Sparse Sensing for Statistical Inference

Sundeep Prabhakar Chepuri
Faculty of Electrical Engineering,
Mathematics, and Computer Science
Delft University of Technology
The Netherlands
s.p.chepuri@tudelft.nl

Geert Leus
Faculty of Electrical Engineering,
Mathematics, and Computer Science
Delft University of Technology
The Netherlands
g.j.t.leus@tudelft.nl

the essence of knowledge
Boston — Delft

Foundations and Trends® in Signal Processing

Published, sold and distributed by:
now Publishers Inc.
PO Box 1024
Hanover, MA 02339
United States
Tel. +1-781-985-4510
www.nowpublishers.com
sales@nowpublishers.com

Outside North America:
now Publishers Inc.
PO Box 179
2600 AD Delft
The Netherlands
Tel. +31-6-51115274

The preferred citation for this publication is

S.P. Chepuri and G. Leus. *Sparse Sensing for Statistical Inference*. Foundations and Trends® in Signal Processing, vol. 9, no. 3-4, pp. 233–386, 2015.

This Foundations and Trends® issue was typeset in LaTeX using a class file designed by Neal Parikh. Printed on acid-free paper.

ISBN: 978-1-68083-236-5

Foundations and Trends® in Signal Processing
Volume 9, Issue 3-4, 2015

Editorial Scope

Topics

Foundations and Trends® in Signal Processing publishes survey and tutorial articles in the following topics:

- Adaptive signal processing
- Audio signal processing
- Biological and biomedical signal processing
- Complexity in signal processing
- Digital signal processing
- Distributed and network signal processing
- Image and video processing
- Linear and nonlinear filtering
- Multidimensional signal processing
- Multimodal signal processing
- Multirate signal processing
- Multiresolution signal processing
- Nonlinear signal processing
- Randomized algorithms in signal processing
- Sensor and multiple source signal processing, source separation
- Signal decompositions, subband and transform methods, sparse representations
- Signal processing for communications
- Signal processing for security and forensic analysis, biometric signal processing
- Signal quantization, sampling, analog-to-digital conversion, coding and compression
- Signal reconstruction, digital-to-analog conversion, enhancement, decoding and inverse problems
- Speech/audio/image/video compression
- Speech and spoken language processing
- Statistical/machine learning
- Statistical signal processing

Information for Librarians

Foundations and Trends® in Signal Processing, 2015, Volume 9, 4 issues. ISSN paper version 1932-8346. ISSN online version 1932-8354. Also available as a combined paper and online subscription.

Foundations and Trends® in Signal Processing
Vol. 9, No. 3-4 (2015) 233–386

DOI: 10.1561/2000000069

Sparse Sensing for Statistical Inference

Sundeep Prabhakar Chepuri
Faculty of Electrical Engineering,
Mathematics, and Computer Science
Delft University of Technology
The Netherlands
s.p.chepuri@tudelft.nl

Geert Leus
Faculty of Electrical Engineering,
Mathematics, and Computer Science
Delft University of Technology
The Netherlands
g.j.t.leus@tudelft.nl

Contents

Abstract

In today's society, we are flooded with massive volumes of data in the order of a billion gigabytes on a daily basis from pervasive sensors. It is becoming increasingly challenging to sense, store, transport, or process (i.e., for inference) the acquired data. To alleviate these problems, it is evident that there is an urgent need to significantly reduce the sensing cost (i.e., the number of expensive sensors) as well as the related memory and bandwidth requirements by developing unconventional sensing mechanisms to extract as much information as possible yet collecting fewer data.

The aim of this monograph is therefore to develop theory and algorithms for smart data reduction. We develop a data reduction tool called sparse sensing, which consists of a deterministic and structured sensing function (guided by a sparse vector) that is optimally designed to achieve a desired inference performance with the reduced number of data samples. We develop sparse sensing mechanisms, convex programs, and greedy algorithms to efficiently design sparse sensing functions, where we assume that the data is not yet available and the model information is perfectly known.

Sparse sensing offers a number of advantages over compressed sensing (a state-of-the-art data reduction method for sparse signal recovery). One of the major differences is that in sparse sensing the underlying signals need not be sparse. This allows for general signal processing tasks (not just sparse signal recovery) under the proposed sparse sensing framework. Specifically, we focus on fundamental statistical inference tasks, like estimation, filtering, and detection. In essence, we present topics that transform classical (e.g., random or uniform) sensing methods to low-cost data acquisition mechanisms tailored for specific inference tasks. The developed framework can be applied to sensor selection, sensor placement, or sensor scheduling, for example.

S.P. Chepuri and G. Leus. *Sparse Sensing for Statistical Inference*. Foundations and Trends® in Signal Processing, vol. 9, no. 3-4, pp. 233–386, 2015.
DOI: 10.1561/2000000069.

1

Introduction

1.1 Pervasive sensors and data deluge

Every day, we are generating data in the order of a billion gigabytes. This massive volume of data comes from omnipresent sensors used in medical imaging (e.g., breast or fetal ultrasound), seismic processing (e.g., for oil or gas field exploration), environmental monitoring (e.g., pollution, temperature, precipitation sensing), radio astronomy (e.g., from radio telescopes like the square kilometre array), power networks (e.g., to monitor wind farms or other distribution grids), smart infrastructures (e.g., to monitor the condition of railway tracks or bridges), localization and surveillance platforms (e.g., security cameras or drones, indoor navigation), and so on.

The acquired data samples are stored locally and then transported to a central location (e.g., a server or cloud) to extract meaningful information (that is, for inference). Due to an unprecedented increase in the volume of the acquired data, it is becoming increasingly challenging to locally store and transport all the data samples to a central location for data/signal processing. This is because the amount of sampled data quickly exceeds the storage and communication capacity by several orders of magnitude. Since the data processing is generally carried out at a central location with ample computing power, mainly the sensing, storage and transportation costs form the main bottleneck. To alleviate these bottlenecks, most of the data is blindly discarded without even being examined in order to limit the memory and communication requirements, causing a serious performance loss.

In this era of data deluge, it is of paramount importance to gather only the informative data needed for a specific task. If we had some prior knowledge about the task we want to perform on the data samples, then just a small portion of that data might be sufficient to reach a desired inference accuracy, thereby significantly reducing the amount of sampled, stored and transported data. That is to say, if the inference task is known beforehand, less data needs to be acquired. Thus, the memory and bandwidth requirements can be seriously curtailed. In addition, the cost of data collection (or sensing) can be significantly reduced, where the major factors that determine the sensing cost are the number of physical sensors (and their economical and energy costs) and the physical space they occupy when installed. So, it is evident that there is an urgent need for developing unconventional and innovative sensing mechanisms tailored for specific inference tasks to extract as much information as possible yet collecting fewer data. This leads to the main question:

> *How can task-cognition be exploited to reduce the costs of sensing as well as the related storage and communications requirements?*

This is different from the classical big data setting in which the data is already available and the question is how to mine information from that large-scale data. Our problem has close similarities to sampling, and is only related to model information, where the data is not yet available. Given the central role of sampling in engineering sciences, answering this question will impact a wide range of applications. The basic question of interest for such applications is, how to design sensing systems in order to minimize the amount of data acquired yet reach a prescribed inference performance. In particular, the design questions that should be answered are related to the optimal sensor placement in space and/or time, data rate, and sampling density to reduce the sensing cost as well as to reduce the storage and communications requirements. We next illustrate our ideas with two specific examples of sensor placement for indoor localization and temperature sensing.

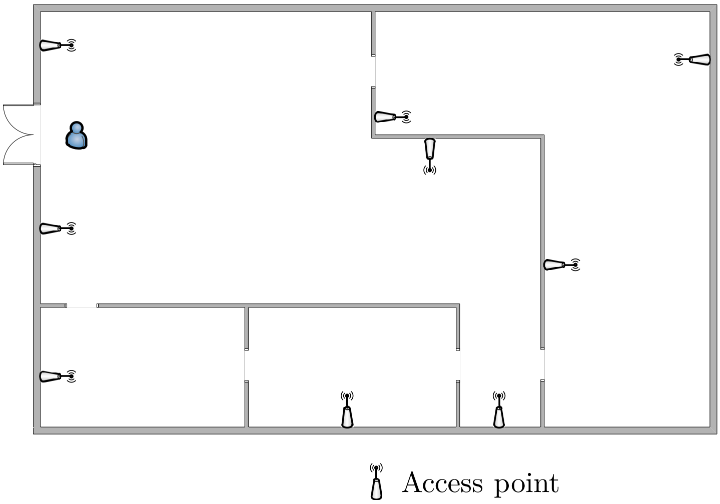
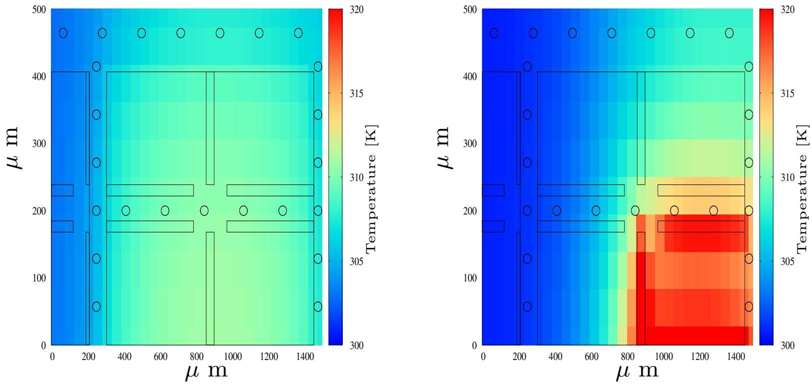


Figure 1.1: Illustration of an indoor localization setup. We show the floor plan of a building (e.g., museum) with candidate locations for installing the access points. The restriction on installing the access points in only certain areas might be for security or ambience purposes.

Example 1.1 (Target localization)**.** Indoor localization is becoming increasingly important in many applications (see [69]). Some examples include: locating people inside a building for rescue operations, monitoring logistics in a production plant, lighting control, and so on. In such environments, global positioning system (GPS) signals are typically unavailable. Thus, other types of measurements such as visual, acoustic or radio waves revealing information about range, bearing, and/or Doppler are used. These measurements are gathered by access points, like cameras, microphones, radars, or wireless transceivers. One such scenario is illustrated in Figure 1.1, where we show an indoor localization setup for navigating a visitor inside a building. An interesting question is, instead of installing many such costly access points randomly, how can we minimize the number of access points (hence, the amount of data), by optimizing their characteristics (e.g., their spatial position, sampling rate) in such a way that a certain localization performance can be guaranteed.



(a) **(b)**

Figure 1.2: Heatmaps of a 32KB data cache (a) without and (b) with a hot spot. Black circles (○) denote the candidate temperature sensor locations—these are the areas with less or no active logic.

Example 1.2 (Field detection)**.** Consider a multi-core processor with a hot spot. A historical question of interest is to estimate the thermal distribution, for instance, by interpolating noisy measurements. In some applications, though, a precise estimation of the temperature field might not be required. Instead, detecting the hot spots (i.e., the areas where the temperature exceeds a certain threshold) would be sufficient for subsequent control actions. Such a scenario is illustrated in Figure 1.2[1], where the image on the right (left) shows a 32 KB data cache with (no) hotspots. An important question of interest for such detection problems then is, how to optimally design spatial samplers (i.e., how to optimize the sensor placement [64]) by exploiting the knowledge of the underlying model, physical space and processing limitations.

Such optimally designed sensing systems can be used to perform a number of inference tasks, such as estimation, filtering, and detection.

[1]We would like to thank the authors of [57] for the heatmaps.

1.2 Outline

This monograph is organized into three parts. In the first part of this monograph (i.e., in Chapter 2), the theory of sparse sensing is discussed in depth. In the second part of this monograph (i.e., in Chapters 3—6) the developed theory in Chapter 2 is applied to basic statistical signal processing problems. Finally, the monograph concludes with the third part (i.e., Chapter 7), where we pose some interesting open problems for future research.

Chapter 2 on *sparse sensing* forms the backbone of this monograph. In order to reduce the sensing and other related costs, it is crucial to tailor the sensing mechanism for the specific inference task that will be performed on the acquired data samples. The tool that we will exploit in this monograph to reduce the cost of sensing is *sparse sensing*, which consists of an optimally designed structured and deterministic sparse (i.e., with many zeros and a few nonzeros) sensing function that is used to acquire the data in order to reach a desired inference performance. Here, the number of nonzeros determines the amount of data samples acquired (thus determines the amount of data reduction). In this chapter, we will model the sparse sensing function as a linear projection operation, where the sensing function is parameterized by a sparse vector. This vector is basically a design parameter that is used as a handle to trade the amount of acquired data samples with the inference performance. We refer to this sparse sensing scheme as *discrete sparse sensing*, as the continuous observation domain is first discretized into grid points and we select (using the sparse vector) the best subset out of those grid points. To harness the full potential of sparse sensing, we need to sample in between the grid points and take samples anywhere in the continuous observation domain. We refer to such sensing mechanisms as *continuous sparse sensing*. We will discuss some applications of the proposed sparse sensing mechanisms and also list major differences with the state of the art in data reduction, that is, compressed sensing. Although the specific inference task is kept abstract in this chapter, the obtained novel unifying view allows us to jointly treat sparse sensing mechanisms for the different inference tasks considered in Chapters 3—6.

Chapter 3 focuses on discrete sparse sensing for a general *nonlinear estimation* problem. In particular, we solve the problem of choosing the best subset of observations that follow known nonlinear models with arbitrary yet independent distributions. We also extend this framework to nonlinear colored Gaussian observations as it occurs frequently when the observations are subject to external noises or interference. The data is acquired using a discrete sparse sensing function, which is guided by a sparse vector. The Cramér-Rao bound (CRB) is used as an inference performance metric and we derive several functions of the CRB that include the sparse vector. To compute the optimal sparse samplers, we propose convex relaxations of the derived inference performance metric and also develop low-complexity solvers. We also discuss greedy algorithms leveraging the submodularity of the inference performance metric. In sum, we can conclude that discrete sparse samplers for nonlinear inverse problems can be computed efficiently (in polynomial or even linear time).

Chapter 4 extends the theory developed in Chapter 3 to *nonlinear filtering* problems, that is, the focus will be on the design of discrete sparse sensing functions for systems that admit a known nonlinear state-space representation. In particular, we solve the problem of choosing the best subset of time-varying observations based on the entire history of measurements up to that point. The posterior CRB is used as the inference performance metric to decide on the best subset of observations. Although this framework is valid for independent observations that follow arbitrary distributions (e.g., non-Gaussian), we also extend it to colored Gaussian observations. Further, we introduce some additional constraints to obtain smooth sensing patterns over time. Finally, we devise sparse sensing mechanisms for *structured* time-varying observations (e.g., for time-varying sparse signals). In all these cases, the discrete sparse samplers can be designed efficiently by solving a convex program or through a greedy algorithm that leverages on submodularity.

Chapter 5 is dedicated to discrete sparse sensing for *statistical detection*. Specifically, the aim is to choose the best subset of observations that are conditioned on the hypothesis, which belongs to a binary set. Naturally, the best subset of observations is the one that results in a prescribed global error probability. Since the numerical optimization of the error probabilities is difficult, we adopt simpler costs related to distance measures between the conditional distributions of the sensor observations. We design sparse samplers for the Bayesian and Neyman-Pearson setting, where we respectively use the Bhattacharyya distance and Kullback-Leibler distance (and J-divergence) as the inference performance metric. For conditionally independent observations, we give an explicit solution, which is optimal in terms of the error exponents. More specifically, the best subset of observations is the one with the smallest local average root-likelihood ratio and largest local average log-likelihood ratio in the Bayesian and Neyman-Pearson setting, respectively. We supplement the proposed framework with a thorough analysis for Gaussian observations, including the case when the sensors are conditionally dependent, and also provide examples for other observation distributions. One of the results shows that, for nonidentical Gaussian sensor observations with uncommon means and common covariances under both hypotheses, the number of sensors required to achieve a desired detection performance reduces significantly as the sensors become more coherent.

Chapter 6 contrasts with the discrete sparse sensing mechanisms that have been considered in Chapter 3 to Chapter 5, where the sparse sensing functions are parameterized by a discrete sparse vector that needs to be optimally designed. This basically means that the continuous observation domain is first discretized into grid points and we have to select the best subset out of those grid points. However, this discretization might be very coarse because of complexity reasons, preventing the system to achieve the best possible compression rates for the considered inference task. Therefore, in this chapter, we introduce *continuous sparse sensing* (or off-the-grid sparse sensing), where it is possible to sample in between the grid points and take samples anywhere in the continuous observation domain. The basic idea is to start from a discretized sampling space and to model every sampling point in the continuous sampling space as a discrete sampling point plus a perturbation. Then, the smallest set of possible discrete sampling points is searched for, along with the best possible perturbations, in order to reach the prescribed inference performance. We will demonstrate this approach for linear inverse problems, that is, for linear estimation problems with additive Gaussian noise, although it can be extended for other inference problems as well.

Chapter 7 contains the conclusions and outlines a number of directions for future research along with some open problems.

References

[1] SM Ali and Samuel D Silvey. A general class of coefficients of divergence of one distribution from another. *Journal of the Royal Stat. Society. Series B (Methodological)*, 28(1):131–142, 1966.

[2] Daniele Angelosante, Georgios B Giannakis, and Emanuele Grossi. Compressed sensing of time-varying signals. In *Proc. of 16th International Conference on Digital Signal Processing*, Jul. 2009, Santorini, (Hellas) Greece, 2009.

[3] Swaroop Appadwedula, Venugopal V Veeravalli, and Douglas L Jones. Decentralized detection with censoring sensors. *IEEE Trans. Signal Process.*, 56(4):1362–1373, 2008.

[4] Randall K Bahr and James A Bucklew. Optimal sampling schemes for the Gaussian hypothesis testing problem. *IEEE Trans. Acoust., Speech, Signal Process.*, 38(10):1677–1686, 1990.

[5] Adam Baig and Ted Urbancic. Microseismic moment tensors: A path to understanding frac growth. *The Leading Edge*, 29(3):320–324, 2010.

[6] D. Bajovic, B. Sinopoli, and J. Xavier. Sensor selection for event detection in wireless sensor networks. *IEEE Trans. Signal Process.*, 59(10):4938–4953, Oct 2011.

[7] Richard G Baraniuk. Compressive sensing. *IEEE Signal Process. Mag.*, 24(4):118–121, 2007.

[8] D. P. Bertsekas. *Nonlinear programming.* Athena Scientific optimization and computation series. Belmont, MA: Athena Scientific, 1999.

[9] Rick S Blum and Brian M Sadler. Energy efficient signal detection in sensor networks using ordered transmissions. *IEEE Trans. Signal Process.*, 56(7):3229–3235, 2008.

[10] S. Boyd and L. Vandenberghe. *Convex Optimization.* Cambridge, U.K.: Cambridge Univ. Press, 2004.

[11] S. Boyd, L. Xiao, and A. Mutapcic. *Subgradient methods.* Lecture notes of EE364b. Stanford Univ., Stanford, CA, USA, 2003.

[12] R. Calderbank, S. Howard, and S. Jafarpour. Construction of a large class of deterministic sensing matrices that satisfy a statistical isometry property. *IEEE J. Sel. Topics Signal Process.*, 4(2):358–374, 2010.

[13] S. Cambanis and E. Masry. Sampling designs for the detection of signals in noise. *IEEE Trans. Inf. Theory*, 29(1):83–104, Jan 1983.

[14] E. Candès and M.B. Wakin. An introduction to compressive sampling. *IEEE Signal Process. Mag.*, 25(2):21–30, 2008.

[15] E. Candès, M.B. Wakin, and S. Boyd. Enhancing sparsity by reweighted ℓ_1 minimization. *Journal of Fourier Analysis and Applications*, 14(5):877–905, 2008.

[16] A. Carmi, P. Gurfil, and D. Kanevsky. Methods for sparse signal recovery using Kalman filtering with embedded pseudo-measurement norms and quasi-norms. *IEEE Trans. Signal Process.*, 58(4):2405–2409, April 2010.

[17] J. Chamberland and V.V. Veeravalli. Wireless sensors in distributed detection applications. *IEEE Signal Process. Mag.*, 24(3):16–25, May 2007.

[18] S. P. Chepuri and G. Leus. Sparse sensing for distributed Gaussian detection. In *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing*, May 2015, Brisbane, Australia, 2015.

[19] S. P. Chepuri and G. Leus. Sparsity-promoting sensor selection for non-linear measurement models. *IEEE Trans. Signal Process.*, 63(3):684–698, Feb. 2015.

[20] S. P. Chepuri, G. Leus, and A.-J. van der Veen. Sparsity-exploiting anchor placement for localization in sensor networks. In *Proc. of the 21st European Signal Processing Conference*, Sept. 2013, Marrakech, Morocco, 2013.

[21] Thomas M Cover and Joy A Thomas. *Elements of information theory.* Hoboken, NJ, USA: John Wiley & Sons, 2012.

[22] David L Donoho. Compressed sensing. *IEEE Trans. Inf. Theory*, 52(4):1289–1306, 2006.

[23] D.L. Donoho. For most large underdetermined systems of linear equations the minimal ℓ_1-norm solution is also the sparsest solution. *Communications on pure and applied mathematics*, 59(6):797–829, 2006.

[24] P. Drineas, M. Mahoney, and S. Muthukrishnan. Sampling algorithms for ℓ_2 regression and applications. In *Proc. of 17th annual ACM-SIAM symp. on Discrete algorithm*, Jan. 2006, Miami, FL, USA, 2006.

[25] C. Ekanadham, D. Tranchina, and E.P. Simoncelli. Recovery of sparse translation-invariant signals with continuous basis pursuit. *IEEE Trans. Signal Process.*, 59(10):4735–4744, Oct 2011.

[26] Dalia El Badawy, Juri Ranieri, and Martin Vetterli. Near-optimal sensor placement for signals lying in a union of subspaces. In *Proc. of 22nd European Signal Processing Conference*, Sept. 2014, Lisbon, Portugal, 2014.

[27] Michael Elad. Optimized projections for compressed sensing. *IEEE Trans. Signal Process.*, 55(12):5695–5702, 2007.

[28] Yonina C Eldar and Gitta Kutyniok. *Compressed sensing: theory and applications*. Cambridge, U.K.: Cambridge Univ. Press, 2012.

[29] S. Farahmand, G. B. Giannakis, G. Leus, and Z. Tian. Tracking target signal strengths on a grid using sparsity. *EURASIP Journal on Advances in Signal Processing*, 2014(1):7, 2014.

[30] I. Ford, D. M. Titterington, and Christos P. Kitsos. Recent advances in nonlinear experimental design. *Technometrics*, 31(1):49–60, 1989.

[31] Jerome Friedman, Trevor Hastie, and Robert Tibshirani. A note on the group LASSO and a sparse group LASSO. *arXiv preprint arXiv:1001.0736*, 2010.

[32] Yinfei Fu, Qing Ling, and Zhi Tian. Distributed sensor allocation for multi-target tracking in wireless sensor networks. *IEEE Trans. Aerosp. Electron. Syst.*, 48(4):3538–3553, 2012.

[33] S. Gezici, Z. Tian, G. B. Giannakis, H. Kobayashi, A. F. Molisch, H. V. Poor, and Z. Sahinoglu. Localization via ultra-wideband radios: a look at positioning aspects for future sensor networks. *IEEE Signal Process. Mag.*, 22(4):70–84, Jul. 2005.

[34] G. H. Golub and C. F. Van Loan. *Matrix Computations*. Johns Hopkins Studies in the Mathematical Sciences. Baltimore, MD, USA: Johns Hopkins Univ. Press, 1996.

[35] M. Grant and S. Boyd. CVX: Matlab software for disciplined convex programming, version 2.0 beta. `http://cvxr.com/cvx`, September 2012.

[36] T Grettenberg. Signal selection in communication and radar systems. *IEEE Trans. Inf. Theory*, 9(4):265–275, 1963.

[37] F. Gustafsson and F. Gunnarsson. Mobile positioning using wireless networks: possibilities and fundamental limitations based on available wireless network measurements. *IEEE Signal Process. Mag.*, 22(4):41 – 53, Jul. 2005.

[38] Jarvis Haupt, Robert Nowak, and Rui Castro. Adaptive sensing for sparse signal recovery. In *Proc. of IEEE 13th Digital Signal Processing Workshop and 5th IEEE Signal Processing Education Workshop*, Jan. 2009, Marco Island, Florida, USA, 2009.

[39] M.L. Hernandez, T. Kirubarajan, and Y. Bar-Shalom. Multisensor resource deployment using Posterior Cramér-Rao bounds. *IEEE Trans. Aerosp. Electron. Syst.*, 40(2):399–416, 2004.

[40] Wassily Hoeffding. Probability inequalities for sums of bounded random variables. *Journal of the American statistical association*, 58(301):13–30, 1963.

[41] Shihao Ji, Ya Xue, and Lawrence Carin. Bayesian compressive sensing. *IEEE Trans. Signal Process.*, 56(6):2346–2356, 2008.

[42] F. Jiang, J. Chen, and A. L. Swindlehurst. Linearly reconfigurable Kalman filtering for a vector process. In *Proc. of IEEE International Conference on Acoustics, Speech and Signal Processing*, May 2013, Vancouver, Canada, May 2013.

[43] S. Joshi and S. Boyd. Sensor selection via convex optimization. *IEEE Trans. Signal Process.*, 57(2):451–462, Feb. 2009.

[44] S.J. Julier and J.J. LaViola. On Kalman filtering with nonlinear equality constraints. *IEEE Trans. Signal Process.*, 55(6):2774–2784, June 2007.

[45] T Kadota and Larry A Shepp. On the best finite set of linear observables for discriminating two Gaussian signals. *IEEE Trans. Inf. Theory*, 13(2):278–284, 1967.

[46] G. Kail, S.P. Chepuri, and G. Leus. Robust censoring using metropolis-hastings sampling. *IEEE J. Sel. Topics Signal Process.*, 10(2):270–283, 2016.

[47] T. Kailath. The divergence and Bhattacharyya distance measures in signal selection. *IEEE Trans. Commun. Technol.*, 15(1):52–60, February 1967.

[48] S. M. Kay. *Fundamentals of Statistical Signal Processing: Estimation Theory.* Englewood Cliffs, NJ, USA: Prentice-Halll, 1993.

[49] V. Kekatos and G. B. Giannakis. From sparse signals to sparse residuals for robust sensing. *IEEE Trans. Signal Process.*, 59(7):3355–3368, 2011.

[50] V. Kekatos, G. B. Giannakis, and B. Wollenberg. Optimal placement of phasor measurement units via convex relaxation. *IEEE Trans. Power Syst.*, 27(3):1521–1530, Aug. 2012.

[51] HISASHI Kobayashi and JOHN B Thomas. Distance measures and related criteria. In *Proc. 5th Annu. Allerton Conf. Circuit and System Theory*, Oct. 1967, Monticello, IL, USA, 1967.

[52] A. Krause and C. Guestrin. Near-optimal observation selection using submodular functions. In *Proc. of Twenty-Second Conference on Artificial Intelligence*, Jul. 2007, Vancouver, Canada, 2007.

[53] A. Krause, A. Singh, and C. Guestrin. Near-optimal sensor placements in Gaussian processes: Theory, efficient algorithms and empirical studies. *The Journal of Machine Learning Research*, 9(2):235–284, Feb. 2008.

[54] Andreas Krause. *Optimizing sensing: Theory and applications*. Ph.D. dissertation, School of Comput. Sci. Carnegie Mellon Univ., Pittsburgh, PA, United States, 2008.

[55] Andreas Krause, H Brendan McMahan, Carlos Guestrin, and Anupam Gupta. Robust submodular observation selection. *The Journal of Machine Learning Research*, 9(12):2761–2801, 2008.

[56] S. Kullback. *Information theory and statistics*. Hoboken, NJ, USA: John Wiley & Sons, 1959.

[57] S.S. Kumar, A. Zjajo, and R. van Leuken. Ctherm: An integrated framework for thermal-functional co-simulation of systems-on-chip. In *Proc. of the 23rd IEEE/Euromicro Conference on Parallel, Distributed and Network-Based Processing*, Mar. 2015, Turku, Finland, 2015.

[58] Sijia Liu, Mohammad Fardad, Pramod K Varshney, and Engin Masazade. Optimal periodic sensor scheduling in networks of dynamical systems. *IEEE Trans. Signal Process.*, 62(12):3055–3068, 2014.

[59] J. Lofberg. YALMIP : A toolbox for modeling and optimization in MATLAB. In *Proc. of International Symposium on Computer Aided Control System Design*, Sept. 2004, Taipei, Taiwan, 2004.

[60] Dmitry M Malioutov, Sujay R Sanghavi, and Alan S Willsky. Sequential compressed sensing. *IEEE J. Sel. Topics Signal Process.*, 4(2):435–444, 2010.

[61] Farokh Marvasti. *Nonuniform sampling: theory and practice*. New York, USA: Kluwer Academic/Plenum Publishers, 2001.

[62] P. Marziliano and M. Vetterli. Reconstruction of irregularly sampled discrete-time bandlimited signals with unknown sampling locations. *IEEE Trans. Signal Process.*, 48(12):3462–3471, 2000.

[63] E. Masazade, M. Fardad, and P.K. Varshney. Sparsity-promoting extended Kalman filtering for target tracking in wireless sensor networks. *IEEE Signal Process. Lett.*, 19(12):845–848, 2012.

[64] Seda Ogrenci Memik, Rajarshi Mukherjee, Min Ni, and Jieyi Long. Optimizing thermal sensor allocation for microprocessors. *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, 27(3):516–527, 2008.

[65] Todd K Moon and Wynn C Stirling. *Mathematical methods and algorithms for signal processing.* Englewood Cliffs, NJ, USA: Prentice-Hall, 2000.

[66] E. J. Msechu and G. B. Giannakis. Sensor-centric data reduction for estimation with WSNs via censoring and quantization. *IEEE Trans. Signal Process.*, 60(1):400–414, 2012.

[67] George L Nemhauser, Laurence A Wolsey, and Marshall L Fisher. An analysis of approximations for maximizing submodular set functions— I. *Mathematical Programming*, 14(1):265–294, 1978.

[68] C. H. Papadimitriou. *Computational complexity.* Hoboken, NJ, USA: John Wiley & Sons, 2003.

[69] N. Patwari, J. N. Ash, S. Kyperountas, A. O. Hero III, R.L. Moses, and N.S. Correal. Locating the nodes: cooperative localization in wireless sensor networks. *IEEE Signal Process. Mag.*, 22(4):54 – 69, Jul. 2005.

[70] Boris Polyak, Mikhail Khlebnikov, and Pavel Shcherbakov. An LMI approach to structured sparse feedback design in linear control systems. In *Proc. of European Control Conference (ECC)*, Jul. 2013, Zurich, Switzerland, 2013.

[71] F. Pukelsheim. *Optimal design of experiments*, volume 50. Philadelphia, PA: SIAM, 1993.

[72] Zhi Quan, William J Kaiser, and Ali H Sayed. Innovations diffusion: A spatial sampling scheme for distributed estimation and detection. *IEEE Trans. Signal Process.*, 57(2):738–751, 2009.

[73] Constantino Rago, Peter Willett, and Yaakov Bar-Shalom. Censoring sensors: A low-communication-rate scheme for distributed detection. *IEEE Trans. Aerosp. Electron. Syst.*, 32(2):554–568, 1996.

[74] J. Ranieri, A. Chebira, and M. Vetterli. Near-optimal sensor placement for linear inverse problems. *IEEE Trans. Signal Process.*, 62(5):1135–1146, Mar. 2014.

[75] Juri Ranieri and Martin Vetterli. Near-optimal source placement for linear physical fields. In *Proc. of IEEE International Conference on Acoustics, Speech and Signal Processing*, May 2014, Florence, Italy, 2014.

[76] S. Rao, S. P. Chepuri, and G. Leus. Greedy sensor selection for non-linear models. In *Proc. of the IEEE Workshop on Comp. Adv. in Multi-Sensor Adaptive Processing*, Dec. 2015, Cancun, Mexico, 2015.

[77] T.J. Rothenberg. Identification in parametric models. *Econometrica: Journal of the Econometric Society*, 39(3):577–591, 1971.

[78] Louis L Scharf. *Statistical signal processing*. Reading, MA, USA: Addison-Wesley, 1991.

[79] M. Shamaiah, S. Banerjee, and H. Vikalo. Greedy sensor selection: Leveraging submodularity. In *Proc. of 49th IEEE Conference on Decision and Control*, Dec. 2010, Atlanta, Georgia, USA, 2010.

[80] K. Slavakis, G. Giannakis, and G. Mateos. Modeling and optimization for big data analytics:(statistical) learning tools for our era of data deluge. *IEEE Signal Process. Mag.*, 31(5):18–31, 2014.

[81] P. Stoica and B. C. Ng. On the Cramér-Rao bound under parametric constraints. *IEEE Signal Process. Lett.*, 5(7):177–179, July 1998.

[82] Jos F Sturm. Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones. *Optimization methods and software*, 11(1-4):625–653, 1999.

[83] Youngchul Sung, Lang Tong, and H Vincent Poor. A large deviations approach to sensor scheduling for detection of correlated random fields. In *Proc. of IEEE International Conference on Acoustics, Speech and Signal Processing*, pages Mar. 2005, Philadelphia, PA, USA, 2005.

[84] Robert Tibshirani, Michael Saunders, Saharon Rosset, Ji Zhu, and Keith Knight. Sparsity and smoothness via the fused LASSO. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 67(1):91–108, 2005.

[85] P. Tichavsky, C.H. Muravchik, and A. Nehorai. Posterior Cramér-Rao bounds for discrete-time nonlinear filtering. *IEEE Trans. Signal Process.*, 46(5):1386–1396, 1998.

[86] PP Vaidyanathan. Generalizations of the sampling theorem: Seven decades after Nyquist. *IEEE Trans. Circuits Syst. I: Fundamental Theory and Applications*, 48(9):1094–1109, 2001.

[87] Alle-Jan van der Veen and Stefan J Wijnholds. Signal processing tools for radio astronomy. In *Handbook of Signal Processing Systems*, pages 421–463. Springer, 2013.

[88] Harry L Van Trees. *Detection, Estimation, and Modulation Theory, Optimum Array Processing*. Hoboken, NJ, USA: John Wiley & Sons, 2004.

[89] Namrata Vaswani. Kalman filtered compressed sensing. In *Proc. of 15th IEEE International Conference on Image Processing*, Oct. 2008, San Diego, CA, USA, 2008.

[90] Noam Wagner, Yonina C Eldar, and Zvi Friedman. Compressed beamforming in ultrasound imaging. *IEEE Trans. Signal Process.*, 60(9):4643–4657, 2012.

[91] T. Wang, G. Leus, and L. Huang. Ranging energy optimization for robust sensor positioning based on semidefinite programming. *IEEE Trans. Signal Process.*, 57(12):4777 –4787, Dec. 2009.

[92] Garrett Warnell, Dikpal Reddy, and Rama Chellappa. Adaptive rate compressive sensing for background subtraction. In *Proc. of IEEE International Conference on Acoustics, Speech and Signal Processing*, Mar. 2012, Kyoto, Japan, 2012.

[93] Lin Xiao and S. Boyd. Fast linear iterations for distributed averaging. *Systems & Control Letters*, 53(1):65–78, 2004.

[94] L. Yao, W.A. Sethares, and D.C. Kammer. Sensor placement for on-orbit modal identification via a genetic algorithm. *The American Institute of Aeronautics and Astronautics Journal*, 31(10):1922–1928, 1993.

[95] Chao-Tang Yu and Pramod K Varshney. Sampling design for Gaussian detection problems. *IEEE Trans. Signal Process.*, 45(9):2328–2337, 1997.

[96] Alan L Yuille and Anand Rangarajan. The concave-convex procedure. *Neural computation*, 15(4):915–936, 2003.

[97] Pengcheng Zhan, D.W. Casbeer, and A.L. Swindlehurst. Adaptive mobile sensor positioning for multi-static target tracking. *IEEE Trans. Aerosp. Electron. Syst.*, 46(1):120–132, 2010.

[98] Hao Zhu, G. Leus, and G. B. Giannakis. Sparsity-cognizant total least-squares for perturbed compressive sampling. *IEEE Trans. Signal Process.*, 59(5):2002–2016, May 2011.

[99] Long Zuo, Ruixin Niu, and P.K. Varshney. Posterior CRLB based sensor selection for target tracking in sensor networks. In *Proc. of IEEE International Conference on Acoustics, Speech and Signal Processing*, Apr. 2007, Honolulu, Hawaii, USA, 2007.