# Part II
# Discrete Time Processes

In the last chapter we considered two basic examples of large deviation theory for discrete time processes. One example was the empirical measure for iid random variables in a Polish space $S$, and the other was the sample mean for the case $S = \mathbb{R}^d$. In the next part of the book we will generalize these two examples in various directions. In all cases, the starting point will be a representation obtained using the chain rule for relative entropy, though the underlying noise models will be more complex than the simple product measures of Chap. 3. One generalization will be relatively straightforward, which is the extension from the empirical measure for iid to the empirical measure for a Markov chain in Chap. 6.

A second generalization will be a "process level" and "state dependent" generalization of Cramér's theorem in Chap. 4. Here we will consider the large deviation principle for a very general stochastic recursive model. A scaling is introduced that makes explicit that the system can be thought of as a small random perturbation of an ordinary differential equation, and indeed included would be models such as the Euler–Maruyama approximation to an SDE with small noise. Our perspective here will be very similar to the one used in Chap. 3 to obtain Cramér's theorem from Sanov's theorem, and in fact, we will prove large deviation properties of a "time-dependent" empirical measure for the driving noises and then view the state of the stochastic recursive system as a mapping on this empirical measure. Under suitable integrability assumptions, the passage from the large deviation of the empirical measure to that of the process will parallel that taking us from Sanov's theorem to Cramér's theorem.

Chapter 7 focuses on models with various special features. One class comprises discrete time dynamical models for occupancy-type problems. These are classical models from combinatorial probability (e.g., the coupon collector's problem), and the particular feature that makes their analysis different is that the probabilities of certain types of jumps tend to zero as a boundary is approached. This "diminishing rates" feature puts them outside the models of Chap. 4, and in fact, a careful construction is needed to establish the large deviation lower bound for trajectories that touch the boundary. On the other hand, as also discussed in this chapter, these models and many related generalizations have the feature that the variational problems one needs to solve to extract information from the rate function on path space have nearly explicit solutions, i.e., solutions that can be identified by solving a low-dimensional constrained convex optimization problem. Also included in this chapter is the formulation of a two-time-scale model and the statement of the corresponding LDP. The proof of the LDP for this model, which combines arguments used in Chaps. 4 and 6, is omitted.

Moderate deviations for the same class of small noise Markov processes as in Chap. 4 are the topic of Chap. 5. What is meant by “moderate deviations” is approximations for events that are closer to the LLN limit then those approximated via standard large deviations. While the starting point is the same relative entropy representation as in Chap. 4, moderate deviations (which can be phrased as large deviations for a suitably centered and rescaled system) are in some ways simpler but in other ways more difficult than the corresponding large deviations. As discussed at some length in Chap. 5, a particular motivation for the moderate deviation approximation is the development of accelerated Monte Carlo schemes for this same class of events. An example of such will be given in Sect. 17.5.