

Spin systems with an external field: a complexity classification

Leslie Ann Goldberg^{*} and Mark Jerrum[†]

^{*}University of Oxford, United Kingdom, and [†]Queen Mary, University of London, United Kingdom.

Submitted to Proceedings of the National Academy of Sciences of the United States of America

We study the computational complexity of approximating the partition function of a q -state spin system with an external field. There are just three possible levels of computational difficulty, depending on the interaction strengths between adjacent spins: (1) efficiently exactly computable, (2) equivalent to the ferromagnetic Ising model, and (3) equivalent to the antiferromagnetic Ising model. Thus, every non-trivial q -state spin system, irrespective of the number q of spins, is computationally equivalent to one of two fundamental 2-state spin systems.

approximation algorithm | computational complexity | Ising model | spin system

Spin systems

Let $Q = \{1, 2, \dots, q\}$ denote a set of *spins*. A q -state spin system is specified by a non-negative real symmetric *interaction matrix* $A \in (\mathbb{R}_{\geq 0})^{q \times q}$. The entries (a_{ij}) of A represent "interaction strengths" between spins in Q . An *instance* of such a spin system is a graph $G = (V, E)$, where V is a set of vertices (sites) and E is a set of edges (bonds) together with a collection $\mathbf{h} = \{h_w : Q \rightarrow \mathbb{R}_{\geq 0} \mid w \in V\}$ of functions, representing the action of an external field. The function h_w represents the effect of the field on vertex w . The *partition function* of the system is then

$$Z_A(G, \mathbf{h}) = \sum_{\sigma: V \rightarrow Q} \prod_{w \in V} h_w(\sigma(w)) \prod_{\{u, v\} \in E} a_{\sigma(u), \sigma(v)}, \quad [1]$$

a weighted sum over *configurations* $\sigma: V \rightarrow Q$.

The above setting encompasses all spin systems with uniform interactions between pairs of sites, and includes many familiar models. For example, the interaction matrices

$$A_{\text{Ising}}^\lambda = \begin{pmatrix} \lambda & 1 \\ 1 & \lambda \end{pmatrix} \quad A_{\text{IS}} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$$

$$A_{3\text{Potts}}^\lambda = \begin{pmatrix} \lambda & 1 & 1 \\ 1 & \lambda & 1 \\ 1 & 1 & \lambda \end{pmatrix} \quad \text{and} \quad A_{\text{WR}} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix}$$

capture the Ising, independent set (hard-core), 3-state Potts and Widom-Rowlinson models, respectively. In the case of the Ising and Potts models, $\lambda > 1$ corresponds to a ferromagnetic system and $\lambda < 1$ to an antiferromagnetic one.

The problem and what is known

We study the following computational problem. Fix an interaction matrix A . Given an instance consisting of a graph G and field \mathbf{h} , approximately evaluate the partition function $Z_A(G, \mathbf{h})$. An important point to note is that the interaction matrix A does not form part of the problem instance. By fixing A , we fix a particular model, say the hard-core model or the q -state ferromagnetic Potts model. We then ask, for that model: what is the computational complexity of approximately evaluating the partition function given the input pair (G, \mathbf{h}) ? We are interested in determining how that complexity depends on A .

It follows from a general result of Chen, Dyer, Goldberg, Jerrum, Lu, McQuillan and Richerby [1] that there are only three possibilities that can arise for an interaction matrix A .

Possibility 1: The partition function Z_A can be evaluated exactly in polynomial time.

Possibility 2: The complexity of approximately evaluating Z_A is equivalent to the complexity of approximately evaluating the partition function of the ferromagnetic Ising model.

Possibility 3: Approximating Z_A is equivalent to approximating the partition function of the antiferromagnetic Ising model, and hence is NP-hard.

The purpose of this paper is to map the space of interaction matrices, delineating which case arises for each interaction matrix A . Before stating our result, we fill in some details about the three possibilities above (and approximation complexity in general), we say a little bit more about the result of [1], and we present some matrix preliminaries.

The first possibility, that the partition function can be evaluated exactly in polynomial time, is straightforward. Unfortunately, as we shall see, it rarely arises. Let us turn to the second possibility. It will be helpful to introduce the problem #BIS, which plays an important role in the complexity theory of approximate counting problems [2]. #BIS is the problem of counting the independent sets of a bipartite graph. It is essentially the same as the problem of evaluating $Z_{A_{\text{BIS}}}(G, \mathbf{h})$ for the interaction matrix

$$A_{\text{BIS}} = \begin{pmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

Significance

Spin systems are a primary object of study in statistical physics. Computational complexity is the rigorous study of the computational resources required to achieve specified computational goals. We examine the problem of estimating the partition function of a q -state spin system from the point of view of computational complexity. There are just three possible levels of computational difficulty, depending on the interaction strengths between adjacent spins. Every non-trivial q -state spin system, irrespective of the number q of spins, is computationally equivalent to one of two fundamental 2-state spin systems. We give a simple characterisation of the classification.

Reserved for Publication Footnotes

in the special case where, for all $w \in V$, h_w is a constant function which assigns equal weight to all spins. A great number of problems have been shown to be equivalent to $\#BIS$ (and hence to each other) with respect to approximation complexity. We will give a formal definition of this equivalence later, after stating our theorem. For now it suffices to know that if two problems are equivalent, then they are roughly equally difficult to approximate. The authors have shown [3] that the problem of evaluating the partition function of $A_{\text{Ising}}^{\lambda=4}$ is $\#BIS$ -equivalent.¹ Thus, for any interaction matrix A for which the second possibility applies, we can conclude that the complexity of approximately evaluating Z_A is equivalent to that of approximating the partition function of the 2-spin matrix $A_{\text{Ising}}^{\lambda=4}$ even though A itself may have many more than two spins.

Finally, let us consider the third possibility, that approximating the partition function Z_A is NP-hard. This means that there is a constant $\alpha > 0$ such that even approximating Z_A within a factor $\exp(\alpha|V(G)|)$ is as hard as the famous NP-complete problems (which include problems such as determining whether a Boolean formula is satisfiable or determining whether a graph is 3-colourable). It follows that the free energy per site, $|V(G)|^{-1} \ln Z_A(G)$, is NP-hard to approximate within $\pm\alpha$.

It is worth mentioning that these partition function evaluation problems have a property called “self-reducibility”. This means [4] that if you can’t efficiently approximate the partition function then it is also infeasible to sample approximately from the corresponding Gibbs measure.

In order to gain a better understanding of the result of [1], consider the following method for deriving a new interaction matrix from the initial matrix $A = A_{\text{Ising}}^{\lambda=2} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$. Take a graph Γ with three vertices, x , y and z and edges connecting x to z , and x to y . Suppose that h_z satisfies $h_z(1) = 1$ and $h_z(2) = 3$ and that $h_x(1) = h_x(2) = h_y(1) = h_y(2) = 1$. The total contribution to the partition function from configurations with $\sigma(x) = \sigma(y) = 1$ is 10 — the contribution is 2×2 (2 for each edge) if $\sigma(z) = 1$, and it is 2×3 if $\sigma(z) = 2$ (2 for the edge between x and y and 3 from $h_z(2)$). Similarly, the total contributions of the three other possibilities for spins $\sigma(x)$ and $\sigma(y)$ are 4, 7 and 14. Thus, we can view the overall effect of this system as being equivalent to a single edge (x, y) with effective interaction matrix

$$A_{\text{eff}} = \begin{pmatrix} 10 & 5 \\ 7 & 14 \end{pmatrix}.$$

Note that A_{eff} is not even symmetric, so we have moved outside of the symmetric systems that we started with. Obviously, we can create an infinite collection of new interaction matrices from A by varying the graph Γ with specified vertices x and y (and allowing any possible field functions on the vertices). The result of [1] states the following for any interaction matrix A .

- Possibility 1 occurs for A if every single one of the (infinitely many) matrices A_{eff} that can be constructed from A is singular or is diagonal or off-diagonal.
- Possibility 2 occurs if possibility 1 does not occur, but every single one of the matrices A_{eff} that can be constructed has a non-negative determinant or is off-diagonal.
- Otherwise, possibility 3 occurs.

Fortunately, it is not necessary to investigate an infinite collection of matrices A_{eff} to decide which of the three possibilities applies to a particular initial matrix A . In [1] it is shown that distinguishing between the possibilities reduces to deciding whether a certain finite system (based on A) has a property called “balance” — a problem which is known [5] to be decidable — and to deciding whether A has an algebraic

invariant called an STP-MJN multimorphism [6]. This can be checked by brute force, given A . Nevertheless, the existence of a decision algorithm for determining which of the possibilities applies to A does not tell us much, qualitatively, about how and why the interaction matrices divide between the three possibilities. In fact, there is a nice clean characterisation, which we describe in this paper.

Our result

We start with some basic facts about matrices. We say that matrix A is *log-supermodular* if every 2×2 submatrix has non-negative determinant. The matrix A is *irreducible* if, for all $i, j \in Q$, there exists an integer L such that $(A^L)_{ij} > 0$. If A is not irreducible then it may be transformed, by simultaneous row and column permutations, into block diagonal form, after which the complexity of Z_A is determined by the maximum complexity of any block. (We can pick out any desired block by setting h_w , for every vertex $w \in V$, to be the characteristic function of that block.) A stronger condition than irreducibility is primitivity. A matrix A is *primitive* if there exists a number L with $A^L > 0$. Irreducibility and primitivity may be understood in terms of the “underlying graph” of A , i.e., the graph with vertex set Q and edge set $\{\{i, j\} : a_{ij} > 0\}$. Specifically, A is irreducible iff the underlying graph is connected, and is primitive iff it is connected and non-bipartite.

The distinction between a primitive and an imprimitive interaction matrix A is important for us. In the imprimitive case, A has the block form

$$A = \begin{pmatrix} 0 & B \\ B^\top & 0 \end{pmatrix} \quad [2]$$

where B is (say) a $q_1 \times q_2$ matrix, with $q_1 + q_2 = q$. Thus the set of spins may be partitioned as $Q = Q_1 \cup Q_2$, where $Q_1 = \{1, \dots, q_1\}$ and $Q_2 = \{q_1 + 1, \dots, q\}$, in such a way that spins in Q_1 can only be adjacent to spins in Q_2 , and v.v.

An example of an imprimitive system is given by the interaction matrix A_{BIS} . This matrix has the appropriate block structure where $B = A_{\text{IS}}$. The imprimitive form of A_{BIS} ensures that the partition function evaluates to 0 if the instance G is not bipartite. If G is bipartite, with bipartition $V = L \cup R$, then the only configurations σ that contribute to the partition function are ones in which $\sigma(L) \subseteq \{1, 2\}$ and $\sigma(R) \subseteq \{3, 4\}$, or the same with L and R interchanged. Interpreting spins 2 and 4 as meaning “in the independent set” and 1 and 3 as meaning “out of the independent set”, we see that the non-zero terms in [1] correspond to independent sets in G (two terms for each independent set).

In stating the main result, we treat the primitive and imprimitive cases separately. We first state the result and give examples, then we give the formal definitions that clarify the notion of “equivalence” that is used in the $\#BIS$ -equivalent case.

Theorem 1. *First suppose A is primitive.*

- If A has rank 1, then there is a polynomial-time algorithm for evaluating Z_A .
- Otherwise, if there is a simultaneous permutation of the rows and columns of A that renders A log-supermodular, then evaluating Z_A is equivalent to $\#BIS$ under approximation-preserving reductions.
- Otherwise, there is a constant $\alpha > 0$ such that approximating Z_A within a factor $\exp(\alpha|V(G)|)$ is NP-hard.

¹There is no significance in the particular choice of λ , beyond it being greater than 1, and being convenient for the proof.

Now suppose A is imprimitive. Write A in the form [2].

- If B has rank 1, then there is a polynomial-time algorithm for evaluating Z_A .
- Otherwise, if there are independent permutations of the rows and columns of B that render B log-supermodular, then evaluating Z_A is equivalent to $\#BIS$ under approximation-preserving reductions.
- Otherwise, there is a constant $\alpha > 0$ such that approximating Z_A within a factor $\exp(\alpha|V(G)|)$ is NP-hard.

In both primitive and imprimitive cases there are three possibilities. In the first, matrix A (or B , as appropriate) has rank 1, and the partition function factorises. Thus, it is trivially possible to evaluate Z_A exactly in polynomial time. In the second, the matrix A has a particular form, which allows the q -state system with interaction matrix A to be coded up as an instance of $\#BIS$ and v.v. Since, in particular, $Z_{A_{\text{Ising}}^{\lambda=4}}$ is in this class, we see that all these spin systems are computationally equivalent to the ferromagnetic Ising model with an external field. Finally, in the third situation, the antiferromagnetic Ising model can be coded up as an instance of a spin system with interaction matrix A . Since the partition function of the antiferromagnetic Ising model is NP-hard to approximate (even finding the ground state is a hard problem), Z_A is hard to approximate for all such spin systems also. The striking feature of this result is that q -state spin systems for arbitrary q are computationally equivalent to very familiar 2-spin systems.

Example applications of the main theorem. Assuming $\lambda > 0$, the interaction matrix $A_{\text{Ising}}^{\lambda}$ for the Ising model is primitive. Its determinant is positive, zero or negative according to whether λ is greater, equal or less than 1. Thus, the theorem shows that $Z_{A_{\text{Ising}}^{\lambda}}$ has a polynomial-time algorithm when $\lambda = 1$, is equivalent to $\#BIS$ when $\lambda > 1$, and is NP-hard to approximate when $\lambda < 1$. Now consider the interaction matrix $A_{3\text{Potts}}^{\lambda}$ for the 3-state Potts model, which is primitive for all $\lambda \geq 0$. The matrix is of course invariant under simultaneous row and column permutations. If $\lambda = 1$ then the matrix has rank 1; if $\lambda > 1$ then the submatrix $\begin{pmatrix} 1 & 1 \\ \lambda & 1 \end{pmatrix}$ has negative determinant; and if $\lambda < 1$ then the submatrix $\begin{pmatrix} \lambda & 1 \\ 1 & \lambda \end{pmatrix}$ has negative determinant. Thus, the theorem shows that $Z_{A_{3\text{Potts}}^{\lambda}}$ is NP-hard to approximate except when $\lambda = 1$, when it is trivial. The theorem also shows that approximating $Z_{A_{\text{IS}}}$, the partition function of the independent set (hard-core) model on a general graph, is NP-hard. Also, approximating $Z_{A_{\text{WR}}}$, the partition function of the Widom-Rowlinson model, is equivalent to $\#BIS$, as its interaction matrix A_{WR} is log-supermodular.

Finally, we consider an imprimitive example. Consider the matrix A_{BIS} . Applying the theorem to this matrix shows us what we already knew – namely, that approximating $Z_{A_{\text{BIS}}}$ is equivalent to $\#BIS$. To apply the theorem, write A_{BIS} in form [2], so that $B = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}$. Now transpose the rows of B , and note that the resulting matrix is log-supermodular.

The formal definition of equivalence. In order to prove Theorem 1 we have to give a formal meaning to the phrase “evaluating Z_A is equivalent to $\#BIS$ under approximation-preserving reductions”.

For this we need the notion of a “Randomised Approximation Scheme” (see Mitzenmacher and Upfal [7, Definition 10.2]). A randomised approximation scheme for Z_A is a (randomised) algorithm that takes both an input (G, \mathbf{h}) and an accuracy parameter ε , and produces an output \hat{Z} which is an

approximation to $Z_A(G, \mathbf{h})$. The algorithm is randomised, so with some probability (say at most $1/4$, to be precise), it may fail, but the requirement is that, with probability at least $3/4$, the output must satisfy

$$(1 - \varepsilon)Z_A(G, \mathbf{h}) \leq \hat{Z} \leq (1 + \varepsilon)Z_A(G, \mathbf{h}).$$

Thus, with high probability (with probability at least $3/4$), the algorithm must give a good approximation (within a factor of $1 \pm \varepsilon$ of the right answer).

The value “ $3/4$ ” in the definition is not very important. It may be changed to any value strictly between $1/2$ and 1 without any complexity-theoretic consequences. Essentially, the algorithm can be repeated, taking the median over various runs in order to decrease the failure probability.

An important kind of randomised approximation scheme is a so-called “fully polynomial” randomised approximation scheme (FPRAS) which is a randomised approximation scheme that runs quickly, both as a function of its input size and of the desired accuracy. In particular, for an n -vertex graph G , the running time must be bounded from above by a polynomial in n and ε^{-1} . Thus, the algorithm is allowed more time as the amount of tolerable error ε gets smaller, but it is only allowed polynomially more time.

By analogy with Z_A , a randomised approximation scheme for $\#BIS$ is a (randomised) algorithm that takes both an input (in this case a bipartite graph H) and an accuracy parameter δ , and produces an output which, with probability at least $3/4$, is within $(1 \pm \delta)$ of the number of independent sets of H . It is an FPRAS if the running time is at most a polynomial in $|V(H)|$ and δ^{-1} .

Now we need the idea of an Approximation-Preserving (AP-reduction) [2] which is a way of showing that one approximation problem, say the problem of approximating Z_A , is at least as easy as another approximation problem, say $\#BIS$. Specifically, an AP-reduction from Z_A to $\#BIS$ is a randomised algorithm \mathcal{A} for approximating Z_A using a subroutine called an “oracle” for approximating $\#BIS$. The input to \mathcal{A} is of course an instance (G, \mathbf{h}) together with an accuracy parameter ε . Using these, algorithm \mathcal{A} computes an approximation \hat{Z} to $Z_A(G, \mathbf{h})$. It is allowed to make sub-routine calls to the $\#BIS$ oracle with inputs of the form (H, δ) where H is a bipartite graph and $0 < \delta < 1$ is an error bound with $\delta^{-1} \leq \text{poly}(n, \varepsilon^{-1})$ where $n = |V(G)|$. The reduction must have the property that \mathcal{A} is a randomised approximation scheme for Z_A whenever the oracle is a randomised approximation scheme for $\#BIS$. Also, the run-time of \mathcal{A} must be a polynomial in n and in ε^{-1} . The main point of all of that is that an AP-reduction from Z_A to $\#BIS$ turns an FPRAS for $\#BIS$ into an FPRAS for Z_A .

We say that the problem of approximating Z_A is *AP-reducible* to $\#BIS$ if an AP-reduction from Z_A to $\#BIS$ exists. We can similarly define an AP-reduction from $\#BIS$ to Z_A . If reductions in both directions exist then we say that Z_A and $\#BIS$ are *equivalent under AP-reductions*. This is the case for all interaction matrices A in our second category.

The proof of Theorem 1

The proof of Theorem 1 is in three parts. First we show how to separate the NP-hard cases from the $\#BIS$ -equivalent cases. Then we show how to separate the $\#BIS$ -equivalent cases from the polynomial-time computable cases. The final section puts the pieces together to complete the proof.

The proof: NP-hardness

In this section, we separate NP-hard spin systems from those that are computationally equivalent to $\#BIS$. As usual in

the theory of NP-completeness, we employ carefully designed gadgets, but in this case they have a simple form.

Gadgets. A gadget Γ is a path of length L together with a set of allowed spins for each vertex that is not an endpoint of the path. Let the vertices of the path be $\{0, 1, \dots, L\}$ and the set of allowed spins at vertex k be $S_k \subseteq Q$, for all $1 \leq k \leq L-1$. Note that restricting the spins at k to a set S_k is equivalent to applying the field h_k to the vertex, where h_k is the characteristic function of the set S_k . So a gadget is an instance of a spin system of the kind we are studying. The matrix A_Γ of effective interaction strengths for Γ is given by

$$A_\Gamma = AD_1AD_2A \cdots AD_{L-1}A, \quad [3]$$

where D_k is the diagonal matrix with 1s at positions $\{(i, i) : i \in S_k\}$ and 0s elsewhere. Thus, for all $i, j \in Q$, the entry $(A_\Gamma)_{ij}$ represents the multiplicative contribution made by the gadget to the partition function when $\sigma(0) = i$ and $\sigma(L) = j$. If $S_k = Q$ for all $1 \leq k \leq L-1$ we say that the gadget is *free*; if $|S_k| = 2$ for all $1 \leq k \leq L-1$ we say that the gadget is *tight*. Note that it doesn't make sense to restrict a tight gadget further: if any S_k is a singleton then A_Γ has rank 1, and there is no correlation between spins $\sigma(0)$ and $\sigma(L)$.

Here is some useful notation. $M[i, i'; j, j']$ is the 2×2 matrix obtained from a matrix M by deleting all rows except i, i' and all columns except j, j' . We regard the indices in the notation $M[i, i'; j, j']$ as ordered; thus the first row of this 2×2 matrix comes from row i of M and the second from row i' .

We are interested in how gadgets interact with (ordered) pairs of spins. We use Q^2 to denote the set of all ordered pairs of *distinct* elements of a set Q . Thus for a primitive matrix A the relevant set of ordered pairs is $P = Q^2$ whereas for imprimitive A the relevant set is $P = Q_1^2 \cup Q_2^2$. We don't consider mismatched pairs that mix elements of Q_1 and Q_2 .

We now give some definitions concerning the properties of gadgets. We say that the gadget Γ is *ferromagnetic* (respectively, *antiferromagnetic*, *neutral*) *between* ordered pairs of spins $(i, i') \in P$ and $(j, j') \in P$ if $\det A_\Gamma[i, i'; j, j'] > 0$ (respectively, < 0 , $= 0$). If, in addition, the matrix $A_\Gamma[i, i'; j, j']$ has at most one zero entry then we say that Γ is *permissive* ferromagnetic, permissive antiferromagnetic, or permissive neutral between (i, i') and (j, j') . Finally, we say that a pair of spins (i, i') is antiferromagnetic if there exists a gadget that is antiferromagnetic between (i, i') and (i, i') . There is no need to define a similar concept of a ferromagnetic pair of spins, since all pairs of spins are "ferromagnetic" via the gadget of length $L = 0$. However, we say that a pair of spins (i, i') is permissive ferromagnetic if there exists a gadget that is permissive ferromagnetic between (i, i') and (i, i') .

Given a tight gadget Γ , with $S_k = \{i_k, i'_k\}$ for all $1 \leq k \leq L-1$, we may write $A_\Gamma[i, i'; j, j']$ as the following product of 2×2 submatrices of A :

$$A_\Gamma[i, i'; j, j'] = A[i, i'; i_1, i'_1] \times A[i_1, i'_1; i_2, i'_2] \times \cdots \times A[i_{L-1}, i'_{L-1}; j, j']; \quad [4]$$

note that the ordering we impose on the elements $\{i_r, i'_r\}$ of each S_r is irrelevant. Then the earlier definitions can be phrased in terms of the determinants of the 2×2 matrices in product [4]. Thus, A_Γ is antiferromagnetic between (i, i') and (j, j') iff none of the matrices in [4] is singular, and an odd number of them have negative determinant. Furthermore, Γ is permissive antiferromagnetic iff, in addition, at least one of the matrices has at most one zero entry. This observation is useful for avoiding calculation.

Here is an illustrative example based on the interaction matrix $A = A_{3\text{Potts}}^{\lambda=2}$ of a ferromagnetic 3-state Potts model. Recall that A has 2s on the diagonal and 1s elsewhere. Consider the gadget Γ of length $L = 3$ with $S_1 = \{1, 3\}$ and $S_2 = \{3, 2\}$. Then

$$\begin{aligned} A_\Gamma[1, 2; 1, 2] &= A[1, 2; 1, 3] \times A[1, 3; 3, 2] \times A[3, 2; 1, 2] \\ &= \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 7 & 10 \\ 5 & 7 \end{pmatrix}. \end{aligned}$$

So $\det(A_\Gamma[1, 2; 1, 2]) = -1$, and the pair $(1, 2)$ is antiferromagnetic. The point to note is that, although we started with a ferromagnetic model, it is possible to use the external field to simulate an edge of an antiferromagnetic 2-spin system. By Lemma 3 below, we will be able to conclude that approximating the partition function of the ferromagnetic 3-state Potts model is intractable (NP-hard).

Our strategy in the remainder of this section is to provide a systematic approach to constructing such antiferromagnetic gadgets, which fails only when the interaction matrix A has a special form. In the following section, we indicate how a spin system whose interaction matrix has this special form can be simulated by an instance of #BIS or, equivalently, an instance of the ferromagnetic Ising model. We start with some technical lemmas concerning gadgets and their properties. The proofs are included in the Supporting Information.

Lemma 1. *If the pair $(i, i') \in P$ is antiferromagnetic, then there is a tight, permissive gadget that is antiferromagnetic between (i, i') and (i, i') .*

Lemma 2. *Suppose that the free gadget of odd length L is neutral between $(i, i') \in P$ and $(j, j') \in P$, but that $\det A[i, i'; j, j'] \neq 0$. Then the pair (i, i') is antiferromagnetic.*

Lemma 3. *If there exists an antiferromagnetic pair $(i, i') \in P$, then there is a constant $\alpha > 0$ such that approximating Z_A within a factor $\exp(\alpha|V(G)|)$ is NP-hard.*

Lemma 3 gives a sufficient condition for it to be NP-hard to approximate Z_A . Note that, by Lemma 1, the condition may be tested by dynamic programming in polynomial time. Roughly, the method is as follows. Fix a pair (i, i') . Let Ω_L^a be the set of pairs (j, j') such that there exists a tight gadget of length less than or equal to L that is antiferromagnetic between (i, i') and (j, j') and let Ω_L^f be the set of pairs (j, j') such that there exists a tight gadget of length less than or equal to L that is ferromagnetic between (i, i') and (j, j') . We can compute $(\Omega_1^a, \Omega_1^f), (\Omega_2^a, \Omega_2^f), \dots$ in turn, using (Ω_i^a, Ω_i^f) as data to compute $(\Omega_{i+1}^a, \Omega_{i+1}^f)$. (In fact, the set Ω_i^f is just the reversal of all pairs in Ω_i^a , so technically, it does not need to be explicitly stored.) The sequence $\Omega_1^a \subseteq \Omega_2^a \subseteq \dots$ is monotonic and must converge before $L = n^2$. We now show that if the condition in Lemma 3 fails, then A has a particular form. There are two cases to our analysis, depending on whether or not A is primitive.

The primitive case. A *tournament* is a complete graph with orientations assigned to the edges. A *Hamiltonian path* in a tournament is a directed path that visits all of the vertices of the tournament exactly once, and which respects the orientations of the edges. It is a standard graph theoretic fact that every tournament has a Hamiltonian path [8] (or see [9, Thm 2.3]).

Lemma 4. *Suppose the matrix A is primitive, and that it does not support an antiferromagnetic pair of spins (i.e., there does not exist a pair $(i, i') \in P$ and a gadget Γ such that $\det A_\Gamma[i, i'; i, i'] < 0$). Then there is a simultaneous permutation of the rows and columns of A such that the resulting matrix is log-supermodular.*

Proof. Let L be the minimum odd number such that $A^L > 0$. Note that A^L is the interaction matrix of the free gadget of length L . Consider the graph with vertex set Q^2 , whose adjacency relation \sim is defined by the condition $(i, i') \sim (j, j')$ iff $\det A^L[i, i'; j, j'] > 0$. Note the obvious symmetry $(i, i') \sim (j, j')$ iff $(i', i) \sim (j', j)$. Note also that if (i, i') and (j, j') are in the same component of (Q^2, \sim) then there is a ferromagnetic gadget between (i, i') and (j, j') . To see this, take a shortest path from (i, i') to (j, j') in (Q^2, \sim) . Each edge in the path corresponds to a ferromagnetic gadget; concatenate these to obtain a ferromagnetic gadget between (i, i') and (j, j') . More formally, suppose the path is $(i, i') = (i_0, i'_0), (i_1, i'_1), \dots, (i_\ell, i'_\ell) = (j, j')$. Then

$$\begin{aligned} \det(A^L[i_0, i'_0; i_1, i'_1] \times \dots \times A^L[i_{\ell-1}, i'_{\ell-1}; i_\ell, i'_\ell]) \\ = \det A^L[i_0, i'_0; i_1, i'_1] \times \dots \times \det A^L[i_{\ell-1}, i'_{\ell-1}; i_\ell, i'_\ell] > 0, \end{aligned}$$

from which a gadget of length ℓL that is ferromagnetic between $(i, i') = (i_0, i'_0)$ and $(j, j') = (i_\ell, i'_\ell)$ may be read off.

By assumption, for every pair $(i, i') \in Q^2$ it is the case that (i, i') and (i', i) are in different components of (Q^2, \sim) . (If there were in the same component, then there would be a ferromagnetic gadget between (i, i') and (i', i) , and this would imply that the pair (i, i') is antiferromagnetic.) Thus the components of (Q^2, \sim) come in symmetric pairs C, C' , where C' contains exactly the reversals of all the pairs in C . We can then construct a tournament on Q by the following procedure. For each pair of matched components C and C' , choose one, say C , and add a directed edge from i to i' for every pair $(i, i') \in C$. Note that this procedure runs into no conflicts, and determines a direction for every edge in the complete graph on vertex set Q . Let the resulting tournament be T . (The construction of T and the use we are about to make of it is a simplified version of a similar line of proof used in a more general situation by Cohen, Cooper and Jeavons [10].)

Take any Hamiltonian path in the tournament T , and form the linear order \sqsubset on Q such that the vertices/spins along the path appear in increasing order in (Q, \sqsubset) . By renumbering Q we may assume that $1 \sqsubset 2 \sqsubset \dots \sqsubset q$; note that this operation is equivalent to applying a permutation simultaneously to the rows and columns of A^L (and A itself). In this permuted matrix, every 2×2 submatrix of the form $A^L[i, i+1; j, j+1]$ has non-negative determinant. (Suppose to the contrary that $\det A^L[i, i+1; j, j+1] < 0$. Then $(i, i+1) \sim (j+1, j)$ and hence either $i \sqsubset i+1$ and $j+1 \sqsubset j$, or $i+1 \sqsubset i$ and $j \sqsubset j+1$, by construction of \sqsubset . Both possibilities contradict the fact that the rows and columns of A^L are ordered by \sqsubset .) Since $A^L > 0$, this implies (a standard fact about finite Monge matrices [11, Observation 2.2]) that every 2×2 submatrix $A^L[i, i'; j, j']$ with $i < i'$ and $j < j'$ has non-negative determinant, i.e., the matrix A^L is log-supermodular. (See the end of the proof for an illustration of this construction.)

Finally we argue that A itself must be log-supermodular. Assume to the contrary that it is not, i.e., that there are $i < i'$ and $j < j'$ such that $\det A[i, i'; j, j'] < 0$. Then the free gadget of length 1 is antiferromagnetic between (i, i') and (j, j') . Recall that the matrix A^L is realised by the free gadget of length L . If $\det A^L[j, j'; i, i'] > 0$ then we immediately obtain a gadget of length $L+1$ that is antiferromagnetic between (i, i') and (i, i') , a contradiction. (Explicitly, the gadget is obtained from the free gadget of length $L+1$ by restricting the spin on the second vertex to be in the set $\{j, j'\}$.) The only other possibility is that $\det A^L[j, j'; i, i'] = 0$, but this is ruled out by Lemma 2. \square

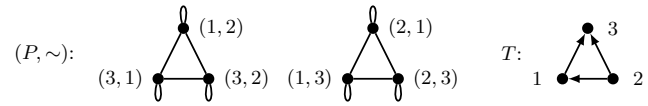


Fig. 1. The graph (P, \sim) and one of two tournaments T arising from it.

An example to illustrate the tournament construction. Consider the Widom-Rowlinson interaction matrix, but permute the rows and columns to hide the log-supermodular structure. To obtain matrix A below we transposed rows 1 and 2, and columns 1 and 2. The third power of A (corresponding to $L = 3$) is strictly positive:

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \quad \text{and} \quad A^L = A^3 = \begin{pmatrix} 7 & 5 & 5 \\ 5 & 4 & 3 \\ 5 & 3 & 4 \end{pmatrix}.$$

The graph (P, \sim) , together with one of the two possible tournaments T arising from it, is depicted in Figure 1.

The (unique) Hamilton path in T is $2 \rightarrow 1 \rightarrow 3$ corresponding to the linear order $2 \sqsubset 1 \sqsubset 3$. This order indicates that we need to transpose rows 1 and 2, and columns 1 and 2; if we do this then we recover the original log-supermodular matrix A_{WR} . (If we take the other tournament, then the linear order is $3 \sqsubset 1 \sqsubset 2$, but we arrive at the same end result.)

The imprimitive case. Recall that in this case A has the block form [2].

Lemma 5. Suppose the matrix A is imprimitive, and that it does not support an antiferromagnetic pair of spins (i.e., there is no pair $(i, i') \in P$ and gadget Γ such that $\det A_\Gamma[i, i'; i, i'] < 0$). Write A in the form [2]. Then there are independent permutations of the rows and columns of B such that the resulting matrix is log-supermodular.

Proof. Let L be the smallest odd integer such that the matrix

$$A^L = \begin{pmatrix} 0 & B^* \\ (B^*)^\top & 0 \end{pmatrix}$$

has the property that $B^* > 0$. (Explicitly, $B^* = B(B^\top B)^{(L-1)/2}$.) We construct a graph on vertex set $P = Q_1^2 \cup Q_2^2$. The adjacency relation \sim of this graph is defined as follows: $(i, i') \sim (j, j')$ iff $\det A^L[i, i'; j, j'] > 0$. (Necessarily $(i, i') \in Q_1^2$ and $(j, j') \in Q_2^2$, or v.v.) Note the obvious symmetry $(i, i') \sim (j, j')$ iff $(i', i) \sim (j', j)$. Note also that if (i, i') and (j, j') are in the same component of (P, \sim) then there is a ferromagnetic gadget between (i, i') and (j, j') .

As in the primitive case, for every pair $(i, i') \in P$ it is the case that (i, i') and (i', i) are in different components of (P, \sim) . Thus the components of (P, \sim) come in symmetric pairs C, C' , where C' contains exactly the reversals of all the pairs in C . Using the same procedure as before, we may construct two tournaments, T_1 on Q_1^2 and T_2 on Q_2^2 .

Also as before, we may use Hamiltonian paths in T_1 and T_2 to define linear orders \sqsubset_1 on Q_1 and \sqsubset_2 on Q_2 . By renumbering Q_1 and Q_2 (within themselves) we may assume that $1 \sqsubset_1 2 \sqsubset_1 \dots \sqsubset_1 q_1$ and $q_1+1 \sqsubset_2 q_1+2 \sqsubset_2 \dots \sqsubset_2 q$; note that this operation is equivalent to applying independent permutations to the rows and columns of B^* (and B). In this permuted matrix, every 2×2 submatrix of the form $B^*[i, i+1; j, j+1]$ has non-negative determinant. Since $B^* > 0$, this implies as before that every 2×2 submatrix $B^*[i, i'; j, j']$ with $i < i'$ and $j < j'$ has non-negative determinant, i.e., the matrix B^* is log-supermodular.

The argument that B itself is log-supermodular goes through exactly as in the primitive case. \square

We have now completed the NP-hardness section. Lemma 3 gave a sufficient condition for it to be NP-hard to approximate Z_A and we showed (shortly after the statement of Lemma 3) how to use Lemma 1 to test the condition in polynomial time. If the condition fails (so there is no antiferromagnetic pair) then Lemmas 4 and 5 show that there is a permutation of rows and columns such that the resulting matrix is log-supermodular. We will soon see that this means that approximating Z_A is #BIS-easy. First, however, we examine how to identify #BIS-hardness.

The proof: equivalence to #BIS

Lemma 6. *Suppose the interaction matrix A satisfies one of the following two conditions:*

- *A is primitive and has rank greater than 1, or*
- *A is imprimitive and has rank greater than 2 (equivalently, the matrix B in the block decomposition [2] of A has rank greater than 1).*

Then there exists a pair of spins $(i, i') \in P$ such that A supports a gadget that is permissive ferromagnetic between (i, i') and (i, i') .

Proof. In the primitive case, let L be the smallest integer such that $A^L > 0$. Since A is symmetric and hence diagonalisable, the rank of A is equal to the number of non-zero eigenvalues of A (counted according to multiplicity). The eigenvalues of A^L are just the L th powers of the eigenvalues of A , and so the rank of A^L is equal to the rank of A , which it is greater than 1 by assumption. We argue that this implies that some principal 2×2 submatrix $A^L[i, i'; i, i']$ of A^L has full rank. Suppose to the contrary that every submatrix $A^L[i, i'; i, i']$ has rank 1. Let the diagonal entries of A^L be $\mathbf{d} = (d_1, \dots, d_n)$. Then by considering the submatrix $A^L[i, i'; i, i']$ we see that $(A^L)_{ii'} = (A^L)_{i'i} = \sqrt{d_i d_{i'}}$, for all i, i' . This gives an explicit expression for A^L as a rank 1 matrix, namely $A^L = (\sqrt{d_1}, \dots, \sqrt{d_n})^T (\sqrt{d_1}, \dots, \sqrt{d_n})$. A submatrix $A^L[i, i'; i, i']$ of full rank provides a gadget that is permissive but not neutral between (i, i') and (i, i') . If the gadget is antiferromagnetic, we can concatenate it with itself to produce a ferromagnetic one.

In the imprimitive case, let L be the smallest even number such that, for some $A_{11} > 0$ and $A_{22} > 0$, A^L has the form

$$A^L = \begin{pmatrix} A_{11} & 0 \\ 0 & A_{22} \end{pmatrix}.$$

Arguing as before, A has rank greater than 2, so either A_{11} or A_{22} has rank greater than 1. It follows that either A_{11} or A_{22} has a principal 2×2 submatrix of full rank. \square

The proof: putting the pieces together

Theorem 1 follows quickly. Consider first the primitive case. If A has rank 1, then the partition function factorizes, giving an easy polynomial-time (exact) algorithm for evaluating Z_A . (This is similar to the trivial direction of the dichotomy result of Bulatov and Grohe [12] — the external fields do not cause any added difficulty.) Otherwise, by Lemma 6, there is a gadget that is permissive ferromagnetic, from which it follows (see [13, Theorem 10.2]) that #BIS is AP-reducible to the evaluation of Z_A . If A can be permuted to be log-supermodular, then, by [1, Theorem 47], evaluating Z_A is AP-reducible to #BIS. (The reduction and its verification are the essential content of [1, Lemma 46].) Finally, if A cannot be so permuted, then there is an antiferromagnetic pair by Lemmas 4 and 5, and evaluating Z_A is NP-hard to approximate by Lemma 3. The argument in the imprimitive case is identical.

Discussion

We have now seen exactly what causes the approximate evaluation of the partition function Z_A to be NP-hard — namely, an antiferromagnetic pair (i, i') which can easily be demonstrated via a tight permissive gadget. If no such gadget exists, then approximating Z_A is #BIS-easy. We have also seen exactly what causes the approximate evaluation of Z_A to be #BIS-hard — namely, the existence of a block with rank greater than 1. Without such a block, Z_A can be evaluated exactly in polynomial time.

Another interesting question is the complexity of approximately evaluating Z_A when the external field acts in a restricted manner. For example, the algorithm of Jerrum and Sinclair [14] gives an FPRAS for $Z_{A_{\text{Ising}}}$ in the ferromagnetic case $\lambda > 1$ as long as the field acts in a consistent manner in the sense that the sign of $h_w(2) - h_w(1)$ is the same for all vertices [15]. In the antiferromagnetic regime it is known that tractability corresponds to the uniqueness threshold when there are just two spins [16]. The ferromagnetic regime is not fully understood [17], even with just two spins, and the multi-spin case is open.

ACKNOWLEDGMENTS. The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013) ERC grant agreement no. 334828. The paper reflects only the authors' views and not the views of the ERC or the European Commission. The European Union is not liable for any use that may be made of the information contained therein.

1. Chen X et al. (2015) The complexity of approximating conservative counting CSPs. *J. Comput. System Sci.* 81(1):311–329.
2. Dyer M, Goldberg LA, Greenhill C, Jerrum M (2004) The relative complexity of approximate counting problems. *Algorithmica* 38(3):471–500. Approximation algorithms.
3. Goldberg LA, Jerrum M (2007) The complexity of ferromagnetic Ising with local fields. *Combin. Probab. Comput.* 16(1):43–61.
4. Jerrum M, Valiant LG, Vazirani VV (1986) Random generation of combinatorial structures from a uniform distribution. *Theor. Comput. Sci.* 43:169–188.
5. Cai J, Chen X, Lu P (2011) Non-negatively Weighted #CSP: An Effective Complexity Dichotomy. pp. 45–54.
6. Kolmogorov V, Zivny S (2013) The complexity of conservative valued csps. *J. ACM* 60(2):10.
7. Mitzenmacher M, Upfal E (2005) Probability and Computing. (Cambridge University Press).
8. Rédei L (1934) Ein kombinatorischer Satz. *Acta Litt. Sci. Szeged* 7:39–43.
9. Bondy JA, Murty USR (2008) Graph theory, Graduate Texts in Mathematics. (Springer, New York) Vol. 244, pp. xii+651.
10. Cohen DA, Cooper MC, Jeavons PG (2008) Generalising submodularity and Horn clauses: tractable optimization problems defined by tournament pair multimorphisms. *Theoret. Comput. Sci.* 401(1-3):36–51.
11. Rudolf R, Woeginger GJ (1995) The cone of Monge matrices: extremal rays and applications. *ZOR—Math. Methods Oper. Res.* 42(2):161–168.
12. Bulatov A, Grohe M (2005) The complexity of partition functions. *Theoret. Comput. Sci.* 348(2-3):148–186.
13. Bulatov AA, Dyer ME, Goldberg LA, Jerrum M, McQuillan C (2013) The expressibility of functions on the boolean domain, with applications to counting csps. *J. ACM* 60(5):32.
14. Jerrum M, Sinclair A (1993) Polynomial-time approximation algorithms for the Ising model. *SIAM J. Comput.* 22(5):1087–1116.
15. Goldberg LA, Jerrum M, Paterson M (2003) The computational complexity of two-state spin systems. *Random Struct. Algorithms* 23(2):133–154.
16. Li L, Lu P, Yin Y (2013) Correlation Decay up to Uniqueness in Spin Systems. pp. 67–84.
17. Liu J, Lu P, Zhang C (2014) The Complexity of Ferromagnetic Two-spin Systems with External Fields. pp. 843–856.