



# Infinite-Volume Gibbs States of the Generalized Mean-Field Orthoptical Model

Kalle Koskinen<sup>1</sup>

Received: 8 April 2024 / Accepted: 5 August 2024  
© The Author(s) 2024

## Abstract

The generalized mean-field orthoptical model is a mean-field model on a space of continuous spins on  $\mathbb{R}^n$  that are constrained to a scaled  $(n - 1)$ -dimensional  $\ell_1$ -sphere, equivalently a scaled  $(n - 1)$ -dimensional orthoplex, and interact through a general interaction function. The finite-volume Gibbs states of this model correspond to singular probability measures. In this paper, we use probabilistic methods to rigorously classify the infinite-volume Gibbs states of this model, and we show that they are convex combinations of product states. The predominant methods utilize the theory of large deviations, relative entropy, and equivalence of ensembles, and the key technical tools utilize exact integral representations of certain partition functions and locally uniform estimates of expectations of certain local observables.

**Keywords** Infinite-volume Gibbs states · Equivalence of ensembles · Equilibrium statistical mechanics · Mean-field models · Curie–Weiss models

## 1 Introduction

The purpose of this paper is to present rigorous probabilistic methods to compute and classify the large  $n$ -limits of integrals of the form

$$\mu_n^g[f \circ \pi_I] := \frac{1}{Q_n(g)} \int_{\mathbb{R}^n} d\phi e^{ng\left(\frac{1}{n} \sum_{i=1}^n \phi_i\right)} \delta\left(\sum_{i=1}^n |\phi_i| - n\right) (f \circ \pi_I)(\phi), \quad (1.0.1)$$

where  $g : \mathbb{R} \rightarrow \mathbb{R}$  is a “sufficiently regular” function which will be referred to as an interaction function,  $d\phi$  is the Lebesgue measure on  $\mathbb{R}^n$ ,  $\delta(\cdot)$  is formally a delta function,  $f \in C_b(\mathbb{R}^I)$ , where  $C_b(\mathbb{R}^I)$  is the space of continuous bounded functions on a finite index set  $I \subset [n] := \{1, 2, \dots, n\}$ ,  $\pi_I : \mathbb{R}^n \rightarrow \mathbb{R}^I$  is the canonical coordinate projection, and  $Q_n(g)$  is a normalization constant, which will be referred to as the partition function, which make  $\mu_n^g$  into a probability measure. The main result in this paper is given in Theorem 3.3.6, and it constitutes a full characterization of the infinite-volume Gibbs states corresponding

---

Communicated by Aernout van Enter.

✉ Kalle Koskinen  
kalle.koskinen@helsinki.fi

<sup>1</sup> Department of Mathematics and Statistics, University of Helsinki, P.O. Box 68, 00014 Helsinki  
yliopisto, Finland

to the models given by the probability measures in Eq. (1.0.1) given some regularity of the interaction function  $g$ . The main result states the infinite-volume Gibbs states of this model are given by convex combinations of particular exponential product states, and the coefficients of the convex combination are explicitly determined by an associated free energy function. Explicit examples of this type of result are given in the associated examples appearing after Theorems 3.2.4 and 3.2.6.

As a guide for the further reading of our introduction, we present an informal version of this main result.

**Theorem** *For sufficiently regular interaction functions  $g$ , it follows that*

$$\lim_{n \rightarrow \infty} \frac{1}{n} \ln Q_n(g) = \sup_{m \in [-1, 1]} \psi^g(m),$$

where  $\psi^g : [-1, 1] \rightarrow \mathbb{R}$  is given by

$$\psi^g(m) := g(m) + 1 + \ln \left( 1 + \sqrt{1 - m^2} \right),$$

and, if we denote the elements of the finite collection  $M^*$  of global maximizing points of  $\psi^g$  by the collection  $(m^*)$ , then it follows that there exists a collection  $(c_{m^*})$  of positive weights summing to unity such that

$$\lim_{n \rightarrow \infty} \mu_n^g = \sum_{m^*} c_{m^*} \eta^{m^*},$$

where  $\eta^{m^*}$  are probability measures corresponding to factorizable product states on  $\mathbb{R}^{\mathbb{N}}$  with single-site marginal densities given by

$$x \mapsto \frac{e^{-\beta x - \mu|x|}}{q(\beta, \mu)},$$

where  $(\beta, \mu) \in \mathbb{R} \times (0, \infty)$  are coefficients satisfying  $|\beta| < \mu$  and  $q(\beta, \mu)$  are normalization constants making the marginals into probability measures.

The dependence on the collection  $(m^*)$  of global maximizing points of the weights  $c_{m^*}$  and of the coefficients  $(\beta, \mu)$  of each  $\eta^{m^*}$  can be explicitly determined.

The rest of this introduction is dedicated to the explanation and motivation of the various objects appearing in this informal statement of the main result. For the actual statement of the main result, one should note that the notation is changed to be more specific and suggestive. We will refer to the probability measure  $\mu_n^g$  as a finite-volume Gibbs state and the infinite-volume limit, i.e. the large  $n$ -limit, when it exists, will be referred to as the infinite-volume Gibbs state. We refer to Sect. 2 for a complete definition and discussion of the notion of infinite-volume Gibbs state in this context.

At a heuristic level, to make such finite-volume Gibbs states rigorous, we use the fact that the constraint function inside the delta function, when restricted to an orthant of  $\mathbb{R}^n$ , precisely defines a uniform measure over a scaled  $(n - 1)$ -dimensional simplex. This method is used in Sect. 4.1. Since the  $(n - 1)$ -dimensional  $\ell_1$ -sphere corresponds to the  $(n - 1)$ -dimensional orthoplex, we refer to this model as the generalized mean-field orthoplicial model. This naming convention is similar to the convention used for the mean-field spherical model, see [1], but the constraint changes from the  $\ell_2$ -sphere to the orthoplex.

Mean-field models of equilibrium statistical mechanics have been studied extensively as toy models of spins on various types of spaces, and the most famous model belonging to this

class is the Curie–Weiss model, see [2–4]. The classical spin- $\frac{1}{2}$ -Curie–Weiss model is a relatively simple exactly solvable model with interesting statistical mechanical phenomena such as phase transitions, anomalous scaling, infinite-volume Gibbs states, etc. In addition, the model can be generalized in a variety of ways while retaining the essential simplicity of the models. Such generalizations are for instance the modifications of the interaction function, see [5], additions of external random fields, see [6], and modifications to the ambient space, see [2]. In this direction, there are also modifications to the entire measure of the ambient spin space, where one changes the product structure to a constrained singular measure such as the uniform measure on a scaled sphere, see [1]. The model here also falls into this category, since the ambient space of spins is assumed to be constrained to scaled spheres in the  $\ell_1$  norm.

The primary motivation for investigating this particular model is that it is a non-trivial exactly solvable model corresponding in a sense to a canonical ensemble probability measure of a thermodynamic system with two constraints, such that the classical ferromagnetic mean-field models are included via a suitable choice of the interaction function. Similar systems have been investigated in [7–14], and there is a wealth of different methodologies that have been used depending on the different models, but it is typical that these models do not utilize a general interaction function if any interaction functional at all. For this particular model, as opposed to the mean-field spherical model in [1, 11], we must employ different more abstract methods to investigate the infinite-volume states, and it is our belief that these more abstract methods can be of use to study other similar problems. In fact, the general solution of this problem does not follow the more standard approach used in [2] which uses the Hubbard–Stratonovich transform, also referred to as Gaussian linearization. We comment on this point in Sect. 2. When comparing this model to the Berlin–Kac model [14], one should note that label permutation invariant models like the models of this paper, have no underlying geometry of the index set. However, since the limiting states of the model are shown to be convex combinations of product states, unless the convex combination is actually trivial i.e. there is just one product state, the limiting state is in general not factorizable.

To our knowledge, the orthoplicial model and the methods presented to solve the various problems associated with the orthoplicial model are novel in the literature, and known methods, such as in [2], are not necessarily applicable. A particular approach, which is quite natural, is to try to swap the delta function for an appropriately parametrized exponential function, and solve rigorously the problem of interchanging such functions. This approach, however, fails, see Sect. 2. Instead, the general method introduced in this paper, presented heuristically in Sect. 2, is to go from a singly delta constrained probability measure to one which is doubly constrained. This particular doubly constrained probability measures is more tractable, and one can characterize both its limiting probability measure, and the uniform parametric convergence it has in the large  $n$ -limit. The limiting measure of the doubly constrained measure has a product structure, and, subject to further analysis, we find the general result that for a wide variety of suitable interaction functions, the limiting states of the model are convex combinations of product states. The suitable swapping between constrained and non-constrained measures is one aspect of the equivalence of ensembles, see [15]. In fact, the approach we mentioned in the beginning of this paragraph, which is also shown to fail, is related to the fact that if we instead study this model as a thermodynamic system with two conserved quantities, namely the Hamiltonian associated to the interaction function, chosen to be quadratic here, and the particle number function corresponding to the  $\ell_1$  norm, then the calculation given in Eq. (2.0.15) shows that the grand canonical ensemble of this model fails to capture the full set of possible microstate values. We can only attain the trivial vanishing energy density with expectations in the grand canonical ensemble. By the result of this paper,

we know that the corresponding canonical ensemble has limiting states corresponding to non-trivial energy densities, and thus the grand canonical ensemble is insufficient to describe all the infinite-volume Gibbs states of this system. This is one form of non-equivalence of ensembles considered in [15]. This point is also important in the sense that it is one of the practical reasons for studying the canonical versions of these models as opposed to the grand canonical versions, and why it is valuable to develop different techniques to studying such models. In this vein, the Berlin–Kac model [14] was one of the first models of this type, and it too exhibits a “trivial” infinite-volume Gibbs state of the grand canonical version of that corresponding model, since it does not capture the phase transition of the standard Berlin–Kac model.

Let us now remark on the main methods and concepts used in this paper in more detail. The first step involves writing the finite-volume Gibbs state as an integral mixture of probability measures such that the mixture acts on two variables which parametrize a doubly constrained measure which we will call a microcanonical probability measure. This step is carried out in Sects. 3 and 4.1. The general strategy is then to utilize a type of generalized dominated convergence theorem where both the integrating measure and the functions we are integrating are varying, see Lemma 3.1.1.

Using relative entropy methods, we are able to show that the difference in expectations of local observables of microcanonical probability measure and the completely unconstrained probability measures, referred to as the grand-canonical probability measure, depend explicitly on their corresponding statistical mechanical entropies, see Lemma 3.1.2. Using large deviations methods, we are able to prove a broadly applicable theorem which allows one to prove locally uniform convergence of the finite-volume microcanonical entropies using the concavity of the microcanonical entropies along with convergence of the grand-canonical entropy, see Theorem 3.1.3. Note that we are using a non-standard terminology by referring to the normalized logarithm of a partition function as an entropy irrespective of the ensemble the partition function comes from. We should emphasize that Theorem 3.1.3 formalizes, at least at this level of regularity, the notion that one can rigorously deduce many properties of the limiting microcanonical entropy from the grand canonical entropy, which is typically far more tractable mathematically.

For the orthoplacial model, one can easily verify some of the conditions of Theorem 3.1.3 that are required, since the corresponding grand-canonical probability measure is a product state. Of some methodological interest is the fact that we use the notion of Lorentzian polynomials, see [16], to prove that the microcanonical entropies are log-concave functions. The end result is that by combining together, Lemmas 3.1.2, 3.1.5, and 3.1.6, we obtain the locally uniform convergence of the difference of expectations of local observables for the microcanonical and grand-canonical probability measure in the form of Corollary 3.1.7.

As for the mixture probability measure, we begin by once again applying the general theorem Theorem 3.1.3, to deduce the entropy of the corresponding canonical model, see Lemma 3.2.1, which is directly related to our model with a linear interaction function. Using tilting, we are then able to show that the mixture probability measures satisfies a large deviations principle, see Corollary 3.2.2. Using large deviations techniques found in Sect. 4.3, we are able to already classify the limiting states of our model for a variety of relevant non-trivial interaction functions, see Example 3.2.5 for the quadratic mean-field interaction with a non-vanishing magnetic field, see Example 3.2.7, for the quadratic mean-field interaction without an external magnetic field, and see Theorems 3.2.4 and 3.2.6 for the rigorous results concerning these two examples.

In order to fully classify the limiting states of the model for more general interaction functions, we need an additional result concerning the microcanonical partition function which comes

in the form of an exact generating function representation, see Lemma 3.3.1. The generating function that we obtain is a modified Bessel function of the first kind, and we utilize a particular integral representation of it. This allows one to fully characterize the weak convergence of the mixture probability measure by relating it to Laplace-type integrals in three variables for which we can exactly deduce their asymptotics, see Lemma 3.3.4. The exact result employs the notions of type, and maximal type, given in [2], adapted to this particular model. The primary pair of results concerning this final result are Theorems 3.3.5 and 3.3.6, which can be summarized by stating that given sufficient regularity of the interaction function  $g$ , which are intimately related to properties of the limiting entropy, one is able to show that the limiting states are convex combinations of products states.

In the literature, the closest works are [1, 11], in which similar results, with entirely different methods, are produced for the so-called mean-field spherical model. Another similar work which considers a Berlin–Kac-type, see [14], model with a spherical constraint is given in [10]. From the pure mathematical perspective, non-interacting continuous models with multiple constraints have been considered in [7, 8]. These works both consider the particular phenomenon of condensation, and their approach could be described as probabilistic ones. For discrete two-constraint models, and formalism for the equivalence of ensembles for such models, see [12]. In terms of methods, in [17], there is an approach to proving a type of uniform convergence between constrained and non-constrained probability measures by adapting a uniform local central limit theorem. A similar approach based on uniform estimates is used in [18] to prove ensemble equivalence of some observables for microcanonical and canonical ensembles. For a random-field model constrained to the sphere, a similar uniform convergence result between constrained and non-constrained probability measures is obtained in [19]. Finally, we should also remark that this paper does not make use of the method of steepest descent, see [14], nor do we rely on characteristic functions in any particular way to complete any of the proofs. In Sect. 2, we use Gaussian linearization, also called the Hubbard–Stratonovich transform, to form a counter-example, but we will otherwise not use this standard tool.

## 1.1 Reading Guide

This paper is primarily organized so that a majority of the concepts and methods without proofs can be gathered by reading the introduction contained in Sect. 1 and the heuristics contained in Sect. 2. These sections do not contain any proofs, but they do contain some definitions and outline the basic approach to the problems in this paper.

The statements of the results, some important intermediate results, short or simple proofs, and relevant expository computations are done in Sect. 3. The more involved proofs or methods are contained in Sect. 4. Note that Sect. 4 also contains an entire subsection devoted to some results in theory of large deviations, see Sect. 4.3, and the basic concepts and properties of relative entropy are given in Sect. 4.2.

## 2 Heuristics

The functions  $f$  used in Eq. (1.0.1) will be referred to as local functions and their associated finite index sets  $I$  will be referred to as local index sets. Such local functions  $f$  are naturally functions on  $\mathbb{R}^n$  for large enough  $n$  by using the coordinate projection  $\pi_I : \mathbb{R}^n \rightarrow \mathbb{R}^I$ , and representing them as a composition  $f \circ \pi_I$ . If one is able to resolve the large- $n$  limits

of integrals of the form given in Eq. (1.0.1), then one is able to specify, in the limit, the “expectations” of a large class of local observables. In doing so, subject to other regularity conditions on this limiting state, one is able to produce a genuine probability measure on  $\mathbb{R}^{\mathbb{N}}$ . From now on, we will omit the coordinate projection  $\pi_I$ , and simply write the expectation with respect to a local function  $f$  without the composition, unless it becomes pertinent for a specified reason. We will use the following definition of weak convergence and limit points of probability measures.

**Definition 2.0.1** A sequence of probability measures  $\mathcal{G} := \{\mu_n\}_{n \in \mathbb{N}}$ , such that each  $\mu_n$  is a probability measure on  $\mathbb{R}^n$ , is said to converge weakly to a probability measure  $\mu_\infty$  on  $\mathbb{R}^{\mathbb{N}}$  if

$$\lim_{n \rightarrow \infty} \mu_n[f] = \mu_\infty[f]$$

for any  $f \in C_b(\mathbb{R}^I)$ .

The set of limit points  $\mathcal{G}_\infty$  of  $\mathcal{G}$  is given by

$$\mathcal{G}_\infty := \left\{ \mu \in \mathcal{P}(\mathbb{R}^{\mathbb{N}}) : \exists \{n_k\}_{k \in \mathbb{N}}, \lim_{k \rightarrow \infty} \mu_{n_k} = \mu \right\},$$

where the limit is understood in the sense of the weak limit given here.

There are simple extensions, see [19] for an extension by “tensoring on 0” to the remaining  $\mathbb{N} \setminus [n]$  components, that make the probability measure  $\mu_n$  in this definition into probability measures on  $\mathbb{R}^{\mathbb{N}}$ , and using these extensions the definitions above are equivalent to the standard definitions of weak convergence of probability measure on Polish spaces, and the notion of limit points is to be understood as limit points with respect to the Lévy–Prokhorov metric. In notation, we would redefine the measures  $\mu'_n := \mu_n \otimes \delta_0^{\mathbb{N} \setminus \{1, 2, \dots, n\}}$ , where  $\delta_0^{\mathbb{N} \setminus \{1, 2, \dots, n\}}$  is the Dirac measure on the 0 vector of the space  $\mathbb{R}^{\mathbb{N} \setminus \{1, 2, \dots, n\}}$ . It is now clear that if  $n$  is large enough, then for any local observable  $f$ , we have  $\mu_n[f] = \mu'_n[f]$ . We see then that this type of redefinition simply extends the probability measures on  $\mathbb{R}^n$  to  $\mathbb{R}^{\mathbb{N}}$ , but since we are predominantly interested in large  $n$ -limits, we might as well work only on the sequence of probability measures  $\{\mu_n\}_{n \in \mathbb{N}}$  since their values coincide for expectations of fixed local observables for large enough  $n$ . For our purposes, understanding that we are predominantly interested in studying the limit of expectations of local observables is sufficient for the contents of this paper.

Using this notation, we are then interested in studying and classifying the structure and content of the sets  $\mathcal{G}^g$ , corresponding to the sequence of probability measures  $\{\mu_n^g\}_{n \in \mathbb{N}}$  specified in their functional form in Eq. (1.0.1), which will be called the collection of finite-volume Gibbs states, and  $\mathcal{G}_\infty^g$ , which will be called the collection of infinite-volume Gibbs states, and their dependence on the interaction function  $g$ .

The prototypical interaction function  $g$  of this paper is based on the Curie–Weiss Hamiltonian  $H_{CW,n}^J : \mathbb{R}^n \rightarrow \mathbb{R}$  given by

$$H_{CW,n}^J(\phi) := -\frac{J}{2n} \sum_{i,j=1}^n \phi_i \phi_j = n \left( -\frac{J}{2} \left( \frac{1}{n} \sum_{i=1}^n \phi_i \right)^2 \right),$$

where  $J > 0$  is a coupling constant, with the associated interaction function  $g^{\beta,J} : \mathbb{R} \rightarrow \mathbb{R}$  given by

$$g^{\beta,J}(m) := \frac{\beta J}{2} m^2,$$

where  $\beta > 0$ . With this interaction function, the probability measure in Eq. (1.0.1) takes the form

$$\mu_n^{\beta, J}[f] := \frac{1}{Q_n(\beta, J)} \int_{\mathbb{R}^n} d\phi e^{\frac{\beta J}{2n} \sum_{i,j=1}^n \phi_i \phi_j} \delta\left(\sum_{i=1}^n |\phi_i| - n\right) f(\phi),$$

and can be seen to contain two competing weights in the integrand: the interaction function gives larger weight to fields  $\phi$  in which the components are of the same sign and as large as possible, this type of behaviour is why we refer to this interaction as ferromagnetic, while the delta function terms constrains the size aspect of the interaction. It is this competition which produces the non-trivial nature of the limiting state.

From this recipe of going from the Hamiltonian to the interaction function  $g$ , we can produce a number of “generalized” interactions such as  $k$ -body interactions corresponding to interaction functions of polynomial-type

$$g(m) := \sum_{j=1}^k \alpha_j m^{2j},$$

where  $\alpha_j$  are some real constants, even convex smooth interactions intended to model non-polynomial ferromagnetic interaction, and countless others which might be of interest.

The problem described here is well understood for models where the delta function is replaced by a product of density functions, see [2]. Let us now remark on the connection between these types of generalized Curie–Weiss models, and the generalized mean-field orthopolical model. Formally, using delta functions, we have

$$\begin{aligned} & \int_{\mathbb{R}^n} d\phi e^{ng\left(\frac{1}{n} \sum_{i=1}^n \phi_i\right)} \delta\left(\sum_{i=1}^n |\phi_i| - n\right) f(\phi) \tag{2.0.1} \\ &= n \int_{-1}^1 dm e^{ng(m)} \int_{\mathbb{R}^n} d\phi \delta\left(\sum_{i=1}^n \phi_i - mn\right) \delta\left(\sum_{i=1}^n |\phi_i| - n\right) f(\phi) \\ &= n \int_{-1}^1 dm e^{ng(m)} Z_n(mn, n) v_n(m, 1)[f], \end{aligned}$$

where

$$v_n(m, \rho)[f] := \frac{1}{Z_n(mn, \rho n)} \int_{\mathbb{R}^n} d\phi \delta\left(\sum_{i=1}^n \phi_i - mn\right) \delta\left(\sum_{i=1}^n |\phi_i| - \rho n\right) f(\phi), \tag{2.0.2}$$

where  $\rho > 0$ ,  $|m| \leq \rho$ , and  $Z_n(mn, \rho n)$  is a normalization constant which makes  $v_n(m, \rho)$  into a probability measure. The values of  $(m, \rho)$  for which the probability measure  $v_n(m, \rho)$  exists in some formal sense are given by pairs satisfying  $\rho > 0$ , and  $|m| \leq \rho$ . These statements can be heuristically guessed “geometrically” by considering the intersection of hyperplanes with the  $\ell_1$ -spheres. For reasons which will become clear later, we will consider the interior of this set of existence, given and denoted by  $\mathcal{A} := \{(m, \rho) : \rho > 0, |m| < \rho\}$ . Returning to Eq. (1.0.1), we see that

$$\mu_n^g[f] = \frac{n}{Q_n(g)} \int_{-1}^1 dm e^{ng(m)} Z_n(m, 1) v_n(m, 1)[f]. \tag{2.0.3}$$

In this form, the finite-volume Gibbs state is written as an integral mixture of another probability measure.

Although the original problem constrained the integrals to the  $\ell_1$  ball of radius  $n$ , we have suggestively modified the notation so as to include the other possible values of the radius. This suggestive notation is due to the principle or phenomenon of the equivalence of ensembles, see [15]. We will refer to the probability measure  $\nu_n(m, \rho)$  given formally in Eq. (2.0.2) as the microcanonical probability measure. This probability measure is constrained by two functions  $M_n, N_n : \mathbb{R}^I \rightarrow \mathbb{R}$  given by

$$M_n(\phi) := \sum_{i=1}^n \phi_i, \quad N_n(\phi) := \sum_{i=1}^n |\phi_i|. \tag{2.0.4}$$

We will refer to these functions as macrostates and the individual functions will be referred to as the magnetization and particle number respectively. In this paper, we will often refer to either ensembles or probability measures when discussing a particular thermodynamic model. Integrals with delta functions of the macrostates are referred to as constrained, and whenever we replace a delta function by some non-singular “function” of a macrostate, we are moving toward a less constrained state. With this perspective in mind, we will focus on the connection between the microcanonical probability measure and the grand canonical probability measure  $\eta(\beta, \mu)$  on  $\mathbb{R}^{\mathbb{N}}$  given by its action on  $f \in C_b(\mathbb{R}^I)$  given by

$$\eta(\beta, \mu)[f \circ \pi_I] := \frac{1}{q(\beta, \mu)^{|I|}} \int_{\mathbb{R}^I} d\phi e^{-\beta \sum_{i \in I} \phi_i - \mu \sum_{i \in I} |\phi_i|} f(\phi), \tag{2.0.5}$$

where  $\mu > 0, |\beta| < \mu$ , and  $q(\beta, \mu)^{|I|}$  is a normalization constant making the finite marginals into probability measures. One can compute, by direct integration, that

$$q(\beta, \mu) := \frac{1}{\mu + \beta} + \frac{1}{\mu - \beta}. \tag{2.0.6}$$

Note that, strictly speaking, the grand canonical probability measure should refer to the probability measure obtained from  $\eta(\beta, \mu)$  by considering its marginal distribution on the index set  $[n]$ .

The equivalence of ensembles principle states that, subject to some yet to be verified properties of the microcanonical and grand canonical partition functions, there are a number of ways in which these two probability measures are the same. For our purposes, we will utilize ideas stemming from the ensemble equivalence principle corresponding in some sense to thermodynamic, macrostate, and measure level equivalence of these probability measures. For a more complete view on the principle of the equivalence of ensembles, see [15]. To that end, we will need the finite- and infinite-volume specific microcanonical entropies  $s_n, s : \mathcal{A} \rightarrow \mathbb{R}$  given respectively by

$$s_n(m, \rho) := \frac{1}{n} \ln Z_n(mn, \rho n), \quad s(m, \rho) := \lim_{n \rightarrow \infty} s_n(m, \rho). \tag{2.0.7}$$

In addition, for the grand canonical ensemble we will need the finite- and infinite-volume specific entropies  $f_n, f : \mathcal{A} \rightarrow \mathbb{R}$  given respectively by

$$f_n(\beta, \mu) := \frac{1}{n} \ln q(\beta, \mu)^n, \quad f(\beta, \mu) := \lim_{n \rightarrow \infty} f_n(\beta, \mu). \tag{2.0.8}$$

Note the sign conventions used here. We will omit the specific part in their naming, and refer simply to entropies. For this particular model, as for all product state models, we trivially have  $f_n(\beta, \mu) = f(\beta, \mu) = \ln q(\beta, \mu)$ .



Using the entropies, we can rewrite Eq. (2.0.3) as

$$\mu_n^g[f] = \frac{n}{Q_n(g)} \int_{-1}^1 dm e^{n(g(m)+s_n(m,1))} v_n(m, 1)[f]. \tag{2.0.9}$$

The first type of equivalence property that we wish to utilize is the following pair of relations

$$\sup_{(m,\rho) \in \mathcal{A}} \{s(m, \rho) - \beta m - \mu \rho\} = f(\beta, \mu), \quad \inf_{(\beta,\mu) \in \mathcal{A}} \{f(\beta, \mu) + \beta m + \mu \rho\} = s(m, \rho). \tag{2.0.10}$$

This relation is practically equivalent to that of two functions being Legendre conjugates, see [20]. Since we already have a closed form for  $f(\beta, \mu)$ , we may extract the form of  $s(m, \rho)$  if this relation holds.

The second equivalence property is the parameter matching scheme given by

$$\eta(\beta, \mu) \left[ \frac{M_n}{n} \right] = m, \quad \eta(\beta, \mu) \left[ \frac{N_n}{n} \right] = \rho. \tag{2.0.11}$$

If for every pair  $(m, \rho) \in \mathcal{A}$  there exists a corresponding pair  $(\beta, \mu) \in \mathcal{A}$  satisfying the above relations and vice versa, then these corresponding pairs of values are the values for which we would expect the probability measures to be the same. We will use the notations  $m(\beta, \mu)$ ,  $\rho(\beta, \mu)$ ,  $\beta(m, \rho)$ , and  $\mu(m, \rho)$  for this bijection. This bijection is intimately connected to the first equivalence property through the Legendre conjugates.

The final form of equivalence is then the rough statement that in the large  $n$ -limit, we have

$$v_\infty(m, \rho)[f] := \lim_{n \rightarrow \infty} v_n(m, \rho)[f] = \eta(\beta(m, \rho), \mu(m, \rho))[f] \tag{2.0.12}$$

for local functions  $f \in C_b(\mathbb{R}^I)$ .

If we now return to Eq. (2.0.3), the heuristic behaviour of the model in the large  $n$ -limit is roughly speaking that

$$\mu_n^g[f] \approx \left( \int_{-1}^1 dm e^{n(g(m)+s(m,1))} \right)^{-1} \int_{-1}^1 dm e^{n(g(m)+s(m,1))} v_\infty(\beta(m, 1), \mu(m, 1))[f], \tag{2.0.13}$$

and using the Laplace method, see [21], one would expect that

$$\begin{aligned} & \left( \int_{-1}^1 dm e^{n(g(m)+s(m,1))} \right)^{-1} \int_{-1}^1 dm e^{n(g(m)+s(m,1))} v_\infty(m, 1)[f] \\ & \approx \int_{M^*(\psi^g)} \alpha(dm) v_\infty(m, 1)[f], \end{aligned} \tag{2.0.14}$$

where  $\alpha$  is a probability measure on  $[-1, 1]$  and  $M^*(\psi^g) \subset (-1, 1)$  is the set of global maximizing points of the mapping  $[-1, 1] \ni m \mapsto \psi^g(m) := g(m) + s(m, 1)$ . This is to be expected since integrands of the form above have an exponential rate concentration to the global maximum points of the given function.

The connection between this model and the generalized Curie–Weiss model is now evident. The limiting states of both models are given by mixtures of product states. However, for this model, one cannot realize these limiting states without the  $\ell_1$  constraint. To see this, let us consider the following integral

$$W_n(\beta, \mu) := \int_{\mathbb{R}^n} d\phi e^{\frac{\beta J}{2n} \sum_{i,j=1}^n \phi_i \phi_j - \mu \sum_{i=1}^n |\phi_i|},$$

where  $\mu > 0$  and  $\beta \leq 0$ . This would be the less constrained grand canonical partition function for which one would hope that an equivalence principle holds. The partition function here is not finite if  $\beta > 0$ . For the allowed values of  $\beta$ , using the Fourier transform of the Gaussian, we have

$$\begin{aligned} W_n(\beta, \mu) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dz e^{-\frac{1}{2}z^2} \left( \int_{-\infty}^{\infty} d\phi e^{i\sqrt{\frac{(-\beta)J}{n}}z\phi - \mu|\phi|} \right)^n \\ &= \sqrt{\frac{n}{2\pi}} \int_{-\infty}^{\infty} dz e^{-\frac{1}{2}nz^2} \left( \frac{2\mu}{\mu^2 + (-\beta)Jz^2} \right)^n \\ &= (2\mu)^n \sqrt{\frac{n}{2\pi}} \int_{-\infty}^{\infty} dz e^{-n\left(\frac{1}{2}z^2 + \ln(\mu^2 + (-\beta)Jz^2)\right)}. \end{aligned}$$

Since the function  $z \mapsto \frac{1}{2}z^2 + \ln(\mu^2 + (-\beta)Jz^2)$  is trivially minimized when  $z = 0$ , by the Laplace method, it follows that

$$\lim_{n \rightarrow \infty} \frac{1}{n} \ln W_n(\beta, \mu) = \ln(2\mu) - \mu^2. \tag{2.0.15}$$

Now, if we include the mixture measure form of this integral, it follows that

$$\begin{aligned} &\frac{1}{W_n(\beta, \mu)} \int_{\mathbb{R}^n} d\phi e^{\frac{\beta J}{2n} \sum_{i=1}^n \phi_i \phi_j - \mu \sum_{i=1}^n |\phi_i|} f(\phi) \\ &= \left( \int_{-\infty}^{\infty} dz e^{-n\left(\frac{1}{2}z^2 + \ln(\mu^2 + (-\beta)Jz^2)\right)} \right)^{-1} \int_{-\infty}^{\infty} dz e^{-n\left(\frac{1}{2}z^2 + \ln(\mu^2 + (-\beta)Jz^2)\right)} \times \eta(i\sqrt{(-\beta)J}z, \mu)[f], \end{aligned}$$

where  $f \in C_b(\mathbb{R}^J)$  is a local function, from which we have

$$\lim_{n \rightarrow \infty} \frac{1}{W_n(\beta, \mu)} \int_{\mathbb{R}^n} d\phi e^{\frac{\beta J}{2n} \sum_{i=1}^n \phi_i \phi_j - \mu \sum_{i=1}^n |\phi_i|} f(\phi) = \eta(0, \mu)[f].$$

As can be seen, the limiting state is trivial in the sense that it is a pure state, i.e. not a convex combination of any other probability measures, and it does not depend on  $\beta \leq 0$ . It is this property why it is desirable to study the  $\ell_1$  constrained model, since the replacement of the product measure, for this particular model, with a delta function reproduces the non-trivial limiting states.

The heuristic is then that the limiting states of the model are mixtures of product states of the form given in Eq. (2.0.5), where the mixture probability measure is determined by the properties of the interaction function  $g$ . This is precisely what we will prove rigorously.

Before presenting the main results and proofs, let us remark on the what exactly is not rigorous, incorrect, or too formal in the above exposition. The delta functions appearing in Eqs. (1.0.1) and (2.0.2) are completely formal objects, and we will rigorously define the micro-canonical probability measure on which we can actually perform non-formal computations. In particular, the formal calculation presented in Eq. (2.0.1) is strictly speaking incorrect. For this particular model, it is important to take into consideration the ‘‘boundary values’’ of the set  $\mathcal{A}$ . That is to say, the admissible pairs which satisfy  $\rho > 0$  and  $|m| = \rho$  produce partition functions which can not be neglected if one wants to verify the formal calculation in Eq. (2.0.1). In addition, the form of equivalence of ensembles we have specified here are vague and unverified. We will verify these forms of equivalence explicitly, and they will be presented as lemmas.

### 3 Main Results

In this section, we present the main results, short or simple proofs, and expository computations concerning the main results.

#### 3.1 Locally Uniform Convergence of Observables and Entropy of the Microcanonical Ensemble

We begin by rigorously defining the microcanonical probability measure  $\nu_n(m, \rho)$  from Eq. (2.0.2) for  $(m, \rho) \in \mathcal{A}$ , and the so-called “boundary values” corresponding to  $\rho > 0$  and  $|m| = \rho$ . This is done by identifying the microcanonical probability measure as a convex combination of products of uniform measures on simplexes. The uniform measures on simplexes are rigorously definable via the so-called flag coordinates, and these uniform measures are computationally tractable. Due to the large number of properties that need to be shown for the microcanonical probability measures, we dedicate an entire section, see Sect. 4.1, to the rigorous definition and methods of use of this particular probability measure. The key definitions are for the microcanonical probability measures  $\nu_n(m, \rho)$  and the microcanonical partition functions  $Z_n(M, N)$ , now defined in Definition 4.1.3.

In this work, we will often refer to Polish spaces and probability measures on them. Whenever we do so without an explicit reference to a  $\sigma$ -algebra, we implicitly mean with respect to the Borel  $\sigma$ -algebra associated with the topology of the Polish space. The basic principle by which we will identify the infinite-volume Gibbs states is presented in the following lemma.

**Lemma 3.1.1** *Let  $X$  be a Polish space. If  $\{\mu_n\}_{n \in \mathbb{N}}$  is a sequence of probability measures on  $X$  converging weakly to a probability measure  $\mu$  on  $X$ ,  $K \subset X$  is a compact continuity set of  $\mu$  such that  $\text{supp}(\mu) \subset K$ , and  $\{f_n\}_{n \in \mathbb{N}}$  is a sequence of uniformly bounded functions on  $X$  converging uniformly on  $K$  to a function  $f$ , then it follows that*

$$\lim_{n \rightarrow \infty} \int_X \mu_n(dx) f_n(x) = \int_K \mu(dx) f(x).$$

The proof, see Sect. 4.2, is an application of conditioning to  $K$  and applying various weak convergence properties.

With reference to Eqs. (1.0.1) and (2.0.3), using the definition and methods of Sect. 4.1, we can write the finite-volume Gibbs states in the following form

$$\mu_n^g = \int_{\mathbb{R}} \kappa_n^g(dm) \nu_n(m, 1),$$

where  $\kappa_n^g$  are probability measures on  $\mathbb{R}$  supported by  $[-1, 1]$  with actions on  $f \in C_b(\mathbb{R})$  given by

$$\begin{aligned} &\kappa_n^g[f] \\ &:= \frac{1}{Q_n(g)} \left( \int_{-1}^1 dm ne^{ng(m)} Z_n(mn, n) f(m) + ne^{ng(1)} Z_n(n, n) f(1) + ne^{ng(-1)} Z_n(-n, n) f(-1) \right), \end{aligned}$$

where the partition function then takes the following form

$$Q_n(g) := \int_{-1}^1 dm ne^{ng(m)} Z_n(mn, n) + ne^{ng(1)} Z_n(n, n) + ne^{ng(-1)} Z_n(-n, n). \quad (3.1.1)$$

In light of Lemma 3.1.1, we have two goals. The first goal is to show that the collection of mixture probability measures  $\{\kappa_n^g\}_{n \in \mathbb{N}}$  converges weakly to some limiting probability measure, and that there exists a compact continuity set of this limiting probability measures which contains the support of the limiting probability measure. The second goal is to show that for a fixed  $f \in C_b(\mathbb{R}^I)$  the collection of functions  $\{v_n(m, 1)[f]\}_{n \in \mathbb{N}}$  understood as a collection of functions on the variable  $m \in [-1, 1]$  is uniformly bounded, which is immediate by the boundedness of  $f$ , and uniformly convergent on the required compact continuity set. In the heuristic sketch in the introduction, we did not pay any particular attention to the modes of convergence of the limiting objects. For this particular model, we are able to obtain locally uniform convergence by relating the rate of convergence of local functions to the rate and mode of convergence of the finite-volume entropies. This connection is described in the following fundamental inequality.

**Lemma 3.1.2** *For any finite index set  $I \subset [n]$  and any pairs of values  $(m, \rho) \in \mathcal{A}$  and  $(\beta, \mu) \in \mathcal{A}$ , we have*

$$\begin{aligned} & \sup_{f \in C_b(\mathbb{R}^I), \|f\|_\infty \leq 1} |v_n(m, \rho)[f] - \eta(\beta, \mu)[f]| \\ & \leq \sqrt{\frac{|I|(n-2)}{2(n-2-|I|)} \left( \beta m + \mu \rho + f(\beta, \mu) - \frac{n}{n-2} s_n(m, \rho) \right)}. \end{aligned}$$

The proof of this result is an application of Pinsker’s inequality for relative entropy, followed by the subadditivity property of relative entropy coupled with the permutation invariance of the microcanonical probability measure. For this model, we can exactly compute the relative entropy of the  $(n - 2)$ :th marginal of the microcanonical probability measure from which we obtain the entropy terms in the above inequality. For the full proof, see Sect. 4.2.

If we were only interested in showing that the microcanonical probability measure converges to the grand canonical probability measure, it can be accomplished by studying the pointwise convergence of the entropies. However, since we want to prove locally uniform convergence, we need some additional regularity. The additional regularity that we will prove is that the sequence of finite-volume microcanonical entropies are pointwise uniformly bounded, and that the microcanonical partition functions are log-concave functions on  $\mathcal{A}$ . By a classical result in convex analysis, see [20, Section 10], once the pointwise limit of the finite-volume microcanonical entropies is deduced, the convergence is immediately elevated to locally uniform convergence.

In some models, the grand canonical entropy is more computationally tractable than the microcanonical entropy. This is the case here as well and we will prove a general result which utilizes the aforementioned regularity properties of the microcanonical partition functions coupled with some additional regularity properties of the grand canonical entropy to prove a result, which might also be of general interest in other models.

**Theorem 3.1.3** *Let  $\{Z_n\}_{n \in \mathbb{N}}$  be a sequence of log-concave functions  $Z_n : n\mathcal{C} \rightarrow (0, \infty)$ , where  $\mathcal{C} \subset \mathbb{R}^m$  is a non-empty open convex set and  $n\mathcal{C} := \{nc : c \in \mathcal{C}\}$ , such that*

$$\sup_{n \in \mathbb{N}} \left| \frac{1}{n} \ln Z_n(nx) \right| < \infty$$

*for any  $x \in \mathcal{C}$ , and there exists a non-empty open convex set  $\mathcal{C}' \subset \mathbb{R}^m$  such that*

$$\int_{n\mathcal{C}} dX e^{-\langle t, X \rangle} Z_n(X) < \infty$$

for all  $t \in \mathcal{C}'$  and all  $n \in \mathbb{N}$ , where  $\langle \cdot, \cdot \rangle$  is the Euclidean inner product.

If the function  $f : \mathbb{R}^m \rightarrow \mathbb{R} \cup \{\pm\infty\}$  given by the mapping

$$f(t) := \lim_{n \rightarrow \infty} \frac{1}{n} \ln \int_{n\mathcal{C}} dX e^{-\langle t, X \rangle} Z_n(X)$$

exists and is a proper convex lower semi-continuous function of Legendre type which satisfies  $\nabla[-f]\mathcal{C}' = \mathcal{C}$  then it follows that

$$\lim_{n \rightarrow \infty} \sup_{x \in K} \left| \frac{1}{n} \ln Z_n(nx) - \inf_{t \in \mathbb{R}^m} \{ \langle t, x \rangle + f(t) \} \right| = 0,$$

for any compact set  $K \subset \mathcal{C}$ .

The proof of this result, see Sect. 4.3, requires definitions and notions from large deviations theory. We have dedicated an entire section, see Sect. 4.3, to the relevant definitions, and results which can be deduced after establishing a large deviations principle. The proof itself uses a relative compactness argument concerning locally uniformly convergent subsequences, and a characterization of the limits of said subsequences using a large deviations principle. To apply this method to this model, we proceed by providing the sufficient regularity of the finite-volume microcanonical entropies.

**Lemma 3.1.4** *The collection of finite-volume microcanonical entropies  $\{s_n\}_{n \in \mathbb{N}}$  is pointwise uniformly bounded and concave on  $\mathcal{A}$ .*

The proof, see Sect. 4.4, of log-concavity proceeds by identifying the microcanonical partition functions  $Z_n$  as a composition of a bivariate Lorentzian polynomial of degree  $n - 2$  and a linear map. To prove the uniform pointwise boundedness, we use the positivity of the relative entropy between the  $(n - 2)$ :th marginal of the microcanonical probability measure and the grand-canonical probability measure.

In light of Theorem 3.1.3, it remains to consider the mapping  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  given by

$$f(\beta, \mu) := \lim_{n \rightarrow \infty} \frac{1}{n} \ln \int_{\mathcal{A}} dM dN e^{-\beta M - \mu N} Z_n(M, N),$$

where, in accordance with Eq. (4.1.3), we have

$$Z_n(M, N) = \frac{1}{2} \sum_{k=1}^{n-1} \binom{n}{k} \frac{\left(\frac{N+M}{2}\right)^{k-1}}{(k-1)!} \frac{\left(\frac{N-M}{2}\right)^{n-k-1}}{(n-k-1)!}$$

for  $(M, N) \in \mathcal{A}$ .

It is immediate that if  $(\beta, \mu) \notin \mathcal{A}$ , then  $f(\beta, \mu) = \infty$ . As for  $(\beta, \mu) \in \mathcal{A}$ , we can directly compute that

$$\begin{aligned} \int_{\mathcal{A}} dM dN e^{-\beta M - \mu N} Z_n(M, N) &= \int_0^\infty dX \int_0^\infty dY e^{-(\mu+\beta)X - (\mu-\beta)Y} \\ &\quad \sum_{k=1}^{n-1} \binom{n}{k} \frac{X^{k-1}}{(k-1)!} \frac{Y^{n-k-1}}{(n-k-1)!} \\ &= \sum_{k=1}^{n-1} \binom{n}{k} \left(\frac{1}{\mu+\beta}\right)^k \left(\frac{1}{\mu-\beta}\right)^{n-k} \\ &= \left(\frac{1}{\mu+\beta} + \frac{1}{\mu-\beta}\right)^n - \left(\frac{1}{\mu+\beta}\right)^n - \left(\frac{1}{\mu-\beta}\right)^n. \end{aligned}$$

Computing the limit, it follows that

$$f(\beta, \mu) = \ln \left( \frac{1}{\mu + \beta} + \frac{1}{\mu - \beta} \right) = \ln q(\beta, \mu).$$

In summary, we have

$$f(\beta, \mu) = \begin{cases} \ln \left( \frac{1}{\mu + \beta} + \frac{1}{\mu - \beta} \right), & (\beta, \mu) \in \mathcal{A} \\ \infty, & (\beta, \mu) \notin \mathcal{A} \end{cases}.$$

We have included this calculation here to emphasize the fact that this calculation is relatively straightforward.

We present the relevant regularity conditions of the map  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  in the following result.

**Lemma 3.1.5** *The mapping  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  is a proper convex lower semi-continuous function of Legendre type.*

*In addition, it follows that  $(-\nabla[f])\mathcal{A} = \mathcal{A}$ , and*

$$\begin{aligned} \inf_{(\beta, \mu) \in \mathbb{R}^2} \{ \beta m + \mu \rho + f(\beta, \mu) \} &= \beta(m, \rho)m + \mu(m, \rho)\rho + f(\beta(m, \rho), \mu(m, \rho)) \\ &= 1 + \ln \left( \left( \sqrt{\frac{\rho + m}{2}} + \sqrt{\frac{\rho - m}{2}} \right)^2 \right), \end{aligned}$$

where  $(\beta, \mu) := (-\nabla[f])^{-1} : \mathcal{A} \rightarrow \mathcal{A}$  is given by

$$\beta(m, \rho) := -\frac{\rho}{m} \frac{1}{\sqrt{\rho^2 - m^2}} + \frac{1}{m}, \quad \mu(m, \rho) := \frac{1}{\sqrt{\rho^2 - m^2}}.$$

For the proof, which is completely computational, see Sect. 4.4.

Combining together the regularity of the finite-volume entropies from Lemma 3.1.4, and the computations and verifications concerning the function  $f$  given in Lemma 3.1.5, we have the following result.

**Lemma 3.1.6** *It follows that*

$$\lim_{n \rightarrow \infty} \sup_{(m, \rho) \in K \subset \mathcal{A}} |s_n(m, \rho) - s(m, \rho)| = 0,$$

for any compact set  $K \subset \mathcal{A}$ , where

$$s(m, \rho) := 1 + \ln \left( \left( \sqrt{\frac{\rho + m}{2}} + \sqrt{\frac{\rho - m}{2}} \right)^2 \right),$$

for any  $(m, \rho) \in \mathcal{A}$ .

Combining together Lemmas 3.1.2, 3.1.5, and 3.1.6, we have the following result concerning the mode of convergence of local observables of the microcanonical probability measures.

**Corollary 3.1.7** *For any finite index set  $I \subset [n]$ , it follows that*

$$\lim_{n \rightarrow \infty} \sup_{(m, \rho) \in K \subset \mathcal{A}} \sup_{f \in C_b(\mathbb{R}^I), \|f\|_\infty \leq 1} |v_n(m, \rho)[f] - \eta(\beta(m, \rho), \mu(m, \rho))[f]| = 0.$$

Having established the compact-open convergence of the microcanonical probability measures, we move on to the weak convergence of the mixture probability measures.

### 3.2 Limiting Entropy and Convergence of Mixture Probability Measures

By the heuristics given, it is evident that the mixture probability measures  $\{\kappa_n(g)\}_{n \in \mathbb{N}}$  should converge, at an exponential rate, to the global maximizing points of some tilting function. This idea can be realized by proving that the mixture probability measures satisfy a large deviations principle. Since the full models have a general interaction function  $g$ , we will first prove a large deviations principle for linear  $g$ , and then use tilting to obtain the full large deviations principle. The following result considers the large deviations principle for a linear  $g$ .

**Lemma 3.2.1** *Let  $\beta \in \mathbb{R}$ ,  $g^\beta(m) := -\beta m$ ,  $Q_n(\beta) := Q_n(g^\beta)$ , and  $\kappa_n^\beta := \kappa_n^{g^\beta}$ . Then, it follows that*

$$\lim_{n \rightarrow \infty} \frac{1}{n} \ln Q_n(\beta) = \sup_{m \in [-1, 1]} \{s(m, 1) - \beta m\}.$$

Moreover,  $\{\kappa_n^\beta\}_{n=1}^\infty$  satisfies a large deviations principle with rate function  $I^\beta : \mathbb{R} \rightarrow [0, \infty]$  given by

$$[-1, 1] \ni m \mapsto I^\beta(m) := \sup_{m \in [-1, 1]} \{s(m, 1) - \beta m\} - (s(m, 1) - \beta m),$$

and  $I^\beta(m) = \infty$  for  $m \notin [-1, 1]$ .

The proof, see Sect. 4.4, follows the same strategy as for the microcanonical entropy. Here, the log-concavity is proved by an application of the Prekopa–Leindler theorem, and pointwise uniform boundedness is a direct calculation.

Since the previous result yields a large deviations principle for the mixture probability measures  $\{\kappa_n^0\}_{n \in \mathbb{N}}$ , corresponding to the choice of  $g$  being identically 0, as direct corollary of tilting, see [22], we have the following large deviations principle for the full mixture probability measures.

**Corollary 3.2.2** *For any  $g \in C_b([-1, 1])$ , it follows that*

$$\lim_{n \rightarrow \infty} \frac{1}{n} \ln Q_n(g) = \sup_{m \in [-1, 1]} \{g(m) + s(m, 1)\}.$$

Moreover,  $\{\kappa_n^g\}_{n=1}^\infty$  satisfies a large deviations principle with rate function  $I^g : \mathbb{R} \rightarrow [0, \infty]$  given by

$$[-1, 1] \ni m \mapsto I^g(m) := \sup_{m \in [-1, 1]} \{g(m) + s(m, 1)\} - (g(m) + s(m, 1)),$$

and  $I^g(m) = \infty$  for  $m \notin [-1, 1]$ .

Whenever a sequence of probability measures satisfies a large deviations principle with some rate function, it is accompanied by a measure concentration result to the kernel of the rate function, see Sect. 4.3. In this vein, consider the function  $\psi^g : [-1, 1] \rightarrow \mathbb{R}$  given by

$$\psi^g(m) := g(m) + s(m, 1).$$

It is clear that if  $I^g(m^*) = 0$ , then the point  $m^*$  corresponds to a global maximum point of  $\psi^g$  by definition, and vice versa. Denote the set of global maximizing points of  $\psi^g$  by  $M^*(\psi^g)$ , and, by the previous observation, we have  $(I^g)^{-1}\{0\} = M^*(\psi^g)$ .

We may now begin the classification of the infinite-volume Gibbs states. As a first partial result, by combining together Lemma 3.1.1, Corollary 3.1.7, and Theorem 4.3.8, we have the following result.

**Lemma 3.2.3** *Let  $g \in C_b([-1, 1])$ , and suppose that  $M^*(\psi^g) \subset (-1, 1)$ . Then, it follows that*

$$\mathcal{G}_\infty^g \subset \left\{ \int_{-1}^1 \kappa(dm) \eta(\beta(m, 1), \mu(\beta, 1)) : \kappa \in \mathcal{M}_1([-1, 1]), \text{supp}(\kappa) \subset M^*(\psi^g) \right\}.$$

The proof, see Sect. 4.4, is a direct combination of the given results.

As a corollary, if we can deduce that there is exactly one global maximizing point of  $\psi^g$  contained in the interval  $(-1, 1)$ , then there is a unique infinite-volume Gibbs state. This follows since the Dirac measure on a single point is the only probability measure supported on a single point.

**Theorem 3.2.4** *Let  $g \in C^1([-1, 1])$ , and suppose that  $\psi^g$  has a unique global maximizing point  $m^* \in (-1, 1)$ .*

*Then, it follows that*

$$\lim_{n \rightarrow \infty} \mu_n^g = \eta(\beta(m^*, 1), \mu(m^*, 1)).$$

There are two prototypical functions  $g$  that fall into this category. One we have already seen which is  $g^\beta(m) = -\beta m$  for  $\beta \in \mathbb{R}$ . Since  $s$  is strictly concave it is easy to check that there is a unique global maximizing point of  $\psi(g^\beta)$ . The other example is related to the Curie–Weiss Hamiltonian with an external field.

**Example 3.2.5** Consider  $g(m) := \frac{\beta J}{2} m^2 + \beta h m$ , where  $J > 0$ ,  $\beta > 0$ , and  $h \neq 0$ . Let us first remark that  $\psi^g$  must attain its maximum on  $[-1, 1]$ . Suppose first that  $h > 0$ . For any point  $m^* < 0$  of  $\psi^g$ , this point cannot be a global maximum point since  $\psi^g(-m^*) > \psi^g(m^*)$ . It follows that if there exists a global maximizing point, then it must be of the same sign as  $h$ . Let us continue now with the case where  $h > 0$ , and note that the other case is analogous. By direct computation, we have

$$\partial[\psi^g](m) = 0 \iff \beta J m + \beta h - \frac{m}{\sqrt{1-m^2}} \frac{1}{1 + \sqrt{1-m^2}} = 0.$$

One can further compute that

$$\partial^3[\psi^g](m) = \frac{2m^5 + 4m^3 - 9m(\sqrt{1-m^2} + 1)}{(\sqrt{1-m^2} + 1)^3 (1-m^2)^{\frac{5}{2}}},$$

and

$$\frac{2m^4 + 4m^2}{9(\sqrt{1-m^2} + 1)} < \frac{2+4}{9} = \frac{6}{9} < 1 \implies 2m^5 + 4m^3 - 9m(\sqrt{1-m^2} + 1) < 0$$

for  $0 < m \leq 1$ . It follows that  $\partial[\psi^g](m) < 0$  on  $(0, 1]$  and  $\partial[\psi^g]$  is thus strictly concave. In addition, we have  $\partial[\psi^g](0) = \beta h > 0$ , and  $\lim_{m \rightarrow 1^-} \partial[\psi^g](m) = -\infty$ . Using these properties, it follows that there must exist a unique point  $m^* \in (0, 1)$  such that  $\partial[\psi^g](m^*) = 0$ . In addition, by strict concavity of  $\partial[\psi^g]$ , it follows that  $\psi^g$  is monotonically increasing on  $(0, m^*)$  and monotonically decreasing on  $(m^*, 1)$  which implies that this  $m^*$  is the unique global maximum point and it is contained on  $(0, 1)$ . A similar argument shows that if  $h < 0$ , then there is a unique global maximum point contained in  $(-1, 0)$ .



For the second type of interaction, we consider even functions  $g \in C_b([-1, 1])$  such that  $g$  has precisely two global maximizing points  $m^+ \in (0, 1)$ , and  $m^- = -m^+ \in (-1, 0)$ . For such even functions, by spin-flip symmetry, or by changing variables  $m \mapsto -m$ , it follows that

$$\frac{\kappa_n^g(B(m^+, \delta))}{\kappa_n^g(B(m^+, \delta)) + \kappa_n^g(B(m^-, \delta))} = \frac{1}{2} = \frac{\kappa_n^g(B(m^-, \delta))}{\kappa_n^g(B(m^+, \delta)) + \kappa_n^g(B(m^-, \delta))},$$

for small enough  $\delta > 0$ . In particular, by Corollary 4.3.10, it follows that

$$\lim_{n \rightarrow \infty} \kappa_n^g = \frac{1}{2} \delta_{m^+} + \frac{1}{2} \delta_{m^-}$$

weakly. By combining this simple result with Lemma 3.1.1, and Corollary 3.1.7, we have the following result.

**Theorem 3.2.6** *Let  $g \in C_b([-1, 1])$  be an even function such that  $M^*(\psi^g) = \{m^+, m^-\}$ , where  $m^+ > 0$ , and  $m^- = -m^+$ .*

*Then, it follows that*

$$\lim_{n \rightarrow \infty} \mu_n^g = \frac{1}{2} \eta(\beta(m^+, 1), \mu(m^+, 1)) + \frac{1}{2} \eta(\beta(m^-, 1), \mu(m^-, 1)).$$

The prototypical example here is the Curie–Weiss Hamiltonian without an external field.

**Example 3.2.7** Consider  $g(m) := \frac{\beta J}{2} m^2$  where  $\beta > 0$ , and  $J > 0$ . We have

$$\begin{aligned} \partial[\psi^g](m) &= m \left( \beta J - \frac{1}{\sqrt{1-m^2}} \frac{1}{1+\sqrt{1-m^2}} \right), \\ \partial^2[\psi^g](m) &= \beta J - \frac{1}{\sqrt{1-m^2}} \frac{1}{1+\sqrt{1-m^2}} - \frac{m^2(2\sqrt{1-m^2}+1)}{(\sqrt{1-m^2}+1)^2(1-m^2)^{\frac{3}{2}}}. \end{aligned}$$

From the form of the first derivative, we see that  $\psi^g$  cannot obtain a maximum at either end of the interval  $[-1, 1]$  and must thus be attained at a critical point in the open interval  $(-1, 1)$ . There are now two options for the critical point, the first is that  $m = 0$ , from which we have

$$\partial[\psi^g](0) = 0, \quad \partial^2[\psi^g](0) = \beta J - \frac{1}{2}.$$

Due to the sign of the second derivative, this fails to be even a local maximum when  $\beta J > \frac{1}{2}$ , and whatever other critical point must be the global maximizing point if we are in this parameter range. The other case is that

$$\beta J = \frac{1}{\sqrt{1-m^{\pm 2}}} \frac{1}{1+\sqrt{1-m^{\pm 2}}} \iff \sqrt{1-m^{\pm 2}} = \frac{1}{2} \left( \sqrt{\frac{4}{\beta J} + 1} - 1 \right),$$

when  $\beta J > \frac{1}{2}$ . For other values of  $\beta J$ , there is no solution to this equation and we must conclude that the other critical point corresponds to the global maximizing point. We can conclude that when  $\beta J \in \mathbb{R}$ , then  $m^* = 0$  is always a critical point, but it cannot be even a local maximizing point when  $\beta J > \frac{1}{2}$ , hence in this regime we must conclude that the pair of solutions  $m^\pm$  given above are the only viable critical points, but since they are the only critical points, and the function must attain its maximum at a critical point, we may conclude that  $m^\pm$  also correspond to global maximizing points of the function. When  $\beta J < \frac{1}{2}$ , the  $m^* = 0$  critical point is the only critical point, and we can again conclude that this must then

be the global maximizing point. If  $\beta J = \frac{1}{2}$ , we can check that both  $m^\pm = 0$ , and thus we again have a single critical point which must be a global maximizing point.

The interactions described here are ones which can be dealt with without any further study of the structure of the function  $\psi^g$ . When there are no symmetries or unique global maximum points, one has to resort to other methods to resolve the limits. We will now present such methods for dealing with sufficiently smooth interaction functions that have multiple global maximizing points.

### 3.3 Exact Integral Representations of the Weights and Full Classification of the Infinite-Volume Gibbs States

We will need a preliminary result concerning the microcanonical partition function in order to have better control of the mixture probability measures. We have the following generating function based representation of the microcanonical partition function.

**Lemma 3.3.1** *Let  $(m, \rho) \in \mathcal{A}$ . Then, it follows that*

$$Z_n(mn, \rho n) = \frac{2^{2n-1} n^{n-2} n!}{(2n)! \sqrt{\rho^2 - m^2} n^2} \binom{2n}{2} \frac{e^{-(n-1)}}{\pi^2} \int_0^\pi d\theta_1 \int_0^\pi d\theta_2 \cos \theta_1 \cos \theta_2 e^{(n-1)s(m, \rho, \theta_1, \theta_2)},$$

where  $s : \mathcal{A} \times [0, 2\pi) \times [0, 2\pi)$  is given by

$$s(m, \rho, \theta_1, \theta_2) := 1 + \ln \left( \left( \sqrt{\frac{\rho + m}{2}} \cos \theta_1 + \sqrt{\frac{\rho - m}{2}} \cos \theta_2 \right)^2 \right).$$

The proof of this representation, see Sect. 4.5, follows by using the convolution structure of the microcanonical partition function and identifying the generating function to be the product of modified Bessel functions of the second kind. The proof is concluded by differentiation of these Bessel functions.

In the previous result, we introduced the overloaded  $s$  function by adding an angular dependence. We will differentiate between these functions by always specifying, in one form or another, the number of arguments the function takes.

In the following, we will specialize to functions  $g$  that are infinitely continuously differentiable, and obtain there finitely many global maximum points in the interval  $(-1, 1)$ . In light of Corollary 4.3.10, our goal is to study quantities of the form

$$\frac{\kappa_n^g(B(m^*, \delta))}{\sum_{m^* \in M^*(\psi^g)} \kappa_n^g(B(m^*, \delta))} = \frac{\int_{m^*-\delta}^{m^*+\delta} dm e^{n(g(m)+s_n(m,1))}}{\sum_{m^* \in M^*(\psi^g)} \int_{m^*-\delta}^{m^*+\delta} dm e^{n(g(m)+s_n(m,1))}}.$$

Using Lemma 3.3.1, it follows that

$$\begin{aligned} & \frac{(2n)! n^2 \pi^2}{2^{2n-1} n^{n-2} n! \binom{2n}{2} e^{-(n-1)}} \int_{m^*-\delta}^{m^*+\delta} dm e^{n(g(m)+s_n(m,1))} & (3.3.1) \\ &= \int_{m^*-\delta}^{m^*+\delta} dm \int_0^\pi d\theta_1 \int_0^\pi d\theta_2 \frac{\cos \theta_1 \cos \theta_2 e^{g(m)}}{\sqrt{1 - m^2}} e^{(n-1)(g(m)+s(m,1,\theta_1,\theta_2))} \\ &= \int_{m^*-\delta}^{m^*+\delta} dm \int_0^\pi d\theta_1 \int_0^\pi d\theta_2 \frac{\cos \theta_1 \cos \theta_2 e^{g(m)}}{\sqrt{1 - m^2}} e^{(n-1)(\psi^g(m,\theta_1,\theta_2))}, \end{aligned}$$

where we have introduced the overloaded function  $\psi^g : (-1, 1) \times [0, 2\pi) \times [0, 2\pi)$  given by

$$\psi^g(m, \theta_1, \theta_2) := g(m) + 1 + \ln \left( \left( \sqrt{\frac{1+m}{2}} \cos \theta_1 + \sqrt{\frac{1-m}{2}} \cos \theta_2 \right)^2 \right).$$

We see that the integral in Eq. (3.3.1) takes the form of a Laplace-type integral in three variables, and we expect that the local structure around the global maximum points of the overloaded function  $\psi^g$  determines the exponential asymptotics of such integrals precisely. To that end, we present the following result which contains the relevant information concerning the structure and local asymptotics of the overloaded  $\psi^g$  function.

**Lemma 3.3.2** *Suppose that  $\psi^g$  has a local maximizing point  $m^*$  contained in the interval  $(m^* - \delta, m^* + \delta)$ , and there exists  $k \in \mathbb{N}$  such that  $\partial^{2k}[\psi^g](m^*) < 0$  and  $\partial^j[\psi^g](m^*) = 0$  for all  $1 \leq j \leq 2k - 1$ .*

*Then, it follows that*

$$\begin{aligned} \psi^g(m^* + m, \theta_1, \theta_2) &= \psi^g(m^*) + \frac{1}{2} \partial_2^2[\psi^g](m^*, 0, 0) \theta_1^2 + \frac{1}{2} \partial_3^2[\psi^g](m^*, 0, 0) \theta_2^2 \\ &+ \frac{1}{(2k)!} \partial^{2k}[\psi^g](m^*) m^{2k} + \sum_{|\alpha|=3, \alpha_1 \notin \{2,3\}} R_\alpha(m, \theta_1, \theta_2) (m, \theta_1, \theta_2)^\alpha \\ &+ R_{(2k+1,0,0)}(m, \theta_1, \theta_2) m^{2k+1}, \end{aligned}$$

where

$$R_\alpha(m, \theta_1, \theta_2) = \frac{|\alpha|}{\alpha!} \int_0^1 dt (1-t)^{|\alpha|-1} \partial_\alpha[\psi^g](m^* + t(m, \theta_1, \theta_2)).$$

In addition,

$$\begin{aligned} \lim_{n \rightarrow \infty} n \left( \psi^g \left( m^* + \frac{m}{n^{2k}}, \frac{\theta_1}{n^{\frac{1}{2}}}, \frac{\theta_2}{n^{\frac{1}{2}}} \right) - \psi^g(m^*) \right) \\ = \frac{1}{2} \partial_2^2[\psi^g](m^*, 0, 0) \theta_1^2 + \frac{1}{2} \partial_3^2[\psi^g](m^*, 0, 0) \theta_2^2 + \frac{1}{(2k)!} \partial^{2k}[\psi^g](m^*) m^{2k}. \end{aligned}$$

The proof of this result, see Sect. 4.5, follows by developing the Taylor polynomial of the overloaded  $\psi^g$  function around the point  $(m^*, 0, 0)$ , and using the fact that odd derivatives of cosines vanish when evaluated at 0. The second statement simply follows by taking the limit.

From the previous result, we see that it is pertinent to introduce the following classification, which is directly adapted from [2], of the global maxima of  $\psi^g$ .

**Definition 3.3.3** A global maximum point  $m^* \in (-1, 1)$  of  $\psi^g$  is said to be of type  $k(m^*) \in \mathbb{N}$  if  $\partial^{2k}[\psi^g](m^*) < 0$  and  $\partial^j[\psi^g](m^*) = 0$  for all  $1 \leq j \leq 2k - 1$ .

For a finite collection of global maximum points  $M^*(\psi^g) \subset (-1, 1)$  of  $\psi^g$ , the maximal type  $k_\infty(\psi^g)$  is given by  $k_\infty(\psi^g) = \max_{m^* \in M^*(\psi^g)} k(m^*)$ . The collection of global maximum points of maximal type  $M_\infty^*(\psi^g)$  is given by  $M_\infty^*(\psi^g) := \{m^* \in (-1, 1) : k(m^*) = k_\infty(\psi^g)\}$ .

Combining together Lemmas 3.3.1 and 3.3.2, and the form given in Eq. (3.3.1), we have the following asymptotic result.

**Lemma 3.3.4** *Suppose that  $\psi^g$  has a single unique maximizing point  $m^* \in (m^* - \delta, m^* + \delta)$  of type  $k \in \mathbb{N}$ . Then, it follows that*

$$\begin{aligned} &\lim_{n \rightarrow \infty} \frac{n^{\frac{1}{2k}+1} \int_{m^*-\delta}^{m^*+\delta} dm e^{n(g(m)+s_n(m,1))}}{e^{n\psi^g(m^*)}} \frac{(2n)!n^2\pi^2}{2^{2n-1}n^{n-2}n! \binom{2n}{2} e^{-(n-1)}} \\ &= \frac{e^{g(m^*)}}{e^{\psi^g(m^*)} \sqrt{1-m^{*2}}} \int_{\mathbb{R}^3} d\theta_1 d\theta_2 dm e^{\frac{1}{2}\partial_2^2[\psi^g](m^*,0,0)\theta_1^2 + \frac{1}{2}\partial_3^2[\psi^g](m^*,0,0)\theta_2^2 + \frac{1}{(2k)!}\partial^{2k}[\psi^g](m^*)m^{2k}}. \end{aligned}$$

The proof of this result, see Sect. 4.5, is a standard application of the multivariate Laplace method.

From the previous result, denote  $W_n(g, m^*, \delta)$  to be the quantity given by

$$W_n^g(m^*, \delta) := \frac{n^{\frac{1}{2k}+1} \int_{m^*-\delta}^{m^*+\delta} dm e^{n(g(m)+s_n(m,1))}}{e^{n\psi^g(m^*)}} \frac{(2n)!n^2\pi^2}{2^{2n-1}n^{n-2}n! \binom{2n}{2} e^{-(n-1)}},$$

and its limit  $W(g, m^*)$  given by

$$\begin{aligned} W^g(m^*) &:= \lim_{n \rightarrow \infty} W_n^g(m^*, \delta) = \frac{e^{g(m^*)}}{e^{\psi^g(m^*)} \sqrt{1-m^{*2}}} \\ &\int_{\mathbb{R}^3} d\theta_1 d\theta_2 dm e^{\frac{1}{2}\partial_2^2[\psi^g](m^*,0,0)\theta_1^2 + \frac{1}{2}\partial_3^2[\psi^g](m^*,0,0)\theta_2^2 + \frac{1}{(2k)!}\partial^{2k}[\psi^g](m^*)m^{2k}}. \end{aligned}$$

To resolve the weak convergence of the mixture measure, using both Lemma 3.3.4 and Corollary 4.3.10, we compute

$$\frac{\kappa_n^g(\overline{B}(m' - \delta, m' + \delta))}{\sum_{m^* \in \mathcal{M}^*(\psi^g)} \kappa_n^g(\overline{B}(m^* - \delta, m + \delta))} = \frac{n^{-\left(\frac{1}{2k(m')} - \frac{1}{2k_\infty}\right)} W_n^g(m', \delta)}{\sum_{m^* \in \mathcal{M}^*(\psi^g)} n^{-\left(\frac{1}{2k(m^*)} - \frac{1}{2k_\infty}\right)} W_n^g(m^*, \delta)},$$

from which it follows that

$$\lim_{n \rightarrow \infty} \frac{\kappa_n^g(\overline{B}(m' - \delta, m' + \delta))}{\sum_{m^* \in \mathcal{M}^*(\psi^g)} \kappa_n^g(\overline{B}(m^* - \delta, m + \delta))} = \begin{cases} \frac{W^g(m')}{\sum_{m^* \in \mathcal{M}_\infty(\psi^g)} W^g(m^*)}, & k(m') = k_\infty(\psi^g) \\ 0, & k(m') < k_\infty(\psi^g) \end{cases}$$

Following this computation, we have the following result.

**Theorem 3.3.5** *Let  $g \in C_b([-1, 1])$  be an infinitely continuously differentiable function such that  $\psi^g$  has finitely many global maximizing points  $M^*(\psi^g) \subset (-1, 1)$  of finite type. Then, it follows that*

$$\lim_{n \rightarrow \infty} \kappa_n^g = \left( \sum_{m^* \in M_\infty(\psi^g)} W^g(m^*) \right)^{-1} \sum_{m^* \in M_\infty(\psi^g)} W^g(m^*) \delta_{m^*}.$$

To finish, we can directly compute the following

$$\partial_2^2[\psi^g](m^*, 0, 0) = -\frac{2\sqrt{\frac{1+m^*}{2}}}{\sqrt{\frac{1+m^*}{2}} + \sqrt{\frac{1-m^*}{2}}}, \quad -\partial_3^2[\psi^g](m^*, 0, 0) = -\frac{2\sqrt{\frac{1-m^*}{2}}}{\sqrt{\frac{1+m^*}{2}} + \sqrt{\frac{1-m^*}{2}}},$$

this implies that the integral containing these terms does not depend on  $g$ , other than through the value of the global maximizing point. In addition, it is immediate that the factor  $e^{\psi^g(m^*)-g(m^*)}$  does not depend on  $g$  either. Furthermore, we immediately have

$$\int_{-\infty}^{\infty} dm e^{\frac{1}{(2k)!} \partial^{2k}[\psi^g](m^*)m^{2k}} = \frac{1}{|\partial^{2k}[\psi^g](m^*)|^{\frac{1}{2k}}} \int_{-\infty}^{\infty} dm e^{-\frac{m^{2k}}{(2k)!}}.$$

We can thus combine all factors not depending functionally on  $g$  into a single function  $C^k : (-1, 1) \rightarrow (0, \infty)$  given by

$$C^k(m^*) := \frac{e^{g(m^*)}}{e^{\psi^g(m^*)} \sqrt{1 - m^{*2}}} \int_{\mathbb{R}^3} d\theta_1 d\theta_2 dm e^{\frac{1}{2} \partial_2^2[\psi^g](m^*, 0, 0) \theta_1^2 + \frac{1}{2} \partial_3^2[\psi^g](m^*, 0, 0) \theta_2^2 - \frac{m^{2k}}{(2k)!}},$$

so that

$$W^g(m^*) = \frac{C^k(m^*)}{|\partial^{2k}[\psi^g](m^*)|^{\frac{1}{2k}}}.$$

Using Corollary 3.1.7, Theorem 3.3.5, Lemma 3.1.1, and the form of the weights  $W^g(m^*)$  given above, we have the final result.

**Theorem 3.3.6** *Let  $g \in C_b([-1, 1])$  be an infinitely continuously differentiable function such that  $\psi^g$  has finitely many global maximizing points  $M^*(\psi^g) \subset (-1, 1)$  of finite type, and let  $k_\infty := k_\infty(\psi^g)$ .*

*Then, it follows that*

$$\lim_{n \rightarrow \infty} \mu_n^g = \left( \sum_{m^* \in M_\infty^*(\psi^g)} \frac{C^{k_\infty}(m^*)}{|\partial^{2k_\infty}[\psi^g](m^*)|} \right)^{-1} \sum_{m^* \in M_\infty^*(\psi^g)} \frac{C^{k_\infty}(m^*)}{|\partial^{2k_\infty}[\psi^g](m^*)|} \eta(\beta(m^*, 1), \mu(m^*, 1)).$$

*This concludes the presentation of the main results of this paper.*

### 4 Intermediate Results and Proofs

This section contains proof of some of the results in Sect. 3, and some collections of intermediate results and theory that are required.

#### 4.1 Microcanonical Probability Measures

To motivate the rigorous definition of the microcanonical ensemble and its associated probability measure, consider the following formal calculation

$$\begin{aligned} & \int_{\mathbb{R}^n} d\phi \delta(M_n(\phi) - mn) \delta(N_n(\phi) - \rho n) f(\phi) \\ &= \sum_{\sigma \in \{-1, 1\}^n} \int_{[0, \infty)^n} d\phi \delta \left( \sum_{i \in \sigma^{-1}\{+1\}} \phi_i - \sum_{i \in \sigma^{-1}\{-1\}} \phi_i - mn \right) \delta \end{aligned}$$

$$\begin{aligned} & \left( \sum_{i \in \sigma^{-1}\{+1\}} \phi_i + \sum_{i \in \sigma^{-1}\{-1\}} \phi_i - \rho n \right) f(\sigma\phi) \\ &= \sum_{\sigma \in \{-1, 1\}^n} \frac{1}{2} \int_{[0, \infty)^n} d\phi \delta \left( \sum_{i \in \sigma^{-1}\{+1\}} \phi_i - \frac{\rho + m}{2} n \right) \delta \left( \sum_{i \in \sigma^{-1}\{-1\}} \phi_i - \frac{\rho - m}{2} n \right) f(\sigma\phi), \end{aligned}$$

where the pair  $(m, \rho) \in \mathcal{A}$ ,  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  is a sufficiently regular function, and  $\sigma\phi$  notation for a multiplication map defined by  $(\sigma\phi)_i := \sigma_i\phi_i$ . Note that the integral in the sum is a product of two integrals since the index sets  $\sigma^{-1}\{+1\}$  and  $\sigma^{-1}\{-1\}$  are trivially disjoint. Note that the primary formal rule we have made use of is the following one

$$\delta(Tx - y) = \frac{1}{|\det(T)|} \delta(x - T^{-1}y)$$

for an invertible linear map  $T : \mathbb{R}^k \rightarrow \mathbb{R}^k$ , and elements  $x, y \in \mathbb{R}^k$ .

To make this formal calculation rigorous, we need to define integrals over scaled simplexes in arbitrary dimensions. To do this, we introduce the so-called flag coordinates  $\phi' : \mathbb{R}^k \rightarrow \mathbb{R}^k$  given by

$$\phi'_i(\phi) := \sum_{j=1}^i \phi_j.$$

Note that  $\phi'([0, \infty)^k) = \{\phi \in [0, \infty)^k : \phi_1 \leq \phi_2 \leq \dots \leq \phi_k\}$ ,  $\det(\phi') = 1$ , and the inverse function of  $\phi'$  is given by

$$\phi'^{-1}(\phi') = \phi'_i - \phi'_{i-1},$$

where we take the convention that  $\phi'_0 := 0$ .

The connection between the flag coordinates and the integrals over simplexes can be seen from the following formal calculation

$$\begin{aligned} & \int_{[0, \infty)^k} d\phi \delta \left( \sum_{i=1}^k \phi_i - r \right) f(\phi) \\ &= \int_{[0, \infty)^k} d\phi \delta(\phi_k - r) \mathbb{1}(\phi_1 \leq \phi_2 \leq \dots \leq \phi_k) f(\phi_1, \phi_2 - \phi_1, \dots, \phi_k - \phi_{k-1}) \\ &= \int_{[0, \infty)^{k-1}} d\phi \mathbb{1}(\phi_1 \leq \phi_2 \leq \dots \leq \phi_{k-1} \leq r) f(\phi_1, \phi_2 - \phi_1, \dots, r - \phi_{k-1}), \end{aligned}$$

where  $r > 0$ , and  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  is a sufficiently regular function.

From this formal calculation, we produce the following definition.

**Definition 4.1.1** For a finite index set  $I$  and  $r > 0$ , the measure  $S_I(r)$  on  $[0, \infty)^I$  corresponding to the integral over an  $(|I| - 1)$ -dimensional  $r$ -scaled simplex on the index set  $I$  is given by its action on  $f \in C_b([0, \infty)^I)$  given by

$$S_I(r)[f] := \int_{[0,\infty)^{k-1}} d\phi \mathbb{1}(\phi_{i_1} \leq \phi_{i_2} \leq \dots \leq \phi_{i_{|I|-1}} \leq r) f(\phi_{i_1}, \phi_{i_2} - \phi_{i_1}, \dots, r - \phi_{i_{|I|-1}}),$$

where  $\{i_k\}_{k=1}^{|I|}$  is some enumeration of  $I$ .

For future use, whenever it is clear that we are either referring to the measure or the normalization constant, we will use the following notation

$$S_I(r) := S_I(r)[1] = \frac{r^{|I|-1}}{(|I| - 1)!},$$

where the right-hand side follows by direct computation. Using dominated convergence, it is also clear that the mapping  $r \mapsto S_I(r)[f]$  is continuous if  $f \in C_b([0, \infty)^I)$  is continuous. To show that Definition 4.1.1 is independent of the enumeration of  $I$  given above, we will use a Lebesgue-absolutely continuous approximation of  $S_I(r)$ . Let  $g : [0, \infty) \rightarrow \mathbb{R}$  be a measurable function such that

$$\int_0^\infty dr |g(r)|r^{|I|-1} < \infty.$$

It follows that

$$\int_{[0,\infty)^I} d\phi g\left(\sum_{i \in I} \phi_i\right) f(\phi) = \int_0^\infty dr g(r)S_I(r)[f],$$

where  $f \in C_b([0, \infty)^I)$ . Now, consider the family  $\{g_\varepsilon\}_{\varepsilon>0}$  given by

$$g_\varepsilon(r) := \frac{\mathbb{1}(|r| < \varepsilon)}{2\varepsilon}. \tag{4.1.1}$$

Fix  $r > 0$ . Since  $f \in C_b([0, \infty)^I)$ , as stated before, one can verify that  $S_I(\cdot)[f] \in C([0, \infty))$ . It follows that

$$S_I(r)[f] = \lim_{\varepsilon \rightarrow 0^+} \int_0^\infty dr' g_\varepsilon(r' - r)S_I(r')[f] = \lim_{\varepsilon \rightarrow 0^+} \int_{[0,\infty)^I} d\phi g_\varepsilon\left(\sum_{i \in I} \phi_i - r\right) f(\phi).$$

We see that the left-hand side of the above equality will inherit properties from the right-hand side limiting term. In particular, the measure given by its action on  $f \in C_b([0, \infty)^I)$  given by

$$f \mapsto \int_{[0,\infty)^I} d\phi g_\varepsilon\left(\sum_{i \in I} \phi_i\right) f(\phi)$$

is independent of any enumeration of  $I$ , and it is label permutation invariant. It follows that the measure  $S_I(r)$  is independent of the given enumeration in the definition, and it is label permutation invariant.

We can now define the microcanonical probability measure using Definition 4.1.1.

**Definition 4.1.2** The measure  $Z_n(M, N)$  is given by its action on  $f \in C_b(\mathbb{R}^n)$  given by

$$Z_n(M, N)[f] := \begin{cases} \frac{1}{2} \sum_{\sigma \in \{-1, 1\}^n} (S_{\sigma^{-1}\{+1\}}\left(\frac{N+M}{2}\right) \otimes S_{\sigma^{-1}\{-1\}}\left(\frac{N-M}{2}\right)) [f \circ \sigma], & (M, N) \in \mathcal{A}, \\ S_n(N)[f], & (M, N) \in \partial\mathcal{A} \setminus \{0\}, \end{cases}$$

where  $\otimes(\cdot)$  is the tensor product of two measures,  $f \circ \sigma$  is the composition of the multiplication map  $\sigma$  with  $f$ , and we take the necessary convention that

$$\left( S_{\sigma^{-1}\{+1\}} \left( \frac{N+M}{2} \right) \otimes S_{\sigma^{-1}\{-1\}} \left( \frac{N-M}{2} \right) \right) [f \circ \sigma] = 0$$

if  $\sigma = \{1, 1, \dots, 1\}$  or  $\sigma = \{-1, -1, \dots, -1\}$  whenever  $(M, N) \in \mathcal{A}$ .

This last convention implies that we do not include the ‘‘first’’ and ‘‘last’’ in the sum, but we have left them in to save space on notation.

To conclude this section, we will, finally, give the definition of the microcanonical probability measure.

**Definition 4.1.3** For  $(m, \rho) \in \overline{\mathcal{A}} \setminus \{0\}$ , the probability measure  $\nu_n(m, \rho)$  on  $\mathbb{R}^n$  corresponding to the microcanonical probability measure is defined by its action on  $f \in C_b(\mathbb{R}^n)$  given by

$$\nu_n(m, \rho)[f] := \frac{Z_n(mn, \rho n)[f]}{Z_n(mn, \rho n)}, \tag{4.1.2}$$

and the microcanonical partition function, acting as the normalization constant  $Z_n(mn, \rho n)$  is given by

$$Z_n(mn, \rho n) := Z_n(mn, \rho n)[1] = \begin{cases} \frac{1}{2} \sum_{k=1}^{n-1} \binom{n}{k} \frac{\left(\frac{\rho n + mn}{2}\right)^{k-1}}{(k-1)!} \frac{\left(\frac{\rho n - mn}{2}\right)^{n-k-1}}{(n-k-1)!}, & (m, \rho) \in \mathcal{A}, \\ \frac{(\rho n)^{n-1}}{(n-1)!}, & (m, \rho) \in \partial \mathcal{A} \setminus \{0\}, \end{cases} \tag{4.1.3}$$

which can be verified by direct computation.

To make the microcanonical probability measure computationally tractable, we will utilize a similar Lebesgue-absolutely continuous approximation as for the integrals over the simplex. However, as opposed to the approximation for the integrals over the simplexes, one must be more careful here. Using the family of functions  $\{g_\varepsilon\}_{\varepsilon>0}$  from Eq. (4.1.1), observe that

$$\begin{aligned} & \int_{[0, \infty)^{\sigma^{-1}\{+1\}} \times [0, \infty)^{\sigma^{-1}\{-1\}}} d\phi \\ & \times g_\varepsilon \left( \sum_{i \in \sigma^{-1}\{+1\}} \phi_i - \frac{\rho n + mn}{2} \right) g_\varepsilon \left( \sum_{i \in \sigma^{-1}\{-1\}} \phi_i - \frac{\rho n - mn}{2} \right) f(\sigma\phi) \\ & = \int_{[0, \infty)^{\sigma^{-1}\{+1\}} \times (-\infty, 0]^{\sigma^{-1}\{-1\}}} d\phi \\ & \times g_\varepsilon \left( \sum_{i=1}^n \frac{|\phi_i| + \phi_i}{2} - \frac{\rho n + mn}{2} \right) g_\varepsilon \left( \sum_{i=1}^n \frac{|\phi_i| - \phi_i}{2} - \frac{\rho n - mn}{2} \right) f(\phi), \end{aligned}$$

where  $(m, \rho) \in \mathcal{A}$ , and  $\sigma$  does not consist of all 1’s or all  $-1$ ’s. Now, if we consider instead the right-hand side first, then it makes sense even when  $\sigma$  consists of all 1’s or  $-1$ ’s. In that instance, the argument of one of the  $g_\varepsilon$  will not integrate over any  $\phi$ -variables, and for small enough  $\varepsilon > 0$  the indicator function vanishes. Summing over the  $\sigma$ , in this case, it then follows that



$$Z_n(mn, \rho n)[f] \tag{4.1.4}$$

$$= \lim_{\varepsilon \rightarrow 0^+} \frac{1}{2} \int_{\mathbb{R}^n} d\phi \, g_\varepsilon \left( \sum_{i=1}^n \frac{|\phi_i| + \phi_i}{2} - \frac{\rho n + mn}{2} \right) g_\varepsilon \left( \sum_{i=1}^n \frac{|\phi_i| - \phi_i}{2} - \frac{\rho n - mn}{2} \right) f(\phi).$$

Returning now to the microcanonical probability measure, we see that it inherits the various properties of the measure with action on  $f \in C_b(\mathbb{R}^n)$  given by

$$f \mapsto \int_{\mathbb{R}^n} d\phi \, g_\varepsilon \left( \sum_{i=1}^n \frac{|\phi_i| + \phi_i}{2} - \frac{\rho n + mn}{2} \right) g_\varepsilon \left( \sum_{i=1}^n \frac{|\phi_i| - \phi_i}{2} - \frac{\rho n - mn}{2} \right) f(\phi).$$

In particular, it is label permutation invariant. Furthermore, this approximation will be used for some calculations related to the microcanonical probability measure.

### 4.2 Relative Entropy and Local Observables

We begin with the proof of a type of generalized dominated convergence theorem.

**Proof of Lemma 3.1.1** The condition that  $K$  is a continuity set of  $\mu$  implies that

$$\lim_{n \rightarrow \infty} \mu_n(K) = \mu(K),$$

and the condition that  $\text{supp}(\mu) \subset K$  implies that  $\mu(K) = 1$ .

Next, we have the following two simple inequalities

$$\left| \int_X \mu_n(dx) f_n(x) - \int_K \mu_n(dx) f_n(x) \right| \leq \mu_n(X \setminus K) \sup_{n \in \mathbb{N}} \sup_{x \in X} |f_n(x)|,$$

and

$$\left| \int_K \mu_n(dx) f_n(x) - \int_K \mu_n(dx) f(x) \right| \leq \mu_n(K) \sup_{x \in K} |f_n(x) - f(x)|.$$

Since  $K$  is a continuity set of  $\mu$ , using the continuity set definition of weak convergence, it follows that  $\mu_n$  conditioned to  $K$  converges weakly to  $\mu$  conditioned to  $K$ . Transitioning to the continuous bounded form of weak convergence, it follows that

$$\lim_{n \rightarrow \infty} \frac{1}{\mu_n(K)} \int_K \mu_n(dx) f(x) = \frac{1}{\mu(K)} \int_K \mu(dx) f(x) = \int_K \mu(dx) f(x).$$

For completeness, we have the following final inequality

$$\left| \int_K \mu_n(dx) f(x) - \int_X \mu(dx) f(x) \right| \leq \frac{\mu_n(X \setminus K)}{\mu_n(K)} \left| \int_X \mu_n(dx) f(x) \right|.$$

Combining together all three inequalities, the result follows. □

We will need the relative entropy between two absolutely continuous probability measures.

**Definition 4.2.1** Let  $X$  be a Polish space, and let  $\mu$  and  $\nu$  be probability measures on  $X$ . If  $\mu$  is absolutely continuous with respect to  $\nu$ , the relative entropy  $\mathcal{H}(\mu||\nu)$  is given by

$$\mathcal{H}(\mu||\nu) := \int_X d\mu \ln \frac{d\mu}{d\nu}.$$

If  $\mu$  is not absolutely continuous with respect to  $\nu$ , we set  $\mathcal{H}(\mu||\nu) = \infty$ .

We will need the following properties of relative entropy.

**Theorem 4.2.2** *Let  $X$  be a Polish space, and let  $\mu$  and  $\nu$  be probability measures on  $X$  such that  $\mu$  is absolutely continuous with respect to  $\nu$ .*

- For any  $\mu$  and  $\nu$  satisfying the assumptions

$$\mathcal{H}(\mu||\nu) \geq 0.$$

- For any  $\mu$  and  $\nu$  satisfying the assumptions

$$\sup_{f \in M_b(X), \|f\|_\infty \leq 1} |\mu[f] - \nu[f]| \leq \sqrt{\frac{\mathcal{H}(\mu||\nu)}{2}},$$

where  $M_b(X)$  is the space of measurable bounded functions on  $X$ .

- If  $X = Y^n$ , where  $Y$  is another Polish space, and  $\nu = \otimes_{k=1}^n \lambda$ , where  $\lambda$  is a probability measure on  $Y$ , it follows that

$$\mathcal{H}_I(\mu||\nu) + \mathcal{H}_J(\mu||\nu) \leq \mathcal{H}_{I \cup J}(\mu||\nu) + \mathcal{H}_{I \cap J}(\mu||\nu),$$

where  $I, J \subset \{1, 2, \dots, n\}$ , and  $\mathcal{H}_I(\mu||\nu)$  is denotes the relative entropy of the  $I$ :th marginal distributions of  $\mu$  and  $\nu$ .

The first and third properties are discussed and given proofs in [23]. The second property is sometimes referred to as Pinsker’s inequality and references to proofs and other details concerning this inequality can be found in [24].

We can now give a proof of the fundamental inequality connecting the constrained and non-constrained ensemble probability measures.

**Proof of Lemma 3.1.2** Using Eq. (4.1.4), we can compute the integral over only the first 2 variables leaving the other  $n - 2$  variables fixed. We compute

$$\begin{aligned} & \int_{\mathbb{R}^2} d\phi \, g_\varepsilon \left( \sum_{i=1}^n \frac{|\phi_i| + \phi_i}{2} - \frac{\rho n + mn}{2} \right) g_\varepsilon \left( \sum_{i=1}^n \frac{|\phi_i| - \phi_i}{2} - \frac{\rho n - mn}{2} \right) \\ &= g_\varepsilon \left( \sum_{i=3}^n \frac{|\phi_i| - \phi_i}{2} - \frac{\rho n - mn}{2} \right) \int_{[0, \infty)^2} d\phi \, g_\varepsilon \left( \phi_1 + \phi_2 - \left( \frac{\rho n + mn}{2} - \sum_{i=3}^n \frac{|\phi_i| + \phi_i}{2} \right) \right) \\ &+ 2 \int_{[0, \infty)^2} d\phi \, g_\varepsilon \left( \phi_1 - \left( \frac{\rho n + mn}{2} - \sum_{i=3}^n \frac{|\phi_i| + \phi_i}{2} \right) \right) g_\varepsilon \left( \phi_2 - \left( \frac{\rho n - mn}{2} - \sum_{i=3}^n \frac{|\phi_i| - \phi_i}{2} \right) \right) \\ &+ g_\varepsilon \left( \sum_{i=3}^n \frac{|\phi_i| + \phi_i}{2} - \frac{\rho n + mn}{2} \right) \int_{[0, \infty)^2} d\phi \, g_\varepsilon \left( \phi_1 + \phi_2 - \left( \frac{\rho n - mn}{2} - \sum_{i=3}^n \frac{|\phi_i| - \phi_i}{2} \right) \right). \end{aligned}$$

Taking the limit, it follows that

$$\begin{aligned} & \lim_{\varepsilon \rightarrow 0^+} \int_{\mathbb{R}^2} d\phi \, g_\varepsilon \left( \sum_{i=1}^n \frac{|\phi_i| + \phi_i}{2} - \frac{\rho n + mn}{2} \right) g_\varepsilon \left( \sum_{i=1}^n \frac{|\phi_i| - \phi_i}{2} - \frac{\rho n - mn}{2} \right) \\ &= 2 \mathbb{1} \left( \frac{\rho n + mn}{2} - \sum_{i=3}^n \frac{|\phi_i| + \phi_i}{2} \geq 0 \right) \mathbb{1} \left( \frac{\rho n - mn}{2} - \sum_{i=3}^n \frac{|\phi_i| - \phi_i}{2} \geq 0 \right). \end{aligned}$$

Accounting for the normalization, the  $(n - 2)$ :th marginal of the microcanonical probability measure is given by

$$\nu_n(m, \rho)(d\phi_{n-2}) = \frac{\mathbb{1} \left( \frac{\rho n + mn}{2} - \sum_{i=3}^n \frac{|\phi_i| + \phi_i}{2} \geq 0 \right) \mathbb{1} \left( \frac{\rho n - mn}{2} - \sum_{i=3}^n \frac{|\phi_i| - \phi_i}{2} \geq 0 \right)}{Z_n(mn, \rho n)} d\phi_{n-2},$$

where  $d\phi_{n-2}$  is the  $(n - 2)$ -dimensional Lebesgue measure. Note the factor of 2 vanishes due to the presence of a factor of  $\frac{1}{2}$  in the partition function. It follows that

$$\frac{dv_n(m, \rho)}{d\eta_n(\beta, \mu)}(\phi_{n-2}) = \frac{Q_{n-2}(\beta, \mu)}{e^{-\beta \sum_{i=3}^n \phi_i - \mu \sum_{i=3}^n |\phi_i|}} \times \frac{\mathbb{1}\left(\frac{\rho n + mn}{2} - \sum_{i=3}^n \frac{|\phi_i| + \phi_i}{2} \geq 0\right) \mathbb{1}\left(\frac{\rho n - mn}{2} - \sum_{i=3}^n \frac{|\phi_i| - \phi_i}{2} \geq 0\right)}{Z_n(mn, \rho n)} d\phi_{n-2}.$$

The relative entropy is then directly computed to be

$$\mathcal{H}_{n-2}(v_n(m, \rho) || \eta_n(\beta, \mu)) = \beta v_n(m, \rho) [M_{n-2}] + \mu v_n(m, \rho) [N_{n-2}] + \ln Q_{n-2}(\beta, \mu) - \ln Z_n(mn, \rho n).$$

Using label permutation invariance, one can directly compute that

$$v_n(m, \rho) [M_{n-2}] = (n - 2)m, \quad v_n(m, \rho) [N_{n-2}] = (n - 2)\rho, \quad \ln Q_{n-2}(\beta, \mu) = (n - 2)f(\beta, \mu).$$

In summary, we have

$$\frac{1}{n - 2} \mathcal{H}_{n-2}(v_n(m, \rho) || \eta_n(\beta, \mu)) = \beta m + \mu \rho + f(\beta, \mu) - \frac{n}{n - 2} s_n(m, \rho).$$

To continue, by Theorem 4.2.2, it follows that

$$\sup_{f \in C_b(\mathbb{R}^I), \|f\|_\infty \leq 1} |v_n(m, \rho)[f] - \eta_n(\beta, \mu)[f]| \leq \sqrt{\frac{\mathcal{H}_I(v_n(m, \rho) || \eta_n(\beta, \mu))}{2}}.$$

By label permutation invariance, it follows that

$$\mathcal{H}_I(v_n(m, \rho) || \eta_n(\beta, \mu)) = \mathcal{H}_{[|I|]}(v_n(m, \rho) || \eta_n(\beta, \mu)).$$

Since  $I$  is finite, it follows that there exists  $k \in \mathbb{N}$  such that  $(k - 1)|I| \leq n - 2 < k|I|$ . Since  $\eta_n(\beta, \mu)$  is a product measure, using Theorem 4.2.2, it follows that

$$\begin{aligned} \mathcal{H}_I(v_n(m, \rho) || \eta_n(\beta, \mu)) &= \frac{1}{k - 1} \sum_{j=1}^{k-1} \mathcal{H}_{[|I|+(j-1)|I]}(v_n(m, \rho) || \eta_n(\beta, \mu)) \\ &\leq \frac{\mathcal{H}_{n-2}(v_n(m, \rho) || \eta_n(\beta, \mu))}{k - 1} \\ &\leq |I| \frac{\mathcal{H}_{n-2}(v_n(m, \rho) || \eta_n(\beta, \mu))}{n - 2 - |I|}. \end{aligned}$$

Combining these inequalities together, it follows that

$$\begin{aligned} &\sup_{f \in C_b(\mathbb{R}^I), \|f\|_\infty \leq 1} |v_n(m, \rho)[f] - \eta_n(\beta, \mu)[f]| \\ &\leq \sqrt{\frac{|I|(n - 2)}{2(n - 2 - |I|)} \left( \beta m + \mu \rho + f(\beta, \mu) - \frac{n}{n - 2} s_n(m, \rho) \right)}, \end{aligned}$$

as desired. □

### 4.3 Large Deviations and Weak Convergence

We begin with the standard key definitions of large deviations theory. Note that these definitions are either the same or slightly modified versions of the same results and definitions found in [22]. In addition, the result concerning convexity are either provided in [22], or we refer to [20] for more detailed analysis of convex objects.

In the following  $\{P_n\}_{n=1}^\infty$  is a sequence of probability measures on a Polish space  $X$ .

**Definition 4.3.1** A function  $I : X \rightarrow [0, \infty]$  is called a rate function if it satisfies the following properties

- $I(x) < \infty$  for all  $x \in X$ .
- $I$  is lower semi-continuous.
- $I$  has compact level sets.

In the following, we use the notation  $I(A) := \inf_{x \in A} I(x)$ .

**Definition 4.3.2** A sequence of probability measure  $\{P_n\}_{n=1}^\infty$  is said to satisfy a large deviations principle with rate function  $I$  if it satisfies the following properties

- For all closed sets  $C \subset X$ , we have

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \ln P_n(C) \leq -I(C).$$

- For all open sets  $O \subset X$ , we have

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \ln P_n(O) \geq -I(O).$$

Now, we specialize to probability distributions on  $\mathbb{R}^d$ . In the following, let  $\{m_n\}_{n=1}^\infty$  be a sequence of random variables on  $\mathbb{R}^d$ , and we set  $P_n(A) := \mathbb{P}(m_n \in A)$ . The moment generating functions  $\varphi_n : \mathbb{R}^d \rightarrow (0, \infty]$  are given by  $\varphi_n(t) := \mathbb{E}e^{\langle t, m_n \rangle}$ . In the following, we assume the existence of a function  $\Lambda : \mathbb{R}^d \rightarrow [-\infty, \infty]$  given by

$$\Lambda(t) := \lim_{n \rightarrow \infty} \frac{1}{n} \ln \varphi_n(nt),$$

and that this function satisfies  $0 \in \text{int}(\mathcal{D}(\Lambda))$  where  $\mathcal{D}(\Lambda) := \{t \in \mathbb{R}^d : \Lambda(t) < \infty\}$ . For such a function, it follows that  $\Lambda$  is convex and  $\Lambda(t) > -\infty$  for all  $t \in \mathbb{R}^d$ . A convex function  $\Lambda : \mathbb{R}^d \rightarrow [-\infty, \infty]$  is called proper if  $\Lambda(t) > -\infty$  for all  $t \in \mathbb{R}^d$ , and there exists at least one point  $t_0 \in \mathbb{R}$  such that  $\Lambda(t_0) < \infty$ . It is clear that when  $\Lambda$  is the limit of the scaled logarithmic moment generating functions, then it is a proper convex function.

We will need the Legendre transform of  $\Lambda$ .

**Definition 4.3.3** The Legendre transform  $\Lambda^* : \mathbb{R}^d \rightarrow [-\infty, \infty]$  of a  $\Lambda : \mathbb{R}^d \rightarrow [-\infty, \infty]$  is given by

$$\Lambda^*(x) := \sup_{t \in \mathbb{R}^d} \{\langle x, t \rangle - \Lambda(t)\}.$$

For  $\Lambda$  given by the limit of the scaled logarithm, it follows that  $\Lambda^*$  is a convex rate function. In particular, we see that the range of  $\Lambda^*$  must be contained in  $[0, \infty)$ .

To specify the form of the Gärtner-Ellis theorem, that we wish to utilize, we need the concept of essential smoothness.

**Definition 4.3.4** A proper convex function  $\Lambda : \mathbb{R}^d \rightarrow (-\infty, \infty]$  is called essentially smooth if it satisfies the following properties

- $\text{int}(\mathcal{D}(\Lambda)) \neq \emptyset$ .
- $\Lambda$  is differentiable on  $\text{int}(\mathcal{D}(\Lambda))$ .
- Either  $\mathcal{D}(\Lambda) = \mathbb{R}^d$  or, for any  $t^* \in \partial\mathcal{D}(\Lambda)$ , it follows that  $\lim_{t \rightarrow t^*} \|\nabla[\Lambda](t)\| = \infty$ .

We can now give the essentially smooth form of the Gärtner-Ellis theorem.

**Theorem 4.3.5** Let  $\Lambda : \mathbb{R}^d \rightarrow (0, \infty]$  be an essentially smooth lower semi-continuous function.

It follows that  $\{P_n\}_{n=1}^\infty$  satisfies a large deviations principle with rate function  $\Lambda^*$ .

It is typical to introduce the notion of strict convexity of a function, but we will instead directly introduce the notion of a Legendre-type function.

**Definition 4.3.6** A proper convex lower semi-continuous function  $\Lambda : \mathbb{R}^d \rightarrow (-\infty, \infty]$  is said to be of Legendre-type if it is both essentially smooth and strictly concave on  $\text{int}(\mathcal{D}(\Lambda))$ .

The primary feature of Legendre-type functions that we will use is that the gradient of such a function  $\Lambda$  is a bijection between  $\text{int}(\mathcal{D}(\Lambda))$  and  $\text{int}(\mathcal{D}(\Lambda^*))$ .

We can now prove the following general theorem.

**Theorem 4.3.7** Let  $\{Z_n\}_{n=1}^\infty$  be a sequence of functions  $Z_n : n\mathcal{A} \rightarrow (0, \infty)$ , where  $\mathcal{A} \subset \mathbb{R}^d$  is a non-empty open convex set such that each  $Z_n$  is log-concave, and

$$\sup_{n \in \mathbb{N}} \left| \frac{1}{n} \ln Z_n(xn) \right| < \infty$$

for each  $x \in \mathcal{A}$ . Denote by  $s_n : \mathcal{A} \rightarrow (-\infty, \infty)$  the function given by

$$s_n(x) := \frac{1}{n} \ln Z_n(xn).$$

In addition, suppose that the function  $f : \mathbb{R}^d \rightarrow [-\infty, \infty]$  given by

$$f(t) := \lim_{n \rightarrow \infty} \frac{1}{n} \ln Q_n(t),$$

exists, where  $Q_n : \mathbb{R}^d \rightarrow (0, \infty]$  are given by

$$Q_n(t) := \int_{n\mathcal{A}} dX e^{-\langle t, X \rangle} Z_n(X),$$

and there exists a non-empty open convex set  $\mathcal{B} \subset \mathbb{R}^d$  such that  $\mathcal{D}(Q_n) = \text{int}(\mathcal{D}(f)) = \mathcal{B}$ .

If  $f$  is a proper convex lower semi-continuous function of Legendre type such that  $-\nabla[f]\mathcal{B} = \mathcal{A}$  then the function  $s : \mathcal{A} \rightarrow \mathbb{R}$  given by the limit

$$s(x) := \lim_{n \rightarrow \infty} \frac{1}{n} \ln Z_n(xn),$$

exists, and satisfies

$$s(x) := \inf_{t \in \mathbb{R}^d} \{ \langle t, x \rangle + f(t) \}, \quad \lim_{n \rightarrow \infty} \sup_{K \subset \mathcal{A}} |s_n(x) - s(x)| = 0$$

for any compact set  $K \subset \mathcal{A}$ .

**Proof** For the first step, let  $t_0 \in \mathcal{B}$  be any base point, and we define the sequence of probability measures  $\{P_n\}_{n=1}^\infty$  on  $\mathbb{R}^d$  by setting

$$P_n(A) := \frac{1}{Q_n(t_0)} \int_{n(A \cap \mathcal{A})} dX e^{-\langle t_0, X \rangle} Z_n(X),$$

where  $A \subset \mathbb{R}^d$  is Borel measurable.

The moment generating function  $\varphi_n : \mathbb{R}^d \rightarrow (0, \infty]$  of the random variable  $m_n$  on  $\mathbb{R}^d$  with distribution given by  $P_n$  is given by

$$\varphi_n(t) := \frac{Q_n(t_0 - \frac{t}{n})}{Q_n(t_0)}.$$

The limit of the scaled logarithm moment generating function  $\Lambda : \mathbb{R}^d \rightarrow [-\infty, \infty]$  is given by

$$\Lambda(t) := \lim_{n \rightarrow \infty} \frac{1}{n} \ln \varphi_n(nt) = f(t_0 - t) - f(t_0).$$

Since  $\Lambda$  inherits its properties from  $f$ , it follows that  $\Lambda$  exists, is a proper convex lower semi-continuous function of Legendre-type, and satisfies  $0 = t_0 - t_0 \in \text{int}(\mathcal{D}(\Lambda)) = t_0 - \text{int}(\mathcal{D}(f)) = t_0 - \mathcal{B}$ . It follows that  $\{P_n\}_{n \in \mathbb{N}}$  satisfies a large deviations principle with rate function  $\Lambda^*$ . Since  $\Lambda$  is of Legendre-type, it follows that  $\text{int}(\mathcal{D}(\Lambda^*)) = \nabla[\Lambda](\text{int}(\mathcal{D}(\Lambda))) = -\nabla[f]\mathcal{B} = \mathcal{A}$ .

Let  $y \in \text{int}(\mathcal{D}(\Lambda^*)) = \mathcal{A}$ . Since  $\Lambda^*$  is convex, it follows that it is continuous on  $\mathcal{A}$  and thus the compact balls  $\overline{B}(y, \delta)$  for small enough  $\delta > 0$  are continuity sets from which it follows that

$$\lim_{n \rightarrow \infty} \frac{1}{n} \ln P_n(\overline{B}(y, \delta)) = -\Lambda^*(\overline{B}(y, \delta)).$$

For the second step, since each  $s_n$  is concave and the collection  $\{s_n\}_{n \in \mathbb{N}}$  is pointwise uniformly bounded, it follows that the collection  $\{s_n\}_{n \in \mathbb{N}}$  is relatively compact in the compact-open topology of continuous functions. Let  $\{s_{n_k}\}_{k=1}^\infty$  be any locally uniformly convergent subsequence with limiting function  $s'$ . Since  $\overline{B}(y, \delta)$  is a compact set, it follows that

$$\begin{aligned} \lim_{k \rightarrow \infty} \frac{1}{n_k} \ln \int_{n_k \overline{B}(y, \delta)} dX e^{-\langle t_0, X \rangle} Z_{n_k}(X) &= \lim_{k \rightarrow \infty} \frac{1}{n_k} \ln \int_{\overline{B}(y, \delta)} dx n_k e^{n_k(s_{n_k}(x) - \langle t_0, x \rangle)} \\ &= \sup_{x \in \overline{B}(y, \delta)} \{s'(x) - \langle t_0, x \rangle\}. \end{aligned}$$

Then we have

$$\lim_{k \rightarrow \infty} \frac{1}{n_k} \ln P_{n_k}(\overline{B}(y, \delta)) = \sup_{x \in \overline{B}(y, \delta)} \{s'(x) - \langle t_0, x \rangle\} - f(t_0).$$

By combining this result with the large deviations principle, we deduce that

$$\sup_{x \in \overline{B}(y, \delta)} \{s'(x) - \langle t_0, x \rangle\} - f(t_0) = -\Lambda^*(\overline{B}(y, \delta)).$$

Now, since both functions inside the supremum and infimum respectively are continuous, letting  $\delta \rightarrow 0^+$ , we obtain

$$s'(y) - \langle \beta_0, y \rangle - f(t_0) = -\Lambda^*(y) \iff s'(y) = \inf_{t \in \mathbb{R}^d} \{ \langle y, t \rangle + f(t) \}.$$

Since  $s'$  was the locally uniform limit of an arbitrary convergent subsequence  $\{s_{n_k}\}_{k=1}^\infty$ , the above result implies that this holds for any such  $s'$ , and thus the limit of any convergent subsequence is the same from which it follows that

$$\lim_{n \rightarrow \infty} s_n(x) = \inf_{t \in \mathbb{R}^d} \{ \langle t, x \rangle + f(t) \},$$

for  $x \in \text{int}(\mathcal{D}(\Lambda^*)) = \mathcal{A}$ , and since the  $s_n$  are concave and pointwise uniformly bounded, this convergence is automatically locally uniform. □

Let us also give a quick proof of the following weak convergence result concerning large deviations principles.

**Theorem 4.3.8** *Let  $\{P_n\}_{n=1}^\infty$  be a sequence of probability measures on  $X$  satisfying a large deviations principle with rate function  $I$ .*

*It follows that*

$$L(\{P_n\}_{n=1}^\infty) \subset \{P \in \mathcal{P}(X) : \text{supp}(P) \subset I^{-1}\{0\}\},$$

where  $\mathcal{P}(X)$  is the space of Borel probability measures on  $X$ .

**Proof** Let us first show that  $I^{-1}\{0\}$  is non-empty and closed. Since  $I$  has compact level sets, it follows that  $I^{-1}[0, c]$  are compact for  $c > 0$ , but possibly empty. If they are not empty, then  $I^{-1}\{0\} = \bigcap_{n=1}^\infty I^{-1}[0, \frac{1}{n}]$ , and it follows directly that  $I^{-1}\{0\}$  is non-empty and compact. However, if  $I^{-1}[0, c]$  is empty for some  $c > 0$ , observe that

$$0 = \lim_{n \rightarrow \infty} I(I^{-1}[0, n]) = \lim_{n \rightarrow \infty} \inf_{x \in I^{-1}[0, n]} I(x) = \lim_{n \rightarrow \infty} \inf_{x \in I^{-1}(c, n]} I(x) > c > 0,$$

which is a contradiction, and thus  $I^{-1}[0, c]$  are non-empty for every  $c > 0$ , and subsequently  $c = 0$ . Note that the first line of the above proof by contradiction follows from the fact that  $\{P_n\}_{n=1}^\infty$  satisfies a large deviations principle.

Let  $y \notin I^{-1}\{0\}$  be such that  $\overline{B}(y, \delta)$  is disjoint from  $I^{-1}\{0\}$  for small enough  $\delta > 0$ . Note that

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \ln P_n(\overline{B}(y, \delta)) \leq -I(\overline{B}(y, \delta)) < 0.$$

The last strict inequality follows since by lower semi-continuity  $I$  attains its minimum on any non-empty compact set, and  $I$  is strictly positive on the set  $\overline{B}(y, \delta)$ . It follows that

$$P_n(B(y, \delta)) \leq P_n(\overline{B}(y, \delta)) \leq e^{n \sup_{k \geq n} \frac{1}{n} \ln P_k(\overline{B}(y, \delta))},$$

so that

$$\lim_{n \rightarrow \infty} P_n(B(y, \delta)) = 0.$$

Since the sequence of probability measures satisfies a large deviations principle, it is exponentially tight which implies that it is uniformly tight in the weak sense. Let  $\{P_{n_k}\}_{k=1}^\infty$  be any weakly convergent subsequence with limiting probability measure  $P$ . Let  $\overline{B}(y, \delta)$  be as before, by weak convergence, it follows that

$$P(\overline{B}(y, \delta)) \leq \liminf_{k \rightarrow \infty} P_{n_k}(B(y, \delta)) = 0.$$

Since  $y \notin I^{-1}\{0\}$  is arbitrary, it follows that

$$(I^{-1}\{0\})^c \subset (\text{supp}(P))^c \iff \text{supp}(P) \subset I^{-1}\{0\}.$$

□

For the purposes of this paper, the most important corollary is the case where  $I^{-1}\{0\}$  consists of a single point.

**Corollary 4.3.9** *Let  $\{P_n\}_{n=1}^\infty$  be a sequence of probability measures on  $X$  satisfying a large deviations principle with rate function  $I$  such that  $I(x^*) = 0$  for exactly one  $x^* \in X$ . It follows that*

$$\lim_{n \rightarrow \infty} P_n = \delta_{x^*}$$

weakly.

The proof of this statement is an application of the previous theorem in combination with Prokhorov’s theorem.

Another important corollary is the following result concerning the case where  $I^{-1}\{0\}$  consists of finitely many points.

**Corollary 4.3.10** *Let  $\{P_n\}_{n=1}^\infty$  be a sequence of probability measures on  $X$  satisfying a large deviations principle with rate function  $I$  such that the set  $M^* := I^{-1}\{0\}$  is finite. It follows that*

$$\int_X P_n(dx) f(x) = \sum_{x^* \in M^*} \frac{P_n(\overline{B}(x^*, \delta))}{P_n(A_\delta)} f(x^*) + o(1)$$

for any  $0 < \delta < \min_{x^*, y^* \in M^*} d(x^*, y^*)$ .

**Proof** Let  $\delta < \min_{x^*, y^* \in M^*} d(x^*, y^*)$ . We decompose  $X$  as follows

$$X = A_\delta \cup A_\delta^c,$$

where

$$A_\delta := \bigcup_{x^* \in M^*} \overline{B}(x^*, \delta).$$

Using this decomposition, we have

$$P_n = P_n(A_\delta) \sum_{x^* \in M^*} \frac{P_n(\overline{B}(x^*, \delta))}{P_n(A_\delta)} P_n|_{\overline{B}(x^*, \delta)} + P_n(A_\delta^c) P_n|_{A_\delta^c}.$$

Using the large deviations principle, it follows that

$$\lim_{n \rightarrow \infty} P_n(A_\delta) = 1,$$

and, by using the previous corollary, it follows that

$$\lim_{n \rightarrow \infty} P_n|_{\overline{B}(x^*, \delta)} = \delta_{x^*}$$

weakly, where  $x^* \in M^*$ . Using these limits together, we obtain

$$\lim_{n \rightarrow \infty} \left| \int_X P_n(dx) f(x) - \sum_{x^* \in M^*} \frac{P_n(\overline{B}(x^*, \delta))}{P_n(A_\delta)} f(x^*) \right| = 0.$$

□



### 4.4 Infinite-Volume Entropies and States

Next, we prove the regularity and boundedness of the finite-volume entropies.

**Proof of Lemma 3.1.4** From Eq. (4.1.3), we see that the microcanonical partition function is a homogeneous bivariate polynomial of degree  $n - 2$ . Let us introduce the change of coordinates  $z : \mathcal{A} \rightarrow (0, \infty)^2$  given by

$$z(M, N) := (x(M, N), y(M, N)) = \left( \frac{N + M}{2}, \frac{N - M}{2} \right).$$

It follows that  $Z_n(M, N) = \frac{1}{2} P_n(z(M, N))$ , where  $P_n : (0, \infty)^2 \rightarrow (0, \infty)$  is given by

$$P_n(x, y) := \sum_{k=1}^{n-1} \binom{n}{k} \frac{x^k}{(k-1)!} \frac{y^{n-k-1}}{(n-k-1)!}.$$

Using the properties of the binomial coefficient, we can manipulate  $P_n$  into the following form

$$P_n(x, y) = n(n-1) \sum_{k=0}^{n-2} \binom{n-2}{k} \frac{x^k}{(k+1)!} \frac{y^{n-2-k}}{(n-1-k)!}.$$

Let us denote the coefficients of the above manipulated polynomial by  $\{c_k\}_{k=0}^{n-2}$ . For  $k \in \mathbb{N}$ , using the simple relation

$$(k+1)!(k-1)! > k!,$$

it follows that

$$\frac{c_k^2}{\binom{n-2}{k}^2} > \frac{c_{k+1}}{\binom{n-2}{k+1}} \frac{c_{k-1}}{\binom{n-2}{k-1}}$$

for  $0 < k < n - 2$ . Using [16, Example 2.3], this implies that the sequence of coefficients  $\{c_k\}_{k=0}^{n-2}$  is ultra log-concave, which yields that  $P_n$  is Lorentzian, which shows that  $P_n$  is log-concave, see [16, Theorem 2.30] and the definition of completely log-concave polynomials due to [25]. Since  $Z_n$  is the composition of an invertible linear map, simple scaling by a factor of 2, and a log-concave polynomial it follows that  $Z_n$  is log-concave.

For boundedness, by Theorem 4.2.2, we have  $\mathcal{H}_{n-2}(v_n(m, \rho) || \eta_n(\beta, \mu)) \geq 0$ , from which it follows that

$$s_n(m, \rho) \leq \frac{n}{n-2} s_n(m, \rho) \leq f(\beta, \mu) + \beta m + \mu \rho,$$

which shows that the family of entropies is pointwise bounded above. As for a lower bound, it is enough to use the following trivial lower bound

$$Z_n(mn, \rho n) \geq \frac{1}{2} \frac{n!}{(n-1)!} \frac{\left(\frac{\rho n + mn}{2}\right)^{n-2}}{(n-2)!},$$

from which we obtain

$$\frac{1}{n} \ln Z_n(mn, \rho n) \geq \frac{1}{n} \ln \frac{1}{2} + \frac{n-2}{n} \ln \frac{\rho + m}{2} + \frac{1}{n} \ln \frac{n^{n-1}}{(n-2)!}.$$

It follows that

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \ln Z_n(mn, \rho n) \geq \ln \frac{\rho + m}{2} + 1,$$

as desired. □

We continue by consider the properties of the limiting entropy  $f(\beta, \mu)$ .

**Proof of Lemma 3.1.6** First, we observe that

$$f(\beta, \mu) = \ln \int_{-\infty}^{\infty} d\phi e^{-\beta\phi - \mu|\phi|}.$$

From this form, it is apparent that  $f$  is strictly convex on  $\mathcal{A}$  and thus is a proper convex function on  $\mathbb{R}^2$ . For lower semi-continuity, if  $(\beta, \mu) \in \mathbb{R}^2 \setminus \mathcal{A}$ , then  $f$  is lower semi-continuous for trivial reasons, in addition, since  $f$  is continuous on  $\mathcal{A}$ , it is also necessarily lower semi-continuous there. For the points in  $(\beta, \mu) \in \overline{\mathcal{A}}$ , it is clear that these points are of the form  $(\pm\mu', \mu')$  for  $\mu' \geq 0$ . It is easy to check that  $\lim_{(\beta, \mu) \rightarrow (\pm\mu', \mu')} f(\beta, \mu) = \infty$ , since  $f(\beta, \mu)$  is either equal to infinity, or it is increasing without bound for points inside  $\mathcal{A}$  approaching  $(\pm\mu', \mu')$ .

As for the other properties, the non-empty interior of the domain of finiteness of  $f$  is given by  $\mathcal{A}$ . The mapping  $f$  is differentiable in  $\mathcal{A}$ . For steepness, which is the third property of being essentially smooth, observe that

$$\|\nabla[f](\beta, \mu)\| = \frac{1}{\frac{1}{\mu+\beta} + \frac{1}{\mu-\beta}} \sqrt{\frac{2}{(\mu+\beta)^4} + \frac{2}{(\mu-\beta)^4}}.$$

Since all norms on  $\mathbb{R}^2$  are equivalent, it follows that there exists a constant  $C > 0$  such that

$$\left( \frac{1}{(\mu+\beta)^4} + \frac{1}{(\mu-\beta)^4} \right)^{\frac{1}{4}} \geq C \left( \frac{1}{\mu+\beta} + \frac{1}{\mu-\beta} \right).$$

Using this estimate, it follows that

$$\|\nabla[f](\beta, \mu)\| \geq \sqrt{2}C \left( \frac{1}{\mu+\beta} + \frac{1}{\mu-\beta} \right).$$

From this estimate it is now clear that if  $(\beta, \mu) \rightarrow (\pm\mu', \mu')$  for  $\mu' \geq 0$  for points inside  $\mathcal{A}$ , then clearly  $\lim_{(\beta, \mu) \rightarrow (\pm\mu', \mu')} \|\nabla[f](\beta, \mu)\| = \infty$ , which shows steepness.

In summary, we find that  $f$  is a proper convex lower semi-continuous function of Legendre type.

For the next few computational steps, it is useful to introduce the change of variables  $g : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  given by  $(\beta, \mu) \mapsto g(\beta, \mu) = (\mu + \beta, \mu - \beta)$  so that for  $(\beta, \mu) \in \mathcal{A}$ , we have

$$f(\beta, \mu) = \ln \left( \frac{1}{g_1(\beta, \mu)} + \frac{1}{g_2(\beta, \mu)} \right).$$

We can now equivalently consider the function  $f' : (0, \infty)^2 \rightarrow \mathbb{R}$  given by

$$f'(g_1, g_2) = \ln \left( \frac{1}{g_1} + \frac{1}{g_2} \right),$$

so that  $f \circ g^{-1} = f'$ . For the function  $f'$  it is easy to verify that

$$-\nabla[f'](g_1, g_2) = \left( \frac{g_2}{g_1(g_2 + g_1)}, \frac{g_1}{g_2(g_2 + g_1)} \right),$$

and the inverse map can be computed from

$$(0, \infty)^2 \ni (a, b) = -\nabla[f'](g_1, g_2)$$

$$\iff (g_1, g_2) = \left( \frac{1}{\sqrt{a}(\sqrt{a} + \sqrt{b})}, \frac{1}{\sqrt{b}(\sqrt{a} + \sqrt{b})} \right) = (-\nabla[f'])^{-1}(a, b).$$

This shows that  $(-\nabla[f'])(0, \infty)^2 = (0, \infty)^2$ . Finally, for  $(a, b) \in (0, \infty)^2$ , one can observe that

$$\inf_{(g_1, g_2) \in (0, \infty)^2} \{ag_1 + bg_2 + f'(g_1, g_2)\} = (-\nabla[f'])_1^{-1}(a, b)a + (-\nabla[f'])_2^{-1}(a, b)b$$

$$+ (f \circ (-\nabla[f']))(a, b)$$

$$= 1 + \ln \left( (\sqrt{a} + \sqrt{b})^2 \right).$$

To return to the function  $f$ , we have

$$(-\nabla[f])\mathcal{A} = (D[g])^T ((-\nabla[f'])g(\mathcal{A})) = (D[g])^T ((-\nabla[f'])(0, \infty)^2)$$

$$= (D[g])^T (0, \infty)^2 = \mathcal{A},$$

where  $D[g]$  is the derivative of the map  $g$ . We can also compute the following

$$\inf_{(\beta, \mu) \in \mathbb{R}^2} \{\beta m + \mu \rho + f(\beta, \mu)\} = \inf_{(\beta, \mu) \in \mathcal{A}} \{\beta m + \mu \rho + f(\beta, \mu)\}$$

$$= \inf_{(g_1, g_2) \in g(\mathcal{A}) = (0, \infty)^2} \{g_1^{-1}(g_1, g_2)m$$

$$+ g_2^{-1}(g_1, g_2)\rho + (f \circ g^{-1})(g_1, g_2)\}$$

$$= \inf_{(g_1, g_2) \in (0, \infty)^2} \left\{ \frac{g_1 - g_2}{2} m + \frac{g_1 + g_2}{2} \rho + f'(g_1, g_2) \right\}$$

$$= \inf_{(g_1, g_2) \in (0, \infty)^2} \left\{ \frac{\rho + m}{2} g_1 + \frac{\rho - m}{2} g_2 + f'(g_1, g_2) \right\}$$

$$= 1 + \ln \left( \left( \sqrt{\frac{\rho + m}{2}} + \sqrt{\frac{\rho - m}{2}} \right)^2 \right).$$

To finish, note that we can simply compute the gradient

$$-\nabla[f](\beta, \mu) = \left( -\frac{2\beta}{\mu^2 - \beta^2}, \frac{\mu^2 + \beta^2}{\mu(\mu^2 - \beta^2)} \right),$$

but its inverse map is simpler to solve from the composite function  $f'$ . Doing so, we obtain

$$\mathcal{A} \ni (m, \rho) = -\nabla[f](\beta, \mu) \iff (\beta, \mu) = \left( -\frac{\rho}{m} \frac{1}{\sqrt{\rho^2 - m^2}} + \frac{1}{m}, \frac{1}{\sqrt{\rho^2 - m^2}} \right)$$

$$= (-\nabla[f])^{-1}(m, \rho).$$

Compiling together all of these results, we find that  $f$  is a proper convex lower semi-continuous function of Legendre type which satisfies  $(-\nabla[f])\mathcal{A} = \mathcal{A}$ , and, for  $(m, \rho) \in \mathcal{A}$ , we have

$$\inf_{(\beta, \mu) \in \mathbb{R}^2} \{\beta m + \mu \rho + f(\beta, \mu)\} = \inf_{(\beta, \mu) \in \mathcal{A}} \{\beta m + \mu \rho + f(\beta, \mu)\}$$

$$\begin{aligned}
 &= \beta(m, \rho)m + \mu(m, \rho)\rho + f(\beta(m, \rho), \mu(m, \rho)) \\
 &= 1 + \ln \left( \left( \sqrt{\frac{\rho+m}{2}} + \sqrt{\frac{\rho-m}{2}} \right)^2 \right),
 \end{aligned}$$

where

$$(\beta(m, \rho), \mu(m, \rho)) = (-\nabla[f])^{-1}(m, \rho) = \left( -\frac{\rho}{m} \frac{1}{\sqrt{\rho^2 - m^2}} + \frac{1}{m}, \frac{1}{\sqrt{\rho^2 - m^2}} \right).$$

□

We begin with the proof of the half-constrained ensemble limiting entropy.

**Proof of Lemma 3.2.1** Fix  $\beta \in \mathbb{R}$ , and consider the mapping  $Q_n(g^\beta, \cdot) : (0, \infty) \rightarrow \mathbb{R}$  given by

$$Q_n(g^\beta, \rho) := \int_{\mathbb{R}^n} d\phi e^{-\beta M_n(\phi)} \delta(N_n(\phi) - \rho n),$$

which, like Eq. (3.1.1), is to be understood as

$$Q_n(g^\beta, \rho) = e^{-\beta\rho n} Z_n(\rho n, \rho n) + e^{\beta\rho n} Z_n(-\rho n, \rho n) + \int_{-\rho}^{\rho} dm n e^{-\beta mn} Z_n(mn, \rho n).$$

By direct computation, using Eq. (4.1.3), it follows that

$$\begin{aligned}
 \lim_{n \rightarrow \infty} \frac{1}{n} \ln(e^{-\beta\rho n} Z_n(\rho n, \rho n)) &= -\beta\rho + \ln \rho + 1, \quad \lim_{n \rightarrow \infty} \frac{1}{n} \ln(e^{\beta\rho n} Z_n(-\rho n, \rho n)) \\
 &= \beta\rho + \ln \rho + 1.
 \end{aligned}$$

As for the mapping

$$\rho \mapsto \int_{-\rho}^{\rho} dm n e^{-\beta mn} Z_n(mn, \rho n) = \int_{\mathbb{R}} dm n \mathbb{1}(|m| < \rho) e^{-\beta mn} Z_n(mn, \rho n),$$

it is enough to notice that the individual mappings in the integrand

$$\mathbb{R}^2 \ni (m, \rho) \mapsto (\mathbb{1}(|m| < \rho), e^{-\beta mn}, Z_n(mn, \rho n))$$

are log-concave functions. To be more precise, the indicator function is the indicator of a convex set and is thus log-concave, the exponential function is trivially log-concave by direct computation, and, finally, the microcanonical partition function, which is to be understood as the microcanonical partition function on  $\mathcal{A}$  extended beyond this set by setting its value to 0, is log-concave by Lemma 3.1.4. It follows that that the mapping

$$\rho \mapsto \int_{-\rho}^{\rho} dm n e^{-\beta mn} Z_n(mn, \rho n)$$

is log-concave by the Prékopa–Leindler inequality or Prékopa’s theorem, see [26, Section 9], since it is the marginal of a log-concave function.

For pointwise uniform boundedness, we begin by observing that

$$e^{-|\beta|\rho n} \int_{-\rho}^{\rho} dm n Z_n(mn, \rho n) \leq \int_{-\rho}^{\rho} dm n e^{-\beta mn} Z_n(mn, \rho n) \leq e^{|\beta|\rho n} \int_{-\rho}^{\rho} dm n Z_n(mn, \rho n)$$

and

$$\int_{-\rho}^{\rho} dm n Z_n(mn, \rho n) = \rho^{n-1} n^{n-1} \int_{-1}^1 dm Z_n(m, 1).$$

We will use the beta function  $B(z_1, z_2)$  given by

$$B(z_1, z_2) := \int_0^1 dt t^{z_1-1} (1-t)^{z_2-1}$$

for  $\text{Re}(z_1), \text{Re}(z_2) > 0$ . By a change of variables, one can see that

$$B(z_1, z_2) = \frac{1}{2} \int_{-1}^1 dt \left(\frac{1+t}{2}\right)^{z_1-1} \left(\frac{1-t}{2}\right)^{z_2-1}.$$

For integer values, we have the following identity

$$B(m, n) = \frac{(m-1)!(n-1)!}{(m+n-1)!}$$

from which it follows that

$$\int_{-1}^1 dm Z_n(m, 1) = \sum_{k=1}^{n-1} \binom{n}{k} \frac{B(k, n-k)}{(k-1)!(n-k-1)!} = \frac{1}{(n-1)!} \sum_{k=1}^{n-1} \binom{n}{k} = \frac{2^n - 2}{(n-1)!}.$$

In summary, we have

$$e^{-|\beta|\rho n} \rho^{n-1} n^{n-1} \frac{2^n - 2}{(n-1)!} \leq \int_{-\rho}^{\rho} dm ne^{-\beta mn} Z_n(mn, \rho n) \leq e^{|\beta|\rho n} \rho^{n-1} n^{n-1} \frac{2^n - 2}{(n-1)!}.$$

Computing the limits, it follows that

$$-\infty < \liminf_{n \rightarrow \infty} \frac{1}{n} \ln \int_{-\rho}^{\rho} dm ne^{-\beta mn} Z_n(mn, \rho n) \leq \limsup_{n \rightarrow \infty} \frac{1}{n} \ln \int_{-\rho}^{\rho} dm ne^{-\beta mn} Z_n(mn, \rho n) < \infty,$$

from which the uniform pointwise boundedness follows.

For  $\mu > |\beta|$ , we can directly compute that

$$\int_0^{\infty} d\rho ne^{-\mu\rho n} \int_{-\rho}^{\rho} dm ne^{-\beta mn} Z_n(mn, \rho n) = \left(\frac{1}{\mu + \beta} + \frac{1}{\mu - \beta}\right)^n - \left(\frac{1}{\mu + \beta}\right)^n - \left(\frac{1}{\mu - \beta}\right)^n.$$

For any other value of  $\mu$ , it is clear that the above integral is infinite. It follows that the limit and subsequent mapping given by

$$\mu \mapsto \lim_{n \rightarrow \infty} \frac{1}{n} \ln \int_0^{\infty} d\rho ne^{-\mu\rho n} \int_{-\rho}^{\rho} dm ne^{-\beta mn} Z_n(mn, \rho n) = f(\beta, \mu),$$

exists and has a domain of finiteness given by the half-infinite interval  $(|\beta|, \infty)$ . By using the properties of the full map  $(\beta, \mu) \mapsto f(\beta, \mu)$ , already verified and computed in Lemma 3.1.5, one can verify that the mapping  $\mu \mapsto f(\beta, \mu)$  for fixed  $\beta$  is a proper convex lower semi-continuous function of Legendre type that satisfies  $-D[f(\beta, \cdot)] = (0, \infty)$ . By Theorem 3.1.3, for any  $\rho > 0$ , it follows that

$$\lim_{n \rightarrow \infty} \frac{1}{n} \ln \int_{-\rho}^{\rho} dm ne^{-\beta mn} Z_n(mn, \rho n) = \inf_{\mu > |\beta|} \{\mu\rho + f(\beta, \mu)\}.$$

To continue, by Lemma 3.1.5, we have

$$f(\beta, \mu) = \inf_{(m, \rho) \in \mathcal{A}} \{\beta m + \mu\rho - s(m, \rho)\} = \inf_{\rho > 0} \left\{ \mu\rho + \inf_{|m| < \rho} \{\beta m - s(m, \rho)\} \right\},$$

so that

$$\inf_{\mu > |\beta|} \{\mu\rho + f(\beta, \mu)\} = - \inf_{|m| < \rho} \{\beta m - s(m, \rho)\} = \sup_{|m| < 1} \{s(m, \rho) - \beta m\}.$$

For the rate function, the scaled logarithmic moment generating function  $\Lambda : \mathbb{R} \rightarrow [-\infty, \infty]$  of a sequence of random variables with distributions given by  $\{\kappa_n^\beta\}_{n \in \mathbb{N}}$  is given by

$$\begin{aligned} \Lambda(t) &:= \lim_{n \rightarrow \infty} \frac{1}{n} \ln \frac{Q_n(\beta - t)}{Q_n(\beta)} = \sup_{|m| < 1} \{s(m, 1) - (\beta - t)m\} - \sup_{|m| < 1} \{s(m, 1) - \beta m\} \\ &= \sup_{|m| < 1} \{tm - (-(s(m, 1) - \beta m))\} - \sup_{|m| < 1} \{s(m, 1) - \beta m\}. \end{aligned}$$

We can identify the first term on the last line as the convex conjugate of the restriction of a proper convex lower semi-continuous function of Legendre type with an interior of the domain of finiteness given by  $(-1, 1)$ . From the form of the function  $s(m, 1)$ , for  $m \in (-1, 1)$ , we immediately see that

$$\lim_{m \rightarrow \pm 1^\mp} s(m, 1) = 1.$$

Defining  $s(\pm 1, 1) = 1$  yields a continuous extension of  $s(m, 1)$  from  $(-1, 1)$  to  $[-1, 1]$ , and we will consider it so from now on. The extended mapping given by

$$\mathbb{R} \ni m \mapsto \begin{cases} s(m, 1), & m \in [-1, 1], \\ -\infty, & m \notin [-1, 1], \end{cases}$$

is upper semi-continuous, and we will consider this the redefinition of  $s(m, 1)$  to be understood now as not necessarily finite function on  $\mathbb{R}$ . Compiling all of this together, it follows that the mapping  $\mathbb{R} \ni m \mapsto -(s(m, 1) - \beta m)$  defines a proper convex lower semi-continuous function of Legendre type, and thus the convex conjugate is involutive from which it follows that

$$\Lambda^*(m) = \sup_{|m| < 1} \{s(m, 1) - \beta m\} - (s(m, 1) - \beta m),$$

which is the rate function of  $\{\kappa_n^\beta\}_{n \in \mathbb{N}}$ . □

We finish by giving the proof of the limit point result.

**Proof of Lemma 3.2.3** Using Theorem 4.3.8, let  $\{\kappa_{n_k}^g\}_{k \in \mathbb{N}}$  be a weakly convergent subsequence with a limit  $\kappa$ . Since  $M^*(\psi^g)$  is a compact subset of  $(-1, 1)$ , it follows that there exists  $a := \min M^*(\psi^g)$  and  $b := \max M^*(\psi^g)$ . There exists  $\delta > 0$  such that  $\text{supp}(\kappa) \subset M^*(\psi^g) \subset [a - \delta, b + \delta] \subset (-1, 1)$ . Since  $\text{supp}(\kappa) \subset [a - \delta, b + \delta]$ , we deduce that  $\kappa([a - \delta, b + \delta]) = 1$ , and, since  $\partial[a - \delta, b + \delta] \cap \text{supp}(\mu) \subset \{a - \delta, b + \delta\} \cap M^*(\psi^g) = \emptyset$ , we see that  $\kappa(\partial[a - \delta, b + \delta]) = 0$ . It follows that  $[a - \delta, b + \delta] \subset (-1, 1)$  is a continuity set of  $\kappa$ , and we can apply Lemma 3.1.1 along this subsequence with Corollary 3.1.7 to obtain the result. □

### 4.5 Asymptotics of the Weights

We first establish the Laplace-type representation of the microcanonical partition function.

**Proof of Lemma 3.3.1** The microcanonical partition function can be written as

$$Z_n(mn, \rho n) = \frac{2n^{n-2}n!}{(\rho^2 - m^2)n^2} \sum_{k=1}^{n-1} \frac{\left(\frac{\rho+m}{2}\right)^k}{(k-1)!k!} \frac{\left(\frac{\rho-m}{2}\right)^{n-k}}{(n-k-1)!(n-k)!},$$

which one can recognize as the convolution of two sequences with some factors in front. We consider the generating function  $G : \mathbb{C} \rightarrow \mathbb{C}$  given by

$$\begin{aligned} G(z) &:= \sum_{n=2}^{\infty} \frac{(\rho^2 - m^2)n^2}{2n^{n-2}n!} Z_n(mn, \rho n) \left(\frac{z^2}{4}\right)^n \\ &= \left( \sum_{n=2}^{\infty} \frac{\left(\frac{1}{4} \left(\sqrt{\frac{\rho+m}{2}} z\right)^2\right)^n}{n!(n-1)!} \right) \left( \sum_{n=2}^{\infty} \frac{\left(\frac{1}{4} \left(\sqrt{\frac{\rho-m}{2}} z\right)^2\right)^n}{n!(n-1)!} \right). \end{aligned}$$

One can verify that the convolution yields a Cauchy product, and that the power series on the right define entire functions with absolutely convergent power series. We have the standard relation between the derivatives of  $G$  and its power series coefficients

$$\frac{G^{(2n)}(0)}{(2n)!} = \frac{(\rho^2 - m^2)n^2}{2n^{n-2}n!4^n} Z_n(mn, \rho n) \iff Z_n(mn, \rho n) = \frac{2^{2n+1}n^{n-2}n!}{(\rho^2 - m^2)n^2} \frac{G^{(2n)}(0)}{(2n)!}.$$

Next, using the modified Bessel function of the first kind  $I_\nu(z)$  given by

$$I_\nu(z) := \left(\frac{1}{2}z\right)^\nu \sum_{n=0}^{\infty} \frac{\left(\frac{z^2}{4}\right)^n}{n!\Gamma(\nu + n + 1)},$$

where  $\nu \in \mathbb{Z}$ , and we have

$$G(z) = \frac{1}{4} \sqrt{\frac{\rho^2 - m^2}{4}} z^2 I_{-1}\left(\sqrt{\frac{\rho+m}{2}} z\right) I_{-1}\left(\sqrt{\frac{\rho-m}{2}} z\right).$$

Using the integral representation, see [27, Chapter 9], given by

$$I_\nu(z) := \frac{1}{\pi} \int_0^\pi d\theta \cos(\nu\theta) e^{z \cos \theta},$$

we see that

$$G(z) = \frac{1}{4} \sqrt{\frac{\rho^2 - m^2}{4}} z^2 \frac{1}{\pi^2} \int_0^\pi d\theta_1 \int_0^\pi d\theta_2 \cos \theta_1 \cos \theta_2 e^{z \left(\sqrt{\frac{\rho+m}{2}} \cos \theta_1 + \sqrt{\frac{\rho-m}{2}} \cos \theta_2\right)}.$$

Taking derivatives, using the general Leibniz rule, we obtain

$$\begin{aligned} G^{(2n)}(0) &= \frac{1}{2} \sqrt{\frac{\rho^2 - m^2}{4}} \binom{2n}{2} \frac{1}{\pi^2} \int_0^\pi d\theta_1 \int_0^\pi d\theta_2 \\ &\quad \times \cos \theta_1 \cos \theta_2 \left(\sqrt{\frac{\rho+m}{2}} \cos \theta_1 + \sqrt{\frac{\rho-m}{2}} \cos \theta_2\right)^{2n-2}, \end{aligned}$$

from which it follows that

$$Z_n(mn, \rho n) = \frac{2^{2n-1} n^{n-2} n!}{(2n)! \sqrt{\rho^2 - m^2} n^2} \binom{2n}{2} \frac{1}{\pi^2} \int_0^\pi d\theta_1 \int_0^\pi d\theta_2 \\ \times \cos \theta_1 \cos \theta_2 \left( \sqrt{\frac{\rho + m}{2}} \cos \theta_1 + \sqrt{\frac{\rho - m}{2}} \cos \theta_2 \right)^{2n-2}.$$

By using the given form of the overloaded  $s$  function and simplifying, we obtain the desired representation. □

We present the proof of the local asymptotics of the overloaded  $\psi^s$  function.

**Proof** By computing the critical points of the overloaded  $\psi^s$  function, we see that there is precisely one critical point in the given set in the assumptions, and it is given by  $(m^*, 0, 0)$ . For this particular critical point, it is easy to see that any odd partial derivative with respect to either  $\theta_1$  or  $\theta_2$  is vanishing.

By developing  $\psi^s$  to second order in  $(\theta_1, \theta_2)$ , and  $(2k)$ :th order in  $m$ , it follows that

$$\psi^s(m^* + m, \theta_1, \theta_2) = \psi^s(m^*) + \frac{1}{2} \partial_2^2 [\psi^s](m^*, 0, 0) \theta_1^2 + \frac{1}{2} \partial_3^2 [\psi^s](m^*, 0, 0) \theta_2^2 \\ + \frac{1}{(2k)!} \partial^{2k} [\psi^s](m^*) m^{2k} + \sum_{|\alpha|=3, \alpha_1 \notin \{2,3\}} R_\alpha(m, \theta_1, \theta_2) (m, \theta_1, \theta_2)^\alpha \\ + R_{(2k+1,0,0)}(m, \theta_1, \theta_2) m^{2k+1},$$

where

$$R_\alpha(m, \theta_1, \theta_2) = \frac{|\alpha|}{\alpha!} \int_0^1 dt (1-t)^{|\alpha|-1} \partial_\alpha [\psi^s]((m^*, 0, 0) + t(m, \theta_1, \theta_2)).$$

□

We can now prove the full Laplace method for the mixture measures.

**Proof** Let us first remark that in the following proof, we will frequently use the statement for small enough  $\delta > 0$  something holds. In the context of this proof, we repeat this to imply that there is a series of finite choice of  $\delta > 0$  small enough such that all the conditions required will hold. In reality this proof should be worked through “backwards” so that the choice of  $\delta > 0$  is clear.

We begin by noting that

$$\frac{(2n)! n^2 \pi^2}{2^{2n-1} n^{n-2} n! \binom{2n}{2} e^{-(n-1)}} \int_{m^*-\delta}^{m^*+\delta} dm e^{n(g(m)+s_n(m,1))} \\ = \int_{m^*-\delta}^{m^*+\delta} dm \int_0^\pi d\theta_1 \int_0^\pi d\theta_2 \frac{\cos \theta_1 \cos \theta_2 e^{g(m)}}{\sqrt{1-m^2}} e^{(n-1)\psi^s(m, \theta_1, \theta_2)}$$

and by using the symmetries of the trigonometric functions, it follows that

$$\int_{m^*-\delta}^{m^*+\delta} dm \int_0^\pi d\theta_1 \int_0^\pi d\theta_2 \frac{e^{g(m)} \cos \theta_1 \cos \theta_2}{\sqrt{1-m^2}} e^{(n-1)\psi^s(m, \theta_1, \theta_2)} \\ = \int_{m^*-\delta}^{m^*+\delta} dm \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} d\theta_1 \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} d\theta_2 \frac{e^{g(m)} \cos \theta_1 \cos \theta_2}{\sqrt{1-m^2}} e^{(n-1)\psi^s(m, \theta_1, \theta_2)}$$



$$\begin{aligned}
 & - \int_{m^*-\delta}^{m^*+\delta} dm \int_0^{\frac{\pi}{2}} d\theta_1 \int_0^{\frac{\pi}{2}} d\theta_2 \frac{e^{g(m)} \sin \theta_1 \cos \theta_2}{\sqrt{1-m^2}} e^{(n-1)\psi^g(m, \theta_1 + \frac{\pi}{2}, \theta_2)} \\
 & - \int_{m^*-\delta}^{m^*+\delta} dm \int_0^{\frac{\pi}{2}} d\theta_1 \int_0^{\frac{\pi}{2}} d\theta_2 \frac{e^{g(m)} \cos \theta_1 \sin \theta_2}{\sqrt{1-m^2}} e^{(n-1)\psi^g(m, \theta_1, \theta_2 + \frac{\pi}{2})}.
 \end{aligned}$$

We want to show that the first integral on the second line of this manipulation is exponentially dominant. To save space, denote the integrals as follows

$$\begin{aligned}
 I_1(n) & := \int_{m^*-\delta}^{m^*+\delta} dm \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} d\theta_1 \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} d\theta_2 \frac{e^{g(m)} \cos \theta_1 \cos \theta_2}{\sqrt{1-m^2}} e^{(n-1)\psi^g(m, \theta_1, \theta_2)}, \\
 I_2(n) & := \int_{m^*-\delta}^{m^*+\delta} dm \int_0^{\frac{\pi}{2}} d\theta_1 \int_0^{\frac{\pi}{2}} d\theta_2 \frac{e^{g(m)} \sin \theta_1 \cos \theta_2}{\sqrt{1-m^2}} e^{(n-1)\psi^g(m, \theta_1 + \frac{\pi}{2}, \theta_2)}, \\
 I_3(n) & := \int_{m^*-\delta}^{m^*+\delta} dm \int_0^{\frac{\pi}{2}} d\theta_1 \int_0^{\frac{\pi}{2}} d\theta_2 \frac{e^{g(m)} \cos \theta_1 \sin \theta_2}{\sqrt{1-m^2}} e^{(n-1)\psi^g(m, \theta_1, \theta_2 + \frac{\pi}{2})}.
 \end{aligned}$$

For the terms  $I_2$  and  $I_3$ , observe that

$$\left| \sqrt{\frac{1+m}{2}} \sin \alpha - \sqrt{\frac{1-m}{2}} \cos \beta \right| \leq \max \left\{ \sqrt{\frac{1+m}{2}}, \sqrt{\frac{1-m}{2}} \right\} < \sqrt{\frac{1+m}{2}} + \sqrt{\frac{1-m}{2}}$$

for any  $\alpha, \beta \in [0, \frac{\pi}{2}]$  and  $m \in (m^* - \delta, m^* + \delta)$ . Using this property, one can check that

$$\begin{aligned}
 M_2(\delta) & := \max_{(m, \theta_1, \theta_2) \in (m^*-\delta, m^*+\delta) \times [0, \frac{\pi}{2}] \times [0, \frac{\pi}{2}]} \psi^g \left( m, \theta_1 + \frac{\pi}{2}, \theta_2 \right) \\
 & \leq \max_{m \in (m^*-\delta, m^*+\delta)} \left\{ g(m) + 1 + \ln \left( \left( \max \left\{ \sqrt{\frac{1+m}{2}}, \sqrt{\frac{1-m}{2}} \right\} \right)^2 \right) \right\}.
 \end{aligned}$$

By continuity of the function inside the maximum, one can check that

$$\lim_{\delta' \rightarrow 0^+} M_2(\delta') < M_1(\delta) := \max_{(m, \theta_1, \theta_2) \in (m^*-\delta, m^*+\delta) \times [-\frac{\pi}{2}, \frac{\pi}{2}] \times [-\frac{\pi}{2}, \frac{\pi}{2}]} \psi^g(m, \theta_1, \theta_2) = \psi^g(m^*),$$

from which it follows that for small enough  $\delta > 0$ , we have  $M_2(\delta) < M_1(\delta)$ . One can verify in the same way that

$$M_3(\delta) := \max_{(m, \theta_1, \theta_2) \in (m^*-\delta, m^*+\delta) \times [0, \frac{\pi}{2}] \times [0, \frac{\pi}{2}]} \psi^g \left( m, \theta_1, \theta_2 + \frac{\pi}{2} \right) < M_1(\delta)$$

for small enough  $\delta > 0$ . For such  $\delta$ , it follows that

$$\lim_{n \rightarrow \infty} \frac{1}{n} \ln I_{2/3}(n) = M_{2/3}(\delta) < M_1(\delta) = \lim_{n \rightarrow \infty} \frac{1}{n} \ln I_1(n),$$

which shows that  $I_1(n)$  exponentially dominates  $I_{2/3}(n)$ .

To continue, we have

$$\begin{aligned}
 \frac{n^{\frac{1}{2k}+1} (I_1(n) - I_2(n) - I_3(n))}{e^{nM_1}} & = \frac{n^{\frac{1}{2k}+1} I_1(n)}{e^{nM_1}} - n^{\frac{1}{2k}+1} e^{n \left( \frac{1}{n} \ln I_2(n) - M_1 \right)} \\
 & \quad - n^{\frac{1}{2k}+1} e^{n \left( \frac{1}{n} \ln I_3(n) - M_1 \right)}.
 \end{aligned}$$

It is now clear that in the limit the terms on the right of the  $I_1(n)$  term vanish since they are exponentially small. As for the limit of the integral  $I_1(n)$ , it is solved by a routine application

of Laplace’s method using the asymptotics developed in Lemma 3.3.2. First, however, we must split the integral  $I_1(n)$  with respect to the angular variables. Denote

$$f(m, \theta_1, \theta_2) := \frac{e^{g(m)} \cos \theta_1 \cos \theta_2}{\sqrt{1 - m^2}}.$$

Since  $\psi^g$  attains its unique maximum at  $(m^*, 0, 0)$ , it follows that

$$\begin{aligned} & \lim_{n \rightarrow \infty} \frac{1}{n} \ln \int_{m^* - \delta}^{m^* + \delta} dm \int_{([- \delta, \delta] \times [- \delta, \delta])^c} d\theta_1 d\theta_2 f(m, \theta_1, \theta_2) e^{(n-1)\psi^g(m, \theta_1, \theta_2)} \\ &= \sup_{m \in [m^* - \delta, m^* + \delta] \times ([- \delta, \delta] \times [- \delta, \delta])^c} \psi^g(m, \theta_1, \theta_2) < M_1. \end{aligned}$$

If we denote

$$I_{1,\delta}(n) := \int_{m^* - \delta}^{m^* + \delta} dm \int_{-\delta}^{\delta} d\theta_1 \int_{-\delta}^{\delta} d\theta_2 f(m, \theta_1, \theta_2) e^{(n-1)\psi^g(m, \theta_1, \theta_2)},$$

we have

$$\frac{n^{\frac{1}{2k}+1} I_1(n)}{e^{nM_1}} = \frac{n^{\frac{1}{2k}+1} I_{1,\delta}(n)}{e^{nM_1}} + n^{\frac{1}{2k}+1} e^{n\left(\frac{1}{n} \ln(I_1(n) - I_{1,\delta}(n)) - M_1\right)}.$$

Again, since the right hand side contains exponentially decreasing terms, the asymptotics will be determined by the first term on the right. Finally, by changing variables, observe that

$$\begin{aligned} & \frac{n^{\frac{1}{2k}+1} I_{1,\delta}(n)}{e^{(n-1)M_1}} \\ &= \int_{-\delta n^{\frac{1}{2k}}}^{\delta n^{\frac{1}{2k}}} dm \int_{-\frac{\pi}{2} n^{\frac{1}{2}}}^{\frac{\pi}{2} n^{\frac{1}{2}}} d\theta_1 \int_{-\frac{\pi}{2} n^{\frac{1}{2}}}^{\frac{\pi}{2} n^{\frac{1}{2}}} d\theta_2 f\left(m^* + \frac{m}{n^{\frac{1}{2k}}}, \frac{\theta_1}{n^{\frac{1}{2}}}, \frac{\theta_2}{n^{\frac{1}{2}}}\right) e^{(n-1)\left(\psi^g\left(m^* + \frac{m}{n^{\frac{1}{2k}}}, \frac{\theta_1}{n^{\frac{1}{2}}}, \frac{\theta_2}{n^{\frac{1}{2}}}\right) - \psi^g(m^*)\right)}. \end{aligned}$$

If one looks at the remainder term displayed in Lemma 3.3.2, one finds that

$$\begin{aligned} & \left| \sum_{|\alpha|=3, \alpha_1 \notin \{2,3\}} R_\alpha(m, \theta_1, \theta_2) (m, \theta_1, \theta_2)^\alpha \right| \\ & \leq \max_{(m, \theta_1, \theta_2) \in [-\delta, \delta]^3, |\alpha|=3, \alpha_1 \notin \{2,3\}} |R_\alpha(m, \theta_1, \theta_2)| \sum_{|\alpha|=3, \alpha_1 \notin \{2,3\}} |(m, \theta_1, \theta_2)^\alpha| \\ & \leq \max_{(m, \theta_1, \theta_2) \in [-\delta, \delta]^3, |\alpha|=3, \alpha_1 \notin \{2,3\}} |R_\alpha(m, \theta_1, \theta_2)| (A|\theta_1|^3 + B\theta_1^2|\theta_2| \\ & \quad + C|\theta_1|\theta_2^2 + D|\theta_2|^3 + E|m||\theta_1||\theta_2|) \\ & \leq \left( \delta F \max_{(m, \theta_1, \theta_2) \in [-\delta, \delta]^3, |\alpha|=3, \alpha_1 \notin \{2,3\}} |R_\alpha(m, \theta_1, \theta_2)| \right) (\theta_1^2 + \theta_2^2), \end{aligned}$$

and

$$\left| R_{(2k+1,0,0)}(m, \theta_1, \theta_2) m^{2k+1} \right| \leq \left( \delta \max_{(m, \theta_1, \theta_2) \in [-\delta, \delta]^3} |R_{(2k+1,0,0)}(m, \theta_1, \theta_2)| \right) m^{2k},$$

where  $A, B, C, D, E, F > 0$  are all positive constants. For  $\delta$  satisfying

$$\delta F \max_{(m, \theta_1, \theta_2) \in [-\delta, \delta]^3, |\alpha|=3, \alpha_1 \notin \{2,3\}} |R_\alpha(m, \theta_1, \theta_2)|$$

$$< \max \left\{ -\frac{1}{2} \partial_2^2 [\psi^g](m^*, 0, 0)m, -\frac{1}{2} \partial_3^2 [\psi^g](m^*, 0, 0) \right\},$$

and

$$\delta \max_{(m, \theta_1, \theta_2) \in [-\delta, \delta]^3} |R_{(2k+1, 0, 0)}(m, \theta_1, \theta_2)| \leq -\frac{1}{(2k)!} \partial^{2k} [\psi^g](m^*).$$

Ultimately, for  $\delta > 0$  chosen small enough so as to satisfy the finite number of conditions given previously, using the error bounds above, by dominated convergence, it follows that

$$\begin{aligned} & \lim_{n \rightarrow \infty} \int_{-\delta n^{\frac{1}{2k}}}^{\delta n^{\frac{1}{2k}}} dm \int_{-\frac{\pi}{2} n^{\frac{1}{2}}}^{\frac{\pi}{2} n^{\frac{1}{2}}} d\theta_1 \int_{-\frac{\pi}{2} n^{\frac{1}{2}}}^{\frac{\pi}{2} n^{\frac{1}{2}}} d\theta_2 f \left( m^* + \frac{m}{n^{\frac{1}{2k}}}, \frac{\theta_1}{n^{\frac{1}{2}}}, \frac{\theta_2}{n^{\frac{1}{2}}} \right) e^{(n-1) \left( \psi^g \left( m^* + \frac{m}{n^{\frac{1}{2k}}}, \frac{\theta_1}{n^{\frac{1}{2}}}, \frac{\theta_2}{n^{\frac{1}{2}}} \right) - \psi^g(m^*) \right)} \\ &= f(m^*, 0, 0) \int_{\mathbb{R}^3} d\theta_1 d\theta_2 dm e^{\frac{1}{2} \partial_2^2 [\psi^g](m^*, 0, 0) \theta_1^2 + \frac{1}{2} \partial_3^2 [\psi^g](m^*, 0, 0) \theta_2^2 + \frac{1}{(2k)!} \partial^{2k} [\psi^g](m^*) m^{2k}}. \end{aligned}$$

Combining all of these results together, it follows that

$$\begin{aligned} & \lim_{n \rightarrow \infty} \frac{n^{\frac{1}{2k}+1} \int_{m^*-\delta}^{m^*+\delta} dm e^{n(g(m)+s_n(m,1))}}{e^{n\psi^g(m^*)}} \frac{(2n)! n^2 \pi^2}{2^{2n-1} n^{n-2} n! \binom{2n}{2} e^{-(n-1)}} \\ &= \frac{e^{g(m^*)}}{e^{\psi^g(m^*)} \sqrt{1-m^{*2}}} \int_{\mathbb{R}^3} d\theta_1 d\theta_2 dm e^{\frac{1}{2} \partial_2^2 [\psi^g](m^*, 0, 0) \theta_1^2 + \frac{1}{2} \partial_3^2 [\psi^g](m^*, 0, 0) \theta_2^2 + \frac{1}{(2k)!} \partial^{2k} [\psi^g](m^*) m^{2k}}. \end{aligned}$$

□

**Acknowledgements** I thank my advisor Jani Lukkarinen for his support and encouragement during this project. For their technical comments and helpful discussions, I thank my colleagues Brecht Donvil, and Gerardo Barrera Vargas. I also thank Gerardo Barrera Vargas for his diligent reading and suggested corrections to this manuscript. The research has been supported by the Academy of Finland, via an Academy project (Project No. 339228) and the Finnish *centre of excellence in Randomness and Structures* (Project No. 346306).

**Funding** Open Access funding provided by University of Helsinki (including Helsinki University Central Hospital). This work was supported by the Academy of Finland, via an Academy project (Project No. 339228) and the Finnish *centre of excellence in Randomness and Structures* (Project No. 346306).

**Data Availability** Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

### Declarations

**Conflict of interest** The author has no conflict of interest to declare that are relevant to the content of this article.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

1. Kastner, M., Schnetz, O.: On the mean-field spherical model. *J. Stat. Phys.* **122**(6), 1195–1214 (2006)
2. Ellis, R.S., Newman, C.M.: The statistics of Curie-Weiss models. *J. Stat. Phys.* **19**(2), 149–161 (1978)
3. Ellis, R.S.: *Entropy, Large Deviations, and Statistical Mechanics*. Springer, Berlin/Heidelberg (2006)
4. Friedli, S., Velenik, Y.: *Statistical Mechanics of Lattice Systems*. Cambridge University Press, Cambridge (2017)
5. Eisele, T., Ellis, R.S.: Multiple phase transitions in the generalized Curie-Weiss model. *J. Stat. Phys.* **52**(1–2), 161–202 (1988)
6. Amaro de Matos, J.M.G., Fernando Perez, J.: Fluctuations in the Curie-Weiss version of the random field Ising model. *J. Stat. Phys.* **62**(3–4), 587–608 (1991)
7. Chatterjee, S.: A note about the uniform distribution on the intersection of a simplex and a sphere. *J. Topol. Anal.* **09**(04), 717–738 (2017)
8. Nam, K.: Large deviations and localization of the microcanonical ensembles given by multiple constraints. *Ann. Probab.* **48**(5), 2525–2564 (2020)
9. Huvneers, F., Theil, E.: Equivalence of ensembles, condensation and glassy dynamics in the Bose-Hubbard Hamiltonian. *J. Stat. Phys.* **177**(5), 917–935 (2019)
10. Lukkarinen, J.: Multi-state condensation in Berlin-Kac spherical models. *Commun. Math. Phys.* **373**(1), 389–433 (2019)
11. Koskinen, K., Lukkarinen, J.: Estimation of local microcanonical averages in two lattice mean-field models using coupling techniques. *J. Stat. Phys.* **180**(1–6), 1206–1251 (2020)
12. Großkinsky, S.: Equivalence of ensembles for two-species zero-range invariant measures. *Stoch. Process. Appl.* **118**(8), 1322–1350 (2008)
13. Szavits-Nossan, J., Evans, M.R., Majumdar, S.N.: Condensation transition in joint large deviations of linear statistics. *J. Phys. A* **47**(45), 455004 (2014)
14. Berlin, T.H., Kac, M.: The spherical model of a ferromagnet. *Phys. Rev.* **86**(6), 821–835 (1952)
15. Touchette, H.: Equivalence and nonequivalence of ensembles: thermodynamic, macrostate, and measure levels. *J. Stat. Phys.* **159**(5), 987–1016 (2015)
16. Brändén, P., Huh, J.: Lorentzian polynomials. *Ann. Math.* **192**(3), 821–891 (2020)
17. Caputo, P.: Uniform poincaré inequalities for unbounded conservative spin systems: the non-interacting case. *Stoch. Process. Appl.* **106**(2), 223–244 (2003)
18. Cancrini, N., Olla, S.: Ensemble dependence of fluctuations: canonical microcanonical equivalence of ensembles. *J. Stat. Phys.* **168**(4), 707–730 (2017)
19. Koskinen, K.: Infinite volume Gibbs states and metastates of the random field mean-field spherical model. *J. Stat. Phys.* **190**(3), 62 (2023)
20. Rockafellar, R.T.: *Convex Analysis*. Princeton University Press, Princeton (1997)
21. Wong, R.: *Asymptotic Approximations of Integrals*. Society for Industrial and Applied Mathematics, Philadelphia (2001)
22. den Hollander, F.: *Large Deviations*. American Mathematical Society, Providence (2008)
23. Georgii, H.-O.: *Gibbs Measures and Phase Transitions*. De Gruyter, Berlin (2011)
24. Sason, I., Verdu, S.:  $f$ -Divergence inequalities. *IEEE Trans. Inf. Theory* **62**(11), 5973–6006 (2016)
25. Anari, N., Gharan, S.O., Vintant, C.: Log-concave polynomials, I: entropy and a deterministic approximation algorithm for counting bases of matroids. *Duke Math. J.* **170**(16), 3459–3504 (2021)
26. Gardner, R.J.: The Brunn-Minkowski inequality. *Bull. Am. Math. Soc.* **39**(03), 355–406 (2002)
27. Abramowitz, M.: *Handbook of Mathematical Functions, With Formulas, Graphs, and Mathematical Tables*. Dover Publications Inc, Mineola (1974)

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.