and $\mathbf{T}$ is a combined random shot selector and time shifting operator. This structure decouples the receiver axis from the source-time axis in the sampling operator allowing $\mathbf{T}$ to operate on the vectorized common receiver gathers. Taking the Kronecker product of $\mathbf{T}$ with $\mathbf{I}$ simply repeats the operation of $\mathbf{T}$ on every available receiver. The operator $\mathbf{T}$ turns the sequential-source recordings into continuous recordings with $N_s$ impulsive sources, and firing at time instances selected uniformly at random from $\{1, \ldots (N_s - 1)N_t\}$ discrete times[1]. Consequently, the operator $\mathbf{T}$ subsamples the $N_s N_t$ samples recorded at each receiver to $m \ll N_s N_t$ samples resulting in a total number $n = mN_r$ compressive samples[2].

In the marine case with air-gun sources, we can only work with binary $(0, 1)$-matrices because we have virtually no control over the source signature and energy output of air-gun arrays. In the conventional sequential acquisition scheme where no overlap exists between the source responses, the operator $\mathbf{T}$ would be a block diagonal matrix of $N_s$ blocks, each block being an $N_t \times N_t$ identity matrix. Each $N_t \times N_t$ identity matrix corresponds to the time taken for a source response to decay. This results in a large identity matrix of size $N_s N_t \times N_s N_t$. The simultaneous-source acquisition scheme destroys the block diagonal structure by placing the $N_t \times N_t$ identity matrices at random positions inside the matrix $\mathbf{T}$. An example of the corresponding operator is shown in Figure 3.2(a). In the case of random time-dithering, the $N_t \times N_t$ identity matrices are situated in an overlapped block diagonal structure as illustrated in Figure 3.2(b). The effect of random time-dithering on the structure of the operator $\mathbf{T}$ can be seen when we look at the operator in Figure 3.2(c) which corresponds to the periodic time-dithering scheme where we simply decrease the intershot time delays.

### 3.5.3 Assessment of the sampling operators

In this section, we limit the assessment to the 2D case due to the large dimensionality of the data in the 3D case. Consequently, the sampling operator $\mathbf{A}$ that we analyze here is constructed as $\mathbf{A} = \mathbf{T}\mathbf{S}^H$, where $\mathbf{S}$ is a 2D curvelet or Fourier transform. The operator $\mathbf{A}$ is then the lowest dimensional nonseparable component of the 3D Kronecker structure. These results constitute worst case bounds for the 3D case since the performance of the Kronecker structure is bounded by the worst case performance of its components (Duarte and Baraniuk, 2011).

#### Mutual-coherence based assessment

The randomized time-dithering operator results in a measurement matrix $\mathbf{A}$ that exhibits a smaller mutual coherence compared to the periodic time-dithering sampling operator (cf. Figures 3.2(b) and Figure 3.2(c)). Notice that in both cases we fixed the number of source experiments and the number of collected samples. Aside from the randomization of the sampling operator, the choice of

---

[1] It is possible to subsample the number of sources such that $n_s < N_s$ shots are selected uniformly at random from $\{1, \ldots N_s\}$ source locations. Such a configuration requires a modified operator $\mathbf{T}$ in which the number of $N_t \times N_t$ identity submatrices is equal to $n_s$.

[2] It is also possible to subsample the receiver axis or equivalently to randomize the locations of the available receivers in order to produce a higher resolution receiver grid. This is achieved by replacing the $N_r \times N_r$ identity matrix in $\mathbf{RM}$ by an $n_r \times N_r$ restriction matrix that selects the $n_r < N_r$ physical receiver coordinates from the high resolution grid. Consequently, the number of collected CS measurements would be $n = mn_r$.

sparsifying transform also determines the mutual coherence. To study this combined effect, let us consider the Gram matrix $\mathbf{G} = \mathbf{A}^H\mathbf{A}$ of deterministic versus random sampling matrices $\mathbf{A}$ using either curvelets or Fourier as sparsifying transforms.

Recall that the mutual coherence $\mu(\mathbf{A})$ is given by the largest off-diagonal element of $\mathbf{G}$. As Figures 3.3 and 3.4 indicate, there is big difference between the coherences for the deterministic versus the randomized acquisitions. In fact, the mutual coherence of the curvelet based operator in this example is 0.695 for random time-dithering compared with 0.835 for periodic time-dithering. Similarly, the mutual coherence of the Fourier based operator is 0.738 for random time-dithering compared with 0.768 for periodic time-dithering. For the Fourier and curvelet-based CS-sampling there is only a slight difference between the coherences despite significant differences in appearances of the Gram matrices. Therefore, the mutual coherence is a crude measure, which is confirmed in our experimental section.

### RIP-based assessment

While the calculation of the mutual coherence is straight forward, it can only give us a pessimistic upper bound on the recoverable sparsity $k$ of a signal (Bruckstein et al., 2009). A better bound (i.e. a guarantee for larger $k$) can be achieved by evaluating the restricted isometry property (RIP) of $\mathbf{A}$. Bounding this RIP constant allows us to guarantee recovery of less sparse (larger $k$) signals than what is guaranteed by the mutual coherence. The RIP constant $\delta_k$ of $\mathbf{A}$ is evaluated for all submatrices of $\mathbf{A}$ of size $n \times k$. Let $\Lambda$ be a set of column indices of $\mathbf{A}$ of size $k$. For any matrix $\mathbf{A}_\Lambda$ the following property holds

$$\sigma_{\min}^2\|\mathbf{u}\|_2^2 \leq \|\mathbf{A}_\Lambda\mathbf{u}\|_2^2 \leq \sigma_{\max}^2\|\mathbf{u}\|_2^2,$$

where $\sigma_{\min}$ and $\sigma_{\max}$ are the smallest and largest singular values of the matrix $\mathbf{A}_\Lambda$, respectively. The RIP constant $\delta_k$ is the smallest upper bound on $\hat{\delta}_\Lambda := \max\{1-\sigma_{\min}, \sigma_{\max}-1\}$, i.e. $\delta_k = \max\limits_{\Lambda} \hat{\delta}_\Lambda$. Consequently, if we can show that for all sets $\Lambda \in \{1, \ldots P\}$, $\hat{\delta}_\Lambda < 1$ then there exists some solver which can recover any $k$-sparse signal. Moreover, if $\hat{\delta}_\Lambda < \sqrt{2} - 1$ or $\hat{\delta}_\Lambda < \frac{a-1}{a+1}$ for some integer $a > 1$, then with high probability we can guarantee that the BP program (4.2) can recover any sparse signal with sparsity less than or equal to $k/2$ or $k/a$, respectively.

Unfortunately, there are $\binom{P}{k}$ combinations of the $\mathbf{A}_\Lambda$ submatrices in $\mathbf{A}$, which makes evaluating $\hat{\delta}_\Lambda$ computationally infeasible in realistic settings. To overcome this difficulty, and since the transform coefficients of seismic images are often not strictly sparse, we first identify the appropriate $k$ as the smallest number of transform coefficients that capture say 90% of the signal energy. This allows us to bound the recovery error in terms of the best $k$-term approximation of the signal. We are unaware of theoretical results that lead to a bound on $\hat{\delta}_\Lambda$ for our particular choice of $\mathbf{A}$. To overcome this, we estimate this quantity by Monte-Carlo sampling over different realizations of the sampling matrix and different realizations of the support. In our simulations, we generate 1000 realizations of the random time-dithering sampling matrix. For each of these matrices, we evaluate $\hat{\delta}_\Lambda$ for 100 random realizations of the set $\Lambda$. The RIP constant $\delta_k$ is estimated as the maximum of the computed values of $\hat{\delta}_\Lambda$.

We plot the results of these simulations in Figure 3.5, which shows the histograms of $\hat{\delta}_\Lambda$ for the curvelet and Fourier transforms. These simulations show that the $\hat{\delta}_\Lambda$ for the curvelet transform are less than one, which means that this matrix has RIP, while the matrix based on Fourier may not have RIP for certain realizations of $\Lambda$. As a consequence, we can expect higher fidelity for curvelet-based recovery. Notice that in the simulations for the curvelet case, the estimate for $\hat{\delta}_\Lambda \approx 0.76$ is larger than the theoretical bound of $\sqrt{2}-1$ which guarantees stable recovery using BP with respect to the best $k/2$-term approximation $\mathbf{x}$. This is mainly due to the choice of $k$ we used in our simulations. By choosing a smaller value for $k$, it would be possible to achieve the $\sqrt{2}-1$ mark at the expense of increasing the best $k$-term approximation error. For example, the condition $\delta_{9s} < 7/9$ guarantees stable recovery with respect to the best $s$-term approximation of the signal, where $8s = k$ and $k = |\Lambda|$ is the same as in the simulations above.

On the other hand, the RIP constant of the Fourier-based operator may exceed one. Therefore, it is not possible to find an $s$ for which the RIP-based recovery conditions hold. We believe that this observation reflects the poorer recovery results of the Fourier-based operator compared to the curvelet-based operator as will be shown in the experimental results section.

### 3.5.4 Economic considerations

Aside from mathematical factors, such as the mutual coherence that determines the recovery quality, there are also economical factors to consider. For this purpose, Berkhout (2008) proposed two performance indicators, which quantify the cost savings associated with simultaneous and continuous acquisition. The first measure compares the number of sources involved in conventional and simultaneous acquisition and is expressed in terms of the source-density ratio

$$\mathrm{SDR} = \frac{\text{number of sources in the simultaneous survey}}{\text{number of sources in the conventional sequential survey}}. \tag{3.11}$$

In the marine data acquisition setting, the $\mathrm{SDR} = \frac{n_s}{N_s}$. Aside from the number of sources, the cost of acquisition is also determined by survey-time ratio

$$\mathrm{STR} = \frac{\text{time of the conventional sequential survey}}{\text{time of the continuous and simultaneous recording}}. \tag{3.12}$$

The survey time ratio is therefore given by $\mathrm{STR} = \frac{N_s N_t}{m}$ which is equal to the aspect ratio of the operator $\mathbf{T}$. The overall economic performance is measured by the product of these two ratios.

## 3.6 Experimental results

We illustrate the effectiveness of our simultaneous source acquisition approach by studying the performance of the three sampling schemes; simultaneous-source acquisition, random time-dithering, and periodic time-dithering, on a seismic line from the Gulf of Suez (Figure 3.6 shows a common-shot gather). The fully sampled sequential data is composed of $N_s = 128$ shots, $N_r = 128$ receivers and $N_t = 512$ time samples with 12.5 m source-receiver sampling interval. Prestack data from sequential sources is recovered using $\ell_1$ minimization with 3D curvelets as the sparsifying transform. For comparison, we perform sparse recovery with a 3D Fourier transform and with the more rudimentary median filtering, which can also be used to suppress the crosstalk. Note that this Fourier transform is a simple 3D transform and not a windowed Fourier transform that is typically used in seismic data processing. We also perform linear recovery using the adjoint of the sampling operator followed by 2D median filtering in the offset domain for additional comparison.

We evaluate the recovery performance in terms of the *signal-to-noise ratio* (S/N) which is computed as follows for a signal $\mathbf{x}$ and its estimate $\tilde{\mathbf{x}}$:

$$\mathrm{S/N}(\mathbf{x}, \tilde{\mathbf{x}}) = -20 \log_{10} \frac{\|\mathbf{x} - \tilde{\mathbf{x}}\|_2}{\|\mathbf{x}\|_2}. \tag{3.13}$$

### 3.6.1 Simultaneous-source acquisition

We simulate simultaneous-source marine acquisition by randomly selecting 128 shots from the total survey time $t = N_s \times N_t$ with a subsampling ratio $\gamma = \frac{m}{N_s \times N_t} = \frac{1}{\mathrm{STR}} = 0.5$. The subsampling is performed only in time so that the length of the "supershot" is half the length of the conventional survey time of sequential-source data. Figure 3.7(a) represents the "supershot" plotted as conventional survey by applying the sampling operator $RM$ to sequential-source data. Notice that this type of "marine" acquisition is physically realizable only with a limited number of simultaneous sources, although truly random positioning of sources may still prove impractical depending on the manoeuvrability of source vessels.

To exploit continuity of the wavefield along all three coordinate axes, we use the 3D curvelet transform Ying et al. (2005). Figures 3.7(b) and 3.7(c) show the recovery and residual results, respectively. The recovered data volume has an S/N of 10.5 dB and was obtained with 200 iterations of solving the BP problem using SPG$\ell_1$.

### 3.6.2 Random time-dithering

To overcome the limitation in the number of simultaneous sources required by the "ideal" simultaneous-source approach, we propose the random time-dithering scheme. Under this scheme, we allow all 128 shots to be fired sequentially with adjacent shots firing before the previous shot fully decays. We impose a random overlap between the shot records created by a random time lag between the firing of each shot. Therefore, the start time of each shot is chosen uniformly at random between the starting time of the previous shot and the time by which the previous shot record decays.

In our simulations, we apply a sampling operator with subsampling ratio $\gamma = 0.5$. A section of the "supershot" obtained by random time-dithering is shown in Figure 3.8(a). Using the 3D curvelet transform, a recovery of 8.06 dB is achieved (Figure 3.8(b)). Figure 3.8(c) shows the corresponding residual plot. Figure 3.9 summarizes the results for recovery based on the 3D Fourier transform. The recovered data volume has an S/N = 6.83 dB, which agrees with our predictions for RIP constants estimated in the previous section.

We also demonstrate the effectiveness of sparse recovery compared with linear recovery using the adjoint of the sampling operator $\mathbf{RM}$ followed by 2D median filtering in the midpoint-offset domain. The recovery results are shown in Figure 3.10. The resulting S/N, 3.92 dB, is considerably lower than the S/Ns achieved by sparse recovery.

### 3.6.3 Periodic time-dithering

The importance of the "randomness" in time-dithering becomes evident when we compare the recovery of this sampling operator with that of a periodic time-dithering operator. Under the same subsampling conditions and sparsifying transform, the periodic time dither operator can only achieve an S/N = 4.80 dB. Figure 3.11 shows the periodic time-dithered "supershot", the recovered data volume and the corresponding residual. This poor performance is consistent with predictions of CS that require randomness in the design of sampling matrices. Furthermore, comparing sparse recovery with linear recovery of 1.26 dB (Figure 3.12), demonstrates the effectiveness of the former.

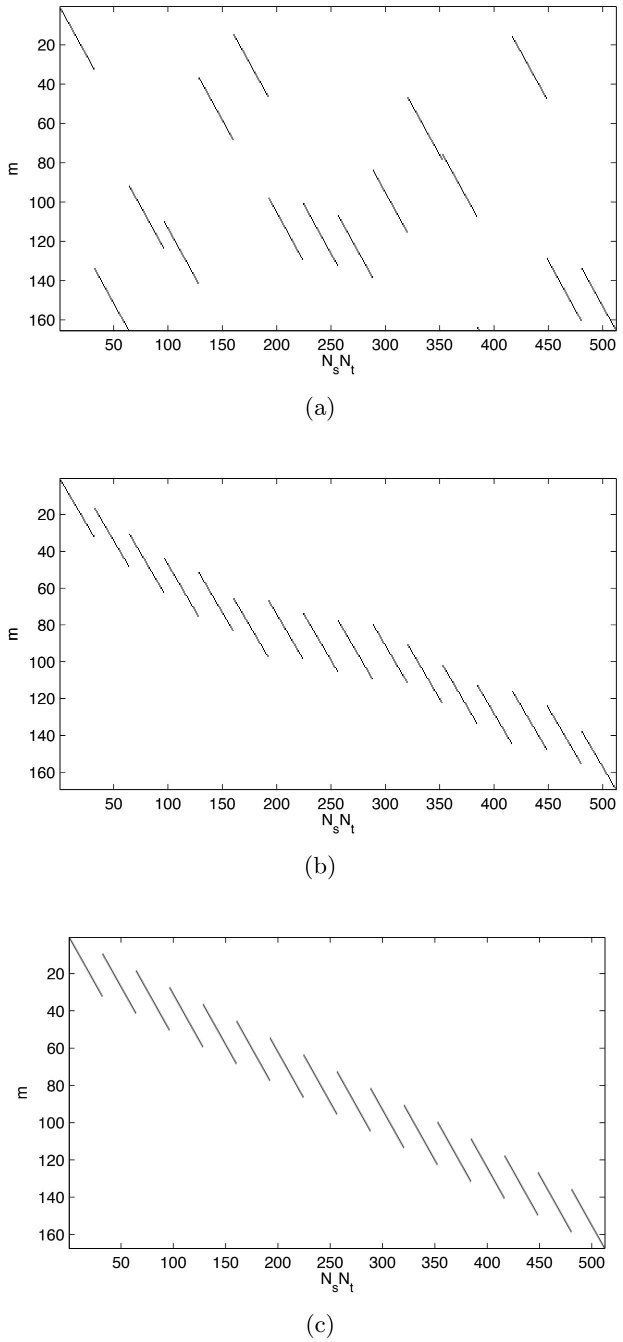
Finally, we illustrate the recovery for five subsampling ratios ($\gamma = 0.75, 0.50, 0.33, 0.25, 0.10$) for each of the schemes described above. Table 3.1 summarizes the S/Ns for the three sampling schemes based on the 3D curvelet and 3D Fourier transforms.

| **Subsample ratio** | **Simultaneous acquisition** | | **Random time-dithering** | | **Periodic time-dithering** | |
|---|---|---|---|---|---|---|
| 1/STR | Curvelet | Fourier | Curvelet | Fourier | Curvelet | Fourier |
| 0.75 | 13.0 | 10.2 | 11.2 | 9.44 | 6.93 | 4.93 |
| 0.50 | 10.5 | 7.06 | 8.06 | 6.83 | 4.80 | 2.42 |
| 0.33 | 8.31 | 4.50 | 5.33 | 4.10 | 7.32 | 1.37 |
| 0.25 | 6.55 | 2.93 | 4.35 | 2.88 | 2.85 | 0.89 |
| 0.10 | 2.82 | 0.27 | 1.14 | 0.20 | 1.60 | 0.19 |

**Table 3.1:** Summary of recovery results (S/N in dB) based on the 3D curvelet and the 3D Fourier transforms for the three sampling schemes.

## 3.7 Conclusions

Recovering single-source prestack data volumes from simultaneously acquired marine data essentially involves removing noise-like crosstalk from coherent seismic responses. Many authors have noticed the important role of sparsity-based recovery for this problem, but few have thoroughly investigated the underlying interaction between acquisition design and reconstruction fidelity, especially in the marine setting. In contrast, we identify simultaneous marine acquisition as a linear subsampling system, which we subsequently analyze by using metrics from Compressive Sensing. We also propose a randomized time-dithering scheme which can match with a single source the performance of simultaneous-source acquisition. With the introduction of methods to calculate the mutual coherence and restricted isometry constants we are able to assert the importance of randomness in the acquisition system in combination with the appropriate choice for the sparsifying transform in the reconstruction. By comparing reconstructions on a real seismic marine line with different sparsifying transforms and sampled with different synthetic acquisitions, we quantitatively verified that more randomness in the acquisition system and more compressible transforms improve the mutual coherence and restricted isometry constants, which predict a higher reconstruction quality. As such this work represents a first step towards a comprehensive theory that predicts the reconstruction quality as a function of the type of acquisition.



(a)

(b)

(c)

**Figure 3.2:** Example of (a) “ideal” simultaneous-source operator defined by a Bernouilli matrix, (b) operator that corresponds to the more realizable Marine acquisition by the random time-dithering, and (c) sampling operator with periodic time-dithering.

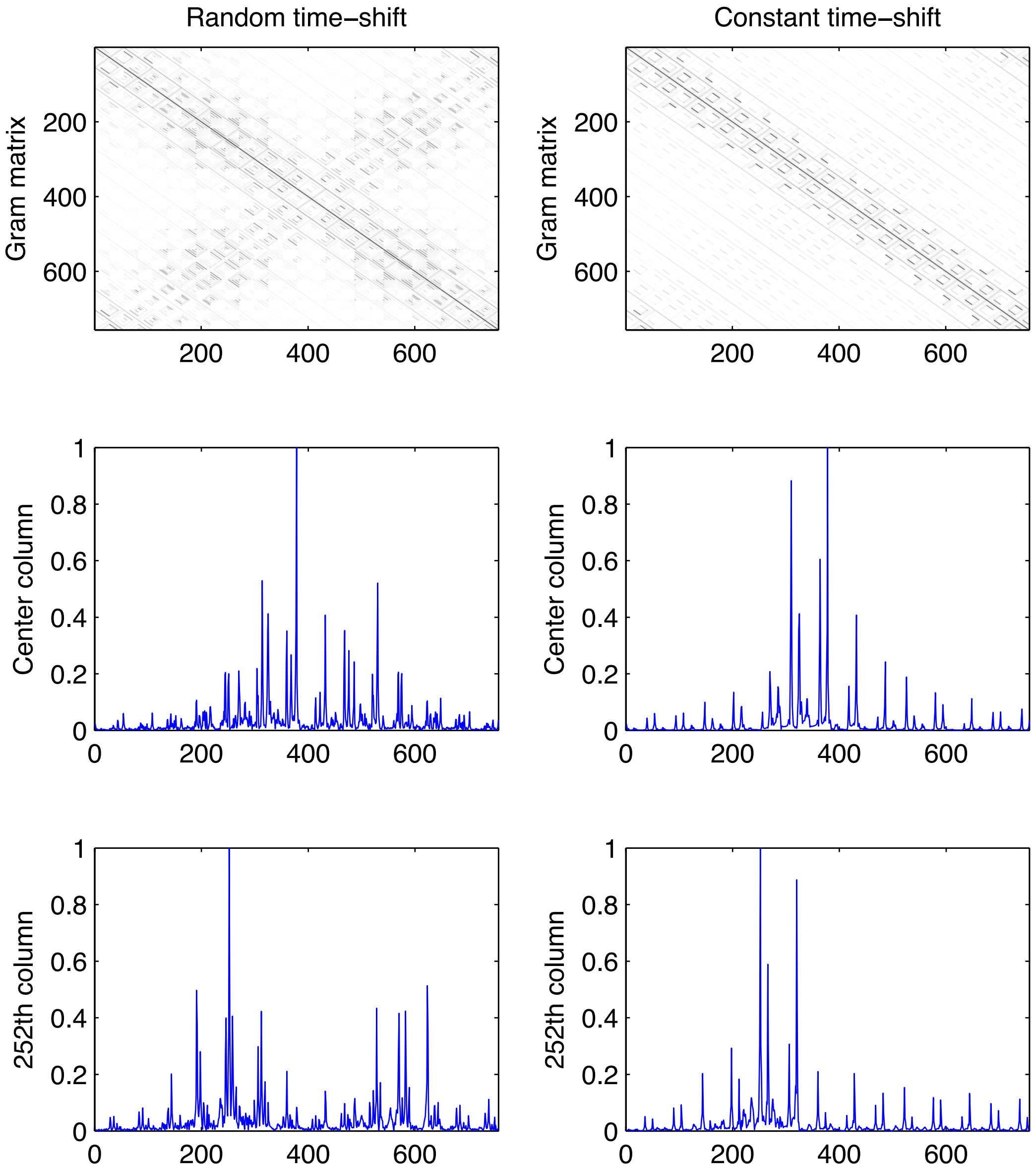


**Figure 3.3:** Gram matrices of example random time-dithering and constant time-dithering operators, top row, with $N_s = 10$ and $N_t = 40$ coupled with a **curvelet** transform. The resulting mutual coherence is 0.695 for random time-dithering compared with 0.835 for periodic time-dithering. The center plots show column the center column of the Gram matrices. The bottom row shows column 252 (one third) of the Gram matrices.

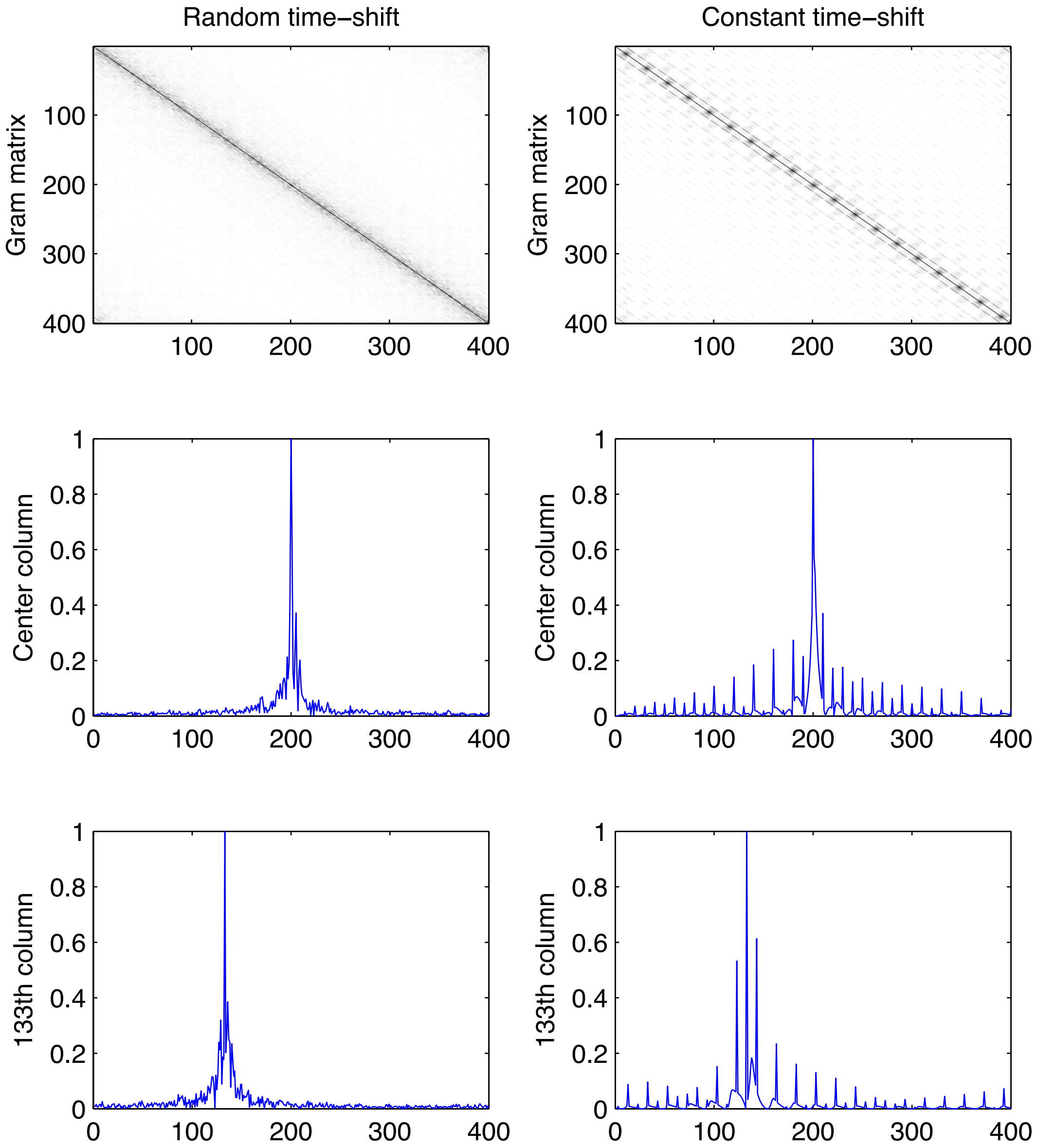
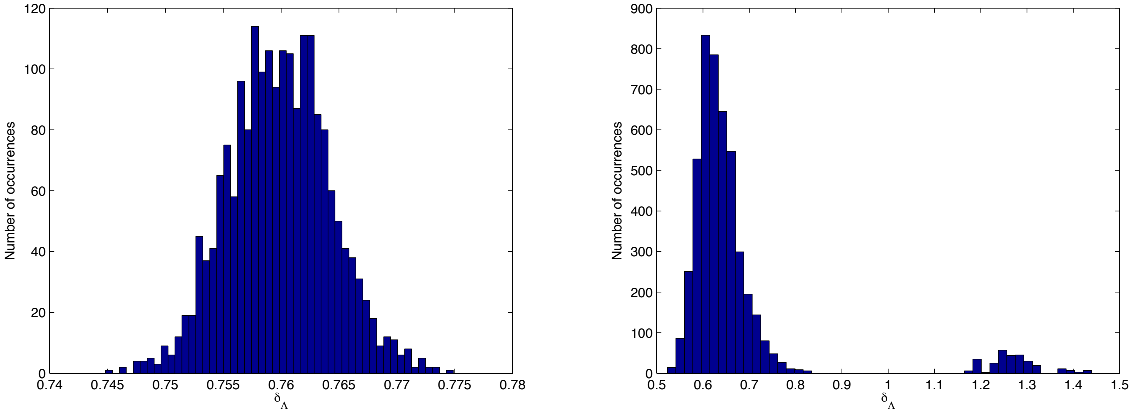


**Figure 3.4:** Gram matrices of example random time-dithering and constant time-dithering operators, top row, with $N_s = 10$ and $N_t = 40$ coupled with a **Fourier** transform. The resulting mutual coherence is 0.738 for random time-dithering compared with 0.768 for periodic time-dithering. The center plots show column the center column of the Gram matrices. The bottom row shows column 133 (one third) of the Gram matrices.



(a) (b)

**Figure 3.5:** Comparison between the histograms of $\hat{\delta}_\Lambda$ from 1000 realizations of $\mathbf{A}_\Lambda$, the random time-dithering sampling matrices $\mathbf{A} = \mathbf{RMS}^H$ restricted to a set $\Lambda$ of size $k$, the size support of the largest transform coefficients of a real (Gulf of Suez) seismic image. The transform $\mathbf{S}$ is (a) the curvelet transform and (b) the nonlocalized 2D Fourier transform. The histograms show that randomized time-shifting coupled with the curvelet transform has better behaved RIP constant ($\hat{\delta}_\Lambda = \max\{1 - \sigma_{\min}, \sigma_{\max} - 1\} < 1$) and therefore promotes better recovery.

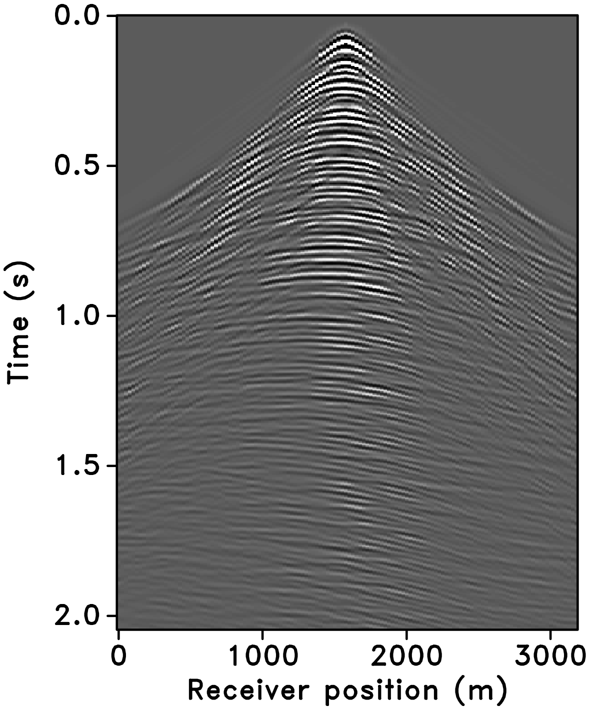
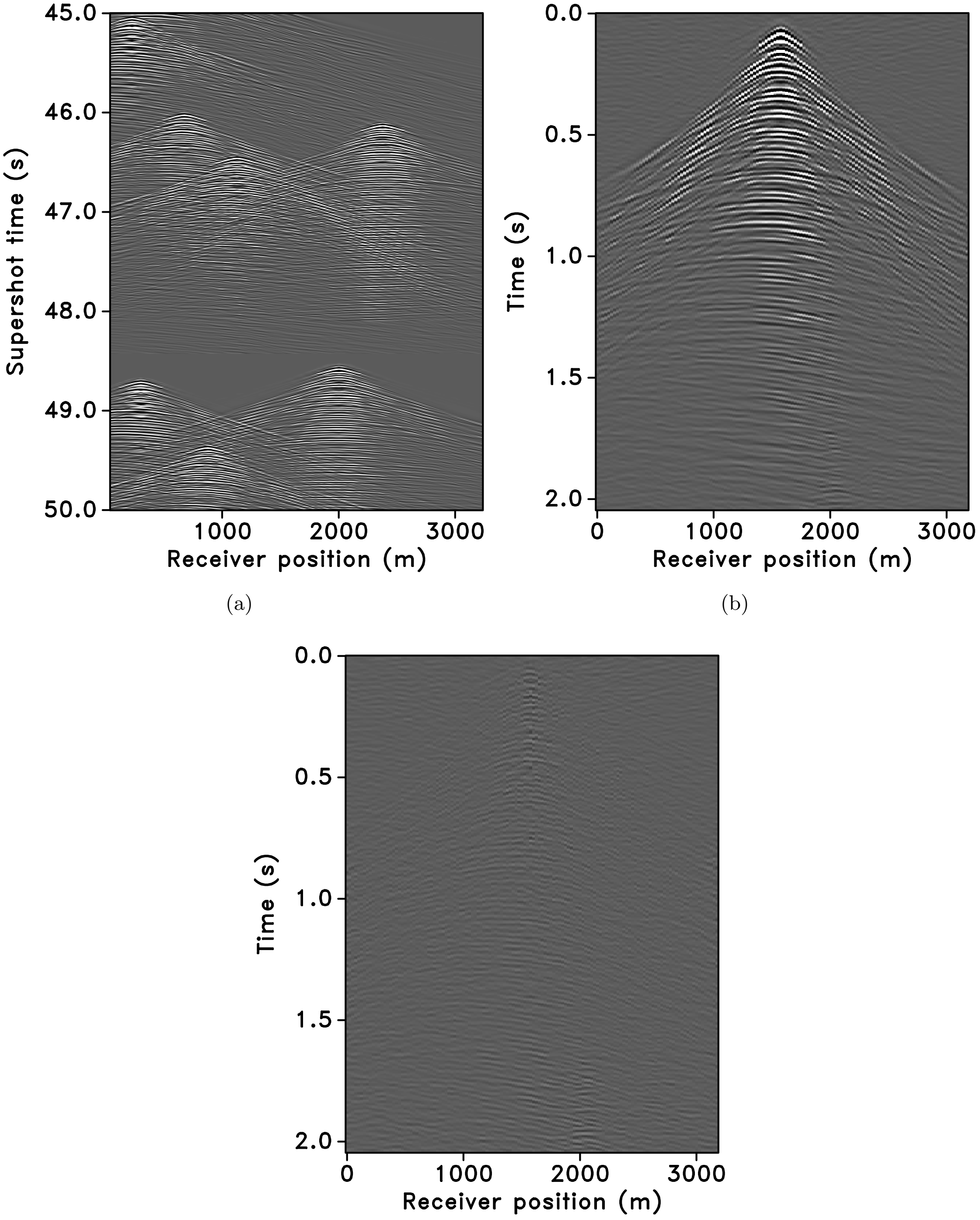


**Figure 3.6:** A common-shot gather from Gulf of Suez data set.

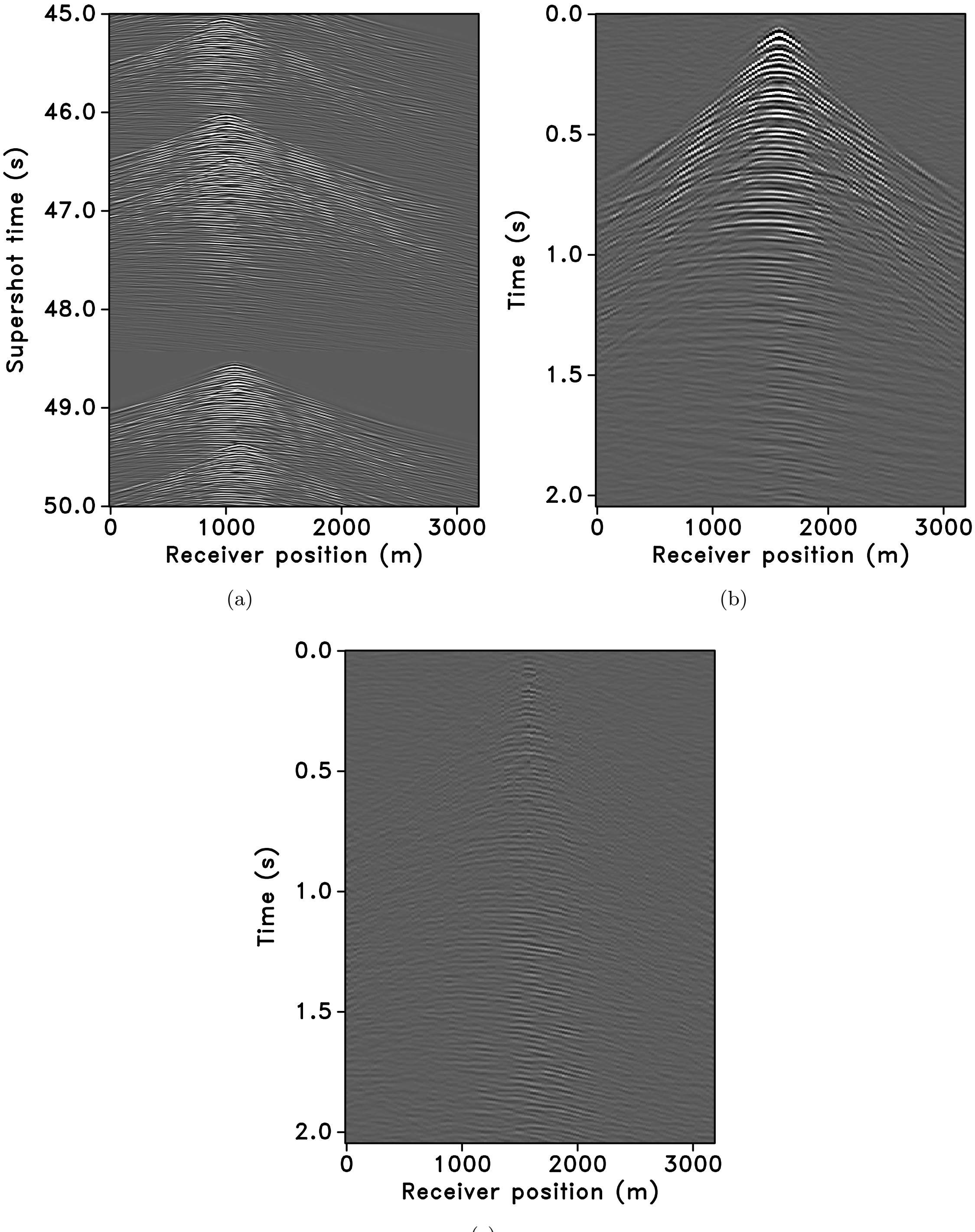


(a)

(b)

(c)

**Figure 3.7:** (a) Simultaneous-source marine data ($\gamma = 0.5$) shown as a section between 45 to 50 seconds of the "supershot". (b) Recovery from simultaneous 'marine' data (S/N = 10.5 dB). (c) The corresponding residual plot.

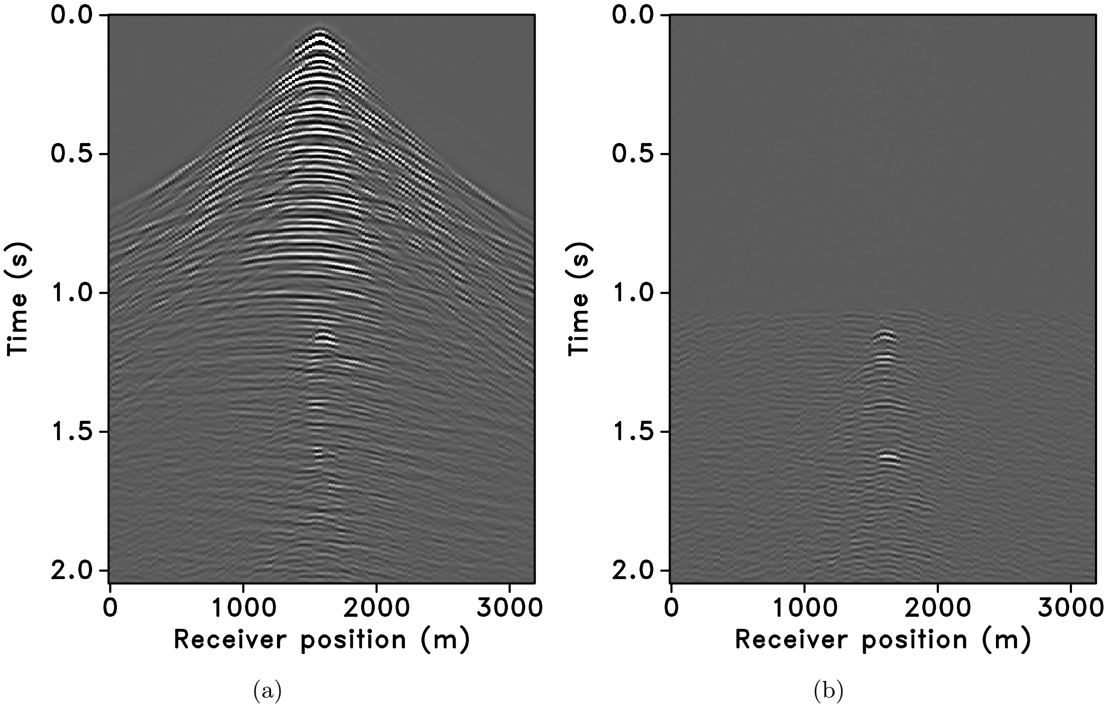


(a)

(b)

(c)

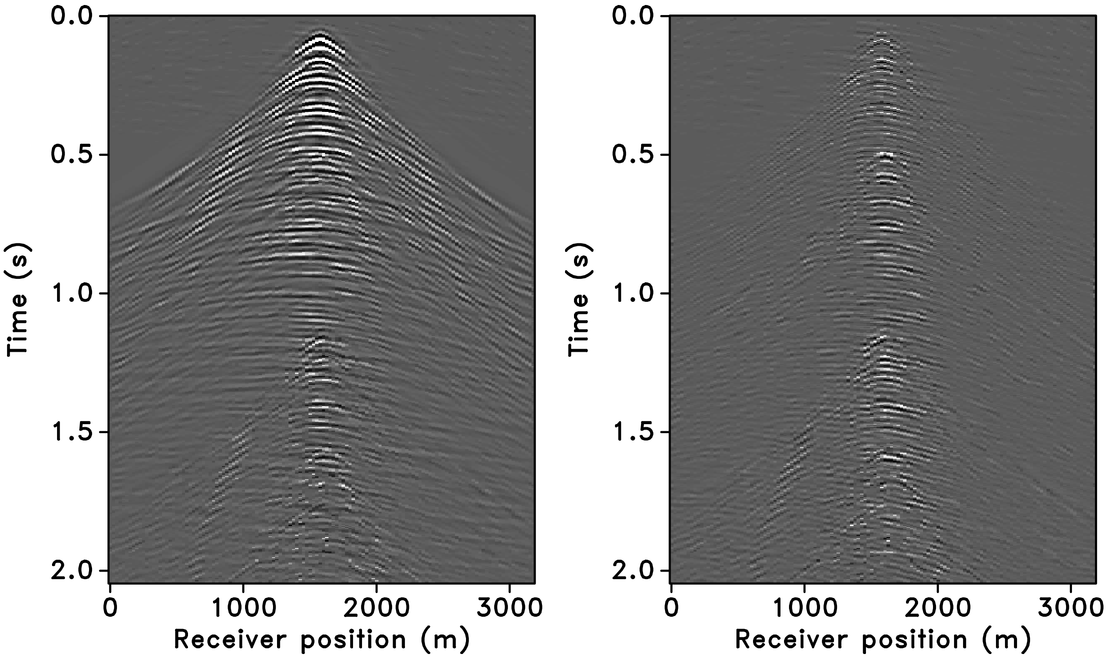
**Figure 3.8:** (a) Random time-dithered "marine" data ($\gamma = 0.5$) shown as a section between 45 and 50 seconds of the "supershot". (b) Sparse recovery with curvelet transform and S/N = 8.06dB. (c) The corresponding residual plot.



(a)

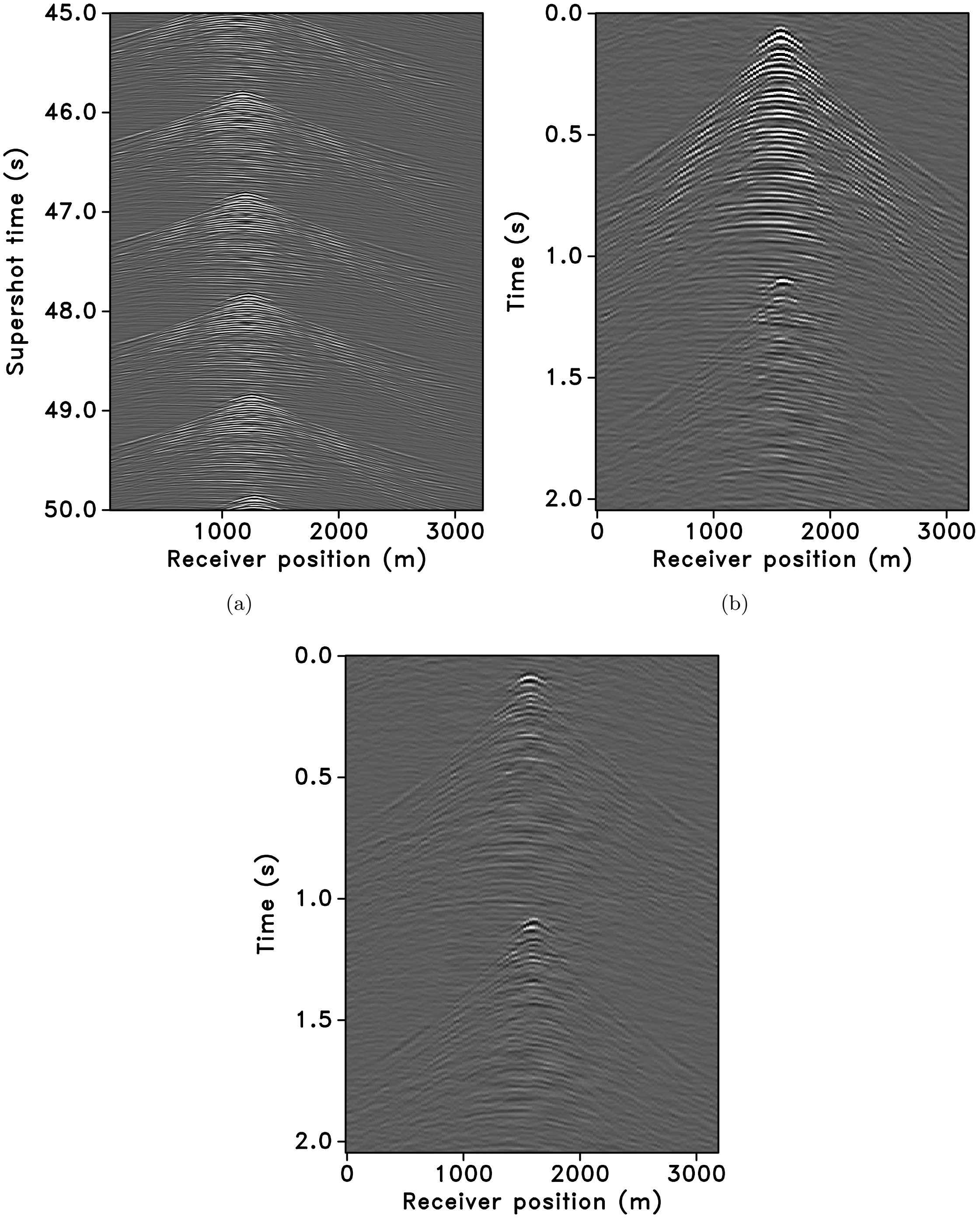
(b)

**Figure 3.9:** (a) Sparse recovery with 3D Fourier transform from the same data shown in Figure 3.8(a), S/N = 6.83dB. (b) The corresponding residual plot.



(a) (b)

**Figure 3.10:** (a) Data recovered by applying adjoint of the sampling operator **RM** and 2D median filtering, from the same data shown in Figure 3.8(a), with S/N = 3.92dB. (b) The corresponding residual plot.

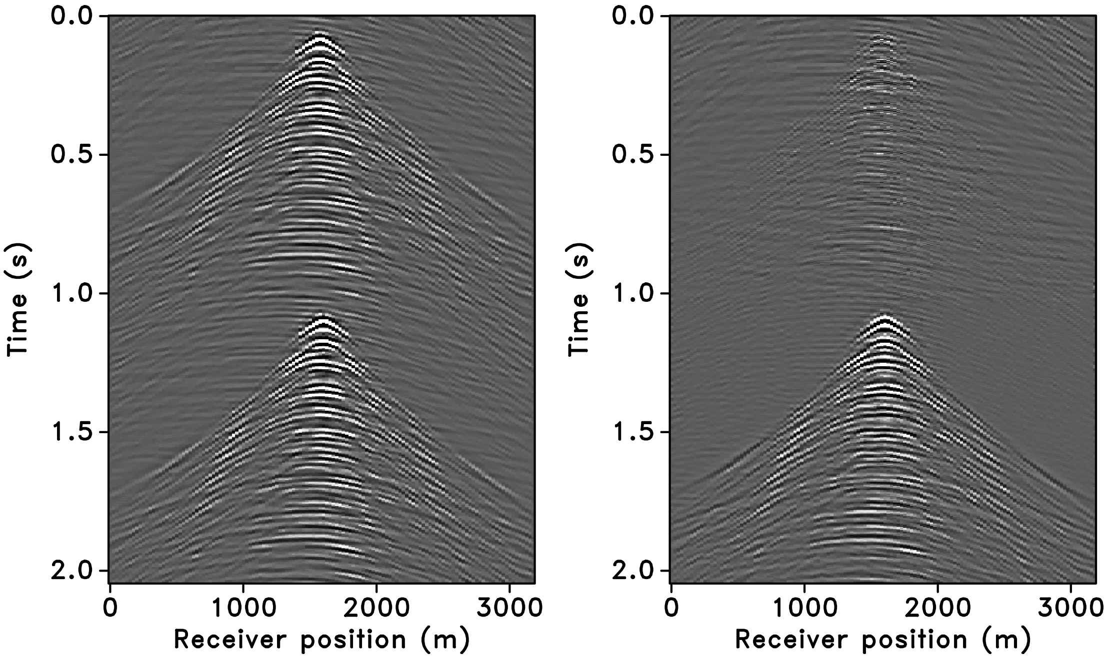


(a)

(b)

(c)

**Figure 3.11:** (a) Periodic time-dithered "marine" data ($\gamma = 0.5$) shown as a section between 45 and 50 seconds of the "supershot". (b) Sparse recovery with curvelet transform and S/N = 4.80dB. (c) The corresponding residual plot.



(a) (b)

**Figure 3.12:** (a) Data recovered by applying adjoint of the sampling operator **RM** and 2D median filtering, from the same data shown in Figure 3.11(a), with S/N = 1.26dB. (b) The corresponding residual plot.

# Chapter 4

# Simultaneous-source time-jittered marine acquisition

## 4.1 Summary

Adapting ideas from the field of compressive sensing, we show how simultaneous- or blended-source acquisition can be setup as a compressive sensing problem. This helps us to design a pragmatic simultaneous-source *time-jittered marine* acquisition scheme where multiple source vessels sail across an ocean-bottom array firing air guns at jittered source locations and instances in time, resulting in better spatial sampling, and speedup acquisition. Furthermore, we can significantly impact the reconstruction quality of conventional seismic data from simultaneous jittered data and demonstrate successful recovery by sparsity promotion. In contrast to random (sub)sampling, acquisition via jittered (sub)sampling helps in controlling the maximum gap size, which is a practical requirement of wavefield reconstruction with localized sparsifying transforms. Results are illustrated with simulations of *time-jittered marine* acquisition on a seismic line, which translates to jittered source locations for a given speed of the source vessel.

## 4.2 Introduction

Constrained by the Nyquist sampling rate, the increasing sizes of the conventionally acquired marine seismic data volumes pose a fundamental shortcoming in the traditional sampling paradigm and make large area acquisition particularly expensive. Physical constraints on the speed of a source vessel during acquisition, on the minimal time interval between adjacent shots (to avoid overlaps), and on the minimal spatial shot sampling further aggravate the acquisition related costs. Several works in the seismic acquisition literature have explored the concept of simultaneous- or blended-source activation to account for these situations (Beasley et al., 1998; de Kok and Gillespie, 2002; Beasley, 2008; Berkhout, 2008; Hampson et al., 2008; Moldoveanu and Fealy, 2010).

For simultaneous-source acquisition, the challenge is to estimate interference-free shot gathers (*source separation* or *deblending*) and recover small subtle late reflections that can be overlaid by interfering seismic responses from other shots. Stefani et al. (2007), Moore et al. (2008) and

---

A version of this chapter has been published in SEG Technical Program Expanded Abstracts, 2013, vol. 32, pp. 1–6.

Akerberg et al. (2008) have observed that the interferences in simultaneous (or blended) data will appear noise-like in specific gather domains such as common offset and common receiver, turning the separation into a typical (random) noise removal procedure. Application to land acquisition is reported in Bagaini and Ji (2010). Subsequent processing techniques vary from vector-median filters (Huo et al., 2009) to inversion-type algorithms (Moore, 2010; Abma et al., 2010; Mahdad et al., 2011). The former are mostly "processing" techniques where the interfering energy (i.e., source crosstalk) is removed and not mapped back to coherent energy, at least not in a single step alone, while the latter (inversion-type algorithms) are designed to take advantage of sparse representations of coherent seismic signals, which is advantageous because they exploit inherent structure in seismic data. In this chapter, we show that the challenge of source separation can be effectively addressed through a combination of tailored simultaneous-source acquisition design and curvelet-based sparsity-promoting recovery, where we map noise-like or incoherent source crosstalk to coherent seismic responses.

Recently, compressive sensing (CS, Donoho, 2006; Candès and Tao, 2006) has emerged as an alternate sampling paradigm in which randomized sub-Nyquist sampling is used to capture the structure of the data with the assumption that it is sparse or compressible in some transform domain. Seismic data consists of wavefronts that exhibit structure across different scales and amongst different directions. With the appropriate data transformation, we can capture this structure by a small number of significant transform coefficients resulting in a sparse representation of data. In our work, we rely on the CS literature to analyze a physically realizable *time-jittered marine* acquisition scheme, and recover the canonical sequential single-source (interference-free/nonsimultaneous) data by solving a sparsity-promoting problem (Mansour et al., 2012a; Wason and Herrmann, 2012). Hence, we develop a relation between simultaneous-source acquisition design and (curvelet-based) sparse recovery, within the CS framework.

## 4.3 Compressive sensing

Compressive sensing is a signal processing technique that allows a signal to be sampled at sub-Nyquist rate and reconstructs it (from relatively few measurements) by utilizing the prior knowledge that the signal is sparse or compressible in some transform domain, i.e., if only a small number $k$ of the transform coefficients are nonzero or if the signal can be well approximated by the $k$ largest-in-magnitude transform coefficients. For high resolution data represented by the $N$-dimensional vector $\mathbf{f}_0 \in \mathbb{R}^N$, which admits a sparse representation $\mathbf{x}_0 \in \mathbb{C}^P$ in some transform domain characterized by the operator $\mathbf{S} \in \mathbb{C}^{P\times N}$ with $P \geq N$, the sparse recovery problem involves solving an underdetermined system of equations

$$\mathbf{b} = \mathbf{A}\mathbf{x}_0, \tag{4.1}$$

where $\mathbf{b} \in \mathbb{C}^n$, $n \ll N \leq P$, represents the compressively sampled data of $n$ measurements, and $\mathbf{A} \in \mathbb{C}^{n\times P}$ represents the measurement matrix. We denote by $\mathbf{x}_0$ a sparse synthesis coefficient vector of $\mathbf{f}_0$. When $\mathbf{x}_0$ is strictly sparse (i.e., only $k < n$ nonzero entries in $\mathbf{x}_0$), sparsity-promoting recovery can be achieved by solving the $\ell_0$ minimization problem, which is a combinatorial problem and quickly becomes intractable as the dimension increases. Instead, the basis pursuit (BP) convex optimization problem

$$\tilde{\mathbf{x}} = \arg\min_{\mathbf{x}\in\mathbb{C}^P} \|\mathbf{x}\|_1 \quad \text{subject to} \quad \mathbf{b} = \mathbf{A}\mathbf{x}, \tag{4.2}$$

can be used to recover $\tilde{\mathbf{x}}$, where $\tilde{\mathbf{x}}$ represents the estimate of $\mathbf{x}_0$, and the $\ell_1$ norm $\|\mathbf{x}\|_1$ is the sum of absolute values of the elements of a vector $\mathbf{x}$. The BP problem typically finds a sparse or (under some conditions) the sparsest solution that explains the measurements exactly. The matrix $\mathbf{A}$ can be composed of the product of a restriction operator (subsampling matrix) $\mathbf{R} \in \mathbb{R}^{n \times N}$, an $N \times N$ mixing matrix $\mathbf{M}$, and the sparsifying operator $\mathbf{S}$ such that $\mathbf{A} := \mathbf{RMS}^H$, here $^H$ denotes the Hermitian transpose. Consequently, the measurements $\mathbf{b}$ are given by $\mathbf{b} = \mathbf{A}\mathbf{x}_0 = \mathbf{RM}\mathbf{f}_0$. A seismic line with $N_s$ sources, $N_r$ receivers, and $N_t$ time samples can be reshaped into an $N$ dimensional vector $\mathbf{f}$, where $N = N_s \times N_r \times N_t$. For simplicity, we assume that all sources see the same receivers, which makes our method applicable to marine acquisition with ocean-bottom cables or nodes (OBC or OBN). We wish to recover a sparse approximation $\tilde{\mathbf{f}}$ of the discretized wavefield $\mathbf{f}$ from measurements $\mathbf{b} = \mathbf{RMf}$ (jittered data). This is done by solving the BP sparsity-promoting program (Equation 4.2), using the SPG$\ell_1$ solver (Berg and Friedlander, 2008), yielding $\tilde{\mathbf{f}} = \mathbf{S}^H\tilde{\mathbf{x}}$.

## 4.4 Time-jittered marine acquisition

The success of CS hinges on randomization of the acquisition, since random subsampling renders coherent aliases (e.g., interferences due to overlapping shot records in simultaneous-source acquisition) into harmless incoherent random noise, effectively turning the interpolation problem, which is also a source separation problem in our case, into a simple denoising problem (Hennenfent and Herrmann, 2008). Given limited control over the source signature of the air guns and their recharge time between shots (typically, a minimal time interval of 10.0 s is required), the only way to invoke randomness is to work with sources that fire at random times that map to random shot locations for a given speed of the source vessel. Unfortunately, random (sub)sampling does not provide control on the maximum gap size between adjacent measurements (Figure 4.1), which is a practical requirement of wavefield reconstruction with localized sparsifying transforms such as curvelets. Jittered (sub)sampling, on the other hand, shares the benefits of random sampling and offers control on the maximum gap size (Figure 4.1) (Hennenfent and Herrmann, 2008). Since we are still on the grid, this is a case of discrete jittering. A jittering parameter, dictated by the type of acquisition and parameters such as the minimum distance (and/or minimum recharge time for the air guns) required between adjacent shots, relates to how close and how far the jittered sampling point can be from the regular coarse grid, effectively controlling the maximum acquisition gap.

The design of the sampling operator $\mathbf{M}$ is critical to the success of the recovery algorithm. Note that we overwrite our notation of the sampling operator from Chapter 3 and define the operator $\mathbf{M}$ as the $n \times N$ acquisition operator. We present a pragmatic marine acquisition scheme wherein a single (and/or multiple) source vessel(s) maps the survey area while firing shots at jittered time instances, which translate to jittered shot locations for a fixed (conventional) speed of the source vessel. Conventional acquisition with one source vessel and two air gun arrays where each air-gun array fires at every alternate periodic location is called flip-flop acquisition. If we wish to acquire 10.0 s-long shot records at every 12.5 m, the speed of the source vessel would have to be reduced to about 1.25 m/s (≈ 2.5 knots). The conventional speed of seismic source vessels is about 2.0–2.5 m/s (≈ 4–5 knots). Figure 4.2(a) illustrates a conventional acquisition scheme with one source vessel travelling at about 1.25 m/s carrying two air-gun arrays, where each air-gun array fires every 20.0 s (or 25.0 m) in a flip-flop manner, resulting in nonoverlapping shot records. In time-jittered acquisition, the source vessel travels at a conventional speed of about 2.5 m/s with air-gun arrays firing every 20.0 s (or 50.0 m) jittered time instances (or shot locations), i.e., the minimum interval between the jittered times (or shots) is maintained at 10.0 s (or 25.0 m, a practical requirement)

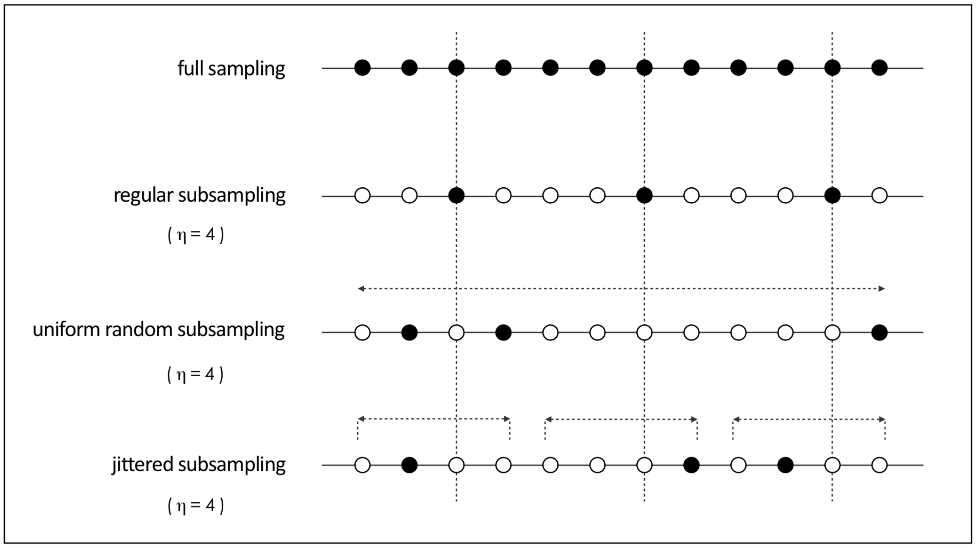


**Figure 4.1:** Schematic comparison between different subsampling schemes. $\eta$ is the subsampling factor. The vertical dashed lines define the regularly subsampled spatial grid.

and the maximum interval is 30.0 s (or 75.0 m). Figure 4.2(b) depicts this scenario resulting in overlapping shot records. This corresponds to ($\eta =$) 2× subsampled jittered acquisition grid for a conventional acquisition with nonoverlapping shot records at every 12.5 m. $\eta$ is the subsampling factor, which is calculated as

$$\eta = \frac{1}{\text{number of air-gun arrays}} \times \frac{\text{jittered spatial grid interval}}{\text{conventional spatial grid interval}} = \frac{1}{2} \times \frac{50.0 \text{ m}}{12.5 \text{ m}} = 2. \tag{4.3}$$

Note that the source vessel travels at a fixed conventional speed during the time-jittered acquisition, i.e., it does *not* accelerate or decelerate while firing at jittered instances in time (cf. Chapter 3), which would render this scenario impractical.

With a fixed (conventional) speed of the source vessel, if conventional acquisition could be carried out at a shot interval of 6.25 m then (following Equation 4.3) acquisition on the 50.0 m jittered grid would be a result of a subsampling factor of 4 (Figures 4.2(c) and 4.2(d)). Hence, in order to recover data at finer source (and/or receiver) sampling intervals of 12.5 m, 6.25 m, etc., from simultaneous jittered data, the recovery problem becomes a joint source separation and interpolation problem. Since subsampling is performed in the source-time domain, the acquisition operator is defined as

$$\mathbf{M} := [\mathbf{I} \otimes \mathbf{T}], \tag{4.4}$$

where $\otimes$ is the Kronecker product, $\mathbf{I}$ is an $N_r \times N_r$ identity matrix, and $\mathbf{T}$ is a combined jittered-shot selector and time-shifting operator. Taking the Kronecker product of $\mathbf{T}$ with $\mathbf{I}$ simply repeats the operation of $\mathbf{T}$ on every available receiver. Note that it is also possible to subsample the receiver axis or equivalently randomize/jitter positions of the ocean-bottom transducers (as in the case of OBN acquisition).

Following the same methods of estimating the RIP (restricted isometry property) constant, $\delta_{ak}$, as used in Chapter 3, we analyze the performance of the proposed physically-realizable (or realistic) time-jittered acquisition scheme by estimating $\delta_{ak}$. In summary, if $\sigma_{min}$ and $\sigma_{max}$ are the smallest and largest singular values of $\mathbf{A}_\Lambda$, respectively, then the RIP constant

$$\delta_{ak} = \sup_{\Lambda \in \{1, \ldots P\}} \max\{1 - \sigma_{\min}, \sigma_{\max} - 1\}. \tag{4.5}$$

That is, the RIP constant is the smallest upper bound on the maximum of $\{1 - \sigma_{\min}, \sigma_{\max} - 1\}$ for all subsets $\Lambda \in \{1, \ldots P\}$ of size $ak$. The RIP constant is difficult to compute since it requires evaluating $\delta_{ak}$ for every subset of $ak$ columns of $\mathbf{A}$ and there are $\binom{P}{ak}$ of such subsets. Therefore, we use Monte Carlo simulations to approximate the value of $\delta_{ak}$. Figure 4.3 shows the histogram of the estimated RIP constant, $\hat{\delta}_\Lambda$ from 100 realizations of $\mathbf{A}_\Lambda$, for the realistic time-jittering matrix $\mathbf{A} = \mathbf{M}\mathbf{S}^H$ restricted to a set $\Lambda$ of size $k$, the size support of the largest transform-domain (i.e., curvelet-domain) coefficients of a seismic image. For each of these 100 realizations, we evaluate $\hat{\delta}_\Lambda$ for 100 random realizations of the set $\Lambda$. For the proposed time-jittering matrix, the estimated RIP constant, $\hat{\delta}_\Lambda = \max\{1 - \sigma_{\min}, \sigma_{\max} - 1\} < 1$, illustrating that the design of this realistic time-jittered marine acquisition will favor recovery by curvelet-domain sparsity promotion.

## 4.5 Experimental results

We illustrate the performance of our time-jittered marine acquisition scheme on simultaneous (time-jittered) data generated from a real seismic line from the Gulf of Suez. We use two sets of conventional data, one sampled at a source (and receiver) sampling of 12.5 m and the other sampled at a source (and receiver) sampling of 6.25 m, with $N_s = 128$ shots, $N_r = 128$ receivers and $N_t = 1024$ time samples each. Figures 4.4(a) and 4.4(b) show a common-receiver and a common-shot gather, respectively, from conventional data sampled at 12.5 m. We recover (conventional) dense periodically-sampled seismic lines from simultaneous data via $\ell_1$ minimization using 2D curvelets (Fast discrete curvelet transform (FDCT), Candès et al. (2006a)) Kroneckered with 1D wavelets as the sparsifying transform. We also compare recoveries with 3D curvelets (Ying et al., 2005). It is well known that seismic data admit sparse representations by curvelets that capture "wavefront sets" efficiently (Smith, 1998; Candès and Demanet, 2005; Candès et al., 2006a; Herrmann et al., 2008).

Figures 4.5(a) and 4.5(b) display 100.0 s of the simultaneous jittered data volumes for the 12.5 m and 6.25 m spatial sampling, respectively, where the periodic coarse 50.0 m grid is jittered using our jitter subsampling scheme (Figure 4.1) resulting in overlapping shot records. If we simply apply the adjoint of the acquisition operator to the simultaneous data, i.e., $\mathbf{M}^H\mathbf{y}$, the interferences (or source crosstalk) due to overlaps in the shot records appear as random noise, i.e., incoherent and nonsparse, in common-receiver gathers (Figures 4.6(a) and 4.6(c)) and coherent overlaps in common-shot gathers (Figures 4.6(b) and 4.6(d)). Our aim is to recover conventional, nonoverlapping shot records from simultaneous data by working with the entire simultaneous data volume, and not on a shot-by-shot basis. We compare recoveries by computing a signal-to-noise ratio:

$$\mathrm{S/N}(\mathbf{f}, \widetilde{\mathbf{f}}) = -20 \log_{10} \frac{\|\mathbf{f} - \widetilde{\mathbf{f}}\|_2}{\|\mathbf{f}\|_2}, \tag{4.6}$$

Sparsity-promoting recovery results in a S/N of 11.5 dB for 2D curvelets and 12.4 dB for 3D curvelets, effectively separating the jittered data and interpolating it to a fine 12.5 m grid. The slight improvement in S/N with 3D curvelets is attributed to exploiting sparse structure of seismic data with curvelets, which represent curve-like singularities optimally (Candès and Demanet, 2005), in all three dimensions, and therefore lead to sparser representations. The improvement in recovery is small potentially due to small size of the data volumes. Moreover, this improvement comes at the expense of increased computational costs because the 3D FDCT is about 24× redundant, in contrast to the 8× redundant 2D FDCT, rendering large-scale processing extremely memory intensive, and hence impractical. Figures 4.7(a) – 4.7(d) show a recovered common-receiver and common-shot gather and the corresponding residual, respectively, for recovery with 2D FDCT. The corresponding recoveries for 3D FDCT are shown in Figures 4.7(e) – 4.7(h). Figures 4.8(a) – 4.8(h) show the corresponding zoom sections where most of the weak late-arriving events are well recovered. Similarly, for data with the spatial sampling of 6.25 m, recovery with 2D FDCT results in a S/N of 4.9 dB while the S/N with 3D FDCT is 5.7 dB. Figures 4.9(a) – 4.9(h) show the corresponding recoveries with the zoom sections shown in Figures 4.10(a) – 4.10(h). The artifacts near the bottom of the zoom sections in Figures 4.10(e) – 4.10(h) can possibly be reduced by running more iterations of the recovery algorithm. Note that the S/N's decrease for increased subsampling factor $\eta = 4$ (for 6.25 m spatial sampling grid). This observation is in accordance with CS theory, where the recovery quality decreases for increased subsampling.

To quantify the cost savings associated with simultaneous-source acquisition, we measure the performance of the proposed acquisition design and subsequent recovery strategy in terms of an improved spatial-sampling ratio (ISSR), defined as

$$\text{ISSR} = \frac{\text{number of shots recovered via sparsity-promoting inversion}}{\text{number of shots in simultaneous-source acquisition}}. \tag{4.7}$$

For time-jittered marine acquisition, a subsampling factor $\eta = 2$, 4, etc., implies a gain in the spatial sampling by a factor of 2, 4, etc. In practice, this corresponds to an improved efficiency of the acquisition by the same factor. Recently, Mosher et al. (2014) have shown that factors of two or as high as ten in efficiency improvement are achievable in the field.

One of the performance indicators proposed by Berkhout (2008) is the survey-time ratio (STR):

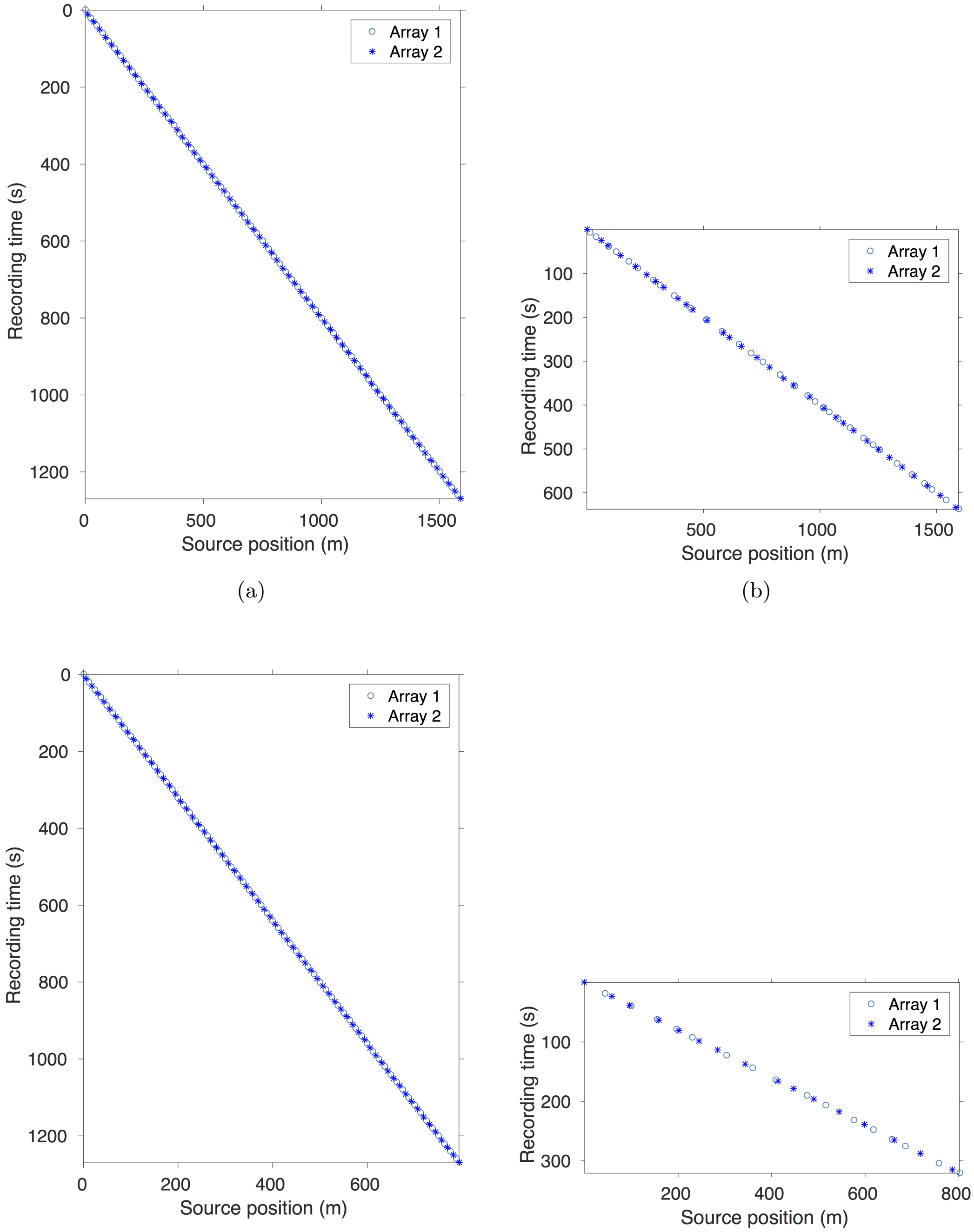
$$\text{STR} = \frac{\text{time of conventional acquisition}}{\text{time of simultaneous-source acquisition}}, \tag{4.8}$$

If we wish to acquire 10.0 s-long shot records at every 12.5 m, the speed of the source vessel would have to be reduced to about 1.25 m (≈ 2.5 knots). As mentioned previously, in simultaneous-source acquisition, speed of the source vessel is approximately maintained at (the standard) 2.5 m/s (≈ 5.0 knots). Therefore, for a subsampling factor of $\eta = 2$, 4, etc., there is an implicit reduction in the survey time by $\frac{1}{\eta}$.

## 4.6 Conclusions

Simultaneous-source time-jittered marine acquisition is an instance of compressive sensing, which shares the benefits of random sampling while offering control on the maximum acquisition gap size. The results vindicate the importance of randomness in the acquisition scheme, wherein the more randomness we have in terms of the air-gun firing times/shot locations (as shown here) and/or receiver locations, the more readily we can adapt ideas from CS to sample data economically (i.e., acquire subsampled data) and recover dense periodically-sampled data via structure promotion.

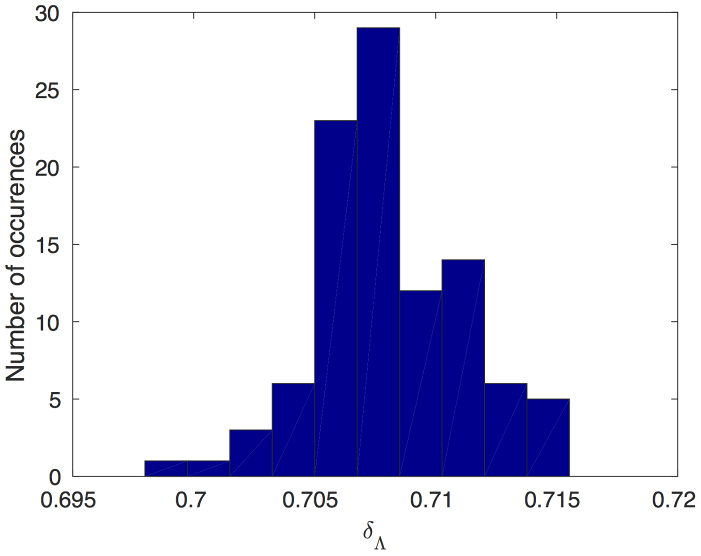
Using 3D FDCT for sparsity-promoting recovery results in slightly improved recoveries compared to 2D FDCT, since it exploits curvelet-domain sparsity in all three dimensions, but at the expense of increased computational costs. The redundancy of 3D FDCT (about 24 ×) renders large-scale processing extremely memory intensive, and hence impractical. The combination of randomized sampling and sparsity-promoting recovery technique aids in improved source separation coupled with interpolation to finer and finer sampling grids, mitigating acquisition-related costs in the increasingly complicated regions of the earth to produce images of desired resolution. Future work includes working with nonuniform sampling grids.



(a) (b)

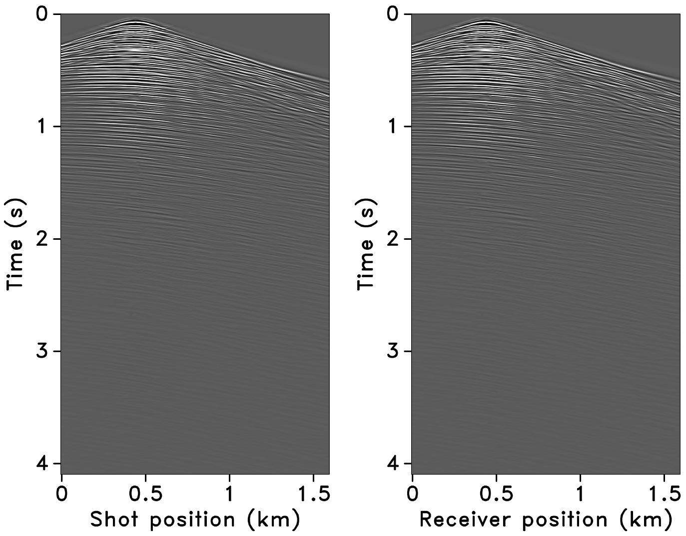
(c) (d)

**Figure 4.2:** Acquisition geometry. (a,c) Conventional marine acquisition with one source vessel and two air-gun arrays for a spatial sampling of 12.5 m and 6.25 m, respectively. (b,d) The corresponding time-jittered marine acquisition with $\eta = 2$ and $\eta = 4$, respectively. Note the acquisition speedup during jittered acquisition, where the recording time is reduced to one-half and one-fourth the recording time of the conventional acquisition, respectively.



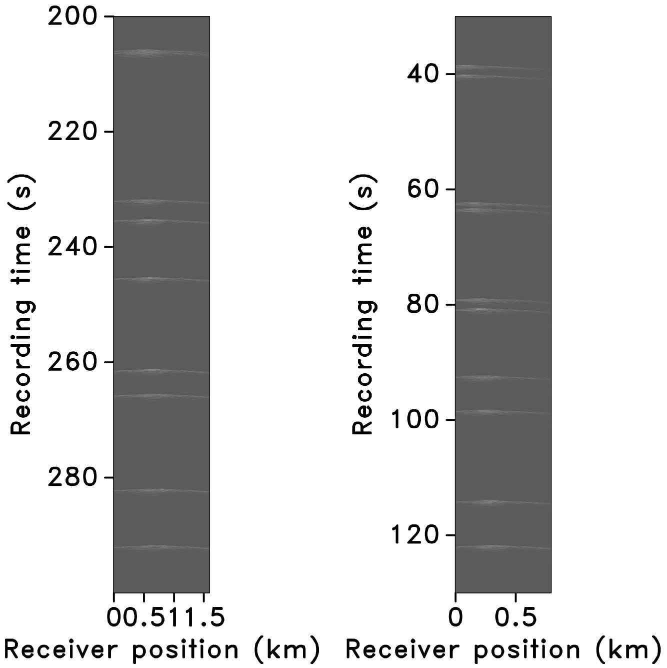
(a)

**Figure 4.3:** Histogram of $\hat{\delta}_\Lambda$ from 100 realizations of $\mathbf{A}_\Lambda$, restricted to a set $\Lambda$ of size $k$, the size support of the largest curvelet-domain coefficients of a real (Gulf of Suez) seismic image.



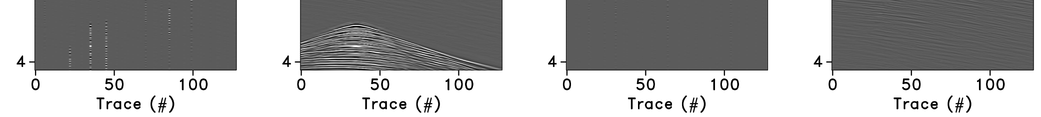
(a) (b)

**Figure 4.4:** Conventional data for a seismic line from the Gulf of Suez. (a) Common-receiver gather spatially sampled at 12.5 m. (b) Common-shot gather spatially sampled at 12.5 m.



(a) (b)

**Figure 4.5:** Simultaneous data for conventional data spatially sampled at (a) 12.5 m and (b) 6.25 m. Note that only 100.0 s of the full simultaneous data volume is shown.



(a) (b) (c) (d)

**Figure 4.6:** Interferences (or source crosstalk) in a (a) common-receiver gather and (b) common-shot gather for data spatially sampled at 12.5 m; and in a (c) common-receiver gather and (d) common-shot gather for data spatially sampled at 6.25 m. Since the subsampling factor $\eta = 2$ and $\eta = 4$ for a spatial sampling of 12.5 m and 6.25 m, respectively, (a) and (c) also have missing traces. The simultaneous data are separated and interpolated to their respective fine sampling grids.

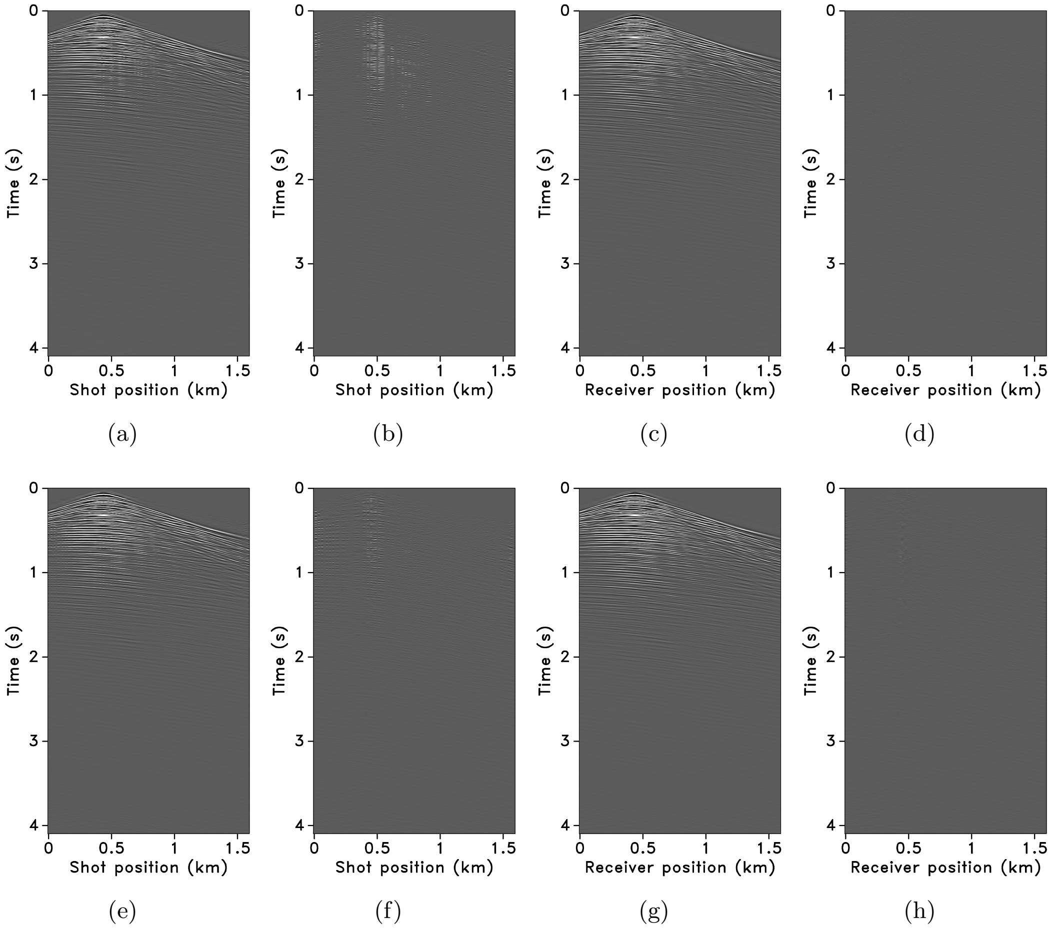


**Figure 4.7:** Recovered data for a subsampling factor $\eta = 2$. (a,e) Common-receiver gathers recovered with 2D FDCT and 3D FDCT, respectively. (b,f) The corresponding difference from conventional data. (c,g) Common-shot gathers recovered with 2D FDCT and 3D FDCT, respectively. (d,h) The corresponding difference from conventional data.

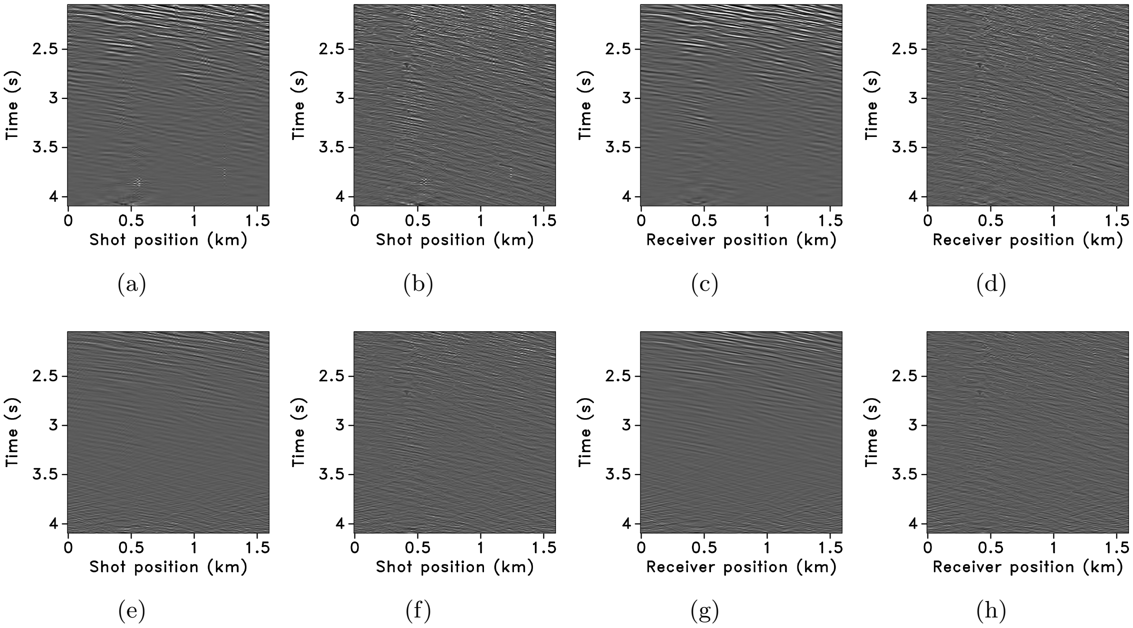
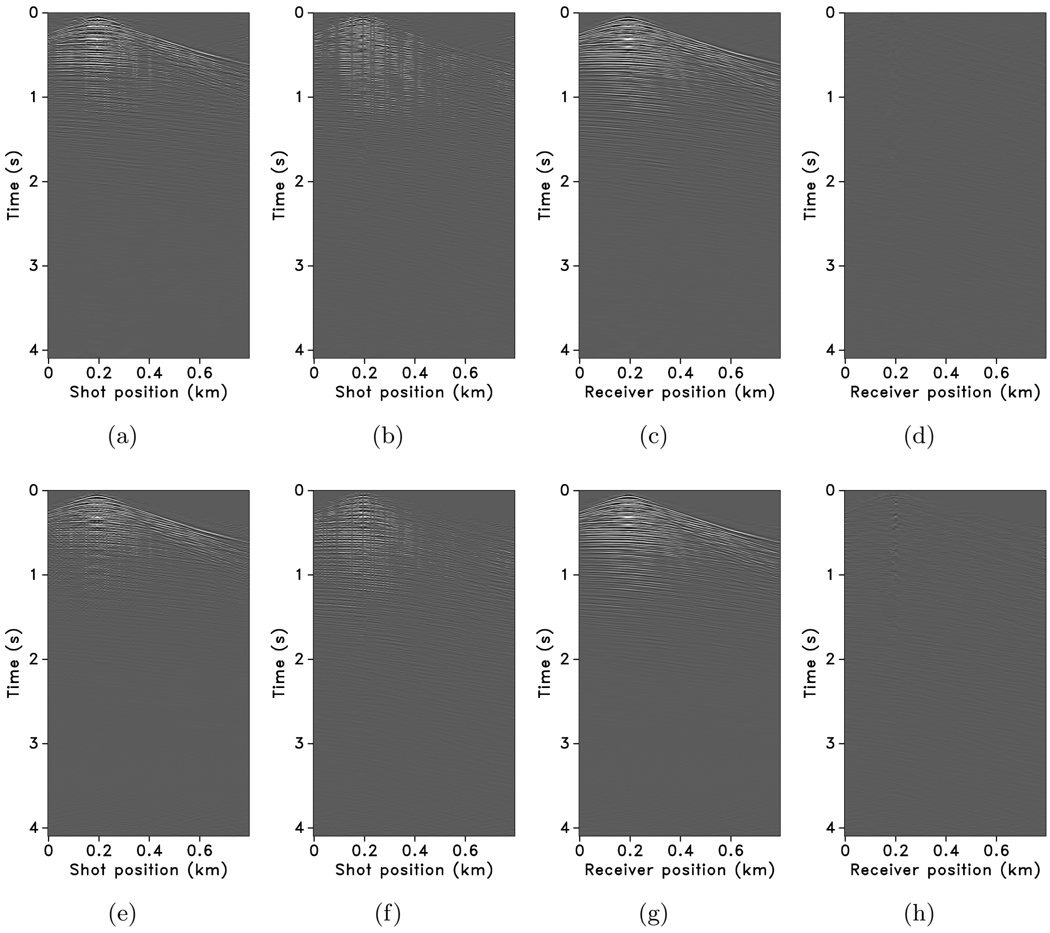


**Figure 4.8:** Zoom sections of recovered data for a subsampling factor $\eta = 2$. Note that the color axis has been clipped at one-tenth the color axis of Figure 4.7. (a,e) Common-receiver gathers recovered with 2D FDCT and 3D FDCT, respectively. (b,f) The corresponding difference from conventional data. (c,g) Common-shot gathers recovered with 2D FDCT and 3D FDCT, respectively. (d,h) The corresponding difference from conventional data.



(a) (b) (c) (d)

(e) (f) (g) (h)

**Figure 4.9:** Recovered data for a subsampling factor $\eta = 4$. (a,e) Common-receiver gathers recovered with 2D FDCT and 3D FDCT, respectively. (b,f) The corresponding difference from conventional data. (c,g) Common-shot gathers recovered with 2D FDCT and 3D FDCT, respectively. (d,h) The corresponding difference from conventional data.

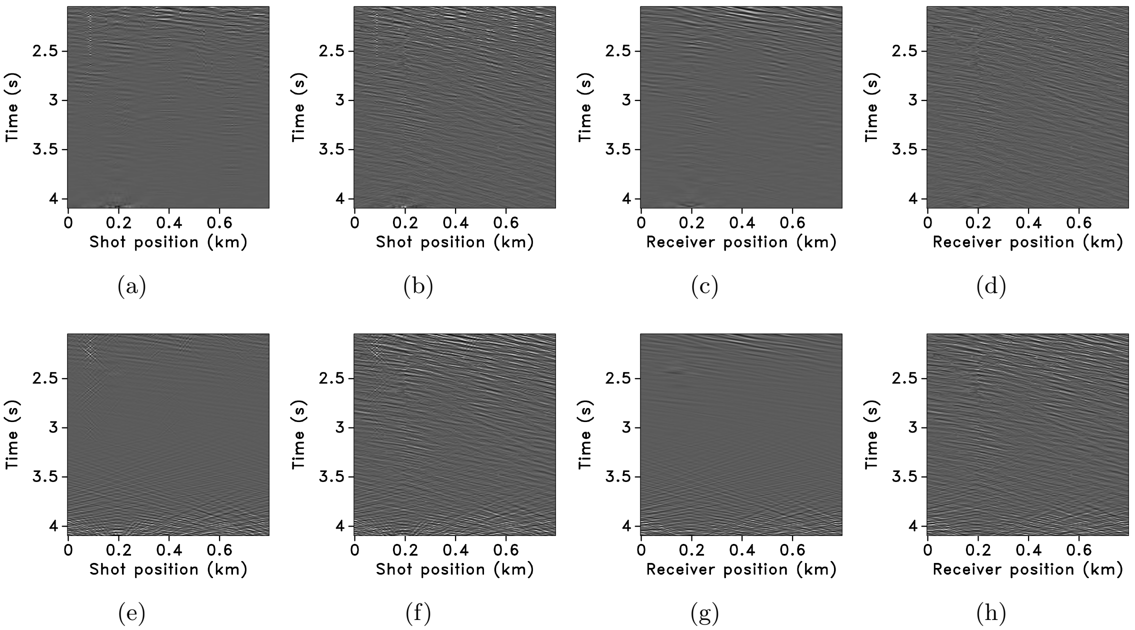


**Figure 4.10:** Zoom sections of recovered data for a subsampling factor $\eta = 4$. Note that the color axis has been clipped at one-tenth the color axis of Figure 4.9. (a,e) Common-receiver gathers recovered with 2D FDCT and 3D FDCT, respectively. (b,f) The corresponding difference from conventional data. (c,g) Common-shot gathers recovered with 2D FDCT and 3D FDCT, respectively. (d,h) The corresponding difference from conventional data.

# Chapter 5

# Low-cost time-lapse seismic with distributed Compressive Sensing — exploiting common information amongst the vintages

## 5.1 Summary

Time-lapse seismic is a powerful technology for monitoring a variety of subsurface changes due to reservoir fluid flow. However, the practice can be technically challenging when one seeks to acquire colocated time-lapse surveys with high degrees of replicability amongst the shot locations. We demonstrate that under "ideal" circumstances, where we ignore errors related to taking measurements off the grid, high-quality prestack data can be obtained from randomized subsampled measurements that are observed from surveys where we choose not to revisit the same randomly subsampled on-the-grid shot locations. Our acquisition is low cost since our measurements are subsampled. We find that the recovered finely sampled prestack baseline and monitor data actually improve significantly when the same on-the-grid shot locations are not revisited. We achieve this result by using the fact that different time-lapse data share information and that nonreplicated (on-the-grid) acquisitions can add information when prestack data are recovered jointly. Whenever the time-lapse data exhibit joint structure—i.e., are compressible in some transform domain and share information—sparsity-promoting recovery of the "common component" and "innovations", with respect to this common component, outperforms independent recovery of both the prestack baseline and monitor data. The recovered time-lapse data are of high enough quality to serve as input to extract poststack attributes used to compute time-lapse differences. Without joint recovery, artifacts—due to the randomized subsampling—lead to deterioration of the degree of repeatability of the time-lapse data. We support our claims by carrying out experiments that collect reliable statistics from thousands of repeated experiments. We also confirm that high degrees of repeatability are achievable for an ocean-bottom cable survey acquired with time-jittered continuous recording.

---

A version of this chapter has been published in Geophysics, 2017, vol. 82, pp. P1–P13.

## 5.2 Introduction

Time-lapse (4D) seismic techniques involve the acquisition, processing and interpretation of multiple 2D or 3D seismic surveys, over a particular time period of production (Lumley, 2001). While this technology has been applied successfully for reservoir monitoring (Koster et al., 2000; Fanchi, 2001) and $CO_2$ sequestration (Lumley, 2010), it remains a challenging and expensive technology because it relies on finely sampled and replicated surveys each of which have their challenges (Lumley and Behrens, 1998). To improve repeatability of the combination of acquisition and processing, various approaches have been proposed varying from more repeatable survey geometries (Beasley et al., 1997; Porter-Hirsche and Hirsche, 1998; Eiken et al., 2003; Brown and Paulsen, 2011; Eggenberger et al., 2014) to tailored processing techniques (Ross and Altan, 1997) such as cross equalization (Rickett and Lumley, 2001), curvelet-domain processing (Beyreuther et al., 2005) and matching (Tegtmeier-Last and Hennenfent, 2013).

We present a new approach that addresses these acquisition- and processing-related issues by explicitly exploiting common information shared by the different time-lapse vintages. To this end, we consider time-lapse acquisition as an inversion problem, which produces finely sampled colocated data from randomly subsampled baseline and monitor measurements. The presented joint recovery method, which derives from distributed compressive sensing (DCS, Baron et al., 2009), inverts for the "common component" and "innovations" with respect to this common component. As during conventional compressive sensing (CS, Donoho, 2006; Candès and Tao, 2006), which has successfully been adapted and applied to various seismic settings (Hennenfent and Herrmann, 2008; Herrmann, 2010; Mansour et al., 2012b; Wason and Herrmann, 2013b) including actual field surveys (see e.g., Mosher et al., 2014), the proposed method exploits transform-based (curvelet) sparsity in combination with the fact that randomized acquisitions break this structure and thereby create favorable recovery conditions.

While the potential advantages of randomized subsampling on individual surveys are relatively well understood (see e.g., Wason and Herrmann, 2013b), the implications of these randomized subsampling schemes on time-lapse seismic have not yet been studied, particularly regarding achievable repeatability of the prestack data after recovery and processing. Since the different surveys contain the common component and their respective innovations, the question is how the proposed joint recovery model performs on the vintages and the time-lapse differences, and what is the importance of replicating the surveys. Our analyses will be carried out assuming our observations lie on a discrete grid so that exact survey replicability is in principle achievable. In this situation, we ignore any errors associated with taking measurements from an irregular grid. Our approach makes our time-lapse acquisition low-cost since our measurements are always subsampled and we do not necessarily replicate the surveys. In the next chapter, we demonstrate how we deal with the effects of non-replicability of the surveys, particularly when we take measurements from an irregular grid. Since the observations are subsampled and on the grid for this chapter (off the grid for the next chapter), the aim is to recover vintages on a colocated fine grid.

We also ignore complicating factors—such as tidal differences and seasonal changes in water temperature—that may adversely affect repeatability of the time-lapse surveys. Since one of the goals of 4D seismic data processing is to obtain excellent 3D seismic images for each data set (Lumley, 2001), and since time-lapse changes are mostly derived from poststack attributes (Landrø, 2001; Spetzler and Kvam, 2006), we will be mainly concerned with the quality of the prestack vintages themselves rather than the prestack time-lapse differences.

The chapter is organized as follows. First, we summarize the main findings of CS, its governing equations, and its main premise that structured signals can be recovered from randomized measurements sampled at a rate below Nyquist. Next, we set up the CS framework for time-lapse surveys, and we discuss an independent recovery strategy, where the baseline and monitor data are recovered independently. We juxtapose this approach with our joint recovery method, which produces accurate estimates for the common component—i.e., the component that is shared amongst all vintages—and innovations with respect to this common component. To study the performance of these two recovery strategies, we conduct a series of stylized experiments for thousands of random realizations that capture the essential features of randomized seismic acquisition. From these experiments, we compute recovery probabilities as a function of the number of measurements and survey replicability, the two main factors that determine the cost of seismic acquisitions. Next, we conduct a series of synthetic experiments that involve time-lapse ocean-bottom surveys with time-jittered continuous recordings and overlapping shots as recently proposed by Wason and Herrmann (2013b). Aside from computing signal-to-noise ratios measured with respect to finely sampled true baseline, monitor, and time-lapse differences and their stacks, we also use Kragh and Christie (2002)'s root-mean-square (NRMS) metric to quantify the repeatability of the recovered data.

## 5.3 Methodology

### 5.3.1 Synopsis of compressive sensing

Compressive sensing (CS) is a sampling paradigm that aims to reconstruct a signal $\mathbf{x} \in \mathbb{R}^N$ ($N$ is the fully sampled ambient dimension) that is *sparse* (only a few of the entries are non-zero) or *compressible* (can be well approximated by a sparse signal) in some transform domain, from few measurements $\mathbf{y} \in \mathbb{R}^n$, with $n \ll N$. According to the theory of CS (Candès et al., 2006c; Donoho, 2006), recovery of $\mathbf{x}$ is attained from $n$ linear subsampled measurements given by

$$\mathbf{y} = \mathbf{A}\mathbf{x}, \tag{5.1}$$

where $\mathbf{A} \in \mathbb{R}^{n \times N}$ is the sampling matrix.

Finding a solution to the above underdetermined system of equations involves solving the following sparsity-promoting convex optimization program :

$$\widetilde{\mathbf{x}} = \arg\min_{\mathbf{x}} \|\mathbf{x}\|_1 := \sum_{i=1}^{N} |x_i| \quad \text{subject to} \quad \mathbf{y} = \mathbf{A}\mathbf{x}. \tag{5.2}$$

where $\widetilde{\mathbf{x}}$ is an approximation of $\mathbf{x}$. In the noise-free case, this ($\ell_1$-minimization ) problem finds amongst all possible vectors $\mathbf{x}$, the vector that has the smallest $\ell_1$-norm and that explains the observed subsampled data. To arrive at this solution, we use the software package SPG$\ell_1$ (Berg and Friedlander, 2008). The main contribution of CS is to design sampling matrices that guarantee solutions to the recovery problem in Equation 6.1, by providing rigorous proofs in specific settings. Furthermore, a key highlight in CS is that favorable conditions for recovery is attained via randomized subsampling rather than periodic subsampling. This is because random subsampling introduces incoherent, and therefore non-sparse, subsampling related artifacts that are removed during sparsity-promoting signal recovery. Basically, CS is an extension of the anti-leakage Fourier transform (Xu et al., 2005; Schonewille et al., 2009), where random sampling in the physical domain renders coherent aliases into incoherent noisy crosstalk (leakage) in the spatial Fourier domain. In this case, the signal is sparse in the Fourier basis.

For details on precise recovery conditions in terms of the number of measurements $n$, allowable recovery error, and construction of measurement/sampling matrices $\mathbf{A}$, we refer to the literature on compressive sensing (Donoho, 2006; Candès and Tao, 2006; Candès and Wakin, 2008). For our application to time-lapse seismic, we follow adaptations of this theory by Herrmann et al. (2008) and Herrmann and Hennenfent (2008), and use curvelets as the sparsifying transform in the seismic examples that involve randomized marine acquisition (Mansour et al., 2012b; Wason and Herrmann, 2013b; Wason et al., 2015). The latter references involve marine acquisition with ocean-bottom nodes and time-jittered time-compressed firing times with single or multiple source vessels. As shown by Wason and Herrmann (2013b), this type of randomized acquisition and processing leads to better wavefield reconstructions than the processing of regularly subsampled data. Furthermore, because of the reduced acquisition time, it is more efficient economically (Mosher et al., 2014).

### 5.3.2 Independent recovery strategy (IRS)

To arrive at a compressive sensing formulation for time-lapse seismic, we describe noise-free time-lapse data acquired from the baseline ($j = 1$) and monitor ($j = 2$) surveys as

$$\mathbf{y}_j = \mathbf{A}_j \mathbf{x}_j \quad \text{for} \quad j = \{1, 2\}. \tag{5.3}$$

In this CS formulation, which can be extended to $J > 2$ surveys, the vectors $\mathbf{y}_1$ and $\mathbf{y}_2$ represent the corresponding subsampled measurement vectors; $\mathbf{A}_1$ and $\mathbf{A}_2$ are the corresponding flat ($n \ll N$) measurement matrices, which are not necessarily equal. As before, finely sampled vintages can in principle be recovered under the right conditions by solving Equation 5.3 with a sparsity-promoting optimization program (cf. Equation 6.1) for each vintage separately. We will refer to this approach as the independent recovery strategy (IRS). In this context, we compute the time-lapse signal by directly subtracting the recovered vintages.

### 5.3.3 Shared information amongst the vintages

Aside from invoking randomizations during subsampling, CS exploits structure residing within seismic data volumes during reconstruction—the better the compression the better the reconstruction becomes for a given set of measurements. If we consider the surveys separately, curvelets are good candidates to reveal this structure because they concentrate the signal's energy into few large-magnitude coefficients and many small coefficients (see left-hand side plot in Figure 5.1). Curvelets have this ability because they decompose seismic data into multiscale and multi-angular localized waveforms. As the cross plot in Figure 5.1 reveals (right-hand side plot), the curvelet transform's ability to compress seismic data and time-lapse difference (left-hand side plot Figure 5.1) is not the only type of structure that we can exploit. The fact that most of the magnitudes of the curvelet coefficients of two common-receiver gathers from a 2D OBS time-lapse survey (see Figure 5.8) nearly coincide indicate that the data from the two vintages shares lots of information in the curvelet domain. Therefore, we can further exploit this complementary structure during time-lapse recovery from randomized subsampling in order to improve the repeatability.

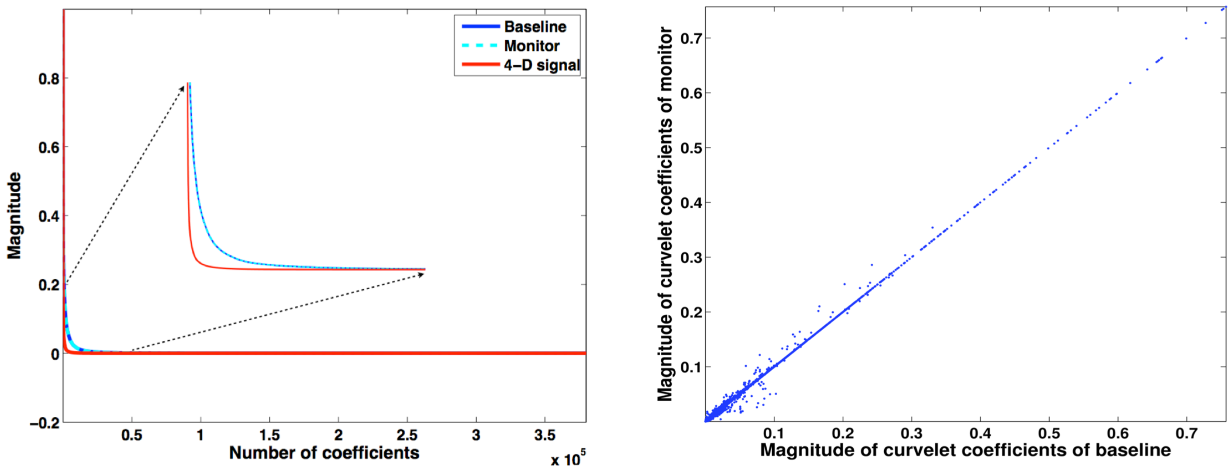


**Figure 5.1:** Left: Decay of curvelet coefficients of time-lapse data and difference. Right: Scatter plot of curvelet coefficients of the baseline and monitor data indicating that they share significant information.

### 5.3.4 Joint recovery method (JRM)

Baron et al. (2009) introduced and analyzed mathematically a model for distributed CS where jointly sparse signals are recovered jointly. Aside from permitting sparse representations individually, jointly sparse signals share information. For instance, sensor arrays aimed at the same object tend to share information (see Xiong et al. (2004) and the references therein) and time-lapse seismic surveys are no exception.

There are different ways to incorporate this shared information amongst the different vintages. We found that we get the best recovery result if we exploit the common component amongst the baseline and monitor data explicitly. This means that for two-vintage surveys we end up with three unknown vectors. One for the common component, denoted by $\mathbf{z}_0$, and two for the innovations $\mathbf{z}_j$ for $j \in 1, 2$ with respect to this common component that is shared by the vintages. In this model, the vectors for the vintages are given by

$$\mathbf{x}_j = \mathbf{z}_0 + \mathbf{z}_j, \qquad j \in 1, 2. \tag{5.4}$$

As we can see, the vintages contain the common component $\mathbf{z}_0$ and the time-lapse difference is contained within the difference between the innovations $\mathbf{z}_j$ for $j \in 1, 2$. Because $\mathbf{z}_0$ is part of both surveys, the observed measurements are now given by

$$\begin{aligned} \begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{bmatrix} &= \begin{bmatrix} \mathbf{A}_1 & \mathbf{A}_1 & \mathbf{0} \\ \mathbf{A}_2 & \mathbf{0} & \mathbf{A}_2 \end{bmatrix} \begin{bmatrix} \mathbf{z}_0 \\ \mathbf{z}_1 \\ \mathbf{z}_2 \end{bmatrix}, \quad \text{or} \\ \mathbf{y} &= \mathbf{A}\mathbf{z}. \end{aligned} \tag{5.5}$$

In this expression, we overloaded the symbol $\mathbf{A}$, which from now on refers to the matrix linking the observations of the time-lapse surveys to the common component and innovations pertaining to the different vintages. The above joint recovery model readily extends to $J > 2$ surveys, yielding a $J \times$ (number of vintages $+ 1$) system.

Contrary to the IRS, which essentially corresponds to setting the common component to zero so there is no communication between the different surveys, both vintages share the common component in Equation 5.5. As a result correlations amongst the vintages will be exploited if we solve instead

$$\widetilde{\mathbf{z}} = \arg\min_{\mathbf{z}} \|\mathbf{z}\|_1 \quad \text{subject to} \quad \mathbf{y} = \mathbf{A}\mathbf{z}. \tag{5.6}$$

As a result, we seek solutions for the common component and innovations that have the smallest $\ell_1$-norm such that the observations explain both the incomplete recordings for both vintages. Estimates for the finely sampled vintages are readily obtained via Equation 6.4 with the recovered $\widetilde{\mathbf{z}}$ while the time-lapse difference is computed via $\widetilde{\mathbf{z}}_1 - \widetilde{\mathbf{z}}_2$.

Albeit recent progress has been made (Li, 2015), precise recovery conditions for JRM are not yet very well studied. Moreover, the JRM was also not designed to compute differences between the innovations. To gain some insight on our formulation, we will first compare the performance of IRS and JRM in cases where the surveys are exactly replicated ($\mathbf{A}_1 = \mathbf{A}_2$), partially replicated ($\mathbf{A}_1$ and $\mathbf{A}_2$ share certain fractions of rows), or where $\mathbf{A}_1$ and $\mathbf{A}_2$ are statistically completely independent. To get reliable statistics on the recovery performance for the different recovery schemes, we repeat a series of small stylized problems thousands of times. These small stylized examples serve as proxies for seismic acquisition problems that we will discuss later.

## 5.4 Stylized experiments

To collect statistics on the performance of the different recovery strategies, we repeat several series of small experiments many times. Each random time-lapse realization is represented by a vector with $N = 50$ elements that has $k = 13$ nonzero entries with Gaussian distributed weights that are located at random locations such that the number of nonzero entries in each innovation is two—i.e., $k_1 = k_2 = 2$. This leaves 11 nonzeros for the common component. For each random experiment, $n = \{10, 11, \cdots, 40\}$ observations $\mathbf{y}_1$ and $\mathbf{y}_2$ are collected using Equation 5.3 for Gaussian matrices $\mathbf{A}_1$ and $\mathbf{A}_2$ that are redrawn for each repeated experiment. These Gaussian matrices have independent identically distributed Gaussian entries and serve as a proxy for randomized acquisitions in the field. An example of the time lapse vectors $\mathbf{z}_0, \mathbf{z}_1, \mathbf{z}_2, \mathbf{x}_1, \mathbf{x}_2$, and $\mathbf{x}_1 - \mathbf{x}_2$ involved in these experiments is included in Figure 5.2. Our goal is to recover estimates for the vintages and time-lapse signals—i.e., we want to obtain the estimates $\widetilde{\mathbf{x}}_1$ and $\widetilde{\mathbf{x}}_2$, and their difference $\widetilde{\mathbf{x}}_1 - \widetilde{\mathbf{x}}_2$ from subsampled measurements $\mathbf{y}_1$ and $\mathbf{y}_2$. When using the joint recovery model, we compute estimates for the jointly sparse vectors via $\widetilde{\mathbf{x}}_1 = \widetilde{\mathbf{z}}_0 + \widetilde{\mathbf{z}}_1$, and $\widetilde{\mathbf{x}}_2 = \widetilde{\mathbf{z}}_0 + \widetilde{\mathbf{z}}_2$, where $\widetilde{\mathbf{z}}$ is found by solving Equation 6.5.

To get reliable statistics on the probability of recovering the vectors representing the vintages and the time-lapse differences, we choose to perform $M = 2000$ repeated time-lapse experiments generating $M$ different realizations for $\mathbf{y}_1$ and $\mathbf{y}_2$ from different realizations of $\mathbf{x}_1$ and $\mathbf{x}_2$. Next, we recover $\widetilde{\mathbf{x}}_1$ and $\widetilde{\mathbf{x}}_2$ from these measurements using the IRS or JRM. From these estimates, we compute empirical probabilities of successful recovery via

$$P(\mathbf{x}) = \frac{\text{Number of times} \dfrac{\|\mathbf{x} - \tilde{\mathbf{x}}\|_2}{\|\mathbf{x}\|_2} < \rho}{M}. \tag{5.7}$$

We set the relative error threshold to $\rho = 0.1$. The vector $\mathbf{x}$ either represents the vintages or the

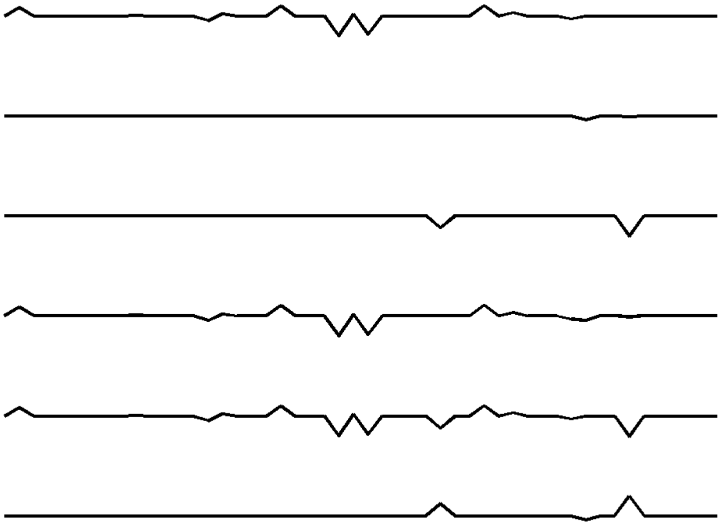


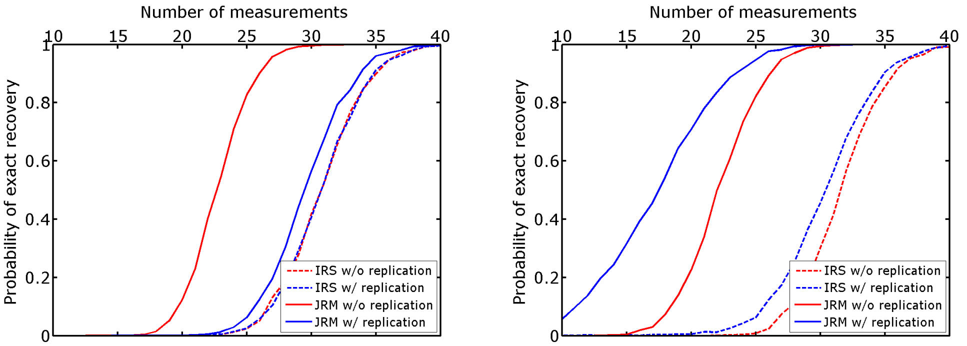
**Figure 5.2:** From top to bottom: $\mathbf{z}_0, \mathbf{z}_1, \mathbf{z}_2, \mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_1 - \mathbf{x}_2$. We are particularly interested in recovering estimates for $\mathbf{x}_1, \mathbf{x}_2$ and $\mathbf{x}_1 - \mathbf{x}_2$ from $\mathbf{y}_1$ and $\mathbf{y}_2$.

difference. In case of the vintages, we multiply the probabilities.

### 5.4.1 Experiment 1—independent versus joint recovery

To reflect current practices in time-lapse acquisition—where people aim to replicate the surveys—we run the experiments by drawing the same random Gaussian matrices of size $n \times N$ for $n = \{10, 11, \cdots, 40\}$ and $N = 50$ for $\mathbf{A}_1$ and $\mathbf{A}_2$—i.e., $\mathbf{A}_1 = \mathbf{A}_2$. We conduct the same experiments where the surveys are not replicated by drawing statistically independent measurement matrices for each repeated experiment, yielding $\mathbf{A}_1 \neq \mathbf{A}_2$. For each series of experiments, we recover estimates $\widetilde{\mathbf{x}}_1$, $\widetilde{\mathbf{x}}_2$, and $\widetilde{\mathbf{x}}_1 - \widetilde{\mathbf{x}}_2$ from which we compute the corresponding recovery probabilities using Equation 5.7. The results are plotted in Figure 5.3 for the recovery of the vintages (Figure 5.3(a)) and time-lapse difference (Figure 5.3(b)).

The results of these experiments indicate that regardless of the number of measurements, JRM leads to improved recovery compared to IRS because it exploits information shared by the two jointly sparse vectors representing the vintages. The recovery probabilities for JRM (solid lines in Figure 5.3) show an overall improvement for both the time-lapse vectors and the time-lapse difference vector—all probability curves are to the left compared to those from IRS meaning that recovery is more likely for fewer measurements. For the time-lapse vectors, this improvement is much more pronounced for measurement matrices that are statistically independent—i.e., not replicated ($\mathbf{A}_1 \neq \mathbf{A}_2$). This observation is consistent with distributed compressive sensing, which predicts significant improvements when the time-lapse vectors share a significant common component. In that case, the shared component benefits most from being observed by both surveys (via



(a)

(b)

**Figure 5.3:** Recovery of (a) the jointly sparse signals $\mathbf{x}_1$ and $\mathbf{x}_2$, (b) $\mathbf{x}_1 - \mathbf{x}_2$; with and without repetition of the measurement matrices, using the independent recovery strategy versus the joint recovery method.

the first column of $\mathbf{A}$, cf. Equation 5.5). The IRS results for the time-lapse vectors are much less affected whether the survey is replicated or not, which makes sense because the processing is done in parallel and independently. This suggests that for time-lapse seismic, independent surveys give additional information on the sparse structure of the vintages that is reflected in their improved recovery quality. Another likely interpretation is that time-lapse data obtained via JRM has better repeatability compared to data obtained via IRS.

While independent surveys improve recovery with JRM, the recovery probability of the time-lapse difference vectors improves drastically when the experiments are replicated exactly. The reason for this is that the JRM simplifies to the recovery of the time-lapse differences alone in cases where the time-lapse measurements are exactly replicated. Since these time-lapse differences are sparser than the vintage vectors themselves, the time-lapse difference vectors are well recovered while the time-lapse vectors themselves are not. This result is not surprising since the error in reconstructing the vintages cancels out in the difference. This means that in CS, if one is interested in the time-lapse difference, exact repetition of the survey is preferred. However, this approach does not provide any additional structural information in the vintages. We will revisit this observation in Experiment 2 to see how the recovery performs when we have varying degrees of repeatability in the measurements.

### 5.4.2 Experiment 2—impact of degree of survey replicability

So far, we explored only two extremes, namely recovery of vintages with absolutely no replication ($\mathbf{A}_1 \neq \mathbf{A}_2$ and statistically independent) or exact replication ($\mathbf{A}_1 = \mathbf{A}_2$). To get a better understanding of how replication factors into the recovery, we repeat the experiments where we vary the degree of dependence between the surveys by changing the number of rows the matrices $\mathbf{A}_1$ and $\mathbf{A}_2$ have in common. When all rows are in common, the survey is replicated and the percentage of overlap between the surveys is a measure for the degree of replicability of the surveys. Since JRM clearly outperformed IRS, we only consider recovery with JRM.

As before we compute recovery probabilities from $M = 2000$ repeated time-lapse experiments

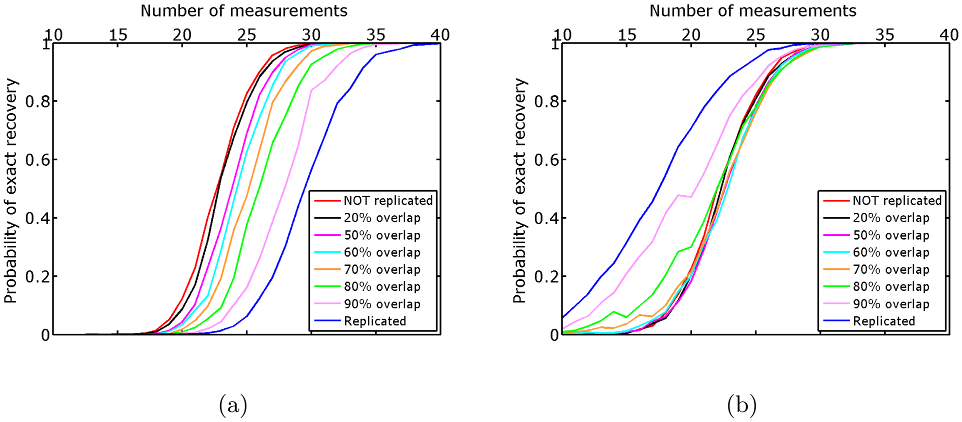


**Figure 5.4:** Recovery as a function of overlap between measurement matrices. Probability of recovering (a) $\mathbf{x}_1$ and $\mathbf{x}_2$, (b) $\mathbf{x}_1 - \mathbf{x}_2$, with joint recovery method.

generating $M$ different realizations for the observations. We summarize the recovery probability curves for varying degrees of overlap in Figure 5.4. These curves confirm that the recovery of the time-lapse vectors improves when the surveys are not replicated. As soon as the surveys are no longer replicated, the recovery probabilities for the time-lapse vectors improve. These improvements become less prominent when large percentages do not overlap and as expected reaches its maximum when the surveys become independent. Recovery of the time-lapse differences on the other hand suffers drastically when the surveys are no longer 100% replicated. When less then 80% of the surveys are no longer replicated, the recovery probabilities no longer benefit from replicating the surveys. Recovery of the time-lapse vectors, on the other hand, already improves significantly at this point.

While these experiments are perhaps too idealized and small to serve as a strict guidance on how to design time-lapse surveys, they lead to the following observations. Firstly, the recovery probabilities improve when we exploit joint sparsity amongst the time-lapse vectors via JRM. Secondly, since the joint component is observed by all surveys recovery of the common component and therefore vintages improves if the surveys are not replicated. Thirdly, the time-lapse differences benefit from high degrees of replication of the surveys. In that case, the JRM degenerates to recovery of the time-lapse difference alone and as a consequence the time-lapse vectors are not well recovered.

Even though the quality of the time-lapse difference is often considered as a good means of quality control, we caution the reader to draw the conclusion that we should aim to replicate the surveys. The reason for this is that time-lapse differences are generally computed from poststack attributes computed from finely sampled, and therefore recovered, prestack baseline and monitor data and not from prestack differences. Therefore, recovery of time-lapse difference alone may not be sufficient to draw firm conclusions. Our observations were also based on very small idealized experiments that did not involve stacking and permit exact replication, which may not be realistic in practice.

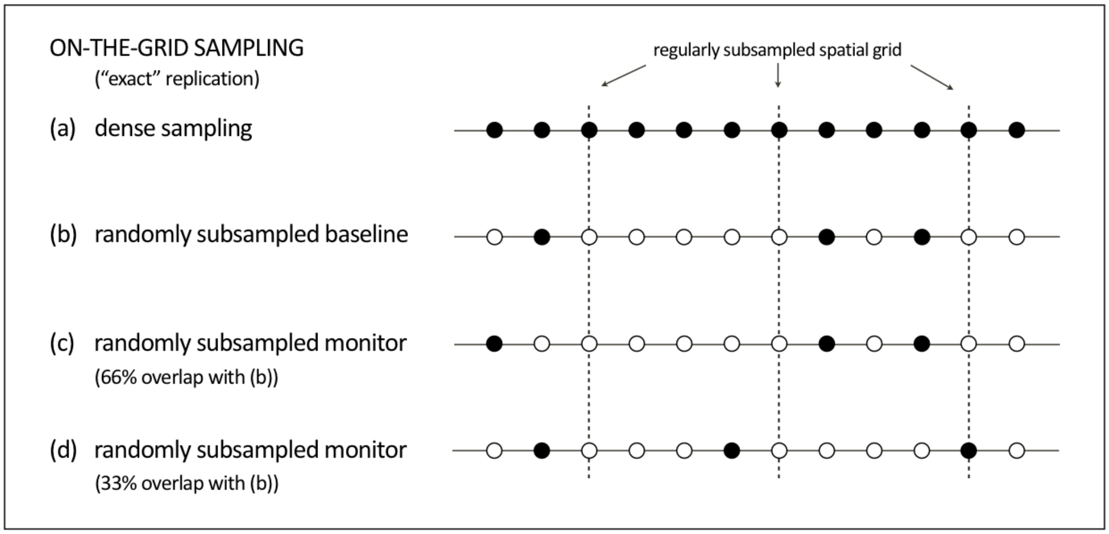
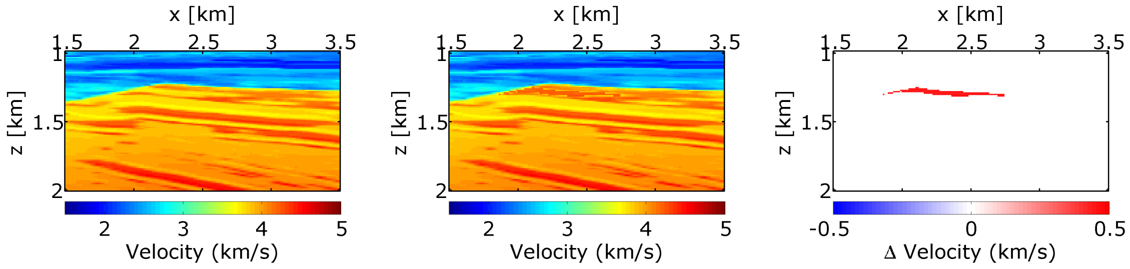


**Figure 5.5:** Schematic comparison between different random realizations of a subsampled grid. The subsampling factor is 3. As illustrated, random samples are taken exactly on the grid. Moreover, the samples are exactly replicated whenever there is an overlap between the time-lapse surveys.

## 5.5 Experimental setup—on-the-grid time-lapse randomized subsampling

One of the main parts of the experimental setup for the synthetic seismic case study is how we define the underlying grid on which samples are taken. In context of this chapter, we assume that the samples are taken on a discrete grid—i.e., samples lie "exactly" on the grid. It is also important to note that we randomly subsample the grid. As mentioned in the compressive sensing section above, randomized subsampling introduces incoherent subsampling related artifacts that are removed during sparsity-promoting signal recovery. Figure 5.5 shows a schematic comparison between different random realizations of a subsampled grid. As illustrated in the schematic, random samples are taken exactly on the grid. We define the term "overlap" as the percentage of on-the-grid shot locations exactly replicated between two (or more) time-lapse surveys. For the synthetic seismic case study, whenever there is an overlap between the surveys (e.g., 50%, 33%, 25%, etc.) the on-the-grid shot locations are exactly replicated for the baseline and monitor surveys. Similarly, for the stylized experiments, when two rows of the Gaussian matrices are the same it can be interpreted as if we hit the same shot location for both the baseline and monitor surveys. Therefore, we either assume that the experimental circumstances are ideal or alternatively we can think of this assumption as ignoring the effects of being off the grid. The next chapter analyses the effects of the more realistic off-the-grid sampling. In summary, we consider the case where measurements are exactly replicated whenever we choose to visit the same shot location for the two surveys. However, because we are subsampled we need not choose to revisit all the shot locations of the baseline survey.



(a) (b) (c)

**Figure 5.6:** Reservoir zoom of the synthetic time-lapse velocity models showing the change in velocity as a result of fluid substitution. (a) Baseline model, (b) monitor model, (c) difference between (a) and (b).

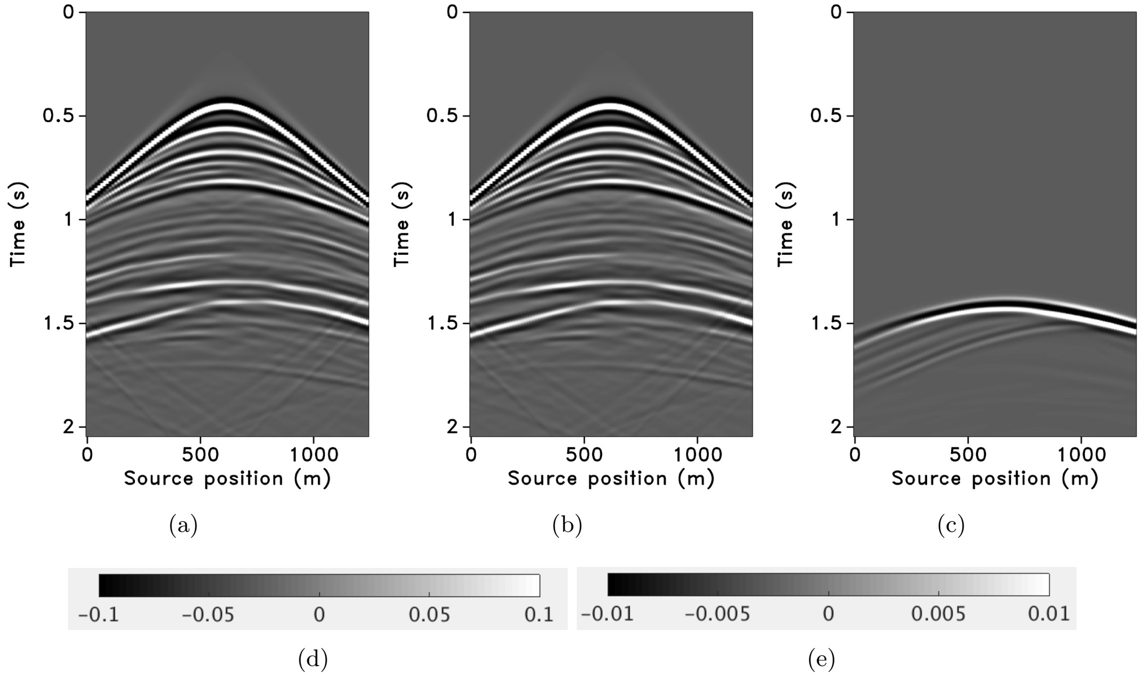
## 5.6 Synthetic seismic case study—time-lapse marine acquisition via time-jittered sources

To study a more realistic example, we carry out a number of experiments on 2D seismic lines generated from a synthetic velocity model—the BG COMPASS model (provided by BG Group). To illustrate the performance of randomized subsamplings—in particular the time-jittered marine acquisition—in time-lapse seismic, we use a subset of the BG COMPASS model (Figure 6.8(a)) for the baseline. We define the monitor model (Figure 6.8(b)) from the baseline via a fluid substitution resulting in a localized time-lapse difference at the reservoir level as shown in Figure 6.8(c).

Using IWAVE (Symes, 2010) time-stepping acoustic simulation software, two acoustic datasets with a conventional source (and receiver) sampling of 12.5 m are generated, one from the baseline model and the other from the monitor model. Each dataset has $N_t = 512$ time samples, $N_r = 100$ receivers, and $N_s = 100$ sources. Subtracting the two datasets yields the time-lapse difference, whose amplitude is small in comparison to the two datasets (about one-tenth). Since no noise is added to the data, the time-lapse difference is simply the time-lapse signal. A receiver gather from the simulated baseline data, the monitor data and the corresponding time-lapse difference is shown in Figure 5.7. In order to make the time-lapse difference visible, the color axis for the figures showing the time-lapse difference is one-tenth the scale of the color axis for the figures showing the baseline and the monitor data. This colormap applies for the remainder of the chapter. Also, the first source position in the receiver gathers—labeled as 0 m in the figures—corresponds to 725 m in the synthetic velocity model.

### 5.6.1 Time-jittered marine acquisition

Wason and Herrmann (2013b) presented a pragmatic single vessel, albeit easily extendable to multiple vessels, simultaneous marine acquisition scheme that leverages CS by invoking randomness in the acquisition via random jittering of the source firing times. As a result, source interferences become incoherent in common-receiver gathers creating a favorable condition for separating the simultaneous data into conventional nonsimultaneous data (also known as "deblending") via curvelet-domain sparsity promotion. Like missing-trace interpolation, the randomization via jittering turns the recovery into a relatively simple "denoising" problem with control over the maximum



(a) (b) (c)

(d) (e)

**Figure 5.7:** A synthetic receiver gather from the conventional (a) baseline survey, (b) monitor survey. (c) The corresponding 4D signal. (d) Color scale of the vintages. (e) Color scale of the 4D signal. Note that (e) is one-tenth the scale of (d). These color scales apply to all the corresponding figures for the vintages and the 4D signal.

gap size between adjacent shot locations (Hennenfent and Herrmann, 2008), which is a practical requirement of wavefield reconstruction with localized sparsifying transforms such as curvelets (Hennenfent and Herrmann, 2008). The basic idea of jittered subsampling is to regularly decimate the interpolation grid and subsequently perturb the coarse-grid sample points on the fine grid. A jittering parameter, dictated by the type of acquisition and parameters such as the minimum distance (or minimum recharge time for the airguns) required between adjacent shots, relates to how close and how far the jittered sampling point can be from the regular coarse grid, effectively controlling the maximum acquisition gap. Since we are still on the grid, this is a case of discrete jittering. In this chapter, we limit ourselves to the discrete case but this technique can relatively easily be taken off the grid as we discuss in the next chapter.

A seismic line with $N_s$ sources, $N_r$ receivers, and $N_t$ time samples can be reshaped into an $N$ dimensional vector $\mathbf{f}$, where $N = N_s \times N_r \times N_t$. For simplicity, we assume that all sources see the same receivers, which makes our method applicable to marine acquisition with ocean-bottom cables or nodes (OBC or OBN). As stated previously, seismic data volumes permit a compressible representation $\mathbf{x}$ in the curvelet domain denoted by $\mathbf{S}$. Therefore, $\mathbf{f} = \mathbf{S}^H\mathbf{x}$, where $^H$ denotes the Hermitian transpose (or adjoint), which equals the inverse curvelet transform. Since curvelets are a redundant frame (an over-complete sparsifying dictionary), $\mathbf{S} \in \mathbb{C}^{P \times N}$ with $P > N$, and $\mathbf{x} \in \mathbb{C}^P$.

With the inclusion of the sparsifying transform, the matrix $\mathbf{A}$ can be factored into the product of a $n \times N$ (with $n \ll N$) acquisition matrix $\mathbf{M}$ and the synthesis matrix $\mathbf{S}^H$. The design of the acquisition matrix $\mathbf{M}$ is critical to the success of the recovery algorithm. From a practical point of view, it is important to note that matrix-vector products with these matrices are matrix free—i.e., these matrices are operators that define the action of the matrix on a vector. Since the marine acquisition is performed in the source-time domain, the acquisition operator $\mathbf{M}$ is a combined jittered-shot selector and time-shifting operator. Note that in this framework it is also possible to randomly subsample the receivers.

Given a baseline data vector $\mathbf{f}_1$ and a monitor data vector $\mathbf{f}_2$, $\mathbf{x}_1$ and $\mathbf{x}_2$ are the corresponding sparse representations—i.e., $\mathbf{f}_1 = \mathbf{S}^H\mathbf{x}_1$, and $\mathbf{f}_2 = \mathbf{S}^H\mathbf{x}_2$. Given the measurements $\mathbf{y}_1 = \mathbf{M}_1\mathbf{f}_1$ and $\mathbf{y}_2 = \mathbf{M}_2\mathbf{f}_2$, and $\mathbf{A}_1 = \mathbf{M}_1\mathbf{S}^H$, $\mathbf{A}_2 = \mathbf{M}_2\mathbf{S}^H$, our aim is to recover sparse approximations $\widetilde{\mathbf{f}}_1$ and $\widetilde{\mathbf{f}}_2$ by solving sparse recovery problems for the scenarios (IRS and JRM) as described above from which the time-lapse signal can be computed.

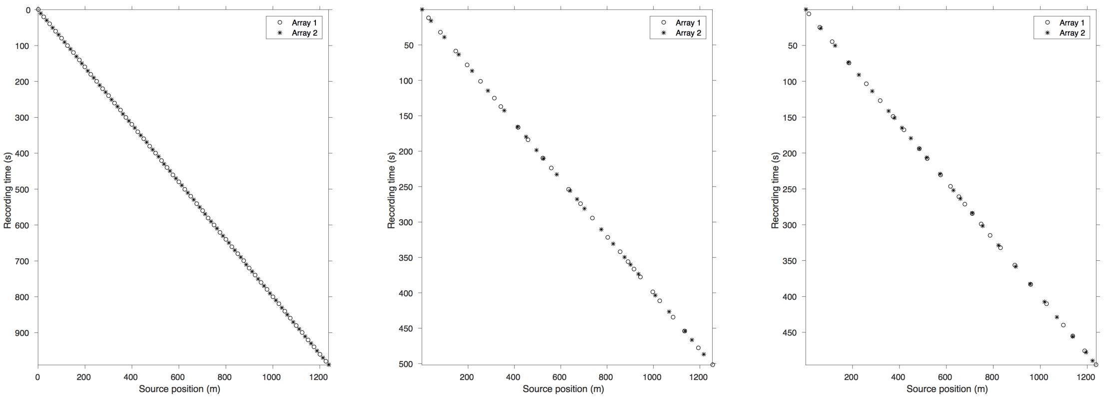
### 5.6.2 Acquisition geometry

In time-jittered marine acquisition, source vessels map the survey area firing shots at jittered time-instances, which translate to jittered shot locations for a given speed of the source vessel. Conventional acquisition with one source vessel and two airgun arrays—where each airgun array fires at every alternate periodic location—is called flip-flop acquisition. If we wish to acquire 10.0 s—long shot records at every 12.5 m, the speed of the source vessel would have to be about 1.25 m/s (approximately 2.5 knots). Figure 5.8(a) illustrates one such conventional acquisition scheme, where each airgun array fires every 20.0 s (or 25.0 m) in a flip-flop manner, traveling at about 1.25 m/s, resulting in nonoverlapping shot records of 10.0 s every 12.5 m. In time-jittered acquisition, Figure 5.8(b), each airgun array fires at every 20.0 s jittered time-instances, traveling at about 2.5 m/s (approximately 5.0 knots), with the receivers (OBC) recording continuously, resulting in overlapping (or blended) shot records (Figure 5.9(a)). Since the acquisition design involves subsampling, the acquired data volume has overlapping shot records and missing shots/traces. Consequently, the jittered flip-flop acquisition might not mimic the conventional flip-flop acquisition where airgun array 1 and 2 fire one after the other—i.e., in Figures 5.8(b) and 5.8(c), a circle (denoting array 1) may be followed by another circle instead of a star (denoting array 2). The minimum interval between the jittered times, however, is maintained at 10.0 s (typical interval required for airgun recharge) and the maximum interval is 30.0 s. For the speed of 2.5 m/s, this translates to jittering a 50.0 m source grid with a minimum (and maximum) interval of 25.0 m (and 75.0 m) between jittered shots. Both arrays fire at the 50.0 m jittered grid independent of each other.

Two realizations of the time-jittered marine acquisition are shown in Figures 5.8(b) and 5.8(c), one each for the baseline and the monitor survey. Acquisition on the 50.0 m jittered grid results in an subsampling factor,

$$\eta = \frac{1}{\text{number of airgun arrays}} \times \frac{\text{jittered spatial grid interval}}{\text{conventional spatial grid interval}} = \frac{1}{2} \times \frac{50.0\text{ m}}{12.5\text{ m}} = 2. \qquad (5.8)$$

Figures 5.9(a) and 5.9(b) show the corresponding randomly subsampled and simultaneous measurements for the baseline and monitor surveys, respectively. Note that only 50.0 s of the continuously recorded data is shown. If we simply apply the adjoint of the acquisition operator to the simultaneous data—i.e., $\mathbf{M}^H\mathbf{y}$, the interferences (or source crosstalk) due to overlaps in the shot records appear as random noise—i.e., incoherent and nonsparse, as illustrated in Figures 5.9(c) and 5.9(d). Our aim is to recover conventional, nonoverlapping shot records from simultaneous data by work-



(a)

(b)

(c)

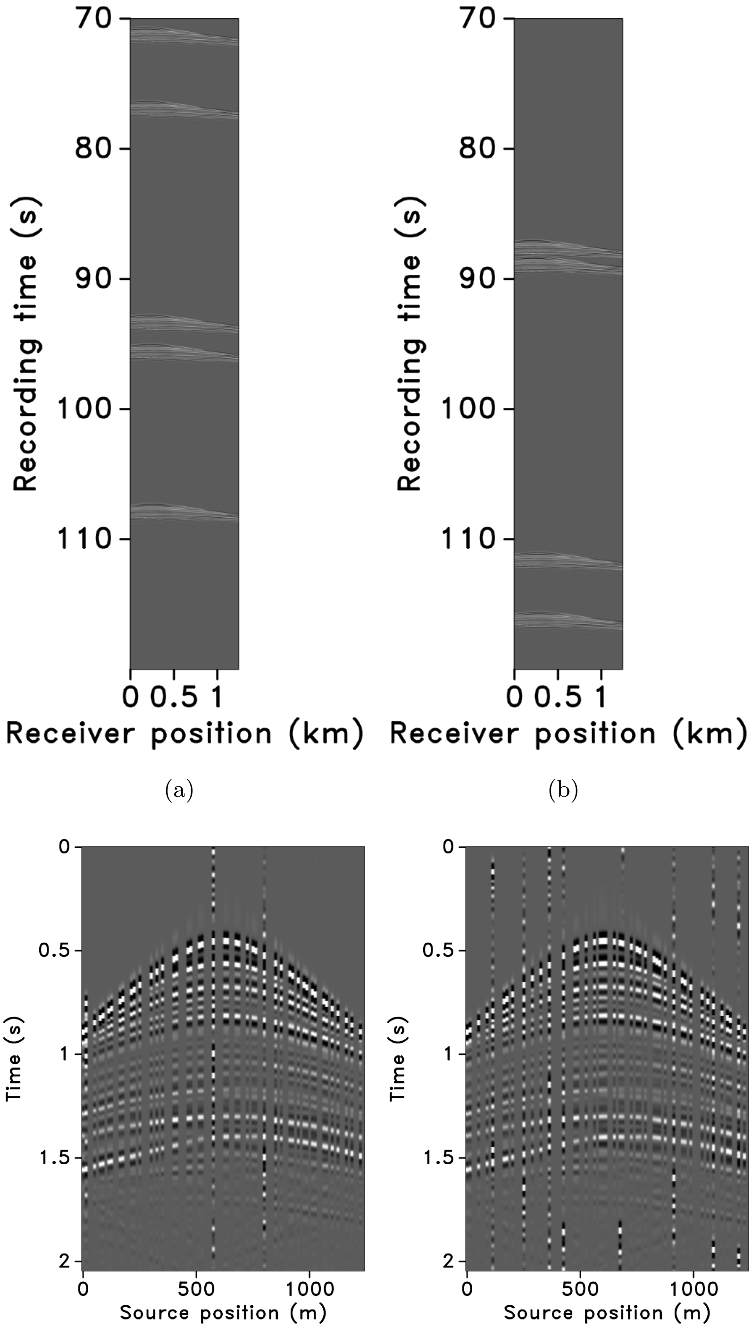
**Figure 5.8:** Acquisition geometry: (a) conventional marine acquisition with one source vessel and two airgun arrays; time-jittered marine acquisition (with $\eta = 2$) for (b) baseline, and (c) monitor. Note the acquisition speedup during jittered acquisition, where the recording time is reduced to one-half the recording time of the conventional acquisition. (b) and (c) are plotted on the same scale as (a) in order to make the jittered locations easily visible.

ing with the entire (simultaneous) data volume, and not on a shot-by-shot basis. For the present scenario, since $\eta = 2$, the recovery problem becomes a joint deblending and interpolation problem. In contrast to conventional acquisition at a source sampling grid of 12.5 m (Figure 5.8(a)), time-jittered acquisition takes half the acquisition time (Figures 5.8(b) and 5.8(c)), and the simultaneous data is separated into its individual shot records along with interpolation to the 12.5 m sampling grid. The recovery problem is solved by applying the independent recovery strategy and the joint recovery method, as we will describe in the next section.

### 5.6.3 Experiments and observations

To analyze the implications of the time-jittered marine acquisition in time-lapse seismic, we follow the same sequence of experiments as conducted for the stylized examples—i.e., we compare the independent (IRS) and joint recovery methods (JRM) for varying degrees of replicability in the acquisition. Given the 12.5 m spatial sampling of the simulated (conventional) time-lapse data, applying the time-jittered marine acquisition scheme results in a subsampling factor, $\eta = 2$ (Equation 5.8). In practice, this corresponds to an improved efficiency of the acquisition with the same factor. Recent work (Mosher et al., 2014) has shown that factors of two or as high as ten in efficiency improvement are achievable in the field. With this subsampling factor, the number of measurements for each experiment is fixed—i.e., $n = N/2$, each for $\mathbf{y}_1$ and $\mathbf{y}_2$ albeit other scenarios are possible.

We simulate different realizations of the time-jittered marine acquisition with 100%, 50%, and 25% overlap between the baseline and monitor surveys. Because we are in a discrete setting, these overlaps translate one-to-one into percentages of replicated on-the-grid shot locations for the surveys. Since $\eta = 2$, and by virtue of the design of the blended acquisition, it is not possible to



(a) (b)

(c) (d)

**Figure 5.9:** Simultaneous data for the (a) baseline and (b) monitor surveys (only 50.0 s of the full data is shown). Interferences (or source crosstalk) in a common-receiver gather for the (c) baseline and (d) monitor surveys, respectively. Since the subsampling factor $\eta = 2$, (c) and (d) also have missing traces. The simultaneous data is separated and interpolated to a sampling grid of 12.5 m.

have two completely different (0% overlap) realizations of the time-jittered acquisition. In all cases, we recover the deblended and interpolated baseline and monitor data from the blended data $\mathbf{y}_1$ and $\mathbf{y}_2$, respectively, using the independent recovery strategy (by solving Equation 6.1) and the joint recovery method (by solving Equation 6.5). As stated previously, the inherent time-lapse difference is computed by subtracting the recovered baseline and monitor data.

We perform 100 experiments for the baseline measurements, wherein each experiment has a different random realization of the measurement matrix $\mathbf{A}_1$. Then, for each experiment, we fix the baseline measurement and subsequently work with different random realizations for the monitor survey, each corresponding to the 50% and the 25% overlap. The purpose of doing this is to examine the impact of degree of replicability of acquisition in time-lapse seismic. Table 5.1 summarizes the recovery results for the stacked sections, in terms of the signal-to-noise ratio defined as

$$\mathrm{S/N}(\mathbf{f}, \widetilde{\mathbf{f}}) = -20 \log_{10} \frac{\|\mathbf{f} - \widetilde{\mathbf{f}}\|_2}{\|\mathbf{f}\|_2}, \tag{5.9}$$

for different overlaps between the baseline and monitor surveys—i.e., measurement matrices $\mathbf{A}_1$ and $\mathbf{A}_2$. Each S/N value is an average of 100 experiments including the standard deviation.

Figure 5.10 shows the recovered receiver gathers and difference plots for the monitor survey (for the different overlaps) using the independent recovery strategy (IRS), and Figure 5.11 shows the corresponding result using the joint recovery method (JRM). As illustrated in these figures, JRM leads to significantly improved recovery of the vintage compared to IRS because it exploits the shared information between the baseline and monitor data. Moreover, the recovery improves with decrease in the overlap. The IRS and JRM recovered time-lapse differences for the different overlaps are shown in Figure 5.12, which shows that recovery via JRM is still significantly better than IRS, however, the recovery is slightly improved with increase in the overlap. The edge artifacts in Figures 5.10, 5.11 and 5.12 are related to missing traces near the edges that curvelets are unable to reconstruct.

The S/Ns for the stacked sections indicate a similar trend in the observations as made from the stylized experiments—i.e., *(i)* JRM performs better than IRS because it exploits information shared between the baseline and monitor data. Note that the S/N value, which is an average of the 100 experiments, for recovery of the baseline dataset via IRS is repeated for all three cases of overlap because we work with the same 100 realizations of the jittered acquisition throughout. However, for each of the 100 experiments, different realizations are drawn for the monitor survey, which explains the variations in the S/Ns for the recovery via IRS. Similar fluctuations were observed by Herrmann (2010). *(ii)* Replication of surveys hardly affects recovery of the vintages via IRS (note similar S/Ns), since the processing is done in parallel and independently. *(iii)* Recovery of the baseline and monitor data with JRM is better when there is a small degree of overlap between the two surveys, and it decreases with increasing degrees of overlap. As explained earlier, this behavior can be attributed to partial independence of the measurement matrices that contribute additional information via the first column of $\mathbf{A}$ in Equation 6.5, i.e., for time-lapse seismic, independent surveys give additional structural information leading to improved recovery quality of the vintages. *(iv)* The converse is true for recovery of the time-lapse difference, wherein it is better if the surveys are exactly replicated. Again, as stated previously, the reason for this is the increased sparsity of the time-lapse difference itself and apparent cancelations of recovery errors due to the exactly replicated geometry.

In addition to the above observations, we find that for 100% overlap, good recovery of the

| Overlap | Baseline | | Monitor | | 4D signal | |
|---|---|---|---|---|---|---|
| | IRS | JRM | IRS | JRM | IRS | JRM |
| 100% | 23.1 ± 1.2 | 24.8 ± 1.2 | 23.1 ± 1.3 | 24.8 ± 1.2 | 21.4 ± 1.8 | 23.4 ± 2.1 |
| 50% | 23.1 ± 1.2 | 32.8 ± 1.6 | 23.4 ± 1.2 | 32.8 ± 1.6 | 9.1 ± 1.2 | 20.2 ± 1.3 |
| 25% | 23.1 ± 1.2 | 35.3 ± 1.5 | 22.0 ± 1.1 | 35.0 ± 1.5 | 7.8 ± 1.3 | 18.0 ± 1.1 |

**Table 5.1:** Summary of recoveries in terms of S/N (in dB) for the stacked sections.

stacks for IRS and JRM is possible with S/Ns that are similar for the time-lapse difference and the vintages themselves. The standard deviations for the two recovery methods are also similar. One could construe that this is the ideal situation but unfortunately it is not easily attained in practice. As we move to more practical acquisition schemes where we decrease the overlap between the surveys, we see a drastic jump downwards in the S/Ns for the time-lapse stack obtained with IRS. The results from JRM, on the other hand, decrease much more gradually with standard deviations that vary slightly from those for IRS, however, drops off with decrease in the overlap. In contrast, we actually see significant improvements for the S/Ns of the stacks of both the baseline and monitor data with slight variations in the standard deviations.

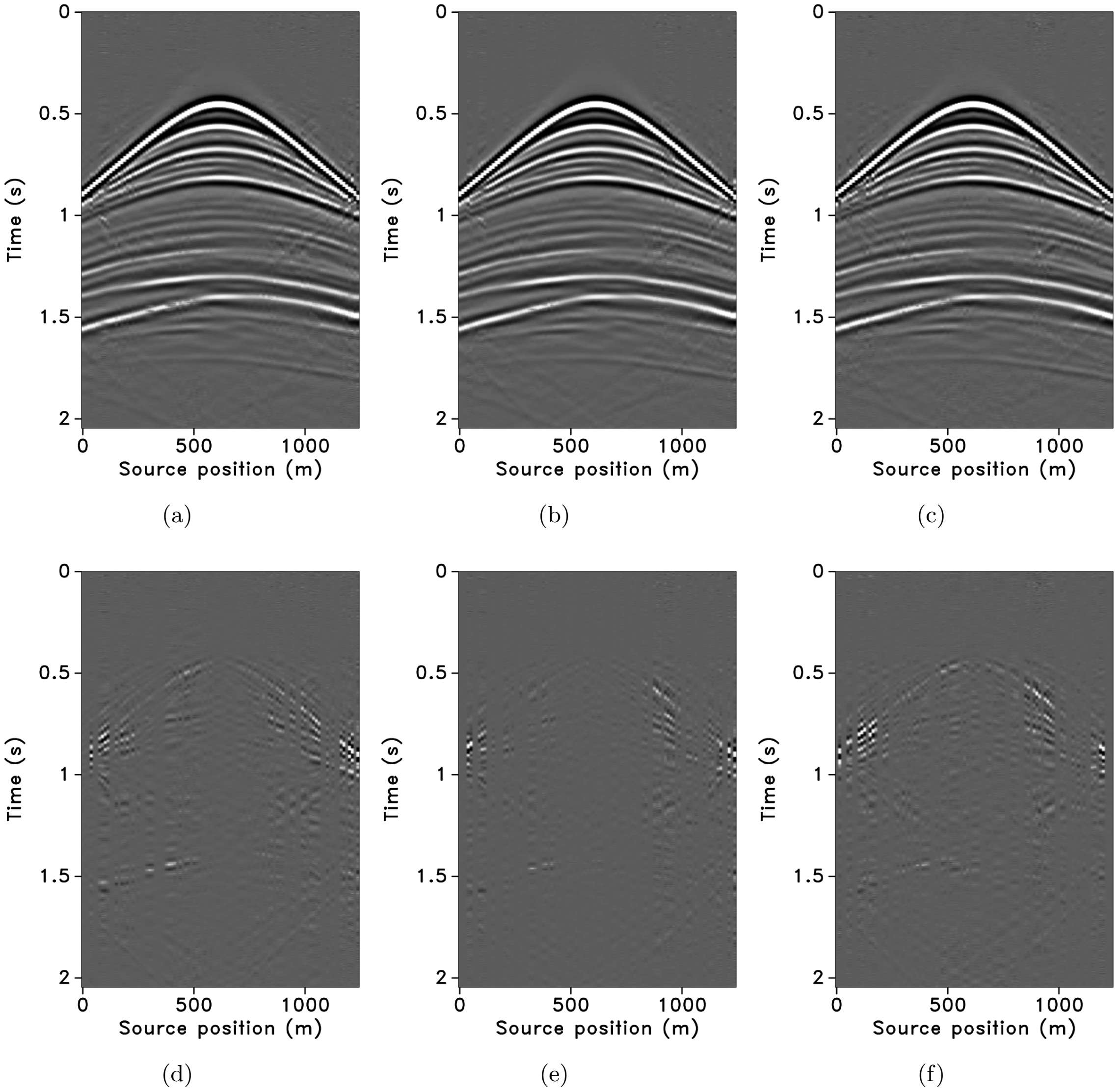
Remember, that the number of measurements is the same for all experiments and the observed differences can be fully attributed to the performance of the recovery method in relation to the overlap between the two surveys encoded in the measurement matrices. Also, the improvements in S/Ns of the vintages are significant as we lower the overlap, which goes at the expense of a relatively small loss in S/N for the time-lapse stack. However, given the context of randomized subsampling, it is important to recover the finely sampled vintages and then the time-lapse difference. In addition, time-lapse differences are often studied via differences in certain poststack attributes computed from the vintages, hence, reinforcing the importance of recovering prestack baseline and monitor data as opposed to recovering the time-lapse difference alone. While some degree of replication seemingly improves the prestack time-lapse difference, we feel that quality of the vintages themselves should prevail in the light of the above discussion. In addition, concentrating on the quality of the vintages gives us the option to compute prestack time-lapse differences in alternative ways (Wang et al., 2008).

All these observations are corroborated by the plots of the recovered (monitor) receiver gathers and their differences from the original (idealized) gather in Figures 5.10 and 5.11, and the recovered time-lapse differences in Figure 5.12. Stacked sections of the IRS and the JRM recovered time-lapse difference are shown in Figure 5.13.

### 5.6.4 Repeatability measure

Aside from measuring S/Ns, researchers have introduced repeatability measures expressing the similarity between prestack and poststack time-lapse datasets. One of the most commonly used metrics, which gives an intuitive understanding of the data repeatability, is the normalized root-mean-square (NRMS, Kragh and Christie, 2002):

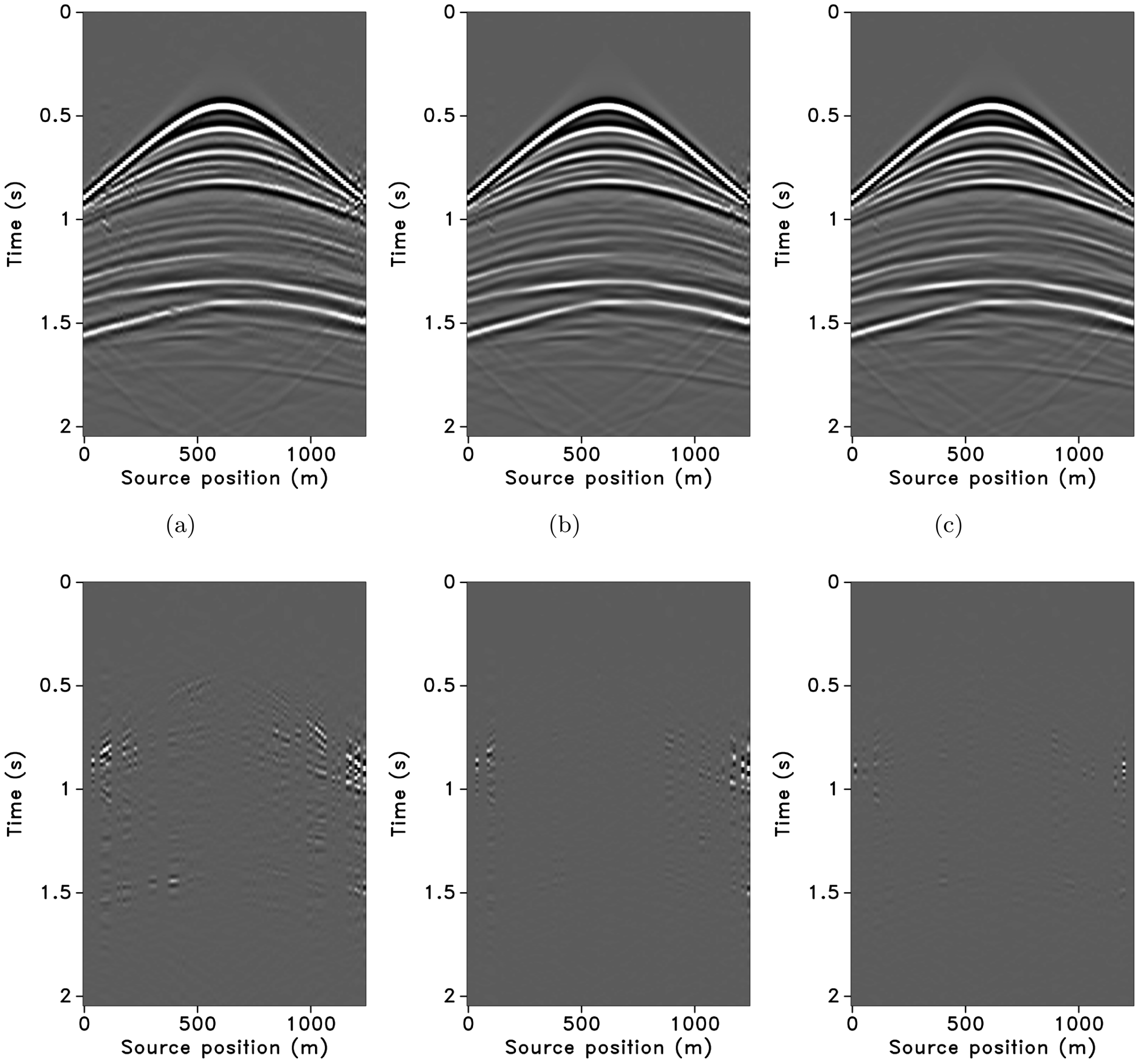
$$\mathrm{NRMS} = \frac{2\,\mathrm{RMS}(\tilde{\mathbf{f}}_2 - \widetilde{\mathbf{f}}_1)}{\mathrm{RMS}(\widetilde{\mathbf{f}}_1) + \mathrm{RMS}(\widetilde{\mathbf{f}}_2)}, \tag{5.10}$$



(a) (b) (c)

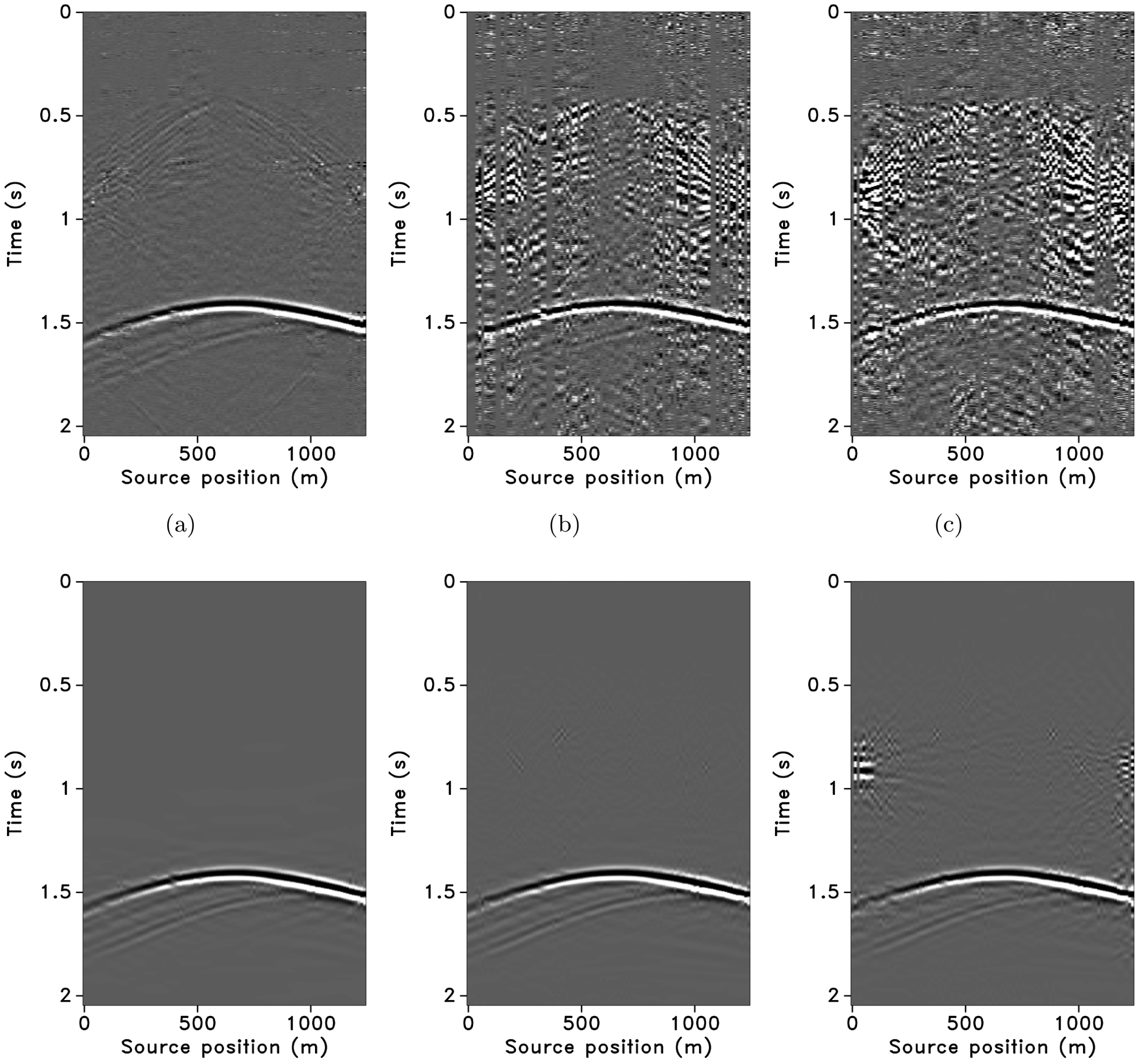
(d) (e) (f)

**Figure 5.10:** Receiver gathers (from monitor survey) recovered via IRS from time-jittered marine acquisition with (a) 100%, (b) 50%, and (c) 25% overlap in the measurement matrices ($\mathbf{A_1}$ and $\mathbf{A_2}$). (d), (e), and (f) Corresponding difference plots from the original receiver gather (5.7(b)).



(a)
(b)
(c)
(d)
(e)
(f)

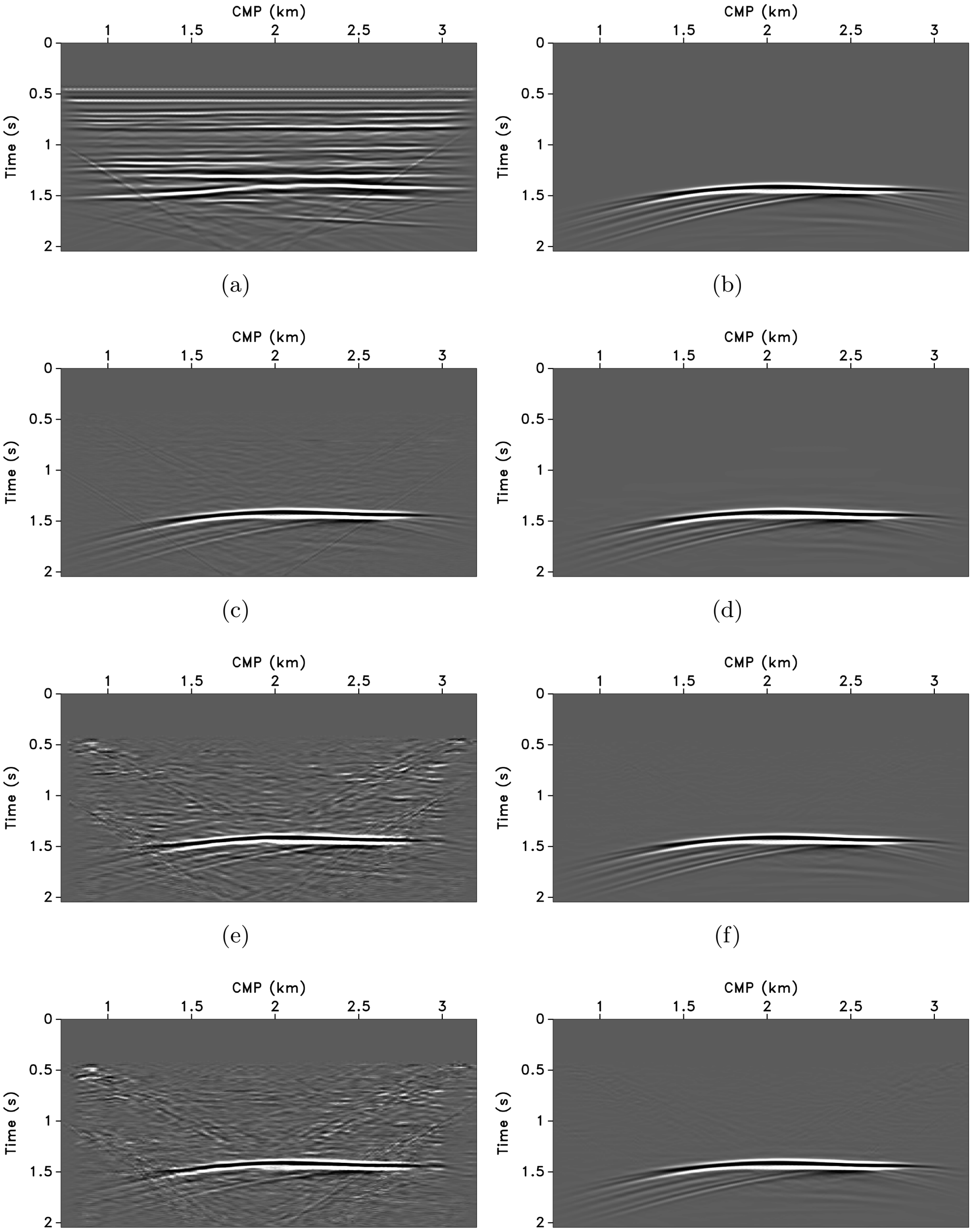
**Figure 5.11:** Receiver gathers (from monitor survey) recovered via JRM from time-jittered marine acquisition with (a) 100%, (b) 50%, and (c) 25% overlap in the measurement matrices ($\mathbf{A_1}$ and $\mathbf{A_2}$). (d), (e), and (f) Corresponding difference plots from the original receiver gather (5.7(b)).



(a) (b) (c)

(d) (e) (f)

**Figure 5.12:** Recovered 4D signal for the (a) 100%, (b) 50%, and (c) 25% overlap. Top row: IRS, bottom row: JRM. Note that the color axis is one-tenth the scale of the color axis for the vintages.



(a) (b) (c) (d) (e) (f) (g) (h)

**Figure 5.13:** Stacked sections. (a) baseline; (b) true 4D signal; reconstructed 4D signals via IRS for 100% (c), 50%(e), and 25% (g) overlap; the reconstructed 4D signals via JRM for 100%(d), 50%(f), and 25% (h) overlap. Notice the improvements for JRM where we see much less deterioration as the overlap between the surveys decreases. Note that the color axis for the time-lapse difference stacks is one-tenth the scale of the color axis for the baseline stack.

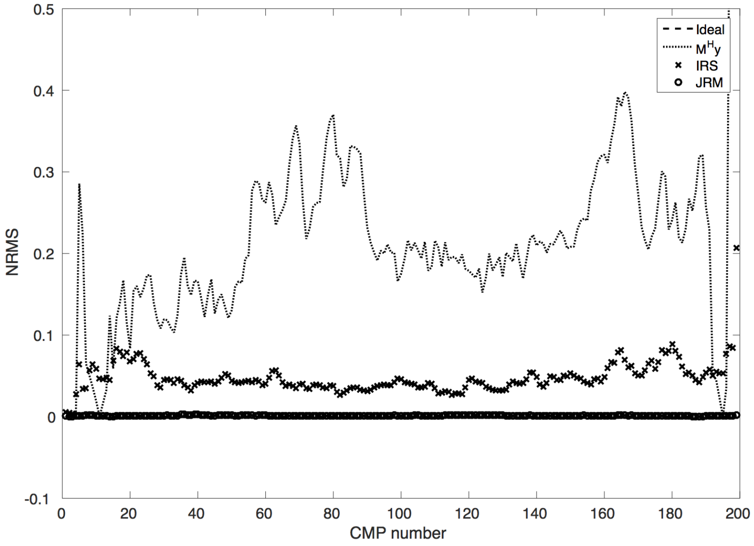
with $\mathrm{RMS}(\widetilde{\mathbf{f}})$ being the root-mean-square of either vintage. This formula implies that the lower the NRMS, the higher the repeatability between the recovered datasets. Usually, lower levels of NRMS are observed for stacked data compared to prestack data since stacking reduces uncorrelated random noise. A NRMS ratio of 0 is achievable only in a perfectly repeatable world. In practice, NRMS ratios between 0.2 and 0.3 are considered as acceptable; ratios less than 0.2 are considered excellent. To further evaluate the results of our synthetic seismic experiment, we compute the NRMS ratios from stacked sections before and after recovery via IRS and JRM.

To compute this quantity, we extract time windows from stacked sections around two-way travel time between 0.5 s and 1.3 s, where we know there is no time-lapse signal present. We obtain the stacked sections before and after processing by either applying the adjoint of the sampling matrix (see discussion under Equation 5.8) to the observed data or by solving a sparsity-promoting program. The former serves as a proxy for acquisition scenarios where one relies on the fold to stack out acquisition related artifacts. Results of this exercise for 50% overlap and 25% overlap are included in Figures 5.14(a) and 5.14(b). These plots clearly show that *(i)* simply applying the adjoint, followed by stacking, leads to poor repeatability, and therefore is unsuitable for time-lapse practices; *(ii)* sparse recovery improves the NRMS; *(iii)* exploiting shared information amongst the vintages leads to near optimal values for the NMRS despite the subsampling; and finally *(iv)* high degrees of repeatability of recovered data are achievable from data collected with small overlaps in the acquisition geometry.

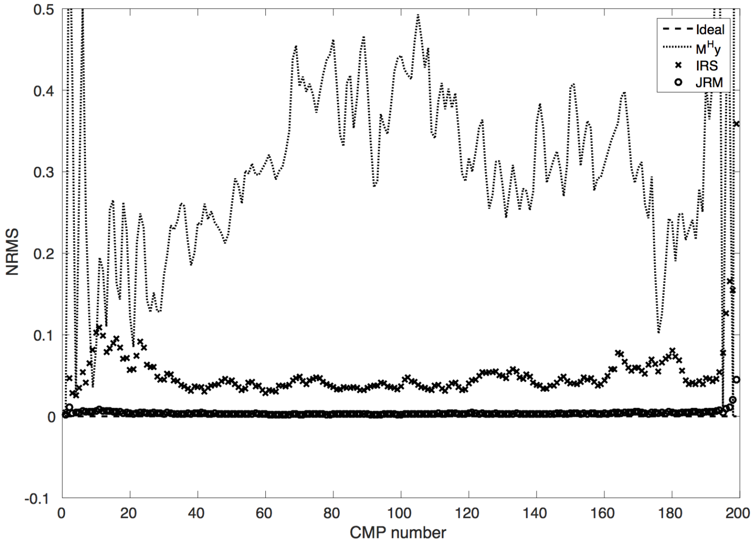
## 5.7 Discussion

Obtaining useful time-lapse seismic is challenging for many reasons, including cost, the need to calibrate the surveys, and the subsequent processing to extract reliable time-lapse information. Meeting these challenges in the field has resulted in acquisitions which aim to replicate the geometry of the previous survey(s) as precisely as possible. Unfortunately, this replication can be both difficult to achieve and expensive. Post acquisition, processing aims to improve the repeatability of the data such that certain (poststack) attributes can be derived reliably from the baseline and monitor surveys. Within this context, our aim is to reduce the cost and improve the quality of the prestack baseline and monitor data without relying on expensive fine sampling and high degrees of replicability of the surveys. Our methodology involves a combination of economical randomized samplings and sparsity-promoting data recovery. The latter exploits (curvelet-domain) sparsity and correlations amongst different vintages. To the authors' knowledge, this approach is among the first to address time-lapse seismic problems in which the common component amongst vintages—and innovations with respect to this shared component—is made explicit.

The presented synthetic seismic case study, supported by the findings from the stylized examples and theoretical results from the distributed compressive sensing literature (Baron et al., 2009), represents a proof of concept for how sharing information amongst the vintages can lead to high-fidelity vintages and 4D signals (with minor trade-offs) in a cost effective manner. This approach creates new possibilities for meeting modern survey objectives, including cost and environmental impact considerations, and improvements in spatial sampling. In this chapter, even though our measurements are taken on the grid, allowing us to ignore errors related to sampling off the grid, our proposed time-lapse acquisition is low-cost since we are always subsampled in the surveys. Our joint recovery model produces finely sampled data volumes from these subsampled, and not necessarily replicable, randomized surveys. These data volumes exhibit better repeatability levels (in terms of NRMS ratios) compared to independent recovery, where correlations amongst the



(a)



(b)

**Figure 5.14:** Normalized root-mean-squares NMRS for each recovered trace of the stacked section for *(a)* 50% and the *(b)* 25% overlap. Vintages obtained with the joint recovery method are far superior to results obtained with the independent recovery strategy and the "unprocessed" stacked data. The latter are unsuitable for time lapse.

vintages are not exploited.

In the next chapter, we demonstrate how we deal with the effects of non-replicability of the surveys when we take measurements from an irregular grid. We demonstrate that errors related to being off the grid cannot be ignored. The "bad news" is that replication is unattainable because small inevitable deviations in the shot locations amongst the time-lapse surveys negate the benefit of replication for the time-lapse signal itself. However, the good news is that a slightly deviated measurement already adds information that improves recovery of the vintages. This implies that an argument can be made to not replicate the surveys as long as we know sufficiently accurately where we fired in the field. Please remember that the claims of this chapter relate to the unnecessary requirement to visit the same randomly subsampled on-the-grid shot locations during the two, or more, surveys.

Furthermore, we did not consider surveys that have been acquired in situations where there are significant variations in the water column velocities amongst the different surveys. As long as these physical changes can be modelled, we do not foresee problems. As expected using standard CS, our recovery method should be stable with respect to noise (Candès et al., 2006c), but this needs to be investigated further. Moreover, recent successes in the application of compressive sensing to actual land and marine field data acquisition (see e.g. Mosher et al. (2014)) support the fact that these technical challenges with noise and calibration can be overcome in practice. Our future research will also involve working with towed-streamer surveys where other challenges like the sparse and irregular crossline sampling will be investigated.

In this study, we concentrated our efforts on producing high-quality baseline and monitor surveys from economic randomized acquisitions. There are areas of application for the joint recovery model that have not yet been explored in detail, such as imaging and full-waveform inversion problems. Early results on these applications suggest that our joint recovery model extends to sparsity-promoting imaging (Tu et al., 2011; Herrmann and Li, 2012) including imaging with surface-related multiples, and time-lapse full-waveform inversion (Oghenekohwo et al., 2015). In all applications, the use of shared information amongst vintages improves the inversion results even for acquisitions with large gaps. Finally, none of the other recently proposed approaches in this research area—e.g., double differences (Yang et al., 2014) and total-variation norm minimization on time-lapse earth models (Maharramov and Biondi, 2014)—use the shared information amongst the vintages explicitly.

## 5.8 Conclusions

We considered the situation of recovering time-lapse data from on-the-grid but randomly subsampled surveys. In this idealized setting, where we ignore the effects of being off the grid, we found that it is better not to revisit the on-the-grid shot locations amongst the time-lapse surveys when the vintages themselves are of prime interest. This result is a direct consequence of introducing a common component, which contains information shared amongst the vintages, as part of our proposed joint recovery method. Compared to independent recoveries of the vintages, we obtain time-lapse data exhibiting a higher degree of repeatability in terms of normalized root-mean-square ratios. Under the above stated idealized setting and ignoring complicating factors such as tidal differences, our proposed method lowers the cost and environmental imprint of acquisition because fewer shot locations are visited. It also allows us to extend the survey area or to increase the data's resolution at the same costs as conventional surveys. Our improvements concern the vintages and not the time-lapse difference itself, which would benefit if we choose to use the same shot locations during the surveys. Because we are generally interested in “poststack” attributes derived from the vintages, their recovery took prevalence. So, we make the argument not to replicate—i.e., revisit on-the-grid shot locations during randomized surveys in cases where poststack time-lapse attributes are of interest only.

# Chapter 6

# Low-cost time-lapse seismic with distributed Compressive Sensing — impact on repeatability

## 6.1 Summary

Irregular or off-the-grid spatial sampling of sources and receivers is inevitable in field seismic acquisitions. Consequently, time-lapse surveys become particularly expensive since current practices aim to replicate densely sampled surveys for monitoring changes occurring in the reservoir due to hydrocarbon production. We demonstrate that under certain circumstances, high-quality prestack data can be obtained from cheap randomized subsampled measurements that are observed from nonreplicated surveys. We extend our time-jittered marine acquisition to time-lapse surveys by designing acquisition on irregular spatial grids that render simultaneous, subsampled and irregular measurements. Using the fact that different time-lapse data share information and that nonreplicated surveys add information when prestack data are recovered jointly, we recover periodic densely sampled and colocated prestack data by adapting the recovery method to incorporate a regularization operator that maps traces from an irregular spatial grid to a regular periodic grid. The recovery method is, therefore, a combined operation of regularization, interpolation (estimating missing fine-grid traces from subsampled coarse-grid data), and source separation (unraveling overlapping shot records). By relaxing the insistence on replicability between surveys, we find that recovery of the time-lapse difference shows little variability for realistic field scenarios of slightly nonreplicated surveys that suffer from unavoidable natural deviations in spatial sampling of shots (or receivers) and pragmatic compressed-sensing based nonreplicated surveys when compared to the "ideal" scenario of exact replicability between surveys. Moreover, the recovered densely sampled prestack baseline and monitor data improve significantly when the acquisitions are not replicated, and hence can serve as input to extract poststack attributes used to compute time-lapse differences. Our observations are based on experiments conducted for an ocean-bottom cable survey acquired with time-jittered continuous recording assuming source equalization (or same source signature) for the time-lapse surveys and no changes in wave heights, water column velocities or temperature and salinity profiles, etc.

---

A version of this chapter has been published in Geophysics, 2017, vol. 82, pp. P15–P30.

## 6.2 Introduction

Simultaneous marine acquisition is being recognized as an economic and environmentally more sustainable way to sample seismic data and speedup acquisition, wherein single or multiple source vessels fire shots at random, compressed times resulting in overlapping shot records (de Kok and Gillespie, 2002; Beasley, 2008; Berkhout, 2008; Hampson et al., 2008; Moldoveanu and Quigley, 2011; Abma et al., 2013), and hence generating compressed seismic data volumes. The aim then is to separate the overlapping shot records into individual shot records, as acquired during conventional acquisition, but with denser source sampling while preserving amplitudes of the late, often weak, arrivals. This leads to recovering densely sampled data economically, which is essential for producing high-resolution images of the subsurface.

Mansour et al. (2012b), Wason and Herrmann (2013b) and Mosher et al. (2014) have showed that compressed sensing (CS, Candès et al., 2006c; Donoho, 2006) is a viable technology to sample seismic data economically with low environmental imprint—by reducing numbers of shots (or injected energy in the subsurface) or compressing survey times. Mansour et al. (2012b) and Wason and Herrmann (2013b) proposed an alternate sampling strategy for simultaneous acquisition ("time-jittered" marine), addressing the separation problem through a combination of tailored (simultaneous) acquisition design and sparsity-promoting recovery via convex optimization using $\ell_1$ objectives. This separation technique interpolates sub-Nyquist jittered shot positions to a fine regular grid while unraveling the overlapping shots. The time-jittered marine acquisition is designed for continuous recording, fixed-receiver (or "static") geometries, which is different from the case of towed-streamer (or "dynamic") geometries, wherein multiple sources fire shots within a time interval of $(0,1)$ or $(0,2)$ s generating overlapping shot records that need to be separated into its constituent sources, i.e., a data volume for each individual source (Kumar et al., 2015b). Our approach for conventional data recovery from simultaneous data from static geometries can equally apply to other settings including static land and other static marine geometries.

The implications of randomization in time-lapse (or 4D) seismic, however, are less well-understood since the current paradigm relies on dense sampling and replicability amongst the baseline and monitor surveys (Lumley and Behrens, 1998). These requirements impose major challenges because the insistence on dense sampling may be prohibitively expensive and variations in acquisition geometries (between the surveys) due to physical constraints do not allow for exact replication of the surveys. In Chapter 5, we presented a new approach (the "joint recovery method") that addresses these acquisition- and processing-related issues by explicitly exploiting common information shared by the different time-lapse vintages. Our analyses were carried out assuming that the observations lied on a discrete grid so that exact survey replicability is in principle achievable. We also assume sources to have the same source signature for the time-lapse surveys. While assuming source equalization in this chapter, we extend our work on simultaneous time-jittered marine acquisition to time-lapse surveys for more realistic field acquisitions that lie on irregular spatial grids, where the notion of exact replicability of the surveys is inexistent. This is because the "real" world suffers from unavoidable deviations between pre- and post-acquisition shot (and receiver) positions, rendering regular, periodic spatial grids irregular, and hence exact replication of the surveys impossible. As mentioned later in the chapter, accounting for the irregularity of seismic data is key to recovering densely sampled data. Moreover, while we do not insist that we actually visit predesigned (irregular) shot positions, but it is important to know these positions to sufficient accuracy after acquisition for high-quality data recovery. Recent successes in the application of compressed sensing to land and marine field data acquisition (see e.g., Mosher et al., 2014) show that this can be achieved in practice.

Simultaneous time-jittered marine acquisition generates compressed and subsampled data volumes, therefore, extending this to time-lapse surveys generates compressed and subsampled baseline and monitor data. Consequently, we are interested in recovering densely sampled vintages and time-lapse difference. Moreover, time-lapse differences are often studied via differences in certain poststack attributes computed from the vintages (Landrø, 2001; Spetzler and Kvam, 2006), hence, we prioritize on recovering the prestack vintages. In this chapter, we push this technology to realistic settings of off-the-grid acquisitions and demonstrate that we actually gain if we relax the insistence to replicate surveys since even the smallest known deviations from the grid can lead to significant improvements in the recovery of the vintages with minimal compromise with the recovery of the time-lapse difference.

### 6.2.1 Motivation: on-the-grid vs. off-the-grid data recovery

Chapter 5 illustrated that the joint recovery method gives better recoveries of time-lapse data and time-lapse difference than the independent recovery strategy, since the former approach exploits the common information shared by the vintages. It also showed that "exact" replication of the baseline and monitor surveys lead to good recovery of the time-lapse difference but not of the vintages. These analyses, however, were carried out assuming that the observations lied on a discrete grid so that exact survey replicability is achievable. Realistic field acquisitions, on the contrary, lie off the grid—i.e., have irregular spatial sampling—where exact replicability of the surveys is inexistent. Figure 6.1 shows a comparison between conventional periodic acquisition which generates data with nonoverlapping shot records, and simultaneous time-jittered acquisition which generates compressed recordings with overlapping shots. Note that the sampling grid for conventional acquisition "in the field" would be slightly irregular, however, this in contrast to the jittered acquisition which by virtue of its design is aperiodic and lies on an irregular sampling grid. Since the time-jittered acquisition scheme leverages compressed sensing—the success of which hinges on randomized subsampling—additional and unavoidable deviations in the field add to the randomization of the designed irregular shot positions, and helps in sparsity-promoting inversion as long as we know the final shot positions to sufficient precision.

Figures 6.2(a)-6.2(c) show receiver gathers from a conventional (synthetic) time-lapse data set and the corresponding time-lapse difference. To recover periodic densely sampled data from simultaneous, compressed and irregular data, we could implicitly rely on binning, however, failure to account for irregularity of seismic traces can adversely affect the recovery as shown in Figure 6.3. This is because binning does not represent accurate positions of irregular traces. Note that this example corresponds to a time-jittered acquisition scheme for the baseline that is exactly replicated for the monitor. The results show that binning offsets all the gains of exact survey replication and also of the joint recovery method. Figure 6.4 illustrates the importance of regularization of irregular traces for high-quality data recovery. In this chapter, therefore, we extend our work on simultaneous time-jittered acquisition to time-lapse surveys by acknowledging the irregularity of field seismic data and incorporating sparsifying transforms that exploit this irregularity to recover periodic densely sampled time-lapse data.

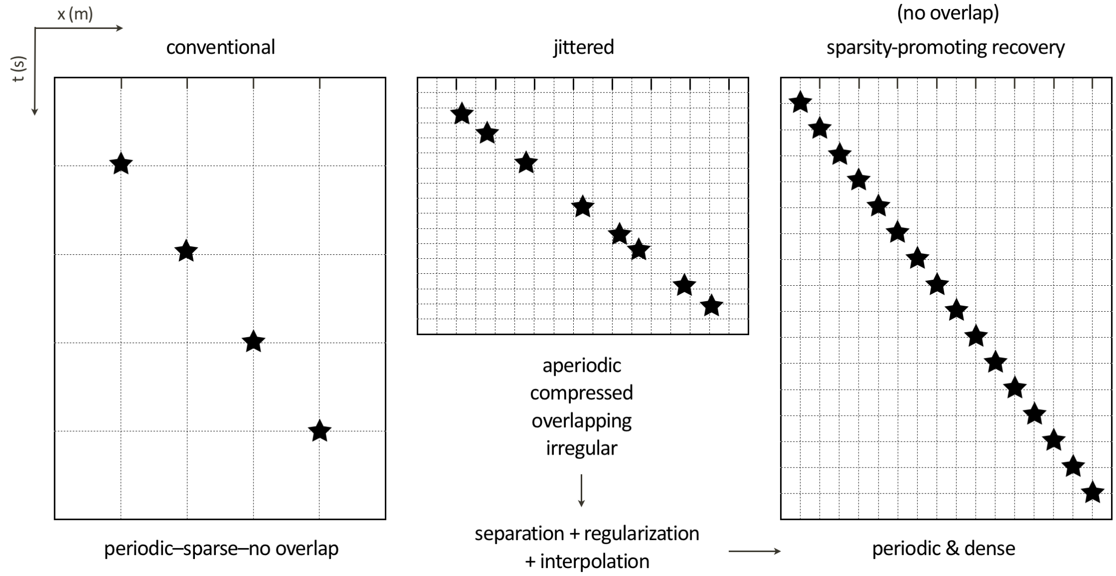


**Figure 6.1:** Schematic of conventional acquisition and simultaneous, compressed (or time-jittered) acquisition. If the source sampling grid for conventional acquisition is 25.0 m (or 50.0 m for flip-flop acquisition), then the time-jittered acquisition jitters (or perturbs) shot positions on a finer grid, which is 1/4 th of the conventional flip-flop sampling grid, for a single air-gun array. Following the same strategy, adding another air-gun array makes the acquisition simultaneous, and hence results in a compressed data volume with overlapping, irregular shots and missing traces. The sparsity-promoting inversion then aims to recover densely sampled data by separating the overlapping shots, regularizing irregular traces and interpolating missing traces.

### 6.2.2 Contributions

The contributions of this work can be summarized as follows. First, we present an extension of our simultaneous time-jittered marine acquisition for time-lapse surveys by working on more realistic field acquisition scenarios by incorporating irregular spatial grids. Second, we leverage ideas from compressed sensing and distributed compressed sensing to develop an algorithm that separates simultaneous data, regularizes irregular traces and interpolates missing traces—all at once. Third, through simulated experiments, we show that insistence on replicability of time-lapse surveys can be relaxed since small known deviations in shot positions from a regular grid (or deviations in shot positions of the monitor survey from those in the baseline survey) lead to significant improvements in the recovery of the vintages, without drastic modifications in the recovery of the time-lapse difference.

### 6.2.3 Outline

The chapter is organized as follows. We begin with the description of the simultaneous time-jittered marine acquisition design, where we explain how subsampled and irregular measurements are generated. Next, we introduce the nonequispaced fast discrete curvelet transform (NFDCT) and its application to recover periodic densely sampled seismic lines from simultaneous and irregular

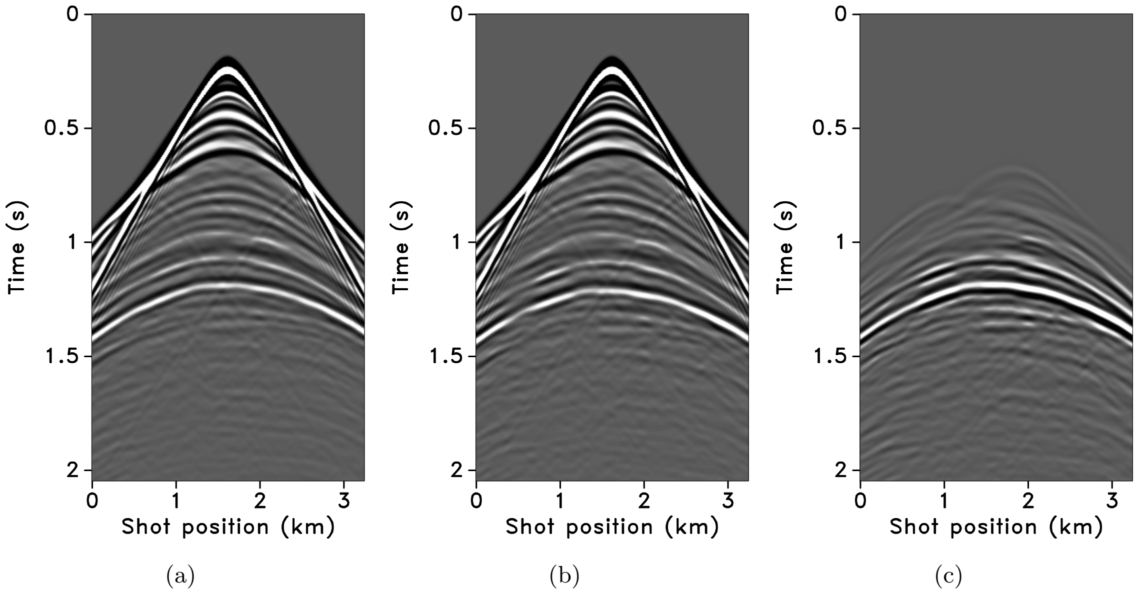
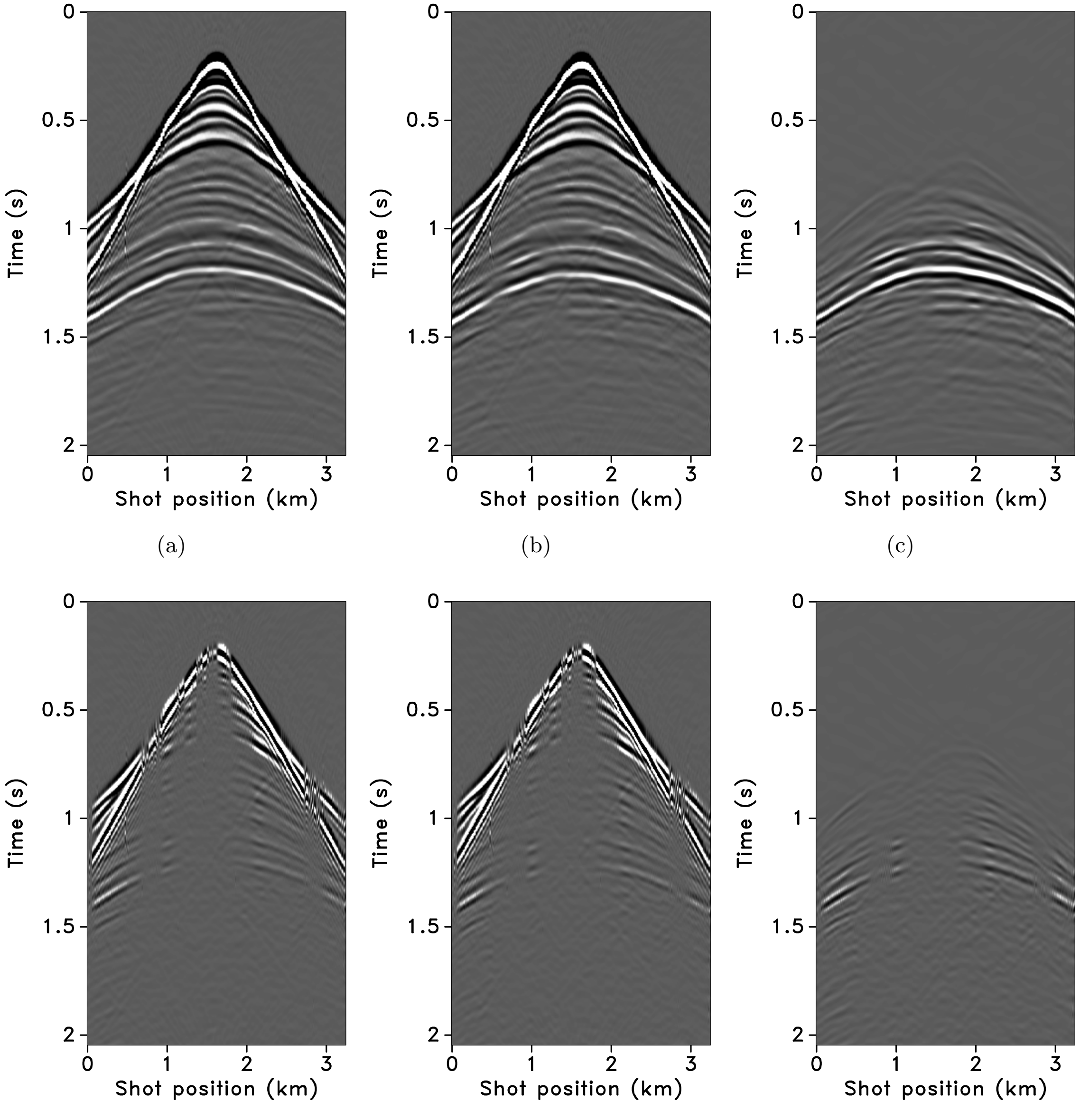


**Figure 6.2:** Synthetic receiver gathers from a conventional (a) baseline survey, (b) monitor survey. (c) Corresponding time-lapse difference.

measurements via sparsity-promoting inversion. We then extend this framework to time-lapse surveys where we modify the measurement matrices in the joint recovery method to include the off-the-grid information—i.e., the irregular shot positions and jittered times. Note that we do not describe the independent recovery strategy since it is clear in Chapter 5 that the joint recovery method outperforms the former approach. We conduct a series of synthetic seismic experiments with different random realizations of the simultaneous time-jittered marine acquisition to assess the effects (or risks) of irregular sampling in the field on time-lapse data and demonstrate that high-quality data recoveries are the norm and not the exception. We show this by generating 2D seismic lines using two different velocity models—one with simple geology and complex time-lapse difference (BG COMPASS model), and the other with complex geology and complex time-lapse difference (SEAM Phase 1 model with simulated time-lapse difference). Aside from computing signal-to-noise ratios measured with respect to densely sampled true baseline, monitor, and time-lapse differences, we also measure the economic and environmental performance of the proposed acquisition design and recovery strategy by computing the improvement in spatial sampling.

## 6.3 Time-jittered marine acquisition

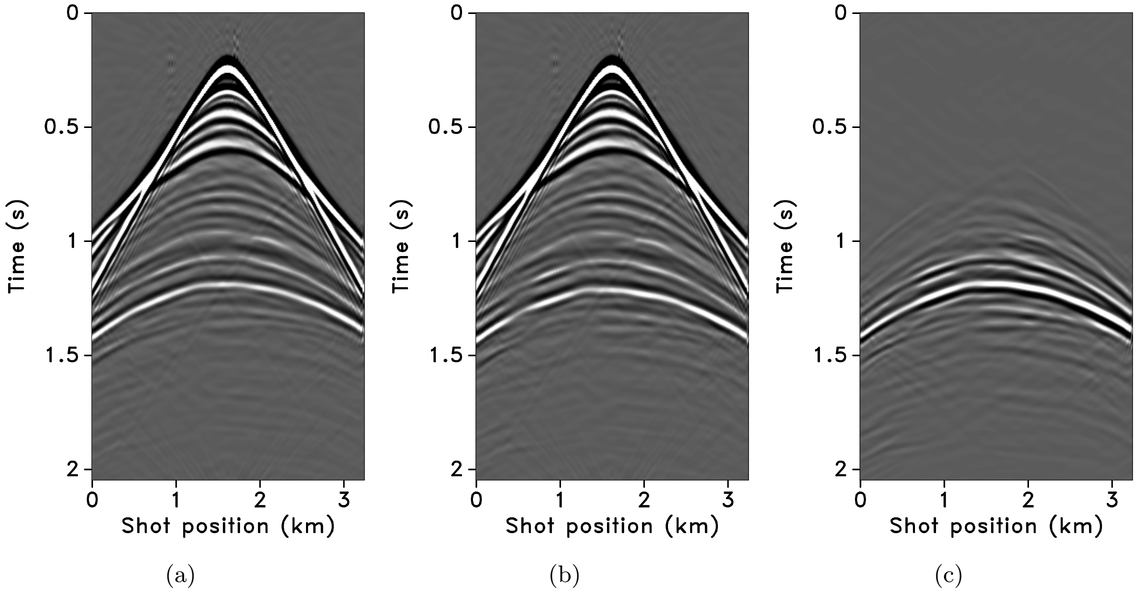
The objective of CS is to recover densely sampled data from (randomly) subsampled data by exploiting sparse structure in the data during sparsity-promoting recovery. Mansour et al. (2012b), Wason and Herrmann (2013b) presented a pragmatic simultaneous marine acquisition scheme, termed as time-jittered marine, that leverages ideas from compressed sensing by invoking randomness and subsampling—i.e., sample randomly with fewer samples than required by Nyquist sampling criteria in the acquisition via random jittering of the source firing times. The success of CS hinges on randomized subsampling since it renders subsampling related artifacts incoherent, and therefore



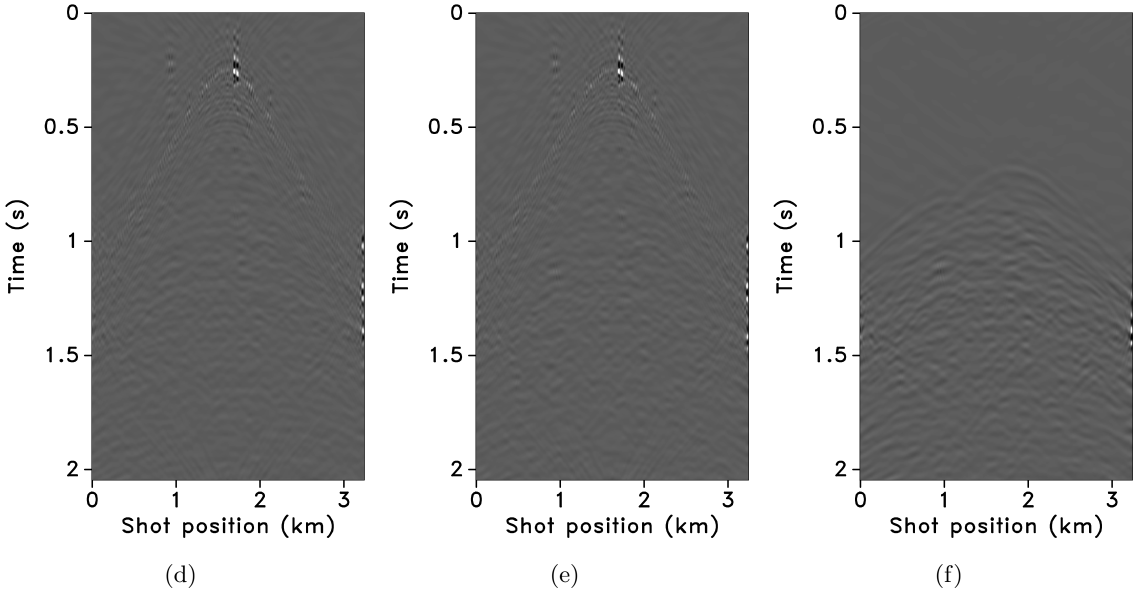
(a) (b) (c)

(d) (e) (f)

**Figure 6.3:** Data recovery via the joint recovery method and binning. (a), (b) Binned vintages and (c) corresponding time-lapse difference. (d), (e), (f) Corresponding difference plots.



(a) (b) (c)



(d) (e) (f)

**Figure 6.4:** Data recovery via the joint recovery method and regularization. (a), (b) Vintages and (c) time-lapse difference recovered via sparsity promotion including regularization of irregular traces. (d), (e), (f) Corresponding difference plots. As illustrated, regularization is imperative for high-quality data recovery.

nonsparse, favouring data recovery via structure-promoting inversion. Consequently, source interferences (in simultaneous acquisition) become incoherent in common-receiver gathers creating a favorable condition for separating simultaneous data into conventional nonsimultaneous data via curvelet-domain sparsity promotion. The CS paradigm, however, assumes signals to be sampled on a periodic discrete grid—i.e., signals with sparse representation in finite discrete dictionaries.

Data volumes collected during seismic acquisition represent discretization of analog finite-energy wavefields in up to five dimensions including time—i.e., we acquire an analog spatiotemporal wavefield $\bar{f}(t,x) \in L^2((0,T] \times [-X,X]^4)$, two dimensions for receivers and two dimensions for sources, with time $T$ in order of seconds and length $X$ in order of kilometers. In an ideal world, signals would perfectly lie on a periodic, regular grid. Hence, with a linear high-resolution analog-to-digital converter $\bar{\Phi}_s$, the discrete signal is represented as $f[q] = \bar{f} \star \bar{\Phi}_s(q)$, for $0 \leq q < N$ (Mallat, 2008), where the samples lie on a grid. Typically, these samples are organized into a vector $\mathbf{f} = f[q]_{q=0,\dots,N-1} \in \mathbb{R}^N$. Signals we encounter in the real world, however, are usually not uniformly regular and do not lie on a regular grid. Therefore, it is imperative to define an irregular sampling adapted to the local signal regularity (Mallat, 2008). For irregular sampling, the discretized irregular signal is represented as $f[q_n] = \bar{f} \star \bar{\Phi}_s(q_n)$, for $n = 0, ..., M-1$ and $M \leq N$, where $q_n$ are irregular points (or nonequispaced nodes) randomly chosen from the set $\{0, ..., N-1\}$. Its vector representation is $\mathbf{f} = f[q_n]_{n=0,\dots,M-1}$.

For a signal $\mathbf{f}_0 \in \mathbb{R}^N$ that admits a sparse representation $\mathbf{x}_0$ in some transform domain—i.e., $\mathbf{f}_0$ is sparse with respect to a basis or redundant frame $\mathbf{S} \in \mathbb{C}^{P \times N}$, with $P \geq N$, such that $\mathbf{f}_0 = \mathbf{S}^H \mathbf{x}_0$ ($\mathbf{x}_0$ sparse), where $^H$ denotes the Hermitian transpose—the goal in CS is to reconstruct the signal $\mathbf{f}_0$ from few random linear measurements, $\mathbf{y} = \mathbf{A}\mathbf{f}_0$, where $\mathbf{A}$ is an $n \times N$ measurement matrix with $n \ll N$. Utilizing prior knowledge that $\mathbf{f}_0$ is sparse with respect to a basis or redundant frame $\mathbf{S}$ and assuming the signal to be sampled on a periodic discrete grid, CS aims to find an estimate $\tilde{\mathbf{x}}$ for the underdetermined system of linear equations: $\mathbf{y} = \mathbf{A}\mathbf{f}_0$. This is done by solving the basis pursuit (BP, Candès et al., 2006c; Donoho, 2006) convex optimization problem:

$$\tilde{\mathbf{x}} = \arg\min_{\mathbf{x}} \|\mathbf{x}\|_1 := \sum_{i=1}^{N} |x_i| \quad \text{subject to} \quad \mathbf{y} = \mathbf{A}\mathbf{x}. \tag{6.1}$$

In the noise-free case, this problem finds amongst all possible vectors $\mathbf{x}$, the vector that has the smallest $\ell_1$-norm and that explains the observed subsampled data.

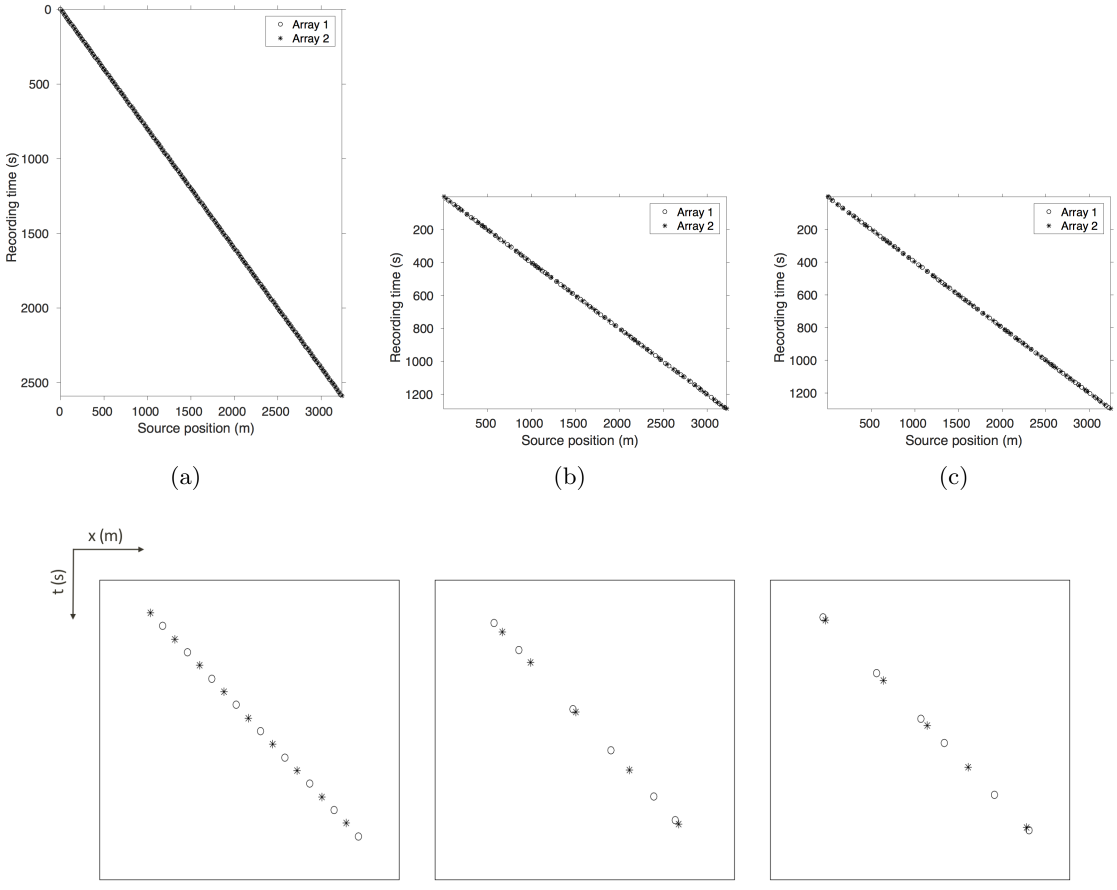
Mathematically, a seismic line with $N_s$ sources, $N_r$ receivers, and $N_t$ time samples can be reshaped into an $N$ dimensional vector $\mathbf{f}$, where $N = N_s \times N_r \times N_t$. Since real-world signals are not exactly sparse but compressible—i.e., can be well approximated by a sparse signal—a compressible representation, $\mathbf{x}$, of the seismic line in the curvelet domain, $\mathbf{S}$, is represented as $\mathbf{f} = \mathbf{S}^H\mathbf{x}$. Since curvelets are a redundant frame (an over-complete sparsifying dictionary), $\mathbf{S} \in \mathbb{C}^{P \times N}$ with $P > N$, and $\mathbf{x} \in \mathbb{C}^P$. With the inclusion of the sparsifying transform, the measurement matrix $\mathbf{A}$ can be factored into the product of a $n \times N$ (with $n \ll N$) acquisition matrix $\mathbf{M}$ and the synthesis matrix $\mathbf{S}^H$—i.e., $\mathbf{A} = \mathbf{M}\mathbf{S}^H$. For the real-world irregular signals, however, we need to account for the acquired unstructured measurements for high-resolution data recovery. We do this by introducing an operator in the recovery algorithm (by modifying the measurement operator $\mathbf{A}$—see details in the next sections) that acknowledges the irregularity of seismic traces and uses this information to render regularized and interpolated data.

### 6.3.1 Acquisition geometry

In time-jittered marine acquisition, source vessels map the survey area firing shots at jittered time instances, which translate to jittered shot positions for a given (fixed) speed of the source vessel. The simultaneous data is time compressed, and therefore acquired economically with low environmental imprint. The recovered separated data is periodic and dense. For simplicity, we assume that all shot positions see the same receivers, which makes our method applicable to marine acquisition with ocean bottom cables or nodes (OBC or OBN). The receivers record continuously resulting in simultaneous shot records. Randomization via jittered subsampling offers control over the maximum gap size (on the acquisition grid), which is a practical requirement of wavefield reconstruction with localized sparsifying transforms such as curvelets (Hennenfent and Herrmann, 2008). For simultaneous time-jittered acquisition, parameters such as the minimum distance required between adjacent shots and minimum recharge time for the air guns help in controlling the maximum acquisition gap while maintaining the minimum realistic acquisition gap.

Conventional acquisition with one source vessel and two air-gun arrays where each air-gun array fires at every alternate periodic location is called flip-flop acquisition. If we wish to acquire 10.0 s-long shot records at every 12.5 m, the speed of the source vessel would have to be about 1.25 m/s ($\approx$ 2.5 knots). Figure 6.5(a) illustrates one such conventional acquisition scheme, where each air-gun array fires every 20.0 s (or 25.0 m) in a flip-flop manner traveling at about 1.25 m/s, resulting in nonoverlapping shot records of 10.0 s every 12.5 m. In time-jittered acquisition, Figures 6.5(b) and 6.5(c), each air-gun array fires on average at every 20.0 s jittered time-instances traveling at about 2.5 m/s ($\approx$ 5.0 knots) with the receivers (OBC/OBN) recording continuously, resulting in overlapping shot records (Figures 6.6(a) and 6.6(b)). Note that the acquisition design involves jittered subsampling—i.e., regular decimation of the (fine) interpolation grid and subsequent perturbation of the coarse-grid points completely *off* the fine grid. The idealized discrete jittered subsampling, by contrast, perturbs the coarse-grid points on the fine grid, as presented in Chapter 5. The subsampling factor is represented by $\eta$. Therefore, the acquired data volume has overlapping shots and missing shots/traces (Figure 6.6(a) and 6.6(b)). For this reason, the jittered flip-flop acquisition might not mimic the conventional flip-flop acquisition where air-gun array 1 and 2 fire one after the other—i.e., in the center and right-hand plots of Figure 6.5(d) a circle (denoting array 1) may be followed by another circle instead of a star (denoting array 2), and vice versa. However, the minimum interval between the jittered times is maintained at 10.0 s (typical interval required for air-gun recharge), while the maximum interval is 30.0 s. For the speed of 2.5 m/s, this translates to jittering a 50.0 m source grid with a minimum (and maximum) interval of 25.0 m (and 75.0 m) between jittered shots. Both arrays fire at the 50.0 m jittered grid independent of each other.

In time-jittered marine acquisition, the acquisition operator $\mathbf{M}$ is a combined jittered-shot selector and time-shifting operator. Since data is acquired on an irregular grid, it is imperative to incorporate operators in the design of the acquisition matrix $\mathbf{M}$ that account for and hence regularize the irregularity in the data. This is critical to the success of the recovery algorithm. The off-the-grid acquisition design is different from that presented by Li et al. (2012), wherein an *interpolated restriction operator* accounts for irregular points by incorporating Lagrange interpolation into the restriction operator—i.e., the measurements are approximated using a $k^{\text{th}}$-order Lagrange interpolation. In time-jittered acquisition, the jittered time instances are put on a time grid (defined by a time-sampling interval) where each jittered time instance is placed on the point closest to it on the regular time grid. The difference between the true jittered time and the regular grid point,



(a)

(b)

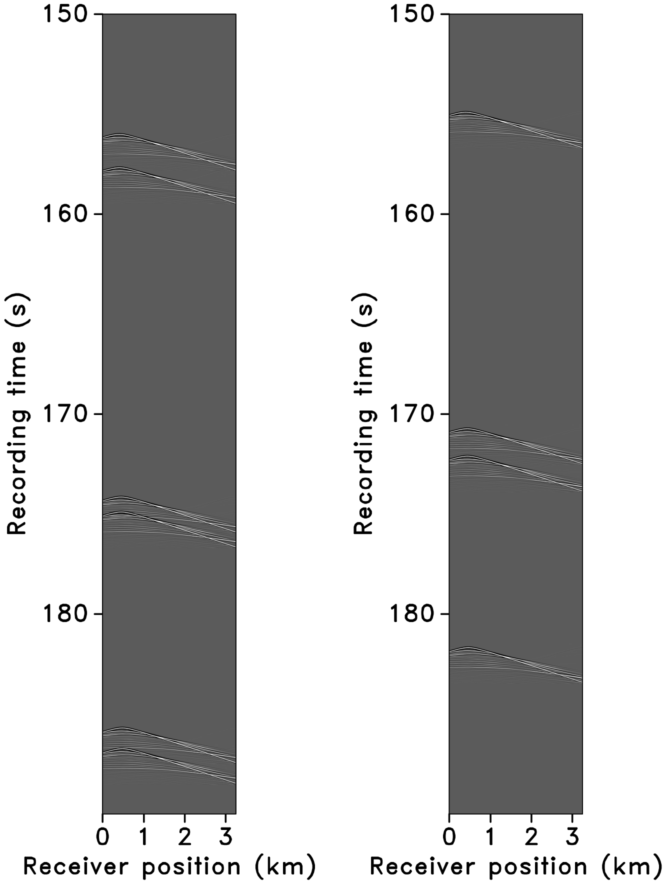
(c)

(d)

**Figure 6.5:** Marine acquisition with one source vessel and two air-gun arrays. (a) Conventional flip-flop acquisition. Time-jittered acquisition with a subsampling factor $\eta = 2$ for the (b) baseline and (c) monitor. Note the acquisition speedup during jittered acquisition, where the recording time is reduced to one-half the recording time of the conventional acquisition. (d) Zoomed sections of (a), (b) and (c), respectively.

$\Delta t$, is corrected by shifting the traces by $e^{-i\omega\Delta t}$, where $\omega$ is the angular frequency. The irregularity in the shot positions is corrected by including the nonequispaced fast Fourier transform, NFFT (Potts et al., 2001; Kunis, 2006), in the sparsifying operator $\mathbf{S}$ (Hennenfent and Herrmann, 2006; Hennenfent et al., 2010), as described in the next section. The NFFT evaluates a Fourier expansion at nonequispaced locations defined by the time-jittered acquisition. Note that in this framework it is also possible to randomly subsample the receivers.

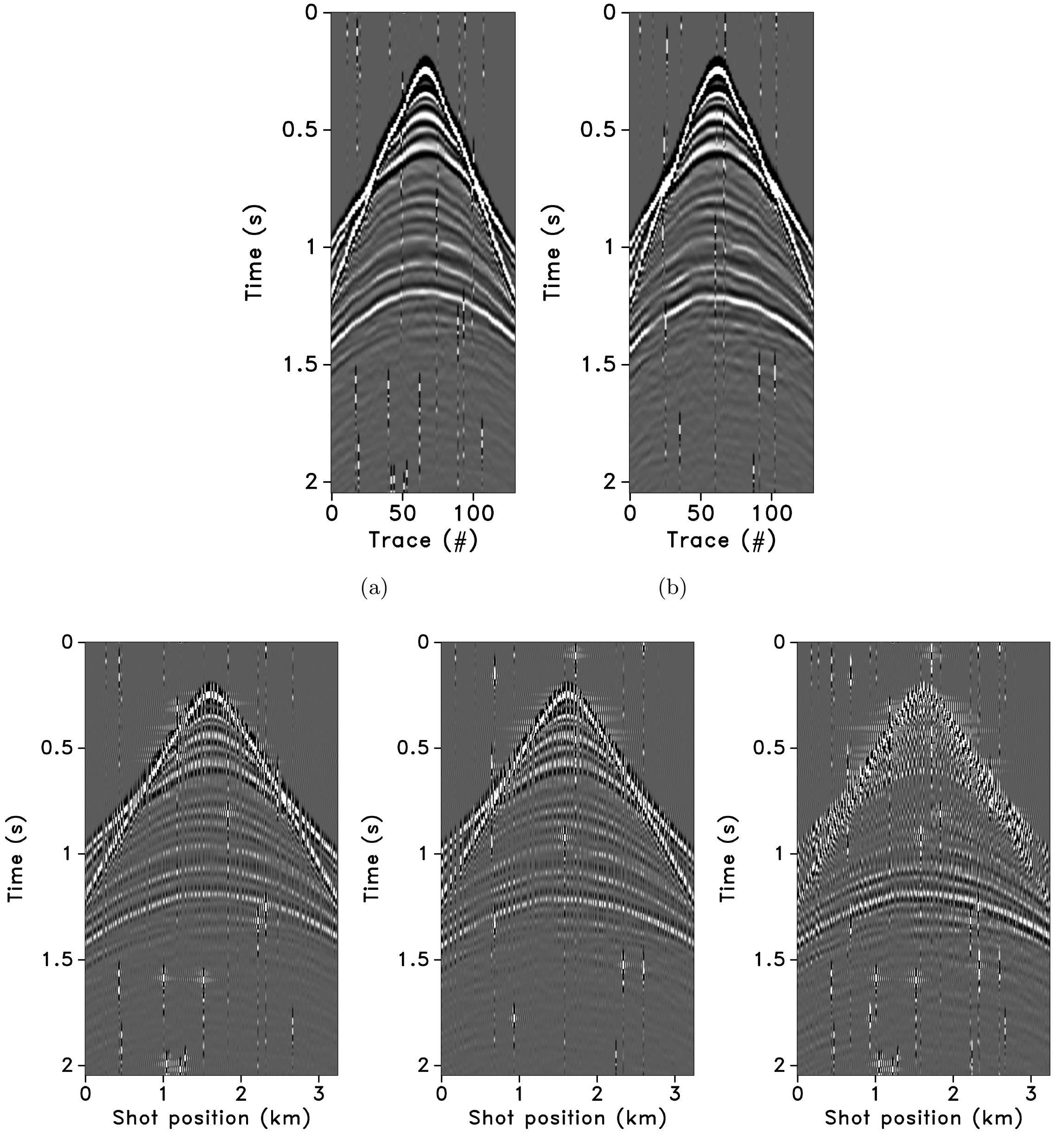
Randomly subsampled and simultaneous measurements for the baseline and monitor surveys are shown in Figures 6.6(a) and 6.6(b), respectively. Note that only 40.0 s of the continuously recorded data is shown. If we simply apply the adjoint of the acquisition operator to the corresponding simultaneous data—i.e., $\mathbf{M}^H\mathbf{y}$—the interferences (or source crosstalk) due to overlapping shots appear as incoherent and nonsparse in the receiver gathers (Figures 6.7(a) and 6.7(b)). Moreover,



(a) (b)

**Figure 6.6:** Simultaneous data for the (a) baseline and (b) monitor surveys. Only 40.0 s of the full data is shown. Time-jittered acquisition generates a simultaneous data volume with overlapping shots and missing shots.

since regularization (and interpolation) is performed by the NFFT inside a nonequispaced curvelet framework (see next section), Figures 6.7(a) and 6.7(b) have $\frac{N_s}{\eta}$ irregular traces, where $\eta > 1$ is the subsampling factor. Since the baseline and monitor surveys have different irregular shot positions, the corresponding time-lapse difference cannot be computed unless both time-lapse data are realigned to a common spatal grid. For this purpose, if we apply the adjoint of a 1D NFFT operator $\mathbf{N}$—i.e., $\mathbf{N}^H\mathbf{M}^H\mathbf{y}$—the time-lapse data are realigned to a common fine spatial grid (Figures 6.7(c) and 6.7(d)). The corresponding time-lapse difference is shown in Figure 6.7(e). As illustrated by these figures, in order to eventually remove the interferences and interpolate missing traces it is important to consider the recovery problem as an inversion problem. Since the time-jittered acquisition generates simultaneous, irregular data with missing traces, the recovery problem becomes a joint source separation, regularization and interpolation problem.



(a) (b)

(c) (d) (e)

**Figure 6.7:** Interferences (or source crosstalk) in a common-receiver gather for the (a) baseline and (b) monitor surveys, respectively. Receiver gathers are obtained via $\mathbf{M}^H\mathbf{y}$ for the time-lapse surveys. For a subsampling factor $\eta$, (a) and (b) have $\frac{N_s}{\eta}$ irregular traces. (c), (d) Common-receiver gathers for the baseline and monitor surveys, respectively, after applying the adjoint of a 1D NFFT operator to (a) and (b). (e) Corresponding time-lapse difference. As illustrated, the recovery problem needs to be considered as a (sparse) structure-promoting inversion problem, wherein the simultaneous data volume is separated, regularized and interpolated to a finer sampling grid rendering interference-free data.

### 6.3.2 From discrete to continuous spatial subsampling

Subsampling schemes that are based on an underlying fine interpolation grid incorporate the discrete (spatial) subsampling schemes, since the subsampling is done on the grid. This situation typically occurs when binning continuous randomly-sampled seismic data into small bins that define the fine grid used for interpolation (Hennenfent and Herrmann, 2008). For such cases, the wrapping-based fast discrete curvelet transform, FDCT via wrapping (Candès et al., 2006a) can be used to recover the fully sampled data since the inherent fast Fourier transform (FFT) assumes regular sampling along all coordinates. For the interested reader, the curvelet transform is a multiscale, multidirectional, and localized transform that corresponds to a specific tiling of the f-k domain into dyadic annuli based on concentric squares centered around the zero-frequency zero-wavenumber point. Each annulus is subdivided into parabolic angular wedges—i.e., length of wedge $\propto$ width$^2$ of wedge. The architecture of the analysis operator (or forward operation) of the FDCT via wrapping is as follows: (1) apply the analysis 2D FFT; (2) form the angular wedges; (3) wrap each wedge around the origin; and (4) apply the synthesis 2D FFT to each wedge. The synthesis/adjoint operator—also the inverse owing to the tight-frame property—is computed by reversing these operations (Candès et al., 2006a).

Seismic data, however, is usually acquired irregularly, typically nonuniformly sampled along the spatial coordinates. Simultaneous time-jittered marine acquisition, mentioned above, is an instance of acquiring seismic data on irregular spatial grids. Hence, binning can lead to a poorly-jittered subsampling scheme, which will not favor wavefield reconstruction by sparsity-promoting inversion. Moreover, failure to account for the nonuniformly sampled data can adversely affect seismic processing, imaging, etc. Therefore, we should work with an extension to the curvelet transform for irregular grids (Hennenfent et al., 2010). Using this extension for the simultaneous time-jittered marine acquisition will produce colocated fine-scale time-lapse data. Continuous random sampling typically leads to improved interpolation results because it does not create coherent subsampling artifacts (Xu et al., 2005).

### 6.3.3 Nonequispaced fast discrete curvelet transform (NFDCT)

For irregularly acquired seismic data, the (FFT inside) FDCT (Candès et al., 2006a) assumes regular sampling along all (spatial) coordinates. Ignoring the nonuniformity of the spatial sampling no longer helps in detecting the wavefronts because of a lack of continuity. Hennenfent and Herrmann (2006) addressed this issue by extending the FDCT to nonuniform (or irregular) grids via the nonequispaced fast Fourier transform, NFFT (Potts et al., 2001; Kunis, 2006). The outcome was the 'first generation NFDCT' (nonequispaced fast discrete curvelet transform), which relied on accurate Fourier coefficients obtained by an $\ell_2$-regularized inversion of the NFFT.

The NFDCT handles irregular sampling, thus, exploring continuity along the wavefronts by viewing seismic data in a geometrically correct way—typically nonuniformly sampled along the spatial coordinates (source and/or receiver). In Hennenfent et al. (2010), the authors introduced a 'second generation NFDCT', which is based on a direct, $\ell_1$-regularized inversion of the operator that links curvelet coefficients to irregular data. Unlike the first generation NFDCT, the second generation NFDCT is lossless by construction—i.e., the curvelet coefficients explain the data at irregular locations exactly. This property is important for processing irregularly sampled seismic data in the curvelet domain and bringing them back to their irregular recording locations with high fidelity. Note that the second generation NFDCT is lossless for regularization not interpolation.

The NFDCT framework as setup in Hennenfent et al. (2010) basically involves a Kronecker product ($\otimes$) of a 1D FFT operator $\mathbf{F}_t$, used along the temporal coordinate, and a 1D NFFT operator $\mathbf{N}_x$, used along the spatial coordinate, followed by the application of the curvelet tilling operator $\mathbf{T}$ that maps curvelet coefficients to the Fourier domain—i.e., $\mathbf{B} \overset{\text{def}}{=} \mathbf{T}(\mathbf{N}_x \otimes \mathbf{F}_t)$. Therefore, $\mathbf{B}$ is the NFDCT operator that links the curvelet coefficients to nonequispaced traces. The 1D NFFT operator ($\mathbf{N}_x$) replaces the 1D FFT operator ($\mathbf{F}_x$) that acts along the spatial coordinate in FDCT. Note that the NFDCT operator described above is written differently than in Hennenfent et al. (2010) because the latter defines the synthesis FFT operator as $\mathbf{F}$, whereas $\mathbf{F}$ is the analysis FFT operator in this chapter. This also ensures consistency of notation and terminology with Chapter 5.

For the proposed simultaneous acquisition, the joint problem of source separation, regularization and interpolation is addressed by using a sparsifying operator ($\mathbf{S}$) that handles the multidimensionality of this problem. Therefore, $\mathbf{S} \overset{\text{def}}{=} \mathbf{C} \otimes \mathbf{W}$, where $\mathbf{C}$ is a 2D NFDCT operator and $\mathbf{W}$ is a 1D wavelet operator. The NFDCT operator is modified as

$$\mathbf{C} \overset{\text{def}}{=} \mathbf{T}(\mathbf{N}_{xs} \otimes \mathbf{F}_{xr}), \tag{6.2}$$

where the 1D NFFT operator $\mathbf{N}_{xs}$ acts along the jittered shot coordinate and the 1D FFT operator $\mathbf{F}_{xr}$ acts along the regular receiver coordinate. The 1D wavelet operator is applied on the time coordinate. As mentioned previously, the measurement matrix $\mathbf{A} = \mathbf{M}\mathbf{S}^H$. From a practical point of view, it is important to note that matrix-vector products with all the matrices are matrix free—i.e., these matrices are operators that define the action of the matrix on a vector, but are never formed explicitly.

In summary, recovery of nonoverlapping, periodic and densely sampled data from simultaneous, irregular and compressed data is achieved by incorporating an NFFT operator inside the curvelet framework that acts along the irregular spatial coordinate(s) and applying time shifts to the traces wherever necessary. Note that the NFFT operator is incorporated in the 2D NFDCT operator $\mathbf{C}$, which is incorporated in the sparsifying operator $\mathbf{S}$, and the time shift $\Delta t$ is incorporated in the acquisition operator $\mathbf{M}$. The NFFT computes (fine grid) 2D Fourier coefficients by mapping the coarse nonuniform spatial grid to a fine uniform grid. The curvelet coefficients are computed directly from the 2D Fourier coefficients.

## 6.4 Time-lapse acquisition via jittered sources

In Chapter 5, we extended the time-jittered marine acquisition to time-lapse surveys where the shot positions were jittered on a discrete periodic grid. In this chapter, we extend the framework to more realistic field acquisition scenarios by incorporating irregular grids. Figure 6.5(a) illustrates a conventional marine acquisition scheme and two realizations of the off-the-grid time-jittered marine acquisition are shown in Figures 6.5(b) and 6.5(c), one each for the baseline and the monitor survey. Remember that these surveys generate simultaneous, irregular and subsampled measurements. We assume no significant variations in the water column velocities, wave heights or temperature and salinity profiles, etc., amongst the different surveys. The source signature is also assumed to be the same.

We describe noise-free time-lapse data acquired from a baseline and a monitor survey as $\mathbf{y}_j = \mathbf{A}_j\mathbf{x}_j$ for $j = \{1, 2\}$, where $\mathbf{y}_1$ and $\mathbf{y}_2$ represent the subsampled, simultaneous measurements for the baseline and monitor surveys, respectively; $\mathbf{A}_1$ and $\mathbf{A}_2$ are the corresponding flat ($n \ll N < P$) measurement matrices. Note that both the measurement matrices incorporate the NFDCT operator, as described above, to account and correct for the irregularity in the observed measurements of the baseline ($\mathbf{y_1}$) and monitor surveys ($\mathbf{y_2}$). Recovering densely sampled vintages for each vintage independently (via Equation 6.1) is referred to as the independent recovery strategy (IRS). Since in Chapter 5 we demonstrated that recovery via IRS is inferior to recovery via the joint recovery method, we work only with the latter in this chapter.

### 6.4.1 Joint recovery method

The joint recovery method (JRM) performs a joint inversion by exploiting shared information between the vintages. The joint recovery model (DCS, Baron et al., 2009) is formulated as

$$\begin{aligned} \begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{bmatrix} &= \begin{bmatrix} \mathbf{A}_1 & \mathbf{A}_1 & \mathbf{0} \\ \mathbf{A}_2 & \mathbf{0} & \mathbf{A}_2 \end{bmatrix} \begin{bmatrix} \mathbf{z}_0 \\ \mathbf{z}_1 \\ \mathbf{z}_2 \end{bmatrix}, \quad \text{or} \\ \mathbf{y} &= \mathbf{A}\mathbf{z}. \end{aligned} \tag{6.3}$$

In this model, the vectors $\mathbf{y}_1$ and $\mathbf{y}_2$ represent observed measurements from the baseline and monitor surveys, respectively. The vectors for the vintages are given by

$$\mathbf{x}_j = \mathbf{z}_0 + \mathbf{z}_j, \qquad j \in 1, 2, \tag{6.4}$$

where the common component is denoted by $\mathbf{z}_0$, and the innovations are denoted by $\mathbf{z}_j$ for $j \in 1, 2$ with respect to this common component that is shared by the vintages. The symbol $\mathbf{A}$ is overloaded to refer to the matrix linking the observations of the time-lapse surveys to the common component and innovations pertaining to the different vintages. The above joint recovery model can be extended to $J > 2$ surveys, yielding a $J \times$ (number of vintages $+ 1$) system.

Since the vintages share the common component in Equation 6.3, solving

$$\widetilde{\mathbf{z}} = \arg\min_{\mathbf{z}} \|\mathbf{z}\|_1 \quad \text{subject to} \quad \mathbf{y} = \mathbf{A}\mathbf{z}, \tag{6.5}$$

will exploit the correlations amongst the vintages. Equation 6.5 seeks solutions for the common component and innovations that have the smallest $\ell_1$-norm such that the observations explain the incomplete recordings for both vintages. The densely sampled vintages are estimated via Equation 6.4 with the recovered $\widetilde{\mathbf{z}}$ and the time-lapse difference is computed via $\widetilde{\mathbf{z}}_1 - \widetilde{\mathbf{z}}_2$.

Given a baseline data vector $\mathbf{f}_1$ and a monitor data vector $\mathbf{f}_2$, $\mathbf{x}_1$ and $\mathbf{x}_2$ are the corresponding sparse representations. Given the measurements $\mathbf{y}_1 = \mathbf{M}_1\mathbf{f}_1$ and $\mathbf{y}_2 = \mathbf{M}_2\mathbf{f}_2$, and $\mathbf{A}_1 = \mathbf{M}_1\mathbf{S}_1^H$, $\mathbf{A}_2 = \mathbf{M}_2\mathbf{S}_2^H$, our aim is to recover the wavefields (or sparse approximations) $\widetilde{\mathbf{f}}_1$ and $\widetilde{\mathbf{f}}_2$ by solving the sparse recovery problem as described above from which the time-lapse signal can be computed. Note that $\mathbf{S} \overset{\text{def}}{=} \mathbf{C} \otimes \mathbf{W}$, where $\mathbf{C}$ is the NFDCT operator (see Equation 6.2) and $\mathbf{W}$ is a 1D wavelet operator. The reconstructed wavefields $\widetilde{\mathbf{f}}_1$ and $\widetilde{\mathbf{f}}_2$ are obtained as: $\widetilde{\mathbf{f}}_1 = \mathbf{S}^H\widetilde{\mathbf{x}}_1$ and $\widetilde{\mathbf{f}}_2 = \mathbf{S}^H\widetilde{\mathbf{x}}_2$, where $\widetilde{\mathbf{x}}_1$ and $\widetilde{\mathbf{x}}_2$ are the recovered sparse representations and the operator $\mathbf{S}$ is overwritten to represent the Kronecker product between the standard FDCT operator and the 1D wavelet operator. The standard FDCT operator is used because the recovered sparse representations $\widetilde{\mathbf{x}}_1$ and $\widetilde{\mathbf{x}}_2$ correspond to the coefficients of the regularized wavefields. Since we are always subsampled in both the baseline and monitor surveys, have irregular traces and cannot exactly repeat, which is inherent of the acquisition design and due to natural environmental constraints, we would like to recover the periodic densely sampled prestack vintages and time-lapse difference. For the given recovery problem, the vintages and time-lapse difference are mapped to one colocated fine regular periodic grid.

## 6.5 Economic performance indicators

To quantify the cost savings associated with simultaneous acquisition, we measure the performance of the proposed acquisition design and recovery scheme in terms of an improved spatial-sampling ratio (ISSR), defined as

$$\mathrm{ISSR} = \frac{\text{number of shots recovered via sparsity-promoting inversion}}{\text{number of shots in simultaneous acquisition}}. \tag{6.6}$$

For time-jittered marine acquisition, a subsampling factor $\eta = 2, 4, ...$, etc., implies a gain in the spatial sampling by factor of $2, 4, ...$, etc. In practice, this corresponds to an improved efficiency of the acquisition by the same factor. Recently, Mosher et al. (2014) have shown that factors of two or as high as ten in efficiency improvement are achievable in the field.

The survey-time ratio (STR)—a performance indicator proposed by Berkhout (2008)—compares the time taken for conventional and simultaneous acquisition:

$$\mathrm{STR} = \frac{\text{time of conventional acquisition}}{\text{time of simultaneous acquisition}}. \tag{6.7}$$

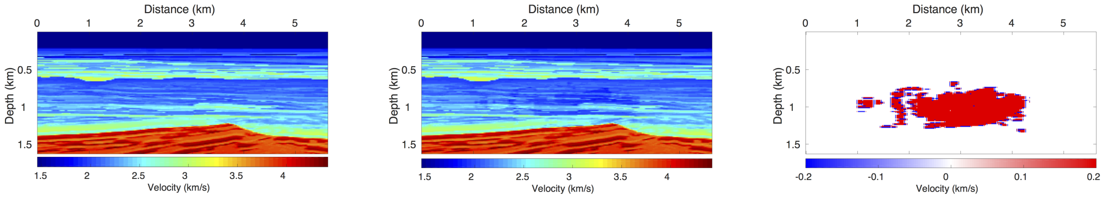
As mentioned previously, if we wish to acquire 10.0 s-long shot records at every 12.5 m, the speed of the source vessel would have to be about 1.25 m/s ($\approx$ 2.5 knots). In simultaneous acquisition, the speed of the source vessel is approximately maintained at (the standard) 2.5 m/s ($\approx$ 5.0 knots). Therefore, for a subsampling factor of $\eta = 2, 4, ...$, etc., there is an implicit reduction in the survey time by $\frac{1}{\eta}$.

## 6.6 Synthetic seismic case study

To illustrate the performance of our proposed joint recovery method for off-the-grid surveys, we carry out a number of experiments on 2D seismic lines generated from two different velocity models—first, the BG COMPASS model (provided by BG Group) that has simple geology with complex time-lapse difference; and second, the SEAM Phase 1 model (provided by HESS) that has complex geology with complex time-lapse difference due to the complexity of the overburden. Note that for the SEAM model, we generate the time-lapse difference via fluid substitution as shown below. Also, the geology of the BG COMPASS model is relatively simpler than the SEAM model, although it does have vertical and lateral complexity.

### 6.6.1 BG COMPASS model—simple geology, complex time-lapse difference

The synthetic BG COMPASS model has a (relatively) simple geology but a complex time-lapse difference. Figures 6.8(a) and 6.8(b) display the baseline and monitor models. Note that this is a subset of the BG COMPASS model, wherein the monitor model includes a gas cloud. The



(a) (b) (c)

**Figure 6.8:** Subset of the BG COMPASS model. (a) Baseline model; (b) monitor model; (c) difference between (a) and (b) showing the gas cloud.

time-lapse difference in Figure 6.8(c) shows the gas cloud.

Using IWAVE (Symes, 2010), a time-stepping simulation software, two acoustic data sets with a conventional source (and receiver) sampling of 12.5 m are generated, one from the baseline model and the other from the monitor model. Each data set has $N_t = 512$ time samples, $N_r = 260$ receivers and $N_s = 260$ sources. The time sampling interval is 0.004 s. Subtracting the two data sets yields the time-lapse difference. Since no noise is added to the data, the time-lapse difference is simply the time-lapse signal. A receiver gather from the simulated baseline data, the monitor data and the corresponding time-lapse difference is shown in Figure 6.2(a), 6.2(b) and 6.2(c), respectively. The first shot position in the receiver gathers—labeled as 0 m in the figures—corresponds to 1.5 km in the synthetic velocity model. Given the spatial sampling of 12.5 m, the subsampling factor $\eta$ for the time-jittered acquisition is 2. Hence, the number of measurements for each experiment is fixed—i.e., $n = N/\eta = N/2$, each for $\mathbf{y}_1$ and $\mathbf{y}_2$. We also conduct experiments for $\eta = 4$.

To reflect current practices in time-lapse acquisition—where people aim to replicate the surveys—we simulate 10 different realizations of the time-jittered marine acquisition with 100% overlap between the baseline and monitor surveys. The term "overlap" refers to the percentage of shot positions from the baseline survey revisited (or replicated exactly) for the monitor survey, and therefore rows in the measurement matrices $\mathbf{A_1}$ and $\mathbf{A_2}$ are exactly the same. Note that these shot positions are irregular, and hence off the grid. However, since exact replication of the surveys in the field is not possible, we conduct experiments to study the impact of deviations in the shot positions that would occur naturally in the field. We introduce small deviations of average $\pm(1, 2, 3)$ m in the shot positions of the baseline surveys to generate the shot positions for the monitor surveys. For instance, given a realization of the time-jittered baseline survey, deviating each shot position by $\approx \pm 1$ m generates shot positions for the corresponding monitor survey. Note that these deviations are average deviations in the sense that for a given realization of the time-jittered baseline survey, the shot positions are deviated by random real numbers resulting in average deviations of $\pm 1$ m, $\pm 2$ m or $\pm 3$ m. One of our aims is to analyze the effects of nonreplication of the time-lapse surveys on time-lapse data—i.e., when $\mathbf{A_1} \neq \mathbf{A_2}$. By virtue of the design of the simultaneous acquisition and based upon the subsampling factor ($\eta$), it is not possible to have two completely different (0% overlap) realizations of the time-jittered acquisition. Therefore, we compare recoveries from the above cases with the acquisition scenarios that have least possible (or unavoidable) overlap between the time-lapse surveys. In all cases, we recover periodic densely sampled baseline and monitor data from the simultaneous data $\mathbf{y}_1$ and $\mathbf{y}_2$, respectively, using the joint recovery method (by solving Equation 6.5). The inherent time-lapse difference is computed by subtracting the recovered baseline and monitor data.

We conduct 10 experiments for the baseline measurements, wherein each experiment has a different random realization of the measurement matrix $\mathbf{A_1}$. Then, for each experiment, we fix the baseline measurement and subsequently work with different realizations of the monitor survey generated by introducing small deviations in the shot positions and jittered firing times from the baseline survey, resulting in slightly different overlaps between the surveys. To get better insight on the effects of nonreplication of the time-lapse surveys, we also conduct experiments for the case of least possible overlap between the surveys. Tables 6.1 and 6.2 summarize the recovery results for the time-lapse data for $\eta = 2$ and 4, respectively, in terms of the signal-to-noise ratio defined as

$$\mathrm{S/N}(\mathbf{f}, \widetilde{\mathbf{f}}) = -20 \log_{10} \frac{\|\mathbf{f} - \widetilde{\mathbf{f}}\|_2}{\|\mathbf{f}\|_2}. \tag{6.8}$$

Each table compares recoveries for different overlaps between the baseline and monitor surveys, with and without position deviations. Each S/N value is an average of 10 experiments including the standard deviation. Note that for time-jittered acquisition with $\eta = 2$, the least possible overlap between the surveys is observed to be greater than 0% and less than 15%. Hence, Table 6.1 shows the S/Ns for the overlap of $< 15\%$. Similarly, for time-jittered acquisition with $\eta = 4$, Table 6.2 shows the S/Ns for the overlap of $< 5\%$.

We recover periodic densely sampled data from simultaneous, subsampled and irregular data by solving Equation 6.5. The recovered time-lapse data is colocated, regularized and interpolated to a fine uniform grid since both the measurement matrices $\mathbf{A_1}$ and $\mathbf{A_2}$ incorporate a 2D nonequispaced fast discrete curvelet transform that handles irregularity of traces by viewing the observed data in a geometrically correct way. The S/Ns of the recovered time-lapse data lead to some interesting observations. First, there is little variability in the recovery of the time-lapse difference from (the ideal) 100% overlap between the surveys to the more realistic scenarios of in-the-field acquisitions that have natural deviations or irregularities in the shot positions. Second, time-lapse difference recovery from the least possible overlap (between the surveys) is similar to the recovery of 100% overlap with and without deviations. This is significant because it indicates a possibility to relax the insistence on replication of the time-lapse surveys, which makes this technology challenging and expensive. The small standard deviations for each case suggest little variability in the recovery for different random realizations. Moreover, the standard deviations are greater for cases other than the minimum overlap. The above observations hold for both subsampling factors, $\eta = 2$ and 4, as illustrated in Figures 6.10 and 6.12.

Third, increasing deviations or irregularities in shot positions improve recovery of the vintages (Figures 6.9(c), 6.9(e), 6.9(g)), with the minimum overlap between surveys giving the best recovery (Figure 6.9(i)). This is due to the (partial) independence of the measurement matrices that contribute additional information via the first column of $\mathbf{A}$ in Equation 6.3 connecting the common component to observations of both vintages—i.e., for time-lapse seismic, independent surveys give additional structural information leading to improved recovery quality of the vintages. The improvement in the recoveries is better visible through the corresponding difference plots in Figures 6.9(d), 6.9(f), 6.9(h), 6.9(j). This observation is important because, as mentioned previously, time-lapse differences are often studied via differences in certain poststack attributes computed from the (recovered) prestack vintages. Hence, as the quality of the recovered prestack vintages improves with decrease in the overlap, they serve as better input to extract the poststack attributes. Moreover, the small standard deviations for each overlap indicate little variability in the recovery from one random realization to another. This is desirable since it offers a possibility to relax the

| Overlap ± avg. deviation | Baseline | Monitor | 4D signal |
|---|---|---|---|
| 100% | 19.8 ± 1.0 | 19.7 ± 1.0 | 11.3 ± 2.2 |
| 100% ± 1.0 m | 19.7 ± 1.0 | 19.6 ± 1.0 | 10.3 ± 1.5 |
| 100% ± 2.0 m | 20.3 ± 1.1 | 20.2 ± 1.0 | 10.7 ± 1.1 |
| 100% ± 3.0 m | 20.8 ± 1.2 | 20.7 ± 1.1 | 11.0 ± 1.4 |
| $< 15\%$ | 23.8 ± 1.4 | 23.6 ± 1.4 | 10.2 ± 1.2 |

**Table 6.1:** Summary of recoveries in terms of S/N (dB) for data recovered via JRM for a subsampling factor $\eta = 2$. The S/Ns show little variability in the time-lapse difference recovery for different overlaps between the surveys offering a possibility to relax insistence on replicability of time-lapse surveys. This is supported by the improved recovery of the vintages as the overlap decreases. Note that the deviations are average deviations.

insistence on replication of the time-lapse surveys along with embracing the naturally occurring random deviations in the field. The standard deviations for different overlaps also do not fluctuate as much as compared to those of the time-lapse difference. Recovery of the vintages and the corresponding difference plots for a subsampling of $\eta = 4$ are shown in Figure 6.11.

An increase in the subsampling factor leads to decrease in the S/Ns of the recovered time-lapse data, however, the recoveries are reasonable as shown in Figures 6.11 and 6.12. This observation is in accordance with the CS theory wherein the recovery quality decreases for increased subsampling. Note that recovery of weak late-arriving events can be further improved by rerunning the recovery algorithm using the residual as input, using weighted one-norm minimization that exploits correlations between locations of significant transform-domain coefficients of different partitions—e.g., shot records, common-offset gathers, or frequency slices—of the acquired data (Mansour et al., 2013), etc. This needs to be carefully investigated. Remember that for a given subsampling factor the number of measurements is the same for all experiments and the observed differences can be fully attributed to the performance of the joint recovery method in relation to the overlap between the two surveys encoded in the measurement matrices. Also, given the context of randomized subsampling and irregularity of the observed data, it is important to recover the densely sampled vintages and then the time-lapse difference. Moreover, as mentioned previously, while we do not insist that we actually visit predesigned irregular (or off-the-grid) shot positions for the time-lapse surveys, however, it is important to know these positions to sufficient accuracy after acquisition for high-quality data recovery. This can be achieved in practice as shown by Mosher et al. (2014).

### 6.6.2 SEAM Phase 1 model—complex geology, complex time-lapse difference

The SEAM model is a 3D deepwater subsalt earth model that includes a complex salt intrusive in a folded Tertiary basin. We select a 2D slice from the 3D model to generate a seismic line. Figure 6.13(a) shows a subset of the 2D slice used as the baseline model. We define the monitor model, Figure 6.13(b), from the baseline model via fluid substitution resulting in a time-lapse difference under the overburden as shown in Figure 6.13(c).

Using IWAVE (Symes, 2010), two acoustic data sets with a conventional source (and receiver) sampling of 12.5 m are generated, one from the baseline model and the other from the monitor model. Each data set has $N_t = 2048$ time samples, $N_r = 320$ receivers and $N_s = 320$ sources. The time sampling interval is 0.004 s. Subtracting the two data sets yields the time-lapse difference.

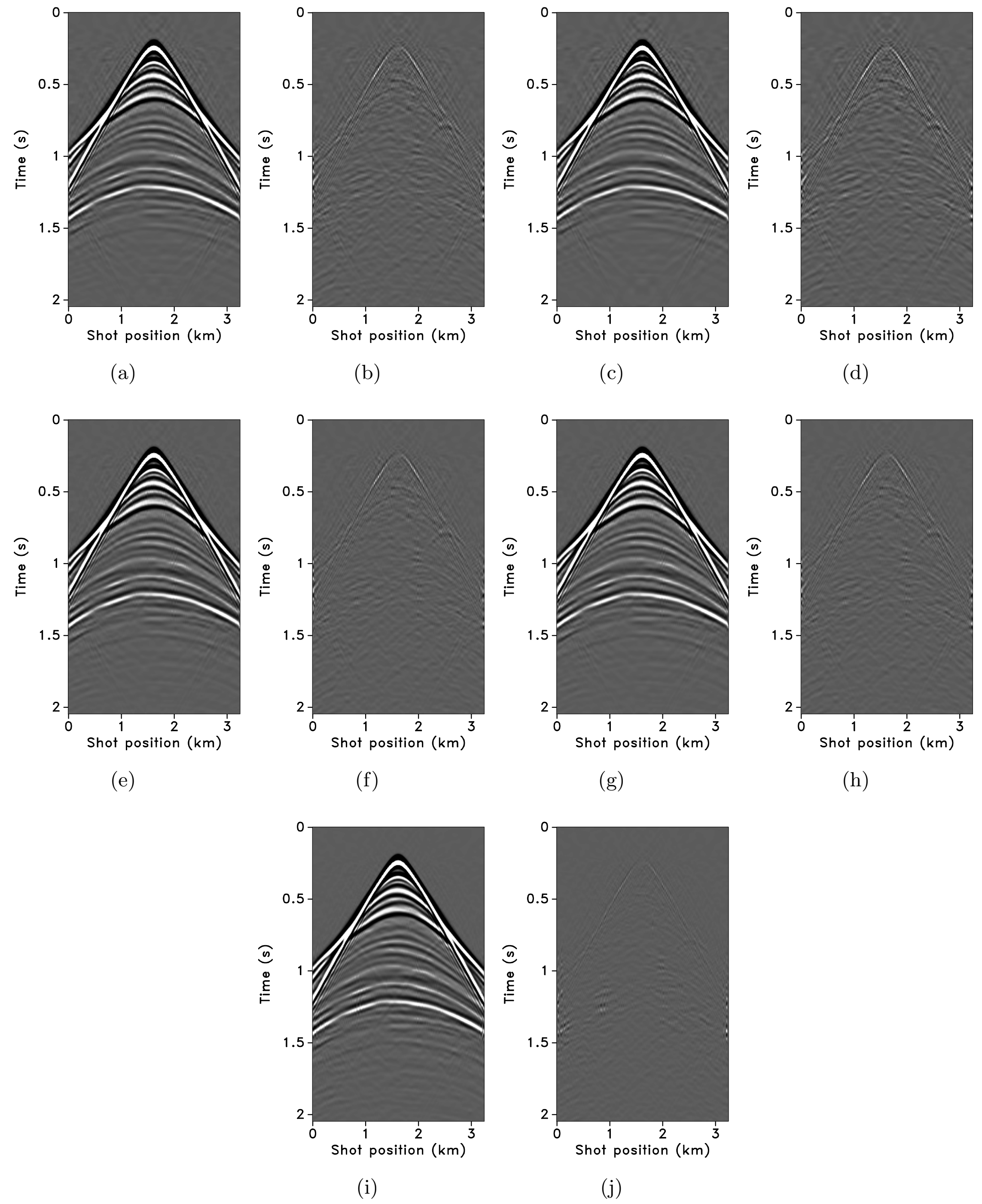
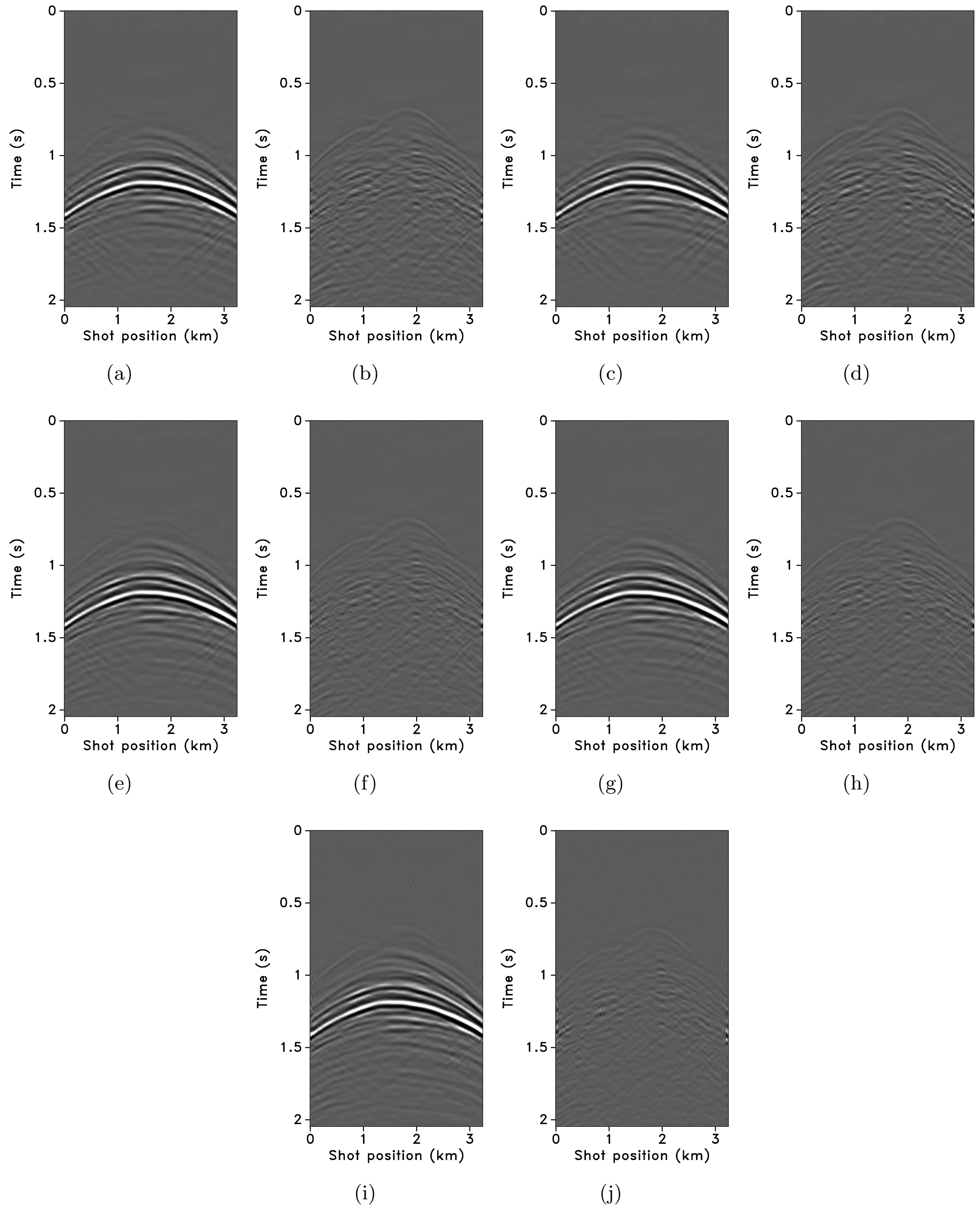


**Figure 6.9:** JRM recovered monitor receiver gathers from the BG COMPASS model for a subsampling factor $\eta = 2$. Recovered monitor data and residual with (a,b) 100% overlap in the measurement matrices ($\mathbf{A_1}$ and $\mathbf{A_2}$); (c,d) 100% overlap and average shot-position deviation of 1 m; (e,f) 100% overlap and average shot-position deviation of 2 m; (g,h) 100% overlap and average shot-position deviation of 3 m; (i,j) $< 15\%$ overlap, respectively.

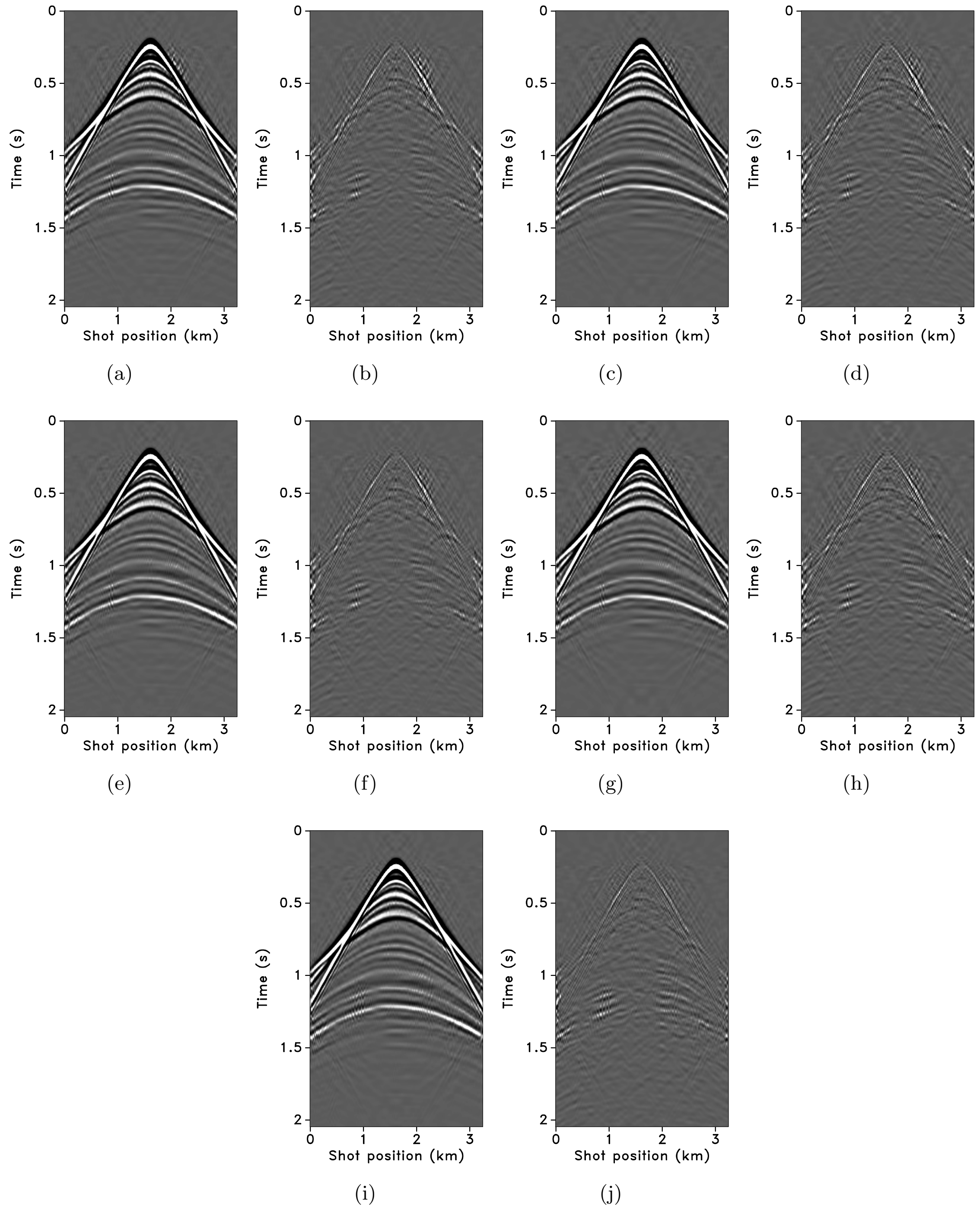


(a) (b) (c) (d)

(e) (f) (g) (h)

(i) (j)

**Figure 6.10:** JRM recovered time-lapse difference receiver gathers from the BG COMPASS model for a subsampling factor $\eta = 2$. Recovered time-lapse difference and residual with (a,b) 100% overlap in the measurement matrices ($\mathbf{A_1}$ and $\mathbf{A_2}$); (c,d) 100% overlap and average shot-position deviation of 1 m; (e,f) 100% overlap and average shot-position deviation of 2 m; (g,h) 100% overlap and average shot-position deviation of 3 m; (i,j) $< 15\%$ overlap, respectively.

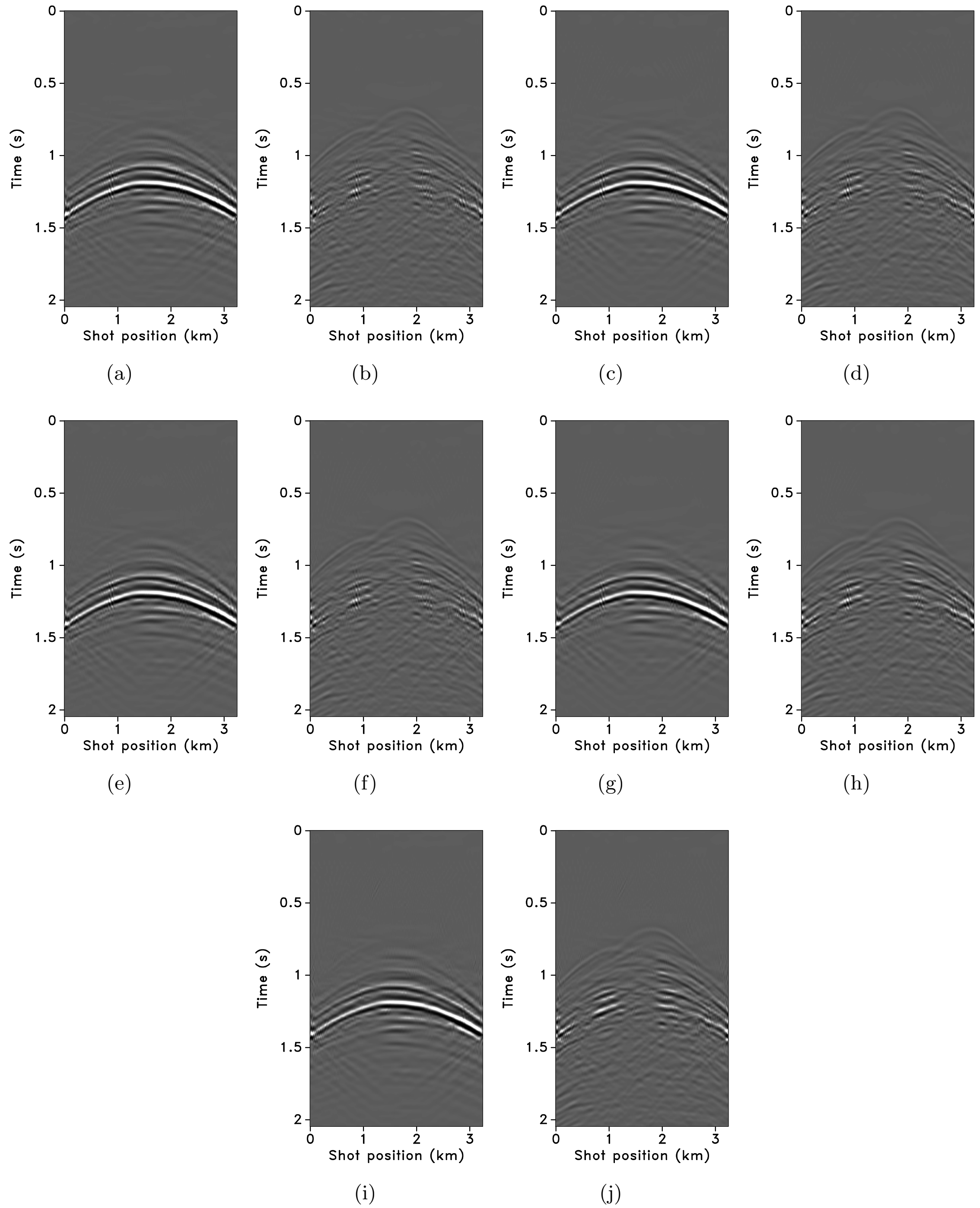


(a) (b) (c) (d)

(e) (f) (g) (h)

(i) (j)

**Figure 6.11:** JRM recovered monitor receiver gathers from the BG COMPASS model for a subsampling factor $\eta = 4$. Recovered monitor data and residual with (a,b) 100% overlap in the measurement matrices ($\mathbf{A_1}$ and $\mathbf{A_2}$); (c,d) 100% overlap and average shot-position deviation of 1 m; (e,f) 100% overlap and average shot-position deviation of 2 m; (g,h) 100% overlap and average shot-position deviation of 3 m; (i,j) $< 5\%$ overlap, respectively.



(a) (b) (c) (d)

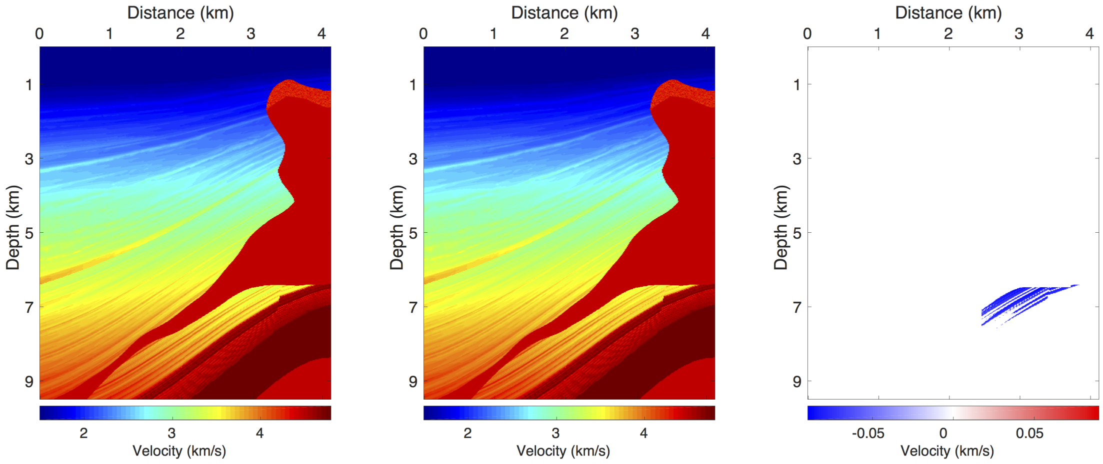
(e) (f) (g) (h)

(i) (j)

**Figure 6.12:** JRM recovered time-lapse difference receiver gathers from the BG COMPASS model for a subsampling factor $\eta = 4$. Recovered time-lapse difference and residual with (a,b) 100% overlap in the measurement matrices ($\mathbf{A_1}$ and $\mathbf{A_2}$); (c,d) 100% overlap and average shot-position deviation of 1 m; (e,f) 100% overlap and average shot-position deviation of 2 m; (g,h) 100% overlap and average shot-position deviation of 3 m; (i,j) $< 5\%$ overlap, respectively.

| Overlap ± avg. deviation | Baseline | Monitor | 4D signal |
|---|---|---|---|
| 100% | 14.3 ± 0.6 | 14.2 ± 0.6 | 6.4 ± 0.7 |
| 100% ± 1.0 m | 14.9 ± 0.8 | 14.8 ± 0.8 | 6.5 ± 1.0 |
| 100% ± 2.0 m | 15.6 ± 1.0 | 15.5 ± 1.0 | 6.4 ± 1.3 |
| 100% ± 3.0 m | 16.4 ± 0.9 | 16.3 ± 0.9 | 6.4 ± 0.7 |
| < 5% | 18.4 ± 0.7 | 18.2 ± 0.7 | 5.8 ± 0.4 |

**Table 6.2:** Summary of recoveries in terms of S/N (dB) for data recovered via JRM for a subsampling factor $\eta = 4$. The S/Ns show little variability in the time-lapse difference recovery for different overlaps between the surveys offering a possibility to relax insistence on replicability of time-lapse surveys. This is supported by the improved recovery of the vintages as the overlap decreases. Note that the deviations are average deviations.



(a) (b) (c)

**Figure 6.13:** Subset of the SEAM model. (a) Baseline model; (b) monitor model; (c) difference between (a) and (b) showing the time-lapse difference.

Since no noise is added to the data, the time-lapse difference is simply the time-lapse signal. A receiver gather from the simulated baseline data, the monitor data and the corresponding time-lapse difference is shown in Figures 6.14(a), 6.14(b) and 6.14(c), respectively. Note that the amplitude of the time-lapse difference is one-tenth the amplitude of the baseline and monitor data. Therefore, in order to make the time-lapse difference visible, the color axis for the figures showing the time-lapse difference is one-tenth the color axis for the figures showing the baseline and monitor data. This colormap applies for the remainder of the chapter. Given the spatial sampling of 12.5 m, the subsampling factor $\eta$ for the time-jittered acquisition is 2. The number of measurements for each experiment is fixed—i.e., $n = N/\eta = N/2$, each for $\mathbf{y}_1$ and $\mathbf{y}_2$.

We simulate a realization of the time-jittered marine acquisition with 100% overlap between the baseline and monitor surveys. Since our main aim is to analyze the effects of nonreplication of the time-lapse surveys on time-lapse data—i.e., when $\mathbf{A_1} \neq \mathbf{A_2}$—we compare recovery from the

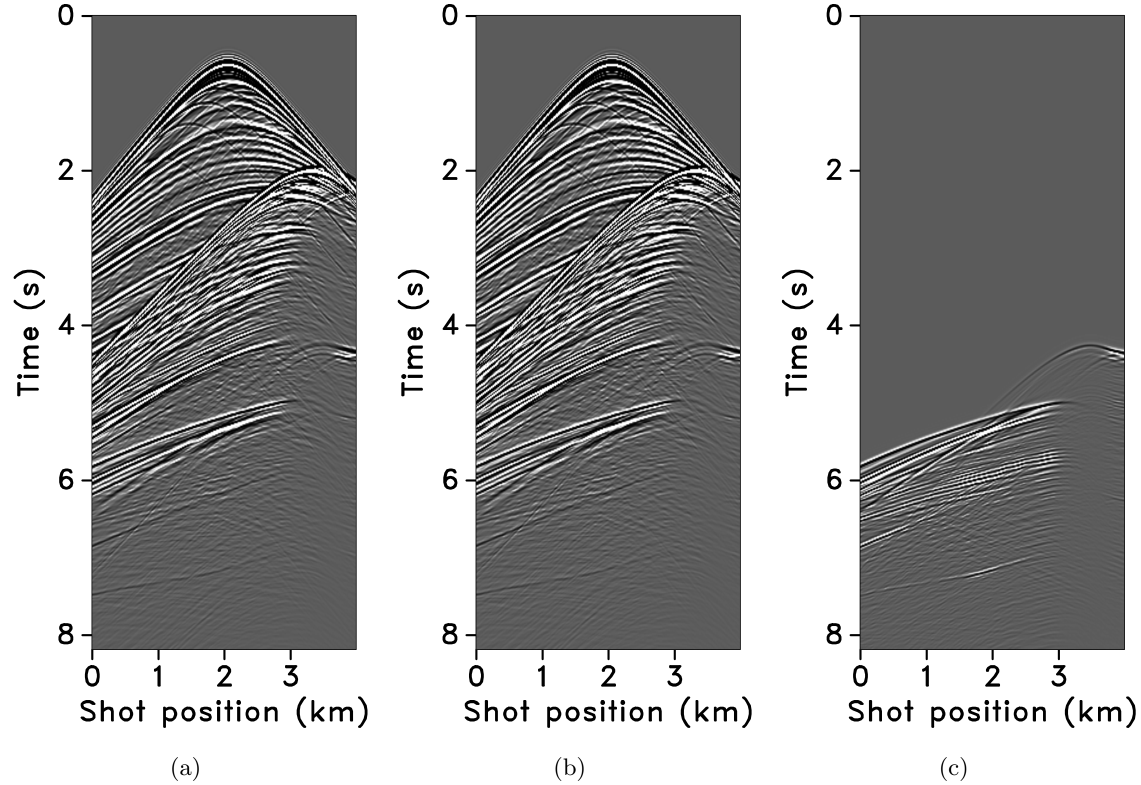


**Figure 6.14:** Synthetic receiver gathers from the conventional SEAM (a) baseline survey, (b) monitor survey. (c) Corresponding time-lapse difference. The amplitude of the time-lapse difference is one-tenth the amplitude of the baseline and monitor data.

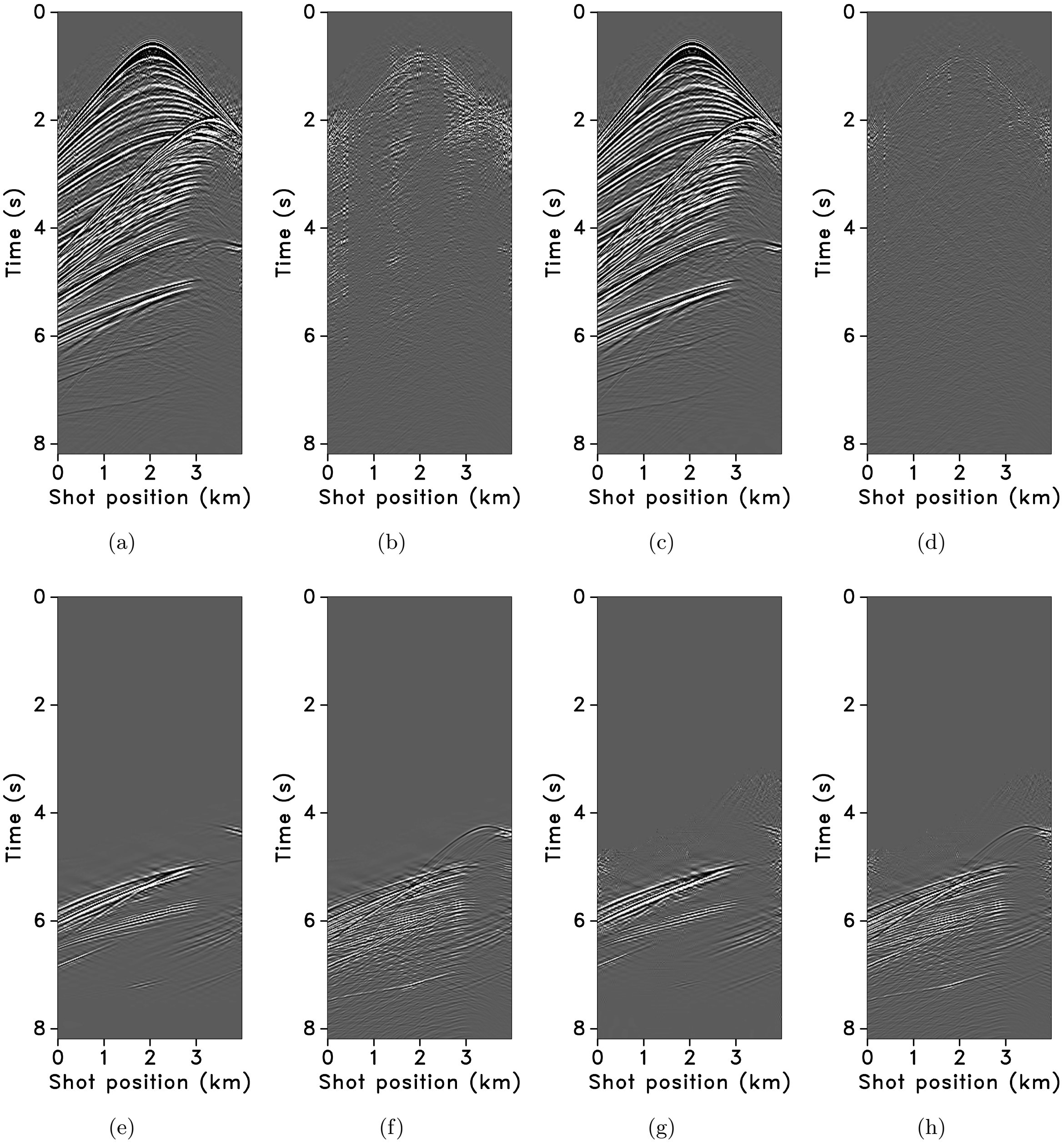
above case with the acquisition scenario that has least possible (or unavoidable) overlap between the time-lapse surveys only. Given the bigger size of the data set and limited computational resources, we restrict ourselves to one experiment for each case and a subsampling of $\eta = 2$. Periodic densely sampled baseline and monitor data is recovered from the simultaneous data $\mathbf{y}_1$ and $\mathbf{y}_2$, respectively, by solving Equation 6.5. The inherent time-lapse difference is computed by subtracting the recovered baseline and monitor data.

The recovered time-lapse data is colocated, regularized and interpolated to a fine uniform grid. We note that all the observations made for the BG COMPASS model, which is a relatively simpler model, hold true for the more complex SEAM model. Minimum overlap (or nonreplication) between time-lapse surveys improves recovery of the vintages since independent surveys give additional structural information. Hence, they serve as better input to extract certain poststack attributes used to study time-lapse differences. Figures 6.15(a), 6.15(b), 6.15(c) and 6.15(d) show the corresponding monitor data recovery. The S/N for the vintage recovery for minimum overlap between the surveys is 30.2 dB—a significant improvement from the 19.5 dB recovery for 100% overlap between the surveys. Moreover, as seen in Figures 6.15(e), 6.15(f), 6.15(g) and 6.15(h), there is little variability in the recovery of the time-lapse difference from (the ideal) 100% overlap between the surveys to the more realistic almost nonreplicated surveys. The corresponding S/Ns for the recovered time-lapse difference are 9.6 dB for 100% overlap and 4.1 dB for minimum overlap between the surveys. We note that the S/N for the minimum overlap between the surveys is biased due the presence of incoherent noise—between 3.5 s to 5.0 s—above the main time-lapse difference. If we compute the S/Ns for the lower-half of the data that contains the time-lapse difference—i.e., after 4.5 s—the S/N for minimum overlap between the surveys increases to 6.8 dB. More importantly, if we look at the plots themselves, we see that there is not much difference in the two recoveries. We are able to recover the primary arrivals and some reverberations below. Recall that the amplitude of the time-lapse difference is one-tenth the amplitude of the vintages. It is quite remarkable that we get good results given the complexity of the model and the low amplitude of the time-lapse difference. Recovery of the vintages and the time-lapse difference for a subsampling of $\eta = 4$ follows the same trend as above.

## 6.7 Discussion

Realistic field seismic acquisitions suffer, amongst other possibly detrimental external factors, from irregular spatial sampling of sources and receivers. This poses technical challenges for the time-lapse seismic technology that currently aims to replicate densely sampled surveys for monitoring changes due to production. The experiments and synthetic results shown in the previous sections demonstrate favourable effects of irregular sampling and nonreplication of surveys on time-lapse data—i.e., decrease in replicability of the surveys leads to improved recovery of the vintages with little variability in the recovery of the time-lapse difference itself—while unraveling overlapping shot records. Note that we do not insist on replicating the irregular spatial positions in the field, however, the above observations hold as long as we know the irregular sampling positions after acquisition to a sufficient degree of accuracy, which is attainable in practice (see e.g., Mosher et al., 2014). Furthermore, we assume that there are no significant variations in the water column velocities, wave heights or temperature and salinity profiles amongst the different surveys while the source signature is also assumed to be the same. As long as these physical changes can be modeled, we do not foresee major problems. For instance, we expect that our approach can relatively easily be combined with source equalization (see e.g., Rickett and Lumley, 2001) and curvelet-domain matched filtering techniques (Beyreuther et al., 2005; Tegtmeier-Last and Hennenfent, 2013).

The proposed methodology involves a combination of economical randomized samplings with low environmental imprint and sparsity-promoting data recovery that aims to reduce cost of surveys and improve quality of the prestack time-lapse data without relying on expensive dense sampling and high degrees of replicability of the surveys. The combined operation of source separation, regularization and interpolation renders periodic densely sampled time-lapse data from time-compressed, and therefore economical, simultaneous, subsampled and irregular data. While the simultaneous data are separated reasonably well, recovery of the weak late-arriving events can be further improved by rerunning the recovery algorithm using the residual as input, using weighted one-norm minimization that exploits correlations between locations of significant transform-domain coefficients of different partitions—e.g., shot records, common-offset gathers, or frequency slices—of the acquired data (Mansour et al., 2013), etc. This needs to be examined in detail. Effects of noise and other physical changes in the environment also need to be carefully investigated. Nevertheless, as expected using standard CS, our recovery method should be stable with respect to noise (Candès et al., 2006c). Moreover, recent successes in the application of compressed sensing to land and marine field data acquisition (see e.g., Mosher et al., 2014) support the fact that technical challenges with noise and calibration can be overcome in practice.



(a) (b) (c) (d)

(e) (f) (g) (h)

**Figure 6.15:** JRM recovered monitor and time-lapse difference receiver gathers from the SEAM model for a subsampling factor $\eta = 2$. Recovered monitor data and residual with (a,b) 100% overlap in the measurement matrices ($\mathbf{A_1}$ and $\mathbf{A_2}$); (c,d) $<$ 15% overlap, respectively. Recovered time-lapse difference and residual with (e,f) 100% overlap in the measurement matrices; (g,h) $<$ 15% overlap, respectively. Note that the amplitude of the time-lapse difference is one-tenth the amplitude of the monitor data.

## 6.8 Conclusions

We present an extension of our simultaneous time-jittered marine acquisition to time-lapse surveys for realistic, off-the-grid acquisitions where the sample points are known but do not coincide with a regular periodic grid. We conduct a series of synthetic seismic experiments with different random realizations of the simultaneous time-jittered marine acquisition to assess the effects of irregular sampling in the field on time-lapse data and demonstrate that dense, high-quality data recoveries are the norm and not the exception. We achieve this by adapting our proposed joint recovery method—a new and economic approach to randomized simultaneous time-lapse data acquisition that exploits transform-domain sparsity and shared information among different time-lapse recordings—to incorporate a regularization operator that maps traces from an irregular grid to a regular periodic grid. The recovery method is a combined operation of source separation, regularization and interpolation, wherein periodic densely sampled and colocated prestack data is recovered from time-compressed, and therefore economical, simultaneous, subsampled and irregular data.

We observe that with decrease in replication between the surveys—i.e., shot points are not replicated amongst the vintages—recovery of time-lapse data improve significantly with little variability in recovery of the time-lapse difference itself. We make this observation assuming source equalization and no significant changes in wave heights, water column velocities or temperature and salinity profiles, etc., amongst the different surveys. We also demonstrate the delicate reliance on exact replicability (between surveys) by showing that known deviations as small as average $\pm(1,2,3)\,\mathrm{m}$ in shot positions of the monitor surveys from the baseline surveys vary recovery quality of the time-lapse difference—expressed as slight decrease or increase in the signal-to-noise ratios—and hence negate the efforts to replicate. Therefore, it would be better to focus on knowing what the shot positions were (post acquisition) than aiming to replicate. Moreover, since irregular spatial sampling is inevitable in the real world, the requirement for replicability in time-lapse surveys can perhaps be relaxed by embracing or better purposefully randomizing the acquisitions to maximize collection of information by effectively doubling the number of measurements for the common component, leading to surveys acquired at low cost and environmental imprint.

# Chapter 7

# Source separation for simultaneous towed-streamer marine acquisition — a compressed sensing approach

## 7.1 Summary

Simultaneous marine acquisition is an economic way to sample seismic data and speedup acquisition, wherein single and/or multiple source vessels fire sources at near-simultaneous or slightly random times, resulting in overlapping shot records. The current paradigm for simultaneous towed-streamer marine acquisition incorporates "low-variability" in source firing times—i.e., $0 \leq 1$ or 2 seconds, since both the sources and receivers are moving. This results in low degree of randomness in simultaneous data, which is challenging to separate (into its constituent sources) using compressed sensing based separation techniques since randomization is the key to successful recovery via compressed sensing. We address the challenge of source separation for simultaneous towed-streamer acquisitions via two compressed sensing based approaches—i.e., sparsity-promotion and rank-minimization. We illustrate the performance of both the sparsity-promotion and rank-minimization based techniques by simulating two simultaneous towed-streamer acquisition scenarios—i.e., over/under and simultaneous long offset. A field data example from the Gulf of Suez for the over/under acquisition scenario is also included. We observe that the proposed approaches give good and comparable recovery qualities of the separated sources, but the rank-minimization technique outperforms the sparsity-promoting technique in terms of the computational time and memory. We also compare these two techniques with the NMO-based median filtering type approach.

## 7.2 Introduction

The benefits of simultaneous source marine acquisition are manifold—it allows the acquisition of improved-quality seismic data at standard (conventional) acquisition turnaround, or a reduced turnaround time while maintaining similar quality, or a combination of both advantages. In simultaneous marine acquisition, a single or multiple source vessels fire sources at near-simultaneous or slightly random times resulting in overlapping shot records (de Kok and Gillespie, 2002; Beasley,

---

A version of this chapter has been published in Geophysics, 2015, vol. 80, pp. WD73–WD88.

2008; Berkhout, 2008; Hampson et al., 2008; Moldoveanu and Quigley, 2011; Abma et al., 2013), as opposed to nonoverlapping shot records in conventional marine acquisition. A variety of simultaneous source survey designs have been proposed for towed-streamer and ocean bottom acquisitions, where small-to-large random time delays between multiple sources have been used (Beasley, 2008; Moldoveanu and Fealy, 2010; Mansour et al., 2012b; Abma et al., 2013; Wason and Herrmann, 2013b; Mosher et al., 2014).

An instance of low-variability in source firing times—e.g., $0 \leq 1$ (or 2) second, is the over/under (or multi-level) source acquisition (Hill et al., 2006; Moldoveanu et al., 2007; Lansley et al., 2007; Long, 2009; Hegna and Parkes, 2012; Hoy et al., 2013). The benefits of acquiring and processing over/under data are clear, the recorded bandwidth is extended at both low and high ends of the spectrum since the depths of the sources produce complementary ghost functions, avoiding deep notches in the spectrum. The over/under acquisition allows separation of the up- and down-going wavefields at the source (or receiver) using a vertical pair of sources (or receivers) to determine wave direction. Simultaneous long offset acquisition (SLO) is another variation of simultaneous towed-streamer acquisition, where an extra source vessel is deployed, sailing one spread-length ahead of the main seismic vessel (Long et al., 2013). The SLO technique is better in comparison to conventional acquisition since it provides longer coverage in offsets, less equipment downtime (doubling the vessel count inherently reduces the streamer length by half), easier maneuvering, and shorter line turns.

Simultaneous acquisition (e.g., over/under and SLO) results in seismic interferences or source crosstalk that degrades quality of the migrated images. Therefore, an effective (simultaneous) source separation technique is required, which aims to recover unblended interference-free data—as acquired during conventional acquisition—from simultaneous data. The challenge of source separation (or deblending) has been addressed by many researchers (Stefani et al., 2007; Moore et al., 2008; Akerberg et al., 2008; Huo et al., 2009), wherein the key observation has been that as long as the sources are fired at suitably randomly dithered times, the resulting interferences (or source crosstalk) will appear noise-like in specific gather domains such as common-offset and common-receiver, turning the separation problem into a (random) noise removal procedure. Inversion-type algorithms (Moore, 2010; Abma et al., 2010; Mahdad et al., 2011; Doulgeris et al., 2012; Baardman and van Borselen, 2013) take advantage of sparse representations of coherent seismic signals. Wason and Herrmann (2013a); Wason and Herrmann (2013b) proposed an alternate sampling strategy for simultaneous acquisition (*time-jittered marine*) that leverages ideas from compressed sensing (CS), addressing the source-separation problem through a combination of tailored (blended) acquisition design and sparsity-promoting recovery via convex optimization using one-norm constraints. This represents a scenario of high-variability in source firing times—e.g., $> 1$ second, resulting in irregular shot locations.

One of the source separation techniques is the normal moveout based median filtering, where the key idea is as follows: *i*) transform the blended data into the midpoint-offset domain, *ii*) perform semblance analysis on common-midpoint gathers to pick the normal moveout (NMO) velocities followed by NMO corrections, *iii*) perform median filtering along the offset directions and then apply inverse NMO corrections. One of the major assumptions in the described workflow is that the seismic events become flat after NMO corrections, however, this can be challenging when the geology is complex and/or with the presence of noise in the data. Therefore, the above process along with the velocity analysis is repeated a couple of times to get a good velocity model to eventually separate simultaneous data.

Recently, rank-minimization based techniques have been used for source separation by Maraschini et al. (2012) and Cheng and Sacchi (2013). The general idea is to exploit the low-rank structure of seismic data when it is organized in a matrix. Low-rank structure refers to the small number of nonzero singular values, or quickly decaying singular values. Maraschini et al. (2012) followed the rank-minimization based approach proposed by Oropeza and Sacchi (2011), who identified that seismic temporal frequency slices organized into a block Hankel matrix, in ideal conditions, is a matrix of rank $k$, where $k$ is the number of different plane waves in the window of analysis. Oropeza and Sacchi (2011) showed that additive random noise increase the rank of the block Hankel matrix and presented an iterative algorithm that resembles seismic data reconstruction with the method of projection onto convex sets, where they use a low-rank approximation of the Hankel matrix via the randomized singular value decomposition (Liberty et al., 2007; Halko et al., 2011) to interpolate seismic temporal frequency slices. While this technique may be effective the approach requires embedding the data into an even larger space where each dimension of size $n$ is mapped to a matrix of size $n \times n$. Consequently, these approaches are applied on small data windows, where one has to choose the size of these windows. Although mathematically desirable due to the seismic signal being stationary in sufficiently small windows, Kumar et al. (2015a) showed that the act of windowing from a matrix-rank point of view degrades the quality of reconstruction in the case of missing-trace interpolation. Choosing window sizes a priori is also a difficult task, as it is not altogether obvious how to ensure that the resulting subvolume is approximately a plane wave.

### 7.2.1 Motivation

The success of CS hinges on randomization of the acquisition, as presented in our previous work on simultaneous source acquisition (Mansour et al., 2012b; Wason and Herrmann, 2013b), which represents a case of high-variability in source firing times—e.g., within a range of 1-20 seconds, resulting in overlapping shot records that lie on irregular spatial grids. Consequently, this made our method applicable to marine acquisition with ocean bottom cables/nodes. Successful separation of simultaneous data by sparse inversion via one-norm minimization, in this high-variability scenario, motivated us to analyze the performance of our separation algorithm for the low-variability, simultaneous towed-streamer acquisitions. We address the challenge of source separation for two types of simultaneous towed-streamer marine acquisition—over/under and simultaneous long offset. We also compare the sparsity-promoting separation technique with separation via rank-minimization based technique, since the latter is relatively computationally faster and memory efficient, as shown by Kumar et al. (2015a) for missing-trace interpolation.

### 7.2.2 Contributions

Our contributions in this work are the following: first, we propose a practical framework for source separation based upon the compressed sensing (CS) theory, where we outline the necessary conditions for separating the simultaneous towed-streamer data using sparsity-promoting and rank-minimization techniques. Second, we show that source separation using the rank-minimization based framework includes a "transform domain" where we exploit the low-rank structure of seismic data. We further establish that in simultaneous towed-streamer acquisition each monochromatic frequency slice of the fully sampled blended data matrix with periodic firing times has low-rank structure in the proposed transform domain. However, uniformly random firing-time delays increase the rank of the resulting frequency slice in this transform domain, which is a necessary condition for successful recovery via rank-minimization based techniques.

Third, we show that seismic frequency slices in the proposed transform domain exhibit low-rank structure at low frequencies, but not at high frequencies. Therefore, in order to exploit the low-rank structure at higher frequencies we adopt the Hierarchical Semi-Separable matrix representation (HSS) method proposed by Chandrasekaran et al. (2006) to represent frequency slices. Finally, we combine the (singular-value-decomposition-free) matrix factorization approach recently developed by Lee et al. (2010) with the Pareto curve approach proposed by Berg and Friedlander (2008). This renders the framework suitable for large-scale seismic data since it avoids the computation of the singular value decomposition (SVD), a necessary step in traditional rank-minimization based methods, which is prohibitively expensive for large matrices.

We simulate two simultaneous towed-streamer acquisitions—over/under and simultaneous long offset, and also use a field data example for over/under acquisition. We compare the recovery in terms of the separation quality, computational time and memory usage. In addition, we also make comparisons with the NMO-based median filtering type technique proposed by Chen et al. (2014).

## 7.3 Theory

Compressed sensing is a signal processing technique that allows a signal to be sampled at sub-Nyquist rate and offers three fundamental principles for successful reconstruction of the original signal from relatively few measurements. The first principle utilizes the prior knowledge that the underlying signal of interest is sparse or compressible in some transform domain—i.e., if only a small number $k$ of the transform coefficients are nonzero or if the signal can be well approximated by the $k$ largest-in-magnitude transform coefficients. The second principle is based upon a sampling scheme that breaks the underlying structure—i.e., decreases the sparsity of the original signal in the transform domain. Once the above two principles hold, a sparsity-promoting optimization problem can be solved in order to recover the fully sampled signal. It is well known that seismic data admit sparse representations by curvelets that capture “wavefront sets” efficiently (see e.g., Smith (1998), Candès and Demanet (2005), Hennenfent and Herrmann (2006) and the references therein).

For high resolution data represented by the $N$-dimensional vector $\mathbf{f}_0 \in \mathbb{R}^N$, which admits a sparse representation $\mathbf{x}_0 \in \mathbb{C}^P$ in some transform domain characterized by the operator $\mathbf{S} \in \mathbb{C}^{P\times N}$ with $P \geq N$, the sparse recovery problem involves solving an underdetermined system of equations:

$$\mathbf{b} = \mathbf{A}\mathbf{x}_0, \tag{7.1}$$

where $\mathbf{b} \in \mathbb{C}^n$, $n \ll N \leq P$, represents the compressively sampled data of $n$ measurements, and $\mathbf{A} \in \mathbb{C}^{n\times P}$ represents the measurement matrix. We denote by $\mathbf{x}_0$ a sparse synthesis coefficient vector of $\mathbf{f}_0$. When $\mathbf{x}_0$ is strictly sparse—i.e., only $k < n$ nonzero entries in $\mathbf{x}_0$, sparsity-promoting recovery can be achieved by solving the $\ell_0$ minimization problem, which is a combinatorial problem and quickly becomes intractable as the dimension increases. Instead, the basis pursuit denoise (BPDN$_\epsilon$) convex optimization problem:

$$\underset{\mathbf{x}\in\mathbb{C}^P}{\text{minimize}} \quad \|\mathbf{x}\|_1 \quad \text{subject to} \quad \|\mathbf{b}-\mathbf{A}\mathbf{x}\|_2 \leq \epsilon, \tag{BPDN$_\epsilon$}$$

can be used to recover $\tilde{\mathbf{x}}$, which is an estimate of $\mathbf{x}_0$. Here, $\epsilon$ represents the error-bound in the least-squares misfit and the $\ell_1$ norm $\|\mathbf{x}\|_1$ is the sum of absolute values of the elements of a vector $\mathbf{x}$. The matrix $\mathbf{A}$ can be composed of the product of an $n \times N$ sampling (or acquisition) matrix $\mathbf{M}$ and the sparsifying operator $\mathbf{S}$ such that $\mathbf{A} := \mathbf{M}\mathbf{S}^H$, here $^H$ denotes the Hermitian transpose. Consequently, the measurements are given by $\mathbf{b} = \mathbf{A}\mathbf{x}_0 = \mathbf{M}\mathbf{f}_0$. A seismic line with $N_s$ sources, $N_r$ receivers, and $N_t$ time samples can be reshaped into an $N$ dimensional vector $\mathbf{f}$, where $N = N_s \times N_r \times N_t$. For simultaneous towed-streamer acquisition, given two unblended data vectors $\mathbf{x}_1$ and $\mathbf{x}_2$ and (blended) measurements $\mathbf{b}$, we can redefine Equation 7.1 as

$$\overbrace{\begin{bmatrix}\mathbf{MT_1S}^H & \mathbf{MT_2S}^H\end{bmatrix}}^{\mathbf{A}} \overbrace{\begin{bmatrix}\mathbf{x}_1 \\ \mathbf{x}_2\end{bmatrix}}^{\mathbf{x}} = \mathbf{b}, \tag{7.2}$$

where $\mathbf{T}_1$ and $\mathbf{T}_2$ are defined as the firing-time delay operators which apply uniformly random time delays to the first and second source, respectively. Note that accurate knowledge of the firing times is essential for successful recovery by the proposed source separation techniques. We wish to recover a sparse approximation $\tilde{\mathbf{f}}$ of the discretized wavefield $\mathbf{f}$ (corresponding to each source) from the measurements $\mathbf{b}$. This is done by solving the BPDN$_\epsilon$ sparsity-promoting program, using the SPG$\ell_1$ solver (see Berg and Friedlander, 2008; Hennenfent et al., 2008, for details), yielding $\tilde{\mathbf{f}} = \mathbf{S}^H\tilde{\mathbf{x}}$ for each source.

Sparsity is not the only structure seismic data exhibits where three- or five-dimensional seismic data is organized as a vector. High-dimensional seismic data volumes can also be represented as matrices or tensors, where the low-rank structure of seismic data can be exploited (Trickett and Burroughs, 2009; Oropeza and Sacchi, 2011; Kreimer and Sacchi, 2012; Silva and Herrmann, 2013; Aravkin et al., 2014). This low-rank property of seismic data leads to the notion of matrix completion theory which offers a reconstruction strategy for an unknown matrix $\mathbf{X}$ from its known subsets of entries (Candès and Recht, 2009; Recht et al., 2010). The success of matrix completion framework hinges on the fact that regularly sampled target dataset should exhibit a low-rank structure in the rank-revealing "transform domain" while subsampling should destroy the low-rank structure of seismic data in the transform domain.

### 7.3.1 Rank-revealing "transform domain"

Following the same analogy of CS, the main challenge in applying matrix completion techniques to the source separation problem is to find a "transform domain" wherein: *i*) fully sampled conventional (or unblended) seismic data have low-rank structure—i.e., quickly decaying singular values; *ii*) blended seismic data have high-rank structure—i.e., slowly decaying singular values. When these properties hold, rank-minimization techniques (used in matrix completion) can be used to recover the source-separated signal. Kumar et al. (2013) showed that the frequency slices of unblended seismic data do not exhibit low-rank structure in the source-receiver (s-r) domain since strong wavefronts extend diagonally across the s-r plane. However, transforming the data into the midpoint-offset (m-h) domain results in a vertical alignment of the wavefronts, thereby reducing the rank of the frequency slice matrix. The midpoint-offset domain is a coordinate transformation defined as:

$$\begin{aligned} x_{midpoint} &= \frac{1}{2}(x_{source} + x_{receiver}), \\ x_{offset} &= \frac{1}{2}(x_{source} - x_{receiver}). \end{aligned}$$

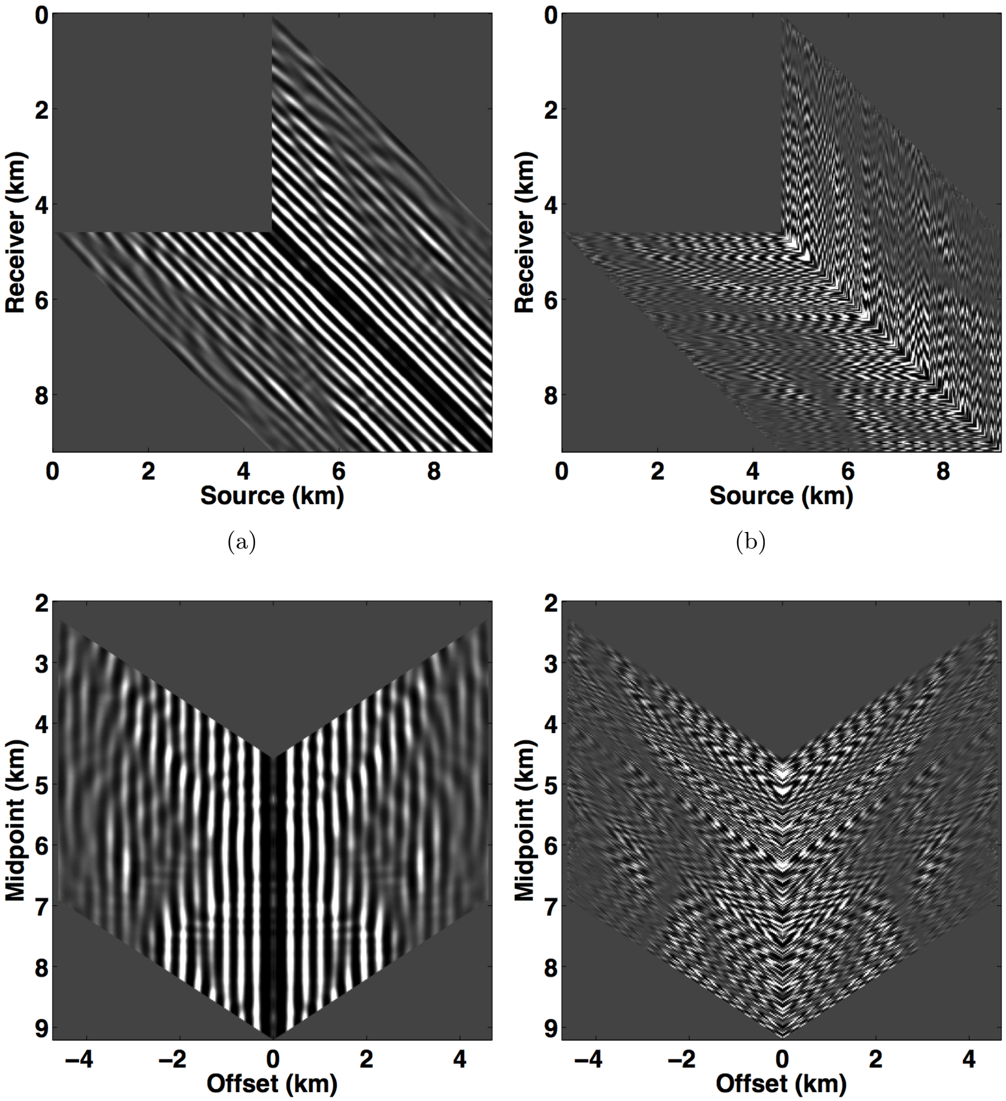
These observations motivate us to exploit the low-rank structure of seismic data in the midpoint-offset domain for simultaneous towed-streamer acquisition. Although the given problem does not contain any missing traces and is a source-separation problem alone, incorporating this scenario in the current framework is straightforward (Kumar et al., 2017). Seismic data processing literature contains numerous works done on missing-trace interpolation in midpoint-offset coordinates (see Trad, 2009; Kreimer, 2013, and references therein).

Figures 7.1(a) and 7.1(c) show a monochromatic frequency slice (at 5 Hz) for simultaneous acquisition with periodic firing times in the source-receiver (s-r) and midpoint-offset (m-h) domains, while Figures 7.1(b) and 7.1(d) show the same for simultaneous acquisition with random firing-time delays. Note that we use the source-receiver reciprocity to convert each monochromatic frequency slice of the towed-streamer acquisition to split-spread type acquisition, which is required by our current implementation of rank-minimization based techniques for 2D seismic acquisition. For 3D seismic data acquisition, where seismic data exhibit 5D structure, we can follow the strategy proposed by Kumar et al. (2015a), where a simple permutation of matricized seismic data is used as a transformation domain to exploit the low-rank structure of seismic data. Here, matricization refers to a process that reshapes a tensor into a matrix along specific dimensions. As shown in Kumar et al. (2015a), 3D seismic data exhibits low-rank structure in the noncanonical matrix representation, where the source and receiver coordinates in the $x-$ and $y-$ direction are lumped together as opposed to the canonical matrix representation, where both the source coordinates (in the $x-$ and $y-$ direction) are lumped together, and similarly for the receiver coordinates. Therefore, in 3D seismic data acquisition we do not have to work in the midpoint-offset domain which removes the requirement of source-receiver reciprocity.

As illustrated in Figure 7.1, simultaneously acquired data with periodic firing times preserves continuity of the waveforms in the s-r and m-h domains, which inherently do not change the rank of blended data compared to unblended data. Introducing random time delays destroys continuity of the waveforms in the s-r and m-h domains, thus increasing the rank of the blended data matrix drastically, which is a necessary condition for rank-minimization based algorithms to work effectively. To illustrate this behaviour, we plot the decay of the singular values of a 5 Hz monochromatic frequency slice extracted from the periodically and randomized simultaneous acquisition in the s-r and m-h domains, respectively in Figure 7.2(a). Note that uniformly random firing-time delays do not noticeably change the decay of the singular values in the source-receiver (s-r) domain, as expected, but significantly slow down the decay rate in the m-h domain.

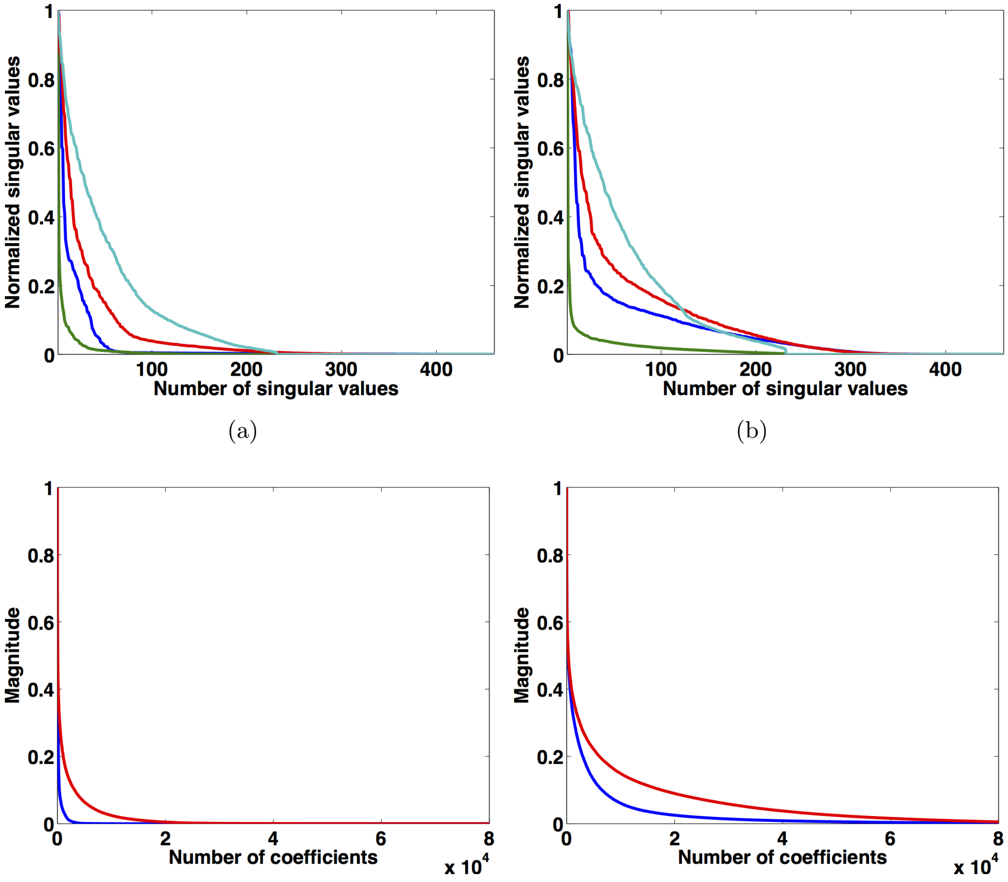
Similar trends are observed for a monochromatic frequency slice at 40 Hz in Figure 7.2(b). Following the same analogy, Figures 7.2(c) and 7.2(d) show how randomization in acquisition destroys the sparse structure of seismic data in the source-channel (or source-offset) domain—i.e., slow decay of the curvelet coefficients, hence, favouring recovery via sparsity-promotion in this domain. Similarly, for simultaneous long offset acquisition, we exploit the low-rank structure of seismic data in the m-h domain, and the sparse structure in the source-channel domain.

Seismic frequency slices exhibit low-rank structure in the m-h domain at low frequencies, but the same is not true for data at high frequencies. This is because in the low-frequency slices, the vertical alignment of the wavefronts can be accurately approximated by a low-rank representation. On the other hand, high-frequency slices include a variety of wave oscillations that increase the rank, even though the energy remains focused around the diagonal (Kumar et al., 2013). To illustrate this phenomenon, we plot a monochromatic frequency slice at 40 Hz in the s-r domain and the m-h domain for over/under acquisition in Figure 7.3. When analyzing the decay of the singular



(a) (b) (c) (d)

**Figure 7.1:** Monochromatic frequency slice at 5 Hz in the source-receiver (s-r) and midpoint-offset (m-h) domain for blended data (a,c) with periodic firing times and (b,d) with uniformly random firing times for both sources.



(a) (b)

(c) (d)

**Figure 7.2:** Decay of singular values for a frequency slice at (a) 5 Hz and (b) 40 Hz of blended data. Source-receiver domain: blue—periodic, red—random delays. Midpoint-offset domain: green—periodic, cyan—random delays. Corresponding decay of the normalized curvelet coefficients for a frequency slice at (c) 5 Hz and (d) 40 Hz of blended data, in the source-channel domain.

values for high-frequency slices in the s-r domain and the m-h domain (Figure 7.2(b)), we observe that the singular value decay is slower for the high-frequency slice than for the low-frequency slice. Therefore, rank-minimization in the high-frequency range requires extended formulations that incorporate the low-rank structure.

To exploit the low-rank structure of high-frequency data, we rely on the Hierarchical Semi-Separable matrix representation (HSS) method proposed by Chandrasekaran et al. (2006) to represent frequency slices. The key idea in the HSS representation is that certain full-rank matrices, e.g., matrices that are diagonally dominant with energy decaying along the off-diagonals, can be represented by a collection of low-rank sub-matrices. Kumar et al. (2013) showed the possibility of finding accurate low-rank approximations of sub-matrices of the high-frequency slices by partitioning the data into the HSS structure for missing-trace interpolation. Jumah and Herrmann (2014) showed that HSS representations can be used to reduce the storage and computational cost

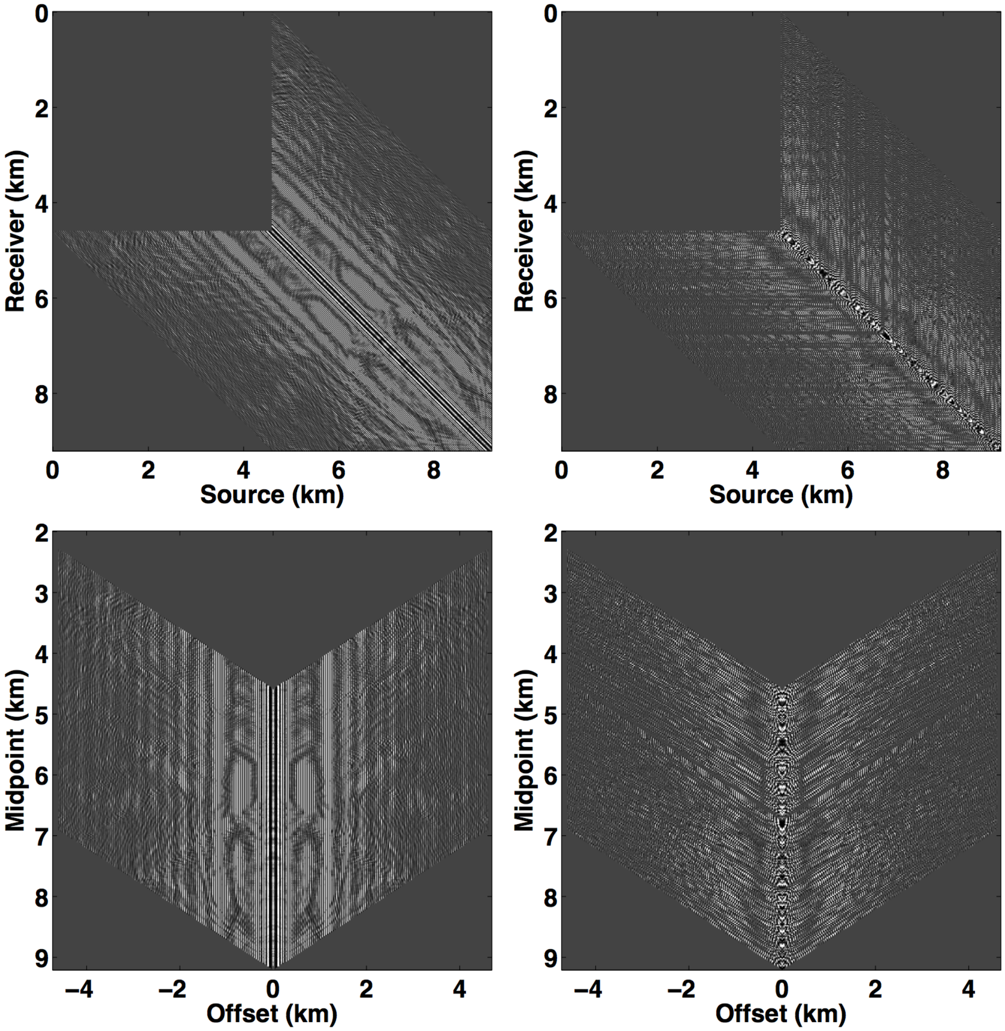
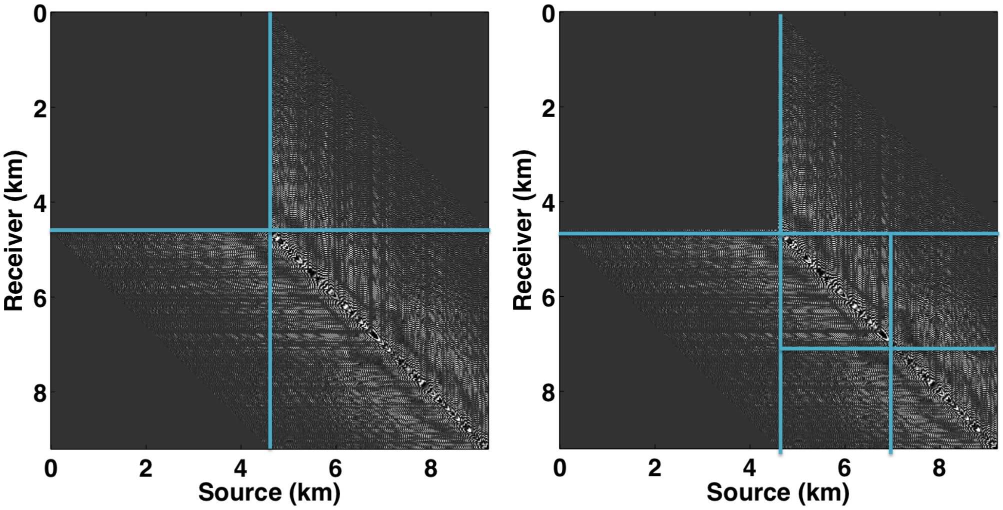


**Figure 7.3:** Monochromatic frequency slice at 40 Hz in the s-r and m-h domain for blended data (a,c) with periodic firing times and (b,d) with uniformly random firing times for both sources.

for the estimation of primaries by sparse inversions. They combined the HSS representation with the randomized SVD proposed by Halko et al. (2011) to accelerate matrix-vector multiplications that are required for sparse inversion.

### 7.3.2 Hierarchical Semi-Separable matrix representation (HSS)

The HSS structure first partitions a matrix into diagonal and off-diagonal sub-matrices. The same partitioning structure is then applied recursively to the diagonal sub-matrices only. To illustrate the HSS partitioning, we consider a 2D monochromatic high-frequency data matrix at 40 Hz in the s-r domain. We show the first-level of partitioning in Figure 7.4(a) and the second-level partitioning in Figure 7.4(b) in their corresponding source-receiver domains. Figures 7.5(a) and 7.5(b) display the first-level off-diagonal sub-blocks, Figure 7.5(c) is the diagonal sub-block, and the corresponding



(a) (b)

**Figure 7.4:** HSS partitioning of a high-frequency slice at 40 Hz in the s-r domain: (a) first-level, (b) second-level, for randomized blended acquisition.

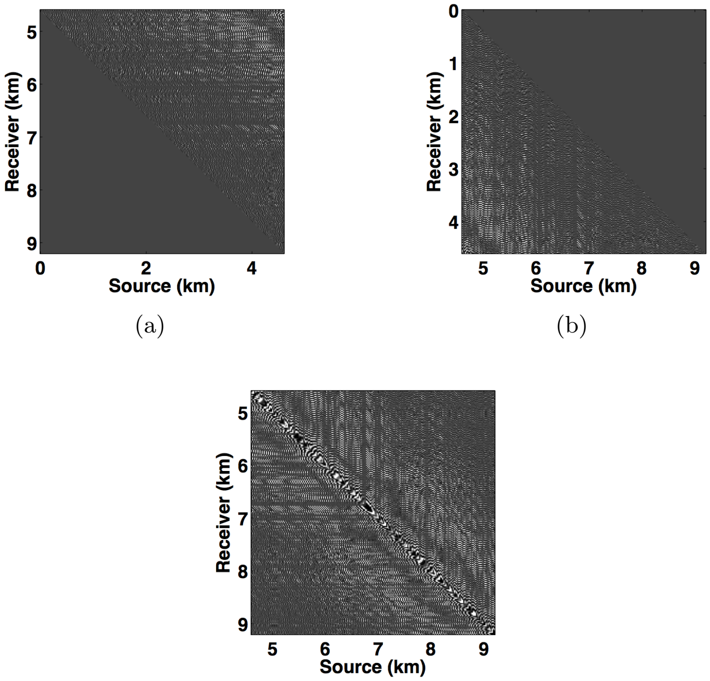
decay of the singular values is displayed in Figure 7.6. We can clearly see that the off-diagonal sub-matrices have low-rank structure, while the diagonal sub-matrices have higher rank. Further partitioning of the diagonal sub-blocks (Figure 7.4(b)) allows us to find better low-rank approximations. The same argument holds for the simultaneous long offset acquisition. Therefore, for low-variability acquisition scenarios, each frequency slice is first partitioned using HSS and then separated in its respective m-h domain, as shown for missing-trace interpolation by Kumar et al. (2013).

One of the limitations of matrix completion type approaches for large-scale seismic data is the nuclear-norm projection, which inherently involves the computation of SVDs. Aravkin et al. (2014) showed that the computation of SVD is prohibitively expensive for large-scale data such as seismic, therefore, we propose a matrix-factorization based approach to avoid the need for expensive computation of SVDs (see Aravkin et al., 2014, for details). In the next section, we introduce the matrix completion framework and explore its necessary extension to separate large-scale simultaneous seismic data.

### 7.3.3 Large-scale seismic data: SPG-LR framework

Let $\mathbf{X}_0$ be a low-rank matrix in $\mathbb{C}^{n\times m}$ and $\mathcal{A}$ be a linear measurement operator that maps from $\mathbb{C}^{n\times m} \rightarrow \mathbb{C}^p$ with $p \ll n \times m$. Under the assumption that the blending process increases the rank of the matrix $\mathbf{X}_0$, the source separation problem is to find the matrix of lowest possible rank that agrees with the above observations. The rank-minimization problem involves solving the following problem for $\mathcal{A}$, up to a given tolerance $\epsilon$:

$$\underset{\mathbf{X}}{\text{minimize}} \quad \text{rank}(\mathbf{X}) \quad \text{subject to} \quad \|\mathcal{A}(\mathbf{X}) - \mathbf{b}\|_2 \leq \epsilon,$$



(a)

(b)

(c)

**Figure 7.5:** (a,b,c) First-level sub-block matrices (from Figure 7.4(a)).

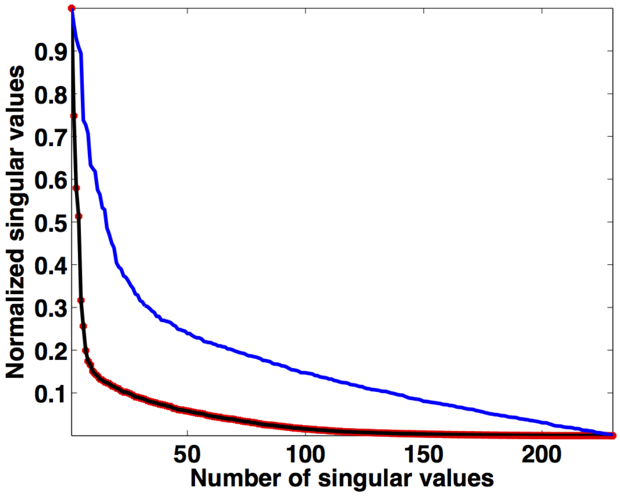


**Figure 7.6:** Decay of singular values of the HSS sub-blocks in s-r domain: red—Figure 7.5(a), black—Figure 7.5(b), blue—Figure 7.5(c).

where rank is defined as the maximum number of linearly independent rows or column of a matrix, $\mathbf{b}$ is a set of blended measurements. For simultaneous towed-streamer acquisition, we follow equation 7.2 and redefine our system of equations as

$$\overbrace{\begin{bmatrix}\mathbf{MT}_1\mathcal{S}^H & \mathbf{MT}_2\mathcal{S}^H\end{bmatrix}}^{\mathcal{A}}\overbrace{\begin{bmatrix}\mathbf{X}_1\\ \mathbf{X}_2\end{bmatrix}}^{\mathbf{X}} = \mathbf{b},$$

where $\mathcal{S}$ is the transformation operator from the s-r domain to the m-h domain. Recht et al. (2010) showed that under certain general conditions on the operator $\mathcal{A}$, the solution to the rank-minimization problem can be found by solving the following nuclear-norm minimization problem:

$$\underset{\mathbf{X}}{\text{minimize}} \quad \|\mathbf{X}\|_* \quad \text{subject to} \quad \|\mathcal{A}(\mathbf{X}) - \mathbf{b}\|_2 \leq \epsilon, \tag{BPDN$_\epsilon$}$$

where $\|\mathbf{X}\|_* = \|\sigma\|_1$, and $\sigma$ is a vector of singular values. Unfortunately, for large-scale data, solving the BPDN$_\epsilon$ problem is difficult since it requires repeated projections onto the set $\|\mathbf{X}\|_* \leq \tau$, which means repeated SVD or partial SVD computations. Therefore, we avoid computing SVDs of the matrices and use an extension of the SPG$\ell_1$ solver (Berg and Friedlander, 2008) developed for the BPDN$_\epsilon$ problem in Aravkin et al. (2013). We refer to this extension as SPG-LR in the rest of the chapter. The SPG-LR algorithm finds the solution to the BPDN$_\epsilon$ problem by solving a sequence of LASSO (least absolute shrinkage and selection operator) subproblems:

$$\underset{\mathbf{X}}{\text{minimize}} \quad \|\mathcal{A}(\mathbf{X}) - \mathbf{b}\|_2 \quad \text{subject to} \quad \|\mathbf{X}\|_* \leq \tau, \tag{LASSO$_\tau$}$$

where $\tau$ is updated by traversing the Pareto curve. The Pareto curve defines the optimal trade-off between the two-norm of the residual and the one-norm of the solution (Berg and Friedlander, 2008). Solving each LASSO subproblem requires a projection onto the nuclear-norm ball $\|\mathbf{X}\|_* \leq \tau$ in every iteration by performing a singular value decomposition and then thresholding the singular values. For large-scale seismic problems, it becomes prohibitively expensive to carry out such a large number of SVDs. Instead, we adopt a recent factorization-based approach to nuclear-norm minimization (Rennie and Srebro, 2005; Lee et al., 2010; Recht and Re, 2013). The factorization approach parametrizes the matrix ($\mathbf{X}_1$, $\mathbf{X}_2$) $\in \mathbb{C}^{n\times m}$ as the product of two low-rank factors ($\mathbf{L}_1$, $\mathbf{L}_2$) $\in \mathbb{C}^{n\times k}$ and ($\mathbf{R}_1$, $\mathbf{R}_2$) $\in \mathbb{C}^{m\times k}$ such that,

$$\mathbf{X} = \begin{bmatrix}\mathbf{L}_1\mathbf{R}_1^H\\ \mathbf{L}_2\mathbf{R}_2^H\end{bmatrix}. \tag{7.3}$$

Here, $k$ represents the rank of the $\mathbf{L}$ and $\mathbf{R}$ factors. The optimization scheme can then be carried out using the factors $(\mathbf{L}_1, \mathbf{L}_2)$ and $(\mathbf{R}_1, \mathbf{R}_2)$ instead of $(\mathbf{X}_1, \mathbf{X}_2)$, thereby significantly reducing the size of the decision variable from $2nm$ to $2k(n+m)$ when $k \ll m, n$. Rennie and Srebro (2005) showed that the nuclear-norm obeys the relationship:

$$\|\mathbf{X}\|_* \quad \leq \quad \frac{1}{2}\left\|\begin{bmatrix}\mathbf{L}_1\\ \mathbf{R}_1\end{bmatrix}\right\|_F^2 + \frac{1}{2}\left\|\begin{bmatrix}\mathbf{L}_2\\ \mathbf{R}_2\end{bmatrix}\right\|_F^2 \quad =: \quad \Phi(\mathbf{L}_1, \mathbf{R}_1, \mathbf{L}_2, \mathbf{R}_2), \tag{7.4}$$

where $\| \cdot \|_F^2$ is the Frobenius norm of the matrix—i.e., sum of the squared entires. Consequently, the LASSO subproblem can be replaced by

$$\underset{\mathbf{L}_1, \mathbf{R}_1, \mathbf{L}_2, \mathbf{R}_2}{\text{minimize}} \quad \|\mathcal{A}(\mathbf{X}) - \mathbf{b}\|_2 \quad \text{subject to} \quad \Phi(\mathbf{L}_1, \mathbf{R}_1, \mathbf{L}_2, \mathbf{R}_2) \leq \tau \ , \tag{7.5}$$

where the projection onto $\Phi(\mathbf{L}_1, \mathbf{R}_1, \mathbf{L}_2, \mathbf{R}_2) \leq \tau$ is easily achieved by multiplying each factor $(\mathbf{L}_1, \mathbf{L}_2)$ and $(\mathbf{R}_1, \mathbf{R}_2)$ by the scalar $\sqrt{2\tau / \Phi(\mathbf{L}_1, \mathbf{R}_1, \mathbf{L}_2, \mathbf{R}_2)}$. Equation 7.4, for each HSS sub-matrix in the m-h domain, guarantees that $\|\mathbf{X}\|_* \leq \tau$ for any solution of 7.5. Once the optimization problem is solved, each sub-matrix in the m-h domain is transformed back into the s-r domain, where we concatenate all the sub-matrices to get the separated monochromatic frequency data matrices. One of the advantages of the HSS representation is that it works with recursive partitioning of a matrix and sub-matrices can be solved in parallel, speeding up the optimization formulation.

## 7.4 Experiments

We perform source separation for two simultaneous towed-streamer acquisition scenarios—over/under and simultaneous long offset, by generating synthetic datasets on complex geological models using the IWAVE (Symes et al., 2011) time-stepping acoustic simulation software, and also use a field dataset from the Gulf of Suez. Source separation for over/under acquisition is tested on two different datasets. The first dataset is simulated on the Marmousi model (Bourgeois et al., 1991), which represents a complex-layer model with steeply dipping reflectors that make the data challenging. With a source (and channel/receiver) sampling of 20.0 m, one dataset is generated with a source-depth of 8.0 m (Figures 7.7(a) and 7.7(d)), while the other dataset has the source at 12.0 m depth (Figures 7.7(b) and 7.7(e)), resulting in 231 sources and 231 channels. The temporal length of each dataset is 4.0 s with a sampling interval of 0.004 s. The second dataset is a field data example from the Gulf of Suez. In this case, the first source is placed at 5.0 m depth (Figures 7.8(a) and 7.8(d)) and the second source is placed at 10.0 m depth (Figures 7.8(b) and 7.8(e)). The source (and channel) sampling is 12.5 m, resulting in 178 sources and 178 channels with a time sampling interval of 0.004 s.

The simultaneous long offset acquisition is simulated on the BP salt model (Billette and Brandsberg-Dahl, 2004), where the presence of salt-bodies make the data challenging. The two source vessels are 6.0 km apart and the streamer length is 6.0 km. Both the datasets (for source 1 and source 2) contain 361 sources and 361 channels with a spatial interval of 12.5 m, where the source and streamer depth is 6.25 m. The temporal length of each dataset is 6.0 s with a sampling interval of 0.006 s. A single shot gather from each dataset is shown in Figures 7.9(a) and 7.9(b) and the corresponding channel gathers are shown in Figures 7.9(d) and 7.9(e). The datasets for each source in both the acquisition scenarios are (simply) summed for simultaneous acquisition with periodic firing times, while uniformly random time delays between 0-1 second are applied to each source for the randomized simultaneous acquisition. Figures 7.7(c), 7.8(c) and 7.9(c) show the randomized blended shot gathers for the Marmousi, the Gulf of Suez and the BP datasets, respectively. As illustrated in the figures, both the sources fire at random times (independent of each other) within the interval of 0-1 second, hence, the difference between the firing times of the sources is always less than 1 second. The corresponding randomized blended channel gathers are shown in Figures 7.7(f), 7.8(f) and 7.9(f). Note that the speed of the vessels in both the acquisition scenarios is no different than the current practical speed of the vessels in the field.

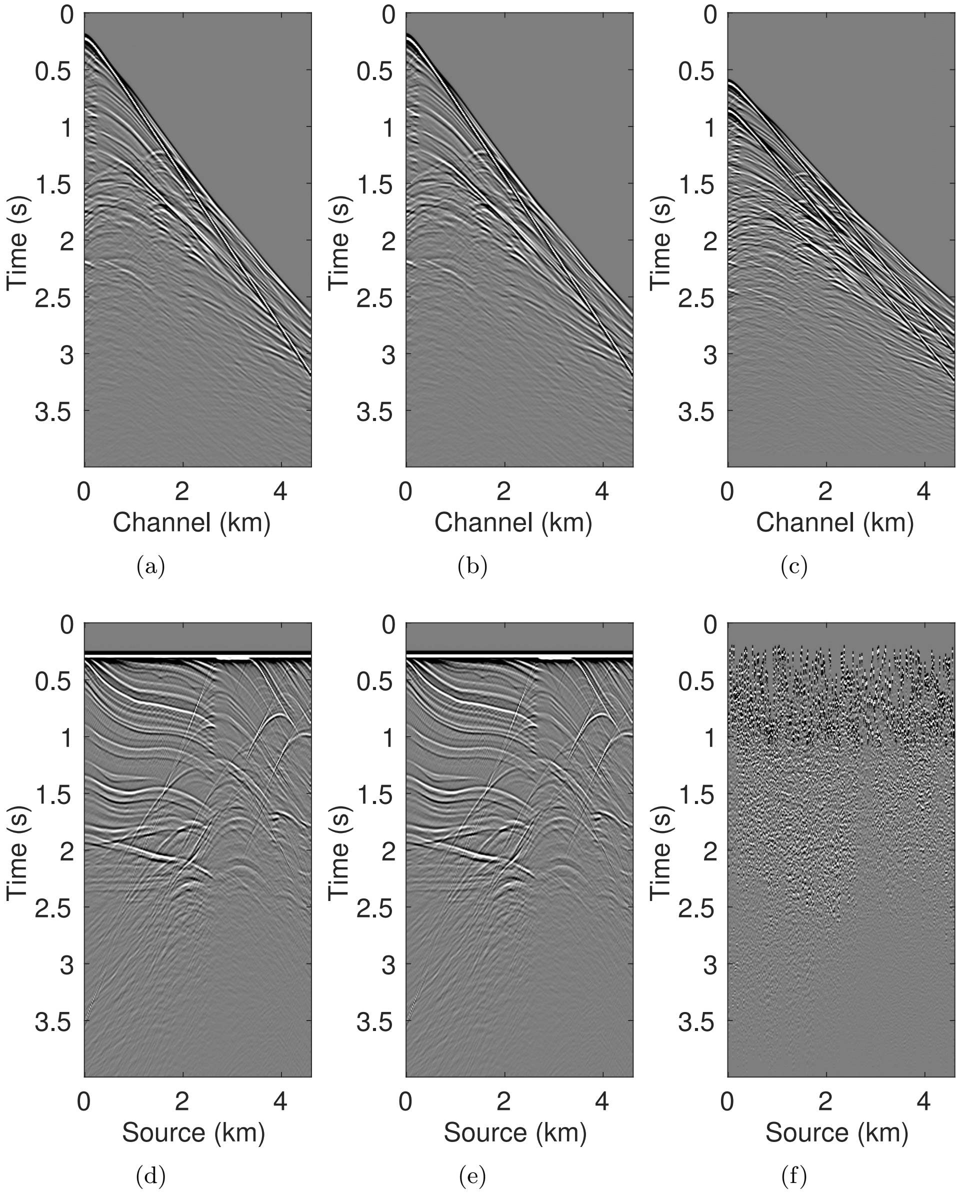
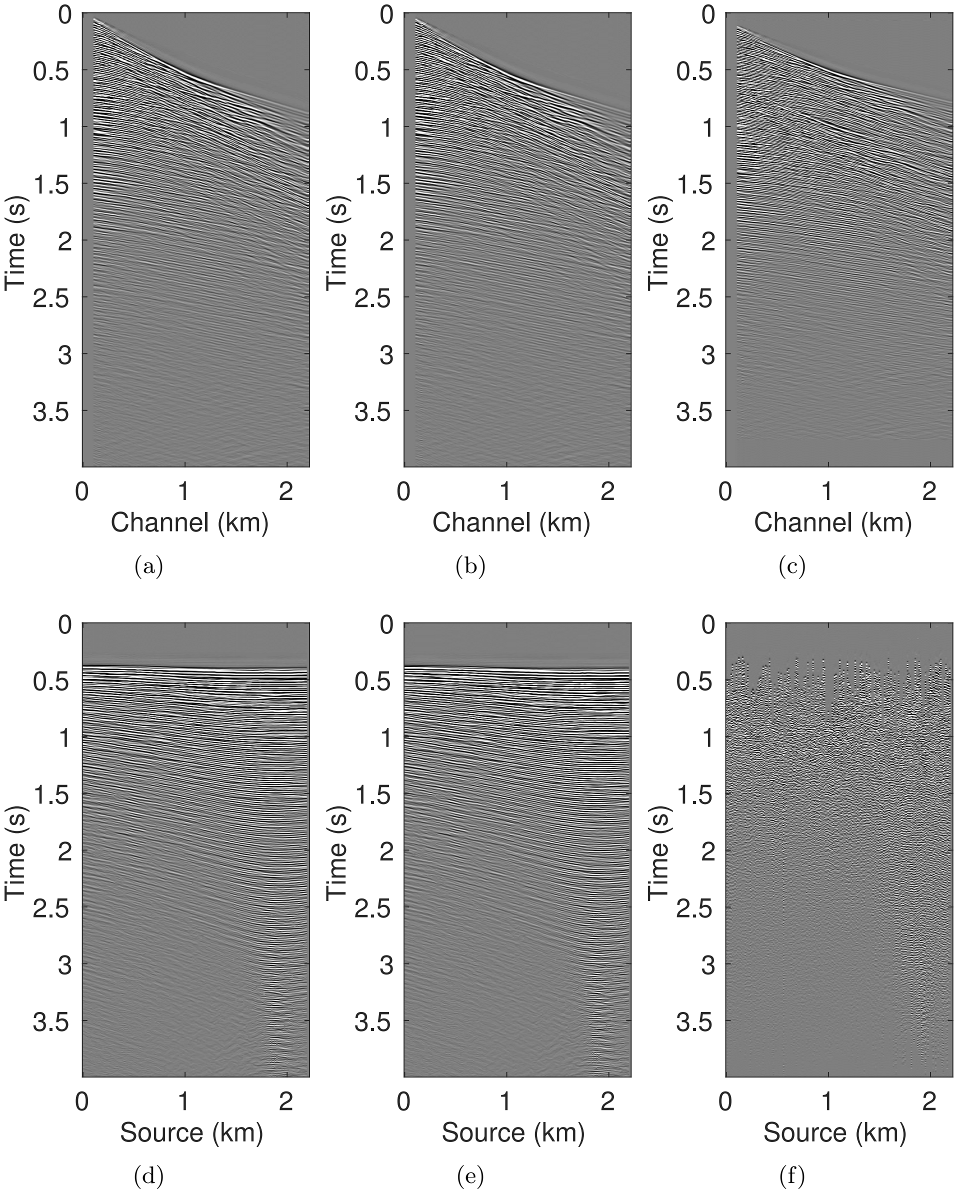


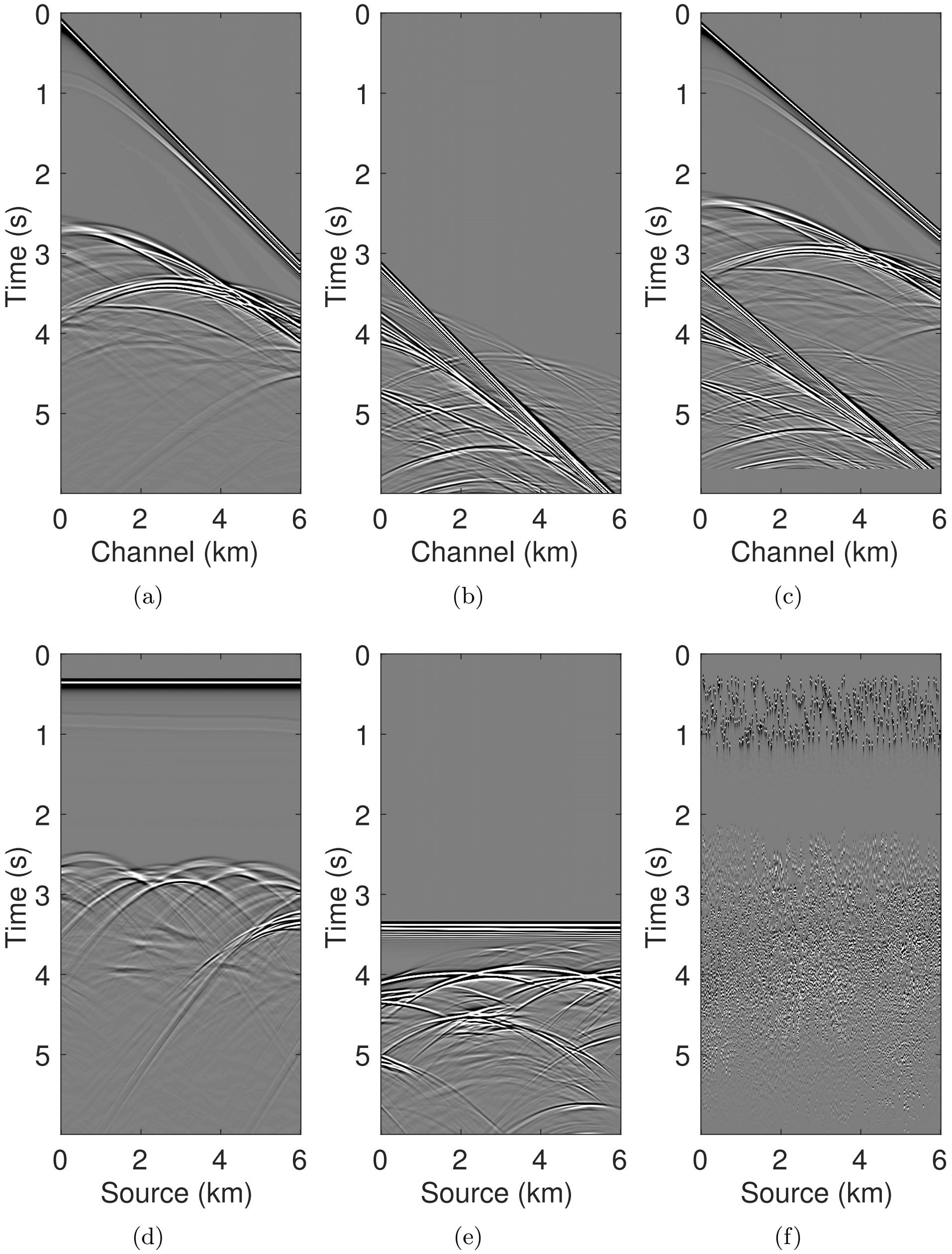
**Figure 7.7:** Original shot gather of (a) source 1, (b) source 2, and (c) the corresponding blended shot gather for simultaneous over/under acquisition simulated on the Marmousi model. (d, e) Corresponding common-channel gathers for each source and (f) the blended common-channel gather.



(a) (b) (c)

(d) (e) (f)

**Figure 7.8:** Original shot gather of (a) source 1, (b) source 2, and (c) the corresponding blended shot gather for simultaneous over/under acquisition from the Gulf of Suez dataset. (d,e) Corresponding common-channel gathers for each source and (f) the blended common-channel gather.



(a) (b) (c)

(d) (e) (f)

**Figure 7.9:** Original shot gather of (a) source 1, (b) source 2, and (c) the corresponding blended shot gather for simultaneous long offset acquisition simulated on the BP salt model. (d, e) Corresponding common-channel gathers for each source and (f) the blended common-channel gather.

For source separation via rank-minimization, second-level of HSS partitioning, on each frequency slice in the s-r domain, was sufficient for successful recovery in both the acquisition scenarios. After transforming each sub-block into the m-h domain, source separation is then performed by solving the nuclear-norm minimization formulation ($\text{BPDN}_\epsilon$) on each sub-block, using 350 iterations of SPG-LR. In order to choose an appropriate rank value, we first perform source separation for frequency slices at 0.2 Hz and 125 Hz. For the over/under acquisition simulated on the Marmousi model, the best rank value is 30 and 80 for each frequency slice, respectively. The best rank values for the Gulf of Suez dataset are 20 and 100, respectively. For simultaneous long offset acquisition, the best rank value is 10 and 90 for frequency slices at 0.15 Hz and 80 Hz, respectively. Hence, we adjust the rank linearly within these ranges when moving from low to high frequencies, for each acquisition scenario. For source separation via sparsity-promotion, we use the $\text{BPDN}_\epsilon$ formulation to minimize the $\ell_1$ norm (instead of the nuclear-norm) where the transformation operator $\mathcal{S}$ is the 2D curvelet operator. Here, we run 350 iterations of $\text{SPG}\ell_1$.

For the over/under acquisition scenario simulated on the Marmousi model, Figures 7.10(a) and 7.10(c) show the separated shot gathers via rank-minimization and Figures 7.10(e) and 7.10(g) show the separated shot gathers via sparsity-promotion, respectively. The separated common-channel gathers via rank-minimization and sparsity-promotion are shown in Figures 7.11(a), 7.11(c) and Figures 7.11(e), 7.11(g), respectively. For the Gulf of Suez field dataset, Figures 7.12 and 7.13 show the separated gathers and difference plots in the common-shot and common-channel domain, respectively. The corresponding separated gathers and difference plots in the common-shot and common-channel domain for the simultaneous long offset acquisition scenario are shown in Figures 7.14 and 7.15.

As illustrated by the results and their corresponding difference plots, both the CS-based approaches of rank-minimization and sparsity-promotion are able to separate the data for the low-variability acquisition scenarios fairly well. In all the three different datasets, the average S/Ns for separation via sparsity-promotion is slightly better than rank-minimization, but the difference plots show that the recovery via rank-minimization is equivalent to the sparsity-promoting based recovery where it is able to recover most of the coherent energy. Also, rank-minimization outperforms the sparsity-promoting technique in terms of the computational time and memory usage as represented in Tables 7.1, 7.2 and 7.3. Both the CS-based recoveries are better for the simultaneous long offset acquisition than the recoveries from the over/under acquisition scenario. A possible explanation for this improvement is the long offset distance that increases randomization in the simultaneous acquisition, which is a more favourable scenario for recovery by CS-based approaches. Figure 7.16 demonstrates the advantage of the HSS partitioning, where the S/Ns of the separated data are significantly improved.

### 7.4.1 Comparison with NMO-based median filtering

We also compare the performance of our CS-based source-separation techniques with deblending using the NMO-based median filtering technique proposed by Chen et al. (2014), where we work on a common-midpoint gather from each acquisition scenario. For the over/under acquisition simulated on the Marmousi model, Figures 7.17(a) and 7.17(e) show the blended common-midpoint gathers and source separation using the median filtering technique is shown in Figures 7.17(b) and 7.17(f). The corresponding separated common-midpoint gathers from the two CS-based techniques are shown in Figures 7.17(c,d,g,h). Figure 7.18 shows the blended and separated common-midpoint gathers for the field data from the Gulf of Suez. We observe that recoveries via the proposed CS-

| Marmousi model | | | |
|---|---|---|---|
| | Time | Memory | S/N |
| Sparsity | 167 | 7.0 | 16.7, 16.7 |
| Rank | 12 | 2.8 | 15.0, 14.8 |

**Table 7.1:** Comparison of computational time (in hours), memory usage (in GB) and average S/N (in dB) using sparsity-promoting and rank-minimization based techniques for the Marmousi model.

| Gulf of Suez | | | |
|---|---|---|---|
| | Time | Memory | S/N |
| Sparsity | 118 | 6.6 | 14.6 |
| Rank | 8 | 2.6 | 12.8 |

**Table 7.2:** Comparison of computational time (in hours), memory usage (in GB) and average S/N (in dB) using sparsity-promoting and rank-minimization based techniques for the Gulf of Suez dataset.

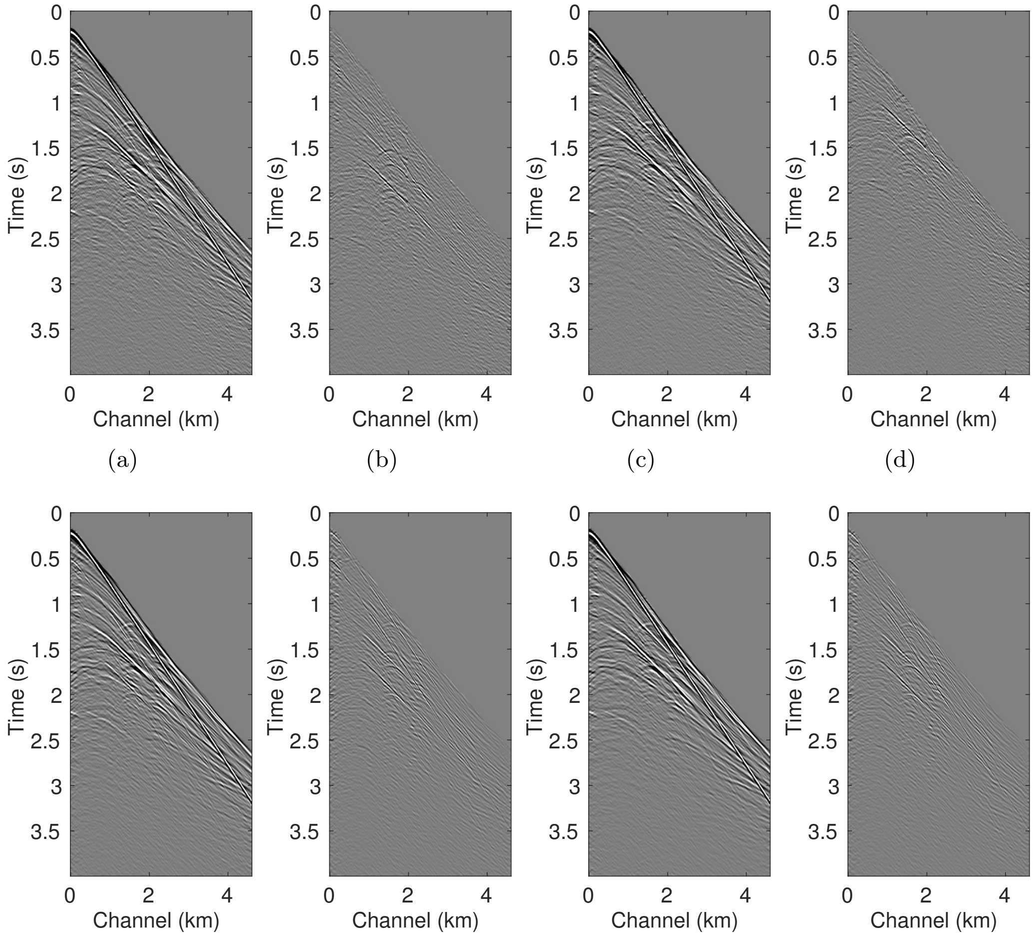
based approaches are comparable to the recovery from the median filtering technique. Similarly, Figure 7.19 shows the results for the simultaneous long offset acquisition simulated on the BP salt model. Here, the CS-based techniques result in slightly improved recoveries.

### 7.4.2 Remark

It is important to note here that we perform the CS-based source separation algorithms only once, however, we can always perform a few more runs of the algorithms where we can first subtract the separated source 1 and source 2 from the acquired blended data and then rerun the algorithms to separate the energy in the residual data. Hence, the recovery can be further improved until necessary. Since separation via rank-minimization is computationally faster than the sparsity based technique, multiple passes through the data is a computationally viable option for the former source-separation technique.

| BP model | | | |
|---|---|---|---|
| | time | memory | S/N |
| Sparsity | 325 | 7.0 | 32.0, 29.4 |
| Rank | 20 | 2.8 | 29.4, 29.0 |

**Table 7.3:** Comparison of computational time (in hours), memory usage (in GB) and average S/N (in dB) using sparsity-promoting and rank-minimization based techniques for the BP model.



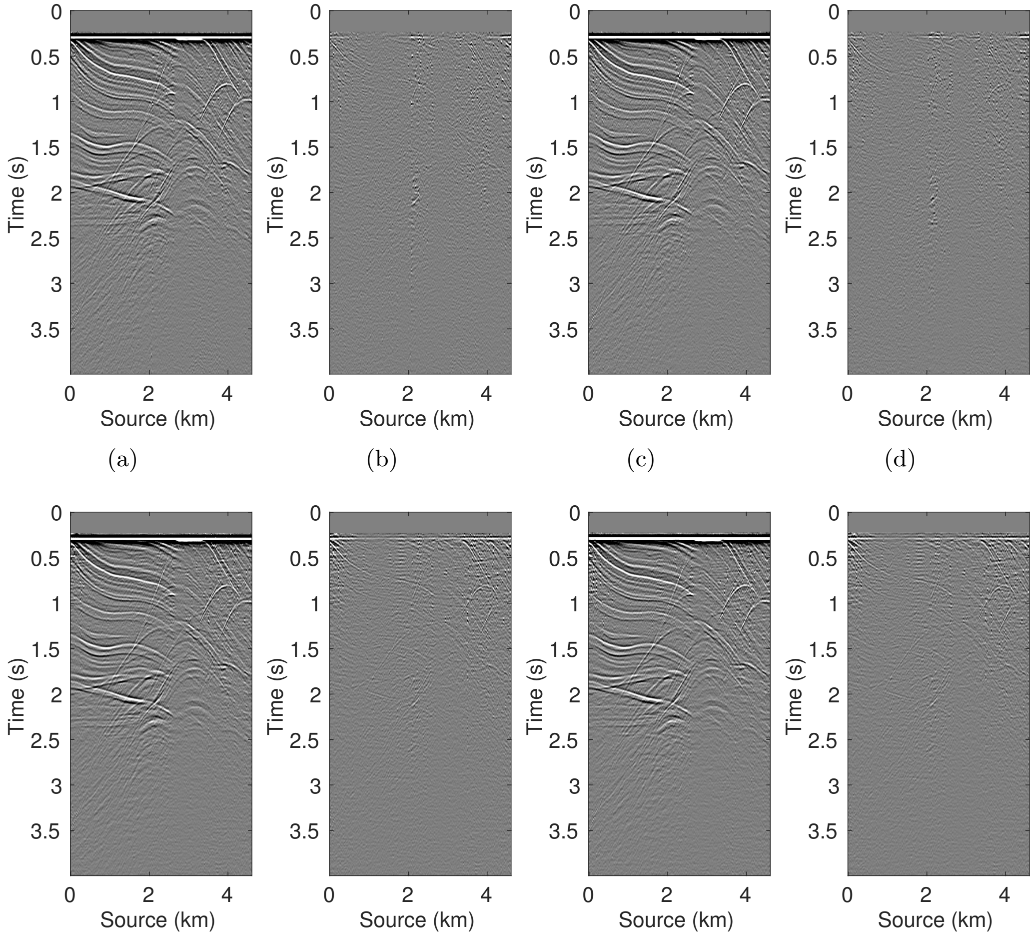
(a) (b) (c) (d)

(e) (f) (g) (h)

**Figure 7.10:** Separated shot gathers and difference plots (from the Marmousi model) of source 1 and source 2: (a,c) source separation using HSS based rank-minimization and (b,d) the corresponding difference plots; (e,g) source separation using curvelet-based sparsity-promotion and (f,h) the corresponding difference plots.

## 7.5 Discussion

The above experiments demonstrate the successful implementation of the proposed CS-based approaches of rank-minimization and sparsity-promotion for source separation in the low-variability, simultaneous towed-streamer acquisitions. The recovery is comparable for both approaches, however, separation via rank-minimization is significantly faster and memory efficient. This is further enhanced by incorporating the HSS partitioning since it allows the exploitation of the low-rank structure in the high-frequency regime, and renders its extension to large-scale data feasible. Note that in the current implementation, we work with each temporal frequency slice and perform the source separation individually. The separation results can further be enhanced by incorporating the information from the previously recovered frequency slice to the next frequency slice, as shown by Mansour et al. (2013) for seismic data interpolation.



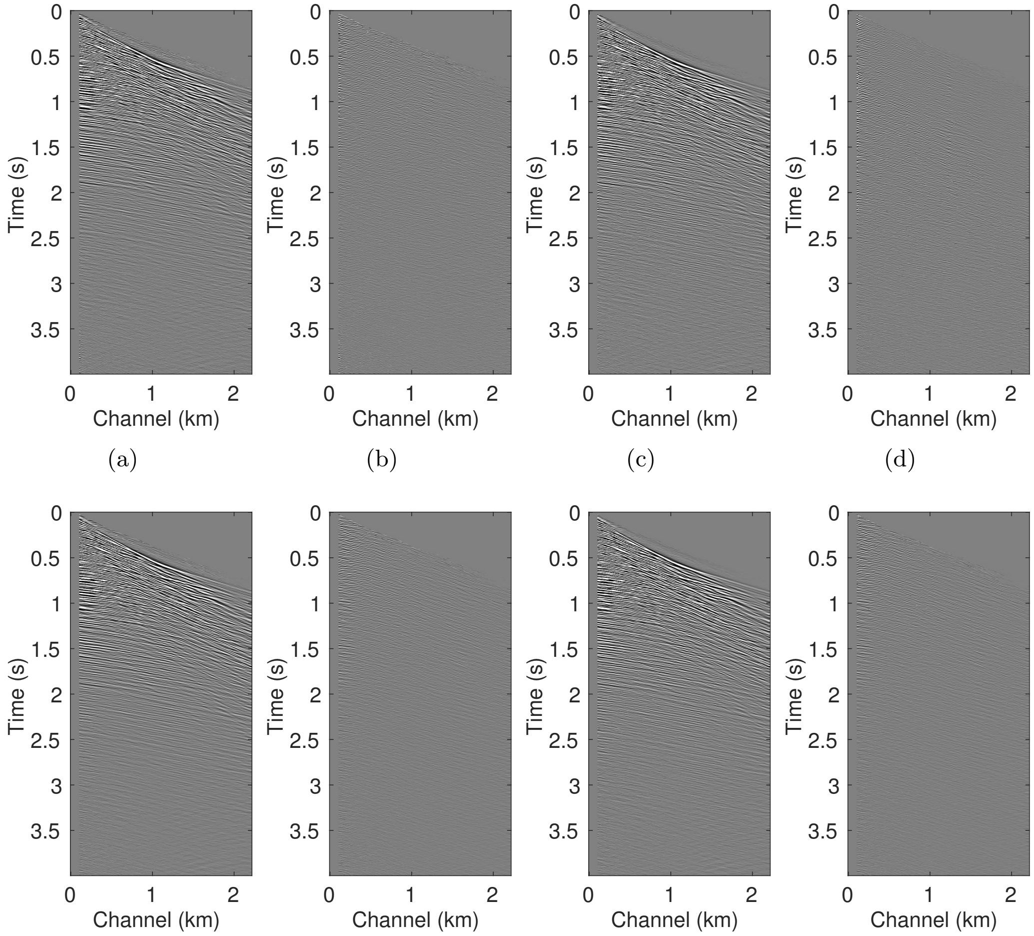
(a) (b) (c) (d)

(e) (f) (g) (h)

**Figure 7.11:** Separated common-channel gathers and difference plots (from the Marmousi model) of source 1 and source 2: (a,c) source separation using HSS based rank-minimization and (b,d) the corresponding difference plots; (e,g) source separation using curvelet-based sparsity-promotion and (f,h) the corresponding difference plots.

The success of CS hinges on randomization of the acquisition. Although, the low degree of randomization (e.g., $0 \leq 1$ second) in simultaneous towed-streamer acquisitions seems favourable for source separation via CS-based techniques, however, high-variability in the firing times enhances the recovery quality of separated seismic data volumes, as shown in Wason and Herrmann (2013a); Wason and Herrmann (2013b) for ocean-bottom cable/node acquisition with continuous recording. One of the advantages of the proposed CS-based techniques is that it does not require velocity estimation, which can be a challenge for data with complex geologies. However, the proposed techniques require accurate knowledge of the random firing times.

So far, we have not considered the case of missing traces (sources and/or receivers), however, incorporating this scenario in the current framework is straightforward. This makes the problem a joint source separation and interpolation problem. We can also extend our methods to separate 3D blended seismic data volumes as shown in Kumar et al. (2017), wherein we address the problem of



(a) (b) (c) (d)

(e) (f) (g) (h)

**Figure 7.12:** Separated shot gathers and difference plots (from the Gulf of Suez dataset) of source 1 and source 2: (a,c) source separation using HSS based rank-minimization and (b,d) the corresponding difference plots; (e,g) source separation using curvelet-based sparsity-promotion and (f,h) the corresponding difference plots.

joint source separation and interpolation for time-lapse seismic. In reality, seismic data are typically irregularly sampled along spatial axes, and therefore future work includes working with nonuniform sampling grids.

## 7.6 Conclusions

We have presented two compressed sensing based methods for source separation for simultaneous towed-streamer type acquisitions, such as the over/under and the simultaneous long offset acquisition. Both the compressed sensing based approaches of rank-minimization and sparsity-promotion give comparable source-separation results, however, the former approach is readily scalable to large-scale blended seismic data volumes and is computationally faster. This can be further enhanced by incorporating the HSS structure with factorization-based rank-regularized optimization formu-

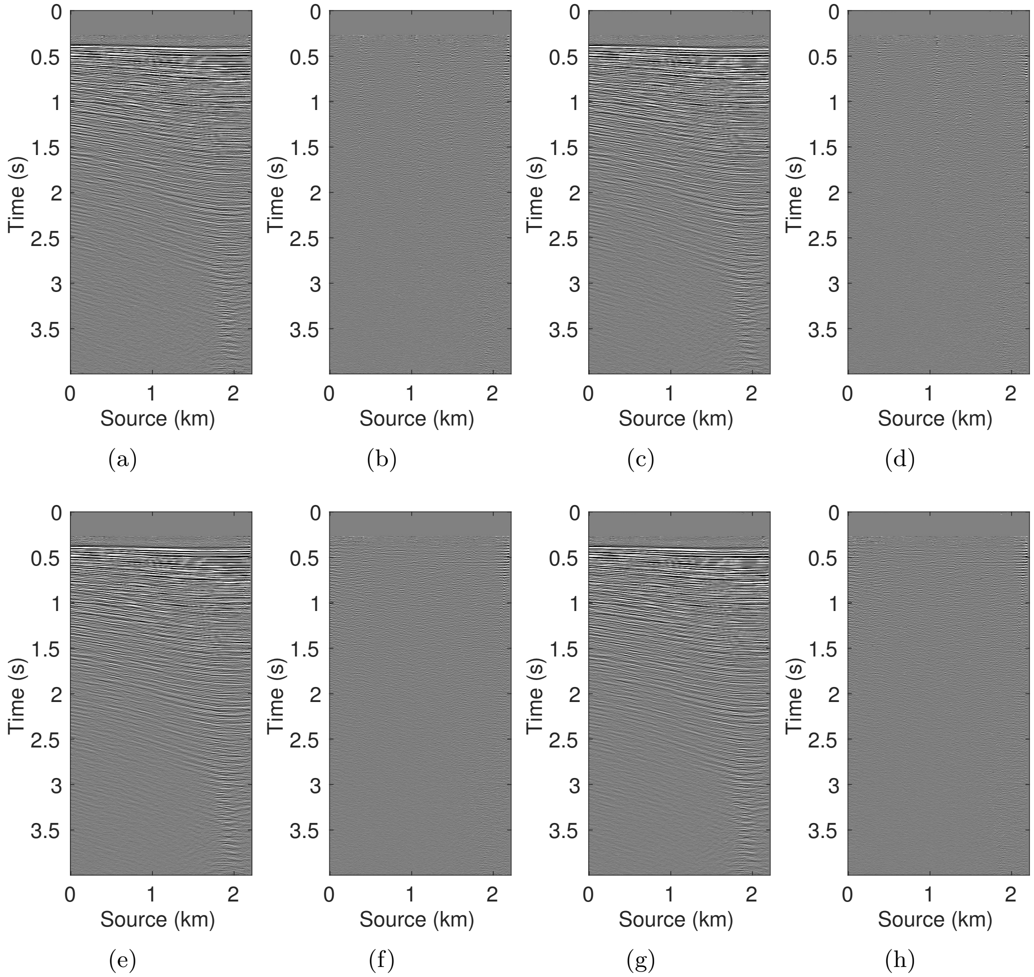


**Figure 7.13:** Separated common-channel gathers and difference plots (from the Gulf of Suez dataset) of source 1 and source 2: (a,c) source separated using HSS based rank-minimization and (b,d) the corresponding difference plots; (e,g) source separation using curvelet-based sparsity-promotion and (f,h) the corresponding difference plots.

lations, along with improved recovery quality of the separated seismic data. We have combined the Pareto curve approach for optimizing BPDN$_\epsilon$ formulations with the SVD-free matrix factorization methods to solve the nuclear-norm optimization formulation, which avoids the expensive computation of SVDs, a necessary step in traditional rank-minimization based methods. We find that our proposed techniques are comparable to the commonly used NMO-based median filtering techniques.

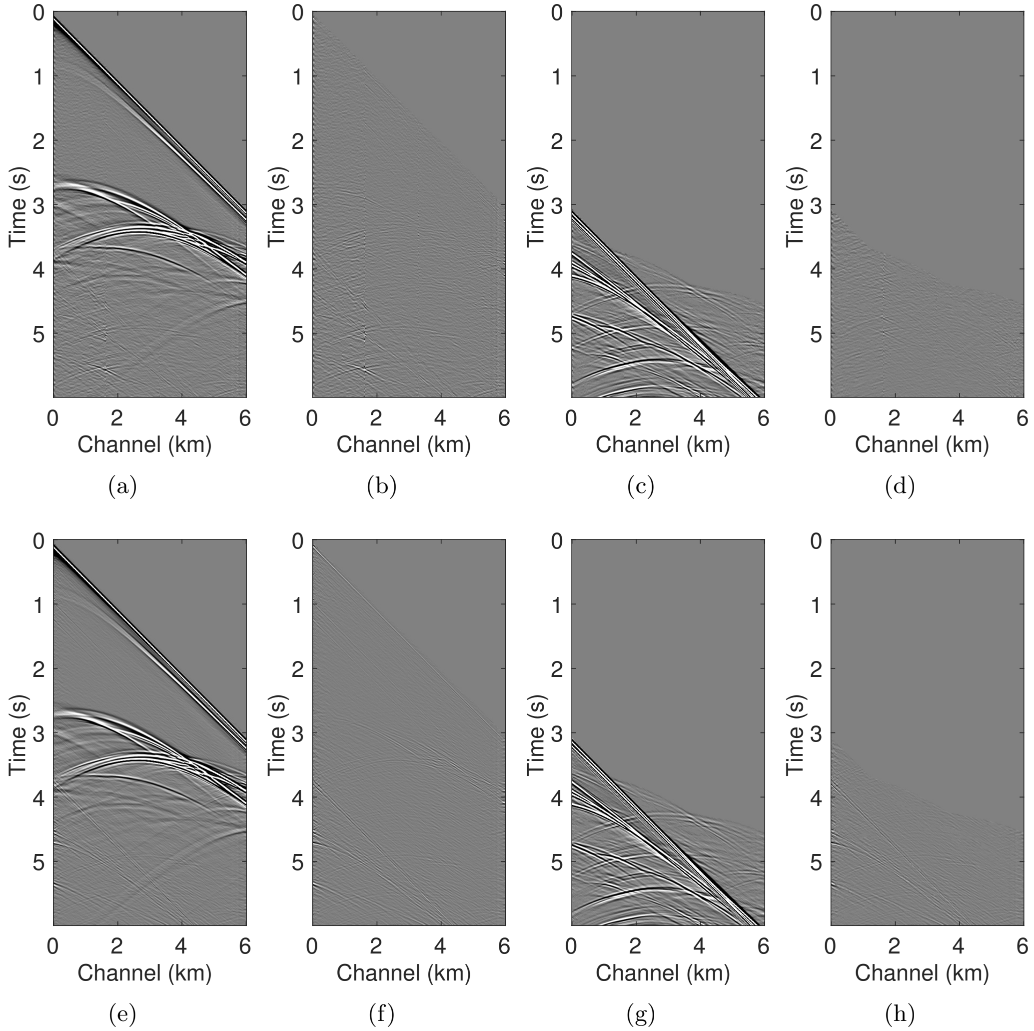
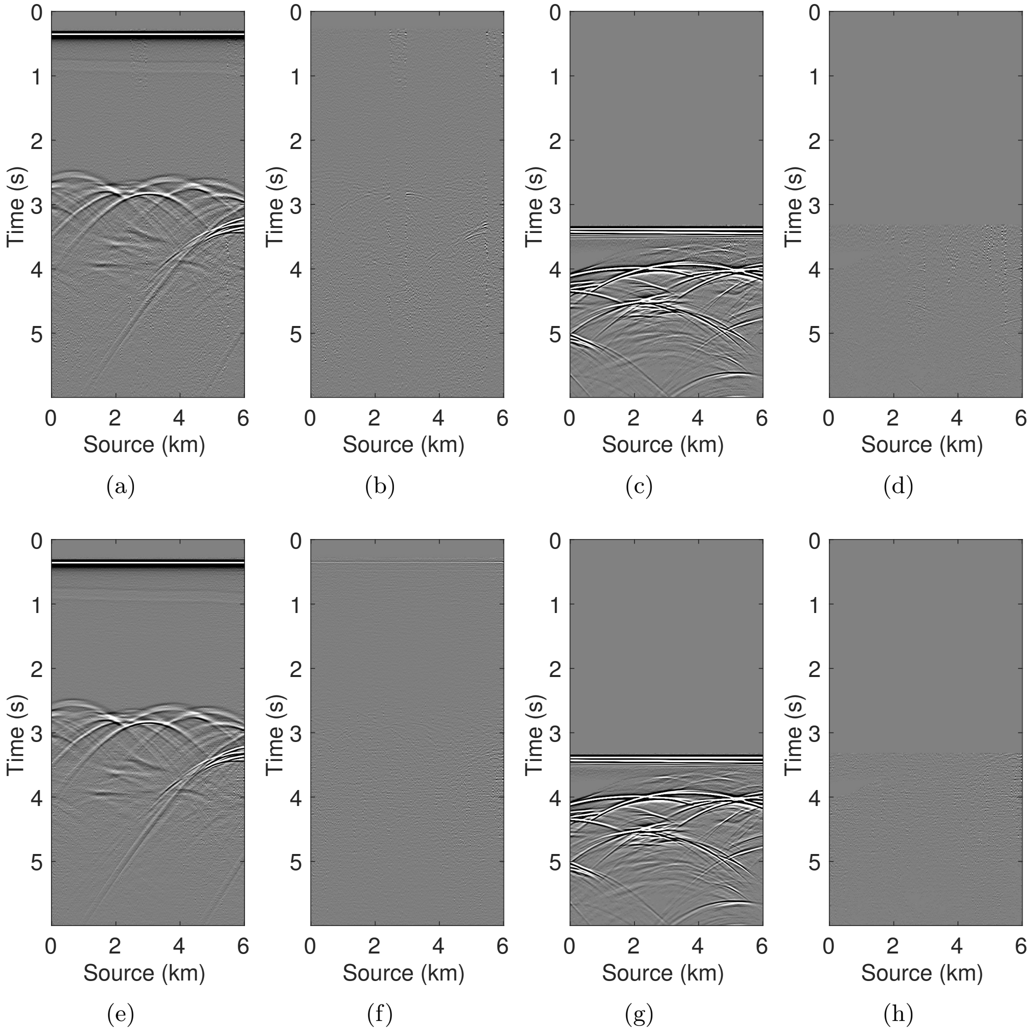


**Figure 7.14:** Separated shot gathers and difference plots (from the BP salt model) of source 1 and source 2: (a,c) source separation using HSS based rank-minimization and (b,d) the corresponding difference plots; (e,g) source separation using curvelet-based sparsity-promotion and (f,h) the corresponding difference plots.



(a) (b) (c) (d)

(e) (f) (g) (h)

**Figure 7.15:** Separated common-channel gathers and difference plots (from the BP salt model) of source 1 and source 2: (a,c) source separation using HSS based rank-minimization and (b,d) the corresponding difference plots; (e,g) source separation using curvelet-based sparsity-promotion and (f,h) the corresponding difference plots.

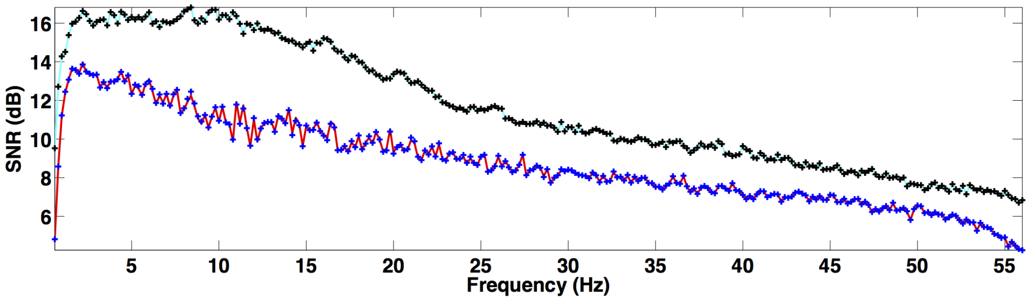
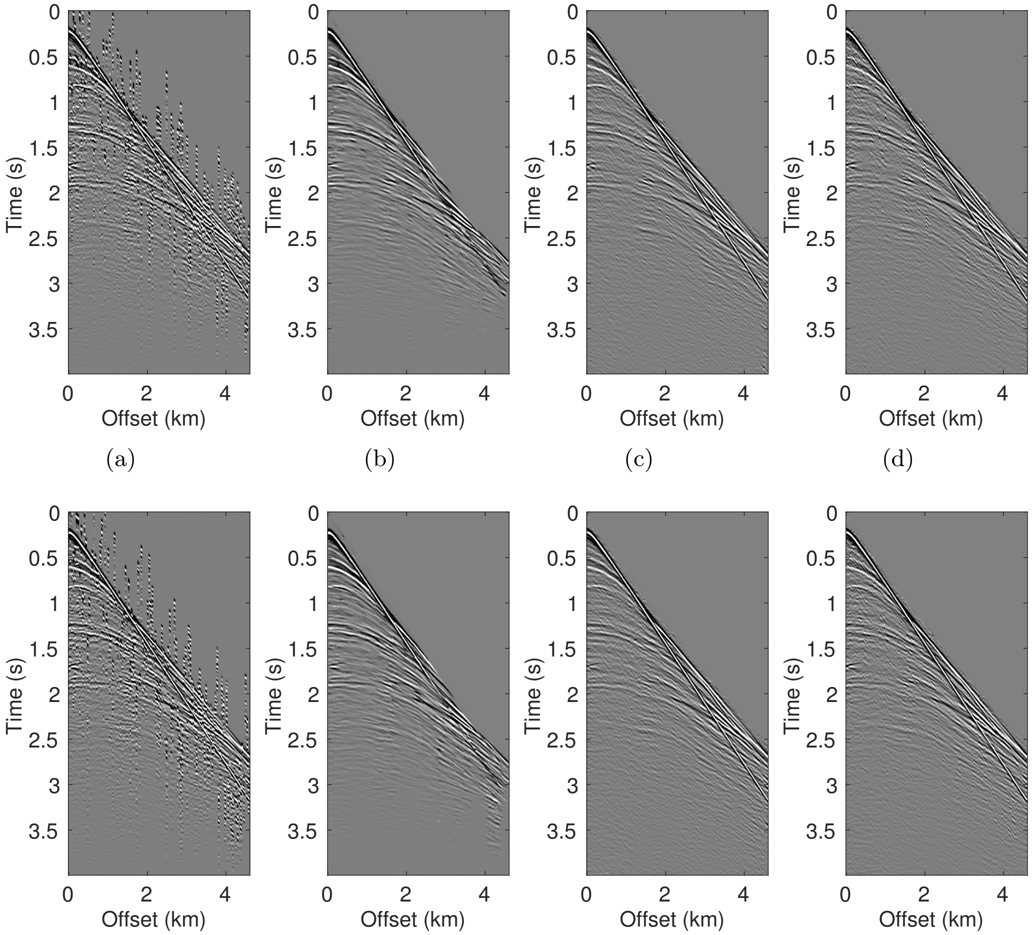


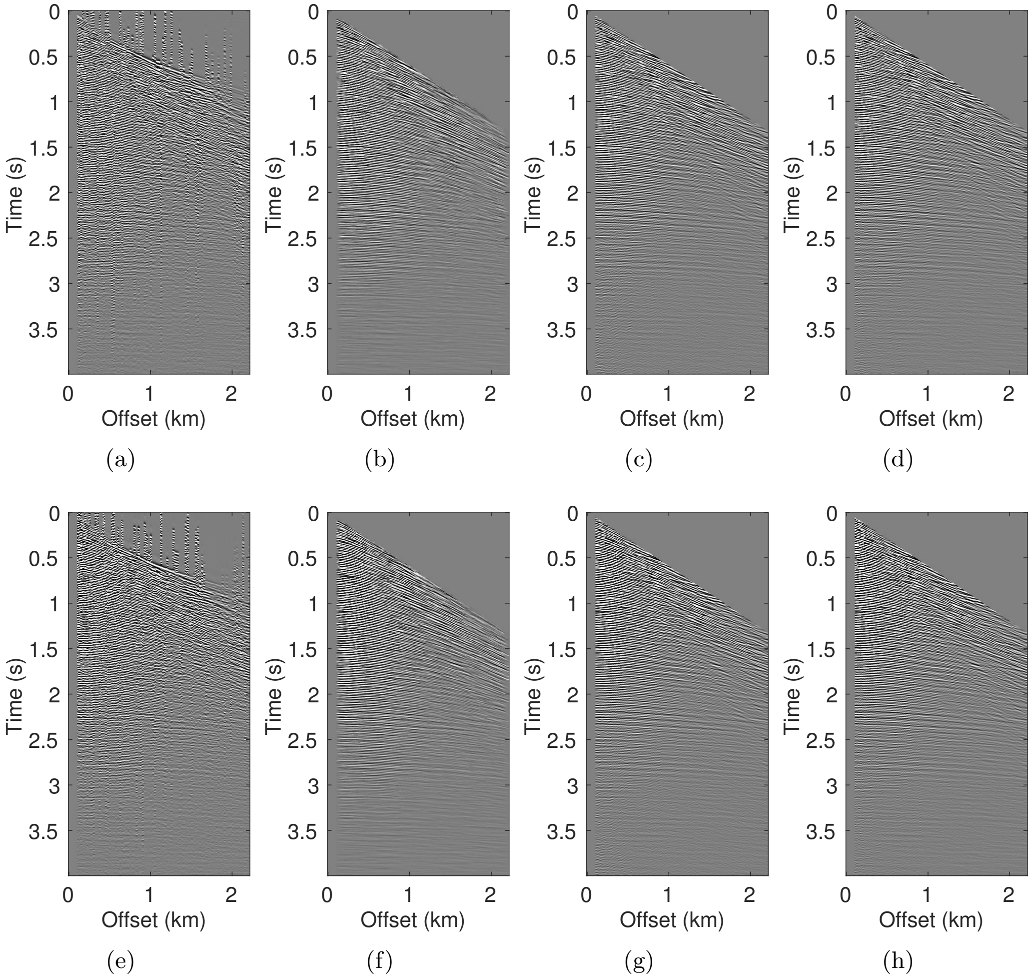
**Figure 7.16:** Signal-to-noise ratio (dB) over the frequency spectrum for the separated data from the Marmousi model. Red, blue curves—source separation without HSS; cyan, black curves—source separation using second-level HSS partitioning. Solid lines—separated source 1, + marker—separated source 2.



(a) (b) (c) (d)

(e) (f) (g) (h)

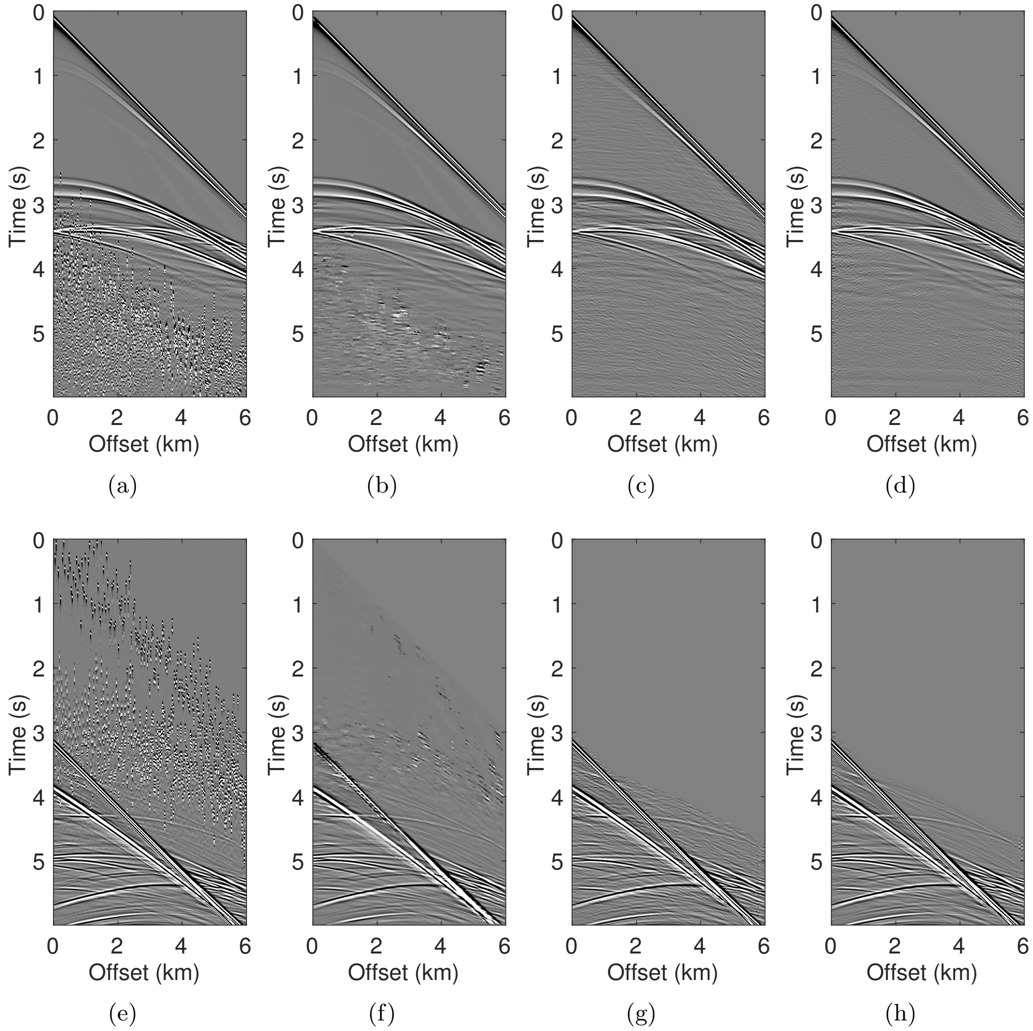
**Figure 7.17:** Blended common-midpoint gathers of (a) source 1 and (e) source 2 for the Marmousi model. Source separation using (b,f) NMO-based median filtering, (c,g) rank-minimization and (d,h) sparsity-promotion.



(a) (b) (c) (d)

(e) (f) (g) (h)

**Figure 7.18:** Blended common-midpoint gathers of (a) source 1, (e) source 2 for the Gulf of Suez dataset. Source separation using (b,f) NMO-based median filtering, (c,g) rank-minimization and (d,h) sparsity-promotion.



(a) (b) (c) (d)

(e) (f) (g) (h)

**Figure 7.19:** Blended common-midpoint gathers of (a) source 1, (e) source 2 for the BP salt model. Source separation using (b,f) NMO-based median filtering, (c,g) rank-minimization and (d,h) sparsity-promotion.

## Chapter 8

# Conclusion

Adapting ideas from the field of compressive sensing leads to new insights into acquiring and processing seismic data where we can fundamentally rethink on how we design acquisition surveys. Compressive randomized simultaneous-source acquisitions provide flexibility in acquisition geometries for better area coverage (i.e., improves data density) in surveys and speedup acquisition. The main contributions of this thesis are summarized below.

## 8.1 Compressive sensing in seismic exploration

We propose an alternative sampling method adapting insights from CS towards seismic acquisition and processing for data that are subsampled. The main outcome of this approach is a new technology where acquisition and processing related costs are decoupled from the stringent Nyquist sampling criterion. Instead, these costs scale with the desired reconstruction error and transform-domain sparsity of the data. By means of carefully designed numerical experiments on synthetic and real data, we establish that CS can indeed successfully be adapted to seismic data acquisition, wherein seismic wavefields can be reconstructed with a controllable error from randomized subsamplings. Specifically, three key components need to be in place: *(i)* a sparsifying signal (i.e., structure revealing) representation that exploits the signal's structure by mapping the energy into a small number of significant transform-domain coefficients; *(ii)* a randomized subsampling scheme that turns subsampling related artifacts into incoherent noise that is not sparse or compressible; and *(iii)* recovery of artifact-free fully sampled data by promoting structure, i.e., sparse recovery via one-norm minimization. We also introduce performance measures for nonlinear approximation and recovery errors and empirically demonstrate that curvelets lead to compressible (real-world signals are not strictly sparse) representation of seismic data compared to wave atoms, wavelets, etc. Hence, we use curvelets for recovery of densely sampled conventional data via sparsity promotion. In a nutshell, compressive sensing offers new perspectives towards the design of land and marine acquisition schemes.

## 8.2 Compressive simultaneous-source marine acquisition

We identify simultaneous-source marine acquisition as a linear subsampling system and analyze it using CS metrics such as mutual coherence and restricted isometry property. We quantitatively investigate the underlying interaction between acquisition design and reconstruction fidelity, and show that more randomness in the acquisition system and more compressible transforms improve the mutual coherence and restricted isometry constants, which predict a higher reconstruction quality. This is also true for our proposed *pragmatic* compressive marine simultaneous-source acquisition scheme, termed *time-jittered marine*, wherein a single (and/or multiple) source vessel(s) sails across an ocean-bottom array firing air guns at jittered-time instances, which translate to jittered shot positions for a given (fixed) speed of the source vessel. The simultaneous data are time compressed with overlapping shot records, and are therefore acquired economically with a small environmental imprint. The proposed acquisition scheme shares the benefits of random sampling while offering control on the maximum acquisition gap size since randomization via jittering turns the recovery into a relatively simple "denoising" problem with control over the maximum gap size between adjacent shot locations (Hennenfent and Herrmann, 2008), which is a practical requirement of wavefield reconstruction with localized sparsifying transforms such as curvelets.

According to CS, a sparsifying transform that is incoherent with the CS matrix can significantly impact the reconstruction quality. In other words, the mutual coherence between the sparsifying transform and the sensing (or measurement) matrix should be small for good signal reconstruction. We show that the CS matrix resulting from our proposed sampling scheme is incoherent with the curvelet transform. Recovering conventional nonsimultaneous prestack data volumes from simultaneous marine data essentially involves mapping noise-like or incoherent source crosstalk to coherent seismic responses. Randomized simultaneous-source acquisitions render source crosstalk (or interferences) incoherent in common-receiver gathers creating favorable conditions for recovering conventional nonsimultaneous data via curvelet-domain sparsity promotion. We recover conventional interference-free data using 2D and 3D FDCT (Fast Discrete Curvelet Transform, (Candès et al., 2006a; Ying et al., 2005)). We observe that the 3D FDCT leads to slightly improved recoveries compared to the 2D FDCT but at the expense of increased computational costs. This is because the 3D FDCT is about 24 × redundant, in contrast to the 8 × redundant 2D FDCT, rendering large-scale processing extremely memory intensive, and hence impractical. The combination of randomized subsampling and sparsity-promoting recovery technique results in high-quality nonsimultaneous data volumes recovered on fine periodic sampling grids. The results vindicate the importance of randomness in the acquisition scheme. Recovery from the more realistic irregular or off-the-grid subsamplings are reported in Chapter 6.

## 8.3 Compressive simultaneous-source time-lapse marine acquisition

We present a first instance of adapting ideas from CS and DCS to assess the effects of random or irregular subsampling (in the field) on time-lapse data, and demonstrate that high-quality data recoveries are the norm and not the exception. The main finding is that compressive randomized time-lapse surveys need not be replicated to attain similar/acceptable levels of data repeatability compared to data acquired from (expensive) dense, periodically sampled and replicated time-lapse surveys. This observation holds true when the vintages themselves are of prime interest, and in an idealized setting where we assume subsampled time-lapse measurements are taken on a discrete regular periodic grid, i.e., samples lie "exactly" on the grid (Chapter 5). This leads to "exact" replication of sampling points between time-lapse surveys when the sampling points overlap (for different percentages of overlap). We achieve this result by using a joint-recovery model (JRM), derived from DCS, that exploits common information shared between time-lapse vintages, and additional structural information provided by nonreplicated (or independent) surveys. Therefore, whenever time-lapse data exhibit joint structure — i.e., they are compressible in some transform domain and share information — sparsity-promoting recovery of the "common component" and "innovations," with respect to this common component, outperforms independent recovery of colocated (prestack) baseline and monitor data. Moreover, jointly recovered colocated (prestack) vintages exhibit a higher degree of repeatability in terms of NRMS ratios compared to independent recovery. Our proposed method lowers acquisition cost and environmental imprint because we have subsampled measurements, i.e., fewer shot locations are visited. This offers a possibility to extend the survey area or to increase the data's resolution at the same costs as conventional surveys.

Our findings in Chapter 6 corroborate the observations made above when we consider a more realistic (field) scenario of taking measurements off the grid, i.e., irregular samples that do not lie on a discrete regular periodic grid. Extending our simultaneous time-jittered marine acquisition to time-lapse surveys for realistic off-the-grid acquisitions, wherein off-the-grid sample points are known, generates simultaneous subsampled and irregular measurements. Consequently, we adapt our proposed JRM to incorporate a regularization operator that maps traces from an irregular grid to a regular periodic grid. The recovery method is a combined operation of source separation, regularization, and interpolation, in which periodic densely sampled and colocated prestack data are recovered from time-compressed, and therefore economical, simultaneous subsampled and irregular data. Similar to the observations made in Chapter 5, joint recovery of the vintages improves significantly when the time-lapse surveys are not replicated, since independent surveys give additional structural information. We also show that realistic and inevitable off-the-grid sampling leads to little variability in recovery of the time-lapse difference for decreasing overlap between the surveys, and hence negates the efforts to replicate.

Using two different geological velocity models, SEAM Phase 1 model that has a relatively complex geology and complex time-lapse difference compared to the BG COMPASS model (simple geology and complex time-lapse difference), we show that the observations are quite universal. These observations are significant because they can potentially change the current paradigm of time-lapse seismic that relies on expensive dense periodic sampling and replication of time-lapse surveys. Although these observations are made assuming source equalization and no significant changes in wave heights, water column velocities or temperature, and salinity profiles among the different surveys, recent successes of randomized surveys in the field (see, e.g., Mosher et al. (2014)) build our confidence in the success of pragmatic (field) compressive randomized simultaneous time-lapse surveys. Moreover, since irregular spatial sampling is inevitable in the real world, it would be better to focus on knowing what the shot positions were (post acquisition) to a sufficient degree of accuracy, than aiming to replicate them. Embracing randomness in surveys, whether natural randomness such as (streamer) cable feathering or randomness by design, maximizes collection of different/independent information leading to surveys acquired at low cost and a small environmental imprint.

## 8.4 Compressive simultaneous-source towed-streamer acquisition

We present two CS-based techniques — sparsity promotion and rank minimization — for source separation for dynamic simultaneous towed-streamer acquisitions, such as over/under and simultaneous-long offset acquisition. Recoveries from both techniques are comparable; however, the latter approach readily scales to large-scale seismic data volumes and is computationally faster. Moreover, the SVD-free matrix factorization method used to solve the nuclear-norm optimization formulation avoids expensive SVDs, a necessary step in traditional rank-minimization-based methods. We address the challenge of processing high-frequency monochromatic slices, which do not exhibit low-rank structure in the midpoint-offset domain due to increase in wave oscillations away from the diagonal compared to low-frequency slices, by incorporating the hierarchical semiseparable structure in the factorization-based rank-minimization framework leading to improved recovery quality of the separated data volumes. We also find that both these techniques are comparable with the commonly used NMO-based median-filtering techniques.

## 8.5 Follow-up work

Motivated by the successful implementation of CS to static and dynamic marine simultaneous-source acquisitions, and DCS to (static) simultaneous-source time-lapse surveys, we extend our work to derive a viable low-cost and low-environmental impact multi-azimuth towed-streamer time-lapse acquisition scheme. Initial findings of a simulation-based feasibility study for 3D randomized towed-streamer time-lapse surveys in a realistic field-scale setting have been reported in Kumar et al. (2017). In this acquisition scheme, we acquire economic, randomly subsampled (about 70%) and simultaneous towed-streamer time-lapse data without the need of replicating the surveys. We recover densely sampled full-azimuth time-lapse data on one and the same periodic grid by using the joint-recovery model (Chapters 5 and 6) coupled with the computationally cheap and scalable rank-minimization technique (Chapter 7). Our findings are consistent with those reported in this thesis that indicate that acquisition efficiency can be improved significantly by adapting the principles of CS. Furthermore, this new paradigm can also provide an appropriate framework for low-cost time-lapse wide-azimuth acquisition with towed arrays and multiple source vessels.

Initial findings on the possible impact of unknown calibration errors, such as unknown deviations between actual and post-plot acquisition geometry, on time-lapse data repeatability have been reported by Oghenekohwo and Herrmann (2017). In this contribution, the authors show that for data acquired via compressive, irregular and nonreplicated surveys, attainable recovery quality and repeatability of time-lapse vintages and difference deteriorates gracefully as a function of increasing calibration errors.

## 8.6 Current limitations

Some limitations of the work presented in this thesis are as follows:

1. Current implementation/code of the curvelet transform — the FDCT based on the wrapping of specially selected Fourier samples (Candès et al., 2006a) — is around 8 × redundant in 2D and around 24 × redundant in 3D (Ying et al., 2005). This precludes tractable higher-dimensional FDCTs. Moreover, it renders large-scale computations infeasible especially when dealing with massive 3D seismic data volumes.

2. For recovery of periodic densely sampled time-lapse data (Chapters 5 and 6), we assume that magnitude of the common component and innovations in the joint-recovery model is more or less similar. However, this is not the case in reality, and therefore certain scaling measures are required that will lead to improved (sparse) recoveries (see point 4 in the next section).

3. One of the limitations of the proposed SVD-free matrix-factorization approach (Chapter 7) is to find the rank parameter $k$ associated with each low-rank factor. The cross-validation

techniques used for the proposed formulation are rendered impractical when dealing with large-scale (subsampled) seismic data volumes since it involves finding an appropriate small volume of data to run the cross-validation techniques — a computationally expensive process. The implementation of the HSS representation for 3D seismic data volumes also remains to be investigated.

## 8.7 Future research directions

Some ideas for future work are as follows:

1. Develop a computationally faster and memory efficient implementation of the 2D and (more importantly) 3D curvelet transforms, since the curvelet transform is designed to represent curve-like singularities optimally (Candès and Demanet, 2005). Curvelets have different frequency content and dips to match wavefronts locally, leading to a sparse — arguably the sparsest — representation of seismic data.
2. Conduct investigations for recovery of weak late-arriving events with high degrees of accuracy during source separation. One approach to improve recovery of weak late-arriving events is to use weighted one-norm minimization (Mansour et al., 2013) that exploits correlations between locations of significant transform-domain coefficients of different partitions —– e.g., shot records, common-offset gathers or frequency slices —– of the acquired data.
3. Develop robust algorithms to use simultaneous-source data directly in imaging and inversion without the need for separating simultaneous data. Some researchers have obtained good synthetic results for direct imaging of simultaneous-source data (Berkhout et al., 2012; Choi and Alkhalifah, 2012; Guitton and Daz, 2012; Gan et al., 2016; Xue et al., 2016) but successful field applications remain challenging.
4. Investigate the effects of imposing a weight $\gamma$ on the common component and innovations that turns the joint recovery via $\ell_1$-norm minimization formulation (Chapters 5 and 6) in to a *$\gamma$-weighted $\ell_1$-norm formulation* (DCS, Baron et al. (2009)):

$$\widetilde{\mathbf{z}} = \arg\min_{\mathbf{z}} \gamma_0 ||\mathbf{z_0}||_1 + \gamma_1 ||\mathbf{z_1}||_1 + \gamma_2 ||\mathbf{z_2}||_1 \quad \text{subject to} \quad \mathbf{y} = \mathbf{A}\mathbf{z}. \tag{8.1}$$

So far we have assumed $\gamma_0 = \gamma_1 = \gamma_2 = 1$. However, in reality, magnitude of the common component and innovations of seismic data vary, and thus call for some enhancements in the reconstruction scheme by means of introducing appropriate weights on each element of $\mathbf{z}$. Similar to improvement in sparse recoveries from weighted $\ell_1$-norm minimization formulations (Mansour et al., 2013), the above weighted-norm formulation for DCS will potentially lead to improved sparse time-lapse data recoveries with a possible reduction in the requisite number of measurements (Li, 2015).

# Bibliography

Abma, R., A. Ford, N. Rose-Innes, H. Mannaerts-Drew, and J. Kommedal, 2013, Continued development of simultaneous source acquisition for ocean bottom surveys: Presented at the 75th EAGE Annual Conference Proceedings. → pages 5, 6, 107, 135

Abma, R., and N. Kabir, 2006, 3D interpolation of irregular data with a POCS algorithm: Geophysics, **71**, E91–E97. → pages 18

Abma, R., T. Manning, M. Tanis, J. Yu, and M. Foster, 2010, High quality separation of simultaneous sources by sparse inversion: Presented at the 72nd EAGE Annual Conference Proceedings. → pages 6, 45, 68, 135

Abma, R., Q. Zhang, A. Arogunmati, and G. Beaudoin, 2012, An overview of BPs marine independent simultaneous source field trials: SEG Technical Program Expanded Abstracts, 1–5. → pages 6

Akerberg, P., G. Hampson, J. Rickett, H. Martin, and J. Cole, 2008, Simultaneous source separation by sparse Radon transform: SEG Technical Program Expanded Abstracts, **27**, 2801–2805. → pages 6, 45, 68, 135

Allen, K. P., M. L. Johnson, and J. S. May, 1998, High Fidelity Vibratory Seismic (HFVS) method for acquiring seismic data: 140–143. → pages 45

Aravkin, A. Y., J. V. Burke, and M. P. Friedlander, 2013, Variational properties of value functions: SIAM Journal on Optimization, **23**, 1689–1717. → pages 145

Aravkin, A. Y., R. Kumar, H. Mansour, B. Recht, and F. J. Herrmann, 2014, Fast methods for denoising matrix completion formulations, with applications to robust seismic data interpolation: SIAM Journal on Scientific Computing, **36**, S237–S266. → pages 138, 143

Baardman, R. H., and R. G. van Borselen, 2013, Method and system for separating seismic sources in marine simultaneous shooting acquisition. → pages 6, 135

Bagaini, C., 2006, Overview of simultaneous vibroseis acquisition methods: SEG Technical Program Expanded Abstracts, 70–74. → pages 45

———, 2010, Acquisition and processing of simultaneous vibroseis data: Geophysical Prospecting, **58**, 81–99. → pages 5

Bagaini, C., and Y. Ji, 2010, Dithered slip-sweep acquisition: SEG Technical Program Expanded Abstracts, **29**, 91–95. → pages 6, 45, 68

Baron, D., M. F. Duarte, M. B. Wakin, S. Sarvotham, and R. G. Baraniuk, 2009, Distributed compressive sensing: CoRR, **abs/0901.3403**. → pages 11, 82, 85, 102, 120, 166

Beasley, C. J., 2008, A new look at marine simultaneous sources: The Leading Edge, **27**, 914–917. → pages 5, 6, 44, 67, 107, 134, 135

Beasley, C. J., R. E. Chambers, and Z. Jiang, 1998, A new look at simultaneous sources: SEG Technical Program Expanded Abstracts, 133–135. → pages 5, 6, 44, 45, 67

Beasley, C. J., R. E. Chambers, R. L. Workman, K. L. Craft, and L. J. Meister, 1997, Repeatability of 3-D ocean-bottom cable seismic surveys: The Leading Edge, **16**, 1281–1286. → pages 82

Beasley, C. J., B. Dragoset, and A. Salama, 2012, A 3D simultaneous source field test processed using alternating projections: a new active separation method: Geophysical Prospecting, **60**, 591–601. → pages 6

Berg, E. v., and M. P. Friedlander, 2008, Probing the Pareto frontier for basis pursuit solutions: SIAM Journal on Scientific Computing, **31**, 890–912. → pages 15, 19, 30, 49, 69, 83, 137, 138, 145

Berkhout, A. J., 2008, Changing the mindset in seismic data acquisition: The Leading Edge, **27**, 924–938. → pages 5, 6, 29, 44, 53, 67, 72, 107, 121, 135

———, 2012, Blended acquisition with dispersed source arrays: Geophysics, **77**, A19–A23. → pages 6

Berkhout, A. J., D. J. Verschuur, and G. Blacquire, 2012, Illumination properties and imaging promises of blended, multiple-scattering seismic data: a tutorial: Geophysical Prospecting, **60**, 713–732. → pages 166

Beyreuther, M., J. Cristall, and F. J. Herrmann, 2005, Computation of time-lapse differences with 3-D directional frames: SEG Technical Program Expanded Abstracts, 2488–2491. → pages 82, 131

Billette, F. J., and S. Brandsberg-Dahl, 2004, The 2004 bp velocity benchmark: Presented at the 67th EAGE Annual Conference Proceedings. → pages 146

Blacquière, G., A. J. Berkhout, and E. Verschuur, 2009, Survey design for blended acquisition: SEG Technical Program Expanded Abstracts, 56–60. → pages 45

Bleistein, N., J. Cohen, and J. Stockwell, 2001, Mathematics of Multidimensional Seismic Imaging, Migration and Inversion: Springer. → pages 18

Bobin, J., J.-L. Starck, and R. Ottensamer, 2008, Compressed sensing in astronomy: IEEE Journal of Selected Topics in Signal Processing, **2**, 718–726. → pages 31

Bourgeois, A., M. Bourget, P. Lailly, M. Poulet, P. Ricarte, and R. Versteeg, 1991, Marmousi data and model, *in* The Marmousi experience: EAGE, **5–9**. → pages 146

Brown, G., and J. O. Paulsen, 2011, Improved marine 4D repeatability using an automated vessel, source and receiver positioning system: First Break, **29**, 49–58. → pages 82

Bruckstein, A. M., D. L. Donoho, and M. Elad, 2009, From sparse solutions of systems of equations to sparse modeling of signals and images: SIAM Review, **51**, 34–81. → pages 48, 52

Caldwell, J., and C. Walker, 2011, An overview of marine seismic operations: Technical Report 448. → pages 1, 2

Candès, E. J., 2008, The restricted isometry property and its implications for compressed sensing: Comptes Rendus Mathématique, **346**, 589–592. → pages 48

Candès, E. J., and L. Demanet, 2005, The curvelet representation of wave propagators is optimally sparse: Communications on Pure and Applied Mathematics, **58**, 1472–1528. → pages 24

———, 2005, The curvelet representation of wave propagators is optimally sparse: Communications on Pure and Applied Mathematics, **58**, 1472–1528. → pages 48, 71, 72, 137, 166

Candès, E. J., L. Demanet, D. L. Donoho, and L. Ying, 2006a, Fast discrete curvelet transforms: Multiscale Modeling and Simulation, **5**, 861–899. → pages 14, 24, 48, 71, 118, 163, 165

Candès, E. J., and B. Recht, 2009, Exact matrix completion via convex optimization:

Foundations of Computational Mathematics, **9**, 717–772. → pages 138

Candès, E. J., J. Romberg, and T. Tao, 2006b, Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information: IEEE Transactions on Information Theory, **52**, 489–509. → pages 22

———, 2006c, Stable signal recovery from incomplete and inaccurate measurements: Communications on Pure and Applied Mathematics, **59**, 1207–1223. → pages 6, 17, 18, 21, 22, 46, 47, 83, 104, 107, 113, 131

Candès, E. J., and T. Tao, 2005, Decoding by linear programming: IEEE Transactions on Information Theory, **51**, 4203–4215. → pages 47

———, 2006, Near-optimal signal recovery from random projections: Universal encoding strategies: IEEE Transactions on Information Theory, **52**, 5406–5425. → pages 44, 49, 68, 82, 84

Candès, E. J., and M. B. Wakin, 2008, An introduction to compressive sampling: IEEE Signal Processing Magazine, **25**, 21–30. → pages 84

Canning, A., and G. H. F. Gardner, 1996, Regularizing 3-D data sets with DMO: Geophysics, **61**, 1103–1114. → pages 18

Chandrasekaran, S., P. Dewilde, M. Gu, W. Lyons, and T. Pals, 2006, A fast solver for HSS representations via sparse matrices: SIAM Journal on Matrix Analysis Applications, **29**, 67–81. → pages 15, 137, 141

Chen, Y., J. Yuan, Z. Jin, K. Chen, and L. Zhang, 2014, Deblending using normal moveout and median filtering in common-midpoint gathers: Journal of Geophysics and Engineering, **11**, 045012. → pages 15, 137, 150

Cheng, J., and M. D. Sacchi, 2013, Separation of simultaneous source data via iterative rank reduction: SEG Technical Program Expanded Abstracts, 88–93. → pages 6, 136

Choi, Y., and T. Alkhalifah, 2012, Application of multi-source waveform inversion to marine streamer data using the global correlation norm: Geophysical Prospecting, **60**, 748–758. → pages 166

de Kok, R. J., and D. Gillespie, 2002, A universal simultaneous shooting technique: Presented at the 64th EAGE Annual Conference Proceedings. → pages 5, 6, 44, 45, 67, 107, 134

Demanet, L., and L. Ying, 2007, Wave atoms and sparsity of oscillatory patterns: Applied and Computational Harmonic Analysis, **23**, 368–387. → pages 24, 25

Donoho, D. L., 2006, Compressed sensing: IEEE Transactions on Information Theory, **52**, 1289–1306. → pages 6, 17, 18, 22, 44, 46, 49, 68, 82, 83, 84, 107, 113

Donoho, D. L., and M. Elad, 2003, Optimally sparse representation in general (nonorthogonal) dictionaries via $ell^1$ minimization: Proceedings of the National Academy of Sciences, **100**, 2197–2202. → pages 46

Donoho, D. L., and X. Huo, 2001, Uncertainty principles and ideal atomic decomposition: IEEE Transactions on Information Theory, **47**, 2845–2862. → pages 21

Donoho, D. L., and B. F. Logan, 1992, Signal recovery and the large sieve: SIAM Journal on Applied Mathematics, **52**, 577–591. → pages 18

Donoho, D. L., A. Maleki, and A. Montanari, 2009, Message passing algorithms for compressed sensing: Proceedings of the National Academy of Sciences. → pages 19, 20

Donoho, D. L., and J. Tanner, 2009, Observed universality of phase transitions in high-dimensional geometry, with implications for modern data analysis and signal processing: Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, **367**, 4273–4293. → pages 20

Donoho, P. L., R. A. Ergas, and R. S. Polzer, 1999a, Development of seismic data compression methods for reliable, low-noise performance: SEG Technical Program Expanded Abstracts, 1903–1906. → pages 18

———, 1999b, Development of seismic data compression methods for reliable, low-noise, performance: SEG Technical Program Expanded Abstracts, 1903–1906. → pages 49

Doulgeris, P., K. Bube, G. Hampson, and G. Blacquiere, 2012, Convergence analysis of a coherency-constrained inversion for the separation of blended data: Geophysical Prospecting, **60**, 769–781. → pages 6, 135

Duarte, M. F., and R. G. Baraniuk, 2011, Kronecker Compressive Sensing: IEEE Transactions on Image Processing, **21**, 494–504. → pages 51

Duey, R., 2012, Simultaneous sources move offshore: Exploration and Production magazine. → pages 6

Eggenberger, K., P. Christie, D.-J. van Manen, and M. Vassallo, 2014, Multisensor streamer recording and its implications for time-lapse seismic and repeatability: The Leading Edge, **33**, 150–162. → pages 82

Eiken, O., G. U. Haugen, M. Schonewille, and A. Duijndam, 2003, A proven method for acquiring highly repeatable towed streamer seismic data: Geophysics, **68**, 1303–1309. → pages 82

Fanchi, J. R., 2001, Time-lapse seismic monitoring in reservoir management: The Leading Edge, **20**, 1140–1147. → pages 10, 82

Fomel, S., 2003, Theory of differential offset continuation: Geophysics, **68**, 718–732. → pages 18

Fomel, S., J. G. Berryman, R. G. Clapp, and M. Prucha, 2002, Iterative resolution estimation in least-squares Kirchhoff migration: Geophysical Prospecting, **50**, 577–588. → pages 18

Gan, S., S. Wang, Y. Chen, S. Qu, and S. Zu, 2016, Velocity analysis of simultaneous-source data using high-resolution semblancecoping with the strong noise: Geophysical Journal International, **204**, 768–779. → pages 166

Guitton, A., and E. Daz, 2012, Attenuating crosstalk noise with simultaneous source full waveform inversion: Geophysical Prospecting, **60**, 759–768. → pages 166

Güntürk, S., A. M. Powell, R. Saab, and O. Yilmaz, 2010, Sobolev duals for random frames and sigma-delta quantization of compressed sensing measurements: CoRR, **abs/1002.0182**. → pages 31

Hale, D., 1995, DMO processing: Geophysics Reprint Series. → pages 18

Halko, N., P. G. Martinsson, and J. A. Tropp, 2011, Finding structure with randomness: Probabilistic algorithms for constructing approximate matrix decompositions: SIAM review, **53**, 217–288. → pages 136, 142

Hampson, G., J. Stefani, and F. Herkenhoff, 2008, Acquisition using simultaneous sources: The Leading Edge, **27**, 918–923. → pages 6, 44, 45, 67, 107, 135

Harlan, W. S., J. F. Claerbout, and F. Rocca, 1984, Signal/noise separation and velocity estimation: Geophysics, **49**, 1869–1880. → pages 18

Hegna, S., and G. E. Parkes, 2012, Method for acquiring and processing marine seismic data to extract and constructively use the up-going and down-going wave-fields emitted by the source(s). (US Patent Application 13/686,408). → pages 9, 135

Hennenfent, G., L. Fenelon, and F. J. Herrmann, 2010, Nonequispaced curvelet transform for seismic data reconstruction: a sparsity-promoting approach: Geophysics, **75**, WB203–WB210. → pages 15, 115, 118, 119

Hennenfent, G., and F. J. Herrmann, 2006, Seismic denoising with non-uniformly sampled curvelets: Computing in Science and Engineering, **8**, 16–25. → pages 115, 118, 137

———, 2008, Simply denoise: wavefield reconstruction via jittered undersampling: Geophysics, **73**, V19–V28. → pages xii, 7, 18, 19, 20, 33, 34, 69, 82, 92, 114, 118, 163

Hennenfent, G., E. van den Berg, M. P. Friedlander, and F. J. Herrmann, 2008, New insights into one-norm solvers from the Pareto curve: Geophysics, **73**, A23–A26. → pages 138

Herrmann, F. J., 2009, Compressive imaging by wavefield inversion with group sparsity: SEG Technical Program Expanded Abstracts, 2337–2341. → pages 26

———, 2010, Randomized sampling and sparsity: Getting more information from fewer samples: Geophysics, **75**, WB173–WB187. → pages 7, 49, 82, 96

Herrmann, F. J., U. Boeniger, and D. J. Verschuur, 2007, Nonlinear primary-multiple separation with directional curvelet frames: Geophysical Journal International, **170**, 781–799. → pages 18, 20, 49

Herrmann, F. J., Y. A. Erlangga, and T. T. Lin, 2009, Compressive simultaneous full-waveform simulation: Geophysics, **74**, A35–A40. → pages 7, 18, 26, 28, 29, 31, 46

Herrmann, F. J., and G. Hennenfent, 2008, Non-parametric seismic data recovery with curvelet frames: Geophysical Journal International, **173**, 233–248. → pages 18, 84

Herrmann, F. J., and X. Li, 2012, Efficient least-squares imaging with sparsity promotion and compressive sensing: Geophysical prospecting, **60**, 696–712. → pages 104

Herrmann, F. J., P. P. Moghaddam, and C. C. Stolk, 2008, Sparsity- and continuity-promoting seismic imaging with curvelet frames: Journal of Applied and Computational Harmonic Analysis, **24**, 150–173. → pages 24, 48, 71, 84

Hill, D., C. Combee, and J. Bacon, 2006, Over/under acquisition and data processing: the next quantum leap in seismic technology?: First Break, **24**, 81–95. → pages 9, 135

Hoy, T., P. E. Dhelie, R. Sorley, and D. Lippett, 2013, A step change in seismic imaging—using a unique ghost free source and receiver system: Presented at the CSEG Annual Conference Proceedings. → pages 9, 135

Huo, S., Y. Luo, and P. Kelamis, 2009, Simultaneous sources separation via multi-directional vector-median filter: SEG Technical Program Expanded Abstracts, **28**, 31–35. → pages 6, 45, 68, 135

Jumah, B., and F. J. Herrmann, 2014, Dimensionality-reduced estimation of primaries by sparse inversion: Geophysical Prospecting, **62**, 972–993. → pages 141

Koster, K., P. Gabriels, M. Hartung, J. Verbeek, G. Deinum, and R. Staples, 2000, Time-lapse seismic surveys in the North Sea and their business impact: The Leading Edge, **19**, 286–293. → pages 82

Kragh, E., and P. Christie, 2002, Seismic repeatability, normalized rms, and predictability: The Leading Edge, **21**, 640–647. → pages 83, 97

Kreimer, N., 2013, Multidimensional seismic data reconstruction using tensor analysis: PhD thesis, University of Alberta. → pages 139

Kreimer, N., and M. Sacchi, 2012, A tensor higher-order singular value decomposition for prestack seismic data noise reduction and interpolation: Geophysics, **77**, V113–V122. → pages 138

Krohn, C. E., and R. Neelamani, 2008, Simultaneous sourcing without compromise: Presented at the 70th EAGE Annual Conference Proceedings. → pages 6, 31

Kumar, R., H. Mansour, A. Y. Aravkin, and F. J. Herrmann, 2013, Reconstruction of seismic wavefields via low-rank matrix factorization in the hierarchical-separable matrix representation: SEG Technical Program Expanded Abstracts, 3628–3633. → pages 138, 139, 141, 143

Kumar, R., C. D. Silva, O. Akalin, A. Y. Aravkin, H. Mansour, B. Recht, and F. J. Herrmann, 2015a, Efficient matrix completion for seismic data reconstruction: Geophysics, **80**, V97–V114. → pages 136, 139

Kumar, R., H. Wason, and F. J. Herrmann, 2015b, Source separation for simultaneous towed-streamer marine acquisition — a compressed sensing approach: Geophysics, **80**, WD73–WD88. → pages 6, 107

Kumar, R., H. Wason, S. Sharan, and F. J. Herrmann, 2017, Highly repeatable 3D compressive full-azimuth towed-streamer time-lapse acquisition — a numerical feasibility study at scale: The Leading Edge, **36**, 677–687. → pages 139, 153, 165

Kunis, S., 2006, Nonequispaced FFT: generalisation and inversion: PhD thesis, Lübeck University. → pages 115, 118

Landrø, M., 2001, Discrimination between pressure and fluid saturation changes from time-lapse seismic data: Geophysics, **66**, 836–844. → pages 82, 108

Lansley, R. M., M. Berraki, and M. M. M. Gros, 2007, Seismic array with spaced sources having variable pressure. (US Patent Application 12/998,723). → pages 9, 135

Lee, J., B. Recht, R. Salakhutdinov, N. Srebro, and J. Tropp, 2010, Practical large-scale optimization for max-norm regularization: Presented at the Advances in Neural Information Processing Systems, 2010. → pages 15, 137, 145

Levy, S., D. Oldenburg, and J. Wang, 1988, Subsurface imaging using magnetotelluric data: Geophysics, **53**, 104–117. → pages 18

Li, C., C. C. Mosher, and S. T. Kaplan, 2012, Interpolated compressive sensing for seismic data reconstruction: SEG Technical Program Expanded Abstracts, 1–6. → pages 114

Li, X., 2015, A weighted $\ell_1$-minimization for distributed compressive sensing: Master's thesis, The University of British Columbia. → pages 86, 166

Liberty, E., F. Woolfe, P. Martinsson, V. Rokhlin, and M. Tygert, 2007, Randomized algorithms for the low-rank approximation of matrices: Proceedings of the National Academy of Sciences, **104**, 20167–20172. → pages 136

Lin, T. T., and F. J. Herrmann, 2009a, Designing simultaneous acquisitions with compressive sensing: Presented at the 71st EAGE Annual Conference and Proceedings. → pages 7, 26, 45

———, 2009b, Unified compressive sensing framework for simultaneous acquisition with primary estimation: SEG Technical Program Expanded Abstracts, 3113–3117. → pages 26

———, 2013, Robust estimation of primaries by sparse inversion via one-norm minimization: Geophysics, **78**, R133–R150. → pages 5

Long, A., 2009, A new seismic method to significantly improve deeper data character and interpretability. → pages 9, 135

Long, A. S., E. von Abendorff, M. Purves, J. Norris, and A. Moritz, 2013, Simultaneous long offset (SLO) towed streamer seismic acquisition: Presented at the 75th EAGE Annual Conference Proceedings. → pages 9, 135

Lumley, D., 2010, 4d seismic monitoring of CO2 sequestration: The Leading Edge, **29**, 150–155. → pages 82

Lumley, D. E., 2001, Time-lapse seismic reservoir monitoring: Geophysics, **66**, 50–53. → pages 10, 82

Lumley, D. E., and R. A. Behrens, 1998, Practical issues of 4D seismic reservoir monitoring: What an engineer needs to know: SPE Reservoir Evaluation & Engineering, **1**, 528–538. → pages 11, 82, 107

Maharramov, M., and B. Biondi, 2014, Robust joint full-waveform inversion of time-lapse seismic data sets with total-variation regularization: arXiv. → pages 104

Mahdad, A., P. Doulgeris, and G. Blacquiere, 2011, Separation of blended data by iterative estimation and subtraction of blending interference noise: Geophysics, **76**, Q9–Q17. → pages 6, 45, 68, 135

Malcolm, A. E., M. V. de Hoop, and J. H. L. Rousseau, 2005, The applicability of DMO/AMO in the presence of caustics: Geophysics, **70**, S1–S17. → pages 18

Mallat, S., 2008, A wavelet tour of signal processing: the sparse way: Academic press. → pages 113

Mallat, S. G., 2009, A Wavelet Tour of Signal Processing: the Sparse Way: Academic Press. → pages 17, 18, 21, 22, 24, 27, 49

Mansour, H., F. J. Herrmann, and O. Yilmaz, 2013, Improved wavefield reconstruction from randomized sampling via weighted one-norm minimization: Geophysics, **78**, V193–V206. → pages 124, 131, 152, 166

Mansour, H., H. Wason, T. T. Lin, and F. J. Herrmann, 2012a, Randomized marine acquisition with compressive sampling matrices: Geophysical Prospecting, **60**, 648–662. → pages 68

———, 2012b, Randomized marine acquisition with compressive sampling matrices: Geophysical Prospecting, **60**, 648–662. → pages 82, 84, 107, 110, 135, 136

Maraschini, M., R. Dyer, K. Stevens, and D. Bird, 2012, Source separation by iterative rank reduction — theory and applications: Presented at the 74th EAGE Annual Conference Proceedings. → pages 6, 136

Moldoveanu, N., 2010, Random sampling: A new strategy for marine acquisition: SEG Technical Program Expanded Abstracts, 51–55. → pages 5, 6, 31

Moldoveanu, N., L. Combee, M. Egan, G. Hampson, L. Sydora, and W. Abriel, 2007, Over/under towed-streamer acquisition: A method to extend seismic bandwidth to both higher and lower frequencies: The Leading Edge, **26**, 41–58. → pages 9, 135

Moldoveanu, N., and S. Fealy, 2010, Multi-vessel coil shooting acquisition. → pages 67, 135

Moldoveanu, N., and J. Quigley, 2011, Random sampling for seismic acquisition: Presented at the 73rd EAGE Annual Conference Proceedings. → pages 107, 135

Moore, I., 2010, Simultaneous sources — processing and applications: Presented at the 72nd EAGE Annual Conference Proceedings. → pages 6, 45, 68, 135

Moore, I., B. Dragoset, T. Ommundsen, D. Wilson, C. Ward, and D. Eke, 2008, Simultaneous source separation using dithered sources: **27**, 2806–2810. → pages 6, 45, 67, 135

Mosher, C., C. Li, L. Morley, Y. Ji, F. Janiszewski, R. Olson, and J. Brewer, 2014, Increasing the efficiency of seismic data acquisition via compressive sensing: The Leading Edge, **33**, 386–391. → pages 5, 6, 7, 13, 72, 82, 84, 94, 104, 107, 121, 124, 131, 135, 164

Neelamani, R., C. E. Krohn, J. R. Krebs, J. K. Romberg, M. Deffenbaugh, and J. E. Anderson, 2008, Efficient seismic forward modeling using simultaneous random sources and sparsity: SEG Technical Program Expanded Abstracts, **27**, 2107–2111. → pages 6, 7, 46

———, 2010, Efficient seismic forward modeling using simultaneous random sources and sparsity: Geophysics, **75**, WB15–WB27. → pages 7, 46

Oghenekohwo, F., and F. J. Herrmann, 2017, Highly repeatable time-lapse seismic with distributed Compressive Sensing—mitigating effects of calibration errors: The Leading Edge, **36**, 688–694. → pages 165

Oghenekohwo, F., R. Kumar, E. Esser, and F. J. Herrmann, 2015, Using common information in compressive time-lapse full-waveform inversion: Presented at the 77th EAGE Annual Conference Proceedings. → pages 104

Oldenburg, D. W., S. Levy, and K. P. Whittall, 1981, Wavelet estimation and deconvolution: Geophysics, **46**, 1528–1542. → pages 18

Oropeza, V., and M. D. Sacchi, 2011, Simultaneous seismic data denoising and reconstruction via multichannel singular spectrum analysis: Geophysics, **76**, V25–V32. → pages 136, 138

Porter-Hirsche, J., and K. Hirsche, 1998, Repeatability study of land data acquisition and processing for time lapse seismic: SEG Technical Program Expanded Abstracts, 9–11. → pages 82

Potts, D., G. Steidl, and M. Tasche, 2001, Modern sampling theory: mathematics and applications: 249–274. → pages 115, 118

Rauhut, H., K. Schnass, and P. Vandergheynst, 2008, Compressed sensing and redundant dictionaries: IEEE Transactions on Information Theory, **54**, 2210–2219. → pages 22, 23

Recht, B., M. Fazel, and P. A. Parrilo, 2010, Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization: SIAM Review, **52**, 471–501. → pages 138, 145

Recht, B., and C. Re, 2013, Parallel stochastic gradient algorithms for large-scale matrix completion: Mathematical Programming Computation, **5**, 201–226. → pages 145

Rennie, J. D. M., and N. Srebro, 2005, Fast maximum margin matrix factorization for collaborative prediction: Proceedings of the 22nd international conference on Machine learning, ACM, 713–719. → pages 145

Rickett, J. E., and D. E. Lumley, 2001, Cross-equalization data processing for time-lapse seismic reservoir monitoring: A case study from the Gulf of Mexico: Geophysics, **66**, 1015–1025. → pages 82, 131

Romberg, J., 2008, Compressive sensing by random convolution: SIAM Journal on Imaging Sciences, **2**, 1098–1128. → pages 28, 29, 31

Ross, C. P., and M. S. Altan, 1997, Time-lapse seismic monitoring: Some shortcomings in nonuniform processing: The Leading Edge, **16**, 931–937. → pages 82

Sacchi, M. D., T. J. Ulrych, and C. Walker, 1998, Interpolation and extrapolation using a high-resolution discrete Fourier transform: IEEE Transactions on Signal Processing, **46**, 31–38. → pages 20

Sacchi, M. D., D. R. Velis, and A. H. Cominguez, 1994, Minimum entropy deconvolution with frequency-domain constraints: Geophysics, **59**, 938–945. → pages 18

Santosa, F., and W. Symes, 1986, Linear inversion of band-limited reflection seismogram: SIAM Journal on Scientific and Statistical Computing, **7**, 1307–1330. → pages 18

Schonewille, M., A. Klaedtke, A. Vigner, J. Brittan, and T. Martin, 2009, Seismic data regularization with the anti-alias anti-leakage fourier transform: First Break, **27**, 85–92. → pages 83

Silva, C. D., and F. J. Herrmann, 2013, Structured tensor missing-trace interpolation in the Hierarchical Tucker format: SEG Technical Program Expanded Abstracts, 3623–3627. → pages 138

Smith, H. F., 1998, A Hardy space for Fourier integral operators: The Journal of Geometric Analysis, **8**, 629–653. → pages 24, 48, 71, 137

Spetzler, J., and Ø. Kvam, 2006, Discrimination between phase and amplitude attributes in time-lapse seismic streamer data: Geophysics, **71**, O9–O19. → pages 82, 108

Spitz, S., 1999, Pattern recognition, spatial predictability, and subtraction of multiple events: The

Leading Edge, **18**, 55–58. → pages 18

Stefani, J., G. Hampson, and E. F. Herkenhoff, 2007, Acquisition using simultaneous sources: Presented at the 69th EAGE Annual Conference Proceedings. → pages 6, 45, 67, 135

Symes, W. W., 2010, IWAVE: a framework for wave simulation. → pages 91, 122, 124

Symes, W. W., D. Sun, and M. Enriquez, 2011, From modelling to inversion: designing a well-adapted simulator: Geophysical Prospecting, **59**, 814–833. → pages 146

Tang, G., J. Ma, and F. J. Herrmann, 2009, Design of two-dimensional randomized sampling schemes for curvelet-based sparsity-promoting seismic data recovery: Technical Report TR-2009-03, The University of British Columbia. → pages 20

Taylor, H. L., S. C. Banks, and J. F. McCoy, 1979, Deconvolution with the $\ell_1$ norm: Geophysics, **44**, 39–52. → pages 18

Tegtmeier-Last, S., and G. Hennenfent, 2013, System and method for processing 4D seismic data. (US Patent App. 13/804,029). → pages 82, 131

Trad, D., 2003, Interpolation and multiple attenuation with migration operators: Geophysics, **68**, 2043–2054. → pages 18

———, 2009, Five-dimensional interpolation: Recovering from acquisition constraints: Geophysics, **74**, V123–V132. → pages 139

Trad, D., T. Ulrych, and M. D. Sacchi, 2003, Latest views of the sparse Radon transform: Geophysics, **68**, 386–399. → pages 18

Trickett, S., and L. Burroughs, 2009, Prestack rank-reducing noise suppression: Theory: Presented at the SEG Technical Program Expanded Abstracts. → pages 138

Tu, N., T. T. Lin, and F. J. Herrmann, 2011, Sparsity-promoting migration with surface-related multiples: Presented at the 73rd EAGE Annual Conference Proceedings. → pages 104

Ulrych, T. J., and C. Walker, 1982, Analytic minimum entropy deconvolution: Geophysics, **47**, 1295–1302. → pages 18

van Groenestijn, G. J. A., and D. J. Verschuur, 2009, Estimating primaries by sparse inversion and application to near-offset data reconstruction: Geophysics, **74**, A23–A28. → pages 5

Verschuur, D. J., A. J. Berkhout, and C. P. A. Wapenaar, 1992, Adaptive surface-related multiple elimination: Geophysics, **57**, 1166–1177. → pages 5

Wang, D., R. Saab, Ö. Yilmaz, and F. J. Herrmann, 2008, Bayesian wavefield separation by transform-domain sparsity promotion: Geophysics, **73**, A33–A38. → pages 97

Wason, H., and F. J. Herrmann, 2012, Only dither: efficient simultaneous marine acquisition: Presented at the 74th EAGE Annual Conference Proceedings. → pages 68

———, 2013a, Ocean bottom seismic acquisition via jittered sampling: Presented at the 75th EAGE Annual Conference Proceedings. → pages 135, 153

———, 2013b, Time-jittered ocean bottom seismic acquisition: SEG Technical Program Expanded Abstracts, 1–6. → pages 82, 83, 84, 91, 107, 110, 135, 136, 153

Wason, H., F. Oghenekohwo, and F. J. Herrmann, 2015, Compressed sensing in 4-D marine—recovery of dense time-lapse data from subsampled data without repetition: Presented at the 77th EAGE Annual Conference Proceedings. → pages 84

Xiong, Z., A. D. Liveris, and S. Cheng, 2004, Distributed source coding for sensor networks: IEEE Signal Processing Magazine, **21**, 80–94. → pages 85

Xu, S., Y. Zhang, D. Pham, and G. Lambare, 2005, Antileakage Fourier transform for seismic data regularization: Geophysics, **70**, V87–V95. → pages 18, 19, 20, 83, 118

Xue, Z., Y. Chen, S. Fomel, and J. Sun, 2016, Seismic imaging of incomplete data and simultaneous-source data using least-squares reverse time migration with shaping regularization: Geophysics, **81**, S11–S20. → pages 166

Yang, D., A. E. Malcolm, and M. C. Fehler, 2014, Time-lapse full waveform inversion and uncertainty analysis with different survey geometries: Presented at the 76th EAGE Conference and Exhibition 2014. → pages 104

Ying, L., L. Demanet, and E. J. Candès, 2005, 3d discrete curvelet transform: Proceedings of SPIE on Wavelets XI, SPIE, 861–899. → pages 14, 54, 71, 163, 165

Zwartjes, P. M., and M. D. Sacchi, 2007, Fourier reconstruction of nonuniformly sampled, aliased seismic data: Geophysics, **72**, V21–V32. → pages 18, 20