Preface

This book arose from courses on large deviations and related topics given by the authors in the Departments of Mathematics at the Ohio State University (1993), at the University of Wisconsin-Madison (2006 and 2013), and at the University of Utah (2008 and 2013).

Our goal has been to create an attractive collection of material for a semester's course which would also serve the broader needs of students from different fields. This goal has had two implications for the book.

(1) We have not aimed at anything like an encyclopedic coverage of different techniques for proving large deviation principles (LDPs). Part I of the book focuses on one classic line of reasoning: (i) upper bound by an exponential Markov-Chebyshev inequality, (ii) lower bound by a change of measure, and (iii) an argument to match the rates from (i) and (ii). Beyond this technique Part I covers Bryc's theorem and proves Cramér's theorem with the subadditive method. Part III of the book covers the Gärtner-Ellis theorem and an approach based on the convexity of a local rate function due to Baxter and Jain.

(2) We have not felt obligated to stay within the boundaries of large deviation theory but instead follow the trail of some interesting material. Large deviation theory is a natural gateway to statistical mechanics. A discussion of statistical mechanics would be incomplete without some study of phase transitions. We prove the phase transition of the Ising model in two different ways: (i) first with classical techniques: the Peierls argument, Dobrushin's uniqueness condition, and correlation inequalities and (ii) the second time with random cluster measures. This means leaving large deviation theory completely behind. Along the way we have the opportunity to learn coupling methods which are central to modern probability theory

but do not get serious application in the typical first graduate course in probability.

We give now a brief overview of the contents of the book.

Part I covers core general large deviation theory, the relevant convex analysis, and the large deviations of independent and identically distributed (i.i.d.) processes on three levels: Cramér's theorem, Sanov's theorem, and the process level LDP for i.i.d. variables indexed by a multidimensional square lattice.

Part II introduces Gibbs measures and proves the Dobrushin-Lanford-Ruelle variational principle that characterizes translation-invariant Gibbs measures. After this we study the phase transition of the Ising model. Part II ends with a chapter on the Fortuin-Kasteleyn random cluster model and the percolation approach to Ising phase transition.

Part III develops the large deviation themes of Part I in several directions. Large deviations of i.i.d. variables are complemented with moderate deviations and with more precise large deviation asymptotics. The Gärtner-Ellis theorem is developed carefully, together with the necessary additional convex analysis beyond the basics covered in Part I. From large deviations of i.i.d. processes we move on to Markov chains, to nonstationary independent random variables, and finally to random walk in a dynamical random environment. The last two topics give us an opportunity to apply another approach to proving large deviation principles, namely the Baxter-Jain theorem. The Baxter-Jain theorem has not previously appeared in textbooks, and its application to random walk in random environment is new.

Here is a guide to the dependencies between the parts of the book. Sections 2.1-2.3 and 3.1-3.2 are foundational for all discussions of large deviations. In addition, we have the following links. Chapter 5 relies on Sections 4.1-4.2, and Chapter 6 relies on Chapter 5. Chapter 8 relies on Chapters 6 and 7. Chapter 9 can be read independently of large deviations after Sections 7.1-7.3 and 7.6. Section 10.2 makes sense only in the context of Chapter 9. Chapters 12 and 14 are independent of each other and both rely on Sections 4.1-4.2. Chapter 13 relies on Chapter 5. Chapter 15 relies on Section 13.1 and Chapter 14. Chapter 16 relies on Chapter 14.

The ideal background for reading this book would be some familiarity with the language of measure-theoretic probability. Large deviation theory does also require a little analysis, point set topology, and functional analysis. For example, readers should be comfortable with lower semicontinuity and the weak topology on probability measures. It should be possible for an instructor to accommodate students with quick lectures on technical prerequisites whenever needed. It is also possible to consider everything in the framework of concrete finite spaces, in which case probability measures become simply probability vectors.

In practice our courses have been populated by students with very diverse backgrounds, many with less than ideal knowledge of analysis and probability. This has turned out less problematic than one might initially fear. Mathematics students are typically fully satisfied only after every theoretical point is rigorously justified. But engineering students are content to set aside much of the theory and focus on the essentials of the phenomenon in question. There is great interest in probability theory among students of economics, engineering, and the sciences. This interest should be encouraged and nurtured with accessible courses.

The appendixes in the back of the book serve two purposes. There is a quick overview of some basic results of analysis and probability without proofs, for the reader who wants a quick refresher. In particular, here the reader can look up textbook tools such as convergence theorems and inequalities that are referenced in the text. The other material in the appendixes consists of specialized results used in the text, such as a minimax theorem and inequalities from statistical mechanics. These are proved.

Since this book evolved in courses where we tried to actively engage the students, the development of the material relies on frequent exercises. We realize that this feature may not appeal to some readers. On the other hand, spelling out all the technical details left as exercises might make for tedious reading. Hopefully an instructor can fill in those details fairly easily if he or she wants to present full details in class. Exercises that are referred to in the text are marked with an asterisk.

One of us (Timo Seppäläinen) first learned large deviations from a course taught by Steven Orey in 1988–1989 at the University of Minnesota. We are greatly indebted to the existing books on the subject, especially those by Amir Dembo and Ofer Zeitouni [**15**], Frank den Hollander [**16**], Jean-Dominique Deuschel and Daniel Stroock [**18**], Richard Ellis [**32**], and Srinivasa Varadhan [**79**].

As a text that combines large deviations with equilibrium statistical mechanics, [**32**] is a predecessor of ours. There is obviously a good degree of overlap but the books are different. Ours is a textbook with a lighter touch while [**32**] is closer to a research monograph, covers more models in detail, and explains much of the physics. We recommend [**32**] to our readers and students for further study. Our phase transition discussion covers the nearest-neighbor Ising model while [**32**] also covers long-range Ising models. On the other hand, [**32**] does not cover Dobrushin's uniqueness theorem, random cluster models, general lattice systems, or their large deviations.

Our literature references are sparse and sometimes do not assign credit to the originators of the ideas. We encourage the reader to consult the superb historical notes and references in the monographs of Dembo-Zeitouni, Ellis, and Georgii.

We thank Jeff Steif for lecture notes that helped shape the proof of Theorem 9.2, Jim Kuelbs for material for Chapter 11, and Chuck Newman for helpful discussions on the liquid-gas phase transition for Chapter 7. We also thank Davar Khoshnevisan for several valuable suggestions. We thank the team at the AMS and especially Ed Dunne for his patience in the face of serial breaches of agreed deadlines, and we thank the several reviewers for valuable suggestions.

Support from the National Science Foundation, the Simons Foundation, and the Wisconsin Alumni Research Foundation is gratefully acknowledged.

Firas Rassoul-Agha
Timo Seppäläinen