

Preface

Contemporary signal processing technologies are frequently required to cope with undersampled, rare, missing or even conflicting measurements. In some cases the amount of available data can be below a threshold which seemingly inhibits plausible inference. Often, in such cases, conventional inference methods fall short of providing reliable solutions. As normally signals of interest can be discerned into only a few fundamental components in some mathematical domain, dedicated inference techniques seek to find solutions of the lowest complexity, a concept which has proved to be extremely useful when dealing with limited data.

This book is aimed at presenting concepts, methods and algorithms able to cope with undersampled and limited data. One such trend that recently gained popularity and to some extent revolutionised signal processing is compressed sensing. Compressed sensing builds upon the observation that many signals in nature are nearly sparse (or compressible, as they are normally referred to) in some domain, and consequently they can be reconstructed to within high accuracy from far fewer observations than traditionally held to be necessary.

Apart from compressed sensing this book contains other related approaches. Each methodology has its own formalities for dealing with such problems. As an example, in the Bayesian approach, sparseness promoting priors such as Laplace and Cauchy are normally used for penalising improbable model variables, thus promoting low complexity solutions. Compressed sensing techniques and homotopy-type solutions, such as the LASSO, utilise l_1 -norm penalties for obtaining sparse solutions using fewer observations than conventionally needed. The book emphasizes on the role of sparsity as a machinery for promoting low complexity representations and likewise, its connections to variable selection and dimensionality reduction in various engineering problems.

This book is intended for researchers, academics and practitioners with interest in various aspects and applications of sparse signal processing.

Book Outline

Each chapter in the present book forms a complete self-contained work that can be read independently of others. The reader who is not acquainted with the subject matter and in particular with compressed sensing is advised to read at least the first half of [Chap. 1](#). A brief description of the content of each chapter is provided below.

- [Chapter 1](#) is a concise exposition to the basic theory of compressed sensing. It assumes no prior knowledge of the subject and gradually builds the theory while elaborating on the basic results. The second half of this chapter is mostly concerned with the application of compressed sensing ideas to dynamic systems and sparse state estimation.
- [Chapter 2](#) is concerned with the geometrical foundations of compressed sensing. The geometric point of view adopted by the author not only underlies many of the initial theoretical developments on which much of the theory of compressed sensing is built, but has also allowed ideas to be extended to much more general recovery problems and structures. A unifying framework is that of non-convex, low-dimensional constraint sets in which the signal to be recovered is assumed to reside. The sparse signal structure of traditional compressed sensing translates into a union of low-dimensional subspaces, each subspace being spanned by a small number of the coordinate axes. The union of subspaces interpretation is readily generalised and many other recovery problems can be seen to fall into this setting. For example, instead of vector data, in many problems, data are more naturally expressed in matrix form (for example a video is often best represented in a pixel by time matrix). A powerful constraint on matrices are constraints on the matrix rank. For example, in low-rank matrix recovery, the goal is to reconstruct a low-rank matrix given only a subset of its entries. Importantly, low-rank matrices also lie in a union of subspaces structure, although now, there are infinitely many subspaces (though each of these is finite dimensional). Many other examples of union of subspaces signal models appear in applications, including sparse wavelet-tree structures (which form a subset of the general sparse model) and finite rate of innovations models, where we can have infinitely many infinite dimensional subspaces. The chapter provides an introduction to these and related geometrical concepts and shows how they can be used to (a) develop algorithms to recover signals with given structures and (b) allow theoretical results that characterise the performance of these algorithmic approaches.
- [Chapter 3](#) extends the basic theory of compressed sensing to the general case of exponential-family noise that includes Gaussian noise as a particular case; the underlying recovery problem is then formulated as l_1 -regularized generalized linear model (GLM) regression. In this chapter it is further shown that, under standard restricted isometry property assumptions on the design matrix, l_1 -minimization can provide stable recovery of a sparse signal in presence of

exponential-family noise. Sufficient conditions on the noise distribution are also provided that guarantee stable recovery.

- **Chapter 4** provides a brief review of some of the state of the art in nuclear norm optimization algorithms. The nuclear norm of a matrix, as the tightest convex surrogate of the matrix rank, has fueled much of the recent research and has proved to be a powerful tool in many areas. In this chapter the authors propose a novel application of the nuclear norm to the linear model recovery problem, as well as a viable algorithm for solution of the recovery problem.
- **Chapter 5** presents very recent developments in the area of non-negative tensor factorization which admit sparse representations. Specifically, it considers the approximate factorization of third and fourth order tensors into non-negative sums of types of outer-products of objects with one dimension less using the so-called t-product. A demonstration on an application in facial recognition shows the potential promise of the overall approach. This chapter also discusses a number of algorithmic options for solving the resulting optimization problems, and modification of such algorithms for increasing the underlying sparsity.
- **Chapter 6** describes the application of compressed sensing and sub-Nyquist sampling to cognitive radio. Cognitive radio has become one of the most promising solutions for addressing the spectral under-utilization problem in wireless communication systems. This chapter pays a special attention to the use of sub-Nyquist sampling and compressed sensing techniques for realizing wideband spectrum sensing. In addition, an adaptive compressed sensing approach is described for wideband spectrum sensing.
- **Chapter 7** presents a few identification algorithms for sparse nonlinear multi input multi output (MIMO) systems. The algorithms are potentially useful in a variety of application areas including digital transmission systems incorporating power amplifier(s) along with multiple antennas, cognitive processing, adaptive control of nonlinear multivariable systems, and multivariable biological systems. Sparsity is a key constraint imposed on the model. The presence of sparsity is often dictated by physical considerations as in wireless fading channel estimation. In other cases it appears as a pragmatic modelling approach that seeks to cope with the curse of dimensionality, particularly acute in nonlinear systems like Volterra type series. The authors discuss three identification approaches: conventional identification based on both input and output samples, semi-blind identification placing emphasis on minimal input resources and blind identification whereby only output samples are available plus a priori information on input characteristics. Based on this taxonomy a variety of algorithms, existing and new, are studied and evaluated by simulations.
- **Chapter 8** is concerned with the optimization formulation of the Kalman filtering and smoothing problems. The authors use this perspective to develop a variety of extensions and applications. They consider various extensions of Kalman smoothing to systems with nonlinear process and measurement models, systems with linear and nonlinear inequality constraints, systems with outliers in the measurements or sudden changes in the state, and systems where the sparsity

of the state sequence must be accounted for. All extensions preserve the computational efficiency of the classic algorithms, and most of the extensions are illustrated with numerical examples, which are part of an open source Kalman smoothing Matlab/Octave package.

- [Chapter 9](#) develops a novel Kalman filtering-based method for estimating the coefficients of sparse, or more broadly, compressible autoregressive models using fewer observations than normally required. The proposed algorithm facilitates sequential processing of observations and is shown to attain a good recovery performance, particularly under substantial deviations from ideal conditions. In the second half of this chapter, a few information-theoretic bounds are derived pertaining to the problem at hand. The obtained bounds establish the relation between the complexity of the autoregressive process and the attainable estimation accuracy through the use of a novel measure of complexity. This measure is suggested as a substitute to the generally incomputable restricted isometric property.
- [Chapter 10](#) introduces selective gossip which is an algorithm that applies the idea of iterative information exchange to vectors of data. Instead of communicating the entire vector and wasting network resources, the derived approach adaptively focuses communication on the most significant entries of the vector. The authors prove that nodes running selective gossip asymptotically reach consensus on these significant entries, and they simultaneously reach an agreement on the indices of entries which are insignificant. The results demonstrate that selective gossip provides significant communication savings in terms of number of scalars transmitted. In the second part of this chapter, a distributed particle filter is derived employing selective gossip. It is then shown that distributed particle filters employing selective gossip provide comparable results to the centralized bootstrap particle filter while decreasing the communication overhead compared to using randomized gossip to distribute the filter computations.
- [Chapter 11](#) describes a recent work on the design and analysis of recursive algorithms for causally reconstructing a time sequence of (approximately) sparse signals from a greatly reduced number of linear projection measurements. The signals are sparse in some transform domain referred to as the sparsity basis and their sparsity patterns (support set of the sparsity basis coefficients) can change with time. The term “recursive” implies that only the previous signal’s estimate and the current measurements are used to get the current signal’s estimate. The authors briefly summarize their exact reconstruction results for the noise-free case and likewise present error bounds and error stability results for the noisy case. Connections with related work are also discussed. A key example application where the underlying recovery problem occurs is dynamic magnetic resonance imaging (MRI) for real-time medical applications such as interventional radiology and MRI-guided surgery, or in functional MRI to track brain activation changes. Cross-sectional images of the brain, heart, larynx or other human organ images are piecewise smooth, and thus approximately sparse in the

wavelet domain. In a time sequence, their sparsity pattern changes with time, but quite slowly. The same is also often true for the nonzero signal values. This simple fact, which was first observed by the authors, is the key reason that the proposed recursive algorithms can achieve provably exact or accurate reconstruction from very few measurements.

- **Chapter 12** considers the problem of reconstructing time-varying sparse signals in a sensor network with limited communication resources. In each time interval, the fusion centre transmits the predicted signal estimate and its corresponding error covariance to a selected subset of sensors. The selected sensors compute quantized innovations and transmit them to the fusion centre. The authors consider the situation where the signal is sparse, i.e. a large fraction of its components is zero-valued. Algorithms are presented for signal estimation in the described scenario, and their complexity is analysed. It is shown that the proposed algorithms maintain near-optimal performance even in the case where sensors transmit a single bit (i.e., the sign of innovation) to the fusion centre.
- **Chapter 13** is concerned with the application of sparsity and compressed sensing ideas in imaging radars, also known as synthetic aperture radars (SARs). The authors provide a brief overview of how sparsity-driven imaging has recently been used in various radar imaging scenarios. They then focus on the problem of imaging from undersampled data, and point to recent work on the exploitation of compressed sensing theory in the context of radar imaging. This chapter considers and describes in detail the geometry and measurement model for multi-static radar imaging, where spatially distributed multiple transmitters and receivers are involved in data collection from the scene to be imaged. The mono-static case, where transmitters and receivers are collocated is treated as a special case. For both the mono-static and the multi-static scenarios the authors examine various ways and patterns of undersampling the data. These patterns reflect spectral and spatial diversity trade-offs. Characterization of the expected quality of the reconstructed images in these scenarios prior to actual data collection is a problem of central interest in task planning for multi-mode radars. Compressed sensing theory argues that the mutual coherence of the measurement probes is related to the reconstruction performance in imaging sparse scenes. With this motivation the authors propose a closely related, but more effective parameter they termed the t %-average mutual coherence as a sensing configuration quality measure and examine its ability to predict reconstruction quality in various monostatic and ultra-narrow band multi-static configurations.
- **Chapter 14** shows how a sparse solution can be obtained for a range of problems in a Bayesian setting by using prior models on sparsity structure. As an example, a model to remove impulse and background noise from audio signals via their representation in time–frequency space using Gabor wavelets is presented. A range of prior models for the sparse structure of the signal in this space is introduced, including simple Bernoulli priors on each coefficient, Markov chains linking neighbouring coefficients in time or frequency and Markov random fields, imposing two-dimensional coherence on the coefficients. The effect of

each of these priors on the reconstruction of a corrupted audio signal is shown. Impulse removal is also covered, with similar sparsity priors being applied to the location of impulse noise in the audio signal. Inference is performed by sampling from the posterior distribution of the model variables using the Gibbs sampler.

- [Chapter 15](#) presents the methods that are currently exploited for sparse optimization in speech. It also demonstrates how sparse representations can be constructed for classification and recognition tasks, and gives an overview of recent results that were obtained with sparse representations.

February 2013

Avishy Y. Carmi
Lyudmila S. Mihaylova
Simon J. Godsill

Compressed Sensing & Sparse Filtering

Carmi, A.Y.; Mihaylova, L.S.; Godsill, S.J. (Eds.)

2014, XII, 502 p. 135 illus., Hardcover

ISBN: 978-3-642-38397-7