

Chapter 4

Coda

The previous two chapters described the main results of this thesis on the auditory system characterization by taking top-down and bottom-up approaches, respectively. Here I will present a general discussion on how we could approach the sensory coding problems. Following a brief summary of the approaches and the results of this thesis, Section 4.1 will discuss further questions and future challenges in close relation to the topics of this thesis. As a concluding remark, Section 4.2 will brainstorm and argue what needs to be kept in mind for conducting theoretical studies on the brain.

4.1 Conclusions and Future Challenges

Sensory signal processing in the brain can be conceptually divided into several stages (see also Section 1.2); sensors at the very periphery, preprocessing at the subcortical levels, signal representations and perceptions at the cortical level, and output motor actions. This dissertation mainly targeted the representation—i.e., the cortical—level, and took two complementary approaches to understand the underlying logics of the auditory signal processing. On the one hand, in Chapter 2 we employed a top-down theoretical approach and considered principles that govern the activities of populations of neurons. Here we focused on the anatomical characteris-

tics that the auditory cortex has by far more neurons than the periphery does, and demonstrated that sparse overcomplete representations could contribute to difficult computations such as to solve the monaural cocktail party problem (Sections 2.2–2.3). We also derived several predictions on neural behaviors to provide ways to experimentally test the theory (Section 2.4). On the other hand, in Chapter 3 we took a bottom-up experimental approach and aimed at single-neuron dynamics. A major goal here was to study response properties of auditory cortex so as to tune up an encoding model, and we specifically assessed the time-course of neurons’ “memory” to limit the window length—and consequently the number of free parameters—of a model. We found that a slow stimulus adaptation over seconds would be the most relevant for response predictions (Section 3.5). However, the model performance did not improve by simply extending a window length, and none of the models we examined could even reach the estimated upper bound of the response prediction performance based on the experimental data.

Two comments are in order about the fundamental caveats we should be aware of. First, I generally assumed serial or sequential stages for the sensory signal processing, but the brain would most likely be a parallel processor. We could overlook this point as long as we focus only on the underlying principles—or the top two levels of the understandings in the Marr’s framework (Marr, 1982; see also Section 1.2). For example, the idea of sparse overcomplete representations was exploited in Chapter 2 as a possible principle of computation in the brain, but one would not be surprised even if the discussed implementations—i.e., feature extractions by non-negative matrix factorization and sparse representations by L_1 -norm minimization—do not correspond to the actual algorithms employed in the nervous system. To fully understand the brain, however, we cannot avoid considering the physical instantiations—or the bottom level in the Marr’s classification (Marr, 1982)—and this would be one of the most important problems from biological viewpoints (see also Section 4.2 below). In the computer science community, parallel or distributed computing is now an area of extensive research because of its various advantages in cost and time savings and because of its ability to solve larger problems. Although it is less studied in the context of neuroscience so far, I believe that insights gained

from the one can be applied to the other (and vice versa), and that interdisciplinary research would open ways to the “parallelism.”

The second caveat is the processing with respect to “time.” From modeling viewpoints, we in general think of auditory signal processing as an analogue of visual signal processing (Massaro, 1972; Shamma, 2001)—i.e., frame-by-frame “auditory image” processing—and the aspects of continuous and/or sequential processing over time have been less studied so far despite a growing awareness of adaptive and/or dynamic models. It is not yet clear if we could treat “time” simply as an extra dimension in much the same way as “space” processing, or if we need a distinct theory for understanding temporal processing (see also Schwartz et al., 2007; Section 1.2). We should also be aware that changes in response properties could range over various time-scales depending on the properties of input sensory signals (or, individual sensory experiences; see e.g., Fritz et al., 2003; Ulanovsky et al., 2004; Weinberger, 2004)—from a rapid change over tens of milliseconds as in forward masking (Bregman, 1990) to a slow adaptation over seconds as shown in Chapter 3, and to the developmental and plastic changes (including cell death) in much longer time-scales—and that such changes can occur in both reversible and irreversible manners. While the former could be explained by changing parameters in a given model framework (e.g., by a gain control), the latter changes might then require different model frameworks—or “concepts”—to make sense of them. Because the system identification approach¹ typically expects that a target system should be “unchanged,” it would thus be critical to look at appropriate time-scales for building plausible models that work well (at least within the given time-scales). It remains to be addressed, however, how many different classes of functions or models are required to fully account for neural behaviors, and if they can in fact be unified by introducing some meta-parameters.

¹Note however that multi-resolution multi-model analysis is now attracting much attention in various fields of science.

4.2 General Discussion

How useful would theory be for understanding biology? A theory (or model) can be made in data- and/or hypothesis-driven manners, but the main contributions of theory would be to provide comprehensive understandings of phenomena and predict outcomes for novel situations. It is then especially important to examine the correspondence between theory and biology by experiments, because biology is the study to identify what exactly nature selected among many possibilities that could all work well for achieving organisms' goals. Therefore, biologically reasonable interpretations are indispensable; we could then design and perform (new) experiments to verify—or falsify—the models accordingly, and also seek theoretically interesting phenomena to improve the models and faithfully mimic mechanisms underlying functions or computations in the brain.

To see how important interpretations would be for a theory in biology, below I discuss some more conceptual issues on two prevailing ideas in neuroscience; *feature extraction* and *optimization*. First, a goal of sensory signal processing is often considered to be extracting features from received signals (Victor, 2005). Here I will not deny this idea but argue that it can be achieved by many different approaches. A straightforward method would be to search for structures of interest from backgrounds as “active” processes. Alternatively, we could also extract features by ignoring irrelevancies or removing uninteresting objects as a “passive” approach. From theoretical viewpoints, it would not matter whether we take active or passive strategies as long as the mathematical formulations are the same—e.g., a band-pass filter is nothing but a band-pass filter, and it does not matter whether it passes a particular bandwidth or attenuates the surroundings. However, this interpretation does matter from biological viewpoints because nature has presumably followed some particular strategies but not the others under certain selection pressures it has faced in evolution, leading to the nervous system organizations as we now have in this world.

The second example is the idea of optimization, e.g., with respect to “efficient” coding or redundancy reduction² (Attneave, 1954; Barlow, 1961, 2001). Again, I am not criticizing the optimization approach itself, but bringing out its usage in the context of biology. The point is that we do not know if the brain has been optimized, or is still in the process of optimization (from both phylogenetic and ontogenetic perspectives). Thus it does not mean a lot to simply measure some “neural correlates,” e.g., how much redundancy in “bits” there is in neural spiking patterns—it is even unclear if “bits” are what neurons care about—but we should compare it across species in different phylogenetic trees (ideally across those in different stages in evolution) and also across different developmental stages (preferably at different processing levels in the brain as well) so as to discuss the meaning of the optimization principle from biological viewpoints and *justify* or *falsify* it. It would also be interesting to think about activities/functions in suboptimal regimes, especially because human behaviors—even those of the “highest” animals—do not seem optimal or logical in many situations. One biologically interesting question would then be to examine when the optimality breaks down—e.g., by exploiting speed-accuracy trade-off or by lesion study approaches—and ask if higher animals are somehow “more optimized” than lower animals.

Finally, I would dare like to pose questions, some of which might even be philosophical and unanswerable by contemporary science (partly because of the lack of a solid—or mathematical—definition, and partly because they are “pseudoscientific” questions by themselves; Popper, 1934).

- To what extent are experimental animals (e.g., rats) similar to humans? Do we work in the same way as they do? How much we could know about us humans by studying experimental animals?

²It should be mentioned that “redundancy” in Barlow’s sense is defined as $1 - I/I_{\max}$, where I and I_{\max} are the observed Shannon information and the channel capacity, respectively (or, defined as: $1 - H/H_{\max}$, where H and H_{\max} are the observed and maximum entropies, respectively; Shannon, 1948) *but not* the difference between the sum of the information from individual cells and the total information (as is often defined in many literatures; see e.g., Schneidman et al., 2003; Nirenberg and Latham, 2003; Latham and Nirenberg, 2005; Averbeck et al., 2006), even though the point of the discussion here is not what “redundancy” means but how we could interpret a theoretically-interesting quantity from biological viewpoints.

- How big are individual differences between experimental animals, and between us? Could all contributions of *nurture* be explained as the results of “learning processes” once we understand everything about *nature*?
- Where does “conscious mind” arise from? Is there any neural correlate of consciousness in humans, and in animals (e.g., apes) as well? How many neurons (computational units) are required to have a conscious mind? Can we implement it in electronic devices?
- To what extent is the brain deterministic, and to what extent is it probabilistic/stochastic? How well can we describe the “the theory of the brain” at the level of classical mechanics, and how about at the level of quantum mechanics? Is everything—even the fact that I wrote this thesis and you are reading it—predetermined, or do we have “free will?”
- Can we build—in theory and even in practice—an “artificial brain?”

It might be nonsense as a biological science to ask any question unless there is a way to test it by (Gedanken) experiments. However, I would like to ask myself such questions as well because “to think is to live” (*vivere est cogitare*; Marcus Tullius Cicero) and because I am a “thinking reed” (*roseau pensant*; Blaise Pascal).