Statistical Physics and Information Theory

Statistical Physics and Information Theory

Neri Merhav

Department of Electrical Engineering
Technion — Israel Institute of Technology
Haifa 32000
Israel
merhav@ee.technion.ac.il

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
USA
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 N. Merhav, Statistical Physics and Information Theory, Foundations and Trends® in Communications and Information Theory, vol 6, nos 1–2, pp 1–212, 2009

ISBN: 978-1-60198-406-7

Foundations and Trends® in Communications and Information Theory

Volume 6 Issues 1–2, 2009

Editorial Scope

Foundations and Trends® in Communications and Information Theory will publish 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, 2009, Volume 6, 6 issues. ISSN paper version 1567-2190. ISSN online version 1567-2328. Also available as a combined paper and online subscription.

Foundations and Trends® in
Communications and Information Theory
Vol. 6, Nos. 1–2 (2009) 1–212

DOI: 10.1561/0100000052

Statistical Physics and Information Theory

Neri Merhav

Department of Electrical Engineering, Technion — Israel Institute of Technology, Haifa 32000, Israel, merhav@ee.technion.ac.il

Abstract

This monograph is based on lecture notes of a graduate course, which focuses on the relations between information theory and statistical physics. The course was delivered at the Technion during the Spring of 2010 for the first time, and its target audience consists of EE graduate students in the area of communications and information theory, as well as graduate students in Physics who have basic background in information theory. Strong emphasis is given to the analogy and parallelism between information theory and statistical physics, as well as to the insights, the analysis tools and techniques that can be borrowed from statistical physics and 'imported' to certain problem areas in information theory. This is a research trend that has been very active in the last few decades, and the hope is that by exposing the students to the meeting points between these two disciplines, their background and perspective may be expanded and enhanced. This monograph is substantially revised and expanded relative to an earlier version posted in arXiv (1006.1565v1 [cs.iT]).

Contents

1

Introduction

This work focuses on some of the relationships and the interplay between information theory and statistical physics — a branch of physics that deals with many-particle systems using probabilistic and statistical methods in the microscopic level.

The relationships between information theory and statistical thermodynamics are by no means new, and many researchers have been exploiting them for many years. Perhaps the first relation, or analogy, that crosses one's mind is that in both fields there is a fundamental notion of *entropy*. Actually, in information theory, the term entropy was coined in the footsteps of the thermodynamic entropy. The thermodynamic entropy was first introduced by Clausius in 1850, and its probabilistic-statistical interpretation was established by Boltzmann in 1872. It is virtually impossible to miss the functional resemblance between the two notions of entropy, and indeed it was recognized by Shannon and von Neumann. The well-known anecdote on this tells that von Neumann advised Shannon to adopt this term because it would provide him with "... *a great edge in debates because nobody really knows what entropy is anyway.*"

But the relationships between the two fields go far beyond the fact that both share the notion of entropy. In fact, these relationships have many aspects. We will not cover all of them in this work, but just to taste the flavor of their scope, we will mention just a few.

The maximum entropy (ME) principle. This is perhaps the oldest concept that ties the two fields and it has attracted a great deal of attention, not only of information theorists, but also that of researchers in related fields like signal processing and image processing. The ME principle evolves around a philosophy, or a belief, which, in a nutshell, is the following: if in a certain problem, the observed data comes from an unknown probability distribution, but we do have some knowledge (that stems, e.g., from measurements) of certain moments of the underlying quantity/signal/random-variable, then assume that the unknown underlying probability distribution is the one with *maximum entropy* subject to (s.t.) moment constraints corresponding to this knowledge. For example, if we know the first and the second moments, then the ME distribution is Gaussian with matching first and second order moments. Indeed, the Gaussian model is perhaps the most common model for physical processes in information theory as well as in signal- and image processing. But why maximum entropy? The answer to this philosophical question is rooted in the *second law of thermodynamics*, which asserts that in an isolated system, the entropy cannot decrease, and hence, when the system reaches thermal equilibrium, its entropy reaches its maximum. Of course, when it comes to problems in information theory and other related fields, this principle becomes quite heuristic, and so, one may question its justification, but nevertheless, this approach has had an enormous impact on research trends throughout the last 50 years, after being proposed by Jaynes in the late fifties of the previous century [45, 46], and further advocated by Shore and Johnson afterward [106]. In the book by Cover and Thomas [13, Section 12], there is a good exposition on this topic. We will not put much emphasis on the ME principle in this work.

Landauer's erasure principle. Another aspect of these relations has to do with a theory whose underlying guiding principle is that information is a physical entity. Specifically, Landauer's erasure principle [63] (see also [6]), which is based on this physical theory of information, asserts that every bit that one erases, increases the entropy of the universe by $k\ln 2$, where k is Boltzmann's constant.

The more comprehensive picture behind Landauer's principle is that "any logically irreversible manipulation of information, such as the erasure of a bit or the merging of two computation paths, must be accompanied by a corresponding entropy increase in non-information bearing degrees of freedom of the information processing apparatus or its environment." (see [6]). This means that each lost information bit leads to the release of an amount $kT\ln 2$ of heat. By contrast, if no information is erased, computation may, in principle, be achieved in a way which is thermodynamically a reversible process, and hence requires no release of heat. This has had a considerable impact on the study of reversible computing. Landauer's principle is commonly accepted as a law of physics. However, there has also been some considerable dispute among physicists on this. This topic is not going to be included either in this work.

Large deviations theory as a bridge between information theory and statistical physics. Both information theory and statistical physics have an intimate relation to large deviations theory, a branch of probability theory which focuses on the assessment of the exponential rates of decay of probabilities of rare events, where one of the most elementary mathematical tools is the Legendre transform, which stands at the basis of the Chernoff bound. This topic will be covered quite thoroughly, mostly in Section 3.2.

Random matrix theory. How do the eigenvalues (or, more generally, the singular values) of random matrices behave when these matrices have very large dimensions or if they result from products of many randomly selected matrices? This is a very active area in probability theory with many applications, both in statistical physics and information theory, especially in modern theories of wireless communication (e.g., MIMO systems). This is again outside the scope of this work, but the interested reader is referred to [115] for a comprehensive introduction on the subject.

Spin glasses and coding theory. As was first observed by Sourlas [109] (see also [110]) and further advocated by many others, it turns out that many problems in channel coding theory (and also to some extent, source coding theory) can be mapped almost verbatim to parallel problems in the field of physics of *spin glasses* — amorphic magnetic materials with a high degree of disorder and very complicated physical behavior, which is customarily treated using statistical-mechanical approaches. It has been many years that researchers have made attempts to "import" analysis techniques rooted in statistical physics of spin glasses and to apply them to analogous coding problems, with various degrees of success. This is one of main subjects of this work and we will study it extensively, at least from some aspects.

The above list of examples is by no means exhaustive. We could have gone much further and add many more examples of these very fascinating meeting points between information theory and statistical physics, but most of them will not be touched upon in this work. Many modern analyses concerning multiuser situations, such as MIMO channels, CDMA, etc., and more recently, also in compressed sensing, are based on statistical-mechanical techniques. But even if we limit ourselves to single-user communication systems, yet another very active problem area under this category is that of codes on graphs, iterative decoding, belief propagation, and density evolution. The main reason for not including it in this work is that it is already very well covered in recent textbooks, such as the one Mézard and Montanari [80] as well as the one by Richardson and Urbanke [98]. Another comprehensive exposition of graphical models, with a fairly strong statistical-mechanical flavor, was written by Wainwright and Jordan [118].

As will be seen, the physics and the information-theoretic subjects are interlaced with each other, rather than being given in two continuous, separate parts. This way, it is hoped that the relations between information theory and statistical physics will be made more apparent. We shall see that, not only these relations between information theory and statistical physics are interesting academically on their own right, but, moreover, they also prove useful and beneficial in that they provide us with new insights and mathematical tools to deal with information-theoretic problems. These mathematical tools sometimes prove a lot more efficient than traditional tools used in information theory, and they may give either simpler expressions for performance analysis, or improved bounds, or both.

Having said that, a certain digression is in order. The reader should not expect to see too many real breakthroughs, which are allowed exclusively by statistical-mechanical methods, but could not have been achieved otherwise. Perhaps one exception to this rule is the replica method of statistical mechanics, which will be reviewed in this work, but not in great depth, because of two reasons: first, it is not rigorous (and so, any comparison to rigorous information-theoretic methods would not be fair), and secondly, because it is already very well covered in existing textbooks, such as [80] and [87]. If one cares about rigor, however, then there are no miracles. Everything, at the end of the day, boils down to mathematics. The point then is which culture, or scientific community, has developed the suitable mathematical techniques and what are the new insights that they provide; in many cases, it is the community of statistical physicists.

There are several examples of such techniques and insights, which are emphasized rather strongly in this work. One example is the use of integrals in the complex plane and the saddle-point method. Among other things, this should be considered as a good substitute to the method of types, with the bonus of lending itself to extensions that include the countable and the continuous alphabet case (rather than just the finite alphabet case). Another example is the analysis technique of error exponents, which stems from the random energy model (see Section 6 and onward), along with its insights about phase transitions. Again, in retrospect, these analyses are just mathematics and therefore could have been carried out without relying on any knowledge in physics. But it is nevertheless the physical point of view that provides the trigger for its use. Moreover, there are situations (see, e.g., Section 7.3), where results from statistical mechanics can be used almost verbatim in order to obtain stronger coding theorems. The point is then that it is not the physics itself that may be useful, it is the way in which physicists use mathematical tools.

One of the main take-home messages, that will hopefully remain with the reader after reading this work, is that whatever the field of statistical mechanics has to offer to us, as information theorists, goes much beyond the replica method. It is believed that this message is timely, because the vast majority of papers at the interface between the two disciplines are about applying the replica method to some information-theoretic problem.

The outline of the remaining part of this work is as follows: in Section 2, we give some elementary background in statistical physics and we relate fundamental thermodynamic potentials, like thermodynamical entropy and free energy with fundamental information measures, like the Shannon entropy and the Kullback–Leibler divergence. In Section 3, we explore a few aspects of physical interpretations of some fundamental results in information theory, like non–negativity of the Kullback–Leibler divergence, the data processing inequality, and the elementary coding theorems of information theory. In Section 4, we review some analysis tools commonly used in statistical physics, like the Laplace integration method, the saddle-point method, and the replica method, all accompanied by examples. Section 5 is devoted to a (mostly descriptive) exposition of systems with interacting particles and phase transitions, both in physics and information theory. Section 6 focuses on one particular model of a disordered physical system with interacting particles — the random energy model, which is highly relevant to the analysis of random code ensembles. Section 7 extends the random energy model in several directions, all relevant to problems in information theory. Finally, Section 8 contains a summary and an outlook on the interplay between information theory and statistical mechanics.

As with every paper published in *Foundations and Trends in Communications and Information Theory*, the reader is, of course, assumed to have some solid background in information theory. Concerning the physics part, prior background in statistical mechanics does not harm, but is not necessary. This work is intended to be self-contained as far as the physics background goes.

In a closing note, it is emphasized again that the coverage of topics, in this work, is by no means intended to be fully comprehensive, nor is it aimed at providing the complete plethora of problem areas, methods and results. The choice of topics, the approach, the flavor, and the style are nothing but the mirror image of the author's personal bias, perspective, and research interests in the field. Therefore, this work should actually be viewed mostly as a monograph, and not quite as a review or a tutorial paper. This is also the reason that a considerable part of the topics, covered in this work, is taken from articles in which the author has been involved.

References

[1] L.-P. Arguin, "Spin glass computations and Ruelle's probability cascades," arXiv:math-ph/0608045v1, August 17 2006.

[2] A. Barg and G. D. Forney, Jr., "Random codes: minimum distances and error exponents," *IEEE Transactions on Information Theory*, vol. 48, no. 9, pp. 2568–2573, September 2002.

[3] G. B. Bağci, "The physical meaning of Rényi relative entropies," arXiv:cond-mat/0703008v1, March 1 2007.

[4] R. J. Baxter, *Exactly Solved Models in Statistical Mechanics*. Academic Press, 1982.

[5] A. H. W. Beck, *Statistical Mechanics, Fluctuations and Noise*. Edward Arnold Publishers, 1976.

[6] C. H. Bennett, "Notes on Landauer's principle, reversible computation and Maxwell's demon," *Studies in History and Philosophy of Modern Physics*, vol. 34, pp. 501–510, 2003.

[7] G. P. Beretta and E. Zanchini, "Rigorous and general definition of thermodynamic entropy," arXiv:1010.0813v1 [quant-ph], October 5 2010.

[8] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

[9] E. Buffet, A. Patrick, and J. V. Pulé, "Directed polymers on trees: A martingale approach," *Journal of Physics A: Mathematical and General*, vol. 26, pp. 1823–1834, 1993.

[10] F. Cousseau, K. Mimura, and M. Okada, "Statistical mechanics of lossy compression for non-monotonic multilayer perceptron," in *Proceedings of ISIT 2008*, pp. 509–513, Toronto, Canada, July 2008.

[11] F. Cousseau, K. Mimura, T. Omori, and M. Okada, "Statistical mechanics of lossy compression for non-monotonic multilayer perceptrons," *Physical Review E*, vol. 78, 021124 2008, arXiv:0807.4009v1 [cond-mat.stat-mech], 25 July 2008.

[12] T. M. Cover and E. Ordentlich, "Universal portfolios with side information," *IEEE Transactions on Information Theory*, vol. IT-42, no. 2, pp. 348–363, March 1996.

[13] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. John Wiley & Sons, Second ed., 2006.

[14] G. E. Crooks, "Beyond Boltzmann–Gibbs statistics: Maximum entropy hyperensembles out of equilibrium," *Physical Review E*, vol. 75, 041119, 2007.

[15] I. Csiszár, "A class of measures of informativity of observation channels," *Periodica Mathematica Hungarica*, vol. 22, no. 1–4, pp. 191–213, 1972.

[16] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Earlier Memoryless Systems*. New York: Academic Press, 1981.

[17] B. Dacorogna and P. Maréchal, "The role of perspective functions in convexity, polyconvexity, rank-one convexity and separate convexity," http://caa.epfl.ch/publications/2008-The_role_of_perspective_functions_in_convexity.pdf.

[18] W.-S. Dai and M. Xie, "The explicit expression of the fugacity for hard-sphere Bose and Fermi gases," arXiv:0906.0952v1 [cond-mat.stat-mech], 4 June 2009.

[19] C. R. Davis and M. E. Hellman, "On tree coding with a fidelity criterion," *IEEE Transactions on Information Theory*, vol. IT-21, no. 4, pp. 373–378, July 1975.

[20] N. G. de Bruijn, *Asymptotic Methods in Analysis*. Dover Publications, 1981.

[21] C. De Dominicis, M. Gabay, T. Garel, and H. Orland, "White and weighted averages over solutions of Thouless Anderson Palmer equations for the Sherrington Kirkpatrick spin glass," *Journal de Physique*, vol. 41, pp. 923–930, 1980.

[22] C. R. de Oliveira and T. Werlang, "Ergodic hypothesis in classical statistical mechanics," *Revista Brasileira de Ensino de Fisíca*, vol. 29, no. 2, pp. 189–201, Also available on–line: http://www.sbfisica.org.br/rbef/pdf/060601.pdf, 2007.

[23] B. Derrida, "The random energy model," *Physics Reports (Review Section of Physics Letters)*, vol. 67, no. 1, pp. 29–35, 1980.

[24] B. Derrida, "Random-energy model: Limit of a family of disordered models," *Physical Review Letters*, vol. 45, no. 2, pp. 79–82, July 1980.

[25] B. Derrida, "Random-energy model: An exactly solvable model for disordered systems," *Physical Review B*, vol. 24, no. 5, pp. 2613–2626, September 1981.

[26] B. Derrida, "A generalization of the random energy model which includes correlations between energies," *Journal de Physique — Lettres*, vol. 46, pp. L-401–L-407, May 1985.

[27] B. Derrida and E. Gardner, "Solution of the generalised random energy model," *Jounal of Physics C: Solid State Physics*, vol. 19, pp. 2253–2274, 1986.

[28] R. J. Dick, T. Berger, and F. Jelinek, "Tree encoding of Gaussian sources," *IEEE Transformations on Information Theory*, vol. IT-20, no. 3, pp. 332–336, May 1974.

[29] M. Doi and S. F. Edwards, *The Theory of Polymer Dynamics*. Oxford University Press, 1986.

[30] V. Dotsenko, "One more discussion on the replica trick: The examples of exact solutions," arXiv:1010.3913v1 [cond-mat.stat-mech], 19 October 2010.

[31] R. S. Ellis, *Entropy, Large Deviations, and Statistical Mechanics*. Springer-Verlag, NY, 1985.

[32] R. Etkin, N. Merhav, and E. Ordentlich, "Error exponents of optimum decoding for the interference channel," *IEEE Transactions on Information Theory*, vol. 56, no. 1, pp. 40–56, January 2010.

[33] G. D. Forney, Jr., "Exponential error bounds for erasure, list, and decision feedback schemes," *IEEE Transactions on Information Theory*, vol. IT-14, no. 2, pp. 206–220, March 1968.

[34] G. D. Forney, Jr. and A. Montanari, "On exponential error bounds for random codes on the DMC," manuscript, http://www.stanford.edu/∼montanar/PAPERS/FILEPAP/dmc.ps, 2001.

[35] R. G. Gallager, *Information Theory and Reliable Communication*. John Wiley & Sons, 1968.

[36] R. G. Gallager, "Tree encoding for symmetric sources with a distortion measure," *IEEE Transactions on Information Theory*, vol. IT-20, no. 1, pp. 65–76, January 1974.

[37] R. M. Gray, *Source Coding Theory*. Kluwer Academic Publishers, 1990.

[38] D. Guo and S. Verdú, "Randomly spread CDMA: Asymptotics via statistical physics," *IEEE Transactions on Information Theory*, vol. 51, no. 6, pp. 1982–2010, June 2005.

[39] M. J. W. Hall, "Universal geometric approach to uncertainty, entropy, and information," *Physical Review A*, vol. 59, no. 4, pp. 2602–2615, April 1999.

[40] J. Honerkamp, *Statistical Physics — An Advanced Approach with Applications*. Springer–Verlag, 2nd ed., 2002.

[41] W. G. Hoover and C. G. Hoover, "Nonequilibrium temperature and thermometry in heat-conducting ϕ^4 models," *Physical Review E*, vol. 77, 041104, 2008.

[42] T. Hosaka and Y. Kabashima, "Statistical mechanical approach to error exponents of lossy data compression," *Journal of the Physical Society of Japan*, vol. 74, no. 1, pp. 488–497, January 2005.

[43] P. A. Humblet, "Generalization of Huffman coding to minimize the probability of buffer overflow," *IEEE Transactions on Information Theory*, vol. 27, pp. 230–232, March 1981.

[44] C. Jarzynski, "Nonequilibrium equality for free energy differences," *Physical Review Letters*, vol. 78, no. 14, pp. 2690–2693, 7 April 1997.

[45] E. T. Jaynes, "Information theory and statistical mechanics," *Physical Review A*, vol. 106, pp. 620–630, May 1957.

[46] E. T. Jaynes, "Information theory and statistical mechanics-II," *Physical Review A*, vol. 108, pp. 171–190, October 1957.

[47] F. Jelinek, "Buffer overflow in variable length coding of fixed rate sources," *IEEE Transactions on Information Theory*, vol. IT-14, no. 3, pp. 490–501, May 1968.

[48] F. Jelinek, "Tree encoding of memoryless time-discrete sources with a fidelity criterion," *IEEE Transactions on Information Theory*, vol. IT-15, no. 5, pp. 584–590, September 1969.

[49] F. Jelinek and J. B. Anderson, "Instrumentable tree encoding of information sources," *IEEE Transactions on Information Theory*, vol. IT-17, no. 1, pp. 118–119, January 1971.

[50] Y. Kabashima, "How could the replica method improve accuracy of performance assessment of channel coding?," in *Proceedings of the International Workshop on Statistical–Mechanical Informatics*, Sendai, Japan. arXiv:0808.0548v1 [cs.IT], September 14–17 2008. 5 August 2008.

[51] Y. Kabashima and T. Hosaka, "Statistical mechanics for source coding with a fidelity criterion," *Progress of Theoretical Physics*, no. 157, pp. 197–204, Supplement, 2005.

[52] Y. Kabashima, K. Nakamura, and J. van Mourik, "Statistical mechanics of typical set decoding," *Physical Review E*, vol. 66, 2002.

[53] Y. Kabashima and D. Saad, "Statistical mechanics of error correcting codes," *Europhysics Letters*, vol. 45, no. 1, pp. 97–103, 1999.

[54] Y. Kabashima and D. Saad, "Statistical mechanics of low-density parity check codes," *Journal of Physics A: Mathamatical and General*, vol. 37, pp. R1–R43, 2004.

[55] M. Kardar, *Statistical Physics of Particles*. Cambridge University Press, 2007.

[56] Y. Kaspi and N. Merhav, "Error exponents of optimum decoding for the degraded broadcast channel using moments of type class enumerators," in *Proceedings of the ISIT 2009*, pp. 2507–2511, Seoul, South Korea, June–July 2009. Full version: available in arXiv:0906.1339.

[57] R. Kawai, J. M. R. Parrondo, and C. V. den Broeck, "Dissipation: The phase-space perspective," *Physical Review Letters*, vol. 98, 080602, 2007.

[58] F. P. Kelly, *Reversibility and Stochastic Networks*. John Wiley & Sons, 1979.

[59] K. Kitagawa and T. Tanaka, "Optimal spreading sequences in large CDMA systems: a Proceedings of ISIT 2008," pp. 1373–1377, Toronto, Canada, July 2008.

[60] C. Kittel, *Elementary Statistical Physics*. John Wiley & Sons, 1958.

[61] R. Kubo, *Statistical Mechanics*. North-Holland, 1961.

[62] L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics — vol. 5: Statistical Physics, Part 1*. Elsevier, 3rd ed., 1980.

[63] R. Landauer, "Irreversibility and heat generation in the computing process," *IBM Journal of Research and Development*, vol. 5, pp. 183–191, 1961.

[64] T. D. Lee and C. N. Yang, "Statistical theory of equations of state and phase transitions. II. Lattice gas and Ising model," *Physical Review*, vol. 87, no. 3, pp. 410–419, August 1952.

[65] H. Löwen, "Fun with hard spheres," in *Spatial Statistics and Statistical Physics*, vol. 554, (K. Mecke and D. Stoyan, eds.), pp. 295–331, Berlin: Springer Lecture Notes in Physics, 2000.

[66] N. Merhav, "Another look at the physics of large deviations with application to rate–distortion theory," http://arxiv.org/PS_cache/arxiv/pdf/0908/0908.3562v1.pdf.

[67] N. Merhav, "Universal coding with minimum probability of code word length overflow," *IEEE Transactions on Information Theory*, vol. 37, no. 3, pp. 556–563, May 1991.

[68] N. Merhav, "Error exponents of erasure/list decoding revisited via moments of distance enumerators," *IEEE Transactions on Information Theory*, vol. 54, no. 10, pp. 4439–4447, October 2008.

[69] N. Merhav, "An identity of Chernoff bounds with an interpretation in statistical physics and applications in information theory," *IEEE Transactions on Information Theory*, vol. 54, no. 8, pp. 3710–3721, August 2008.

[70] N. Merhav, "The random energy model in a magnetic field and joint source–channel coding," *Physica A: Statistical Mechanics and Its Applications*, vol. 387, no. 22, pp. 5662–5674, September 15 2008.

[71] N. Merhav, "The generalized random energy model and its application to the statistical physics of ensembles of hierarchical codes," *IEEE Transactions on Information Theory*, vol. 55, no. 3, pp. 1250–1268, March 2009.

[72] N. Merhav, "Relations between random coding exponents and the statistical physics of random codes," *IEEE Transactions on Information Theory*, vol. 55, no. 1, pp. 83–92, January 2009.

[73] N. Merhav, "Data processing theorems and the second law of thermodynamics," submitted to *IEEE Transactions on Information Theory*, 2010.

[74] N. Merhav, "On the physics of rate–distortion theory," in *Proceedings of ISIT 2010*, pp. 71–75, Austin, Texas, U.S.A., June 2010.

[75] N. Merhav, "On the statistical physics of directed polymers in a random medium and their relation to tree codes," *IEEE Transactions on Information Theory*, vol. 56, no. 3, pp. 1345–1350, March 2010.

[76] N. Merhav, "Physics of the Shannon limits," *IEEE Transactions on Information Theory*, vol. 56, no. 9, pp. 4274–4285, September 2010.

[77] N. Merhav, "Rate–distortion function via minimum mean square error estimation," accepted *IEEE Transactions on Information Theory*, November 2010.

[78] N. Merhav, "Threshold effects in parameter estimation as phase transitions in statistical mechanics," submitted to *IEEE Transactions on Information Theory*, 2010.

[79] N. Merhav, D. Guo, and S. S. (Shitz), "Statistical physics of signal estimation in Gaussian noise: Theory and examples of phase transitions," *IEEE Transactions on Information Theory*, vol. 56, no. 3, pp. 1400–1416, March 2010.

[80] M. Mézard and A. Montanari, *Information, Physics and Computation*. Oxford University Press, 2009.

[81] A. Montanari, "Turbo codes: The phase transition," *The European Physical Journal B*, vol. 18, p. 121, E-print: cond-mat/0003218, 2000.

[82] A. Montanari, "The glassy phase of Gallager codes," *The European Physical Journal B — Condensed Matter and Complex Systems*, vol. 23, no. 1, pp. 121–136, E-print: cond-mat/0104079, 2001.

[83] T. Mora and O. Rivoire, "Statistical mechanics of error exponents for error-correcting codes," arXiv:cond-mat/0606696, June 2006.

[84] T. Murayama, "Statistical mechanics of the data compression theorem," *Journal of the Physics A: Mathamatical General*, vol. 35, pp. L95–L100, 2002.

[85] K. R. Narayanan and A. R. Srinivasa, "On the thermodynamic temperature of a general distribution," arXiv:0711.1460v2 [cond-mat.stat-mech], November 10 2007.

[86] K. R. Narayanan and A. R. Srinivasa, "A Shannon entropy-based non-equilibrium temperature of a general distribution with application to Langevin dynamics," preprint, May 2009.

[87] H. Nishimori, *Statistical Physics of Spin Glasses and Information Processing: An Introduction.* (International Series of Monographs on Physics, no. 111): Oxford University Press, 2001.

[88] L. Onsager, "Crystal statistics. I. A two-dimensional model with an order–disorder transition," *Physical Review*, vol. 65, no. 2, pp. 117–149, 1944.

[89] G. Parisi and F. Zamponi, "Mean field theory of the glass transition and jamming hard spheres," arXiv:0802.2180v2 [cond-mat.dis-nn], 18 December 2008.

[90] P. Pradhan, Y. Kafri, and D. Levine, "Non-equilibrium fluctuation theorems in the presence of local heating," arXiv:0712.0339v2 [cond-mat.stat-mech], 3 April 2008.

[91] A. Procacci and B. Scoppola, "Statistical mechanics approach to coding theory," *Journal of Statistical Physics*, vol. 96, no. 3/4, pp. 907–912, 1999.

[92] H. Qian, "Relative entropy: Free energy associated with equilibrium fluctuations and nonequilibrium deviations," *Physical Review E*, vol. 63, 042103, 2001.

[93] F. Reif, *Fundamentals of Statistical and Thermal Physics.* McGraw-Hill, 1965.

[94] H. Reiss, "Thermodynamic-like transformations in Information Theory," *Journal of Statistical Physics*, vol. 1, no. 1, pp. 107–131, 1969.

[95] H. Reiss, H. L. Frisch, and J. L. Lebowitz, "Statistical mechanics of rigid spheres," *Journal of Chemical Physics*, vol. 31, no. 2, pp. 369–380, August 1959.

[96] H. Reiss and C. Huang, "Statistical thermodynamic formalism in the solution of information theory problems," *Journal of Statistical Physics*, vol. 3, no. 2, pp. 191–211, 1971.

[97] H. Reiss and P. Schaaf, "Hard spheres: Thermodynamics and geometry," *Journal of Chemical Physics*, vol. 91, no. 4, pp. 2514–2524, 15 August 1989.

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

[99] K. Rose, "A mapping approach to rate-distortion computation and analysis," *IEEE Transactions on Information Theory*, vol. 40, no. 6, pp. 1939–1952, November 1994.

[100] D. Ruelle, *Statistical Mechanics: Rigorous Results.* Addison-Wesley, 1989.

[101] P. Ruján, "Finite temperature error-correcting codes," *Physics Review Letters*, vol. 70, no. 19, pp. 2968–2971, May 1993.

[102] F. W. Sears, M. W. Zemansky, and H. D. Young, *University Physics.* Addison–Wesley, 1976.

[103] J. P. Sethna, *Statistical Mechanics: Entropy, Order Parameters, and Complexity.* Oxford University Press, 2007.

[104] O. Shental and I. Kanter, "Shannon capacity of infinite-range spin-glasses," Technical report, Bar Ilan University, 2005.

[105] O. Shental, N. Shental, S. S. (Shitz), I. Kanter, A. J. Weiss, and Y. Weiss, "Discrete input two-dimensional Gaussian channels with memory: Estimation and information rates via graphical models and statistical mechanics," *IEEE Transactions on Information Theory*, vol. 54, no. 4, April 2008.

[106] J. E. Shore and R. W. Johnson, "Axiomatic derivation of the principle of maximum entropy and the principle of minimum cross-entropy," *IEEE Transactions on Information Theory*, vol. IT-26, no. 1, pp. 26–37, January 1980.

[107] N. Simanyi, "The Boltzmann–Sinai ergodic hypothesis in full generality," arXiv:1007.1206v2 [math.DS], 10 August 2010.

[108] A. Somekh-Baruch and N. Merhav, "Exact random coding exponents for erasure decoding," in *Proceedings of ISIT 2010*, pp. 260–264, Austin, Texas, U.S.A., June 2010.

[109] N. Sourlas, "Spin-glass models as error-correcting codes," *Nature*, vol. 339, pp. 693–695, June 1989.

[110] N. Sourlas, "Spin glasses, error-correcting codes and finite-temperature decoding," *Europhysics Letters*, vol. 25, pp. 159–164, 1994.

[111] K. Takeuchi, M. Vehpaperä, T. Tanaka, and R. R. Müller, "Replica analysis of general multiuser detection in MIMO DS-CDMA channels with imperfect CSI," in *Proceedings of ISIT 2008*, pp. 514–518, Toronto, Canada, July 2008.

[112] T. Tanaka, "Statistical mechanics of CDMA multiuser demodulation," *Europhysics Letters*, vol. 54, no. 4, pp. 540–546, 2001.

[113] T. Tanaka, "A statistical-mechanics approach to large-system analysis of CDMA multiuser detectors," *IEEE Transactions on Information Theory*, vol. 48, no. 11, pp. 2888–2910, November 2002.

[114] H. Touchette, "Methods for calculating nonconcave entropies," arXiv:1003.0382v1 [cond-mat.stat-mech], 1 March 2010.

[115] A. M. Tulino and S. Verdú, "Random matrix theory and wireless communications," *Foundations and Trends in Communications and Information Theory*, vol. 1, no. 1, 2004.

[116] J. J. M. Varbaarschot and M. R. Zirnbauer, "Critique of the replica trick," *Journal of Physics A: Mathamatical and General*, vol. 17, pp. 1093–1109, 1985.

[117] A. J. Viterbi and J. K. Omura, *Principles of Digital Communication and Coding*. McGraw–Hill, 1979.

[118] M. J. Wainwright and M. I. Jordan, "Graphical models, exponential families, and variational inference," *Foundations and Trends in Machine Learning*, vol. 1, no. 1–2, 2008.

[119] J. M. Wozencraft and I. M. Jacobs, *Principles of Communication Engineering*. John Wiley & Sons, 1965. Reissued by Waveland Press, 1990.

[120] A. D. Wyner, "On the probability of buffer overflow under an arbitrary bounded input-output distribution," *SIAM Journal on Applied Mathematics*, vol. 27, no. 4, pp. 544–570, December 1974.

[121] C. N. Yang and T. D. Lee, "Statistical theory of equations of state and phase transitions. I. Theory of condensation," *Physical Review*, vol. 87, no. 3, pp. 404–409, August 1952.

[122] J. S. Yedidia, *Quenched Disorder: Understanding Glasses Using a Variational Principle and the Replica Method*. lectures delivered at the Santa Fe Summer School on Complex Systems, June 15–19 1992.

[123] J. S. Yedidia, W. T. Freeman, and Y. Weiss, "Constructing free energy approximations and generalized belief propagation algorithms," *IEEE Transactions on Information Theory*, vol. 51, no. 7, pp. 2282–2312, July 2005.

[124] R. W. Yeung, "A first course in information theory," in *Information Theory and Network Coding*, (R. W. Yeung, ed.), New York, 2002: Kluwer Academic/Plenum Publishers, 2008. Springer.

[125] M. Zakai and J. Ziv, "A generalization of the rate-distortion theory and applications," in *Information Theory New Trends and Open Problems*, (G. Longo, ed.), pp. 87–123, Springer-Verlag, 1975.

[126] M. R. Zirnbauer, "Another critique of the replica trick," arXiv:cond-mat/9903338v1 [cond-mat.mes-hall], 22 March 1999.

[127] J. Ziv and M. Zakai, "On functionals satisfying a data-processing theorem," *IEEE Transactions on Information Theory*, vol. IT-19, no. 3, pp. 275–283, May 1973.