Bidirectional Inference Counts on Forests and Cycles

Alper Ferudun*

29 September 2026

Abstract

For a binary restricted Boltzmann machine on a fixed labeled bipartite graph, consider the pair of conditional maximum-a-posteriori maps in both directions, with the same weights used in the two maps and no ties. We count such pairs exactly on forests and on even cycles. On a forest, local threshold functions can share symmetric weights if and only if their monotonicity directions agree on every input edge essential at both endpoints. Positive rescaling gives a constructive proof and an incidence partition function evaluable by a tree recurrence. For paths, the counts satisfy $p_1 = 2$, $p_2 = 14$, and $p_n = 10p_{n-1} + 6p_{n-2}$. For even cycles of length $n \geq 4$, the count is $(5+\sqrt{31})^n + (5-\sqrt{31})^n - 2^{n+1}$; the subtraction removes exactly two cyclic families of impossible magnitude comparisons. Two trees with identical degree lists in each bipartition class have different bidirectional counts. This invariant is distinct from the classical one-way inference count, which factors over output degrees.

1 The counting problem

Let $G = (V \sqcup H, E)$ be a finite labeled simple bipartite graph. A binary restricted Boltzmann machine (RBM) assigns joint mass proportional to

$$\exp\left(\sum_{v\in V} b_v x_v + \sum_{h\in H} b_h y_h + \sum_{vh\in E} w_{vh} x_v y_h\right), \qquad (x,y) \in \{0,1\}^{V\sqcup H}.$$

For either type of vertex u, its conditional MAP update is

$$f_u(z) = \mathbf{1}\left\{b_u + \sum_{v\in N(u)} w_{uv} z_v > 0\right\}. \tag{1}$$

We require every displayed affine form to be nonzero on every local binary input. All parameters are real and unconstrained in magnitude. The same $w_{uv} = w_{vu}$ is used at both endpoints.

Let $\mathcal{B}(G)$ be the number of tuples $(f_u)_{u\in V\sqcup H}$ obtained as these strict parameters vary. This is equivalently the number of ordered pairs of conditional MAP maps $\{0,1\}^V \to \{0,1\}^H$ and $\{0,1\}^H \to \{0,1\}^V$. We count functions, not parameter vectors, and do not identify tuples under graph automorphisms or state relabeling. Connected components contribute independently. In particular, $\mathcal{B}(G_1 \sqcup G_2) = \mathcal{B}(G_1)\mathcal{B}(G_2)$ and an isolated vertex contributes 2.

Write $T(d)$ for the number of Boolean linear threshold functions on d labeled inputs, including constants. The standard one-way RBM count is

$$\prod_{h\in H} T(\deg h). \tag{2}$$

*Mercury Software GmbH. Email: alper@mercurycodelab.com. GitHub: AlperTheKing.

Indeed each hidden coordinate is an arbitrary threshold function, and different hidden vertices use disjoint sets of parameters. The reverse one-way count is the analogous product over V. This follows directly from the classical description of Cueto, Morton and Sturmfels [4, Proposition 5.1]; see also [5, Section 4.2] for discrete inference and fan slicings. Formula (2) is not a new result claimed here. It does not count pairs constrained to share weights: even a single edge has 14 pairs, rather than $T(1)^2 = 16$.

The 2018 AIM workshop asks about inference under fixed connectivity [1]. We study the explicitly bidirectional extension just defined, not an asserted reinterpretation or full resolution of that source item. Symmetric threshold networks and their energy and symmetrization questions have a long history; see [3, 2]. The steady-state transformation of [2, Theorem 1] can change the graph, whereas our forest realization below preserves every specified local truth table on its original graph. We make no priority claim for positive diagonal rescaling as a general method.

2 Exact realization and counting on forests

An input of a Boolean function is *essential* if changing it can change the output while fixing the other inputs. A threshold function is *unate*: it is either increasing or decreasing in each input. The direction is unique on every essential input, and agrees with the sign of its coefficient in any strict affine representation.

Theorem 1. *Suppose G is a forest. Specify a threshold function f_u on $N(u)$ at every vertex. The tuple has a shared-weight realization (1) if and only if, on every edge essential at both endpoints, the two unateness directions agree.*

Proof. Necessity follows from the common coefficient on an edge. For sufficiency, choose an edge sign forced by an essential endpoint if there is one, and choose either sign otherwise. The hypothesis makes these choices consistent.

Choose a strict affine representation of each function on its essential inputs and initially set its other coefficients to zero. Such a representation exists by fixing the irrelevant inputs in any representation. Its finitely many input values have a positive minimum absolute value. We may therefore perturb each zero coefficient to a sufficiently small nonzero number of the assigned edge sign without changing any output. Write $a_{u,e}$ for the resulting local coefficients. The coefficients at the two ends of each edge have the same sign.

Root each tree and put $\lambda_r = 1$ at its root. For an edge e from parent u to child v, define

$$\lambda_v = \lambda_u \frac{a_{u,e}}{a_{v,e}} > 0.$$

The unique paths in a tree make these assignments consistent. Multiply the entire local affine form at u, including its bias, by λ_u. Its outputs are unchanged, and the two coefficients on e now both equal $\lambda_u a_{u,e} = \lambda_v a_{v,e}$. These rescaled forms are the desired shared-weight realization. Isolated vertices cause no constraint. □

The proof also shows that requiring every allowed edge weight to be nonzero does not change the count. No bound on the sizes of the constructed weights is asserted.

Let $E(k)$ count threshold functions essential in all k inputs, and let $M(k)$ count the increasing ones among them. Decomposing by the essential set and flipping decreasing inputs gives

$$E(k) = \sum_{j=0}^{k} (-1)^{k-j} \binom{k}{j} T(j) = 2^k M(k). \tag{3}$$

For $k = 0$ there are two constants, so $M(0) = 2$. The standard initial values $T(0), \ldots, T(4) = 2, 4, 14, 104, 1882$ [4] give $M(0), \ldots, M(4) = 2, 1, 2, 9, 96$.

Theorem 2. *For every forest G, let $a_{u,e} \in \{0,1\}$ be a binary variable for each vertex-edge incidence, and put $k_u = \sum_{e \ni u} a_{u,e}$. Then*

$$\mathcal{B}(G) = \sum_{a \in \{0,1\}^{\{(u,e):u \in e\}}} \prod_u M(k_u) \prod_{e=uv} W_{a_{u,e},a_{v,e}}, \qquad W = \begin{pmatrix} 1 & 2 \\ 2 & 2 \end{pmatrix}. \tag{4}$$

Proof. Interpret $a_{u,e} = 1$ as essentiality at that incidence. Once essential sets and their directions have been specified, vertex u admits exactly $M(k_u)$ threshold functions. An edge essential at neither endpoint has no observable direction and contributes 1. An edge essential at one or both endpoints has exactly two consistent choices of direction and contributes 2. Theorem 1 realizes every counted tuple. Conversely, each realized tuple uniquely determines these essential sets, their observable directions and the local functions, so no tuple is counted twice. □

Here is a bottom-up evaluation of (4). At a nonroot vertex u, let $m_u(a)$ denote its descendant contribution conditional on the parent incidence being a, omitting the parent edge factor. For each child v, set

$$U_v = m_v(0) + 2m_v(1), \qquad V_v = 2m_v(0) + 2m_v(1), \qquad \prod_{v \text{ child of } u} (U_v + V_v z) = \sum_k c_k z^k.$$

Then

$$m_u(a) = \sum_k c_k M(k + a). \tag{5}$$

At the root use $\sum_k c_k M(k)$ and multiply over components. A leaf message is $(2, 1)$. Multiplying the linear factors by elementary polynomial arithmetic takes $O(\sum_u \deg(u)^2)$ arithmetic operations, given $T(0), \ldots, T(\Delta)$, where Δ is the maximum degree. For fixed Δ this is linear in the number of vertices. This assertion is not a degree-uniform bit-complexity bound or an algorithm for arbitrary $T(d)$.

Corollary 3. *The star with d leaves has*

$$\mathcal{B}(K_{1,d}) = \sum_{k=0}^{d} \binom{d}{k} 3^k T(k). \tag{6}$$

For $d = 0, 1, 2, 3, 4$, the counts are $2, 14, 152, 3224, 164480$.

Proof. For a fixed central function essential in r inputs, an essential leaf admits three unary functions: its two constants and its compatible nonconstant function. An inessential leaf admits all four unary functions. Thus the count is

$$\sum_{r=0}^{d} \binom{d}{r} E(r) 3^r 4^{d-r}.$$

Substitute (3). The coefficient of $T(k)$ becomes $\binom{d}{k} 3^k \sum_{j=0}^{d-k} \binom{d-k}{j} (-3)^j 4^{d-k-j} = \binom{d}{k} 3^k$, proving (6). □

3 Paths and the cycle obstruction

Put

$$A = \begin{pmatrix} 4 & 6 \\ 5 & 6 \end{pmatrix}, \qquad \lambda_\pm = 5 \pm \sqrt{31}.$$

For an internal degree-two vertex with one child, (5) acts on the column $(m(0), m(1))^\mathsf{T}$ by A. At an endpoint root the functional is $(4, 6)$. Hence for the path P_n on n vertices,

$$p_1 = 2, \quad p_2 = 14, \quad p_n = 10p_{n-1} + 6p_{n-2} \quad (n \geq 3), \qquad p_n = \mathcal{B}(P_n). \tag{7}$$

For example the next terms are $152, 1604, 16952, 179144$. The recurrence follows from $A^2 = 10A + 6I$; for $n = 3$ it also follows directly from $(4, 6)A(2, 1)^\mathsf{T} = 152$. Equivalently,

$$p_n = \left(1 + \frac{2}{\sqrt{31}}\right)\lambda_+^{n-1} + \left(1 - \frac{2}{\sqrt{31}}\right)\lambda_-^{n-1}. \tag{8}$$

Sign compatibility alone is no longer sufficient on a cycle. Its exact failure is captured by the following elementary observation.

Lemma 4. *For a two-input threshold function with prescribed compatible coefficient signs, let $\alpha, \beta > 0$ be their magnitudes. A constant function can be realized with any such magnitudes. A function essential in both inputs can also be realized with any such magnitudes. A function essential only in the first input can be realized if and only if $\alpha > \beta$.*

Proof. Flip each negatively weighted input. This makes its coefficient positive and merely changes the local bias. Increasing Boolean functions essential in both inputs are AND and OR. A threshold between $\max(\alpha, \beta)$ and $\alpha + \beta$ realizes AND; one between 0 and $\min(\alpha, \beta)$ realizes OR. Constant functions are obtained by putting the threshold outside the whole range. For a function essential only in the first input, the values in its output-zero layer range from 0 to β, and the output-one layer ranges from α to $\alpha + \beta$. Strict separation is equivalent to $\beta < \alpha$. The second-input case is symmetric. $\square$

Theorem 5. *For every even integer $n \geq 4$,*

$$\mathcal{B}(C_n) = \operatorname{tr}(A^n) - 2^{n+1} = \lambda_+^n + \lambda_-^n - 2^{n+1}. \tag{9}$$

In particular, $\mathcal{B}(C_4) = 12440$ and $\mathcal{B}(C_6) = 1392704$.

Proof. First count tuples with compatible essential signs, before imposing any magnitude constraints. At a degree-two vertex, its two incidence states a, b have vertex weight $M(a + b)$, giving the matrix

$$K = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}.$$

Summing the products of alternating vertex factors K and edge factors W around the labeled cycle gives $\operatorname{tr}((KW)^n)$. Since $KW = A$, this is $\operatorname{tr}(A^n)$. It counts each tuple once, for the same observable-sign reason as in Theorem 2; no realizability is assumed yet.

By Lemma 4, the only restrictions on positive edge magnitudes come from vertices having exactly one essential incidence. Each requires the essential edge to have strictly larger magnitude than its other edge. Regard these as directed comparisons between consecutive edges of C_n. They can be satisfied whenever their directed graph is acyclic: assign strictly increasing positive magnitudes in a topological order, and then choose each bias as in the lemma.

The underlying comparison graph is itself an n-cycle. A directed cycle can therefore occur only when every vertex contributes a comparison and all comparisons run consistently around the cycle. There are exactly two such orientations. In each one, every original edge is essential at exactly one endpoint, so it has two independent observable sign choices. These give 2^n distinct tuples per orientation, and the orientations are disjoint. All and only these 2^{n+1} tuples are impossible. Finally, the eigenvalues of A are $\lambda_\pm$, proving (9). $\square$

Combining components, (7) and (9) count every finite simple bipartite graph of maximum degree two. The same cycle argument is valid for odd cycles as symmetric binary threshold networks, but odd cycles are not RBM graphs and are excluded from the theorem's RBM interpretation.

4 Equal degree profiles need not give equal counts

Let $S(a,b,c)$ be the tree with three paths of edge lengths a,b,c attached at one common center. Color the center visible. The two trees $S(4,1,1)$ and $S(3,2,1)$ each have seven vertices and share visible degree list $(1,2,3)$ and hidden degree list $(1,1,2,2)$. Thus both have one-way hidden count $T(1)^2T(2)^2 = 3136$ and reverse count $T(1)T(2)T(3) = 5824$. In contrast,

$$\mathcal{B}(S(4,1,1)) = 3707552, \qquad \mathcal{B}(S(3,2,1)) = 3610352. \tag{10}$$

Here is a reproducible integer calculation. An arm of length ℓ has child message $A^{\ell-1}(2,1)^{\mathsf{T}}$, contributing a root factor $U_\ell + V_\ell z$. The needed pairs are

$$(U_1,V_1) = (4,6), \quad (U_2,V_2) = (46,60), \quad (U_3,V_3) = (484,636), \quad (U_4,V_4) = (5116,6720).$$

Expand the three factors and take the coefficient dot product with $(M(0),M(1),M(2),M(3)) = (2,1,2,9)$ to obtain (10). This distinguishes bidirectional capacity from the degree-profile invariance of the standard one-way count.

5 Exact checks and limitations

There is a second, general counting description. In parameter space $\mathbb{R}^{|V|+|H|+|E|}$, form all hyperplanes

$$b_u + \sum_{e=uv} w_e z_v = 0, \qquad z \in \{0,1\}^{N(u)}.$$

A strict sign pattern is exactly a local-function tuple. Its parameter set is an intersection of open half-spaces, hence convex when nonempty, so the tuples correspond bijectively to chambers. The classical characteristic-polynomial formula of Zaslavsky [6] therefore provides an independent computational route to the finite counts.

The accompanying standard-library Python checks compute exact characteristic polynomials for ten small arrangements, including paths through five vertices, a three-leaf star and the four-cycle. Ranks are computed modulo the prime 2147483647 after verifying that it exceeds a Hadamard bound for every integer minor. Thus these ranks agree with rational ranks. The Whitney sums collectively represent 156690 subsets; their chamber counts agree with the formulas above.

A separate constructive checker enumerates 95546 local-table tuples on nine small graphs. It constructs shared rational weights for every sign-compatible forest tuple and for every feasible cycle tuple, then checks every local binary input, totaling 552634 input checks. For C_4 it finds 12472 sign-compatible tuples, precisely 32 impossible comparison cycles, and 12440 rationally realized tuples. Both checkers give identical reports under Python 3.9.6 and 3.13.5 apart from the runtime label. The odd three-cycle is explicitly a non-RBM diagnostic. These computations supplement the proofs; they do not establish the arbitrary-size theorems by extrapolation.

The note does not count ties, automorphism orbits, multistate inference, joint modes, or general higher-degree cyclic graphs in closed form. It does not supply new exact values of $T(d)$ for arbitrary d. The standard one-way result and the general symmetrization and arrangement frameworks are prior work. A bounded literature comparison did not locate the displayed bidirectional formulas; this is not an absolute-priority claim. The original live AimPL item was unavailable, so the frozen source wording is not treated as evidence for a bidirectional convention.

This unrefereed manuscript was developed with AI assistance in proof construction, literature search, drafting and exact computation. Its arguments were self-audited in the originating research workflow; independent human review and proof-assistant certification are not claimed.

References

[1] American Institute of Mathematics, *Boltzmann machines*, workshop, 17–21 September 2018; problem list, section 3, item 3.3. Workshop page and report; frozen problem transcription in *UnsolvedMath*, version 1.6.0, record AIM-PROBABILITY-0009.

[2] J. A. Aledo, E. Goles, M. Montalva-Medel, P. Montealegre and J. C. Valverde, Symmetrizable Boolean networks, *Information Sciences* **626** (2023), 787–804. doi:10.1016/j.ins.2023.01.082.

[3] M. Cosnard and E. Goles, A characterization of the existence of energies for neural networks, in *ICALP 1995*, Lecture Notes in Computer Science **944**, Springer, 1995, 570–580. Publisher record.

[4] M. A. Cueto, J. Morton and B. Sturmfels, Geometry of the Restricted Boltzmann Machine, *Contemporary Mathematics* **516** (2010), 135–153. arXiv:0908.4425.

[5] G. Montúfar and J. Morton, Discrete Restricted Boltzmann Machines, *Journal of Machine Learning Research* **16**(21) (2015), 653–672. Journal version.

[6] T. Zaslavsky, *Facing Up to Arrangements: Face-Count Formulas for Partitions of Space by Hyperplanes*, Memoirs of the American Mathematical Society **1**, no. 154, 1975. doi:10.1090/memo/0154.