An exactly solvable random satisfiability problem

# An exactly solvable random satisfiability problem

**Sergio Caracciolo**[1,2] **and Andrea Sportiello**[3]

[1] Dipartimento di Fisica and INFN, Università degli Studi di Milano, via Celoria 16,
I-20133 Milano, Italy
[2] NEST-INFM, I-56126 Pisa, Italy
[3] Scuola Normale Superiore and INFN – Sez. di Pisa, Piazza dei Cavalieri 7, I-56100 Pisa, Italy

E-mail: Sergio.Caracciolo@mi.infn.it and Andrea.Sportiello@sns.it

**Abstract**
We introduce a new model for the generation of random satisfiability problems.
It is an extension of the hyper-SAT model of Ricci-Tersenghi, Weigt and
Zecchina, which is a variant of the famous $K$-SAT model: it is extended to
$q$-state variables and relates to a different choice of the statistical ensemble.

The model has an exactly solvable statistic: the critical exponents
and scaling functions of the SAT/UNSAT transition are calculable at zero
temperature, with no need of replicas, also with exact finite-size corrections.

We also introduce an exact duality of the model, and show an analogy of
thermodynamic properties with the random energy model of disordered spin
system theory. Relations with error correcting codes are also discussed.

PACS numbers: 89.80.+h, 75.10.Nr

## 1. Introduction

### 1.1. Motivations

Complexity theory concerns the classification of combinatorial optimization problems,
according to the computational cost required for their solution [1–3]. A central role is
occupied by the class of hard problems named NP (non-deterministic polynomial time). For
problems in this class, a potential solution can be checked in polynomial time, whereas finding
one solution may require, in the worst case, a time growing exponentially with system size.

A sub-class of the NP class exists, called NP-complete, such that for each problem in this
sub-class no other NP problem is more than a polynomial factor harder. Informally, a problem
is NP-complete if answers can be verified quickly (NP requirement), and a quick algorithm
to solve this problem can be used to solve all other NP problems quickly (completeness
requirement). NP-complete problems are particularly important as, in some sense, they

manifest all the potential difficulties of any NP problem, and are a sort of 'language' for properly describing every possible problem in NP class.

Often, the crude worst-case analysis of ordinary complexity theory hides the mechanism which leads to NP complexity. A statistical mechanical analysis of the ensemble of possible instances of the problem (we could say a 'typical case' analysis, eventually dependent on some macroscopic parameter of the theory) would have the ambitious challenge of deepening the understanding of computational complexity, and improving the algorithm design for real world applications.

For example, in a variety of cases, the problem admits a space of parameters, and a certain order parameter exists, showing a critical behaviour. In some of these cases, a widespread conjecture is that the computationally hard instances appear with a significant probability only when generated near 'phase boundaries' of the ensemble, with respect to this phase transition.

An exemplar case of this phenomenon is the random Boolean satisfiability problem ($K$-SAT [4–6]). In this case, a peculiar transition, called *SAT/UNSAT transition*, occurs: when the density of clauses is higher than a certain critical value, the system is highly constrained, and the average probability of having a solution goes to 0 in the thermodynamic limit. Otherwise, the system is slightly constrained, and the average probability of having a solution goes to 1 in the thermodynamic limit. Difficult instances can be found only in a narrow region on the density axes, in correspondence with the critical density. This fact has a clear interpretation: heuristic algorithms for randomly searching a solution are efficient when the density is low enough, while algorithms for randomly searching a certificate of unsatisfiability are efficient when the density is high enough. At the phase boundary, however, both the average probability of being satisfiable or unsatisfiable are of order 1. Then heuristic (incomplete) search algorithms have no way to disentangle, in a given finite time, the unsatisfiable instances from those which have very few solutions. So, in this region there appears an exponential critical slowing down which makes the search inefficient for any practical purpose.

The $K$-SAT problem is an NP-complete problem for $K \geqslant 3$. For its simplicity in formulation and versatility in encoding other NP problems, it has acquired a paradigmatic role for the study of NP complexity. It has been studied both directly (see e.g. [7, 8] or, for a review, [5] and references therein) or by the introduction of more 'handlable' variants which reproduce some of the peculiarities of $K$-SAT problems. One step in this direction is the introduction of the *hyper-SAT* model by Ricci-Tersenghi, Weight and Zecchina [9, 10].

Recently, a very powerful tool, derived from the theory of disordered systems (namely the cavity method at zero temperature, see [11]), has opened new perspectives in the study of these models, and the $K$-SAT problem in particular, leading to the determination of the phase diagram, and to the formulation of a new promising class of algorithms [12].

## 1.2. Presentation of the paper

In this paper, we shall discuss a very simple and exactly solvable model for the generation of random combinatorial problems. The model has deep analogies with the hyper-SAT model, and inherits from it some important characteristics:

- it is hard to handle by local search methods in the whole region of the parameter space;
- the problem may be solved in polynomial time by a simple global method (Gauss reduction) and therefore belongs to the class P;
- in a variant (unfrustrated model) one solution may be superimposed by construction.

The most important feature of our model is the fact that the statistical properties of the model are also exactly known, at finite and infinite size, thus yielding to a wide amount of analytical

results. For example, we rederive the exact expression for the probability distribution of the rank of an $M \times N$ random matrix in $GF(q)$, which is a classical topic of combinatorial theory (see [13, 14]), with a more 'physical' exact replica method. This result contains all the information about the average 'satisfiability' probability and much more. The methods used in the paper can be extended in order to calculate other interesting quantities, among which we remark the probability distribution of the magnetization.

Furthermore, the simple underlying geometry of the model allows for a clarifying graph-theoretical interpretation of the problem, which leads to the formulation of an exact duality of the system.

The paper is organized as follows:

(i) In this introductory section, we shortly describe some problems related to our model: random linear systems, the $K$-SAT model and the hyper-SAT model.
(ii) In section 2 we discuss some related fields of application, mainly error correcting codes theory and hypergraphs theory.
(iii) In section 3 we introduce the model, and describe its main properties with particular attention to the duality and relation derived from it.
(iv) In section 4 we exactly calculate the free energy distribution function at zero temperature, solving a realization of the 'classical moment problem'.
(v) In section 5 we rederive the results of the previous section, by means of a direct analysis of the Gauss algorithm which allows for a depiction of the structure underlying the ensemble of instances.
(vi) In section 6 we draw the conclusions of our work, and anticipate the directions of our forthcoming research.

### 1.3. A prologue: systems of random linear equations in $GF(q)$

We will start with a small prologue of elementary linear algebra, as it will provide a nice setting for our model.

Systems of linear equations, beyond the ordinary theory on $\mathbb{R}$ or $\mathbb{C}$, can be studied on more generic structures. For example, in the textbook [15], following the original work of Kovalenko and collaborators [16, 17], a certain number of questions about systems of linear equations on the Galois field $GF(q)$ are answered.

For each integer $q$, if it is not the power of a prime number, there exists no finite field of order $q$. Elsewhere, there exists exactly one finite field of order $q$ up to isomorphisms, named the *Galois field $GF(q)$*. We can concentrate on the easier case in which $q$ is a prime number. In this case the field $GF(q)$ simply consists of the integers in $\mathbb{Z}_q$ (i.e. the integers from 0 to $q - 1$), and sum and product are defined modulo $q$

$$(x_1 + x_2)_{GF(q)} \equiv x_1 + x_2 \quad (\text{mod } q) \qquad (x_1 + x_2)_{GF(q)} \in \mathbb{Z}_q \qquad (1a)$$
$$(x_1 \cdot x_2)_{GF(q)} \equiv x_1 \cdot x_2 \quad (\text{mod } q) \qquad (x_1 \cdot x_2)_{GF(q)} \in \mathbb{Z}_q. \qquad (1b)$$

The prototype of a linear system $A\boldsymbol{x} = \boldsymbol{v}$, or of a homogeneous linear system $A\boldsymbol{x} = \boldsymbol{0}$, in $GF(q)$, when the above definitions are assumed, maintains the same formal structure. It also preserves all of the fundamental results of linear algebra, like the existence of a Gauss algorithm for triangulating a matrix, or the fact that the rank plus the dimension of the kernel is equal to the number of rows of the matrix, which are derived in general for any field.

Statistical mechanics plays a role when we consider averages over a certain ensemble of matrices and vectors, in a sort of 'thermodynamic limit' on their dimension. For example, given two integers $N$ and $M$ and a prime number $q$, a good statistical ensemble is the set of

the couples $(A, v)$, with $A$ an $M \times N$ matrix and $v$ an $M$-dimensional vector, with elements randomly chosen in $\mathbb{Z}_q$. A statistical method would calculate, for a certain fixed value of $\gamma = N/M$, in the limit $N \to \infty$, the average satisfiability of this problem, i.e. for a random 'extraction' of an instance $(A, v)$ from the ensemble, which is the probability that the linear system in $GF(q)$, $Ax = v$, has at least one solution. In more detail, we could calculate the probability distribution for the dimension of the kernel of the linear application $A$, in the limit of infinite $N$ and $M$, with $N - M$ fixed.

This last point is exactly the main question answered by the Russian school in [15], which will be rederived as a preliminary result in this paper.

### 1.4. Nomenclature

In the framework of statistical mechanics, one frequently works with spin variables $\sigma = \pm 1$, while in the framework of random satisfiability, it is customary to work with Boolean variables $b \in \{\text{TRUE}, \text{FALSE}\}$. We will often switch between the two formulations, with the table of correspondences

$$
\begin{array}{|c|c|}
\hline
b & \sigma \\
\hline
\text{TRUE} & 1 \\
\text{FALSE} & -1 \\
\hline
\end{array}
\tag{2}
$$

So, we will consider a set of $N$ spin variables $\sigma_j$, with $j = 1, \ldots, N$ or, alternatively, a set of $N$ Boolean variables $b_j$. Each of the $2^N$ choices of the variables, denoted by $\sigma$ or $\mathbf{b}$, will be called a *configuration*.

A random satisfiability model is defined in terms of an ensemble of random *instances*, that is a collection of $M = N/\gamma$ Boolean conditions on the variables (called *clauses*), randomly chosen from a certain set. This set of possible clauses will primarily distinguish the different models, so it will be discussed further case by case.

Given a certain instance, if we have at least one configuration which simultaneously satisfies all the clauses[4], the instance is a *satisfiable* (SAT) one, else it is an *unsatisfiable* (UNSAT) one.

We will introduce a fictitious Hamiltonian for each instance. Our choice is to assign to a configuration an energy equal to the number of unsatisfied clauses. The low-temperature limit of the partition function calculated with this Hamiltonian is the number of solutions for a certain instance.

We will adopt the graphical convention for the logical operators $\wedge$ (and), $\vee$ (or) and $\dot{\vee}$ (xor).

### 1.5. The K-SAT model

In this model, given an integer $K$, the set of clauses $\{C\}$ is the set of all the $K$-uples of variable indices $n_1, \ldots, n_K$, with a Boolean assignment $s_j$ attached to each index. An instance in the ensemble is a subset $\mathcal{I}$ of the set of clauses, with cardinality $M$ and with no repetitions admitted.

A certain configuration will be a solution if

$$
\mathbf{b} \text{ solution} \quad \Longleftrightarrow \quad \bigwedge_{C_i \in \mathcal{I}} \left( \bigvee_{j \leqslant K} \left( s_j^{(i)} \dot{\vee} b_{n_j^{(i)}} \right) \right) = \text{TRUE}.
\tag{3}
$$

---

[4] It will be 'one non-trivial configuration' in the model with a superimposed solution.

If we call $\sigma_j$ the spin variables associated with the Boolean variables $b_j$, and $\tau_j^{(i)}$ the spin value associated with the Boolean assignments $s_j^{(i)}$, both following the table (2), a possible spin Hamiltonian which counts the unsatisfied clauses is

$$\mathcal{H}_{\mathcal{I}} = \frac{1}{2^K} \sum_{C_i \in \mathcal{I}} \prod_{j=1}^{K} \left( 1 - \tau_j^{(i)} \sigma_{n_j^{(i)}} \right). \tag{4}$$

Interpreted as a spin model Hamiltonian, the system has random diluted quenched couplings with a complicated pattern of interactions: for each $K$-uple of variables, if we have a $K$-spin coupling between them, then we also have couplings between all the subsets of the $K$ variables, with a choice of signs causing the $2^K$ levels of the $K$-variables 1-clause subsystem to split into a $2^K - 1$ multiplet of energy $-1$ and one level of energy $2^K - 1$.

### 1.6. The hyper-SAT model

This model, introduced by Ricci-Tersenghi, Weigt and Zecchina in [9, 10] as a variant of the $K$-SAT with a clearer spin structure, is defined as follows.

Given an integer $K$, the set of clauses $\{C\}$ is the set of all the $K$-uples of variable indices $n_1, \ldots, n_K$, with a single Boolean assignment $s$ attached. An instance in the ensemble is a subset $\mathcal{I}$ of this new set of clauses, with cardinality $M$ and no repetitions admitted.

A certain configuration will be a solution if

$$\mathbf{b} \text{ solution} \iff \bigwedge_{C_i \in \mathcal{I}} \left( \left( \overset{.}{\bigvee_{j \leqslant K}} b_{n_j^{(i)}} \right) \dot{\vee} s^{(i)} \right) = \text{TRUE}. \tag{5}$$

If we call $\tau^{(i)}$ the spin value associated with the Boolean assignments $s^{(i)}$, again following the table (2), a possible spin Hamiltonian which counts the unsatisfied clauses is

$$\mathcal{H}_{\mathcal{I}} = \frac{1}{2} \sum_{C_i \in \mathcal{I}} \left( 1 - \tau^{(i)} \prod_{j=1}^{K} \sigma_{n_j^{(i)}} \right). \tag{6}$$

This ensemble of Hamiltonians has the form of a diluted mean field $K$-spin disordered system, that is a spin model with $N$ variables and $M$ quenched $K$-spin couplings, randomly chosen between all the $K$-uples of variables, and with values randomly chosen in $\{-1, 1\}$.

A variant of the problem, named the 'unfrustrated' model, is the restriction of the clauses set only to the clauses with $s = \text{TRUE}$ (that is a ferromagnetic diluted $K$-spin system). In this case, we will have the obvious 'ferromagnetic' solution $b_j = \text{TRUE}$ (or $\sigma_i = +1$) for each instance.

### 1.7. The hyper-SAT model as a random linear system

The hyper-SAT model can be re-interpreted in the framework of random linear systems we introduced in the prologue. Given a hyper-SAT instance with $N$ variables and $M$ clauses, we can introduce an $M \times N$ matrix $J_{ij}$ in the field $GF(2)$, such that $J_{ij} = 1$ if the variable $b_j$ appears in the $i$th clause, and $J_{ij} = 0$ elsewhere. With the expanded table of correspondences

| $b$ | $\sigma$ | $x$ |
|-------|------|-----|
| TRUE | 1 | 0 |
| FALSE | $-1$ | 1 |

$$\tag{7}$$

we have the composition correspondences

$$b_1 \dot{\vee} b_2 = \sigma_1 \sigma_2 = (x_1 + x_2)_{GF(2)}.$$

So, for each clause $C_i$,

$$\dot{\bigvee_{j|J_{ij}=1}} b_j = \prod_j \sigma_j^{J_{ij}} = \left(\sum_j J_{ij}x_j\right)_{GF(2)}.$$

Therefore, at the end, we can analyse the problem in the more handlable form $Jx = v$ (frustrated model) or $Jx = 0$ (unfrustrated model).

The shift to $K$-spin models will be given by correspondence (7), so the Hamiltonian in the spin formulation will be

$$\mathcal{H}_{J,v}(\sigma) = \frac{1}{2}\sum_i\left(1 - (-1)^{v_i}\prod_j \sigma_j^{J_{ij}}\right) \tag{8}$$

while in the linear system formulation it will be

$$\mathcal{H}_{J,v}(x) = \sum_i\left(1 - \delta_2\left(\sum_j J_{ij}x_j - v_i\right)\right) \tag{9}$$

with $\delta_2(n) = 1$ if $n$ is even, and 0 if $n$ is odd. As in the following we will generalize the system to $q$-state variables, we will introduce a $\delta_q$ function with the definition

$$\delta_q(n) = \begin{cases} 1 & n \equiv 0 \pmod{q} \\ 0 & n \equiv 0 \pmod{q}. \end{cases} \tag{10}$$

Beyond the correspondence between a specific hyper-SAT instance and a specific random linear system, we remark that the choice of the ensemble is crucial for the statistical properties of the system. The statistical ensemble of the hyper-SAT is very different to the random ensemble we suggested as an example in the prologue; as in the hyper-SAT we have exactly $K$ non-zero terms in each row of $J$.

## 2. Random satisfiability, graph theory and random linear codes

### 2.1. Linear codes in information theory

The main problem in information theory consists in signal reconstruction when a sequence of bits is transmitted through a 'noisy' channel, that is a channel where transmission errors can occur with a certain probability. Just to fix the ideas, we will concentrate on the so-called *symmetric noise channel*, in which each bit has a probability $(1 - p)$ of being transmitted unaltered and a probability $p$ of being transmitted as its negate:

|   | 0 | 1 |
|---|---|---|
| 0 | $1-p$ | $p$ |
| 1 | $p$ | $1-p$ |

When $p = 0$ no error can occur, and transmission is taken for granted. When $p = \frac{1}{2}$ the output signal is totally decorrelated from the input signal, and no inference is possible even in principle. In the intermediate cases, the task is to use the information transmitted through the channel in the form of the partial correlation between input and output signals, to obtain the safe transmission of a shorter signal.

This problem can be easily generalized to a generic 'alphabet', with characters in $\mathbb{Z}_q$ (i.e. instead of having bits in $\{0, 1\}$, we have variables in $\{0, 1, \ldots, q-1\}$). A generalization of the symmetric noise channel defined above could be as follows: each character has a probability

$(1 - p)$ of being transmitted unaltered and a probability $\frac{p}{q-1}$ of being converted to each of the other characters. For example, when $q = 3$ we have

|   | 0 | 1 | 2 |
|---|---|---|---|
| 0 | $1 - p$ | $p/2$ | $p/2$ |
| 1 | $p/2$ | $1 - p$ | $p/2$ |
| 2 | $p/2$ | $p/2$ | $1 - p$ |

In this case, the totally decorrelated case occurs for $p = \frac{q-1}{q}$.

A *code* is an algorithm which fulfils the task above, replacing the interesting signal with a redundant but error-robust encoding of it. The ratio between the length of the original and of the encoded signals is called the *rate* of the code. A critical rate exists such that, in the limit of infinite length of the message, the decoding procedure almost always succeeds in reconstructing the original signal from the noise-distorted transmitted stream. In most cases, the probability of success as a function of the rate jumps to 1 in an abrupt way, typical of critical phenomena (reconstructable–unreconstructable transition).

The quality of a code is mainly given by two parameters: the critical *rate* at given error probability and the computational complexity class of the encoding and decoding algorithms.

A family of codes, called *linear codes* [18], is closely related to random satisfiability problems and certain spin glass models in statistical mechanics [19], and a deep insight into the reconstructable–unreconstructable transition can be derived from the study of the phase diagrams of the related spin glass models [20]. A well-known example of linear code, called *Varshamov's linear code* or *RLC* (*random linear code*) [21, 22] was the first discovered code to reach the *Gilbert–Varshamov bound* on the rate of the code. The Varshamov code is closely related to our model.

Suppose we have a random linear system $Jx = v$ in $GF(q)$, with $M$ equations, in $N > M$ variables, and suppose the kernel of the linear application $J$ has the minimum possible dimension $L = N - M$. This linear system can be used for the error correcting transmission of a message of length $L$, on an alphabet of $q$ characters. First, we introduce some definitions:

(i) The possible original messages are the vectors $y$ of the linear space $\mathbb{Z}_q{}^L$.
(ii) The encoding map $\mathcal{E}(\mathbb{Z}_q{}^L \to \mathbb{Z}_q{}^N)$, which encodes the original $L$-bits message in the redundant $N$-bits codeword, is defined with the aid of the linear application $J$. If we choose a basis for the kernel

$$\ker(J) = \mathrm{span}(s_{(1)}, \ldots, s_{(L)}) \tag{11}$$

the map will be defined as

$$\mathcal{E} : y \to x(y) = \sum_{i=1}^{L} y_i s_{(i)}. \tag{12}$$

(iii) A distance on the codeword space $\mathbb{Z}_q{}^N$ will be introduced as

$$d(x, x') = \sum_{j=1}^{N} [1 - \delta(x_j - x'_j)]. \tag{13}$$

Then, for the transmission of a certain message, we follow the procedure:

(i) The original message is a vector $y_{\mathrm{mes}} \in \mathbb{Z}_q{}^L$.
(ii) We use the map $\mathcal{E}$ for the encoding. The original codeword is $x_{\mathrm{mes}} = \mathcal{E}(y_{\mathrm{mes}})$.

(iii) The codeword is transmitted through the noisy channel. The transmitted codeword will
      be a vector $\boldsymbol{x}_{\text{trans}}$. The probability distribution for the transmitted codeword depends on
      the kind of noise we consider. In the case of symmetric noise, we have

$$p(\boldsymbol{x}) = (1-p)^N \left( \frac{p}{(1-p)(q-1)} \right)^{d(\boldsymbol{x}, \boldsymbol{x}_{\text{mes}})}. \tag{14}$$

(iv) For the reconstruction of the signal could be used, for example, the minimum distance
     decoding criterion, that is, if we call $\boldsymbol{x}_{\text{rec}}$ the minimum of the function $d(\boldsymbol{x}, \boldsymbol{x}_{\text{trans}})$ in
     the image of the encoding map $\mathcal{E}(\mathbb{Z}_q{}^L)$, we choose as the reconstructed message $\boldsymbol{y}_{\text{rec}}$ the
     vector $\boldsymbol{y}_{\text{rec}} = \mathcal{E}^{-1}(\boldsymbol{x}_{\text{rec}})$.

The interesting question is, in the 'thermodynamic' limit of long signals $N \to \infty$, with the
*rate* $R = L/N$ fixed, what is the probability that the reconstructed message $\boldsymbol{y}_{\text{rec}}$ is identical
to the original message $\boldsymbol{y}_{\text{mes}}$, as a function of the noise parameter $p$ and of the rate $R$.

A statistical study [20] shows the tricky fact that this probability tends to 1 if $R < R_c(p)$,
and it tends to 0 if $R > R_c(p)$, and that the critical line $R_c(p)$ is exactly the Shannon bound
for the capacity of a symmetric noise channel [18, 22].

## 2.2. Random graphs and statistics

Another mathematical framework in which this kind of satisfiability problem is well depicted
is hypergraph theory. In these sections, we will shortly introduce some notation which will be
useful in the following, first starting with ordinary graph theory, then jumping to hypergraphs
(for an introductory bibliography, see [23, 24]).

Given

$$V = \{p_1, \ldots, p_N\}$$

a finite set of cardinality $N$, called the set of 'vertices', we define the set of all the non-ordered
couples of vertices in $V$,

$$\mathcal{E} = \left\{ \left( p_{n_1}, p_{n_2} \right) \right\} \qquad n_1 \neq n_2$$

as the set of 'edges'. For each subset $E \subseteq \mathcal{E}$ of cardinality $M$, the datum of the set of vertices
and of this set is called the *graph* $\Gamma(V, E)$, with $N$ vertices and $M$ edges.

Given a vertex $p_i$, its *coordination* (or *degree*) is defined as the number of edges $E_j \in E$
such that $p_i \in E_j$. A *path* between the vertices $p_a$ and $p_b$ is a sequence of vertices and edges
in $E$,

$$(p_a, E_1, p_1, E_2, p_2, \ldots, E_L, p_b)$$

such that $p_i \in E_i, E_{i+1}$; the number of edges $L$ is called its *length*. A graph is said to be
*connected* if, for each couple of vertices, there exists a path which joins them.

A non-trivial closed path is usually called a *loop*. More generically, we will adopt the
'physical' convention of using the term loop for any subset of the set of edges $E_L \subseteq E$ such
that the graph, restricted to this set $\Gamma(V, E_L)$, has only vertices of even (eventually zero)
coordination.

In the limit of large $N$ and $M$, we can introduce an ensemble of graphs, giving a probability
measure over the possible sets of edges, and we can calculate statistical averages of interesting
quantities. For example, an ensemble could be that in which only connected graphs with $M$
edges are admitted, and no repetitions of edges are allowed.

### 2.3. Hypergraphs and K-SAT problems

A generalization of graph theory concerns objects in which the set of possible edges $\mathcal{E}$ is replaced by the set of all non-ordered $K$-uples of vertices (typically $K \geqslant 3$), called *hyper-edges*:

$$\mathcal{E}_K = \left\{ \left( p_{n_1}, p_{n_2}, \ldots, p_{n_K} \right) \right\} \qquad n_1 \neq n_2 \neq \cdots \neq n_K.$$

The datum of the set of vertices and of a subset of admitted hyper-edges will be called a *hypergraph*.

The notions of coordination, path, connectivity and loop, as we defined them in the previous section, immediately generalize to this case. In particular, the notion of *hyperloop*, as a subset $E_L \subseteq E$ such that the relative hypergraph $\Gamma(V, E_L)$ has only vertices of even coordination, will be crucial in the following.

The ensemble of connected hypergraphs with $M$ hyper-edges, with repetitions allowed, is the proper graphical scheme for describing the topological features of ordinary $K$-SAT problems, and also of the hyper-SAT, although, as we have already mentioned in the previous section, we should also imagine some Boolean assignments attached to the hyper-edges (one for each vertex of each hyper-edge in the case of $K$-SAT, just one per hyper-edge in the case of hyper-SAT and no assignments at all in the case of unfrustrated hyper-SAT[5]).

Another possible variant could be that in which hyper-edges with any number of vertices are admitted

$$\mathcal{E}_\forall = \bigcup_{K=0}^{N} \mathcal{E}_K.$$

The ensemble of all possible hypergraphs of $M$ hyper-edges in $\mathcal{E}_\forall$, with homogeneous measure and no repetitions allowed, is the one which describes the model we will define in the next section.

### 2.4. Group structure of the hyper-SAT and the role of hyperloops in satisfiability

In the hyper-SAT model, the configurations show a $\mathbb{Z}_2^N$ group structure. The group operation is the term-by-term sum in the linear system formulation or the term-by-term product in the spin formulation, that is

$$x_i^{(ab)} \equiv x_i^{(a)} + x_i^{(b)} \quad (\text{mod } 2) \qquad \sigma_i^{(ab)} = \sigma_i^{(a)} \sigma_i^{(b)}. \tag{15}$$

In the unfrustrated model, the set of solutions is a subgroup $\mathbb{Z}_2^g$, with $g$ generators, of the configuration group.

The fact that in the unfrustrated hyper-SAT the set of solutions is a subgroup is an immediate consequence of the fact that the kernel of a linear application in $GF(2)$ is a subspace $\mathbb{Z}_2^g$ of the vector space $\mathbb{Z}_2^N$.

Let us go back to the formulation of a hyper-SAT instance in terms of a system of linear equations in $GF(2)$:

$$\sum_{j=1}^{K} x_{n_j^{(i)}} = v_i.$$

We can graphically identify each variable with a vertex and each clause with a hyper-edge joining the vertices which appear in it.

---

[5] This immediate correspondence with the hypergraph interpretation is the reason for the name.

Each subset of equations in the system can be identified with the sub-hypergraph (in the following just hypergraph) of the original one, obtained by restriction to the corresponding hyper-edges. The sum of the equations in the subset will be called the *equation associated with the hypergraph*.

Note that hypergraphs also have a group structure $\mathbb{Z}_2{}^M$: given two hypergraphs, the group operation is the disjoint union of the two hypergraphs, from the graphical point of view, and the sum of the two corresponding equations, from the algebraic point of view. The identity is the empty set.

The set of hyperloops is a subgroup $\mathbb{Z}_2{}^{g'}$ of the hypergraph group. When the subgraph is a hyperloop, the associated equation is a relation involving no variables: in the unfrustrated case, it is just a trivial identity; in the frustrated case, it is a configuration-independent characteristic of the instance which could eventually determine *a priori* unsatisfiability. Because of randomness in the choice of the $s_{(j)}$ for instance, this will happen with $\frac{1}{2}$ probability for each linearly independent hyperloop. The number of linearly independent elements in a group $\mathbb{Z}_2{}^{g'}$ is just $g'$.

So the sentence in [9] which explains the role of hyperloops in the frustrated model,

> "As soon as one hyperloop arises in the hypergraph, half the formulas become unsatisfiable . . . "

can also be interpreted as 'As soon as the average number of generators for the hyperloop group, $g'$, increases by 1, the average number of solutions would decrease by a factor $\frac{1}{2}$ (if we were working with the frustrated model), so the average number of generators of the group of solutions, $g$, decreases by 1'.

We can see that an intriguing symmetry of roles between solutions and hyperloops seems to arise. This symmetry of roles is clearer in the linear system formulation, as we have that a homogeneous solution is a solution of $J\boldsymbol{x} = \boldsymbol{0}$, while a hyperloop is a solution of $J^T\boldsymbol{y} = \boldsymbol{0}$. In our model, the analytical solvability will provide a clarifying re-interpretation of these heuristic arguments.

## 3. The model

One difficulty in an exact analytic approach to the hyper-SAT is that the ensemble of formulae is *not* symmetric: for an admitted clause, the number of 1 in each row must be exactly $K$, while there is no such constraint on the columns. This fact introduces difficult-to-handle correlations on the values of the entries in the grid formulation.

In this section, we introduce a model defined on a larger ensemble, which cancels correlations between the entries of the matrix, and fully exploits the symmetry of roles between solutions and hyperloops we suggested in section 2.4.

### 3.1. Description of the model

The model we define here is a random satisfiability problem. So it will be defined by an ensemble of *instances*, with a probability measure on it (the possible realizations of the problem, with the associated probability that such a realization occurs), an ensemble of *configurations* (the potential solutions) and a notion of *satisfiability*.

Given the integer parameter $q$, which must be a prime number, and the 'noise' parameter $p \in [0, 1]$ (as it will have a role similar to that of the noise parameter of error correcting codes, see section 2.1), the ensemble of the instances is the product set of the set of all the $M \times N$ matrices $J_{ij}$, filled with integers in $\mathbb{Z}_q$, each matrix being equiprobable, and the $M$-dimensional

vector space $\mathbb{Z}_q{}^M$ where the probability for a vector $v$ to be extracted is dumped according to the number of its non-zero entries. More precisely

$$\mu(J, v) = q^{-MN} \cdot (1 - p)^M \left( \frac{p}{(1 - p)(q - 1)} \right)^{\sum_{i=1}^{M}(1 - \delta_q(v_i))}$$

where we used the short notation of equation (10); the parameter $p$ has been chosen to reproduce the average fraction of non-zero elements in $v$. The parameter $\gamma = N/M$ will determine the average solvability of the problem: large values of $\gamma$ correspond to weakly constrained systems, and a typical instance will be solvable while, in contrast, small values of $\gamma$ correspond to highly constrained systems, and a typical instance will be unsolvable (or will have only the superimposed solution in the unfrustrated case, see below).

The ensemble of the configurations is the $N$-dimensional vector space $\mathbb{Z}_q{}^N$. A certain configuration $x$ is said to be a solution if it solves the linear system

$$J x = v.$$

The particular case of $p = 0$ is called *unfrustrated* as we have a superimposed 'ferromagnetic' solution $x$ for each instance $J$. When $p \neq 0$ we will deal with the *frustrated* version. When $p = (q - 1)/q$, the model will be called *totally frustrated*, as the probability measure is flat also for the inhomogeneous terms $v$.

A particular realization of the problem (*instance*) will be called *satisfiable* (SAT) if we have at least one solution in the frustrated model, and at least one more solution, after the superimposed one, in the unfrustrated model. Elsewhere, the instance will be called *unsatisfiable* (UNSAT).

Although in this part we only study the average solvability of the problem, it is customary for a statistical mechanical approach to introduce a fictitious Hamiltonian, such that the solutions correspond to the states with zero energy, and all the other configurations have a rigorously positive energy, study the thermodynamic of this new system, and recover the original model in the zero temperature limit.

Obviously, there is a large arbitrarity in the choice of the Hamiltonian. Just to fix the ideas, we will concentrate on the simplest choice, the one in which the Hamiltonian just 'counts' the number of violated clauses:

$$\mathcal{H}_J(x) = \sum_{i=1}^{M} \left( 1 - \delta_q \left( \sum_{j=1}^{N} J_{ij} x_j - v_i \right) \right). \tag{16}$$

A more familiar spin reformulation is

$$\mathcal{H}_J(\sigma) = \frac{1}{2} \sum_{i=1}^{M} \left( 1 - \delta \left( s_i \prod_{j=1}^{N} \sigma_j^{J_{ij}} - 1 \right) \right) \tag{17}$$

where the correspondence is

$$\sigma_j = e^{\frac{2\pi i}{q} x_j} \qquad s_i = e^{\frac{2\pi i}{q} v_i}.$$

In the homogeneous case every $s_i = 1$.

This model shares with the random energy model [25–29], the property that the energy of different states is (almost always) uncorrelated. More precisely, it is true that, if $(E, E_1, \ldots, E_k)$ are respectively the energies of the states $(\sigma, \sigma^{(1)}, \ldots, \sigma^{(k)})$, calculated with a certain instance, for each couple of functions $f(x), g(x_1, \ldots, x_k)$ we have

$$\overline{f(E) g(E_1, \ldots, E_k)} = \overline{f(E)} \cdot \overline{g(E_1, \ldots, E_k)}$$

if $\sigma \notin \mathrm{span}(\sigma^{(1)}, \ldots, \sigma^{(k)})$, and the fraction of states for which this last condition is verified is about $1 - q^{-N+k}$.

### 3.2. The underlying group structure

As we anticipated in section 2.4, a nice feature of our model, which is common also to the hyper-SAT model and to the random linear system problem, is that the configurations show a $G_\sigma = \mathbb{Z}_q{}^N$ group structure.

The group operation in the linear system formulation is the term-by-term sum of the configuration entries:

$$x_i^{(ab)} \equiv x_i^{(a)} + x_i^{(b)} \pmod{q}. \tag{18}$$

In the unfrustrated model, the set of solutions is a subgroup $G_S = \mathbb{Z}_q{}^g$, with $g$ generators, of the configuration group.

In the frustrated model, given an instance $(J, \boldsymbol{v})$, we define the *associated homogeneous problem* in which we search for solutions of the homogeneous system $J\boldsymbol{x} = \boldsymbol{0}$ (that is the problem in which we turn all the Boolean assignments $s_{(j)}$ to TRUE). As we mentioned above, the set of solutions of this homogeneous problem is a subgroup $G_S = \mathbb{Z}_q{}^g$. If we have at least one solution $\widehat{\boldsymbol{x}}$ for the original problem, the set of solutions of the original problem is the lateral class $\widehat{\boldsymbol{x}} + G_S$.

These statements have a clarifying correspondence in the random linear system on $GF(q)$ framework: the space of admitted configurations $\boldsymbol{x}$ is the discrete torus $\mathbb{Z}_q{}^N$ with , as a group operation, the sum of discrete vectors quotiented by periodicity. The fact that in the unfrustrated hyper-SAT the set of solutions is a subgroup, and in the frustrated case it is a lateral class, is in correspondence with the fact that the set of solutions of a linear system on $GF(q)$ is a subspace $\mathbb{Z}_q{}^g$ of the vector space $\mathbb{Z}_q{}^N$ in the homogeneous case, and an affine subspace in the inhomogeneous case.

The model naturally shows also a gauge invariance. As the Hamiltonian is of the form

$$\mathcal{H}_{J,v}(\boldsymbol{x}) = \mathcal{H}(J\boldsymbol{x} - \boldsymbol{v}) = \mathcal{H}(\boldsymbol{w})$$

it will remain unchanged for transformations which keep the vector $\boldsymbol{w}$ unchanged, that are the changes of basis and the translations on the torus of configurations $\mathbb{Z}_q{}^N$. The more general gauge transformation $\mathcal{G}$ of this kind is

$$\mathcal{G}(B, \xi) : \begin{cases} \boldsymbol{x} \to \boldsymbol{x}' = B(\boldsymbol{x} + \xi) \\ J \to J' = JB^{-1} \\ \boldsymbol{v} \to \boldsymbol{v}' = \boldsymbol{v} + J\xi \end{cases} \tag{19}$$

where $B \in GL(N, GF(q))$ (change of basis) and $\xi \in \mathbb{Z}_q{}^N$ (translation).

This gauge structure is purely a consequence of the vectorial form of the combination $\boldsymbol{w} = J\boldsymbol{x} - \boldsymbol{v}$. We can derive a larger gauge group from the peculiar form of the Hamiltonian

$$\mathcal{H}(\boldsymbol{w}) = \sum_{i=1}^{M} (1 - \delta_q(w_i)).$$

From this form, we deduce that vectors $\boldsymbol{w}_1, \boldsymbol{w}_2$, differing only for a permutation of the elements and for the product of each entry for a non-null value, have the same energy. The group $\{1, 2, \ldots, q-1\}$, with the product of $GF(q)$ ($q$ a prime number), is called $\mathbb{Z}_q^*$. We have $M$ copies acting independently on the $M$ entries of $\boldsymbol{w}$. Then we have the permutation group of $M$ elements, $\mathfrak{S}_M$. The group whose elements are a permutation of the $M$ entries, followed by the product of each entry for an element of $\mathbb{Z}_q^*$, is the group[6]

$$(\mathbb{Z}_q^*)^{\otimes M} \ltimes \mathfrak{S}_M.$$

---

[6] The symbol $\ltimes$ means that $(\mathbb{Z}_q^*)^{\otimes M}$ is a normal subgroup of the whole group.

We call $q$-Perm $(M, GF(q))$ the set of $M \times M$ matrices which is the $M$-dimensional representation of this group (that is the matrices with exactly one non-zero element in each row and in each column. Note that the normal subgroup is the subgroup of diagonal matrices). So the most general gauge transformation $\mathcal{G}$ is

$$\mathcal{G}(S, B, \xi) : \begin{cases} \boldsymbol{x} \to \boldsymbol{x}' = B(\boldsymbol{x} + \xi) \\ J \to J' = SJB^{-1} \\ \boldsymbol{v} \to \boldsymbol{v}' = S\boldsymbol{v} + SJ\xi \end{cases} \tag{20}$$

where $S \in q$-Perm $(M, GF(q))$ (permutation and product on the entries of $\boldsymbol{w}$), $B \in GL(N, GF(q))$ (change of basis) and $\xi \in \mathbb{Z}_q{}^N$ (translation).

In the case of the unfrustrated model, in which we impose the vector $\boldsymbol{v}$ to be the null vector, we will restrict to $\xi = \boldsymbol{0}$ gauge transformations.

A common tool when dealing with systems with quenched disorder is the replica trick. This method leads to the calculation of annealed averages over a replicated system, that is, if the number of replicas is $m$, the space of configurations is the (tensor) product of $m$ copies of the original one, and the Hamiltonian of the replicated system $\mathcal{H}^{(m)}$ is the sum of the Hamiltonians calculated on the different configurations of each replica:

$$\mathcal{H}_J^{(m)}(\boldsymbol{x}_1, \dots, \boldsymbol{x}_m) = \sum_{a=1}^m \mathcal{H}_J(\boldsymbol{x}_a). \tag{21}$$

We are interested in the 'annealed' partition function of this system, that is the sum of the Gibbs factor $e^{-\beta \mathcal{H}_J^{(m)}(\boldsymbol{x}_1, \dots, \boldsymbol{x}_m)}$ on all the configurations $(\boldsymbol{x}_1, \dots, \boldsymbol{x}_m)$ *and* all the instances $J$.

In this frame, and in our peculiar case, a new symmetry arises between replicas and clauses, which enlarges the gauge group. If we call

$$x_j{}^{(a)} = \widetilde{J}_{ja}$$

and defining

$$W_{ia} = \sum_{j=1}^N J_{ij} \widetilde{J}_{ja}$$

we have for the Hamiltonian the symmetric expression

$$\mathcal{H}^{(m)}(W) = \sum_{a=1}^m \sum_{i=1}^M (1 - \delta_q(W_{ia})). \tag{22}$$

From this expression, we see that we could permutate the $mM$ elements of the matrix $W$ in an arbitrary way, and multiply each of them for an independent element in $\mathbb{Z}_q^*$. Such a general class of transformations does not have a clear action on the physical variables $(J, \widetilde{J})$. We could restrict ourselves to the action of a permutation matrix for the clauses $S \in q$-Perm $(M, GF(q))$ on the left, and a permutation matrix for the replicas $S' \in q$-Perm $(m, GF(q))$ on the right. So we have that, as a consequence of

$$\mathcal{H}^{(m)}(W) = \mathcal{H}^{(m)}(SWS')$$

the general gauge transformation is of the form $\mathcal{G}$,

$$\mathcal{G}(S, B, S') : \begin{cases} J^T \to J'^T = SJ^T B^{-1} \\ \widetilde{J} \to \widetilde{J}' = B\widetilde{J}S' \end{cases} \tag{23}$$

where $S \in q$-Perm $(M, GF(q))$ (permutation and product on the clauses), $B \in GL(N, GF(q))$ (change of basis on the variables space) and $S' \in q$-Perm $(m, GF(q))$ (permutation and product on the replicas).

### 3.3. The clause–variable (C–V) duality

For the unfrustrated model, we can state a notion of duality which formalizes the heuristic arguments of section 2.4 on the symmetry of roles between clauses and variables. We will call this duality the *C–V duality*.

The C–V duality maps an $M \times N$ instance of the problem in an $N \times M$ instance. The matrix $J'$ associated with the new instance is simply the transpose of the original one $J' = J^T$. As the duality brings the density parameter $\gamma$ in $\gamma^{-1}$, it has a 'fixed point' in this parameter space at $\gamma_c = 1$. We will see that this density is also the critical density for the SAT–UNSAT transition.

We anticipate that this duality is indeed much more general, and extends to a wide class of statistical mechanics models [30], which includes the finite-temperature treatment of our model (see [31]).

In a graphical representation, the dual of a generic configuration is a generic hypergraph, and the dual of a solution is a hyperloop.

More generally, we can consider all the degeneracy numbers of the energy levels in the two problems

$$S^{(n)} = \sum_{\sigma} \delta(\mathcal{H}_J(\sigma), n) \qquad H^{(m)} = \sum_{\tau} \delta(\mathcal{H}_{J^T}(\tau), m) \tag{24}$$

which will define sort of two dual partition functions associated with each instance,

$$Z(J) = \sum_{n} S^{(n)} e^{-\beta n} \qquad Z^{(T)}(J) = \sum_{m} H^{(m)} e^{-\beta m}. \tag{25}$$

A first easy statement derived from duality could be

$$Z(J^T) = Z^{(T)}(J) \qquad Z^{(T)}(J^T) = Z(J). \tag{26}$$

But the duality provides much stronger relations between arbitrary functions in the operators $S^{(n)}$ and $H^{(m)}$,

$$f(S^{(n)}(J), H^{(m)}(J)) = f(H^{(n)}(J^T), S^{(m)}(J^T)) \tag{27}$$

and then, averaging over the ensemble of instances,

$$\overline{f(S^{(n)}, H^{(m)})}_{N,M} = \overline{f(H^{(n)}, S^{(m)})}_{M,N} \qquad \text{finite-size} \tag{28}$$

$$\overline{f(S^{(n)}, H^{(m)})}_{\gamma} = \overline{f(H^{(n)}, S^{(m)})}_{\gamma^{-1}} \qquad \text{infinite size limit.} \tag{29}$$

In the following, we will often omit the superscripts in $S^{(0)}$ and $H^{(0)}$, referring to them simply as $S$ and $H$.

### 3.4. Expectation values of S and H moments

In what follows, we will give elementary combinatorial derivations of the expectation values for a certain set of polynomials in $S$ and $H$.

As the formal definition of the number of solutions $S$ is

$$S(J) = \sum_{\sigma} \delta(\mathcal{H}_J(\sigma), 0)$$

a nice feature of this operator is that averages over $\sigma$ and $J$ easily commute, so that

$$\overline{S} = \overline{\sum_{\sigma} \delta(\mathcal{H}_J(\sigma), 0)} = \sum_{\sigma} \overline{\delta(\mathcal{H}_J(\sigma), 0)}. \tag{30}$$

We will call $\mathcal{N}_k(q^g)$ the number of subsets of order $k$ of linearly independent elements in a group $\mathbb{Z}_q{}^g$. This quantity can be deduced from a simple inductive argument: if we already have $k-1$ linearly independent elements, we can choose the $k$th element between the $q^g$ elements of the whole $\mathbb{Z}_q{}^g$, unless it is already one of the $q^{k-1}$ elements of the $\mathbb{Z}_q{}^{k-1}$ generated subgroup, so we must have

$$\mathcal{N}_k(q^g) = \mathcal{N}_{k-1}(q^g) \cdot (q^g - q^{k-1}).$$

For the case $k = 0$, we have only the empty set, so $\mathcal{N}_0(q^g) = 1$. This implies that

$$\mathcal{N}_k(q^g) = \prod_{i<k}(q^g - q^i). \tag{31}$$

We will introduce a short notation for the recurrent products,

$$\mathcal{R}_{a,b} = \prod_{j=a+1}^{b}(1 - q^{-j}) \qquad \begin{bmatrix} n \\ m \end{bmatrix} = \frac{\mathcal{R}_{n-m,n}}{\mathcal{R}_{0,m}} = \frac{\mathcal{R}_{0,n}}{\mathcal{R}_{0,m}\mathcal{R}_{0,n-m}} \tag{32}$$

so that $\mathcal{N}_k(q^g) = q^{kg}\mathcal{R}_{g-k,g}$. The expressions $\begin{bmatrix} n \\ m \end{bmatrix}$ are known in the literature as 'Gauss polynomials'; in this specific case they are polynomials in the variable $q^{-1}$ (see appendix A).

A nice feature of our ensemble is that a change of basis does not affect the probability distribution of the ensemble of $J$ grids, as it is flat over all the possible grids.

Given a certain set of $k$ linearly independent configurations $\{\sigma^{(1)}, \ldots, \sigma^{(k)}\}$, we can calculate the probability that all of them are solutions of a given instance, averaged over the instances.

Using the stability of the ensemble under change of basis, we could pass to a new basis in which the first $k$ elements are the $\sigma^{(1)}, \ldots, \sigma^{(k)}$. As they are linearly independent, this change of basis is not singular. It is easily seen that they are all solutions of the instance if and only if the first $k$ rows are filled with zeros, so we have

$$\overline{\prod_{i=1,\ldots,k} \delta(\mathcal{H}_J(\sigma^{(i)}), 0)} = (q^{-M})^k. \tag{33}$$

Collecting equations (30), (31) and (33) we get

$$\overline{\prod_{i<k}(S - q^i)} = \overline{\mathcal{N}_k(S)} = \overline{\sum_{\{\sigma^{(1)},\ldots,\sigma^{(k)}\}} \prod_i \delta(\mathcal{H}_J(\sigma^{(i)}), 0)}$$

$$= \sum_{\{\sigma^{(1)},\ldots,\sigma^{(k)}\}} \overline{\prod_i \delta(\mathcal{H}_J(\sigma^{(i)}), 0)}$$

$$= \mathcal{N}_k(q^N)q^{-kM} = q^{k(N-M)}\mathcal{R}_{N-k,N}. \tag{34}$$

The dual of equation (34) is

$$\overline{\prod_{i<h}(H - q^i)} = \overline{\mathcal{N}_h(H)} = \mathcal{N}_h(q^M)q^{-hN} = q^{-h(N-M)}\mathcal{R}_{M-h,M}. \tag{35}$$

A more complex question we can answer is the probability that $k$ linearly independent configurations $\{\sigma^{(1)}, \ldots, \sigma^{(k)}\}$ are solutions of a certain instance $J$, and $h$ linearly independent configurations $\{\tau^{(1)}, \ldots, \tau^{(h)}\}$ are solutions of its dual instance $J^T$, at the same time. With a similar 'double' change of basis, we find that both the first $k$ rows and the first $h$ columns should be filled with zeros, so

$$\overline{\prod_{i=1,\ldots,k} \delta(\mathcal{H}_J(\sigma^{(i)}), 0) \prod_{j=1,\ldots,h} \delta(\mathcal{H}_{J^T}(\tau^{(j)}), 0)} = q^{kh-kM-hN}. \tag{36}$$

If we collect equations (30), (31) and (36) we obtain that

$$\overline{\prod_{i<k}(S-q^i)\prod_{j<h}(H-q^j)} = \overline{\mathcal{N}_k(S)\mathcal{N}_h(H)} = \mathcal{N}_k(q^N)\mathcal{N}_h(q^M)q^{kh-kM-hN}$$

$$= q^{kh}q^{(k-h)(N-M)}\mathcal{R}_{N-k,N}\mathcal{R}_{M-h,M}. \tag{37}$$

Note that $\mathcal{R}_{N-k,N}$ and $\mathcal{R}_{M-h,M}$ are exponentially depressed finite-size corrections in the thermodynamic limit. It is also worth noting that relation (37) for $h=k=1$ reads

$$\overline{(S-1)(H-1)} = q(1-q^{-N})(1-q^{-M}) \simeq q$$

that is the average of the product between the numbers of non-trivial solutions and of non-trivial hyperloops, in the thermodynamic limit, tends to a constant, independent of the variable-to-clauses ratio $\gamma = N/M$. This is the precise formalization of the heuristic solution–hyperloop correspondence of section 2.4.

### 3.5. Expectation values for excited state degeneracy

In a way similar to the previous calculation, we could find the expectation value for expressions involving the number of excited states $S^{(n)}$, with $n$ violated clauses, or by duality, the number of hypergraphs $H^{(n)}$ with $n$ variables appearing in the associated equation.

A quantity which has been considered important in the literature [9] is $H^{(1)}$, as such a kind of hypergraph fixes the value of a variable, which will be part of the so-called 'backbone', defined as the subset of variables which take the same value on each solution. As one trivial solution is the ferromagnetic one, the size of the backbone is in some sense the magnetization of the system.

The average probability that a certain variable configuration is an $n$-level excited state is exactly $\binom{M}{n}(q-1)^n q^{-M}$, as, after the change of basis which makes this configuration the first basis vector, it will be an $n$-level excited state if and only if we have exactly $n$ non-zero entries in the first column of the grid:

$$\overline{S^{(n)}} = \binom{M}{n}(q-1)^n(q^{(N-M)}-q^{-M}) \tag{38}$$

$$\overline{H^{(n)}} = \binom{N}{n}(q-1)^n(q^{(M-N)}-q^{-N}). \tag{39}$$

In particular, the expression for $\overline{H^{(1)}}$ shows analytically the conjecture, supported by numerical simulations on traditional hyper-SAT [9], that is by varying the density of clauses, level 1 hypergraphs arise at the same threshold of hyperloops, but in a discontinuous way.

## 4. Derivation of $p_{N,M}(R)$ from integer moments of $S$

In this section, we will calculate the exact probability distribution $p_{N,M}(R)$ for the rank of $M \times N$ matrices on $GF(q)$, first by calculating the expectation value of integer moments of the number of solutions for the related homogeneous system $J\boldsymbol{x} = \boldsymbol{0}$, then, solving the 'moments problem' [32] derived from them.

The order parameter for the SAT/UNSAT phase transition is the probability that an instance has no other solution beyond the imposed one, that is, the probability that the rank $R$ of the matrix $J$ is equal to the number of variables $N$, so it is simply the value of $p_{N,M}(N)$.

### 4.1. Calculation of $\overline{S^k}$

Equation (34), in the infinite size limit, reads

$$\overline{\prod_{i<k}(S - q^i)} = q^{-k\Delta} \tag{40}$$

where, as we mentioned above, $\Delta = M - N$. Given the linear system of the first $k$ equations, we could try to solve it, and find an expression for each moment $\overline{S^k}$. The linear system is

$$\sum_h A_{kh}\overline{S^h} = q^{-k\Delta} \tag{41}$$

where the entries of the matrix $A_{kh}$ are defined by

$$\prod_{h=0}^{k-1}(S - q^h) = \sum_{h=0}^{k} A_{kh}S^h.$$

This product is analogous to the expression of theorem 2 in appendix A, so we obtain

$$\prod_{h=0}^{k-1}(S - q^h) = (-1)^k q^{\frac{k(k-1)}{2}} \prod_{h=0}^{k-1}(1 - Sq^{-h}) = \sum_{h=0}^{k}(-1)^{k-h}S^h q^{\frac{k(k-1)}{2} - \frac{h(h-1)}{2}}\begin{bmatrix} k \\ h \end{bmatrix}$$

and, from the definition, we recognize

$$A_{kh} = (-1)^{k-h}q^{\frac{k(k-1)}{2} - \frac{h(h-1)}{2}}\begin{bmatrix} k \\ h \end{bmatrix}. \tag{42}$$

Note that $A_{kh}$ is a triangular matrix, as obvious, as in the $i$th equation only moments up to $\overline{S^i}$ appear. This property is encoded in equation (42), as in $\mathcal{R}_{k-h,k}$ a zero factor appears when $h > k$.

If we could find the inverse matrix $B_{lk}$ such that

$$\sum_{k=0}^{n} B_{lk}A_{kh} = \delta_{lh} \qquad n \geqslant l, h \tag{43}$$

we would solve the system. We have an ansatz for this matrix:

*Ansatz 4.1*

$$B_{lk} = q^{k(l-k)}\begin{bmatrix} l \\ k \end{bmatrix}. \tag{44}$$

We should verify that

$$\sum_{k=h}^{l} q^{k(l-k)}\begin{bmatrix} l \\ k \end{bmatrix} \cdot (-1)^{k-h}q^{\frac{k(k-1)}{2} - \frac{h(h-1)}{2}}\begin{bmatrix} k \\ h \end{bmatrix} = \delta_{lh}. \tag{45}$$

When $l < h$ this is trivial, as we have no summands. When $l = h$ it is still easy, as we have only one summand, which is trivially 1. When $l > h$ we need a bit of algebra,

$$\sum_{k=h}^{l} q^{k(l-k)}\begin{bmatrix} l \\ k \end{bmatrix}(-1)^{k-h}q^{\frac{k(k-1)}{2} - \frac{h(h-1)}{2}}\begin{bmatrix} k \\ h \end{bmatrix} = \sum_{i=0}^{j} q^{hj}\begin{bmatrix} j+h \\ h \end{bmatrix}q^{-\frac{i(i-1)}{2}}(-q^{j-1})^i\begin{bmatrix} j \\ i \end{bmatrix}$$

where $i = k - h$ and $j = l - h$. We can recognize again the expression of theorem 2, so the last quantity is

$$\sum_{i=0}^{j} q^{-\frac{i(i-1)}{2}}(-q^{j-1})^i\begin{bmatrix} j \\ i \end{bmatrix} = \prod_{n=0}^{j-1}(1 - q^{j-1}q^{-n}) = 0. \tag{46}$$

Therefore, at the end we find the matrix which solves the linear system:

$$\overline{S^k} = B_{kh} q^{-h\Delta} = \sum_{h=0}^{k} q^{-h\Delta} q^{h(k-h)} \begin{bmatrix} k \\ h \end{bmatrix}. \tag{47}$$

### 4.2. Derivation of $p_\Delta(g)$

If we write the expression for $B_{kh}$ as a polynomial calculated in $q^k$, we have

$$\mathcal{B}_h(x) = \prod_{i=0}^{h-1} \frac{x - q^i}{q^h - q^i} \qquad B_{kh} = \mathcal{B}_h(q^k). \tag{48}$$

As, by definition, $\overline{S^k} = \overline{(q^g)^k} = \sum_g p(g)(q^k)^g$, we have

$$\sum_{g} p_\Delta(g)(q^k)^g = \sum_{h=0}^{k} q^{-h\Delta} \mathcal{B}_h(q^k). \tag{49}$$

Furthermore, as $\mathcal{B}_h(q^k) = 0$ for $h > k$, we can extend summation to infinity. In this form, we could make an ansatz of the analytical extension of the property above, and write

*Ansatz 4.2*

$$\sum_{g} p_\Delta(g)x^g = \sum_{h=0}^{\infty} q^{-h\Delta} \mathcal{B}_h(x). \tag{50}$$

If this property is formally true, and not only valid for the discrete set of points $x = q^k$, we could find $p_\Delta(g)$ just by comparing the series coefficients. Again from theorem 2 in appendix A, we have

$$\prod_{i=0}^{h-1}(x - q^i) = (-1)^h q^{\frac{h(h-1)}{2}} \prod_{i=0}^{h-1}(1 - xq^{-i}) = \sum_{g=0}^{h}(-1)^{h-g} x^g q^{\frac{h(h-1)}{2} - \frac{g(g-1)}{2}} \begin{bmatrix} h \\ g \end{bmatrix}$$

$$\prod_{i=0}^{h-1} \frac{1}{q^h - q^i} = q^{-h^2} \frac{1}{\mathcal{R}_{0,h}}.$$

So, calling $l = h - g$, we have

$$p_\Delta(g) = \frac{q^{g(g+\Delta)}}{\mathcal{R}_{0,g}} \sum_{l=0}^{\infty} q^{-\frac{l(l-1)}{2}} (-q^{-(g+\Delta+1)})^l \frac{1}{\mathcal{R}_{0,l}}. \tag{51}$$

With a further use of theorem 2 we finally have

$$p_\Delta(g) = q^{-g(g+\Delta)} \frac{\mathcal{R}_{0,\infty}}{\mathcal{R}_{0,g} \mathcal{R}_{0,g+\Delta}}. \tag{52}$$

Note that, as asymptotically the moments diverge faster than exponentially ($\overline{S^k} \sim C q^{(k-\Delta)^2/4}$), the heuristic ansatz 4.2 has no mathematical justification [32].

### 4.3. Finite-size corrections

In this procedure, it is not difficult to introduce finite-size corrections. In fact, if we go back to the exact form of equation (34), we could write

$$\overline{\prod_{i<k}(S - q^i)} = q^{-k\Delta} \mathcal{R}_{N-k,N}. \tag{53}$$

So the linear system has the same matrix $A_{kh}$ but a different inhomogeneous term:

$$A_{kh}\overline{S^h} = q^{-k\Delta}\mathcal{R}_{N-k,N}. \tag{54}$$

All the procedures of inversion of $A_{kh}$ are still valid, and we just need to replace equations (47) and (49) with the correct ones

$$\overline{S^k} = \sum_g p_\Delta(g)(q^k)^g = \sum_{h=0}^k q^{-h\Delta}\mathcal{R}_{N-h,N}\mathcal{B}_h(q^k). \tag{55}$$

Now, beyond a finite-size cut-off $k \gg 2N + \Delta$ (as $\mathcal{R}_{N-h,N} = 0$ for $h > N$), moments diverge as $\overline{S^k} \sim Cq^{N(k-N-\Delta)}$, and the finite-size version of ansatz 4.2 is actually a precise mathematical statement:

*Ansatz 4.3*

$$\sum_g p_\Delta(g)x^g = \sum_{h=0}^\infty q^{-h\Delta}\mathcal{R}_{N-h,N}\mathcal{B}_h(x). \tag{56}$$

This is a rare case in which a careful finite-size regularization demonstrates the validity of a heuristic analytic continuation.

Now just some algebra follows: we can correct equation (52) in

$$p_{N,M}(g) = \frac{q^{g(g+\Delta)}}{\mathcal{R}_{0,g}} \sum_{l=0}^{N-g} q^{-\frac{l(l-1)}{2}}(-q^{-(g+\Delta+1)})^l \frac{\mathcal{R}_{N-g-l,N}}{\mathcal{R}_{0,l}} = q^{-g(g+\Delta)}\frac{\mathcal{R}_{0,N+\Delta}\mathcal{R}_{N-g,N}}{\mathcal{R}_{0,g}\mathcal{R}_{0,g+\Delta}}. \tag{57}$$

Introducing the rank of the matrix $R = N - g$ to symmetrize the notations

$$p_{N,M}(\mathrm{rank}(J) = R) = q^{-(N-R)(M-R)}\mathcal{R}_{0,R}\begin{bmatrix} N \\ R \end{bmatrix}\begin{bmatrix} M \\ R \end{bmatrix}. \tag{58}$$

### 4.4. Finite-size scaling function of the SAT/UNSAT order parameter

Equation (58) provides us all the information about the behaviour of our model, at zero temperature with respect to the fictitious Hamiltonian (16). If we want to compare our results with the standard 'open questions' of random satisfiability problems, we should concentrate on the order parameter for the SAT/UNSAT transition, that is

$$\phi(\gamma, M) = p_{N,M}(\mathrm{rank}(J) = N)$$

where it is understood that $N = \gamma M$. The results of equation (58) lead to

$$\phi(\gamma, M) = \mathcal{R}_{0,N}\begin{bmatrix} M \\ N \end{bmatrix}.$$

The thermodynamic limit $M \to \infty$, at fixed density $\gamma$, allows us to determine the critical density $\gamma_c = 1$:

$$\phi(\gamma, \infty) = \lim_{M\to\infty} \frac{\mathcal{R}_{0,M}}{\mathcal{R}_{0,(1-\gamma)M}} = \begin{cases} 1 & \gamma < 1 \\ \mathcal{R}_{0,\infty} & \gamma = 1 \\ 0 & \gamma > 1. \end{cases} \tag{59}$$

We expect a non-trivial $M \to \infty$ limit when we keep a certain combination of density and size of the system fixed, $(\gamma - \gamma_c)M^{1/\nu}$. Again from equation (58) it is clear that the proper scaling window is obtained for $\nu = 1$. Introducing the parameter $\Delta = (\gamma_c - \gamma)M$, and performing the thermodynamic limit at fixed $\Delta$, we re-obtain the result implicit in equation (52):

$$\phi(\Delta) = \lim_{M\to\infty} \frac{\mathcal{R}_{0,M}}{\mathcal{R}_{0,\Delta}} = \frac{\mathcal{R}_{0,\infty}}{\mathcal{R}_{0,\Delta}}. \tag{60}$$
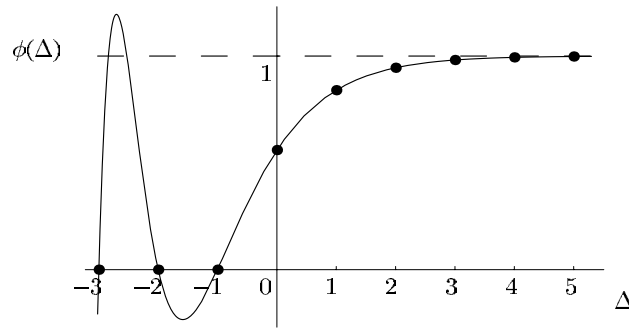
**Figure 1.** Order parameter $\phi(\Delta)$ as a function of $\Delta = M - N$, for $q = 3$. The 'physical points' are for integer values of $\Delta$, the continuous curve corresponds to the analytical continuation (61).

Although this result makes sense only for an integer number $\Delta$, analytical extension to an arbitrary complex number $\Delta$ is possible if we use the property

$$\mathcal{R}_{0,n} = \frac{\mathcal{R}_{0,\infty}}{\mathcal{R}_{n,\infty}}$$

to restate equation (60) as

$$\phi(\Delta) = \mathcal{R}_{\Delta,\infty} = \prod_{k=1}^{+\infty}(1 - q^{-\Delta-k}). \tag{61}$$

A plot of this function for a typical value of $q$, and $\Delta$ a real number, is shown in figure 1.

## 5. Derivation of $p_{N,M}(R)$ from the Gauss algorithm

As we have seen in section 3.2, we can perform a gauge transformation which, on the space of configurations, is essentially a change of basis. We can perform a gauge fixing to make clear the rank of an instance matrix $J$, and, from the analysis of this procedure, rederive the probability distribution for the rank, $p_{N,M}(R)$, which we have already found in a purely algebraic way in equation (58).

   We will define our gauge fix in an algorithmic way, which we call the *Gauss algorithm*, as it is essentially the standard algorithm to triangulate a matrix. We remark that we should only admit 'moves' on the columns, as they correspond to a change of basis in the space of configurations ($J \to JB^{-1}$, with $B \in GL(N)$), while we only admit simple permutations on the rows ($J \to SJ$, with $S \in q\text{-Perm}(M)$).

### 5.1. Step zero

The Gauss algorithm takes an $M \times N$ matrix $J$, and reduces it to a 'normal form'

$$J_{\text{norm}} = \begin{array}{|c|c|} \hline \begin{matrix} 1 & & & \\ & 1 & 0 & \\ & & \cdot & \\ & & & 1 \\ & J_{\text{red}} & & \end{matrix} & \begin{matrix} 0 & 0 & \cdot \\ 0 & & \\ 0 & & \\ \cdot & & \end{matrix} \\ \hline \end{array} \tag{62}$$

where the central triangular matrix has side $R = N - g$, corresponding to the rank of the matrix.

Suppose we know the probability distribution $p_{N,M}(R)$. Now consider the following two inductive steps.

### 5.2. Induction on columns (add one variable)

Take an $M \times (N + 1)$ matrix $J$, and apply the Gauss algorithm neglecting the last column $c$. The new matrix will be in a 'pre-normal form'

$$J_{\text{norm}}^{(1)} = \begin{array}{|c|c|c|} \hline \end{array} \qquad (63)$$

(note that the probability distribution for the last column is flat). Now reduce to zero the first $R$ terms of the column via the Gauss algorithm, obtaining

$$J_{\text{norm}}^{(2)} = \begin{array}{|c|c|c|} \hline \end{array} \qquad (64)$$

(note that the probability distribution for the remaining part of the column is still flat!). Now we have two ways to get further:

$c'$ *contains only zeros* (this happens with probability $q^{-M+R}$): the matrix is already in normal form, as we have simply added an empty row, so $R' = R$.

$c'$ *contains some non-zero element* (this happens with probability $1 - q^{-M+R}$): we will bring the column on position $R+1$, and eventually permute two of the last $M - R$ rows in order to have a 1 in the diagonal term; again a matrix in normal form. At the end, we have $R' = R + 1$.

The recursion law derived from these reasonings is resumed in the diagram below:
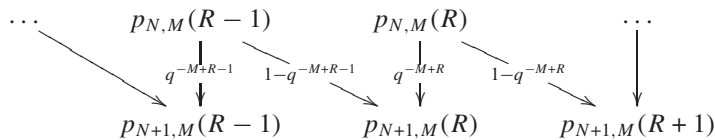


The analysis of the diagram leads to the equation

$$p_{N+1,M}(R) = (1 - q^{-M+R-1})p_{N,M}(R - 1) + q^{-M+R}p_{N,M}(R). \qquad (65)$$

### 5.3. Induction on rows (add one clause)

Take an $(M + 1) \times N$ matrix $J$, and apply the Gauss algorithm neglecting the last row $r$. The new grid will be in a 'pre-normal form'
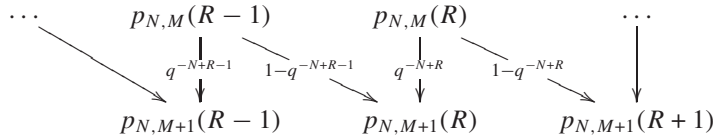
$$
J_{\text{norm}}^{(1)} =
\begin{array}{|c|c|}
\hline
\begin{matrix} 1 & & \\ & 1 & 0 \\ & & \cdot \\ & & 1 \\ J_{\text{red}} \end{matrix} &
\begin{matrix} 0 & 0 & \cdot \\ 0 & & \\ 0 & & \\ \cdot & & \end{matrix} \\
\hline
\multicolumn{2}{|c|}{r} \\
\hline
\end{array}
\tag{66}
$$

We call $r'$ the last $N - R$ terms of the row $r$, in analogy to the previous section. Note that the probability distribution for these terms is again flat. Now we have two ways to get further:

> $r'$ *contains only zeros* (this happens with probability $q^{-N+R}$): the matrix is already in normal form, with $R' = R$.
>
> $r'$ *contains some non-zero element* (this happens with probability $1 - q^{-N+R}$): we will bring the row on position $R + 1$, and apply the Gauss algorithm to the last $N - R$ rows, in order to have just a single 1 left in the first term (the diagonal one); again a matrix in normal form. At the end, we have $R' = R + 1$.

Again the recursion law derived from these reasonings is summed up in a diagram:

$$
\begin{array}{cccccc}
\cdots & p_{N,M}(R-1) & & p_{N,M}(R) & & \cdots \\
& \downarrow_{q^{-N+R-1}} \;\; {}_{1-q^{-N+R-1}} & & \downarrow_{q^{-N+R}} \;\; {}_{1-q^{-N+R}} & & \downarrow \\
& p_{N,M+1}(R-1) & & p_{N,M+1}(R) & & p_{N,M+1}(R+1)
\end{array}
$$

The analysis of this diagram leads to a second equation

$$
p_{N,M+1}(R) = (1 - q^{-N+R-1})p_{N,M}(R-1) + q^{-N+R}p_{N,M}(R).
\tag{67}
$$

### 5.4. Derivation of the probability distribution

The two equations (65) and (67) are linear and homogeneous. The quantities involved depend on three independent parameters: $N$, $M$ and $R$. We will show *a posteriori* that the quantity $p_{N,M}(R)$ has a good thermodynamic limit $M \to \infty$ when the limit is performed at $\Delta = M - N$ and $g = N - R$ fixed. We will call this limit function $p_\Delta(g)$.

This assumption allows us to write equations (65) and (67) as a system of two linear equations for $p_\Delta(g)$, in the two variables $g$ and $\Delta$:

$$
\begin{cases}
p_{\Delta-1}(g) = (1 - q^{-g-1})p_\Delta(g+1) + q^{-g}p_\Delta(g) \\
p_{\Delta+1}(g) = (1 - q^{-g-1+\Delta})p_\Delta(g+1) + q^{-g+\Delta}p_\Delta(g).
\end{cases}
\tag{68}
$$

An easy to verify solution is

$$
p_\Delta(g) = p_0(0)\frac{q^{-g(g+\Delta)}}{\mathcal{R}_{0,g}\mathcal{R}_{0,g+\Delta}}.
\tag{69}
$$

We can find the right normalization from the limit

$$
\lim_{\Delta \to +\infty} p_\Delta(0) = 1
\tag{70}
$$

which leads to

$$p_\Delta(g) = q^{-g(g+\Delta)} \frac{\mathcal{R}_{0,+\infty}}{\mathcal{R}_{0,g} \mathcal{R}_{0,g+\Delta}} \tag{71}$$

which is again equation (52).

### 5.5. Finite-size corrections

In this subsection, we will show how, with a careful use of the theory of partitions (see [33] and, for a summary of the main result, appendix A), only equation (65) (or in a similar way only equation (67)) leads to the exact expression for $p_{N,M}(R)$, calculated directly from the initial conditions, with no need for any assumption.

Equation (65) was

$$p_{N+1,M}(R) = (1 - q^{-M+R-1})p_{N,M}(R-1) + q^{-M+R}p_{N,M}(R).$$

It is more suitable to switch to the variable $S = N - R$ instead of $N$ for some steps:

$$p_M(S, R) = (1 - q^{-M+R-1})p_M(S, R-1) + q^{-M+R}p_M(S-1, R). \tag{72}$$

Imagine expanding this equation recursively until we reconduce $p_M(S, R)$ to the initial conditions

$$p_M(0, 0) = 1 \qquad p_M(S, R) = 0 \qquad \text{if} \quad S < 0, R < 0.$$

This would produce a prolification of terms. Each of these terms can be labelled by a trajectory on an $S \times R$ grid, starting from the point $(0, 0)$ at the top-left corner and stopping at the point $(S, R)$ at the bottom-right corner. This trajectory is directed: we only admit steps to the right nearest neighbour $(s, r) \rightarrow (s, r + 1)$ or to the down nearest neighbour $(s, r) \rightarrow (s + 1, r)$. The weights of these steps, determined by equation (72), are

$$(s, r) \xrightarrow{1-q^{-M+r}} (s, r+1)$$
$$\downarrow q^{-M+r}$$
$$(s+1, r)$$
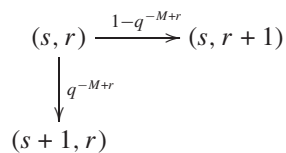
An interesting fact about these weights is that they only depend on $r$. For this reason, the contribution derived from the horizontal steps is the same in all the trajectories:

$$\prod_{r=0}^{R-1}(1 - q^{-M+r}) = \mathcal{R}_{M-R,M} = \mathcal{R}_{0,R}\begin{bmatrix} M \\ R \end{bmatrix}. \tag{73}$$

The contribution derived from the vertical steps is more complicated, and needs some results from the theory of partitions. Consider the trajectory as a string of $S$ integer numbers $(r_1, \ldots, r_S)$, corresponding to the length of the rows of the diagram individuated by the trajectory. As the trajectory is directed, we should have $0 \leqslant r_i \leqslant r_{i+1} \leqslant R$:
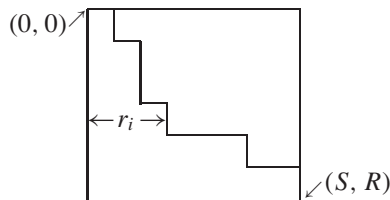
The vertical step contribution for a certain trajectory will be $\prod_{i=1}^{S} q^{-M+r_i}$. As a result of theorem 1 in appendix A we could write the generating function in the number of rows,

$$f(z) = \sum_{S} z^S \sum_{(r_1,\ldots,r_S)} \prod_{i=1}^{S} q^{-M+r_i} = \prod_{r=0}^{R} \frac{1}{1 - zq^{-M+r}} \tag{74}$$

and from theorem 3 we have

$$\prod_{r=0}^{R} \frac{1}{1 - zq^{-M+r}} = \sum_{k=0}^{\infty} z^k q^{-k(M-R)} \begin{bmatrix} R + k \\ R \end{bmatrix}$$

and, as we are interested in the term $z^S$, for $S = N - R$, we finally have

$$q^{-(N-R)(M-R)} \begin{bmatrix} N \\ R \end{bmatrix}$$

which, combined with the contribution derived from the horizontal steps, formula (73), gives again the result of equation (58).

## 6. Conclusions

We introduced a variant of the hyper-SAT model, by relaxing the constraint on the length of the clauses. This makes the model exactly solvable, so that we could not only compute the probability for an instance of being satisfiable, but also the probability distribution for the number of configurations which satisfy all the constraints. This probability distribution can also be evaluated for a finite number of variables $N$, and of clauses $M$. The SAT/UNSAT transition occurs in the thermodynamic limit, for which $N$ and $M$ go to infinity with their ratio fixed to 1. This model enjoys a dual formulation in which the role of clauses and variables is interchanged, in agreement with the previous remark that the phase transition occurs when the number of clauses is equal to the number of instances. The whole analysis is soon extended to the general $GF(q)$ case. Finite-size scaling functions are easily derived from the general solutions.

The reasons for our work were mainly three:

(i)  Random matrices on $GF(q)$, with $N > M$, are connected to random linear codes of the error correcting codes theory, in particular to the random codeword model introduced by Montanari in [20]. The unexpected existence of an exact solution of the zero-temperature thermodynamic, and some other exact results also for finite temperature, open new perspectives in the study of this model.

(ii)  The random satisfiability of sparse linear systems on $GF(q)$ is a polynomial problem of computation complexity $N^3$. A careful average-case complexity analysis shows both an SAT/UNSAT transition (discontinuity in the order parameter (probability for an instance to have at least one solution) as a function of the number of variables to the number of clauses ratio) and a 'dynamical' transition in the complexity class, from linear to cubic complexity [34]. These characteristics make the hyper-SAT model an ideal toy model to investigate the underlying mechanisms of average-case complexity in the $K$-SAT problem. Our model is a further idealization of hyper-SAT, where the dynamical transition is lost, which has the advantage of an exact analytical control and also of the replica approach.

(iii)  The random satisfiability of sparse linear systems on $GF(2)$ can be restated as a combinatorial problem in random hypergraph theory. Our results are thus also relevant in this area.

The analysis of our model at finite temperature will be presented elsewhere [31].

## Acknowledgments

## Appendix. Results from theory of partitions

In this appendix, we summarize some results from analytic number theory of partitions, mainly due to the ancient work of Euler and Gauss on the generating function formalism. Our main reference for these results is chapter 12 of Rademaker's textbook [33].

It is worth noting that the mathematical results of this appendix are the groundwork for important results of number theory known as *Rogers–Schur–Ramanujan identities*, which are shown to be connected to the Virasoro characters of conformal minimal models (see e.g. the papers by McCoy *et al* [36, 37]). They are also known as *q-calculus*, and are connected with one-dimensional asymmetric exclusion processes (ASEP) (see e.g. the work of Sasamoto [35]).

Given an integer $n$, a *partition* $[p]_n$ is a list of ordered integers $(n_1, n_2, \ldots, n_k)$ such that their sum is $n$:

$$[p]_n = (n_1, n_2, \ldots, n_k) \qquad 1 \leqslant n_i \leqslant n_{i+1} \qquad \sum n_i = n.$$

A common tool consists in introducing formal power series in which the $n$th series coefficient is the sum, on all the partitions $[p]_n$ of the integer $n$, of a certain weight $W([p]_n)$. In the more general framework, the weight function could be valued in $\mathbb{C}$. In the case of 'well-behaved' positive real weights, we can think of the power series as a grand-canonical partition function on the set of partitions.

An analytic expression for these power series could be found for a wide class of weight functions. We will restrict ourselves to weights which factorize on the elements of the list, that is, given a certain function $w : \mathbb{N} \to \mathbb{C}$, the weight of a partition is

$$W([p]_n) = \prod_{i=1}^{k} w(n_i).$$

We can state the first easy result as follows:

**Theorem 1.** *The generating function for the partitions, with a factorized weight $W([p]_n) = \prod_{i=1}^{k} w(n_i)$, is*

$$\sum_n z^n \sum_{[p]_n} W([p]_n) = \prod_{i=1}^{\infty} \frac{1}{1 - zw(i)}. \tag{A.1}$$

In the most common case, in which we just want to count all the partitions of a certain integer, the proper weight would be $w(n) = q^{-n}$, so we will be concerned with products of the kind $\prod_i (1 - zq^{-i})$. It will be useful, as in equations (32), to introduce the notation

$$\mathcal{R}_{m,n} = \prod_{j=m+1}^{n} (1 - q^{-j}) \qquad \begin{bmatrix} n \\ m \end{bmatrix} = \frac{\mathcal{R}_{n-m,n}}{\mathcal{R}_{0,m}}. \tag{A.2}$$

A more common notation, but more tedious to handle in our context, is the following,

$$(a; x)_n = \prod_{i=0}^{n-1}(1 - ax^i) \qquad \begin{bmatrix} n \\ m \end{bmatrix}_x = \frac{(x; x)_n}{(x; x)_m(x; x)_{n-m}} \tag{A.3}$$

where the quantity on the right-hand side is known as the $(n, m)$ Gaussian polynomial in the variable $x$. In this paper, we have always used $q^{-1}$ as the argument of $\mathcal{R}_{m,n}$ products. A few easy relations between the Gaussian polynomials are

$$\begin{bmatrix} n \\ m \end{bmatrix} = \begin{bmatrix} n \\ n-m \end{bmatrix} \tag{A.4a}$$

$$\begin{bmatrix} n \\ m \end{bmatrix} = \begin{bmatrix} n-1 \\ m-1 \end{bmatrix} + q^{-m}\begin{bmatrix} n-1 \\ m \end{bmatrix} = q^{-(n-m)}\begin{bmatrix} n-1 \\ m-1 \end{bmatrix} + \begin{bmatrix} n-1 \\ m \end{bmatrix} \tag{A.4b}$$

$$\begin{bmatrix} n \\ m \end{bmatrix}\begin{bmatrix} m \\ l \end{bmatrix} = \begin{bmatrix} n \\ l \end{bmatrix}\begin{bmatrix} n-l \\ m-l \end{bmatrix}. \tag{A.4c}$$

Furthermore, at fixed $m$, in the limit $n \to \infty$ we have

$$\lim_{n\to\infty}\begin{bmatrix} n \\ m \end{bmatrix} = \frac{1}{\mathcal{R}_{0,m}}.$$

Now consider the formal power series

$$P_n(x, z) = \prod_{i=1}^{n}(1 - zx^i) = \sum_{m=0}^{n} a_{m,n}(x)z^m.$$

From the relation

$$P_n(x, xz) = \frac{1 - zx^{n+1}}{1 - zx}P_n(x, z)$$

we find that

$$a_{m,n}(x) = -x^m\frac{1 - x^{n-m+1}}{1 - x^m}a_{m-1,n}(x)$$

and from the initial condition $a_{0,n}(x) = 1$ we find

$$a_{m,n}(x) = \prod_{m'=1}^{m} -x^{m'}\frac{1 - x^{n-m'+1}}{1 - x^{m'}}$$

from which we recognize

**Theorem 2.**

$$\prod_{i=1}^{n}(1 - zq^{-i}) = \sum_{m=0}^{n}(-z)^m q^{-\frac{m(m-1)}{2}}\begin{bmatrix} n \\ m \end{bmatrix} \tag{A.5a}$$

$$\prod_{i=1}^{\infty}(1 - zq^{-i}) = \sum_{m=0}^{\infty}(-z)^m q^{-\frac{m(m-1)}{2}}\frac{1}{\mathcal{R}_{0,m}}. \tag{A.5b}$$

Analogously, if we consider the formal power series

$$Q_n(x, z) = \prod_{i=0}^{n}\frac{1}{1 - zx^i} = \sum_{m=0}^{\infty} b_{m,n}(x)z^m$$

from the relation

$$Q_n(x, xz) = \frac{1 - z}{1 - zx^{n+1}} Q_n(x, z)$$

we find that

$$b_{m,n}(x) = \frac{1 - x^{m+n}}{1 - x^m} b_{m-1,n}(x)$$

and from the initial condition $b_{0,n}(x) = 1$ we find

$$b_{m,n}(x) = \prod_{m'=1}^{m} \frac{1 - x^{m'+n}}{1 - x^{m'}} = \begin{bmatrix} m + n \\ m \end{bmatrix}$$

from which we recognize

**Theorem 3.**

$$\prod_{i=0}^{n} \frac{1}{1 - zq^{-i}} = \sum_{m=0}^{\infty} z^m \begin{bmatrix} m + n \\ m \end{bmatrix} \tag{A.6a}$$

$$\prod_{i=0}^{\infty} \frac{1}{1 - zq^{-i}} = \sum_{m=0}^{\infty} z^m \frac{1}{\mathcal{R}_{0,m}}. \tag{A.6b}$$

## References

[1] Garey M and Johnson D S 1979 *Computers and Intractability; A Guide to the Theory of NP-Completeness* (San Francisco, CA: Freeman)
  Papadimitriou C 1994 *Computational Complexity* (Reading, MA: Addison-Wesley)
[2] Goldreich O 1999 *Introduction to Complexity Theory* (*Lecture Notes Series of ECCC*); webpage http://www.eccc.uni-trier.de/eccc-local/ECCC-LectureNotes/IntroComplTh
[3] Wilf H S 1986 *Algorithms and Complexity* (Englewood Cliffs, NJ: Prentice-Hall) exp ch 5; webpage http://www.cis.upenn.edu/~wilf/AlgComp2.html
[4] Martin O C, Monasson R and Zecchina R 2001 Statistical mechanics methods and phase transitions in optimization problems *Theor. Computer Sci.* **265** 3–67 (*Preprint* cond-mat/0104428)
[5] Monasson R, Zecchina R, Kirkpatrick S, Selman B and Troyansky L 1999 Typical-case complexity *Nature* **400** 133
[6] Bollobás B, Borgs C, Chayes J T, Kim J H and Wilson D B 1999 *Preprint* arXiv:math. CO/9909031
[7] Achlioptas D 2001 Lower bounds for random 3-SAT via differential equations *Theor. Comput. Sci.* **265** 159–85
[8] Achlioptas D and Sorkin G B 2000 Optimal myopic algorithms for random 3-SAT *FOCS 2000* pp 590–600
[9] Ricci-Tersenghi F, Weigt M and Zecchina R 2001 Simplest random K-satisfiability problem *Phys. Rev.* E **63** 26702 (*Preprint* cond-mat/0011181)
[10] Leone M, Ricci-Tersenghi F and Zecchina R 2001 Phase coexistence and finite-size scaling in random combinatorial problems *J. Phys. A: Math. Gen.* **34** 4615 (*Preprint* cond-mat/0103200)
[11] Mézard M and Parisi G 2002 The cavity method at zero temperature *Preprint* cond-mat/0207121
[12] Mézard M and Zecchina R 2002 The random K-satisfiability problem: from an analytic solution to an efficient algorithm *Preprint* cond-mat/0207194
[13] Landsberg G 1893 Über eine Anzahlbestimmung un eine damit Zusammenhängende Reihe *J. Reine Angewandte Math.* **111** 87
[14] Goldman J and Rota G C 1970 On the foundations of combinatorial theory IV: Finite vector spaces and Eulerian generating functions *Stud. Appl. Math.* **49** 239
[15] Kolchin V F 1999 Random graphs *Encyclopedia of Mathematics and its Applications* vol 53 (Cambridge: Cambridge University Press) ch 3
[16] Kovalenko I N 1967 On the limit distribution of the number of solutions of a random system of linear equations in the class of Boolean functions *Theory Probab. Appl.* **12** 51–61 (in Russian)

[17] Kovalenko I N, Levitskaya A A and Savchuk M N 1986 *Selected Problems of Probabilistic Combinatorics* (Kiev: Naukova Dumka) (in Russian)
[18] Roman S 1992 *Coding and Information Theory* (New York: Springer)
[19] Sourlas N 1999 Statistical mechanics and error-correction codes *Proc. Marseille Satellite Colloquium 'Mathematical Results in Stat Mechanics'* (*Preprint* cond-mat/9811406)
[20] Montanari A 2001 The glassy phase of Gallager codes *Eur. Phys. J.* B **23** 121 (*Preprint* cond-mat/ 0104079)
[21] Varshamov R R 1957 Estimate of the number of signals in error correcting codes *Dokl. Akad. Nauk* **117** 739
[22] Sudan M 2001 *Algorithmic Introduction to Coding Theory (Notes of Lectures held at MIT in fall 2001)* exp lecture 5; webpage http://theory.lcs.mit.edu/~madhu/FT01/course.html
[23] Bollobás B 1985 *Random Graphs* (London: Academic)
[24] Berge C 1973 *Graphs and Hypergraphs* (Amsterdam: North-Holland)
[25] Mézard M, Parisi G and Virasoro M A 1987 *Spin Glass Theory and Beyond* (Singapore: World Scientific)
[26] Ruelle D 1987 A mathematical reformulation of Derrida's REM and GREM *Commun. Math. Phys.* **108** 225–39
[27] Derrida B 1981 Random energy model: an exactly solvable model of disordered systems *Phys. Rev.* B **24** 2613
[28] Gross D and Mézard M 1984 The simplest spin glass *Nucl. Phys.* B **240** 431
[29] Gardner E 1985 *Nucl. Phys.* B **257** 747
[30] Caracciolo S and Sportiello A 2002 General duality for statistical mechanics model *Preprint* IFUM-751-FT
[31] Caracciolo S and Sportiello A 2002 Finite-temperature statistical mechanics of a random satisfiability problem *Preprint* IFUM-750-FT
[32] Akhiezer N I 1965 *The Classical Moment Problem* ed D E Rutherford (Edinburgh: Oliver and Boyd) ch 1 and 2
[33] Rademacher H 1973 *Topics in Analytic Number Theory* (Berlin: Springer)
[34] Braunstein A, Leone M, Ricci-Tersenghi F and Zecchina R 2002 Complexity transitions in global algorithms for sparse linear systems over finite fields *Preprint* cond-mat/0203613
[35] Sasamoto T 1999 One-dimensional partially asymmetric simple exclusion process with open boundaries: orthogonal polynomials approach *J. Phys. A: Math. Gen.* **32** 7109–31
[36] Berkovich A, McCoy B M and Schilling A 1998 Rogers–Schur–Ramanujan type identities for the $M(p, p')$ minimal models of conformal field theory *Commun. Math. Phys.* **191** 325 (*Preprint* q-alg/9607020)
[37] Berkovich A and McCoy B M 1998 Rogers–Ramanujan identities: a century of progress from mathematics to physics *Proc. ICM98 Documenta Mathematica* vol 3 p 163 extra volume ICM 1998