

References

- [1] J. Felsenstein. *Inferring Phylogenies*. Sinauer, Sunderland, MA, 2004.
- [2] D. Gusfield. *Algorithms on Strings, Trees, and Sequences*. Cambridge University Press, New York, 1997.
- [3] C. Semple and M. Steel. *Phylogenetics*. Oxford University Press, New York, 2003.
- [4] A. W. F. Edwards and L. L. Cavalli-Sforza. The reconstruction of evolution. *Annals of Human Genetics*, 27:105–106 (1963).
- [5] H. Taketomi, Y. Ueda, and N. Gō. Studies on protein folding, unfolding and fluctuations by computer simulation. *International Journal of Peptide and Protein Research*, 7:445–459 (1975).
- [6] H. S. Chan and K. A. Dill. Energy landscape and the collapse dynamics of homopolymers. *Journal of Chemical Physics*, 99:2116–2127 (1993).
- [7] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. Equation of state calculation by fast computing machines. *Journal of Chemical Physics*, 21:1087–1092 (1953).
- [8] R. J. Folz and J. I. Gordon. Computer-assisted predictions of signal peptidase processing sites. *Biochemistry and Biophysics Research Communications*, 146:870–877 (1987).
- [9] K.-C. Chou. Review: Prediction of HIV protease cleavage sites in proteins. *Analytical Biochemistry*, 233:1–14 (1996).
- [10] P. Saxová, S. Buus, S. Brunak, and C. Keşmir. Predicting proteasomal cleavage sites: A comparison of available methods. *International Immunology*, 15(7):781–787 (July 2003).
- [11] L. Excoffier and P. E. Smouse. Using allele frequencies and geographic subdivision to reconstruct gene trees within a species: Molecular variance parsimony. *Genetics*, 136:343–359 (1994).
- [12] H.-J. Bandelt, P. Forster, B. C. Sykes, and M. B. Richards. Mitochondrial portraits of human populations using median networks. *Genetics*, 141:743–753 (1995).
- [13] H.-J. Bandelt, P. Forster, and A. Röhl. Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution*, 16(1):37–48 (1999).
- [14] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein. *Introduction to Algorithms*, 2nd ed. MIT Press, Cambridge, MA, 2001.
- [15] A. V. Aho, J. E. Hopcroft, J. D. Ullman. *The Design and Analysis of Computer Algorithms*. Addison-Wesley, Reading, MA, 1974.
- [16] D. Knuth. *The Art of Computer Programming*, vol. 1, *Fundamental Algorithms*, 3rd ed. Addison-Wesley, Reading, MA, 1997.
- [17] D. Kozen. *The Design and Analysis of Algorithms*. Springer, New York, 1992.
- [18] J. B. Kruskal. On the shortest spanning subtree of a graph and the traveling salesman problem. *Proceedings of the American Mathematical Society*, 7:48–50 (1956).
- [19] R. C. Prim. Shortest connection networks and some generalizations. *Bell System Technical Journal*, 36:1389–1401 (1957).

- [20] E. W. Dijkstra. A note on two problems in connexion with graphs. *Numerische Mathematik*, 1:269–271 (1959).
- [21] R. Bellman. On a routine problem. *Quarterly Journal of Applied Mathematics*, 16:87–90 (1958).
- [22] L. R. Ford and D. R. Fulkerson. *Flows in Networks*. Princeton University Press, Princeton, NJ, 1962.
- [23] R. W. Floyd. Algorithm 97 (shortest path). *Communications of the ACM*, 5(6):345 (1962).
- [24] D. B. Johnson. Efficient algorithms for shortest paths in sparse networks. *Journal of the ACM*, 24(1):1–13 (1977).
- [25] J. Edmonds and R. M. Karp. Theoretical improvements in the algorithmic efficiency for network flow problems. *Journal of the ACM*, 19:248–264 (1972).
- [26] H. W. Kuhn. The Hungarian method for the assignment problem. *Naval Research Logistics Quarterly*, 2:83–97 (1955).
- [27] H. N. Gabow. *Implementation of Algorithms for Maximum Matching on Non-bipartite Graphs*. Ph.D. thesis, Department of Computer Science, Stanford University, 1974.
- [28] S. Needleman and C. Wunsch. A general method applicable to the search for similarities in the amino acid sequences of two proteins. *Journal of Molecular Biology*, 48:443–453 (1970).
- [29] T. Smith and M. Waterman. Identification of common molecular subsequences. *Journal of Molecular Biology*, 147(1):195–197 (1981).
- [30] P. Weiner. Linear pattern matching algorithms. *Proceedings of the 14th IEEE Symposium on Switching and Automata Theory*, pp. 1–11. IEEE Computer Society Press, Silver Spring, MD, 1973.
- [31] E. Ukkonen. On-line construction of suffix trees. *Algorithmica*, 14(3):249–260 (1995).
- [32] G. C. L. Johnson, L. Esposito, B. J. Barratt, A. Smith, J. Heward, G. Di Genova, H. Ueda, H. Cordell, I. Eaves, F. Dudbridge, R. C. Twells, F. Payne, W. Hughes, S. Nutland, H. Stevens, P. Carr, E. Tuomilehto-Wolf, J. Tuomilehto, S. C. Gough, D. G. Clayton, and J. A. Todd. Haplotype tagging for the identification of common disease genes. *Nature Genetics*, 29:233–237 (2001).
- [33] M. R. Garey and D. S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. W. H. Freeman, San Francisco, 1979.
- [34] P. Crescenzi and V. Kann, eds. *A compendium of NP optimization problems*. 1998. <http://www.nada.kth.se/~viggo/problemlist/compendium.html>.
- [35] G. Ausiello, P. Crescenzi, G. Gambosi, V. Kann, A. Marchetti-Spaccamela, and M. Protasi. *Complexity and Approximation*. Springer, New York, 1999.
- [36] S. Cook. The complexity of theorem-proving procedures. *Journal of the ACM*, 18:4–18 (1971).
- [37] R. M. Karp. Reducibility among combinatorial problems. In *Complexity of Computer Computations: Proceedings*, R. E. Miller and J. W. Thatcher, eds., pp. 85–103. Plenum Press, New York, 1972.
- [38] G. B. Dantzig, W. O. Blattner, and M. R. Rao. All shortest routes from a fixed origin in a graph. In *Theory of Graphs: International Symposium*, pp. 85–90. Gordon and Breach, New York, 1967.
- [39] O. Goldschmidt and D. S. Hochbaum. Polynomial algorithm for the k-cut problem. In *Proceedings of the 29th Annual Symposium on the Foundations of Computer Science*, pp. 444–451. Association for Computing Machinery, White Plains, NY, 1988.
- [40] M. Yannakakis. *The Node Deletion Problem for Hereditary Properties*. Technical report. Princeton University, Princeton, NJ: 1978.
- [41] M. Yannakakis. Node- and edge-deletion NP-complete problems. In *Proceedings of the 10th Annual ACM Symposium on the Theory of Computing*, pp. 253–264. Association for Computing Machinery, San Diego, CA, 1978.
- [42] J. M. Lewis. On the complexity of the maximum subgraph problem. In *Proceedings of the 10th Annual ACM Symposium on the Theory of Computing*, pp. 265–274. 1978.
- [43] D. Maier. The complexity of some problems on subsequences and supersequences. *Journal of the ACM*, 25:322–336 (1978).
- [44] D. Maier and J. A. Storer. *A Note on the Complexity of the Superstring Problem*. Technical report. Princeton University, Princeton, NJ: 1977.

- [45] D. J. Rosenkrantz, R. E. Stearns, and P. M. Lewis. An analysis of several heuristics for the traveling salesman problem. *SIAM Journal on Computing*, 6:563–581 (1977).
- [46] N. Christofides. *Worst-Case Analysis of a New Heuristic for the Traveling Salesman Problem*. Technical report. Carnegie Mellon University, Pittsburgh, PA, 1976.
- [47] S. Arora. Polynomial time approximation scheme for Euclidean TSP and other geometric problems. In *Proceedings of the 37th Annual IEEE Symposium on Foundations of Computer Science*, pp. 2–11. IEEE Computer Society, Burlington, VT, 1996.
- [48] B. Monien and E. Speckenmeyer. Ramsey numbers and an approximation algorithm for the vertex cover problem. *Acta Informatica*, 22:115–123 (1985).
- [49] R. Bar-Yehuda and S. Even. A local-ratio theorem for approximating the weighted vertex cover problem. In *Analysis and Design of Algorithms for Combinatorial Problems*, G. Ausiello and M. Lucertiori, eds., pp. 27–46. North-Holland, Amsterdam, 1985.
- [50] E. Halperin. Improved approximation algorithms for the vertex cover problem in graphs and hypergraphs. In *Proceedings of the 11th Annual ACM-SIAM Symposium on Discrete Algorithms*. Association for Computing Machinery, San Francisco, CA, 2000.
- [51] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi. Optimization by simulated annealing. *Science*, 4598:671–680 (May 13, 1983).
- [52] V. Cerny. A thermodynamical approach to the travelling salesman problem: An efficient simulation algorithm. *Journal of Optimization Theory and Applications*, 45:41–51 (1985).
- [53] N. A. Barricelli. Esempi numerici di processi di evoluzione. *Methodos*, 6:45–68 (1954).
- [54] M. Sipser. *Introduction to the Theory of Computation*, 2nd ed. Thompson Course Technology, Boston, 2005.
- [55] C. H. Papadimitriou. *Computational Complexity*. Addison-Wesley, Reading, MA, 1994.
- [56] D. S. Hochbaum, ed. *Approximation Algorithms for NP-Hard Problems*. PWS, Boston, 1997.
- [57] L. Stryer. *Biochemistry*, 4th ed. W. H. Freeman, New York, 1995.
- [58] A. M. Maxam and W. Gilbert. A new method for sequencing DNA. *Proceedings of the National Academy of Sciences USA*, 74(2):560–564 (1977).
- [59] F. Sanger and A. R. Coulson. A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *Journal of Molecular Biology*, 94:441–448 (1975).
- [60] L. M. Smith, J. Z. Sanders, R. J. Kaiser, P. Hughes, C. Dodd, C. R. Connell, C. Heiner, S. B. H. Kent, and L. E. Hood. Fluorescence detection in automated DNA sequence analysis. *Nature*, 321:674–679 (June 12, 1986).
- [61] H. Drossman, J. A. Luckey, A. J. Kostichka, J. D’Cunha, and L. M. Smith. High-speed separations of DNA sequencing reactions by capillary electrophoresis. *Analytical Chemistry*, 62:900–903 (1990).
- [62] W. Bender, P. Spierer, and D. S. Hogness. Chromosomal walking and jumping to isolate DNA from the Ace and rosy loci and the bithorax complex in *Drosophila melanogaster*. *Journal of Molecular Biology*, 168(1):17–33 (July 1983).
- [63] International Human Genome Sequencing Consortium. Initial sequencing and analysis of the human genome. *Nature*, 409:860–921 (2001).
- [64] P. Pevzner. *Computational Molecular Biology: An Algorithmic Approach*. MIT Press, Cambridge, MA, 2000.
- [65] N. C. Jones and P. Pevzner. *An Introduction to Bioinformatics Algorithms*. MIT Press, Cambridge, MA, 2004.
- [66] W. Bains and G. C. Smith. A novel method for nucleic acid sequence determination. *Journal of Theoretical Biology*, 135:303–307 (1988).
- [67] E. Southern. United Kingdom Patent Application GB8810400. 1988.
- [68] I. P. Lysov, V. L. Florent’ev, A. A. Khorlin, K. R. Khrapko, and V. V. Shik. Determination of the nucleotide sequence of DNA using hybridization with oligonucleotides. A new method. *Doklady Akademii Nauk SSSR, Ser. Biol.*, 303:1508–1511 (1988).

- [69] R. Drmanac, I. Labat, I. Brukner, and R. Crkvenjakov. Sequencing of megabase plus DNA by hybridization: Theory of the method. *Genomics*, 4:114–128 (1989).
- [70] R. Drmanac et al. DNA sequence determination by hybridization: A strategy for efficient large-scale sequencing. *Science*, 260(5114):1649–1652 (1993).
- [71] P. A. Pevzner, *l*-tuple sequencing: Computer analysis. *Journal of Biomolecular Structure and Dynamics*, 7(1):63–73 (1989).
- [72] R. M. Idury and M. S. Waterman. A new algorithm for DNA sequence assembly. *Journal of Computational Biology*, 2(2):291–306 (1995).
- [73] R. D. Fleischmann et al. Whole-genome random sequencing and assembly of *Haemophilus influenzae* rd. *Science*, 269(5223):496–498, 507–512 (1995).
- [74] A. Edwards, H. Voss, P. Rice, A. Civitello, J. Stegemann, C. Schwager, J. Zimmerman, H. Erfle, C. T. Caskey, and W. Ansorge. Automated DNA sequencing of the human HPRT locus. *Genomics*, 6:593–608 (1990).
- [75] J. L. Weber and E. W. Myers. Whole-genome shotgun sequencing. *Genome Research*, 7(5):401–409 (May 1997).
- [76] E. W. Myers et al. A whole-genome assembly of *Drosophila*. *Science*, 287(5461):2196–2204 (March 24, 2000).
- [77] D. Huson et al. Design of a compartmentalized shotgun assembler for the human genome. *Bioinformatics*, 17, S1:S132–S139 (2001).
- [78] S. Batzoglou et al. ARACHNE: A whole-genome shotgun assembler. *Genome Research*, 12:177–189 (2002).
- [79] M. Akeson, D. Branton, J. J. Kasianowicz, E. Brandin, and D. W. Deamer. Microsecond time-scale discrimination among polycytidylic acid, polyadenylic acid, and polyuridylic acid as homopolymers or as segments within single RNA molecules. *Biophysical Journal*, 77(6):3227–3233 (1999).
- [80] D. W. Deamer and M. Akeson. Nanopores and nucleic acids: Prospects for ultrarapid sequencing. *Trends in Biotechnology*, 18(4):147–151 (2000).
- [81] J. Shendure, R. D. Mitra, C. Varma, and G. M. Church. Advanced sequencing technologies: Methods and goals. *Nature Reviews Genetics*, 5:335–344 (May 2004).
- [82] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. *Numerical Recipes in C: The Art of Scientific Computing*. Cambridge University Press, Cambridge, 1992.
- [83] J. E. Dennis and R. B. Schnabel. *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*. Society for Industrial and Applied Mathematics, Philadelphia, 2006.
- [84] A. P. Ruszczyński. *Nonlinear Optimization*. Princeton University Press, Princeton, NJ, 2006.
- [85] R. Fletcher. *Practical Methods of Optimization*. Wiley, New York, 2000.
- [86] L. N. Trefethen and D. Bau. *Numerical Linear Algebra*. Society for Industrial and Applied Mathematics, Philadelphia, 1997.
- [87] *Wikipedia: The Free Encyclopedia*. <http://www.wikipedia.org>.
- [88] I. Newton. *De Analysi per Aequationes Numero Terminorum Infinitas*. William Jones, London, 1711.
- [89] J. Wallis. *A Treatise of Algebra Both Historical and Practical*. J. Playford for R. Davis, London, 1685.
- [90] J. Raphson. *Analysis Aequationum Universalis*. A. and I. Churchill, London, 1690.
- [91] M. R. Hestenes and E. Stiefel. Method of conjugate gradients for solving linear systems. *Journal of Research of the National Bureau of Standards*, 49:409–436 (1952).
- [92] K. Levenberg. A method for the solution of certain non-linear problems in least squares. *Quarterly of Applied Mathematics*, 2:164–168 (1944).
- [93] D. Marquardt. An algorithm for least-squares estimation of nonlinear parameters. *SIAM Journal of Applied Mathematics*, 11:431–441 (1963).
- [94] S. G. Nash and A. Sofer. *Linear and Nonlinear Programming*. McGraw-Hill, New York, 1996.

- [95] G. Dantzig. Maximization of a linear function of variables subject to linear inequalities. In Cowles Commission for Research in Economics, *Activity Analysis of Production and Allocation*. Wiley, New York, 1951.
- [96] L. G. Khachiyan. A polynomial algorithm in linear programming. *Doklady Akademii Nauk SSSR*, 244:1093–1096 (1979).
- [97] A. S. Nemirovskii and D. B. Iudin. Optimization methods adapting to the “significant” dimension of the problem. *Automatika i Telemekhanika*, 38:75–87 (1977).
- [98] N. Karmarkar. A new polynomial time algorithm for linear programming. *Combinatorica*, 4:373–395 (1984).
- [99] *Computational Infrastructure for Operations Research (COIN-OR)*. <http://www.coin-or.org/>.
- [100] M. Berkelaar, J. Dirks, K. Eikland, and P. Notebaert. *lp_solve*. <http://lpsolve.sourceforge.net>.
- [101] A. Makhorin. *GLPK (GNU Linear Programming Kit)*. <http://www.gnu.org/software/glpk>.
- [102] S. M. Ross. *Introduction to Probability Models*, 8th ed. Academic Press, San Diego, 2003.
- [103] S. M. Ross. *Simulation*, 3rd ed. Academic Press, San Diego, 2002.
- [104] D. Knuth. *The Art of Computer Programming*. Addison-Wesley, Reading, MA, 1981.
- [105] R. Eckhardt. Stan Ulam, John von Neumann, and the Monte Carlo method. *Los Alamos Science*, 15:131–137 (1987).
- [106] J. von Neumann. Various techniques used in connection with random digits. In *Monte Carlo Methods*, pp. 36–38. Applied Math Series, 12. National Bureau of Standards, Washington, DC, 1951.
- [107] G. E. P. Box and M. E. Müller. A note on the generation of random normal deviates. *Annals of Mathematical Statistics*, 29:610–611 (1958).
- [108] Y. A. Rozanov. *Probability Theory: A Concise Course*. Dover, New York, 1977.
- [109] D. T. Gillespie. *Markov Processes*. Academic Press, Boston, 1992.
- [110] D. J. Wilkinson. *Stochastic Modelling for Systems Biology*, Chapman & Hall/CRC, Boca Raton, FL, 2006.
- [111] A. A. Markov. Extension of the law of large numbers to dependent quantities. *Izvestia Fiziko-Matematicheskikh Obshchestva Kazan University*, 15:135–136 (1906).
- [112] L. Wasserman. *All of Statistics*. Springer, New York, 2004.
- [113] S. Geman and D. Geman. Stochastic relaxation, Gibbs distribution, and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6:721–741 (1984).
- [114] G. M. Torrie and J. P. Valleau. Monte-Carlo free energy estimates using non-Boltzmann sampling. *Chemical Physics Letters*, 28:578–581 (1974).
- [115] A. J. Sinclair and M. R. Jerrum. Approximate counting, uniform generation and rapidly mixing Markov chains. *Information and Computation*, 82(1):93–133 (1989).
- [116] A. Sinclair. Improved bounds for mixing rates of Markov chains and multicommodity flow. *Combinatorics, Probability and Computing*, 1:351–370 (1992).
- [117] S. M. Ross. *Stochastic Processes*. Wiley, New York, 1983.
- [118] J. L. Doob. *Stochastic Processes*. Wiley, New York, 1953.
- [119] D. R. Cox and H. D. Miller. *The Theory of Stochastic Processes*. Wiley, New York, 1965.
- [120] S. Karlin. *A First Course in Stochastic Processes*. Academic Press, New York, 1968.
- [121] S. I. Resnick. *Adventures in Stochastic Processes*. Birkhäuser, Boston, 1992.
- [122] N. G. van Kampen. *Stochastic Processes in Physics and Chemistry*. North-Holland, New York, 1981.
- [123] G. B. Benedek and F. M. H. Villars. *Physics with Illustrative Examples from Medicine and Biology: Statistical Physics*, 2nd ed. Springer, New York, 2000.
- [124] D. Graur and W.-H. Li. *Fundamentals of Molecular Evolution*, 2nd ed. Sinauer, Sunderland, MA, 2000.

- [125] M. Nordborg. Coalescent theory. In *Handbook of Statistical Genetics*, pp. 179–212. Wiley, Chichester, UK, 2001.
- [126] D. L. Hartl and A. G. Clark. *Principles of Population Genetics*, 3rd ed. Sinauer, Sunderland, MA, 1997.
- [127] T. H. Jukes and C. R. Cantor. Evolution of protein molecules. In *Mammalian Protein Metabolism*, vol. 3, pp. 21–123. Academic Press, New York, 1969.
- [128] M. Kimura. A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution*, 16:111–120 (1980).
- [129] J. F. C. Kingman. The coalescent. *Stochastic Processes and Applications*, 13:235–248 (1982).
- [130] D. C. Rapaport. The event scheduling problem in molecular dynamics simulation. *Journal of Computational Physics*, 32:184–201 (1980).
- [131] R. Brown. Calendar queues: A fast $O(1)$ priority queue implementation for the simulation event set problem. *Communications of the ACM*, 31(10):1220–1227 (1998).
- [132] J. Stoer and R. Bulirsch. *Introduction to Numerical Analysis*, 2nd ed. Springer, New York, 1996.
- [133] P. K. Kytte and M. R. Schäferkotter. *Handbook of Computational Methods for Integration*. Chapman & Hall/CRC, Boca Raton, FL, 2004.
- [134] R. W. Hamming. *Numerical Methods for Scientists and Engineers*, 2nd ed. Dover, New York, 1986.
- [135] R. J. LeVeque. *Numerical Methods for Conservation Laws*, 2nd ed. Birkhäuser Boston, 1992.
- [136] D. Gottlieb and S. A. Orszag. *Numerical Analysis of Spectral Methods: Theory and Applications*. Society for Industrial and Applied Mathematics, Philadelphia, 1977.
- [137] J. Strickwerda. *Finite Difference Schemes and Partial Differential Equations*. Wordsworth & Brooks, Pacific Grove, CA, 1989.
- [138] A. Turing. The chemical basis of morphogenesis. *Philosophical Transactions of the Royal Society of London*, B237:37–72 (1952).
- [139] D. J. Higham. An algorithmic introduction to numerical simulation of stochastic differential equations. *SIAM Review*, 43(3):525–546 (2001).
- [140] P. E. Kloeden and E. Platen. *Numerical Solution of Stochastic Differential Equations*. Springer, Berlin, 1999.
- [141] P. E. Protter. *Stochastic Integration and Differential Equations*. Springer, New York, 2004.
- [142] D. Lamberton and B. Lapeyre. *Introduction to Stochastic Calculus Applied to Finance*. CRC Press, Boca Raton, FL, 2000.
- [143] B. Øksendal. *Stochastic Differential Equations: An Introduction with Applications*, 6th ed. Springer, New York, 2005.
- [144] P. Mendes. GEPASI: A software package for modeling the dynamics, steady states and control of biochemical and other systems. *Computer Applications in the Biosciences [now Bioinformatics]*, 9:563–571 (1993).
- [145] M. Tomita, K. Hashimoto, K. Takahashi, T. Simon Shimizu, Y. Matsuzaki, F. Miyoshi, K. Saito, S. Tanida, K. Yugi, J. C. Venter, and C. A. Hutchinson. E-cell: Software environment for whole-cell simulation. *Bioinformatics*, 15:72–84 (1999).
- [146] P. Mendes. Biochemistry by numbers: Simulation of biochemical pathways with Gepasi 3. *Trends in Biochemical Sciences*, 22:361–363 (1997).
- [147] J. Schaff, C. C. Fink, B. Slepcheckno, J. H. Carson, and L. M. Loew. A general computational framework for modeling cellular structure and function. *Biophysical Journal*, 73:1135–1146 (September 1997).
- [148] P. Mendes and D. B. Kell. MEG (Model Extender for GEPASI): A program for the modelling of complex, heterogeneous, cellular systems. *Bioinformatics*, 17(3):288–289 (2001).
- [149] D. T. Gillespie. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *Journal of Computational Physics*, 22(4):403–434 (1976).

- [150] N. Le Novère and T. S. Shimizu. StochSim: Modelling of stochastic biomolecular processes. *Bioinformatics*, 17(6): 575–576 (2001).
- [151] M. A. Gibson and J. Bruck. Efficient exact stochastic simulation of chemical systems with many species and channels. *Journal of Physical Chemistry*, A104: 1876–1889 (2000).
- [152] Y. Cao, H. Li, and L. Petzold. Efficient formulation of the stochastic simulation algorithm for chemically reacting systems. *Journal of Chemical Physics*, 121(9): 4059–4067 (September 2004).
- [153] F. Jamalyaria, R. Rohlf, and R. Schwartz. Queue-based method for efficient simulation of biological self-assembly systems. *Journal of Computational Physics*, 204: 100–120 (2005).
- [154] J. R. Stiles, D. Van Helden, T. M. Bartol, Jr., E. E. Salpeter, and M. M. Salpeter. Miniature end-plate current rise times $< 100 \mu\text{s}$ from improved dual recordings can be modeled with passive acetylcholine diffusion from a synaptic vesicle. *Proceedings of the National Academy of Sciences USA*, 93: 5747–5752 (June 1996).
- [155] J. R. Stiles and T. M. Bartol, Jr. Monte Carlo methods for simulating realistic synaptic microphysiology using MCell. In *Computational Neuroscience: Realistic Modeling for Experimentalists*, pp. 87–127. CRC Press, Boca Raton, FL, 2001.
- [156] A. B. Stundzia and C. J. Lumsden. Stochastic simulation of coupled reaction-diffusion processes. *Journal of Computational Physics*, 127: 196–207 (1996).
- [157] J. Elf, A. Dončić, M. Ehrenberg. Mesoscopic reaction-diffusion in intracellular signaling. *Proceedings of SPIE*, 5510: 114–124 (2003).
- [158] M. Ander, P. Beltrao, B. Di Ventura, J. Ferkinghoff-Borg, M. Foglierini, A. Kaplan, C. Lemerle, I. Tomas-Oliveira, and L. Serrano. SmartCell, a framework to simulate cellular processes that combines stochastic approximation with diffusion and localisation: Analysis of simple networks. *IEEE Systems Biology*, 1: 129–138 (2004).
- [159] D. T. Gillespie. Approximate accelerated stochastic simulation of chemically reacting systems. *Journal of Chemical Physics*, 115(4): 1716–1733 (July 22, 2001).
- [160] S. Hoops, S. Sahle, R. Gauges, C. Lee, J. Pahle, N. Simus, M. Singhal, L. Xu, P. Mendes, and U. Kummer. COPASI—a COMplex PATHway SIMulator. *Bioinformatics*, 22: 3067–3074 (2006).
- [161] K. Takahashi, K. Kaizu, B. Hu, M. Tomita. A multi-algorithm, multi-timescale method for cell simulation. *Bioinformatics*, 20(4): 538–546 (2004).
- [162] M. L. Blinov, J. R. Faeder, B. Goldstein, and W. S. Hlavacek. BioNetGen: Software for rule-based modeling of signal transduction based on the interaction of molecular domains. *Bioinformatics*, 20: 3289–3291 (June 24, 2004).
- [163] L. Lok and R. Brent. Automatic generation of cellular reaction networks with Molecuizer 1.0. *Nature Biotechnology*, 23: 131–136 (2005).
- [164] T. Zhang, R. Rohlf, and R. Schwartz. Implementation of a discrete event simulator for biological self-assembly. In *Proceedings of the Winter Simulation Conference*, pp. 2223–2231. Institute for Operations Research and the Management Sciences, Orlando, FL, 2005.
- [165] D. L. Ermak. A computer simulation of charged particles in solution. I. Technique and equilibrium properties. *Journal of Chemical Physics*, 62(10): 4189–4196 (1975).
- [166] J. S. van Zon and P. R. ten Wolde. Simulating biochemical networks at the particle level and in time and space: Green's function reaction dynamics. *Physical Review Letters*, 94: 128103 (2005).
- [167] M. Hucka, A. Finney, H. M. Sauro, H. Bolouri, J. C. Doyle, H. Kitano, et al. The Systems Biology Markup Language (SBML): A medium for representation and exchange of biochemical network models. *Bioinformatics*, 19(4): 524–531 (2003).
- [168] M. Hucka, A. Finney, S. Hoops, S. Keating, and N. Le Novère. *Systems Biology Markup Language (SBML) level 2: Structures and facilities for model definitions*. 2007. <http://www.sbml.org>.
- [169] A. A. Cuellar, P. F. Nielsen, D. P. Bullivant, D. Nickerson, W. Hedley, M. Nelson, and C. M. Lloyd. *CellML specification 1.1. draft*. 2003. <http://www.cellml.org>.
- [170] C. N. Lloyd, M. D. B. Halstead, and P. F. Nielsen. CellML: Its future, present, and past. *Progress in Biophysics and Molecular Biology*, 85(2–3): 433–450 (June–July 2004).

- [171] *E-cell*. 2007. <http://www.e-cell.org>.
- [172] *National Resource for Cell Analysis and Modeling*. 2007. <http://www.vcell.org>.
- [173] *COPASI: Complex Pathway Simulator*. 2007. <http://www.copasi.org/tiki-index.php>.
- [174] Center for Quantitative Biological Simulation. *Microphysiology Gateway*. 2007. <http://www.mcell.psc.edu>.
- [175] C. P. Fall, E. S. Marland, J. M. Wagner, and J. T. Tyson, eds. *Computational Cell Biology*. Springer, New York, 2002.
- [176] T. Niu, Z. S. Qin, X. Xu, and J. S. Liu. Bayesian haplotype inference for multiple linked single-nucleotide polymorphisms. *American Journal of Human Genetics*, 70:157–169 (2002).
- [177] T. M. Mitchell. *Machine Learning*. McGraw-Hill, New York, 1997.
- [178] L. E. Baum, T. Petries, G. Soules, and N. Weiss. A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains. *Annals of Mathematical Statistics*, 41:164–171 (1970).
- [179] A. P. Dempster, N. M. Laird, and D. B. Rubin. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society*, B39:1–38 (1977).
- [180] E. L. L. Sonnhammer, S. R. Eddy, and R. Durbin. Pfam: A comprehensive database of protein domain families based on seed alignments. *Proteins*, 28:405–420 (1997).
- [181] L. R. Rabiner. A tutorial on hidden Markov models and selected applications in speech recognition. *Proceedings of the IEEE*, 77:257–286 (February 1989).
- [182] R. Durbin, S. Eddy, A. Krogh, and G. Mitchison. *Biological Sequence Analysis*. Cambridge University Press, Cambridge, 2003.
- [183] D. W. Mount. *Bioinformatics: Sequence and Genome Analysis*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 2001.
- [184] A. J. Viterbi. Error bounds for convolutional codes and an asymptotically optimum decoding algorithm. *IEEE Transactions on Information Theory*, 13(2):260–269 (1967).
- [185] L. E. Baum and J. A. Eagon. An inequality with applications to statistical estimation for probabilistic functions of Markov processes and to a model for ecology. *Bulletin of the American Meteorological Society*, 73:360–363 (1967).
- [186] C. Burge and S. Karlin. Prediction of complete gene structures in human genomic DNA. *Journal of Molecular Biology*, 268:78–94 (1997).
- [187] M. Delorenzi and T. Speed. An HMM model for coiled-coil domains and a comparison with PSSM-based predictions. *Bioinformatics*, 8(4):617–625 (2002).
- [188] A. Krogh, M. Brown, I. S. Mian, K. Sjölander, and D. Haussler. Hidden Markov models in computational biology: Applications to protein modeling. *Journal of Molecular Biology*, 235:1501–1531 (1994).
- [189] S. R. Eddy. Hidden Markov models. *Current Opinion in Structural Biology*, 6:361–365 (1996).
- [190] S. R. Eddy. Profile hidden Markov models. *Bioinformatics*, 14:755–763 (1998).
- [191] S. R. Eddy. What is a hidden Markov model? *Nature Biotechnology*, 22:1315–1316 (2004).
- [192] G. Strang. *Linear Algebra and Its Applications*, 3rd ed. Harcourt Brace Jovanovich, San Diego, 1988.
- [193] Y. Saad. *Iterative Methods for Sparse Linear Systems*. PWS, Boston, 1996.
- [194] P. Kuzmic. Program DYNAFIT for the analysis of enzyme kinetic data: Application to HIV protease. *Analytical Biochemistry*, 237:260–273 (1996).
- [195] D. F. Walnut. *An Introduction to Wavelet Analysis*. Birkhäuser, Boston, 2004.
- [196] J. S. Walker. *A Primer on Wavelets and Their Scientific Applications*. Chapman & Hall/CRC, Boca Raton, FL, 1999.
- [197] M. B. Eisen, P. T. Spellman, P. O. Brown, and D. Botstein. Cluster analysis and display of genome-wide expression patterns. *Proceedings of the National Academy of Sciences USA*, 95(25):14863–14868 (December 8, 1998).

- [198] U. Alon, N. Barkai, D. A. Notterman, K. Gish, S. Ybarra, D. Mack, and A. J. Levine. Broad patterns of gene expression revealed by clustering analysis of tumor and normal colon tissues probed by oligonucleotide arrays. *Proceedings of the National Academy of Sciences USA*, 96:6745–6750 (June 1999).
- [199] E. Mjolsness, T. Mann, R. Castaño, and B. Wold. *From Coexpression to Coregulation: An Approach to Inferring Transcriptional Regulation Among Gene Classes from Large-Scale Expression Data*. Technical report. Jet Propulsion Laboratory, Pasadena, CA, 1999.
- [200] J. MacQueen. Some methods for classification and analysis of multivariate observation. In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, vol. 1, pp. 281–297. University of California Press, Berkeley, 1967.
- [201] T. Kohonen. The self-organizing map. *Proceedings of the IEEE*, 78:1464–1480 (1990).
- [202] P. Tamayo, D. Slonim, J. Mesirov, Q. Zhu, S. Kitareewan, E. Dmitrovsky, E. S. Lander, and T. R. Golub. Interpreting patterns of gene expression with self-organizing maps: Methods and application to hematopoietic differentiation. *Proceedings of the National Academy of Sciences USA*, 96:2907–2912 (March 16, 1999).
- [203] N. Friedman, M. Linial, I. Nachman, and D. Pe’er. Using Bayesian networks to analyze expression data. *Journal of Computational Biology*, 7:601–620 (2000).
- [204] J. Yu, V. A. Smith, P. P. Wang, A. J. Hartemink, and E. D. Jarvis. Advances to Bayesian network inference for generating causal networks from observational biological data. *Bioinformatics*, 20:3594–3603 (2004).
- [205] D. Madigan and J. York. Bayesian graphical models for discrete data. *International Statistical Review*, 63:215–232 (1995).
- [206] B. Efron and R. J. Tibshirani. *An Introduction to the Bootstrap*. Chapman & Hall, New York, 1993.
- [207] N. Friedman, M. Goldszmidt, and A. Wyner. Data analysis with Bayesian networks: A bootstrap approach. In *Proceedings of the Fifteenth Conference on Uncertainty in Artificial Intelligence (UAI)*, pp. 206–215. Morgan Kaufmann, San Francisco, 1999.
- [208] N. Friedman, I. Nachman, and D. Pe’er. Learning Bayesian network structure from massive datasets: The “sparse candidate” algorithm. In *Proceedings of the Fifteenth Conference on Uncertainty in Artificial Intelligence (UAI)*, pp. 196–205. Morgan Kaufmann, San Francisco, 1999.
- [209] S. Tavazoie, J. D. Hughes, M. J. Campbell, R. J. Cho, and G. M. Church. Systematic determination of genetic network architecture. *Nature Genetics*, 22:281–285 (1999).
- [210] T. Ideker, O. Ozier, B. Schwikowski, and A. F. Siegel. Discovering regulatory and signalling circuits in molecular interaction networks. *Bioinformatics*, 18(supp. 12002):S233–S240 (2002).
- [211] Z. Tu, L. Wang, M. N. Arbeitman, T. Chen, and F. Sun. An integrative approach for causal gene identification and gene regulatory pathway inference. *Bioinformatics*, 22(142006):e489–e496 (2006).
- [212] T. S. Gardner, D. di Bernardo, D. Lorenz, and J. J. Collins. Inferring genetic networks and identifying compound mode of action via expression profiling. *Science*, 301(5629):102–105 (July 4, 2003).
- [213] D. di Bernardo, M. J. Thompson, T. S. Gardner, S. E. Chobot, E. L. Eastwood, A. P. Wojtovich, S. J. Elliott, S. E. Schaus, and J. J. Collins. Chemogenomic profiling on a genome-wide scale using reverse-engineered gene networks. *Nature Biotechnology*, 23(3):377–383 (March 3, 2005).
- [214] I. Nachman, A. Regev, and N. Friedman. Inferring quantitative models of regulatory networks from expression data. *Bioinformatics*, 20(supp. 1):i248–i256 (2004).
- [215] P. D’haeseleer, S. Liang, and R. Somogyi. Genetic network inference: From co-expression clustering to reverse engineering. *Bioinformatics*, 16(8):707–726 (2000).
- [216] H. De Jong. Modeling and simulation of genetic regulatory systems: A review. *Journal of Computational Biology*, 9:67–103 (2002).
- [217] M. Bansal, V. Belcastro, A. Ambesi-Impiombato, and D. di Bernardo. How to infer gene networks from expression profiles. *Molecular Systems Biology*, 3(78):1–10 (2007).
- [218] H. Causton, J. Quackenbush, and A. Brazma. *Microarray Gene Expressions Data Analysis: A Beginner’s Guide*. Blackwell Science, Malden, MA, 2003.

- [219] D. P. Berrar, W. Dubitzky, and M. Granzow, eds. *A Practical Approach to Microarray Data Analysis*. Kluwer Academic, Boston, 2003.
- [220] A. Zhang. *Advanced Analysis of Gene Expression Microarray Data*. World Scientific, Singapore, 2006.
- [221] P. D'haeseleer. How does gene expression clustering work? *Nature Biotechnology*, 23:1499–1501 (2005).
- [222] A. K. Jain, M. N. Murty, and P. J. Flynn. Data clustering: A review. *ACM Computing Surveys*, 31(3):264–323 (1999).
- [223] P. Congdon. *Applied Bayesian Modelling*. Wiley, Chichester, UK, 2003.
- [224] A. Gelman, J. B. Carlin, H. S. Stern, and D. B. Rubin. *Bayesian Data Analysis*. Chapman & Hall/CRC Press, Boca Raton, FL, 2003.
- [225] R. E. Neapolitan. *Learning Bayesian Networks*. Pearson Prentice-Hall, Upper Saddle River, NJ, 2004.
- [226] R. C. Campbell. *Statistics for Biologists*, 3rd ed. Cambridge University Press, New York, 1989.
- [227] L. Fleck. *Genesis and Development of a Scientific Fact*. University of Chicago Press, Chicago, 1979.
- [228] T. Kuhn. *The Structure of Scientific Revolutions*. University of Chicago Press, Chicago, 1962.
- [229] P. Kosso. *Reading the Book of Nature: An Introduction to the Philosophy of Science*. Cambridge University Press, New York, 1992.
- [230] A. Rosenberg. *The Philosophy of Science: A Contemporary Introduction*. Routledge, New York, 2000.