

## Research Article

# Large Deviation Analysis of a Droplet Model Having a Poisson Equilibrium Distribution

Richard S. Ellis<sup>1</sup> and Shlomo Ta'asan<sup>2</sup>

<sup>1</sup>Department of Mathematics and Statistics, University of Massachusetts, Amherst, MA 01003, USA

<sup>2</sup>Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh, PA 15213, USA

Correspondence should be addressed to Richard S. Ellis; [rsellis@math.umass.edu](mailto:rsellis@math.umass.edu)

Received 5 March 2015; Accepted 20 August 2015

Academic Editor: Lukasz Stettner

Copyright © 2015 R. S. Ellis and S. Ta'asan. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In this paper we use large deviation theory to determine the equilibrium distribution of a basic droplet model that underlies a number of important models in material science and statistical mechanics. Given  $b \in \mathbb{N}$  and  $c > b$ ,  $K$  distinguishable particles are placed, each with equal probability  $1/N$ , onto the  $N$  sites of a lattice, where  $K/N$  equals  $c$ . We focus on configurations for which each site is occupied by a minimum of  $b$  particles. The main result is the large deviation principle (LDP), in the limit  $K \rightarrow \infty$  and  $N \rightarrow \infty$  with  $K/N = c$ , for a sequence of random, number-density measures, which are the empirical measures of dependent random variables that count the droplet sizes. The rate function in the LDP is the relative entropy  $R(\theta \mid \rho^*)$ , where  $\theta$  is a possible asymptotic configuration of the number-density measures and  $\rho^*$  is a Poisson distribution with mean  $c$ , restricted to the set of positive integers  $n$  satisfying  $n \geq b$ . This LDP implies that  $\rho^*$  is the equilibrium distribution of the number-density measures, which in turn implies that  $\rho^*$  is the equilibrium distribution of the random variables that count the droplet sizes.

## 1. Introduction

This paper is motivated by a natural question for a basic model of a droplet. Given  $b \in \mathbb{N}$  and  $c > b$ ,  $K$  distinguishable particles are placed, each with equal probability  $1/N$ , onto the  $N$  sites of a lattice  $\Lambda_N = \{1, 2, \dots, N\}$ . Under the assumption that  $K/N = c$  and that each site is occupied by a minimum of  $b$  particles, what is the equilibrium distribution, as  $N \rightarrow \infty$ , of the number of particles per site? We prove in Corollary 3 that this equilibrium distribution is a Poisson distribution, with mean  $c$ , restricted to the set of positive integers  $n$  satisfying  $n \geq b$ . As we explain near the end of the Introduction, this equilibrium distribution has important applications to technologies using sprays and powders.

As in many other models in statistical mechanics, we can identify the equilibrium distribution by exhibiting it as the unique minimum point of a rate function in a large deviation principle (LDP). Other models for which this procedure can be implemented are discussed at the end of the Introduction.

For the droplet model we prove the LDP for a sequence of random probability measures, called number-density measures, which are the empirical measures of a sequence of

dependent random variables that count the droplet sizes. This LDP is stated in Theorem 1. Our proof is self-contained and starts from first principles, using techniques that are familiar in applied mathematics and statistical mechanics. For example, the proof of the local large deviation estimate in Theorem 5, a key step in the proof of the LDP for the number-density measures, is based on combinatorics, Stirling's formula, and Laplace asymptotics.

Our use of combinatorial methods goes back to Boltzmann in his work on the discrete ideal gas. He calculated the Maxwell-Boltzmann equilibrium distribution for this system by analyzing the asymptotic behavior of a particular multinomial coefficient [1]. Starting with Boltzmann's work, combinatorial methods have remained an important tool in both statistical mechanics and in the theory of large deviations, offering insights into a wide variety of physical and mathematical phenomena via techniques that are elegant, powerful, and often elementary. In applications to statistical mechanics, this state of affairs is explained by the observation that “many fundamental questions ... are inherently combinatorial, ... including the Ising model, the Potts model,

monomer-dimer systems, self-avoiding walks and percolation theory" [2]. For the two-dimensional Ising model and other exactly soluble models, [3, 4] are recommended.

A similar situation holds in the theory of large deviations. For example, Section 2.1 of [5] discusses combinatorial techniques for finite alphabets and points out that because of the concreteness of these applications the LDPs are proved under much weaker conditions than the corresponding results in the general theory, into which the finite-alphabet results give considerable insight. The text [6] devotes several early sections to large deviation results for i.i.d. random variables having a finite state space and proved by combinatorial methods, including a sophisticated, level-3 result for the empirical pair measure.

In order to formulate the LDP for the number-density measures in our droplet model, a standard probabilistic model is introduced. The configuration space is the set  $\Omega_N = \Lambda_N^K$  consisting of all  $\omega = (\omega_1, \omega_2, \dots, \omega_K)$ , where  $\omega_i$  denotes the site in  $\Lambda_N$  occupied by the  $i$ th particle. The cardinality of  $\Omega_N$  equals  $N^K$ . Denote by  $P_N$  the uniform probability measure that assigns equal probability  $1/N^K$  to each of the  $N^K$  configurations  $\omega \in \Omega_N$ . The asymptotic analysis of the droplet model involves the two random variables, which are functions of the configuration  $\omega \in \Omega_N$ : for  $\ell \in \Lambda_N$ ,  $K_\ell(\omega)$  denotes the number of particles occupying the site  $\ell$  in the configuration  $\omega$ ; for  $j \in \mathbb{N} \cup \{0\}$ ,  $N_j(\omega)$  denotes the number of sites  $\ell \in \Lambda_N$  for which  $K_\ell(\omega) = j$ .

We focus on the subset of  $\Omega_N$  consisting of all configurations  $\omega$  for which every site of  $\Lambda_N$  is occupied by at least  $b$  particles. Because of this restriction  $N_j(\omega)$  is indexed by  $j \in \mathbb{N}_b = \{n \in \mathbb{Z} : n \geq b\}$ . It is useful to think of each particle as having one unit of mass and of the set of particles at each site  $\ell$  as defining a droplet. With this interpretation, for each configuration  $\omega$ ,  $K_\ell(\omega)$  denotes the mass or size of the droplet at site  $\ell$ . The  $j$ th droplet class has  $N_j(\omega)$  droplets and mass  $jN_j(\omega)$ . Because the number of sites in  $\Lambda_N$  equals  $N$  and the sum of the masses of all the droplet classes equals  $K$ , the following conservation laws hold for such configurations:

$$\begin{aligned} \sum_{j \in \mathbb{N}_b} N_j(\omega) &= N, \\ \sum_{j \in \mathbb{N}_b} jN_j(\omega) &= K. \end{aligned} \quad (1)$$

In addition, since the total number of particles is  $K$ , it follows that  $\sum_{\ell \in \Lambda_N} K_\ell = K$ . These equality constraints show that the random variables  $N_j$  and  $K_\ell$  are not independent.

In order to carry out the asymptotic analysis of the droplet model, we introduce a quantity  $m = m(N)$  that converges to  $\infty$  sufficiently slowly with respect to  $N$ ; specifically, we require that  $m(N)^2/N \rightarrow 0$  as  $N \rightarrow \infty$ . In terms of  $b$  and  $m$  we define the subset  $\Omega_{N,b,m}$  of  $\Omega_N$  consisting of all configurations  $\omega$  for which every site of  $\Lambda_N$  is occupied by at least  $b$  particles and at most  $m$  of the quantities  $N_j(\omega)$  are positive. This second condition is a useful technical device that allows us to control the errors in several estimates. In Appendix D of [7] we present evidence supporting the conjecture that this condition can be eliminated. The discussion in that appendix

involves a number of interesting topics including Stirling numbers of the second kind (see [8, pp. 96-97] and [9, §5.4]) and their asymptotic behavior [10, Example 5.4].

The random quantities in the droplet model for which we formulate an LDP are the number-density measures  $\Theta_{N,b}$ . For  $\omega \in \Omega_{N,b,m}$  these random probability measures assign to  $j \in \mathbb{N}_b$  the probability  $N_j(\omega)/N$ , which is the number density of the  $j$ th droplet class. Because of the two conservation laws in (1) and because  $K/N = c$ , for  $\omega \in \Omega_{N,b,m}$ ,  $\Theta_{N,b}(\omega)$  is a probability measure on  $\mathbb{N}_b = \{n \in \mathbb{Z} : n \geq b\}$  having mean  $c$ . Thus  $\Theta_{N,b}$  takes values in  $\mathcal{P}_{\mathbb{N}_b,c}$ , which is defined to be the set of probability measures on  $\mathbb{N}_b$  having mean  $c$ .

The probability measure  $P_{N,b,m}$  defining the droplet model is obtained by restricting the uniform measure  $P_N$  to the set of configurations  $\Omega_{N,b,m}$ . Thus  $P_{N,b,m}$  equals the conditional probability  $P_N(\cdot \mid \Omega_{N,b,m})$ . In the language of statistical mechanics  $P_{N,b,m}$  defines a microcanonical ensemble that incorporates the conservation laws for number and mass expressed in (1).

A natural question is to determine two equilibrium distributions: the equilibrium distribution  $\rho^*$  of the number-density measures and the equilibrium distribution  $\rho^{**} = \sum_{j \in \mathbb{N}_b} \rho_j^{**} \delta_j$  of the droplet-size random variables  $K_\ell$ . These distributions are defined by the following two limits: for any  $\varepsilon > 0$ , any  $\ell \in \Lambda_N$ , and all  $j \in \mathbb{N}_b$

$$\begin{aligned} \lim_{N \rightarrow \infty} P_{N,b,m}(\Theta_{N,b} \in B(\rho^*, \varepsilon)) &\rightarrow 1, \\ \lim_{N \rightarrow \infty} P_{N,b,m}(K_\ell = j) &= \rho_j^{**}, \end{aligned} \quad (2)$$

where  $B(\rho^*, \varepsilon)$  denotes the open ball with center  $\rho^*$  and radius  $\varepsilon$  defined with respect to an appropriate metric on  $\mathcal{P}_{\mathbb{N}_b,c}$ . As we prove, the equilibrium distributions of  $\Theta_{N,b}$  and  $K_\ell$  coincide. As in many models in statistical mechanics, an efficient way to determine the equilibrium distribution  $\Theta_{N,b}$  is to prove an LDP for  $\Theta_{N,b}$ , which we carry out in Theorem 1. This theorem is the main result in the paper.

The content of Theorem 1 is the following: as  $N \rightarrow \infty$ , the sequence of number-density measures  $\Theta_{N,b}$  satisfies the LDP on  $\mathcal{P}_{\mathbb{N}_b,c}$  with respect to the measures  $P_{N,b,m}$ . The rate function is the relative entropy  $R(\theta \mid \rho_{b,\alpha})$  of  $\theta \in \mathcal{P}_{\mathbb{N}_b,c}$  with respect to the Poisson distribution  $\rho_{b,\alpha}$  on  $\mathbb{N}_b$  having components  $\rho_{b,\alpha;j} = [Z_b(\alpha)]^{-1} \cdot \alpha^j/j!$  for  $j \in \mathbb{N}_b$ . In this formula  $Z_b(\alpha)$  is the normalization that makes  $\rho_{b,\alpha}$  a probability measure, and  $\alpha$  equals the unique value  $\alpha_b(c)$  for which  $\rho_{b,\alpha_b(c)}$  has mean  $c$  [Theorem A.2]. Using the fact that  $R(\theta \mid \rho_{b,\alpha_b(c)})$  equals 0 at the unique measure  $\theta = \rho_{b,\alpha_b(c)}$ , we apply the LDP for  $\Theta_{N,b}$  to conclude in Theorem 2 that  $\rho_{b,\alpha_b(c)}$  is the equilibrium distribution of  $\Theta_{N,b}$ . Corollary 3 then implies that  $\rho_{b,\alpha_b(c)}$  is also the equilibrium distribution of  $K_\ell$ .

The space  $\mathcal{P}_{\mathbb{N}_b,c}$  is the most natural space on which to formulate the LDP for  $\Theta_{N,b}$  in Theorem 1. Not only is  $\mathcal{P}_{\mathbb{N}_b,c}$  the smallest convex set of probability measures containing the range of  $\Theta_{N,b}$  for all  $N \in \mathbb{N}$ , but also the union over  $N \in \mathbb{N}$  of the range of  $\Theta_{N,b}$  is dense in  $\mathcal{P}_{\mathbb{N}_b,c}$ . As we explain in part (a) of Theorem 4,  $\mathcal{P}_{\mathbb{N}_b,c}$  is not a complete, separable metric space, a situation that prevents us from directly applying general

results in the theory of large deviations that require the setting of a complete, separable metric space.

The droplet model is defined in Section 2. Step 1 in the proof of the LDP for  $\Theta_{N,b}$  is to derive the local large deviation estimate in part (b) of Theorem 5. This local estimate, one of the centerpieces of the paper, gives information not available in the LDP for  $\Theta_{N,b}$ , which involves global estimates. Step 2 is to lift the local large deviation estimate to the large deviation limit for  $\Theta_{N,b}$  lying in open balls and certain other subsets of  $\mathcal{P}_{\mathbb{N}_b,c}$  while Step 3 is to lift the large deviation limit for open balls and certain other subsets to the LDP for  $\Theta_{N,b}$  stated in Theorem 1. Steps 2 and 3 are explained in Section 4.

Details of Steps 2 and 3 as well as other routine proofs are omitted from the present paper. They appear in the unpublished companion paper [7], which also contains additional background material. The paper [1] explores how our work on the droplet model was inspired by the work of Ludwig Boltzmann on a simple model of a discrete ideal gas. The main connection is via the local large deviation estimate in part (b) of Theorem 5. When  $b = 0$ , the LDP for a path version of  $\Theta_{n,0}$  with  $K = tN$  and  $t > 0$  varying appears in [11, 12].

The main application of the results in this paper is to technologies using sprays and powders, which are ubiquitous in many fields, including agriculture, the chemical and pharmaceutical industries, consumer products, electronics, manufacturing, material science, medicine, mining, paper making, the steel industry, and waste treatment. In this paper we focus on sprays; our theory also applies to powders with only changes in terminology [13]. The behavior of sprays might be complex depending on various parameters including evaporation, temperature, and viscosity. Our goal here is to consider the simplest model where the only assumption is made on the average size of droplets in the spray. In many situations it is important to have good control over the sizes of the droplets, which can be translated into properties of probability distributions. The size distributions are important because they determine reliability and safety in each particular application.

Interestingly, there does not seem to be a rigorous theory that predicts the equilibrium distribution of droplet sizes, analogous to the Maxwell-Boltzmann distribution of energy levels in a discrete ideal gas [14, 15]. Our goal in the present paper is to provide such a theory. We do so by focusing on one aspect of the problem related to the relative entropy, an approach that characterizes the equilibrium distribution of droplet sizes as being a Poisson distribution restricted to  $\mathbb{N}_b$ . We expect that this distribution will be important in experimental observations. A full understanding of droplet behavior under dynamic conditions requires treating many other aspects and is beyond the scope of this paper. We plan to apply the ideas in this paper to understand the entropy of dislocation networks [16].

The importance of predicting droplet size can be seen from the wide range of applications utilizing sprays [17, 18]. Because of the importance of this problem, novel approaches for measuring size distribution of droplet size in sprays have been developed [19–23]. What makes the problem of predicting droplet size particularly interesting is the complexity of droplet-size distribution, which is attributed to

many factors such as temperature and viscosity. As [24] shows, even the nozzle plays a significant role in the outcome. Many theoretical tools used to understand the distribution of droplet size in sprays include entropy [25], which also plays a key role in the present paper.

We end the Introduction by expanding on a comment made at the beginning of this section. This comment concerns one of the main applications of large deviation theory in statistical mechanics, which is to identify the equilibrium distribution or distributions of a model as the minimum point(s) of the rate function in an LDP for the model. This procedure is also useful to study phase transitions in the model, which concern how the structure of the set of equilibrium distributions changes as the parameters defining the model change. There are numerous other models for which this procedure has been used. They include the following three lattice spin models: the Curie-Weiss spin system, the Curie-Weiss-Potts model, and the mean-field Blume-Capel model, which is also known as the mean-field BEG model. As explained in the respective Sections 6.6.1, 6.6.2, and 6.6.3 of [26], the large deviation analysis shows that each of these three models has a different phase transition structure. Details of the analysis for the three models are given in the references [6, §IV.4], [27–29]. Section 9 of [30] outlines how large deviation theory can be applied to determine equilibrium structures in statistical models of two-dimensional turbulence. Details of this analysis are given in [31].

## 2. Definition of Droplet Model and Main Theorem

After defining the droplet model, we state the main theorem in the paper, Theorem 1. The content of this theorem is the LDP for the sequence of random, number-density measures, which are the empirical measures of a sequence of dependent random variables that count the droplet sizes in the model. As we show in Theorem 2 and in Corollary 3, the LDP enables us to identify a Poisson distribution as the equilibrium distribution both of the number-density measures and of the droplet-size random variables. In Theorem 4 we prove a number of properties of two spaces of probability measures in terms of which the LDP for the number-density measures is formulated.

We start by fixing parameters  $b \in \mathbb{N} \cup \{0\}$  and  $c \in (b, \infty)$ . The droplet model is defined by a probability measure  $P_{N,b}$  parameterized by  $N \in \mathbb{N}$  and the nonnegative integer  $b$ . The measure depends on two other positive integers,  $K$  and  $m$ , where  $2 \leq m \leq N < K$ . Both  $K$  and  $m$  are functions of  $N$  in the large deviation limit  $N \rightarrow \infty$ . In this limit we take  $K \rightarrow \infty$  and  $N \rightarrow \infty$ , where  $K/N$ , the average number of particles per site, equals  $c$ . Thus  $K = Nc$ . In addition, we take  $m \rightarrow \infty$  sufficiently slowly by choosing  $m$  to be a function  $m(N)$  satisfying  $m(N) \rightarrow \infty$  and  $m(N)^2/N \rightarrow 0$  as  $N \rightarrow \infty$ ; for example,  $m(N) = N^\delta$  for some  $\delta \in (0, 1/2)$ . Throughout this paper we fix such a function  $m(N)$ . The parameter  $b$  and the function  $m = m(N)$  first appear in the definition of the set



of configurations  $\Omega_{N,b,m}$  in (3), where these quantities will be explained.

Because  $K$  and  $N$  are integers,  $c$  must be a rational number. This in turn imposes a restriction on the values of  $N$  and  $K$ . If  $c$  is a positive integer, then  $N \rightarrow \infty$  along the positive integers and  $K \rightarrow \infty$  along the subsequence  $K = cN$ . If  $c = x/y$ , where  $x$  and  $y$  are relatively prime, positive integers with  $y \geq 2$ , then  $N \rightarrow \infty$  along the subsequence  $N = yn$  for  $n \in \mathbb{N}$  and  $K \rightarrow \infty$  along the subsequence  $K = cN = xn$ . Throughout this paper, when we write  $N \in \mathbb{N}$  or  $N \rightarrow \infty$ , it is understood that  $N$  and  $K$  satisfy the restrictions discussed here.

In the droplet model  $K$  distinguishable particles are placed, each with equal probability  $1/N$ , onto the sites of the lattice  $\Lambda_N = \{1, 2, \dots, N\}$ . This simple description corresponds to a simple probabilistic model. The configuration space is the set  $\Omega_N = \Lambda_N^K$  consisting of all sequences  $\omega = (\omega_1, \omega_2, \dots, \omega_K)$ , where  $\omega_i \in \Lambda_N$  denotes the site in  $\Lambda_N$  occupied by the  $i$ th particle. Let  $\rho^{(N)}$  be the measure on  $\Lambda_N$  that assigns equal probability  $1/N$  to each site in  $\Lambda_N$ , and let  $P_N = (\rho^{(N)})^K$  be the product measure on  $\Omega_N$  with equal one-dimensional marginals  $\rho^{(N)}$ . Thus  $P_N$  is the uniform probability measure that assigns equal probability  $1/N^K$  to each of the  $N^K$  configurations  $\omega \in \Omega_N$ ; for subsets  $A$  of  $\Omega_N$  we have  $P_N(A) = \text{card}(A)/N^K$ , where  $\text{card}$  denotes cardinality.

The asymptotic analysis of the droplet model involves two random variables. For  $\ell \in \Lambda_N$  and  $\omega \in \Omega_N$ ,  $K_\ell(\omega)$  denotes the number of particles occupying site  $\ell$  in the configuration  $\omega$ . For  $j \in \mathbb{N} \cup \{0\}$  and  $\omega \in \Omega_N$ ,  $N_j(\omega)$  denotes the number of sites  $\ell \in \Lambda_N$  for which  $K_\ell(\omega) = j$ . The dependence of  $K_\ell(\omega)$  and  $N_j(\omega)$  on  $N$  is not indicated in the notation. Because the distributions of both random variables depend on  $N$ , both  $K_\ell$  and  $N_j$  form triangular arrays.

We now specify the role played by the nonnegative integer  $b$ , first focusing on the case where  $b$  is a positive integer. The case where  $b = 0$  is discussed later. For  $\omega \in \Omega_N$ , in general there exist sites  $\ell \in \Lambda_N$  for which  $K_\ell(\omega) = 0$ ; that is, sites that are occupied by 0 particles. The next step in the definition of the droplet model is to restrict to a subset  $\Omega_{N,b,m}$  of configurations  $\omega \in \Omega_N$  for which every site is occupied by at least  $b$  particles and the following constraint holds: for any configuration  $\omega \in \Omega_{N,b,m}$  at most  $m$  of the components  $N_j(\omega)$  are positive, where  $m = m(N) \rightarrow \infty$  and  $m(N)^2/N \rightarrow 0$  as  $N \rightarrow \infty$ . Because for  $\omega \in \Omega_{N,b,m}$  every site  $\ell \in \Lambda_N$  is occupied by at least  $b$  particles, we have  $K_\ell(\omega) \geq b$  and  $N_j(\omega)$  is indexed by  $j \in \mathbb{N}_b = \{n \in \mathbb{Z} : n \geq b\}$ . We denote by  $N(\omega)$  the sequence  $\{N_j(\omega), j \in \mathbb{N}_b\}$  and define  $|N(\omega)|_+ = \text{card}\{j \in \mathbb{N}_b : N_j(\omega) \geq 1\}$ . In terms of this notation

$$\Omega_{N,b,m} = \{\omega \in \Omega_N : K_\ell(\omega) \geq b, \forall \ell \in \Lambda_N, |N(\omega)|_+ \leq m = m(N)\}. \quad (3)$$

The constraint restricting the number of positive components of  $N(\omega)$  is a useful technical device that allows us to control the errors in several estimates. In Appendix D of [7] we give evidence supporting the conjecture that this restriction can be eliminated.

When  $b$  is a positive integer, for each  $\omega \in \Omega_{N,b,m}$ , each site in  $\Lambda_N$  is occupied by at least  $b$  particles. In this case it is useful to think of each particle as having one unit of mass and of the set of particles at each site  $\ell$  as defining a droplet. With this interpretation, for each configuration  $\omega$ ,  $K_\ell(\omega)$  denotes the mass or the size of the droplet at site  $\ell$ . The  $j$ th droplet class has  $N_j(\omega)$  droplets and mass  $jN_j(\omega)$ . Because the number of sites in  $\Lambda_N$  equals  $N$  and the sum of the masses of all the droplet classes equals  $K$ , it follows that the quantities  $N_j(\omega)$  satisfy the two conservation laws in (1) for all  $\omega \in \Omega_{N,b,m}$ .

We now consider the modifications that must be made in these definitions when  $b = 0$ . In this case the first constraint in the definition of  $\Omega_{N,b,m}$  disappears because we allow sites to be occupied by 0 particles, and therefore  $N_j(\omega)$  is indexed by  $j \in \mathbb{N}_0 = \mathbb{N} \cup \{0\}$ . On the other hand, we retain the second constraint in the definition of  $\Omega_{N,0,m}$ , which requires that for any configuration  $\omega \in \Omega_{N,0,m}$  at most  $m$  of the components  $N_j(\omega)$  for  $j \in \mathbb{N}_0$  are positive. When  $b = 0$ , the definition of  $\Omega_{N,0,m}$  becomes  $\Omega_{N,0,m} = \{\omega \in \Omega_N : |N(\omega)|_+ \leq m = m(N)\}$ . Because the choice  $b = 0$  allows sites to be empty, we lose the interpretation of the set of particles at each site as being a droplet. However, for  $\omega \in \Omega_{N,0,m}$  the two conservation laws in (1) continue to hold.

For the remainder of this paper we work with any fixed nonnegative integer  $b$ . The probability measure  $P_{N,b,m}$  defining the droplet model is obtained by restricting the uniform measure  $P_N$  to the set  $\Omega_{N,b,m}$ . Thus  $P_{N,b,m}$  equals the conditional probability  $P_N(\cdot | \Omega_{N,b,m})$ . For subsets  $A$  of  $\Omega_{N,b,m}$ ,  $P_{N,b,m}(A)$  takes the form

$$\begin{aligned} P_{N,b,m}(A) &= P_N(A | \Omega_{N,b,m}) \\ &= \frac{1}{\text{card}(\Omega_{N,b,m})} \cdot \text{card}(A). \end{aligned} \quad (4)$$

Having defined the droplet model, we introduce the random probability measures whose large deviations we will study. For  $\omega \in \Omega_{N,b,m}$  these measures are the number-density measures  $\Theta_{N,b}$  that assign to  $j \in \mathbb{N}_b$  the probability  $N_j(\omega)/N$ . This ratio represents the number density of droplet class  $j$ . Thus for any subset  $A$  of  $\mathbb{N}_b$

$$\begin{aligned} \Theta_{N,b}(\omega, A) &= \sum_{j \in A} \Theta_{N,b;j}(\omega), \\ \text{where } \Theta_{N,b;j}(\omega) &= \frac{N_j(\omega)}{N}. \end{aligned} \quad (5)$$

By the two formulas in (1)  $\sum_{j \in \mathbb{N}_b} \Theta_{N,b;j}(\omega) = 1$  and  $\sum_{j \in \mathbb{N}_b} j\Theta_{N,b;j}(\omega) = K/N = c$ . Thus  $\Theta_{N,b}(\omega)$  is a probability measure on  $\mathbb{N}_b$  having mean  $c$ .

We next introduce several spaces of probability measures that arise in the large deviation analysis of the droplet model.  $\mathcal{P}_{\mathbb{N}_b}$  denotes the set of probability measures on  $\mathbb{N}_b = \{n \in \mathbb{Z} : n \geq b\}$ . Thus  $\theta \in \mathcal{P}_{\mathbb{N}_b}$  has the form  $\sum_{j \in \mathbb{N}_b} \theta_j \delta_j$ , where the components  $\theta_j$  satisfy  $\theta_j \geq 0$  and  $\sum_{j \in \mathbb{N}_b} \theta_j = 1$ . We say that a sequence of measures  $\{\theta^{(n)}, n \in \mathbb{N}\}$  in  $\mathcal{P}_{\mathbb{N}_b}$  converges weakly to  $\theta \in \mathcal{P}_{\mathbb{N}_b}$ , and write  $\theta^{(n)} \Rightarrow \theta$ , if, for any bounded function  $f$  mapping  $\mathbb{N}_b$  into  $\mathbb{R}$ ,  $\int_{\mathbb{N}_b} f d\theta^{(n)} \rightarrow \int_{\mathbb{N}_b} f d\theta$  as  $n \rightarrow \infty$ .

$\mathcal{P}_{\mathbb{N}_b}$  is topologized by the topology of weak convergence. There is a standard technique for introducing a metric structure on  $\mathcal{P}_{\mathbb{N}_b}$  for which we quote the main facts. Because  $\mathbb{N}_b$  is a complete, separable metric space with metric  $d(x, y) = |x - y|$ , there exists a metric  $\pi$  on  $\mathcal{P}_{\mathbb{N}_b}$  called the Prohorov metric with the following two properties: (1) convergence with respect to the Prohorov metric is equivalent to weak convergence [32, Thm. 3.3.1]; (2) with respect to the Prohorov metric,  $\mathcal{P}_{\mathbb{N}_b}$  is a complete, separable metric space [32, Thm. 3.1.7].

We denote by  $\mathcal{P}_{\mathbb{N}_b, c}$  the set of measures in  $\mathcal{P}_{\mathbb{N}_b}$  having mean  $c$ . Thus  $\theta \in \mathcal{P}_{\mathbb{N}_b, c}$  has the form  $\sum_{j \in \mathbb{N}_b} \theta_j \delta_j$ , where the components  $\theta_j$  satisfy  $\theta_j \geq 0$ ,  $\sum_{j \in \mathbb{N}_b} \theta_j = 1$ , and  $\sum_{j \in \mathbb{N}_b} j \theta_j = c$ . The number-density measures  $\Theta_{N, b}$  defined in (5) take values in  $\mathcal{P}_{\mathbb{N}_b, c}$ .

According to part (a) of Theorem 4,  $\mathcal{P}_{\mathbb{N}_b, c}$  is not a closed subset of  $\mathcal{P}_{\mathbb{N}_b}$ . Hence it is natural to introduce the closure of  $\mathcal{P}_{\mathbb{N}_b, c}$  in  $\mathcal{P}_{\mathbb{N}_b}$ . As we prove in part (b) of Theorem 4, the closure of  $\mathcal{P}_{\mathbb{N}_b, c}$  in  $\mathcal{P}_{\mathbb{N}_b}$  equals  $\mathcal{P}_{\mathbb{N}_b, [b, c]}$ , which is the set of measures in  $\mathcal{P}_{\mathbb{N}_b}$  having mean lying in the closed interval  $[b, c]$ . Being the closure of the relatively compact, separable metric space  $\mathcal{P}_{\mathbb{N}_b, c}$ ,  $\mathcal{P}_{\mathbb{N}_b, [b, c]}$  is a compact, separable metric space with respect to the Prohorov metric. This space appears in the formulation of the large deviation upper bound in part (c) of Theorem 1.

We next state Theorem 1, which is the LDP for the sequence of distributions  $P_{N, b, m}(\Theta_{N, b} \in d\theta)$  on  $\mathcal{P}_{\mathbb{N}_b, c}$  as  $N \rightarrow \infty$ . The rate function in the LDP is the relative entropy of  $\theta$  with respect to the Poisson distribution  $\rho_{b, \alpha_b(c)} = \sum_{j \in \mathbb{N}_b} \rho_{b, \alpha_b(c); j} \delta_j$  defined in (7), where each  $\rho_{b, \alpha_b(c); j} > 0$ . Thus any  $\theta \in \mathcal{P}_{\mathbb{N}_b, c}$  is absolutely continuous with respect to  $\rho_{b, \alpha_b(c)}$ . For  $\theta \in \mathcal{P}_{\mathbb{N}_b, c}$  the relative entropy of  $\theta$  with respect to  $\rho_{b, \alpha_b(c)}$  is defined by

$$R(\theta \mid \rho_{b, \alpha_b(c)}) = \sum_{j \in \mathbb{N}_b} \theta_j \log \left( \frac{\theta_j}{\rho_{b, \alpha_b(c); j}} \right). \quad (6)$$

If  $\theta_j = 0$ , then  $\theta_j \log(\theta_j / \rho_{b, \alpha_b(c); j}) = 0$ . For  $j \in \mathbb{N}_b$  the components of the measure  $\rho_{b, \alpha_b(c)}$  appearing in the LDP have the form

$$\rho_{b, \alpha_b(c); j} = \frac{1}{Z_b(\alpha_b(c))} \cdot \frac{[\alpha_b(c)]^j}{j!}, \quad (7)$$

where  $\alpha_b(c) \in (0, \infty)$  is chosen so that  $\rho_{b, \alpha_b(c)}$  has mean  $c$  and  $Z_b(\alpha_b(c))$  is the normalization making  $\rho_{b, \alpha_b(c)}$  a probability measure; thus  $Z_0(\alpha_0(c)) = e^{\alpha_0(c)}$  and, for  $b \in \mathbb{N}$ ,  $Z_b(\alpha_b(c)) = e^{\alpha_b(c)} - \sum_{j=0}^{b-1} [\alpha_b(c)]^j / j!$ . As we show in Theorem A.2, there exists a unique value of  $\alpha_b(c)$ .

As a consequence of the fact that  $\mathcal{P}_{\mathbb{N}_b, c}$  is not closed in  $\mathcal{P}_{\mathbb{N}_b}$ , the large deviation upper bound takes two forms depending on whether the subset  $F$  of  $\mathcal{P}_{\mathbb{N}_b, c}$  is compact or whether  $F$  is closed. When  $F$  is compact, in part (b) we obtain the standard large deviation upper bound for  $F$ . When  $F$  is closed, in part (c) we obtain a variation of the standard large deviation upper bound, which, when  $F$  is compact, coincides with the upper bound in part (b). The refinement in part (c) is important. It is applied in the proof of Theorem 2 to show that

$\rho_{b, \alpha_b(c)}$  is the equilibrium distribution of the number-density measures  $\Theta_{N, b}$ . In turn, Theorem 2 is applied in the proof of Corollary 3 to show that  $\rho_{b, \alpha_b(c)}$  is the equilibrium distribution of the droplet-size random variables  $K_\ell$ .

In the next theorem we assume that  $m$  is the function  $m(N)$  appearing in the definition of  $\Omega_{N, b, m}$  in (3) and satisfying  $m(N) \rightarrow \infty$  and  $m(N)^2/N \rightarrow 0$  as  $N \rightarrow \infty$ . The assumption that  $m(N)^2/N \rightarrow 0$  is used to control error terms in Lemmas 6 and 7 in the present paper and in Lemma B.3 in [7]. This assumption on  $m(N)$  is optimal in the sense that it is a minimal assumption guaranteeing that error terms in parts (a) and (b) of Lemma B.3 in [7] converge to 0. In the next theorem, for  $A$  a subset of  $\mathcal{P}_{\mathbb{N}_b, c}$  or  $\mathcal{P}_{\mathbb{N}_b, [b, c]}$  we denote by  $R(A \mid \rho_{b, \alpha_b(c)})$  the infimum of  $R(\theta \mid \rho_{b, \alpha_b(c)})$  over  $\theta \in A$ .

**Theorem 1.** Fix a nonnegative integer  $b$  and a rational number  $c \in (b, \infty)$ . Let  $m$  be the function  $m(N)$  appearing in the definition of  $\Omega_{N, b, m}$  in (3) and satisfying  $m(N) \rightarrow \infty$  and  $m(N)^2/N \rightarrow 0$  as  $N \rightarrow \infty$ . Let  $\rho_{b, \alpha_b(c)} \in \mathcal{P}_{\mathbb{N}_b, c}$  be the distribution having the components defined in (7). Then as  $N \rightarrow \infty$ , with respect to the measures  $P_{N, b, m}$ , the sequence  $\Theta_{N, b}$  satisfies the LDP on  $\mathcal{P}_{\mathbb{N}_b, c}$  with rate function  $R(\theta \mid \rho_{b, \alpha_b(c)})$  in the following sense.

(a)  $R(\theta \mid \rho_{b, \alpha_b(c)})$  maps  $\mathcal{P}_{\mathbb{N}_b, c}$  into  $[0, \infty]$  and has compact level sets in  $\mathcal{P}_{\mathbb{N}_b, c}$ ; that is, for any  $M < \infty$  the set  $\{\theta \in \mathcal{P}_{\mathbb{N}_b, c} : R(\theta \mid \rho_{b, \alpha_b(c)}) \leq M\}$  is compact.

(b) For any compact subset  $F$  of  $\mathcal{P}_{\mathbb{N}_b, c}$  we have the large deviation upper bound

$$\limsup_{N \rightarrow \infty} \frac{1}{N} \log P_{N, b, m}(\Theta_{N, b} \in F) \leq -R(F \mid \rho_{b, \alpha_b(c)}). \quad (8)$$

(c) For any closed subset  $F$  of  $\mathcal{P}_{\mathbb{N}_b, c}$ , let  $\bar{F}$  denote the closure of  $F$  in  $\mathcal{P}_{\mathbb{N}_b, [b, c]}$ . We have the large deviation upper bound

$$\limsup_{N \rightarrow \infty} \frac{1}{N} \log P_{N, b, m}(\Theta_{N, b} \in F) \leq -R(\bar{F} \mid \rho_{b, \alpha_b(c)}). \quad (9)$$

(d) For any open subset  $G$  of  $\mathcal{P}_{\mathbb{N}_b, c}$  we have the large deviation lower bound

$$\liminf_{N \rightarrow \infty} \frac{1}{N} \log P_{N, b, m}(\Theta_{N, b} \in G) \geq -R(G \mid \rho_{b, \alpha_b(c)}). \quad (10)$$

The properties of  $R(\theta \mid \rho_{b, \alpha_b(c)})$  in part (a) are proved in [33, Lem. 1.4.1] and part (a) of Theorem A.1. The basic step in proving the large deviation bounds in parts (b)–(d) is the local large deviation estimate in part (b) of Theorem 5. As explained in Section 4, this local estimate is lifted to large deviation limits involving open balls stated in Theorem 8, which in turn are used to derive the bounds in parts (b)–(d) of Theorem 1.

In the next theorem we use the large deviation upper bound in part (c) of Theorem 1 to prove that the Poisson distribution  $\rho_{b, \alpha_b(c)}$  is the equilibrium distribution of the number-density measures  $\Theta_{N, b}$ . In this theorem  $[B_\pi(\rho_{b, \alpha_b(c)}, \varepsilon)]^c$  denotes the complement in  $\mathcal{P}_{\mathbb{N}_b, c}$  of the open

ball  $B_\pi(\rho_{b,\alpha_b(c)}, \varepsilon) = \{\nu \in \mathcal{P}_{\mathbb{N}_b, c} : \pi(\rho_{b,\alpha_b(c)}, \nu) < \varepsilon\}$ .  $[\widehat{B}_\pi(\rho_{b,\alpha_b(c)}, \varepsilon)]^c$  denotes the complement in  $\mathcal{P}_{\mathbb{N}_b, [b, c]}$  of the open ball  $\widehat{B}_\pi(\rho_{b,\alpha_b(c)}, \varepsilon) = \{\nu \in \mathcal{P}_{\mathbb{N}_b, [b, c]} : \pi(\rho_{b,\alpha_b(c)}, \nu) < \varepsilon\}$ .

**Theorem 2.** *One assumes the hypotheses of Theorem 1. The following results hold for any  $\varepsilon > 0$ .*

- (a) *The quantity  $x^* = \inf\{R(\theta \mid \rho_{b,\alpha_b(c)}) : \theta \in [\widehat{B}_\pi(\rho_{b,\alpha_b(c)}, \varepsilon)]^c\}$  is strictly positive.*
- (b) *For any number  $y$  in the interval  $(0, x^*)$  and all sufficiently large  $N$*

$$P_{N,b,m}(\Theta_{N,b} \in [B_\pi(\rho_{b,\alpha_b(c)}, \varepsilon)]^c) \leq \exp[-Ny] \quad (11)$$

$$\longrightarrow 0 \text{ as } N \longrightarrow \infty.$$

*This upper bound implies that, as  $N \rightarrow \infty$ ,  $P_{N,b,m}(\Theta_{N,b} \in B_\pi(\rho_{b,\alpha_b(c)}, \varepsilon)) \rightarrow 1$  and for any bounded, continuous function  $g$  mapping  $\mathcal{P}_{\mathbb{N}_b, c}$  into  $\mathbb{R}$*

$$\lim_{N \rightarrow \infty} \int_{\Omega_{N,b,m}} g(\Theta_{N,b}) dP_{N,b,m} = g(\rho_{b,\alpha_b(c)}). \quad (12)$$

*These two limits allow us to interpret the Poisson distribution  $\rho_{b,\alpha_b(c)}$  as the equilibrium distribution of the number-density measures  $\Theta_{N,b}$  with respect to  $P_{N,b,m}$ .*

*Proof.* The starting point is the large deviation upper bound in part (c) of Theorem 1 applied to the closed set  $[B_\pi(\rho_{b,\alpha_b(c)}, \varepsilon)]^c$ , which is a subset of  $[\widehat{B}_\pi(\rho_{b,\alpha_b(c)}, \varepsilon)]^c$ . We denote the closure of  $[B_\pi(\rho_{b,\alpha_b(c)}, \varepsilon)]^c$  in  $\mathcal{P}_{\mathbb{N}_b, [b, c]}$  by  $[\widehat{B}_\pi(\rho_{b,\alpha_b(c)}, \varepsilon)]^c$ . Since  $[B_\pi(\rho_{b,\alpha_b(c)}, \varepsilon)]^c \subset [\widehat{B}_\pi(\rho_{b,\alpha_b(c)}, \varepsilon)]^c$ , the large deviation upper bound in part (c) of Theorem 1 takes the form

$$\begin{aligned} \limsup_{N \rightarrow \infty} \frac{1}{N} \log P_{N,b,m}(\Theta_{N,b} \in [B_\pi(\rho_{b,\alpha_b(c)}, \varepsilon)]^c) \\ \leq -R\left(\overline{[B_\pi(\rho_{b,\alpha_b(c)}, \varepsilon)]^c} \mid \rho_{b,\alpha_b(c)}\right) \\ \leq -R\left([\widehat{B}_\pi(\rho_{b,\alpha_b(c)}, \varepsilon)]^c \mid \rho_{b,\alpha_b(c)}\right). \end{aligned} \quad (13)$$

We now prove part (a) of Theorem 2. Since  $R(\theta \mid \rho_{b,\alpha_b(c)})$  is lower semicontinuous on  $\mathcal{P}_{\mathbb{N}_b, [b, c]}$  and has compact level sets in  $\mathcal{P}_{\mathbb{N}_b, [b, c]}$  [33, Lem. 1.4.3(b)–(c)], it attains its infimum  $x^*$  on the closed set  $[\widehat{B}_\pi(\rho_{b,\alpha_b(c)}, \varepsilon)]^c$ . If  $x^* = 0$ , then there would exist  $\theta \in [\widehat{B}_\pi(\rho_{b,\alpha_b(c)}, \varepsilon)]^c$  such that  $R(\theta \mid \rho_{b,\alpha_b(c)}) = 0$ . But on  $\mathcal{P}_{\mathbb{N}_b, [b, c]}$ ,  $R(\theta \mid \rho_{b,\alpha_b(c)})$  attains its infimum of 0 at the unique measure  $\theta = \rho_{b,\alpha_b(c)}$  [33, Lem. 1.4.1]. This contradicts the fact that  $\rho_{b,\alpha_b(c)} \notin [\widehat{B}_\pi(\rho_{b,\alpha_b(c)}, \varepsilon)]^c$ , completing the proof of part (a). The inequality in part (b) is an immediate consequence of part (a) and the large deviation upper bound (13). This inequality yields the limit  $P_{N,b,m}(\Theta_{N,b} \in B_\pi(\rho_{b,\alpha_b(c)}, \varepsilon)) \rightarrow 1$ , which in turn implies (12). The proof of Theorem 2 is complete.  $\square$

We now apply Theorem 2 to prove that  $\rho_{b,\alpha_b(c)}$  is also the equilibrium distribution of the random variables  $K_\ell$ , which

count the droplet sizes at the sites of  $\Lambda_N$ . This is the content of the next corollary. A fact needed in the proof is that  $\Theta_{N,b}$  is the empirical measure of these random variables; that is, for  $\omega \in \Omega_{N,b,m}$ ,  $\Theta_{N,b}(\omega)$  assigns to subsets  $A$  of  $\mathbb{N}_b$  the probability  $\Theta_{N,b}(\omega, A) = N^{-1} \sum_{\ell=1}^N \delta_{K_\ell(\omega)}(A)$ . This representation is valid because both  $\Theta_{N,b}(\omega)$  and the empirical measure assign to  $j \in \Lambda_N$  the probability  $N_j(\omega)/N$ .

**Corollary 3.** *One assumes the hypotheses of Theorem 1. Then for any site  $\ell \in \Lambda_N$  and any  $j \in \mathbb{N}_b$*

$$\begin{aligned} \lim_{N \rightarrow \infty} P_{N,b,m}(K_\ell = j) &= \rho_{b,\alpha_b(c);j} \\ &= \frac{1}{Z_b(\alpha_b(c))} \cdot \frac{[\alpha_b(c)]^j}{j!}. \end{aligned} \quad (14)$$

*Proof.* Since the random variables  $K_\ell$  are identically distributed, it suffices to prove the corollary for  $\ell = 1$ . For fixed  $j \in \mathbb{N}_b$ , the limit (12) with  $g(\theta) = \theta_j$  yields

$$\begin{aligned} \lim_{N \rightarrow \infty} P_{N,b,m}(K_1 = j) \\ = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{\ell=1}^N \int_{\Omega_{N,b,m}} 1_j(K_\ell) dP_{N,b,m} \\ = \lim_{N \rightarrow \infty} \int_{\Omega_{N,b,m}} \Theta_{N,b;1} dP_{N,b,m} = \rho_{b,\alpha_b(c);j}. \end{aligned} \quad (15)$$

This completes the proof.  $\square$

The last theorem in this section proves several properties of  $\mathcal{P}_{\mathbb{N}_b, c}$  and  $\mathcal{P}_{\mathbb{N}_b, [b, c]}$  with respect to the Prohorov metric that are needed in the paper.

**Theorem 4.** *Fix a nonnegative integer  $b$  and a real number  $c \in (b, \infty)$ . The metric spaces  $\mathcal{P}_{\mathbb{N}_b, c}$  and  $\mathcal{P}_{\mathbb{N}_b, [b, c]}$  have the following properties.*

- (a)  $\mathcal{P}_{\mathbb{N}_b, c}$ , the set of probability measures on  $\mathbb{N}_b$  having mean  $c$ , is a relatively compact, separable subset of  $\mathcal{P}_{\mathbb{N}_b}$ . However,  $\mathcal{P}_{\mathbb{N}_b, c}$  is not a closed subset of  $\mathcal{P}_{\mathbb{N}_b}$  and thus is not a compact subset or a complete metric space.
- (b)  $\mathcal{P}_{\mathbb{N}_b, [b, c]}$ , the set of probability measures on  $\mathbb{N}_b$  having mean lying in the closed interval  $[b, c]$ , is the closure of  $\mathcal{P}_{\mathbb{N}_b, c}$  in  $\mathcal{P}_{\mathbb{N}_b}$ .  $\mathcal{P}_{\mathbb{N}_b, [b, c]}$  is a compact, separable subset of  $\mathcal{P}_{\mathbb{N}_b}$ .

*Proof.* (a) For  $\xi \in \mathbb{N}$  satisfying  $\xi \geq b$  let  $\Psi_\xi$  denote the compact subset  $\{b, b+1, \dots, \xi\}$  of  $\mathbb{N}_b$ , and let  $[\Psi_\xi]^c$  denote its complement. For any  $\theta \in \mathcal{P}_{\mathbb{N}_b, c}$

$$c = \sum_{j \in \mathbb{N}_b} j\theta_j \geq \sum_{j \geq \xi+1} j\theta_j \geq \xi \sum_{j \geq \xi+1} \theta_j = \xi\theta([\Psi_\xi]^c). \quad (16)$$

It follows that  $\mathcal{P}_{\mathbb{N}_b, c}$  is tight; that is, for any  $\varepsilon > 0$  there exists  $\xi \in \mathbb{N}$  such that  $\theta([\Psi_\xi]^c) < \varepsilon$  for all  $\theta \in \mathcal{P}_{\mathbb{N}_b, c}$ . Prohorov's theorem implies that  $\mathcal{P}_{\mathbb{N}_b, c}$  is relatively compact [32, Thm. 3.2.2]. The separability of  $\mathcal{P}_{\mathbb{N}_b, c}$  is proved in Corollary B.2 in [7].



We now prove that  $\mathcal{P}_{\mathbb{N}_b, c}$  is not a closed subset of  $\mathcal{P}_{\mathbb{N}_b}$  by exhibiting a sequence  $\theta^{(n)} \in \mathcal{P}_{\mathbb{N}_b, c}$  having a weak limit that does not lie in  $\mathcal{P}_{\mathbb{N}_b, c}$ . Let  $\theta$  be any measure in  $\mathcal{P}_{\mathbb{N}_b}$  with mean  $\beta \in [b, c]$ ; thus  $\theta \notin \mathcal{P}_{\mathbb{N}_b, c}$ . The sequence

$$\theta^{(n)} = \frac{n-c}{n-\beta}\theta + \frac{c-\beta}{n-\beta}\delta_n \quad \text{for } n \in \mathbb{N}, n > c \quad (17)$$

has the property that  $\theta^{(n)} \in \mathcal{P}_{\mathbb{N}_b, c}$  and that  $\theta^{(n)} \Rightarrow \theta \notin \mathcal{P}_{\mathbb{N}_b, c}$ . This completes the proof of part (a).

(b) Since  $\mathcal{P}_{\mathbb{N}_b, c}$  is a separable subset of  $\mathcal{P}_{\mathbb{N}_b}$  and  $\mathcal{P}_{\mathbb{N}_b, c}$  is dense in  $\mathcal{P}_{\mathbb{N}_b, [b, c]}$ , it follows that  $\mathcal{P}_{\mathbb{N}_b, [b, c]}$  is separable. We prove that  $\mathcal{P}_{\mathbb{N}_b, [b, c]}$  is the closure of  $\mathcal{P}_{\mathbb{N}_b, c}$  in  $\mathcal{P}_{\mathbb{N}_b}$ . Let  $\theta^{(n)}$  be a sequence in  $\mathcal{P}_{\mathbb{N}_b, c}$  converging weakly to  $\theta \in \mathcal{P}_{\mathbb{N}_b}$ . Since  $\theta^{(n)} \Rightarrow \theta$  implies that  $\theta_j^{(n)} \rightarrow \theta_j$  for each  $j \in \mathbb{N}_b$ , Fatou's lemma implies that  $c = \liminf_{n \rightarrow \infty} \langle \theta^{(n)} \rangle \geq \langle \theta \rangle$ , where  $\langle \theta^{(n)} \rangle$  and  $\langle \theta \rangle$  denote the means of  $\theta^{(n)}$  and  $\theta$ . Since for any  $\theta \in \mathcal{P}_{\mathbb{N}_b}$  we have  $\langle \theta \rangle \geq b$ , it follows that  $c \geq \langle \theta \rangle \geq b$ . This shows that the closure of  $\mathcal{P}_{\mathbb{N}_b, c}$  in  $\mathcal{P}_{\mathbb{N}_b}$  is a subset of  $\mathcal{P}_{\mathbb{N}_b, [b, c]}$ . We next prove that  $\mathcal{P}_{\mathbb{N}_b, [b, c]}$  is a subset of the closure of  $\mathcal{P}_{\mathbb{N}_b, c}$  in  $\mathcal{P}_{\mathbb{N}_b}$  by showing that for any  $\theta \in \mathcal{P}_{\mathbb{N}_b, [b, c]}$  there exists a sequence  $\theta^{(n)} \in \mathcal{P}_{\mathbb{N}_b, c}$  such that  $\theta^{(n)} \Rightarrow \theta$ . If  $\langle \theta \rangle = c$ , then we choose  $\theta^{(n)} = \theta$  for all  $n \in \mathbb{N}$ . If  $\langle \theta \rangle = \beta \in [b, c)$ , then we use the sequence  $\theta^{(n)}$  in (17), which converges weakly to  $\theta$ . We conclude that  $\theta$  lies in the closure of  $\mathcal{P}_{\mathbb{N}_b, c}$  and thus that  $\mathcal{P}_{\mathbb{N}_b, [b, c]}$  is a subset of the closure of  $\mathcal{P}_{\mathbb{N}_b, c}$  in  $\mathcal{P}_{\mathbb{N}_b}$ . This completes the proof of part (b). The proof of Theorem 4 is done.  $\square$

In the next section we present the local large deviation estimate that will be used in Section 4 to prove the LDP for  $\Theta_{N, b}$  in Theorem 1.

### 3. Local Large Deviation Estimate Yielding Theorem 1

The main result needed to prove the LDP in Theorem 1 is the local large deviation estimate stated in part (b) of Theorem 5. The first step is to introduce a set  $A_{N, b, m}$  that plays a central role in this paper. Fix a nonnegative integer  $b$  and a rational number  $c \in (b, \infty)$ . Given  $N \in \mathbb{N}$  define  $K = Nc$  and let  $m$  be the function appearing in the definition of  $\Omega_{N, b, m}$  in (3) and satisfying  $m(N) \rightarrow \infty$  and  $m(N)^2/N \rightarrow 0$  as  $N \rightarrow \infty$ . Define  $\mathbb{N}_b = \{n \in \mathbb{Z} : n \geq b\}$ ; thus  $\mathbb{N}_0$  is the set of nonnegative integers. Let  $\nu$  be a sequence  $\{\nu_j, j \in \mathbb{N}_b\}$  for which each  $\nu_j \in \mathbb{N}_0$ ; thus  $\nu \in \mathbb{N}_0^{\mathbb{N}_b}$ . We define  $A_{N, b, m}$  to be the set of  $\nu \in \mathbb{N}_0^{\mathbb{N}_b}$  satisfying

$$\begin{aligned} \sum_{j \in \mathbb{N}_b} \nu_j &= N, \\ \sum_{j \in \mathbb{N}_b} j \nu_j &= K, \end{aligned} \quad (18)$$

$$|\nu|_+ \leq m = m(N),$$

where  $|\nu|_+ = \text{card}\{j \in \mathbb{N}_b : \nu_j \geq 1\}$ . Because  $\nu_j \in \mathbb{N}_0$ , the two sums involve only finitely many terms.

For  $\omega \in \Omega_{N, b, m}$  the components  $\Theta_{N, b, j}(\omega)$  of the number-density measure defined in (5) are  $N_j(\omega)/N$  for  $j \in \mathbb{N}_b$ , where  $N_j(\omega)$  denotes the number of sites in  $\Lambda_N$  containing  $j$  particles in the configuration  $\omega$ . We denote by  $N(\omega)$  the sequence  $\{N_j(\omega), j \in \mathbb{N}_b\}$ . By definition, for every  $\omega \in \Omega_{N, b, m}$  each site  $\ell \in \Lambda_N$  is occupied by at least  $b$  particles, and  $|N(\omega)|_+ \leq m = m(N)$ . It follows that  $A_{N, b, m}$  is the range of  $N(\omega)$  for  $\omega \in \Omega_{N, b, m}$ ; the two sums involving  $\nu_j$  in (18) correspond to the two sums involving  $N_j(\omega)$  in (1).

Since the range of  $N(\omega)$  is  $A_{N, b, m}$ , for  $\omega \in \Omega_{N, b, m}$  the range of  $\Theta_{N, b}(\omega)$  is the set of probability measures  $\theta_{N, b, \nu}$  whose components for  $j \in \mathbb{N}_b$  have the form  $\theta_{N, b, \nu, j} = \nu_j/N$  for  $\nu \in A_{N, b, m}$ . By (18)  $\theta_{N, b, \nu}$  takes values in  $\mathcal{P}_{\mathbb{N}_b, c}$ , the set of probability measures on  $\mathbb{N}_b$  having mean  $c$ . It follows that the set

$$\begin{aligned} B_{N, b, m} &= \left\{ \theta \in \mathcal{P}_{\mathbb{N}_b, c} : \theta_j = \frac{\nu_j}{N} \text{ for } j \in \mathbb{N}_b \text{ for some } \nu \right. \\ &\quad \left. \in A_{N, b, m} \right\} \end{aligned} \quad (19)$$

is the range of  $\Theta_{N, b}(\omega)$  for  $\omega \in \Omega_{N, b, m}$ .

In part (b) of the next theorem we state the local large deviation estimate for the event  $\{\Theta_{N, b} = \theta_{N, b, \nu}\}$ . In part (a) we introduce the Poisson distribution  $\rho_{b, \alpha_b(c)}$  that appears in the local estimate;  $\rho_{b, \alpha_b(c)}$  is defined in terms of a parameter  $\alpha_b(c)$  guaranteeing that it has mean  $c$ .

In part (a) of Theorem C.2 in [7] we give the straightforward proof of the existence of  $\alpha_b(c)$  for  $b = 1$ . The proof of the existence of  $\alpha_b(c)$  for general  $b \in \mathbb{N}$  is much more subtle than the proof for  $b = 1$ . The proof for general  $b \in \mathbb{N}$  is given in Theorem A.2 in the present paper.

**Theorem 5.** (a) Fix a nonnegative integer  $b$  and a real number  $c \in (b, \infty)$ . For  $\alpha \in (0, \infty)$  let  $\rho_{b, \alpha}$  be the measure on  $\mathbb{N}_b$  having components  $\rho_{b, \alpha, j} = [Z_b(\alpha)]^{-1} \cdot \alpha^j / j!$  for  $j \in \mathbb{N}_b$ , where  $Z_{0, \alpha} = e^\alpha$ , and, for  $b \in \mathbb{N}$ ,  $Z_b(\alpha) = e^\alpha - \sum_{j=0}^{b-1} \alpha^j / j!$ . Then there exists a unique value  $\alpha_b(c) \in (0, \infty)$  such that  $\rho_{b, \alpha_b(c)}$  lies in the set  $\mathcal{P}_{\mathbb{N}_b, c}$  of probability measures on  $\mathbb{N}_b$  having mean  $c$ . If  $b = 0$ , then  $\alpha_0(c) = c$ . If  $b \in \mathbb{N}$ , then  $\alpha_b(c)$  is the unique solution in  $(0, \infty)$  of  $\alpha Z_{b-1}(\alpha) / Z_b(\alpha) = c$ .

(b) Fix a nonnegative integer  $b$  and a rational number  $c \in (b, \infty)$ . Let  $m$  be the function  $m(N)$  appearing in the definition of  $\Omega_{N, b, m}$  in (3) and satisfying  $m(N) \rightarrow \infty$  and  $m(N)^2/N \rightarrow 0$  as  $N \rightarrow \infty$ . For any  $\nu \in A_{N, b, m}$  we define  $\theta_{N, b, \nu} \in \mathcal{P}_{\mathbb{N}_b, c}$  to have the components  $\theta_{N, b, \nu, j} = \nu_j/N$  for  $j \in \mathbb{N}_b$ . Then

$$\frac{1}{N} \log P_{N, b, m}(\Theta_{N, b} = \theta_{N, b, \nu})$$

$$= -R(\theta_{N, b, \nu} \mid \rho_{b, \alpha_b(c)}) + \varepsilon_N(\nu).$$

$R(\theta_{N, b, \nu} \mid \rho_{b, \alpha_b(c)})$  is finite because it involves only finitely many components of  $\theta_{N, b, \nu}$ , and  $\varepsilon_N(\nu) \rightarrow 0$  uniformly for  $\nu \in A_{N, b, m}$  as  $N \rightarrow \infty$ .

We now prove the local large deviation estimate in part (b) of Theorem 5. This proof is based on a combinatorial argument that is reminiscent of and is as natural as

the combinatorial argument used to prove Sanov's theorem for empirical measures defined in terms of i.i.d. random variables having a finite state space [1, §3]. Part (b) of Theorem 5 is proved by analyzing the asymptotic behavior of the product of two multinomial coefficients that we now introduce.

Given  $\nu \in A_{N,b,m}$ , our goal is to estimate the probability  $P_{N,b,m}(\Theta_{N,b} = \theta_{N,b,\nu})$ , where  $\theta_{N,b,\nu}$  has the components  $\theta_{N,b,\nu;j} = \nu_j/N$  for  $j \in \mathbb{N}_b$ . A basic observation is that  $\{\omega \in \Omega_{N,b,m} : \Theta_{N,b}(\omega) = \theta_{N,b,\nu}\}$  coincides with

$$\Delta_{N,b,m;\nu} = \{\omega \in \Omega_{N,b,m} : N_j(\omega) = \nu_j \text{ for } j \in \mathbb{N}_b\}. \quad (21)$$

It follows that

$$\begin{aligned} P_{N,b,m}(\Theta_{N,b} = \theta_{N,b,\nu}) &= P_{N,b,m}(\Delta_{N,b,m;\nu}) \\ &= \frac{1}{\text{card}(\Omega_{N,b,m})} \cdot \text{card}(\Delta_{N,b,m;\nu}). \end{aligned} \quad (22)$$

Our first task is to determine the asymptotic behavior of  $\text{card}(\Delta_{N,b,m;\nu})$ . In determining the asymptotic behavior of  $\text{card}(\Omega_{N,b,m})$ , we will use the fact that  $\Omega_{N,b,m}$  can be written as the disjoint union

$$\Omega_{N,b,m} = \bigcup_{\nu \in A_{N,b,m}} \Delta_{N,b,m;\nu}. \quad (23)$$

Let  $\nu \in A_{N,b,m}$  be given. We start by expressing the cardinality of  $\text{card}(\Delta_{N,b,m;\nu})$  as a product of two multinomial coefficients. For each configuration  $\omega \in \Delta_{N,b,m;\nu}$ ,  $K$  particles are distributed onto the  $N$  sites of the lattice  $\Lambda_N$  with  $j$  particles going onto  $\nu_j$  sites for  $j \in \mathbb{N}_b$ . We carry this out in two stages. In stage one  $K$  particles are placed into  $N$  bins,  $\nu_j$  of which have  $j$  particles for  $j \in \mathbb{N}_b$ . The number of ways of making this placement equals the multinomial coefficient  $K!/\prod_{j \in \mathbb{N}_b} (j!)^{\nu_j}$ . This multinomial coefficient is well-defined since  $\sum_{j \in \mathbb{N}_b} j\nu_j = K$ . Given this placement of  $K$  particles into  $N$  bins, the number of ways of moving the particles from the bins onto the sites  $1, 2, \dots, N$  of the lattice  $\Lambda_N$  equals the multinomial coefficient  $N!/\prod_{j \in \mathbb{N}_b} \nu_j!$ . This second multinomial coefficient is well-defined since  $\sum_{j \in \mathbb{N}_b} \nu_j = N$ . We conclude that the cardinality of  $\Delta_{N,b,m;\nu}$  is given by the product of these two multinomial coefficients:

$$\text{card}(\Delta_{N,b,m;\nu}) = \frac{N!}{\prod_{j \in \mathbb{N}_b} \nu_j!} \cdot \frac{K!}{\prod_{j \in \mathbb{N}_b} (j!)^{\nu_j}}. \quad (24)$$

Since  $|\nu|_+ \leq m$ , at most  $m$  of the components  $\nu_j$  are positive. Such a product of multinomial coefficients is well known in combinatorial analysis [8, Thm. 2.10]. A related version of this formula is derived in Example III.23 of [34]. See also [35, p. 115] and formula (2) in [36, p. 36].

The next two steps in the proof of the local estimate given in part (b) of Theorem 5 are to prove the asymptotic formula for  $\text{card}(\Delta_{N,b,m;\nu})$  in Lemma 6 and the asymptotic formula for  $\text{card}(\Omega_{N,b,m})$  in part (b) of Lemma 7. The proof of Lemma 6 is greatly simplified by a substitution in line 4 of (34). This substitution involves a parameter  $\alpha \in (0, \infty)$ , which, we emphasize, is arbitrary in this lemma. The substitution in line

4 of (34) allows us to express the asymptotic behavior of both  $\text{card}(\Delta_{N,b,m;\nu})$  in Lemma 6 and  $\text{card}(\Omega_{N,b,m})$  in Lemma 7 directly in terms of the relative entropy  $R(\theta_{N,b,\nu} \mid \rho_{b,\alpha})$ , where  $\rho_{b,\alpha}$  is the probability measure on  $\mathbb{N}_b$  having the components defined in part (a) of Theorem 5. One of the major issues in the proof of part (b) of Theorem 5 is to show that the arbitrary parameter  $\alpha$  appearing in Lemmas 6 and 7 must take the value  $\alpha_b(c)$ , which is the unique value of  $\alpha$  guaranteeing that  $\rho_{b,\alpha} \in \mathcal{P}_{\mathbb{N}_b,c}$  [Theorem 5(a)]. We show that  $\alpha$  must equal  $\alpha_b(c)$  after the statement of Lemma 7.

**Lemma 6.** Fix a nonnegative integer  $b$  and a rational number  $c \in (b, \infty)$ . Let  $\alpha$  be any real number in  $(0, \infty)$ , and let  $m$  be the function  $m(N)$  appearing in the definition of  $\Omega_{N,b,m}$  in (3) and satisfying  $m(N) \rightarrow \infty$  and  $m(N)^2/N \rightarrow 0$  as  $N \rightarrow \infty$ . We define

$$f(\alpha, b, c, K) = \log Z_b(\alpha) - c \log \alpha + c \log K - c. \quad (25)$$

For any  $\nu \in A_{N,b,m}$ , we define  $\theta_{N,b,\nu} \in \mathcal{P}_{\mathbb{N}_b,c}$  to have the components  $\theta_{N,b,\nu;j} = \nu_j/N$  for  $j \in \mathbb{N}_b$ . Then

$$\begin{aligned} \frac{1}{N} \log \text{card}(\Delta_{N,b,m;\nu}) \\ = -R(\theta_{N,b,\nu} \mid \rho_{b,\alpha}) + f(\alpha, b, c, K) + \zeta_N(\nu). \end{aligned} \quad (26)$$

The quantity  $\zeta_N(\nu) \rightarrow 0$  uniformly for  $\nu \in A_{N,b,m}$  as  $N \rightarrow \infty$ .

*Proof.* The proof is based on a weak form of Stirling's approximation, which states that, for all  $N \in \mathbb{N}$  satisfying  $N \geq 2$  and for all  $n \in \mathbb{N}$  satisfying  $1 \leq n \leq N$ ,  $1 \leq \log(n!) - (n \log n - n) \leq 2 \log N$ . We summarize the last formula by writing

$$\begin{aligned} \log(n!) &= n \log n - n + O(\log N), \\ \forall N \in \mathbb{N}, N \geq 2, \forall n \in \{1, 2, \dots, N\}. \end{aligned} \quad (27)$$

The term denoted by  $O(\log N)$  satisfies  $1 \leq O(\log N) \leq 2 \log N$ .

To simplify the notation, we rewrite (24) in the form  $\text{card}(\Delta_{N,b,m;\nu}) = M_1(N, \nu) \cdot M_2(K, \nu)$ , where  $M_1(N, \nu)$  denotes the first multinomial coefficient on the right side of (24), and  $M_2(K, \nu)$  denotes the second multinomial coefficient on the right side of (24). We have

$$\begin{aligned} \frac{1}{N} \log \text{card}(\Delta_{N,b,m;\nu}) \\ = \frac{1}{N} \log M_1(N, \nu) + \frac{1}{N} \log M_2(K, \nu). \end{aligned} \quad (28)$$

The asymptotic behavior of the first term on the right side of the last display is easily calculated. Since  $\nu \in A_{N,b,m}$ , there are  $|\nu|_+ \in \{1, 2, \dots, m\}$  positive components  $\nu_j$ . Because of this restriction on the number  $|\nu|_+$  of positive components of  $\nu$ , we are able to control the error in line 3 of (29). We define  $\Psi_N(\nu) = \{j \in \mathbb{N}_b : \nu_j \geq 1\}$ . For each  $j \in \Psi_N(\nu)$ , since the components  $\nu_j$  satisfy  $1 \leq \nu_j \leq N$ , we have



$\log(\nu_j!) = \nu_j \log \nu_j - \nu_j + O(\log N)$  for all  $N \geq 2$ . Using the fact that  $\sum_{j \in \Psi_N(\nu)} \nu_j = N$ , we obtain

$$\begin{aligned} \frac{1}{N} \log M_1(N, \nu) &= \frac{1}{N} \log(N!) - \frac{1}{N} \sum_{j \in \Psi_N(\nu)} \log(\nu_j!) \\ &= \frac{1}{N} (N \log N - N + O(\log N)) \\ &\quad - \frac{1}{N} \sum_{j \in \Psi_N(\nu)} (\nu_j \log \nu_j - \nu_j + O(\log N)) \\ &= - \sum_{j \in \mathbb{N}_b} \theta_{N,b,\nu;j} \log \theta_{N,b,\nu;j} + \zeta_N^{(1)} - \zeta_N^{(2)}(\nu), \end{aligned} \quad (29)$$

where  $\zeta_N^{(1)} = [O(\log N)]/N \rightarrow 0$  as  $N \rightarrow \infty$  and  $\zeta_N^{(2)}(\nu) = N^{-1} \sum_{j \in \Psi_N(\nu)} O(\log N)$ . By the inequality noted after (27) and the fact that  $|\nu|_+ \leq m$

$$\begin{aligned} 0 \leq \max_{\nu \in A_{N,b,m}} \zeta_N^{(2)}(\nu) &\leq \max_{\nu \in A_{N,b,m}} \frac{2}{N} \sum_{j \in \Psi_N(\nu)} \log N \\ &\leq \frac{2m \log N}{N}. \end{aligned} \quad (30)$$

Since  $(m \log N)/N \rightarrow 0$  as  $N \rightarrow \infty$ , we conclude that  $\zeta_N^{(2)}(\nu) \rightarrow 0$  uniformly for  $\nu \in A_{N,b,m}$  as  $N \rightarrow \infty$ .

We now study the asymptotic behavior of the second term on the right side of (28). Since  $K = Nc$ , we obtain for all  $K \geq 2$

$$\begin{aligned} \frac{1}{N} \log M_2(K, \nu) &= \frac{1}{N} \log(K!) - \frac{1}{N} \sum_{j \in \mathbb{N}_b} \nu_j \log(j!) \\ &= c \log K - c - \sum_{j \in \mathbb{N}_b} \theta_{N,b,\nu;j} \log(j!) \\ &\quad + \zeta_N^{(3)}, \end{aligned} \quad (31)$$

where  $0 \leq \zeta_N^{(3)} = O(\log K)/N = O(\log N)/N \rightarrow 0$  as  $N \rightarrow \infty$ . The weak form of Stirling's formula is used to rewrite the term  $\log(K!)$  in the last display, but not to rewrite the terms  $\log(j!)$ , which we leave untouched.

Substituting (29) and (31) into (28), we obtain

$$\begin{aligned} \frac{1}{N} \log \text{card}(\Delta_{N,b,m;\nu}) &= \frac{1}{N} \log M_1(N, \nu) + \frac{1}{N} \log M_2(K, \nu) \\ &= - \sum_{j \in \mathbb{N}_b} \theta_{N,b,\nu;j} \log(\theta_{N,b,\nu;j} j!) + c \log K - c \\ &\quad + \zeta_N(\nu). \end{aligned} \quad (32)$$

In this formula  $\zeta_N(\nu) = \zeta_N^{(1)} - \zeta_N^{(2)}(\nu) + \zeta_N^{(3)}$ . As  $N \rightarrow \infty$ ,

$$\max_{\nu \in A_{N,b,m}} |\zeta_N(\nu)| \leq \zeta_N^{(1)} + \max_{\nu \in A_{N,b,m}} \zeta_N^{(2)}(\nu) + \zeta_N^{(3)} \rightarrow 0. \quad (33)$$

We conclude that  $\zeta_N(\nu) \rightarrow 0$  uniformly for  $\nu \in A_{N,b,m}$  as  $N \rightarrow \infty$ .

Now comes the key step, the purpose of which is to express the sum in the next-to-last line of (32) as the relative entropy  $R(\theta_{N,b,\nu;j} \mid \rho_{b,\alpha})$ , where  $\alpha \in (0, \infty)$  is arbitrary. To express the sum in the next-to-last line of (32) as  $R(\theta_{N,b,\nu} \mid \rho_{b,\alpha})$ , we rewrite the sum as shown in line 4 of the next display:

$$\begin{aligned} \frac{1}{N} \log \text{card}(\Delta_{N,b,m;\nu}) &= - \sum_{j \in \mathbb{N}_b} \theta_{N,b,\nu;j} \log(\theta_{N,b,\nu;j} j!) + c \log K - c \\ &\quad + \zeta_N(\nu) \\ &= - \sum_{j \in \mathbb{N}_b} \theta_{N,b,\nu;j} \log\left(\frac{\theta_{N,b,\nu;j}}{\alpha^j / (Z_b(\alpha) \cdot j!)} \cdot \frac{\alpha^j}{Z_b(\alpha)}\right) \\ &\quad + c \log K - c + \zeta_N(\nu) \\ &= -R(\theta_{N,b,\nu} \mid \rho_{b,\alpha}) + \log Z_b(\alpha) - c \log \alpha + c \log K \\ &\quad - c + \zeta_N(\nu) \\ &= -R(\theta_{N,b,\nu} \mid \rho_{b,\alpha}) + f(\alpha, b, c, K) + \zeta_N(\nu). \end{aligned} \quad (34)$$

The facts that  $\sum_{j \in \mathbb{N}_b} \theta_{N,b,\nu;j} = 1$  and  $\sum_{j \in \mathbb{N}_b} j \theta_{N,b,\nu;j} = c$  are used to derive the next-to-last equality. The proof of Lemma 6 is complete.  $\square$

The next step in the proof of the local large deviation estimate in part (b) of Theorem 5 is to prove the asymptotic formula for  $\text{card}(\Omega_{N,b,m})$  stated in part (b) of the next lemma. The proof of this lemma uses Lemma 6 in a fundamental way. After the statement of this lemma we show how to apply it and Lemma 6 to prove part (b) of Theorem 5.

**Lemma 7.** Fix a nonnegative integer  $b$  and a rational number  $c \in (b, \infty)$ . The following conclusions hold:

- (a)  $\lim_{N \rightarrow \infty} N^{-1} \log \text{card}(A_{N,b,m}) = 0$ .
- (b) Let  $\alpha$  be the positive real number in Lemma 6, and let  $m$  be the function  $m(N)$  appearing in the definition of  $\Omega_{N,b,m}$  in (3) and satisfying  $m(N) \rightarrow \infty$  and  $m(N)^2/N \rightarrow 0$  as  $N \rightarrow \infty$ . We define  $f(\alpha, b, c, K) = \log Z_b(\alpha) - c \log \alpha + c \log K - c$ . Then  $R(\theta \mid \rho_{b,\alpha})$  attains its infimum over  $\theta \in \mathcal{P}_{\mathbb{N}_b, c}$  and

$$\begin{aligned} \frac{1}{N} \log \text{card}(\Omega_{N,b,m}) &= f(\alpha, b, c, K) \\ &\quad - \min_{\theta \in \mathcal{P}_{\mathbb{N}_b, c}} R(\theta \mid \rho_{b,\alpha}) + \eta_N. \end{aligned} \quad (35)$$

The quantity  $\eta_N \rightarrow 0$  as  $N \rightarrow \infty$ .

Before proving Lemma 7, we derive the local large deviation estimate in part (b) of Theorem 5 by applying Lemmas 6 and 7. An integral part of the proof is to show how the arbitrary value of  $\alpha \in (0, \infty)$  appearing in these lemmas is replaced by the specific value  $\alpha_b(c)$  appearing in Theorem 5. As in the statement of part (b) of Theorem 5, let  $\nu$  be any

vector in  $A_{N,b,m}$  and define  $\theta_{N,b,\nu} \in \mathcal{P}_{\mathbb{N}_b,c}$  to have the components  $\theta_{N,b,\nu;j} = \nu_j/N$  for  $j \in \mathbb{N}_b$ . By (22)

$$\begin{aligned} & \frac{1}{N} \log P_{N,b,m}(\Theta_{N,b} = \theta_{N,b,\nu}) \\ &= \frac{1}{N} \log \text{card}(\Delta_{N,b,m;\nu}) - \frac{1}{N} \log \text{card}(\Omega_{N,b,m}). \end{aligned} \quad (36)$$

Substituting the asymptotic formula for  $\log \text{card}(\Delta_{N,b,m;\nu})$  derived in Lemma 6 and the asymptotic formula for  $\log \text{card}(\Omega_{N,b,m})$  given in part (b) of Lemma 7 yields

$$\begin{aligned} & \frac{1}{N} \log P_{N,b,m}(\Theta_{N,b} = \theta_{N,b,\nu}) \\ &= -R(\theta_{N,b,\nu} \mid \rho_{b,\alpha}) + \min_{\theta \in \mathcal{P}_{\mathbb{N}_b,c}} R(\theta \mid \rho_{b,\alpha}) + \varepsilon_N(\nu). \end{aligned} \quad (37)$$

The error term  $\varepsilon_N(\nu)$  equals  $\zeta_N(\nu) - \eta_N$ ;  $\zeta_N(\nu)$  is the error term in Lemma 6, and  $\eta_N$  is the error term in Lemma 7. As  $N \rightarrow \infty$ ,  $\zeta_N(\nu) \rightarrow 0$  uniformly for  $\nu \in A_{N,b,m}$ , and  $\eta_N \rightarrow 0$ . It follows that  $\varepsilon_N(\nu) \rightarrow 0$  uniformly for  $\nu \in A_{N,b,m}$  as  $N \rightarrow \infty$ .

We now consider the first two terms on the right side of (37). By part (b) of Theorem A.1 applied to  $\theta = \theta_{N,b,\nu} \in \mathcal{P}_{\mathbb{N}_b,c}$ , for any  $\alpha \in (0, \infty)$

$$\begin{aligned} & R(\theta_{N,b,\nu} \mid \rho_{b,\alpha}) - \min_{\theta \in \mathcal{P}_{\mathbb{N}_b,c}} R(\theta \mid \rho_{b,\alpha}) \\ &= R(\theta_{N,b,\nu} \mid \rho_{b,\alpha_b(c)}). \end{aligned} \quad (38)$$

With this step we have succeeded in replacing the relative entropy  $R(\theta_{N,b,\nu} \mid \rho_{b,\alpha})$  with respect to  $\rho_{b,\alpha}$ , which appears in Lemma 6, by the relative entropy  $R(\theta_{N,b,\nu} \mid \rho_{b,\alpha_b(c)})$  with respect to  $\rho_{b,\alpha_b(c)}$ , which appears in Theorem 5. Substituting the last equation into (37) gives

$$\begin{aligned} & \frac{1}{N} \log P_{N,b,m}(\Theta_{N,b} = \theta_{N,b,\nu}) \\ &= -R(\theta_{N,b,\nu} \mid \rho_{b,\alpha_b(c)}) + \varepsilon_N(\nu), \end{aligned} \quad (39)$$

where  $\varepsilon_N(\nu) \rightarrow 0$  uniformly for  $\nu \in A_{N,b,m}$  as  $N \rightarrow \infty$ . This is the conclusion of part (b) of Theorem 5.

We now complete the proof of part (b) of Theorem 5 by proving Lemma 7.

*Proof of Lemma 7.* (a) We write  $A_{N,b,m} \subset \bigcup_{k=1}^m \{\nu \in \mathbb{N}_0^N : \sum_{j \in \mathbb{N}_b} \nu_j = N, |\nu|_+ = k\}$ . By [8, Cor. 2.5] the number of elements in the set indexed by  $k$  equals the binomial coefficient  $C(N-1, k-1)$ . Since by assumption  $m/N \rightarrow 0$  as  $N \rightarrow \infty$ , for all sufficiently large  $N$ , the quantities  $C(N-1, k-1)$  are increasing and are maximal when  $k = m$ . Since  $C(N-1, k-1) \leq C(N, k)$ , it follows that

$$\begin{aligned} \text{card}(A_{N,b,m}) &\leq \sum_{k=1}^m C(N, k) \leq mC(N, m) \\ &= m \frac{N!}{m!(N-m)!}. \end{aligned} \quad (40)$$

An application of the weak form of Stirling's formula yields for all  $m \geq 2$  and all  $N \geq m+2$

$$\begin{aligned} 0 &\leq \frac{1}{N} \log \text{card}(A_{N,b,m}) \\ &\leq \frac{\log m}{N} - \frac{m}{N} \log \frac{m}{N} - \left(1 - \frac{m}{N}\right) \log \left(1 - \frac{m}{N}\right) \\ &\quad + \frac{O(\log N)}{N}. \end{aligned} \quad (41)$$

Since  $m/N \rightarrow 0$  as  $N \rightarrow \infty$ , we conclude that  $0 \leq N^{-1} \log \text{card}(A_{N,b,m}) \rightarrow 0$  as  $N \rightarrow \infty$ . This completes the proof of part (a).

(b) The starting point is (23), which states that  $\Omega_{N,b,m} = \bigcup_{\nu \in A_{N,b,m}} \Delta_{N,b,m;\nu}$ . For distinct  $\nu \in A_{N,b,m}$  the sets  $\Delta_{N,b,m;\nu}$  are disjoint. Hence

$$\begin{aligned} \frac{1}{N} \log \text{card}(\Omega_{N,b,m}) &= \frac{1}{N} \log \sum_{\nu \in A_{N,b,m}} \text{card}(\Delta_{N,b,m;\nu}) \\ &= \frac{1}{N} \log \left( \max_{\nu \in A_{N,b,m}} \text{card}(\Delta_{N,b,m;\nu}) \right) + \delta_N, \end{aligned} \quad (42)$$

where

$$\begin{aligned} 0 &< \delta_N \\ &= \frac{1}{N} \log \left( \sum_{\nu \in A_{N,b,m}} \frac{\text{card}(\Delta_{N,b,m;\nu})}{\max_{\nu \in A_{N,b,m}} \text{card}(\Delta_{N,b,m;\nu})} \right) \\ &\leq \frac{1}{N} \log \text{card}(A_{N,b,m}). \end{aligned} \quad (43)$$

It follows from part (a) that  $\delta_N \rightarrow 0$  as  $N \rightarrow \infty$ .

We continue with the estimation of  $\text{card}(\Omega_{N,b,m})$ . By Lemma 6

$$\begin{aligned} & - \min_{\nu \in A_{N,b,m}} R(\theta_{N,b,\nu} \mid \rho_{b,\alpha}) + f(\alpha, b, c, K) \\ & - \max_{\nu \in A_{N,b,m}} |\zeta_N(\nu)| \\ & \leq \frac{1}{N} \log \left( \max_{\nu \in A_{N,b,m}} \text{card}(\Delta_{N,b,m;\nu}) \right) \\ & \leq - \min_{\nu \in A_{N,b,m}} R(\theta_{N,b,\nu} \mid \rho_{b,\alpha}) + f(\alpha, b, c, K) \\ & \quad + \max_{\nu \in A_{N,b,m}} |\zeta_N(\nu)|. \end{aligned} \quad (44)$$

As proved in Lemma 6,  $\max_{\nu \in A_{N,b,m}} |\zeta_N(\nu)| \rightarrow 0$  as  $N \rightarrow \infty$ . Hence by (42)

$$\begin{aligned} & - \min_{\nu \in A_{N,b,m}} R(\theta_{N,b,\nu} \mid \rho_{b,\alpha}) + f(\alpha, b, c, K) \\ & - \max_{\nu \in A_{N,b,m}} |\zeta_N(\nu)| + \delta_N \leq \frac{1}{N} \log \text{card}(\Omega_{N,b,m}) \\ & \leq - \min_{\nu \in A_{N,b,m}} R(\theta_{N,b,\nu} \mid \rho_{b,\alpha}) + f(\alpha, b, c, K) \\ & \quad + \max_{\nu \in A_{N,b,m}} |\zeta_N(\nu)| + \delta_N. \end{aligned} \quad (45)$$

Under the assumption that  $R(\cdot \mid \rho_{b,\alpha})$  attains its infimum over  $\mathcal{P}_{\mathbb{N}_b,c}$ , we define

$$\eta_N = \frac{1}{N} \log \text{card}(\Omega_{N,b,m}) - f(\alpha, b, c, K) + \min_{\theta \in \mathcal{P}_{\mathbb{N}_b,c}} R(\theta \mid \rho_{b,\alpha}). \quad (46)$$

In the last two paragraphs of this proof, we show that  $\eta_N \rightarrow 0$  as  $N \rightarrow \infty$ . Given this fact, the last equation yields the asymptotic formula (35) in part (b).

We now prove that  $\eta_N \rightarrow 0$  as  $N \rightarrow \infty$ . To do this, we use (45) to write

$$|\eta_N| \leq \left( \min_{\nu \in A_{N,b,m}} R(\theta_{N,b,\nu} \mid \rho_{b,\alpha}) - \min_{\theta \in \mathcal{P}_{\mathbb{N}_b,c}} R(\theta \mid \rho_{b,\alpha}) \right) + \max_{\nu \in A_{N,b,m}} |\zeta_N(\nu)| + \delta_N. \quad (47)$$

Like the second and third terms on the right side, the first term on the right side is nonnegative because  $A_{N,b,m}$  is a subset of  $\mathcal{P}_{\mathbb{N}_b,c}$ . Since  $\max_{\nu \in A_{N,b,m}} |\zeta_N(\nu)| \rightarrow 0$  and  $\delta_N \rightarrow 0$  as  $N \rightarrow \infty$ , it will follow that  $\eta_N \rightarrow 0$  if we can show that  $R(\cdot \mid \rho_{b,\alpha})$  attains its infimum over  $\mathcal{P}_{\mathbb{N}_b,c}$  and that

$$\lim_{N \rightarrow \infty} \min_{\nu \in A_{N,b,m}} R(\theta_{N,b,\nu} \mid \rho_{b,\alpha}) = \min_{\theta \in \mathcal{P}_{\mathbb{N}_b,c}} R(\theta \mid \rho_{b,\alpha}). \quad (48)$$

We now prove (48).  $R(\cdot \mid \rho_{b,\alpha})$  is lower semicontinuous on  $\mathcal{P}_{\mathbb{N}_b}$  [33, Lem. 1.4.3(b)] and thus on  $\mathcal{P}_{\mathbb{N}_b,c}$ . Since  $R(\cdot \mid \rho_{b,\alpha})$  has compact level sets in  $\mathcal{P}_{\mathbb{N}_b,c}$  [Theorem A.1(a)], it attains its infimum over  $\mathcal{P}_{\mathbb{N}_b,c}$  at some measure  $\theta^*$ . We apply Theorem B.1 in [7] to  $\theta = \theta^*$ , obtaining a sequence  $\theta^{(N)}$  with the following properties: (1) for  $N \in \mathbb{N}$ ,  $\theta^{(N)} \in B_{N,b,m}$  has components  $\theta_j^{(N)} = \nu_j^{(N)}/N$  for  $j \in \mathbb{N}_b$ , where  $\nu^{(N)}$  is an appropriate sequence in  $A_{N,b,m}$ ; (2)  $\theta^{(N)} \Rightarrow \theta^*$  as  $N \rightarrow \infty$ ; (3)  $R(\theta^{(N)} \mid \rho_{b,\alpha}) \rightarrow R(\theta^* \mid \rho_{b,\alpha})$  as  $N \rightarrow \infty$ . The limit in (48) follows from the inequalities

$$\min_{\theta \in \mathcal{P}_{\mathbb{N}_b,c}} R(\theta \mid \rho_{b,\alpha}) \leq \min_{\nu \in A_{N,b,m}} R(\theta_{N,b,\nu} \mid \rho_{b,\alpha}) \leq R(\theta^{(N)} \mid \rho_{b,\alpha}) \quad (49)$$

and the limit  $R(\theta^{(N)} \mid \rho_{b,\alpha}) \rightarrow R(\theta^* \mid \rho_{b,\alpha}) = \min_{\theta \in \mathcal{P}_{\mathbb{N}_b,c}} R(\theta \mid \rho_{b,\alpha})$  as  $N \rightarrow \infty$ . This completes the proof of Lemma 7 and thus the proof of the local estimate in part (b) of Theorem 5.  $\square$

In the next section we explain how the local large deviation estimate in part (b) of Theorem 5 yields the LDP in Theorem 1.

#### 4. Proof of Theorem 1 from Part (b) of Theorem 5

In Theorem 1 we state the LDP for the sequence  $\Theta_{N,b}$  of number-density measures. This sequence takes values in

$\mathcal{P}_{\mathbb{N}_b,c}$ , which is the set of probability measures on  $\mathbb{N}$  having mean  $c \in (b, \infty)$ . The purpose of the present section is to explain how the local large deviation estimate in part (b) of Theorem 5 yields the LDP for  $\Theta_{N,b}$ . All details appear in Section 4 of [7]. The basic idea is first to prove the large deviation limit for  $\Theta_{N,b}$  lying in open balls in  $\mathcal{P}_{\mathbb{N}_b,c}$  and in other subsets defined in terms of open balls and then to use this large deviation limit to prove the LDP in Theorem 1.

In Theorem 8 we state the large deviation limit for open balls and other subsets defined in terms of open balls. Two types of open balls are considered. Let  $\theta$  be a measure in  $\mathcal{P}_{\mathbb{N}_b,c}$ , and take  $r > 0$ . Part (a) states the large deviation limit for open balls  $B_\pi(\theta, r) = \{\mu \in \mathcal{P}_{\mathbb{N}_b,c} : \pi(\theta, \mu) < r\}$ , where  $\pi$  denotes the Prohorov metric on  $\mathcal{P}_{\mathbb{N}_b,c}$ . This limit is used to prove the large deviation upper bound for compact subsets of  $\mathcal{P}_{\mathbb{N}_b,c}$  in part (b) of Theorem 1 and the large deviation lower bound for open subsets of  $\mathcal{P}_{\mathbb{N}_b,c}$  in part (d) of Theorem 1. Now let  $\theta$  be a measure in  $\mathcal{P}_{\mathbb{N}_b,[b,c]}$ . Part (b) states the large deviation limit for sets of the form  $\hat{B}_\pi(\theta, r) \cap \mathcal{P}_{\mathbb{N}_b,c}$ , where  $\hat{B}_\pi(\theta, r) = \{\mu \in \mathcal{P}_{\mathbb{N}_b,[b,c]} : \pi(\theta, \mu) < r\}$ . This limit is used to prove the large deviation upper bound for closed subsets in part (c) of Theorem 1. If  $\theta \in \mathcal{P}_{\mathbb{N}_b,c}$ , then  $B_\pi(\theta, r) = \hat{B}_\pi(\theta, r) \cap \mathcal{P}_{\mathbb{N}_b,c}$ , and the conclusions of parts (a) and (b) of the next theorem coincide.

**Theorem 8.** Fix a nonnegative integer  $b$  and a rational number  $c \in (b, \infty)$ . Let  $m$  be the function  $m(N)$  appearing in the definitions of  $\Omega_{N,b,m}$  in (3) and satisfying  $m(N) \rightarrow \infty$  and  $m(N)^2/N \rightarrow 0$  as  $N \rightarrow \infty$ . The following conclusions hold:

- (a) Let  $\theta$  be a measure in  $\mathcal{P}_{\mathbb{N}_b,c}$  and take  $r > 0$ . Then for any open ball  $B_\pi(\theta, r)$  in  $\mathcal{P}_{\mathbb{N}_b,c}$ ,  $R(B_\pi(\theta, r) \mid \rho_{b,\alpha_b(c)})$  is finite, and one has the large deviation limit

$$\lim_{N \rightarrow \infty} \frac{1}{N} \log P_{N,b,m}(\Theta_{N,b} \in B_\pi(\theta, r)) = -R(B_\pi(\theta, r) \mid \rho_{b,\alpha_b(c)}). \quad (50)$$

- (b) Let  $\theta$  be a measure in  $\mathcal{P}_{\mathbb{N}_b,[b,c]}$  and take  $r > 0$ . Then the set  $\hat{B}_\pi(\theta, r) \cap \mathcal{P}_{\mathbb{N}_b,c}$  is nonempty,  $R(\hat{B}_\pi(\theta, r) \cap \mathcal{P}_{\mathbb{N}_b,c} \mid \rho_{b,\alpha_b(c)})$  is finite, and one has the large deviation limit

$$\lim_{N \rightarrow \infty} \frac{1}{N} \log P_{N,b,m}(\Theta_{N,b} \in \hat{B}_\pi(\theta, r) \cap \mathcal{P}_{\mathbb{N}_b,c}) = -R(\hat{B}_\pi(\theta, r) \cap \mathcal{P}_{\mathbb{N}_b,c} \mid \rho_{b,\alpha_b(c)}). \quad (51)$$

We prove Theorem 8 by applying the local large deviation estimate in part (b) of Theorem 5. A key step is to approximate probability measures in  $B_\pi(\theta, \varepsilon)$  and in  $\hat{B}_\pi(\theta, r) \cap \mathcal{P}_{\mathbb{N}_b,c}$  by appropriate sequences of probability measures in the range of  $\Theta_{N,b}$ . This procedure allows one to show in part (a) that the infimum  $R(B_\pi(\theta, \varepsilon) \mid \rho_{b,\alpha_b(c)})$  can be approximated by the infimum of  $R(\theta \mid \rho_{b,\alpha_b(c)})$  over  $\theta$  lying in the intersection of  $B_\pi(\theta, \varepsilon)$  and the range of  $\Theta_{N,b}$ ; a similar statement holds for the infimum in part (b). A set of hypotheses that allow one to carry out this approximation procedure is given

in Theorem 4.2 in [7], a general formulation that yields Theorem 8 as a special case.

Theorem 1 states the LDP for the number-density measures  $\Theta_{N,b}$ . In order to complete the proof of Theorem 1, we must lift the large deviation limits in Theorem 8 to the large deviation upper bound for compact sets and for closed sets and the large deviation lower bound for open sets. The large deviation lower bound for open sets is immediate from the limit in part (a). To prove the large deviation upper bound for compact sets, we cover the compact set by open balls and use the limit in part (a); the large deviation upper bound for closed sets follows by a similar procedure involving part (b). The details of this procedure are carried out as an application of general formulation in Theorem 4.3 in [7].

In the Appendix we prove two properties of the relative entropy and prove the existence of the quantity  $\alpha_b(c)$  appearing in part (a) of Theorem 5.

## Appendix

### Properties of Relative Entropy and Existence of $\alpha_b(c)$

We fix a nonnegative integer  $b$  and a real number  $c \in (b, \infty)$ . Given  $\theta$  a probability measure on  $\mathbb{N}_b = \{n \in \mathbb{Z} : n \geq b\}$ , the mean  $\int_{\mathbb{N}_b} x \theta(dx)$  of  $\theta$  is denoted by  $\langle \theta \rangle$ . In Theorem A.1 we present two properties of the relative entropy  $R(\theta \mid \rho_{b,\alpha})$  and  $R(\theta \mid \rho_{b,\alpha_b(c)})$  for  $\theta$  in each of the following three spaces, which are introduced in Section 2:  $\mathcal{P}_{\mathbb{N}_b}$ , the set of probability measures on  $\mathbb{N}_b$ ;  $\mathcal{P}_{\mathbb{N}_b,c}$ , the set of  $\theta \in \mathcal{P}_{\mathbb{N}_b}$  satisfying  $\langle \theta \rangle = c$ ; and  $\mathcal{P}_{\mathbb{N}_b,[b,c]}$ , the set of  $\theta \in \mathcal{P}_{\mathbb{N}_b}$  satisfying  $\langle \theta \rangle \in [b, c]$ .

We recall that, for  $\alpha \in (0, \infty)$ ,  $\rho_{b,\alpha}$  denotes the Poisson distribution on  $\mathbb{N}_b$  having components  $\rho_{b,\alpha;j} = [Z_b(\alpha)]^{-1} \cdot \alpha^j / j!$  for  $j \in \mathbb{N}_b$ , where  $Z_0(\alpha) = e^\alpha$ , and, for  $b \in \mathbb{N}$ ,  $Z_b(\alpha) = e^\alpha - \sum_{j=0}^{b-1} \alpha^j / j!$ . According to part (a) of Theorem 5 there exists a unique value  $\alpha = \alpha_b(c)$  for which  $\langle \rho_{b,\alpha_b(c)} \rangle = c$ ; thus  $\rho_{b,\alpha_b(c)}$  lies in  $\mathcal{P}_{\mathbb{N}_b,c}$ . In Theorem A.2 we prove the existence of  $\alpha_b(c)$ . In part (a) of the next theorem we show that  $R(\theta \mid \rho_{b,\alpha})$  has compact level sets in  $\mathcal{P}_{\mathbb{N}_b}$ ,  $\mathcal{P}_{\mathbb{N}_b,[b,c]}$ , and  $\mathcal{P}_{\mathbb{N}_b,c}$ . After the statement of Lemma 7 we use part (b) of the next theorem to show that the arbitrary parameter  $\alpha$  in Lemmas 6 and 7 must have the value  $\alpha_b(c)$ .

**Theorem A.1.** Fix a nonnegative integer  $b$  and a real number  $c \in (b, \infty)$ . For any  $\alpha \in (0, \infty)$  the relative entropy  $R(\theta \mid \rho_{b,\alpha}) = \sum_{j \in \mathbb{N}_b} \theta_j \log(\theta_j / \rho_{b,\alpha;j})$  has the following properties:

- (a)  $R(\cdot \mid \rho_{b,\alpha})$  has compact level sets in  $\mathcal{P}_{\mathbb{N}_b}$ ,  $\mathcal{P}_{\mathbb{N}_b,[b,c]}$ , and  $\mathcal{P}_{\mathbb{N}_b,c}$ .
- (b) For any  $\theta \in \mathcal{P}_{\mathbb{N}_b,c}$ ,  $R(\theta \mid \rho_{b,\alpha}) - \min_{\theta \in \mathcal{P}_{\mathbb{N}_b,c}} R(\theta \mid \rho_{b,\alpha}) = R(\theta \mid \rho_{b,\alpha_b(c)})$ .

*Proof.* (a) The fact that  $R(\cdot \mid \rho_{b,\alpha_b(c)})$  has compact level sets in  $\mathcal{P}_{\mathbb{N}_b}$  is proved in part (c) of Lemma 1.4.3 in [33]. Since  $\mathcal{P}_{\mathbb{N}_b,[b,c]}$  is a compact subset of  $\mathcal{P}_{\mathbb{N}_b}$  [Theorem 4(d)],  $R(\cdot \mid \rho_{b,\alpha})$  also has compact level sets in  $\mathcal{P}_{\mathbb{N}_b,[b,c]}$ . Because  $\mathcal{P}_{\mathbb{N}_b,c}$  is not a closed subset of  $\mathcal{P}_{\mathbb{N}_b,[b,c]}$  [Theorem 4(a)], the proof that

$R(\cdot \mid \rho_{b,\alpha})$  has compact level sets in  $\mathcal{P}_{\mathbb{N}_b,c}$  is more subtle. If  $\theta^{(n)}$  is any sequence in  $\mathcal{P}_{\mathbb{N}_b,c}$  satisfying  $R(\theta^{(n)} \mid \rho_{b,\alpha}) \leq M < \infty$ , then since  $\theta^{(n)} \in \mathcal{P}_{\mathbb{N}_b}$  and  $R(\cdot \mid \rho_{b,\alpha})$  has compact level sets in  $\mathcal{P}_{\mathbb{N}_b}$ , there exist  $\theta \in \mathcal{P}_{\mathbb{N}_b}$  and a subsequence  $\theta^{(n')}$  such that  $\theta^{(n')} \Rightarrow \theta$  and  $R(\theta \mid \rho_{b,\alpha}) \leq M$ . To complete the proof that  $R(\cdot \mid \rho_{b,\alpha})$  has compact level sets in  $\mathcal{P}_{\mathbb{N}_b,c}$ , we must show that  $\theta \in \mathcal{P}_{\mathbb{N}_b,c}$ ; that is,  $\langle \theta \rangle = c$ . By Fatou's lemma  $\langle \theta \rangle \leq \liminf_{N \rightarrow \infty} \langle \theta^{(n')} \rangle = c$ . In addition, for any  $w \in (0, \infty)$

$$\begin{aligned} \int_{\mathbb{N}_b} e^{wx} \rho_{b,\alpha}(dx) &= \frac{1}{Z_b(\alpha)} \cdot \sum_{j \in \mathbb{N}_b} e^{wj} \frac{\alpha^j}{j!} \\ &\leq \frac{1}{Z_b(\alpha)} \cdot \exp(\alpha e^w) < \infty. \end{aligned} \quad (\text{A.1})$$

Lemma 5.1 in [37] shows that the sequence  $\theta^{(n')}$  is uniformly integrable, implying that  $c = \lim_{n' \rightarrow \infty} \langle \theta^{(n')} \rangle = \langle \theta \rangle$  [32, Appendix, Prop. 2.3]. This completes the proof that  $R(\cdot \mid \rho_{b,\alpha})$  has compact level sets in  $\mathcal{P}_{\mathbb{N}_b,c}$ . The proof of part (a) is finished.

(b) We define  $g(\alpha, b, c) = \log Z_b(\alpha) - c \log \alpha - (\log Z_b(\alpha_b(c)) - c \log \alpha_b(c))$ . Step 1 is to prove that for any  $\theta \in \mathcal{P}_{\mathbb{N}_b,c}$

$$R(\theta \mid \rho_{b,\alpha}) = R(\theta \mid \rho_{b,\alpha_b(c)}) + g(\alpha, b, c). \quad (\text{A.2})$$

For any  $\theta \in \mathcal{P}_{\mathbb{N}_b,c}$  we have  $\sum_{j \in \mathbb{N}_b} \theta_j = 1$  and  $\sum_{j \in \mathbb{N}_b} j \theta_j = c$ . Hence

$$\begin{aligned} R(\theta \mid \rho_{b,\alpha}) &= \sum_{j \in \mathbb{N}_b} \theta_j \log \left( \frac{\theta_j}{\rho_{b,\alpha;j}} \right) \\ &= \sum_{j \in \mathbb{N}_b} \theta_j \log \left( \frac{\theta_j}{\rho_{b,\alpha_b(c);j}} \right) + \sum_{j \in \mathbb{N}_b} \theta_j \log \left( \frac{\rho_{b,\alpha_b(c);j}}{\rho_{b,\alpha;j}} \right) \\ &= R(\theta \mid \rho_{b,\alpha_b(c)}) \\ &\quad + \sum_{j \in \mathbb{N}_b} \theta_j \log \left( \frac{[\alpha_b(c)]^j}{Z_b(\alpha_b(c)) j!} \cdot \frac{Z_b(\alpha) j!}{\alpha^j} \right) \\ &= R(\theta \mid \rho_{b,\alpha_b(c)}) + \log \left( \frac{Z_b(\alpha)}{Z_b(\alpha_b(c))} \right) \\ &\quad + c \log \left( \frac{\alpha_b(c)}{\alpha} \right). \end{aligned} \quad (\text{A.3})$$

Since the last two lines equal  $R(\theta \mid \rho_{b,\alpha_b(c)}) + g(\alpha, b, c)$ , the proof of (A.2) is complete. Step 2 is to prove that  $R(\theta \mid \rho_{b,\alpha})$  attains its infimum over  $\theta \in \mathcal{P}_{\mathbb{N}_b,c}$  at the measure  $\theta = \rho_{b,\alpha_b(c)}$ , and

$$\min_{\theta \in \mathcal{P}_{\mathbb{N}_b,c}} R(\theta \mid \rho_{b,\alpha}) = R(\rho_{b,\alpha_b(c)} \mid \rho_{b,\alpha}) = g(\alpha, b, c). \quad (\text{A.4})$$

Given these two assertions part (b) of the theorem follows by substituting  $g(\alpha, b, c) = \min_{\theta \in \mathcal{P}_{\mathbb{N}_b,c}} R(\theta \mid \rho_{b,\alpha})$  into (A.2).



We now prove the two assertions in Step 2.  $R(\cdot \mid \rho_{b,\alpha})$  is lower semicontinuous on  $\mathcal{P}_{\mathbb{N}_b}$  [33, Lem. 1.4.3(b)] and thus on  $\mathcal{P}_{\mathbb{N}_b,c}$ . Since  $R(\cdot \mid \rho_{b,\alpha})$  has compact level sets in  $\mathcal{P}_{\mathbb{N}_b,c}$ , it attains its infimum over  $\mathcal{P}_{\mathbb{N}_b,c}$ . The relative entropy  $R(\cdot \mid \rho_{b,\alpha_b(c)})$  attains its minimum value of 0 over  $\mathcal{P}_{\mathbb{N}_b,c}$  at the unique measure  $\rho_{b,\alpha_b(c)}$  [33, Lem. 1.4.1]. Hence (A.2) implies that the minimum value of  $R(\cdot \mid \rho_{b,\alpha})$  over  $\mathcal{P}_{\mathbb{N}_b,c}$  equals

$$\begin{aligned} \min_{\theta \in \mathcal{P}_{\mathbb{N}_b,c}} R(\theta \mid \rho_{b,\alpha}) &= \min_{\theta \in \mathcal{P}_{\mathbb{N}_b,c}} R(\theta \mid \rho_{b,\alpha_b(c)}) + g(\alpha, b, c) \\ &= g(\alpha, b, c) \\ &= R(\rho_{b,\alpha_b(c)} \mid \rho_{b,\alpha_b(c)}) + g(\alpha, b, c) \\ &= R(\rho_{b,\alpha_b(c)} \mid \rho_{b,\alpha}). \end{aligned} \quad (\text{A.5})$$

The last equality follows by applying (A.2) with  $\theta = \rho_{b,\alpha_b(c)}$ . This display shows that  $R(\cdot \mid \rho_{b,\alpha})$  attains its infimum over  $\mathcal{P}_{\mathbb{N}_b,c}$  at  $\rho_{b,\alpha_b(c)}$  and yields (A.4). The proof of part (b) is finished, completing the proof of the theorem.  $\square$

We now prove that there exists a unique value of  $\alpha_b(c)$  for which  $\langle \rho_{b,\alpha_b(c)} \rangle = c$ . The conclusion of the next theorem is part (a) of Theorem C.1 in [7]. In part (b) of that theorem we derive two sets of bounds on  $\alpha_b(c)$  and use these bounds to show that  $\alpha_b(c)$  is asymptotic to  $c$  as  $c \rightarrow \infty$ . In part (d) of Theorem C.1 in [7] we make precise the relationship between  $\rho_{b,\alpha_b(c)}$  and a Poisson random variable having parameter  $\alpha_b(c)$ .

**Theorem A.2.** Fix a nonnegative integer  $b$  and a real number  $c \in (b, \infty)$ . There exists a unique value  $\alpha_b(c) \in (0, \infty)$  such that  $\rho_{b,\alpha_b(c)}$  lies in the set  $\mathcal{P}_{\mathbb{N}_b,c}$  of probability measures on  $\mathbb{N}_b$  having mean  $c$ . If  $b = 0$ , then  $\alpha_0(c) = c$ . If  $b \in \mathbb{N}$ , then  $\alpha_b(c)$  is the unique solution in  $(0, \infty)$  of  $\alpha Z_{b-1}(\alpha)/Z_b(\alpha) = c$ .

According to this theorem, for  $b \in \mathbb{N}$ ,  $\alpha_b(c)$  is the unique solution of  $\alpha Z_{b-1}(\alpha)/Z_b(\alpha) = c$ . The heart of the proof of Theorem A.2, and its most subtle step, is to prove that the function  $\gamma_b(\alpha) = \alpha Z_{b-1}(\alpha)/Z_b(\alpha)$  satisfies  $\gamma'_b(\alpha) > 0$  for  $\alpha \in (0, \infty)$  and thus is monotonically increasing on this interval. This fact is proved in the next lemma.

**Lemma A.3.** Fix a positive integer  $b$  and a real number  $c \in (b, \infty)$ . For  $\alpha \in (0, \infty)$  the function  $\gamma_b(\alpha) = \alpha Z_{b-1}(\alpha)/Z_b(\alpha)$  satisfies  $\gamma'_b(\alpha) > 0$ .

*Proof.* For  $b \in \mathbb{N}$  and for  $\alpha \in (0, \infty)$ , we have  $Z'_b(\alpha) = Z_{b-1}(\alpha)$ . Thus  $\gamma_b(\alpha) = \alpha(\log Z_b(\alpha))'$ . The key to proving that  $\gamma'_b(\alpha) > 0$  is to represent  $\log Z_b(\alpha)$  in terms of the moment generating function of a probability measure. We do this by first expressing  $Z_b(\alpha)$  in terms of the upper incomplete gamma function via the formula  $Z_b(\alpha) = [e^\alpha/(b-1)!] \int_0^\alpha x^{b-1} e^{-x} dx$ . As suggested in [38], we now make the change of variables  $x = \gamma\alpha$ , obtaining the representation

$$\begin{aligned} Z_b(\alpha) &= \frac{e^\alpha}{b!} \alpha^b g_b(\alpha), \\ \text{where } g_b(\alpha) &= \int_{-1}^0 e^{\alpha y} b(-y)^{b-1} dy. \end{aligned} \quad (\text{A.6})$$

The function  $g_b$  is the moment generating function of the probability measure on  $\mathbb{R}$  having the density  $h_b(y) = b(-y)^{b-1}$  on  $[-1, 0]$ . For  $\alpha \in (0, \infty)$  let  $\sigma_{b,\alpha}$  be the probability measure on  $\mathbb{R}$  having the density  $e^{\alpha y} h_b(y)/g_b(\alpha)$  on  $[-1, 0]$ . A straightforward calculation shows that

$$\begin{aligned} (\log g_b)'(\alpha) &= \int_{\mathbb{R}} y \sigma_{b,\alpha}(dy), \\ (\log g_b)''(\alpha) &= \int_{\mathbb{R}} [y - (\log g_b)'(\alpha)]^2 \sigma_{b,\alpha}(dy). \end{aligned} \quad (\text{A.7})$$

It follows that  $(\log g_b)''(\alpha) > 0$  for all  $\alpha \in (0, \infty)$ .

Using (A.6) and the formulas  $Z_{b-1}(\alpha) = \sum_{j=b-1}^\infty \alpha^j/j!$  and  $Z_b(\alpha) = \sum_{j=b}^\infty \alpha^j/j!$ , we calculate

$$\begin{aligned} \gamma'_b(\alpha) &= (\log Z_b(\alpha))' + \alpha (\log Z_b(\alpha))'' \\ &= (\log Z_b(\alpha))' + \alpha \left[ \log \left( \frac{e^\alpha}{b!} \alpha^b g_b(\alpha) \right) \right]'' \\ &= \frac{Z_{b-1}(\alpha)}{Z_b(\alpha)} - \frac{b}{\alpha} + \alpha (\log g_b(\alpha))'' \\ &= \frac{\alpha Z_{b-1}(\alpha) - b Z_b(\alpha)}{\alpha Z_b(\alpha)} + \alpha (\log g_b(\alpha))'' \\ &= \frac{1}{Z_b(\alpha)} \sum_{j=b}^\infty \frac{j-b}{j!} \alpha^{j-1} + \alpha (\log g_b(\alpha))'' > 0. \end{aligned} \quad (\text{A.8})$$

This completes the proof of the lemma.  $\square$

We are now ready to prove Theorem A.2.

*Proof of Theorem A.2.* We first consider  $b = 0$ . In this case  $\rho_{0,\alpha}$  is a standard Poisson distribution on  $\mathbb{N}_0$  having mean  $\alpha$ . It follows that  $\alpha_0(c) = c$  is the unique value for which  $\rho_{0,\alpha_0(c)}$  has mean  $c$  and thus lies in  $\mathcal{P}_{\mathbb{N}_0,c}$ . This completes the proof for  $b = 0$ .

We now consider  $b \in \mathbb{N}$ . In this case  $\rho_{b,\alpha}$  is a probability measure on  $\mathbb{N}_b$  having mean

$$\begin{aligned} \sum_{j \in \mathbb{N}_b} j \rho_{b,\alpha;j} &= \frac{1}{Z_b(\alpha)} \cdot \sum_{j \in \mathbb{N}_b} \frac{\alpha^j}{(j-1)!} \\ &= \frac{1}{Z_b(\alpha)} \cdot \alpha Z_{b-1}(\alpha). \end{aligned} \quad (\text{A.9})$$

Thus  $\rho_{b,\alpha}$  has mean  $c$  if and only if  $\alpha$  satisfies  $\gamma_b(\alpha) = c$ , where  $\gamma_b(\alpha) = \alpha Z_{b-1}(\alpha)/Z_b(\alpha)$ . We prove the theorem by showing that  $\gamma_b(\alpha) = c$  has a unique solution  $\alpha_b(c) \in (0, \infty)$  for all  $b \in \mathbb{N}$  and any  $c > b$ . This assertion is a consequence of the following three steps: (1)  $\lim_{\alpha \rightarrow 0^+} \gamma_b(\alpha) = b$ ; (2)  $\lim_{\alpha \rightarrow \infty} \gamma_b(\alpha) = \infty$ ; (3) for all  $\alpha \in (0, \infty)$ ,  $\gamma'_b(\alpha) > 0$ . Steps 1 and 2 follow immediately from the definition of  $\gamma_b(\alpha)$ , and Step 3 is proved in Lemma A.3.

We have proved the theorem for all  $b \in \mathbb{N}$ . Since we also validated the conclusion of the theorem for  $b = 0$ , the proof for all nonnegative integers  $b$  is done.  $\square$

## Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

## Acknowledgments

The research of Shlomo Ta'asan is supported in part by a grant from the National Science Foundation (NSF-DMS-1216433). Richard S. Ellis thanks Jonathan Machta for sharing his insights into statistical mechanics and Michael Sullivan for his generous help with a number of topological issues arising in this paper. Both authors thank the referee for a careful reading of the paper and for suggesting a number of references.

## References

- [1] R. S. Ellis and S. Ta'asan, "The Boltzmann-Sanov large deviation principle and applications to statistical mechanics," 2014, 48 pages, <http://people.math.umass.edu/~rsellis/pdf-files/boltzmann-sanov-applications.pdf>.
- [2] D. Randall, "Combinatorial Methods for Statistical Physics Models," Special Topics Course, 1999, <http://people.math.gatech.edu/~randall/topics2.html>.
- [3] R. J. Baxter, *Exactly Solved Models in Statistical Mechanics*, Academic Press, New York, NY, USA, 1982.
- [4] B. McCoy and T. T. Wu, *The Two-Dimensional Ising Model*, Harvard University Press, Cambridge, Mass, USA, 1973.
- [5] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications*, Springer, New York, NY, USA, 2nd edition, 1998.
- [6] R. S. Ellis, *Entropy, Large Deviations, and Statistical Mechanics*, Classics of Mathematics, Springer, New York, NY, USA, 1985.
- [7] R. S. Ellis and S. Ta'asan, "Detailed large deviation analysis of a droplet model having a Poisson equilibrium distribution," <http://arxiv.org/abs/1405.5091v4>.
- [8] C. A. Charalambides, *Enumerative Combinatorics*, Chapman & Hall, CRC Press, Boca Raton, Fla, USA, 2002.
- [9] L. Comtet, *Advanced Combinatorics: The Art of Finite and Infinite Expansions*, D. Reidel Publishing Company, Dordrecht, The Netherlands, 1974, Translated by: J. W. Nienhuys.
- [10] E. A. Bender, "Central and local limit theorems applied to asymptotic enumeration," *Journal of Combinatorial Theory, Series A*, vol. 15, no. 1, pp. 91–111, 1973.
- [11] S. Boucheron, F. Gamboa, and C. Léonard, "Bins and balls: large deviations of the empirical occupancy process," *Annals of Applied Probability*, vol. 12, no. 2, pp. 607–636, 2002.
- [12] P. Dupuis, C. Nuzman, and P. Whiting, "Large deviation asymptotics for occupancy problems," *Annals of Probability*, vol. 32, no. 3, pp. 2765–2818, 2004.
- [13] J. Duran, *Sands, Powders, and Grains: An Introduction to the Physics of Granular Materials*, Springer, New York, NY, USA, 2000, Translated by Axel Reisinger.
- [14] R. A. Mugele and H. D. Evans, "Droplet size distribution in sprays," *Industrial & Engineering Chemistry*, vol. 43, no. 6, pp. 1317–1324, 1951.
- [15] R. W. Sellens and T. A. Brzustowski, "A prediction of the drop size distribution in a spray from first principles," *Atomisation and Spray Technology*, vol. 1, pp. 89–102, 1985.
- [16] A. J. C. Ladd and W. G. Hoover, "Energy and entropy of interacting dislocations," *Physical Review B*, vol. 26, no. 10, pp. 5469–5479, 1982.
- [17] A. Datta and S. K. Som, "Effects of spray characteristics on combustion performance of a liquid fuel spray in a gas turbine combustor," *International Journal of Energy Research*, vol. 23, no. 3, pp. 217–228, 1999.
- [18] C. O. M. Miller, "A mathematical model of aerial deposition of pesticides from aircraft," *Environmental Science & Technology*, vol. 14, no. 7, pp. 824–831, 1980.
- [19] M. Alderliesten, "Mean particle diameters. Part I: evaluation of definition systems," *Particle & Particle Systems Characterization*, vol. 7, no. 4, pp. 233–241, 1990.
- [20] M. Alderliesten, "Mean particle diameters. Part II: standardization of nomenclature," *Particle & Particle Systems Characterization*, vol. 8, no. 3, pp. 237–241, 1991.
- [21] X. Li, R. S. Tankin, and M. Renksizbulut, "Calculated characteristics of droplet size and velocity distributions in liquid sprays," *Particle & Particle Systems Characterization*, vol. 7, no. 2, pp. 54–59, 1990.
- [22] L. Ma and R. K. Hanson, "Measurement of aerosol size distribution functions by wavelength-multiplexed laser extinction," *Applied Physics B: Lasers and Optics*, vol. 81, no. 4, pp. 567–576, 2005.
- [23] N. C. Megoulas and M. A. Koupparis, "Twenty years of evaporative light scattering detection," *Critical Reviews in Analytical Chemistry*, vol. 35, no. 4, pp. 301–316, 2005.
- [24] C.-D. Schegk, F. Loeffler, and H. Umhauer, "Drop-size distributions produced by flat-spray nozzles," *Particle & Particle Systems Characterization*, vol. 3, no. 1, pp. 14–19, 1986.
- [25] C. Dumouchel, "A new formulation of the maximum entropy formalism to model liquid spray drop-size distribution," *Particle & Particle Systems Characterization*, vol. 23, no. 6, pp. 468–479, 2007.
- [26] R. S. Ellis, "The theory of large deviations and applications to statistical mechanics," in *Long-Range Interacting Systems: Les Houches 2008 Session XC*, T. Dauxois, S. Ruffo, and L. F. Cugliandolo, Eds., pp. 227–277, Oxford University Press, New York, NY, USA, 2010, <http://people.math.umass.edu/~rsellis/pdf-files/Les-Houches-paper.pdf>.
- [27] R. S. Ellis and K. Wang, "Limit theorems for the empirical vector of the Curie-Weiss-Potts model," *Stochastic Processes and Their Applications*, vol. 35, no. 1, pp. 59–79, 1990.
- [28] M. Costeniuc, R. S. Ellis, and H. Touchette, "Complete analysis of phase transitions and ensemble equivalence for the Curie-Weiss-Potts model," *Journal of Mathematical Physics*, vol. 46, Article ID 063301, 2005.
- [29] R. S. Ellis, P. T. Otto, and H. Touchette, "Analysis of phase transitions in the mean-field Blume-Emery-Griffiths model," *Annals of Applied Probability*, vol. 15, no. 3, pp. 2203–2254, 2005.
- [30] R. S. Ellis, "The theory of large deviations: from Boltzmann's 1877 calculation to equilibrium macrostates in 2D turbulence," *Physica D: Nonlinear Phenomena*, vol. 133, no. 1–4, pp. 106–136, 1999.
- [31] C. Boucher, R. S. Ellis, and B. Turkington, "Derivation of maximum entropy principles in two-dimensional turbulence via large deviations," *Journal of Statistical Physics*, vol. 98, no. 5–6, pp. 1235–1278, 2000.
- [32] S. N. Ethier and T. G. Kurtz, *Markov Processes: Characterization and Convergence*, John Wiley & Sons, New York, NY, USA, 1986.

- [33] P. Dupuis and R. S. Ellis, *A Weak Convergence Approach to the Theory of Large Deviations*, John Wiley & Sons, New York, NY, USA, 1997.
- [34] P. Flajolet and R. Sedgewick, *Analytic Combinatorics*, Cambridge University Press, Cambridge, UK, 2009.
- [35] N. L. Johnson and S. Kotz, *Urn Models and Their Application: An Approach to Modern Discrete Probability Theory*, John Wiley & Sons, New York, NY, USA, 1977.
- [36] V. F. Kolchin, B. A. Sevast'yanov, and V. P. Chistyakov, *Random Allocations*, edited by: A. V. Balakrishnan, V. H. Winston & Sons, Washington, DC, USA, 1978.
- [37] M. D. Donsker and S. R. S. Varadhan, "Asymptotic evaluation of certain Markov process expectations for large time—III," *Communications on Pure and Applied Mathematics*, vol. 29, no. 4, pp. 389–461, 1976.
- [38] E. Neuman, "Inequalities and bounds for the incomplete gamma function," *Results in Mathematics*, vol. 63, no. 3-4, pp. 1209–1214, 2013.

