

OpenAI’s proof of the Haldane gap in the $S = 1$ antiferromagnetic Heisenberg chain: Discrete-time version

Hal Tasaki*

We give a self-contained account of the analytic argument in OpenAI’s proof of a uniform spectral gap immediately above the ground-state energy for the $S = 1$ antiferromagnetic Heisenberg chain on all sufficiently large even periodic rings. Remarkably, the criterion for the existence of a gap is expressed solely in terms of the standard partition function of the model. We also explain how the computer-assisted estimates in the original paper establish the finite criterion; the underlying computations are not reproduced or independently checked here.

This is an expository note, not a research article, published on the author’s personal GitHub repository.¹ For completeness, we discuss some extensions not contained in the original OpenAI paper, but make no claim of priority for them.²

This is one of two versions of this expository note. The companion “Continuous-time version” follows the functional-analytic construction of the spatial transfer operator in the original paper. Here we replace that construction by a discrete-time checkerboard expansion and finite transfer matrices. The proof of the analytic criterion uses only undergraduate calculus and linear algebra; the criterion and its conclusions are the same in both versions.

Contents

1	Introduction	2
2	The model	4
3	The partition function and its space-time representation	6
3.1	The Lie-product formula and the checkerboard expansion	6
3.2	Decomposition into bond histories	7
3.3	Site traces and signed history weights	9
3.4	The finite spatial transfer matrix	9
4	An elementary inequality for purity	11
5	Temporal and spatial purities	13
5.1	Temporal purity: doubling the inverse temperature	13
5.2	Spatial purity: doubling the system size	14
5.3	Coupling the two improvements	14
6	The gap proof along dyadic lengths	15
6.1	From a Gibbs-purity bound to a gap	15
6.2	The finite criterion and its iteration	16
7	Discussion	17

*Department of Physics, Gakushuin University, Mejiro, Toshima-ku, Tokyo 171-8588, Japan.

¹The latest versions of both accounts are available at https://haltasaki.github.io/writings/AI/OpenAI_Haldane.html.

²I expect that other researchers who asked suitable questions of sufficiently advanced AI reached these extensions soon after the OpenAI paper became available.

A	General even lengths and other choices of constants	20
A.1	Interpolation between dyadic lengths	20
A.2	Other constants and separate error bounds	21
B	Connection with the finite numerical certificates	22
B.1	The constant energy shift	22
B.2	The auxiliary twists	23
B.3	What is actually computed	24
C	Other spin chains	26
C.1	The antiferromagnetic Heisenberg chain with arbitrary S	26
C.2	The $S = 1$ bilinear-biquadratic chain	27
C.3	The bond-alternating $S = 1$ chain	29
D	A sufficient condition for the converse	30

1 Introduction

The aim of this note is to explain the mathematical argument in OpenAI’s paper [1] in a form that I, and hopefully other readers, can readily understand. The Haldane-gap argument presented here is an exposition of that proof, not a new proof.³

On October 6, 2026, OpenAI released a collection of several hundred mathematical manuscripts reporting results obtained by an internal model [2, 3]. Among them are the paper [1], dated September 24, 2026, on the periodic $S = 1$ Haldane gap, and its companion [4], on chains with endpoint magnetic fields. The present note concerns the former, which proves a standard periodic formulation of the fundamental conjecture proposed by Duncan Haldane in the early 1980s [5, 6, 7]: the $S = 1$ antiferromagnetic Heisenberg chain has a strictly positive lower bound on the energy gap immediately above its ground-state energy, uniform for all sufficiently large even chain lengths. We do not review the physical background of the Haldane conjecture or the related developments here; see, e.g., [8, Part II].

Given that the Haldane conjecture had remained unresolved for over forty years despite considerable efforts by mathematical physicists and mathematicians, including myself, the analytic core of the proof in [1] is astonishingly short and simple; I would also call it elegant. It is also robust and naturally extends to a large class of quantum spin chains. Rather surprisingly, however, the direct extension of the transfer construction does *not* cover the additional positive biquadratic coupling of the $S = 1$ Affleck-Kennedy-Lieb-Tasaki (AKLT) chain [9], for which a gap has long been proved.⁴ The proof establishes the gap, while leaving open the physical question of *why* the $S = 1$ Heisenberg antiferromagnetic chain is gapped. The conditional theorem also applies to models that may be gapless; a gap follows only when its two finite inequalities, involving solely the standard partition function $Z_L(\beta)$, are verified at some size L and inverse temperature β . For the Heisenberg chain, that verification uses the finite computations in the original paper. See also Discussion.

³The finite numerical inputs are taken from [1]. Appendix B explains their role; the underlying computations are not reproduced or independently checked here.

⁴As we shall see in Appendix C.2, the analytic criterion in the OpenAI proof extends to the $S = 1$ bilinear-biquadratic Hamiltonian $H_\mu = \sum_j \{\mathbf{S}_j \cdot \mathbf{S}_{j+1} + \mu(\mathbf{S}_j \cdot \mathbf{S}_{j+1})^2\}$ for $\mu \leq 0$, whereas the AKLT value is $\mu = 1/3$.

OpenAI’s proof starts from a space-time representation of the Gibbs factor of a quantum spin system, a standard idea already present in Ginibre’s work in the 1960s [10, 11]. The particular ordered bond-event formulation used in the proof is presented by Aizenman and Nachtergaele [12]. The other essential ingredient is transfer in the spatial direction, as in the quantum transfer methods developed in work by Araki [13], Suzuki [14], and Betsuyaku [15]. The original paper constructs a self-adjoint spatial transfer operator in continuous time. In this version, an even–odd construction using the same arrangement of bonds as Suzuki’s checkerboard decomposition [16] gives a finite spatial transfer matrix at each discretization parameter N . The continuous-time construction is explained in the companion account [17].

The essential observation is that the partition function can be studied in two directions. The usual expression (3.1) gives the canonical distribution over physical energy eigenvectors. In the spatial direction, the expression (3.22) writes each finite approximation as the sum of the L th powers of the nonnegative singular values of a transfer matrix, and therefore gives another probability distribution. An elementary inequality for its purity survives the limit $N \uparrow \infty$. Coupling the resulting spatial inequality with the corresponding Gibbs-purity inequality shows that both purities approach one when the length and inverse temperature are doubled, provided that they are sufficiently close to one initially. This yields a finite-size, finite-inverse-temperature criterion for a uniform gap expressed solely in terms of the partition function. As we already noted, the criterion alone does not prove a gap for the Heisenberg chain: its hypotheses must be verified for that model. The original paper supplies this final step through certified finite computations.

In the present article, we reorganize the analytic argument to separate its core from the ingredients needed only for the finite computations.⁵ The energy shift a and the auxiliary twists g , for example, are absent from the main proof; their role in the numerical initialization is explained in Appendix B. We explain the discrete space-time representation and the spatial transfer matrix in detail, and write the induction in terms of lower bounds on the purities themselves.⁶ The exposition of the analytic criterion is self-contained apart from elementary finite-dimensional linear algebra and calculus. In particular, we neither construct a limiting transfer operator nor assume convergence of individual spatial eigenvalues. Reading Sections 2–6 gives the core argument, ending with the gap proof along dyadic lengths. The routine extension to all sufficiently large even lengths is deferred to Appendix A. The Discussion addresses the scope and history of the criterion, its relation to topological indices, and the question of a converse. Appendix C treats arbitrary spin S , the bilinear-biquadratic chain, and the bond-alternating chain; Appendix D proves a supplementary sufficient condition for a converse.

How this article was prepared: I began reading [1] on October 7 in Japan, soon after its release, alongside my usual work as a full-time university faculty member. Finding the paper not easy to read, I asked ChatGPT (OpenAI, GPT-6 Astra Pro) for a reading guide and reconstructed the proof myself with pencil and paper, occasionally asking it questions. The formulation of the purity recursion used below was worked out during these exchanges. After working through the analytic argument, I decided on the structure of this note and the presentation of the space-time representation, and prepared handwritten

⁵We use our own notation and organization, although the correspondence with [1] should be straightforward. In particular, the essential quantities $S(L, \beta)$ and $T(L, \beta)$ are defined exactly as in the original paper.

⁶Readers interested in how an AI-generated proof is presented may wish to compare our recursion using h with the original one using f . The former is much easier for me to follow, although the two are mathematically equivalent.

notes and drawings. I then asked ChatGPT to prepare a LaTeX draft from this material, complete the detailed derivations along the lines of our discussion, and prepare Figure 3. I have since been substantially rewriting and checking the draft, with further assistance from ChatGPT. At my request, it also found the supplementary sufficient condition for a converse proved in Appendix D; this is not part of the original Haldane-gap proof. I completed the first continuous-time draft on October 14 and circulated successive revisions to close colleagues for comments before publication. On October 15, I outlined the checkerboard construction in handwritten notes and prepared the new space-time figure for this discrete-time version. I asked ChatGPT to check the derivation, develop the finite-matrix argument, and adapt the rest of the manuscript consistently. The two versions share the same purity iteration and gap criterion and have been revised together. On October 17, 2026, Hidetoshi Nishimori arranged independent AI checks of both versions using Claude Opus 5.5, Claude Fable 5.1, and Claude Opus 5.⁷ The resulting comments helped improve the exposition. I take full responsibility for the statements and exposition in this note.

2 The model

Let $S^{(1)}, S^{(2)}, S^{(3)}$ be the spin angular-momentum operators with quantum number $S = 1$, acting on $\mathbb{C}^3$. Thus they are self-adjoint matrices satisfying

$$[S^{(\alpha)}, S^{(\beta)}] = i \sum_{\gamma=1}^3 \epsilon_{\alpha\beta\gamma} S^{(\gamma)}, \quad \sum_{\alpha=1}^3 (S^{(\alpha)})^2 = 2I_3, \quad (2.1)$$

where $\epsilon_{\alpha\beta\gamma}$ is the Levi-Civita symbol, or, equivalently, the completely antisymmetric tensor with $\epsilon_{123} = 1$. We define

$$G^{(\alpha)} = iS^{(\alpha)} \quad (\alpha = 1, 2, 3). \quad (2.2)$$

Then $\|G^{(\alpha)}\| = 1$. The matrices $G^{(\alpha)}$ need not be real, but traces of their products have the following property. Here tr denotes the ordinary, unnormalized trace of a single-spin operator.

Lemma 2.1 (Reality of single-spin traces) *For every integer $m \geq 0$ and every choice of labels $\alpha_1, \dots, \alpha_m \in \{1, 2, 3\}$, we have*

$$\text{tr}[G^{(\alpha_1)} \dots G^{(\alpha_m)}] \in \mathbb{R}. \quad (2.3)$$

For $m = 0$, the product is understood to be I_3 .

Proof. The trace of an ordered product is unchanged by a common unitary change of the spin basis. We may therefore use any representation of the spin-one operators. Following the original OpenAI paper [1], we choose the Cartesian representation, in which

$$G_{\mu\nu}^{(\alpha)} = \epsilon_{\alpha\mu\nu}. \quad (2.4)$$

Explicitly, these matrices are

$$G^{(1)} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}, \quad G^{(2)} = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad G^{(3)} = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \quad (2.5)$$

⁷He submitted the discrete-time manuscript alone to Claude Opus 5.5 and the continuous-time manuscript alone to Claude Fable 5.1, asking each model to examine the manuscript critically. He then asked Claude Opus 5 to examine the report on the continuous-time version, and Claude Fable 5.1 to examine the report on the discrete-time version. These checks did not include an independent verification of the original numerical certificates.

The matrices $S^{(\alpha)} = -iG^{(\alpha)}$ satisfy (2.1). The corresponding basis is often denoted by p_x, p_y, p_z when this representation is realized by orbital angular momentum. All three $G^{(\alpha)}$ are real in this basis, so every product of them has a real trace. $\square$

Remark. Lemma 2.1 in fact holds for every spin $S = 1/2, 1, 3/2, \dots$. It can be proved by using time-reversal symmetry, without choosing a particular spin basis. Consequently, the transfer construction and the conditional gap theorem extend to every S , with the same purity thresholds. The required bounds change only through the single-spin dimension $2S + 1$ and the norm $\|G^{(\alpha)}\| = S$. Appendix C.1 gives the proof and the details of the extension. Except in Appendix C.1, we concentrate on $S = 1$ throughout this note.

Consider a periodic chain with sites $1, \dots, L$, where $L \geq 4$. Our gap statements concern even L . The physical Hilbert space is $\mathcal{H}_L = (\mathbb{C}^3)^{\otimes L}$. We write $S_j^{(\alpha)}$ and $G_j^{(\alpha)}$ for the corresponding matrices acting on site j , and set $S_{L+1}^{(\alpha)} = S_1^{(\alpha)}$ and $G_{L+1}^{(\alpha)} = G_1^{(\alpha)}$. The Hamiltonian is

$$H_L = \sum_{j=1}^L \sum_{\alpha=1}^3 S_j^{(\alpha)} S_{j+1}^{(\alpha)} = - \sum_{j=1}^L \sum_{\alpha=1}^3 G_j^{(\alpha)} G_{j+1}^{(\alpha)}. \quad (2.6)$$

Let

$$E_0(L) \leq E_1(L) \leq \dots \leq E_{3L-1}(L) \quad (2.7)$$

be the eigenvalues of H_L , counted with multiplicity, and put

$$\gamma_L = E_1(L) - E_0(L). \quad (2.8)$$

With this convention, $\gamma_L = 0$ if the ground energy is degenerate. Counting multiplicities throughout will be important.

For the reader's convenience, we state the main theorem here, leaving the precise finite criterion to Section 6.

Theorem 2.2 (Uniform periodic gap) *Suppose that the hypotheses of Theorem 6.2, including the two finite inequalities (6.6), hold for some even $L_0 \geq 4$ and inverse temperature $\beta_0 > 0$. Assume also the scalar conditions (A.5) in Appendix A. Then there is a constant $\delta > 0$, independent of L , such that*

$$\gamma_L > \delta \quad \text{for every even } L \geq L_0. \quad (2.9)$$

The theorem follows from Theorems 6.2 and A.1; an explicit value of δ is given in (A.6). The finite inequalities involve only partition functions at finitely many sizes and inverse temperatures. The original paper [1] supplies a computer-assisted verification, which we outline in Appendix B. Since the finite inequalities are statements about physical partition functions, their use here does not depend on the representation used to establish them.

With our convention of counting energies with multiplicity, the theorem proves both uniqueness of the ground state and a uniform positive gap immediately above its energy.

3 The partition function and its space-time representation

For $\beta > 0$, the partition function is

$$Z_L(\beta) = \text{Tr}_{\mathcal{H}_L} [e^{-\beta H_L}] = \sum_{k=0}^{3^L-1} e^{-\beta E_k(L)}. \quad (3.1)$$

Throughout this section, $L \geq 4$ is even. We shall construct approximations $Z_{L,N}(\beta)$ using a finite spatial transfer matrix $\mathbf{K}_{\beta,N}$. The eigenvalues of the real symmetric matrix $\mathbf{K}_{\beta,N} \mathbf{K}_{\beta,N}^t$ will be denoted by $\lambda_i(\beta, N)^2$, with $\lambda_i(\beta, N) \geq 0$. The superscript t denotes the transpose. Readers who wish to reach the core argument first may skip the construction, read (3.21) and (3.23) together with the concluding Remark, and then proceed to Section 4.

The original paper [1] and our continuous-time account [17] start from the ordered bond-event expansion presented by Aizenman and Nachtergaele [12]. Here we use an even-odd construction following the arrangement of bonds in Suzuki's *checkerboard* decomposition [16, Section 5(ii), Figure 1(b)], which turns out to be particularly well suited to an elementary reconstruction of OpenAI's proof. At every fixed N , the construction uses only finite sums and finite matrices. We then take $N \uparrow \infty$ at fixed L and β .

3.1 The Lie-product formula and the checkerboard expansion

Write

$$B_{j,\alpha} = G_j^{(\alpha)} G_{j+1}^{(\alpha)}, \quad -H_L = \sum_{j=1}^L \sum_{\alpha=1}^3 B_{j,\alpha}, \quad (3.2)$$

and divide the bonds into the two sets

$$\Lambda_o = \{1, 3, \dots, L-1\}, \quad \Lambda_e = \{2, 4, \dots, L\}. \quad (3.3)$$

Here j labels the bond $(j, j+1)$, with $L+1$ identified with 1. For $\varepsilon > 0$, we define

$$P_o(\varepsilon) = \prod_{j \in \Lambda_o} \left(I + \varepsilon \sum_{\alpha=1}^3 B_{j,\alpha} \right), \quad P_e(\varepsilon) = \prod_{j \in \Lambda_e} \left(I + \varepsilon \sum_{\alpha=1}^3 B_{j,\alpha} \right), \quad (3.4)$$

and

$$W_{L,N}(\beta) = \left\{ P_o\left(\frac{\beta}{N}\right) P_e\left(\frac{\beta}{N}\right) \right\}^N. \quad (3.5)$$

The factors within either product in (3.4) commute because the corresponding bonds are disjoint. The following is a variant of the classical Lie-product formula for finite-dimensional matrices.

Lemma 3.1 *For every fixed even $L \geq 4$ and $\beta > 0$,*

$$\lim_{N \uparrow \infty} W_{L,N}(\beta) = e^{-\beta H_L}. \quad (3.6)$$

Proof. Fix L and β , and put $U_N = P_o(\beta/N) P_e(\beta/N)$ and $V_N = e^{-\beta H_L/N}$. Expanding the finite products gives $U_N = I - \beta H_L/N + O(N^{-2})$, and the same expansion holds for V_N . Hence there

are constants $c_0, c_1 > 0$, independent of sufficiently large N , such that $\|U_N - V_N\| \leq c_0 N^{-2}$ and $\|U_N\|, \|V_N\| \leq 1 + c_1/N$. The telescoping identity

$$U_N^N - V_N^N = \sum_{r=0}^{N-1} U_N^r (U_N - V_N) V_N^{N-1-r} \quad (3.7)$$

therefore gives $\|U_N^N - V_N^N\| \leq (c_0/N)(1 + c_1/N)^{N-1} \leq c_0 e^{c_1}/N$. Since $U_N^N = W_{L,N}(\beta)$ and $V_N^N = e^{-\beta H_L}$, this proves (3.6). The constants may depend on L and β ; no uniformity in either is required. $\square$

Define the approximate partition function

$$Z_{L,N}(\beta) = \text{Tr}_{\mathcal{H}_L} W_{L,N}(\beta), \quad Z_L(\beta) = \lim_{N \uparrow \infty} Z_{L,N}(\beta). \quad (3.8)$$

We now expand the finite product defining $W_{L,N}$. Label its $2N$ successive layers by $s = 1, \dots, 2N$, counting from left to right as

$$W_{L,N}(\beta) = \overbrace{P_o\left(\frac{\beta}{N}\right)}^{s=1} \overbrace{P_e\left(\frac{\beta}{N}\right)}^{s=2} \overbrace{P_o\left(\frac{\beta}{N}\right)}^{s=3} \cdots \overbrace{P_e\left(\frac{\beta}{N}\right)}^{s=2N}. \quad (3.9)$$

Odd layers contain odd bonds, and even layers contain even bonds. The label s records the order of multiplication, not imaginary-time evolution of an individual $B_{j,\alpha}$. One pair of layers approximates evolution by β/N under the full Hamiltonian, so the coefficient in (3.4) is β/N , not $\beta/(2N)$.

At each allowed pair (s, j) , choose either the identity or one of the three terms $(\beta/N)B_{j,\alpha}$. A *space-time history* records the nonidentity choices:

$$\omega = ((s_1, j_1, \alpha_1), \dots, (s_m, j_m, \alpha_m)), \quad 0 \leq m \leq NL. \quad (3.10)$$

The empty history is included. The times obey $1 \leq s_1 \leq \dots \leq s_m \leq 2N$, and s_r and j_r have the same parity. If $s_r = s_{r+1}$, we require $j_r < j_{r+1}$. There is at most one event at each allowed pair (s, j) , and every label α_r lies in $\{1, 2, 3\}$. Several events can occur in the same layer, but they act on disjoint bonds, so ordering them by j introduces no ambiguity. See Figure 1 for an example.

Let $\mathcal{W}_{L,N}$ denote this finite set of histories and let $m(\omega)$ be the number of events. The expansion is exactly

$$W_{L,N}(\beta) = \sum_{\omega \in \mathcal{W}_{L,N}} \left(\frac{\beta}{N}\right)^{m(\omega)} B_{j_1, \alpha_1} \cdots B_{j_m, \alpha_m}, \quad (3.11)$$

with the empty product equal to the identity. There are NL allowed space-time positions, so $|\mathcal{W}_{L,N}| = 4^{NL}$. There are no time integrals or factorials in this finite expansion.

3.2 Decomposition into bond histories

For a fixed bond $(j, j+1)$, collect its events and retain their layer order. The resulting *bond history* has the form

$$x = ((t_1, \alpha_1), \dots, (t_k, \alpha_k)), \quad 1 \leq t_1 < \dots < t_k \leq 2N, \quad \alpha_r \in \{1, 2, 3\}, \quad (3.12)$$

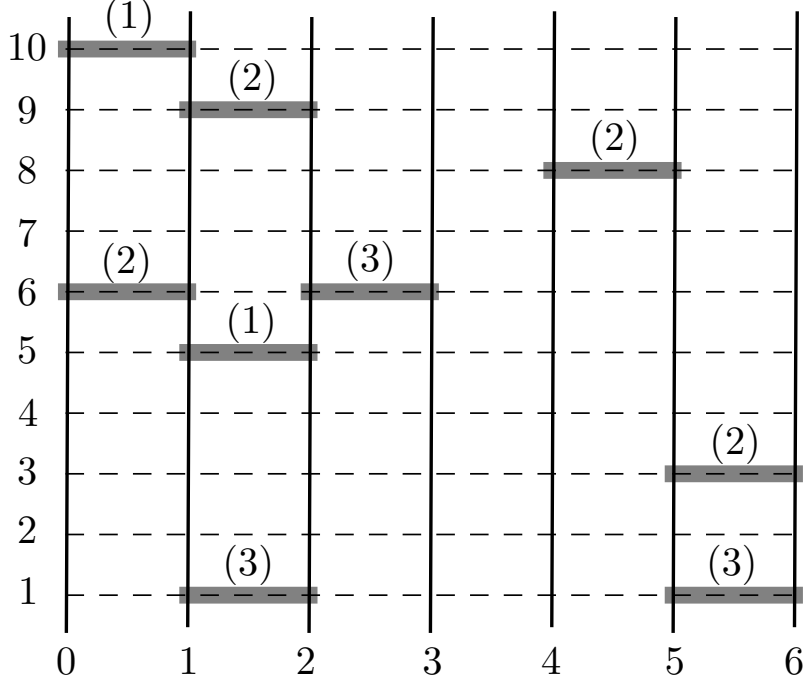


Figure 1: A discrete-time history for $L = 6$ and $N = 5$, with nine events. The vertical labels $1, \dots, 10$ are layer numbers; site 0 is identified with site 6. A segment labeled (α) on bond $(j, j + 1)$ in layer s represents the term $(\beta/N)B_{j,\alpha}$. Every event obeys the checkerboard constraint $s \equiv j \pmod{2}$. In our convention, the complete history is $\omega = ((1, 1, 3), (1, 5, 3), (3, 5, 2), (5, 1, 1), (6, 2, 3), (6, 6, 2), (8, 4, 2), (9, 1, 2), (10, 6, 1))$. Note that the segment drawn on $(0, 1)$ is on bond $j = 6$. For example, the histories on bonds $(1, 2)$ and $(2, 3)$ are $x = ((1, 3), (5, 1), (9, 2))$ and $y = ((6, 3))$; their merge at site 2 has labels $3, 1, 3, 2$ in that order. Hence $k_{\beta,N}(x, y) = (\beta/N)^2 \text{tr}[G^{(3)}G^{(1)}G^{(3)}G^{(2)}]$.

with $0 \leq k \leq N$. All t_r are odd when j is odd, and even when j is even. Let $\mathcal{B}_o$ and $\mathcal{B}_e$ be the sets of allowed bond histories for odd and even bond numbers j , respectively; both include the empty history. Their dependence on N is suppressed in the notation. At each of the N permitted layers there are four choices, so

$$|\mathcal{B}_o| = |\mathcal{B}_e| = 4^N. \quad (3.13)$$

Write x_j for the history on bond $(j, j + 1)$, and set $x_0 = x_L$. Each global history uniquely determines a collection of bond histories satisfying

$$x_j \in \mathcal{B}_o \quad (j = 1, 3, \dots, L - 1), \quad x_j \in \mathcal{B}_e \quad (j = 2, 4, \dots, L). \quad (3.14)$$

Conversely, any such collection determines a unique global history. We attach the bond number j to each entry of x_j , merge the lists, and arrange all entries in increasing order of the layer number s . Within each layer, we arrange the entries in increasing order of the bond number j . Thus the sum over ω is exactly the sum over all independent choices of the bond histories $x_1, \dots, x_L$ satisfying (3.14).

3.3 Site traces and signed history weights

Consider site j . Only the bond histories x_{j-1} and x_j contribute matrices acting there. One of these histories has odd layer numbers and the other has even layer numbers, so they never have coincident events. For $x \in \mathcal{B}_o$ and $y \in \mathcal{B}_e$, merge the two lists to obtain

$$((\tau_1, \delta_1), \dots, (\tau_{k+\ell}, \delta_{k+\ell})), \quad \tau_1 < \dots < \tau_{k+\ell}, \quad (3.15)$$

where $k = |x|$ and $\ell = |y|$ denote the numbers of events. Define

$$k_{\beta,N}(x, y) = \left(\frac{\beta}{N}\right)^{(k+\ell)/2} \text{tr}[G^{(\delta_1)} \dots G^{(\delta_{k+\ell})}], \quad (3.16)$$

where tr is the ordinary, unnormalized trace of a 3-by-3 matrix. Use the same definition when the arguments are given in the opposite order. The factor β/N attached to an event is divided equally between its two endpoints: each contributes $\sqrt{\beta/N}$. This explains the exponent $(k + \ell)/2$ in (3.16). See Figure 1 for an example. For two empty histories, $k_{\beta,N}(\emptyset, \emptyset) = 3$.

Although matrices on the same site need not commute, operators on different sites do commute. We can therefore collect the factors at each site while retaining their original order. Taking the physical trace in (3.11) gives

$$Z_{L,N}(\beta) = \sum_{\substack{x_j \in \mathcal{B}_o \ (j=1,3,\dots,L-1) \\ x_j \in \mathcal{B}_e \ (j=2,4,\dots,L)}} \prod_{j=1}^L k_{\beta,N}(x_{j-1}, x_j), \quad x_0 = x_L. \quad (3.17)$$

Indeed, every bond history occurs in exactly two site factors, so the powers of β/N multiply to $(\beta/N)^{m(\omega)}$. All sums are finite, and no convergence argument is needed for this regrouping.

It is important to stress that the weight $\prod_j k_{\beta,N}(x_{j-1}, x_j)$ need not be nonnegative. Some histories have strictly negative weights.⁸ Dividing these weights by $Z_{L,N}(\beta)$ does not give a probability distribution on histories. In this respect, (3.17) differs from a nonnegative stochastic-geometric representation. What matters for the proof is instead that the site traces are real and unchanged when the two incident histories are exchanged.

3.4 The finite spatial transfer matrix

The definition gives

$$k_{\beta,N}(x, y) \in \mathbb{R}, \quad k_{\beta,N}(x, y) = k_{\beta,N}(y, x). \quad (3.18)$$

Reality follows from Lemma 2.1. Exchanging the two histories does not change their merged list; in particular, no matrix product is reversed.

Let $\mathbf{K}_{\beta,N}$ be the 4^N -by- 4^N matrix with rows indexed by $\mathcal{B}_o$, columns indexed by $\mathcal{B}_e$, and entries

$$(\mathbf{K}_{\beta,N})_{x,y} = k_{\beta,N}(x, y). \quad (3.19)$$

⁸For $L = 4$ and $N \geq 3$, take the history $\omega = ((1, 1, 1), (2, 2, 1), (3, 3, 2), (4, 4, 1), (5, 3, 3))$. Every event satisfies the checkerboard constraint. At sites 1 and 2, the matrix trace is $\text{tr}[(G^{(1)})^2] = -2$. At sites 3 and 4, the chronological products have traces $\text{tr}[G^{(1)}G^{(2)}G^{(3)}] = 1$ and $\text{tr}[G^{(2)}G^{(1)}G^{(3)}] = -1$. Thus the complete weight is $-4(\beta/N)^5 < 0$.

The row and column sets are different. Thus (3.18) does not assert that $\mathbf{K}_{\beta,N}$ is a symmetric matrix after any particular identification of these sets. Rather, the transfer matrix for the opposite parity of site is its transpose. Summing over the even-bond histories gives

$$\mathbf{Q}_{\beta,N} = \mathbf{K}_{\beta,N} \mathbf{K}_{\beta,N}^t, \quad (\mathbf{Q}_{\beta,N})_{x,z} = \sum_{y \in \mathcal{B}_e} k_{\beta,N}(x,y) k_{\beta,N}(z,y), \quad (3.20)$$

with $x, z \in \mathcal{B}_o$. The remaining cyclic sum is precisely a matrix trace:

$$Z_{L,N}(\beta) = \text{Tr}_{\mathcal{H}_L} W_{L,N}(\beta) = \text{Tr}_{\mathcal{B}_o} [\mathbf{Q}_{\beta,N}^{L/2}]. \quad (3.21)$$

The subscript on the last trace means that its matrix indices are the histories in $\mathcal{B}_o$. Each factor $\mathbf{Q}_{\beta,N}$ advances two sites, which accounts for the exponent $L/2$.

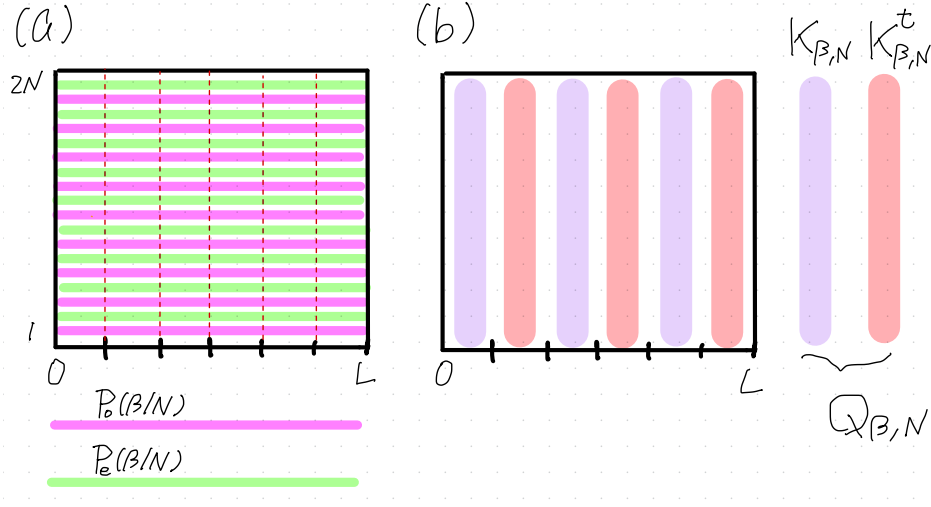


Figure 2: Two directions of transfer in the discrete-time representation. (a) Successive horizontal strips represent $P_o(\beta/N)$ and $P_e(\beta/N)$, acting on the physical spin space; their N pairs give $W_{L,N}(\beta)$. The checkerboard restriction within each strip is displayed explicitly in Figure 1. (b) Successive vertical strips carry $\mathbf{K}_{\beta,N}$ and $\mathbf{K}_{\beta,N}^t$; a pair gives the two-site matrix $\mathbf{Q}_{\beta,N}$. Both directions give exactly $Z_{L,N}(\beta)$ after taking the trace. The vertical coordinate labels layers $1, \dots, 2N$, not a continuous time variable.

The matrix $\mathbf{Q}_{\beta,N}$ is real symmetric and positive semidefinite, since $v^t \mathbf{Q}_{\beta,N} v = \|\mathbf{K}_{\beta,N}^t v\|^2 \geq 0$. It is not the zero matrix because $(\mathbf{K}_{\beta,N})_{\emptyset, \emptyset} = 3$. Let its eigenvalues, including multiplicities and zeros, be $\lambda_i(\beta, N)^2$, with $\lambda_i(\beta, N) \geq 0$ and $i = 1, \dots, 4^N$. Equivalently, the $\lambda_i(\beta, N)$ are the singular values of $\mathbf{K}_{\beta,N}$. Diagonalizing the finite matrix in (3.21) gives

$$Z_{L,N}(\beta) = \sum_{i=1}^{4^N} \lambda_i(\beta, N)^L > 0, \quad L \geq 4 \text{ even}. \quad (3.22)$$

Consequently,

$$Z_L(\beta) = \lim_{N \uparrow \infty} \sum_{i=1}^{4^N} \lambda_i(\beta, N)^L, \quad L \geq 4 \text{ even}. \quad (3.23)$$

For each fixed β, N , the same finite list of singular values is used at every even length L . We do not need convergence of the individual singular values, or a limiting transfer matrix.

Remark. The two traces in (3.21) are ordinary matrix traces on different spaces: the physical spin space and the space indexed by odd bond histories. We call them the *physical trace* and the *history-space trace*. The product $P_o(\beta/N)P_e(\beta/N)$ approximates an imaginary-time step of length β/N , whereas each factor $Q_{\beta,N}$ advances two sites in space. See Figure 2. Although the rows and columns of $Q_{\beta,N}$ are indexed by time histories, $Q_{\beta,N}$ is therefore a *spatial* transfer matrix. This distinction will explain our terms temporal purity and spatial purity.

4 An elementary inequality for purity

We temporarily leave the spin chain and consider any finite or countably infinite probability distribution (p_i) . Thus $p_i \geq 0$ and $\sum_i p_i = 1$. Its *purity*⁹ is

$$P = \sum_i p_i^2. \quad (4.1)$$

Clearly $0 < P \leq 1$. The purity is one precisely when all the probability is concentrated at one entry.

Square the probabilities and normalize them to obtain a new probability distribution

$$\tilde{p}_i = \frac{p_i^2}{P}. \quad (4.2)$$

Let $\tilde{P} = \sum_i \tilde{p}_i^2$ be its purity. The following is Lemma 4.1 of [1], written as a lower bound on the purity itself rather than an upper bound on $1 - \tilde{P}$.

Proposition 4.1 *For the two probability distributions defined above,*

$$\tilde{P} \geq h(P), \quad h(P) = 1 - \frac{(1 - P)^2}{2P^2}. \quad (4.3)$$

Proof. Since $\sum_i \tilde{p}_i = 1$, we have

$$1 - \tilde{P} = \left(\sum_i \tilde{p}_i \right)^2 - \sum_i \tilde{p}_i^2 = 2 \sum_{i < j} \tilde{p}_i \tilde{p}_j = \frac{2}{P^2} \sum_{i < j} p_i^2 p_j^2. \quad (4.4)$$

All the numbers $p_i p_j$ are nonnegative, so the square of their sum is at least the sum of their squares:

$$\sum_{i < j} p_i^2 p_j^2 \leq \left(\sum_{i < j} p_i p_j \right)^2. \quad (4.5)$$

Finally,

$$1 = \left(\sum_i p_i \right)^2 = \sum_i p_i^2 + 2 \sum_{i < j} p_i p_j = P + 2 \sum_{i < j} p_i p_j, \quad (4.6)$$

⁹Following the original paper, we use “purity” for $P = \sum_i p_i^2$. This is also the collision probability, or inverse participation ratio; the participation ratio is $1/P$. The order-two Rényi entropy is $-\log P$. Our argument concerns probabilities and requires no interpretation in terms of quantum entanglement.

and hence $\sum_{i < j} p_i p_j = (1 - P)/2$. Combining these identities and inequalities gives

$$1 - \tilde{P} \leq \frac{2}{P^2} \left(\frac{1 - P}{2} \right)^2 = \frac{(1 - P)^2}{2P^2}, \quad (4.7)$$

which is equivalent to (4.3). For countable distributions, the same calculation is justified by nonnegativity and convergence of the sums. $\square$

The function h is increasing on $(0, 1]$, since

$$h'(P) = \frac{1 - P}{P^3} \geq 0. \quad (4.8)$$

It satisfies $h(1) = 1$ and, for small $u \geq 0$,

$$h(1 - u) = 1 - \frac{u^2}{2} + O(u^3). \quad (4.9)$$

Thus, when the original purity is close to one, its defect is reduced quadratically by normalized squaring. Figure 3 shows h on $[1/2, 1]$. In particular, $h(1/2) = 1/2$ and $h(P) > P$ for $1/2 < P < 1$. For smaller purities the lower bound in (4.3) can be weak or even negative; we will use it only when its arguments are close to one.

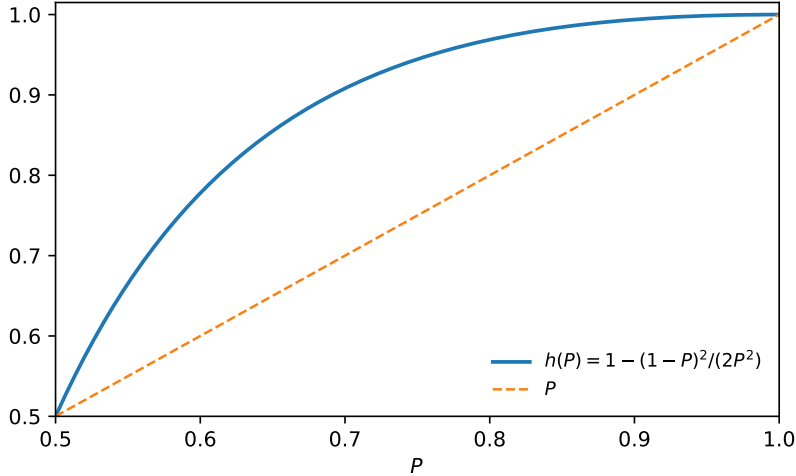


Figure 3: The lower-bound function $h(P) = 1 - (1 - P)^2 / (2P^2)$ and the diagonal P . The bound improves a purity greater than $1/2$, and the derivative of h vanishes at $P = 1$. The latter property is responsible for the quadratic improvement of the purity defect.

We will need the following slightly different estimate. The cube in its argument will arise when we couple three factors of purity.

Lemma 4.2 *There are positive constants u_0 and C such that $Cu_0 < 1$, $C(1 - u_0) > 1$, and*

$$h((1 - u)^3) \geq 1 - Cu^2 \quad (0 \leq u \leq u_0). \quad (4.10)$$

For example, one may choose

$$(u_0, C) = \left(\frac{1}{16}, 8 \right) \quad \text{or} \quad (u_0, C) = \left(\frac{1}{8}, \frac{39}{5} \right). \quad (4.11)$$

Proof. For $(u_0, C) = (1/16, 8)$, note that

$$1 - (1 - u)^3 \leq 3u, \quad (1 - u)^3 \geq 1 - 3u \geq \frac{13}{16} > \frac{3}{4} \quad (4.12)$$

when $0 \leq u \leq 1/16$. Consequently,

$$1 - h((1 - u)^3) = \frac{[1 - (1 - u)^3]^2}{2(1 - u)^6} \leq \frac{9u^2}{2(3/4)^2} = 8u^2. \quad (4.13)$$

To obtain the less restrictive threshold $u_0 = 1/8$, retain the exact coefficient:

$$1 - h((1 - u)^3) = G(u)u^2, \quad G(u) = \frac{1}{2} \left\{ \frac{1}{1 - u} + \frac{1}{(1 - u)^2} + \frac{1}{(1 - u)^3} \right\}^2. \quad (4.14)$$

Indeed, $1 - (1 - u)^3 = u\{1 + (1 - u) + (1 - u)^2\}$. Every term in braces in (4.14) is increasing in u on $[0, 1)$, and therefore G is increasing. For $0 \leq u \leq 1/8$,

$$G(u) \leq G(1/8) = \frac{913952}{117649} < \frac{39}{5}. \quad (4.15)$$

This proves the second estimate. Finally, the two products Cu_0 are $1/2$ and $39/40$, respectively, both strictly less than one. The corresponding products $C(1 - u_0)$ are $15/2$ and $273/40$, both strictly greater than one. $\square$

The first pair has simpler constants, whereas the second requires less accurate initial purities. Nothing else in the proof distinguishes them.

5 Temporal and spatial purities

We now apply Proposition 4.1 to the physical Gibbs distribution and to the finite spatial distributions from the preceding section. For the latter, we pass to the limit only after deriving the probability inequality. Throughout this section, $L \geq 4$ is even and $\beta > 0$.

5.1 Temporal purity: doubling the inverse temperature

The usual Gibbs probabilities, with one probability for each energy eigenvector, are

$$p_k(L, \beta) = \frac{e^{-\beta E_k(L)}}{Z_L(\beta)}, \quad k = 0, \dots, 3^L - 1. \quad (5.1)$$

Their purity is

$$T(L, \beta) := \sum_k p_k(L, \beta)^2 = \frac{Z_L(2\beta)}{Z_L(\beta)^2}. \quad (5.2)$$

We call this the *temporal purity*, or simply the *Gibbs purity*. It is also the ordinary density-matrix purity $\text{Tr}_{\mathcal{H}_L}[\rho_{L,\beta}^2]$, where $\rho_{L,\beta} = e^{-\beta H_L}/Z_L(\beta)$.

The normalized squares of (5.1) are

$$\tilde{p}_k = \frac{p_k(L, \beta)^2}{T(L, \beta)} = \frac{e^{-2\beta E_k(L)}}{Z_L(2\beta)} = p_k(L, 2\beta). \quad (5.3)$$

Thus normalized squaring doubles the inverse temperature and leaves the system size unchanged. Proposition 4.1 gives

$$T(L, 2\beta) \geq h(T(L, \beta)). \quad (5.4)$$

5.2 Spatial purity: doubling the system size

For the spatial direction, first keep N introduced in Section 3.1 finite. The second expression (3.22) for the approximate partition function gives the probabilities

$$q_i(L, \beta, N) = \frac{\lambda_i(\beta, N)^L}{Z_{L,N}(\beta)}, \quad i = 1, \dots, 4^N, \quad \sum_i q_i(L, \beta, N) = 1. \quad (5.5)$$

Their purity is

$$S_N(L, \beta) := \sum_i q_i(L, \beta, N)^2 = \frac{Z_{2L,N}(\beta)}{Z_{L,N}(\beta)^2}. \quad (5.6)$$

The same singular-value list occurs at L and $2L$, since $K_{\beta,N}$ does not depend on the length. There is no claim that the q_i describe physical spin configurations.

The normalized squares are

$$\tilde{q}_i = \frac{q_i(L, \beta, N)^2}{S_N(L, \beta)} = \frac{\lambda_i(\beta, N)^{2L}}{Z_{2L,N}(\beta)} = q_i(2L, \beta, N). \quad (5.7)$$

Thus normalized squaring doubles the system size, leaving both β and N unchanged. Proposition 4.1 gives the exact finite- N inequality

$$S_N(2L, \beta) \geq h(S_N(L, \beta)). \quad (5.8)$$

Now use (3.8) at L , $2L$, and $4L$, all at the same fixed β . Since the limiting partition functions are strictly positive, the ratios converge, and we define

$$S(L, \beta) := \frac{Z_{2L}(\beta)}{Z_L(\beta)^2} = \lim_{N \uparrow \infty} S_N(L, \beta). \quad (5.9)$$

We call this the *spatial purity*, following the continuous-time account; here it is obtained as a limit of finite-distribution purities. In particular, $0 < S(L, \beta) \leq 1$. Continuity of h at this positive value permits passage to the limit in (5.8), giving

$$S(2L, \beta) \geq h(S(L, \beta)). \quad (5.10)$$

Together with the Gibbs-purity bound, we have

$$0 < S(L, \beta) \leq 1, \quad 0 < T(L, \beta) \leq 1. \quad (5.11)$$

These are exactly the functions and inequalities used in the continuous-time proof. From this point onward, all partition functions, purities, and recursion relations are exact: the approximation index N has disappeared.

5.3 Coupling the two improvements

The two purities are expressed in terms of one and the same partition function. This gives exact identities relating an improvement in one direction to quantities in the other direction. First, direct cancellation yields

$$\begin{aligned} S(L, 2\beta) &= S(L, \beta)^2 \frac{T(2L, \beta)}{T(L, \beta)^2} \\ &\geq S(L, \beta)^2 T(2L, \beta). \end{aligned} \quad (5.12)$$

For clarity, the equality on the first line reads

$$\frac{Z_{2L}(2\beta)}{Z_L(2\beta)^2} = \frac{Z_{2L}(\beta)^2}{Z_L(\beta)^4} \frac{Z_{2L}(2\beta)}{Z_{2L}(\beta)^2} \frac{Z_L(\beta)^4}{Z_L(2\beta)^2}. \quad (5.13)$$

The inequality in (5.12) follows because $T(L, \beta) \leq 1$. Similarly,

$$\begin{aligned} T(4L, \beta) &= T(2L, \beta)^2 \frac{S(2L, 2\beta)}{S(2L, \beta)^2} \\ &\geq T(2L, \beta)^2 S(2L, 2\beta), \end{aligned} \quad (5.14)$$

where the equality reads

$$\frac{Z_{4L}(2\beta)}{Z_{4L}(\beta)^2} = \frac{Z_{2L}(2\beta)^2}{Z_{2L}(\beta)^4} \frac{Z_{4L}(2\beta)}{Z_{2L}(2\beta)^2} \frac{Z_{2L}(\beta)^4}{Z_{4L}(\beta)^2}. \quad (5.15)$$

We used $S(2L, \beta) \leq 1$ for the inequality in (5.14).

Apply the spatial squaring inequality (5.10) at inverse temperature 2β . Using (5.12) and the monotonicity of h , we find

$$\begin{aligned} S(2L, 2\beta) &\geq h(S(L, 2\beta)) \\ &\geq h(S(L, \beta)^2 T(2L, \beta)). \end{aligned} \quad (5.16)$$

Next apply temporal squaring (5.4) at length $4L$, and use (5.14), to obtain

$$\begin{aligned} T(4L, 2\beta) &\geq h(T(4L, \beta)) \\ &\geq h(T(2L, \beta)^2 S(2L, 2\beta)). \end{aligned} \quad (5.17)$$

Equations (5.16) and (5.17) are the two recursive inequalities that drive the proof. They are Lemma 4.2 of [1], expressed in terms of the purities themselves.

The order of use matters. We first improve the spatial purity by (5.16). The resulting spatial purity at $(2L, 2\beta)$ is then used in (5.17) to improve the temporal purity.

6 The gap proof along dyadic lengths

We first prove the gap on a sequence of lengths that double at each step. This contains the essential argument. The elementary interpolation needed for general even lengths is deferred to Appendix A.

6.1 From a Gibbs-purity bound to a gap

The following finite-dimensional observation will be used at the end of the iteration.

Lemma 6.1 *Let H be self-adjoint on a finite-dimensional Hilbert space of dimension at least two, and let $E_0 \leq E_1 \leq \dots$ list its eigenvalues with multiplicity. At inverse temperature $\beta > 0$, suppose that the Gibbs purity satisfies*

$$\sum_k p_k^2 \geq 1 - \varepsilon, \quad 0 < \varepsilon < \frac{1}{2}, \quad p_k = \frac{e^{-\beta E_k}}{\sum_\ell e^{-\beta E_\ell}}. \quad (6.1)$$

Then the ground state is unique, and its gap $\gamma = E_1 - E_0$ satisfies

$$e^{-\beta\gamma} \leq \frac{\varepsilon}{1 - \varepsilon}, \quad \gamma \geq \frac{1}{\beta} \log \frac{1 - \varepsilon}{\varepsilon}. \quad (6.2)$$

Proof. The Gibbs probabilities are ordered as $p_0 \geq p_1 \geq \dots$. Since $p_k^2 \leq p_0 p_k$ for every k ,

$$1 - \varepsilon \leq \sum_k p_k^2 \leq p_0 \sum_k p_k = p_0. \quad (6.3)$$

In particular, $p_0 > 1/2$. If the ground energy were degenerate, then $p_1 = p_0 > 1/2$, contradicting $\sum_k p_k = 1$. This proves uniqueness.

By (6.3), the total probability outside the ground state is $1 - p_0 \leq \varepsilon$. Since $p_1 \leq 1 - p_0$, it follows that

$$e^{-\beta\gamma} = \frac{p_1}{p_0} \leq \frac{1 - p_0}{p_0} \leq \frac{\varepsilon}{1 - \varepsilon}. \quad (6.4)$$

There is no degeneracy factor in the first equality: p_1 is the probability of one eigenvector, not the sum over its entire eigenspace. Taking logarithms and then multiplying by $-1/\beta$ reverses the inequality and gives the gap bound in (6.2). $\square$

6.2 The finite criterion and its iteration

For any constants u_0, C satisfying the conditions in Lemma 4.2 and an initial inverse temperature $\beta_0 > 0$, put

$$r = Cu_0 \in (0, 1), \quad \delta = \frac{\log(1/r)}{\beta_0}. \quad (6.5)$$

Theorem 6.2 (Finite criterion along dyadic lengths) *Suppose that there are an even integer $L_0 \geq 4$ and an inverse temperature $\beta_0 > 0$ such that*

$$S(L_0, \beta_0) \geq 1 - u_0, \quad T(2L_0, \beta_0) \geq 1 - u_0, \quad (6.6)$$

where u_0 and C satisfy the conditions in Lemma 4.2. Then, for every $n = 1, 2, \dots$, the Hamiltonian H_{L_n} , with $L_n = 2^n L_0$, has a unique ground state and

$$\gamma_{L_n} > \delta = \frac{\log(1/(Cu_0))}{\beta_0}. \quad (6.7)$$

The two concrete choices of (u_0, C) in (4.11) require initial purities at least $15/16$ and $7/8$, respectively. Their gap bounds are $\log 2/\beta_0$ and $\log(40/39)/\beta_0$. In the second case the criterion is easier to meet, but the resulting lower bound is smaller at a fixed β_0 .

Proof. Set

$$L_n = 2^n L_0, \quad \beta_n = 2^n \beta_0, \quad S_n = S(L_n, \beta_n), \quad T_n = T(2L_n, \beta_n). \quad (6.8)$$

Notice the factor of two in the length used for T_n . The coupled inequalities (5.16) and (5.17) become

$$S_{n+1} \geq h(S_n^2 T_n), \quad T_{n+1} \geq h(T_n^2 S_{n+1}). \quad (6.9)$$

Define the error bounds by

$$u_{n+1} = Cu_n^2. \quad (6.10)$$

Since $Cu_0 = r < 1$, they decrease and stay in $(0, u_0]$. Solving the recursion gives

$$u_n = \frac{1}{C} r^{2^n}. \quad (6.11)$$

We claim that

$$S_n \geq 1 - u_n, \quad T_n \geq 1 - u_n \quad \text{for every } n \geq 0. \quad (6.12)$$

The case $n = 0$ is (6.6). Suppose the claim holds at n . The first inequality in (6.9), the monotonicity of h , and Lemma 4.2 give

$$S_{n+1} \geq h((1 - u_n)^3) \geq 1 - Cu_n^2 = 1 - u_{n+1}. \quad (6.13)$$

In particular, $S_{n+1} \geq 1 - u_n$, since $u_{n+1} \leq u_n$. We can now use the second inequality in (6.9):

$$\begin{aligned} T_{n+1} &\geq h((1 - u_n)^2 S_{n+1}) \\ &\geq h((1 - u_n)^2 (1 - u_{n+1})) \\ &\geq h((1 - u_n)^3) \\ &\geq 1 - Cu_n^2 = 1 - u_{n+1}. \end{aligned} \quad (6.14)$$

This completes the induction. The replacement of u_{n+1} by u_n in the third line is deliberate: it lets the same sequence control both defects. It does not destroy the quadratic improvement. Retaining the improved spatial bound in the temporal update can, however, be useful for the finite numerical initialization; see (A.11) and Appendix B.

It remains to extract the physical gap. At length $L_{n+1} = 2L_n$ and inverse temperature β_n , the Gibbs purity is $T_n \geq 1 - u_n$. The conditions $Cu_0 < 1 < C(1 - u_0)$ imply $u_0 < 1/2$. Since $u_n \leq u_0$, Lemma 6.1 proves uniqueness and gives

$$\begin{aligned} e^{-\beta_n \gamma_{L_{n+1}}} &\leq \frac{u_n}{1 - u_n} = \frac{1}{C(1 - u_n)} r^{2^n} \\ &\leq \frac{1}{C(1 - u_0)} r^{2^n} < r^{2^n}. \end{aligned} \quad (6.15)$$

Taking logarithms and using $\beta_n = 2^n \beta_0$, we obtain

$$\gamma_{L_{n+1}} > \frac{2^n \log(1/r)}{\beta_n} = \frac{\log(1/r)}{\beta_0} = \delta. \quad (6.16)$$

Relabeling $n + 1$ as n completes the proof. $\square$

The essential matching of scales is worth emphasizing. By (6.11) and (6.5),

$$u_n = \frac{1}{C} \exp[-\delta \beta_n]. \quad (6.17)$$

The purity defect is exponentially small in the inverse temperature, not merely small as n increases. That is exactly what gives a positive, length-independent gap after taking the logarithm of a Gibbs-weight ratio.

7 Discussion

We have reviewed the analytic argument in OpenAI's proof of the Haldane gap for the $S = 1$ antiferromagnetic Heisenberg chain [1]. Theorem 6.2 gives a uniform lower bound on the gap at lengths $L_n = 2^n L_0$, $n = 1, 2, \dots$, under the finite-size, finite-inverse-temperature condition (6.6). Appendix A gives the routine extension to all sufficiently large even lengths; for both concrete choices in (4.11), the same bound holds for every even $L \geq L_0$.

The original paper also verifies the finite condition by computer-assisted estimates. The resulting bounds, quoted in (B.14) and (B.15), should be compared with (6.6). Since $S(L, \beta)$ and $T(L, \beta)$ are ratios of partition functions, evaluating the relevant partition functions with sufficient accuracy would in principle suffice. The actual verification is more indirect: short-chain thermal traces and variational bounds from explicit matrix-product states are used to establish the initial purities at larger scales. Appendix B explains how these estimates give the two initial purity bounds.

Together, the analytic criterion and its computer-assisted verification in [1] prove a nonzero uniform gap above the unique ground state on all sufficiently large even periodic chains. Any subsequential local limit of these ground states is then a locally-unique gapped ground state, by the standard argument reviewed in the End Matter of [18] (Theorem 12) and in Section 2.2 of [19]. This does not establish uniqueness among all ground states of the infinite chain.

The analytic core is remarkably general and robust. It uses neither frustration freeness, an explicit ground-state wave function, nor a preliminary proof of decay of correlations. Individual history weights and entries of the spatial transfer matrix need not be nonnegative, as noted after (3.17). What supplies nonnegative spectral weights here is the Gram matrix $K_{\beta, N} K_{\beta, N}^t$, not positivity of those entries. Once the spatial transfer representation (3.23) is established, only the two initial purity bounds are needed to obtain a uniform gap. The criterion therefore applies well beyond the $S = 1$ Heisenberg antiferromagnet.¹⁰ Examples include antiferromagnetic Heisenberg chains with arbitrary spin S in Appendix C.1, the bilinear-biquadratic chain with nonpositive biquadratic coupling in Appendix C.2, and the bond-alternating $S = 1$ chain in Appendix C.3.¹¹

The space-time expansion, the checkerboard decomposition, and the spatial transfer-matrix viewpoint have a long history [10, 11, 12, 13, 14, 15, 16]. Normalized squaring is also a familiar concentration procedure. At my request, ChatGPT located earlier sharp comparisons of power sums [21, 22] that imply the quantitative bound in Proposition 4.1.¹² The essential step in the OpenAI argument is to apply concentration estimates to two distributions associated with the same partition function and to couple their improvements in space and imaginary time. A closely related two-direction argument already appears in Section 2 of OpenAI's $O(4)$ mass-bound paper [23], dated September 23, 2026, one day before the Haldane-gap manuscript. The order of these manuscript dates does not establish the order of discovery.

In hindsight, it is tempting to think that the analytic criterion could have been found

¹⁰At fixed finite size and inverse temperature, the partition function is continuous in the interaction parameters. The strict initial inequalities (B.14) therefore persist in a sufficiently small neighborhood of the isotropic point. Since the same transfer construction applies for positive J_α , this also proves a unique ground state and a uniform gap on all sufficiently large even periodic rings for the anisotropic chain $H_L = \sum_j \sum_{\alpha=1}^3 J_\alpha S_j^{(\alpha)} S_{j+1}^{(\alpha)}$ whenever (J_1, J_2, J_3) is sufficiently close to $(1, 1, 1)$. No explicit size of this neighborhood is asserted here. I find this direct consequence for gap stability remarkable. By contrast, adding a uniform magnetic field does not preserve the reality of the site traces used here, so a direct extension of the present argument does not work. Recall, however, that the $S = 1$ antiferromagnetic Heisenberg chain is strongly believed to remain gapped in a sufficiently weak uniform magnetic field [20].

¹¹These examples do not exhaust the possibilities. Beyond the usual Gibbs representation, it suffices to have approximations of the form (3.23), with a common finite nonnegative list at each β, N for all allowed lengths and convergence to the exact partition functions at each fixed length. An exact spatial representation of the type used in the continuous-time account is another sufficient starting point. In other models, suitable symmetries may provide the required positivity.

¹²For $P = \sum_i p_i^2 \in [1/2, 1]$, these results give the sharp lower bound on $\sum_i p_i^4$ attained by a two-point distribution $(t, 1-t)$. Since $t(1-t) = (1-P)/2$, its fourth power sum is $P^2 - (1-P)^2/2$. For $P \leq 1/2$, Jensen's inequality gives $\tilde{P} \geq P \geq h(P)$. Countable distributions follow by finite approximation.

already in the 1980s. This does not mean that the full proof was then within reach. The numerical certification uses matrix-product states and methods developed in subsequent research on quantum many-body systems, motivated in part by the Haldane problem itself [9, 24].

As stressed in the Introduction, the applicability of the criterion must not be confused with satisfaction of its hypotheses. The distinction is not merely that the application to the Heisenberg chain requires numerical input. The same representation and conditional theorem also apply to gapless models. If their periodic gaps satisfy $\liminf_{L \uparrow \infty, L \text{ even}} \gamma_L = 0$, the two finite inequalities cannot both hold at any admissible pair of scales. A notable example is the antiferromagnetic Heisenberg chain with half-odd-integer S : the finite criterion remains valid, as shown in Appendix C.1, but the Lieb–Schultz–Mattis and Affleck–Lieb theorems imply that $\gamma_L \rightarrow 0$ along even lengths [25, 26]; see also [8, Theorem 6.3]. Further examples where gapless behavior is expected are the bilinear-biquadratic chain (Appendix C.2) at the Takhtajan–Babujian point $\mu = -1$ [27, 28] and the bond-alternating chain at its transition (Appendix C.3).

A related question concerns the converse: if a chain admits the spatial transfer representation and has unique ground states with a uniform gap on all sufficiently large periodic rings, must the finite condition eventually be satisfied for a suitable pair (L_0, β_0) ? Appendix D proves a sufficient condition involving the finite-size dependence of the ground energy. This observation was found by ChatGPT in response to my question; it is not part of the original proof. We do not establish the converse under the gap assumption alone.

Even when verified, the criterion does not distinguish a topologically trivial gapped ground state from a nontrivial one. Determining the topological character of the ground state requires a separate argument. The bond-alternating $S = 1$ chain treated in Appendix C.3 provides a prototypical example. This is why OpenAI’s companion paper [4], which treats open $S = 1$ Heisenberg chains with endpoint magnetic fields, is important. It verifies the boundary-field gap assumption in [18], whose theorem then gives a boundary-selected infinite-volume ground state with twist index -1 ; the index was introduced in [29]. Building on these results, we¹³ have proved in version 2 of [30] that the infinite $S = 1$ antiferromagnetic Heisenberg chain admits a rotationally invariant, locally-unique gapped ground state whose Ogata index for on-site rotations [31, Appendix B] is -1 . This is an astonishingly strong rigorous result in topological condensed matter physics—one I could not have imagined before October 6, 2026.

Finally, I return to the deep (or naive?) question that the OpenAI proof [1] leaves unanswered: why should integer-spin antiferromagnetic Heisenberg chains have unique gapped ground states?¹⁴ The conditional theorem applies to a large class of one-dimensional quantum many-body models, including the gapless half-odd-integer-spin antiferromagnetic Heisenberg chains. Yet the direct extension of the spatial transfer construction does not cover the $S = 1$ AKLT chain [9], whose gap has long been proved and whose valence-bond ground state is believed to share essential features with that of the Heisenberg chain; see footnote 4 and Appendix C.2. After more than forty years of collective work on this problem, I therefore close with a distinction that the proof does not erase: establishing the existence of the gap is not the same as explaining its origin. The fundamental physical mechanism underlying the Haldane gap in integer-spin antiferromagnetic Heisenberg chains remains to be understood.

¹³Here “we” means “Hal Tasaki and ChatGPT Astra”, so the plural is intentional.

¹⁴As shown in Appendix C.1, the analytic criterion applies to every spin S . Its numerical verification is a separate task, expected to become more demanding as the gap decreases and the correlation length grows. The $S = 2$ chain appears to be a plausible target for a further computer-assisted proof, although the feasibility of obtaining the required rigorous finite estimates remains to be investigated. Larger spins would provide a more stringent test of the numerical methods.

Appendices

A General even lengths and other choices of constants

A.1 Interpolation between dyadic lengths

The proof so far controls the physical purities only at lengths $2L_n$. We now fill the intervals between dyadic scales.

For any nonnegative list (a_i) with $0 < \sum_i a_i^p < \infty$, and any $q \geq p > 0$, one has

$$\sum_i a_i^q \leq \left(\sum_i a_i^p \right)^{q/p}. \quad (\text{A.1})$$

To see this, set $w_i = a_i^p / \sum_j a_j^p$. Then (w_i) is a probability distribution and $\sum_i w_i^{q/p} \leq \sum_i w_i = 1$. This proves (A.1) without a dimension-dependent constant.

Let ℓ and L be even and satisfy $\ell \leq L \leq 2\ell$. At fixed N , apply (A.1) to $a_i = \lambda_i(\beta, N)$, with $p = \ell$ and $q = L$. Using (3.22), we get $Z_{L,N}(\beta) \leq Z_{\ell,N}(\beta)^{L/\ell}$. Taking $N \uparrow \infty$ at these fixed lengths gives

$$Z_L(\beta) \leq Z_\ell(\beta)^{L/\ell}, \quad (\text{A.2})$$

by (3.8). At inverse temperature 2β , apply the same finite- N argument with $p = L$ and $q = 2\ell$, and then take $N \uparrow \infty$. This gives $Z_{2\ell}(2\beta) \leq Z_L(2\beta)^{2\ell/L}$, or equivalently

$$Z_L(2\beta) \geq Z_{2\ell}(2\beta)^{L/(2\ell)}. \quad (\text{A.3})$$

The directions differ because L is the larger exponent in (A.2) and the smaller exponent in (A.3), not because the temperatures differ.

Combining the two estimates, we get

$$T(L, \beta) = \frac{Z_L(2\beta)}{Z_L(\beta)^2} \geq \frac{Z_{2\ell}(2\beta)^{L/(2\ell)}}{Z_\ell(\beta)^{2L/\ell}} = \{T(2\ell, \beta)S(\ell, \beta)^2\}^{L/(2\ell)}. \quad (\text{A.4})$$

The last equality follows by cancelling the factors of $Z_{2\ell}(\beta)$ in the definitions (5.2) and (5.9) of the purities.

For either of the pairs in Lemma 4.2, we also have

$$u_0 < \frac{1}{6}, \quad C(1 - 3u_0) > 3. \quad (\text{A.5})$$

The latter product is $13/2$ for the first pair and $39/8$ for the second. These slightly stronger conditions allow us to absorb the loss in the interpolation estimate.

Theorem A.1 (Uniform periodic gap) *Under the hypotheses of Theorem 6.2, assume also (A.5). Then H_L has a unique ground state and*

$$\gamma_L > \delta = \frac{\log(1/(Cu_0))}{\beta_0} \quad (\text{A.6})$$

for every even $L \geq L_0$.

Proof. Choose $n \geq 0$ so that $L_n \leq L \leq 2L_n$. Use (A.4) with $\ell = L_n$ and $\beta = \beta_n$. The bracket on its right is in $(0, 1]$ and the exponent $L/(2L_n)$ is at most one. Therefore, by (6.12),

$$T(L, \beta_n) \geq \{T_n S_n^2\}^{L/(2L_n)} \geq T_n S_n^2 \geq (1 - u_n)^3 \geq 1 - 3u_n. \quad (\text{A.7})$$

Here the final inequality follows by expanding the cube. Since $3u_n \leq 3u_0 < 1/2$, Lemma 6.1 gives a unique ground state and

$$e^{-\beta_n \gamma_L} \leq \frac{3u_n}{1 - 3u_n} = \frac{3}{C(1 - 3u_n)} r^{2^n} \leq \frac{3}{C(1 - 3u_0)} r^{2^n} < r^{2^n}. \quad (\text{A.8})$$

We used $C(1 - 3u_0) > 3$ from (A.5) in the last step. Taking logarithms proves (A.6), just as in (6.16). The intervals $[L_n, 2L_n]$ cover all lengths $L \geq L_0$. For a fixed L , we use one appropriate value of n ; we do not take an inverse-temperature limit at fixed L . $\square$

A.2 Other constants and separate error bounds

The two pairs in Lemma 4.2 are convenient examples, not essential features of the argument. For comparison with the original paper, we record its general form.

Corollary A.2 *Theorems 6.2 and A.1 remain valid for any u_0, C satisfying*

$$0 < u_0 < \frac{1}{6}, \quad G(u_0) \leq C, \quad r := Cu_0 < 1, \quad C(1 - 3u_0) > 3, \quad (\text{A.9})$$

where G is defined in (4.14).

Proof. Monotonicity of G gives $h((1 - u)^3) \geq 1 - Cu^2$ for $0 \leq u \leq u_0$. The condition $Cu_0 < 1$ makes the sequence (6.10) decrease. The other two conditions ensure that the interpolated purity is greater than $1/2$ and that the final prefactor in (A.8) is strictly less than one. These are all the scalar properties used in the preceding proofs. $\square$

These are precisely the scalar hypotheses in Proposition 4.3 of [1]. In particular, $u_0 < 1/6$ alone is not an admissible-error threshold. One must also have $u_0 G(u_0) < 1$ in order to choose C with $G(u_0) \leq C$ and $Cu_0 < 1$. For example,

$$1 - h((1 - 1/6)^3) = \frac{8281}{31250} > \frac{1}{6}, \quad (\text{A.10})$$

so a common-error contraction certainly does not start at $u_0 = 1/6$. The original paper's first initialization uses $u_0 = 1/8$. Our simpler choice $1/16$ only makes the proof of the scalar estimate shorter.

When preparing the initial values numerically, it can be useful to retain the improved spatial bound in the temporal update, rather than make the common-error replacement in (6.14). For lower bounds $s, t \in (0, 1]$, one may retain the successive updates

$$s_+ = h(s^2 t), \quad t_+ = h(t^2 s_+), \quad (\text{A.11})$$

as long as the lower bounds being used remain positive. For example, starting from $s = t = 1 - u$ gives

$$1 - s_+ = \frac{9}{2}u^2 + O(u^3), \quad 1 - t_+ = 2u^2 + O(u^3). \quad (\text{A.12})$$

Thus the second defect can be smaller than the first. Retaining that improvement can help the next finite step. This is the reason for the separate, rounded updates in the original paper's numerical initialization. For indefinite continuation, the common sequence u_n , as used in the present account, is simpler and sufficient.

B Connection with the finite numerical certificates

For this appendix only, we use the continuous-time notation of the original paper, as explained in Section 3 of the companion account [17]. In particular, K_β acts on the history space $\mathcal{K}_\beta = L^2(\Omega_\beta, \mu_\beta; \mathbb{C})$ defined there. Its real, possibly negative eigenvalues are distinct from the nonnegative singular values $\lambda_i(\beta, N)$ of our finite checkerboard matrices. The remainder of this appendix is identical to Appendix B of that account and explains the numerical inputs taken from the original paper. The continuous-time account is used here to explain notation, not as an additional premise of the finite-matrix criterion. The original certification also uses odd-length power sums, which are not supplied by our checkerboard construction.

The main text and Appendix A give a complete analytic proof of a sufficient criterion. What remains for the particular Hamiltonian (2.6) is to prove that the two finite inequalities (6.6) hold somewhere. In this appendix we outline how the original paper does this, and quote the resulting inputs. We do not reproduce or independently check the computer-assisted certificates. All references to sections, lemmas, and appendices of the original paper below refer to [1].

B.1 The constant energy shift

The original paper uses the shifted partition function

$$\widehat{Z}_L(\beta) = \text{Tr}_{\mathcal{H}_L}[e^{-\beta(H_L + aLI)}] = e^{-a\beta L} Z_L(\beta), \quad a = \frac{700741}{500000} = 1.401482, \quad (\text{B.1})$$

The seemingly mysterious value of a is a convenient normalization: $-a$ is close to the numerically estimated ground-state energy density and lies just above the certified variational energy densities used in the original paper. No exact value of the infinite-volume energy density is needed. Its spatial transfer operator is

$$X_\beta = e^{-a\beta} K_\beta. \quad (\text{B.2})$$

Both purities are unchanged:

$$\frac{\widehat{Z}_{2L}(\beta)}{\widehat{Z}_L(\beta)^2} = S(L, \beta), \quad \frac{\widehat{Z}_L(2\beta)}{\widehat{Z}_L(\beta)^2} = T(L, \beta). \quad (\text{B.3})$$

Thus the functions S and T in this note really are the same as those in the original article, despite our omission of the shift in the main proof.

The normalization is useful for the finite estimates. The original paper's explicit variational vectors yield

$$E_0(72) + 72a < 0, \quad E_0(120) + 120a < 0 \quad (\text{B.4})$$

(Lemma C.2). Hence one physical Boltzmann factor in each of $\widehat{Z}_{72}(\beta)$ and $\widehat{Z}_{120}(\beta)$ exceeds one, for every $\beta > 0$. This provides a convenient lower bound to combine with upper bounds obtained from short-chain partition functions. The printed thermal values, spectral filters, and variational comparisons are matched to this normalization. Changing a does not change the purities, but one must rescale all those auxiliary estimates consistently; one cannot simply change a while leaving the printed numerical inequalities untouched.

B.2 The auxiliary twists

The auxiliary twists are another ingredient of the original paper that does not occur in our main proof. They are ordinary spin rotations inserted at the closing bond of a periodic chain. Their purpose is to separate symmetry sectors of the spatial transfer operator and thereby improve the finite estimates.

For the symmetry calculations in this subsection, choose the Cartesian representation (2.5), and denote its basis vectors by e_1, e_2, e_3 . Only two nonidentity rotations are needed in the finite calculations:

$$R_2 = \text{diag}(1, -1, -1), \quad R_3 : e_1 \mapsto e_2, \quad e_2 \mapsto e_3, \quad e_3 \mapsto e_1. \quad (\text{B.5})$$

The first is a rotation by π about the first axis; the second is a rotation by $2\pi/3$ about the $(1, 1, 1)$ axis. They are denoted by P and C in the original paper. We use R_2, R_3 to avoid confusion with the purity P and the scalar constant C . The choice $g = I_3$ means no twist. In the Cartesian spin-one representation, each of these real orthogonal matrices is also the unitary matrix implementing that spin rotation.

For $g = I_3, R_2, R_3$, define

$$H_L(g) = \sum_{j=1}^{L-1} \sum_{\alpha=1}^3 S_j^{(\alpha)} S_{j+1}^{(\alpha)} + \sum_{\alpha=1}^3 S_L^{(\alpha)} (g S^{(\alpha)} g^{-1})_1. \quad (\text{B.6})$$

Only the interaction between sites L and 1 has changed. For example, $g = R_2$ replaces this bond by $S_L^{(1)} S_1^{(1)} - S_L^{(2)} S_1^{(2)} - S_L^{(3)} S_1^{(3)}$. The origin in the original paper is site 1 in our labeling.

To describe the symmetry decomposition, it is convenient to allow g to be any proper signed permutation matrix. This means that g permutes the Cartesian basis vectors, possibly changing their signs, and has determinant one. Write

$$g e_\alpha = \epsilon_g(\alpha) e_{\sigma_g(\alpha)}, \quad \epsilon_g(\alpha) \in \{-1, 1\}. \quad (\text{B.7})$$

The same twisted Hamiltonian is defined by (B.6) for every such g . On a history $x = ((t_1, \alpha_1), \dots, (t_k, \alpha_k))$, let $\sigma_g x$ denote the history obtained by replacing each label α_r by $\sigma_g(\alpha_r)$, without changing its time. Set

$$s_g(x) = \prod_{r=1}^k \epsilon_g(\alpha_r), \quad (U_g f)(x) = s_g(\sigma_g^{-1} x) f(\sigma_g^{-1} x). \quad (\text{B.8})$$

The empty product is one. Since label permutations preserve μ_β , U_g is unitary on the history space $\mathcal{K}_\beta$. Rotation covariance of the single-site trace shows that U_g commutes with K_β , and hence with $X_\beta = e^{-a\beta} K_\beta$. The ordered expansion at the twisted bond gives

$$\hat{Z}_L(\beta, g) := \text{Tr}_{\mathcal{H}_L} [e^{-\beta(H_L(g) + aLI)}] = \text{Tr}_{\mathcal{K}_\beta} [X_\beta^L U_g] \quad (\text{B.9})$$

(Proposition 3.1 of the original paper).

The eigenvalues discussed in the rest of this subsection are those of the *shifted spatial operator* X_β , acting on $\mathcal{K}_\beta$. They are the eigenvalues $\lambda_i(\beta)$ of K_β multiplied by $e^{-a\beta}$, not physical energies of H_L . The symmetry decomposition in Lemma 3.2 of the original paper gives three real eigenvalue lists, say (ι_i) , (o_i) , and (ν_i) , for restrictions of X_β to invariant

subspaces. Counted with multiplicity, the full eigenvalue list of X_β consists of one copy of (ι_i) , two copies of (o_i) , and three copies of (ν_i) . Put

$$I_m = \sum_i \iota_i^m, \quad O_m = \sum_i o_i^m, \quad N_m = \sum_i \nu_i^m. \quad (\text{B.10})$$

These sums converge absolutely for $m \geq 2$; the inverse temperature is suppressed from the notation. The trace identities are

$$\begin{aligned} \widehat{Z}_m(\beta) &= I_m + 2O_m + 3N_m, \\ \widehat{Z}_m(\beta, R_2) &= I_m + 2O_m - N_m, \\ \widehat{Z}_m(\beta, R_3) &= I_m - O_m \quad (m \geq 4). \end{aligned} \quad (\text{B.11})$$

Thus the three finite thermal traces determine the power sums in the separate symmetry sectors. For even m , all three power sums are nonnegative. Their bounds help isolate one dominant eigenvalue of X_β and control the contributions of all the others, which is the information needed for the spatial purity. Every physical trace on the left still includes the full physical Hilbert space; the sector decomposition is made only on the auxiliary side.

For example, let (ξ_i) be any one of these lists of real eigenvalues of X_β . For a polynomial $F(x) = \sum_{m \geq 4} f_m x^m$ that is nonnegative on $\mathbb{R}$, absolute convergence gives

$$\sum_i F(\xi_i) = \sum_m f_m \sum_i \xi_i^m. \quad (\text{B.12})$$

Bounds on the power sums give an upper bound on this nonnegative sum and therefore exclude an eigenvalue wherever F exceeds that bound. The polynomial F is general in this observation, but the actual calculation uses specific polynomials chosen to make the resulting bounds useful. For example, the proof of Proposition 5.5 in [1, Section 5.3] gives two explicit degree-twelve polynomials, F_J and F_N , each a product of x^4 and a squared polynomial. Their coefficients are chosen so that the available short-chain traces bound the right-hand side of (B.12), while the polynomials are large on the spectral regions to be excluded. The original paper combines these filters with estimates of higher power sums and the variational bounds (B.4). The first initialization uses R_2 and R_3 ; the second uses only R_2 , in both cases together with the identity. The twists are auxiliary tools for certifying the initial purities, not modifications of the Hamiltonian in the final gap theorem.

B.3 What is actually computed

The physical systems used for the thermal calculations are small. Appendix A of [1] gives rigorous enclosures of the shifted thermal traces for lengths 6, 8, 10, at inverse temperature $21/2$, with g equal to the identity or either of the two twists. Appendix B of that paper uses lengths 4, 5, $\dots$, 12, at inverse temperature $49/4$, with the identity and the half-turn. Odd lengths in this second table supply odd power sums of the spatial eigenvalues for the polynomial filters; these sums are not used to define probability distributions.

The matrix-exponential calculations are accompanied by explicit truncation and rounding bounds. The original algorithms use integer or rational arithmetic to enclose the traces, rather than taking floating-point eigenvalue estimates as premises. The variational inputs in Appendix C of [1] come from explicit periodic matrix-product states, represented by unnormalized vectors with auxiliary dimension 26, at lengths 72 and 120. Their norms and energy

expectations are computed by finite integer matrix contractions. Only an energy upper bound is needed from these trial vectors; no assumption that a trial vector is an exact ground state enters the analytic criterion.

We now describe how these finite calculations lead to the two initial purity bounds. For the first initialization, Section 5.2 of [1] starts at $\beta_* = 21/2$. The sector bounds and polynomial filters give upper bounds on the shifted partition functions at lengths 30 and 120, including the required twists at length 30. The inverse temperature is then increased to $\beta_0 = (5/2)\beta_* = 105/4$. This step uses the physical Boltzmann factors: for each fixed length and twist, writing them as $z_k > 0$ at β_* gives factors $z_k^{5/2}$ at β_0 , and $\sum_k z_k^{5/2} \leq (\sum_k z_k)^{5/2}$. It does not assume that the spatial purity is monotone in inverse temperature.

At β_0 , write (ξ_i) for the full eigenvalue list of X_{β_0} and set $w_i = |\xi_i|^{30}$. The resulting upper bounds, together with $\widehat{Z}_{120}(\beta_0) > 1$ from (B.4), isolate one weight $w_0 > 0.998$ and bound the remainder by $\sum_{i \neq 0} w_i^2 \leq 0.068404$. The multiplicities of the three sector lists are used to sharpen this remainder bound. Since $S(60, \beta_0) = \sum_i w_i^4 / (\sum_i w_i^2)^2$, retaining w_0^4 in the numerator yields

$$S(60, \beta_0) > \left(1 + \frac{0.068404}{0.998^2}\right)^{-2} > \frac{7}{8}. \quad (\text{B.13})$$

For the temporal purity, the length-120 estimate at β_* and its variational ground-energy bound are used separately. After raising the physical Boltzmann factors to the power 5/2, their total weight apart from one largest factor, divided by that factor, is less than 0.0684. Its normalized probability therefore exceeds $(1 + 0.0684)^{-1}$, and $T(120, \beta_0) > (1 + 0.0684)^{-2} > 7/8$.

For the second initialization, Section 5.3 of [1] works first at $\beta_* = 49/4$. The degree-twelve filters and the variational bound at length 72 isolate a spatial eigenvalue of modulus greater than 0.999 and bound the remaining power sums. Retaining the dominant spatial contribution gives $1 - S(36, 49/4) < 0.047916$; combining an upper bound on $\widehat{Z}_{72}(49/4)$ with $\widehat{Z}_{72}(49/2) > 1$ gives $1 - T(72, 49/4) < 0.181521$. These preliminary bounds do not yet meet the convenient common-defect criterion. The original paper applies the separate updates (A.11) six times, using exact rational initial bounds and rounding each new defect upward to a multiple of 10^{-7} . Both scales are doubled at each step, giving $36 \cdot 2^6 = 2304$ and $(49/4)2^6 = 784$. The final spatial and temporal defects are at most 0.0084513 and 0.0054986, respectively, both less than 1/100.

Thus the original calculations give the following two finite initializations:

$$S(60, 105/4) > \frac{7}{8}, \quad T(120, 105/4) > \frac{7}{8}, \quad (\text{B.14})$$

$$S(2304, 784) \geq 1 - \frac{1}{100}, \quad T(4608, 784) \geq 1 - \frac{1}{100}. \quad (\text{B.15})$$

These are Propositions 5.4 and 5.5 of the original paper. Neither pair requires a full physical matrix at the large lengths appearing in these inequalities. The inequalities (B.14) and (B.15) are the numerical inputs taken from the original paper.

Using Corollary A.2, the original choices of parameters are

L_0	β_0	u_0	C	$r = Cu_0$
60	105/4	1/8	79/10	79/80
2304	784	1/100	5	1/20

(B.16)

The first row is admissible because (4.15) also gives $G(1/8) < 79/10$, and $(79/10)(1 - 3/8) = 79/16 > 3$. For the second row, $1 - (1 - u)^3 \leq 3u$ and $(1 - u)^3 \geq 1 - 3u$ imply

$$G(u) \leq \frac{9}{2(1 - 3u)^2} \leq \frac{45000}{9409} < 5 \quad (0 \leq u \leq 1/100), \quad (\text{B.17})$$

and $5(1 - 3/100) > 3$. Both rows have $r < 1$. The first row uses the original paper's $79/10$ rather than our slightly smaller convenient constant $39/5$, so that the following bounds agree with the printed theorem.

Substituting the finite inputs (B.14) and (B.15) into the criterion proved here recovers the original conclusions:

$$\begin{aligned} \gamma_L &> \frac{4}{105} \log \frac{80}{79} \quad \text{for every even } L \geq 60, \\ \gamma_L &> \frac{\log 20}{784} \quad \text{for every even } L \geq 2304, \end{aligned} \quad (\text{B.18})$$

with a unique ground state at each of these lengths. Since the ground state is unique at these lengths, the gap in (B.18) has the same meaning here as in the original paper.¹⁵ This completes the connection between the analytic criterion explained here and its computer-assisted application to the Heisenberg chain in the original paper.

C Other spin chains

C.1 The antiferromagnetic Heisenberg chain with arbitrary S

We prove the extension stated in the Remark after Lemma 2.1. In particular, we give a time-reversal proof of the reality of the single-site traces for every spin $S = 1/2, 1, 3/2, \dots$

For every spin S , time reversal is represented by an antiunitary operator Θ satisfying

$$\Theta^{-1} S^{(\alpha)} \Theta = -S^{(\alpha)} \quad (\alpha = 1, 2, 3). \quad (\text{C.1})$$

These are standard properties; see, e.g., [8, Section 2.3 and Appendix A.4.3] for definitions and details. Since Θ is antilinear, it conjugates scalar coefficients, and hence

$$\Theta^{-1} G^{(\alpha)} \Theta = (-i)(-S^{(\alpha)}) = G^{(\alpha)}. \quad (\text{C.2})$$

Conjugation by Θ preserves the order of multiplication: $\Theta^{-1}(AB)\Theta = (\Theta^{-1}A\Theta)(\Theta^{-1}B\Theta)$. Thus it fixes every ordered product $A = G^{(\alpha_1)} \dots G^{(\alpha_m)}$. On the other hand, for any linear operator A , antiunitarity gives

$$\text{tr}(\Theta^{-1}A\Theta) = \overline{\text{tr } A}. \quad (\text{C.3})$$

This follows by evaluating the trace in an orthonormal basis and noting that Θ maps it to another orthonormal basis. We conclude that

$$\text{tr}[G^{(\alpha_1)} \dots G^{(\alpha_m)}] \in \mathbb{R} \quad (m \geq 0). \quad (\text{C.4})$$

Although $\Theta^2 = (-1)^{2S}I$, the sign of Θ^2 was not used anywhere in this argument. For $S = 1$, this is an alternative proof of Lemma 2.1, without choosing any particular spin basis.

¹⁵The original paper defines the gap above the entire ground-state eigenspace, using the next distinct energy level. We instead list energies with multiplicity in (2.7), so our γ_L is zero when the ground energy is degenerate. The conventions coincide whenever the ground state is unique, as both arguments establish at the lengths in (B.18).

Let $H_L^{(S)} = \sum_{j=1}^L \sum_{\alpha=1}^3 S_j^{(\alpha)} S_{j+1}^{(\alpha)}$ be the periodic spin- S Hamiltonian, with $S_{L+1}^{(\alpha)} = S_1^{(\alpha)}$, and set $Z_L^{(S)}(\beta) = \text{Tr} e^{-\beta H_L^{(S)}}$. The bond identity $-\sum_{\alpha} S^{(\alpha)} \otimes S^{(\alpha)} = \sum_{\alpha} G^{(\alpha)} \otimes G^{(\alpha)}$ is unchanged. Thus the same history expansion applies, with traces taken over $d = 2S + 1$ dimensions. Equation (C.4) makes every site trace real, and exchanging the two incident histories still leaves their merged list unchanged.

For the discrete-time construction, use the same history sets $\mathcal{B}_o$ and $\mathcal{B}_e$ as in Section 3. Each set still has 4^N elements: the three event labels have not changed. Replace the single-site trace in (3.16) by the d -dimensional trace, and denote the resulting real matrix by $\mathbf{K}_{\beta,N}^{(S)}$. Then

$$\mathbf{Q}_{\beta,N}^{(S)} = \mathbf{K}_{\beta,N}^{(S)} (\mathbf{K}_{\beta,N}^{(S)})^t \geq 0. \quad (\text{C.5})$$

The Lie-product argument applies at each fixed L and β , since all the spin matrices are finite dimensional. The same regrouping of the finite sum gives

$$Z_{L,N}^{(S)}(\beta) = \text{Tr}[(\mathbf{Q}_{\beta,N}^{(S)})^{L/2}] = \sum_{i=1}^{4^N} \lambda_i^{(S)}(\beta, N)^L, \quad Z_L^{(S)}(\beta) = \lim_{N \uparrow \infty} Z_{L,N}^{(S)}(\beta), \quad (\text{C.6})$$

for even $L \geq 4$, where $\lambda_i^{(S)}(\beta, N) \geq 0$ are the singular values of $\mathbf{K}_{\beta,N}^{(S)}$. At fixed β, N , the same singular-value list is used for every allowed L .

Define the temporal and spatial purities by (5.2) and (5.9), using $Z_L^{(S)}$ in place of Z_L . The inequalities in Section 5, the recursion in Section 6, and the interpolation in Appendix A now apply without any change of constants. This proves the extension of the conditional criterion, not the satisfaction of its hypotheses. For half-odd-integer S , the Lieb–Schultz–Mattis and Affleck–Lieb theorems [25, 26] imply that the periodic gap tends to zero as $L \uparrow \infty$ through even values; see [8, Theorem 6.3]. Hence the two initial inequalities cannot both hold at any admissible finite pair of scales in that case.

C.2 The $S = 1$ bilinear-biquadratic chain

Consider the periodic Hamiltonian

$$H_L^{\text{bb}}(\mu) = \sum_{j=1}^L \{Q_j + \mu Q_j^2\}, \quad Q_j = \mathbf{S}_j \cdot \mathbf{S}_{j+1}, \quad \mu \in \mathbb{R}. \quad (\text{C.7})$$

The Heisenberg interaction corresponds to $\mu = 0$. The AKLT interaction has $\mu = 1/3$ [9]. We explain why the direct extension of the spatial transfer construction in [1] works for $\mu \leq 0$, but not for this positive coefficient.

For the calculations in this subsection, we use the Cartesian representation (2.5). On two sites, write $Q = \mathbf{S}_1 \cdot \mathbf{S}_2$ and $v_\mu = Q + \mu Q^2$. Since $S^{(\alpha)} = -iG^{(\alpha)}$, we have

$$-Q = \sum_{\alpha=1}^3 G^{(\alpha)} \otimes G^{(\alpha)}, \quad Q^2 = \sum_{\alpha,\eta=1}^3 (G^{(\alpha)} G^{(\eta)}) \otimes (G^{(\alpha)} G^{(\eta)}). \quad (\text{C.8})$$

The bond operator inserted in the expansion of $e^{-\beta H_L^{\text{bb}}(\mu)}$ is therefore

$$-v_\mu = \sum_{\alpha=1}^3 G^{(\alpha)} \otimes G^{(\alpha)} - \mu \sum_{\alpha,\eta=1}^3 (G^{(\alpha)} G^{(\eta)}) \otimes (G^{(\alpha)} G^{(\eta)}). \quad (\text{C.9})$$

The minus sign multiplying the biquadratic term is the essential point.

For $\mu \leq 0$, define

$$A^{(\alpha,\eta)} = \sqrt{-\mu} G^{(\alpha)} G^{(\eta)}. \quad (\text{C.10})$$

Then (C.9) is a sum of identical tensor factors, all of which are real: the three matrices $G^{(\alpha)}$ and the nine matrices $A^{(\alpha,\eta)}$. Use these twelve event types in the checkerboard expansion, with $I - \varepsilon v_\mu$ on each bond and $\varepsilon = \beta/N$. There are thirteen choices at each allowed bond-layer position, including the identity, so each bond-history set now has 13^N elements. The site matrix entry is $\varepsilon^{|x|+|y|/2}$ times the trace of the merged product of event matrices, exactly as in (3.16). All entries are real, and interchanging the two incident histories leaves the merged product unchanged. The matrices $A^{(\alpha,\eta)}$ need not be antisymmetric; their reality is enough.

Denote the matrix with odd-history rows and even-history columns by $\mathsf{K}_{\beta,N,\mu}^{\text{bb}}$. The same finite regrouping gives

$$\begin{aligned} Z_{L,N}^{\text{bb}}(\beta, \mu) &= \text{Tr} \left[\left(\mathsf{K}_{\beta,N,\mu}^{\text{bb}} (\mathsf{K}_{\beta,N,\mu}^{\text{bb}})^t \right)^{L/2} \right], \\ Z_L^{\text{bb}}(\beta, \mu) &:= \text{Tr}[e^{-\beta H_L^{\text{bb}}(\mu)}] = \lim_{N \uparrow \infty} Z_{L,N}^{\text{bb}}(\beta, \mu). \end{aligned} \quad (\text{C.11})$$

The product inside the trace is positive semidefinite. Applying the finite-dimensional probability inequality and then taking $N \uparrow \infty$ proves the same spatial purity inequality as before. The entire conditional argument applies, with unchanged thresholds. This construction alone does not verify the finite criterion (6.6) at any particular coupling. Using the strict seed (B.14) at $\mu = 0$, however, we can apply the continuity argument in footnote 10. For some $\varepsilon_0 > 0$, it gives a unique ground state and a common positive gap bound for all sufficiently large even rings and all $-\varepsilon_0 < \mu \leq 0$.

For $\mu > 0$, absorbing the coefficient into identical factors instead gives

$$A^{(\alpha,\eta)} = i\sqrt{\mu} G^{(\alpha)} G^{(\eta)}. \quad (\text{C.12})$$

These factors are imaginary, and the site traces need not be real. Take one history to be empty and the other to contain one event of type (α, α) in an allowed layer t . The corresponding matrix entry is

$$k_{\beta,N,\mu}(\emptyset, ((t, (\alpha, \alpha)))) = \sqrt{\frac{\beta}{N}} \text{tr} A^{(\alpha,\alpha)} = -2i\sqrt{\frac{\beta\mu}{N}}. \quad (\text{C.13})$$

Here $\text{tr}[(G^{(\alpha)})^2] = -2$ follows directly from (2.5). Exchanging the histories gives the same entry, not its complex conjugate. The spatial contraction still pairs a matrix with its transpose, but for a complex matrix K , the product KK^t is not a Gram matrix KK^* . Thus the preceding argument no longer guarantees nonnegative eigenvalues, and we cannot infer a spatial probability distribution in this way.

This is a limitation of this direct factorization and transfer construction, not a proof that no other representation could make the abstract criterion applicable to models with $\mu > 0$, including the AKLT chain. Negative individual matrix entries or history weights are not the obstruction. What is lost is the real-matrix structure that made the two-site transfer matrix positive semidefinite. Frustration freeness plays no role in this distinction.

C.3 The bond-alternating $S = 1$ chain

This example illustrates why applicability of the finite criterion does not by itself imply a gap, and why a verified gap does not determine the topological phase. Consider the periodic Hamiltonian on $2M$ sites, with $M \geq 2$,

$$H_{2M}^{\text{alt}} = \sum_{j=1}^M \{ \mathbf{S}_{2j-1} \cdot \mathbf{S}_{2j} + J \mathbf{S}_{2j} \cdot \mathbf{S}_{2j+1} \}, \quad (\text{C.14})$$

where $\mathbf{S}_k = (S_k^{(1)}, S_k^{(2)}, S_k^{(3)})$, $\mathbf{S}_{2M+1} = \mathbf{S}_1$, and $0 \leq J \leq 1$. At $J = 1$ this is the uniform Heisenberg chain; at $J = 0$ it consists of independent dimers with a product of spin singlets as its unique ground state. Grouping the two sites of each dimer into one cell makes the latter a noninteracting system with a gap of one. Perturbation theory about such a product state gives a unique periodic ground state and a uniform gap for all sufficiently small J . See, e.g., [32, Theorem 1]. Along this symmetry-preserving gapped path from $J = 0$, the ground state remains topologically trivial by invariance of the index [31]. Numerical studies support a transition to the Haldane phase through a gapless critical point at an intermediate J [33]. Nevertheless, the transfer construction underlying our criterion applies throughout the family, as we shall see below.

Use the same finite history sets as in Section 3. An event on an even bond now has an additional factor J . Split this factor equally between its two endpoints and define

$$(\mathbf{K}_{\beta,N,J})_{x,y} = J^{|y|/2} k_{\beta,N}(x,y), \quad x \in \mathcal{B}_o, \quad y \in \mathcal{B}_e. \quad (\text{C.15})$$

At $J = 0$, use $J^0 = 1$ for the empty history; all columns corresponding to nonempty even-bond histories vanish. The matrix is real for every $J \geq 0$. Summing over the even-bond histories again gives the positive semidefinite two-site matrix

$$\mathbf{Q}_{\beta,N,J} = \mathbf{K}_{\beta,N,J} \mathbf{K}_{\beta,N,J}^t. \quad (\text{C.16})$$

If $\mathcal{Z}_{M,N}(\beta, J)$ denotes the checkerboard approximation for $2M$ physical sites, then

$$\mathcal{Z}_{M,N}(\beta, J) = \text{Tr}[\mathbf{Q}_{\beta,N,J}^M], \quad \mathcal{Z}_M(\beta, J) := \text{Tr}[e^{-\beta H_{2M}^{\text{alt}}}] = \lim_{N \uparrow \infty} \mathcal{Z}_{M,N}(\beta, J). \quad (\text{C.17})$$

Below we suppress J in the partition functions. The spatial probability argument applies to the nonnegative eigenvalues of $\mathbf{Q}_{\beta,N,J}$, with length measured in two-site cells. Passing to the limit gives the same criterion and thresholds for the exact $\mathcal{Z}_M$. Here every integer cell number $M \geq 2$ is allowed: no evenness restriction on M is needed because $\mathbf{Q}_{\beta,N,J}$ is positive semidefinite. The interpolation argument works for cell numbers in the same way. This is an extension of the conditional theorem, not a numerical certification supplied by [1].

At a parameter where the periodic gaps have limit inferior zero as $M \uparrow \infty$, the two sufficient initial inequalities cannot both hold at any admissible finite pair of scales. At $J = 0$, on the other hand, let $z(\beta)$ be the partition function of one dimer. Then $\mathcal{Z}_M(\beta) = z(\beta)^M$, so

$$\frac{\mathcal{Z}_{2M}(\beta)}{\mathcal{Z}_M(\beta)^2} = 1, \quad \frac{\mathcal{Z}_{2M}(2\beta)}{\mathcal{Z}_M(\beta)^2} = \left\{ \frac{z(2\beta)}{z(\beta)^2} \right\}^{2M} \longrightarrow 1 \quad (\beta \uparrow \infty) \quad (\text{C.18})$$

for every fixed M . The limit follows from the uniqueness of the dimer ground state. The criterion is therefore satisfied in this trivially gapped limit as well: it certifies a uniform gap, not membership in the Haldane phase.

The continuity argument in footnote 10 also gives local gap stability at both $J = 0$ and $J = 1$, without invoking a separate perturbative stability theorem. At $J = 0$, fix any admissible cell number M_0 and then choose a finite β_0 so that both ratios in (C.18) are strictly greater than $7/8$. The finite-dimensional partition functions are continuous and strictly positive as functions of J , so these strict inequalities persist for all sufficiently small $J \geq 0$. The finite criterion gives a common positive lower bound on the gap throughout this neighborhood, uniformly in all sufficiently large cell numbers.

At $J = 1$, we have $\mathcal{Z}_M(\beta) = Z_{2M}(\beta)$. Thus the two initial inequalities in cell notation, with $M_0 = L_0/2$, are exactly the original inequalities for $S(L_0, \beta_0)$ and $T(2L_0, \beta_0)$. For example, the strict seed (B.14) corresponds to $M_0 = 30$ and $\beta_0 = 105/4$. Using that certified seed, continuity gives the same conclusion for $J \leq 1$ sufficiently close to one. Although the auxiliary transfer matrices differ, no comparison between their spectra is needed: the finite criterion is written entirely in terms of the exact physical partition functions.

D A sufficient condition for the converse

Suppose a family of spin chains admits the spatial transfer representation (3.23) and has unique ground states with a uniform gap on all sufficiently large periodic rings. Must there exist finite L_0 and β_0 for which both initial purity inequalities (6.6) hold? In other words, does the criterion eventually detect every such gapped family?

A low-temperature limit at fixed length does not answer this question. Although the temporal purity tends to one, the spatial purity can tend to zero, as we shall see below. The length and inverse temperature must be chosen together. We show that an additional condition on the finite-size ground energies suffices to make both purities tend to one along a suitable sequence of scales.¹⁶ We do not derive that condition from the gap alone.

Consider finite-dimensional self-adjoint Hamiltonians H_L on the allowed even periodic lengths $L \geq 4$. Write d_L for their Hilbert-space dimensions and assume $2 \leq d_L \leq \nu^L$ for a fixed $\nu > 1$. Suppose their partition functions have a spatial transfer representation of the form

$$Z_L(\beta) = \text{Tr}[e^{-\beta H_L}] = \lim_{N \uparrow \infty} \sum_{i=1}^{r_N} \lambda_i(\beta, N)^L, \quad r_N < \infty, \quad \lambda_i(\beta, N) \geq 0, \quad (\text{D.1})$$

for every $\beta > 0$, with the same finite list at each β, N for all allowed even lengths. As in Section 5, finite-sum inequalities pass to the limit in the partition functions; no limiting spectral list is assumed. Define S and T by (5.9) and (5.2). As before, $0 < S(L, \beta), T(L, \beta) \leq 1$.

Suppose first that the ground states at L and $2L$ are unique. List all energies with multiplicity and define the excitation contribution

$$R_L(\beta) = \sum_{k=1}^{d_L-1} \exp\{-\beta[E_k(L) - E_0(L)]\}. \quad (\text{D.2})$$

Then

$$Z_L(\beta) = e^{-\beta E_0(L)}[1 + R_L(\beta)]. \quad (\text{D.3})$$

¹⁶This supplementary observation was found by ChatGPT in response to my question during the preparation of this note, and is not part of the original proof in [1].

Substitution into the definitions of the purities gives the identities

$$T(L, \beta) = \frac{1 + R_L(2\beta)}{[1 + R_L(\beta)]^2}, \quad (\text{D.4})$$

and

$$S(L, \beta) = e^{-\beta D_L} \frac{1 + R_{2L}(\beta)}{[1 + R_L(\beta)]^2}, \quad D_L = E_0(2L) - 2E_0(L). \quad (\text{D.5})$$

For fixed L , uniqueness gives $R_L(\beta), R_{2L}(\beta) \rightarrow 0$ as $\beta \uparrow \infty$. Since $S(L, \beta) \leq 1$, identity (D.5) implies $D_L \geq 0$; a negative D_L would make its right-hand side diverge.¹⁷ Moreover,

$$S(L, \beta) \sim e^{-\beta D_L} \quad (\beta \uparrow \infty \text{ at fixed } L). \quad (\text{D.6})$$

For $D_L > 0$, the spatial purity tends to zero while the temporal purity tends to one, so taking β sufficiently large at fixed L fails to meet the finite criterion. If $D_L = 0$, both purities instead tend to one.

If, on the other hand, there is a sequence of even lengths $L \uparrow \infty$ and inverse temperatures β_L for which

$$R_L(\beta_L) \rightarrow 0, \quad R_{2L}(\beta_L) \rightarrow 0, \quad \beta_L D_L \rightarrow 0, \quad (\text{D.7})$$

one necessarily has $S(L, \beta_L) \rightarrow 1$ and $T(2L, \beta_L) \rightarrow 1$. For the latter conclusion, one can use $0 \leq R_{2L}(2\beta_L) \leq R_{2L}(\beta_L)$ in (D.4). The two requirements are therefore to suppress physical excitations and to control the difference between ground energies in distinct volumes. This leads to the following proposition.

Proposition D.1 *Assume (D.1) and $d_L \leq \nu^L$. Suppose that, for all sufficiently large even L , H_L has a unique ground state and*

$$E_1(L) - E_0(L) \geq \Delta > 0, \quad D_L = E_0(2L) - 2E_0(L) = o(L^{-1}). \quad (\text{D.8})$$

Then, for every $c > 2 \log \nu / \Delta$,

$$\lim_{\substack{L \uparrow \infty \\ L \in 2\mathbb{N}}} S(L, cL) = \lim_{\substack{L \uparrow \infty \\ L \in 2\mathbb{N}}} T(2L, cL) = 1. \quad (\text{D.9})$$

In particular, for every prescribed $u_0 > 0$, the two finite inequalities (6.6) are satisfied at $L_0 = L$, $\beta_0 = cL$ for all sufficiently large even L .

Proof. For all sufficiently large even L , every excited energy is at least Δ above the ground energy. The number of excited eigenvectors is at most $\nu^L - 1$, so

$$0 \leq R_L(\beta) \leq (\nu^L - 1)e^{-\beta \Delta}. \quad (\text{D.10})$$

¹⁷We have thus proved $E_0(2L) \geq 2E_0(L)$. In fact, this inequality holds without assuming uniqueness or a gap. The spatial transfer representation gives $Z_{2L}(\beta) \leq Z_L(\beta)^2$; taking logarithms, multiplying by $-1/\beta$, and using $-\beta^{-1} \log Z_L(\beta) \rightarrow E_0(L)$ proves the assertion. This is a consequence of the assumed representation, not a claim for arbitrary periodic Hamiltonians. For the Heisenberg chain, the inequality can also be obtained by the reflection-positivity method of Dyson, Lieb, and Simon [34]. The proof given here uses only the spatial transfer representation with real eigenvalues and also applies to models admitting such a representation without the usual quantum bond-reflection positivity.

Fix $c > 2 \log \nu / \Delta$ and let

$$\eta_L = \nu^{2L} e^{-cL\Delta} = e^{-\{c\Delta - 2 \log \nu\}L}. \quad (\text{D.11})$$

The condition on c implies that η_L decays exponentially as L grows. Now set $\beta_L = cL$. For all sufficiently large even L , applying (D.10) at L and $2L$ gives

$$R_L(\beta_L) \leq \eta_L, \quad R_{2L}(\beta_L) \leq \eta_L, \quad R_{2L}(2\beta_L) \leq \eta_L. \quad (\text{D.12})$$

Since $D_L \geq 0$, (D.5) and $S \leq 1$ imply

$$1 \geq S(L, \beta_L) \geq \frac{e^{-cLD_L}}{(1 + \eta_L)^2}. \quad (\text{D.13})$$

Similarly, retaining the ground contribution in the numerator of (D.4) gives

$$1 \geq T(2L, \beta_L) \geq \frac{1}{(1 + \eta_L)^2}. \quad (\text{D.14})$$

Now $LD_L \rightarrow 0$ by hypothesis and $\eta_L \rightarrow 0$, proving both limits. $\square$

Acknowledgments

I thank OpenAI for making these papers available. I thank ChatGPT (OpenAI, GPT-6 Astra Pro) for extensive help in reading and understanding the OpenAI paper, for discussions of the reformulation of the proof, and for its help in preparing and revising this article from my handwritten notes. I also thank Hosho Katsura for discussions that prompted me to write this discrete-time version, Tohru Koma for useful comments on spatial transfer matrices, and the late Professor Masuo Suzuki for introducing me to space-time representations of quantum spin systems when I was his student. I thank Hidetoshi Nishimori for sharing the results of independent AI checks of both versions. I often had *Symphonie fantastique* playing in the background while working on this note; I thank Hector Berlioz and Seiji Ozawa for providing an uncannily fitting accompaniment to this unusual undertaking.

The present project is supported in part by JSPS Grants-in-Aid for Scientific Research No. 25K07171.

References

- [1] OpenAI, *The periodic spin-one Haldane gap*, OpenAI Math Release preprint (September 24, 2026), 30 pages. The section and proposition numbers in this note refer to that manuscript.
Public manuscript on GitHub.
- [2] OpenAI, *Sharing AI progress in mathematics*, announcement (October 6, 2026).
<https://openai.com/index/sharing-ai-progress-in-mathematics/>
- [3] OpenAI, *Mathematics manuscripts and supporting proof artifacts*, public repository (consulted October 12, 2026).
<https://github.com/openai/math>

- [4] OpenAI, *A boundary-field gap for the spin-one Heisenberg chain*, OpenAI Math Release preprint (September 24, 2026).
Public manuscript on GitHub.
- [5] F.D.M. Haldane, *Ground State Properties of Antiferromagnetic Chains with Unrestricted Spin: Integer Spin Chains as Realisations of the $O(3)$ Non-Linear Sigma Model*, ILL preprint SP-81/95 (1981).
<https://arxiv.org/abs/1612.00076>
- [6] F.D.M. Haldane, *Continuum dynamics of the 1-D Heisenberg antiferromagnet: identification with the $O(3)$ nonlinear sigma model*, Phys. Lett. **93A**, 464–468 (1983).
- [7] F.D.M. Haldane, *Nonlinear field theory of large-spin Heisenberg antiferromagnets: semi-classically quantized solitons of the one-dimensional easy-axis Néel state*, Phys. Rev. Lett. **50**, 1153–1156 (1983).
<https://journals.aps.org/prl/abstract/10.1103/PhysRevLett.50.1153>
- [8] H. Tasaki, *Physics and Mathematics of Quantum Many-Body Systems*, Graduate Texts in Physics (Springer, 2020).
Author’s book page, with corrections and addenda.
- [9] I. Affleck, T. Kennedy, E.H. Lieb, and H. Tasaki, *Valence bond ground states in isotropic quantum antiferromagnets*, Commun. Math. Phys. **115**, 477–528 (1988).
Public full text at Project Euclid.
- [10] J. Ginibre, *Reduced density matrices of the anisotropic Heisenberg model*, Commun. Math. Phys. **10**, 140–154 (1968).
Public full text at Project Euclid.
- [11] J. Ginibre, *Existence of phase transitions for quantum lattice systems*, Commun. Math. Phys. **14**, 205–234 (1969).
Public full text at Project Euclid.
- [12] M. Aizenman and B. Nachtergaele, *Geometric aspects of quantum spin states*, Commun. Math. Phys. **164**, 17–63 (1994).
<https://arxiv.org/abs/cond-mat/9310009>
- [13] H. Araki, *Gibbs states of a one dimensional quantum lattice*, Commun. Math. Phys. **14**, 120–157 (1969).
Public full text at Project Euclid.
- [14] M. Suzuki, *Transfer-matrix method and Monte Carlo simulation in quantum spin systems*, Phys. Rev. B **31**, 2957–2965 (1985).
- [15] H. Betsuyaku, *Study of one-dimensional quantum spin systems by the transfer-matrix method*, Prog. Theor. Phys. **73**, 319–331 (1985).
<https://academic.oup.com/ptp/article/73/2/319/1848731>
- [16] M. Suzuki, *Relationship between d -dimensional quantal spin systems and $(d + 1)$ -dimensional Ising systems: Equivalence, critical exponents and systematic approximants of the partition function and spin correlations*, Prog. Theor. Phys. **56**, 1454–1469 (1976).
- [17] H. Tasaki, *OpenAI’s proof of the Haldane gap in the $S = 1$ antiferromagnetic Heisenberg chain: Continuous-time version*, companion expository note (2026).
https://haltasaki.github.io/writings/AI/OpenAI_Haldane_CT.pdf

- [18] H. Tasaki, *The ground state of the $S = 1$ antiferromagnetic Heisenberg chain is topologically nontrivial if gapped*, Phys. Rev. Lett. **134**, 076602 (2025).
<https://arxiv.org/abs/2407.17041v4>
- [19] H. Tasaki, *The Lieb–Schultz–Mattis theorem: a topological point of view*, in R.L. Frank, A. Laptev, M. Lewin, and R. Seiringer (eds.), *The Physics and Mathematics of Elliott Lieb*, Vol. 2, pp. 405–446 (EMS Press, 2022).
<https://arxiv.org/abs/2202.06243v3>
- [20] L. Campos Venuti, E. Ercolessi, G. Morandi, P. Pieri, and M. Roncaglia, *Spin chains in an external magnetic field: closure of the Haldane gap and effective field theories*, Int. J. Mod. Phys. B **16**, 1363 (2002).
<https://arxiv.org/abs/cond-mat/9908044>
- [21] P. Harremoës, *Joint range of Rényi entropies*, Kybernetika **45**, 901–911 (2009).
<https://www.kybernetika.cz/content/2009/6/901/paper.pdf>
- [22] Y. Sakai and K. Iwata, *Sharp bounds between two Rényi entropies of distinct positive orders*, preprint (2016).
<https://arxiv.org/abs/1605.00019>
- [23] OpenAI, *Sharp mass bounds for the two-dimensional $O(4)$ model*, OpenAI Math Release preprint (September 23, 2026), 102 pages.
Public manuscript on GitHub.
- [24] S. Östlund and S. Rommer, *Thermodynamic limit of density matrix renormalization*, Phys. Rev. Lett. **75**, 3537–3540 (1995).
<https://arxiv.org/abs/cond-mat/9503107>
- [25] E. Lieb, T. Schultz, and D. Mattis, *Two soluble models of an antiferromagnetic chain*, Ann. Phys. **16**, 407–466 (1961).
Public full text.
- [26] I. Affleck and E.H. Lieb, *A proof of part of Haldane’s conjecture on spin chains*, Lett. Math. Phys. **12**, 57–69 (1986).
- [27] L.A. Takhtajan, *The picture of low-lying excitations in the isotropic Heisenberg chain of arbitrary spins*, Phys. Lett. A **87**, 479–482 (1982).
- [28] H.M. Babujian, *Exact solution of the one-dimensional isotropic Heisenberg chain with arbitrary spins S* , Phys. Lett. A **90**, 479–482 (1982).
- [29] H. Tasaki, *Topological phase transition and $\mathbb{Z}_2$ index for $S = 1$ quantum spin chains*, Phys. Rev. Lett. **121**, 140604 (2018).
<https://arxiv.org/abs/1804.04337>
- [30] H. Tasaki, *The equivalence of topological indices for $U(1) \rtimes \mathbb{Z}_2$ -symmetric quantum spin chains*, preprint (2026), version 2.
<https://arxiv.org/abs/2609.34119>
- [31] Y. Ogata, *A $\mathbb{Z}_2$ -index of symmetry protected topological phases with time reversal symmetry for quantum spin chains*, Commun. Math. Phys. **374**, 705–734 (2020).
<https://arxiv.org/abs/1810.01045>
- [32] D.A. Yarotsky, *Ground states in relatively bounded quantum perturbations of classical lattice systems*, Commun. Math. Phys. **261**, 799–819 (2006).
<https://arxiv.org/abs/math-ph/0412040>

- [33] W. Chen, K. Hida, and B.C. Sanctuary, *Critical properties of spin-1 antiferromagnetic Heisenberg chains with bond alternation and uniaxial single-ion-type anisotropy*, J. Phys. Soc. Jpn. **69**, 237 (2000). The linked version includes the subsequent erratum.
<https://arxiv.org/abs/cond-mat/9908125v2>
- [34] F.J. Dyson, E.H. Lieb, and B. Simon, *Phase transitions in quantum spin systems with isotropic and nonisotropic interactions*, J. Stat. Phys. **18**, 335–383 (1978).