8.581 Entropy, information, and the brain

References (revised 6 January 1998); papers marked by * are included in the course reader.

Introduction & foundations.

- C. E. Shannon, A mathematical theory of communciation, *Bell Sys. Tech. J.* **27,** 379–423, 623–656 (1948). Reprinted in Shannon and Weaver (1949).
- *C. E. Shannon, Communication in the presence of noise, *Proc. I. R. E.* **37,** 10–21 (1949).
- C. E. Shannon and W. Weaver, *The Mathematical Theory of Communication* (University of Illinois Press, Urbana, 1949).
- L. Brillouin, *Science and Information Theory* (Academic Press, New York, 1962).
- *R. R. de Ruyter van Steveninck and S. B. Laughlin, The rate of information transfer at graded–potential synapses, *Nature* **379,** 642–645 (1996).
- F. Rieke, D. Warland, R. de Ruyter van Steveninck, and W. Bialek, *Spikes: Exploring the Neural Code* (MIT Press, Cambridge, 1997).

Electrical & chemical signals in the brain.

- A. L. Hodgkin and A. F. Huxley, Currents carried by sodium and potassium ions through the membrane of the giant axon of *Loligo, J. Physiol.* **116,** 449–472 (1952a).
- A. L. Hodgkin and A. F. Huxley, The components of membrane conductance in the giant axon of *Loligo, J. Physiol.* **116,** 473–496 (1952b).
- A. L. Hodgkin and A. F. Huxley, The dual effect of membrane potential on sodium conductance in the giant axon of *Loligo, J. Physiol.* **116,** 497–506 (1952c).
- *A. L. Hodgkin and A. F. Huxley, A quantitative description of membrane current and its application to conduction and excitation in nerve, *J. Physiol.* **117,** 500–544 (1952d).
- B. Katz, *Nerve, Muscle, and Synapse.* (McGraw–Hill, New York, 1966). Note: this is a brief and delightful book by one of the great masters.

- *E. Neher and B. Sakmann, Single-channel currents recorded from membranes of denervated frog muscle fibres, *Nature* **260,** 799–802 (1976).
- B. Hille, *Ionic Channels of Excitable Membranes,* 2d ed. (Sinauer Associates, Sunderland MA, 1992).
- L. Y. Jan and Y. N. Jan (1994). Potassium channels and their evolving gates, *Nature* **371,** 119–122.
- D. Johnston and S.-M. Wu, *Foundations of Cellular Neurophysiology* (MIT Press, Cambridge, 1995).

Reproducibility and variability in neural responses.

- D. MacKay and W. S. McCulloch, The limiting information capacity of a neuronal link, *Bull. Math. Biophys.* **14,** 127–135 (1952).
- H. B. Barlow, Summation and inhibition in the frog's retina, *J. Physiol. (Lond.)* **119,** 69–88 (1953).
- S. W. Kuffler, Discharge patterns and functional organization of mammalian retina, *J. Neurophys.* **16,** 37–68 (1953).
- *D. H. Hubel and T. N. Wiesel, Receptive fields, binocular interaction and functional architecture in the cat's visual cortex, *J. Physiol. (Lond.)***160,** 106–154 (1962).
- *Z. F. Mainen and T. J. Sejnowski, Reliability of spike timing in neocortical neurons, *Science* **268,** 1503–1506 (1995).
- *R. de Ruyter van Steveninck, G. Lewen, S. Strong, R. Koberle, and W. Bialek, Reproducibility and variability in neural spike trains, *Science* **275,** 1805–1808 (1997).
- *M. J. Berry II, D. K. Warland, and M. Meister, The structure and precision of retinal spike trains, *Proc. Nat. Acad. Sci. (USA)* **94,** 5411–5416 (1997).
- *S. P. Strong, R. Koberle, R. de Ruyter van Steveninck, and W. Bialek, Entropy and information in neural spike trains, *Phys. Rev. Lett.* in press.
- *N. Brenner, S. P. Strong, R. Koberle, R. de Ruyter van Steveninck, and W. Bialek, The economy of impulses and the rigidity of spike trains, preprint.
- G. Buracas, A. Zador, M. DeWeese, and T. Albright, Efficient discrimination of temporal patterns by motion sensitive neurons in primate visual cortex, preprint.

Decoding the information in spike trains.

- *R. de Ruyter van Steveninck and W. Bialek, Real–time performance of a movement sensitive neuron in the blowfly visual system: Coding and information transfer in short spike sequences, *Proc. R. Soc. Lond. Ser. B* **234,** 379–414 (1988).
- *W. Bialek and A. Zee, Coding and computation with neural spike trains, *J. Stat. Phys.* **59,** 103–115 (1990).
- *W. Bialek, F. Rieke, R. R. de Ruyter van Steveninck, and D. Warland, Reading a neural code, *Science* **252,** 1854–1857 (1991).
- H. S. Seung and H. Sompolinsky, Simple models for the reading neuronal population codes, *Proc. Nat. Acad. Sci. (USA)* **90,** 10749–10753 (1993).
- F. Rieke, D. Warland, and W. Bialek, Coding efficiency and information rates in sensory neurons, *Europhys. Lett.* **22,** 151–156 (1993).
- E. Salinas and L. Abbott, Vector reconstruction from firing rates, *J. Comp. Neurosci.* **1,** 89–107 (1994).
- F. Rieke, D. Bodnar, and W. Bialek, Naturalistic stimuli increase the rate and efficiency of information transmission by primary auditory neurons, *Proc. R. Soc. Lond. Ser. B* **262,** 259–265 (1995).
- D. K. Warland, P. Reinagel, and M. Meister, Decoding visual information from a population of retinal ganglion cells, *J. Neurophysiol.* in press.

Note: Much of what is discussed in the preceding references is also described in *Spikes: Exploring the Neural Code.*

- E. N. Brown, L. M. Frank, D. Tang, M. C. Quirk, and M. A. Wilson, A statistical paradigm for neural spike train decoding applied to position prediction from ensemble firing patterns of rat hippocampal place cells, preprint.
- W. Bialek and R. R. de Ruyter van Steveninck, What do motion sensitive visual neurons compute?, in preparation.

The search for efficient representations.

- *H. B. Barlow, Possible principles underlying the transformation of sensory messages. In *Sensory Communication,* W. Rosenblith, ed., pp. 217–234 (MIT Press, Cambridge, 1961).[1]

[1] This is a widely cited paper, usually taken as the origin of the ideas about the efficiency of neural representation that are discussed in the lecture. In fact it appears in the proceedings

- H. B. Barlow, Single units and sensation: A neuron doctrine for perception, *Perception* **1,** 371–394 (1972).

- H. B. Barlow, The absolute efficiency of perceptual decisions, *Philos. Trans. R. Soc. Lond. Ser. B* **290,** 71–82 (1980).

- H. B. Barlow, Critical limiting factors in the design of the eye and visual cortex, *Proc. R. Soc. Lond. Ser. B* **212,** 1–34 (1981).

- S. B. Laughlin, A simple coding procedure enhances a neuron's information capacity, *Z. Naturforsch.* **36c,** 910–912 (1981).

- D. Field, Relations between the statistics of natural images and the response properties of cortical cells, *J. Opt. Soc. Am. A* **4,** 2379–2394 (1987).

- *K. Rose, E. Gurewitz, and G. Fox, Statistical mechanics and phase transitions in clustering, *Phys. Rev. Lett.* **65,** 945–948 (1990).

- J. J. Atick and A. N. Redlich, Towards a theory of early visual processing, *Neural Comp.* **2,** 308–320 (1990).

- J. J. Atick, Z. Li, and A. N. Redlich, Understanding retinal color coding from first principles, *Neural Comp.* **4,** 559–572 (1992).

- *J. J. Atick, Could information theory provide an ecological theory of sensory processing?. In *Princeton Lectures on Biophysics,* W. Bialek, ed., pp. 223–289 (World Scientific, Singapore, 1992).

- *F. Pereira, N. Tishby, and L. Lee, Distributional clustering of english words, *30th Annual Meeting of the Association for Computational Linguistics,* pp. 183–190 (1993).

- D. L. Ruderman and W. Bialek, Statistics of natural images: Scaling in the woods, *Phys. Rev. Lett.* **73,** 814–817 (1994).

- *N. Barkai, H. S. Seung, and H. Sompolinsky, Scaling laws in learning of classification tasks, *Phys. Rev. Lett.* **70,** 3167–3170 (1993). See also N. Barkai and H. Sompolinsky, Statistical mechanics of the maximum likelihood density estimation, *Phys. Rev. E* **50,** 1766–1769 (1994).

of a conference, but there was an earlier conference (24–27 November, 1958, at the National Physical Laboratory of the U. K.) where Barlow presented a similar view with even a more broad context. In rummaging through files at the last minute, I couldn't find the 1961 paper (the book is surely in several MIT libraries) but I discovered that my colleague Rob de Ruyter had a copy of the earlier paper, apparently taken from a set of preprints distributed before the conference; this is what I include in the reader, and it remains, I think, stimulating reading almost 40 years later. There is a published version, courtesy of H. M. Stationary Office (!), and I shall try to find the proper reference.

- M. DeWeese, *Optimization Principles for the Neural Code* (Dissertation, Princeton University, 1995).
- S. Becker, Mutual information maximization: Models of cortical self-organization, *Network* **7,** 7–31 (1996).
- S. Smirnakis, M. Berry, D. Warland, W. Bialek and M. Meister, Adaptation of retinal processing to image contrast and spatial scale, *Nature* **386,** 69–73 (1997).

Gathering information and learning rules.

- E. Levin, N. Tishby, and S. A. Solla, A statistical approach to learning and generalization in layered neural networks, *Proc. I. E. E. E.* **78,** 1568–1574 (1990).
- *H. S. Seung, H. Sompolinsky, H., and N. Tishby, Statistical mechanics of learning from examples, *Phys. Rev. A* **45,** 6056–6091 (1992).
- D. J. C. MacKay, Bayesian interpolation, *Neural Comp.* **4,** 415–447 (1992).
- D. Haussler, M. Kearns, H. S. Seung, and N. Tishby, Rigorous learning curve bounds from statistical mechanics, *Machine Learning* **25,** 195–236 (1996).
- *W. Bialek, C. G. Callan, and S. P. Strong, Field theories for learning probability distributions, *Phys. Rev. Lett.* **77,** 4693–4697 (1996).
- *V. Balasubramanian, Statistical inference, Occam's razor, and statistical mechanics on the space of probability distributions, *Neural Comp.* **9,** 349–368 (1997).

Predictability, complexity, and meaning (maybe).

- *C. E. Shannon, Prediction and entropy of printed English, *Bell Sys. Tech. J.* **30,** 50–64 (1951).
- *P. Grassberger, Toward a quantitative theory of self-generated complexity, *Int. J. Theor. Phys.* **25,** 907–938 (1986).
- C. H. Bennett, How to define complexity in physics, and why, in *Complexity, Entropy and the Physics of Information,* W. H. Zurek, ed., pp. 137–148 (Addison-Wesley, Redwood City, 1990).
- M. Li and P. Vitányi, *An Introduction to Kolmogorov Complexity and its Applications* (Springer-Verlag, New York, 1993).

- A. S. Weigend and N. A. Gershenfeld, eds., *Time series prediction: Forecasting the future and understanding the past* (Addison–Wesley, Reading MA, 1994).
- *W. Bialek, Notes on predictive information, unpublished.