A Bibliography

[AgSa05] P.M. Agrawal, A.N.A. Samadh, L.M. Raff, M. Hagan, S. T. Bukkapatnam, and R. Komanduri, "Prediction of molecular-dynamics simulation results using feedforward neural networks: Reaction of a C2 dimer with an activated diamond (100) surface," *The Journal of Chemical Physics* 123, 224711, 2005. (Chapter 24)

[Albe72] A. Albert, *Regression and the Moore-Penrose Pseudoinverse*, New York: Academic Press, 1972. (Chapter 7)

[AmMu97] S. Amari, N. Murata, K.-R. Muller, M. Finke, and H. H. Yang, "Asymptotic Statistical Theory of Overtraining and Cross-Validation," *IEEE Transactions on Neural Networks*, vol. 8, no. 5, 1997. (Chapter 13)

[Ande72] J. A. Anderson, "A simple neural network generating an interactive memory," *Mathematical Biosciences*, vol. 14, pp. 197–220, 1972. (Chapter 1, 15, 21)

[AnRo88] J. A. Anderson and E. Rosenfeld, *Neurocomputing: Foundations of Research*, Cambridge, MA: MIT Press, 1989. (Chapter 1, 10)

[AnSi77] J. A. Anderson, J. W. Silverstein, S. A. Ritz and R. S. Jones, "Distinctive features, categorical perception, and probability learning: Some applications of a neural model," *Psychological Review*, vol. 84, pp. 413–451, 1977. (Chapter 21)

[Barn92] E. Barnard, "Optimization for training neural nets," *IEEE Transactions on Neural Networks*, vol. 3, no. 2, pp. 232–240, 1992. (Chapter 12)

[BaSu83] A. R. Barto, R. S. Sutton and C. W. Anderson, "Neuronlike adaptive elements that can solve difficult learning control problems," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 13, pp. 834–846, 1983. (Chapter 4)

[Batt92] R. Battiti, "First and second order methods for learning: Between steepest descent and Newton's method," *Neural Computation*, vol. 4, no. 2, pp. 141–166, 1992. (Chapter 9, 12)

[Bish91] C. M. Bishop, "Improving the generalization properties of radial basis function neural networks," *Neural Computation*, vol. 3, no. 4, pp. 579-588, 1991. (Chapter 17)

[Bish95] C.M. Bishop, *Neural Networks for Pattern Recognition*, Oxford University Press,1995. (Chapter 22)

[BlDe99] J. A. Blackard and D. J. Dean, "Comparative Accuracies of Artificial Neural Networks and Discriminant Analysis in Predicting Forest Cover Types from Cartographic Variables," Computers and Electronics in Agriculture, vol. 24, pp. 131-151, 1999. (Chapter 26)

[BoJe94] G.E.P. Box, G.M. Jenkins, and G.C. Reinsel, *Time Series Analysis: Forecasting and Control*, 4th Edition, John Wiley & Sons, 2008. (Chapter 22, 27)

[BrLo88] D.S. Broomhead and D. Lowe, "Multivariable function interpolation and adaptive networks," *Complex Systems*, vol.2, pp. 321-355, 1988. (Chapter 17)

[Brog91] W. L. Brogan, *Modern Control Theory*, 3rd Ed., Englewood Cliffs, NJ: Prentice-Hall, 1991. (Chapter 4, 5, 6, 8, 9, 20)

[CaGr87a] G. A. Carpenter and S. Grossberg, "A massively parallel architecture for a self-organizing neural pattern recognition machine," *Computer Vision, Graphics, and Image Processing*, vol. 37, pp. 54–115, 1987. (Chapter 19)

[CaGr87b] G. A. Carpenter and S. Grossberg, "ART2: Self-organization of stable category recognition codes for analog input patterns," *Applied Optics*, vol. 26, no. 23, pp. 4919–4930, 1987. (Chapter 19)

[CaGr90] G. A. Carpenter and S. Grossberg, "ART3: Hierarchical search using chemical transmitters in self-organizing pattern recognition architectures," *Neural Networks*, vol. 3, no. 23, pp. 129–152, 1990. (Chapter 19)

[CaGrMa92] G. A. Carpenter, S. Grossberg, N. Markuzon, J. H. Reynolds and D. B. Rosen, "Fuzzy ARTMAP: A neural network architecture for incremental learning of analog multidimensional maps," *IEEE Transactions on Neural Networks*, vol. 3, pp. 698–713, 1992. (Chapter 19)

[CaGrRe91] G.A. Carpenter, S. Grossberg and J. Reynolds, "ARTMAP: Supervised real-time learning and classification of nonstationary data by a self-organizing neural network," *Neural Networks*, vol. 4, pp. 565–588, 1991. (Chapter 19)

[Char92] C. Charalambous, "Conjugate gradient algorithm for efficient training of artificial neural networks," *IEEE Proceedings*, vol. 139, no. 3, pp. 301–310, 1992. (Chapter 12)

[ChCo91] S. Chen, C.F.N. Cowan, and P.M. Grant, "Orthogonal least squares learning algorithm for radial basis function networks," *IEEE Transactions on Neural Networks*, vol.2, no.2, pp.302-309, 1991. (Chapter 17)

[ChCo92] S. Chen, P. M. Grant, and C. F. N. Cowan, "Orthogonal least squares algorithm for training multioutput radial basis function networks," *Proceedings of the Institute of Electrical Engineers*, vol. 139, Pt. F, no. 6, pp. 378–384, 1992. (Chapter 17)

[ChCh96] S. Chen, E. S. Chng, and K. Alkadhimi, "Regularised orthogonal least squares algorithm for constructing radial basis function networks," *International Journal of Control*, vol. 64, no. 5, pp. 829–837, 1996. (Chapter 17)

[ChCo99] S. Chen, C.F.N. Cowan, and P.M. Grant, "Combined Genetic Algorithm Optimization and Regularized Orthogonal Least Squares Learning for Radial Basis Function Networks," *IEEE Transactions on Neural Networks*, vol.10, no.5, pp.302-309, 1999. (Chapter 17)

[CoGr83] M. A. Cohen and S. Grossberg, "Absolute stability of global pattern formation and parallel memory storage by competitive neural networks," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 13, no. 5, pp. 815–826, 1983. (Chapter 20, 21)

[DARP88] *DARPA Neural Network Study*, Lexington, MA: MIT Lincoln Laboratory, 1988. (Chapter 1)

[DeHa07] O. De Jesús and M. Hagan, "Backpropagation Algorithms for a Broad Class of Dynamic Networks," *IEEE Transactions on Neural Networks*, vol. 18, no. 1, pp., 2007. (Chapter 14)

[Dubi00] D. Dubin, *Rapid Interpretation of EKG's*, Sixth Edition, Tampa, FL: COVER, 2000. (Chapter 25)

[Fahl89] S. E. Fahlman, "Fast learning variations on back-propagation: An empirical study," in *Proceedings of the 1988 Connectionist Models Summer School*, D. Touretzky, G. Hinton and T. Sejnowski, eds., San Mateo, CA: Morgan Kaufmann, pp. 38–51, 1989. (Chapter 12)

[FoGi07] L. Fortuna, P. Giannone, S. Graziani, M. G. Xibilia, "Virtual Instruments Based on Stacked Neural Networks to Improve Product Quality Monitoring in a Refinery," *IEEE Transactions on Instrumentation and Measurement*, vol. 56, no. 1, pp. 95–101, 2007. (Chapter 23)

[FoHa97] D. Foresee and M. Hagan, "Gauss-Newton Approximation to Bayesian Learning," *Proceedings of the 1997 International Joint Conference on Neural Networks*, vol. 3, pp. 1930 - 1935, 1997. (Chapter 13)

[FrSk91] J. Freeman and D. Skapura, *Neural Networks: Algorithms, Applications, and Programming Techniques*, Reading, MA: Addison-Wesley, 1991. (Chapter 16)

[Gill81] P. E. Gill, W. Murray and M. H. Wright, *Practical Optimization*, New York: Academic Press, 1981. (Chapter 8, 9)

[GoLa98] C. Goutte and J. Larsen, "Adaptive Regularization of Neural Networks Using Conjugate Gradient," *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, vol. 2, pp. 1201-1204, 1998. (Chapter 13)

[GrMi89] S. Grossberg, E. Mingolla and D. Todorovic, "A neural network architecture for preattentive vision," *IEEE Transactions on Biomedical Engineering*, vol. 36, no. 1, pp. 65–84, Jan. 1989. (Chapter 18)

[Gros67] S. Grossberg, "Nonlinear difference-differential equations in prediction and learning theory," *Proceedings of the National Academy of Sciences*, vol. 58, pp. 1329–1334, 1967. (Chapter 21)

[Gros68] S. Grossberg, "Some physiological and biochemical consequences of psychological postulates," *Proceedings of the National Academy of Sciences*, vol. 60, pp. 758–765, 1968. (Chapter 15)

[Gros76] S. Grossberg, "Adaptive pattern classification and universal recoding: I. Parallel development and coding of neural feature detectors," *Biological Cybernetics*, vol. 23, pp. 121–134, 1976. (Chapter 1, 18, 19)

[Gros80] S. Grossberg, "How does the brain build a cognitive code?" *Psychological Review*, vol. 88, pp. 375–407, 1980. (Chapter 1)

[Gros82] S. Grossberg, *Studies of Mind and Brain*, Boston: D. Reidel Publishing Co., 1982. (Chapter 15, 18, 19)

[Gros90] S. Grossberg, "Neural networks: From foundations to applications," Short-Course Notes, Boston University, Boston, MA, May 6–11, 1990. (Chapter 18)

[HaBo07] L. Hamm, B. W. Brorsen and M. T. Hagan, "Comparison of Stochastic Global Optimization Methods to Estimate Neu-

ral Network Weights," *Neural Processing Letters*, vol. 26, no. 3, December 2007. (Chapter 22)

[HaDe02] M. Hagan, H. Demuth, O. De Jesus, "An Introduction to the Use of Neural Networks in Control Systems," *International Journal of Robust and Nonlinear Control*, vol. 12, no. 11, pp. 959-985, 2002. (Chapter 27)

[HaMe94] M. T. Hagan and M. Menhaj, "Training feedforward networks with the Marquardt algorithm," *IEEE Transactions on Neural Networks*, vol. 5, no. 6, pp. 989–993, 1994. (Chapter 12)

[HeBa99] S. Hettich and S. D. Bay, *The UCI KDD Archive* [http://kdd.ics.uci.edu], Irvine, CA: University of California, Department of Information and Computer Science, 1999. (Chapter 26)

[Hebb 49] D. O. Hebb, *The Organization of Behavior*, New York: Wiley, 1949. (Chapter 1, 7, 15)

[Hech90] R. Hecht-Nielsen, *Neurocomputing*, Reading, MA: Addison-Wesley, 1990. (Chapter 16)

[HeOh97] B. Hedén, H. Öhlin, R. Rittner, L. Edenbrandt, "Acute Myocardial Infarction Detected in the 12-Lead ECG by Artificial Neural Networks," *Circulation*, vol. 96, pp. 1798–1802, 1997. (Chapter 22)

[Himm72] D. M. Himmelblau, *Applied Nonlinear Programming*, New York: McGraw-Hill, 1972. (Chapter 8, 9)

[Hopf82] J. J. Hopfield, "Neural networks and physical systems with emergent collective computational properties," *Proceedings of the National Academy of Sciences*, vol. 79, pp. 2554–2558, 1982. (Chapter 1, 21)

[Hopf84] J. J. Hopfield, "Neurons with graded response have collective computational properties like those of two-state neurons," *Proceedings of the National Academy of Sciences*, vol. 81, pp. 3088–3092, 1984. (Chapter 21)

[HoTa85] J. J. Hopfield and D. W. Tank, " 'Neural' computation of decisions in optimization problems," *Biological Cybernetics*, vol. 52, pp. 141–154, 1985. (Chapter 21)

[HoSt89] K. M. Hornik, M. Stinchcombe and H. White, "Multilayer feedforward networks are universal approximators," *Neural Networks*, vol. 2, no. 5, pp. 359–366, 1989. (Chapter 11)

[Hube88] D. H. Hubel, *Eye, Brain, and Vision*, New York: Scientific American Library, 1988. (Chapter 18)

[Jaco88] R. A. Jacobs, "Increased rates of convergence through learning rate adaptation," *Neural Networks*, vol. 1, no. 4, pp 295–308, 1988. (Chapter 12)

[John01] G. L. Johnson, "Contributions to the comparative anatomy of the mammalian eye, chiefly based on ophthalmoscopic examination," *Philosophical Transactions of the Royal Society of London*, Series B., vol. 194, pp. 1–82, Plate 1, 1901. (Chapter 18)

[Joll02] I.T. Jolliffe, *Principal Component Analysis*, Springer Series in Statistics, 2nd ed., Springer, NY, 2002. (Chapter 22)

[Koho72] T. Kohonen, "Correlation matrix memories," *IEEE Transactions on Computers*, vol. 21, pp. 353–359, 1972. (Chapter 1, 15, 21)

[Koho87] T. Kohonen, *Self-Organization and Associative Memory*, 2nd Ed., Berlin: Springer-Verlag, 1987. (Chapter 15, 16)

[Koho93] T. Kohonen, "Things you haven't heard about the Self-Organizing Map," *Proceedings of the International Conference on Neural Networks* (ICNN), San Francisco, pp. 1147-1156, 1993. (Chapter 26)

[Koho95] T. Kohonen, Self-Organizing Map, 2nd ed., Springer-Verlag, Berlin, 1995. (Chapter 26)

[LaSa67] J. P. LaSalle, "An invariance principle in the theory of stability," in *Differential Equations and Dynamic Systems*, J. K. Hale and J. P. Lasalle, eds., New York: Academic Press, pp. 277–286, 1967. (Chapter 20)

[LeCu85] Y. Le Cun, "Une procedure d'apprentissage pour reseau a seuil assymetrique," *Cognitiva*, vol. 85, pp. 599–604, 1985. (Chapter 11)

[LeCu98] Y. LeCun, L. Bottou, G. B. Orr, K.-R. Mueller, "Efficient BackProp," *Lecture Notes in Computer Science*, vol. 1524, 1998. (Chapter 22)

[Leib90] D. Lieberman, *Learning, Behavior and Cognition*, Belmont, CA: Wadsworth, 1990. (Chapter 15)

[LiMi89] J. Li, A. N. Michel and W. Porod, "Analysis and synthesis of a class of neural networks: Linear systems operating on a closed hypercube," *IEEE Transactions on Circuits and Systems*, vol. 36, no. 11, pp. 1405–1422, November 1989. (Chapter 21)

[Lowe89] D. Lowe, "Adaptive radial basis function nonlinearities, and the problem of generalization," *Proceedings of the First IEE International Conference on Artificial Neural Networks*, pp. 171 - 175, 1989. (Chapter 17)

[MacK92] D. J. C. MacKay, "Bayesian Interpolation," *Neural Computation*, vol. 4, pp. 415-447, 1992. (Chapter 13)

[MaNe99] J.R. Magnu and H. Neudecker, *Matrix Differential Calculus*, John Wiley & Sons, Ltd., Chichester, 1999. (Chapter 14)

[MaGa00] E. A. Maguire, D. G. Gadian, I. S. Johnsrude, C. D. Good, J. Ashburner, R. S. J. Frackowiak, and C. D. Frith, "Navigation-related structural change in the hippocampi of taxi drivers," Proceedings of the National Academy of Sciences, vol. 97, no. 8, pp. 4398-4403, 2000. (Chapter 1)

[McPi43] W. McCulloch and W. Pitts, "A logical calculus of the ideas immanent in nervous activity," *Bulletin of Mathematical Biophysics*, vol. 5, pp. 115–133, 1943. (Chapter 1, 4, 21)

[Mill90] A.J. Miller, *Subset Selection in Regression*. Chapman and Hall, N.Y., 1990. (Chapter 17)

[Mill93] M.F. Miller, "A scaled conjugate gradient algorithm for fast supervised learning," Neural Networks, vol. 6, pp. 525-533, 1993. (Chapter 24)

[MoDa89] J. Moody and C.J. Darken, "Fast Learning in Networks of Locally-Tuned Processing Units," *Neural Computation*, vol. 1, pp. 281–294, 1989. (Chapter 17)

[Moll93] M. Moller, "A scaled conjugate gradient algorithm for fast supervised learning," *Neural Networks*, vol. 6, pp. 525-533, 1993. (Chapter 22)

[MoMa01] G.B. Moody, R.G. Mark, and A.L. Goldberger, "PhysioNet: a Web-based resource for the study of physiologic signals," *IEEE Transactions on Engineering in Medicine and Biology*, vol. 20, no. 3, pp: 70-75, 2001. (Chapter 25)

[MiPa69] M. Minsky and S. Papert, *Perceptrons*, Cambridge, MA: MIT Press, 1969. (Chapter 1, 4)

[NaMu97] Narendra, K.S.; Mukhopadhyay, S., "Adaptive control using neural networks and approximate models," *IEEE Transactions on Neural Networks*, vol. 8, no. 3, pp. 475 - 485, 1997. (Chapter 27)

[NaPa90] K. S. Narendra and K. Parthasarathy, "Identification and control of dynamical systems using neural networks," *IEEE Transactions on Neural Networks*, vol. 1, no. 1, pp. 4–27, 1990. (Chapter 27)

[NgWi90] D. Nguyen and B. Widrow, "Improving the learning speed of 2-layer neural networks by choosing initial values of the adaptive weights," *Proceedings of the IJCNN*, vol. 3, pp. 21–26, July 1990. (Chapter 12, 22)

[OrHa00] M. J. Orr, J. Hallam, A. Murray, and T. Leonard, "Assessing rbf networks using delve," IJNS, 2000. (Chapter 17)

[Park85] D. B. Parker, "Learning-logic: Casting the cortex of the human brain in silicon," Technical Report TR-47, Center for Computational Research in Economics and Management Science, MIT, Cambridge, MA, 1985. (Chapter 11)

[PaSa93] J. Park and I.W. Sandberg, "Universal approximation using radial-basis-function networks," *Neural Computation*, vol. 5, pp. 305-316, 1993. (Chapter 17)

[PeCo93] M. P. Perrone and L. N. Cooper, "When networks disagree: Ensemble methods for hybrid neural networks," in *Neural Networks for Speech and Image Processing*, R. J. Mammone, Ed., Chapman-Hall, pp. 126-142, 1993. (Chapter 22)

[PhHa13] M. Phan and M. Hagan, "Error Surface of Recurrent Networks," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 24, no. 11, pp. 1709 - 1721, October, 2013. (Chapter 14)

[Powe87] M.J.D. Powell, "Radial basis functions for multivariable interpolation: a review," *Algorithms for Approximation*, pp. 143-167, Oxford, 1987. (Chapter 17)

[PuFe97] G.V. Puskorius and L.A. Feldkamp, "Extensions and enhancements of decoupled extended Kalman filter training," *Proceedings of the 1997 International Conference on Neural Networks*, vol. 3, pp. 1879-1883, 1997. (Chapter 22)

[RaMa05] L.M. Raff, M. Malshe, M. Hagan, D.I. Doughan, M.G. Rockley, and R. Komanduri, "*Ab initio* potential-energy surfaces for complex, multi-channel systems using modified novelty sampling and feedforward neural networks," *The Journal of Chemical Physics*, vol. 122, 2005. (Chapter 22, 24)

[RiIr90] A. K. Rigler, J. M. Irvine and T. P. Vogl, "Rescaling of variables in back propagation learning," *Neural Networks*, vol. 3, no. 5, pp 561–573, 1990. (Chapter 12)

[Rose58] F. Rosenblatt, "The perceptron: A probabilistic model for information storage and organization in the brain," *Psychological Review*, vol. 65, pp. 386–408, 1958. (Chapter 1, 4)

[Rose61] F. Rosenblatt, *Principles of Neurodynamics*, Washington DC: Spartan Press, 1961. (Chapter 4)

[RuHi86] D. E. Rumelhart, G. E. Hinton and R. J. Williams, "Learning representations by back-propagating errors," *Nature*, vol. 323, pp. 533–536, 1986. (Chapter 11)

[RuMc86] D. E. Rumelhart and J. L. McClelland, eds., *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, vol. 1, Cambridge, MA: MIT Press, 1986. (Chapter 1, 11, 16)

[Sarle95] W. S. Sarle, "Stopped training and other remedies for overfitting," In *Proceedings of the 27th Symposium on Interface*, 1995. (Chapter 13)

[Scal85] L. E. Scales, *Introduction to Non-Linear Optimization*, New York: Springer-Verlag, 1985. (Chapter 8, 9, 12)

[ScSm99] B. Schölkopf, A. Smola, K.-R. Muller, "Kernel Principal Component Analysis," in B. Schölkopf, C. J. C. Burges, A. J. Smola (Eds.), *Advances in Kernel Methods-Support Vector Learning*, MIT Press Cambridge, MA, USA, pp. 327-352, 1999. (Chapter 22)

[Shan90] D. F. Shanno, "Recent advances in numerical techniques for large-scale optimization," in *Neural Networks for Control*, Miller, Sutton and Werbos, eds., Cambridge, MA: MIT Press, 1990. (Chapter 12)

[SjLj94] J. Sjoberg and L. Ljung, "Overtraining, regularization and searching for minimum with application to neural networks," Linkoping University, Sweden, Tech. Rep. LiTH-ISY-R-1567, 1994. (Chapter 13)

[SlLi91] J.-J. E. Slotine and W. Li, *Applied Nonlinear Control*, Englewood Cliffs, NJ: Prentice-Hall, 1991. (Chapter 20)

[StDo84] W. D. Stanley, G. R. Dougherty and R. Dougherty, *Digital Signal Processing*, Reston VA: Reston Publishing Co., 1984. (Chapter 10)

[Stra76] G. Strang, *Linear Algebra and Its Applications*, New York: Academic Press, 1980. (Chapter 5, 6)

[TaHo86] D. W. Tank and J. J. Hopfield, "Simple 'neural' optimization networks: An A/D converter, signal decision circuit and a linear programming circuit," *IEEE Transactions on Circuits and Systems*, vol. 33, no. 5, pp. 533–541, 1986. (Chapter 21)

[TeSi00] J. B. Tenenbaum, V. de Silva, J. C. Langford, "A Global Geometric Framework for Nonlinear Dimensionality Reduction," *Science*, vol. 290, pp. 2319-2323, 2000. (Chapter 25)

[Tikh63] A. N. Tikhonov, "The solution of ill-posed problems and the regularization method," *Dokl. Acad. Nauk USSR*, vol. 151, no. 3, pp. 501-504, 1963. (Chapter 13)

[Toll90] T. Tollenaere, "SuperSAB: Fast adaptive back propagation with good scaling properties," *Neural Networks*, vol. 3, no. 5, pp. 561–573, 1990. (Chapter 12)

[vanT75] H. F. J. M. van Tuijl, "A new visual illusion: Neonlike color spreading and complementary color induction between subjective contours," *Acta Psychologica*, vol. 39, pp. 441–445, 1975. (Chapter 18)

[VoMa88] T. P. Vogl, J. K. Mangis, A. K. Zigler, W. T. Zink and D. L. Alkon, "Accelerating the convergence of the backpropagation method," *Biological Cybernetics*, vol. 59, pp. 256–264, Sept. 1988. (Chapter 12)

[vond73] C. von der Malsburg, "Self-organization of orientation sensitive cells in the striate cortex," *Kybernetic*, vol. 14, pp. 85–100, 1973. (Chapter 18)

[WaVe94] C. Wang, S. S. Venkatesh, and J. S. Judd, "Optimal Stopping and Effective Machine Complexity in Learning," *Advances in Neural Information Processing Systems*, J. D. Cowan, G. Tesauro, and J. Alspector, Eds., vol. 6, pp. 303-310, 1994. (Chapter 13)

[Werbo74] P. J. Werbos, "Beyond regression: New tools for prediction and analysis in the behavioral sciences," Ph.D. Thesis, Harvard University, Cambridge, MA, 1974. Also published as *The Roots of Backpropagation*, New York: John Wiley & Sons, 1994. (Chapter 11)

[Werb90] P. J. Werbos, "Backpropagation through time: What it is and how to do it," *Proceedings of the IEEE*, vol. 78, pp. 1550–1560, 1990. (Chapter 14)

[WeTe84] J. F. Werker and R. C. Tees, “Cross-language speech perception: Evidence for perceptual reorganization during the first year of life,” Infant Behavior and Development, vol. 7, pp. 49-63, 1984. (Chapter 1)

[WhSo92] D. White and D. Sofge, eds., *Handbook of Intelligent Control*, New York:Van Nostrand Reinhold, 1992. (Chapter 4)

[WiHo60] B. Widrow, M. E. Hoff, “Adaptive switching circuits,”*1960 IRE WESCON Convention Record*, New York: IRE Part 4, pp. 96–104, 1960. (Chapter 1, 10)

[WiSt 85] B. Widrow and S. D. Stearns, *Adaptive Signal Processing*, Englewood Cliffs, NJ: Prentice-Hall, 1985. (Chapter 10)

[WiWi 88] B. Widrow and R. Winter, “Neural nets for adaptive filtering and adaptive pattern recognition,” *IEEE Computer Magazine*, March 1988, pp. 25–39. (Chapter 10)

[WiZi89] R. J. Williams and D. Zipser, “A learning algorithm for continually running fully recurrent neural networks,” *Neural Computation*, vol. 1, pp. 270–280, 1989. (Chapter 14)