Sharp Threshold and Scaling Window for the Integer Partitioning Problem

Christian Borgs
Microsoft Research
One Microsoft Way
Redmond, WA 98052
borgs@microsoft.com

Jennifer T. Chayes
Microsoft Research
One Microsoft Way
Redmond, WA 98052
jchayes@microsoft.com

Boris Pittel*
Department of Mathematics
Ohio State University
Columbus, OH 43210
bgp@math.ohio-state.edu

ABSTRACT

We consider the problem of partitioning n integers chosen randomly between 1 and 2^m into two subsets such that the discrepancy, the absolute value of the difference of their sums, is minimized. A partition is called *perfect* if the optimum discrepancy is 0 when the sum of all n integers in the original set is even, or 1 when the sum is odd. Parameterizing the random problem in terms of $\kappa = m/n$, we prove that the problem has a sharp threshold at $\kappa = 1$, in the sense that for $\kappa < 1$, there are many perfect partitions with probability tending to 1 as $n \to \infty$, while for $\kappa > 1$, there are no perfect partitions with probability tending to 1. Moreover, we show that the derivative of the so-called entropy is discontinuous at $\kappa = 1$.

We also determine the scaling window about the transition point: $\kappa_n = 1 - (2n)^{-1} \log_2 n + \lambda_n/n$, by showing that the probability of a perfect partition tends to 0, 1, or some explicitly computable $p(\lambda) \in (0,1)$, depending on whether λ_n tends to $-\infty$, ∞, or $\lambda \in (-\infty, \infty)$, respectively. For $\lambda_n \to -\infty$ fast enough, we show that the number of perfect partitions is Gaussian in the limit. For $\lambda_n \to \infty$, we prove that with high probability the optimum partition is unique, and that the optimum discrepancy is $\Theta(\lambda_n)$. Within the window, i.e., if $|\lambda_n|$ is bounded, we prove that the optimum discrepancy is bounded. Both for $\lambda_n \to \infty$ and within the window, the limiting distribution of the (scaled) discrepancy is found.

1. INTRODUCTION

There has recently been much interest in the study of phase transitions in random combinatorial problems. A combinatorial phase transition is an abrupt change in the qualitative behavior of the problem as an appropriately defined parameter is varied. The classic combinatorial phase transition occurs in the random graph model of Erdös and Rényi [8, 9]. There one considers a graph on n vertices with edge occupation probability α/n. As the parameter α passes through 1, the model undergoes a phase transition in the sense that the size of the largest connected component changes from order $\log n$ to order n. More recently, there has been much study of the phase transition in the random k-SAT model, both by heuristic and rigorous methods; see [3] and references therein. Here, the relevant parameter is $\alpha = m/n$, where m is the number of clauses and n is the number of variables. For fixed $k \geq 2$, the model undergoes a sharp transition from solvability to insolvability as α passes through a particular k-dependent value [13].

Phase transitions occur only in the limit of infinite systems. Finite-size scaling describes the "broadening" of the transition point into a "scaling window" in a finite system, and the behavior of the relevant functions in the scaling window. Finite-size scaling results are known for both the random graph model [2, 19, 16] and the 2-SAT problem [3]; in both cases, the window is of width $n^{-1/3}$. But the question of finite-size scaling is still open for k-SAT with $k \geq 3$.

The integer partitioning problem is a classic NP-complete problem of combinatorial optimization. In the random version considered here, an instance is a given a set of n m-bit integers drawn uniformly at random from the set $[M] = \{1, 2, \ldots, M\}$ with $M = 2^m$. The problem is to partition the given set into two subsets in order to minimize the absolute value of the difference between the sum of the integers in the two subsets, the so-called *discrepancy*. Clearly, the smallest possible discrepancy is 0 when the sum of all of the integers is even, and 1 when the sum is odd; a partition with this discrepancy is called *perfect*. In this work, we prove that the optimum partitioning problem undergoes a sharp transition as a function of the parameter $\kappa = m/n$, characterized by a dramatic change in the probability of a perfect partition. For m and n tending to infinity in the limiting ratio κ, the probability of a perfect partition tends to 0 for $\kappa < 1$, while the probability tends to 1 for $\kappa > 1$.

We also derive the finite-size scaling of the system about the transition point $\kappa = 1$. Namely, in terms of the more detailed parameterization $m = \kappa_n n$ with

$$\kappa_n = 1 - \frac{\log_2 n}{2n} + \frac{\lambda_n}{n}, \qquad (1)$$

*Research of B. Pittel supported by Microsoft and by NSF grant# DMS-9803410.

STOC'01, July 6-8, 2001, Hersonissos, Crete, Greece.

the probability of a perfect partition tends to 1, 0, or a computable λ-dependent constant strictly between 0 and 1, depending on whether λ_n tends to $-\infty$, ∞, or $\lambda \in (-\infty, \infty)$, respectively. To our knowledge, this is the first rigorous analysis of finite-size scaling in an NP-complete problem. Equation (1) is the analogue of the scaling $\alpha_n = 1+\lambda_n/n^{1/3}$ in the random graph problem [2, 19] and the 2-SAT problem [3]. Here the scaling window is much smaller than it is in the random graph or 2-SAT, namely it is of width $\Theta(1/n)$ rather than $\Theta(1/n^{1/3})$. Also, in contrast to the random graph and 2-SAT, the center of the scaling window here is shifted from its limiting value by an amount which is larger than the width of the window itself, namely $(2n)^{-1}\log_2 n$ versus $\Theta(1/n)$.

Finally, we derive the limiting distributions of some of the fundamental quantities in the system. For $\lambda_n \to -\infty$, we get the distribution of the number of perfect partitions, which gives us the *entropy*. Both for $\lambda_n \to \infty$ and within the window, we get the detailed asymptotics of the distribution of the minimum discrepancy.

The random optimum partitioning problem has been studied previously by both rigorous and nonrigorous methods. A great deal of rigorous work has been done for the partitioning problem with random numbers drawn from a compact interval in $\mathbb{R}$, which is analogous to the integer partitioning problem with $m \gg n$. Karmarkar and Karp [17] gave a linear time algorithm for a suboptimal solution with a typical discrepancy of order $O(n^{-c\log n})$ for some constant $c > 0$. The optimum solution was studied by Karmarkar, Karp, Lueker and Odlyzko [18] who proved that the typical minimum discrepancy is much smaller, namely of order $O(2^{-n}\sqrt{n})$. More recently, Lueker [20] proved similar bounds for the expected minimum discrepancy. Note that all of these results correspond to $m \gg n$, and hence $\kappa \to \infty$, well above the phase transition studied here.

There have also been (nonrigorous) studies of optimum partitioning in the theoretical physics and artificial intelligence communities, where the possibility of a phase transition was studied. Fu [14] noted that the minimum discrepancy is analogous to the ground state energy of an infinite-range, random antiferromagnetic spin model, but concluded incorrectly that the model did not have a phase transition. Gent and Walsh [15] examined the problem numerically, introduced the parameter $\kappa = m/n$, and estimated that a transition occurs at $\kappa = 0.96$. Ferreira and Fontanari studied the random spin model of Fu, and used statistical mechanical methods to get estimates of the optimum partition [10] and to evaluate the average performance of simple heuristics [11]. Our work was motivated by the paper of Mertens [21], who used statistical mechanical methods and the parameterization of [15] to derive a compelling argument for a phase transition. In a later work, Mertens [22] analyzed Fu's model by mapping it into Derrida's random energy model [7], and heuristically obtained the distribution of the discrepancy. Here we give a rigorous analogue of this result.

It is worth noting that the optimum partitioning problem is closely related to several other classic problems of combinatorial optimization. The first is the "multi-way" partition problem in which a set of "weights" is to be partitioned into $N \geq 3$ subsets (parts), so that the sums of the weights in the N parts are as close to equal as possible. Graham developed a linear-time $\frac{4}{3}$-approximation algorithm for a version of this problem in which the goal is to minimize the weight of the heaviest part [R. Graham, private communication]. The multi-way problem was also considered by Karmarkar, Karp, Lueker and Odlyzko [18], who noted that their analysis of the minimum discrepancy would extend in a natural way to this case also. A second related problem is the so-called subset sum problem, in which one tries to find subsets of a given set of integers which sum to (or near to) a prescribed target number. This problem reduces to a study of solutions of linear equations of the form $\sum_i s_i X_i = T$, where X_i are the numbers in the given set, $s_i \in \{0, 1\}$ represents whether or not X_i is included in a particular subset, and T is the target number. A key idea is to express the total number of solutions to these equations via a Fourier-type inversion integral, a paradigm championed by Freiman [12]; see also Alon and Freiman [1], Chaimovich and Freiman [6]. We will use an analogous integral representation in our study of the integer partitioning problem. Some of the methods and results presented here can be used to obtain stronger results for the subset sum problem, but we will not pursue this here.

2. STATEMENT OF RESULTS

Let us begin with a little notation. The instances of the problem are sets of n integers $X_1, \ldots, X_n$ chosen independently and uniformly from $[M] = \{1, 2, \ldots, M\}$ with $M = 2^m$. We will generally fix m to be some function of n (e.g., by taking $m = \kappa n$). The probability measure induced by the random variables $\mathbb{X} = \{X_1, \ldots, X_n\}$ will be denoted by $\mathbb{P}_n$, and expectation by $\mathbb{E}_n$. When no confusion arises, we will drop the subscript n. The event that "$\sum_{j=1}^n X_j$ is even" will be denoted by $\mathcal{E}_n$, while the event that the sum is odd will be denoted by $\mathcal{O}_n$. As usual we will say that an event happens with high probability (w.h.p) if the probability that it happens goes to one as $n \to \infty$. Finally, X will denote a generic random variable distributed uniformly on $[M]$.

There are 2^n ways to form an *ordered* partition of n integers $X_1, \ldots, X_n$ into two sets. Each such partition can be labelled by $\boldsymbol{\sigma} = (\sigma_1, \ldots \sigma_n)$ with $\sigma_j \in \{-1, 1\}$, so that, say, the first set is $\{X_j : \sigma_j = -1\}$, and the second is $\{X_j : \sigma_j = 1\}$. The discrepancy of the partition with label $\boldsymbol{\sigma}$ is $|\boldsymbol{\sigma}\cdot\mathbb{X}| = \left|\sum_{j=1}^n \sigma_j X_j\right|$. Let d_n denote the optimum discrepancy of $\mathbb{X}$ over all $\boldsymbol{\sigma}$:

$$d_n = d_n(\mathbb{X}) = \min_{\boldsymbol{\sigma}} |\boldsymbol{\sigma}\cdot\mathbb{X}|. \tag{2}$$

Clearly d_n is even on $\mathcal{E}_n$, and odd on $\mathcal{O}_n$. A partition with $|\boldsymbol{\sigma}\cdot\boldsymbol{X}| \leq 1$ (i.e., $|\boldsymbol{\sigma}\cdot\boldsymbol{X}| = 0$ on $\mathcal{E}_n$ and $|\boldsymbol{\sigma}\cdot\boldsymbol{X}| = 1$ on $\mathcal{O}_n$) is called *perfect* , and a partition with $|\boldsymbol{\sigma}\cdot\boldsymbol{X}| = d_n$ is called an optimum or minimum partition. Let $Z_n = Z_n(\mathbb{X})$ and $\widetilde{Z}_n = \widetilde{Z}_n(\mathbb{X})$ denote the number of perfect and optimum partitions of $\mathbb{X}$, respectively. Of course, $Z_n = \widetilde{Z}_n$ iff $d_n = 0$ or $d_n = 1$. Note that a partition with label $\boldsymbol{\sigma}$ has the same discrepancy as that with label $-\boldsymbol{\sigma}$. The random variables $Z_n(\mathbb{X})$ and $\widetilde{Z}_n(\mathbb{X})$ therefore take values in the even nonnegative integers.

Our first result shows that the model has a sharp transition at $\kappa = 1$.

THEOREM 1. *Let $m = \kappa_n n$, and assume that $\lim_{n\to\infty} \kappa_n = \kappa$ exists in $[-\infty, \infty]$. Then*

$$\lim_{n\to\infty} \mathbb{P}_n(\exists \text{ a perfect partition}) = \begin{cases} 1 & \text{if } \kappa < 1 \\ 0 & \text{if } \kappa > 1. \end{cases} \tag{3}$$

Our next result uses the more sensitive parameterization (1) to strengthen Theorem 1, and, in particular, to establish the existence of a scaling window.

THEOREM 2. *Let $m = \kappa_n n$, with κ_n as in* (1)*, and assume that* $\lim_{n\to\infty} \lambda_n = \lambda$ *exists. Then*

$$\lim_{n\to\infty} \mathbb{P}_n(\exists \text{ a perfect partition}) =$$
$$= \begin{cases} 1 & \text{if } \lambda = -\infty \\ 1 - \frac{1}{2}r(\lambda)\big(r(\lambda)+1\big) & \text{if } \lambda \in (-\infty,\infty) \\ 0 & \text{if } \lambda = \infty, \end{cases} \tag{4}$$

where $r(\lambda) = \exp\left(-\sqrt{\frac{3}{2\pi}}2^{-\lambda}\right)$.

Our next result gives detailed information on the distribution of the number of perfect and optimum partitions, Z_n and $\widetilde{Z}_n$, and therefore also on the *entropy* , defined as

$$S_n = \log_2 \widetilde{Z}_n. \tag{5}$$

Note that S_n is well-defined and non-negative for all $\boldsymbol{X}$, since $\widetilde{Z}_n \geq 1$. In contrast, an "entropy" defined as $\log_2 Z_n$ can be negative infinity, which led to some apparent contradictions in [21].

THEOREM 3. *Let $m = \kappa_n n$, with κ_n as in* (1)*, and define*

$$c_M = \mathbb{E}\Big(\frac{X^2}{M^2}\Big) = \frac{1}{3} + \frac{1}{2M} + \frac{1}{6M^2}. \tag{6}$$

i) If $\lambda_n \to -\infty$, then

$$\left(\frac{2^{1+|\lambda_n|}}{\sqrt{2\pi c_M}}\right)^{-1} Z_n \to \begin{cases} 1 & \text{on } \mathcal{E}_n \\ 2 & \text{on } \mathcal{O}_n \end{cases} \tag{7}$$

in probability and in mean,

$$S_n - |\lambda_n| + \frac{1}{2}\log_2 c_M \to \begin{cases} \frac{1}{2}\log_2(2/\pi) & \text{on } \mathcal{E}_n \\ \frac{1}{2}\log_2(8/\pi) & \text{on } \mathcal{O}_n \end{cases} \tag{8}$$

in probability, and

$$n^{-1}\Big(S_n - |\lambda_n|\Big) \to 0 \tag{9}$$

in expectation.

ii) If $\lambda_n \to \lambda \in (-\infty,\infty)$, then S_n is bounded in probability, so that in particular

$$n^{-1}S_n \to 0 \tag{10}$$

in probability. More precisely, on the event $\mathcal{O}_\setminus$ the entropy S_n converges (in distribution) to $1 + \log_2 P(\mu)$ where $P(\mu)$ is Poisson with parameter $\mu = 2^{-\lambda}\sqrt{6/\pi}$ conditioned on $\{P(\mu) \geq 1\}$; on the event $\mathcal{E}_\setminus$ the entropy S_n converges to $1 + \log_2 Q(\mu)$, where $Q(\mu) = P(\mu/2)$ with probability $1 - e^{-\mu/2}$ and $Q(\mu) = P(\mu)$ with probability $e^{-\mu/2}$.

iii) If $\lambda_n \to \infty$ with $\lambda_n = O(n)$, then with probability tending to 1, the optimum partition is unique up to the symmetry $\boldsymbol{\sigma} \to -\boldsymbol{\sigma}$. In particular, $\mathbb{P}(S_n = 1) \to 1$ as $n \to \infty$.

COROLLARY 1. *Assume that m/n converges to some $\kappa < \infty$. Then the entropy per variable, $s_n = n^{-1}S_n$, converges in probability to (the deterministic function)*

$$s(\kappa) = \max\{0, 1-\kappa\}, \tag{11}$$

so that, in particular, the limiting entropy per variable has a discontinuous derivative at $\kappa = 1$.

Remark 1. The reader will note that the statements below the window in Theorem 1 and 2 are immediate corollaries of Theorem 3(i), equation (7), which strengthens the statement $Z_n > 0$ w.h.p. by giving a law of large numbers for Z_n.

Remark 2. If the condition of Theorem 3(i) is slightly strengthened to $\lambda_n + \log_2 n \to -\infty$, we can prove even more, namely a central limit theorem stating that, in the limit, Z_n has a Gaussian distribution with mean implicit in (7), and standard deviation roughly equal to the mean times $n^{-1/2}$; see [4]. This allows us to show that S_n is also Gaussian in the limit with mean implicit in (8), and standard deviation again roughly equal to the mean times $n^{-1/2}$.

Remark 3. In statistical physics, phase transitions are characterized by non-analyticities in derivatives of thermodynamic potentials. These non-analyticities may be discontinuities or smoother non-analyticities. First-order phase transitions are characterized by a discontinuity in a first derivative of a thermodynamic potential (but not necessarily in all first derivatives of all thermodynamic potentials). By contrast, all first derivatives of thermodynamic potentials are continuous at second-order phase transitions; the corresponding second derivatives usually diverge. In the optimum partitioning problem, the entropy—which is a first derivative of a thermodynamic potential—is continuous, but its derivative is discontinuous. This is analogous to the behavior of the entropy of the Ising model in a magnetic field, which has a first-order phase transition as the magnetic field passes through zero.

Another characteristic which can be used to distinguish first- and second-order phase transitions is the width of the scaling window. In a first-order phase transition, such as the Ising model in a field, the scaling window of a system of size n is of width n^{-1}; see [5]. By contrast, second-order phase transitions have scaling windows of width n^{-b} for some $b < 1$, as has been established for the random graph [2, 19] and the 2-SAT problem [3]. In the optimum partitioning problem, the scaling window of is width n^{-1}. Hence we conclude that the problem has a first-order phase transition at $\kappa = 1$.

Our final theorem gives detailed distributional estimates of the discrepancy d_n defined in (2).

THEOREM 4. *Let $m = \kappa_n n$, with κ_n as in* (1)*.*

i) If $\lambda_n \to -\infty$, then

$$d_n \to \begin{cases} 0 & \text{in probability on } \mathcal{E}_n \\ 1 & \text{in probability on } \mathcal{O}_n. \end{cases} \tag{12}$$

ii) If $\lim_{n\to\infty} \lambda_n \in (-\infty,\infty)$, then d_n is bounded in probability. More precisely, in the limit, d_n has a geometric distribution: for $\ell \geq 1$,

$$\lim_{n\to\infty} \mathbb{P}_n\{d_n \geq \ell\} = \frac{1+r}{2}r^{\ell-1}, \tag{13}$$

with $r = r(\lambda)$ as defined in Theorem 1.

iii) If $\lambda_n \to \infty$, then $d_n/2^{\lambda_n}$ and its inverse are bounded in probability. If furthermore $\lambda_n = O(n)$, then in the limit, $d_n/2^{\lambda_n}$ has the following exponential distribution: for $a > 0$,

$$\lim_{n\to\infty} \mathbb{P}_n\left(\frac{d_n}{2^{\lambda_n}} > a\right) = \exp\left(-\sqrt{\frac{3}{2\pi}}a\right). \tag{14}$$

The reader will notice that Theorem 2 is a corollary of Theorem 4, and that Theorem 4(i) follows from Theorem 3(i), equation (7). The proof of Theorems 1–4 is therefore reduced to that of Theorem 3 and Theorem 4(ii)–(iii). This is accomplished by detailed calculations using an integral representation to be described in the next section.

3. SCHEME OF THE PROOFS

Here we give a brief outline of parts of the proofs of Theorems 1–4. The complete proofs are rather complicated; the reader is referred to [4] for the details. Our proofs are based on an integral representation of $Z_{n,\ell}$, the number of partitions with discrepancy ℓ. To derive this representation, we first write $Z_{n,\ell}$ as

$$Z_{n,\ell} = \sum_{\boldsymbol{\sigma}} \mathbb{I}(|\boldsymbol{\sigma}\cdot\boldsymbol{X}| = \ell), \tag{15}$$

where we use $\mathbb{I}(A)$ to denote the indicator of an event A, and then use the identity

$$\mathbb{I}(\boldsymbol{\sigma}\cdot\boldsymbol{X} = \ell) = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{i(\sigma\cdot\boldsymbol{X}-\ell)x}dx \tag{16}$$

to sum over all 2^n configurations $\boldsymbol{\sigma}$. This gives the representation

$$Z_{n,\ell} = 2^n I_{n,\ell} \times \begin{cases} 1 & \text{if } \ell = 0 \\ 2 & \text{if } \ell > 0, \end{cases} \tag{17}$$

where $I_{n,\ell} = I_{n,\ell}(\boldsymbol{X})$ is the random integral

$$I_{n,\ell} = \frac{1}{2\pi}\int_{-\pi}^{\pi} \cos(\ell x)\prod_{j=1}^{n}\cos(xX_j)\,dx. \tag{18}$$

The first set of results, namely Theorem 1, Theorem 2 outside the window, and Theorem 3(i), follow from estimates on the first and second moments of $I_{n,\ell}$. Below we will state these estimates, prove the first moment estimate, and show how the estimates imply the theorems mentioned above. The central limit theorem, referred to in Remark 2 following Theorem 3, is a consequence of detailed estimates on the *random* integral $I_{n,\ell}$, rather than just on a few of its moments. The reader is referred to [4] for details.

Proposition 3.1. *Let $M = M(n)$ be an arbitrary function of n, let*

$$\gamma_n = \frac{1}{M\sqrt{2\pi n c_M}} \tag{19}$$

with c_M as in (6), and assume that ℓ and ℓ' are integers between 0 and M. Then

$$\mathbb{E}[I_{n,\ell}] = \gamma_n(1 + O(n^{-1})). \tag{20}$$

Furthermore

$$\begin{aligned}\mathbb{E}[I_{n,\ell}I_{n,\ell'}] =& 2\gamma_n^2\Big(1 + O(n^{-1}) + O\Big(\frac{n^{-1}}{\gamma_n 2^n}\Big)\Big) \\ &+ 2^{-n}\gamma_n\Big(\delta_{\ell+\ell',0} + \delta_{\ell-\ell',0}\Big)\end{aligned} \tag{21}$$

if ℓ and ℓ' are both even or both odd, while $\mathbb{E}[I_{n,\ell}I_{n,\ell'}] = 0$ if one of them is even and the other is odd. In (20) and (21), the bounds implicit in the O-symbols are uniform in M.

To prove Proposition 3.1, we first use (18) and the fact that the X_j are independent to get

$$\mathbb{E}[I_{n,\ell}] = \frac{1}{2\pi}\int_{-\pi}^{\pi}\cos(\ell x) f^n(x)\,dx \tag{22}$$

and

$$\begin{aligned}&\mathbb{E}\big[I_{n,\ell}I_{n,\ell'}\big] = \\ &= \frac{1}{(2\pi)^2}\iint_{x_1,x_2\in(-\pi,\pi]}\cos(x_1\ell)\cos(x_2\ell')f^n(x_1,x_2)\,dx_1dx_2\end{aligned} \tag{23}$$

where

$$\begin{aligned} f(x) :&= \mathbb{E}[\cos(xX)] = M^{-1}\sum_{j=1}^{M}\cos(jx) \\ &= M^{-1}\left[\frac{\sin((M+1/2)x)}{2\sin(x/2)} - \frac{1}{2}\right],\end{aligned} \tag{24}$$

$$\begin{aligned} f(x_1,x_2) :&= \mathbb{E}\big[\cos(x_1X)\cos(x_2X)\big] \\ &= \frac{1}{2}(f(x_1+x_2) + f(x_1-x_2))\end{aligned} \tag{25}$$

and $f^n(x)$ stands for $[f(x)]^n$. In our analysis of the asymptotics of these integrals, we use that the major contributions come from values of the integration variables near the maxima of $|f(x)|$ and $|f(x_1,x_2)|$. Note, however, that the function f depends on the parameter M, which can grow as $n\to\infty$. Fortunately, a careful treatment of error terms will allow us to apply a variation of the standard saddle point method to get the desired asymptotics. Here we present this analysis for the first moment. The (more involved) analysis for the second moment requires a saddle point for an integral in two variables; it can be found in [4].

Proof of Equation (20)

When M is bounded, the proof is a straightforward application of standard saddle point methods, which we leave to the reader. The arguments below establish (20) for M larger than some M_0, to be determined in the course of the proof.

Pick $1 < a < b$. If $x \in [-\pi,\pi]$ is such that $|2\sin(x/2)| \geq b/M$, then

$$|f(x)| \leq \frac{1}{b} + \frac{1}{2M} \leq \frac{1}{a} \tag{26}$$

for M large enough. We will also use

$$|f(x)| \leq \frac{1}{M|\sin(x/2)|} \leq \frac{C}{M|x|}, \tag{27}$$

uniformly for $|x| \in (0,\pi]$, a direct consequence of (24). Here and below, C,C_1, C', etc., stand for absolute positive constants. (In (27), C can be chosen as $\pi/2$.)

Notice that $x \in (-\pi,\pi)$ satisfies $|2\sin(x/2)| \leq b/M$ iff $|x| \leq b_0/M$, where $b_0 = b_0(M)$ is defined for M large enough by the condition $2\sin(b_0/(2M)) = b/M$ with $b_0/(2M) \in (0,\pi/2)$. Clearly $b_0(M)\to b$ as $M\to\infty$. By (26),

$$|f(x)| \leq \frac{1}{a} \quad \text{if} \quad |x| \geq b_0/M. \tag{28}$$

Consider now $|x| \leq b_0/M$, and set $x = y/M$, i.e., $|y| \leq b_0$.

Then, for $y \neq 0$,

$$f(x) = M^{-1}\left[\frac{\sin\big(y + y/(2M)\big)}{2\sin\big(y/(2M)\big)} - \frac{1}{2}\right] = \frac{\sin y}{2M\tan\big(y/(2M)\big)} + \frac{\cos y - 1}{2M}, \tag{29}$$

and, since $\tan z > z$ on $(0, \pi/2)$,

$$|f(x)| \leq \left|\frac{\sin y}{y}\right| + \frac{1-\cos y}{2M}. \tag{30}$$

Hence there exist a small enough $y_0 \in (0, b_0)$ and $q \in (0,1)$ so that for M large enough

$$|f(x)| \leq \begin{cases} e^{-C_1 y^2} & \text{if } |y| \leq y_0, \\ q, & \text{if } y_0 \leq |y| \leq b_0. \end{cases} \tag{31}$$

For $|y|$ sufficiently small, we also have

$$\begin{aligned} f(x) &= \mathbb{E}\big(\cos(y(X/M))\big) = \mathbb{E}\left(1 - \frac{y^2}{2}\cdot\frac{X^2}{M^2} + O(y^4)\right) \\ &= 1 - \frac{c_M}{2}y^2 + O(y^4) = \exp\left(-\frac{c_M}{2}y^2 + O(y^4)\right), \end{aligned} \tag{32}$$

with c_M as in (6), so that

$$f^n(x) = \exp\left(-n\frac{c_M}{2}y^2\right)\left(1 + O(ny^4)\right). \tag{33}$$

Since $\cos(\ell x) = \cos(y\ell/M) = 1 + O(y^2)$ for $\ell = O(M)$, we can use (31) and (33) to get:

$$\begin{aligned} &\frac{1}{2\pi}\int\limits_{|x|\leq b_0/M} \cos(\ell x) f^n(x)\,dx = \\ &= \frac{1}{2\pi M}\int\limits_{|y|\leq\frac{\log_2 n}{\sqrt{n}}} \left(1 + O(y^2) + O(ny^4)\right)e^{-\frac{nc_M}{2}y^2}\,dy \\ &\quad + O\left(M^{-1}\int_{|y|\geq\frac{\log_2 n}{\sqrt{n}}} e^{-nC_1 y^2}\,dy\right) + O(q^n/M) \\ &= \frac{1}{M\sqrt{2\pi n c_M}}\left(1 + O(n^{-1}) + O\big(e^{-C'\log_2^2 n}\big)\right) \\ &= \gamma_n(1 + O(n^{-1})), \end{aligned} \tag{34}$$

with γ_n as in (19). Besides, by (26) and the definition of b_0,

$$\begin{aligned} &\int\limits_{|x|\in[b_0/M,\pi]} |f(x)|^n\,dx \leq \\ &\leq \int\limits_{|x|\in[b_0/M,\pi]} \left[\min\left\{a^{-1}, \frac{C}{Mx}\right\}\right]^n dx \\ &\leq \frac{2Ca}{M}a^{-n} + \int\limits_{|x|\geq Ca/M} \left(\frac{C}{Mx}\right)^n dx \\ &\leq \frac{4Ca}{M}a^{-n} = O(a^{-n}M^{-1}). \end{aligned} \tag{35}$$

Thus, for $\ell = O(M)$ and every $a > 1$,

$$\begin{aligned} \mathbb{E}I_{n,\ell}(\mathbb{X}) &= \gamma_n(1 + O(n^{-1})) + O(a^{-n}M^{-1}) \\ &= \gamma_n(1 + O(n^{-1})), \end{aligned} \tag{36}$$

where the constant implicit in $O(n^{-1})$ (and similar error terms below) depends on a. This proves the estimate (20) for all sufficiently large M. $\square$

Proof of Theorem 1, Theorem 2 outside the window, and Theorem 3(i)

To apply Proposition 3.1, we first observe that $Z_{n,0} = Z_n$ on $\mathcal{E}_n$ and $Z_{n,0} = 0$ on $\mathcal{O}_n$. Consequently

$$\mathbb{E}[Z_{n,0}] = \mathbb{P}(\mathcal{E}_n)\mathbb{E}(Z_n|\mathcal{E}_n), \quad \mathbb{E}[Z_{n,0}^2] = \mathbb{P}(\mathcal{E}_n)\mathbb{E}(Z_n^2|\mathcal{E}_n). \tag{37}$$

Observing that $\mathbb{P}_n(\mathcal{E}_n) \to 1/2$ as $n \to \infty$, with an error that is exponentially small in n, it follows from (17) and the statements of Proposition 3.1 for $\ell = \ell' = 0$ that

$$\mathbb{E}(Z_n|\mathcal{E}_n) = 2^{n+1}\gamma_n(1 + O(n^{-1})), \tag{38}$$

and

$$\mathbb{E}(Z_n^2|\mathcal{E}_n) = \big(\mathbb{E}^2(Z_n|\mathcal{E}_n) + 2\mathbb{E}(Z_n|\mathcal{E}_n)\big)(1 + O(n^{-1})). \tag{39}$$

It follows from (1) that $M = 2^m = 2^{n+\lambda_n}/\sqrt{n}$, so that $2^{n+1}\gamma_n = 2^{1-\lambda_n}/\sqrt{2\pi c_M} \to \infty$ if $\lambda_n \to -\infty$. The second moment method then immediately gives equation (7) on $\mathcal{E}_n$. In a similar way, the corresponding statements on $\mathcal{O}_n$ follow from Proposition 3.1 for $\ell = \ell' = 1$. Equation (7) implies in particular that w.h.p. $Z_n > 0$, whence $Z_n = \tilde{Z}_n$ and $S_n = \log_2 Z_n$. This observation and the convergence of $Z_n/(2^{n+1}\gamma_n)$ gives (8).

To complete the proof of Theorem 3(i), we note that by (8), $\Delta s_n := n^{-1}(S_n - |\lambda_n|)$ goes to zero in probability. Also $|\Delta s_n| \leq 1$, since $0 \leq S_n \leq n$ and $|\lambda_n| \leq n$. So, by the bounded convergence theorem, $\Delta s_n \to 0$ in expectation as well, completing the proof of Theorem 3(i), and hence the statements of Theorem 1 and 2 below the window. To see that the statements of Theorem 1 and 2 above the window follow from Proposition 3.1 as well, we use that the probability of finding a perfect partition is equal to the probability that $Z_n > 0$, which in turn is bounded above by the expectation of Z_n. But this expectation goes to zero above the window by (38) and its analogue on the event $\mathcal{O}_n$. $\square$

Moment estimates can also be used to obtain some statements in probability (but not any distributional statements) on S_n and d_n, namely the "in probability" statements of Theorems 3(ii) and 4(ii) and part of those in Theorem 4(iii). These are summarized in the following proposition.

Proposition 3.2. *(i) If $\lambda_n \to \lambda \in (-\infty, \infty)$, then both d_n and S_n are bounded in probability.*
(ii) If $\lambda_n \to \infty$, then $2^{\lambda_n}/d_n$ is bounded in probability.

Proof (i) Let $Z_{n,\leq\omega} = \sum_{\ell\leq\omega} Z_{n,\ell}$, and let $\omega(n) \leq M$ be a sequence of integers which goes to infinity. Use (17) and both the $\ell = \ell'$ and the $\ell \neq \ell'$ relations in Proposition 3.1 to calculate the first and second moments of the sum $Z_{n,\leq\omega}$. Then the second moment method implies that, inside the window,

$$\frac{Z_{n,\leq\omega(n)}}{\omega(n)\gamma_n 2^{n+1}} \to 1 \tag{40}$$

in probability. (Since $\gamma_n 2^n$ is bounded above, it is crucial that $\omega(n) \to \infty$ to control the error terms.) Observing that $\gamma_n 2^n$ is bounded away from zero inside the window, and noting that $Z_{n,\le\omega(n)} > 0$ implies $d_n \le \omega(n)$, the bound (40) proves in particular that d_n is bounded in probability. Since $Z_{n,\le\omega(n)} > 0$ implies $\widetilde{Z}_n \le Z_{n,\le\omega(n)}$, the bound (40) also implies that inside the window $\widetilde{Z}_n$ and hence S_n is bounded in probability.

(ii) Let $\omega(n) \to \infty$ as $n \to \infty$, and set $k_n = 2^{\lambda_n}/\omega(n)$. Since $k_n \le M$ for all sufficiently large n, we can use (17) and (20) to conclude that

$$\begin{aligned}\mathbb{P}(d_n \le k_n) &\le \sum_{k=0}^{k_n} \mathbb{E}[Z_{n,k}] \\ &= (1+2k_n)2^n\gamma_n(1+O(n^{-1})) \\ &= O\big(\omega^{-1}(n)\big),\end{aligned} \tag{41}$$

implying that d_n goes to infinity at least as fast as 2^{λ_n}, which is the desired result. □

We have not yet shown all of the "in probability" statements. In particular, for Theorem 4(iii), we need the following.

Proposition 3.3. *If $\lambda_n \to \infty$, then $d_n/2^{\lambda_n}$ is bounded in probability.*

In order to prove Proposition 3.3, by analogy to the proof of Proposition 3.2(ii), we would like to show that (40) holds for some $\omega(n)$ of the form $\omega(n) = 2^{\lambda_n} r_n$ with both λ_n and r_n going to infinity. Unfortunately, our bounds in Proposition 3.1 are not strong enough to prove such a statement using the second moment method. Instead we use a very different technique, inspired both by [18] and by the Gibbs distribution on integer partitions used in [21]. The key idea is to replace $Z_{n,\le\omega(n)}$ by

$$Z_{n,\le\omega(n)}(\beta) = \sum_{\boldsymbol{\sigma}:|\boldsymbol{\sigma}\cdot\boldsymbol{X}|\le\omega(n)} e^{-\beta|\boldsymbol{\sigma}\cdot\boldsymbol{X}|} \tag{42}$$

and then to choose β appropriately. In physics language, this amounts to replacing the "ground state partition function" by a partition function at positive temperature β^{-1}. Mathematically, this amounts to a smoothing technique which regularizes the sum $Z_{n,\le\omega(n)}$. Using this idea, we can now outline the proof of Proposition 3.3.

Sketch of Proof of Proposition 3.3

Let us choose $\beta = \beta_n$ in (42) so that $\beta\omega(n) \to 0$ as $n \to \infty$. Then

$$Z_{n,\le\omega(n)} = Z_{n,\le\omega(n)}(\beta)\Big(1 + O(\beta\omega(n))\Big), \tag{43}$$

so it is enough to prove the analogue of (40) for $Z_{n,\le\omega(n)}(\beta)$. We will sketch how this is done on the event $\mathcal{E}_n$. On $\mathcal{E}_n$, $\boldsymbol{\sigma}\cdot\boldsymbol{X}$ is necessarily even, implying that

$$Z_{n,\le\omega(n)}(\beta) = \sum_{\mu=-(\mu_n-1)}^{\mu_n-1} e^{-2\beta\mu} \sum_{\boldsymbol{\sigma}} \mathbb{I}(\boldsymbol{\sigma}\cdot\boldsymbol{X} = 2\mu) \tag{44}$$

where $\mu_n = \min\{\mu\colon 2\mu > \omega(n)\}$. Using once more that $\boldsymbol{\sigma}\cdot\boldsymbol{X}$ is even, which implies that the indicator function in (44) can now be rewritten as an integral over the smaller interval $[-\pi/2, \pi/2]$, we get

$$\begin{aligned}Z_{n,\le\omega(n)}(\beta) &= \\ &= \frac{2^n}{\pi}\int_{-\pi/2}^{\pi/2}\Bigg(\sum_{|\mu|<\mu_n} e^{-\beta|2\mu|}e^{i2\mu x}\Bigg)\prod_{j=1}^{n}\cos(xX_j)\,dx \\ &= \frac{2^n}{\pi}\int_{-\pi/2}^{\pi/2} g_n(x)\prod_{j=1}^{n}\cos(xX_j)\,dx,\end{aligned} \tag{45}$$

where $g_n(x)$ is an explicitly computable function.

To estimate the random integral in (45), we split it into two parts: the integral over $|x| \le b/M$, where $b < \pi$ is chosen in the course of the proof, and the integral over $|x| \ge b/M$. While the latter can be estimated by a bound on its second moment involving more saddle point estimates, the former requires precise estimates of a random integral. The main idea is to replace the product $\prod_{j=1}^{n}\cos(xX_j)$ in the integrand by

$$\exp\Big(-\frac{x^2}{2}\sum_{i=1}^{n} X_i^2\Big) = \exp\Big(-\frac{y^2}{2}\sum_{i=1}^{n}\frac{X_i^2}{M^2}\Big), \tag{46}$$

where, as before, $y = xM$. While the error terms cannot be controlled in general, it turns out that they can be controlled when the sum $\sum_{i=1}^{n} X_i^2/M^2$ is near to its expectation. Let A_n be the event that the ratio between $\sum_{i=1}^{n} X_i^2/M^2$ and its expectation is between 1/2 and 3/2. Since, by the law of large numbers, the sum converges to its expectation, the probability of A_n tends to one. Therefore, statements about convergence in probability can safely be made by restricting the random integers $\{X_j\}$ to the event $\mathcal{E}_n \cap A_n$. On this event, we finally do a saddle point analysis which requires careful tuning of several constants, including the constant b and the n-dependence of β, thus completing the estimate of the random integral. The proof is quite involved, and the interested reader is encouraged to consult [4] for details. Of course, an analogous proof applies on the event $\mathcal{O}_n$. □

We are left with proving the statement that the optimum partition is unique above the scaling window (see Theorem 3(iii)) and the distributional statements on d_n and S_n in Theorem 3(ii) and Theorems 4(ii) and 4(iii). To establish the statements above the window, we prove that the continuous-time process $(\hat{Z}_n(t))$, where $\hat{Z}_n(t) = \frac{1}{2}Z_{n,\le tb_n}$ with $b_n = 2^{\lambda_n+1}$, converges to a Poisson process with parameter $\sqrt{6/\pi}$. The statements inside the window are proved by showing that the random variables $\hat{Z}_{n,\ell} = \frac{1}{2}Z_{n,\ell}$ are independent Poisson random variables in the limit. These proofs entail estimates on the factorial moments of the relevant random variables. At the core of our argument is a version of the multidimensional local limit theorem for non-identically distributed vector-summands. Once more, the reader is referred to [4] for details.

4. REFERENCES

[1] N. Alon and G. Freiman. On sums of subsets of a set of integers. *Combin.*, 8:297–306, 1988.

[2] B. Bollobás. The evolution of random graphs. *Trans. Amer. Math. Soc.*, 286:257–274, 1984.

[3] B. Bollobas, C. Borgs, J. Chayes, J.-H. Kim, and D. Wilson. The scaling window of the 2-sat transition. *Rand. Struc. Alg.*, 18:201–256, 2001.

[4] C. Borgs, J. Chayes, and B. Pittel. Phase transition and finite-size scaling for the optimum partitioning problem. Technical report, Microsoft Research, http://www.research.microsoft.com/pubs/, 2000.

[5] C. Borgs and R. Kotecký. A rigorous theory of finite-size scaling at first-order transitions. *J. Statist. Phys.*, 61:79–110, 1990.

[6] M. Chaimovich and G. Freiman. Solving dense subset-sum problems by using analytical number theory. *J. of Complexity*, 5:271–282, 1989.

[7] B. Derrida. Random-energy model: An exactly solvable model of disordered systems. *Phys. Rev. B (3)*, 24:2613–2626, 1981.

[8] P. Erdös and A. Rényi. On the evolution of random graphs. *Magyar Tud. Akad. Mat. Kutató Int. Közl.*, 5:17–61, 1960.

[9] P. Erdös and A. Rényi. On the evolution of random graphs. *Bull. Inst. Internat. Statist.*, 38:343–347, 1961.

[10] F. Ferreira and J. Fontanari. Probabilistic analysis of the number partitioning problem. *J. Phys. A: Math. Gen.*, 31:3417–3428, 1998.

[11] F. Ferreira and J. Fontanari. Statistical mechanics analysis of the continuous number partitioning problem. *Physica A*, 269:54–60, 1999.

[12] G. Freiman. An analytical method of analysis of linear boolean equations. *Ann. N.Y. Acad. Sci.*, 337:97–102, 1980.

[13] E. Friedgut. Sharp thresholds of graph properties, and the k-sat problem. *J. Amer. Math. Soc.*, 12:1017–1054, 1999. With an appendix by J. Bourgain.

[14] Y. Fu. The use and abuse of statistical mechanics in computational complexity. In *Lectures in the Science 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.

[15] I. Gent and T. Walsh. In *Proc. of the 12th European Conference on Artificial Intelligence, Budapest, Hungary, 1996, edited by W. Wahlster*, pages 170–174. John Wiley & Sons, New York, NY, 1996.

[16] S. Janson, D. Knuth, T. Łuczak, and B. Pittel. The birth of the giant component. *Rand. Struc. Alg.*, 4:231–358, 1994.

[17] N. Karmarkar and R. Karp. The differencing method of set partitioning. Technical Report UCB/CSD 82/113, Computer Science Division (EECS), University of California, Berkeley, 1982.

[18] N. Karmarkar, R. Karp, G. Lueker, and A. Odlyzko. Probabilistic analysis of optimum partitioning. *J. Appl. Prob.*, 23:626–645, 1986.

[19] T. Łuczak. Component behavior near the critical point of the random graph process. *Rand. Struc. Alg.*, 1:287–310, 1990.

[20] G. Lueker. Exponentially small bounds on the expected optimum of the partition and subset sum problem. *Rand. Struc. Alg.*, 12:51–62, 1998.

[21] S. Mertens. Phase transition in the number partitioning problem. *Phys. Rev. Lett.*, 81:4281–4284, 1998.

[22] S. Mertens. Random costs in combinatorial optimization. *Phys. Rev. Lett.*, 84:1347–1350, 2000.