

Bibliography

- [1] K. Sayood, *Introduction to Data Compression, Fourth Edition*. Morgan Kaufmann, Oct. 2012.
- [2] D. Salomon, *Data Compression: The Complete Reference*. Springer, Feb. 2004.
- [3] A. Said, “Introduction to arithmetic coding-theory and practice,” *Hewlett Packard Laboratories Report*, 2004.
- [4] T. Bell, I. H. Witten, and J. G. Cleary, “Modeling for text compression,” *ACM Computing Surveys (CSUR)*, vol. 21, no. 4, pp. 557–591, 1989.
- [5] J. Rissanen and G. Langdon, Jr., “Universal modeling and coding,” *IEEE Transactions on Information Theory*, vol. 27, no. 1, pp. 12–23, Jan. 1981.
- [6] C. E. Shannon, “A mathematical theory of communication,” *Bell System Technical Journal*, vol. 27, pp. 379–423, 623–656, Oct. 1948.
- [7] R. Gray and L. Davisson, “Source coding theorems without the ergodic assumption,” *IEEE Transactions on Information Theory*, vol. 20, no. 4, pp. 502–516, Jul. 1974.
- [8] D. Slepian and J. Wolf, “Noiseless coding of correlated information sources,” *IEEE Transactions on Information Theory*, vol. 19, no. 4, pp. 471–480, Jul. 1973.
- [9] A. Hitron and U. Erez, “Optimality of linear codes over PAM for the modulo-additive Gaussian channel,” in *2012 IEEE International Symposium on Information Theory Proceedings (ISIT)*, Jul. 2012, pp. 1742–1746.
- [10] A. G. Sahebi and S. S. Pradhan, “On the capacity of abelian group codes over discrete memoryless channels,” in *Information Theory Proceedings (ISIT), 2011 IEEE International Symposium on*. IEEE, 2011, pp. 1743–1747.
- [11] R. Dobrushin, “Asymptotic optimality of group and systematic codes for some channels,” *Theory Probab. Appl.*, vol. 8, no. 1, pp. 47–60, Jan. 1963.
- [12] J.-P. Barthelmy, G. Cohen, and A. Lobstein, *Algorithmic Complexity and Telecommunication Problems*. CRC Press, Jan. 1997.
- [13] D. J. C. MacKay, *Information Theory, Inference and Learning Algorithms*. Cambridge University Press, Oct. 2003.

Bibliography

- [14] R. Gallager, “Low-density parity-check codes,” *IRE Transactions on Information Theory*, vol. 8, no. 1, pp. 21–28, Jan. 1962.
- [15] E. Hof, I. Sason, and S. Shamai, “Performance bounds for nonbinary linear block codes over memoryless symmetric channels,” *IEEE Transactions on Information Theory*, vol. 55, no. 3, pp. 977–996, Mar. 2009.
- [16] A. Kakhaki, H. Abadi, P. Pad, H. Saeedi, F. Marvasti, and K. Alishahi, “Capacity achieving linear codes with random binary sparse generating matrices over the binary symmetric channel,” in *2012 IEEE International Symposium on Information Theory Proceedings (ISIT)*, Jul. 2012, pp. 621–625.
- [17] S. Kumar, A. Young, N. Maoris, and H. Pfister, “A proof of threshold saturation for spatially-coupled LDPC codes on BMS channels,” in *2012 50th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, Oct. 2012, pp. 176–184.
- [18] S. Kudekar, C. Measson, T. Richardson, and R. Urbanke, “Threshold saturation on BMS channels via spatial coupling,” in *2010 6th International Symposium on Turbo Codes and Iterative Information Processing (ISTC)*, Sep. 2010, pp. 309–313.
- [19] I. Sason and B. Shuval, “On universal LDPC code ensembles over memoryless symmetric channels,” *IEEE Transactions on Information Theory*, vol. 57, no. 8, pp. 5182–5202, Aug. 2011.
- [20] J. Lu and J. Moura, “Linear time encoding of LDPC codes,” *IEEE Transactions on Information Theory*, vol. 56, no. 1, pp. 233–249, Jan. 2010.
- [21] M. H. Taghavi, “Decoding linear codes via optimization and graph-based techniques,” Ph.D. dissertation, University of California, San Diego, 2008.
- [22] I. Csiszár, “Linear codes for sources and source networks: Error exponents, universal coding,” *IEEE Transactions on Information Theory*, vol. 28, no. 4, pp. 585–592, Jul. 1982.
- [23] J. L. Massey, “Joint source and channel coding,” 1977.
- [24] G. Caire, S. Shamai, and S. Verdú, “Noiseless data compression with low-density parity-check codes,” *DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, vol. 66, pp. 263–284, 2004.
- [25] J. H. Bae and A. Anastasopoulos, “Capacity-achieving codes for finite-state channels with maximum-likelihood decoding,” *Selected Areas in Communications, IEEE Journal on*, vol. 27, no. 6, pp. 974–984, 2009.

- [26] S. Kudekar and K. Kasai, “Threshold saturation on channels with memory via spatial coupling,” in *2011 IEEE International Symposium on Information Theory Proceedings (ISIT)*, Jul. 2011, pp. 2562–2566.
- [27] M. J. Wainwright and M. I. Jordan, “Graphical models, exponential families, and variational inference,” *Foundations Trends in Machine Learning*, vol. 1, pp. 1–305, 2008.
- [28] P. Clifford, “Markov random fields in statistics,” *Disorder in physical systems*, pp. 19–32, 1990.
- [29] C. M. Bishop, *Pattern Recognition and Machine Learning*. Springer, Oct. 2007.
- [30] D. Koller and N. Friedman, *Probabilistic Graphical Models: Principles and Techniques*. The MIT Press, Jul. 2009.
- [31] B. Potetz and T. S. Lee, “Efficient belief propagation for higher-order cliques using linear constraint nodes,” *Computer Vision and Image Understanding*, vol. 112, no. 1, pp. 39–54, Oct. 2008.
- [32] Y. Weiss and W. Freeman, “On the optimality of solutions of the max-product belief-propagation algorithm in arbitrary graphs,” *IEEE Transactions on Information Theory*, vol. 47, no. 2, pp. 736–744, Feb. 2001.
- [33] N. Shulman, “Communication over an unknown channel via common broadcasting,” Ph.D. dissertation, Tel-Aviv University, 2003.
- [34] N. C. Wormald, “Models of random regular graphs,” *London Mathematical Society Lecture Note Series*, pp. 239–298, 1999.
- [35] A. Aaron, S. D. Rane, E. Setton, and B. Girod, “Transform-domain Wyner-Ziv codec for video,” in *Proceedings of SPIE*, vol. 5308, 2004, pp. 520–528.
- [36] M. Johnson, P. Ishwar, V. Prabhakaran, D. Schonberg, and K. Ramchandran, “On compressing encrypted data,” *IEEE Transactions on Signal Processing*, vol. 52, no. 10, pp. 2992–3006, Oct. 2004.
- [37] D. Schonberg, S. Draper, and K. Ramchandran, “On compression of encrypted images,” in *Proceedings of the International Conference on Image Processing*, 2006, pp. 269–272.
- [38] S. Jalali, S. Verdú, and T. Weissman, “A universal scheme for Wyner-Ziv coding of discrete sources,” *IEEE Transactions on Information Theory*, vol. 56, no. 4, pp. 1737–1750, Apr. 2010.
- [39] D. Donoho, “Compressed sensing,” *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

Bibliography

- [40] L. He and L. Carin, “Exploiting structure in wavelet-based Bayesian compressive sensing,” *IEEE Transactions on Signal Processing*, vol. 57, no. 9, pp. 3488–3497, 2009.
- [41] T. Richardson and R. Urbanke, *Modern Coding Theory*. Cambridge University Press, Mar. 2008.
- [42] M. Wainwright, E. Maneva, and E. Martinian, “Lossy source compression using low-density generator matrix codes: Analysis and algorithms,” *IEEE Transactions on Information Theory*, vol. 56, no. 3, pp. 1351–1368, Mar. 2010.
- [43] V. Chandar, D. Shah, and G. W. Wornell, “A simple message-passing algorithm for compressed sensing,” in *2010 IEEE International Symposium on Information Theory Proceedings (ISIT)*, Jun. 2010, pp. 1968–1972.
- [44] V. B. Chandar, “Sparse graph codes for compression, sensing, and secrecy,” Ph.D. dissertation, Massachusetts Institute of Technology, 2010.
- [45] M. G. Reyes, “Cutset based processing and compression of Markov random fields,” Ph.D. dissertation, The University of Michigan, 2011.
- [46] H. D. Pfister, “On the capacity of finite state channels and the analysis of convolutional accumulate- m codes,” Ph.D. dissertation, University of California, San Diego, 2003.
- [47] G. Colavolpe, “On LDPC codes over channels with memory,” *IEEE Transactions on Wireless Communications*, vol. 5, no. 7, pp. 1757–1766, 2006.
- [48] R. Koetter, A. Singer, and M. Tüchler, “Turbo equalization,” *IEEE Signal Processing Magazine*, vol. 21, no. 1, pp. 67–80, Jan. 2004.
- [49] G. Langdon, Jr., “An introduction to arithmetic coding,” *IBM Journal of Research and Development*, vol. 28, no. 2, pp. 135–149, Mar. 1984.
- [50] J. Ziv and A. Lempel, “A universal algorithm for sequential data compression,” *IEEE Transactions on Information Theory*, vol. 23, no. 3, pp. 337–343, May 1977.
- [51] ———, “Compression of individual sequences via variable-rate coding,” *IEEE Transactions on Information Theory*, vol. 24, no. 5, pp. 530–536, Sep. 1978.
- [52] F. M. J. Willems, Y. Shtarkov, and T. Tjalkens, “The context-tree weighting method: basic properties,” *IEEE Transactions on Information Theory*, vol. 41, no. 3, pp. 653–664, May 1995.
- [53] P. Howard, F. Kossentini, B. Martins, S. Forchhammer, and W. Rucklidge, “The emerging JBIG2 standard,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 8, no. 7, pp. 838–848, Nov. 1998.
- [54] M. Kardar, “8.334 statistical mechanics II notes, lecture 18.” MIT OCW, 2008.

- [55] D. Cimaioni, “A generalized Kac-Ward formula,” *J. Stat. Mech.*, 2010.
- [56] A. Galluccio, “New algorithm for the ising problem: Partition function for finite lattice graphs,” 1999.
- [57] T. J. Richardson, M. A. Shokrollahi, and R. L. Urbanke, “Design of capacity-approaching irregular low-density parity-check codes,” *Information Theory, IEEE Transactions on*, vol. 47, no. 2, pp. 619–637, 2001.
- [58] S. Kudekar, T. Richardson, and R. Urbanke, “Threshold saturation via spatial coupling: Why convolutional LDPC ensembles perform so well over the BEC,” *IEEE Transactions on Information Theory*, vol. 57, no. 2, pp. 803–834, Feb. 2011.
- [59] D. Truhachev, D. G. Mitchell, M. Lentmaier, and D. J. Costello, Jr., “New codes on graphs constructed by connecting spatially coupled chains,” in *Information Theory and Applications Workshop (ITA), 2012*, 2012, pp. 392–397.
- [60] H. Pfister, I. Sason, and R. Urbanke, “Capacity-achieving ensembles for the binary erasure channel with bounded complexity,” *IEEE Transactions on Information Theory*, vol. 51, no. 7, pp. 2352–2379, Jul. 2005.
- [61] H. Pishro-Nik and F. Fekri, “On decoding of low-density parity-check codes over the binary erasure channel,” *IEEE Transactions on Information Theory*, vol. 50, no. 3, pp. 439–454, Mar. 2004.
- [62] I. Andriyanova and J. Tillich, “Designing a good low-rate sparse-graph code,” *IEEE Transactions on Communications*, vol. 60, no. 11, pp. 3181–3190, Nov. 2012.
- [63] D. Burshtein and G. Miller, “An efficient maximum-likelihood decoding of LDPC codes over the binary erasure channel,” *IEEE Transactions on Information Theory*, vol. 50, no. 11, pp. 2837–2844, 2004.
- [64] S. ten Brink, “Designing iterative decoding schemes with the extrinsic information transfer chart,” *AEU Int. J. Electron. Commun.*, vol. 54, no. 6, pp. 389–398, 2000.
- [65] C. Measson, A. Montanari, and R. Urbanke, “Maxwell construction: The hidden bridge between iterative and maximum a posteriori decoding,” *IEEE Transactions on Information Theory*, vol. 54, no. 12, pp. 5277–5307, Dec. 2008.
- [66] I. Csiszár and J. Körner, *Information theory: coding theorems for discrete memoryless systems*. Academic Press, 1981.
- [67] Y. Choi and W. Szpankowski, “Compression of graphical structures: Fundamental limits, algorithms, and experiments,” *IEEE Transactions on Information Theory*, vol. 58, no. 2, pp. 620–638, Feb. 2012.

Bibliography

- [68] Y. Baryshnikov, J. Duda, and W. Szpankowski, “Markov field types and tilings,” in *2014 IEEE International Symposium on Information Theory (ISIT)*, Jun. 2014, pp. 2639–2643.
- [69] M. Li, *An Introduction to Kolmogorov Complexity and Its Applications*. Springer, Nov. 2008.
- [70] R. G. Downey, D. R. Hirschfeldt, and G. L. Forte, “Randomness and reductibility,” in *Mathematical Foundations of Computer Science 2001*. Springer Berlin Heidelberg, Jan. 2001, no. 2136, pp. 316–327.
- [71] G. J. Chaitin, “Information-theoretic computation complexity,” *IEEE Transactions on Information Theory*, vol. 20, no. 1, pp. 10–15, Jan. 1974.
- [72] T. M. Cover and J. A. Thomas, *Elements of information theory*. Wiley, 1991.
- [73] J. G. Cleary and I. Witten, “Data compression using adaptive coding and partial string matching,” *IEEE Transactions on Communications*, vol. 32, no. 4, pp. 396–402, Apr. 1984.
- [74] G. V. Cormack and R. N. S. Horspool, “Data compression using dynamic Markov modelling,” *The Computer Journal*, vol. 30, no. 6, pp. 541–550, 1987.
- [75] J. Rissanen, “A universal data compression system,” *IEEE Transactions on Information Theory*, vol. 29, no. 5, pp. 656–664, Sep. 1983.
- [76] S. Deorowicz, “Universal lossless data compression algorithms,” Ph.D. dissertation, Silesian University of Technology, 2003.
- [77] T. Matsuta, T. Uyematsu, and R. Matsumoto, “Universal Slepian-Wolf source codes using low-density parity-check matrices,” in *2010 IEEE International Symposium on Information Theory Proceedings (ISIT)*, Jun. 2010, pp. 186–190.
- [78] Z. Ghahramani, “Graphical models: parameter learning,” *Handbook of Brain Theory and Neural Networks*, vol. 2, pp. 486–490, 2002.
- [79] P. Abbeel, D. Koller, and A. Y. Ng, “Learning factor graphs in polynomial time and sample complexity,” *The Journal of Machine Learning Research*, vol. 7, pp. 1743–1788, 2006.
- [80] L. Ruschendorf, “Convergence of the iterative proportional fitting procedure,” *Ann. Statist.*, vol. 23, no. 4, pp. 1160–1174, Aug. 1995.
- [81] S. Scholl, F. Kienle, M. Helmling, and S. Ruzika, “ML vs. BP decoding of binary and non-binary LDPC codes,” in *Turbo Codes and Iterative Information Processing (ISTC), 2012 7th International Symposium on*. IEEE, 2012, pp. 71–75.

- [82] N. Chang, “Rate adaptive non-binary LDPC codes with low encoding complexity,” in *2011 Conference Record of the Forty Fifth Asilomar Conference on Signals, Systems and Computers (ASILOMAR)*, Nov. 2011, pp. 664–668.
- [83] Y. Zhao and J. Garcia-Frias, “Data compression of correlated non-binary sources using punctured turbo codes,” in *Data Compression Conference, 2002. Proceedings. DCC 2002*, 2002, pp. 242–251.
- [84] D. H. Schonberg, “Practical distributed source coding and its application to the compression of encrypted data,” Ph.D. dissertation, University of California, Berkeley, 2007.
- [85] Y. Xu, “Potts model and generalizations: Exact results and statistical physics,” Ph.D. dissertation, Stony Brook University, 2012.
- [86] M. Sakalli, W. Pearlman, and M. Farshchian, “SPIHT algorithms using depth first search algorithm with minimum memory usage,” in *2006 40th Annual Conference on Information Sciences and Systems*, Mar. 2006, pp. 1158–1163.
- [87] T. Berger, *Rate Distortion Theory: Mathematical Basis for Data Compression*. Prentice Hall, Oct. 1971.
- [88] A. Wyner and J. Ziv, “The rate-distortion function for source coding with side information at the decoder,” *IEEE Transactions on Information Theory*, vol. 22, no. 1, pp. 1–10, Jan. 1976.
- [89] A. D. Wyner, “The rate-distortion function for source coding with side information at the decoder-II: General sources,” *Information and Control*, vol. 38, no. 1, pp. 60–80, Jul. 1978.
- [90] S. Pradhan and K. Ramchandran, “Geometric proof of rate-distortion function of Gaussian sources with side information at the decoder,” in *IEEE International Symposium on Information Theory, 2000. Proceedings*, 2000, pp. 351–.
- [91] P. Chou, T. Lookabaugh, and R. Gray, “Entropy-constrained vector quantization,” *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 37, no. 1, pp. 31–42, Jan. 1989.
- [92] S. Lloyd, “Least squares quantization in PCM,” *IEEE Transactions on Information Theory*, vol. 28, no. 2, pp. 129–137, Mar. 1982.
- [93] M. J. Sabin and R. Gray, “Global convergence and empirical consistency of the generalized Lloyd algorithm,” *IEEE Transactions on Information Theory*, vol. 32, no. 2, pp. 148–155, Mar. 1986.

Bibliography

- [94] P. Chou, T. Lookabaugh, and R. Gray, "Optimal pruning with applications to tree-structured source coding and modeling," *IEEE Transactions on Information Theory*, vol. 35, no. 2, pp. 299–315, Mar. 1989.
- [95] E. Riskin and R. Gray, "A greedy tree growing algorithm for the design of variable rate vector quantizers," *IEEE Transactions on Signal Processing*, vol. 39, no. 11, pp. 2500–2507, Nov. 1991.
- [96] M. R. Carbonara, J. E. Fowler, and S. C. Ahalt, "Compression of digital video data using artificial neural network differential vector quantization," in *Proceedings of SPIE*, vol. 1709, Sep. 1992, pp. 422–433.
- [97] A. H. Dekker, "Kohonen neural networks for optimal colour quantization," *Network: Computation in Neural Systems*, vol. 5, no. 3, pp. 351–367, 1994.
- [98] A. Aaron, E. Setton, and B. Girod, "Towards practical Wyner-Ziv coding of video," in *2003 International Conference on Image Processing, 2003. ICIP 2003. Proceedings*, vol. 3, Sep. 2003, pp. III–869–72 vol.2.
- [99] D. Rebollo-Monedero and B. Girod, "Generalization of the rate-distortion function for Wyner-Ziv coding of noisy sources in the quadratic-Gaussian case," in *Data Compression Conference, 2005. Proceedings. DCC 2005*. IEEE, 2005, pp. 23–32.
- [100] D. Kubasov, K. Lajnef, and C. Guillemot, "A hybrid encoder/decoder rate control for Wyner-Ziv video coding with a feedback channel," in *IEEE 9th Workshop on Multimedia Signal Processing, 2007. MMSP 2007*, Oct. 2007, pp. 251–254.
- [101] D.-k. He, A. Jagmohan, L. Lu, and V. Sheinin, "Wyner-Ziv video compression using rateless LDPC codes," in *Proc. VCIP*, vol. 8, 2008.
- [102] A. Buzo, A. Gray, Jr., R. Gray, and J. Markel, "Speech coding based upon vector quantization," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 28, no. 5, pp. 562–574, Oct. 1980.
- [103] L. R. Rabiner, M. M. Sondhi, and S. E. Levinson, "Note on the properties of a vector quantizer for LPC coefficients," *Bell System Technical Journal*, vol. 62, no. 8, pp. 2603–2616, Oct. 1983.
- [104] S. Mehrotra, W.-g. Chen, and K. Kotteri, "Low bitrate audio coding using generalized adaptive gain shape vector quantization across channels," in *Acoustics, Speech and Signal Processing, 2009. ICASSP 2009. IEEE International Conference on*. IEEE, 2009, pp. 9–12.
- [105] T. Sreenivas and M. Dietz, "Vector quantization of scale factors in advanced audio coder (AAC)," in *Proceedings of the 1998 IEEE International Conference on Acoustics, Speech and Signal Processing, 1998*, vol. 6, May 1998, pp. 3641–3644 vol.6.

- [106] T. Painter and A. Spanias, “Perceptual coding of digital audio,” *Proceedings of the IEEE*, vol. 88, no. 4, pp. 451–515, 2000.
- [107] A. Skodras, C. Christopoulos, and T. Ebrahimi, “The JPEG 2000 still image compression standard,” *IEEE Signal Processing Magazine*, vol. 18, no. 5, pp. 36–58, Sep. 2001.
- [108] H. Malvar, A. Hallapuro, M. Karczewicz, and L. Kerofsky, “Low-complexity transform and quantization in H.264/AVC,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 13, no. 7, pp. 598–603, Jul. 2003.
- [109] E. Martinian, “Dynamic information and constraints in source and channel coding,” Ph.D. dissertation, Massachusetts Institute of Technology, 2004.
- [110] M. Wainwright and E. Maneva, “Lossy source encoding via message-passing and decimation over generalized codewords of LDGM codes,” in *International Symposium on Information Theory, 2005. ISIT 2005. Proceedings*, Sep. 2005, pp. 1493–1497.
- [111] E. Martinian and M. J. Wainwright, “Analysis of LDGM and compound codes for lossy compression and binning,” *arXiv preprint cs/0602046*, 2006.
- [112] G. Demay, V. Rathi, and L. K. Rasmussen, “Optimality of LDGM-LDPC compound codes for lossy compression of binary erasure source,” in *Information Theory and its Applications (ISITA), 2010 International Symposium on*. IEEE, 2010, pp. 589–594.
- [113] A. Gupta and S. Verdú, “Nonlinear sparse-graph codes for lossy compression,” *IEEE Transactions on Information Theory*, vol. 55, no. 5, pp. 1961–1975, May 2009.
- [114] L. Cappellari, “Lossy source compression of non-uniform binary sources using GQ-LDGM codes,” in *Information Theory Workshop (ITW), 2010 IEEE*. IEEE, 2010, pp. 1–5.
- [115] Z. Sun, M. Shao, J. Chen, K. M. Wong, and X. Wu, “Achieving the rate-distortion bound with low-density generator matrix codes,” *IEEE Transactions on Communications*, vol. 58, no. 6, pp. 1643–1653, Jun. 2010.
- [116] P. Indyk and R. Motwani, “Approximate nearest neighbors: Towards removing the curse of dimensionality,” in *Proceedings of the Thirtieth Annual ACM Symposium on Theory of Computing*. ACM, 1998, pp. 604–613.
- [117] P. T. Boufounos and S. Rane, “Efficient coding of signal distances using universal quantized embeddings,” in *Data Compression Conference (DCC), 2013*. IEEE, 2013, pp. 251–260.
- [118] S. Rane, P. Boufounos, and A. Vetro, “Quantized embeddings: an efficient and universal nearest neighbor method for cloud-based image retrieval,” in *Proceedings of SPIE*, vol. 8856, 2013, p. 885609.

Bibliography

- [119] P. Boufounos, “Universal rate-efficient scalar quantization,” *IEEE Transactions on Information Theory*, vol. 58, no. 3, pp. 1861–1872, Mar. 2012.
- [120] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression*. Springer, Nov. 1991.
- [121] A. Dembo and I. Kontoyiannis, “Source coding, large deviations, and approximate pattern matching,” *IEEE Transactions on Information Theory*, vol. 48, no. 6, pp. 1590–1615, Jun. 2002.
- [122] Z. Sun, T. Tan, Y. Wang, and S. Li, “Ordinal palmprint representation for personal identification,” in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Jun. 2005, pp. 279–284.
- [123] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, “Gradient-based learning applied to document recognition,” *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, Nov. 1998.
- [124] P. Fränti and E. Ageenko, “On the use of context tree for binary image compression,” in *Proceedings of the International Conference on Image Processing*, vol. 3, 1999, pp. 752–756.
- [125] Y. Weiss and W. T. Freeman, “What makes a good model of natural images?” in *Computer Vision and Pattern Recognition, 2007. CVPR’07. IEEE Conference on*, 2007, pp. 1–8.
- [126] D. Zoran and Y. Weiss, “Natural images, Gaussian mixtures and dead leaves,” in *Advances in Neural Information Processing Systems*, 2012, pp. 1736–1744.
- [127] R. L. Rivest, A. Shamir, and L. Adleman, “A method for obtaining digital signatures and public-key cryptosystems,” *Communications of the ACM*, vol. 21, no. 2, pp. 120–126, 1978.
- [128] R. J. McEliece, “A public-key cryptosystem based on algebraic coding theory,” 1978.
- [129] G. Chechik, A. Globerson, N. Tishby, and Y. Weiss, “Information bottleneck for Gaussian variables,” *J. Mach. Learn. Res.*, vol. 6, pp. 165–188, Dec. 2005.
- [130] O. Bursalioglu, G. Caire, and D. Divsalar, “Joint source-channel coding for deep-space image transmission using rateless codes,” *IEEE Transactions on Communications*, vol. 61, no. 8, pp. 3448–3461, Aug. 2013.
- [131] O. Y. Bursalioglu, M. Fresia, G. Caire, and H. V. Poor, “Lossy joint source-channel coding using raptor codes,” *International Journal of Digital Multimedia Broadcasting*, vol. 2008, p. e124685, Sep. 2008.
- [132] J. Gantz and D. Reinsel, “The digital universe in 2020: Big data, bigger digital shadows, and biggest growth in the far east,” *IDC iView: IDC Analyze the Future*, 2012.