A Continuous-Discontinuous Second-Order Transition in the Satisfiability of Random Horn-SAT Formulas

Cristopher Moore

Department of Computer Science, University of New Mexico, Albuquerque NM, USA
Santa Fe Institute, Santa Fe NM, USA
`email: moore@cs.unm.edu`

Gabriel Istrate

CCS-5, Los Alamos National Laboratory, Los Alamos NM, USA
`email: istrate@lanl.gov`

Demetrios Demopoulos

Archimedean Academy, 12425 SW Sunset Drive, Miami FL, USA
`email: demetrios_demopoulos@archimedean.org`

Moshe Y. Vardi

Department of Computer Science, Rice University, Houston TX, USA
`email: vardi@rice.edu`

ABSTRACT

We compute the probability of satisfiability of a class of random Horn-SAT formulae, motivated by a connection with the nonemptiness problem of finite tree automata. In particular, when the maximum clause length is three, this model displays a curve in its parameter space along which the probability of satisfiability is discontinuous, ending in a second-order phase transition where it is continuous but its derivative diverges. This is the first case in which a phase transition of this type has been rigorously established for a random constraint satisfaction problem.

1. INTRODUCTION

In the past decade, sharp thresholds, or *phase transitions*, have been studied intensively in combinatorial problems. Although the idea of thresholds in a combinatorial context was introduced as early as 1960 [16], in recent years it has been a major subject of research in the communities of theoretical computer science, artificial intelligence and statistical physics. Phase transitions in which the probability of satisfiability jumps from 1 to 0 when the density of constraints exceeds a critical threshold have been observed in numerous constraint satisfaction problems.

The problem at the center of this research is, of course, 3-SAT. An instance of 3-SAT is a Boolean formula, consisting of a conjunction of clauses, where each clause is a disjunction of three literals. The goal is to find a truth assignment that satisfies all the clauses and thus the entire formula. The *density* of a 3-SAT instance is the ratio of the number of clauses to the number of variables. We call the number of variables the *size* of the instance. Experimental studies [10, 30, 31] show a dramatic shift in the probability of satisfiability of random 3-SAT instances, from 1 to 0, located at a critical density $r_c \approx 4.26$. However, in spite of compelling arguments from statistical physics [26, 27] and rigorous upper and lower bounds on the threshold if it exists [12, 19, 24], there is still no mathematical proof that a phase transition takes place at that density. For a few variants of SAT the existence and location of phase transitions have been established rigorously, in particular for 2-SAT [9, 18], 3-XORSAT [13, 8], and 1-in-k SAT [3].

In this paper we prove the existence of a more elaborate type of phase transition, where a curve of discontinuities in a two-dimensional parameter space ends in a *second-order* transition, where the probability of satisfiability is continuous but nonanalytic. We focus on a particular variant of 3-SAT, namely Horn-SAT. A Horn clause is a disjunction of literals of which *at most one* is positive, and a Horn-SAT formula is a conjunction of Horn clauses. Unlike 3-SAT, Horn-SAT is a tractable problem; the complexity of Horn-SAT is linear in the size of the formula [11]. This tractability allows one to study random Horn-SAT formulae for much larger input sizes that we can achieve using complete algorithms for 3-SAT.

An additional motivation for studying random Horn-SAT comes from the fact that Horn formulae are connected to several other areas of computer science and mathematics [25]. In particular, Horn formulae are connected to automata theory, as the transition relation, the starting state, and the set of final states of an automaton can be described using Horn clauses. For example, if we consider an automaton on binary trees, then Horn clauses of length three can be used to describe its transition relation, while Horn clauses of length one can describe the starting state and the set of the final states of the automaton. (We elaborate on this below). Then the question of the emptiness of the language of the automaton can be translated to a question about the satisfiability of the formula. Since automata-theoretic techniques have recently been applied in automated reasoning [32, 33], the behavior of random Horn formulae might shed light on these applications.

Threshold properties of random Horn-SAT problems have been studied under a number of probabilistic models. The probability of satisfiability of random Horn formulae under two related *variable*-clause-length models was fully characterized in [21, 22]; in those models random Horn formulae have a *coarse* threshold, meaning that the probability of satisfiability is a continuous function of the parameters of the model. The variable-clause-length model used there is ideally suited to studying Horn formulae in connection with knowledge-based systems [25]. Bench-Capon and Dunne [5] studied a *fixed*-clause-length model, in which each Horn clause has precisely k literals, and proved a sharp threshold with respect to assignments that have at least $k-1$ true variables.

Motivated by the connection between the automata emptiness problem and Horn satisfiability, Demopoulos and Vardi [15] studied the satisfiability of two types of fixed-clause-length random Horn-SAT formulae. They considered 1-2-Horn-SAT, where formulae consist of clauses of length one and two only, and 1-3-Horn-SAT, where formulae consist of clauses of length one and three only. These two classes can be viewed as the Horn analogue of 2-SAT and 3-SAT. For 1-2-Horn-SAT, they showed experimentally that there is a coarse transition (see Figure 4), which can be explained and analyzed in terms of random digraph connectivity [23]. The situation for 1-3-Horn-SAT is less clear cut. On one hand, recent results on random undirected hypergraphs [14] fit the experimental data of [15] quite well. On the other, a scaling analysis of the data suggested that transition between the mostly-satisfiable and mostly-unsatisfiable regions (the "waterfall" in Figure 1) is steep but continuous, rather than a step function. It was therefore not clear if the model exhibits a phase transition, in spite of experimental data for instances with tens of thousands of variables.

In this paper we generalize the fixed-clause-length model of [15] and offer a complete analysis of the probability of satisfiability in this model. For a finite $k > 0$ and a vector $\mathbf{d}$ of k nonnegative real numbers $d_1, d_2, \ldots, d_k$ with $d_1 < 1$, let the random Horn-SAT formula $H^k_{n,\mathbf{d}}$ be the conjunction of

- a single negative literal $\overline{x}_1$,
- $d_1 n$ positive literals chosen uniformly without replacement from $x_2, \ldots, x_n$, and
- for each $2 \le j \le k$, $d_j n$ clauses chosen uniformly with replacement from the $j\binom{n}{j}$ possible Horn clauses with j variables where one literal is positive.

Thus, the classes studied in [15] are H^2_{n,d_1,d_2} and $H^3_{n,d_1,0,d_3}$, or 1-2-Horn-SAT and 1-3-Horn-SAT respectively. For instance, a typical 1-3-Horn-SAT formula might be

$$\overline{x}_1 \wedge \underbrace{x_3 \wedge x_5 \wedge x_{17}}_{d_1 n \text{ literals}} \wedge \underbrace{(\overline{x}_5 \vee \overline{x}_{17} \vee x_{29}) \wedge \cdots}_{d_3 n \text{ clauses}}$$

(note that this formula would immediately imply x_{29}).

With this model in hand, we settle the question of sharp thresholds for 1-3-Horn-SAT. In particular, we show that there are sharp thresholds in some regions of the (d_1, d_3) plane in the probability of satisfiability, although not from 1 to 0. We start with the following general result for the $H^k_{n,\mathbf{d}}$ model.

Theorem 1.1. *Let t_0 be the smallest positive root of the equation*

$$\ln \frac{1-t}{1-d_1} + \sum_{j=2}^{k} d_j t^{j-1} = 0 \ . \tag{1.1}$$

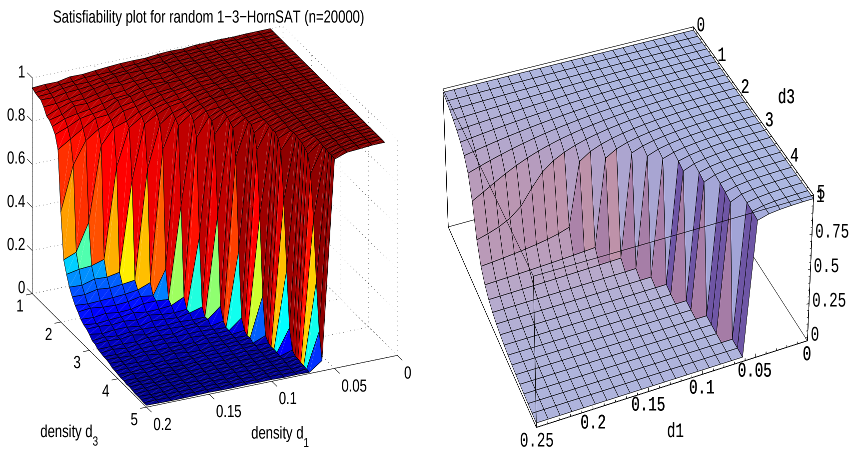

Fig. 1. Satisfiability probability of random 1-3-Horn formulae of size 20000. Left, the experimental results of [15]; right, our analytic results.

Call t_0 simple *if it is not a double root of* (1.1)*, i.e., if the derivative of the left-hand-side of* (1.1) *with respect to* t *is nonzero at* t_0*. If* t_0 *is simple, the probability that a random formula from* $H^k_{n,\mathbf{d}}$ *is satisfiable in the limit* $n \to \infty$ *is*

$$\Phi(\mathbf{d}) := \lim_{n\to\infty} \Pr[H^k_{n,\mathbf{d}} \textit{ is satisfiable}] = \frac{1-t_0}{1-d_1} \quad . \tag{1.2}$$

Specializing this result to the case $k = 2$ yields an exact formula that matches the experimental results in [15]:

Proposition 1.2. *The probability that a random formula from* H^2_{n,d_1,d_2} *is satisfiable in the limit* $n \to \infty$ *is*

$$\Phi(d_1, d_2) := \lim_{n\to\infty} \Pr[H^2_{n,d_1,d_2} \textit{ is satisfiable}] = -\frac{W\big(-(1-d_1)d_2 e^{-d_2}\big)}{(1-d_1)d_2} \quad . \tag{1.3}$$

Here $W(\cdot)$ is Lambert's function, defined as the principal root y of $ye^y = x$.

For the case $k = 3$ and $d_2 = 0$, we do not have a closed-form expression for the probability of satisfiability, though numerically Figure 1 shows a very good fit to the experimental results of [15]. In addition, we find an interesting phase transition behavior in the (d_1, d_3) plane, described by the following proposition.

Proposition 1.3. *The probability of satisfiability* $\Phi(d_1, d_3)$ *that a random formula from* $H^3_{n,d_1,0,d_3}$ *is satisfiable is continuous for* $d_3 < 2$ *and discontinuous for* $d_3 > 2$*. Its* *discontinuities are given by a curve* Γ *in the* (d_1, d_3) *plane described by the equation*

$$d_1 = 1 - \frac{\exp\left(\frac{1}{4}\left(\sqrt{d_3} - \sqrt{d_3 - 2}\right)^2\right)}{d_3 - \sqrt{d_3(d_3 - 2)}} \quad . \tag{1.4}$$

This curve consists of the points (d_1, d_3) *at which* t_0 *is a double root of* (1.1)*, and ends at the critical point*

$$(1 - \sqrt{\mathrm{e}}/2, 2) = (0.1756..., 2) \tag{1.5}$$

at which Φ *is continuous but* $\partial\Phi/\partial d_3$ *diverges.*

The curve Γ of discontinuities described in Proposition 1.3 can be seen in the right part of Figure 1. The drop at the "waterfall" decreases as we approach the critical point (1.5), and at the critical point the probability of satisfiability is continuous but nonanalytic. We can also see this in Figure 2, which shows a contour plot; the contours are bunched at the discontinuity, and "fan out" at the critical point. In both cases our calculations closely match the experimental results of [15].

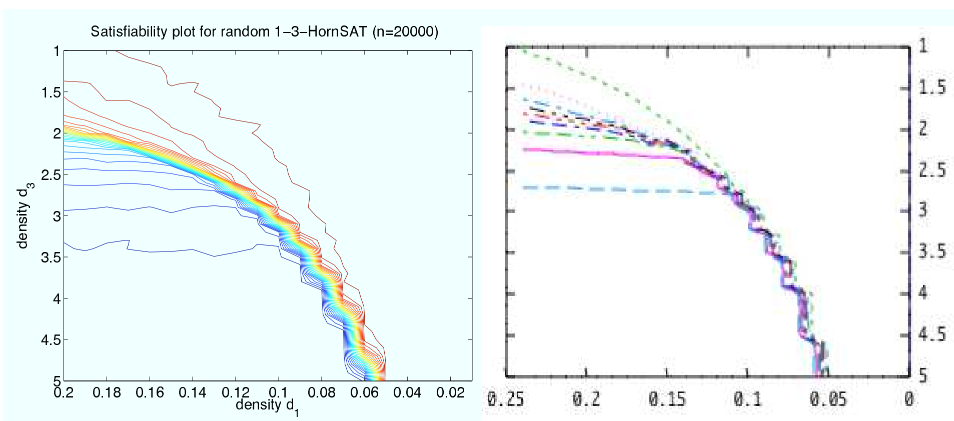

Fig. 2. Contour plots of the probability of satisfiability of random 1-3-Horn formulae. Left, the experimental results of [15]. Right, our analytical results.

In statistical physics, phase transitions are classified as *pth-order* if the $(p-1)$st derivative of the order parameter Φ (in this case, the probability of satisfiability) is the first one which is discontinuous. Thus we would say that Γ is a curve of *first-order* transitions, at which Φ itself is discontinuous. At the tip of this curve, i.e., at the critical point (1.5), Φ is continuous but its derivative is not, giving a *second-order* transition. Finally, beyond the tip Φ is analytic. A similar situation exists in the Ising model of magnetism, where the two parameters are the temperature T and the external field H. In the (T, H) plane there is a line of discontinuities at $H = 0$, ending at the transition temperature T_c. At $(T_c, 0)$, the magnetization is continuous but its derivative is not, giving a second-order transition, and for $T > T_c$ the magnetization is analytic (see e.g. [6]).

To our knowledge, this is the first time that a continuous-discontinuous transition of this type has been established rigorously in a model of random constraint satisfaction problems. We note that a similar phenomenon is believed to take place for $(2+p)$-SAT model at the triple point $p = 2/5$; here the order parameter is the size of the backbone, i.e., the number of variables that take fixed values in every truth assignment [4, 28]. Indeed, in our model the probability of satisfiability is closely related to the size of the backbone, as we see below.

2. HORN-SAT AND AUTOMATA

Our main motivation for studying the satisfiability of Horn formulae is the unusually rich type of phase transition described above, and the fact that its tractability allows us to perform experiments on formulae of very large size. The original motivation of [15] that led to the present work is the fact that Horn formulae can be used to describe finite automata on words and trees (see also [29]).

A finite automaton A is a 5-tuple $A = (S, \Sigma, \delta, s, F)$, where S is a finite set of states, Σ is an alphabet, s is a starting state, $F \subseteq S$ is the set of final (accepting) states and δ is a transition relation. In a word automaton, δ is a function from $S \times \Sigma$ to 2^S, while in a binary-tree automaton δ is a function from $S \times \Sigma$ to $2^{S \times S}$. Intuitively, for word automata δ provides a set of successor states, while for binary-tree automata δ provides a set of successor state pairs. A run of an automaton on a word $a = a_1 a_2 \cdots a_n$ is a sequence of states $s_0 s_1 \cdots s_n$ such that $s_0 = s$ and $(s_{i-1}, a_i, s_i) \in \delta$. A run is succesful if $s_n \in F$: in this case we say that A accepts the word a. A run of an automaton on a binary tree t labeled with letters from Σ is a binary tree r labeled with states from S such that $\text{root}(r) = s$ and for a node i of t, $(r(i), t(i), r(\text{left-child-of-}i), r(\text{right-child-of-}i)) \in \delta$. Thus, each pair in $\delta(r(i), t(i))$ is a possible labeling of the children of i. A run is succesful if for all leaves l of r, $r(l) \in F$: in this case we say that A accepts the tree t. The language $L(A)$ of a word automaton A is the set of all words a for which there is a successful run of A on a. Likewise, the language $L(A)$ of a tree automaton A is the set of all trees t for which there is a successful run of A on t. An important question in automata theory, which is also of great practical importance in the field of formal verification [32, 33], is: given an automaton A, is $L(A)$ non-empty? We now show how the problem of non-emptiness of automata languages translates to Horn satisfiability. Thus, getting a better understanding of the satisfiability of Horn formulae would tell us about the typical answer to random automaton nonemptiness problems.

Consider first a word automaton $A = (S, \Sigma, \delta, s_0, F)$. Construct a Horn formula ϕ_A over the set S of variables as follows: create a clause $(\overline{s}_0)$, for each $s_i \in F$ create a clause (s_i), for each element (s_i, a, s_j) of δ create a clause $(\overline{s}_j, s_i)$, where $(s_i, \cdots, s_k)$ represents the clause $s_i \vee \cdots \vee s_k$ and $\overline{s}_j$ is the negation of s_j. Note that the formula ϕ_A consists of unary and binary clauses. Similarly to the word automata case, we can show how to construct a Horn formula from a binary-tree automaton. Let $A = (S, \Sigma, \delta, s_0, F)$ be a binary-tree automaton. Then we can construct a Horn formula ϕ_A using the construction above with the only difference that since δ in this case is a function from $S \times \{\alpha\}$ to $S \times S$, for each element (s_i, α, s_j, s_k) of δ we create a clause $(\overline{s}_j, \overline{s}_k, s_i)$. Note that the

```
Algorithm PUR:
while (φ contains positive unit clauses)
    choose a random positive unit clause x
    remove all other clauses in which x occurs positively in φ
    shorten all clauses in which x appears negatively
    label x as "implied"
if no contradictory clause was created
    accept φ
else
    reject φ.
```

Fig. 3. Positive Unit Resolution.

formula ϕ_A now consists of unary and ternary clauses. This explains why the formula classes studied in [15] are H^2_{n,d_1,d_2} and $H^3_{n,d_1,0,d_3}$.

Proposition 2.1 [15]. *Let A be a word or binary tree automaton and ϕ_A the Horn formula constructed as described above. Then $L(A)$ is non-empty if and only if ϕ_A is unsatisfiable.*

3. MAIN RESULT

The proof of Theorem 1.1 follows immediately from the following theorem, which establishes the size of the backbone of the formula with the single negative literal $\overline{x}_1$ removed: that is, the set of positive literals implied by the positive unit clauses and the clauses of length greater than 1. Then ϕ is satisfiable as long as x_1 is not in this backbone.

Lemma 3.1. *Let ϕ be a random Horn-SAT formula $H^k_{n,\mathbf{d}}$ with $d_1 > 0$. Denote by t_0 the smallest positive root of Equation* (1.1)*, and suppose that t_0 is simple. Then, for any $\epsilon > 0$, the number $N_{\mathbf{d},n}$ of implied positive literals, including the $d_1 n$ initially positive literals, satisfies w.h.p. the inequality*

$$(t_0 - \epsilon) \cdot n < N_{\mathbf{d},n} < (t_0 + \epsilon) \cdot n, \tag{3.1}$$

Proof. First, we give a heuristic argument, analogous to the branching process argument for the size of the giant component in a random graph. The number m of clauses of length j with a given literal x as their implicate (i.e., in which x appears positively) is Poisson-distributed with mean d_j. If any of these clauses have the property that all their literals whose negations appear are implied, then x is implied as well. In addition, x is implied if it is one of the $d_1 n$ initially positive literals. Therefore, the probability that x is *not* implied is the probability that it is not one of the initially positive literals, and that, for all j, for all m clauses c of length j with x as their implicate, at least one of the $j - 1$ literals whose negations appear in c is not implied. Assuming all these events are independent, if t is the fraction of literals that are implied, we have

$$
\begin{aligned}
1-t &= (1-d_1)\prod_{j=2}^{k}\left(\sum_{m=0}^{\infty}\frac{\mathrm{e}^{-d_j}d_j^m}{m!}(1-t^{j-1})^m\right)\\
&= (1-d_1)\prod_{j=2}^{k}\exp(-d_j t^{j-1})) = (1-d_1)\,\exp\left(-\sum_{j=2}^{k}d_j t^{j-1}\right)
\end{aligned}
$$

yielding Equation (1.1).

To make this rigorous, we use a standard technique for proving results about threshold properties, namely analysis of algorithms via differential equations [34] (see [1] for a review). Specifically, we analyze the Positive Unit Resolution (PUR) algorithm given in Figure 3, which is complete for Horn-SAT. We analyze its while loop by considering a series of *stages*, indexed by the number of literals that are labeled "implied." After T stages of this process, T variables are labeled as implied. At each stage the resulting formula consists of a set of Horn clauses of length j for $1 \le j \le k$ on the $n-T$ unlabeled variables. Let the number of distinct clauses of length j in this formula be $S_j(T)$; we rely on the fact that, at each stage T, conditioned on the values of $S_j(T)$ the formula is uniformly random. This follows from an easy induction argument which is standard for problems of this type (see e.g. [22]).

Now, the variables appearing in the clauses present at stage T are chosen uniformly from the $n-T$ remaining variables, so the probability that the chosen variable x appears in a given clause of length j is $j/(n-T)$, and the probability that a given clause of length $j+1$ is shortened to one of length j (as opposed to removed) is $j/(n-T)$. A newly shortened clause is distinct from all the others with probability $1-o(1)$ unless $j=1$, in which case it is distinct with probability $(n-T-S_1)/(n-T)$. Finally, each stage labels the variable in one of the $S_1(T)$ unit clauses as implied. Writing $\Delta S_j(T) = S_j(T+1) - S_j(T)$, the expected effect of each step is then

$$
\begin{aligned}
\mathrm{E}[\Delta S_j(T+1)] &= j\,\frac{S_{j+1}(T)-S_j(T)}{n-T} + o(1) \quad \text{for all } 2 \le j \le k\\
\mathrm{E}[\Delta S_1(T+1)] &= \left(\frac{n-T-S_1}{n-T}\right)\left(\frac{S_2(T)}{n-T}\right) - 1 + o(1)
\end{aligned}
\tag{3.2}
$$

Let us rescale these difference equations by setting $T = t\cdot n$ and $S_j(T) = s_j(t)\cdot n$, obtaining the following system of differential equations:

$$
\begin{aligned}
\frac{\mathrm{d}s_j}{\mathrm{d}t} &= j\,\frac{s_{j+1}(t)-s_j(t)}{1-t} \quad \text{for all } 2 \le j \le k\\
\frac{\mathrm{d}s_1}{\mathrm{d}t} &= \left(\frac{1-t-s_1(t)}{1-t}\right)\left(\frac{s_2(t)}{1-t}\right) - 1\ .
\end{aligned}
\tag{3.3}
$$

Note that the right-hand side of these equations are continuous functions of the s_j. In addition, the fact that w.h.p. no variable appears in more than $O(\log n)$ clauses implies a simple tail bound on the $\Delta S_j(T)$. These observations satisfy the conditions of Wormald's theorem [34], which implies that for any constant $\delta > 0$, for all t such that $s_1 \geq \delta$, w.h.p. we have $S_j(t \cdot n) = s_j(t) \cdot n + o(n)$ where $s_j(t)$ is the solution to the system (3.3). With the initial conditions $s_j(0) = d_j$ for $1 \leq j \leq k$, a little work shows that this solution is

$$\begin{aligned} s_j(t) &= (1-t)^j \sum_{\ell=j}^{k} \binom{\ell-1}{j-1} d_\ell t^{\ell-j} \quad \text{for all } 2 \leq j \leq k \\ s_1(t) &= 1 - t - (1-d_1) \exp\left(-\sum_{j=2}^{k} d_j t^{j-1} \right) \quad . \end{aligned} \tag{3.4}$$

Note that $s_1(t)$ is continuous, $s_1(0) = d_1 > 0$, and $s_1(1) < 0$ since $d_1 < 1$. Thus $s_1(t)$ has at least one root in the interval $(0, 1)$. Since PUR halts when there are no unit clauses, i.e., when $S_1(T) = 0$, we expect the stage at which it halts to be $T = t_0 n + o(n)$ where t_0 is the smallest positive root of $s_1(t) = 0$, or equivalently, dividing by $1 - d_1$ and taking the logarithm, of Equation (1.1).

We do not run the differential equations all the way up to stage $t_0 n$, since once there is a significant probability that $S_1 = 0$ and the algorithm has already halted, the difference equations (3.2) no longer hold. Therefore, we choose small constants $\epsilon, \delta > 0$ such that $s_1(t_0 - \epsilon) = \delta$ and run the algorithm until stage $(t_0 - \epsilon)n$. At this point $(t_0 - \epsilon)n$ literals are already implied, proving the lower bound of (3.1).

To prove the upper bound of (3.1), recall that by assumption t_0 is a simple root of (1.1), i.e., the second derivative of the left-hand side of (1.1) with respect to t is nonzero at t_0. It is easy to see that this is equivalent to $\mathrm{d}s_1/\mathrm{d}t < 0$ at t_0. Since $\mathrm{d}s_1/\mathrm{d}t$ is analytic, there is a constant $c > 0$ such that $\mathrm{d}s_1/\mathrm{d}t < 0$ for all $t_0 - c \leq t \leq t_0 + c$. Set $\epsilon < c$; the number of literals implied during these stages is bounded above by a subcritical branching process whose initial population is w.h.p. $\delta n + o(n)$. Since the probability an individual in a subcritical branching process has ℓ descendants is bounded by $(1-\eta)^\ell$ for some $\eta > 0$ (see e.g. [2], we can bound the total progeny during this stage to be w.h.p. at most $\epsilon' n$ for any $\epsilon' > 0$ by taking δ small enough. ■

It is easy to see that the backbone of implied positive literals is a uniformly random subset of $\{x_1, \ldots, x_n\}$ of size $N_{\mathbf{d},n}$. Since x_1 is guaranteed to not be among the $d_1 n$ initially positive literals, the probability that x_1 is not in this backbone is

$$\frac{n - N_{\mathbf{d},n}}{(1-d_1) \cdot n}.$$

By completeness of positive unit resolution for Horn satisfiability, this is precisely the probability that the ϕ is satisfiable. Applying Lemma 3.1 and taking $\epsilon \to 0$ proves equation (1.2) and completes the proof of Theorem 1.1.

We make several observations. First, if we set $k = 2$ and take the limit $d_1 \to 0$, Theorem 3.1 recovers Karp's result [23] on the size of the reachable component of a random directed graph with mean out-degree $d = d_2$, namely the root of $\ln(1-t) + dt = 0$.

Secondly, as we will see below, the condition that t_0 is simple is essential. Indeed, for the 1-3-Horn-SAT model studied in [15], the curve Γ of discontinuities, where the probability of satisfiability drops in the "waterfall" of Figure 1, consists exactly of those (d_1, d_3) where t_0 is a double root, which implies $\mathrm{d}s_1/\mathrm{d}t = 0$ at t_0.

Finally, we note that Theorem 3.1 is very similar to Darling and Norris's work [14] on a type of reachability in random undirected hypergraphs called "identifiability." If the number of hyperedges of length j is Poisson-distributed with mean β_j, their result for the fraction t of identifiable vertices is

$$\ln(1-t) + \sum_{j=1}^{k} j\beta_j t^{j-1} = 0 \ .$$

We can make this match (1.1) as follows. First, since each hyperedge of length j corresponds to j Horn clauses, we set $d_j = j\beta_j$ for all $j \geq 2$. Then, since edges are chosen with replacement in their model, the expected number of distinct clauses of length 1 (i.e., positive literals) is $d_1 n$ where $d_1 = 1 - \mathrm{e}^{-\beta_1}$.

4. APPLICATION TO 1-2-HORN-SAT

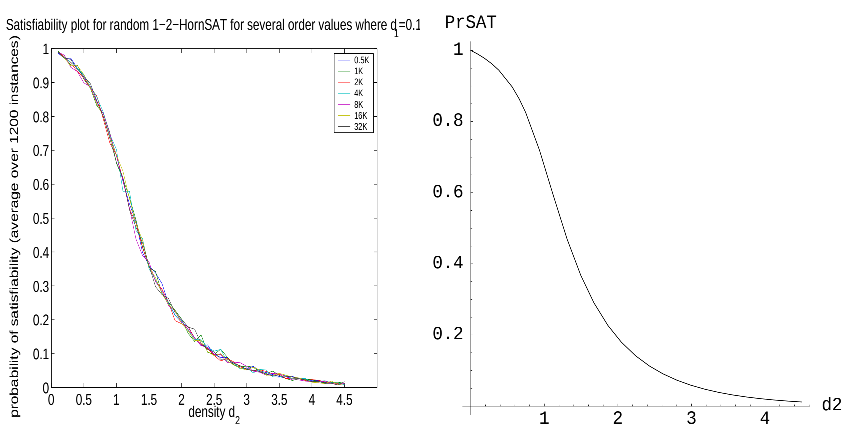

Fig. 4. The probability of satisfiability for 1-2-Horn formulae as a function of d_2, where $d_1 = 0.1$. Left, the experimental data of [15]; right, our analytic results.

For H^2_{n,d_1,d_2}, Equation (1.1) can be rewritten

$$1 - t = (1 - d_1) \cdot \mathrm{e}^{-d_2 t} \ . \tag{4.1}$$

Denoting $y = d_2(t-1)$, this implies

$$y\mathrm{e}^y = d_2(t-1)\cdot \mathrm{e}^{d_2(t-1)} = -d_2(1-d_1)\cdot \mathrm{e}^{-d_2 t}\cdot \mathrm{e}^{d_2(t-1)} = -(1-d_1)d_2\mathrm{e}^{-d_2}.$$

Solving this yields

$$t_0 = 1 + \frac{1}{d_2}W\left(-(1-d_1)d_2\mathrm{e}^{-d_2}\right) \ . \tag{4.2}$$

To show that this is a simple root, note that the derivative of the left-hand side of (1.1) with respect to t is zero if and only if

$$1 - t = \frac{1}{d_2} \tag{4.3}$$

and substituting this into (4.1) gives

$$\frac{1}{d_2} = (1-d_1)\mathrm{e}^{1-d_2} < \mathrm{e}^{1-d_2} \ .$$

This is a contradiction, since $\mathrm{e}^{1-d_2} \le 1/d_2$ for all $d_2 \ge 1$, and if $d_2 < 1$ then (4.3) implies that $t < 0$.

Finally, substituting (4.2) into (1.2) proves Equation (1.3) and Proposition 1.2. In Figure 4 we plot the probability of satisfiability $\Phi(d_1, d_2)$ as a function of d_2 for $d_1 = 0.1$ and compare with the experimental results of [15]; the agreement is excellent.

5. FIRST AND SECOND-ORDER TRANSITIONS FOR 1-3-HORN-SAT

For the random model $H^3_{n,d_1,0,d_3}$ studied in [15], Eq. (1.1) becomes

$$\ln\frac{1-t}{1-d_1} + d_3 t^2 = 0 \ . \tag{5.1}$$

An analytic solution to this equation does not seem to exist. To find its solutions graphically, it is useful to rewrite it as

$$1 - t = f(t) := (1-d_1)\mathrm{e}^{-d_3 t^2} \ . \tag{5.2}$$

We claim that for some values of d_1 and d_3 there is a phase transition in the roots of (5.2). For instance, consider the plot of $f(t)$ shown in Figure 5 for $d_1 = 0.1$ and $d_3 = 3$. Here $f(t)$ is tangent to $1-t$, so there is a bifurcation as we vary either parameter; for $d_3 = 2.9$, for instance, $f(t)$ crosses $1-t$ three times and there is a root of (5.2) at $t = 0.185$, but for $d_3 = 3.1$ the unique root is at $t = 0.943$. This causes the probability of satisfiability to jump discontinuously but from 0.905 to 0.064.

The set of pairs (d_1, d_3) for which this tangency occurs is exactly the curve Γ on which the smallest positive root t_0 of (5.1) is a double root, giving the “waterfall” of Figure 1. To

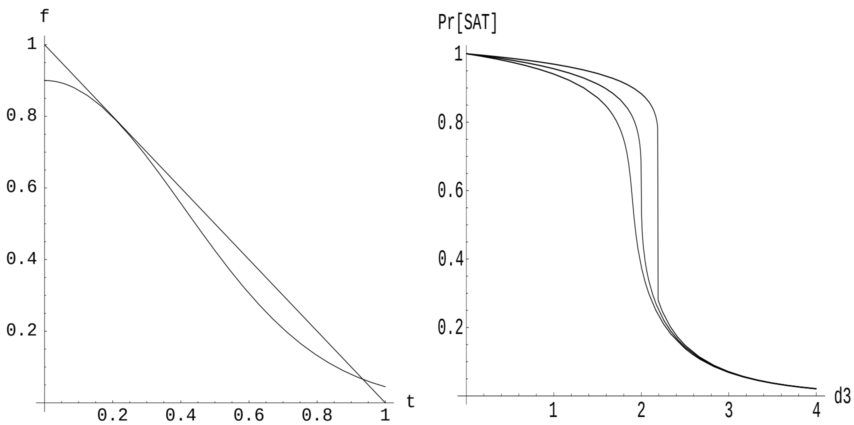

Fig. 5. Left, the function $f(t)$ of (5.2) when $d_1 = 0.1$ and $d_3 = 3$. Right, the probability of satisfiability $\Phi(d_1, d_3)$ with d_1 equal to 0.15 (continuous), 0.1756 (critical), and 0.2 (discontinuous).

see this, set the derivative of the left-hand side of (5.1) with respect to t to zero, obtaining

$$-\frac{1}{1-t} - 2d_3 t = 0 \tag{5.3}$$

This is equivalent to

$$-\frac{1}{1-t} - \frac{f'}{f} = -\frac{1}{1-t} - \frac{f'}{1-t} = 0$$

and so $f' = 1$, which is precisely when $f(t)$ is tangent to $1 - t$. The smallest solution to (5.3) is

$$t_0 = \frac{1}{2}\left(1 - \sqrt{1 - \frac{2}{d_3}}\right) \ . \tag{5.4}$$

and (5.1) then gives

$$d_1 = 1 - \frac{\mathrm{e}^{d_3 t_0^2}}{2 d_3 t_0} \ . \tag{5.5}$$

Combining (5.4) with (5.5) and simplifying gives (1.4), proving Proposition 1.3.

The fact that the root t_0 of (5.3) is only real for $d_3 \geq 2$ explains why Γ ends at $d_3 = 2$. At this extreme case we have

$$d_1 = 1 - \frac{\sqrt{\mathrm{e}}}{2} \approx 0.1756 \ \text{ and } \ \frac{\partial d_1}{\partial d_3} = -\frac{\sqrt{\mathrm{e}}}{8} \ .$$

For $d_3 < 2$, the root t_0 of (5.1) is unique and simple, and therefore the probability of satisfiability $\Phi(d_1, d_3)$ is an analytic function of d_1 and d_3. To illustrate this, in the right part of Figure 5 we plot $\Phi(d_1, d_3)$ as a function of d_3 with three values of d_1. For $d_1 =$

0.15, Φ is continuous; for $d_1 = 0.2$, it is discontinuous; and for $d_1 = 0.1756...$, the critical value at the second-order transition, it is continuous but has an infinite derivative at $d_3 = 2$.

ACKNOWLEDGMENTS

We thank Mark Newman for helpful conversations on the classification of phase transitions. CM would also like to acknowledge Tracy Conrad and Rosemary Moore for helpful conversations, and Makalawena Beach on the Big Island of Hawai'i where much of this work was done. This work was supported by the U.S. Department of Energy under contract W-705-ENG-36, by NSF grants PHY-0200909, EIA-0218563, CCR-0220070, CCR-0124077, CCR-0311326, IIS-9908435, IIS-9978135, eIA-0086264, and ANI-0216467, by BSF grant 9800096, and by Texas ATP grant 003604-0058-2003.

REFERENCES

[1] D. Achlioptas. Lower Bounds for Random 3-SAT via Differential Equations. *Theoretical Computer Science* 265 (1-2), 159–185, 2001.

[2] D. Achlioptas and C. Moore. Almost all graphs of degree 4 are 3-colorable. *J. Comput. System Sci.*, **67** (2003) 441–471, invited paper in special issue for STOC 2002.

[3] D. Achlioptas, A. Chtcherba, G. Istrate, and C. Moore. The phase transition in 1-in-k SAT and NAE 3-SAT. In *Proc. 12th ACM-SIAM Symp. on Discrete Algorithms*, 721–722, 2001.

[4] D. Achlioptas, L.M. Kirousis, E. Kranakis, and D. Krizanc. Rigorous results for random $(2+p)$-SAT. *Theor. Comput. Sci.* 265(1-2): 109-129 (2001).

[5] T. Bench-Capon and P. Dunne. A sharp threshold for the phase transition of a restricted satisfiability problem for Horn clauses. *Journal of Logic and Algebraic Programming*, 47(1):1–14, 2001.

[6] J. J. Binney, N. J. Dowrick, A. J. Fisher and M. E. J. Newman. *The Theory of Critical Phenomena.* Oxford University Press (1992).

[7] B. Bollobás. *Random Graphs*. Academic Press, 1985.

[8] S. Cocco, O. Dubois, J. Mandler, and R. Monasson. Rigorous decimation-based construction of ground pure states for spin glass models on random lattices. *Phys. Rev. Lett.*, 90, 2003.

[9] V. Chvátal and B. Reed. Mick gets some (the odds are on his side). In *Proc. 33rd IEEE Symp. on Foundations of Computer Science*, 620–627, 1992.

[10] J. M. Crawford and L. D. Auton. Experimental results on the crossover point in random 3-SAT. *Artificial Intelligence*, 81(1-2):31–57, 1996.

[11] W. F. Dowling and J. H. Gallier. Linear-time algorithms for testing the satisfiability of propositional Horn formulae. *Logic Programming. (USA) ISSN: 0743-1066*, 1(3):267–284, 1984.

[12] O. Dubois, Y. Boufkhad, and J. Mandler. Typical random 3-SAT formulae and the satisfiability theshold. In *Proc. 11th ACM-SIAM Symp. on Discrete Algorithms*, 126–127, 2000.

[13] O. Dubois and J. Mandler. The 3-XORSAT threshold. In *Proc. 43rd IEEE Symp. on Foundations of Computer Science*, 769–778, 2002.

[14] R. Darling and J.R. Norris. Structure of large random hypergraphs. *Annals of Applied Probability*, 15(1A), 2005.

[15] D. Demopoulos and M. Vardi. The phase transition in random 1-3 Hornsat problems. In A. Percus, G. Istrate, and C. Moore, editors, *Computational Complexity and Statistical Physics*, Santa Fe Institute Lectures in the Sciences of Complexity. Oxford University Press, 2005. Available at `http://www.cs.rice.edu/~vardi/papers/`.

[16] P. Erdös and A. Rényi. On the evolution of random graphs. *Publications of the Mathematical Institute of the Hungarian Academy of Science*, 5:17–61, 1960.

[17] E. Friedgut. Necessary and sufficient conditions for sharp threshold of graph properties and the k-SAT problem. *J. Amer. Math. Soc.*, 12:1917–1054, 1999.

[18] A. Goerdt. A threshold for unsatisfiability. *J. Comput. System Sci.*, 53(3):469–486, 1996.

[19] M. Hajiaghayi and G.B. Sorkin. The satisfiability threshold for random 3-SAT is at least 3.52. IBM Technical Report , 2003.

[20] T. Hogg and C. P. Williams. The hardest constraint problems: A double phase transition. *Artificial Intelligence*, 69(1-2):359–377, 1994.

[21] G. Istrate. The phase transition in random Horn satisfiability and its algorithmic implications. *Random Structures and Algorithms*, 4:483–506, 2002.

[22] G. Istrate. On the satisfiability of random k-Horn formulae. In P. Winkler and J. Nesetril, editors, *Graphs, Morphisms and Statistical Physics*, volume 64 of *AMS-DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, 113–136, 2004.

[23] R. Karp. The transitive closure of a random digraph. *Random Structures and Algorithms*, 1:73–93, 1990.

[24] A. Kaporis, L. M. Kirousis, and E. Lalas. Selecting complementary pairs of literals. In *Proceedings of LICS'03 Workshop on Typical Case Complexity and Phase Transitions*, June 2003.

[25] J. A. Makowsky. Why Horn formulae matter in Computer Science: Initial structures and generic examples. *JCSS*, 34(2-3):266–292, 1987.

[26] M. Mézard and R. Zecchina. Random k-satisfiability problem: from an analytic solution to an efficient algorithm. *Phys. Rev. E*, 66:056126, 2002.

[27] M. Mézard, G. Parisi, and R. Zecchina. Analytic and algorithmic solution of random satisfiability problems. *Science*, 297:812–815, 2002.

[28] R. Monasson, R. Zecchina, S. Kirkpatrick, B. Selman, and L. Troyansky. $2+p$-SAT: Relation of typical-case complexity to the nature of the phase transition. *Random Structures and Algorithms*, 15(3–4):414–435, 1999.

[29] F. Nielson, H. R. Nielson, and H. Seidl. Normalizable Horn clauses, strongly recognizable relations and Spi. In 9th Static Analysis Symposium (SAS). Springer Verlag, LNCS 2477, 2002.

[30] B. Selman, D. G. Mitchell, and H. J. Levesque. Generating hard satisfiability problems. *Artificial Intelligence*, 81(1-2):17–29, 1996.

[31] B. Selman and S. Kirkpatrick. Critical behavior in the computational cost of satisfiability testing. *Artificial Intelligence*, 81(1-2):273–295, 1996.

[32] M.Y. Vardi and P. Wolper. Automata-Theoretic Techniques for Modal Logics of Programs *J. Comput. System Sci.* 32:2(1986), 181–221, 1986.

[33] M.Y. Vardi and P. Wolper. An automata-theoretic approach to automatic program verification (preliminary report). In *Proc. 1st IEEE Symp. on Logic in Computer Science*, 332–344, 1986.

[34] N. Wormald. Differential equations for random processes and random graphs. *Annals of Applied Probability*, 5(4):1217–1235, 1995.

Received (date)
Revised (date)
Accepted (date)