Sparse Regression Codes

Other titles in Foundations and Trends® in Communications and Information Theory

Fundamentals of Index Coding
Fatemeh Arbabjolfaei and Young-Han Kim
ISBN: 978-1-68083-492-5

Community Detection and Stochastic Block Models
Emmanuel Abbe
ISBN: 978-1-68083-476-5

Redundancy of Lossless Data Compression for Known Sources by Analytic Methods
Michael Drmota and Wojciech Szpankowski
ISBN: 978-1-68083-284-6

Multiterminal Secrecy by Public Discussion
Prakash Narayan and Himanshu Tyagi
ISBN: 978-1-68083-186-3

Channel Coding Methods for Non-Volatile Memories
Lara Dolecek and Frederic Sala
ISBN: 978-1-68083-102-3

Sparse Regression Codes

Ramji Venkataramanan
University of Cambridge
ramji.v@eng.cam.ac.uk

Sekhar Tatikonda
Yale University
sekhar.tatikonda@yale.edu

Andrew Barron
Yale University
andrew.barron@yale.edu

now
the essence of knowledge
Boston — Delft

Foundations and Trends® in Communications and Information Theory

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

R. Venkataramanan, S. Tatikonda and A. Barron. *Sparse Regression Codes*. Foundations and Trends® in Communications and Information Theory, vol. 15, no. 1-2, pp. 1–195, 2019.

ISBN: 978-1-68083-581-6

Foundations and Trends® in Communications and Information Theory

Volume 15, Issue 1-2, 2019

Editorial Board

Editorial Scope

Topics

Foundations and Trends® in Communications and Information Theory publishes survey and tutorial articles in the following topics:

- Coded modulation
- Coding theory and practice
- Communication complexity
- Communication system design
- Cryptology and data security
- Data compression
- Data networks
- Demodulation and Equalization
- Denoising
- Detection and estimation
- Information theory and statistics
- Information theory and computer science
- Joint source/channel coding
- Modulation and signal design
- Multiuser detection
- Multiuser information theory
- Optical communication channels
- Pattern recognition and learning
- Quantization
- Quantum information processing
- Rate-distortion theory
- Shannon theory
- Signal processing for communications
- Source coding
- Storage and recording codes
- Speech and Image Compression
- Wireless Communications

Information for Librarians

Foundations and Trends® in Communications and Information Theory, 2019, Volume 15, 4 issues. ISSN paper version 1567-2190. ISSN online version 1567-2328 . Also available as a combined paper and online subscription.

Contents

Sparse Regression Codes

Ramji Venkataramanan[1], Sekhar Tatikonda[2] and Andrew Barron[3]

[1] *University of Cambridge; ramji.v@eng.cam.ac.uk*

[2] *Yale University; sekhar.tatikonda@yale.edu*

[3] *Yale University; andrew.barron@yale.edu*

ABSTRACT

Developing computationally-efficient codes that approach the Shannon-theoretic limits for communication and compression has long been one of the major goals of information and coding theory. There have been significant advances towards this goal in the last couple of decades, with the emergence of turbo codes, sparse-graph codes, and polar codes. These codes are designed primarily for discrete-alphabet channels and sources. For Gaussian channels and sources, where the alphabet is inherently continuous, *Sparse Superposition Codes* or *Sparse Regression Codes* (SPARCs) are a promising class of codes for achieving the Shannon limits.

This monograph provides a unified and comprehensive over-view of sparse regression codes, covering theory, algorithms, and practical implementation aspects. The first part of the monograph focuses on SPARCs for AWGN channel coding, and the second part on SPARCs for lossy compression (with squared error distortion criterion). In the third part, SPARCs are used to construct codes for Gaussian multi-terminal channel and source coding models such as broadcast channels, multiple-access channels, and source and channel coding with side information. The monograph concludes with a discussion of open problems and directions for future work.

Ramji Venkataramanan, Sekhar Tatikonda and Andrew Barron (2019), "Sparse Regression Codes", Foundations and Trends® in Communications and Information Theory: Vol. 15, No. 1-2, pp 1–195. DOI: 10.1561/0100000092.

1

Introduction

Developing computationally-efficient codes that approach the Shannon-theoretic limits for communication and compression has long been one of the major goals of information and coding theory. There have been significant advances towards this goal in the last couple of decades, with the emergence of turbo and sparse-graph codes in the '90s [20, 28, 92], and more recently polar codes and spatially-coupled LDPC codes [4, 68, 73]. These codes are primarily designed for channels with discrete input alphabet, and for discrete-alphabet sources.

There are many channels and sources of practical interest where the alphabet is inherently continuous, e.g., additive white Gaussian noise (AWGN) channels, and Gaussian sources. This monograph discusses a class of codes for such Gaussian models called *Sparse Superposition Codes* or *Sparse Regression Codes* (SPARCs). These codes were introduced by Barron and Joseph [15, 63] for efficient communication over AWGN channels, but have since also been used for lossy compression [112, 113] and multi-terminal communication [114]. Our goal in this monograph is to provide a unified and comprehensive view of SPARCs, covering theory, algorithms, as well as practical implementation aspects.

To motivate the construction of SPARCs, let us begin with the standard AWGN channel. The goal is to construct codes with computationally efficient encoding and decoding that *provably* achieve the channel capacity $\mathcal{C} = \frac{1}{2}\log_2(1+\mathsf{snr})$ bits/transmission, where snr denotes the signal-to-noise ratio. In particular, we are interested in codes whose encoding and decoding complexity grows no faster than a low-order polynomial in the block length n.

Though it is well known that rates approaching $\mathcal{C}$ can be achieved with Gaussian codebooks, this has been largely avoided in practice because of the high decoding complexity of unstructured Gaussian codes. Instead, the popular approach has been to separate the design of the coding scheme into two steps: *coding* and *modulation*. State-of-the-art coding schemes for the AWGN channel such as coded modulation [43, 50, 22] use this two-step design, and combine binary error-correcting codes such as LDPC and turbo codes with standard modulation schemes such as Quadrature Amplitude Modulation (QAM). Though such schemes have good empirical performance, they have not been proven to be capacity-achieving for the AWGN channel. With sparse regression codes, we step back from the coding/modulation divide and instead use a structured codebook to construct low-complexity, capacity-achieving schemes tailored to the AWGN channel.

There have been several lattice based schemes [40, 122] proposed for communication over the AWGN channel, including low density lattice codes [101] and polar lattices [118, 2]. The reader is referred to the cited works for details of the performance vs. complexity trade-offs of these codes.

In the rest of this chapter, we describe the sparse regression codebook, and give a brief overview of the topics covered in the later chapters. First, we lay down some notation that will be used throughout the monograph.

Notation The Gaussian distribution with mean μ and variance σ^2 is denoted by $\mathcal{N}(\mu, \sigma^2)$. For a positive integer L, we use $[L]$ to denote the set $\{1, \dots, L\}$. The Euclidean norm of a vector x is denoted by $\|x\|$. The indicator function of an event $\mathcal{E}$ is denoted by $\mathbf{1}\{\mathcal{E}\}$. The transpose of a matrix A is denoted by A^*. The $n \times n$ identity matrix is denoted by I_n, with the subscript dropped when it is clear from context.

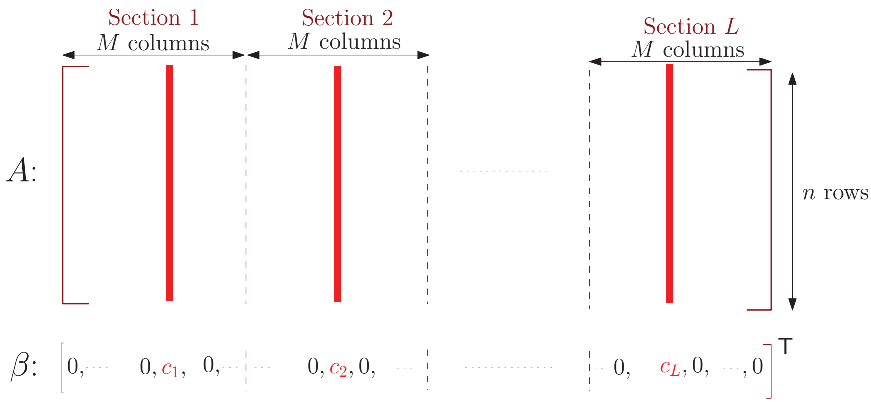

Figure 1.1: A Gaussian sparse regression codebook of block length n: A is a design matrix with independent Gaussian entries, and β is a sparse vector with one non-zero in each of L sections. Codewords are of the form $A\beta$, i.e., linear combinations of the columns corresponding to the non-zeros in β. The message is indexed by the *locations* of the non-zeros, and the values $c_1, \dots, c_L$ are fixed a priori.

Both log and ln are used to denote the natural logarithm. Logarithms to the base 2 are denoted by $\log_2$. For most of the theoretical analysis, we will find it convenient to use natural logarithms. Therefore, rate is measured in *nats*, unless otherwise specified. Throughout, we use n for the block length of the code.

For random vectors X, Y defined on the same probability space, we write $X \overset{d}{=} Y$ to indicate that X and Y have the same distribution.

1.1 The Sparse Regression Codebook

As shown in Fig. 1.1, a SPARC is defined in terms of a 'dictionary' or design matrix A of dimension $n \times ML$, whose entries are chosen i.i.d. $\sim \mathcal{N}(0, \frac{1}{n})$. Here n is the block length, and M, L are integers whose values are specified below in terms of n and the rate R. We think of the matrix A as being composed of L sections with M columns each. The variance of the entries ensures that the lengths of the columns of A are close to 1 for large n. [1]

[1] In some papers, the entries of A are assumed to be $\sim_{i.i.d} \mathcal{N}(0, 1)$. For consistency, throughout this monograph we will assume that the entries are $\sim_{i.i.d.} \mathcal{N}(0, 1/n)$.

Each codeword is a linear combination of L columns, with exactly one column chosen per section. Formally, a codeword can be expressed as $A\beta$, where $\beta = (\beta_1, \dots, \beta_{ML})^*$ is a length ML message vector with the following property: there is exactly one non-zero β_j for $1 \le j \le M$, one non-zero β_j for $M+1 \le j \le 2M$, and so forth. We denote the set of valid message vectors by $\mathcal{B}_{M,L}$. Since each of the L sections contains M columns, the size of this set is

$$|\mathcal{B}_{M,L}| = M^L. \tag{1.1}$$

The non-zero value of β in section $\ell \in [L]$ is set to c_ℓ, where the coefficients $\{c_\ell\}$ are specified a priori. Since the entries of A are i.i.d. $\mathcal{N}(0, \frac{1}{n})$, the entries of the codeword $A\beta$ are i.i.d. $\mathcal{N}(0, \frac{1}{n}\sum_{\ell=1}^{L} c_\ell^2)$. In the case of AWGN channel coding, the variance $\frac{1}{n}\sum_{\ell=1}^{L} c_\ell^2$ is equal to the average symbol power.

Rate: Since each of the L sections contains M columns, the total number of codewords is M^L. To obtain a rate R code, we need

$$M^L = e^{nR} \quad \text{or} \quad L \log M = nR. \tag{1.2}$$

There are several choices for the pair (M, L) which satisfy (1.2). For example, $L = 1$ and $M = e^{nR}$ recovers the Shannon-style random codebook in which the number of columns in A is e^{nR}. For most of our constructions, we will often choose M equal to L^{a}, for some constant $\mathsf{a} > 0$. In this case, (1.2) becomes

$$\mathsf{a} L \log L = nR. \tag{1.3}$$

Thus $L = \Theta(\frac{n}{\log n})$, and the size of the design matrix A (given by $n \times ML = n \times L^{\mathsf{a}+1}$) grows polynomially in n. In our numerical simulations, typical values for L are 512 or 1024.

We note that the SPARC is a non-linear code with pairwise dependent codewords. Indeed, two codewords $A\beta$ and $A\beta'$ are dependent whenever the underlying message vectors β, β' share one or more common non-zero entries.

Subset superposition coding The SPARC described above has a partitioned structure, i.e., the message vector contains exactly one non-zero in each of the L sections, with each section having M entries. One could also define a non-partitioned SPARC, where a message can be indexed by *any* subset of L entries of the length-ML vector β. The number of codewords in this case would be $\binom{ML}{L}$, compared to M^L for the partitioned case. For a given pair (M, L), the non-partitioned SPARC has a larger number of codewords. However, using Stirling's formula we find that

$$\frac{\log \binom{ML}{L}}{\log M^L} = 1 + O\left(\frac{1}{\log M}\right).$$

Hence the ratio of the rates tends to 1 as M grows large. Though subset based (non-partitioned) superposition codes have a small rate advantage for finite M, we focus on the partitioned structure in this monograph as it facilitates the design and analysis of efficient coding algorithms.

1.2 Organization of the monograph

In Part I, we focus on communication over the AWGN channel. The performance of SPARCs with optimal (least-squares) decoding is analyzed in Chapter 2. Though optimal decoding is infeasible, its performance provides a benchmark for the computationally efficient decoders described in the next chapter. It is shown that SPARCs with optimal encoding achieve the AWGN capacity with an error exponent of the same order as Shannon's random coding ensemble. Similar results are also obtained for SPARCs defined via Bernoulli dictionaries rather than Gaussian ones.

In Chapter 3, we describe three efficient iterative decoders. These decoders generate an updated estimate of the message vector in each iteration based on a test statistic. The first decoder makes hard decisions, decoding a few sections of the message vector β in each iteration. The other two decoders are based on soft-decisions, and generate new estimates of the whole message vector in each iteration. All three efficient decoders are asymptotically capacity-achieving, but the soft-decision decoders have better finite length error performance.

In Chapter 4, we turn our attention to techniques for improving the decoding performance at moderate block lengths. We observe that

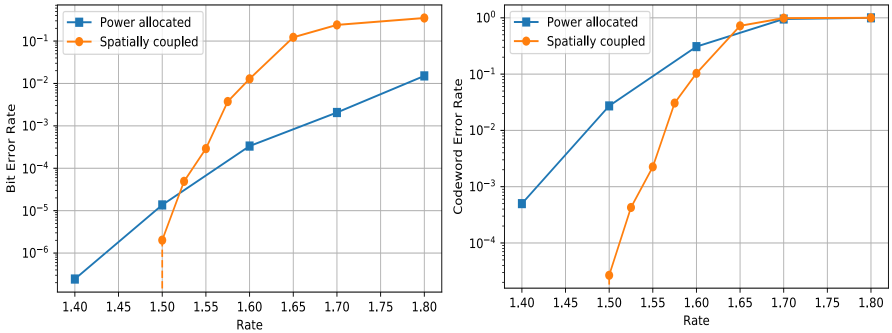

Figure 1.2: Average bit error rate (left) and codeword error rate (right) vs. rate for SPARC over an AWGN channel with snr $= 15$, $\mathcal{C} = 2$ bits. The SPARC parameters are $M = 512$, $L = 1024$, $n \in [5100, 7700]$. Curves are shown for for power allocated SPARC (Chapter 4) and spatially coupled SPARC (Chapter 5). The different ways of measuring error rate performance in a SPARC are discussed in Chapter 2 (p.12). The SPARC is decoded using the Approximate Message Passing (AMP) algorithm described in Chapters 3 and 5.

the power allocation (choice of the non-zero coefficients $\{c_\ell\}$) has a crucial effect on the finite length error performance. We describe an algorithm to determine a good power allocation, provide guidelines on choosing the parameters of the design matrix, and compare the empirical performance with coded modulation using LDPC codes from the WiMAX standard. In Chapter 5, we discuss spatially coupled SPARCs, which consist of several smaller SPARCs chained together in a band-diagonal structure. An attractive feature of spatially coupled SPARCs is that they are asymptotically capacity-achieving and have good finite length performance without requiring a tailored power allocation. Figure 1.2 shows the finite length error rate performance of power allocated SPARCs and spatially coupled SPARCs over an AWGN channel. The figure is discussed in detail in Sec. 5.4.

In Part II of the monograph, we use SPARCs for lossy compression with the squared error distortion criterion. In Chapter 6, we analyze compression with optimal (least-squares) encoding, and show that SPARCs attain the optimal rate-distortion function and the optimal excess-distortion exponent for i.i.d. Gaussian sources. We then describe an efficient successive cancellation encoder in Chapter 7, and show that it achieves the optimal Gaussian rate-distortion function, with the probability of excess distortion decaying exponentially in the block length.

In Part III, we design rate optimal coding schemes using SPARCs for a few canonical models in multiuser information theory. In Chapter 8, we show how SPARCs designed for point-to-point AWGN channels can be combined to construct rate-optimal superposition coding schemes for the AWGN broadcast and multiple-access channels. In Chapter 9, we show how to implement random binning using SPARCs. Using this, we can nest the channel coding and source coding SPARCs constructed in Parts I and II to construct rate-optimal schemes for a variety of problems in multiuser information theory. We conclude in Chapter 10 with a discussion of open problems and directions for future work.

Proofs or proof sketches for the main results in a chapter are given at the end of the chapter. The proofs of some intermediate lemmas are omitted, with pointers to the relevant references. The goal is to describe the key technical ideas in the proofs, while not impeding the flow within the chapter.

References

[1] *131.0-B-2 TM Synchonization and Channel Coding*. 2011. CCSDS. URL: https://public.ccsds.org/Pubs/131x0b2ec1.pdf.

[2] Abbe, E. and A. Barron. 2011. "Polar coding schemes for the AWGN channel". In: *Proc. IEEE Int. Symp. Inf. Theory*. IEEE. 194–198.

[3] Aref, V., N. Macris, and M. Vuffray. 2015. "Approaching the rate-distortion limit with spatial coupling, belief propagation, and decimation". *IEEE Trans. Inf. Theory*. 61(7): 3954–3979.

[4] Arikan, E. 2009. "Channel Polarization: A Method for Constructing Capacity-Achieving Codes for Symmetric Binary-Input Memoryless Channels". *IEEE Trans. Inf. Theory*. 55(7): 3051–3073.

[5] Ashikhmin, A., G. Kramer, and S. ten Brink. 2004. "Extrinsic information transfer functions: model and erasure channel properties". *IEEE Trans. Inf. Theory*. 50(11): 2657–2673.

[6] Bahadur, R. R. and R. R. Rao. 1960. "On Deviations of the Sample Mean". *The Annals of Mathematical Statistics*. 31(4).

[7] Barbier, J., M. Dia, and N. Macris. 2017. "Universal Sparse Superposition Codes with Spatial Coupling and GAMP Decoding". arXiv:1707.04203. July.

[8] Barbier, J., M. Dia, and N. Macris. 2016a. "Proof of threshold saturation for spatially coupled sparse superposition codes". In: *Proc. IEEE Int. Symp. Inf. Theory*.

[9] Barbier, J., M. Dia, and N. Macris. 2016b. "Threshold saturation of spatially coupled sparse superposition codes for all memoryless channels". In: *IEEE Inf. Theory Workshop*.

[10] Barbier, J. and F. Krzakala. 2014. "Replica Analysis and Approximate Message Passing Decoder for Sparse Superposition Codes". In: *Proc. IEEE Int. Symp. Inf. Theory*.

[11] Barbier, J. and F. Krzakala. 2017. "Approximate message passing decoder and capacity-achieving sparse superposition codes". *IEEE Trans. Inf. Theory*. 63(8): 4894–4927.

[12] Barbier, J., C. Schülke, and F. Krzakala. 2015a. "Approximate message-passing with spatially coupled structured operators, with applications to compressed sensing and sparse superposition codes". *Journal of Statistical Mechanics: Theory and Experiment*. (5).

[13] Barbier, J., C. Schülke, and F. Krzakala. 2015b. "Approximate message-passing with spatially coupled structured operators, with applications to compressed sensing and sparse superposition codes". *Journal of Statistical Mechanics: Theory and Experiment*. 2015(5): P05013.

[14] Barron, A. and A. Joseph. 2010. "Least Squares Superposition Codes of Moderate Dictionary Size, Reliable at Rates up to Capacity". Arxiv:1712.06866.

[15] Barron, A. and A. Joseph. 2012. "Least Squares Superposition Codes of Moderate Dictionary Size Are Reliable at Rates up to Capacity". *IEEE Trans. Inf. Theory*. 58(5): 2541–2557.

[16] Barron, A. R. and S. Cho. 2012. "High-rate sparse superposition codes with iteratively optimal estimates". In: *Proc. IEEE Int. Symp. Inf. Theory*.

[17] Barron, R., B. Chen, and G. Wornell. 2003. "The duality between information embedding and source coding with side information and some applications". *IEEE Trans. Inf. Theory*. 49(5): 1159–1180.

[18] Bayati, M. and A. Montanari. 2012. "The LASSO Risk for Gaussian Matrices". *IEEE Trans. Inf. Theory*. 58(4): 1997–2017. ISSN: 0018-9448.

[19] Bayati, M. and A. Montanari. 2011. "The Dynamics of Message Passing on Dense Graphs, with Applications to Compressed Sensing". *IEEE Trans. Inf. Theory*: 764–785.

[20] Berrou, C. and A. Glavieux. 1996. "Near optimum error correcting coding and decoding: turbo-codes". *IEEE Trans. Commun.* 44(10): 1261–1271.

[21] Blahut, R. E. 2003. *Algebraic codes for data transmission*. Cambridge University Press.

[22] Böcherer, G., F. Steiner, and P. Schulte. 2015. "Bandwidth efficient and rate-matched low-density parity-check coded modulation". *IEEE Trans. Commun.* 63(12): 4651–4665.

[23] Chen, B. and G. Wornell. 2001. "Quantization index modulation: a class of provably good methods for digital watermarking and information embedding". *IEEE Trans. Inf. Theory*. 47(4): 1423–1443.

[24] Cho, S. 2014. "High-dimensional regression with random design, including sparse superposition codes". *PhD thesis*. Yale University.

[25] Cho, S. and A. Barron. 2013. "Approximate Iterative Bayes Optimal Estimates for High-Rate Sparse Superposition Codes". In: *Sixth Workshop on Information-Theoretic Methods in Science and Engineering*.

[26] Coja-Oghlan, A. and L. Zdeborová. 2012. "The condensation transition in random hypergraph 2-coloring". In: *Proc. 23rd Annual ACM-SIAM Symp. on Discrete Algorithms (SODA)*. 241–250.

[27] Costa, M. 1983. "Writing on dirty paper (Corresp.)" *IEEE Trans. Inf. Theory*. 29(3): 439–441.

[28] Costello, D. J. and G. D. Forney. 2007. "Channel coding: The road to channel capacity". *Proc. IEEE*. 95(6): 1150–1177.

[29] Cover, T. 1972. "Broadcast channels". *IEEE Trans. Inf. Theory*. 18(1): 2–14.

[30] Cover, T. M. and J. A. Thomas. 2012. *Elements of Information Theory*. John Wiley and Sons.

[31] David, H. and H. Nagaraja. 2003. *Order Statistics*. John Wiley & Sons.

[32] Den Hollander, F. 2008. *Large deviations*. Vol. 14. Amer. Mathematical Society.

[33] Donoho, D. L., A. Javanmard, and A. Montanari. 2013. "Information-Theoretically Optimal Compressed Sensing via Spatial Coupling and Approximate Message Passing". *IEEE Trans. Inf. Theory*. (11): 7434–7464.

[34] Donoho, D. L., A. Maleki, and A. Montanari. 2009. "Message-passing algorithms for compressed sensing". *Proceedings of the National Academy of Sciences*. 106(45): 18914–18919.

[35] El Gamal, A. and Y.-H. Kim. 2011. *Network Information Theory*. Cambridge University Press.

[36] Equitz, W. and T. Cover. 1991. "Successive refinement of information". *IEEE Trans. Inf. Theory*. 37(2): 269–275.

[37] Erez, U., S. Shamai, and R. Zamir. 2005a. "Capacity and lattice strategies for canceling known interference". *IEEE Trans. Inf. Theory*. 51(11): 3820–3833.

[38] Erez, U. and S. ten Brink. 2005. "A close-to-capacity dirty paper coding scheme". *IEEE Trans. Inf. Theory*. 51(10): 3417–3432.

[39] Erez, U., S. Litsyn, and R. Zamir. 2005b. "Lattices which are good for (almost) everything". *IEEE Trans. Inf. Theory*. 51(10): 3401–3416.

[40] Erez, U. and R. Zamir. 2004. "Achieving $\frac{1}{2}\log(1+\mathsf{snr})$ on the AWGN channel with lattice encoding and decoding". *IEEE Trans. Inf. Theory*. 50(10): 2293–2314.

[41] Felstrom, A. J. and K. S. Zigangirov. 1999. "Time-varying periodic convolutional codes with low-density parity-check matrix". *IEEE Trans. Inf. Theory*. 45(6): 2181–2191.

[42] Fengler, A., P. Jung, and G. Caire. 2019. "SPARCs for Unsourced Random Access". arXiv:1901.06234.

[43] Forney, G. D. and G. Ungerboeck. 1998. "Modulation and coding for linear Gaussian channels". *IEEE Trans. Inf. Theory*. 44(6): 2384–2415.

[44] Gallager, R. G. 1968. *Information theory and reliable communication*. Vol. 2. Springer.

[45] Gelfand, S. and M. Pinsker. 1980. "Coding for a channel with random parameters". *Problems of Control and Information*. 9(Jan.): 19–31.

[46] Girod, B., A. M. Aaron, S. Rane, and D. Rebollo-Monedero. 2005. "Distributed Video Coding". *Proceedings of the IEEE*. 93(1): 71–83.

[47] Gish, H. and J. Pierce. 1968. "Asymptotically efficient quantizing". *IEEE Trans. Inf. Theory*. 14(5): 676–683.

[48] Goldsmith, A. 2005. *Wireless communications*. Cambridge University Press.

[49] Greig, A. and R. Venkataramanan. 2018. "Techniques for improving the finite length performance of sparse superposition codes". *IEEE Transactions on Communications*. 66(3): 905–917.

[50] Guillén i Fàbregas, A., A. Martinez, and G. Caire. 2008. *Bit-interleaved coded modulation*. Now Publishers Inc.

[51] Gupta, A. and S. Verdù. 2009. "Nonlinear Sparse-Graph Codes for Lossy Compression". *IEEE Trans. Inf. Theory*. 55(5): 1961–1975.

[52] Gupta, A., S. Verdù, and T. Weissman. 2008. "Rate-distortion in near-linear time". In: *Proc. IEEE Int. Symp. on Inf. Theory*.

[53] Guruswami, V. and P. Xia. 2015. "Polar codes: Speed of polarization and polynomial gap to capacity". *IEEE Trans. Inf. Theory*. 61(1): 3–16.

[54] Hall, P. 1979. "On the rate of convergence of normal extremes". *Journal of Applied Probability*. 16(2): 433–439.

[55] Hamkins, J. and K. Zeger. 2002. "Gaussian source coding with spherical codes". *IEEE Trans. Inf. Theory,* 48(11): 2980–2989.

[56] Herzet, C., A. Drémeau, and C. Soussen. 2016. "Relaxed recovery conditions for OMP/OLS by exploiting both coherence and decay". *IEEE Trans. Inf. Theory*. 62(1): 459–470.

[57] Hsieh, K., C. Rush, and R. Venkataramanan. 2018. "Spatially coupled sparse regression codes: Design and state evolution analysis". In: *Proc. IEEE Int. Symp. Inf. Theory*. IEEE.

[58] Ihara, S. and M. Kubo. 2000. "Error exponent for coding of memoryless Gaussian sources with a fidelity criterion". *IEICE Trans. Fundamentals*. E83-A(10).

[59] Ingber, A. and Y. Kochman. 2011. "The Dispersion of Lossy Source Coding". In: *Data Compression Conference (DCC)*. 53–62.

[60] Jalali, S. and T. Weissman. 2010. "Rate-Distortion via Markov Chain Monte Carlo". In: *Proc. IEEE Int. Symp. on Inf. Theory*.

[61] Janson, S. 2000. *Random Graphs*. Wiley.

[62] Joseph, A. 2012. "Achieving information-theoretic limits with high-dimensional regression". *PhD thesis*. Yale University.

[63] Joseph, A. and A. R. Barron. 2014. "Fast Sparse Superposition Codes Have Near Exponential Error Probability for $R < \mathcal{C}$". *IEEE Trans. Inf. Theory*. 60(2): 919–942.

[64] Kontoyiannis, I. 1999. "An implementable lossy version of the Lempel-Ziv algorithm-I: Optimality for memoryless sources". *IEEE Trans. Inf. Theory*. 45(7): 2293–2305.

[65] Kontoyiannis, I. and C. Gioran. 2009. "Efficient random codebooks and databases for lossy compression in near-linear time". In: *IEEE Inf. Theory Workshop*. 236–240.

[66] Kontoyiannis, I., K. Rad, and S. Gitzenis. 2010. "Sparse superposition codes for Gaussian vector quantization". In: *IEEE Inf. Theory Workshop*. 1.

[67] Korada, S. and R. Urbanke. 2010a. "Polar codes for Slepian-Wolf, Wyner-Ziv, and Gelfand-Pinsker". In: *IEEE Inf. Theory Workshop*.

[68] Korada, S. and R. Urbanke. 2010b. "Polar Codes are Optimal for Lossy Source Coding". *IEEE Trans. Inf. Theory*. 56(4): 1751–1768.

[69] Kostina, V. and S. Verdú. 2012. "Fixed-Length Lossy Compression in the Finite Blocklength Regime". *IEEE Trans. on Inf. Theory*. 58(6): 3309–3338.

[70] Krzakala, F., M. Mézard, F. Sausset, Y. Sun, and L. Zdeborová. 2012. "Probabilistic reconstruction in compressed sensing: algorithms, phase diagrams, and threshold achieving matrices". *Journal of Statistical Mechanics: Theory and Experiment*. (8).

[71] Kudekar, S. and H. D. Pfister. 2010. "The effect of spatial coupling on compressive sensing". In: *Proc. 48th Annual Allerton Conference on Communication, Control, and Computing*. 347–353.

[72] Kudekar, S., T. J. Richardson, and R. L. Urbanke. 2011. "Threshold saturation via spatial coupling: Why convolutional LDPC ensembles perform so well over the BEC". *IEEE Trans. Inf. Theory*. 57(2): 803–834.

[73] Kudekar, S., T. Richardson, and R. L. Urbanke. 2013. "Spatially Coupled Ensembles Universally Achieve Capacity under Belief Propagation". *IEEE Trans. Inf. Theory*. 59(12): 7761–7813.

[74] Kumar, S., A. J. Young, N. Macris, and H. D. Pfister. 2014. "Threshold saturation for spatially coupled LDPC and LDGM codes on BMS channels". *IEEE Trans. Inf. Theory*. 60(12): 7389–7415.

[75] Lapidoth, A. 1997. "On the role of mismatch in rate distortion theory". *IEEE Trans. Inf. Theory*. 43(1): 38–47.

[76] Lentmaier, M., A. Sridharan, D. J. Costello, and K. S. Zigangirov. 2010. "Iterative decoding threshold analysis for LDPC convolutional codes". *IEEE Trans. Inf. Thy*. 56(10): 5274–5289.

[77] Liang, S., J. Ma, and L. Ping. 2017. "Clipping Can Improve the Performance of Spatially Coupled Sparse Superposition Codes". *IEEE Commun. Letters*. 21(12): 2578–2581.

[78] Lin, S. and D. J. Costello. 2004. *Error control coding*. Vol. 2. Prentice Hall Englewood Cliffs.

[79] Marton, K. 1974. "Error exponent for source coding with a fidelity criterion". *IEEE Trans. Inf. Theory*. 20(2): 197–199.

[80] Mitchell, D. G., M. Lentmaier, and D. J. Costello. 2015. "Spatially coupled LDPC codes constructed from protographs". *IEEE Trans. Inf. Theory*. 61(9): 4866–4889.

[81] Montanari, A. 2012. "Graphical models concepts in compressed sensing". In: *Compressed Sensing*. Ed. by Y. C. Eldar and G. Kutyniok. Cambridge University Press. 394–438.

[82] Moulin, P. and R. Koetter. 2005. "Data-Hiding Codes". *Proc. IEEE*. 93(12): 2083–2126.

[83] No, A. and T. Weissman. 2016. "Rateless lossy compression via the extremes". *IEEE Trans. Inf. Theory*. 62(10): 5484–5495.

[84] O'Sullivan, J., P. Moulin, and J. Ettinger. 1998. "Information theoretic analysis of steganography". In: *IEEE Int. Symp. on Inf. Theory*.

[85] Polyanskiy, Y., H. V. Poor, and S. Verdú. 2010. "Channel coding rate in the finite blocklength regime". *IEEE Trans. Inf. Theory*. 56(5): 2307–2359.

[86] Pradhan, S., J. Chou, and K. Ramchandran. 2003. "Duality between source coding and channel coding and its extension to the side information case". *IEEE Trans. Inf. Theory*. 49(5): 1181–1203.

[87] Pradhan, S. and K. Ramchandran. 2003. "Distributed source coding using syndromes (DISCUS): design and construction". *IEEE Trans. Inf. Theory*. 49(3): 626–643.

[88] Puri, R., A. Majumdar, and K. Ramchandran. 2007. "PRISM: A Video Coding Paradigm With Motion Estimation at the Decoder". *IEEE Trans. Image Process.* 16(10): 2436–2448.

[89] *Python script for SPARC with AMP decoding*. URL: https://github.com/sigproc/sparc-amp/blob/master/sparc_amp.ipynb (accessed on 04/26/2017).

[90] Rangan, S. 2011. "Generalized approximate message passing for estimation with random linear mixing". In: *Proc. IEEE Int. Symp. Inf. Theory*. 2168–2172.

[91] Rebollo-Monedero, D., R. Zhang, and B. Girod. 2003. "Design of optimal quantizers for distributed source coding". In: *Data Compression Conference*. 13–22.

[92] Richardson, T. and R. Urbanke. 2008. *Modern Coding Theory*. Cambridge University Press.

[93] Rimoldi, B. 1994. "Successive refinement of information: characterization of the achievable rates". *IEEE Trans. Inf. Theory*. 40(1): 253–259.

[94] Rush, C., A. Greig, and R. Venkataramanan. 2017. "Capacity-achieving Sparse Superposition Codes via Approximate Message Passing Decoding". *IEEE Trans. Inf. Theory*. 63(3): 1476–1500.

[95] Rush, C., K. Hsieh, and R. Venkataramanan. 2018. “Capacity-achieving sparse regression codes via spatial coupling”. In: *Proc. IEEE Inf. Theory Workshop*.

[96] Rush, C. and R. Venkataramanan. 2017. “The Error Probability of Sparse Superposition Codes with Approximate Message Passing Decoding”. Arxiv:1712.06866.

[97] Sakrison, D. 1968. “A geometric treatment of the source encoding of a Gaussian random variable”. *IEEE Trans. Inf. Theory*. 14(3): 481–486.

[98] Sakrison, D. 1969. “The rate distortion function for a class of sources”. *Information and Control*. 15(2): 165–195.

[99] Sakrison, D. 1970. “The rate of a class of random processes”. *IEEE Trans. Inf. Theory*. 16(1): 10–16.

[100] Shanks, J. L. 1969. “Computation of the Fast Walsh-Fourier Transform”. *IEEE Trans. Comput.* 18(5): 457–459.

[101] Sommer, N., M. Feder, and O. Shalvi. 2008. “Low-Density Lattice Codes”. *IEEE Trans. on Inf. Theory*. 54(4): 1561–1585.

[102] Spencer, Q., C. Peel, A. Swindlehurst, and M. Haardt. 2004. “An introduction to the multi-user MIMO downlink”. *IEEE Commun. Mag.* 42(10): 60–67.

[103] Sun, Y., Y. Yang, A. Liveris, V. Stankovic, and Z. Xiong. 2009. “Near-Capacity Dirty-Paper Code Design: A Source-Channel Coding Approach”. *IEEE Trans. Inf. Theory*. 55(7): 3013–3031.

[104] Takeishi, Y., M. Kawakita, and J. Takeuchi. 2014. “Least squares superposition codes with Bernoulli dictionary are still reliable at rates up to capacity”. *IEEE Trans. Inf. Theory*. 60: 2737–2750.

[105] Takeishi, Y. and J. Takeuchi. 2016. “An improved upper bound on block error probability of least squares superposition codes with unbiased Bernoulli dictionary”. In: *IEEE Int. Symp. on Inf. Theory*. 1168–1172.

[106] ten Brink, S. 1999. “Convergence of iterative decoding”. *Electronics letters*. 35(13): 1117–1119.

[107] ten Brink, S., G. Kramer, and A. Ashikhmin. 2004. “Design of low-density parity-check codes for modulation and detection”. *IEEE Trans. Commun.* 52(4): 670–678.

[108] *The Coded Modulation Library*. 2008. URL: http://www.iterativesolutions.com/Matlab.htm (accessed on 01/16/2017).

[109] Tibshirani, R. 1996. "Regression shrinkage and selection via the LASSO". *Journal of the Royal Statistical Society. Series B (Methodological)*: 267–288.

[110] Tibshirani, R., M. Wainwright, and T. Hastie. 2015. *Statistical learning with sparsity: the LASSO and generalizations*. Chapman and Hall/CRC.

[111] Tse, D. and P. Viswanath. 2005. *Fundamentals of wireless communication*. Cambridge University Press.

[112] Venkataramanan, R., A. Joseph, and S. Tatikonda. 2014a. "Lossy Compression via Sparse Linear Regression: Performance under Minimum-distance Encoding". *IEEE Trans. Inf. Thy.* 60(6): 3254–3264.

[113] Venkataramanan, R., T. Sarkar, and S. Tatikonda. 2014b. "Lossy Compression via Sparse Linear Regression: Computationally Efficient Encoding and Decoding". *IEEE Trans. Inf. Theory*. 60(6): 3265–3278.

[114] Venkataramanan, R. and S. Tatikonda. 2012. "Sparse Regression Codes for Multi-terminal Source and Channel Coding". In: *50th Allerton Conf. on Commun., Control, and Computing*.

[115] Venkataramanan, R. and S. Tatikonda. 2017. "The Rate-Distortion Function and Excess-Distortion Exponent of Sparse Regression Codes with Optimal Encoding". *IEEE Trans. Inf. Theory*. 63(8): 5228–5243.

[116] Wainwright, M., E. Maneva, and E. Martinian. 2010. "Lossy Source Compression Using Low-Density Generator Matrix Codes: Analysis and Algorithms". *IEEE Trans. Inf. Theory*. 56(3): 1351–1368.

[117] Wyner, A. D. and J. Ziv. 1976. "The rate-distortion function for source coding with side information at the decoder". *IEEE Trans. Inf. Theory*. 22(Jan.): 1–10.

[118] Yan, Y., L. Liu, C. Ling, and X. Wu. 2014. "Construction of capacity-achieving lattice codes: Polar lattices". arXiv:1411.0187.

[119] Yang, Y., S. Cheng, Z. Xiong, and W. Zhao. 2009. "Wyner-Ziv coding based on TCQ and LDPC codes". *IEEE Trans. Commun.* 57(2): 376–387.

[120] Yedla, A., Y.-Y. Jian, P. S. Nguyen, and H. D. Pfister. 2014. "A simple proof of Maxwell saturation for coupled scalar recursions". *IEEE Trans. Inf. Theory*. 60(11): 6943–6965.

[121] Zamir, R., S. Shamai, and U. Erez. 2002. "Nested linear/lattice codes for structured multiterminal binning". *IEEE Trans. Inf. Theory*. 48(6): 1250–1276.

[122] Zamir, R. 2014. *Lattice Coding for Signals and Networks: A Structured Coding Approach to Quantization, Modulation, and Multiuser Information Theory*. Cambridge University Press.