Proof of the Local REM Conjecture for Number Partitioning. I: Constant Energy Scales

Christian Borgs,[1] Jennifer Chayes,[1] Stephan Mertens,[2] Chandra Nair[3,*]
[1] *Microsoft Research, One Memorial Drive, Cambridge, Massachusetts; e-mail: jchayes@microsoft.com*
[2] *Institut für Theoretische Physik, Otto-von-Guericke Universität, PF 4120, 39016 Magdeburg, Germany*
[3] *Department of Information Engineering, The Chinese University of Hong Kong, Hong Kong, China*

Received 7 February 2005; accepted 28 January 2006; received in final form 12 June 2008
Published online 15 December 2008 in Wiley InterScience (www.interscience.wiley.com).
DOI 10.1002/rsa.20255

ABSTRACT: In this article we consider the number partitioning problem (NPP) in the following probabilistic version: Given n numbers $X_1, \ldots, X_n$ drawn i.i.d. from some distribution, one is asked to find the partition into two subsets such that the sum of the numbers in one subset is as close as possible to the sum of the numbers in the other set. In this probabilistic version, the NPP is equivalent to a mean-field antiferromagnetic Ising spin glass, with spin configurations corresponding to partitions, and the energy of a spin configuration corresponding to the weight difference. Although the energy levels of this model are a priori highly correlated, a surprising recent conjecture of Bauke, Franz, and Mertens asserts that the energy spectrum of number partitioning is locally that of a random energy model (REM): the spacings between nearby energy levels are uncorrelated. More precisely, it was conjectured that the properly scaled energies converge to a Poisson process, and that the spin configurations corresponding to nearby energies are asymptotically uncorrelated. In this article, we prove these two claims, collectively known as the local REM conjecture. © 2008 Wiley Periodicals, Inc. Random Struct. Alg., 34, 217–240, 2009

Keywords: poisson convergence; combinatorial optimization; random energy model

Correspondence to: J. Chayes
*This work was done while C.N. was a student at Stanford University and a post-doc at Microsoft Research.

1. INTRODUCTION

The study of typical properties of random instances of combinatorial problems has recently been the focus of much interest in the theoretical computer science, discrete mathematics, and statistical physics communities. Many of these problems turn out to be closely related to disordered problems in statistical physics [12, 21]—a connection which has motivated a host of interesting conjectures. In this article, we establish one of these conjectures: the local REM property of the random number partitioning problem (NPP).

The non-random NPP is one of the classic NP-complete problems of combinatorial optimization, closely related to other classic problems such as bin packing, multiprocessor scheduling, quadratic programming, and knapsack problems [1, 15]. In addition to its theoretical significance, the NPP has many applications including task scheduling and the minimization of VLSI circuit size and delay [10, 24], public key cryptography [18, 22], and, more amusingly, choosing teams in children's baseball games [17].

A fixed instance of the NPP is defined as follows: Given n numbers $X_1, X_2, \ldots, X_n$, we seek a partition of these numbers into two subsets such that the sum of numbers in one subset is as close as possible to the sum of numbers in the other subset. Each of the 2^n partitions can be encoded as $\boldsymbol{\sigma} \in \{-1, +1\}^n$, where $\sigma_i = 1$ if X_i is put in one subset and $\sigma_i = -1$ if X_i is put in the other subset; in the physics literature, such partitions $\boldsymbol{\sigma}$ are identified with *Ising spin configurations*. The cost function to be minimized over all spin configurations $\boldsymbol{\sigma}$ is therefore the *energy*

$$E(\boldsymbol{\sigma}) = \frac{1}{\sqrt{n}} \left| \sum_{s=1}^{n} \sigma_s X_s \right|, \tag{1.1}$$

where we have inserted a factor $1/\sqrt{n}$ to simplify the equations in the rest of the paper. (The unnormalized quantity $|\sum_{s=1}^{n} \sigma_s X_s|$ is called the discrepancy).

Note that the spin configurations $\boldsymbol{\sigma}$ and $-\boldsymbol{\sigma}$ correspond to the same partition and therefore of course have the same energy. Thus there are $N = 2^{n-1}$ distinct partitions and at most N distinct energies. The lowest of these N energies is the *ground state energy* of the model. The *energy spectrum* is the sorted increasing sequence $E_1, \ldots, E_N$ of the energy values corresponding to these N distinct partitions. Let $\boldsymbol{\sigma}^{(1)}, \ldots, \boldsymbol{\sigma}^{(N)}$ be the configurations corresponding to these ordered energies. The *overlap* between the configurations $\boldsymbol{\sigma}^{(i)}$ and $\boldsymbol{\sigma}^{(j)}$ is defined as

$$q(\boldsymbol{\sigma}^{(i)}, \boldsymbol{\sigma}^{(j)}) = \frac{1}{n} \sum_{s=1}^{n} \sigma_s^{(i)} \sigma_s^{(j)}. \tag{1.2}$$

One often studies random instances of the NPP where the n numbers $X_1, \ldots, X_n$ are taken to be independently and identically distributed according to some density $\rho(X)$. In most cases studied so far, the X_i are taken to be drawn uniformly from a bounded domain, say integer values drawn uniformly from $\{1, \ldots, 2^m\}$ or real values drawn uniformly from $[0, 1]$. The statistical mechanics of this model has been discussed by several authors [13, 14, 19, 23].

When $\rho(X)$ is the uniform distribution on $\{1, \ldots, 2^m\}$, it turns out that the typical properties of random instances depend on the ratio $\kappa = m/n$. Numerical simulations suggested that in the limit $n, m \to \infty$ with κ fixed, this system had a sharp transition at $\kappa = 1$ between a phase in which there are exponentially many optimal solutions with discrepancy 0 or 1, and a phase where the optimal solution is unique (except for trivial symmetry) and has an energy which grows like $2^{n(\kappa-1)}$ [16]. This was supported by a statistical physics approach [19] and confirmed by rigorous analysis [7].

For the random NPP, the costs of two partitions $\boldsymbol{\sigma}$ and $\boldsymbol{\sigma}'$ are *a priori* highly correlated random variables. In [20], one of the authors made a rather surprising "random cost approximation," in which the correlations of energies near the ground state were neglected. Within this approximation, it is easy to calculate the statistics of the ground state and the first excitations. Remarkably, the results of these calculations were later confirmed by rigorous analysis [7], which therefore suggested that there might be a mathematical basis for this approximation.

Numerical simulation and heuristic arguments led to an even stronger conjecture, namely, that the statistical independence of nearby levels is not restricted to energies close to the ground state but extends to all fixed "typical" energies [2]. These authors also conjectured that the overlaps corresponding to these energies are uncorrelated. These two claims were collectively called the *local REM conjecture* [2], as the proposed behavior of nearby energies was analogous to that of the random energy model (REM) in spin glass theory [11]. In this article, we prove the local REM conjecture for the NPP with a general distribution of the X_i.

In physical terms, the optimal partitions of the NPP are precisely analogous to the ground states of a mean-field antiferromagnetic Ising spin system with Mattis-like couplings $J_{ij} = -X_iX_j$, defined by the Hamiltonian

$$H(\boldsymbol{\sigma}) = E^2(\boldsymbol{\sigma}) = \frac{1}{n}\sum_{ij} X_iX_j\sigma_i\sigma_j = -\frac{1}{n}\sum_{ij} J_{ij}\sigma_i\sigma_j. \tag{1.3}$$

Similarly, the energy spectrum and overlaps of the NPP are analogous to those of the mean-field antiferromagnetic Mattis spin glass. Our results therefore also establish the REM conjecture for this spin glass.

2. STATEMENT OF RESULTS

Let $X_1, \dots, X_n$ be independent random variables distributed according to a common density function $\rho(x)$. We assume that ρ has finite second moment τ^2 and satisfies the bound

$$\int_{-\infty}^{\infty} \rho(x)^{1+\epsilon}\mathrm{d}x < \infty \tag{2.1}$$

for some $\epsilon > 0$. Note that this includes, in particular, all bounded density functions with finite second moment. We use the symbol $\mathbb{P}_n(\cdot)$ to denote the probability with respect to the joint probability distribution of $X_1, \dots, X_n$.

As in the introduction, we represent the 2^n partitions of the integers $\{1, .., n\}$ as spin configurations $\boldsymbol{\sigma} \in \{-1, +1\}^n$, define the energy of $\boldsymbol{\sigma}$ as in (1.1), and denote by $E_1, \dots, E_N$ the increasing spectrum of the energy values corresponding to the $N = 2^{n-1}$ distinct partitions. We also denote by $\boldsymbol{\sigma}^{(1)}, \dots, \boldsymbol{\sigma}^{(N)}$ the configurations corresponding to these ordered energies.

Fix a constant $\alpha \geq 0$, let r_n be the random variable defined by $E_{r_n} < \alpha \leq E_{r_n+1}$, and let

$$e_i = \frac{1}{\sqrt{2\pi\tau^2}} 2^n e^{-\alpha^2/(2\tau^2)}(E_{r_n+i} - \alpha). \tag{2.2}$$

We call the sequence $e_1, e_2, \dots$ the rescaled energy spectrum above α, and say that this rescaled energy spectrum converges weakly to a Poisson process if for every finite $l \geq 1$, the random vector $(e_1, e_2, \ldots, e_l)$ converges weakly to $(w_1, w_1+w_2, \ldots, w_1+w_2+\cdots+w_l)$, where w_i are i.i.d. random variables distributed exponentially with mean 1. For $j > i > 0$, we define the rescaled overlap

$$Q_{ij} = \frac{1}{\sqrt{n}} \sum_{s=1}^{n} \sigma_s^{(r_n+i)} \sigma_s^{(r_n+j)}. \tag{2.3}$$

The rescaled overlaps converges weakly to a standard normal if

$$\lim_{n\to\infty} \mathbb{P}_n(Q_{ij} \geq \beta) = \frac{1}{\sqrt{2\pi}} \int_{\beta}^{\infty} e^{-\frac{x^2}{2}} \, \mathrm{d}x \tag{2.4}$$

for all $1 \leq i < j$ and all $\beta \in \mathbb{R}$.

Bauke et al. [2] conjectured the following behavior for the energy level and overlap statistics of the NPP with X_i uniformly distributed in $[0, 1]$:

Conjecture 2.1. *Let $\alpha \geq 0$, and let $X_1, X_2, \ldots, X_n$ be i.i.d. random variables distributed uniformly in* $[0, 1]$. *Then the rescaled energy spectrum above α converges weakly to a Poisson process, and the spin configurations corresponding to different energy levels become asymptotically uncorrelated in the sense that Q_{ij} converges weakly to a standard normal for all $i < j$.*

For $\alpha = 0$, the part of the conjecture concerning the energies was already rigorously established in [7]. In this article, we prove that the full Conjecture 2.1 for fixed $\alpha > 0$ holds not only for the uniform distribution, but also for any distribution ρ on $(-\infty, \infty)$ which has finite second moment and satisfies (2.1).

Theorem 2.2. *Let $\alpha \geq 0$, and let $X_1, \ldots, X_n$ be i.i.d. random variables with probability density ρ. If ρ has finite second moment and satisfies the condition* (2.1)*, then the rescaled energy spectrum above α converges weakly to a Poisson process, and the rescaled overlaps converge weakly to a standard normal.*

Remark 2.3.

1. Taking ρ to be the uniform distribution over $[0, 1]$, with second moment $\tau^2 = \frac{1}{3}$, we see that Theorem 2.2 implies Conjecture 2.1.
2. Having established the original REM Conjecture 2.1, the question naturally arises whether analogous results hold for energy scales α which grow with n. Indeed, the authors of [2] suggested that the conjecture might extend to values of α that grow slowly enough with n, although computational limitations prevented them from supporting this stronger claim by simulations. In a second article [6], we will show that, under suitable additional assumptions on the distribution ρ, the conjecture does indeed hold provided $\alpha = o(n^{1/4})$.

In addition to immediately implying the analogous results for the mean-field antiferromagnetic Mattis spin glass [see Eq. (1.3)], our theorem on the energy spectrum of the NPP also gives the energy spectrum of the one-dimensional Edwards-Anderson (1-d EA) spin glass model away from the ground state. The 1-d EA model has energies

$E(\sigma) = \sum_i J_i\sigma_i\sigma_{i+1}$. Consider the transformation $\tau_i = \sigma_i\sigma_{i+1}$ and take the boundary condition $\sigma_{n+1} = 1$. Then $E(\sigma) = \sum_i J_i\tau_i$, so that, up to a multiplicative factor of $\sqrt{n}$, the energy of the NPP with random variables X_i is the same as the absolute value of the energy of the 1-d EA model with random variables J_i. Note that the energy spectrum of the NPP lies in $[0, E_{\max}]$, with $E_{\max} = \theta(\sqrt{n})$, while that of the 1-d EA model lies in $[-\sqrt{n}E_{\max}, \sqrt{n}E_{\max}]$.

Our theorem says that properly scaled energies of the NPP converge to a Poisson process. By the above transformation, this result applies also to the 1-d EA model except for energies about zero, which are correlated by symmetry. In particular, the result applies to the 1-d EA model in energy intervals of the form $[\sqrt{n}\alpha, \sqrt{n}(\alpha + \theta(e^{-n}))]$ for any bounded $\alpha \geq 0$ or their reflections about 0. If the interval includes the origin as an internal point, the positive and negative energies separately converge to a Poisson processes, with the two obviously related by a spin-flip symmetry.

3. PROOF OF THEOREM 2.2

3.1. Outline of the Proof

Before we proceed with our proof, note that we may assume without loss of generality that the second moment τ^2 is equal to 1 and that ρ is symmetric, $\rho(x) = \rho(-x)$. Indeed, considering the rescaled random variables $\tilde{X}_i = \tau^{-1}X_i$, we immediately see that the statements of theorem for general τ follow from those for $\tau = 1$. Next, consider the random variables $Y_1, \ldots, Y_n$ where each Y_i is obtained as X_i w.p. $\frac{1}{2}$ or $-X_i$ w.p. $\frac{1}{2}$. It is easy to see that the energy spectrum of the $Y_1, \ldots, Y_n$ is identical to that of the $X_1, \ldots, X_n$. Further, from convexity of $|x|^{1+\epsilon}$ and Jensen's inequality, it follows that $\rho_Y(y) = \frac{1}{2}(\rho(x) + \rho(-x))$ also satisfies (2.1). Therefore, w.l.o.g. we can assume that $\rho(x) = \rho(-x)$ as claimed, and in particular that X_i has zero first moment. For simplicity of notation, we omit the subscript n, and denote the probability with respect to the joint distribution of $X_1, \ldots, X_n$ by $\mathbb{P}(\cdot)$, and the expectation with respect to this distribution by $\mathbb{E}(\cdot)$.

Let $Z_n(a, b)$ be the number of points of the energy spectrum that lie in the interval $[a, b]$, and let $N_n(t)$ be the number of points in the energy spectrum that fall into the (shifted and) re-scaled interval $[\alpha, \alpha + t\xi_n]$, where

$$\xi_n = \sqrt{2\pi}2^{-n}e^{\alpha^2/2}. \tag{3.1}$$

We must show that $N_n(t)$ converges to a Poisson process with parameter one. To this end, we will show that for any family of non-overlapping intervals $[c_1, d_1], \ldots, [c_m, d_m]$ with $d_i > c_i \geq 0$, the variables $Z_n(\alpha + c_i\xi_n, \alpha + d_i\xi_n)$ converge in distribution to the *increments* of a Poisson process with parameter one. We establish this by showing the convergence of the multidimensional factorial moments, i.e., by proving the following theorem.

Theorem 3.1. *Let $\alpha \geq 0$, let m be a positive integer, and let $[c_1, d_1], \ldots, [c_m, d_m]$ be a family of non-overlapping intervals. For $\ell = 1, \ldots, m$, set $a_n^\ell = \alpha + c_\ell\xi_n$ and $b_n^\ell = \alpha + d_\ell\xi_n$. Given an m-tuple $(k_1, .., k_m)$ of positive integers, we then have*

$$\lim_{n\to\infty} \mathbb{E}\left[\prod_{\ell=1}^m \left(Z_n\left(a_n^\ell, b_n^\ell\right)\right)_{k_\ell}\right] = \prod_{\ell=1}^m \gamma_\ell^{k_\ell}, \tag{3.2}$$

where $\gamma_\ell = d_\ell - c_\ell$ and, as usual, $(Z)_k = Z(Z-1)\ldots(Z-k+1)$.

Theorem 3.1 establishes that $N_n(t)$ converges to a Poisson with rate one. As we will see later, the asymptotic independence of spin configurations corresponding to nearby energy levels is an immediate corollary to the proof of this theorem.

The proof of Theorem 3.1 proceeds as follows: First, we rewrite the left hand side of equation (3.2) in the form

$$\mathbb{E}\left[\prod_{\ell=1}^{m}\left(Z_n\left(a_n^\ell, b_n^\ell\right)\right)_{k_\ell}\right] = \sum_{\pm\boldsymbol{\sigma}^{(1)}\neq\cdots\neq\pm\boldsymbol{\sigma}^{(k)}} \mathbb{E}[I_k(\boldsymbol{\sigma}^{(1)},\ldots,\boldsymbol{\sigma}^{(k)})] \tag{3.3}$$

where $k = \sum_\ell k_\ell$, the sum goes over configurations $\boldsymbol{\sigma}^{(1)},\ldots,\boldsymbol{\sigma}^{(k)} \in \{-1,+1\}^n$ such that $\boldsymbol{\sigma}^{(i)}$ is different from both $\boldsymbol{\sigma}^{(j)}$ and $-\boldsymbol{\sigma}^{(j)}$ for all $j \neq i$, and $I_k(\boldsymbol{\sigma}^{(1)},\ldots,\boldsymbol{\sigma}^{(k)})$ is proportional to the indicator function that the vector-valued random variable $[E(\boldsymbol{\sigma}^{(1)}),\ldots,E(\boldsymbol{\sigma}^{(k)})]$ lies in an appropriately chosen neighborhood near the point $(\alpha,\ldots,\alpha) \in \mathbb{R}^k$, see (3.6) and (3.10) below for the precise definition of $I_k(\cdot)$.

One might therefore want to prove Theorem 3.1 by establishing a local limit theorem for the terms on the right hand side of (3.3). For $k = 1$, this is not too difficult. Indeed, for $k = 1$, the required local limit theorem can be seen to be a special case of Theorem 19.1 in [5]. But for higher k, the situation is more complicated. Indeed, the local limit theorem does not hold for all terms on the right hand side of (3.3). Instead, it only holds for "most" terms. We will therefore have to distinguish "good" and "bad" sets of configurations $\boldsymbol{\sigma}^{(1)},\ldots,\boldsymbol{\sigma}^{(k)}$. For the good ones, we will establish a new local limit theorem, and for the bad ones, we need to establish bounds which are strong enough to guarantee that the sum over all bad configurations is negligible in the limit $n \to \infty$.

This strategy is similar in structure to that employed in [7], but here the control of the "bad" terms is more complicated. This is partly simply due to the fact that now $\alpha \neq 0$. But additional complications also arise because we need to establish error bounds which are good enough for the applications in Part 2 of this article, where we allow α to grow with n up to the boundary of the region were the local REM-conjecture starts to fail.

3.2. Integral Representation and First Moment

Next we derive the representation (3.3). Let

$$\mathrm{rect}(x) = \begin{cases} 1 & -\frac{1}{2} \leq x \leq \frac{1}{2} \\ 0 & \text{otherwise,} \end{cases} \tag{3.4}$$

let $t_n^\ell = \frac{a_n^\ell + b_n^\ell}{2}$ denote the center of the interval $[a_n^\ell, b_n^\ell]$ and let $q_{n,\ell} = \gamma_\ell \xi_n \sqrt{n}$. Then $Z_n(a_n^\ell, b_n^\ell)$ can be written as

$$Z_n\left(a_n^\ell, b_n^\ell\right) = \sum_{\boldsymbol{\sigma}} I^{(\ell)}(\boldsymbol{\sigma}) \tag{3.5}$$

where

$$I^{(\ell)}(\boldsymbol{\sigma}) = \frac{1}{2}\left[\mathrm{rect}\left(\frac{\sum_{s=1}^n \sigma_s X_s - t_n^\ell\sqrt{n}}{q_{n,\ell}}\right) + \mathrm{rect}\left(\frac{\sum_{s=1}^n \sigma_s X_s + t_n^\ell\sqrt{n}}{q_{n,\ell}}\right)\right]. \tag{3.6}$$

Note the factor $\frac{1}{2}$, which arises from the fact that each partition is counted only once in $Z_n(a_n^\ell, b_n^\ell)$, while the two configurations $\boldsymbol{\sigma}$ and $-\boldsymbol{\sigma}$ correspond to the same partition of $\{1,\ldots,n\}$.

Next we write the k^{th} factorial moment of $Z_n(a_n^\ell, b_n^\ell)$ as a sum over sequences of k distinct configurations $\boldsymbol{\sigma}^{(1)}, \dots, \boldsymbol{\sigma}^{(k)}$. Let

$$I_k^{(\ell)}(\boldsymbol{\sigma}^{(1)}, \dots, \boldsymbol{\sigma}^{(k)}) = \prod_{j=1}^{k} I^{(\ell)}(\boldsymbol{\sigma}^{(j)}). \tag{3.7}$$

A simple counting argument shows that

$$\left(Z_n\left(a_n^\ell, b_n^\ell\right)\right)_k = \sum_{\pm\boldsymbol{\sigma}^{(1)} \neq \cdots \neq \pm\boldsymbol{\sigma}^{(k)}} I_k^{(\ell)}(\boldsymbol{\sigma}^{(1)}, \dots, \boldsymbol{\sigma}^{(k)}) \tag{3.8}$$

where the sum runs over distinct partitions. Using linearity of expectation this implies that

$$\mathbb{E}\left[\left(Z_n\left(a_n^\ell, b_n^\ell\right)\right)_k\right] = \sum_{\pm\boldsymbol{\sigma}^{(1)} \neq \cdots \neq \pm\boldsymbol{\sigma}^{(k)}} \mathbb{E}\left[I_k^{(\ell)}(\boldsymbol{\sigma}^{(1)}, \dots, \boldsymbol{\sigma}^{(k)})\right]. \tag{3.9}$$

To obtain a formula for the multi-dimensional factorial moments, let us consider two disjoint intervals $[a_n^\ell, b_n^\ell]$ and $[a_n^{\ell'}, b_n^{\ell'}]$, and two sequences of configurations $\boldsymbol{\sigma}^{(1)}, \dots, \boldsymbol{\sigma}^{(k_\ell)}$ and $\tilde{\boldsymbol{\sigma}}^{(1)}, \dots, \tilde{\boldsymbol{\sigma}}^{(k_{\ell'})}$ contributing to $(Z_n(a_n^\ell, b_n^\ell))_{k_\ell}$ and $(Z_n(a_n^{\ell'}, b_n^{\ell'}))_{k_{\ell'}}$, respectively. Recall that $E(\boldsymbol{\sigma}) = E(-\boldsymbol{\sigma})$. As $I^{(\ell)}(\boldsymbol{\sigma}) = 0$ unless $E(\boldsymbol{\sigma}) \in [a_n^\ell, b_n^\ell]$ and $I^{(\ell')}(\tilde{\boldsymbol{\sigma}}) = 0$ unless $E(\tilde{\boldsymbol{\sigma}}) \in [a_n^{\ell'}, b_n^{\ell'}]$, we see that $I^{(\ell)}(\boldsymbol{\sigma})I^{(\ell')}(\tilde{\boldsymbol{\sigma}}) = 0$ if $\boldsymbol{\sigma}$ and $\tilde{\boldsymbol{\sigma}}$ are not distinct partitions. The combined sequence $\boldsymbol{\sigma}^{(1)}, \dots, \boldsymbol{\sigma}^{(k_\ell)}, \tilde{\boldsymbol{\sigma}}^{(1)}, \dots, \tilde{\boldsymbol{\sigma}}^{(k_{\ell'})}$ therefore only contributes to the product $(Z_n(a_n^\ell, b_n^\ell))_{k_\ell}(Z_n(a_n^{\ell'}, b_n^{\ell'}))_{k_{\ell'}}$ if $\boldsymbol{\sigma}^{(j)} \neq \pm\tilde{\boldsymbol{\sigma}}^{(\ell')}$ for all $\ell \neq \ell'$. As a consequence, the multi-dimensional factorial moment in Theorem 3.1 is itself given as sum over sequences of pairwise distinct partitions, as claimed in (3.3). More explicitly, let $k = \sum_{\ell=1}^m k_\ell$, and for $j = 1, \dots, k$, let $\ell(j) = 1$ if $j = 1, \dots, k_1$, $\ell(j) = 2$ if $j = k_1 + 1, \dots, k_1 + k_2$, and so on. Setting

$$I_k(\boldsymbol{\sigma}^{(1)}, \dots, \boldsymbol{\sigma}^{(k)}) = \prod_{j=1}^{k} I^{(\ell(j))}(\boldsymbol{\sigma}^{(j)}), \tag{3.10}$$

we obtain (3.3).

Having derived Eq. (3.3), we will now want to establish a local limit theorem for the typical term on the right hand side of (3.10). As in the derivation of standard local limit theorems, it will be convenient to consider the Fourier transform

$$\hat{\rho}(f) = \mathbb{E}[e^{2\pi i f X}]. \tag{3.11}$$

We will use several properties of the Fourier transform in our proof, which we summarize now. All of them follow from the fact that the density $\rho(x)$ has finite second moment, satisfies equation (2.1), and is symmetric under the transformation $x \to -x$.

i. For any $\mu_1 > 0$, there exists $c_1 > 0$, possibly depending on μ_1, such that whenever $|f| \geq \mu_1$, we have $\left|\hat{\rho}(f)\right| < e^{-c_1}$.
ii. For any $n \geq n_o$, where n_0 is the solution of $\frac{1}{1+\epsilon} + \frac{1}{n_o} = 1$ with ϵ as in (2.1), we have

$$\int_{-\infty}^{\infty} |\hat{\rho}(f)|^n \leq \int_{-\infty}^{\infty} |\hat{\rho}(f)|^{n_0} = C_0 < \infty. \tag{3.12}$$

iii. $\hat{\rho}(f) \to 0$ as $|f| \to \infty$.
iv. There exists a $c_2 > 0$ such that, for $\mu_1 > 0$ small enough, whenever $|f| \leq \mu_1$, we have $|\hat{\rho}(f)| \leq e^{-c_2 f^2}$.

The next lemma gives the desired local limit theorem for the first moment.

Lemma 3.2 (First Moment).

$$\lim_{n\to\infty} \mathbb{E}\big[Z_n\big(a_n^\ell, b_n^\ell\big)\big]\big] = \gamma_\ell. \tag{3.13}$$

Proof. We use the local limit theorem in the form[1] of Theorem 19.1 of [5]. By this theorem, Condition (3.12) implies that $\mathbb{E}[I^{(\ell)}(\boldsymbol{\sigma})] = 2^{-n}\gamma_\ell(1 + o(1))$, which in turn implies Lemma 3.2. Note that Theorem 19.1 of [5] is slightly stronger than the standard local limit theorem, which does not allow for intervals $[a_n^\ell, b_n^\ell]$ whose width shrinks with n. But luckily for us, no such restriction is needed for Theorem 19.1 of [5]. ∎

3.3. Higher Moments

In this subsection, we analyze the higher moments. Bearing in mind the similar structure of the representations (3.3) and (3.9), we first consider the one-dimensional factorial moments $\mathbb{E}[(Z_n(a_n^\ell, b_n^\ell))_k]$. We omit the index ℓ, and write a_n, b_n, γ and q_n for a_n^ℓ, b_n^ℓ, γ_ℓ and $q_{n,\ell}$, respectively. We also write $I(\cdot)$ and $I_k(\cdot)$ instead of $I^{(\ell)}(\cdot)$ and $I_k^{(\ell)}(\cdot)$.

We want to show that in the limit $n \to \infty$, the factorial moment $\mathbb{E}[(Z_n(a_n, b_n))_k]$ is equal to γ^k. As the sum in (3.9) contains $(2^n)_k = 2^{nk}(1 + o(1))$ terms, one might therefore try to show that $\mathbb{E}[I_k(\boldsymbol{\sigma}^{(1)}, \dots, \boldsymbol{\sigma}^{(k)})]$ is asymptotically equal to $\gamma^k 2^{-nk}$ by applying vector forms of local limit theorems. But as alluded to earlier, this strategy does not work for all $\boldsymbol{\sigma}^{(1)}, \dots, \boldsymbol{\sigma}^{(k)}$. Instead, we proceed as follows:

First we formulate a condition on $\boldsymbol{\sigma}^{(1)}, \dots, \boldsymbol{\sigma}^{(k)}$ which is strong enough to prove a local limit theorem for the vector-valued random variable $(E(\boldsymbol{\sigma}^{(1)}), \dots, E(\boldsymbol{\sigma}^{(k)}))$, and weak enough to hold for all but a vanishing fraction of the configurations $\boldsymbol{\sigma}^{(1)}, \dots, \boldsymbol{\sigma}^{(k)}$. As a consequence, the total contribution of these configurations will be $\gamma^k + o(1)$. Having extracted the leading behavior, we then estimate the contributions of all other configurations $\boldsymbol{\sigma}^{(1)}, \dots, \boldsymbol{\sigma}^{(k)}$ and show that they do not contribute in the limit $n \to \infty$.

In a preliminary step, we derive an integral representation for $\mathbb{E}[I_k(\boldsymbol{\sigma}^{(1)}, \dots, \boldsymbol{\sigma}^{(k)})]$ when $\boldsymbol{\sigma}^{(1)}, \dots, \boldsymbol{\sigma}^{(k)}$ are linearly independent.

Lemma 3.3. *Let k be a positive integer, and let $\boldsymbol{\sigma}^{(1)}, .., \boldsymbol{\sigma}^{(k)}$ be linearly independent configurations in $\{-1, +1\}^n$. Then*

$$\mathbb{E}[I_k(\boldsymbol{\sigma}^{(1)}, \dots, \boldsymbol{\sigma}^{(k)})] = q_n^k \iiint_{-\infty}^{\infty} \prod_{s=1}^{n} \hat{\rho}(v_s) \prod_{j=1}^{k} \operatorname{sinc}(f_j q_n) \cos(2\pi f_j t_n \sqrt{n}) \mathrm{d}f_j, \tag{3.14}$$

where $\operatorname{sinc}(f) = (\sin \pi f)/(\pi f)$ *and*

$$v_s = \sum_{j=1}^{k} \sigma_s^{(j)} f_j, \quad 1 \leq s \leq n. \tag{3.15}$$

[1] We thank the referee for pointing out this reference.

Proof. Recall that the Fourier transform of $\mathrm{rect}(x)$ is equal to $\mathrm{sinc}(f)$, and let $\mathrm{rect}_B(x) = \int_{-B}^{B} \mathrm{sinc}(f) e^{2\pi i f x} \mathrm{d}f$. Since $\mathrm{rect}_B(x)$ is bounded uniformly in B and converges pointwise to $\mathrm{rect}(x)$ except for $x = \pm 1/2$, we have that

$$
\begin{aligned}
&I_k^{(\le B)}(\boldsymbol{\sigma}^{(1)}, \dots, \boldsymbol{\sigma}^{(k)}) \\
&= q_n^k \iiint_{-B}^{B} \prod_{j=1}^{k} e^{2\pi i f_j \sum_{s=1}^{n} X_s \sigma_s^{(j)}} \mathrm{sinc}(f_j q_n) \cos(2\pi f_j t_n \sqrt{n}) \mathrm{d}f_j \\
&= q_n^k \iiint_{-B}^{B} \prod_{s=1}^{n} e^{2\pi i X_s v_s} \prod_{j=1}^{k} \mathrm{sinc}(f_j q_n) \cos(2\pi f_j t_n \sqrt{n}) \mathrm{d}f_j
\end{aligned} \tag{3.16}
$$

is a random variable that is bounded uniformly in B and converges almost everywhere to $I_k(\boldsymbol{\sigma}^{(1)}, \dots, \boldsymbol{\sigma}^{(k)})$. We thus may use dominated convergence to conclude that

$$
\mathbb{E}[I_k(\boldsymbol{\sigma}^{(1)}, \dots, \boldsymbol{\sigma}^{(k)})] = \lim_{B \to \infty} \mathbb{E}\big[I_k^{(\le B)}(\boldsymbol{\sigma}^{(1)}, \dots, \boldsymbol{\sigma}^{(k)})\big]. \tag{3.17}
$$

Next we use Fubini's theorem to rewrite

$$
\mathbb{E}\big[I_k^{(\le B)}(\boldsymbol{\sigma}^{(1)}, \dots, \boldsymbol{\sigma}^{(k)})\big] = q_n^k \iiint_{-B}^{B} \prod_{s=1}^{n} \hat{\rho}(v_s) \prod_{j=1}^{k} \mathrm{sinc}(f_j q_n) \cos(2\pi f_j t_n \sqrt{n}) \mathrm{d}f_j. \tag{3.18}
$$

To complete the proof of the lemma, we will want to take the limit $B \to \infty$ in (3.18). To justify this step, we will again use dominated convergence theorem.

Let $M_k = [\sigma_s^{(j)}]_{j \le k}, s \le n$. As $\boldsymbol{\sigma}^{(1)} \dots, \boldsymbol{\sigma}^{(k)}$ are linearly independent, the matrix M_k has rank k. Relabeling, if necessary, let us assume that $[\sigma_1^{(j)}]_{j \le k}, \dots, [\sigma_k^{(j)}]_{j \le k}$ form a basis of the row space. Hölder's inequality, the fact that $|\hat{\rho}(v_s)| \le 1$, and a change of variables from $f_1, \dots, f_k$ to $v_1, \dots, v_k$, then leads to the bound

$$
\begin{aligned}
&\iiint_{-B}^{B} \left| \prod_{s=1}^{n} \hat{\rho}(v_s) \prod_{j=1}^{k} \mathrm{sinc}(f_j q_n) \cos(2\pi f_j t_n \sqrt{n}) \right| \mathrm{d}f_j \\
&\le \left(\iiint_{-\infty}^{\infty} \prod_{j=1}^{k} |\mathrm{sinc}(f_j q_n)|^{1+\epsilon} \mathrm{d}f_j \right)^{1/(1+\epsilon)} \left(\iiint_{-\infty}^{\infty} \left| \prod_{s=1}^{k} |\hat{\rho}(v_s)|^{n_0} \prod_{j=1}^{k} \mathrm{d}f_j \right)^{1/n_0} \\
&= \left(\iiint_{-\infty}^{\infty} \prod_{j=1}^{k} |\mathrm{sinc}(f_j q_n)|^{1+\epsilon} \mathrm{d}f_j \right)^{1/(1+\epsilon)} \left(J_k \iiint_{-\infty}^{\infty} \left| \prod_{s=1}^{k} |\hat{\rho}(v_s)|^{n_0} \mathrm{d}v_s \right)^{1/n_0} \\
&< \infty,
\end{aligned} \tag{3.19}
$$

where n_0 and ϵ are as in (3.12) and J_k is the Jacobian of the change of variables from $f_1, \dots, f_k$ to $v_1, \dots, v_k$. By dominated convergence, we can therefore take the limit $B \to \infty$ in (3.18). Putting everything together, this gives (3.14). ∎

Next we would like to prove that for a "typical set of configurations" $\boldsymbol{\sigma}^{(1)}, \dots, \boldsymbol{\sigma}^{(k)}$, the integral on the right hand side of (3.14) is equal to $(2\pi n)^{-k/2} e^{-k\alpha^2/2}(1 + o(1))$. Here the meaning of typical is best formulated in terms of the matrix formed by the row vectors $\boldsymbol{\sigma}^{(1)}, \dots, \boldsymbol{\sigma}^{(k)}$. More generally, for $u \le k$ and $\boldsymbol{\sigma}^{(1)}, \dots, \boldsymbol{\sigma}^{(u)} \in \{-1, +1\}^n$, let M_u be the matrix with matrix elements $\sigma_s^{(j)}$, where $1 \le j, s \le u$. Given this matrix and a vector $\boldsymbol{\delta} \in \{-1,1\}^u$, let

$$n_{\boldsymbol{\delta}} = n_{\boldsymbol{\delta}}(\boldsymbol{\sigma}^{(1)},\ldots,\boldsymbol{\sigma}^{(u)}) = \left|\left\{j \le n : \left(\sigma_j^{(1)},\ldots,\sigma_j^{(u)}\right) = \boldsymbol{\delta}\right\}\right| \tag{3.20}$$

be the number of times the column vector $\boldsymbol{\delta}$ appears in the matrix M_u.

If one were to choose configurations $\boldsymbol{\sigma}^{(1)},\ldots,\boldsymbol{\sigma}^{(u)} \in \{-1,+1\}^n$ independently and uniformly at random, then for all $\delta \in \{-1,+1\}^u$, the expectation of $n_{\boldsymbol{\delta}}$ is clearly equal to $n2^{-u}$. By a standard Martingale argument, for most configurations, the difference between $n_{\boldsymbol{\delta}}$ and $n2^{-u}$ is then not much larger than $\sqrt{n}$, see Lemma 3.7 below. Let us therefore assume for the moment that

$$\max_{\boldsymbol{\delta}} \left| n_{\boldsymbol{\delta}}\left(\boldsymbol{\sigma}^{(1)},\ldots,\boldsymbol{\sigma}^{(u)}\right) - \frac{n}{2^u}\right| \le \sqrt{n}\lambda_n \tag{3.21}$$

for some $\lambda_n \to \infty$ sufficiently slowly. For concreteness, we choose $\lambda_n = \log n$, even though the proof works for much larger class of sequences. The next lemma shows that, under the condition (3.21), the right hand side of (3.14) behaves as desired.

Lemma 3.4. *Let $\lambda_n = \log n$, let k be a positive integer, and let $\boldsymbol{\sigma}^{(1)},\ldots,\boldsymbol{\sigma}^{(k)}$ be a sequence of configurations of rank k that satisfies* (3.21). *Then*

$$2^{nk}\mathbb{E}\left[I_k\left(\boldsymbol{\sigma}^{(1)},\ldots,\boldsymbol{\sigma}^{(k)}\right)\right] = \gamma^k + o(1), \tag{3.22}$$

where the constant implicit in the o-symbol depends on k.

Note that this lemma is a local limit theorem for the vector-valued random variable $(E(\boldsymbol{\sigma}^{(1)}),\ldots,E(\boldsymbol{\sigma}^{(k)}))$. Under slightly stronger conditions on the moments of ρ, such a limit theorem follows from Theorem 19.3 in Chapter 4 of [5].

Proof of Lemma 3.4. In view of Lemma 3.3 we will have to estimate the expression

$$2^{nk}q_n^k \iiint_{-\infty}^{\infty} \prod_{s=1}^{n} \hat{\rho}(v_s) \prod_{j=1}^{k} \operatorname{sinc}(f_j q_n)\cos(2\pi f_j t_n \sqrt{n})\mathrm{d}f_j. \tag{3.23}$$

Define μ_1, c_1, c_2, n_0, and C_0 to satisfy the assumptions about the Fourier transform around (3.12). We assume that n is large enough so that $\frac{n}{2^k} - \sqrt{n}\lambda_n > n_0$. In a first step, we want to show that the contribution of the region where $|v_s| > \mu_1$ for at least one s is negligible.

Thus consider the event that one of the $|v_s|$'s, say $|v_{t_1}|$, is larger than μ_1. Let I_3 be the restriction of the integral in (3.23) to this region, let $\boldsymbol{\delta}^1 = \{\sigma_{t_1}^{(1)},..,\sigma_{t_1}^{(k)}\}$, and let $\boldsymbol{\delta}^2,\ldots,\boldsymbol{\delta}^k$ be vectors such that the rank of $\{\boldsymbol{\delta}^1,..,\boldsymbol{\delta}^k\}$ is k. Let $\{v_{t_2},\ldots,v_{t_k}\}$ be defined by

$$v_{t_i} = \sum_{j=1}^{k} \delta_k^j f_j.$$

Since the vectors $\{\boldsymbol{\delta}^1,..,\boldsymbol{\delta}^k\}$ have rank k, we can change the variables of integration from f_j to v_{t_j}. Let the Jacobian of this transformation be J_k. The Jacobian J_k is bounded above by the largest determinant, J_{max}, of a matrix of size k whose entries are ± 1. We now bound the integral I_3 as follows:

$$
\begin{aligned}
|I_3| &= \left| \iiint_{-\infty}^{\infty} \int_{|v_{t_1}|>\mu_1} J_k \prod_{s=1}^{n} \hat{\rho}(v_s) \prod_{j=1}^{k} \operatorname{sinc}(f_j q_n) \cos(2\pi f_j t_n \sqrt{n}) \mathrm{d}v_{t_j} \right| \\
&\le J_{\max} \iiint_{-\infty}^{\infty} \prod_{j=2}^{k} |\hat{\rho}(v_{t_j})|^{n_{\delta^j}} \mathrm{d}v_{t_j} \times \int_{|v_{t_1}|>\mu_1} |\hat{\rho}(v_{t_1})|^{n_{\delta^1}} \mathrm{d}v_{t_1} \\
&\le J_{\max}(C_0)^{k-1} \int_{|v_{t_1}|>\mu_1} |\hat{\rho}(v_{t_1})|^{n_{\delta^1}} \mathrm{d}v_{t_1} \le J_{\max} C_0^k e^{-c_1(n_{\delta^1}-n_0)}.
\end{aligned} \tag{3.24}
$$

Observing that $2^{kn} q_n^k$ only grows polynomially in n, the number of choices for $\boldsymbol{\delta}^{t_1}$ is bounded by 2^k, and that $n_{\boldsymbol{\delta}^1} = n2^{-k} + o(n)$ by the bound (3.21), we conclude that the total contribution of the regions where at least one of the $|v_s|$'s is larger than μ_1 is exponentially small in n.

Consider now the region where all $|v_s|$'s are bounded by μ_1. In this region, we again would like to approximate the sinc factors in (3.23) by one. To this end, we first note that

$$
\sum_{j=1}^{k} |f_j| = \max_{s \le n} |v_s|. \tag{3.25}
$$

Indeed, by the triangle inequality, we clearly have that $\max_s |v_s| \le \sum_j |f_j|$. To prove the opposite inequality, we use that $n_{\boldsymbol{\delta}} = \frac{n}{2^k}(1 + o(1)) > 0$ for every $\boldsymbol{\delta} \in \{-1, +1\}^n$, implying that there exists a v_{s_0} that is evaluated as $\sum_{j=1}^{k} |f_j|$. In the region where all $|v_s|$'s are bounded by μ_1, we therefore have that all f_j's are bounded by μ_1, so that $\operatorname{sinc}(q_n f_j) = 1 + O(q_n^2)$. Furthermore, by the fact that $t_n\sqrt{n} = \alpha\sqrt{n} + O(q_n)$, we have $\cos(2\pi f_j t_n \sqrt{n}) = \cos(2\pi f_j \alpha \sqrt{n}) + O(q_n)$. Thus the error obtained by replacing the sinc factors by 1, and the product of $\cos(2\pi f_j t_n \sqrt{n})$ by $\cos(2\pi f_j \alpha \sqrt{n})$, can be bounded by $2^{nk} q_n^k O(q_n)$, which again goes to zero exponentially in n.

We thus have shown that, up to an error which is exponentially small in n, the left hand side of (3.22) is equal to

$$
2^{nk} q_n^k \iiint_{-\mu_1}^{\mu_1} \prod_{s=1}^{n} \hat{\rho}(v_s) \prod_{j=1}^{k} \cos(2\pi f_j \alpha \sqrt{n}) \mathrm{d}f_j. \tag{3.26}
$$

Next we show that we can further restrict the range of integration to $|v_s| \le \omega_n/\sqrt{n}$ for all s, as long as ω_n is a sequence that obeys the conditions $\omega_n \to \infty$ as $n \to \infty$. To this end, let us consider the integral where $\omega_n n^{-1/2} \le |v_{s_1}| \le \mu_1$, while v_s can be an arbitrary number in $[-\mu_1, \mu_1]$ for all other s. We will then have to bound the integral

$$
\tilde{I}_3 = \iiint_{-\mu_1}^{\mu_1} \int_{\omega_n/\sqrt{n} \le |v_{t_1}| \le \mu_1} J_k \prod_{s=1}^{n} \hat{\rho}(v_s) \prod_{j=1}^{k} \cos(2\pi f_j \alpha \sqrt{n}) \mathrm{d}v_{t_j}.
$$

Using the fact that $\hat{\rho}(v) \leq e^{-c_2 v^2}$ for $|v| \leq \mu_1$, this can be easily accomplished, leading to the bound

$$|\tilde{I}_3| \leq \iiint_{-\infty}^{\infty} \prod_{j=2}^{k} e^{-n_{\delta^j} c_2 v_{t_j}^2} \mathrm{d}v_{t_j} \times \int_{|v_{t_1}| > \omega_n/\sqrt{n}} e^{-n_{\delta^1} c_2 v_{t_1}^2} \mathrm{d}v_{t_1}$$
$$= O(n^{-k/2} e^{-c_2 \omega_n^2 (n_{\delta^1}/n)}). \tag{3.27}$$

Using the facts that $\omega_n \to \infty$, $n_{\delta^1}/n = 2^{-k} + o(1)$ and $2^{kn} q_n^k = O(n^{k/2})$, this implies that the contribution of the regions where $|v_{s_1}|$ is larger than $\omega_n/\sqrt{n}$ is negligible. However, as there are at most 2^k different possibilities for $|v_s|$, we see that the contribution of the regions where any one of the $|v_s|$'s is larger than $\omega_n/\sqrt{n}$ is negligible.

Next we show that the sequence ω_n can be chosen in such a way that for $|v_s| \leq \omega_n/\sqrt{n}$, we have

$$\hat{\rho}(v_s) = \exp\big(- 2\pi^2 v_s^2 + o(1/n)\big).$$

To see this, we use that ρ has a finite second moment, which implies that its Fourier transform is twice continuously differentiable. As a consequence $\hat{\rho}(v) = 1 - 2\pi^2(1 + o(1))v^2 = e^{-2\pi^2 v^2 (1+o(1))}$ as $v \to 0$. In other words, we can write $\hat{\rho}(v)$ in the form $\hat{\rho}(v) = e^{-2\pi^2 v^2 + g(v) v^2}$ where $g(v) \to 0$ as $v \to 0$. Choose a sequence which goes to zero as $n \to \infty$, say $\log n/\sqrt{n}$. We then define

$$\epsilon_n = \sup_{v : |v| \leq \log n/\sqrt{n}} |g(v)| \quad \text{and} \quad \omega_n = \min\big\{ \log n, \epsilon_n^{-1/3} \big\}.$$

For $|v| \leq \omega_n/\sqrt{n}$, we then have $|nv^2 g(v)| \leq \omega_n^2 \epsilon_n \leq \epsilon_n^{1/3}$, implying that $\hat{\rho}(v) = e^{-2\pi^2 v^2 + o(1/n)}$ as claimed. Furthermore, since $\epsilon_n \to 0$ as $n \to \infty$, we have $\omega_n \to \infty$ as required in the previous argument to restrict the region of integration of v_s to $|v_s| \leq \omega_n/\sqrt{n}$. Note finally that with the above choice, we have that $\lambda_n \omega_n^2 = o(\sqrt{n})$, a fact we will use in a moment.

Using the shorthand $n_{\boldsymbol{\delta}}$ for the quantity $n_{\boldsymbol{\delta}}(\boldsymbol{\sigma}^{(1)}, \ldots, \boldsymbol{\sigma}^{(k)})$, and defining $v_{\boldsymbol{\delta}}$ as $\sum_{j=1}^{k} \delta_j f_j$, we then rewrite

$$\prod_{s=1}^{n} \hat{\rho}(v_s) = \prod_{\boldsymbol{\delta}} \hat{\rho}(v_{\boldsymbol{\delta}})^{n_{\boldsymbol{\delta}}} = \exp\left(-2\pi^2 \sum_{\boldsymbol{\delta}} n_{\boldsymbol{\delta}} v_{\boldsymbol{\delta}}^2 + o(1) \right).$$

We would like to approximate the sum in the exponent by $f^2 = \sum_{j=1}^{k} f_j^2$. To this end, we first note that

$$\sum_{\boldsymbol{\delta} \in \{-1,+1\}^k} \sum_{j_1 j_2} \delta_{j_1} f_{j_1} \delta_{j_2} f_{j_2} = \sum_{\boldsymbol{\delta} \in \{-1,+1\}^k} \sum_{j} f_j^2 = 2^k \sum_{j} f_j^2 = 2^k f^2.$$

If $n_{\boldsymbol{\delta}}$ were *equal* to $2^{-k} n$ for all $\boldsymbol{\delta}$, the sum in the exponent would therefore be equal to nf^2, but for general $\boldsymbol{\delta}$ we get the bound

$$\left| \sum_{\boldsymbol{\delta}} n_{\boldsymbol{\delta}} v_{\boldsymbol{\delta}}^2 - nf^2 \right| = \left| \sum_{\boldsymbol{\delta}} (n_{\boldsymbol{\delta}} - 2^{-k} n) v_{\boldsymbol{\delta}}^2 \right| \leq \big(\max_{\boldsymbol{\delta}} |n_{\boldsymbol{\delta}} - 2^{-k} n| \big) \sum_{\boldsymbol{\delta}} v_{\boldsymbol{\delta}}^2. \tag{3.28}$$

Using the condition (3.21), and the fact that $|v_{\boldsymbol{\delta}}| \leq \omega_n/\sqrt{n}$, we bound the right hand side by $\lambda_n 2^k \omega_n^2/\sqrt{n} = o(1)$. We thus have shown that for $|v_s| \leq \omega_n/\sqrt{n}$,

$$\prod_{s=1}^{n} \hat{\rho}(v_s) = (1 + o(1)) \exp(-2\pi^2 n f^2).$$

Combining the bounds proven so far, we conclude that, up to an error which is negligible as $n \to \infty$, the expression in (3.23) is equal to

$$2^{nk} q_n^k \iiint \prod_{j=1}^{k} \exp\left(-2\pi^2 n f_j^2\right) \cos(2\pi f_j \alpha \sqrt{n}) \mathrm{d}f_j, \tag{3.29}$$

where the integral goes the region where $|v_s| \leq \omega_n/\sqrt{n}$ for all s. As, by an argument very similar to the argument leading to (3.24) and (3.27), the integral of $\prod_{j=1}^{k} \exp(-2\pi^2 n f_j^2)$ over a region in which $|v_s| > \omega_n/\sqrt{n}$ for at least one s is negligible, we therefore have shown that

$$\begin{aligned} &2^{nk}\mathbb{E}[I_k(\boldsymbol{\sigma}^{(1)},\ldots,\boldsymbol{\sigma}^{(k)})] \\ &\quad = 2^{nk} q_n^k \iiint_{-\infty}^{\infty} \prod_{j=1}^{k} \exp\left(-2\pi^2 n f_j^2\right) \cos(2\pi f_j t_n \sqrt{n}) \mathrm{d}f_j + o(1) \\ &\quad = \gamma^k + o(1), \end{aligned} \tag{3.30}$$

as desired. ∎

We need to estimate the contribution coming from terms that do not satisfy the conditions of Lemma 3.4. To this end we establish an upper bound (Lemma 3.6 below) on the expectation of $I_k(\boldsymbol{\sigma}^{(1)},\ldots,\boldsymbol{\sigma}^{(k)})$ that does not rely on the condition (3.21). To formulate this bound, we introduce the following notation.

Definition 3.5. *Let n_0 be such that $1/n_0 + 1/(1+\epsilon) = 1$ where ϵ is the constant from assumption* (2.1). *We say that the configurations $\boldsymbol{\sigma}^{(1)},\ldots,\boldsymbol{\sigma}^{(u)}$ has n_0-rank u_0 if the maximum number of linearly independent column vectors $\boldsymbol{\delta} \in \{-1,+1\}^u$ such that*

$$n_{\boldsymbol{\delta}}(\boldsymbol{\sigma}^{(1)},\ldots,\boldsymbol{\sigma}^{(u)}) \geq n_0 \tag{3.31}$$

is equal to u_0.

Lemma 3.6. *Given a positive integer u, there exists a constant $C_2 = C_2(u)$ such that for all sets of linearly independent row vectors $\boldsymbol{\sigma}^{(1)},\ldots,\boldsymbol{\sigma}^{(u)} \in \{-1,+1\}^n$ that have n_0-rank u_0, we have*

$$\left|\mathbb{E}\left[I_u(\boldsymbol{\sigma}^{(1)},\ldots,\boldsymbol{\sigma}^{(u)})\right]\right| \leq C_2 q_n^{u_0+(u-u_0)/n_0}. \tag{3.32}$$

Proof. Let $A_{\boldsymbol{\delta}} \subset \{1,\ldots,n\}$ be the set of indices i such that the column vector $(\sigma_i^{(1)},\ldots,\sigma_i^{(u)})$ is equal to $\boldsymbol{\delta}$, and let $\tilde{X}_{\boldsymbol{\delta}}$ be the random variable

$$\tilde{X}_{\boldsymbol{\delta}} = \sum_{i \in A_{\boldsymbol{\delta}}} X_i. \tag{3.33}$$

Recalling the definition (3.7), we then rewrite $I_u(\boldsymbol{\sigma}^{(1)},\ldots,\boldsymbol{\sigma}^{(u)})$ as

$$I_u(\boldsymbol{\sigma}^{(1)},\ldots,\boldsymbol{\sigma}^{(u)}) = 2^{-u} \sum_{\boldsymbol{\tau} \in \{-1,+1\}^u} \prod_{j=1}^{u} \operatorname{rect}\left(\frac{\sum_{\boldsymbol{\delta}\in\Delta} \delta_j \tilde{X}_{\boldsymbol{\delta}} - \tau_j t_n \sqrt{n}}{q_n}\right) \tag{3.34}$$

where Δ is the set of vectors $\boldsymbol{\delta} \in \{-1,+1\}$ such that $n_{\boldsymbol{\delta}} \geq 1$.

Choose u linearly independent vectors $\boldsymbol{\delta}^{(1)}, \ldots, \boldsymbol{\delta}^{(u)} \in \Delta$ such that the vectors $\boldsymbol{\delta}^{(1)}, \ldots, \boldsymbol{\delta}^{(u_0)}$ satisfy the condition (3.31). Let $\Delta_0 = \{\boldsymbol{\delta}^{(1)}, \ldots, \boldsymbol{\delta}^{(u_0)}\}$, and let $\Delta_u = \{\boldsymbol{\delta}^{(1)}, \ldots, \boldsymbol{\delta}^{(u)}\}$. Denoting the k-fold convolution of ρ with itself by ρ_k, we then write the expectation of a typical term on the right hand side of (3.34) as

$$\mathbb{E}\left[\prod_{j=1}^{u} \operatorname{rect}\left(\frac{\sum_{\delta\in\Delta} \delta_j \tilde{X}_{\delta} - \tau_j t_n \sqrt{n}}{q_n}\right)\right] = \iiint K_u(x_{\Delta\setminus\Delta_u}) \prod_{\delta\in\Delta\setminus\Delta_u} \rho_{n_\delta}(x_\delta) \mathrm{d}x_\delta \tag{3.35}$$

where $x_{\Delta\setminus\Delta_u}$ is a shorthand for the collection of variables x_δ, $\boldsymbol{\delta} \in \Delta \setminus \Delta_u$, and $K_u(x_{\Delta\setminus\Delta_u})$ is the integral

$$K_u(x_{\Delta\setminus\Delta_u}) = \iiint \prod_{j=1}^{u} \operatorname{rect}\left(\frac{\sum_{\delta\in\Delta} \delta_j x_\delta - \tau_j t_n \sqrt{n}}{q_n}\right) \prod_{\delta\in\Delta_u} \rho_{n_\delta}(x_\delta) \mathrm{d}x_\delta.$$

As $n_\delta \geq 1$ for all $\boldsymbol{\delta} \in \Delta$, we have

$$\iiint \prod_{\delta\in\Delta\setminus\Delta_u} \rho_{n_\delta}(x_\delta) \mathrm{d}x_\delta = 1,$$

so by (3.34) and (3.35), we have that

$$\left|\mathbb{E}\left[I_u(\boldsymbol{\sigma}^{(1)}, \ldots, \boldsymbol{\sigma}^{(u)})\right]\right| \leq \sup_{x_{\Delta\setminus\Delta_u} \in \mathbb{R}^{|\Delta\setminus\Delta_u|}} K_u(x_{\Delta\setminus\Delta_u}). \tag{3.36}$$

It is therefore enough to bound $K_u(x_{\Delta\setminus\Delta_u})$ uniformly in $x_{\Delta\setminus\Delta_u}$.

Let $\alpha_j = \tau_j t_n \sqrt{n} - \sum_{\delta\in\Delta\setminus\Delta_u} \delta_j x_\delta$. Noting that α_j does not depend on the variables which are integrated over in K_u, we then rewrite K_u as

$$K_u(x_{\Delta\setminus\Delta_u}) = \iiint \prod_{j=1}^{u} \operatorname{rect}\left(\frac{\sum_{\delta\in\Delta_u} \delta_j x_\delta - \alpha_j}{q_n}\right) \prod_{\delta\in\Delta_u} \rho_{n_\delta}(x_\delta) \mathrm{d}x_\delta.$$

Let $\tilde{M}$ be the matrix with matrix elements $\tilde{M}_{ji} = \delta_j^{(i)}$. The product of the rect-functions in the above integral then ensures that

$$\max_{j=1,\ldots,u} \left|\sum_{i=1}^{u} M_{ji} x_{\delta^{(i)}} - \alpha_j\right| \leq \frac{1}{2} q_n. \tag{3.37}$$

As the vectors in Δ_u are linearly independent, the matrix $\tilde{M}$ is invertible. Let $\beta_i = \sum_{j=1}^{u} (\tilde{M}^{-1})_{ij} \alpha_j$, and let $\|\tilde{M}^{-1}\|$ be the norm of $\tilde{M}^{-1}$ as an operator from ℓ_∞ to ℓ_∞. The bound (3.37) then implies that

$$\max_{i=1,\ldots,u} |x_{\delta^{(i)}} - \beta_i| \leq \frac{1}{2} \tilde{q}_n, \tag{3.38}$$

where $\tilde{q}_n = \|\tilde{M}^{-1}\| q_n$. As a consequence, the integral K_u is bounded by

$$\begin{aligned} K_u(x_{\Delta\setminus\Delta_u}) &\le \iiint \prod_{i=1}^{u} \operatorname{rect}\left(\frac{x_{\delta^{(i)}} - \beta_i}{\tilde{q}_n}\right) \prod_{\delta\in\Delta_u} \rho_{n_\delta}(x_\delta)\mathrm{d}x_\delta \\ &= \prod_{i=1}^{u} \int \operatorname{rect}\left(\frac{x-\beta_i}{\tilde{q}_n}\right) \rho_{n_i}(x)\mathrm{d}x \\ &= \tilde{q}_n^u \prod_{i=1}^{u} \int \hat{\rho}^{n_i}(f)\operatorname{sinc}(q_n f) e^{2\pi i\beta_i f}\mathrm{d}f, \end{aligned} \tag{3.39}$$

where we used the shorthand $n_i = n_\delta^{(i)}$. For $i = 1, \ldots, u_0$, we use that $n_i \ge n_0$ to bound the integral on the right by

$$\int \hat{\rho}^{n_i}(f)\operatorname{sinc}(q_n f)e^{2\pi i\beta_i f}\mathrm{d}f \le \int |\hat{\rho}^{n_0}(f)|\mathrm{d}f = C_0, \tag{3.40}$$

while for $i = u_0 + 1, \ldots u$, we use $n_i \ge 1$ and Hölder's inequality to obtain the bound

$$\begin{aligned} \int \hat{\rho}^{n_i}(f)\operatorname{sinc}(q_n f)e^{2\pi i\beta_i f}\mathrm{d}f &\le \int |\hat{\rho}(f)\operatorname{sinc}(\tilde{q}_n f)|\mathrm{d}f \\ &\le \left[\int |\hat{\rho}^{n_0}(f)|\mathrm{d}f\right]^{1/n_0} \left[\int |\operatorname{sinc}(\tilde{q}_n f)|^{1+\epsilon}\mathrm{d}f\right]^{1/(1+\epsilon)} \\ &= C_1 \tilde{q}_n^{-1/(1+\epsilon)}. \end{aligned} \tag{3.41}$$

Here

$$C_1 = \left[\int |\hat{\rho}^{n_0}(f)|\mathrm{d}f\right]^{1/n_0} \left[\int |\operatorname{sinc}(f)|^{1+\epsilon}\mathrm{d}f\right]^{1/(1+\epsilon)} < \infty \tag{3.42}$$

is independent of u, u_0 and n. Observing that $1 - 1/(1+\epsilon) = 1/n_0$, we thus get the bound

$$K_u(x_{\Delta\setminus\Delta_u}) \le \max(C_0, C_1)^u \tilde{q}_n^{u_0 + (u-u_0)/n_0}. \tag{3.43}$$

As there are only a finite number of choices for a set Δ of u linearly independent vectors in $\{-1, +1\}^u$, the ratio $\tilde{q}_n/q_n = \|\tilde{M}^{-1}\|$ is bounded by a constant that depends only on u, implying the existence of a constant $C_2 = C_2(u)$ such that

$$K_u(x_{\Delta\setminus\Delta_u}) \le C_2 q_n^{u_0 + (u-u_0)/n_0}. \tag{3.44}$$

Combined with (3.34) and (3.35), this proves the lemma. ∎

To bound the contribution in equation (3.9) coming from the terms where the vectors $\boldsymbol{\delta}^{(1)}, \ldots, \boldsymbol{\delta}^{(k)}$ have rank $u < k$ or do not satisfy condition (3.21), we need the following lemma, whose main statements were already proved in [7].

Lemma 3.7.

1. *Given $u \le k$ linearly independent row vectors $\boldsymbol{\sigma}^{(1)}, \ldots, \boldsymbol{\sigma}^{(u)}$, there are at most $2^{u(k-u)}$ ways to choose $\boldsymbol{\sigma}^{(u+1)}, \ldots, \boldsymbol{\sigma}^{(k)}$ such that the matrix M formed by the row vectors $\boldsymbol{\sigma}^{(1)}, \ldots, \boldsymbol{\sigma}^{(k)}$ has rank u.*

2. *Given u and n_0, there are constants $c_3 = c_3(u, n_0)$ and $C_3 = C_3(u, n_0)$ such that there are at most $C_3 n^{c_3} 2^{nu_0}$ ways to choose u linearly independent configurations $\boldsymbol{\sigma}^{(1)}, \dots, \boldsymbol{\sigma}^{(u)}$ that have n_0-rank u_0.*
3. *Let $u < \infty$, let $C_4 = C_4(u) = 2^{u+1}$, and let λ_n be a sequence of positive number such that $\lambda_n/\sqrt{n} \to 0$ as $n \to \infty$. Then the number of configurations $\boldsymbol{\sigma}^{(1)}, \dots, \boldsymbol{\sigma}^{(u)}$ that violate condition* (3.21) *is bounded by $C_4 2^{nu} e^{-\frac{1}{2}\lambda_n^2}$.*
4. *Let $\boldsymbol{\sigma}^{(1)}, \dots, \boldsymbol{\sigma}^{(k)}$ be distinct spin configurations, assume that rank $M < k$, and let $\boldsymbol{\sigma}^{(1)}, \dots, \boldsymbol{\sigma}^{(u)}$ be linearly independent. Then $n_{\boldsymbol{\delta}}(\boldsymbol{\sigma}^{(1)}, \dots, \boldsymbol{\sigma}^{(u)}) = 0$ for at least one $\boldsymbol{\delta} \in \{-1, +1\}^u$, implying in particular that, for n sufficiently large, $\boldsymbol{\sigma}^{(1)}, \dots, \boldsymbol{\sigma}^{(u)}$ violate condition* (3.21).
5. *Given u, let $\boldsymbol{\sigma}^{(1)}, \dots, \boldsymbol{\sigma}^{(u)} \in \{-1, +1\}^n$ be an arbitrary set of row vectors satisfying* (3.21). *Then*

$$q(\boldsymbol{\sigma}^{(a)}, \boldsymbol{\sigma}^{(b)}) \leq 2^u \frac{\lambda_n}{\sqrt{n}} \tag{3.45}$$

whenever $a \neq b$. For n sufficiently large, condition (3.21) *therefore implies that $\boldsymbol{\sigma}^{(1)}, \dots, \boldsymbol{\sigma}^{(u)}$ are linearly independent.*

Proof. Except for the second statement, the lemma mainly summarizes the relevant results from Section 6 of [7]. More explicitly: statement (1) is proved in the paragraph following (6.10), and for $\lambda = \log n$, statement (3) is proved in the paragraphs around (6.12), statement (4) is proved in the paragraph around the second and third unnumbered equation after (6.12), and statement (5) is equivalent to the bound (6.14).

It is not hard to see that the arguments in Section 6 of [7] can be generalized to arbitrary sequences λ_n of positive number, as long as $\lambda_n/\sqrt{n} \to 0$ as $n \to \infty$. Indeed, starting with statement (3), let us consider n independent trials with 2^u equally likely outcomes, and use Chebychev's inequality to bound the probability that $|n_{\boldsymbol{\delta}}(\boldsymbol{\sigma}^{(1)}, \dots, \boldsymbol{\sigma}^{(u)}) - n2^{-u}| \geq \sqrt{n}\lambda_n$ by $2e^{-\frac{1}{2}\lambda_n^2}$. Combined with the union bound for the 2^u different random variables $n_{\boldsymbol{\delta}}(\boldsymbol{\sigma}^{(1)}, \dots, \boldsymbol{\sigma}^{(u)})$, $\boldsymbol{\delta} \in \{-1, +1\}^u$, this gives statement (3). The first part of (4) does not involve the value of λ_n, and the second follows from the first whenever $\lambda n/\sqrt{n} \to 0$ as $n \to \infty$. To prove the bound (3.45) in statement (5), we rewrite the overlap $q(\boldsymbol{\sigma}^{(a)}, \boldsymbol{\sigma}^{(b)})$ as

$$q(\boldsymbol{\sigma}^{(a)}, \boldsymbol{\sigma}^{(b)}) = \frac{1}{n} \left(\sum_{\substack{\boldsymbol{\delta} \in \{-1,+1\}^u \\ \delta_a = \delta_b}} n_{\boldsymbol{\delta}}(\boldsymbol{\sigma}^{(1)}, \dots, \boldsymbol{\sigma}^{(u)}) - \sum_{\substack{\boldsymbol{\delta} \in \{-1,+1\}^u \\ \delta_a \neq \delta_b}} n_{\boldsymbol{\delta}}(\boldsymbol{\sigma}^{(1)}, \dots, \boldsymbol{\sigma}^{(u)}) \right). \tag{3.46}$$

Noting that each sum contains 2^{u-1} terms, we see that the bound (3.21) implies the bound (3.45). Finally, the last statement of (5) is a direct consequence of (3.45) and the fact that $\lambda_n/\sqrt{n} \to 0$ as $n \to \infty$.

We are left with the proof of (2). To this end, let us consider the matrix $\tilde{M}_u$ obtained from M_u by omitting all columns $\sigma_i^{(1)}, \dots, \sigma_i^{(u)}$ that are equal to a vector $\boldsymbol{\delta} \in \{-1, +1\}^u$ with $n_{\boldsymbol{\delta}}(\boldsymbol{\sigma}^{(1)}, \dots, \boldsymbol{\sigma}^{(u)}) < n_0$. Note that the number of columns n' of $\tilde{M}_u$ is at least $n - 2^u n_0$ and at most n. Fixing n', for the moment, and noting that the rank of $\tilde{M}_u$ is u_0, we now use statement (1) to conclude that there are at most

$$\binom{u}{u_0} 2^{n' u_0} 2^{u_0(u-u_0)} \leq 2^{u+u^2/2} 2^{nu_0} \tag{3.47}$$

ways to choose $\tilde{M}_u$. Given $\tilde{M}_u$ we need to insert $n - n'$ columns in $\{-1, +1\}^u$ to obtain the matrix M_u. Including the number of choices for the positions of these $n - n'$ columns, this gives an extra factor of

$$\binom{n}{n-n'} 2^{(n-n')u} \leq \frac{1}{(n-n')!} n^{n-n'} 2^{n-n'} \leq \frac{1}{(n-n')!} n^{n_0 2^u} 2^{n_0 2^u}. \tag{3.48}$$

Combining the two factors and summing over $n' \in \{n - n_0 2^u, \dots, n\}$, we get a bound of the form $C_3 n^{c_3} 2^{u_0 n}$ where C_3 and c_3 depend only on u and n_0. ∎

Having Lemmas 3.3, 3.4, 3.6 and 3.7 in hand, we are now ready to prove Theorem 3.1.

3.3.1. Proof of Theorem 3.1. We start with the case $m = 1$, i.e., the one-dimensional factorial moment $\mathbb{E}[(Z_n(a_n, b_n))_k]$. As in Lemma 3.4, we choose $\lambda_n = \log n$. Consider the sum over all sequences $\boldsymbol{\sigma}^{(1)}, \dots, \boldsymbol{\sigma}^{(k)} \in \{-1, +1\}^n$ that satisfy the bound (3.21). By Lemma 3.7 (3), this sum contains $2^{nk}(1 + o(1))$ terms, and by Lemma 3.7 (5), the matrix formed by the row vectors $\boldsymbol{\sigma}^{(1)}, \dots, \boldsymbol{\sigma}^{(k)}$ has rank k if n is large enough. With the help of Lemma 3.4, we conclude that the sum over all these terms gives a contribution to the k^{th} factorial moment which is equal to $\gamma^k + o(1)$.

To prove Theorem 3.1, we have to bound the contribution of the remaining terms. To this end, we group the remaining terms in the sum (3.9) into four classes:

1. Sequences of distinct configurations $\boldsymbol{\sigma}^{(1)}, \dots, \boldsymbol{\sigma}^{(k)}$ of rank k and n_0-rank $u_0 < k$ that violate the condition (3.21);
2. Sequences of distinct configurations $\boldsymbol{\sigma}^{(1)}, \dots, \boldsymbol{\sigma}^{(k)}$ of rank k and n_0-rank k that violate the condition (3.21);
3. Sequences of distinct configurations $\boldsymbol{\sigma}^{(1)}, \dots, \boldsymbol{\sigma}^{(k)}$ of rank $u < k$ such that there is a subsequence of linearly independent configurations $\tilde{\boldsymbol{\sigma}}^{(1)}, \dots, \tilde{\boldsymbol{\sigma}}^{(u)}$ of n_0-rank $u_0 < u$;
4. Sequences of distinct configurations $\boldsymbol{\sigma}^{(1)}, \dots, \boldsymbol{\sigma}^{(k)}$ of rank $u < k$ such that all subsequences of linearly independent configurations $\tilde{\boldsymbol{\sigma}}^{(1)}, \dots, \tilde{\boldsymbol{\sigma}}^{(u)}$ have of n_0-rank $u_0 = u$;

By Lemma 3.7 (4), the configurations $\tilde{\boldsymbol{\sigma}}^{(1)}, \dots, \tilde{\boldsymbol{\sigma}}^{(u)}$ in class (4) must violate condition (3.21). Relaxing the constraint that the configurations in class (1) violate condition (3.21), it is therefore enough to bound the following two error terms:

- The sum $R^{<}_{n,k}$ of all sequences of configurations $\boldsymbol{\sigma}^{(1)}, \dots, \boldsymbol{\sigma}^{(k)}$ of rank $u \leq k$ containing a subsequence of linearly independent configurations $\tilde{\boldsymbol{\sigma}}^{(1)}, \dots, \tilde{\boldsymbol{\sigma}}^{(u)}$ of n_0-rank $u_0 < u$, and
- The sum $R^{=}_{n,k}$ of all sequences of configurations $\boldsymbol{\sigma}^{(1)}, \dots, \boldsymbol{\sigma}^{(k)}$ of rank $u \leq k$ such that all subsequence of linearly independent configurations $\tilde{\boldsymbol{\sigma}}^{(1)}, \dots, \tilde{\boldsymbol{\sigma}}^{(u)}$ obey condition (3.21) and have n_0-rank $u_0 = u$.

Before bounding these two error terms, we note that

$$I_k(\boldsymbol{\sigma}^{(1)}, \dots, \boldsymbol{\sigma}^{(k)}) = \prod_{i=1}^{k} I(\boldsymbol{\sigma}^{(i)}) \leq \prod_{i=1}^{u} I(\tilde{\boldsymbol{\sigma}}^{(i)}) = I_u(\tilde{\boldsymbol{\sigma}}^{(1)}, \dots, \tilde{\boldsymbol{\sigma}}^{(u)}), \tag{3.49}$$

implying that

$$\mathbb{E}\left[I_k(\boldsymbol{\sigma}^{(1)},\dots,\boldsymbol{\sigma}^{(k)})\right] \le \mathbb{E}\left[I_u(\tilde{\boldsymbol{\sigma}}^{(1)},\dots,\tilde{\boldsymbol{\sigma}}^{(u)})\right] \tag{3.50}$$

whenever $\tilde{\boldsymbol{\sigma}}^{(1)},\dots,\tilde{\boldsymbol{\sigma}}^{(u)}$ is a subsequence of $\boldsymbol{\sigma}^{(1)},\dots,\boldsymbol{\sigma}^{(k)}$.

To bound $R^{<}_{n,k}$, we now use (3.50) and Lemma 3.6 to bound the expectation of $I_k(\boldsymbol{\sigma}^{(1)},\dots,\boldsymbol{\sigma}^{(k)})$ by $C_5 q_n^{u_0} q_n^{(u-u_0)/n_0}$, where $C_5 = \max_{u\le k} C_2(u)$. Using Lemma 3.7 (2) to bound the number of sequences $\tilde{\boldsymbol{\sigma}}^{(1)},\dots,\tilde{\boldsymbol{\sigma}}^{(u)}$ of n_0-rank u_0 by $C_6 n^{c_6} 2^{nu_0}$, where $C_6 = \max_{u\le k} C_3(u,n_0)$ and $c_6 = \max_{u\le k} c_3(u,n_0)$, and Lemma 3.7 (1) to bound the number of ways $\boldsymbol{\sigma}^{(1)},\dots,\boldsymbol{\sigma}^{(k)}$ can be obtained from $\tilde{\boldsymbol{\sigma}}^{(1)},\dots,\tilde{\boldsymbol{\sigma}}^{(u)}$, we obtain the following upper bound

$$R^{<}_{n,k} \le C_6 C_5 n^{c_6} \sum_{\substack{u_0,u:\\ u_0<u\le k}} \binom{k}{u} 2^{u(k-u)} (2^n q_n)^{u_0} q_n^{(u-u_0)/n_0} \tag{3.51}$$

Substituting $q_n = \gamma\sqrt{n}\xi_n$, and using that $\xi_n \le 1$ and $\sqrt{n}\xi_n 2^n \ge 1$, this gives

$$\begin{aligned} R^{<}_{n,k} &\le C_6 C_5 n^{c_7} (2^n\xi_n)^k \xi_n^{1/n_0} \sum_{\substack{u_0,u:\\ u_0<u\le k}} \binom{k}{u} 2^{u(k-u)} \gamma^{u_0+(u-u_0)/n_0} \\ &= C_7 n^{c_7} (2^n\xi_n)^k \xi_n^{1/n_0} \end{aligned} \tag{3.52}$$

where $c_7 = c_6 + k(1/2 + 1/2n_0)$ and $C_7 = C_7(\gamma,k)$ is a constant that depends on γ and k. Observing that $2^n\xi_n = \sqrt{2\pi}\, e^{\alpha^2/2}$ is bounded uniformly in n, while ξ_n falls exponentially with n, we conclude that $R^{<}_{n,k} = o(1)$ as $n\to\infty$.

The error term $R^{=}_{n,k}$ can be bounded in a similar way. We again use (3.50) and Lemma 3.6 to bound the expectation of $I_k(\boldsymbol{\sigma}^{(1)},\dots,\boldsymbol{\sigma}^{(k)})$, but now we use part (3) of Lemma 3.7 to bound the number of sequences $\tilde{\boldsymbol{\sigma}}^{(1)},\dots,\tilde{\boldsymbol{\sigma}}^{(u)}$. Using again Lemma 3.7 (1) to bound the number of ways $\boldsymbol{\sigma}^{(1)},\dots,\boldsymbol{\sigma}^{(k)}$ can be obtained from $\tilde{\boldsymbol{\sigma}}^{(1)},\dots,\tilde{\boldsymbol{\sigma}}^{(u)}$, we now obtain the upper bound

$$\begin{aligned} R^{=}_{n,k} &\le C_5 C_8 \sum_{\substack{u_0,u:\\ u_0<u\le k}} \binom{k}{u} 2^{u(k-u)} e^{-\lambda_n^2/2} (2^n q_n)^u \\ &\le C_5 C_8 n^{k/2} (2^n\xi_n)^k e^{-\lambda_n^2/2} \sum_{\substack{u_0,u:\\ u_0<u\le k}} \binom{k}{u} 2^{u(k-u)} \gamma^u \\ &= C_9 n^{k/2} (2^n\xi_n)^k e^{-\lambda_n^2/2} \end{aligned} \tag{3.53}$$

where $C_8 = \max_{u\le k} C_4(u) = 2^{k+1}$ and $C_9 = C_9(\gamma,k)$ is a constant that depends only on γ and k. As $e^{-\lambda_n^2/2} = e^{-\log^2 n/2}$ decays faster than any power of n, we get that $R^{=}_{n,k} = o(1)$ as $n\to\infty$, as desired.

This completes the proof that $\mathbb{E}[(Z_n(a_n,b_n))_k] \to \gamma^k$ as $n\to\infty$. To prove the convergence of the higher-dimensional factorial moments, we need to generalize Lemmas 3.3, 3.4, and 3.6. But, except for notational inconveniences, this causes no problems. Indeed, comparing the representations (3.7) and (3.10), we see that the only difference is the appearance of several distinct intervals $[a_n^{\ell(j)}, b_n^{\ell(j)}]$ for the energy of the configuration $\boldsymbol{\sigma}^{(j)}$, instead of the same interval $[a_n,b_n]$ for all of them.

As a consequence, the statement of Lemma 3.3 has to be modified, with the right hand side of (3.14) replaced by

$$\prod_{\ell=1}^{k} q_{n,\ell}^{k_\ell} \iiint_{-\infty}^{\infty} \prod_{s=1}^{n} \hat{\rho}(v_s) \prod_{j=1}^{k} \operatorname{sinc}(f_j q_{n,\ell(j)}) \cos\left(2\pi f_j t_n^{\ell(j)} \sqrt{n}\right) \mathrm{d}f_j. \tag{3.54}$$

But the proof remains unchanged, as it never used that $q_{n,\ell(j)}$ or $t_n^{\ell(j)}$ is constant.

In a similar way, the proof of Lemma 3.4 needs only notational changes: the arguments leading to (3.26) now give a prefactor $2^{nk}\prod_j q_{\ell(j)}$ instead of $2^{nk}q_n^k$, but the integral multiplying this prefactor (and therefore the rest of the proof) remains unchanged, proving that under the conditions of Lemma 3.4,

$$2^{nk}\mathbb{E}\left[I_k(\boldsymbol{\sigma}^{(1)},\ldots,\boldsymbol{\sigma}^{(k)})\right] = \prod_j \gamma_{\ell(j)} + o(1). \tag{3.55}$$

Turning finally to the proof of Lemma 3.6, we note that its proof goes through if we replace $\tilde{q}_n$ by $\|\tilde{M}^{-1}\| \max_{\ell=1,\ldots,m} q_{n,\ell}$. As a consequence, the bound (3.32) has to be modified to

$$\left|\mathbb{E}\left[I_u(\boldsymbol{\sigma}^{(1)},\ldots,\boldsymbol{\sigma}^{(u)})\right]\right| \leq C_2(u)(\max_{\ell=1,\ldots,m} q_{n,\ell})^{u_0+(u-u_0)/n_0}, \tag{3.56}$$

which does not change the n-dependence of the bound.

Using these generalizations of Lemmas 3.3, 3.4, and 3.6, it is easy to see that the bounds (3.52) and (3.53) remain unchanged, except for the fact that the constants C_7 and C_9 now depend on $\max_\ell \gamma_\ell$ instead of γ. This completes the convergence proof for the multi-dimensional factorial moments, and hence the proof of Theorem 3.1.

3.4. Overlap Estimates

To complete the proof of the Theorem 2.2, we need to show that the rescaled overlaps converge to a standard normal. Defining $R(\beta)$ to be the tail of the standard Gaussian,

$$R(\beta) = \int_{\beta}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \mathrm{d}x,$$

we therefore have to show that for any $\beta \in R$ and any $j > i > 0$, we have

$$\mathbb{P}(Q_{ij} \geq \beta) \to R(\beta) \tag{3.57}$$

as $n \to \infty$.

Let E_{r_n+i} and E_{r_n+j} be the ith and jth energy above α, respectively, and let $\lambda_0 > 0$. Having established the convergence of the rescaled energies (2.3), we note that the probability that both E_{r_n+i} and E_{r_n+j} fall into the interval $[\alpha, \alpha+\lambda_0\xi_n]$ can be made arbitrary close to one by choosing λ_0 and n large enough. Consider further a discretization scale η such that λ_0/η is an integer. If both E_{r_n+i} and E_{r_n+j} fall into the interval $[\alpha, \alpha+\lambda_0\xi_n]$, each of them must fall into one of the λ_0/η intervals $[\alpha, \alpha+\eta\xi_n]$, $[\alpha+\eta\xi_n, \alpha+2\eta\xi_n], \ldots, [\alpha+(\gamma_0-\eta)\xi_n, \alpha+\lambda_0\xi_n]$. By choosing η sufficiently small and n sufficiently large, the probability that both fall into the same interval, or that one of the other energies between α and $\alpha+\lambda_0\xi_n$ falls into the same interval as E_{r_n+i} or E_{r_n+j}, can be made arbitrarily close to one as well.

It is therefore enough to consider the intersection of the event $Q_{ij} \geq \beta$, the event that E_{r_n+i} and E_{r_n+j} fall into two different intervals of the form $[\alpha+(m-1)\eta\xi_n, \alpha+m\eta\xi_n]$,

$m = 1, \ldots, \lambda_0/\eta$, and the event that both of them are the only energies that fall into these intervals. Denote the intersection of these events by $A_{ij}(\beta)$. Decomposing the event $A_{ij}(\beta)$ according to the spin configurations $\boldsymbol{\sigma}^{(r_n+i)}$ and $\boldsymbol{\sigma}^{(r_n+j)}$ corresponding to the ith and jth energy above α and the particular intervals containing these energies, we then rewrite the probability of $A_{ij}(\beta)$ as

$$\mathbb{P}(A_{ij}(\beta)) = \sum_{m_i<m_j} \sum_{\boldsymbol{\sigma},\tilde{\boldsymbol{\sigma}}}^{(\beta)} \mathbb{P}\big[A_{m_i}(\boldsymbol{\sigma}) \cap A_{m_j}(\tilde{\boldsymbol{\sigma}}) \cap \big\{Z_n^{(1)} = i-1\big\} \cap \big\{Z_n^{(2)} = 1\big\} \\ \cap \big\{Z_n^{(3)} = j-i-1\big\} \cap \big\{Z_n^{(4)} = 1\big\}\big]. \quad (3.58)$$

Here the second sum runs over pairs of distinct configurations $\boldsymbol{\sigma}, \tilde{\boldsymbol{\sigma}}$ with rescaled overlap larger than β, the first sum runs over integers m_i, m_j with $0 < m_i < m_j \leq \lambda_0/\eta$, the symbol $A_m(\boldsymbol{\sigma})$ denotes the event that the energy of the configuration $\boldsymbol{\sigma}$ falls into the interval $[\alpha+(m-1)\eta\xi_n, \alpha+m\eta\xi_n]$, and the random variables $Z_n^{(\ell)}$ are equal to the number of points in the spectrum that lie in the intervals $[a_n^{(\ell)}, b_n^{(\ell)}]$ where $a_n^{(1)} = \alpha$, $b_n^{(1)} = a_n^{(2)} = \alpha+(m_1-1)\eta\xi_n$, $b_n^{(2)} = a_n^{(3)} = \alpha + m_1\eta\xi_n$, $b_n^{(3)} = a_n^{(4)} = \alpha + (m_2-1)\eta\xi_n$, and $b_n^{(4)} = \alpha + m_2\eta\xi_n$. Defining $I^{(\ell)}(\cdot)$ as before, let

$$(Z_n)_2^{(\beta)} = \sum_{\boldsymbol{\sigma},\tilde{\boldsymbol{\sigma}}}^{(\beta)} I^{(2)}(\boldsymbol{\sigma}) I^{(4)}(\tilde{\boldsymbol{\sigma}}) \quad (3.59)$$

be the number of distinct pairs of configurations $\boldsymbol{\sigma}, \tilde{\boldsymbol{\sigma}}$ with rescaled overlap at least β such that the energy of $\boldsymbol{\sigma}$ falls into the interval $[a_n^{(2)}, b_n^{(2)}]$, and the energy of $\tilde{\boldsymbol{\sigma}}$ falls into the interval $[a_n^{(4)}, b_n^{(4)}]$. We then rewrite the probability $\mathbb{P}(A_{ij}(\beta))$ as

$$\mathbb{P}(A_{ij}(\beta)) = \sum_{m_i<m_j} \mathbb{E}\big[(Z_n)_2^{(\beta)} \mathbb{I}\big(Z_n^{(1)} = i-1\big)\mathbb{I}\big(Z_n^{(2)} = 1\big) \\ \times \mathbb{I}\big(Z_n^{(3)} = j-i-1\big)\mathbb{I}\big(Z_n^{(4)} = 1\big)\big], \quad (3.60)$$

where $\mathbb{I}(A)$ denotes the indicator function of the event A.

Let $N_n(\beta)$ be the number of distinct pairs $\boldsymbol{\sigma}, \tilde{\boldsymbol{\sigma}}$ with rescaled overlap at least β. Combining the methods of the last section with the standard central limit theorem, we now easily establish that

$$\mathbb{E}\big[(Z_n)_2^{(\beta)}\big] = \eta^2 2^{-2n} N_n(\beta)(1+o(1)) = \eta^2 R(\beta)(1+o(1)). \quad (3.61)$$

To analyze the right hand side of (3.60), we would like first to factor the expectation on the right hand side, and then use (3.61) and Poisson convergence of the random variables $Z_n^{(\ell)}$ to analyze the resulting terms. In the process, we will have to analyze the factorial moments

$$\mathbb{E}\big[(Z_n)_2^{(\beta)} \big(Z_n^{(1)}\big)_{k_1} \big(Z_n^{(2)}\big)_{k_2} \big(Z_n^{(3)}\big)_{k_3} \big(Z_n^{(4)}\big)_{k_4}\big]. \quad (3.62)$$

Unfortunately, the methods of the last section cannot be directly applied to these factorial moments since the sum over configurations representing the above expression is not a sum over pairwise distinct configurations: comparing, e.g., the sum over $\boldsymbol{\sigma}$ in (3.59) and the representation of the random variable $Z_n^{(2)}$ as a sum over configurations,

$$Z_n^{(2)} = \sum_{\boldsymbol{\sigma}'} I^{(2)}(\boldsymbol{\sigma}'), \quad (3.63)$$

we see that both involve configurations whose energy lies in the interval $[a_n^{(2)}, b_n^{(2)}]$. But this problem can be easily overcome by considering the random variables $Z_n^{(2)} - 1$ and $Z_n^{(4)} - 1$ instead of $Z_n^{(2)}$ and $Z_n^{(4)}$. We therefore consider the expression

$$\mathbb{E}\big[(Z_n)_2^{(\beta)}\big(Z_n^{(1)}\big)_{k_1}\big(Z_n^{(2)}-1\big)_{k_2}\big(Z_n^{(3)}\big)_{k_3}\big(Z_n^{(4)}-1\big)_{k_4}\big] \\ = \sum_{\sigma,\tilde{\sigma}}^{(\beta)} \mathbb{E}\big[I^{(2)}(\boldsymbol{\sigma})I^{(4)}(\tilde{\boldsymbol{\sigma}})\big(Z_n^{(1)}\big)_{k_1}\big(Z_n^{(2)}-1\big)_{k_2}\big(Z_n^{(3)}\big)_{k_3}\big(Z_n^{(4)}-1\big)_{k_4}\big]. \quad (3.64)$$

We claim that this expression can again be be expressed as a double sum over distinct configurations, allowing us to apply the methods of the last section. Indeed, let us first consider the product

$$I^{(2)}(\boldsymbol{\sigma})\big(Z_n^{(2)}-1\big)_{k_2} = I^{(2)}(\boldsymbol{\sigma})\big(Z_n^{(2)}-1\big)\big(Z_n^{(2)}-2\big)\cdots\big(Z_n^{(2)}-k_2\big). \quad (3.65)$$

Proceeding as in the proof of (3.9), we now rewrite this product as a sum of configurations $\boldsymbol{\sigma}^{(1)}, \ldots, \boldsymbol{\sigma}^{(k_2)}$ which are mutually distinct and distinct from $\boldsymbol{\sigma}$. In a similar way, the product $I^{(4)}(\tilde{\boldsymbol{\sigma}})(Z_n^{(4)}-1)_{k_4}$ can be expressed as a sum over mutually distinct configurations which are distinct from $\tilde{\boldsymbol{\sigma}}$. Using these facts, we now proceed as before to obtain the bound

$$\mathbb{E}\big[(Z_n)_2^{(\beta)}\big(Z_n^{(1)}\big)_{k_1}\big(Z_n^{(2)}-1\big)_{k_2}\big(Z_n^{(3)}\big)_{k_3}\big(Z_n^{(4)}-1\big)_{k_4}\big] \\ = \eta^2\gamma_1^{k_1}\eta^{k_2}\gamma_3^{k_3}\eta^{k_4}2^{-2n}N_n(\beta)(1+o(1)), \quad (3.66)$$

where $\gamma_1 = \eta(m_1-1)$ and $\gamma_3 = \eta(m_2-m_1-1)$.

Consider the four random variables $Z_n^{(1)}$, $Z_n^{(2)}-1$, $Z_n^{(3)}$ and $Z_n^{(4)}-1$, together with the probability distribution μ defined by

$$\mu\big(Z_n^{(1)}=i_1, Z_n^{(2)}-1=i_2, Z_n^{(3)}=i_3, Z_n^{(4)}-1=i_4\big) \\ = \frac{\mathbb{E}\big[(Z_n)_2^{(\beta)}\mathbb{I}\big(Z_n^{(1)}=i_1\big)\mathbb{I}\big(Z_n^{(2)}-1=i_2\big)\mathbb{I}\big(Z_n^{(3)}=i_3\big)\mathbb{I}\big(Z_n^{(4)}-1=i_4\big)\big]}{\mathbb{E}\big[(Z_n)_2^{(\beta)}\big]}. \quad (3.67)$$

The bounds (3.61) and (3.66) then establish that in the measure μ, the four random variables $Z_n^{(1)}, Z_n^{(2)}-1, Z_n^{(3)}$ and $Z_n^{(4)}-1$ converge to four independent Poisson random variables with rates η, γ_2, η, and γ_4, respectively. Using once more the bound (3.61), we conclude that the expectation in the sum in (3.60) can be approximated as

$$\mathbb{E}\big[(Z_n)_2^{(\beta)}\mathbb{I}\big(Z_n^{(1)}=i-1\big)\mathbb{I}\big(Z_n^{(2)}-1=0\big)\mathbb{I}\big(Z_n^{(3)}=j-i-1\big)\mathbb{I}\big(Z_n^{(4)}-1=0\big)\big] \\ = \eta^2\frac{\gamma_1^{i-1}}{(i-1)!}\frac{\gamma_3^{j-i-1}}{(j-i-1)!}e^{-(2\eta+\gamma_1+\gamma_3)}R(\beta)(1+o(1)). \quad (3.68)$$

Inserted into (3.60) this gives the bound

$$\mathbb{P}(A_{ij}(\beta)) = K_\eta(\lambda_0)R(\beta)(1+o(1)), \quad (3.69)$$

where

$$K_\eta(\lambda_0) = \eta^2\sum_{m_1<m_2}\frac{((m_1-1)\eta)^{i-1}}{(i-1)!}\frac{((m_2-m_1-1)\eta)^{j-i-1}}{(j-i-1)!}e^{-\eta m_2} \quad (3.70)$$

is the Riemann-sum approximation to the integral

$$K(\lambda_0) = \int_0^{\lambda_0} d\gamma_1 \frac{\gamma_1^{i-1}}{(i-1)!} e^{-\gamma_1} \int_0^{\lambda_0} d\gamma_3 \frac{\gamma_3^{j-i-1}}{(j-i-1)!} e^{-\gamma_3}. \tag{3.71}$$

As $\eta \to 0$, the Riemann sum $K_\eta(\lambda_0)$ converges to the integral $K(\lambda_0)$, and as $\lambda_0 \to \infty$, the integral $K(\lambda_0)$ converges to 1. Choosing first λ_0 large enough, then η small enough, and then n large enough, the normal distribution function $R(\beta)$ is therefore an arbitrarily good approximation to $\mathbb{P}(A_{ij}(\beta))$, which in turn can be made arbitrary close to $\mathbb{P}(Q_{ij} \geq \beta)$, again by first choosing λ_0 sufficiently large, then η sufficiently small, and then n sufficiently large. This establishes (2.4) and hence the remaining statements of Theorem 2.2.

4. GENERALIZATIONS AND OPEN PROBLEMS

4.1. Generalizations of the NPP

The NPP has a natural generalization: Divide a set $\{X_1, X_2, \ldots, X_n\}$ of numbers into q subsets such that the sums in all q subsets are as equal as possible. This is known as multi-way partitioning or multiprocessor scheduling problem [4]. The latter name refers to the problem of distributing n tasks with running times $\{X_1, X_2, \ldots, X_n\}$ on q processors of a parallel computer such that the overall running time is minimized. Bovier and Kurkova [9] considered the restricted multi-way partitioning problem where the cardinality of each subset is fixed to n/q. For this model they could prove the "energy part" of the local REM hypothesis at $\alpha = 0$, i.e., the convergence of the properly scaled near optimal solutions to a Poisson point process. The local REM (including the "overlap part") is conjectured to be valid for all $\alpha \geq 0$ for the multi-way partitioning problem in the unrestricted case (i.e., for n/q not necessarily fixed) [2]. This generalization is still open.

4.2. Universality

In [3] it is conjectured that the local REM is a property of discrete, disordered systems well beyond number partitioning and its relatives. Since this conjecture represents a fascinating open problem for the rigorous community, we briefly review the heuristic argument of [3]: Consider a model with an energy function of the form

$$E(\boldsymbol{\sigma}) = \sum_{i=1}^{n} \sigma_i X_i, \tag{4.1}$$

where the $\boldsymbol{\sigma}$ is an n-dimensional vector with binary entries $\sigma_i = \pm 1$ or $\sigma_i \in \{0, 1\}$ and the X_i are real random numbers from the unit interval. In case of the NPP (or the 1-d Edwards-Anderson model), any vector $\boldsymbol{\sigma}$ is a feasible configuration. If we add more restrictions, we could write the cost function of many optimization problems in the form (4.1). For example, in the traveling salesman problem, we would take $\sigma_i \in \{0, 1\}$, where $\sigma_i = 1$ means that the distance X_i is part of the tour, and the σ_i would have to fulfill the constraint to encode a valid itinerary. In higher-dimensional spin glasses, the $\sigma_i = \pm 1$ encode satisfied or unsatisfied edges, and are correlated due to loops in the graph. In all cases we have an exponential number of valid configurations, with an exponential number of energy values $E(\boldsymbol{\sigma})$. Since the range of energies scales only linearly with n, it should follow that adjacent levels will be separated by exponentially small distances. The *precise* value of each gap will be determined by the *least significant bits* in the X_i's, however. The dynamical variables σ_i can only control the n most significant bits of the energy. Ref. 3 argue that the residual entropy of the least significant bits then gives rise to the Poisson nature of adjacent energy levels and to the full local REM property. This very heuristic argument has been supported by extensive numerical simulations in various spin glass models (Edwards-Anderson model, Sherrington-Kirkpatrick model, Potts glasses) and in optimization problems (TSP, minimum spanning tree, shortest path) [3].

In this article, we rigorously established the local REM conjecture for a particular model, the NPP. In a recent article [8], submitted shortly after the present one, Bovier and Kurkova showed that the local REM conjecture holds for many types of spin glasses as well, in particular the Edwards-Anderson model and the Sherrington-Kirkpatrick model. Their approach is based on a general theorem establishing Poisson convergence for an abstract class of models, with *conditions* that are very similar to the *statements* of our Lemmas 3.4 and 3.7 in an abstract setting.

4.3. Phase Transition

According to the heuristic argument above, the bit-entropy of the disorder X_i is the essential property that leads to the local REM: if it is larger than the entropy of the configurations, the local REM should apply. If it is lower than the configurational entropy, the distances between adjacent energy levels are multiples of a fixed, smallest distance. In this case, each energy level is populated by an exponential number of configurations. An indicator for the transition between the two regimes is the maximum overlap between two configurations with adjacent energy levels. If the entropy of the disorder is larger than the configurational entropy, this overlap should be 0 (the local REM). If the entropy of the disorder is much smaller than the configurational entropy, this overlap should be $1 - \Theta(n^{-1})$. Numerical simulations in [3] indicate that there is a sharp transition at the point at which these entropies are the same. A canonical problem in which such a transition has been rigorously investigated is the phase transition of the NPP [7]. It is proposed in [3] that a transition of this type may be as universal as the local REM. A proof of the universality of this transition poses yet another challenge for the rigorous community.

ACKNOWLEDGMENTS

During the beginning of this work, C.N. was supported by the Microsoft Graduate Fellowship. S.M. was supported in part by the German Science Council (grant ME2044/1-1). S.M. would also like to thank Microsoft Research for its hospitality.

REFERENCES

[1] G. Ausiello, P. Crescenzi, G. Gambosi, V. Kann, A. Marchetti-Spaccamela, and M. Protassi, Complexity and approximation, Springer-Verlag, Berlin, 1999.

[2] H. Bauke, S. Franz, and S. Mertens, Number partitioning as random energy model, J Stat Mech Theor Exp, 2004, P04003; http://www.iop.org/EJ/article/1742-5468/2004/04/P04003/jstat4_04_p04003.html.

[3] H. Bauke and S. Mertens, Universality in the level statistics of disordered systems, Phys Rev E 70 (2004), 025102(R).

[4] H. Bauke, S. Mertens, and A. Engel, Phase transition in multiprocessor scheduling, Phys Rev Lett 90 (2003), 158701.

[5] R. N. Bhattacharya and R. Ranga Rao, Normal approximation and asymptotic expansions, Wiley, New York, 1976.

[6] C. Borgs, J. Chayes, S. Mertens, and C. Nair, Proof of the local REM conjecture for number partitioning. II. Growing energy scales, Random Struct Algorithms 34 (2009), 241–284.

[7] C. Borgs, J. Chayes, and B. Pittel, Phase transition and finite-size scaling for the integer partitioning problem, Random Struct Algorithms 19 (2001), 247–288.

[8] A. Bovier and I. Kurkova, Local energy statistics in disorderd systems: A proof of the local REM conjecture, J Commun Math Phys 263 (2006), 513–533.

[9] A. Bovier and I. Kurkova, Poisson convergence in the restricted k-partioning problem, Random Struct Algorithms 30 (2007), 505–531.

[10] E. G. Coffman and G. S. Lueker, Probabilistic analysis of packing and partitioning algorithms, Wiley, New York, 1991.

[11] B. Derrida, Random-energy model: An exactly solvable model of disordered systems, Phys Rev B 24 (1981), 2613–2626.

[12] O. C. Martin, R. Monasson, and R. Zecchina, Statistical mechanics methods and phase transitions in optimization problems, Theor Comput Sci 265 (2001), 3–67.

[13] F. F. Ferreira and J. F. Fontanari, Probabilistic analysis of the number partitioning problem, J Phys A 31 (1998), 3417–3428.

[14] Y. Fu, The use and abuse of statistical mechanics in computational complexity, Lectures in the sciences of complexity, Proceedings of the 1988 Complex Systems Summer School, Santa Fe, New Mexico, 1988, edited by D.L. Stein, Addison-Wesley, Reading, MA, 1989, Vol. 1, pp. 815–826.

[15] M. R. Garey and D. S. Johnson, Computers and intractability. A guide to the theory of NP-completeness, W.H. Freeman, New York, 1997.

[16] I. P. Gent and T. Walsh, Phase transitions and annealed theories: Number partitioning as a case study, In the Proceedings of ECAI-96, W. Wahlster, Wiley, New York, 1996, pp. 170–174.

[17] B. Hayes, Computing science: The easiest hard problem, Am Scientist 90 (2002), 113–117.

[18] R. C. Merkle and M. E. Hellman, Hiding informations and signatures in trapdoor knapsacks, IEEE Trans Info Theory 24 (1978), 525–530.

[19] S. Mertens, Phase transition in the number partitioning problem, Phys Rev Lett 81 (1998), 4281–4284.

[20] S. Mertens, Random costs in combinatorial optimization, Phys Rev Lett 84 (2000), 1347–1350.

[21] M. Mézard, Passing messages between disciplines, Science 301 (2003), 1685–1686.

[22] A. M. Odlyzko, The rise and fall of knapsack cryptosystems, in: Cryptology and computational number theory (Boulder, CO, 1989), PSAM: Proceedings of the 42nd Symposium in Applied Mathematics, American Mathematical Society, Providence, RI, 1990. pp. 75–152.

[23] T. Sasamoto, T. Toyoizumi, and H. Nishimori, Statistical mechanics of an NP-complete problem: Subset sum, J Phys A 34 (2001), 9555–9567.

[24] L.-H. Tsai, Asymptotic analysis of an algorithm for balanced parallel processor scheduling, SIAM J Comput 21 (1992), 59–64.