PERTURBATIVE METHODS FOR MOSTLY MONOTONIC PROBABILISTIC SATISFIABILITY PROBLEMS

STEPHEN EUBANK*, MADHURIMA NATH†, YIHUI REN‡, AND ABHIJIN ADIGA*

Abstract. The probabilistic satisfiability of a logical expression is a fundamental concept known as the partition function in statistical physics and field theory, an evaluation of a related graph's Tutte polynomial in mathematics, and the Moore-Shannon network reliability of that graph in engineering. It is the crucial element for decision-making under uncertainty. Not surprisingly, it is provably hard to compute exactly or even to approximate. Many of these applications are concerned only with a subset of problems for which the solutions are monotonic functions. Here we extend the weak- and strong-coupling methods of statistical physics to heterogeneous satisfiability problems and introduce a novel approach to constructing lower and upper bounds on the approximation error for monotonic problems. These bounds combine information from both perturbative analyses to produce bounds that are tight in the sense that they are saturated by some problem instance that is compatible with all the information contained in either approximation.

Key words. satisfiability, weak-coupling, strong-coupling, network reliability, complex networks

AMS subject classifications. 68R07 60C05 05C20 05C21 05C22 05C31 05C75 05C82 05C85 60K35 82M36

1. Introduction.

At the heart of some of the thorniest problems in physics, computer science, engineering, and combinatorics lies a question that can be stated simply: given the probabilities that any of a large set of events occur, what is the overall probability that a logical statement about combinations of those events is true? We can think of the statement, $\mathcal{E}(c)$, as an assertion that a dynamical system in configuration c has a certain property. For example, given a lattice of magnets free to flip their orientations, what is the probability that two magnets separated by k lattice sites point in the same direction? Or given the presence of an electron at position x_0 at time t_0, what is the probability of observing an electron at position x_1 at time t_1? Or given the probability of failure for "crummy" [11] relays in an electrical circuit, what is the probability of establishing a current from one terminal to another? This question is known as the probabilistic satisfiability[6] (PSAT) problem. Hard enough in the homogeneous case, when the probabilities of every event are the same, it becomes nightmarish in the more important heterogenous case, when the probability of each event can be different. Heterogeneous systems arise if, for example, the lattice of magnets is stretched anisotropically, or each relay in the circuit is different. A general method for answering it would find immediate technological applications in designing networks to avoid cascading failure or optimizing vaccination strategies, among many others. Here we synthesize the explicit symmetries of Max Flow / Min Cut properties with the mathematical framework of network reliability and the perturbative analyses of statistical physics and field theory to produce a hierarchy of approximate solutions with controllable, bounded error.

A system's *configuration* is an assignment of states to each of its elements. The system is defined by a probability distribution, p, over configurations. In a system with a finite number of configurations, the probability that $\mathcal{E}$ is true is simply the sum of the probabilities of all configurations for which $\mathcal{E}$ is true, or, in terms of the indicator function δ which is 1 if its argument is true and 0 else, $p(\mathcal{E}) = \sum_{c\in\mathcal{C}} \delta(\mathcal{E}(c))p(c)$. Partitioning the configurations into equiprobable equivalence classes, $\mathcal{K}$, this can be written as $p(\mathcal{E}) = \sum_{k\in\mathcal{K}} n(k)p_k$, where $n(k)$ is the fraction of configurations

*Biocomplexity Institute, University of Virginia (eubank@virginia.edu, abhijin@virginia.edu).
†Slalom, (mnath@vt.edu).
‡Brookhaven National Laboratory, (yren@bnl.gov).

that have probability p_k for which $\mathcal{E}$ is true. The function $n(k)$ is known as the *density of states* function.

When only the relative probability $\widehat{p}$ of different configurations is known, properly normalized probabilities can be obtained as $p = \widehat{p}/Z$, where the normalization constant $Z = \sum_{k \in \mathcal{K}} n(k)\widehat{p}_k$ is known as the *partition function*. For a large class of physical systems, p is the Boltzmann distribution, an exponential distribution that depends on the energy $E(c)$ of a configuration and the inverse of the system's temperature, β, i.e., $p(c) \propto e^{-\beta E(c)}$. In this case, the normalization "constant" is a function of temperature:

$$Z(\beta) = \sum_{k=1}^{K} n(k) e^{-\beta E_k}, \tag{1.1}$$

assuming every configuration has one of K distinct energies, E_k. The partition function is key to understanding statistical systems because its logarithm is the moment generating function. For example:

$$-\frac{\partial}{\partial \beta} \ln Z = \frac{\sum_{k=1}^{K} E_k n(k) e^{-\beta E_k}}{Z} = \langle E \rangle. \tag{1.2}$$

Unfortunately, since evaluating the partition function is equivalent to solving PSAT or solving a network reliability problem, it is in the complexity class #P [16]. In practice, the solution is computationally infeasible for the foreseeable future even for fairly small sets of events. Exact solution requires summing over each of 2^N possible configurations. Fortunately, when $\mathcal{E}$ is *monotonic* – i.e., the sense of each event can be chosen so that $\mathcal{E}$ contains no negations – the oracle does not need to be called for every configuration. Typical expressions in physical systems exhibit regularities that allow some shortcuts, such as explicit forms for the density of states, but this is not generally the case in other domains.

There are several approaches to approximating Z for a monotonic $\mathcal{E}$ when $\widehat{p}$ is known:

1. *Probabilistic methods* and *simulations* approximate $n(k)$. They rely on algorithms that count approximately how many configurations are in each equivalence class in $\mathcal{K}$ and how many of those satisfy $\mathcal{E}$.
2. *Renormalization* produces an *effective* theory with fewer degrees of freedom and thus fewer configurations in the sum. Renormalization recursively "integrates out" some degrees of freedom – e.g., alternate sites in a lattice – to define a new problem with the same solution. Indeed, PSAT is equivalent to an evaluation of a Tutte polynomial, which can be defined by recursive renormalization [1].
3. *Perturbative methods* compute leading terms in a Taylor series expansion of the solution for a system whose parameters have been perturbed slightly away from a system with a known solution.

Roth has shown that approximately counting solutions is also hard [15], but his arguments apply to *relative*, not absolute, error:

> The notion of approximation we [Roth] use is that of relative approximation. . . . For example, there exists a polynomial time randomized algorithm that approximates the number of satisfying assignments of a DNF formula within any constant ratio. It is possible, though, for a #P-complete problem, even if its underlying decision problem is easy, to resist even an efficient approximate solution. . . . We prove, for various propositional languages for which solving satisfiability is easy,

that it is NP-hard to approximate the number of satisfying assignments even in a very weak sense.

Probabilistic satisfiability places the problem in a continuous rather than discrete context, where notions of smoothness make sense. Perturbative approaches, called weak- and strong-coupling expansions in statistical physics, take advantage of smoothness to provide excellent approximations for many problems [3]. Indeed, perturbative approximation of the "propagator" in quantum electrodynamics (which is an infinite PSAT problem) resulted in what is generally acknowledged to be the most careful comparison between experiment and theoretical implications of a physical theory ever made [5].

Here we develop practically useful approximations for finite, monotonic PSAT using a novel combination of perturbative and probabilistic methods with some elements of renormalization and simulation. Furthermore, we show how to interpolate between two perturbative approximations to enforce constraints on unitarity and monotonicity that the truncated Taylor series do not obey. Finally, we use the interpolated perturbation series to construct bounds on the approximation error, which improve on bounds constructed using only one or the other [9]. The interpolation relies on three principles – positivity, unitarity, and duality – that emphasize different aspects of the simple observationthat all probabilities lie in the interval $[0, 1]$. The approach provides a nested hierarchy of perturbative approximations with bounded error giving a controllable trade-off between computational complexity and approximation error. The approximations are controlled by two parameters: S, the number of samples generated in the probabilistic part, and D, the depth of expansion in the perturbative part. For large enough S and D, the approximation becomes exact. When D is small, the quality of approximation depends on the problem instance. When S is sufficiently large, the bounds on approximation error are tight in the sense that there are expressions $\mathcal{E}'$ that are consistent with the perturbative approximations and saturate the bounds. When S is small, the bounds are not strict for $\mathcal{E}$, but they are tight bounds on a simpler problem $\mathcal{E}'$ that contains all the information about $\mathcal{E}$ that is available in the sample.

Section 2 gives a formal statement of the problem. In particular, subsection 2.4 reconciles the claims for our approximations with the fact that even approximating the satisfiability is NP-hard. Section 3 maps the problem to a graphical setting analogous to Feynman diagrams, where the satisfiability becomes the propagator, the 2-terminal Moore-Shannon network reliability [11] or, equivalently, an evaluation of the Tutte polynomial. The approximation methods are described in section 4 for the case of identically distributed variables. A fully worked, nontrivial example is explained in section 5. The extension to non-identically distributed variables is developed in section 6 and applied to several variants of the example in section 7.

2. Problem statement.

2.1. Events.

2.1. Events. Consider a set of N events E_i which might or might not occur in any instance of a random process. We assume the events occur independently of each other. Sometimes the events must be defined carefully to ensure they are truly independent. For example, in simple models of disease transmission, whether one person is infected depends on who else is infected, whereas whether one infected person transmits to a susceptible person is independent of other transmission events. The former is the kind of compound event that is captured by the expression $\mathcal{E}$, but it is not an atomic event E. With each event E_i for $i \in \{1, \ldots, N\}$, we associate a Bernoulli random variable $e_i \in \{0, 1\}$ and a probability $\widetilde{x}_i \in [0, 1]$ that $e_i = 1$, meaning that E_i occurred. Until section 6, we assume a homogeneous system with $\widetilde{x}_i = x$ for all i.

2.2. Satisfiability Expression. Suppose there is a monotonic Boolean expression $\mathcal{E}$ in disjunctive normal form, i.e.,

$$\mathcal{E} = \bigvee_{i=1}^{m} \left(\bigwedge_{j \in L_i} e_j \right), \tag{2.1}$$

where $L_i \subseteq \{1, \ldots, N\}$. To avoid trivial cases we assume that $|L_i| \geq 1$ and $m \geq 1$. Given an assignment of values to every e_i – i.e., a *system configuration* – the expression $\mathcal{E}$ evaluates deterministically to either true or false. We say the expression is *satisfied* if it evaluates to true, and the configuration is a *solution*. We refer to each of the m conjunctions in Equation (2.1) as a *clause*. Each clause is defined by a set L_i of events that must all occur in order for the clause to be satisfied. We also refer to the set of *integers* L_i as *clause* i when it causes no ambiguity.

2.3. Satisfiability Problems. The deterministic satisfiability literature focuses on counting solutions. *Probabilistic* satisfiability asks instead: What is $\Xi(\mathcal{E}, \widetilde{\boldsymbol{x}})$, the probability that the expression $\mathcal{E}$ is satisfied given the probability of individual events $\widetilde{x}_i$? $\Xi(\mathcal{E}, \widetilde{\boldsymbol{x}})$ is a sum of probabilities over all solutions; evaluating it is thus at least as hard as counting the solutions. Indeed, when every event occurs with probability $1/2$, the satisfiability reduces to the ratio of the number of solutions to the number of configurations, 2^N.

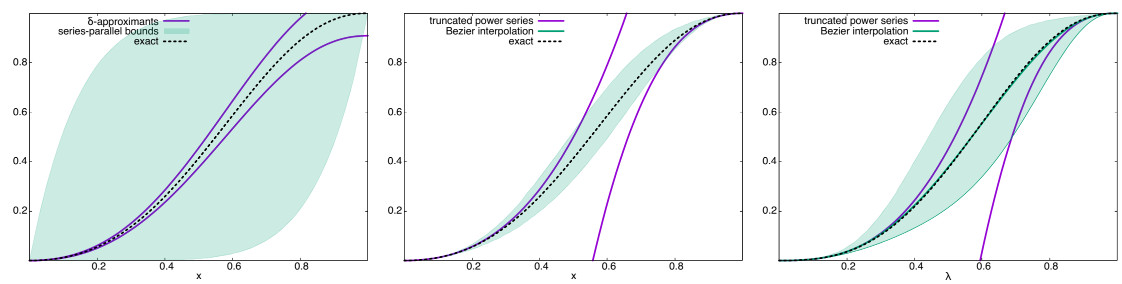

FIG. 2.1. *Satisfiability for the example expression in (5.1), as a function of $x = p(e_2)$. (Left) Upper and lower bounds on (hypothetical) δ-approximants to the satisfiability, with $\delta = 0.1$, along with the strict, but not tight, bounds of (2.10) based only on the number of events in the expression. (Center) Truncated Taylor series for the homogeneous example to $O(x^3)$ at $x = 0$ and $O((1-x)^2)$ at $x = 1$, along with the bounds and interpolation developed here. (Right) The same as the center panel, but for the heterogeneous problem described in section 7 with $(\widetilde{x}_1, \widetilde{x}_6) = (1/2, 1/4)$.*

2.4. On the nature of approximation. It is important to distinguish between the notions of "approximation" used here and in algorithmic complexity. Specifically, Roth considers both *relative approximation* – the result M' is a δ-approximation to M if and only if $M'/(1+\delta) \leq M \leq M'(1+\delta)$ for $\delta \geq 0$ – and the use of *approximate theories* – expressions $\mathcal{E}'$ and $\mathcal{E}''$ for which $p(\mathcal{E}') \leq p(\mathcal{E}) \leq p(\mathcal{E}'')$. Relative approximation is known to be NP-hard for PSAT, but approximate renormalization and perturbative methods produce approximate theories. The left panel of Figure 2.1 illustrates differences in the character of results for the example in sections 5 and 7. The simplest perturbative approximations, based only on the number of events in $\mathcal{E}$ as in (2.10), constrain the exact solution to lie in the green band in the figure; a relative approximation to the exact answer (dashed curve) with $\delta = 0.1$ would lie between the two solid curves. Clearly, at this order, the perturbative approximation is almost useless, and Roth correctly judged the use of approximate theories to be ineffective for the decision problem he studied. However, notice that relative approximation does not automatically incorporate constraints such as unitarity or duality, and that, of course, the absolute error can be large where the exact answer is large. By including more information about the expression in a perturbative analysis, it is possible to "squeeze" the bounds dramatically, as illustrated in the center and right panels of the figure. The resulting approximation respects unitarity and duality, and provides small absolute errors at both ends of the domain. However, the bounds cannot be squeezed much in the middle, where, as we show below, there is little information, so the relative error cannot be reduced uniformly below an arbitrary threshold δ.

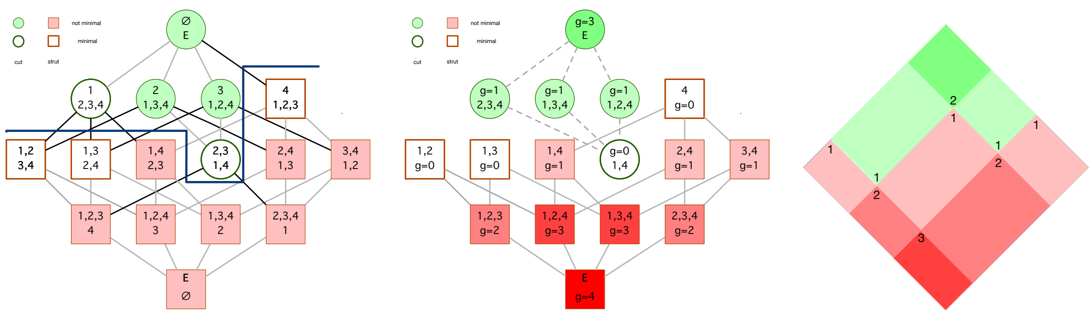

FIG. 2.2. *Perspectives on the partially ordered sets corresponding to the expression* $\mathcal{E} = (e_1 \wedge e_2) \vee (e_1 \wedge e_3) \vee e_4$. *Each configuration is labelled (redundantly) with the indices of the events with* $e_i = 1$ *on the top line and those with* $e_i = 0$ *on the bottom. The* i*-th row, starting from* $i = 0$ *at the top, contains all configurations with exactly* i *true events. Cuts are indicated by circles; struts by squares. Filled circles or squares are not minimal. Left: A separatrix between cuts and struts in the lattice of all possible configurations for a monotonic expression over* $N = 4$ *events. Darker edges indicate those that connect a strut to a cut. The separatrix is the thick, orthogonal curve from the left to the right edge of the lattice. All minimal struts and cuts are adjacent to the separatrix, but the converse is not true. Center: The descendants of minimal cuts and struts and the degeneracy of each configuration (the number of direct ancestors that are cuts or struts). Only the edges that do not cross the separatrix are shown, for clarity. Right: A notional depiction of the "light cones" of minimal struts (red) and cuts (green). The degeneracy of configurations within a cone is indicated by the color's shade and labeled at the minimum configuration of the region. If there are more than two minimal struts, the partition into cones is not necessarily planar, as illustrated in the center panel. Here, the minimal struts are either on the boundary or adjacent to cuts with degeneracy 2 and vice versa.*

2.5. Monotonicity, minimality, and total probability. The set of all configurations is the *power set* of the set of events. Any configuration can be labeled by the set of events for which $e_i = 1$. It is convenient to impose the partial order of set inclusion on the set of all configurations and to represent the graded partial order, with the rank of configuration labeled s given by $|s|$, as the lattice depicted in Figure 2.2. We define $D(c)$ to be the set of all descendants of c in the lattice. That is, $D(c) \equiv \{c' \in \mathcal{C} | c' \supseteq c\}$.

The *total probability* p_t of a configuration c is the probability that the associated clause is true, i.e., the sum of probabilities of all its descendants that satisfy $\mathcal{E}$. Clearly, $p_t(c) \geq p(c)$. For a monotonic system, if c satisfies $\mathcal{E}$, so do all its descendants. If the events are independent of each other, the probability of a configuration factorizes into the product of probabilities of the individual events. Thus, for a configuration c that satisfies $\mathcal{E}$,

$$p_t(c) = p\left(\bigwedge_{i\in c} e_i\right) = \sum_{c'\in D(c)} p(c')\delta(\mathcal{E}|c') \tag{2.2}$$

$$\xrightarrow{monotonic} \sum_{c'\in D(c)} p(c') \xrightarrow{independent} \prod_{i\in c} \widetilde{x}_i \xrightarrow{homogeneous} x^{|c|}. \tag{2.3}$$

Because the descendants of any two configurations c_1 and c_2 are not disjoint, the probability of a disjunction of clauses is sub-additive. The Inclusion-Exclusion principle gives

$$p(L_i \vee L_j) = p_t(L_i) + p_t(L_j) - p_t(L_i \cup L_j) \tag{2.4}$$

A finite, monotonic system admits the notion of *minimal* solutions. A configuration c is a minimal solution if it is a solution, but no proper subset of c is a solution. See examples in section 5. Without loss of generality, we will assume that each clause in $\mathcal{E}$ is a different minimal solution, i.e., for distinct i and j, $L_i \nsubseteq L_j$. Under this condition,

$$\sum_{c\in\mathcal{L}} p(c) \le p(\mathcal{E}) \le \sum_{c\in\mathcal{L}} p_t(c). \tag{2.5}$$

These simplistic bounds could be tightened recursively, but it will be easier to use the duality introduced in subsection 2.6.

The descendants of $L_i \cup L_j$ are all included in both the first and second terms of (2.4); the third term serves to correct this overcounting. In general, we define a configuration's *degeneracy* as the number of distinct minimal solutions it has among its ancestors, including itself. The right panel of Figure 2.2 illustrates the degeneracy in a simple example. Just as there is a set of minimal solutions in a monotonic problem, there are sets of minimal configurations with degeneracy $1 \le k \le L$, which we denote by $\mathcal{M}_k$. (Notice that $\mathcal{M}_1 = \mathcal{L}$.) One way to construct $\mathcal{M}_k$ is first to construct all possible unions of exactly k elements of $\mathcal{M}_1$, then to remove those that are supersets of others. Iterating the Inclusion-Exclusion principle in (2.4), we can reorganize the sum into a form that produces a power series:

$$p(\mathcal{E}) = \sum_{c\in\mathcal{C}} \delta(\mathcal{E}|c)p(c) = \sum_{k=1}^{L}(-1)^{k+1} \sum_{c\in\mathcal{M}_k} p_t(c) \xrightarrow[homogeneous]{independent} \sum_{k=1}^{L}(-1)^{k+1} \sum_{c\in\mathcal{M}_k} x^{|c|}. \tag{2.6}$$

2.6. Duality, struts, and cuts. The logical complement of the expression $\mathcal{E}$, denoted $\overline{\mathcal{E}}$, can be expressed in disjunctive normal form using the complements of each random variable, $\overline{e}$:

$$\overline{\mathcal{E}} = \overline{\bigvee_{i=1}^{m}\left(\bigwedge_{j\in L_i} e_j\right)} = \bigwedge_{i=1}^{m}\left(\bigvee_{j\in L_i} \overline{e}_j\right) = \bigvee_{i=1}^{\overline{m}}\left(\bigwedge_{j\in \overline{L}_i} \overline{e}_j\right). \tag{2.7}$$

The clauses $\overline{L}_1, \dots, \overline{L}_{\overline{m}}$ and $\overline{m}$ are related to $L_1, \dots, L_m$ and m in a complicated way. Notice that $\overline{\mathcal{E}}$ is monotonic in terms of the negated variables $\overline{e}$,. The probability that $\overline{e}_i = 1$ is $1 - \widetilde{x}_i$, hence

$$\Xi(\mathcal{E}, \boldsymbol{x}) = 1 - \Xi(\overline{\mathcal{E}}, 1 - \boldsymbol{x}) \tag{2.8}$$

The complement $\overline{\mathcal{E}}$ induces a partial ordering on the power set of $\{1, \ldots, N\}$ that is the dual of the one induced by $\mathcal{E}$, as illustrated in Figure 2.2. This duality is at the heart of the Max Cut / Min Flow relationship [7]. To emphasize this duality, we refer to solutions of $\mathcal{E}$ as *struts* and solutions of $\overline{\mathcal{E}}$ as *cuts*. The set of struts, as their name suggests, are the support of $\Xi(\mathcal{E}, \boldsymbol{x})$; the cuts, as we show below, are the usual cut sets in a graph induced by $\mathcal{E}$.

2.7. Upper and lower bounds: parallel and series expressions. We refer to the conjunction $\mathcal{E}_s = \bigwedge_{i=1}^{k} e_i$ as the *series* expression formed from these events, and the disjunction $\mathcal{E}_p = \bigvee_{i=1}^{k} e_i$ as the *parallel* expression. These expressions are duals of each other, with

$$\Xi(\mathcal{E}_s, x) = x^k \text{ and } \Xi(\mathcal{E}_p, x) = 1 - (1 - x)^k \tag{2.9}$$

The identity Equation (2.8) is easily verified for this case. Moreover, the series and parallel expressions bound the satisfiability of any non-trivial, finite, monotonic, homogeneous expression on exactly N variables. Considering nothing about $\mathcal{E}$ except that it contains N distinct events, we still know there must be at least one solution, and the smallest probability it can have in the homogeneous case is x^N. Applying the same logic to the dual expression and using (2.5) and (2.8) gives

$$\Xi(\mathcal{E}_s, x) = x^N \leq \Xi(\mathcal{E}, x) \leq 1 - (1 - x)^N = \Xi(\mathcal{E}_p, x). \tag{2.10}$$

2.8. Implicit PSAT. The perturbative approximations here are predicated on explicit expressions for *both* $\mathcal{E}$ and $\overline{\mathcal{E}}$ in terms of minimal struts and cuts, respectively. Life rarely provides either, much less both, and almost never in terms of minimal clauses. This section describes a simulation method for sampling the minimal clauses when they are unavailable *a priori*. This step adds another layer of approximation to the problem, and it is not yet clear how this affects the overall approximation error. In principle, analyzing a small sample of minimal struts or cuts is equivalent to analyzing a simpler expression $\mathcal{E}$.

Although the individual events e_i in a PSAT problem are not deterministic, deciding whether a particular set of events forms a solution is. We will assume that $\mathcal{E}$ is embodied in a deterministic, binary oracle instead of an extensional disjunctive normal form expression. A monotonic problem admits a continuous separatrix splitting the lattice of possible outcomes into cuts and struts, as shown in Figure 2.2. Any path from the known cut at the top of the lattice to the known strut at the bottom must intersect the separatrix exactly once, so it can be located by a binary search in $O(\log N)$ steps. However, not every configuration bordering the separatrix is minimal – it is necessary to test for minimality, which could require up to $O(N)$ tests, i.e., calls to the oracle. If the configuration is not minimal, a minimal one can be found by walking along the separatrix, which requires at most an additional $O(N)$ steps. In the worst case, it is possible that the process requires $O(N^2 \log N)$ calls to the oracle. Average-case complexity is not obvious. The search can be biased, for example towards or away from finding completely disjoint minimal sets or towards finding the smallest minimal sets. Bespoke methods for specific classes of graphs or expressions, e.g., path-finding, may perform substantially better.

3. Relation to network reliability. Every monotonic expression can be mapped into a weighted, directed graph with two special vertices, S and T, and every solution to the associated SAT problem corresponds to a path on that graph from S to T.[1] The partial order lattice itself

[1] This construction can also be used to map any monotonic network reliability problem into an S-T version, demonstrating that S-T reliability is universal.

is one such graph, where the vertices representing each clause and their descendants are identified and labeled T. The mapping is not unique, but there is a single choice that is arguably the most natural representative, constructed as follows. Beginning with the partial order lattice with vertices identified as above, recursively identify first, all the vertices with the same sets of incoming edges, and then, all the vertices with the same sets of outgoing edges. A detailed algorithm is given in Appendix A. The satisfiability Ξ is the probability that a random walker starting at S will reach T or, equivalently, the probability that a random subgraph of $\mathcal{G}$ constructed by choosing each edge with probability equal to its weight includes a path from S to T. This is the "S-T" or "two-point" network reliability introduced by Moore and Shannon[11]. The paths are the Feynman diagrams for the perturbative approximation to Ξ. It is trivial to compute Ξ on a tree, but loops, even acyclic ones, induce dependencies between different paths that are hard to deal with.

Notice that this construction can handle events that occur in both positive and negative senses as independent events, as long as we require that clauses that include both variables are ignored. The resulting graphs can be split into two subgraphs that intersect only at S and T. Of course, every such event requires another split, so that, if there are k such events, there will be 2^k separate subgraphs. The methods described here are appropriate when the expression $\mathcal{E}$ is *mostly* monotonic, i.e., when $k \ll N$. In principle, it is possible that more complicated constraints on the joint probability of two events could be handled, but that is beyond the scope of this work. The inverse process – constructing a SAT expression from a graph – is exactly identifying minimal solutions.

For a monotonic system, $\Xi(\mathcal{E}, \boldsymbol{x})$ is a monotonic polynomial with integer coefficients mapping the unit interval to itself. Since each variable e_i appears at most once in any clause or path, the reliability in the homogeneous case is a polynomial of degree at most N. In the *thermodynamic limit*, $N \to \infty$, for many systems the partition function exhibits a discontinuity at a critical value of x, indicating a phase transition. For finite N, there can be no discontinuity, but there can be a "shadow" of a discontinuity, i.e., an abrupt, nonlinear change in value over a small range of x, called the transition region. This behavior limits the utility of Taylor expansions for Ξ, as indicated in the center panel of Figure 2.1.

4. Perturbative methods.

As discussed in subsection 2.5, the reliability is not just the sum of total probabilities for each clause, because this overcounts the contribution of many configurations. The Inclusion-Exclusion expansion correctly handles all the contributions, but only at the cost of increasing the number of terms to as many as $2^{|\mathcal{L}|}$. Moreover, the terms form an alternating series with combinatorially large coefficients. Applying the expansion to $\mathcal{E}$ (resp., $\overline{\mathcal{E}}$) produces a power series in x (resp., $1-x$) – i.e., a Taylor series at $x = 0$ (resp., $x = 1$) – for the satisfiability. These are the weak- and strong-coupling expansions of statistical physics. Truncating either expansion at depth D has the effect of truncating the associated Taylor series. The truncated Taylor series do not respect unitarity, monotonicity, or duality, nor do they separately provide either an upper or lower bound on the answer. However, combining the two truncated Taylor series using duality and imposing monotonicity on the result restricts the space of possible solutions and allows us to identify upper and lower bounds.

4.1. Taylor series expansion(s).

We can truncate the Inclusion-Exclusion expansion at any desired depth D to obtain the first $\kappa(D)$ Taylor coefficients exactly, and thus an $O(x^{\kappa(D)})$ approximation, where $\kappa(D)$ is 1 less than the size of the smallest union of $D+1$ sets. $\kappa(D)$ depends on $\mathcal{E}$, as illustrated in section 5:

$$\kappa(D) = \min_{D+1-\text{tuples } t} \left| \bigcup_{i=1}^{D+1} L_{t_i} \right| - 1. \tag{4.1}$$

Minimality of the clauses guarantees that $\kappa(D) \geq D$. In practice, it may be the case that $\kappa(D) \gg D$.

4.2. Bezier polynomial representation. Applying the Inclusion-Exclusion expansion to minimal configurations results in degree-N polynomials in x or $1-x$, i.e.,

$$\Xi(\mathcal{E}, x) = \sum_{k=0}^{N} \alpha_k x^k = \sum_{k=0}^{N} \overline{\alpha}_k (1-x)^k. \tag{4.2}$$

The number of configurations with exactly k events occurring is $\binom{N}{k}$ and the probability of each is $x^k(1-x)^{N-k}$. Hence, Moore and Shannon [11] suggested writing

$$\Xi(\mathcal{E}, x) = \sum_{k=0}^{N} \beta_k \binom{N}{k} x^k (1-x)^{N-k}. \tag{4.3}$$

The transformation from $\boldsymbol{\alpha}$ to $\boldsymbol{\beta}$ is a change of basis in the vector space of polynomials of degree N from the *power* basis to the *Bernstein* basis, whose basis elements are: [4, 10]

$$B(N, k, x) \equiv \binom{N}{k} x^k (1-x)^{N-k}. \tag{4.4}$$

As summarized in the commutative diagram of Figure 4.1, a Taylor series can be thought of as a linear operator Y from $\mathbb{R}^{N+1}$ to polynomials of degree N on $[0, 1]$. A Bezier polynomial is a (different) linear operator Z from $\mathbb{R}^{N+1}$ to polynomials of degree N on $[0, 1]$. The representation in the Bernstein basis has many useful, well-known properties. Here we will make use of the following:

- Bernstein basis functions $B(N, k, x)$ are strongly localized around the point k/N. Hence they are kernel density estimators for functions on the unit interval.
- The Bernstein basis functions are invariant under the simultaneous operations $x \leftrightarrow 1-x$ and $k \leftrightarrow N-k$. Hence, if $\boldsymbol{S_\beta}$ is a reflection, i.e.,

  $$\boldsymbol{S_\beta}(\beta_0, \beta_1, \ldots, \beta_{N-1}, \beta_N) \equiv (\beta_N, \beta_{N-1}, \ldots, \beta_1, \beta_0), \tag{4.5}$$

  then

  $$Z(\boldsymbol{\beta}, x) = Z(\boldsymbol{S_\beta}(\boldsymbol{\beta}, 1-x)). \tag{4.6}$$

- The transformation from the power basis to the Bernstein basis is accomplished using a matrix closely related to Pascal's triangle. Specifically, if

  $$\boldsymbol{\beta} = \boldsymbol{T^{(N)}} \boldsymbol{\alpha}, \quad \text{where } T^{(N)}_{k,j} \equiv \binom{k}{j} \Big/ \binom{N}{j} = \binom{N-j}{k-j} \Big/ \binom{N}{k}, \tag{4.7}$$

  then $Z(\boldsymbol{\beta}, x) = Y(\boldsymbol{\alpha}, x)$.

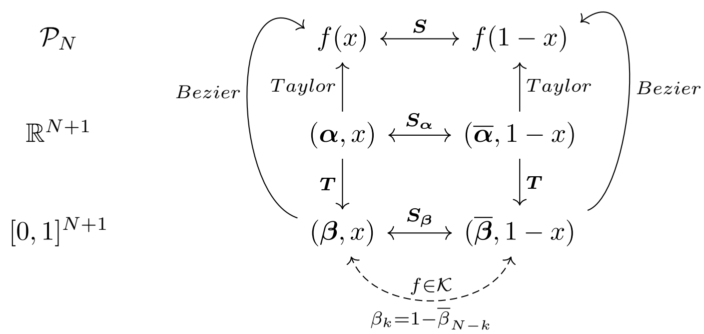

FIG. 4.1. *A commutative diagram illustrating relationships among: (1) $\boldsymbol{\alpha}$, Taylor coefficients at $x = 0$; (2) $\overline{\boldsymbol{\alpha}}$, Taylor coefficients at $x = 1$; (3) $\boldsymbol{\beta}$ and $\overline{\boldsymbol{\beta}}$, Bezier coefficients; and (4) the reflection S which maps $x \leftrightarrow (1 - x)$. The function $f(x)$ is a Taylor series in x constructed from the coefficients $\boldsymbol{\alpha}$ or, equivalently, a Bezier polynomial in x constructed from the coefficients $\boldsymbol{\beta}$. When f is skew-symmetric about $x = 1/2$, e.g., $f(x) = 1 - f(1 - x)$, then $\beta_k = 1 - \overline{\beta}_{N-k}$, as indicated by the dashed curve. The resulting redundance in the Bezier coefficients can be used to bound the approximation error when not all Taylor coefficients are known.*

- The curve $(x, Z(\boldsymbol{\beta}, x))$ is contained within the convex hull of the points $(k/N, \beta_k)$. Hence if $\beta_0 = 0$, $\beta_N = 1$, and $0 \le \beta_k \le 1$ for all other k, then $Z(\boldsymbol{\beta}, x)$ is also in the interval $[0, 1]$. The analogous constraint on Taylor coefficients is not simple – in fact it is most easily derived by way of $\boldsymbol{T}^{-1}$.
- If the coefficients β_k are monotonic in k, then $Z(\boldsymbol{\beta}, x)$ is also monotonic. Once again, the analogous constraint on Taylor coefficients is not obvious. Although the converse is not necessarily true in general, the semantics of the satisfiability's coefficients places a monotonicity constraint on them as shown in subsection 4.3.
- The derivative of a Bernstein polynomial at $x = 0$ or 1 is

$$\frac{d^m}{dx^m} B(N, k, x)\Big|_0 = \begin{cases} \frac{N!}{(N-m)!}(-1)^{m-k}, & k \le m \le N \\ 0 & \text{else} \end{cases} \tag{4.8}$$

- de Casteljau's algorithm for evaluating Bezier polynomials is numerically stable and does not require explicitly constructing the binomial coefficients.

This basis was introduced by Bernstein to prove the Weierstrass approximation theorem [4] by constructing the convergent single-parameter family of approximations

$$f^{(N)} \equiv \sum_{k=0}^{N} f\,(k/N)\, B(N, k, x). \tag{4.9}$$

Members of this family of approximants are referred to as *Bernstein polynomials* for f. Here we instead construct an approximation whose first m (resp., $\overline{m}$) derivatives at $x = 0$ (resp., $x = 1$) match those of f. To emphasize the difference, we call these *Bezier polynomials*[8]. They are also sometimes referred to as "polynomials in Bernstein form".

The symmetry of Bernstein basis functions shown in (4.6) makes them ideal for representing the duality in PSAT problems. Notice that if $\boldsymbol{\alpha}$ and $\overline{\boldsymbol{\alpha}}$ are, respectively, the Taylor coefficients of

$f(x)$ and $1-f(x)$ at $x=0$ and 1, then the symmetry $f(x)=1-f(1-x)$ allows us to write

$$Y(\boldsymbol{\alpha},x) = 1 - Y(\boldsymbol{\alpha},1-x) = Y(\overline{\boldsymbol{\alpha}},1-x) \tag{4.10a}$$

$$Z(\boldsymbol{T\alpha},x) = Z(\overline{\boldsymbol{T}}\overline{\boldsymbol{\alpha}},1-x) \tag{4.10b}$$

$$Z(\boldsymbol{\beta},x) = Z(\overline{\boldsymbol{\beta}},1-x) = Z(\boldsymbol{S}\overline{\boldsymbol{\beta}},x). \tag{4.10c}$$

That is $\boldsymbol{\beta} = \boldsymbol{S}\overline{\boldsymbol{\beta}}$ or

$$\beta_k = \overline{\beta}_{N-k}, \tag{4.11}$$

as illustrated in the commutative diagram.

4.3. Bounds and Interpolation. The transformation matrix $\boldsymbol{T}$ is lower triangular. Hence β_k is completely determined by the values of α_j for $j \le k$. From the first κ coefficients of the Taylor series at $x=0$, we obtain the *first* κ coefficients of the Bernstein representation; from the first $\overline{\kappa}$ coefficients of the Taylor series at $x=1$, we obtain the *last* $\overline{\kappa}$ Bernstein coefficients using Equation (4.10c). When $\kappa+\overline{\kappa}<N$, the coefficients $\beta_\kappa,\ldots,\beta_{N-\overline{\kappa}}$ are undetermined. Nonetheless, monotonicity allows us to place tight upper and lower bounds on β_{k+1} given β_k. Although the convex hull property of Bezier polynomials ensures that monotonically increasing coefficients produce a monotonically increasing polynomial, the converse is not necessarily true. The argument for the converse in this case provides insight into how $\boldsymbol{\beta}$ characterizes scale-dependent structure.

The number of solutions to the satisfiability problem in which exactly k variables are true – i.e., the number of struts in level k of the partial order lattice – is given by $n_k \equiv \beta_k\binom{N}{k}$ (which is thus the "density of states" function of statistical mechanics). By monotonicity, each solution of size k generates $N-k$ solutions of size $k+1$. Of course, these solutions are not necessarily distinct. Indeed, the coefficient β_k encodes not only the number of minimal solutions of size k, but also how all the smaller minimal solutions overlap. Since each vertex in level $k+1$ of the partial order lattice has only $k+1$ incoming edges,

$$\beta_{k+1}\binom{N}{k+1} = n_{k+1} \ge n_k\frac{N-k}{k+1} = \beta_k\binom{N}{k+1} \Longleftrightarrow \beta_{k+1}\ge\beta_k. \tag{4.12}$$

On one hand, this bound is tight in the sense that there is an expression $\mathcal{E}'$ for which $\beta'_{k+1}=\beta'_k=\beta_k$; on the other hand, it does not take advantage of all the known information about how minimal solutions overlap that is contained in $\{\beta_0,\ldots,\beta_{k-1}\}$.

An upper bound can be obtained from the lower bound of the dual problem. Together, these constrain $\beta_\kappa \le \beta_k \le \beta_{N-\overline{\kappa}}$ for $\kappa<k<N-\overline{\kappa}$. These bounds are shown in the center panel of Figure 2.1.

There are several intuitively appealing interpolants between the bounds, including:

1. linear interpolation, $\widehat{\beta}_k = \beta_\kappa + \frac{k-\kappa}{N-\overline{\kappa}-\kappa}(\beta_{N-\overline{\kappa}}-\beta_\kappa)$;
2. logarithmic interpolation, $\ln\widehat{\beta}_k = \ln\beta_\kappa + (\frac{k-\kappa}{N-\overline{\kappa}-\kappa})(\ln\beta_{N-\overline{\kappa}}-\ln\beta_\kappa)$;
3. the expected value of n_k under an assumption that the graph is "structureless" at these scales – for example, lacking any minimal solutions of sizes $\kappa<k<N-\overline{\kappa}$, and whose minimal solutions outside that range overlap randomly. See Appendix B for more details.

4.4. Hybrid estimation. The upper and lower bounds developed in subsection 2.7 define feasible regions that are narrowest near $x=0$ and 1. Details of the region near $0<x<1$ are determined by Bezier coefficients β_k for $k\sim xN$. Estimates of β_k from any source can dramatically narrow the uncertainty in Ξ, although estimates alone do not change the bounds. One such estimate can be provided by Monte Carlo simulation. By definition, a fraction β_k of the subsets of exactly k events satisfies $\mathcal{E}$. Hence, we can estimate β_k for any k to any desired confidence by evaluating $\mathcal{E}$ on a random sample of subsets of events. Another estimate, for the special point $\widetilde{x} = 1/2$ where each event is equally likely to occur or not, is provided by $2^{-N}\mathcal{N}$, where $\mathcal{N}$ is the number of solutions of the corresponding deterministic satisfiability problem.

4.5. Sensitivity analysis. The satisfiability $\Xi(\mathcal{E}, \boldsymbol{x})$ is a multi-affine function of the event probabilities, since multiple appearances of the same event within a single conjunction can be reduced to a single occurrence. That is, it can be written as $\prod_{i=1}^{N}(a_i + b_i\widetilde{x}_i)$. Maximizing or minimizing the satisfiability is thus, in principle, not difficult once the coefficients have been determined. However, a common problem is to optimize satisfiability under correlated constraints on the event probabilities, such as constraining them to lie in a subspace of $\Re^N$.

Restricted to a subspace of dimension M, the satisfiability can be thought of as a smooth (because it is a multinomial) M-dimensional manifold. An approximation allows us to apply standard tools such as sensitivity analysis or differential geometry to this manifold. For example, suppose the probability of each event depends linearly on a finite resource such as energy, bandwidth, vaccine, or human time and effort. Re-allocating resources from one set of events to another can be modeled as a perturbation in the corresponding direction. The partial derivatives $\frac{\partial}{\partial x_i}\Xi(\mathcal{E}, \boldsymbol{x})$ are a differential form of the leave-one-out or Birnbaum importance [2] of event E_i. In the graphical representation of the problem, they define a notion of graph derivatives.

5. Homogeneous Example. The following simple, analytically tractable example illustrates an important point: the relative contribution of different events to the overall satisfiability depends on the probability of the individual events, *even in a homogeneous system*. Concretely, think of two strains of an infectious disease with different transmissibilities spreading over a human contact network in which all contacts are identical. Obviously, the more transmissible strain is more likely to infect any given person; less obviously, the particular contacts whose removal most reduces the probability of infecting that person may differ.

Consider the set of $N = 7$ events and the expression

$$\mathcal{E} = (e_1 \wedge e_2 \wedge e_3) \vee (e_1 \wedge e_4 \wedge e_5) \vee (e_6 \wedge e_7). \tag{5.1}$$

The clauses (and hence the minimal struts) are defined by the sets $L_1 = \{1, 2, 3\}$; $L_2 = \{1, 4, 5\}$; and $L_3 = \{6, 7\}$. An equivalent network reliability problem is the probability of reaching T from S on the graph in the left panel of Figure 5.1, constructed using the algorithm in Appendix A. The three minimal struts are the simple paths from S to T in this graph; the 10 minimal cuts are easily seen to be those given in Table 5.1, which are (S, T) cut sets for this graph. The right panel of Figure 5.1 shows a graph for the dual expression. The minimal struts and cuts for the graph in the left panel are the minimal cuts and struts, respectively, for the graph in the right panel.

TABLE 5.1
The ten minimal cuts for the example expression in Equation (5.1).

$\{1, 6\}$	$\{2, 4, 6\}$	$\{2, 5, 6\}$	$\{3, 4, 6\}$	$\{3, 5, 6\}$
$\{1, 7\}$	$\{2, 4, 7\}$	$\{2, 5, 7\}$	$\{3, 4, 7\}$	$\{3, 5, 7\}$

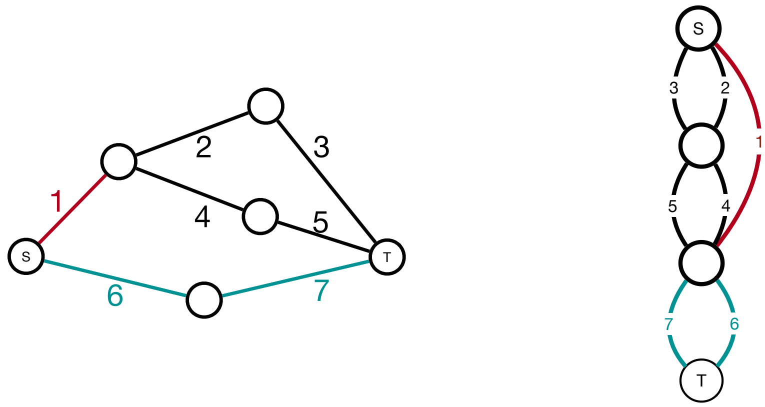

FIG. 5.1. *Graphical representations of the expression $\mathcal{E}$ in Equation (5.1) (left panel) and its dual (right panel). Edges with the same contribution to the satisfiability, as evident by symmetry, are colored the same.*

5.1. Satisfiability. Consider the first clause, $e_1 \wedge e_2 \wedge e_3$. The probability of the minimal strut corresponding to this clause is $p(\{1,2,3\}) = x^3(1-x)^4$; its total probability (including all its descendants) is $p_t(\{1,2,3\}) = x^3$.

Because there are only three minimal struts, the Inclusion-Exclusion expansion for $\Xi(\mathcal{E}, x)$ contains only $2^3 - 1$ terms and can easily be written down by inspection:

$$\begin{aligned}\Xi(\mathcal{E}) = {} & p(e_1 \wedge e_2 \wedge e_3) + p(e_1 \wedge e_4 \wedge e_5) + p(e_6 \wedge e_7) \\ & - p(e_1 \wedge e_2 \wedge e_3 \wedge e_4 \wedge e_5) - p(e_1 \wedge e_2 \wedge e_3 \wedge e_6 \wedge e_7) - p(e_1 \wedge e_4 \wedge e_5 \wedge e_6 \wedge e_7) \\ & + p(e_1 \wedge e_2 \wedge e_3 \wedge e_4 \wedge e_5 \wedge e_6 \wedge e_7).\end{aligned}$$

Substituting in the total probability for each term yields the power series in x (i.e., the Taylor series at $x = 0$):

$$\Xi(\mathcal{E}, x) = x^2 + 2x^3 - 3x^5 + x^7. \tag{5.2}$$

There are $2^{10} - 1 = 1023$ terms in the Inclusion-Exclusion expansion in terms of minimal cuts, so it is more difficult to write down,[2] but it yields the following power series in $y = 1 - x$ (i.e., the Taylor series at $x = 1$):

$$\Xi(\mathcal{E}, 1-x) = 1 - \overline{\Xi}(\mathcal{E}, y) = 1 - 2y^2 - 7y^3 + 20y^4 - 18y^5 + 7y^6 - y^7 \tag{5.3}$$

Equivalently,

$$\boldsymbol{\alpha} = (0, 0, 1, 2, 0, -3, 0, 1) \tag{5.4a}$$

$$\overline{\boldsymbol{\alpha}} = (1, 0, -2, -7, 20, -18, 7, -1). \tag{5.4b}$$

Using Equation (4.7), we find for this expression

$$\boldsymbol{\beta} = (0, 0, 1/21, 1/5, 18/35, 19/21, 1, 1) \tag{5.5a}$$

$$\overline{\boldsymbol{\beta}} = (1, 1, 19/21, 18/35, 1/5, 1/21, 0, 0). \tag{5.5b}$$

[2]Obviously, since this is more than the number of possible distinct configurations, the expansion can be greatly simplified.

Notice that, although the Taylor coefficients are not related in any obvious way, the Bernstein coefficients satisfy Equation (4.6).

TABLE 5.2
Degree of largest exact term in an Inclusion-Exclusion expansion of the example expression (5.1) *truncated at depth D.*

$\kappa(d)$	4	6	7							
d	1	2	3	4	5	6	7	8	9	10
$\overline{\kappa}(d)$	2	4	4	5	5	5	5	6	6	7

The functions $\kappa(d)$ and $\overline{\kappa}(d)$ defined in (4.1) for this expression are tabulated in Table 5.2. Now suppose we truncate the Inclusion-Exclusion expansions at depth $D = 1$, i.e., not even considering pairs of minimal struts or cuts. The resulting Taylor expansion at $x = 0$ is exact up to and including terms of order $x^{\kappa(1)} = x^4$, while the Taylor expansion at $x = 1$ is exact up to and including terms of order $(1-x)^{\overline{\kappa}(1)} = (1-x)^2$. These estimates are shown in the center panel of Figure 2.1. Although each approximation tracks the exact function poorly through the transition region, the Bernstein representation produced by combining the two Taylor expansions determines the value of every coefficient − and thus, the function itself − exactly, since $\kappa(1) + \overline{\kappa}(1) = N - 1$. Suppose, for the sake of argument, that $\kappa(1)$ were smaller, say 3. Monotonicity and linear interpolation between known coefficients define three possible values for β_4: a lower bound of $1/5$, an upper bound of $19/21$, and an interpolated value of $58/105$. The resulting bounds and interpolation are shown in the right panel of Figure 2.1, along with the exact result.

6. Heterogeneous Systems. Suppose we are given a problem instance, i.e., an expression $\mathcal{E}$ with specified event probabilities $\widetilde{\boldsymbol{x}}$. Parameterized forms of the probabilities are often, but not always, part of the problem specification. For example, the events $\boldsymbol{E}$ may be generated by a Poisson process operating for a time τ or, equivalently, by interactions with a coupling constant τ. The transmission of infectious disease from one host to another is often modeled as a Poisson process whose probability depends on the overall transmissibility of the pathogen, τ, and the duration of contact between the hosts, ρ. More generally, any dynamical system whose configurations are probabilistically distributed (in time or across ensembles of identically prepared systems) as an exponential of a property of the configuration can be thought of as a collection of Poisson events with heterogeneous rates.[3] Such systems include statistical mechanical systems governed by Boltzmann distributions and field-theoretical systems governed by a least-action principle.

Applying the duality symmetry to the weak- and strong-coupling perturbation series relied on extending a particular problem instance into a one-parameter family whose solution smoothly interpolated between 0 and 1. This parameter completes the transition from satisfiability as a binary function of binary variables to a continuous function of a continuous variable. For the heterogeneous problem, we proceed analogously, first constructing a mapping to a two-parameter representation in which the two degrees of freedom contained in $\widetilde{x}_1$ and $\widetilde{x}_N$ are replaced with a "location" λ for a homogeneous case and an envelope ε capturing the extent of heterogeneity. This mapping is convenient for distinguishing the effects of stronger interactions from those of heterogeneity through standard sensitivity analyses. Then, in order to apply perturbative methods, we consider curves in this two-dimensional space along which we can ensure the satisfiability is monotonic.

[3]This does not imply that the dynamics of such systems are Poisson processes, only that steady-state or equilibrium distributions can be modeled using them.

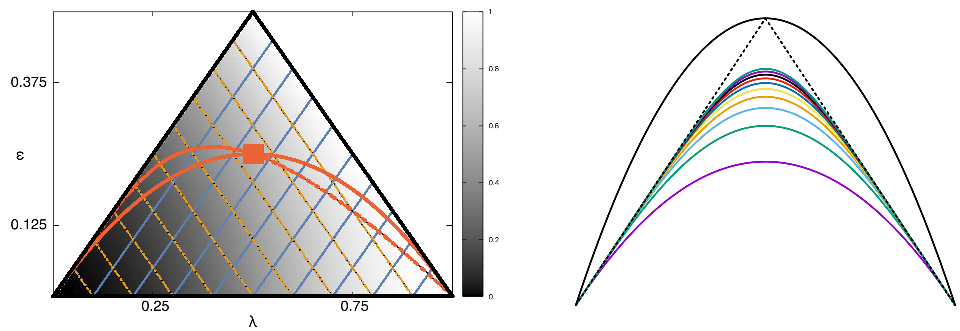

FIG. 6.1. *(Left) The coordinate mapping from* $\widetilde{x}_1, \widetilde{x}_N$ *to* λ, ε *defined by* (6.1)*. Lines of constant* $\widetilde{x}_1$ *are mapped into blue solid curves; lines of constant* $\widetilde{x}_N$ *are mapped into yellow dashed curves. The background is shaded by the satisfiability for the example expression when* $\widetilde{x}_B$ *is the mean of* $\widetilde{x}_A$ *and* $\widetilde{x}_C$*. The red curve is the parametric curve* $(\lambda(t), \varepsilon(t))$ *defined under the polynomial decomposition of* subsection 6.1 *for the case* $\widetilde{x}_1 = 1/4$, $\widetilde{x}_N = 3/4$*. The square on the curve is the location of the example case. Notice that if* ε *were any larger, the parameterized curve would violate the unitarity constraints* $\varepsilon \le \min(\lambda, 1-\lambda)$*. The dashed red curve is the parametric curve defined under the logarithmic decomposition of* subsection 6.2 *for this case. (Right) The first 10 Bezier polynomial approximants defined in* (4.9) *to the unit triangle are shown inscribed in the triangle. The quadratic that passes through the points* $(0,0)$*,* $(1/2, 1)$*, and* $(1,0)$ *is shown in black for comparison.*

Generally, the probabilities $\widetilde{x}_i$ are not unrelated, even when they are not identical. That is, we can consider them all as functions $x_i(t)$, of a single parameter $t \in [0,1]$ with the special values $x_i(0) = 0$, $x_i(1) = 1$, and $x_i(t^*) = \widetilde{x}_i$, for some t^* independent of i. Colloquially, we may say that E_1 is "twice as likely as" E_2, for example. Such a statement does not specify a unitary parameterization over the whole domain, since it cannot be true for $x_2 > 1/2$.

If the parameterization is not specified, we are free to create an arbitrary one. In subsection 6.1, we describe a parameterization that is applicable in the absence of a parametric form for $\widetilde{x}_i$; in subsection 6.2, a different parameterization for a specific, important case when the events are generated by Poisson processes with different rates.

6.1. Polynomial decomposition. A useful polynomial decomposition is illustrated in Figure 6.1:

$$x_i = \lambda + w_i\varepsilon \text{ and } w_i = (\widetilde{x}_i - \widetilde{\lambda})/\widetilde{\varepsilon}, \text{ where} \tag{6.1a}$$

$$\widetilde{\lambda} \equiv (\widetilde{x}_1 + \widetilde{x}_N)/2 \text{ and } \widetilde{\varepsilon} \equiv (\widetilde{x}_N - \widetilde{x}_1)/2. \tag{6.1b}$$

This mapping results in a particularly simple transformation between the original problem and its dual, for which $\overline{\widetilde{x}} = 1 - \widetilde{x}$ and the ordering of events is reversed (i.e., $\overline{\widetilde{x}}_N = 1 - \widetilde{x}_1$):

$$\overline{\lambda} = 1 - \lambda; \quad \overline{\varepsilon} = \varepsilon; \quad \overline{w}_i = -w_{N+1-i}. \tag{6.2}$$

Repeating the analysis of subsection 4.2 shows that the satisfiability can be written as a finite Taylor series in two variables analogous to (4.2):

$$\Xi(\mathcal{E}, \lambda, \varepsilon, \boldsymbol{w}) = \sum_{k=0}^{N} \sum_{\ell=0}^{N-k} \alpha_{k,\ell} \lambda^{\ell} \varepsilon^{k} \tag{6.3a}$$

or, in a mixed Bernstein-Taylor basis:

$$\Xi(\mathcal{E},\lambda,\varepsilon,\boldsymbol{w}) = \sum_{k=0}^{N}\sum_{\ell=0}^{N-k}\gamma_{k,\ell}(\boldsymbol{w})B(N-k,\ell,\lambda)\varepsilon^k. \tag{6.3b}$$

Furthermore, the gradient of the satisfiability at any point in the $\lambda-\varepsilon$ plane is easily computed in terms of $\boldsymbol{\gamma}$:

$$\frac{\partial\Xi(\mathcal{E},\lambda,\varepsilon)}{\partial\lambda} = \sum_{k=0}^{N}\sum_{\ell=0}^{N-k-1}(\gamma_{k+1,\ell}-\gamma_{k,\ell})B(N-k-\ell-1,\ell,\lambda) \tag{6.4a}$$

$$\frac{\partial\Xi(\mathcal{E},\lambda,\varepsilon)}{\partial\varepsilon} = \sum_{k=1}^{N}\sum_{\ell=0}^{N-k-1}k\gamma_{k,\ell}B(N-k-\ell,\ell,\lambda)\varepsilon^{k-1}. \tag{6.4b}$$

Duality provides a generalization of the reflection symmetry previously obtained for homogeneous systems (4.11):

$$\overline{\gamma}_{k,\ell}(-\boldsymbol{w}) = \gamma_{k,N-k-\ell}(\boldsymbol{w}) \tag{6.5}$$

Hence, order by order in ε, we can use methods of subsection 2.6 to evaluate both the lowest and highest order coefficients in λ. Interpolation is not as easy, because Ξ is not necessarily monotonic order by order in ε. Instead, we introduce parametric curves in the $\lambda-\varepsilon$ plane. As long as $x_i(t)$ is monotonic, the satisfiability must also be monotonic in t, so we can apply the interpolation scheme developed for the homogeneous case in subsection 4.3 to $\Xi(\mathcal{E},t)$ to develop bounds on the satisfiability that are exact up to some order in t.

To simplify the perturbative analysis, we choose $\lambda(t)=t$, which implies $t^*=\lambda(t^*)=\widetilde{\lambda}$, and normalize the weights $w_1=-w_N=1$, which implies $\varepsilon(t^*)=\widetilde{\varepsilon}$. Then we design a function $\varepsilon(t)$ that

1. satisfies unitarity: $\forall t\in[0,1]$, $0\le\varepsilon(t)\le\min(t,1-t)\le 1/2$;
2. remains invariant under duality: $t\leftrightarrow 1-t$;
3. maintains the semantics of λ and ε as the midpoint and half-width of $[x_1(t),x_N(t)]$;

Along the parameterized curve, (6.3a) becomes

$$\Xi(\mathcal{E},t,\boldsymbol{w}) = \sum_{k=0}^{N}\sum_{\ell=0}^{N-k}\alpha_{k,\ell}\lambda(t)^{\ell}\,\varepsilon(t)^k. \tag{6.6}$$

Because of the first condition above, $\varepsilon(t)$ vanishes at the points $t=0$ and 1 where the exact solution to the homogeneous problem is known, and the Taylor expansion can be readily determined at both points.

An appealing choice for $\varepsilon(t)$ is a low-degree polynomial. When $\varepsilon(t)$ is a degree-m polynomial, the multinomial Taylor series reduces to a degree-mN polynomial in t. However, it may require extremely large m to maintain unitarity while simultaneously representing the full range of probabilities $\widetilde{x}_N-\widetilde{x}_1$. For example, as illustrated in the left panel of Figure 6.1, a quadratic $\varepsilon(t)$ meeting these constraints cannot represent problems in which $\widetilde{x}_N-\widetilde{x}_1>1/2$. The Bernstein approximants defined in (4.9) for the unit triangular function, $T(t)=1-2|t-1/2|$, provide a convenient way of describing the lowest-order polynomial that can represent the full range of probabilities. That is,

$$\varepsilon_{2n}(t) = \sum_{k=0}^{2n}T(k/n)B(n,k,t) = 1-\frac{1}{n}\sum_{j=-n}^{n}|j|\left[B(2n,n+j,t)+B(2n,n-j,t)\right]. \tag{6.7}$$

where n is the smallest integer such that $\widetilde{x}_N - \widetilde{x}_1 \leq 2\varepsilon(1/2) = 1 - 2^{1-2n}$. For $n = 1$, this reduces to $\varepsilon_2(t) = 2t(1-t)$, and, as noted above, $\widetilde{x}_N - \widetilde{x}_1$ must be less than or equal to $1/2$. The functions ε_{2n} for $n \leq 10$ are illustrated in the right panel of Figure 6.1. Substituting $x_i(t) = t + w_i\varepsilon_2(t)$ into the Taylor series (6.3a) gives

$$\Xi(\mathcal{E}, t) = \sum_{k=0}^{3N} \alpha'_k t^k, \text{ where } \alpha'_k = \sum_{u=0}^{N} \sum_{r=0}^{N-u} 2^r (-1)^{k-u-r} \binom{k}{k-u-r} \alpha_{r,u}. \tag{6.8}$$

6.2. Logarithmic decomposition. When parametric forms of the probabilities $\boldsymbol{x}$ are specified, it is more natural to use them than to pick an arbitrary parameterization as in subsection 6.1. Here we develop a decomposition and related parameterization appropriate to the important case of independent Poisson processes with heterogeneous rates, i.e., $x_i(\tau) = 1 - e^{-\rho_i\tau}$. In this case, it is more useful to perform the decomposition in a logarithmic space, using the geometric, instead of the arithmetic, mean. Choosing $\tau^* = 1$, the analogues of (6.1) become:

$$\widetilde{x}_i(\tau) = 1 - e^{-(\rho + w_i\Delta)\tau} \text{ and } w_i = -\left[\ln(1 - \widetilde{x}_i) - \rho\right]/\Delta, \text{ where} \tag{6.9a}$$

$$\rho \equiv (\rho_1 + \rho_N)/2 = -\ln\sqrt{(1-\widetilde{x}_1)(1-\widetilde{x}_N)} \text{ and } \Delta \equiv (\rho_1 - \rho_N)/2 = \ln\sqrt{\frac{1-\widetilde{x}_N}{1-\widetilde{x}_1}}. \tag{6.9b}$$

When the rates ρ_i are all rationally related – as is common – the exponent can be written as the ratio $\rho_i/\rho_N = m_i/n$ where n is the greatest common divisor of ρ_i and m_i is an integer greater than or equal to 1. Then rescaling the parameter gives

$$x_i(t) = 1 - t^{m_i}, \text{ where } t(\tau) = e^{-\rho_1\tau/n}. \tag{6.10}$$

The satisfiability can thus be written as a degree-M polynomial in t, where $M = \sum_i m_i$.

7. Heterogeneous example. Consider the same example as in section 5, except with heterogeneous probabilities $\widetilde{\boldsymbol{x}}$.[4] Suppose events in each equivalence class have the same probability:

$$\widetilde{x}_i = \begin{cases} x_A = 1 - y_A & i = 1 \\ x_B = 1 - y_B & 2 \leq i \leq 5 \\ x_C = 1 - y_C & 6 \leq i \leq 7. \end{cases} \tag{7.1}$$

Then a straightforward analysis of the three minimal struts gives

$$\Xi(\mathcal{E}, \boldsymbol{x}) = x_C^2 + (1 - x_C^2)x_A x_B^2(2 - x_B^2) \tag{7.2}$$

and, from the 10 minimal cuts:

$$\Xi(\mathcal{E}, \boldsymbol{x}) = 1 - y_C(2 - y_C)\left[y_A + (1 - y_A)y_B^2(2 - y_B)^2\right] \tag{7.3}$$

In this simple example, these polynomials in three variables are amenable to direct analysis. Even here, though, if all the probabilities $\widetilde{x}_i$ were distinct, Ξ and $\overline{\Xi}$ would be multinomials in seven dimensions, with potentially as many as 128 terms.

[4] Code for reproducing these results in Mathematica is available at ....

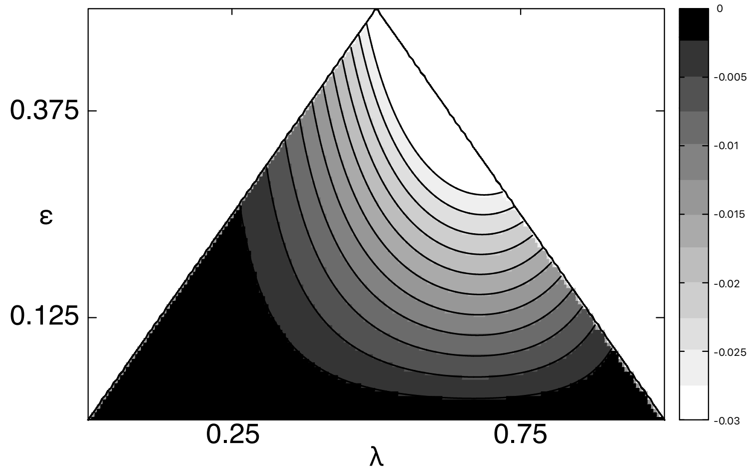

FIG. 7.1. *Approximation error in the satisfiability* Ξ *for the heterogeneous example of* section 7 *with* $(\widetilde{x}_A, \widetilde{x}_B, \widetilde{x}_C) = 1/4, 3/8, 1/2$*, using the decomposition described in* subsection 6.1 *and a linear interpolation between the Bezier coefficients determined by a depth-2 expansion at* $\lambda = 0$ *and a depth-1 expansion at* $\lambda = 1$.

As discussed in subsection 4.5, the derivative of $\Xi(\widetilde{x}_B + 2\alpha, \widetilde{x}_B, \widetilde{x}_B - \alpha)$ evaluated at $\alpha = 0$ gives the effects of redistributing resources among the interactions that enable the single event in class A and the two events in class C. For this example, we find

$$\left.\frac{d\Xi(\widetilde{x}_B + \alpha, \widetilde{x}_B, \widetilde{x}_B - 2\alpha)}{d\alpha}\right|_{\alpha=0} = -2\widetilde{x}_B\left(\widetilde{x}_B^3 - 2\widetilde{x}_B + 1\right). \tag{7.4}$$

This function changes sign at $\widetilde{x}_B \approx 0.61$, signifying that the relative importance of events in class A and C depends on the probability of the events in class B.

7.1. Transforming to $\lambda - \varepsilon$ space. Expressions for the elements of the matrix γ, defined in (6.3b), in terms of weights a, b, and c defined in (6.1) for the corresponding probabilities $\widetilde{x}_A$,$\widetilde{x}_B$, and $\widetilde{x}_C$ are given in Table 8.1. The first row of $\boldsymbol{\gamma}$ comprises the elements of order $O(\varepsilon^0)$. It is thus independent of a, b, and c, and reduces to the Bernstein coefficients for the homogeneous case given in (5.5a). The terms given exactly by a depth 2 Inclusion-Exclusion expansion at $\lambda = 0$ and a depth-1 expansion at $\lambda = 1$ are indicated by the shading in Table 8.1 and in the table below. For ease of explication, here we consider only cases with $\widetilde{x}_B = (\widetilde{x}_A + \widetilde{x}_C)/2$, i.e. $a = -c = \pm 1$ and $b = 0$. In this restricted example, the Bernstein transformations of the expressions in Table 8.1

reduce to:

$$\boldsymbol{\gamma} = \begin{pmatrix} 0 & 0 & \frac{1}{21} & \frac{1}{5} & \frac{18}{35} & \frac{19}{21} & 1 & 1 \\ 0 & 0 & -\frac{2}{15} & -\frac{2}{5} & -\frac{17}{30} & -\frac{1}{6} & 0 & 0 \\ 0 & \frac{1}{8} & \frac{1}{4} & \frac{9}{40} & -\frac{1}{10} & 0 & 0 & 0 \\ -\frac{1}{16} & -\frac{1}{16} & -\frac{1}{96} & \frac{3}{32} & 0 & 0 & 0 & 0 \\ 0 & -\frac{3}{256} & -\frac{3}{128} & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{512} & \frac{1}{512} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{7.5}$$

The shading shows the coefficients that can be determined exactly using the truncated Taylor series. Examining the third row, for example, illustrates that the rows of $\boldsymbol{\gamma}$ are not necessarily monotonic, and thus interpolations and bounds developed in subsection 4.3 are not immediately applicable. Plots of Ξ along the two parametric curves described in subsection 6.1 and subsection 6.2 along with the bounds and interpolation available along the curve are shown in Figure 7.2 for the first case. Of course, the heterogeneous case must also lie between the homogeneous cases for $x(t) = x_A(t)$ and $x_C(t)$, and these bounds are also shown in the figure.

Table 8.2 displays results for several specific choices of $\widetilde{\boldsymbol{x}}$, considering only the depth-2 Inclusion-Exclusion expansion for minimal struts and the depth-1 expansion for minimal cuts. These specific cases were chosen to illustrate the effects of scaling and translating the probabilities or replacing them with their complements (i.e., transforming to the dual problem), and to compare the polynomial and logarithmic decompositions.

8. Discussion. We have described a novel approximation technique for mostly monotonic probabilistic satisfiability problems with event-specific probabilities. Its first stage computes the weak- and strong-coupling perturbative expansions of statistical physics – which are seen to be Inclusion-Exclusion expansions – in terms of certain problem-specific minimal sets. There may be up to $m = 2^N$ of these sets; if they are not given explicitly, they may be sampled in at worst $O(N^2 \log N)$ steps for each of S samples. The expansions can be truncated at an arbitrary depth $D \leq N$, producing $\sum_{i=1}^{D} \binom{M}{i}$ terms at a cost of order S^D. The truncated expansions separately have undesirable properties. In particular, they do not respect the important constraints of monotonicity and unitarity. Imposing unitarity allows them to be combined into a Bezier polynomial that provides a good approximation – in absolute, if not necessarily relative, terms – across the entire domain. Imposing monotonicity allows the construction of problem-specific, tight upper and lower bounds on the approximation error due to truncation. The bounds are tight in the sense that expressions with the same truncated expansions can be constructed whose satisfiabilities saturate both bounds, whereas any estimator that violates the bounds can be shown to be either non-monotonic or non-unitary. It remains to be seen how sampling affects the quality of the bounds.

In this work, we have explored the weak- and strong-coupling perturbation series, which are Taylor series expansions around the distinguished points at $\lambda = 0$ and 1 and at $\varepsilon = 0$. In the heterogeneous case, in addition to these points the point $(\lambda, \varepsilon) = (1/2, 0)$ and the boundary $(\lambda, \varepsilon) = (\lambda, \min(\lambda, 1-\lambda))$ admit interesting simplifications. The satisfiability at $(1/2, 0)$ is simply related to the deterministic solution; along the boundary, the probability of the most or least probable events e_i is 1 or 0, respectively. When $\widetilde{x}_i = 0$, clauses that include e_i can be ignored, because they never

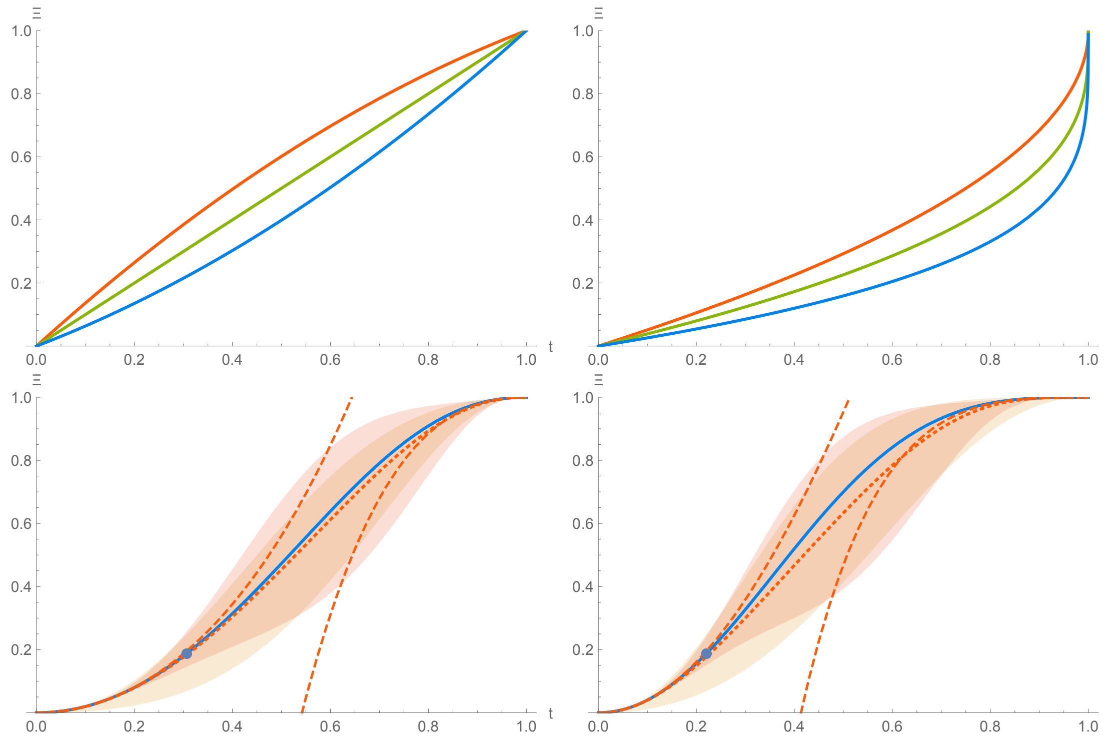

FIG. 7.2. *Results for the heterogeneous example of section 7 with* $(\widetilde{x}_A, \widetilde{x}_B, \widetilde{x}_C) = (1/4, 3/8, 1/2)$ *when no parametric form for* $\widetilde{x}$ *is specified, as in subsection 6.1 (left panels) or when the events result from Poisson processes, as in subsection 6.2 (right panels). (Top) The parameterized probabilities (from bottom) xa, xb, and xc. (Bottom) The exact satisfiability* Ξ *(solid curve) and interpolation (dotted curve) derived here along the parametric curve, along with truncated Taylor expansions for a depth-2 approximation at* $\lambda = 0$ *and depth-1 at* $\lambda = 1$ *(dashed lines). Notice that the Taylor series do not respect unitarity. The lightly shaded region is bounded by the exact homogeneous solutions for* x_a *and* x_c*; the darker shading uses the monotonicity constraint with the approximate solution.*

contribute to the satisfiability; when $\widetilde{x}_i = 1$, the event itself can be ignored, because it is always true. In the dual representation, the same holds true for $\overline{\widetilde{x}}_i = 1 - \widetilde{x}_i$. In principle, it is possible to develop perturbation series in ε around this point and the boundary segments. This leads, in turn, to a step-wise renormalization approach that peels off the most and/or least probable events at each step. Unfortunately, exploring these approaches is beyond the scope of this work.

The approximations and bounds developed here have been applied in the context of network reliability to a range of dynamical systems, including infectious disease transmission over a contact network [12], crop pest movements over a commodity trade network [13], and the Ising model in the presence of an external field [14]. We hope that the synthesis presented here provides both a new perspective on some old problems and a tool that others will find useful in an even greater range of applications.

Appendix A. Transforming a satisfiability problem into S-T reliability on a graph.

Require: $\mathcal{L}$, a set of sets of integers in [1,. . . ,N], i.e., clauses in the expression's DNF form

TABLE 8.1

(Top) The matrix of Taylor coefficients at $(\lambda, \varepsilon) = (0, 0)$, $\boldsymbol{\alpha}$ defined in (6.3a); (bottom) the corresponding matrix $\overline{\boldsymbol{\alpha}}$ at (1,1) for the heterogenous example. Shading in the upper left (respectively, bottom right) indicates coefficients whose values are given exactly by the depth-2 perturbative expansion at $(0, 0)$ (respectively, the depth-1 expansion at $(1, 1)$). Notice that the coefficients along the skew-diagonal match, and that the first row of each matrix, corresponding to $\varepsilon = 0$, agrees with the Taylor series for the homogeneous example, (5.4a) and (5.4b).

$$\begin{pmatrix} 0 & 0 & 1 & 2 & 0 & -3 & 0 & 1 \\ 0 & 2c & 2(a+2b) & 0 & -3a-4(2b+c) & 0 & a+4b+2c & 0 \\ c^2 & 2b(2a+b) & 0 & -2(2b+c)(2a+2b+c) & 0 & 2a(2b+c)+6b^2+8bc+c^2 & 0 & 0 \\ 2ab^2 & 0 & -2\left(a(2b+c)^2+2b\left(b^2+bc+c^2\right)\right) & 0 & a\left(6b^2+8bc+c^2\right)+4b\left(b^2+3bc+c^2\right) & 0 & 0 & 0 \\ 0 & -b\left(4a\left(b^2+bc+c^2\right)+b^3+2bc^2\right) & 0 & b\left(4a\left(b^2+3bc+c^2\right)+b\left(b^2+8bc+6c^2\right)\right) & 0 & 0 & 0 & 0 \\ -ab^2\left(b^2+2c^2\right) & 0 & b^2\left(a\left(b^2+8bc+6c^2\right)+2bc(b+2c)\right) & 0 & 0 & 0 & 0 & 0 \\ 0 & b^3c(2a(b+2c)+bc) & 0 & 0 & 0 & 0 & 0 & 0 \\ ab^4c^2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & -2 & -7 & 20 & -18 & 7 & -1 \\ 0 & -2(a+c) & a-16b-6c & 8(a+6b+3c) & -2(6a+13(2b+c)) & 6(a+4b+2c) & -a-2(2b+c) & 0 \\ -2ac & c(2a+c)-8b^2-16bc & 4\left(2a(2b+c)+9b^2+14bc+c^2\right) & -4\left(4a(2b+c)+13b^2+18bc+2c^2\right) & 5\left(2a(2b+c)+6b^2+8bc+c^2\right) & -2a(2b+c)-6b^2-8bc-c^2 & 0 & 0 \\ c\left(ac-8b^2\right) & 8b\left(a(b+2c)+b^2+5bc+c^2\right) & -4\left(a\left(7b^2+10bc+c^2\right)+b\left(5b^2+17bc+5c^2\right)\right) & 4\left(a\left(6b^2+8bc+c^2\right)+4b\left(b^2+3bc+c^2\right)\right) & -a\left(6b^2+8bc+c^2\right)-4b\left(b^2+3bc+c^2\right) & 0 & 0 & 0 \\ 4b^2c(2a+2b+c) & -2b\left(4a\left(b^2+4bc+c^2\right)+b\left(b^2+12bc+8c^2\right)\right) & 3b\left(4a\left(b^2+3bc+c^2\right)+b\left(b^2+8bc+6c^2\right)\right) & -b\left(4a\left(b^2+3bc+c^2\right)+b\left(b^2+8bc+6c^2\right)\right) & 0 & 0 & 0 & 0 \\ -2b^2c(2a(2b+c)+b(b+2c)) & 2b^2\left(a\left(b^2+8bc+6c^2\right)+2bc(b+2c)\right) & -b^2\left(a\left(b^2+8bc+6c^2\right)+2bc(b+2c)\right) & 0 & 0 & 0 & 0 & 0 \\ b^3c(2a(b+2c)+bc) & -b^3c(2a(b+2c)+bc) & 0 & 0 & 0 & 0 & 0 & 0 \\ -ab^4c^2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

TABLE 8.2

Summary of approximation quality for several instances of $\widetilde{\boldsymbol{x}}$ in the heterogeneous example discussed in section 7. Ξ is the true value of the satisfiability. Estimates are constructed by interpolating between the left and right expansions for missing coefficients in $\boldsymbol{\theta}$. The polynomial decomposition (labeled "P") uses the linear interpolation defined in subsection 4.3; the logarithmic one (labeled "L"), the logarithmic interpolation defined there. The bounds are determined by monotonicity in $\boldsymbol{\theta}$. Results for the polynomial parameterization are not shown in the last column, where $\widetilde{x}_N - \widetilde{x}_1 > 1/2$, and it does not respect unitarity.

$(\widetilde{x}_A, \widetilde{x}_B, \widetilde{x}_C)$	(0.22,0.31,0.39)	$(\frac{1}{4},\frac{3}{8},\frac{1}{2})$	$(\frac{1}{2},\frac{3}{8},\frac{1}{4})$	$(\frac{1}{2},\frac{5}{8},\frac{3}{4})$	$(\frac{1}{2}$,0.65,$\frac{3}{4})$	-
(ρ_A, ρ_B, ρ_C)	$(\frac{1}{4},\frac{3}{8},\frac{1}{2})$	-	-	-	$(\ln 2, \frac{3}{2}\ln 2, \ln 4)$	$(\frac{1}{4},\frac{17}{8},4)$
$\Xi(\mathcal{E}, \widetilde{\boldsymbol{x}})$	0.190	0.299	0.185	0.700	0.707	0.971
estimate P	0.185	0.288	0.190	0.666	0.668	-
upper P	0.258	0.413	0.343	0.878	0.879	-
lower P	0.147	0.213	0.097	0.434	0.438	-
rel. diff. P	0.022	0.037	−0.025	0.049	0.055	-
estimate L	0.199	-	-	-	0.758	1.000
upper L	0.228	-	-	-	0.867	1.000
lower L	0.162	-	-	-	0.366	0.362
rel. diff. L	−0.051	-	-	-	−0.072	−0.029

```
function SAT2GRAPH(ℒ)
    S ← {1, …, N}; T ← ∅; V ← {S, T}; E ← ∅        ▷ Initialize vertex and edge sets, V and E
    PUSH(stack, (S, ℒ))
    while stack is not empty do
        (v, ℓ) ← POP(stack)
        s ← ⋂_{c∈ℓ} c                 ▷ s contains only the events that appear in every clause in ℓ
        for s_i ∈ s do                ▷ Add an incoming edge to v and move to its source
```

```
            v′ ← v \ s_i
            V ← V ∪ v′; E ← E ∪ (v′, v); Labels(v′, v) ← s_i
            v ← v′
            for ℓ_j ∈ ℓ do ℓ_j ← ℓ_j \ s_i end for      ▷ Remove this event from further consideration
        end for
        s ← ∪_{c∈ℓ} c                                  ▷ s contains events that appear in any clause in ℓ
        for s_i ∈ s do                                 ▷ Partition ℓ into clauses that do or don't contain s_i
            ℓ′ ← {c ∈ ℓ | s_i ∈ c}; ℓ ← ℓ \ ℓ′
            if ℓ′ ≠ ∅ then                             ▷ Add an incoming edge to v and move to its source
                v′ ← (∪_{c∈ℓ′} c) \ s_i
                V ← V ∪ v′; E ← E ∪ (v′, v); Labels(v′, v) ← s_i
                for ℓ_j ∈ ℓ′ do ℓ_j ← ℓ_j \ s_i end for ▷ Remove this event from further consideration
                PUSH(stack, (v′, ℓ′))
            end if
        end for
    end while
    return (V, E, Labels)
end function
```

Appendix B. The expected value of n_{k+1} given n_k. Any given configuration at level $k+1$ is connected to $k+1$ configurations at level k, so the probability it is connected to any *specific* configuration is $(k+1)/\binom{N}{k}$. Hence, under the assumption that connections are independent, the probability that a configuration at level $k+1$ is not connected to any of n_k configurations at level k is $\left[1-(k+1)/\binom{N}{k}\right]^{n_k}$, and the expected value of n_{k+1} given n_k is

$$\langle n_{k+1}\rangle = \left[1-\left(1-\frac{k+1}{\binom{N}{k}}\right)^{n_k}\right]\binom{N}{k+1} \tag{B.1}$$

In the limit $\beta_k \ll 1$, this reduces to

$$\langle n_{k+1}\rangle \xrightarrow[\beta_k \ll 1]{} (N-k)n_k, \tag{B.2}$$

as it should. In the example of section 5, the difference between this interpolation and linear interpolation is negligible. In larger systems, with more undetermined coefficients, however, the overlap is significant even for fairly small β_k. Further analysis is beyond the scope of this work.

Acknowledgments. This material is based upon work supported by the following:

- the National Science Foundation under Grant No. CCF-1918656;
- the National Institute of General Medical Sciences of the National Institutes of Health under Award Number R01GM109718;
- the Defense Threat Reduction Agency (DTRA) under Contract No. HDTRA1-19-D-0007;
- the United States Agency for International Development and the generous support of the American people through USAID Leader Award No. AID-OAA-L-15-00001;
- and the Food and Agriculture Cyberinformatics and Tools Initiative Grant No. 2019-67021-29933 from the USDA National Institute of Food and Agriculture.

The content is solely the responsibility of the authors and does not necessarily represent the official views of any of these organizations.

REFERENCES

[1] L. Beaudin, J. Ellis-Monaghan, G. Pangborn, and R. Shrock, *A little statistical mechanics for the graph theorist*, Discrete Mathematics, 310 (2010), pp. 2037–2053.

[2] Z. Birnbaum, *On the importance of different components in a multicomponent system*, in Multivariate Analysis – II, Proceedings of the 2nd International Symposium on Multivariate Analysis, P. Krishnaiah, ed., New York, 1969, Academic Press, pp. 581–592.

[3] C. Domb, *Order-disorder statistics. ii. a two-dimensional model*, Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences, 199 (1949), pp. 199–221.

[4] R. T. Farouki, *The Bernstein polynomial basis: A centennial retrospective*, Computer Aided Geometric Design, 29 (2012), pp. 379–419.

[5] R. P. Feynman, *QED: the strange theory of light and matter*, Princeton University Press, 2014.

[6] M. Finger and G. De Bona, *Probabilistic satisfiability: Logic-based algorithms and phase transition*, in IJCAI, 2011, pp. 528–533.

[7] L. Ford and D. Fulkerson, *Flows in Networks*, Princeton University Press, 1962.

[8] A. R. Forrest, *Interactive interpolation and approximation by Bézier polynomials*, The Computer Journal, 15 (1972), pp. 71–79.

[9] V. Krivoulets and V. Polesskii, *What is the theory of bounds for network reliability?*, Information Transmission in Computeer Networks, 1 (2001), pp. 199–203.

[10] G. G. Lorentz, *Bernstein polynomials*, American Mathematical Soc., 2013.

[11] E. Moore and C. Shannon, *Reliable circuits using less reliable relays*, Journal of the Franklin Institute, 262 (1956), pp. 191–208.

[12] M. Nath, Y. Ren, Y. Khorramzadeh, and S. Eubank, *Determining whether a class of random graphs is consistent with an observed contact network*, Journal of Theoretical Biology, 440 (2018), pp. 121–132, https://doi.org/https://doi.org/10.1016/j.jtbi.2017.12.021, https://www.sciencedirect.com/science/article/pii/S002251931730560X.

[13] M. Nath, S. Venkatramanan, B. Kaperick, S. Eubank, M. V. Marathe, A. Marathe, and A. Adiga, *Using network reliability to understand international food trade dynamics*, in Complex Networks and Their Applications VII, L. M. Aiello, C. Cherifi, H. Cherifi, R. Lambiotte, P. Lió, and L. M. Rocha, eds., Cham, 2019, Springer International Publishing, pp. 524–535.

[14] Y. Ren, S. Eubank, and M. Nath, *From network reliability to the Ising model: A parallel scheme for estimating the joint density of states*, Physical Review E, 94 (2016), p. 042125.

[15] D. Roth, *On the hardness of approximate reasoning*, Artificial Intelligence, 82 (1996), pp. 273–302, https://www.sciencedirect.com/science/article/pii/0004370294000921.

[16] L. G. Valiant, *The complexity of enumeration and reliability problems*, SIAM Journal on Computing, 8 (1979), pp. 410–421.