

Mutual information and breakdown of the Perron-Frobenius scenario in zero-temperature triangular Ising antiferromagnets on cylinders

Amir Nourhani ^{1,2,3,*} Vincent H. Crespi ^{4,5,6} and Paul E. Lammert ^{4,†}

¹*Department of Mechanical Engineering, University of Akron, Akron, Ohio 44325, USA*

²*Biomimicry Research and Innovation Center, University of Akron, Akron, Ohio 44325, USA*

³*Departments of Biology, Mathematics, and Chemical, Biomolecular and Corrosion Engineering, University of Akron, Akron, Ohio 44325, USA*

⁴*Department of Physics, Pennsylvania State University, University Park, Pennsylvania 16802, USA*

⁵*Department of Materials Science and Engineering, Pennsylvania State University, University Park, Pennsylvania 16802, USA*

⁶*Department of Chemistry, Pennsylvania State University, University Park, Pennsylvania 16802, USA*

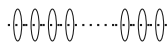
 (Received 23 August 2021; revised 11 February 2022; accepted 4 March 2022; published 5 April 2022)

A nominally two-dimensional spin model wrapped onto a cylinder can profitably be viewed, especially for long cylinders, as a one-dimensional chain. Each site of such a chain is a ring of spins with a complex state space. Traditional correlation functions are inadequate for the study of correlations in such a system and need to be replaced with something like mutual information. Being induced purely by frustration, the disorder of a cylindrical zero-temperature triangular Ising antiferromagnet (TIAFM) and attendant correlations have a chance of evading the consequences of the Perron-Frobenius theorem which describes and constrains correlations in thermally disordered one-dimensional systems. Correlations in such TIAFM systems and the aforementioned evasion are studied here through a fermionic representation. For cylindrical TIAFM models with open boundary conditions, we explain and derive the following characteristics of end-to-end mutual information: period-three oscillation of the decay length, halving of the decay length compared to what Perron-Frobenius predicts on the basis of transfer matrix eigenvalues, and subexponential decay—inverse square in the length—for certain systems.

DOI: [10.1103/PhysRevE.105.044105](https://doi.org/10.1103/PhysRevE.105.044105)

I. INTRODUCTION

Traditionally, probabilistic dependence among the elementary degrees of freedom in statistical mechanical models is studied by means of correlation functions. This is natural on a regular lattice of any dimension. However, a spin model wrapped on a cylinder with a length much greater than its circumference is also naturally viewed as a one-dimensional chain, except that the elementary constituents at the sites of the chain are rings of many spins each having many internal states:



How do we measure the probabilistic dependence among these constituents? Mutual information [1–4], which has been of increasing interest in classical statistical mechanics [5–7], as well as quantum information theory [8], is a good answer. Mutual information precisely quantifies dependence between two random variables of arbitrary complexity. In this paper, we apply this information-theoretic tool to the zero-temperature triangular lattice Ising antiferromagnet (TIAFM) model on cylinders, giving details of results which were briefly reported in a previous publication [9], thoroughly explaining some curious phenomena observed in numerical

experiments. The TIAFM is an archetype of frustration—the presence of incompatible but equally strong elementary interactions. The importance of understanding frustration-induced disorder is indicated by the enormous range of systems in which it occurs, from water ice [10,11] to spin systems [12–17], artificial spin ice [18–20], colloidal assemblies [21,22], Coulomb liquids [23], lattice gases [24], ferroelectrics [25], coupled lasers [26], and self-assembled lattices of microscopic chemical reactors [27].

The frustration-induced disorder of the zero-temperature TIAFM on a cylinder superficially resembles thermal disorder, but with a subtle and important difference. Correlations in thermally disordered one-dimensional systems are described by the Perron-Frobenius theorem [28–32]. Behavior of the cylindrical TIAFM is incompatible with what that theorem describes in a variety of ways, in particular in the rate with which the mutual information between configurations at the ends of a long cylinder with open boundary conditions falls off with cylinder length. Whereas the Perron-Frobenius behavior is a decay length $2 \ln |\lambda_1/\lambda_0|$, where λ_0 and λ_1 are the two largest eigenvalues of the transfer matrix, TIAFM cylinders show a decay length half this, except in cases where the mutual information does not even fall off exponentially, but only as the inverse square of the length. In addition, the decay length oscillates with circumference with a period of three. This last feature is not incompatible with the Perron-Frobenius scenario, but it is strange, and its elucidation, as we shall see, is bound up with that of the features

*nourhani@uakron.edu

†lammert@psu.edu

which are so. Vis-à-vis violation of the Perron-Frobenius scenario, the point of frustration is that it provides a kind of disorder which might lead to such behavior. Interesting analogs of the phenomena studied in this paper might also be found in dynamical systems, where the Perron-Frobenius theorem has a significant role [33] with time taking the place of the spatial dimension considered here, or possibly in systems displaying phyllotaxis-like phenomena [34].

Section II reviews transfer matrices, the Perron-Frobenius theorem, and its significant consequences, indicating how the zero-temperature cylindrical TIAFM deviates. Section III demonstrates the equivalence of the zero-temperature TIAFM with a fermionic model, constructs the transfer matrix in the fermion language, and works out some of its spectral properties. Section IV works out the crucial connectivity properties of the configuration space under powers of the transfer matrix, and asymptotic behavior of matrix elements. Section V returns to the Perron-Frobenius scenario, giving a thorough discussion of mutual information in that setting, exposing “normal” behavior under conditions of thermal disorder. Finally, Sec. VI works out details for the asymptotic end-to-end mutual information of TIAFM cylinders, explains the period-three oscillation of decay rate by a very prosaic fact about energy gaps in systems of free fermions and the inverse-square decay by the presence of zero-energy fermion modes, and shows how the data-processing inequality gives easy evaluation of decay rates. Comparison of the asymptotic results with exact numerical calculations shows that a length only about twice the circumference is already in the asymptotic regime. We have labeled some results Proposition or Lemma. This should not be understood as indicating relative level of mathematical rigor or importance (though Proposition IV.1 is certainly important). Rather, it is an organizational device, used when convenient, which facilitates reference and aids the reader who wishes to skip technical proofs by clearly marking out their beginnings and endings.

II. THE PERRON-FROBENIUS SCENARIO

A. Transfer matrices

Transfer matrices are a basic tool [35] of statistical mechanics. Their use is brought to a high art in the study of solvable two-dimensional models [36], where one takes the limit of both dimensions tending to infinity. Here, we stay with the one-dimensional case. To fix ideas, consider the simple Ising chain of length L . The spins $s_1, \dots, s_L$ take values in the configuration space $\mathcal{X} = \{-1, +1\}$, and the energy of the chain is

$$E(s) = -J \sum_{n=1}^{L-1} s_{n+1} s_n - h \sum_{n=1}^{L-1} s_n. \quad (1)$$

Then, defining the quantities

$$\mathbb{T}_{s's} = e^{\beta(Js's + hs)}, \quad (2)$$

indexed by elements of $\mathcal{X}$, the partition function with fixed boundary conditions $s_1, s_L \in \mathcal{X}$ can be written

$$Z(s_L | s_1) = \sum_{s_2, \dots, s_{L-1}} e^{-\beta E(s)} = \sum_{s_2, \dots, s_{L-1}} \mathbb{T}_{s_L, s_{L-1}} \cdots \mathbb{T}_{s_2, s_1}.$$

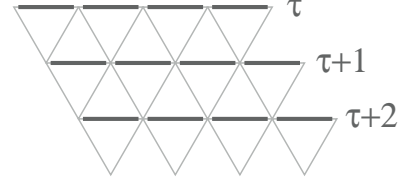


FIG. 1. Part of an unrolled TIAFM cylinder. The thick black bonds go around the cylinder circumference (see also Fig. 2).

Taking $\mathbb{T}_{ss'}$ as components of a matrix

$$\begin{pmatrix} e^{\beta(J+h)} & e^{\beta(-J+h)} \\ e^{\beta(-J-h)} & e^{\beta(J-h)} \end{pmatrix}, \quad (3)$$

the partition function becomes simply

$$Z(s_L | s_1) = (\mathbb{T}^L)_{s_L s_1}. \quad (4)$$

Naturally, we are inclined to think of $\mathbb{T}$ now as a linear operator on a free vector space over $\mathcal{X}$, or (better) on a Hilbert space $\mathcal{H}(\mathcal{X})$ with orthonormal basis $\{|s\rangle | s \in \mathcal{X}\}$. Nothing stops us from so doing, but we should pause to ask whether the vectors in $\mathcal{H}(\mathcal{X})$ are physically meaningful, that is, in the original statistical mechanical context. *Some* of them are. Consider for instance

$$|\bar{1}\rangle := |-1\rangle + |+1\rangle. \quad (5)$$

The partition function with open boundary conditions is $\langle \bar{1} | \mathbb{T}^L | \bar{1} \rangle$. More generally, any linear combination of $|-1\rangle$ and $|+1\rangle$ with nonnegative coefficients represents a (possibly unnormalized) probabilistic mixture, with a clear statistical mechanical meaning.

Figure 1 shows a bit of a cylindrical TIAFM system, unwrapped. The details of an appropriate transfer matrix method for this at zero temperature are different from the simple Ising chain, though the spirit is the same. The role of sites is played by the circumferential rings of bonds, and the configuration space $\mathcal{X}$ is now the space of all bond configurations (satisfied/unsatisfied) on such a ring. Configurations of the noncircumferential bonds are implicit in the transfer matrix $\mathbb{T}$.

At zero temperature, the partition function should simply count ground microstates, once the ground-state energy is set to zero by addition of a constant. However, for a ring-to-ring transfer matrix to exist, the property of being a ground microstate needs to be appropriately local. This is the role of making our cylinders from rings of down-pointing triangles (∇). That ensures that every bond is in one and only one ∇ , and the ground microstates are precisely those in which each ∇ has two satisfied bonds. In turn, this ensures the existence of a ring-to-ring transfer matrix; Sec. III constructs it explicitly.

B. Perron-Frobenius Theorem

An appropriate general abstract setting in which to consider both the zero-temperature cylindrical TIAFM and the systems to which it will be contrasted, is a chain of sites, site i hosting the configuration variable X_i taking value in some finite set $\mathcal{X}$. The partition function might be a matrix element of $\mathbb{T}^L$ (fixed boundary conditions), a sum of matrix elements

(open boundary conditions), or the trace (periodic). As before, it is convenient to identify $\mathcal{X}$ with an orthonormal basis in Hilbert space $\mathcal{H}(\mathcal{X})$, so that the transfer *matrix* becomes a transfer *operator*.

1. The theorem

By nature, the entries of the transfer matrix $\mathbb{T}$ are nonnegative in the $\mathcal{X}$ basis, since they represent statistical weights. The Perron-Frobenius theorem addresses the situation where it has the stronger property

PF: There is an N , such that every element of $\mathbb{T}^n$ is strictly positive when $n \geq N$.

This says that, regardless of the configuration at a given site, at any site sufficiently far away, any configuration can occur. One generally expects this condition to hold for a disordered system, and it certainly holds at nonzero temperature.

Theorem II.1 (Perron-Frobenius). When the condition **PF** holds, $\mathbb{T}$ has a unique eigenvalue λ_0 of largest modulus, which is nondegenerate (real) and greater than zero. Also, every component of the associated left or right eigenvector is strictly positive with appropriate choice of overall phase.

For a proof, see, for example, Proposition 5.6.3 of Ref. [31], or Theorem II.5.1 of Ref. [32]). The theorem implies that $\mathbb{T}$ can be written as

$$\mathbb{T} = \lambda_0(Q_0 + S), \quad (6)$$

with the following characteristics:

(1) Q_0 is a rank-one projection (generally nonorthogonal), hence can be written as

$$Q_0 = |e_0\rangle\langle\theta^0|, \quad \langle\theta^0|e_0\rangle = 1. \quad (7)$$

(2) $\forall x \in \mathcal{X}, \quad \langle e_0|x\rangle > 0, \quad \langle\theta^0|x\rangle > 0.$

(3) $SQ_0 = Q_0S = 0.$

(4) The spectral radius of S is $|\lambda_1|/\lambda_0 < 1$, where the subleading eigenvalue(s) of $\mathbb{T}$ have modulus $|\lambda_1|$.

Generally, we cannot assume that $\mathbb{T}$ is hermitian; one must be on alert against ingrained habits attuned to that case.

We will be interested in asymptotics as the separation L between sites of a chain system (possibly the ends) tends to infinity, and therefore establish some convenient notation now. We use $f(L) \lesssim g(L)$ as a synonym for Landau big- $\mathcal{O}$: $f = \mathcal{O}(g)$ when $|f(L)/g(L)|$ is bounded for large enough L . This is more convenient than $\mathcal{O}$ when the expression for g is long, and includes the case $f = o(g)$, when the bound can be taken as small as desired. If $f \lesssim g$ and $f \gtrsim g$, then we write $f \sim g$, as is common notation in the physics literature. Finally, $f(L) \approx g(L)$ means that $|f(L) - g(L)| = o(g)$.

2. The PF scenario

Now we list some implications of the Perron-Frobenius, which we collectively and somewhat loosely refer to as the Perron-Frobenius (PF) scenario.

(1) For all $x, y \in \mathcal{X}$,

$$\langle y|\mathbb{T}^L|x\rangle \approx \lambda_0^L \langle y|Q_0|x\rangle > 0. \quad (8)$$

(2) Regardless of boundary conditions, in the thermodynamic limit there is a unique bulk state and its free entropy density $\lim L^{-1} \ln Z$ is equal to $\ln \lambda_0$.

(3) In the thermodynamic limit, and for site 0 in the bulk (i.e., the ends recede infinitely far from that site in the limit),

$$\langle f(X_0) \rangle = \sum_{x \in \mathcal{X}} f(x) \langle x|Q_0|x\rangle. \quad (9)$$

(4) In both the cases that sites 0 and L are in the bulk of an infinite system, or they are the end sites of a chain with open boundary conditions, the connected correlations obey

$$\begin{aligned} \langle f(X_0); g(X_L) \rangle &:= \langle f(X_0)g(X_L) \rangle - \langle f(X_0) \rangle \langle g(X_L) \rangle \\ &\lesssim q(L) \left(\frac{|\lambda_1|}{\lambda_0} \right)^L, \end{aligned}$$

where the subleading eigenvalue(s) of $\mathbb{T}$ have modulus $|\lambda_1|$. $q(L)$ is a polynomial; this complication is needed in case some $|\lambda_1|$ eigenspace has algebraic multiplicity greater than its geometric multiplicity.

The preceding aspects are widely known. We add a formula for the asymptotic behavior of mutual information between the configurations on sites X_0 and X_L . (Mutual information will be reviewed in Sec. V)

(5) With sites 0 and L again either in the bulk or the ends of a chain with open boundary conditions, the mutual information between their configurations obeys

$$I(X_0 : X_L) \lesssim q(L) \left(\frac{|\lambda_1|}{\lambda_0} \right)^{2L}. \quad (10)$$

C. Oddities of the zero-temperature cylindrical TIAFM

It is common for a one-dimensional system to violate the hypothesis of the Perron-Frobenius theorem, at zero temperature. However, the result is usually a very simple situation in that there are a few ground states, e.g., for the simple Ising chain, all spins up or all spins down. The cylindrical TIAFM differs by exhibiting disorder which looks very similar to ordinary thermal disorder. It is therefore *a priori* plausible that the Perron-Frobenius scenario applies. Of course, there is a local constraint (each ∇ has exactly two satisfied bonds), but it would be a mistake to jump to conclusions from that. Consider, for instance, a clock model with an antiferromagnetic interaction, allowed directions for the spins being spaced by 15° . Imposing the constraint that neighboring spins can differ by at most 30° makes no qualitative difference.

(1) There are multiple bulk states in the thermodynamic limit. They are labeled by the number $\mathcal{N}$ of satisfied bonds on each circumferential ring, and have differing entropy densities.

(2) In each such bulk state, connected correlation functions and ring-to-ring mutual information behave as in items 4 and 5 above, but the relevant eigenvalues are a subset of those for the full transfer matrix $\mathbb{T}$, specific to $\mathcal{N}$. All the eigenvalues of $\mathbb{T}$ are relevant for end-to-end correlations, however.

(3) Despite that normal aspect of end-to-end mutual information, it behaves very peculiarly in other respects:

(i) If the circumference is a multiple of three and $\mathcal{N}$ is odd, then the end-to-end mutual information decays not exponentially, but as L^{-2} .

(ii) In other cases, the exponent $2L$ in Eq. (10) is replaced by L .

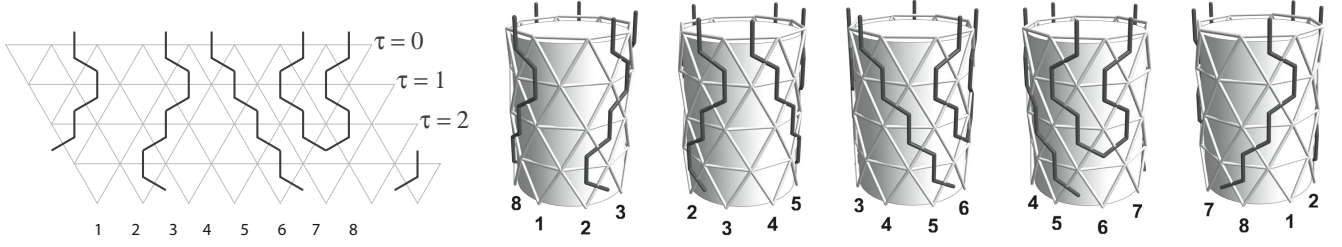


FIG. 2. On this diagram of a cylindrical system, there are four strings. In the fermionic picture, there are five fermions at time 0. The doubling back of one string between $\tau = 2$ and $\tau = 3$ is interpreted as annihilation of a pair of fermions.

(iii) On top of a general increase with circumference, the rate of decay oscillates with period three.

These aspects are all somewhat interrelated. In fact, the period-three oscillation for even- $\mathcal{N}$ systems has no direct relation to violation of the Perron-Frobenius condition, but we need to uncover the whole story to see that clearly. In the following sections, we will construct a very convenient fermionic representation of the transfer operator, and use it to explain these features. The coefficient for the end-to-end mutual information will be calculated, and shown to agree with direct numerical calculations.

D. Non-Perron-Frobenius in general

As before, consider a chain of sites with a finite single-site state space $\mathcal{X}$ and transfer matrix $\mathbb{T}$. We introduce a relation of mutual reachability: For $Y, Y' \in \mathcal{X}$, if there is some N such that both $\langle Y' | \mathbb{T}^m | Y \rangle$ and $\langle Y | \mathbb{T}^m | Y' \rangle$ are nonzero whenever $m \geq N$, Y and Y' are mutually reachable, denoted $Y \sim Y'$. This relation is evidently an equivalence relation, so $\mathcal{X}$ is partitioned into equivalence classes $\{\mathcal{X}_\alpha\}$ of mutually reachable states. In graph theory terminology, these equivalence classes are *strongly connected components*. The directed graph $\mathcal{G}$ with these components as nodes and a directed edge from $\mathcal{X}_\alpha$ to $\mathcal{X}_\beta$ when $\langle Y_\beta | \mathbb{T} | Y_\alpha \rangle$ for some $Y_\alpha \in \mathcal{X}_\alpha$ and $Y_\beta \in \mathcal{X}_\beta$ is directed acyclic (there is no directed path from any node back to it). The Perron-Frobenius condition holds when the graph has only one node. As mentioned in Sec. II B, thermal disorder always produces this case. The graphs $\mathcal{G}$ for zero-temperature TIAFM cylinders have one of the simplest nontrivial structures: nodes correspond to the number $\mathcal{N}$ of satisfied circumferential bonds, and $\mathbb{T}$ connects a state only states with a lesser or equal value of $\mathcal{N}$. Frustration provides the nonthermal disorder making this possible. Since the cylindrical TIAFM can be studied in great detail through the fermionic representation, it is a good model system for this sort of non-Perron-Frobenius behavior.

III. CONSTRUCTION OF TRANSFER MATRIX FOR $T = 0$ CYLINDRICAL TIAFM

We construct the transfer matrix $\mathbb{T}$ for the zero-temperature cylindrical TIAFM in three stages [9]: See how to represent bond configurations by *string diagrams*, identify the elementary processes of which they are composed, and find appropriate fermionic representations of those processes.

We work with configurations of satisfied and unsatisfied bonds rather than directly with spins. The simple two-to-one

correspondence between spin and bond configurations means there is no loss in doing so. In a system composed of ∇ 's (see Figs. 2 and 3), ground states are all and only those bond configurations with exactly two satisfied bonds on each ∇ , with one caveat to be removed in the next paragraph. To convert a bond configuration into a system of strings, mark a perpendicular across each *satisfied* circumferential bond (horizontal in the figures) and across each *unsatisfied* noncircumferential bond (diagonal), as illustrated at the left in Fig. 2. The only motifs allowed in a ∇ are $\nearrow$, $\nwarrow$, and the empty motif ∇ . $\nwarrow$ corresponds to zero satisfied bonds, ruled out by the ground-state constraint, and anything else is topologically impossible. It is very natural to interpret these pictures as a sort of spacetime diagram (Fig. 3) with the strings being worldlines of particles. These particles are conserved, except for the events depicted as $\nwarrow$, which represents annihilation of neighboring particles.

For the ordinary TIAFM, there must be an even number $\mathcal{N}$ of satisfied bonds along any circumferential loop, because these correspond to spin flips. Therefore, systems with odd circumference C are circumferentially frustrated. We are going to decouple the frustration status of circumferential loops from the parity of C . The important point is that a triangle with one antiferromagnetic bond is precisely as frustrated as one with three such. If all bonds crossed by a string running the length of the cylinder are changed from antiferromagnetic to ferromagnetic, then this alters none of the local characteristics but reverses the frustration status of circumferential loops. Allowing this variant, the parities of $\mathcal{N}$ and of C become independent.

A. Particle-conserving submodel

To construct the transfer matrix, we will first ban pair annihilation, work out a fermionic representation for the resulting particle-conserving submodel, and restore pair annihilation in the following subsection.

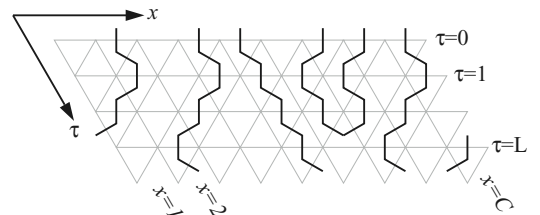


FIG. 3. Spacetime interpretation of a string diagram.

The nature of our particle worldlines is such that they cannot intersect. To implement this constraint automatically, we take the particles to be fermions. The configuration $X = (x_1, \dots, x_N)$ is the locations of satisfied circumferential bonds, in ascending order: $1 \leq x_1 < x_2 < \dots < x_N \leq C$. This configuration is identified with the fermion Fock space vector

$$|X\rangle := c_{x_1}^\dagger \cdots c_{x_N}^\dagger |\emptyset\rangle. \quad (11)$$

The particle preserving transfer matrix $\mathbb{T}_0$ is determined by the conditions

$$\mathbb{T}_0 |\emptyset\rangle = |\emptyset\rangle, \quad \mathbb{T}_0 c_x^\dagger = (c_{x-1}^\dagger + c_x^\dagger) \mathbb{T}_0, \quad (12)$$

the point being that a particle at x at time τ can be at either x or $x-1$ at time $\tau+1$ (see Fig. 3). In k space, the condition Eq. (12) translates to

$$\begin{aligned} \mathbb{T}_0 c(q)^\dagger &= C^{-1/2} \sum_{x \in \mathbb{Z}_C} e^{iqx} \mathbb{T}_0 c_x^\dagger \\ &= \left(2 \cos \frac{q}{2}\right) e^{iq/2} c(q)^\dagger \mathbb{T}_0. \end{aligned} \quad (13)$$

This is solved by

$$\mathbb{T}_0 = e^{iP/2} e^{-H_0}, \quad (14)$$

where

$$H_0 = \sum_{q \in \text{BZ}} \varepsilon(q) n(q), \quad P = \sum_{q \in \text{BZ}} q n(q), \quad (15)$$

have the interpretation of a Hamiltonian and total momentum operator, respectively, and BZ [see Eq. (17)] stands for allowed momentum values in the Brillouin zone. $n(q) = c(q)^\dagger c(q)$ counts the number (0 or 1) of fermions in the mode of momentum q with energy

$$\varepsilon(q) = -\ln \left(2 \cos \frac{q}{2}\right). \quad (16)$$

The allowed fermion momentum modes depend on the parity of the number of particles N , and are given by

$$\text{BZ} = \begin{cases} \frac{2\pi}{C} \mathbb{Z} \cap (-\pi, \pi], & N \text{ odd}, \\ \frac{2\pi}{C} (\mathbb{Z} + \frac{1}{2}) \cap (-\pi, \pi], & N \text{ even}. \end{cases} \quad (17)$$

1. Zero-energy modes

The zeros of the dispersion relation $\varepsilon(k)$ are $k = \pm \frac{2\pi}{3}$. According to Eq. (17), the systems for which those momenta are in BZ are precisely those with $C \in 3\mathbb{Z}$ and odd N . Systems with zero-energy modes turn out to behave very differently from others and will generally require separate treatment in the following.

2. Eigenvectors

$\mathbb{T}_0$ is a hermitian operator, so its eigenvectors comprise a complete orthonormal basis. The eigenvectors have a very simple nature; they are constructed simply by occupying some particular set of k -states, subject to the parity constraint on N . [We speak in the singular, implicitly referring to systems with some particular circumference C , and N parity (even or odd)]. We will follow two different conventions for labeling

eigenvalues and eigenvectors, as convenience dictates. The first way is

$$\mathbb{T}_0 |\varphi_i\rangle = \lambda_i |\varphi_i\rangle, \quad (18)$$

where the eigenvalues are in descending order $|\lambda_0| \geq |\lambda_1| \geq \dots$. (We will never have any significant need to handle degeneracies systematically.) The other convention is to specify the particle number and order the eigenvalues of all N -particle states:

$$\mathbb{T}_0 |\varphi_{N,i}\rangle = \lambda_{N,i} |\varphi_{N,i}\rangle, \quad (19)$$

$$\mathbb{T}_0 = \sum_i \lambda_i |\varphi_i\rangle \langle \varphi_i|, \quad (20)$$

with $|\lambda_{N,0}| \geq |\lambda_{N,1}| \geq \dots$. Eigenstates of $\mathbb{T}_0$ are also eigenstates of P , which is why the factor $e^{iP/2}$ is not very important. In fact, the important eigenvectors for asymptotic properties are the $\varphi_{N,0}$. These are just filled Fermi seas of a density depending on N , and they have $P=0$. As N increases, $\lambda_{N,0}$ first increases monotonically, while modes of negative energy are being filled, then decreases monotonically. The number

$$N_0 := |\{k \in \text{BZ} | \varepsilon(k) \leq 0\}| \quad (21)$$

of nonpositive energy modes is important. $\lambda_{N_0,0}$ is the largest eigenvalue λ_0 of $\mathbb{T}_0$. If there are no zero-energy modes, then it is nondegenerate; otherwise, $\lambda_{N_0-2,0} = \lambda_{N_0,0}$.

B. Pair annihilation

Figure 2 depicts a single pair annihilation event. The transfer matrix “chooses” how the configuration will evolve from time τ to time $\tau+1$. And that can be understood as proceeding in two steps: first, select neighboring pairs for annihilation, then move or leave in place the remaining particles. This annihilation step is implemented by the operator

$$\mathbb{T}_{\text{pr}} = \prod_{i \in \mathbb{Z}_C} (1 + c_{i+1} c_i), \quad (22)$$

which selects neighboring pairs in all possible ways from the ring. Note that the operators in the product commute with each other, so it is unambiguous. Since the operators $c_{i+1} c_i$ square to zero, $1 + c_{i+1} c_i = \exp(c_{i+1} c_i)$. And, since they commute with each other, $\mathbb{T}_{\text{pr}} = \prod_i \exp(c_{i+1} c_i)$. Applying commutativity again produces $\mathbb{T}_{\text{pr}} = e^{-H_{\text{pr}}}$, where

$$H_{\text{pr}} = - \sum_i c_i c_{i+1} = \sum_{0 < q \in \text{BZ}} 2i(\sin q) c(-q) c(q) \quad (23)$$

after Fourier transformation. The exponentiated expression $e^{-H_{\text{pr}}}$ re-expands to

$$\mathbb{T}_{\text{pr}} = \prod_{0 < q \in \text{BZ}} [1 + 2i(\sin q) b(q)], \quad (24)$$

where

$$b(k) := c(k) c(-k) \quad (25)$$

destroys a pair of fermions. The complete transfer matrix is then

$$\mathbb{T} = \mathbb{T}_0 \mathbb{T}_{\text{pr}} = e^{iP/2} e^{-H_0} e^{-H_{\text{pr}}}. \quad (26)$$

Now, P commutes with H_0 and $b(k)$, but $b(k)$ does not commute with $\mathbb{T}_0$. Instead,

$$b(k)\mathbb{T}_0 = e^{-2\varepsilon(k)}\mathbb{T}_0 b(k). \quad (27)$$

Evidently, zero-energy modes are special, since $b(\frac{2\pi}{3})$ commutes with $\mathbb{T}_0$, whereas the other pair annihilation operators do not.

1. Case: No zero-energy modes

When there are no zero-energy modes, we use Eq. (27) to write $\mathbb{T}$ as a similarity transformation of $\mathbb{T}_0$. Since

$$[1 + \alpha b(k)][1 + \beta b(k)] = [1 + (\alpha + \beta)b(k)], \quad (28)$$

for any complex α , we have

$$\begin{aligned} \mathbb{T}_0[1 + \alpha b(k)] &= [1 + \tilde{\alpha}b(k)]\mathbb{T}_0[1 - \tilde{\alpha}b(k)] \\ &= [1 + \tilde{\alpha}b(k)]\mathbb{T}_0[1 + \tilde{\alpha}b(k)]^{-1}, \end{aligned} \quad (29)$$

where

$$\tilde{\alpha} = \frac{\alpha}{e^{-2\varepsilon(k)} - 1}. \quad (30)$$

Using this observation, we can write the full transfer matrix as

$$\mathbb{T} = B(e^{iP/2}\mathbb{T}_0)B^{-1}, \quad (31)$$

with the definitions

$$B := \prod_{0 < q \in \text{BZ}} [1 + i\eta(q)b(q)], \quad (32)$$

and

$$\eta(k) := \frac{2 \sin k}{e^{-2\varepsilon(k)} - 1}. \quad (33)$$

2. Eigenvectors

Corresponding to the eigenvector $|\varphi_i\rangle$ of $\mathbb{T}_0$, with eigenvalue λ_i , $\mathbb{T}$ has a right eigenvector

$$|e_i\rangle = B|\varphi_i\rangle = \prod_{0 < q \in \text{BZ}} [1 + i\eta(q)b(q)]|\varphi_i\rangle \quad (34)$$

and a left eigenvector

$$|\theta^i\rangle = B^{-\dagger}|\varphi_i\rangle = \prod_{0 < q \in \text{BZ}} [1 - i\eta(q)b(q)]|\varphi_i\rangle, \quad (35)$$

with the same eigenvalue [$B^{-\dagger}$ is short for $(B^\dagger)^{-1}$]. B removes $(k, -k)$ pairs in all possible ways, with varying weight; $B^{-\dagger}$ adds them.

These satisfy the biorthogonality relation

$$\langle e_i | \theta_j \rangle = \delta_{ij}, \quad (36)$$

so that we can also write the transfer matrix as

$$\mathbb{T} = \sum_i \lambda_i |e_i\rangle \langle \theta_i|. \quad (37)$$

We can also write this in the form Eq. (38) introduced in Sec. II:

$$\mathbb{T} = \lambda_0(Q_0 + S), \quad (38)$$

where $Q_0 = |e_0\rangle \langle \theta_0|$ and S has spectral radius less than one. In contrast to the Perron-Frobenius scenario, however, it is

not the case that $\langle y | Q_0 | x \rangle > 0$ for all y and x . This is discussed later, and is very important for the behavior of mutual information.

3. Case: Zero-energy modes

As noted in Sec. III A 1, the system has zero-energy modes $k = \frac{2\pi}{3}$ exactly when $\mathcal{N}$ is odd and $C \in 3\mathbb{Z}$. $b(\frac{2\pi}{3})$ commutes with $\mathbb{T}_0$, as well as with P and all the other $b(k)$'s. We cannot deal with them in the same way as the others. If we temporarily hold them in reserve and write

$$B := \prod_{\substack{0 < q \in \text{BZ} \\ \varepsilon(q) \neq 0}} [1 + i\eta(q)b(q)], \quad (39)$$

and

$$\tilde{\mathbb{T}} = B(e^{iP/2}\mathbb{T}_0)B^{-1}, \quad (40)$$

then the full transfer matrix can be written as

$$\mathbb{T} = \tilde{\mathbb{T}}(1 + A_0) = (1 + A_0)\tilde{\mathbb{T}}, \quad (41)$$

where

$$A_0 := i\sqrt{3}b\left(\frac{2\pi}{3}\right). \quad (42)$$

While the situation is in certain respects more complicated than for systems without zero-energy modes, there are some compensations: A_0 commutes with $\mathbb{T}$ and squares to zero. Therefore $\mathbb{T}^L = \tilde{\mathbb{T}}^L(1 + LA_0)$.

We can construct a biorthogonal system of eigenvectors for $\tilde{\mathbb{T}}$ just as was done for systems without zero-energy modes:

$$|\tilde{e}_i\rangle = B|\varphi_i\rangle, \quad |\tilde{\theta}_i\rangle = B^{-\dagger}|\varphi_i\rangle. \quad (43)$$

Then,

$$\mathbb{T}|\tilde{e}_i\rangle = (\lambda_i + A_0)|\tilde{e}_i\rangle, \quad (44)$$

and

$$\mathbb{T} = \sum_i (\lambda_i + A_0) |\tilde{e}_i\rangle \langle \tilde{\theta}_i|. \quad (45)$$

Eigenstates $|\varphi_i\rangle$ of $\mathbb{T}_0$ can be classified according to whether the number of occupied zero-energy mode is one (annihilated by A_0 and $A_0^\dagger$), zero (annihilated by A_0 but not $A_0^\dagger$) or two (annihilated by $A_0^\dagger$, but not by A_0). This classification survives the passage to $|\tilde{e}_i\rangle$ and $|\tilde{\theta}_i\rangle$ because B does nothing to the zero-energy modes. Every eigenvector of $\tilde{\mathbb{T}}$ with no zero-energy modes has a partner with two, and they belong to the same eigenvalue. However, the vectors with both zero-energy modes occupied are not eigenvectors of the full transfer matrix $\mathbb{T}$. Indeed, it is not diagonalizable.

IV. STATES OF LONG TIAFM CYLINDERS

A. Bulk states of particle-conserving submodel

The particle-conserving submodel, $\mathbb{T}_0$, fails to satisfy the Perron-Frobenius condition rather trivially. Since $\mathbb{T}_0$ commutes with particle number $\mathcal{N}$, which is also the number of satisfied circumferential bonds, $\mathbb{T}_0^L$ can never connect configurations x and y if $\mathcal{N}(y) \neq \mathcal{N}(x)$. However, there is an equally trivial fix: restrict to the subspace $\mathcal{X}_{\mathcal{N}}$ of configurations

with N particles, for some fixed N . By use of the string representation, it is easy to see that for any $x, y \in \mathcal{X}_N$, $\langle y | \mathbb{T}_0^L | x \rangle > 0$ for all L large enough. The full Perron-Frobenius scenario is recovered for this restricted model: connected correlations decay as $(|\lambda_{N,1}|/\lambda_{N,0})^L$, and so forth.

Now, consider a finite cylinder with open boundary conditions. The partition function is $\langle \bar{1} | \mathbb{T}_0 | \bar{1} \rangle$, where, recall,

$$|\bar{1}\rangle = \sum_{x \in \mathcal{X}} |x\rangle.$$

Since we are working in the particle-conserving submodel, every allowed configuration of the cylinder has the same number of particles $\mathcal{N}$ at every point along the length. Writing

$$\langle \bar{1} | \mathbb{T}_0^L | \bar{1} \rangle = \sum_{N,i} \langle \bar{1} | \varphi_{N,i} \rangle \lambda_{N,i}^L \langle \varphi_{N,i} | \bar{1} \rangle, \quad (46)$$

we see that, as $L \rightarrow \infty$, the partition function is dominated by configurations with N^* particles, where $\lambda_{N^*,0}$ is largest possible. This value is unique and equal to N_0 , except for systems with zero-energy modes, in which case $\lambda_{N_0-2,0} = \lambda_{N,0}$.

B. A-expansion

The full model is much more complicated, and some crucial information is to be obtained by developing a view of the long cylinder and its partition function which does not work directly with the eigenvectors $|e_i\rangle$ and $|\theta_i\rangle$. Instead, decompose the pair annihilation part $\mathbb{T}_{\text{pr}}$ of $\mathbb{T}$ given in Eq. (22) or Eq. (24) as

$$\mathbb{T}_{\text{pr}} = 1 + A. \quad (47)$$

Here, A removes *one or more* pairs of particles. Then, $\mathbb{T} = \mathbb{T}_0(1 + A)$ and $\mathbb{T}^L$ is expanded as

$$\begin{aligned} \mathbb{T}^L &= \mathbb{T}_0^L + \sum \mathbb{T}_0^{L_2} A \mathbb{T}_0^{L_1} + \sum \mathbb{T}_0^{L_3} A \mathbb{T}_0^{L_2} A \mathbb{T}_0^{L_1} \\ &+ \cdots + \sum \left\{ \text{terms with } \lfloor \frac{L}{2} \rfloor A\text{'s} \right\}. \end{aligned} \quad (48)$$

The restrictions on the L_i are that: they must add up to L and only the first or last is allowed to be zero. In brief, at each “time step,” there is a choice to annihilate some pairs, or not. The main point is that there is a strictly limited number of A ’s, so they must become sparser and sparser on average as $L \rightarrow \infty$. The regions between them are characterized by constant values of $\mathcal{N}$, dropping by two at each A . Such a region of constant $\mathcal{N}$ looks like the corresponding bulk state of the $\mathcal{N}$ -conserving submode.

This way of looking at matters is very helpful in deducing the large- L asymptotics of the partition function $\langle y | \mathbb{T}^L | x \rangle$ for fixed boundary conditions x and y at the top and bottom of the cylinder, with $\mathcal{N}(y) \leq \mathcal{N}(x)$, as given in the following Proposition. States with $\mathcal{N}$ in the interval $[\mathcal{N}(y), \mathcal{N}(x)]$ are accessible, and as $L \rightarrow \infty$, the bulk will be in the fixed- $\mathcal{N}$ phase which maximizes the entropy density.

Proposition IV.1. If $\mathcal{N}(y) > \mathcal{N}(x)$, then $\langle y | \mathbb{T}^L | x \rangle = 0$. Otherwise, $\langle y | \mathbb{T}^L | x \rangle > 0$ for all sufficiently large L , and the following cases obtain:

Without zero-energy modes,

$$\langle y | \mathbb{T}^L | x \rangle \approx \langle y | e_{N^*,0} \rangle \langle \theta_{N^*,0} | x \rangle \lambda_{N^*,0}^L, \quad (49)$$

where N^* is the value of N in $[\mathcal{N}(y), \mathcal{N}(x)]$ which maximizes $\lambda_{N^*,0}$. Explicitly,

$$N^* = \begin{cases} N_0 & \mathcal{N}(y) \leq N_0 \leq \mathcal{N}(x), \\ \mathcal{N}(x) & \mathcal{N}(y) \leq \mathcal{N}(x) \leq N_0, \\ \mathcal{N}(y) & N_0 \leq \mathcal{N}(y) \leq \mathcal{N}(x). \end{cases} \quad (50)$$

With zero-energy modes, the preceding holds except in the case $\mathcal{N}(y) \leq N_0 - 2 < N_0 \leq \mathcal{N}(x)$. For such y, x ,

$$\langle y | \mathbb{T}^L | x \rangle \approx \langle y | \tilde{e}_{N_0-2,0} \rangle \langle \tilde{\theta}_{N_0,0} | x \rangle L \lambda_{N_0,0}^L, \quad (51)$$

Proof. That $\langle y | \mathbb{T}^L | x \rangle = 0$ when $\mathcal{N}(x) < \mathcal{N}(y)$ is clear, since $\mathbb{T}$ cannot add particles.

The A -expansion shows that for L' large enough, there is some $z \in \mathcal{X}_{\mathcal{N}(y)}$ with $\langle z | \mathbb{T}^{L'} | x \rangle > 0$. Since $\mathbb{T}$ acts irreducibly on $\mathcal{X}_{\mathcal{N}(y)}$, also $\langle y | \mathbb{T}^{L''} | z \rangle > 0$ for sufficiently large L'' . Therefore, $L \geq L' + L''$ will guarantee that $\langle y | \mathbb{T}^L | x \rangle > 0$.

Now, with top and bottom boundary conditions fixed at x and y , as $L \rightarrow \infty$, the partition function is dominated by configurations where the bulk has the allowed particle number [between $\mathcal{N}(y)$ and $\mathcal{N}(x)$], which maximizes entropy density, namely, N^* . This shows that $\langle y | \mathbb{T}^L | x \rangle \sim \lambda_{N^*,0}^L$. Equation (49) follows immediately, since

$$\langle y | \mathbb{T}^L | x \rangle = \sum_{N,i} \langle y | e_{N,i} \rangle \langle \theta_{N,i} | x \rangle \lambda_{N,i}^L.$$

For a system with zero-energy modes, the essentially new case is $\mathcal{N}(y) \leq N_0 - 2 < N_0 \leq \mathcal{N}(x)$. With the A -expansion picture, we see that the bulk of the system will be dominated by the N_0 “phase” or the $N_0 - 2$ phase. These have the same entropy density, so a transition from N_0 to $N_0 - 2$ comes at no penalty of entropy density, but can occur at any of L locations. Hence, $\langle y | \mathbb{T}^L | x \rangle \sim L \lambda_{N_0,0}^L$. Now we make the same kind of comparison as before, using $\mathbb{T}^L = \tilde{\mathbb{T}}^L(1 + LA_0)$ and $A_0 \tilde{e}_{N_0,0} = i\sqrt{3} \tilde{e}_{N_0-2,0}$:

$$\begin{aligned} \langle y | \mathbb{T}^L | x \rangle &= \langle y | (1 + LA_0) \tilde{\mathbb{T}}^L | x \rangle \\ &= \sum_{N,i} \langle y | 1 + LA_0 | \tilde{e}_{N,i} \rangle \langle \tilde{\theta}_{N,i} | x \rangle \lambda_{N,i}^L. \end{aligned} \quad (52)$$

■

V. MUTUAL INFORMATION IN THE PERRON-FROBENIUS SCENARIO

Traditional one- and two-spin correlation functions tell us everything there is to know about the distribution of a single spin conditioned on the value of one other. For complex random variables, such as a bond configuration on a circumferential ring, dependencies can be much more difficult to pin down. *Mutual information* provides a measure of the dependence in such situations. It does not characterize the nature of correlations, ferromagnetic versus antiferromagnetic, for example, but neither does it require insight into that nature. Correspondingly, it cannot miss anything we did not know to look for; it is in that sense a complete measure. Another important property of the mutual information which is not shared by ordinary correlation functions is reparametrization invariance. The mutual information $I(Y : X)$ between random variables X and Y is the same as that between X and $Y' = 3Y$

or $Y' = Y^3$; invertible transformations of the ranges have no effect.

Section **VA** is a quick review of important information theory concepts [1–4]. Section **VB** relates mutual information to the more traditional tool of connected correlation functions and gives a basic abstract estimate. Asymptotic formulas for ring-to-ring and end-to-end mutual information in the Perron-Frobenius scenario are worked out Secs. **VC** and **VD**, respectively. The details of the formulas in these sections will not be needed later. It is primarily the qualitative behavior that we wish to contrast with that of the zero-temperature cylindrical TIAFM.

A. Generalities about mutual information

The entropy of a discrete random variable Y is given by

$$H(Y) = - \sum_y P_Y(y) \ln P_Y(y), \quad (53)$$

where P_Y denotes the probability distribution of Y . A traditional interpretation is in the context of sending signals to communicate the results of repeated independent samples of Y . $H(Y)$ is proportional to the average number of bits per sample required to reliably encode the outcomes. Thus, the entropy reduces to a single number the uncertainty about Y codified by its probability distribution.

If X is a second random variable, then conditional entropy and mutual information are of interest. The entropy of Y conditional on X having value x is simply the entropy of the conditional distribution

$$H(Y|X=x) = - \sum_y P_{Y|X=x}(y) \ln P_{Y|X=x}(y), \quad (54)$$

$P_{Y|X=x}$ being the conditional distribution of Y , given that X takes value x . The entropy of Y conditional on X (not a particular outcome) is then the expectation of Eq. (54):

$$\begin{aligned} H(Y|X) &= \sum_x P_X(x) H(Y|X=x) \\ &= - \sum_{x,y} P_{Y,X}(y,x) \ln P_{Y|X=x}(y). \end{aligned} \quad (55)$$

In the coding context, if (X, Y) is sampled jointly and the X outcome is transmitted, then we need an additional $H(Y|X)$ bits per sample to reliably communicate Y as well. It is intuitively clear then that $H(Y) \geq H(Y|X)$, and the difference $H(Y) - H(Y|X)$ ought to represent the average amount of information X carries about Y .

We turn that into a definition. The mutual information between X and Y is

$$\begin{aligned} I(X:Y) &= H(Y) - H(Y|X) \\ &= H(X) + H(Y) - H(X, Y) \\ &= \sum_{y,x} P_{Y,X}(y,x) \ln \frac{P_{Y,X}(y,x)}{P_Y(y)P_X(x)} \\ &= \sum_{y,x} P(y)P(x) \frac{P(y,x)}{P(y)P(x)} \ln \frac{P(y,x)}{P(y)P(x)}. \end{aligned} \quad (56)$$

Manipulation of the definitions produces the second and third right-hand expressions, from which it is clear that mutual information is actually symmetric in its arguments, justifying the name.

In the following, we will use simplified notation as much as possible, as illustrated in the final line of Eq. (56), relying on context for the proper reading. Mostly, we are concerned with two variables X and Y , or X_0 and X_L , two variables along a chain, and will systematically use the dummy variables x and y for their values. We will instead use the subscript position P_L to indicate separation along a chain, or its length.

B. Mutual information and connected correlators

Mutual information has not been a concern of statistical mechanics nearly as much, or for such a long time, as correlation functions and connected correlation functions. We prove here a general abstract result about asymptotics of mutual information which will be applied in the following subsections.

Let P be a probability distribution for a pair (Y, X) of discrete random variables, and define Δ by

$$\frac{P(y,x)}{P(y)P(x)} = 1 + \Delta(y,x). \quad (57)$$

Then, using the notation $\delta_x(X)$ for the function which evaluates to one in case $X = x$, otherwise zero,

$$\begin{aligned} C(y,x) &:= \Delta(y,x)P(y)P(x) \\ &= P(y,x) - P(y)P(x) \\ &= \langle \delta_y(Y) \delta_x(X) \rangle - \langle \delta_y(Y) \rangle \langle \delta_x(X) \rangle \\ &= \langle \delta_y(Y); \delta_x(X) \rangle \end{aligned} \quad (58)$$

is a connected correlation function; the final rewriting explicitly claims as much. In fact, all connected correlation functions can be built from these:

$$\langle f(X); g(Y) \rangle = \sum_{y,x} f(x)g(y)C(y,x).$$

1. General asymptotics

Now, assume a family of probability distributions P_L , where L can be distance between two sites in a chain, or the length of a chain (among other things), and suppose the connected correlators Eq. (58) tend to zero as $L \rightarrow \infty$. This does not imply that $\Delta(y,x)$ tends to zero unless $P(x)$ and $P(y)$ remain bounded away from zero. That condition, however, is met in the Perron-Frobenius scenario, and then $C(y,x)$ and $\Delta(y,x)$ have the same asymptotic behavior.

Lemma V.1. Let $(P_L : L \in \mathbb{N})$ be a family of probability distributions for the pair (Y, X) , with corresponding Δ_L according to Eq. (57), and assume that $\Delta_L(y,x) \rightarrow 0$ for every x and y . Then,

$$I(Y:X) \approx \frac{1}{2} \langle \Delta_L(Y, X)^2 \rangle. \quad (59)$$

The expectation here, $\langle \rangle$, can be with respect to either P_L or the product of the marginals, since the difference is $\mathcal{O}(\Delta)$.

Proof. From the definition Eq. (57) of Δ , and Eq. (56) for mutual information,

$$I(X:Y) = \sum P(y)P(x) \{ [1 + \Delta(y,x)] \ln [1 + \Delta(y,x)] \}.$$

Now, observe that

$$[1 + \Delta(y, x)] \ln[1 + \Delta(y, x)] = \Delta + \frac{1}{2}\Delta^2 + \dots,$$

while

$$\sum_{y,x} P(y)P(x)\Delta(y, x) = \sum_{y,x} P(y, x) - \sum_{y,x} P(y)P(x) = 0.$$

■

C. Ring-to-ring mutual information

Now we consider the mutual information between sites separated by L in an infinite chain, under the Perron-Frobenius condition. We use notation introduced in Sec. II.

The probability of configuration x at a site is

$$P(x) = \langle \theta_0 | x \rangle \langle x | e_0 \rangle = \langle x | Q_0 | y \rangle. \quad (60)$$

For the product, we can do a swap:

$$P(y)P(x) = \langle y | Q_0 | y \rangle \langle x | Q_0 | x \rangle = \langle y | Q_0 | x \rangle \langle x | Q_0 | y \rangle. \quad (61)$$

To see this equality, recall that $Q_0 = |e_0\rangle\langle\theta^0|$, write $\langle y | Q_0 | x \rangle = \langle y | e_0 \rangle \langle \theta^0 | x \rangle$, and so on. The joint probability of configurations x and y at separation L is

$$\begin{aligned} P_L(y, x) &:= P(X_{n+L} = y \& X_n = x) \\ &= \langle \theta^0 | y \rangle \frac{\langle y | T^L | x \rangle}{\lambda_0^L} \langle x | e_0 \rangle \\ &= \langle x | Q_0 | y \rangle \langle y | Q_0 + S^L | x \rangle. \end{aligned} \quad (62)$$

Putting these together, the connected correlation is

$$\begin{aligned} C_L(y, x) &= P_L(y, x) - P(y)P(x) \\ &= \langle x | Q_0 | y \rangle \langle y | S^L | x \rangle. \end{aligned} \quad (63)$$

Also,

$$\Delta_L(y, x) = \frac{\langle y | S^L | x \rangle}{\langle y | Q_0 | x \rangle}. \quad (64)$$

Finally, applying Lemma V.1,

$$I(X_L : X_0) \approx \frac{1}{2} \sum_{y,x \in \mathcal{X}} \frac{\langle x | Q_0 | y \rangle}{\langle y | Q_0 | x \rangle} \langle y | S^L | x \rangle^2. \quad (65)$$

D. End-to-end mutual information

End-to-end mutual information with open boundary conditions in the Perron-Frobenius scenario requires more intricate manipulation, but it is in principle straightforward. Here, L is the length of the chain. Apart from the decomposition $\mathbb{T} = Q_0 + S$, the critical point in the calculations is that $\langle y | Q_0 | x \rangle > 0$ for all $y, x \in \mathcal{X}$. For example, the probability that the configuration is x at the initial or “top” end of the chain is (recall $|\bar{1}\rangle = \sum_{x \in \mathcal{X}} |x\rangle$)

$$\begin{aligned} P_L(x) &= \frac{\langle \bar{1} | T^L | x \rangle}{\langle \bar{1} | T^L | \bar{1} \rangle} = \frac{\langle \bar{1} | Q_0 | x \rangle + \langle \bar{1} | S^L | x \rangle}{\langle \bar{1} | Q_0 | \bar{1} \rangle + \langle \bar{1} | S^L | \bar{1} \rangle} \\ &= \frac{\langle \bar{1} | Q_0 | x \rangle}{\langle \bar{1} | Q_0 | \bar{1} \rangle} \left(1 + \frac{\langle \bar{1} | S^L | x \rangle}{\langle \bar{1} | Q_0 | x \rangle} - \frac{\langle \bar{1} | S^L | \bar{1} \rangle}{\langle \bar{1} | Q_0 | \bar{1} \rangle} \right) + \mathcal{O}(S^{2L}), \end{aligned}$$

where S in $\mathcal{O}(S^{2L})$ stands in for the spectral radius of S . In similar fashion, we obtain the probability that the terminal, or bottom, end has configuration y as

$$P_L(y) = \frac{\langle y | Q_0 | \bar{1} \rangle}{\langle \bar{1} | Q_0 | \bar{1} \rangle} \left(1 + \frac{\langle y | S^L | \bar{1} \rangle}{\langle y | Q_0 | \bar{1} \rangle} - \frac{\langle \bar{1} | S^L | \bar{1} \rangle}{\langle \bar{1} | Q_0 | \bar{1} \rangle} \right) + \mathcal{O}(S^{2L}),$$

and the joint probability of x and y at the ends as

$$\begin{aligned} P_L(y, x) &= \frac{\langle y | T^L | x \rangle}{\langle \bar{1} | T^L | \bar{1} \rangle} \\ &= \frac{\langle y | Q_0 | x \rangle}{\langle \bar{1} | Q_0 | \bar{1} \rangle} \left(1 + \frac{\langle y | S^L | x \rangle}{\langle y | Q_0 | x \rangle} - \frac{\langle \bar{1} | S^L | \bar{1} \rangle}{\langle \bar{1} | Q_0 | \bar{1} \rangle} \right) + \mathcal{O}(S^{2L}). \end{aligned}$$

For $\Delta(y, x)$, the ratio of the leading terms in $P_L(y, x)$ and $P_L(y)P_L(x)$ is needed. With the same trick as for Eq. (61), it is

$$\frac{\langle y | Q_0 | x \rangle \langle \bar{1} | Q_0 | \bar{1} \rangle}{\langle y | Q_0 | \bar{1} \rangle \langle \bar{1} | Q_0 | x \rangle} = 1. \quad (66)$$

Therefore,

$$\Delta_L(y, x) \approx \frac{\langle y | S^L | x \rangle}{\langle y | Q_0 | x \rangle} - \frac{\langle y | S^L | \bar{1} \rangle}{\langle y | Q_0 | \bar{1} \rangle} - \frac{\langle \bar{1} | S^L | x \rangle}{\langle \bar{1} | Q_0 | x \rangle} + \frac{\langle \bar{1} | S^L | \bar{1} \rangle}{\langle \bar{1} | Q_0 | \bar{1} \rangle}.$$

Finally, applying Lemma V.1 again,

$$I(X_L : X_0) \approx \frac{1}{2} \sum_{y,x} \frac{\langle y | Q_0 | x \rangle}{\langle \bar{1} | Q_0 | \bar{1} \rangle} \Delta_L(y, x)^2. \quad (67)$$

No possibility for general simplification is evident here, and the result is considerably more complicated than that for ring-to-ring mutual information in an infinite chain.

VI. END-TO-END MUTUAL INFORMATION ON FINITE-LENGTH CYLINDERS

The previous section showed that end-to-end mutual information $I(X_L : X_0)$ is asymptotically proportional to $|\lambda_1/\lambda_0|^{2L}$ in the Perron-Frobenius scenario. Finally, in this section, we get down to the business of understanding its behavior for the zero-temperature cylindrical TIAFM. Barring systems with zero-energy modes, for which λ_0 has algebraic multiplicity two, we show that $I(X_L : X_0) \approx A(\lambda_1/\lambda_0)^L$, for some amplitude A . Section VIB explains why MI cannot fall off any faster than this, and Sec. VIC presents the more technical calculation of the amplitude A , in the process showing that $(\lambda_1/\lambda_0)^L$ gives the precise decay rate. These results are compared to direct numerical calculations which reveal that the asymptotic regime is reached already at L equal to two or three times the circumference C . Section VID then gives an argument like that of Sec. VIB, showing that $I(X_L : X_0) \sim L^{-2}$ for systems with zero-energy modes and the conclusion is again confirmed with direct numerical calculation.

Before any of that, we need to understand λ_0 and λ_1 . Since $\mathbb{T}$ has the same eigenvalues as $\mathbb{T}_0$, this can be addressed by an examination of noninteracting fermions. Section VIA shows that the lowest-energy excitation are either of two-particle (pp) or two-hole (hh) type, rather than particle-hole type. This provides a simple explanation of the period-three oscillation of the decay rate (see Table I). If, somehow, the lowest energy excitations were of particle-hole type, then the Perron-Frobenius behavior $|\lambda_1/\lambda_0|^{2L}$ (note, $2L$ rather than L) would result.

TABLE I. fundamental energy gaps ΔE of $\mathbb{T}_0$ and corresponding excitation types. “hh” and “pp” indicate excitations involving removal (addition) of two particles. Energies are reported in units of $v_F \cdot \delta k = \pi\sqrt{3}/C$, according to a linearized approximation. Due to the strict convexity of $\varepsilon(k)$, energies reported for hh (pp) excitations are overestimates (underestimates), improving as $C \rightarrow \infty$. There is only one cylinder with $C \geq 3$ for which the linearized approximation (with convexity tie-breaker) leads to an incorrect conclusion: $C = 5$ and odd $\mathcal{N}$ has exactly degenerate hh and pp excitations.

$\mathcal{N}$ parity	Even			Odd		
$C \bmod 3$	0	1	2	0	1	2
Excitation type	hh	pp	hh	hh/pp	hh	pp
$C\Delta E/(\pi\sqrt{3})$	1	1/3	1/3	0	2/3	2/3

A. Energy gaps

Recall that according to Eq. (17), allowed single-particle momenta satisfy

$$\frac{k}{\delta} \in \begin{cases} \mathbb{Z}, & \mathcal{N} \text{ odd}, \\ \mathbb{Z} + \frac{1}{2}, & \mathcal{N} \text{ even}, \end{cases} \quad (68)$$

where $\delta = 2\pi/C$ is the finesse of the momentum spectrum. The maximum eigenvalue λ_0 of $\mathbb{T}_0$ is e^{-E_0} , where E_0 is the energy of the filled Fermi sea, with every negative-energy mode occupied. Linearizing the spectrum near the Fermi momentum $\pm k_F = \pm \frac{2\pi}{3}$ gives a minimum particle-hole excitation energy $v_F \delta k$, where $v_F = d\varepsilon/dk|_{k_F} = \sqrt{3}/2$ is the Fermi velocity. For some $0 < \alpha < 1$, k_F is $\alpha \delta k$ above the highest negative-energy mode. A minimal energy two-hole excitation (removal of two particles) costs energy $2\alpha v_F \delta k$, while a minimal two-particle excitation costs $2(1 - \alpha)v_F \delta k$. As long as $\alpha \neq \frac{1}{2}$, one of these costs less energy than a particle-hole excitation, specifically, the two-hole (hh) excitation if $\alpha < \frac{1}{2}$, and the two-particle (pp) excitation, if $\alpha > \frac{1}{2}$. If $\alpha = \frac{1}{2}$, then there is a three-way tie, in linearized approximation. Considering the positive curvature of the dispersion curve $\varepsilon(k)$, however, the hh excitation gains an advantage. The possibility of α very slightly greater than $\frac{1}{2}$ need not concern us, as it will not arise. The linearized approximation is *a priori* appropriate for large circumference C . Its predictions fail only in the case $C = 5$, odd $\mathcal{N}$, where the hh and pp excitations are degenerate.

Now we will examine which case arises according to C and $\mathcal{N}$ parity. Write

$$C = 3m + p, \quad m, p \in \mathbb{N}. \quad (69)$$

Then, for odd $\mathcal{N}$, Eq. (68) tells us that α in the above discussion is

$$\frac{k_F}{\delta k} - m = \begin{cases} 0 & p = 0 \\ \frac{1}{3} & p = 1 \\ \frac{2}{3} & p = 2 \end{cases} \quad (70)$$

This gives us three entries in Table I. The first case is the special one of zero-energy modes, with a degenerate largest

eigenvalue. For even $\mathcal{N}$, α is

$$\frac{k_F}{\delta k} - \left(m + \frac{1}{2}\right) = \begin{cases} -\frac{1}{2} & p = 0 \\ -\frac{1}{6} & p = 1, \\ +\frac{1}{6} & p = 2 \end{cases} \quad (71)$$

providing the other three entries.

B. Lower bound from data-processing inequality

Combining the information we now have about the nature of the eigenstates of $\mathbb{T}$ with a basic tool of information theory called the data-processing inequality [4], we can show that $I(X_L : X_0)$, the end-to-end mutual information with open boundary conditions, obeys

$$I(X_L : X_0) \gtrsim (\lambda_1/\lambda_0)^L. \quad (72)$$

It should seem plausible, and is true, that “ $\gtrsim$ ” can be replaced by “ $\sim$,” but that requires a more delicate analysis carried out in the following subsection.

The data-processing inequality says that if X' is a function of X , then $I(X' : Y) \leq I(X : Y)$. The intuitive ground for this is simple: $I(X' : Y)$ measures the information X' carries about Y , but X carries at least as much, since it determines X' . Now, we define random variables S_0 and S_L , taking values in $\{\text{LT}, \text{EQ}, \text{GT}\}$. S_0 is determined by X_0 according to

$$S_0 = \begin{cases} \text{LT} & \mathcal{N}(X_0) < N_0 \\ \text{EQ} & \mathcal{N}(X_0) = N_0 \\ \text{GT} & \mathcal{N}(X_0) > N_0 \end{cases} \quad (73)$$

and S_L in the same way from X_L . Then, by two applications of the data-processing inequality [4], and using symmetry of the mutual information,

$$I(X_0 : X_L) \geq I(X_0 : S_L) \geq I(S_0 : S_L). \quad (74)$$

We now show how to get Eq. (72) from this.

Consider the case that the eigenvalue λ_1 of $\mathbb{T}_0$ corresponds to a two-hole excitation, and use the first form for mutual information in Eq. (56) to obtain

$$\begin{aligned} I(S_0 : S_L) &= \sum_x \mathbf{P}(S_0 = x) [H(S_L) - H(S_L | S_0 = x)] \\ &\geq \mathbf{P}(S_0 = \text{LT}) [H(S_L) - H(S_L | S_0 = \text{LT})] \\ &= \mathbf{P}(S_0 = \text{LT}) H(S_L) \sim \left(\frac{\lambda_1}{\lambda_0}\right)^L. \end{aligned}$$

Passage to the final line uses that $\mathcal{N}(X_L) \leq \mathcal{N}(X_0)$. For the last step, $H(S_L)$ tends to some nonzero value as $L \rightarrow \infty$, and $\mathbf{P}(S_0 = \text{LT}) \sim (\lambda_1/\lambda_0)^L$ because λ_1 corresponds to a two-hole excitation. Therefore, we have the claimed bound in this case.

The case that λ_1 corresponds to a two-particle excitation proceeds similarly, but with the roles of the ends swapped, because we need to use that $S_L = \text{GT}$ implies $S_0 = \text{GT}$:

$$I(S_0 : S_L) \geq \mathbf{P}(S_L = \text{GT}) H(S_0) \sim \left(\frac{\lambda_1}{\lambda_0}\right)^L.$$

C. Open boundary end-to-end mutual information: No zero-energy modes

Now we show that

$$I(X_L : X_0) \approx A \left(\frac{\lambda_1}{\lambda_0} \right)^L \quad (75)$$

and calculate A .

1. Categories of end configurations

In the treatment of the asymptotic behavior of mutual information

$$I(X_L : X_0) = \sum_{y,x} P_L(y, x) \ln \frac{P_L(y, x)}{P_L(y)P_L(x)} \quad (76)$$

in the Perron-Frobenius scenario, given in Sec. V, a crucial role was played by the fact that for every x and y ,

$$\Delta(y, x) = \frac{P_L(y, x)}{P_L(y)P_L(x)} - 1 \quad (77)$$

tended to zero as $L \rightarrow \infty$. Now, for the zero-temperature cylindrical TIAFM, this continues to hold in case $\mathcal{N}(y) \leq N_0 \leq \mathcal{N}(x)$, because $\langle y|Q_0|x \rangle \neq 0$ in that case. The argument is just the same as for Eq. (66).

However, in case $N_0 < \mathcal{N}(y) \leq \mathcal{N}(x)$, $\langle y|Q_0 = 0$, so that

$$P_L(y, x) \approx \frac{\langle y|S^L|x \rangle}{\langle \bar{1}|Q_0|\bar{1} \rangle} \approx \frac{\langle y|e_{\mathcal{N}(y),0} \rangle \langle \theta_{\mathcal{N}(y),0}|x \rangle}{\langle \bar{1}|Q_0|\bar{1} \rangle} \left(\frac{\lambda_{\mathcal{N}(y),0}}{\lambda_0} \right)^L, \quad (78)$$

$$\frac{P_L(y, x)}{P_L(y)} \approx \frac{\langle y|S^L|x \rangle}{\langle y|S^L|\bar{1} \rangle} \approx \frac{\langle \theta_{\mathcal{N}(y),0}|x \rangle}{\langle \theta_{\mathcal{N}(y),0}|\bar{1} \rangle}, \quad (79)$$

and

$$P_L(x) \approx \frac{\langle \bar{1}|Q_0|x \rangle}{\langle \bar{1}|Q_0|\bar{1} \rangle} = \frac{\langle \theta_0|x \rangle}{\langle \theta_0|\bar{1} \rangle}. \quad (80)$$

Together, Eqs. (79) and (80) yield

$$\frac{P_L(y, x)}{P_L(y)P_L(x)} \approx \frac{\langle \theta_{\mathcal{N}(y),0}|x \rangle}{\langle \theta_{\mathcal{N}(y),0}|\bar{1} \rangle} \frac{\langle \theta_0|\bar{1} \rangle}{\langle \theta_0|x \rangle}. \quad (81)$$

Similarly, in case $\mathcal{N}(y) \leq \mathcal{N}(x) < N_0$, $Q_0|x \rangle = 0$,

$$P_L(y, x) \approx \frac{\langle y|e_{\mathcal{N}(x),0} \rangle \langle \theta_{\mathcal{N}(x),0}|x \rangle}{\langle \bar{1}|Q_0|\bar{1} \rangle} \left(\frac{\lambda_{\mathcal{N}(x),0}}{\lambda_0} \right)^L, \quad (82)$$

and

$$\frac{P_L(y, x)}{P_L(y)P_L(x)} \approx \frac{\langle y|e_{\mathcal{N}(x),0} \rangle \langle \bar{1}|e_0 \rangle}{\langle \bar{1}|e_{\mathcal{N}(x),0} \rangle \langle y|e_0 \rangle}. \quad (83)$$

Since Eqs. (81) and (83) are nonzero and bounded, their logarithms are bounded. Therefore, in these cases, the contribution to the sum in Eq. (76) is of the same order as $P_L(y, x)$. Thus, if we wish to calculate $I(X_L : X_0)$ to order $(\lambda_1/\lambda_0)^L$, then we need only consider y with $\langle y|e_1 \rangle \neq 0$ or x with $\langle \theta_1|x \rangle \neq 0$, in addition to the first class $\mathcal{N}(y) \leq N_0 \leq \mathcal{N}(x)$. However, the contribution of this class vanishes to this order, as we show next.

Lemma VI.1.

$$\sum_{\mathcal{N}(y) \leq N_0 \leq \mathcal{N}(x)} P_L(y, x) \ln \frac{P_L(y, x)}{P_L(y)P_L(x)} \sim \left(\frac{\lambda_1}{\lambda_0} \right)^{2L}.$$

Proof. As in Lemma V.1, up to order $(\lambda_1/\lambda_0)^{2L}$, the sum is

$$\begin{aligned} & \sum_{\mathcal{N}(y) \leq N_0 \leq \mathcal{N}(x)} P_L(y)P_L(x)\Delta_L(y, x) \\ &= \sum [P_L(y, x) - P_L(y)P_L(x)] \\ &= P_L[\{\mathcal{N}(y) \leq N_0\} \cap \{N_0 \leq \mathcal{N}(x)\}] \\ &\quad - P_L[\mathcal{N}(y) \leq N_0]P_L[N_0 \leq \mathcal{N}(x)]. \end{aligned}$$

Now, we appeal a small general probabilistic identity:

$$P(A \cap B) - P(A)P(B) = P(A^c \cap B^c) - P(A^c)P(B^c). \quad (84)$$

Given that, our sum is asymptotically equal to

$$\begin{aligned} & P_L[(N_0 < \mathcal{N}(y)) \cap \{\mathcal{N}(x) < N_0\}] \\ &\quad - P_L[N_0 < \mathcal{N}(y)]P_L[\mathcal{N}(x) < N_0]. \end{aligned}$$

However, the first term here is zero, while both $P_L[N_0 < \mathcal{N}(y)]$ and $P_L[\mathcal{N}(x) < N_0]$ are of order $(\lambda_1/\lambda_0)^L$.

To finish, we just need to demonstrate the identity Eq. (84):

$$\begin{aligned} P(A \cap B) - P(A)P(B) &= P(B) - P(A^c \cap B) - P(B)(1 - P(A^c)) \\ &= -[P(A^c \cap B) - P(A^c)P(B)]. \end{aligned}$$

Repeat, with B and A^c in the roles of A and B .

2. The amplitude

To recapitulate: Section VIB showed that $I(X_L : X_0)$ is of order $(\lambda_1/\lambda_0)^L$ at least, while Lemma VI.1 shows that pairs (y, x) with $\mathcal{N}(y) \leq N_0 \leq \mathcal{N}(x)$ contribute nothing to that order. Estimates Eqs. (81) and (83) show that the contribution of other pairs is of the same order as $P_L(y, x)$. Eq. (78) or (80) shows which (y, x) pairs to retain.

In case λ_1 corresponds to a two-particle excitation, the relevant contributions come from $N_0 + 2 = \mathcal{N}(y) \leq \mathcal{N}(x)$. We have only to put together the expressions from Eqs. (78) and (81):

$$A = \sum_{\substack{\mathcal{N}(y)=N_0+2 \\ \mathcal{N}(x) \geq \mathcal{N}(y)}} \frac{\langle y|e_1 \rangle \langle \theta_1|x \rangle}{\langle \bar{1}|e_0 \rangle \langle \theta_0|\bar{1} \rangle} \ln \frac{\langle \theta_1|x \rangle \langle \theta_0|\bar{1} \rangle}{\langle \theta_1|\bar{1} \rangle \langle \theta_0|x \rangle}. \quad (85)$$

In case λ_1 corresponds to a two-hole excitation, use Eqs. (82) and (83) instead:

$$A = \sum_{\substack{\mathcal{N}(x)=N_0-2 \\ \mathcal{N}(y) \leq \mathcal{N}(x)}} \frac{\langle y|e_1 \rangle \langle \theta_1|x \rangle}{\langle \bar{1}|e_0 \rangle \langle \theta_0|\bar{1} \rangle} \ln \frac{\langle y|e_1 \rangle \langle \bar{1}|e_0 \rangle}{\langle \bar{1}|e_1 \rangle \langle y|e_0 \rangle}. \quad (86)$$

These results for the amplitude A for C up to 15 are plotted (data points) in Fig. 4. Within each congruence class of C modulo 3 and parity of $\mathcal{N}$, the calculated values appear to behave nearly exponentially in C . In contrast to the dependence of λ_1/λ_0 on $C \bmod 3$, we have no explanation for this.

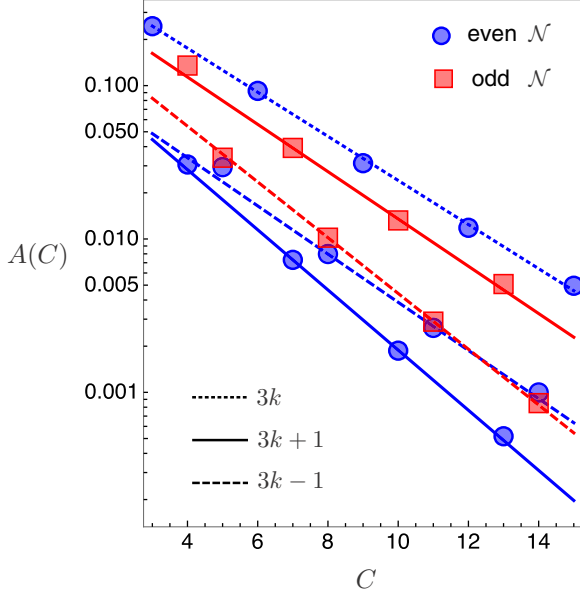


FIG. 4. Amplitudes $A(C)$ [see Eqs. (85) and (86)] of the asymptotic decay of end-to-end mutual information [9]. The lines are empirical fits revealing an approximately exponential decay of A with circumference C for fixed residue class $C \bmod 3$ and $\mathcal{N}$ parity.

3. Exact numerical calculations

Compared to Exact numerical calculations of the end-to-end mutual information are compared to the asymptotic formulas in Fig. 5. The plots show that the asymptotic regime is attained already when L is two or three times C . The data show that $I(X_L : X_0)$ always approaches the asymptotic value from above. Since terms in Eq. (76) can be negative as well as positive, it is not clear why this is the case.

D. Systems with zero-energy modes

Finally, we analyze systems with zero-energy modes. The exceptional nature of these systems with regard to end-to-end mutual information is demonstrated in Fig. 6, which shows numerical results for the rescaled mutual information. Evidently, $I(X_L : X_0)$ decays not exponentially, but as L^{-2} . The task of this section is to derive this behavior.

According to Proposition IV.1, the partition function of a length- L system with open boundary conditions is of order $L\lambda_0^L$. $P_L(y, x)$ tends to zero exponentially if $\mathcal{N}(y) \leq \mathcal{N}(x) < N_0 - 2$ or $N_0 < \mathcal{N}(y) \leq \mathcal{N}(x)$, tends to zero as $1/L$ if $\mathcal{N}(y) = \mathcal{N}(x)$ is $N_0 - 2$ or N_0 , and tends to a nonzero value if $\mathcal{N}(y) \leq N_0 - 2 < N_0 \leq \mathcal{N}(x)$. Analysis of the first case is very similar to that for systems without zero modes, so we shall not go into details. For such a pair (y, x) , $\frac{P_L(y, x)}{P_L(y)P_L(x)}$ is bounded, hence its contribution to the mutual information is of order $P_L(y, x)$, that is, exponentially small.

The important configurations, then, are those where $\mathcal{N}(y) \leq N_0$ and $N_0 - 2 \leq \mathcal{N}(x)$. As we did for systems without zero-energy modes, we introduce reduced variables

$$S_0 = \begin{cases} \text{LO} & \mathcal{N}(X_0) \leq N_0 - 2 \\ \text{HI} & \mathcal{N}(X_0) \geq N_0 \end{cases} \quad (87)$$

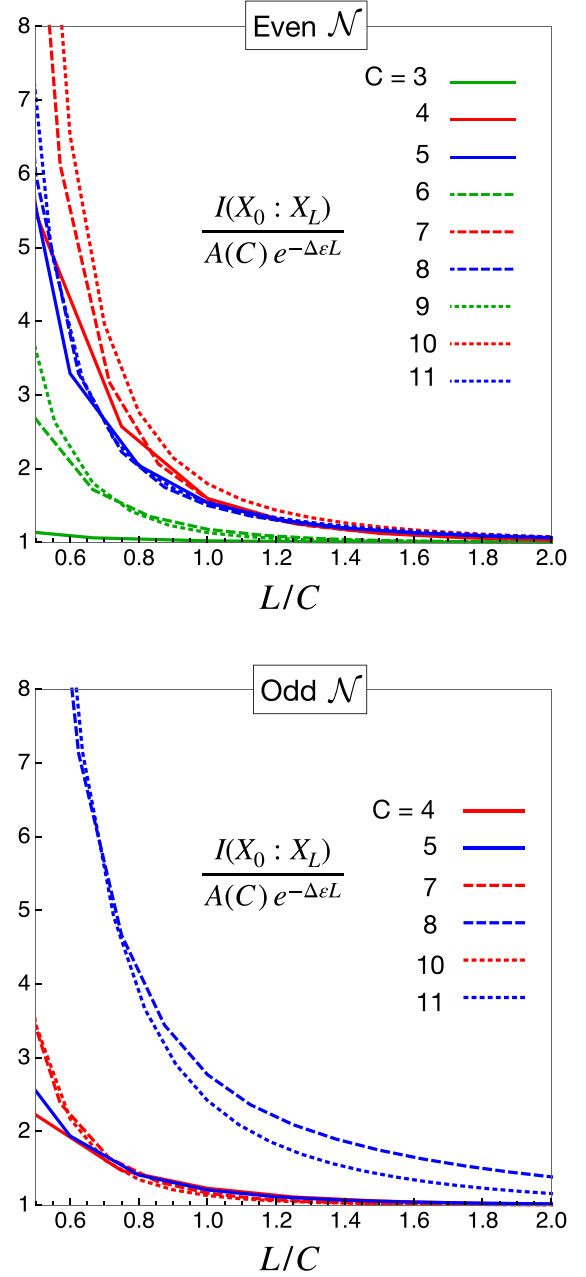


FIG. 5. Ratio of the end-to-end mutual information $I(X_L : X_0)$ to the leading behavior $A(C)e^{-\Delta\epsilon L}$, for cases without zero-energy modes and $3 \leq C \leq 11$.

and similarly S_L defined in terms of X_L . Since the data processing inequality assures that

$$I(X_L : X_0) \geq I(S_L : S_0), \quad (88)$$

we aim to show that $I(S_L : S_0) \sim L^{-2}$. To that end, we need to determine the asymptotic joint distribution of the random variables S_L and S_0 . The partition function is dominated by configurations with $\mathcal{N}$ equal in the bulk to either N_0 or $N_0 - 2$. These have the same entropy density, that is, correspond to the same eigenvalue of $\mathbb{T}_0$. However, one gains additional entropy with the possibility of inserting a transition (“domain wall”) between the two. It can go in any of order L locations, not

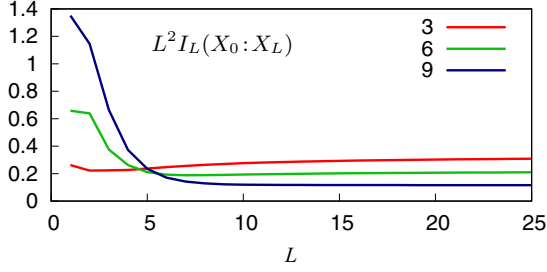


FIG. 6. $L^2 I_L(X_0 : X_L)$, the end-to-end mutual information multiplied by L^2 for the exceptional cases of odd- N , $C \in 3\mathbb{N}$, which have zero-energy modes [9].

be present at all, in which case $\mathcal{N}(y) = \mathcal{N}(x)$ is either N_0 or $N_0 - 2$. Simply counting possibilities therefore leads to

$$\begin{aligned} P_L(S_L = \text{HI}, S_0 = \text{LO}) &= o(L^{-1}), \\ P_L(S_L = \text{HI}, S_0 = \text{HI}) &= \frac{c}{L} + o(L^{-1}), \\ P_L(S_L = \text{LO}, S_0 = \text{LO}) &= \frac{c}{L} + o(L^{-1}), \\ P_L(S_L = \text{LO}, S_0 = \text{HI}) &= 1 - 2\frac{c}{L} + o(L^{-1}). \end{aligned} \quad (89)$$

According to the following Lemma, this leads to the conclusion $I(X_L : X_0) \geq I(S_L : S_0) \approx (c/L)^2$, as required.

Lemma VI.2. Suppose a pair (U, V) of dichotomous random variables has joint and marginal probabilities as in the table

δ	$1 - 2\delta$	$1 - \delta$
0	δ	δ
δ	$1 - \delta$	

Then,

$$I(U : V) = \delta^2 + O(\delta^3). \quad (90)$$

Proof. Compute:

$$\begin{aligned} I(U : V) &= \sum_{u,v} P(u, v) \ln \frac{P(u, v)}{P(u)P(v)} \\ &= 2\delta \ln \frac{1}{1 - \delta} + (1 - 2\delta) \ln \frac{1 - 2\delta}{(1 - \delta)^2} \\ &= (1 - 2\delta) \ln(1 - 2\delta) - 2(1 - \delta) \ln(1 - \delta). \end{aligned}$$

■

VII. CONCLUSION

The three main inter-related themes in this paper have been breakdown of the Perron-Frobenius scenario in one-dimensional systems dominated by nonthermal disorder, the use of mutual information to study correlations between complicated elementary degrees of freedom, and the use of a

fermionic representation to extract detailed information about the zero-temperature triangular Ising antiferromagnet on long cylinders.

The Perron-Frobenius scenario characterizes such basic statistical mechanical properties of thermally disordered one-dimensional systems as asymptotic behavior of correlation functions, and as we have shown, mutual information. Qualitatively different properties can result when the Perron-Frobenius condition on the transfer matrix fails. A natural way to obtain nontrivial transfer matrices for which that failure happens is through frustration-induced disorder. Thermal disorder is completely excluded at zero temperature, but frustration can still be the dominant effect over relatively long length scales for low enough temperature. Cylindrical TIAFM systems serve as an excellent model on which to study the breakdown of the Perron-Frobenius scenario. The TIAFM is of great interest for its own sake, and can be studied in great detail by use of a powerful mapping to a system of fermions. Considered as a one-dimensional chain, each site of which is a ring of spins, the “elementary” degrees of freedom are very complicated. Traditional correlation function techniques are not adequate for this situation. We therefore use mutual information instead as the primary tool with which to describe correlations. Coupled with the fermionic representation, we are able to obtain asymptotic behavior of end-to-end mutual information. Phenomena which are contrary to the Perron-Frobenius (PF) scenario are thereby uncovered, such as decay lengths half what that scenario predicts, or even infinite. Other features, which are not contrary to the PF scenario, but are nevertheless highly surprising, such as oscillation of the mutual information decay length with a period of three in the cylinder circumference, are also completely explained. Some puzzles remain in the details of the results, such as the clustering seen in the curves of Fig. 4 and the fact that the asymptote is always approached from above in Fig. 5.

Direct experimental study of mutual information is very demanding, since one needs access to microstate details. Systems where the relevant degrees of freedom are molecular or atomic scale are therefore very difficult to access. Mesoscopic systems such as artificial spin ice or colloidal systems, however, seem relatively promising. It may even be possible to fabricate fairly precise realizations of the specific model studied here.

ACKNOWLEDGMENTS

This project was funded by the U.S. Department of Energy, Office of Basic Energy Sciences, Materials Sciences and Engineering Division under Grant No. DE-SC0010778, and by the National Science Foundation under Awards No. DMR-1420620 and No. DMR-2011839. A.N. acknowledges funding from the University of Akron.

- [1] C. E. Shannon and W. Weaver, *The Mathematical Theory of Communication* (University of Illinois Press, Urbana, IL, 1949).
- [2] P. Billingsley, *Ergodic Theory and Information* (John Wiley & Sons, New York/London/Sydney, 1965).

- [3] I. Csiszár and J. Körner, *Information Theory* (Academic Press, New York/London, 1981).
- [4] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, Wiley Series in Telecommunications (John Wiley & Sons, New York, 1991).

- [5] H. W. Lau and P. Grassberger, *Phys. Rev. E* **87**, 022128 (2013).
- [6] J. Wilms, M. Troyer, and F. Verstraete, *J. Stat. Mech. Theory Exper.* (2011) P10011.
- [7] O. Melchert and A. K. Hartmann, *Phys. Rev. E* **87**, 022107 (2013).
- [8] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2000).
- [9] A. Nourhani, V. H. Crespi, and P. E. Lammert, *Phys. Rev. E* **98**, 032107 (2018).
- [10] L. Pauling, *J. Am. Chem. Soc.* **57**, 2680 (1935).
- [11] W. F. GIAUQUE and J. W. Stout, *J. Am. Chem. Soc.* **58**, 1144 (1936).
- [12] G. Toulouse, *Commun. Phys.* **2**, 115 (1977).
- [13] R. Moessner, *Can. J. Phys.* **79**, 1283 (2001).
- [14] B. Normand, *Contemp. Phys.* **50**, 533 (2009).
- [15] M. J. P. Gingras and P. A. McClarty, *Rep. Prog. Phys.* **77**, 056501 (2014).
- [16] O. A. Starykh, *Rep. Prog. Phys.* **78**, 052502 (2015).
- [17] B. Schmidt and P. Thalmeier, *Phys. Rep. Rev. Section Phys. Lett.* **703**, 1 (2017).
- [18] R. F. Wang, C. Nisoli, R. S. Freitas, J. Li, W. McConville, B. J. Cooley, M. S. Lund, N. Samarth, C. Leighton, V. H. Crespi, and P. Schiffer, *Nature (London)* **439**, 303 (2006).
- [19] S. Zhang, J. Li, I. Gilbert, J. Bartell, M. J. Erickson, Y. Pan, P. E. Lammert, C. Nisoli, K. K. Kohli, R. Misra, V. H. Crespi, N. Samarth, C. Leighton, and P. Schiffer, *Phys. Rev. Lett.* **109**, 087201 (2012).
- [20] Y. Perrin, B. Canals, and N. Rougemaille, *Nature (London)* **540**, 410 (2016).
- [21] P. Tierno, *Phys. Rev. Lett.* **116**, 038303 (2016).
- [22] Y. Han, Y. Shokef, A. M. Alsayed, P. Yunker, T. C. Lubensky, and A. G. Yodh, *Nature (London)* **456**, 898 (2008).
- [23] S. Mahmoudian, L. Rademaker, A. Ralko, S. Fratini, and V. Dobrosavljević, *Phys. Rev. Lett.* **115**, 025701 (2015).
- [24] M. Weigt and A. Hartmann, *Europhys. Lett.* **62**, 533 (2003).
- [25] N. Choudhury, L. Walizer, S. Lisenkov, and L. Bellaiche, *Nature (London)* **470**, 513 (2011).
- [26] M. Nixon, E. Ronen, A. A. Friesem, and N. Davidson, *Phys. Rev. Lett.* **110**, 184102 (2013).
- [27] A. L. Wang, J. M. Gold, N. Tompkins, M. Heymann, K. I. Harrington, and S. Fraden, *European Phys. J. Special Topics* **225**, 211 (2016).
- [28] D. A. Lavis and G. M. Bell, *Statistical Mechanics of Lattice Systems I*, 2nd ed. (Springer-Verlag, Berlin, 1999).
- [29] B. M. McCoy, *Advanced Statistical Mechanics*, International Series of Monographs on Physics, Vol. 146 (Oxford University Press, Oxford, UK, 2010), pp. xvi+624.
- [30] G. Mussardo, *Statistical Field Theory*, Oxford Graduate Texts (Oxford University Press, Oxford, UK, 2010), pp. xxii+755.
- [31] D. Ruelle, *Statistical Mechanics: Rigorous Results* (W. A. Benjamin, Inc., New York/Amsterdam, 1969).
- [32] B. Simon, *The Statistical Mechanics of Lattice Gases*, Vol. I, Princeton Series in Physics (Princeton University Press, Princeton, NJ, 1993).
- [33] A. Katok and B. Hasselblatt, *Introduction to the Modern Theory of Dynamical Systems*, Encyclopedia of Mathematics and its Applications, Vol. 54 (Cambridge University Press, Cambridge, UK, 1995), pp. xviii+802.
- [34] C. Nisoli, N. M. Gabor, P. E. Lammert, J. D. Maynard, and V. H. Crespi, *Phys. Rev. Lett.* **102**, 186103 (2009).
- [35] D. C. Mattis, *Statistical Mechanics Made Simple* (World Scientific Publishing, River Edge, NJ, 2003).
- [36] R. J. Baxter, *Exactly Solved Models in Statistical Mechanics* (Dover, Mineola, NY, 2007).