Elements of Information Theory
Thomas M. Cover, Joy A. Thomas
Copyright © 1991 John Wiley & Sons, Inc.
Print ISBN 0-471-06259-6 Online ISBN 0-471-20061-1

Bibliography

[1] N. M. Abramson. *Information Theory and Coding*. McGraw-Hill, New York, 1963.

[2] R. L. Adler, D. Coppersmith, and M. Hassner. Algorithms for sliding block codes—an application of symbolic dynamics to information theory. *IEEE Trans. Inform. Theory*, IT-29: 5–22, 1983.

[3] R. Ahlswede. Multi-way communication channels. In *Proc. 2nd. Int. Symp. Information Theory (Tsahkadsor, Armenian S.S.R.)*, pages 23–52, Prague, 1971. Publishing House of the Hungarian Academy of Sciences.

[4] R. Ahlswede. The capacity region of a channel with two senders and two receivers. *Ann. Prob.*, 2: 805–814, 1974.

[5] R. Ahlswede, Elimination of correlation in random codes for arbitrarily varying channels. *Zeitschrift für Wahrscheinlichkeitstheorie and verwandte Gebiete*, 33: 159–175, 1978.

[6] R. Ahlswede and T. S. Han. On source coding with side information via a multiple access channel and related problems in multi-user information theory. *IEEE Trans. Inform. Theory*, IT-29: 396–412, 1983.

[7] R. Ahlswede and J. Körner. Source coding with side information and a converse for the degraded broadcast channel. *IEEE Trans. Inform. Theory*, IT-21: 629–637, 1975.

[8] P. Algoet and T. M. Cover. A sandwich proof of the Shannon-McMillan-Breiman theorem. *Annals of Probability*, 16: 899–909, 1988.

[9] P. Algoet and T. M. Cover. Asymptotic optimality and asymptotic equipartition property of log-optimal investment. *Annals of Probability*, 16: 876–898, 1988.

[10] S. Amari. *Differential-Geometrical Methods in Statistics*. Springer-Verlag, New York, 1985.

[11] S. Arimoto. An algorithm for calculating the capacity of an arbitrary

discrete memoryless channel. *IEEE Trans. Inform. Theory*, IT-18: 14–20, 1972.

[12] S. Arimoto. On the converse to the coding theorem for discrete memoryless channels. *IEEE Trans. Inform. Theory*, IT-19: 357–359, 1973.

[13] R. B. Ash. *Information Theory*. Interscience, New York, 1965.

[14] J. Axzél and Z. Daróczy. *On Measures of Information and Their Characterization*. Academic Press, New York, 1975.

[15] A. Barron. Entropy and the Central Limit Theorem. *Annals of Probability*, 14: 336–342, April 1986.

[16] A. Barron and T. M. Cover. A bound on the financial value of information. *IEEE Trans. Inform. Theory*, IT-34: 1097–1100, 1988.

[17] A. R. Barron. Logically smooth density estimation. PhD Thesis, Department of Electrical Engineering, Stanford University, 1985.

[18] A. R. Barron. The strong ergodic theorem for densities: Generalized Shannon-McMillan-Breiman theorem. *Annals of Probability*, 13: 1292–1303, 1985.

[19] W. Beckner. Inequalities in Fourier Analysis. *Annals of Mathematics*, 102: 159–182, 1975.

[20] R. Bell and T. M. Cover. Competitive Optimality of Logarithmic Investment. *Mathematics of Operations Research*, 5: 161–166, 1980.

[21] R. Bell and T. M. Cover. Game-theoretic optimal portfolios. *Management Science*, 34: 724–733, 1988.

[22] T. C. Bell, J. G.Cleary, and I. H. Witten, *Text Compression*. Prentice-Hall, Englewood Cliffs, NJ, 1990.

[23] R. Bellman. Notes on Matrix Theory—IV: An inequality due to Bergstrøm. *Am. Math. Monthly*, 62: 172–173, 1955.

[24] C. H. Bennett. Demons, Engines and the Second Law. *Scientific American*, 259(5): 108–116, 1987.

[25] C. H. Bennett and R. Landauer. The fundamental physical limits of computation. *Scientific American*, 255: 48–56, 1985.

[26] R. Benzel. The capacity region of a class of discrete additive degraded interference channels. *IEEE Trans. Inform. Theory*, IT-25: 228–231, 1979.

[27] T. Berger. *Rate Distortion Theory: A Mathematical Basis for Data Compression*. Prentice-Hall, Englewood Cliffs, NJ, 1971.

[28] T. Berger. Multiterminal Source Coding. In G. Longo, editor, *The Information Theory Approach to Communications*. Springer-Verlag, New York, 1977.

[29] T. Berger. Multiterminal source coding. In *Lecture notes presented at the 1977 CISM Summer School, Udine, Italy*, pages 569–570, Prague, 1977. Princeton University Press, Princeton, NJ.

[30] T. Berger and R. W. Yeung. Multiterminal source encoding with one distortion criterion. *IEEE Trans. Inform. Theory*, IT-35: 228–236, 1989.

[31] P. Bergmans. Random coding theorem for broadcast channels with degraded components. *IEEE Trans. Inform. Theory*, IT-19: 197–207, 1973.

[32] E. R. Berlekamp. *Block Coding with Noiseless Feedback*. PhD Thesis, MIT, Cambridge, MA, 1964.

[33] M. Bierbaum and H. M. Wallmeier. A note on the capacity region of the multiple access channel. *IEEE Trans. Inform. Theory*, IT-25: 484, 1979.

[34] N. Blachman. The convolution inequality for entropy powers. *IEEE Trans. Inform. Theory*, IT-11: 267–271, 1965.

[35] D. Blackwell, L. Breiman, and A. J. Thomasian. The capacity of a class of channels. *Ann. Math. Stat.*, 30: 1229–1241, 1959.

[36] D. Blackwell, L. Breiman, and A. J. Thomasian. The capacities of certain channel classes under random coding. *Ann. Math. Stat.*, 31: 558–567, 1960.

[37] R. Blahut. Computation of Channel capacity and rate distortion functions. *IEEE Trans. Inform. Theory*, IT-18: 460–473, 1972.

[38] R. E. Blahut. Information bounds of the Fano-Kullback type. *IEEE Trans. Inform. Theory*, IT-22: 410–421, 1976.

[39] R. E. Blahut. *Principles and Practice of Information Theory*. Addison-Wesley, Reading, MA, 1987.

[40] R. E. Blahut. Hypothesis testing and Information theory. *IEEE Trans. Inform. Theory*, IT-20: 405–417, 1974.

[41] R. E. Blahut. *Theory and Practice of Error Control Codes*. Addison-Wesley, Reading, MA, 1983.

[42] R. C. Bose and D. K. Ray-Chaudhuri. On a class of error correcting binary group codes. *Inform. Contr.*, 3: 68–79, 1960.

[43] H. J. Brascamp and E. J. Lieb. Best constants in Young's inequality, its converse and its generalization to more than three functions. *Advances in Mathematics*, 20: 151–173, 1976.

[44] L. Breiman. The individual ergodic theorems of information theory. *Ann. Math. Stat.*, 28: 809–811, 1957.

[45] L. Breiman. Optimal Gambling systems for favourable games. In *Fourth Berkeley Symposium on Mathematical Statistics and Probability*, pages 65–78, Prague, 1961. University of California Press, Berkeley, CA.

[46] Leon. Brillouin. *Science and information theory*. Academic Press, New York, 1962.

[47] J. P. Burg. *Maximum entropy spectral analysis*. PhD Thesis, Department of Geophysics, Stanford University, Stanford, CA, 1975.

[48] A. Carleial. Outer bounds on the capacity of the interference channel. *IEEE Trans. Inform. Theory*, IT-29: 602–606, 1983.

[49] A. B. Carleial. A case where interference does not reduce capacity. *IEEE Trans. Inform. Theory*, IT-21: 569–570, 1975.

[50] G. J. Chaitin. On the length of programs for computing binary sequences. *J. Assoc. Comp. Mach.*, 13: 547–569, 1966.

[51] G. J. Chaitin. Information theoretical limitations of formal systems. *J. Assoc. Comp. Mach.*, 21: 403–424, 1974.

[52] G. J. Chaitin. Randomness and mathematical proof. *Scientific American*, 232: 47–52, 1975.

[53] G. J. Chaitin. Algorithmic information theory. *IBM Journal of Research and Development*, 21: 350–359, 1977.

[54] G. J. Chaitin. *Algorithmic Information Theory*. Cambridge University Press, Cambridge, 1987.

[55] H. Chernoff. A measure of the asymptotic efficiency of tests of a hypothesis based on a sum of observations. *Ann. Math. Stat.*, 23: 493–507, 1952.

[56] B. S. Choi and T. M. Cover. An information-theoretic proof of Burg's Maximum Entropy Spectrum. *Proc. IEEE*, 72: 1094–1095, 1984.

[57] K. L. Chung. A note on the ergodic theorem of information theory. *Ann. Math. Statist.*, 32: 612–614, 1961.

[58] M. Costa and T. M. Cover. On the similarity of the entropy power inequality and the Brunn-Minkowski inequality. *IEEE Trans. Inform. Theory*, IT-30: 837–839, 1984.

[59] T. M. Cover and J. A. Thomas. Determinant inequalities via information theory. *SIAM Journal of Matrix Analysis and its Applications*, 9: 384–392, 1988.

[60] T. M. Cover. Broadcast channels. *IEEE Trans. Inform. Theory*, IT-18: 2–14, 1972.

[61] T. M. Cover. Enumerative source encoding. *IEEE Trans. Inform. Theory*, IT-19: 73–77, 1973.

[62] T. M. Cover. An achievable rate region for the broadcast channel. *IEEE Trans. Inform. Theory*, IT-21: 399–404, 1975.

[63] T. M. Cover. A proof of the data compression theorem of Slepian and Wolf for ergodic sources. *IEEE Trans. Inform. Theory*, IT-22: 226–228, 1975.

[64] T. M. Cover. An algorithm for maximizing expected log investment return. *IEEE Trans. Inform. Theory*, IT-30: 369–373, 1984.

[65] T. M. Cover. Kolmogorov complexity, data compression and inference. In Skwirzynski, J., editor, *The Impact of Processing Techniques on Communications*. Martinus-Nijhoff Publishers, Dodrecht, 1985.

[66] T. M. Cover. Universal Portfolios. *Math. Finance*, 16: 876–898, 1991.

[67] T. M. Cover and A. El Gamal. Capacity theorems for the relay channel. *IEEE Trans. Inform. Theory*, IT-25: 572–584, 1979.

[68] T. M. Cover and A. El Gamal. An information theoretic proof of Hadamard's inequality. *IEEE Trans. Inform. Theory*, IT-29: 930–931, 1983.

[69] T. M. Cover, A. El Gamal, and M. Salehi. Multiple access channels with arbtirarily correlated sources. *IEEE Trans. Inform. Theory*, IT-26: 648–657, 1980.

[70] T. M. Cover, P. Gács, and R. M. Gray, Kolmogorov's contributions to information theory and algorithmic complexity. *Annals of Probability*, 17: 840–865, 1989.

[71] T. M. Cover and B. Gopinath. *Open Problems in Communication and Computation*. Springer-Verlag, New York, 1987.

[72] T. M. Cover and R. King. A convergent gambling estimate of the entropy of English. *IEEE Trans. Inform. Theory*, IT-24: 413–421, 1978.

[73] T. M. Cover and C. S. K. Leung. An achievable rate region for the multiple access channel with feedback. *IEEE Trans. Inform. Theory*, IT-27: 292–298, 1981.

[74] T. M. Cover and S. K. Leung. Some equivalences between Shannon entropy and Kolmogorov complexity. *IEEE Trans. Inform. Theory*, IT-24: 331–338, 1978.

[75] T. M. Cover, R. J. McEliece, and E. Posner. Asynchronous multiple access channel capacity. *IEEE Trans. Inform. Theory*, IT-27: 409–413, 1981.

[76] T. M. Cover and S. Pombra. Gaussian feedback capacity. *IEEE Trans. Inform. Theory*, IT-35: 37–43, 1989.

[77] H. Cramér. *Mathematical Methods of Statistics*. Princeton University Press, Princeton, NJ, 1946.

[78] I. Csiszár. Information type measures of difference of probability distributions and indirect observations. *Studia Sci. Math. Hungar.*, 2: 229–318, 1967.

[79] I. Csiszár. On the computation of rate distortion functions. *IEEE Trans. Inform. Theory*, IT-20: 122–124, 1974.

[80] I. Csiszár. Sanov property, generalized I-projection and a conditional limit theorem. *Annals of Probability*, 12: 768–793, 1984.

[81] I. Csiszár, T. M. Cover, and B. S. Choi. Conditional limit theorems under Markov conditioning. *IEEE Trans. Inform. Theory*, IT-33: 788–801, 1987.

[82] I. Csiszár and J. Körner. Towards a general theory of source networks. *IEEE Trans. Inform. Theory*, IT-26: 155–165, 1980.

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

[84] I. Csiszár and G. Longo. *On the error exponent for source coding and for testing simple statistical hypotheses*. Hungarian Academy of Sciences, Budapest, 1971.

[85] I. Csiszár and G. Tusnády. Information geometry and alternating minimization procedures. *Statistics and Decisions*, Supplement Issue 1: 205–237, 1984.

[86] L. D. Davisson. Universal noiseless coding. *IEEE Trans. Inform. Theory*, IT-19: 783–795, 1973.

[87] A. Dembo. Information inequalities and uncertainty principles. Technical Report, Department of Statistics, Stanford University, 1990.

[88] A. P. Dempster, N. M. Laird, and D. B. Rubin. Maximum likelihood from incomplete data via the EM algorithm. *Journal Royal Stat. Soc., Series B*, 39: 1–38, 1977.

[89] R. L. Dobrushin. General formulation of Shannon's main theorem of information theory. *Usp. Math. Nauk.*, 14: 3–104, 1959. Translated in *Am. Math. Soc. Trans.*, 33: 323–438.

[90] G. Dueck. The capacity region of the two-way channel can exceed the inner bound. *Inform. Contr.*, 40: 258–266, 1979.

[91] G. Dueck. Partial feedback for two-way and broadcast channels. *Inform. Contr.*, 46: 1–15, 1980.

[92] G. Dueck and J. Körner. Reliability function of a discrete memoryless channel at rates above capacity. *IEEE Trans. Inform. Theory*, IT-25: 82–85, 1979.

[93] G. Dueck and L. Wolters. Ergodic theory and encoding of individual sequences. *Problems Contr. Inform. Theory*, 14: 329–345, 1985.

[94] P. M. Ebert. The capacity of the Gaussian channel with feedback. *Bell Sys. Tech. Journal*, 49: 1705–1712, October 1970.

[95] H. G. Eggleston. *Convexity*. Cambridge University Press, Cambridge, UK, 1969.

[96] A. El Gamal. The feedback capacity of degraded broadcast channels. *IEEE Trans. Inform. Theory*, IT-24: 379–381, 1978.

[97] A. El Gamal. The capacity region of a class of broadcast channels. *IEEE Trans. Inform. Theory*, IT-25: 166–169, 1979.

[98] A. El Gamal and T. M. Cover. Multiple user information theory. *Proc. IEEE*, 68: 1466–1483, 1980.

[99] A. El Gamal and T. M. Cover. Achievable rates for multiple descriptions. *IEEE Trans. Inform. Theory*, IT-28: 851–857, 1982.

[100] A. El Gamal and E. C. Van der Meulen. A proof of Marton's coding theorem for the discrete memoryless broadcast channel. *IEEE Trans. Inform. Theory*, IT-27: 120–122, 1981.

[101] P. Elias. Error-free Coding. *IRE Trans. Inform. Theory*, IT-4: 29–37, 1954.

[102] P. Elias. Coding for noisy channels. In *IRE Convention Record, Pt. 4*, pages 37–46, 1955.

[103] Ky Fan. On a Theorem of Weyl concerning the eigenvalues of linear transformations II. *Proc. National Acad. Sci. U.S.*, 36: 31–35, 1950.

[104] Ky Fan. Some inequalities concerning positive-definite matrices. *Proc. Cambridge Phil. Soc.*, 51: 414–421, 1955.

[105] R. M. Fano. Class notes for Transmission of Information, Course 6.574. MIT, Cambridge, MA, 1952.

[106] R. M. Fano. *Transmission of Information: A Statistical Theory of Communication*. Wiley, New York, 1961.

[107] A. Feinstein. A new basic theorem of information theory. *IRE Trans. Inform. Theory*, IT-4: 2–22, 1954.

[108] A. Feinstein. *Foundations of Information Theory*. McGraw-Hill, New York, 1958.

[109] A. Feinstein. On the coding theorem and its converse for finite-memory channels. *Inform. Contr.*, 2: 25–44, 1959.

[110] W. Feller. *An Introduction to Probability Theory and Its Applications*. Wiley, New York, 1957.

[111] R. A. Fisher. On the mathematical foundations of theoretical statistics. *Philos. Trans. Roy. Soc., London, Sec. A*, 222: 309–368, 1922.

[112] R. A. Fisher. Theory of Statistical Estimation. *Proc. Cambridge Phil. Society*, 22: 700–725, 1925.

[113] L. R. Ford and D. R. Fulkerson. *Flows in Networks*. Princeton University Press, Princeton, NJ, 1962.

[114] G. D. Forney. Exponential error bounds for erasure, list and decision feedback schemes. *IEEE Trans. Inform. Theory*, IT-14: 549–557, 1968.

[115] G. D. Forney. Information Theory. Unpublished course notes. Stanford University, 1972.

[116] P. A. Franaszek. On synchronous variable length coding for discrete noiseless channels. *Inform. Contr.*, 15: 155–164, 1969.

[117] T. Gaarder and J. K. Wolf. The capacity region of a multiple-access discrete memoryless channel can increase with feedback. *IEEE Trans. Inform. Theory*, IT-21: 100–102, 1975.

[118] R. G. Gallager. A simple derivation of the coding theorem and some applications. *IEEE Trans. Inform. Theory*, IT-11: 3–18, 1965.

[119] R. G. Gallager. Capacity and coding for degraded broadcast channels. *Problemy Peredaci Informaccii*, 10(3): 3–14, 1974.

[120] R. G. Gallager. *Information Theory and Reliable Communication*. Wiley, New York, 1968.

[121] E. W. Gilbert and E. F. Moore. Variable length binary encodings. *Bell Sys. Tech. Journal*, 38: 933–967, 1959.

[122] S. Goldman. *Information Theory*. Prentice-Hall, Englewood Cliffs, NJ, 1953.

[123] R. M. Gray. Sliding block source coding. *IEEE Trans. Inform. Theory*, IT-21: 357–368, 1975.

[124] R. M. Gray. *Entropy and Information Theory*. Springer-Verlag, New York, 1990.

[125] R. M. Gray and A. Wyner. Source coding for a simple network. *Bell Sys. Tech. Journal*, 58: 1681–1721, 1974.

[126] U. Grenander and G. Szego. *Toeplitz forms and their applications*. University of California Press, Berkeley, 1958.

[127] B. Grünbaum. *Convex Polytopes*. Interscience, New York, 1967.

[128] S. Guiasu. *Information Theory with Applications*. McGraw-Hill, New York, 1976.

[129] R. V. Hamming. Error detecting and error correcting codes. *Bell Sys. Tech. Journal*, 29: 147–160, 1950.

[130] T. S. Han. Nonnegative entropy measures of multivariate symmetric correlations. *Inform. Contr.*, 36: 133–156, 1978.

[131] T. S. an and M. H. M. Costa. Broadcast channels with arbitrarily correlated sources. *IEEE Trans. Inform. Theory*, IT-33: 641–650, 1987.

[132] T. S. Han and K. Kobayashi. A new achievable rate region for the interference channel. *IEEE Trans. Inform. Theory*, IT-27: 49–60, 1981.

[133] R. V. Hartley. Transmission of information. *Bell Sys. Tech. Journal*, 7: 535, 1928.

[134] P. A. Hocquenghem. Codes correcteurs d'erreurs. *Chiffres*, 2: 147–156, 1959.

[135] J. L. Holsinger. Digital communication over fixed time-continuous channels with memory, with special application to telephone channels. Technical Report, M. I. T., 1964.

[136] J. E. Hopcroft and J. D. Ullman. *Introduction to Automata Theory, Formal Languages and Computation*. Addison Wesley, Reading, MA, 1979.

[137] Y. Horibe. An improved bound for weight-balanced tree. *Inform. Contr.*, 34: 148–151, 1977.

[138] D. A. Huffman. A method for the construction of minimum redundancy codes. *Proc. IRE*, 40: 1098–1101, 1952.

[139] E. T. Jaynes. Information theory and statistical mechanics. *Phys. Rev.*, 106: 620, 1957.

[140] E. T. Jaynes. Information theory and statistical mechanics. *Phys. Rev.*, 108: 171, 1957.

[141] E. T. Jaynes. Information theory and statistical mechanics I. *Phys. Rev.*, 106: 620–630, 1957.

[142] E. T. Jaynes. On the rationale of maximum entropy methods. *Proc. IEEE*, 70: 939–952, 1982.

[143] E. T. Jaynes. *Papers on Probability, Statistics and Statistical Physics* . Reidel, Dordrecht, 1982.

[144] F. Jelinek. Buffer overflow in variable length encoding of fixed rate sources. *IEEE Trans. Inform. Theory*, IT-14: 490–501, 1968.

[145] F. Jelinek. Evaluation of expurgated error bounds. *IEEE Trans. Inform. Theory*, IT-14: 501–505, 1968.

[146] F. Jelinek. *Probabilistic Information Theory*. McGraw-Hill, New York, 1968.

[147] J. Justesen. A class of constructive asymptotically good algebraic codes. *IEEE Trans. Inform. Theory*, IT-18: 652–656, 1972.

[148] T. Kailath and J. P. M. Schwalkwijk. A coding scheme for additive noise channels with feedback—Part I: No bandwidth constraints. *IEEE Trans. Inform. Theory*, IT-12: 172–182, 1966.

[149] J. Karush. A simple proof of an inequality of McMillan. *IRE Trans. Inform. Theory*, IT-7: 118, 1961.

[150] J. Kelly. A new interpretation of information rate. *Bell Sys. Tech. Journal*, 35: 917–926, 1956.

[151] J. H. B. Kemperman. *On the Optimum Rate of Transmitting Information* pp. 126–169. Springer-Verlag, New York, 1967.

[152] M. Kendall and A. Stuart. *The Advanced Theory of Statistics*. MacMillan, New York, 1977.

[153] A. Ya. Khinchin. *Mathematical Foundations of Information Theory*. Dover, New York, 1957.

[154] J. C. Kieffer. A simple proof of the Moy-Perez generalization of the Shannon-McMillan theorem. *Pacific J. Math.*, 51: 203–206, 1974.

[155] D. E. Knuth and A. C. Yao. The complexity of random number generation. In J. F. Traub, editor, *Algorithms and Complexity: Recent Results and New Directions. Proceedings of the Symposium on New Directions and Recent Results in Algorithms and Complexity, Carnegie-Mellon University, 1976.* Academic Press, New York, 1976.

[156] A. N. Kolmogorov. On the Shannon theory of information transmission in the case of continuous signals. *IRE Trans. Inform. Theory*, IT-2: 102–108, 1956.

[157] A. N. Kolmogorov. A new invariant for transitive dynamical systems. *Dokl. An. SSR*, 119: 861–864, 1958.

[158] A. N. Kolmogorov. Three approaches to the quantitative definition of information. *Problems of Information Transmission*, 1: 4–7, 1965.

[159] A. N. Kolmogorov. Logical basis for information theory and probability theory. *IEEE Trans. Inform. Theory*, IT-14: 662–664, 1968.

[160] J. Körner and K. Marton. The comparison of two noisy channels. In Csiszár, I. and Elias, P. , editor, *Topics in Information Theory*. Coll. Math. Soc. J. Bolyai, No. 16, North Holland, Amsterdam, 1977.

[161] J. Körner and K. Marton. General broadcast channels with degraded message sets. *IEEE Trans. Inform. Theory*, IT-23: 60–64, 1977.

[162] J. Körner and K. Marton. How to encode the modulo 2 sum of two binary sources. *IEEE Trans. Inform. Theory*, IT-25: 219–221, 1979.

[163] V. A. Kotel'nikov. *The Theory of Optimum Noise Immunity*. McGraw-Hill, New York, 1959.

[164] L. G. Kraft. A device for quanitizing, grouping and coding amplitude modulated pulses. Master's Thesis, Department of Electrical Engineering, MIT, Cambridge, MA, 1949.

[165] S. Kullback. *Information Theory and Statistics*. Wiley, New York, 1959.

[166] S. Kullback. A lower bound for discrimination in terms of variation. *IEEE Trans. Inform. Theory*, IT-13: 126–127, 1967.

[167] S. Kullback and R. A. Leibler. On information and sufficiency. *Ann. Math. Stat.*, 22: 79–86, 1951.

[168] H. J. Landau and H. O. Pollak. Prolate spheroidal wave functions, Fourier analysis and uncertainty: Part III. *Bell Sys. Tech. Journal*, 41: 1295–1336, 1962.

[169] H. J. Landau and H. O. Pollak. Prolate spheroidal wave functions, Fourier analysis and uncertainty: Part II. *Bell Sys. Tech. Journal*, 40: 65–84, 1961.

[170] G. G. Langdon. An introduction to arithmetic coding. *IBM Journal of Research and Development*, 28: 135–149, 1984.

[171] G. G. Langdon and J. J. Rissanen. A simple general binary source code. *IEEE Trans. Inform. Theory*, IT-28: 800, 1982.

[172] H. A. Latané. Criteria for choice among risky ventures. *Journal of Political Economy*, 38: 145–155, 1959.

[173] H. A. Latané and D. L. Tuttle. Criteria for portfolio building. *Journal of Finance*, 22: 359–373, 1967.

[174] E. L. Lehmann and H. Scheffé. Completeness, similar regions and unbiased estimation. *Sankhya*, 10: 305–340, 1950.

[175] A. Lempel and J. Ziv. On the complexity of finite sequences. *IEEE Trans. Inform. Theory*, IT-22: 75–81, 1976.

[176] L. A. Levin. On the notion of a random sequence. *Soviet Mathematics Doklady*, 14: 1413–1416, 1973.

[177] L. A. Levin and A. K. Zvonkin. The complexity of finite objects and the development of the concepts of information and randomness by means of the theory of algorithms. *Russian Mathematical Surveys*, 25/6: 83–124, 1970.

[178] H. Liao. Multiple access channels. PhD Thesis, Department of Electrical Engineering, University of Hawaii, Honolulu, 1972.

[179] S. Lin and D. J. Costello, Jr. *Error Control Coding: Fundamentals and Applications*. Prentice-Hall, Englewood Cliffs, NJ, 1983.

[180] Y. Linde, A. Buzo. and R. M. Gray. An algorithm for vector quantizer design. *IEEE Transactions on Communications*, COM-28: 84–95, 1980.

[181] S. P. Lloyd. Least squares quantization in PCM. Bell Laboratories Technical Note, 1957.

[182] L. Lovasz. On the Shannon capacity of a graph. *IEEE Trans. Inform. Theory*, IT-25: 1–7, 1979.

[183] R. W. Lucky. *Silicon Dreams: Information, Man and Machine*. St. Martin's Press, New York, 1989.

[184] B. Marcus. Sofic systems and encoding data. *IEEE Trans. Inform. Theory*, IT-31: 366–377, 1985.

[185] A. Marshall and I. Olkin. *Inequalities: Theory of Majorization and Its Applications*. Academic Press, New York, 1979.

[186] A. Marshall and I. Olkin. A convexity proof of Hadamard's inequality. *Am. Math. Monthly*, 89: 687–688, 1982.

[187] P. Martin-Löf. The definition of random sequences. *Inform. Contr.*, 9: 602–619, 1966.

[188] K. Marton. Error exponent for source coding with a fidelity criterion. *IEEE Trans. Inform. Theory*, IT-20: 197–199, 1974.

[189] K. Marton. A coding theorem for the discrete memoryless broadcast channel. *IEEE Trans. Inform. Theory*, IT-25: 306–311, 1979.

[190] R. A. McDonald and P. M. Schultheiss. Information Rates of Gaussian signals under criteria constraining the error spectrum. *Proc. IEEE*, 52: 415–416, 1964.

[191] R. J. McEliece. *The Theory of Information and Coding*. Addison-Wesley, Reading, MA, 1977.

[192] B. McMillan. The basic theorems of information theory. *Ann. Math. Stat.*, 24: 196–219, 1953.

[193] B. McMillan. Two inequalities implied by unique decipherability. *IEEE Trans. Inform. Theory*, IT-2: 115–116, 1956.

[194] R. C. Merton and P. A. Samuelson. Fallacy of the log-normal approximation to optimal portfolio decision-making over many periods. *Journal of Financial Economics*, 1: 67–94, 1974.

[195] H. Minkowski. Diskontinuitätsbereich für arithmetische Äquivalenz. *Journal für Math.*, 129: 220–274, 1950.

[196] L. Mirsky. On a generalization of Hadamard's determinantal inequality due to Szasz. *Arch. Math.*, VIII: 274–275, 1957.

[197] S. C. Moy. Generalizations of the Shannon-McMillan theorem. *Pacific Journal of Mathematics*, 11: 705–714, 1961.

[198] J. Neyman and E. S. Pearson. On the problem of the most efficient tests of statistical hypotheses. *Phil. Trans. Roy. Soc., London, Series A*, 231: 289–337, 1933.

[199] H. Nyquist. Certain factors affecting telegraph speed. *Bell Sys. Tech. Journal*, 3: 324, 1924.

[200] J. Omura. A coding theorem for discrete time sources. *IEEE Trans. Inform. Theory*, IT-19: 490–498, 1973.

[201] A. Oppenheim. Inequalities connected with definite Hermitian forms. *J. London Math. Soc.*, 5: 114–119, 1930.

[202] S. Orey. On the Shannon-Perez-Moy theorem. *Contemp. Math.*, 41: 319–327, 1985.

[203] D. S. Ornstein. Bernoulli shifts with the same entropy are isomorphic. *Advances in Math.*, 4: 337–352, 1970.

[204] L. H. Ozarow. The capacity of the white Gaussian multiple access channel with feedback. *IEEE Trans. Inform. Theory*, IT-30: 623–629, 1984.

[205] L. H. Ozarow and C. S. K. Leung. An achievable region and an outer bound for the Gaussian broadcast channel with feedback. *IEEE Trans. Inform. Theory*, IT-30: 667–671, 1984.

[206] H. Pagels. *The Dreams of Reason: The Computer and the Rise of the Sciences of Complexity*. Simon and Schuster, New York, 1988.

[207] R. Pasco. Source coding algorithms for fast data compression. PhD Thesis, Stanford University, 1976.

[208] A. Perez. Extensions of Shannon-McMillan's limit theorem to more general stochastic processes. In *Trans. Third Prague Conference on Information Theory, Statistical Decision Functions and Random Processes*, pages 545–574, Prague, 1964. Czechoslovak Academy of Sciences.

[209] J. T. Pinkston. An application of rate-distortion theory to a converse to the coding theorem. *IEEE Trans. Inform. Theory*, IT-15: 66–71, 1969.

[210] M. S. Pinsker. Talk at Soviet Information Theory Meeting, 1969.

[211] M. S. Pinsker. The capacity region of noiseless broadcast channels. *Prob. Inform. Trans.*, 14: 97–102, 1978.

[212] M. S. Pinsker. *Information and Stability of Random Variables and Processes*. Izd. Akad. Nauk, 1960.

[213] L. R. Rabiner and R. W. Schafer. *Digital Processing of Speech Signals*. Prentice-Hall, Englewood Cliffs, NJ, 1978.

[214] C. R. Rao. Information and accuracy obtainable in the estimation of statistical parameters. *Bull. Calcutta Math. Soc.*, 37: 81–91, 1945.

[215] F. M. Reza. *An Introduction to Information Theory*. McGraw-Hill, New York, 1961.

[216] S. O. Rice. Communication in the presence of noise—Probability of error for two encoding schemes. *Bell Sys. Tech. Journal*, 29: 60–93, 1950.

[217] J. Rissanen. Generalized Kraft inequality and arithmetic coding. *IBM J. Res. Devel.* 20: 198, 1976.

[218] J. Rissanen. Modelling by shortest data description. *Automatica*, 14: 465–471, 1978.

[219] J. Rissanen. A universal prior for integers and estimation by minimum description length. *Ann. Stat.*, 11: 416–431, 1983.

[220] J. Rissanen. Universal coding, information, prediction and estimation. *IEEE Trans. Inform. Theory*, IT-30: 629–636, 1984.

[221] J. Rissanen. Stochastic complexity and modelling. *Ann. Stat.*, 14: 1080–1100, 1986.

[222] J. Rissanen. Stochastic complexity (with discussions). *Journal of the Royal Statistical Society*, 49: 223–239, 252–265, 1987.

[223] J. Rissanen. *Stochastic Complexity in Statistical Inquiry*. World Scientific, New Jersey, 1989.

[224] P. A. Samuelson. Lifetime portfolio selection by dynamic stochastic programming. *Rev. of Economics and Statistics*, 1: 236–239, 1969.

[225] P. A. Samuelson. The 'fallacy' of maximizing the geometric mean in long sequences of investing or gambling. *Proc. Nat. Acad. Science*, 68: 214–224, 1971.

[226] P. A. Samuelson. Why we should not make mean log of wealth big though years to act are long. *Journal of Banking and Finance*, 3: 305–307, 1979.

[227] I. N. Sanov. On the probability of large deviations of random variables. *Mat. Sbornik*, 42: 11–44, 1957.

[228] A. A. Sardinas and G. W. Patterson. A necessary and sufficient condition for the unique decomposition of coded messages. In *IRE Convention Record, Part 8*, pages 104–108, 1953.

[229] H. Sato. The capacity of the Gaussian interference channel under strong interference. *IEEE Trans. Inform. Theory*, IT-27: 786–788, 1981.

[230] H. Sato, On the capacity region of a discrete two-user channel for strong interference. *IEEE Trans. Inform. Theory*, IT-24: 377–379, 1978.

[231] H. Sato and M. Tanabe. A discrete two-user channel with strong interference. *Trans. IECE Japan*, 61: 880–884, 1978.

[232] J. P. M. Schalkwijk. The binary multiplying channel—a coding scheme that operates beyond Shannon's inner bound. *IEEE Trans. Inform. Theory*, IT-28: 107–110, 1982.

[233] J. P. M. Schalkwijk. On an extension of an achievable rate region for the binary multiplying channel. *IEEE Trans. Inform. Theory*, IT-29: 445–448, 1983.

[234] C. P. Schnorr. Process, complexity and effective random tests. *Journal of Computer and System Sciences*, 7: 376–388, 1973.

[236] C. P. Schnorr. A surview on the theory of random sequences. In Butts, R. and Hinitikka, J. , editor, *Logic, Methodology and Philosophy of Science*. Reidel, Dodrecht, 1977.

[237] G. Schwarz. Estimating the dimension of a model. *Ann. Stat.*, 6: 461–464, 1978.

[238] C. E. Shannon. A mathematical theory of communication. *Bell Sys. Tech. Journal*, 27: 379–423, 623–656, 1948.

[239] C. E. Shannon. The zero-error capacity of a noisy channel. *IRE Trans. Inform. Theory*, IT-2: 8–19, 1956.

[240] C. E. Shannon. Communication in the presence of noise. *Proc. IRE*, 37: 10–21, 1949.

[241] C. E. Shannon. Communication theory of secrecy systems. *Bell Sys. Tech. Journal*, 28: 656–715, 1949.

[242] C. E. Shannon. Prediction and entropy of printed English. *Bell Sys. Tech. Journal*, 30: 50–64, 1951.

[243] C. E. Shannon. Certain results in coding theory for noisy channels. *Inform. Contr.*, 1: 6–25, 1957.

[244] C. E. Shannon. Channels with side information at the transmitter. *IBM J. Res. Develop.*, 2: 289–293, 1958.

[245] C. E. Shannon. Coding theorems for a discrete source with a fidelity criterion. *IRE National Convention Record, Part 4*, pages 142–163, 1959.

[246] C. E. Shannon. Two-way communication channels. In *Proc. 4th Berkeley Symp. Math. Stat. Prob.*, pages 611–644, University of California Press, Berkeley, 1961.

[247] C. E. Shannon, R. G. Gallager, and E. R. Berlekamp. Lower bounds to error probability for coding in discrete memoryless channels. I. *Inform. Contr.*, 10: 65–103, 1967.

[248] C. E. Shannon, R. G. Gallager, and E. R. Berlekamp. Lower bounds to error probability for coding in discrete memoryless channels. II. *Inform. Contr.*, 10: 522–552, 1967.

[249] C. E. Shannon and W. W. Weaver. *The Mathematical Theory of Communication*. University of Illinois Press, Urbana, IL, 1949.

[250] W. F. Sharpe. *Investments*. Prentice-Hall, Englewood Cliffs, NJ, 1985.

[251] J. E. Shore and R. W. Johnson. Axiomatic derivation of the principle of maximum entropy and the principle of minimum cross-entropy. *IEEE Trans. Inform. Theory*, IT-26: 26–37, 1980.

[252] D. Slepian. *Key Papers in the Development of Information Theory*. IEEE Press, New York, 1974.

[253] D. Slepian and H. O. Pollak. Prolate spheroidal wave functions, Fourier

analysis and uncertainty: Part I. *Bell Sys. Tech. Journal*, 40: 43–64, 1961.

[254] D. Slepian and J. K. Wolf. A coding theorem for multiple access channels with correlated sources. *Bell Sys. Tech. Journal*, 52: 1037–1076, 1973.

[255] D. Slepian and J. K. Wolf. Noiseless coding of correlated information sources. *IEEE Trans. Inform. Theory*, IT-19: 471–480, 1973.

[256] R. J. Solomonoff. A formal theory of inductive inference. *Inform. Contr.*, 7: 1–22, 224–254, 1964.

[257] A. Stam. Some inequalities satisfied by the quantities of information of Fisher and Shannon. *Inform. Contr.*, 2: 101–112, 1959.

[258] D. L. Tang and L. R. Bahl. Block codes for a class of constrained noiseless channels. *Inform. Contr.*, 17: 436–461, 1970.

[259] S. C. Tornay. *Ockham: Studies and Selections*. Open Court Publishers, La Salle, IL, 1938.

[260] J. M. Van Campenhout and T. M. Cover. Maximum entropy and conditional probability. *IEEE Trans. Inform. Theory*, IT-27: 483–489, 1981.

[261] E. Van der Meulen. Random coding theorems for the general discrete memoryless broadcast channel. *IEEE Trans. Inform. Theory*, IT-21: 180–190, 1975.

[262] E. C. Van der Meulen. A survey of multi-way channels in information theory. *IEEE Trans. Inform. Theory*, IT-23: 1–37, 1977.

[263] E. C. Van der Meulen. Recent coding theorems and converses for multi-way channels. Part II: The multiple access channel (1976–1985). Department Wiskunde, Katholieke Universiteit Leuven, 1985.

[264] E. C. Van der Meulen. Recent coding theorems for multi-way channels. Part I: The broadcast channel (1976–1980). In J. K. Skwyrzinsky, editor, *New Concepts in Multi-user Communication (NATO Advanced Study Insititute Series)*. Sijthoff & Noordhoff International, 1981.

[265] A. C. G. Verdugo Lazo and P. N. Rathie. On the entropy of continuous probability distributions. *IEEE Trans. Inform. Theory*, IT-24: 120–122, 1978.

[266] A. J. Viterbi and J. K. Omura. *Principles of Digital Communication and Coding*. McGraw-Hill, New York, 1979.

[267] V. V. V'yugin. On the defect of randomness of a finite object with respect to measures with given complexity bounds. *Theory Prob. Appl.*, 32: 508–512, 1987.

[268] A. Wald. *Sequential Analysis*. Wiley, New York, 1947.

[269] T. A. Welch. A technique for high-performance data compression. *Computer*, 17: 8–19, 1984.

[270] N. Wiener. *Cybernetics*. MIT Press, Cambridge MA, and Wiley, New York, 1948.

[271] N. Wiener. *Extrapolation, Interpolation and Smoothing of Stationary Time Series*. MIT Press, Cambridge, MA and Wiley, New York, 1949.

[272] H. J. Wilcox and D. L. Myers. *An introduction to Lebesgue integration and Fourier series*. R. E. Krieger, Huntington, NY, 1978.

[273] F. M. J. Willems. The feedback capacity of a class of discrete memoryless multiple access channels. *IEEE Trans. Inform. Theory*, IT-28: 93–95, 1982.

[274] F. M. J. Willems and A. P. Hekstra. Dependence balance bounds for single-output two-way channels. *IEEE Trans. Inform. Theory*, IT-35: 44–53, 1989.

[275] I. H. Witten, R. M. Neal, and J. G. Cleary. Arithmetic coding for data compression. *Communications of the ACM*, 30: 520–540, 1987.

[276] J. Wolfowitz. The coding of messages subject to chance errors. *Illinois Journal of Mathematics*, 1: 591–606, 1957.

[277] J. Wolfowitz. *Coding Theorems of Information Theory*. Springer-Verlag, Berlin, and Prentice-Hall, Englewood Cliffs, NJ, 1978.

[278] P. M. Woodward. *Probability and Information Theory with Applications to Radar*. McGraw-Hill, New York, 1953.

[279] J. M. Wozencraft and I. M. Jacobs. *Principles of Communication Engineering*. Wiley, New York, 1965.

[280] A. Wyner. A theorem on the entropy of certain binary sequences and applications II. *IEEE Trans. Inform. Theory*, IT-19: 772–777, 1973.

[281] A. Wyner. The common information of two dependent random variables. *IEEE Trans. Inform. Theory*, IT-21: 163–179, 1975.

[282] A. Wyner. On source coding with side information at the decoder. *IEEE Trans. Inform. Theory*, IT-21: 294–300, 1975.

[283] A. Wyner and J. Ziv. A theorem on the entropy of certain binary sequences and applications I. *IEEE Trans. Inform. Theory*, IT-19: 769–771, 1973.

[284] A. Wyner and J. Ziv. The rate distortion function for source coding with side information at the receiver. *IEEE Trans. Inform. Theory*, IT-22: 1–11, 1976.

[285] A. Wyner and J. Ziv. On entropy and data compression. Submitted to *IEEE Trans. Inform. Theory*.

[286] A. D. Wyner. The capacity of the band-limited Gaussian channel. *Bell Sys. Tech. Journal*, 45: 359–371, 1965.

[287] Z. Zhang, T. Berger, and J. P. M. Schalkwijk. New outer bounds to capacity regions of two-way channels. *IEEE Trans. Inform. Theory*, IT-32: 383–386, 1986.

[288] J. Ziv. Coding of sources with unknown statistics—Part II: Distortion relative to a fidelity criterion. *IEEE Trans. Inform. Theory*, IT-18: 389–394, 1972.

[289] J. Ziv. Coding theorems for individual sequences. *IEEE Trans. Inform. Theory*, IT-24: 405–412, 1978.

[290] J. Ziv and A. Lempel. A universal algorithm for sequential data compression. *IEEE Trans. Inform. Theory*, IT-23: 337–343, 1977.

[291] J. Ziv and A. Lempel. Compression of individual sequences by variable rate coding. *IEEE Trans. Inform. Theory*, IT-24: 530–536, 1978.

[292] W. H. Zurek. Algorithmic randomness and physical entropy. *Phys. Rev. A*, 40: 4731–4751, October 15, 1989.

[293] W. H. Zurek. Thermodynamic cost of computation, algorithmic complexity and the information metric. *Nature*, 341: 119–124, September 14, 1989.

[294] W. H. Zurek, editor. *Complexity, Entropy and the Physics of Information. Proceedings of the 1988 Workshop on the Complexity, Entropy and the Physics of Information.* Addison-Wesley, Reading, MA, 1990.