



HAL
open science

Entropy constraints for ground energy optimization

Hamza Fawzi, Omar Fawzi, Samuel Scalet

► **To cite this version:**

Hamza Fawzi, Omar Fawzi, Samuel Scalet. Entropy constraints for ground energy optimization. Journal of Mathematical Physics, 2024, 65 (3), pp.032201. 10.1063/5.0159108 . hal-04682903

HAL Id: hal-04682903

<https://hal.science/hal-04682903v1>

Submitted on 8 Oct 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

RESEARCH ARTICLE | MARCH 01 2024

Entropy constraints for ground energy optimization

Hamza Fawzi ; Omar Fawzi ; Samuel O. Scalet  



J. Math. Phys. 65, 032201 (2024)

<https://doi.org/10.1063/5.0159108>



Articles You May Be Interested In

Approximating the set of separable states using the positive partial transpose test

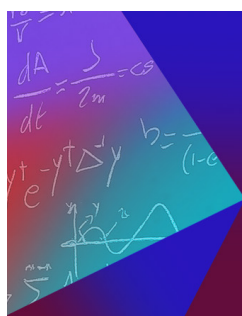
J. Math. Phys. (April 2010)

Extending quantum operations

J. Math. Phys. (October 2012)

Bounds on localizable information via semidefinite programming

J. Math. Phys. (August 2005)



Journal of Mathematical Physics

Young Researcher Award:
Recognizing the Outstanding Work
of Early Career Researchers

[Learn More!](#)

Entropy constraints for ground energy optimization

Cite as: J. Math. Phys. 65, 032201 (2024); doi: 10.1063/5.0159108

Submitted: 20 May 2023 • Accepted: 29 January 2024 •

Published Online: 1 March 2024



Hamza Fawzi,¹ Omar Fawzi,² and Samuel O. Scalet^{1,a)}

AFFILIATIONS

¹Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Cambridge, United Kingdom

²Inria, ENS Lyon, UCBL, LIP, Univ Lyon, Lyon, France

^{a)} Author to whom correspondence should be addressed: sos25@cam.ac.uk

ABSTRACT

We study the use of von Neumann entropy constraints for obtaining lower bounds on the ground energy of quantum many-body systems. Known methods for obtaining certificates on the ground energy typically use consistency of local observables and are expressed as semidefinite programming relaxations. The local marginals defined by such a relaxation do not necessarily satisfy entropy inequalities that follow from the existence of a global state. Here, we propose to add such entropy constraints that lead to tighter convex relaxations for the ground energy problem. We give analytical and numerical results illustrating the advantages of such entropy constraints. We also show limitations of the entropy constraints we construct: they are implied by doubling the number of sites in the relaxation and as a result they can at best lead to a quadratic improvement in terms of the matrix sizes of the variables. We explain the relation to a method for approximating the free energy known as the Markov Entropy Decomposition method.

© 2024 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>). <https://doi.org/10.1063/5.0159108>

I. INTRODUCTION

A fundamental computational problem in quantum many-body theory is to compute the ground energy of local Hamiltonians. Consider a multipartite Hilbert space $\mathcal{H} = \otimes_{v \in V} \mathbb{C}^d$ with local dimension d , on a finite set of sites V . A k -local Hamiltonian, is a Hermitian operator on \mathcal{H} defined as

$$H = \sum_{A \in \binom{V}{k}} h_A, \quad (1)$$

where each of the h_A is a Hermitian operator acting nontrivially only on the set $A \subset V$, and $\binom{V}{k}$ is the set of subsets of V of size k . In this paper, we will be mostly interested in 2-local Hamiltonians, where the interaction can be modeled by a graph $G = (V, E)$ on the set of sites V , and where a Hamiltonian term h_{ij} is attached to each edge $ij \in E$:

$$H = \sum_{ij \in E} h_{ij}. \quad (2)$$

The ground energy of H is defined as its smallest eigenvalue. Due to the special structure of H , its matrix representation is generally sparse and thus one can apply standard methods such as Lanczos iterations¹ to compute its minimal eigenvalue. However, since the dimension of H grows exponentially with $|V|$, this is only feasible for moderate values of $|V|$. It is of considerable theoretical and practical interest to find efficient algorithms, that scale polynomially in $|V|$, to approximate the ground energy of local Hamiltonians.^{2,3}

The smallest eigenvalue of H admits the following variational formulation:

$$\lambda_{\min}(H) = \min_{\psi \in \mathcal{H}} \frac{\langle \psi, H\psi \rangle}{\langle \psi, \psi \rangle}. \quad (3)$$

Variational methods posit a certain form for the state $\psi = \psi_\theta$, and find the value of the parameters θ that minimize the objective function of (3). As such, these methods provide upper bounds on $\lambda_{\min}(H)$. A prominent example are tensor network states,^{4,5} which have been extremely successful and in particular give provably efficient algorithms for gapped systems in one dimension.⁶⁻⁸

Another class of methods that have been studied in the literature are based on convex relaxations and provide lower bounds on $\lambda_{\min}(H)$. For 2-local Hamiltonians H of the form (2), computing the energy $\langle \psi, H\psi \rangle$ only requires knowledge of the two-body marginals ρ_{ij} of $|\psi\rangle\langle\psi|$ for $ij \in E$. If we denote by \mathcal{C} the set of two-body marginals that are consistent with a global state on V , i.e.,

$$\mathcal{C} = \mathcal{C}_{d,V,E} = \left\{ (\rho_{ij})_{ij \in E} : \exists \rho \in \mathcal{D}(\otimes_{v \in V} \mathbb{C}^d), \text{ s.t. } \rho_{ij} = \text{Tr}_{V \setminus \{i,j\}} \rho \right\} \quad (4)$$

where $\mathcal{D}(\mathcal{H})$ denotes the set of density operators on a Hilbert space \mathcal{H} , then one can write the ground energy problem for a 2-local Hamiltonian (2) as a linear optimization problem over \mathcal{C} :

$$\lambda_{\min}(H) = \min_{(\rho_{ij}) \in \mathcal{C}} \sum_{ij \in E} \text{Tr} [h_{ij} \rho_{ij}]. \quad (5)$$

To make this approach tractable, it is required to have a computationally efficient representation of the convex set \mathcal{C} . Unfortunately, it is highly likely that \mathcal{C} does not have any simple representation, e.g., it is known that the problem of checking membership in $\mathcal{C}_{2,[n],\binom{[n]}{2}}$ is QMA-hard.⁹

Rather than aiming to describe \mathcal{C} , we are interested in constructing efficient outer relaxations of \mathcal{C} , i.e., tractable convex sets $\tilde{\mathcal{C}}$ such that $\mathcal{C} \subset \tilde{\mathcal{C}}$. Replacing \mathcal{C} by $\tilde{\mathcal{C}}$ in (5) would then yield a lower bound on $\lambda_{\min}(H)$. Such relaxations $\tilde{\mathcal{C}}$ can be constructed by identifying *necessary conditions* that any set of marginals (ρ_{ij}) which are globally consistent must satisfy. Most relaxations that have been constructed in the literature are based on semidefinite programming (SDP). We describe here the most popular approaches:

- A simple relaxation can be obtained by simply imposing that the two-body marginals are consistent on the intersection of their supports, i.e., one can take

$$\begin{aligned} \tilde{\mathcal{C}}_E^{\text{loc}} = \{ (\rho_{ij})_{ij \in E} : \rho_{ij} \geq 0, \text{Tr } \rho_{ij} = 1 \quad \forall ij \in E \\ \text{and } \text{Tr}_j \rho_{ij} = \text{Tr}_{j'} \rho_{ij'} \quad \forall ij, ij' \in E \}. \end{aligned} \quad (6)$$

This relaxation can be made tighter by introducing higher-order marginals of ρ , namely one can consider

$$\begin{aligned} \tilde{\mathcal{C}}_l^{\text{loc}} = \left\{ (\rho_{ij})_{ij \in E} : \exists (\rho_S)_{|S| \leq l}, \quad \rho_S \geq 0, \text{Tr } \rho_S = 1 \quad \forall S \in \binom{V}{l} \right. \\ \left. \text{and } \text{Tr}_{S \setminus S'} \rho_S = \text{Tr}_{S' \setminus S} \rho_{S'} \quad \forall S, S' \in \binom{V}{l} \right\}. \end{aligned} \quad (7)$$

It is clear that $\mathcal{C} = \tilde{\mathcal{C}}_N^{\text{loc}} \subset \tilde{\mathcal{C}}_{N-1}^{\text{loc}} \subset \dots \subset \tilde{\mathcal{C}}_2^{\text{loc}} \subset \tilde{\mathcal{C}}_E^{\text{loc}}$, where $N = |V|$.

- The Lasserre/sum-of-squares relaxation¹⁰⁻¹⁴ stems from the observation that if $|\psi\rangle$ is a global state on \mathcal{H} , then $\langle \psi, O^\dagger O \psi \rangle \geq 0$ for any observable O acting on \mathcal{H} . In particular if O is a l -local operator, then $O^\dagger O$ is at most $2l$ -local, and $\langle \psi, O^\dagger O \psi \rangle$ is linear in the expectation values $m_F = \langle \psi, F \psi \rangle$ of $2l$ -local observables F . It turns out that the infinite family of constraints

$$\langle \psi, O^\dagger O \psi \rangle \geq 0 \quad \forall O \text{ } l\text{-local observable on } \mathcal{H} \quad (8)$$

can be encoded as a single positive semidefinite (psd) constraint of a matrix whose entries are linear in the expectation values (m_F) . More precisely the matrix is indexed by strings of product operators, each taken from an operator basis, up to weight l . For two such operators O, P , the (O, P) entry of the matrix is then given by the expectation value of the product $O^\dagger P$. Requiring this matrix to be psd corresponds to the constraint (8). The corresponding relaxation $\tilde{\mathcal{C}}_l^{\text{sos}}$ can then be expressed as

$$\tilde{\mathcal{C}}_l^{\text{sos}} = \left\{ (\rho_{ij})_{ij \in E} : \text{Tr} [\rho_{ij}] = 1, \rho_{ij} = \sum_{\alpha, \beta} m_{\sigma_i^\alpha \sigma_j^\beta} \sigma_i^\alpha \otimes \sigma_j^\beta \text{ and psd constraint (8)} \right\}$$

where σ^α is an orthonormal basis of operators acting on \mathbb{C}^d . Since it allows more flexible choices of operators to be included in the basis, the Lasserre hierarchy is a customary choice for settings without an underlying lattice geometry such as problems from quantum chemistry.

Main contributions. In this paper, we are interested in relaxations for the convex set \mathcal{C} that go beyond semidefinite programming. In particular, we are interested in relaxations that use *entropies* of the local marginals. Recall that the von Neumann entropy of the system A for the state ρ is defined by

$$S(A)_\rho = -\text{Tr } \rho_A \log \rho_A,$$

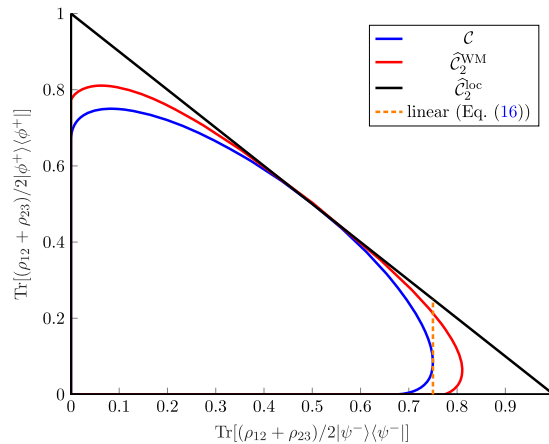


FIG. 1. We compare the sets of valid states on three sites (blue, inner curved line), 2-body marginal relaxations (black triangle), and the entropy constrained marginals fulfilling $S(2|1) + S(2|3) \geq 0$ (red, outer curved line). We also depict the linear inequality mentioned in Eq. (16) (orange, dashed).

where ρ_A denotes the reduced state of ρ on the system A . In addition, the conditional entropy is defined by

$$S(A|B)_\rho = S(AB)_\rho - S(B)_\rho.$$

An important property about the latter is that it is concave in ρ_{AB} ; this follows from the identity

$$S(A|B)_\rho = -D(\rho_{AB} \| \mathbb{1}_A \otimes \rho_B)$$

where $D(\rho \| \sigma) = \text{Tr}[\rho(\log \rho - \log \sigma)]$ is the relative entropy function, which is jointly convex in (ρ, σ) .¹⁵

The main contribution of this paper is to study two families of entropy constraints that yield new strengthened relaxations for the ground energy problem. These relaxations are obtained by imposing inequalities on the conditional entropies of local marginals of the global state, and can be combined with any of the existing semidefinite relaxations. By virtue of the concavity of the conditional entropy function, these relaxations are all convex and can be solved efficiently using tools from convex optimization.^{16,17}

The first family of entropy constraints come from weak monotonicity, see Sec. II, and the second family of constraints, that we call Markov Entropy Decomposition (MED) constraints, are motivated by the work of Poulin and Hastings¹⁸ and are based on combining the chain rule together with strong subadditivity, see Sec. III. In Sec. IV, we consider both families in the special case of infinite lattice systems. Though weak monotonicity constraints are in many cases stronger than MED constraints, we show that in general the two families are not comparable. Our main message is that for many natural Hamiltonians, imposing entropy constraints can lead to significantly tighter bounds compared to simple consistency conditions captured by $\widehat{\mathcal{C}}_l^{\text{loc}}$. Families of examples are presented in Fig. 1 and detailed in Sec. II A and then the Heisenberg XXZ-chain is discussed in Sec. V. We also show limitations on the gains that can be obtained using weak monotonicity constraints (and also MED in many settings): entropy constraints involving l sites are implied by consistency constraints on $2l - 1$ sites, see Eq. (12). As a result, as the size of the matrix variables involved in $\widehat{\mathcal{C}}_l^{\text{loc}}$ is exponential in l , entropy constraints can at most lead to a quadratic improvement in terms of the size of the matrix variables.

We note that weak monotonicity has been mentioned previously as a necessary condition for global consistency in Ref. 19. Another related work considered entropic constraints to the number of orthogonal pure state extensions in a quantum marginal problem.²⁰

An intriguing open question arising from our work is to construct other entropy constraints that could lead to tighter relaxations. This question is related to obtaining inequalities for the so-called quantum entropy cone (see e.g., Ref. 21) though it differs in several respects: in our case, the dimension of the subsystems is fixed, the number of systems involved is bounded by l , and in order to obtain convex relaxations, we look for expressions that are concave in the state ρ , e.g., conical linear combinations of conditional entropies.

II. WEAK MONOTONICITY CONSTRAINTS

We start by recalling the following well-known entropy inequality, also known as weak monotonicity.

Lemma 2.1 (Weak monotonicity). For any state ρ_{ABC} on systems ABC , we have

$$S(A|B)_\rho + S(A|C)_\rho \geq 0. \tag{9}$$

Proof. Let ρ_{ABCD} be a purification of ρ_{ABC} , i.e., ρ_{ABCD} is rank-one and $\text{Tr}_D(\rho_{ABCD}) = \rho_{ABC}$. We have, using the fact that ρ_{ABCD} is pure,

$$S(A|B)_\rho = S(AB)_\rho - S(B)_\rho = S(CD)_\rho - S(ACD)_\rho = -S(A|CD)_\rho \geq -S(A|C)_\rho,$$

where we used strong subadditivity, i.e., the property that $S(A|C) \geq S(A|CD)$. \square

There are two important features of this inequality that we want to highlight: the first one is that it does not involve the global state of ρ on ABC , but only the marginals of ρ on AB and AC . The second important aspect is that the inequality (9) defines a convex region in the space of marginals (ρ_{AB}, ρ_{AC}) . This is a consequence of the concavity of the conditional entropy function. For these reasons, the inequality can be used to strengthen the semidefinite relaxations $\widehat{\mathcal{C}}$ defined earlier, as follows. For example, the set of 2-body locally consistent marginals $\widehat{\mathcal{C}}_2^{\text{loc}}$ can be strengthened by adding the following scalar inequalities:

$$S(i|j)_\rho + S(i|k)_\rho \geq 0 \quad \forall 1 \leq i, j, k \leq n \text{ distinct.} \quad (10)$$

In general, if one considers relaxations involving marginals on $l \geq 2$ sites, one can include all weak monotonicity inequalities (9) for all disjoint sets $A, B, C \subset V$ such that $|AB| \leq l$ and $|AC| \leq l$. The next lemma shows that it is indeed sufficient to consider only inequalities where $|A| = 1$, and $|B| = |C| = l - 1$, i.e.,

$$\begin{aligned} \widehat{\mathcal{C}}_l^{\text{WM}} = & \left\{ (\rho_{ij})_{ij \in E} : \exists (\rho_S)_{|S| \leq l}, \quad \rho_S \geq 0, \text{Tr } \rho_S = 1 \quad \forall S \in \binom{V}{l} \right. \\ & \text{and } \text{Tr}_{S \setminus S'} \rho_S = \text{Tr}_{S' \setminus S} \rho_{S'} \quad \forall S, S' \in \binom{V}{l} \\ & \left. \text{and } S(A|B) + S(A|C) \geq 0 \quad \forall |A| = 1, B, C \in \binom{V \setminus A}{l-1} \text{ disjoint} \right\}. \end{aligned} \quad (11)$$

Lemma 2.2. *Let $|V| \geq 2l - 1$ and $(\rho_S)_{S \in \binom{V}{l}}$ be a set of locally consistent marginals satisfying weak monotonicity $S(A|B)_\rho + S(A|C)_\rho \geq 0$ for any disjoint A, B and C with $|A| = 1$ and $|B| = |C| = l - 1$. Then the same weak monotonicity inequality holds for any disjoint sets A, B and C of any size (as long as it is defined, i.e., $|AB|, |AC| \leq l$).*

Proof. For $|A| = 1$ and B, C of any size at most $l - 1$, the proof directly follows from the data-processing inequality for conditional entropy. If $|A| = m > 1$, let the elements of A be $A_1 \dots A_m$ with $A_i \in V$ and use the chain rule to expand $S(A|B)$ and $S(A|C)$ in a telescoping sum that cancels all but the first and last terms

$$\begin{aligned} S(A|B)_\rho + S(A|C)_\rho &= \sum_{i=1}^m S(A_i|A_1 \dots A_{i-1}B)_\rho + \sum_{i=m}^1 S(A_i|A_{i+1} \dots A_mC)_\rho \\ &= \sum_{i=1}^m S(A_i|A_1 \dots A_{i-1}B)_\rho + S(A_i|A_{i+1} \dots A_mC)_\rho. \end{aligned}$$

Using the fact that $|AB| \leq l$, we have that $|A_1 \dots A_{i-1}B| \leq l - 1$ for any $i \in \{1, \dots, m\}$ and similarly for C , together with the fact that $A_1 \dots A_{i-1}B$ and $A_{i+1} \dots A_mC$ are disjoint we obtain that $S(A|B)_\rho + S(A|C)_\rho \geq 0$. \square

We remark that the Lemma still holds if the marginal relaxation is not defined for *all* l -site subsets, as long as it contains the marginals and entropy inequalities that appear in the proof.

Notice that if ρ_{AB} and ρ_{AC} are classical probability distributions, or more generally if ρ_{AB} and ρ_{AC} are separable quantum states, then conditional entropies are nonnegative and thereby Eq. (9) is automatically satisfied, even without assuming the existence of a global state ρ_{ABC} . As a result, imposing the weak monotonicity inequality is useless for classical Hamiltonians. Furthermore, we note that the weak monotonicity inequalities at level l are automatically implied by the level $2l - 1$ local consistency relaxation, i.e.,

$$\widehat{\mathcal{C}}_{2l-1}^{\text{loc}} \subset \widehat{\mathcal{C}}_l^{\text{WM}} \subset \widehat{\mathcal{C}}_l^{\text{loc}}. \quad (12)$$

This is simply because the level $2l - 1$ ensures that the marginals are consistent with a valid quantum state on ABC , for which weak monotonicity is known to hold.

While this means weak monotonicity constraints cannot help more than doubling the number of sites in the marginal relaxation, it should be noted that this corresponds to squaring the size of each variable in the semidefinite program.

Graphical illustration. We end this section by graphically illustrating the effect of the entropy constraints. Consider the problem of characterizing the $\{1, 2\}$ and $\{2, 3\}$ -body marginals of a 3-site qubit state $\rho = \rho_{123}$, i.e.,

$$\{(\rho_{12}, \rho_{23}) : \rho_{ij} = \text{Tr}_{\setminus \{i,j\}} \rho_{123} \quad \text{for some } \rho_{123} \in \mathcal{D}((\mathbb{C}^2)^{\otimes 3})\}.$$

Figure 1 shows a particular two-dimensional projection of this set on Bell states, namely

$$(\rho_{12}, \rho_{23}) \mapsto (\text{Tr}[(\rho_{12} + \rho_{23})/2|\psi^-\rangle\langle\psi^-|], \text{Tr}[(\rho_{12} + \rho_{23})/2|\phi^+\rangle\langle\phi^+|]).$$

A. Quantitative advantage

In this section we analyze quantitatively the advantage that entropy constraints can provide for estimating the ground energy of certain local Hamiltonians.

A Hamiltonian on three sites. We start by looking at a Hamiltonian defined on a 3-node graph given by

$$H = h_{12} + h_{23},$$

where

$$h_{12} = -\frac{1}{2}|\psi^-\rangle\langle\psi^-|_{12} \quad h_{23} = -\frac{1}{2}|\psi^-\rangle\langle\psi^-|_{23}$$

and $|\psi^-\rangle = (|01\rangle - |10\rangle)/\sqrt{2}$. The ground energy of this Hamiltonian is $\lambda_{\min}(H) = -3/4$. Using the relaxation $\widehat{\mathcal{C}}_2^{\text{loc}}$, we get the (trivial) lower bound -1 : indeed a valid locally consistent assignment of 2-body marginals is

$$\rho_{12} = \rho_{23} = |\psi^-\rangle\langle\psi^-| \tag{13}$$

as all 1-body marginals are consistently equal to the maximally mixed state $\mathbb{1}/2$. However, this violates the monogamy of entanglement property as for a valid global state, 2 cannot be maximally entangled with 1 and 3. This violation of monogamy is to some extent captured by entropy constraints coming from weak monotonicity. Indeed, the assignment (13) violates the inequality $S(2|1)_\rho + S(2|3)_\rho \geq 0$ since the left-hand side in this case is equal to -2 . In fact, by optimizing the Hamiltonian H over the relaxation $\widehat{\mathcal{C}}_2^{\text{WM}}$ we get a value that is $\sim -0.811 > -1$, i.e.,

$$\min_{(\rho_{ij}) \in \widehat{\mathcal{C}}_2^{\text{WM}}} \text{Tr}[h_{12}\rho_{12}] + \text{Tr}[h_{23}\rho_{23}] \geq -0.811. \tag{14}$$

Graphically this value can be seen in Fig. 1 as the x -component of the rightmost point of the red convex set, whereas the true value -0.75 is for the blue convex set.

Remark 2.3 (Exploiting symmetries). In the above example, one can actually use the symmetries of the Hamiltonian to simplify the entropy constraints. As the Hamiltonian commutes with the unitary operation of exchanging subsystems 1 and 3, the minimizer can be chosen to obey the same symmetry. For that state we have $S(2|1) = S(2|3)$, and so the weak monotonicity inequality becomes simply $S(2|1) \geq 0$. We will see more examples of similar arguments when we consider translation-invariant Hamiltonians in Sec. IV.

Larger graphs. We now show that the advantage of entropy constraints extends to Hamiltonians on arbitrary graphs. Consider a Hamiltonian defined on a general connected graph $G = (V, E)$ of the form

$$H = \sum_{ij \in E} h_{ij} \quad h_{ij} = -|\psi^-\rangle\langle\psi^-|_{ij} \quad \forall ij \in E. \tag{15}$$

While the simple relaxation $\widehat{\mathcal{C}}_2^{\text{loc}}$ yields a trivial lower bound of $-|E|$ on $\lambda_{\min}(H)$, one can show that the relaxation incorporating the weak monotonicity constraints $S(ij) + S(ik) \geq 0$ will have a value $\gtrsim -0.811|E|$. More precisely, we prove:

Theorem 2.4. For the Hamiltonian (15) defined on a connected graph $G = (V, E)$, we have

$$-|E| = \min_{(\rho_e) \in \widehat{\mathcal{C}}_E^{\text{loc}}} \sum_{e \in E} \text{Tr}[h_e \rho_e] \leq -0.811|E| - 1 \leq \min_{(\rho_e) \in \widehat{\mathcal{C}}_E^{\text{WM}}} \sum_{e \in E} \text{Tr}[h_e \rho_e].$$

The key to proving this theorem is the following proposition, which allows us to decompose the edges E of the graph into disjoint pairs of adjacent edges.

Proposition 2.5. For any connected graph $G = (V, E)$, there exists a set $P = \{\{f_1, g_1\}, \dots, \{f_m, g_m\}\}$ of disjoint pairs of adjacent edges (i.e., $f_i \in E$ and $g_i \in E$ share a node for each i , and $\{f_i, g_i\} \cap \{f_j, g_j\} = \emptyset$ for all $i \neq j$) that covers all edges of the graph for an even number of edges and all but one in the case of an odd number of edges.

Proof. See Appendix A. □

Proof of Theorem 2.4. We use the decomposition of the edge set P from Proposition 2.5 and write

$$H = \sum_{\{f,g\} \in P} (h_f + h_g) + \delta_{\text{odd}} h_e$$

where $\delta_{odd} = 1$ if the graph has an odd number of edges with e the unmatched edge and $\delta_{odd} = 0$ otherwise. We estimate

$$\begin{aligned} \min_{(\rho_e) \in \mathcal{C}_E^{\text{WMM}}} \sum_{e \in E} \text{Tr}[h_e \rho_e] &\geq \left(\sum_{\{f, g\} \in P} \min_{(\rho_e) \in \mathcal{C}_E^{\text{WMM}}} \text{Tr}[h_f \rho_f] + \text{Tr}[h_g \rho_g] \right) + \delta_{odd} \\ &\geq -0.811(|E| - \delta_{odd}) - \delta_{odd} \end{aligned}$$

where we bounded for each term the objective by the optimal value for the entropy constrained value of the 3-node graph from Eq. (14). \square

We remark that for this problem, we chose a smaller set of marginals and entropy constraints motivated by the structure of the problem rather than simply all possible constraints for the given size of marginals.

These examples are unweighted instances of the quantum version of the Max-Cut problem that has been widely studied in the literature.^{22–27} These works consider semidefinite relaxations of this problem and then round the solutions of the semidefinite program to a valid global quantum state. In order to do this, for example in Ref. 23, it is shown that any (ρ_{12}, ρ_{23}) coming from the relaxation $\widehat{\mathcal{C}}_2^{\text{sos}}$ satisfies the linear inequality

$$\frac{1}{2} (\text{Tr}[\rho_{12} |\psi^-\rangle\langle\psi^-|_{12}] + \text{Tr}[\rho_{23} |\psi^-\rangle\langle\psi^-|_{23}]) \leq 0.75. \quad (16)$$

This can be interpreted in terms of Fig. 1: it shows that Eq. (16) leads to a tight bound when optimizing the linear function defined by $|\psi^-\rangle\langle\psi^-|$ on sites 12 and 23. The entropy constraint performs worse in the directions of the Bell states as seen in Fig. 1 and in Eq. (14), but it is better in some other directions.

Larger marginals. We extend the above example to show how nontrivial entropy constraints can be easily constructed for any choice $k \geq 2$ of number of sites considered by the marginal. Let us consider a k -local Hamiltonian, where each term is given by a projector on the GHZ-state $h_e = -|\text{GHZ}\rangle\langle\text{GHZ}|$, which is given by $|\text{GHZ}\rangle = (|0 \dots 0\rangle + |1 \dots 1\rangle)/\sqrt{2}$ and e is a subset of the sites of size k . Again, for the k -site relaxation the optimal solution is to assign a projector onto the GHZ-state to every marginal $\rho_e = |\text{GHZ}\rangle\langle\text{GHZ}|$, which are all locally consistent. However, the conditional entropy for the GHZ state $S(1 \dots i | i+1 \dots k) = -1$ is negative. As a result, for two distinct subsets e and f of size k that have a nonempty intersection, if we write $e = AB$ and $f = AC$, the weak monotonicity constraint $S(A|B) + S(A|C) \geq 0$ is violated. For concreteness, let $V = 1 \dots (2l-1)$ and consider hyperedges $e = 1 \dots l, f = l+1 \dots (2l-1)$ and consider the Hamiltonian $h_e + h_f$. For the relaxation we specify the marginals ρ_e, ρ_f subject to local consistency on site l and the weak monotonicity constraint

$$S(l | 1 \dots l-1) + S(l | l+1 \dots 2l-1) \geq 0$$

In fact, due to the symmetry of the problem, this can be reduced to the optimization

$$\min_{\rho_e, S(l | 1 \dots l-1) \geq 0} \text{Tr}[h_e \rho_e]. \quad (17)$$

For any l , this problem is related to the problem at $l = 2$ by the local isometry $U_{1 \rightarrow 1, \dots, l-1} = |0 \dots 0\rangle\langle 0| + |1 \dots 1\rangle\langle 1|$. Since this local isometry leaves the conditional entropy invariant the optimal value of Eq. (17) does not depend on l and is thereby equal to ~ -0.811 as in the previously encountered example Eq. (14).

Lasserre/SoS-hierarchy. We omit a detailed study of the enhancement of the Lasserre hierarchy mentioned in the introduction via entropy constraints, but comment on some previous results and future directions. The Lasserre hierarchy at level l provides expectation values on up to $2l$ sites, but only guarantees that they are consistent with a density matrix for up to l sites. Thereby, it is not possible to define any entropy constraints (involving at least two sites) at the first level. To circumvent the lack of positivity, in Ref. 23 a level 1.5 of the Lasserre hierarchy is introduced that strengthens the level 1 by an explicit positivity constraint on the 2-site marginals. In fact, it is used for the analysis of star graphs including the one in this section, for which it yields an optimal value of 0.862. Since this is a worse approximation compared to the entropy constrained value we obtain this shows that there is room for improvement using the entropy constraint, which can be defined for the Lasserre 1.5. It is a future direction to find explicit improvements also at higher levels of the Lasserre hierarchy. They can be imposed whenever the constraints of the hierarchy enforce the positivity of the marginals involved in the entropy constraints.

III. MARKOV ENTROPY DECOMPOSITION

In this section we derive another family of entropy constraints that one can impose on the set of globally consistent marginals. We call these inequalities Markov Entropy Decomposition (MED) inequalities, because they have appeared in Ref. 18 in the context of deriving lower bounds on the free energy of a Hamiltonian H . The next lemma briefly describes the family of inequalities. We explain the connection to the free energy later.

Lemma 3.1 (Markov entropy decomposition inequalities). Let ρ be a density operator acting on the Hilbert space corresponding to the sites V . Consider an ordering $1, \dots, N = |V|$ of the sites in V , and for each $i \in \{1, \dots, N\}$, let $\mathcal{N}_i \subset \{1, \dots, i-1\}$ be a subset of the sites appearing

before i (in the chosen order) with $|\mathcal{N}_i| \leq l - 1$. Then for any $1 \leq k \leq N$

$$\sum_{i=1}^k S(i|\mathcal{N}_i)_\rho \geq 0. \tag{18}$$

Proof. We have, using the chain rule and strong subadditivity:

$$0 \leq S(1 \dots k)_\rho = S(1)_\rho + S(2|1)_\rho + S(3|12)_\rho + \dots + S(k|1 \dots k-1)_\rho \leq \sum_{i=1}^k S(i|\mathcal{N}_i)_\rho.$$

□

We see that the inequalities (18) only involve the marginals of ρ on at most l sites (since $|\mathcal{N}_i| \leq l - 1$) and that they define a convex region in the space of these marginals. Thus these inequalities can be used to strengthen the relaxations $\widehat{C}_l^{\text{loc}}$. We note that there is an MED inequality for each choice of ordering of the sites in V , and each choice of *Markov shields* $\mathcal{N}_i \subset \{1, \dots, i - 1\}$.

A. Relationship with weak monotonicity

A natural question is whether the MED inequalities are equivalent to the weak monotonicity inequalities derived in the previous section. The answer in general is no, and we show in this section that the two relaxations are incomparable.

However, we first observe that some simple MED inequalities are implied by the weak monotonicity inequalities. For example consider the case where the Markov shields \mathcal{N}_i consist of the (at most) $l - 1$ sites that are immediately preceding the site i in the chosen order, i.e., $\mathcal{N}_i = \mathcal{N}_i^- = \{\max(1, i - l + 1), \dots, i - 1\}$. Then one can show that the MED inequality is implied by the weak monotonicity inequalities. Indeed, let \mathcal{N}_i^+ be the (at most) $l - 1$ sites that are immediately following the site i in the chosen order, i.e., $\mathcal{N}_i^+ = \{i + 1, \dots, \min(k, i + l - 1)\}$. The MED inequality states

$$\sum_{i=1}^k S(i|\mathcal{N}_i^-) \geq 0. \tag{19}$$

One can check that $\sum_{i=1}^k S(i|\mathcal{N}_i^-) = \sum_{i=1}^k S(i|\mathcal{N}_i^+)$, and so the inequality above can be equivalently written as

$$\sum_{i=1}^k S(i|\mathcal{N}_i^-) + S(i|\mathcal{N}_i^+) \geq 0.$$

Now it suffices to observe that for each i , $S(i|\mathcal{N}_i^-) + S(i|\mathcal{N}_i^+) \geq 0$ is implied by weak monotonicity.

We now give two examples showing that MED and weak monotonicity constraints are in general inequivalent.

Weak monotonicity does not imply MED. We construct a set of 2-body locally consistent states on N sites (for N large enough) that violate a certain MED inequality but are consistent with all possible weak monotonicity constraints of the form (10). The 2-body marginals are defined as follows (for $i, j \geq 2$):

$$\begin{aligned} \rho_{1i} &= \lambda |\psi^-\rangle \langle \psi^-| + (1 - \lambda) (\mathbb{1}/2 \otimes |0\rangle \langle 0|) \\ \rho_{ij} &= (\lambda \mathbb{1}/2 + (1 - \lambda) |0\rangle \langle 0|)^{\otimes 2}. \end{aligned}$$

Here, $|\psi^-\rangle = (|01\rangle - |10\rangle)/\sqrt{2}$ as before. The entropies of the single-body and two-body marginals are

$$S(1) = 1 \quad \text{and} \quad S(i) = h\left(1 - \frac{\lambda}{2}\right) \quad \text{and} \quad S(ij) = 2h\left(1 - \frac{\lambda}{2}\right) \quad i, j \geq 2$$

where h is the binary entropy function, which is concave, continuous and satisfies $h(1) = 0$ and $h(1/2) = 1$. We omit an explicit formula for the entropy of the two-body marginal $S(1i)$, but it is also concave, continuous, and equal to 1 if $\lambda = 0$, and 0 if $\lambda = 1$.

This means we can pick $0 < \lambda < 1$ such that $S(1i) + S(1j) - 2S(1i) - 2h(1 - \lambda/2) = 0$. Furthermore,

$$\begin{aligned} S(ij) + S(i|1) &= h(1 - \lambda/2) + S(i|1) - 1 \\ &\geq \lambda + (1 - \lambda) - 1 = 0, \end{aligned}$$

where we applied concavity of h and the entropy function. Finally, $S(ij) + S(i|k) \geq 0$ for $i, j, k \geq 2$, because both marginals are product states, so all weak monotonicity constraints that can be defined for 2-site marginals are satisfied.

However, let us consider the MED inequality

$$0 \leq S(1) + \sum_{i=1}^N S(i|1) = 1 + N(S(1i) - 1).$$

Since $0 < S(1i) = h(1 - \lambda/2) < 1$ (h is strictly decreasing on $[1/2, 1]$), we see that the right-hand side becomes negative for sufficiently large N ($N = 8$ is sufficient).

MED does not imply weak monotonicity. We now construct a set of 2-body locally consistent states on $N = 3$ sites that satisfy all the MED inequalities but violate some weak monotonicity inequalities. We choose for all ij

$$\rho_{ij} = \lambda|\psi^-\rangle\langle\psi^-| + (1 - \lambda)\mathbb{1}/4 \tag{20}$$

with λ such that $-1/2 \leq S(ij) < 0$. This is possible due to the continuity in λ and the fact that it takes the values 1 and -1 for $\lambda = 0$ and $\lambda = 1$ respectively. This choice violates weak monotonicity as all conditional entropies are negative. Due to the permutation invariance, the only relevant MED constraints to be checked are

$$\begin{aligned} S(1) + S(2|1) + S(3|2) &\geq 0 \\ S(1) + S(2|1) + S(3|1) &\geq 0. \end{aligned}$$

These are satisfied since $S(1) = 1$ and each $S(i|j) \geq -1/2$.

Remark 3.2. We note that on three sites, the 2-body MED constraints are all implied by weak monotonicity. Indeed, the MED constraints are all of the form

$$S(i) + S(j|i) + S(k|j) \geq 0$$

for all choices of i, j, k distinct. The only other choice $S(i) + S(j|i) + S(k|i) = S(j) + S(i|j) + S(k|i)$ is in fact equivalent. This is precisely of the form (19) where the sites are ordered as $i < j < k$, which have been shown to be implied by the weak monotonicity constraints. The above example (20) thus shows that the weak monotonicity constraints are strictly stronger than the MED constraints on three sites.

B. Connection with the free energy

The expression appearing in the MED inequalities was used in Ref. 18 to derive lower bounds on the free energy $F(T)$ of the Hamiltonian H on sites V at temperature T . Recall that the free energy is given by the following variational formula

$$F(T) = \min_{\rho} \{ \text{Tr} [H\rho] - TS(V)_{\rho} \},$$

where the minimization is over all density operators ρ acting on the Hilbert space of all the sites in V . As opposed to the ground energy problem, the objective does not only depend on few-site marginals as the entropy function involves the global density matrix and its spectrum. The entropy however can be decomposed into conditional entropies which can be upper bounded using the strong subadditivity inequality: Considering an order of the sites $1, \dots, N$, we write

$$\begin{aligned} S(V)_{\rho} &= S(1)_{\rho} + S(2|1)_{\rho} + \dots + S(N|1 \dots N-1) \\ &\leq \sum_{i=1}^N S(i|\mathcal{N}_i)_{\rho} \end{aligned}$$

again using Markov shields $\mathcal{N}_i \subset \{1, \dots, i-1\}$ defining the sites that are taken into account in the conditioning. The intuition in choosing these is that sites that are far away in the hypergraph distance induced by the interaction hypergraph have only small correlation as measured by the conditional mutual information (CMI), see for example Ref. 28 for a result in 1D. The CMI is also equal to the error made in the approximation using strong subadditivity. This relaxation can now again be computed by just the marginals on $\{i\} \cup \mathcal{N}_i$ and the supports of the Hamiltonian terms and therefore allows for a further relaxation as explained in the previous section:

$$\begin{aligned} F(T) &= \min_{\rho} \{ \text{Tr} [H\rho] - TS(\rho) \} \\ &\geq \min_{(\rho_e) \in \mathcal{C}_l^{\text{loc}}} \left\{ \sum_e \text{Tr} [h_e \rho_e] - T \sum_{i=1}^N S(i|\mathcal{N}_i) \right\} := \text{MED}(T) \end{aligned} \tag{21}$$

Here, l is chosen sufficiently large to define all conditional entropies and interaction terms.

Lower bounds for the ground energy. It has been pointed out in Ref. 18 that while the free energy $F(T)$ is a decreasing function of T , its MED approximation $\text{MED}(T)$ need not be. At every temperature, however, we have $\text{MED}(T) \leq F(T) \leq F(0)$, which is to say that the MED lower bounds the ground energy $F(0)$. We can maximize this lower bound over T to obtain the best lower bound on the ground energy at that level. It is easy to see that maximizing the expression for $\text{MED}(T)$ in (21) corresponds to adding the MED term as an inequality, i.e.,

$$F(0) \geq \max_T \text{MED}(T) = \min_{\substack{(\rho_e) \in \mathcal{C}_l^{\text{loc}} \\ \sum_{i=1}^N S(i|\mathcal{N}_i) \geq 0}} \sum_e \text{Tr} [h_e \rho_e].$$

IV. INFINITE SYSTEMS

A practically relevant case of the above methods are lattice systems. We start this section by discussing the application of weak monotonicity to one-dimensional translation-invariant systems.

One-dimensional systems. We consider a translation-invariant Hamiltonian H acting on a chain of length m of the form

$$H_{[1,m]} = \sum_{i=1}^m h_{i,i+1}$$

where the interaction terms $h_{i,i+1}$ act on sites i and $i + 1$, and are translates of the same Hermitian operator h acting on two sites $\mathbb{C}^d \otimes \mathbb{C}^d$. We use periodic boundary conditions, i.e., we identify $m + 1$ with 1. We can define the ground energy per site of the infinite system as the limit

$$e_0(h) = \lim_{m \rightarrow \infty} \frac{1}{m} \lambda_{\min}(H_{[1,m]}) = \lim_{m \rightarrow \infty} \frac{1}{m} \min_{\rho_m \in \mathcal{D}(\otimes_{v \in [1,m]} \mathbb{C}^d)} \text{Tr}[H_{[1,m]}\rho_m]. \quad (22)$$

The existence of the limit above is shown in [Appendix B](#); in fact we show that in the minimization problems in (22), it is equivalent to restrict to just a single interaction term while taking ρ_m translation-invariant, i.e., satisfying $\text{Tr}_1[\rho_m] = \text{Tr}_m[\rho_m]$, since for such translation-invariant states the two-body marginals $\rho_{12}, \dots, \rho_{m-1,m}$ are all equal. This allows us to express the ground energy density $e_0(h)$ in the following way

$$e_0(h) = \min_{\rho \in \mathcal{C}^{\text{TI}}} \text{Tr}[h\rho],$$

where \mathcal{C}^{TI} is the set of two-body density matrices that are extendible to an infinite translation-invariant system, i.e.,

$$\mathcal{C}^{\text{TI}} = \{\rho \in \mathcal{D}(\mathbb{C}^d \otimes \mathbb{C}^d) : \forall m \exists \rho_m \in \mathcal{D}(\otimes_{v \in [1,m]} \mathbb{C}^d), \rho = \text{Tr}_{[3,m]}[\rho_m], \text{Tr}_1[\rho_m] = \text{Tr}_m[\rho_m]\}. \quad (23)$$

A “local” relaxation of this set can be defined in a similar way as discussed previously, and leads to

$$\widehat{\mathcal{C}}_l^{\text{loc, TI}} = \{\rho \in \mathcal{D}(\mathbb{C}^d \otimes \mathbb{C}^d) : \exists \rho_l \in \mathcal{D}(\otimes_{v \in [1,l]} \mathbb{C}^d), \rho = \text{Tr}_{[3,l]}[\rho_l], \text{Tr}_1[\rho_l] = \text{Tr}_l[\rho_l]\}. \quad (24)$$

We prove in [Appendix B](#) that for any h , the value

$$\min_{\rho \in \widehat{\mathcal{C}}_l^{\text{loc, TI}}} \text{Tr}[h\rho]$$

converges to $e_0(h)$ as $l \rightarrow \infty$ at the rate $O(1/l)$.

Using weak monotonicity and translation-invariance, one can strengthen the relaxation (24) by adding the inequality $S(l|1 \dots l-1) \geq 0$, leading to:

$$\begin{aligned} \widehat{\mathcal{C}}_l^{\text{WM, TI}} = \{ & \rho \in \mathcal{D}(\mathbb{C}^d \otimes \mathbb{C}^d) : \exists \rho_l \in \mathcal{D}(\otimes_{v \in [1,l]} \mathbb{C}^d), \\ & \rho = \text{Tr}_{[3,l]}[\rho_l], \text{Tr}_1[\rho_l] = \text{Tr}_l[\rho_l], S(l|1 \dots l-1) \geq 0\}. \end{aligned} \quad (25)$$

The next theorem shows that the above is a valid relaxation, and moreover that the resulting relaxation cannot be better than $\widehat{\mathcal{C}}_{2l-1}^{\text{loc, TI}}$.

Theorem 4.1.

$$\widehat{\mathcal{C}}_{2l-1}^{\text{loc, TI}} \subset \widehat{\mathcal{C}}_l^{\text{WM, TI}} \subset \widehat{\mathcal{C}}_l^{\text{loc, TI}}$$

Proof. The second inclusion is immediate. To prove the first one we start from a translation-invariant state on sites $1, \dots, 2l - 1$. We know from weak monotonicity that $S(l|1 \dots l-1) + S(l|l+1 \dots 2l-1) \geq 0$. However, translation-invariance tells us that

$$\begin{aligned} S(l|l+1 \dots 2l-1) &= S(l \dots 2l-1) - S(l+1 \dots 2l-1) \\ &= S(1 \dots l) - S(1 \dots l-1) \\ &= S(l|1 \dots l-1), \end{aligned}$$

so $2S(l|1 \dots l-1) \geq 0$, which concludes the proof. □

In fact, the entropy constraint $S(l|1 \dots l-1) \geq 0$ can as well be derived from the MED inequalities (Lemma 3.1) by using an increasing order and a Markov shield equal to the previous $l - 1$ sites.

Higher dimensions. We now consider a lattice system on \mathbb{Z}^D for $D \geq 1$. Let h be a Hamiltonian term on the origin and its nearest neighbours in positive axis directions (i.e., a Hermitian matrix of size $d^{D+1} \times d^{D+1}$), and for $v \in \mathbb{Z}^D$, let h_v be its translate to the site $v \in \mathbb{Z}^D$, so that it acts on $\{v, v + e_1, \dots, v + e_D\}$. Let us consider Hamiltonians of the form:

$$H_{[1,m]^D} = \sum_{v \in [1,m]^D} h_v,$$

with periodic boundary conditions, i.e., identifying $m + 1$ with 1. The ground energy density of the system is defined by the limit (see Appendix B for existence).

$$e_0(h) = \lim_{m \rightarrow \infty} \frac{1}{m^D} \lambda_{\min}(H_{[1,m]^D}).$$

Using the same considerations as in the 1D case, we can also write $e_0(h)$ as a linear optimization problem over the set of density matrices defined on $\{0, e_1, \dots, e_D\}$ that are extendible to an infinite translation-invariant system, i.e.,

$$\mathcal{C}^{\text{TI}} = \bigcap_{m \in [2, \infty)} \tilde{\mathcal{C}}_{[1,m]^D}^{\text{loc, TI}}$$

where for any finite subset $A \subset \mathbb{Z}^D$ such that $\{0, e_1, \dots, e_D\} \subset A$ we define

$$\begin{aligned} \tilde{\mathcal{C}}_A^{\text{loc, TI}} = \{ \rho \in \mathcal{D}((\mathbb{C}^d)^{D+1}) : \exists \rho_A \in \mathcal{D}(\otimes_{v \in A} \mathbb{C}^d) \text{ s.t. } \rho = \text{Tr}_{A \setminus \{0, e_1, \dots, e_D\}}[\rho_A] \\ \text{Tr}_{A \setminus A+t}[\rho_A] = \text{Tr}_{A+t \setminus A}[\tau_t(\rho_A)] \quad \forall t \in \mathbb{Z}^D \}, \end{aligned}$$

where τ_t defines the translation operator shifting the state by t sites.

We investigate what entropy constraints arising from weak monotonicity and the MED can be imposed to strengthen the relaxation $\tilde{\mathcal{C}}_A^{\text{loc, TI}}$ of \mathcal{C}^{TI} . The weak monotonicity inequality allows us to strengthen the relaxation above by adding constraints of the form

$$S(i|B) + S(i|C) \geq 0 \tag{26}$$

for any choice of $i \in A$ and $B, C \subset A$ disjoint.

We now consider MED inequalities. Consider a translation-invariant state defined on $[-m, m]^D$ for large m . We fix an order of the sites $1, \dots, N = (2m + 1)^D$ that decreases in each coordinate direction and iterates over the coordinates (see Fig. 2). Let \mathcal{N} be some fixed region of size $l - 1$ independent of m , which is a subset of the sites preceding the origin [a common choice would be the intersection of a ball $B(0, r)$ with the sites preceding the origin, see Fig. 2]. We define the Markov shields \mathcal{N}_i relative to the site i by translating \mathcal{N} to the site numbered i , i.e., abusing notations:

$$\mathcal{N}_i = (i + \mathcal{N}) \cap [-m, m]^D.$$

The MED inequality (18) in this case takes the form

$$0 \leq \sum_{i=1}^N S(i|\mathcal{N}_i).$$

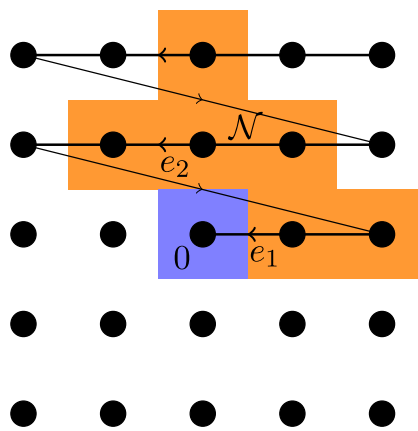


FIG. 2. Markov shield with order of sites in 2D.

For sites i that are not close to the boundary of the region $[-m, m]^D$, the regions $\{i\} \cup \mathcal{N}_i$ are all translates of each other, and so the terms in the equation above are equal, by translation invariance. The number of sites i that are close to the boundary is $o(N)$, and so dividing by N and letting $N \rightarrow \infty$, the inequality above yields

$$0 \leq S(0|\mathcal{N}) \tag{27}$$

for a state in \mathcal{C}^{TI} . This suggests the relaxation

$$\begin{aligned} \mathcal{C}^{\text{TI}} \subset \widehat{\mathcal{C}}_{0 \cup \mathcal{N}}^{\text{MED, TI}} = \{ \tilde{\rho} \in \mathcal{D}((\mathbb{C}^d)^{D+1}) : \exists \rho \in \mathcal{D}(\otimes_{v \in 0 \cup \mathcal{N}} \mathbb{C}^d) \text{ s.t.} \\ \text{Tr}_{(0 \cup \mathcal{N}) \setminus (0 \cup \mathcal{N}) + t} [\rho] = \text{Tr}_{(0 \cup \mathcal{N}) + t \setminus (0 \cup \mathcal{N})} [\tau_t(\rho)] \quad \forall t \in \mathbb{Z}^D \\ S(0|\mathcal{N}) \geq 0 \}. \end{aligned}$$

In general, the MED inequality (27) is different from the weak monotonicity inequality (26): (27) shows the nonnegativity of a particular conditional entropy $S(0|\mathcal{N})$ where \mathcal{N} has to satisfy the conditions of a Markov shield described above (i.e., compatibility with well-chosen order). Weak monotonicity on the other hand, asserts the nonnegativity of the sum of two conditional entropies (26) with the only constraint that B and C are disjoint.

Under some mild conditions however, one can recover the MED inequality (27) as a consequence of weak monotonicity. This is the case if the state ρ is reflection symmetric, which can be assumed if the Hamiltonian itself is reflection symmetric (i.e., $h = \sum_{i=1}^D h_{0, e_i}^i$ where h_{0, e_i}^i acts on sites $0, e_i$ and is symmetric under exchange of the two sites). One can check that with the order defined earlier, the reflection of \mathcal{N} about the origin is disjoint from \mathcal{N} , i.e., $-\mathcal{N} \cap \mathcal{N} = \emptyset$. Furthermore, by reflection symmetry we have $S(0|\mathcal{N}) = S(0|-\mathcal{N})$ and so the weak monotonicity inequality $S(0|\mathcal{N}) + S(0|-\mathcal{N}) \geq 0$ recovers (27). A corollary of this observation is stated in the next theorem.

Theorem 4.2. *Under the reflection symmetry and translation-invariance assumptions explained above with a Markov shield \mathcal{N} of size $l-1$ the MED relaxation is no better than a semidefinite optimization in one $d^{2l-1} \times d^{2l-1}$ dimensional variable, more precisely*

$$\min_{\rho \in \widehat{\mathcal{C}}_{0 \cup \mathcal{N}}^{\text{loc, TI}}} \text{Tr}[h\rho] \leq \min_{\rho \in \widehat{\mathcal{C}}_{0 \cup \mathcal{N}}^{\text{MED, TI}}} \text{Tr}[h\rho] \leq \min_{\rho \in \widehat{\mathcal{C}}_{0 \cup \mathcal{N} \cup -\mathcal{N}}^{\text{loc, TI}}} \text{Tr}[h\rho] \leq e_0(h).$$

Proof. The first inequality is immediate. We start with the value

$$\text{SDP} = \min_{\rho \in \widehat{\mathcal{C}}_{0 \cup \mathcal{N} \cup -\mathcal{N}}^{\text{loc, TI}}} \text{Tr}[h\rho].$$

Due to the translation-invariance this is the same as

$$\min_{\rho \in \widehat{\mathcal{C}}_{0 \cup \mathcal{N} \cup -\mathcal{N}}^{\text{loc, TI}}} \text{Tr}[\tilde{h}\rho].$$

with $\tilde{h} = \sum_{i=1}^D (h_{0, e_i}^i + h_{-e_i, 0}^i)/2$ and the variable over the same system. This problem is intrinsically reflection symmetric, which means that the optimizer ρ^* can be chosen reflection symmetric. Thereby we also have $S(0|\mathcal{N})_{\rho^*} = S(0|-\mathcal{N})_{\rho^*}$ and by using weak monotonicity we deduce

$$S(0|\mathcal{N})_{\rho^*} = \frac{S(0|\mathcal{N})_{\rho^*} + S(0|-\mathcal{N})_{\rho^*}}{2} \geq 0,$$

which holds because \mathcal{N} and $-\mathcal{N}$ are disjoint. This shows that any optimizer of the SDP on $2l-1$ sites gives a feasible point with same objective value and satisfying the MED inequality. This concludes the proof. \square

V. NUMERICAL EXPERIMENTS

In this section we present numerical experiments for the relaxations with entropy constraints. The model we consider is the XXZ-Hamiltonian with the nearest-neighbour interaction

$$h = -\sigma_x \otimes \sigma_x - \sigma_y \otimes \sigma_y - \Delta \sigma_z \otimes \sigma_z \tag{28}$$

on an infinite chain. The model is gapless for $|\Delta| \leq 1$ and gapped otherwise.²⁹ For $\Delta = 0$ this model is also called the XY-Hamiltonian. We consider the values obtained by the relaxation $\widehat{\mathcal{C}}_l^{\text{loc, TI}}$ and its strengthening $\widehat{\mathcal{C}}_l^{\text{WM, TI}}$ based on entropy constraint $S(l|1 \dots l-1) \geq 0$, see Eqs. (24) and (25).

The resulting convex optimization problems can be solved using different algorithms. One approach is to rely on semidefinite programming solvers, via approximations of the (conditional) entropy function.³⁰ Another approach is to use interior-point methods that support the quantum relative entropy cone.^{17,31,32} Finally, another approach is to use custom first-order splitting methods such as Ref. 33, see also Ref. 34. We have used the latter for the experiments in this section.

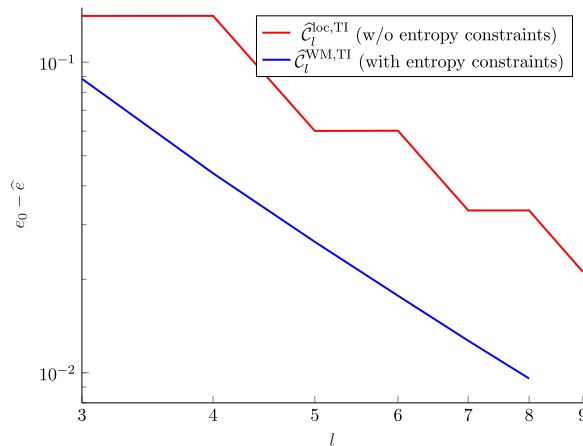


FIG. 3. Accuracy of the relaxations with and without entropy constraints for the XY-Hamiltonian on an infinite 1D chain. We verify that the relaxation with entropy constraint gives a value which is not better than the standard SDP relaxation at level $2l - 1$. It is surprising to note however that it is better than the value of the SDP for all levels up to $2l - 2$.

We start by comparing the convergence with l for the XY-Hamiltonian, see Fig. 3. We go up to $l = 8$ for the problem with entropy constraints which takes about 900s (about 250s at $l = 7$) and $l = 9$ for the standard relaxation $\hat{C}_l^{loc, TI}$ in about 200s (about 20s at $l = 8$) on a laptop computer. The graph shows a significant improvement to the optimal value by using entropy constraints and for a fixed runtime, the entropy-constrained relaxation value outperforms the simpler SDP.

For comparison, we also add equivalent plots for two gapped models, the XXZ-model at $\Delta = -2$ and the transverse field Ising (TFI) model

$$h = -\sigma_z \otimes \sigma_z - g\sigma_x$$

at $g = 0.5$ in Fig. 4. While gapped models are typically computationally easier, it is not *a priori* clear whether the marginal relaxations perform better for these models. We do, however, observe much faster convergence for the TFI model (while convergence for the XXZ model is comparable to the XY model). Interestingly, the roughly quadratic speedup in terms of variable sizes persists even in this case.

Returning to the XXZ model we consider the geometry of the set of marginals. Due to the symmetry of the Hamiltonian, one can show that the two-body marginal of the ground state must be of the form³⁹

$$\rho(x, z) = \frac{1}{4}(1 \otimes 1 + x(\sigma_x \otimes \sigma_x + \sigma_y \otimes \sigma_y) + z\sigma_z \otimes \sigma_z). \tag{29}$$

In Fig. 5 we compare the convex sets \mathcal{C}^{TI} with $\hat{\mathcal{C}}_l^{loc, TI}$ and $\hat{\mathcal{C}}_l^{WM, TI}$ on the two-dimensional slice (29), as done in Ref. 35, i.e., we plot the set of valid states

$$\{(x, z) \in \mathbb{R}^2 : \rho(x, z) \in \mathcal{C}^{TI}\},$$

which we obtain from the analytic solution for infinite systems³⁶ with its relaxations

$$\{(x, z) \in \mathbb{R}^2 : \rho(x, z) \in \hat{\mathcal{C}}\}.$$

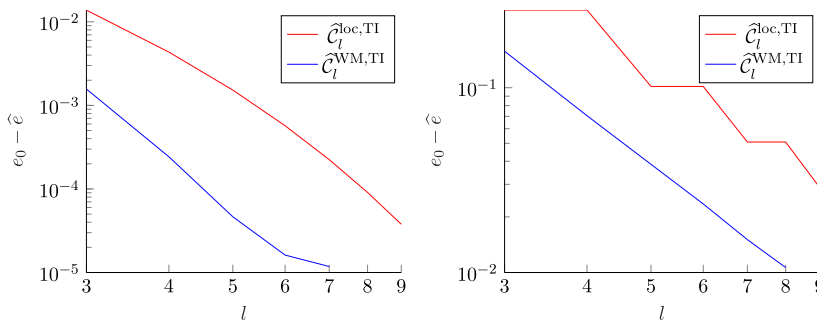


FIG. 4. Accuracy of the TFI model at $g = 0.5$ (left) and the XXZ model at $\Delta = -2$ (right). Both models are gapped. We observe a similar speedup to the XY-model, even for the much faster convergence rate in case of the TFI.

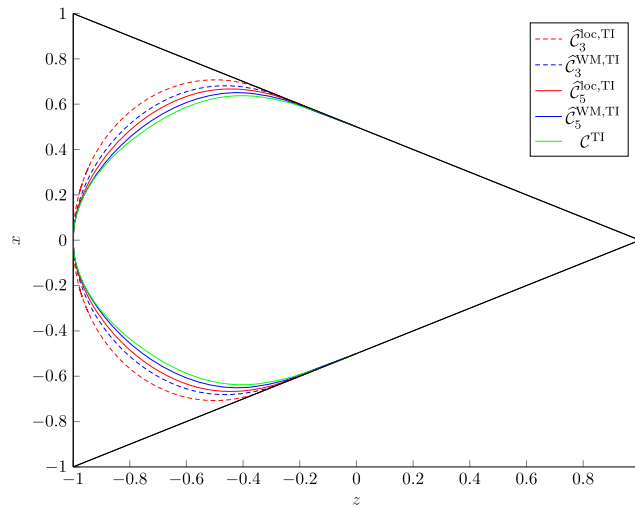


FIG. 5. Two-dimensional slice (29) of sets of feasible 2-body marginals. The blue lines are for the MED, the red lines for the unconstrained marginal relaxation, and the green line is for the exact analytic solution. As the size of the marginals increases the outer approximations become better. Note that we find numerically that $\widehat{C}_3^{\text{loc, TI}} = \widehat{C}_4^{\text{loc, TI}}$. We obtain almost the same optimal speedups as in Fig. 3.

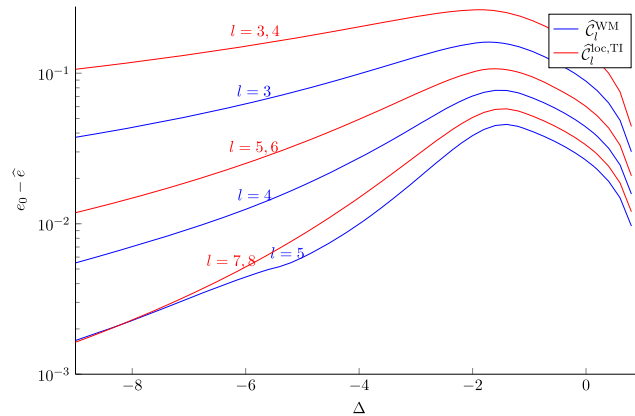


FIG. 6. The energy errors corresponding to the nontrivial part of the boundary of the state sets in Fig. 5.

We depict the approximation errors in Fig. 6.

ACKNOWLEDGMENTS

We would like to thank Matt Hastings for bringing the paper¹⁸ to our attention as well as an anonymous referee for multiple comments that improved the paper. H.F. acknowledges funding from UK Research and Innovation (UKRI) under the UK government’s Horizon Europe funding guarantee Grant No. EP/X032051/1. O.F. acknowledges funding by the European Research Council (ERC Grant AlgoQIP, Agreement No. 851716). S.O.S. acknowledges support from the UK Engineering and Physical Sciences Research Council (EPSRC) under Grant No. EP/W524141/1.

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

Hamza Fawzi: Writing – original draft (equal). **Omar Fawzi:** Writing – original draft (equal). **Samuel O. Scalet:** Writing – original draft (equal).

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

APPENDIX A: EDGE COVERING FOR CONNECTED GRAPHS

We give the Proof of Proposition 2.5.

Proof. The proof is by induction over the number of unmatched edges and the minimum distance between any pair of unmatched edges. We start from a set $P = \{\{e_1, e'_1\}, \dots, \{e_m, e'_m\}\}$ of pairs of edges, where e_i is adjacent to e'_i for all i and the $e_1, e'_1, e_2, \dots, e'_m$ are all distinct. We denote by \tilde{P} the set of all edges contained in elements of P . If $E \setminus \tilde{P}$ contains less than two edges, we are done. Otherwise, consider the pair of edges $f, g \in E \setminus \tilde{P}$ of minimal distance. If the distance is 0, add $\{f, g\}$ to P and repeat. If the distance is greater zero, there is a shortest path f_1, \dots, f_l from f to g , with $f_i \in \tilde{P}$ (otherwise if $f_i \notin \tilde{P}$ then the pair f_i, g has smaller distance). For some i , $f_1 = e_i$ (without loss of generality we can exchange e_i and e'_i). We remove $\{e_i, e'_i\}$ from P and add $\{e'_i, f\}$ if they are adjacent and $\{e_i, f\}$ otherwise. The number of unmatched edges is unchanged but the unmatched edges e'_i or e_i and g are now connected by the path f_2, \dots, f_l , so the minimum distance of unmatched edges has decreased by at least one. We repeat this procedure until the distance goes to zero, after which we can add a pair of edges and repeat until only one edge is unmatched. \square

APPENDIX B: DEFINITION OF THE TRANSLATION-INVARIANT GROUND ENERGY

We prove the existence of the limit ground energy by standard arguments.^{37,38}

Theorem B.1. *The following limits exist and are equal*

$$\lim_{m \rightarrow \infty} \frac{1}{m^D} \lambda_{\min}(H_{[1,m]^D}) = \lim_{m \rightarrow \infty} \min_{\rho_m \text{ TI}} \text{Tr}[\rho_m h] = \min_{\rho \in \mathcal{C}^{\text{TI}}} \text{Tr}[\rho h], \quad (\text{B1})$$

where the minimization in the middle expression goes over all translation-invariant states in $\mathcal{D}(\otimes_{v \in [1,m]^D} \mathbb{C}^d)$.

Proof. We first note that the existence of the limit in the middle expression is trivial since it is an increasing sequence. Indeed, a translation-invariant state on $[1, m]^D$ defines a translation-invariant state on $[1, m-1]^D$.

We prove the first equality. We rewrite the left-hand side as a minimization:

$$\frac{1}{m^D} \lambda_{\min}(H_{[1,m]^D}) = \frac{1}{m^D} \min_{\rho'_m} \text{Tr}[\rho'_m H_{[1,m]^D}]$$

over all states $\rho'_m \in \mathcal{D}(\otimes_{v \in [1,m]^D} \mathbb{C}^d)$. Clearly, the minimizer ρ'_m can be chosen periodic

$$\frac{1}{m^D} \min_{\rho'_m} \text{Tr}[\rho'_m H_{[1,m]^D}] = \frac{1}{m^D} \min_{\rho'_m \text{ periodic}} \text{Tr}[\rho'_m H_{[1,m]^D}] = \min_{\rho'_m \text{ periodic}} \text{Tr}[\rho'_m h]$$

meaning that ρ'_m is invariant under a permutation of sites mapping $v \rightarrow v + t$ modulo $m \forall t \in [1, m]^D$, which is a stronger condition than translation invariance. Therefore,

$$\frac{1}{m^D} \lambda_{\min}(H_{[1,m]^D}) \geq \min_{\rho_m \text{ TI}} \text{Tr}[\rho_m h].$$

For the reverse direction we note that

$$\min_{\rho_m \text{ TI}} \text{Tr}[\rho_m h] = \frac{1}{(m-1)^D} \min_{\rho_m \text{ TI}} \text{Tr}[\rho_m H_{[1,m-1]^D}]$$

due to translation-invariance and

$$\begin{aligned} \left\| \frac{1}{m^D} H_{[1,m]^D} - \frac{1}{(m-1)^D} H_{[1,m-1]^D} \right\| &\leq \left\| \frac{1}{m^D} H_{[1,m]^D} - \frac{1}{m^D} H_{[1,m-1]^D} \right\| \\ &\quad + \left\| \frac{1}{m^D} H_{[1,m-1]^D} - \frac{1}{(m-1)^D} H_{[1,m-1]^D} \right\| \\ &\leq \|h\| \left(1 - \left(\frac{m-1}{m} \right)^D \right) + \|H_{[1,m-1]^D}\| \left(\frac{1}{(m-1)^D} - \frac{1}{m^D} \right) \\ &= 2\|h\| \left(1 - \left(\frac{m-1}{m} \right)^D \right). \end{aligned}$$

We conclude the following bound

$$\begin{aligned} \min_{\rho_m \text{ TI}} \text{Tr} \left[\rho_m \frac{H_{[1,m-1]^D}}{(m-1)^D} \right] &\geq \min_{\rho_m \text{ TI}} \text{Tr} \left[\rho_m \frac{H_{[1,m]^D}}{m^D} \right] - \left\| \frac{H_{[1,m]^D}}{m^D} - \frac{H_{[1,m-1]^D}}{(m-1)^D} \right\| \\ &\geq \frac{1}{m^D} \lambda_{\min}(H_{[1,m]^D}) - 2\|h\| \left(1 - \left(\frac{m-1}{m} \right)^D \right), \end{aligned}$$

where the first inequality follows from lower bounding the objective, and in the second inequality we drop the translation-invariance constraint. Since the last term decays with m , the existence of limits and the first equality follow.

To prove the second equality in (B1), we note that one inequality is again trivial

$$\min_{\rho \in \mathcal{C}^{\text{TI}}} \text{Tr}[\rho h] \geq \lim_{m \rightarrow \infty} \min_{\rho_m \text{ TI}} \text{Tr}[\rho_m h]$$

as the minimizer in \mathcal{C}^{TI} defines finite translation-invariant states of any size and thereby feasible points for every element of the sequence.

To prove the reverse inequality we need to construct an infinite translation-invariant state from a finite one ρ_m that achieves the same objective value up to some error that decays with m . Let us consider the sequence of states

$$\rho_{m,n} = \frac{1}{m^D} \sum_{i \in [1,m]^D} \tau_i \left(\bigotimes_{j=1}^n \rho_m \right)$$

defined on $\mathcal{D}(\bigotimes_{i \in [1, nm]^D} \mathbb{C}^d)$. τ_i is the operator translating the system by i sites (with periodic boundary conditions). Since these states are all periodic and hence translation-invariant and have identical marginals on sites $[1, m]$ for any $n \geq 2$, it defines an element of \mathcal{C}^{TI} . Furthermore, for the objective value we have, using the notation $\vec{1} = (1, \dots, 1)$,

$$\begin{aligned} |\text{Tr}[\rho_{m,n} h] - \text{Tr}[\rho_m h]| &= \left| \frac{1}{m^D} \sum_{i \in [1,m]^D} \text{Tr} \left[\left(\tau_i \left(\bigotimes_{j=1}^n \rho_m \right) - \rho_m \right) h \right] \right| \\ &= \left| \frac{1}{m^D} \sum_{i=1}^D \text{Tr} \left[\left(\tau_{\vec{1}+e_i} \left(\bigotimes_{j=1}^n \rho_m \right) - \rho_m \right) h \right] \right| \\ &\leq \frac{2D}{m^D} \|h\| \end{aligned}$$

where the only nonzero term in the sum appear, where the marginal on sites $\vec{1}$ and $\vec{1} + e_i$ come from two different copies of ρ_m . □

The above proof also implies a convergence rate of $O(1/m)$ for both limits.

REFERENCES

- ¹Golub, G. H. and Van Loan, C. F., *Matrix Computations* (JHU Press, 2013).
- ²Gharibian, S., Huang, Y., Landau, Z., Shin, S. W. *et al.*, “Quantum Hamiltonian complexity,” *Found. Trends Theor. Comput. Sci.* **10**(3), 159–282 (2015).
- ³Wu, D., Rossi, R., Vicentini, F., Astrakhansev, N., Becca, F., Cao, X., Carrasquilla, J., Ferrari, F., Georges, A., Hibat-Allah, M. *et al.*, “Variational benchmarks for quantum many-body problems,” [arXiv:2302.04919](https://arxiv.org/abs/2302.04919) (2023).
- ⁴Bridgeman, J. C. and Chubb, C. T., “Hand-waving and interpretive dance: An introductory course on tensor networks,” *J. Phys. A: Math. Theor.* **50**, 223001 (2017).
- ⁵Cirac, J. I., Pérez-García, D., Schuch, N., and Verstraete, F., “Matrix product states and projected entangled pair states: Concepts, symmetries, theorems,” *Rev. Mod. Phys.* **93**, 045003 (2021).
- ⁶White, S. R., “Density matrix formulation for quantum renormalization groups,” *Phys. Rev. Lett.* **69**, 2863–2866 (1992).

- ⁷Hastings, M. B., “An area law for one-dimensional quantum systems,” *J. Stat. Mech.: Theory Exp.* **2007**, P08024.
- ⁸Landau, Z., Vazirani, U., and Vidick, T., “A polynomial time algorithm for the ground state of one-dimensional gapped local Hamiltonians,” *Nat. Phys.* **11**, 566–569 (2015).
- ⁹Liu, Y.-K., “Consistency of local density matrices is QMA-complete,” APPROX’06/RANDOM’06: Proceedings of the 9th International Conference on Approximation Algorithms for Combinatorial Optimization Problems, and 10th international conference on Randomization and Computation, August 2006 (2006), pp. 438–449.
- ¹⁰Mazziotti, D. A., “Quantum chemistry without wave functions: Two-electron reduced density matrices,” *Acc. Chem. Res.* **39**(3), 207–215 (2006).
- ¹¹Pironio, S., Navascués, M., and Acín, A., “Convergent relaxations of polynomial optimization problems with noncommuting variables,” *SIAM J. Optim.* **20**, 2157–2180 (2010).
- ¹²Baumgratz, T. and Plenio, M. B., “Lower bounds for ground states of condensed matter systems,” *New J. Phys.* **14**, 023027 (2012).
- ¹³Barthel, T. and Hübener, R., “Solving condensed-matter ground-state problems by semidefinite relaxations,” *Phys. Rev. Lett.* **108**, 200404 (2012).
- ¹⁴Brandao, F. G. and Harrow, A. W., “Product-state approximations to quantum ground states,” in *Proceedings of the Forty-Fifth Annual ACM Symposium on Theory of Computing* (ACM, 2013), pp. 871–880.
- ¹⁵Lindblad, G., “Expectations and entropy inequalities for finite quantum systems,” *Commun. Math. Phys.* **39**(2), 111–119 (1974).
- ¹⁶Nemirovski, A. and Nesterov, Y., “Interior point polynomial methods in convex programming: Theory and applications,” in *Studies in Applied and Numerical Mathematics* (SIAM, 1994).
- ¹⁷Fawzi, H. and Saunderson, J., “Optimal self-concordant barriers for quantum relative entropies,” *SIAM J. Optim.* **33**(4), 2858–2884 (2023).
- ¹⁸Poulin, D. and Hastings, M. B., “Markov entropy decomposition: A variational dual for quantum belief propagation,” *Phys. Rev. Lett.* **106**, 080403 (2011).
- ¹⁹Carlen, E. A., Lebowitz, J. L., and Lieb, E. H., “On an extension problem for density matrices,” *J. Math. Phys.* **54**, 062103 (2013).
- ²⁰Osborne, T. J., “Entropic bounds for the quantum marginal problem,” [arXiv:0806.2962](https://arxiv.org/abs/0806.2962) (2008).
- ²¹Linden, N. and Winter, A., “A new inequality for the von Neumann entropy,” *Commun. Math. Phys.* **259**, 129–138 (2005).
- ²²Gharibian, S. and Parekh, O., “Almost optimal classical approximation algorithms for a quantum generalization of max-cut,” in *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2019)* (Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik GmbH, Wadern, Saarbruecken, Germany, 2019).
- ²³Parekh, O. and Thompson, K., “Application of the level-2 quantum Lasserre hierarchy in quantum approximation algorithms,” in *48th International Colloquium on Automata, Languages, and Programming (ICALP 2021)* (Schloss Dagstuhl–Leibniz-Zentrum für Informatik, 2021).
- ²⁴Anshu, A., Gosset, D., and Morenz, K., “Beyond product state approximations for a quantum analogue of max cut,” in *15th Conference on the Theory of Quantum Computation, Communication and Cryptography (TQC 2020)* (Schloss Dagstuhl–Leibniz-Zentrum für Informatik, 2020).
- ²⁵Parekh, O. and Thompson, K., “Beating random assignment for approximating quantum 2-local Hamiltonian problems,” in *29th Annual European Symposium on Algorithms (ESA 2021)* (Schloss Dagstuhl–Leibniz-Zentrum für Informatik, 2021).
- ²⁶King, R., “An improved approximation algorithm for quantum max-cut on triangle-free graphs,” *Quantum* **7**, 1180 (2023).
- ²⁷Lee, E., “Optimizing quantum circuit parameters via SDP,” in *33rd International Symposium on Algorithms and Computation (ISAAC 2022)*, Leibniz International Proceedings in Informatics (LIPIcs) (Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2022), Vol. 248, pp. 48:1–48:16.
- ²⁸Kato, K. and Brandão, F. G. S. L., “Quantum approximate Markov chains are thermal,” *Commun. Math. Phys.* **370**, 117–149 (2019).
- ²⁹Franchini, F., “The XXZ chain,” in *An Introduction to Integrable Techniques for One-Dimensional Quantum Systems, Lecture Notes in Physics* (Springer International Publishing, 2017), pp. 71–92.
- ³⁰Fawzi, H., Saunderson, J., and Parrilo, P. A., “Semidefinite approximations of the matrix logarithm,” *Found. Comput. Math.* **19**, 259–296 (2019).
- ³¹Coey, C., Kapelevich, L., and Vielma, J. P., “Solving natural conic formulations with Hypatia.jl,” *INFORMS J. Comput.* **34**(5), 2686–2699 (2022).
- ³²Karimi, M. and Tunçel, L., “Primal–dual interior-point methods for domain-driven formulations,” *Math. Oper. Res.* **45**(2), 591–621 (2020).
- ³³Chambolle, A. and Pock, T., “A first-order primal-dual algorithm for convex problems with applications to imaging,” *J. Math. Imaging Vision* **40**, 120–145 (2011).
- ³⁴Faust, O., Fawzi, H., and Saunderson, J., “A Bregman divergence view on difference-of-convex optimization,” in *AISTATS*, 2023.
- ³⁵Verstraete, F. and Cirac, J. I., “Matrix product states represent ground states faithfully,” *Phys. Rev. B* **73**, 094423 (2006).
- ³⁶Yang, C. N. and Yang, C. P., “One-dimensional chain of anisotropic spin-spin interactions. II. Properties of the ground-state energy per lattice site for an infinite system,” *Phys. Rev.* **150**, 327–339 (1966).
- ³⁷Ohya, M. and Petz, D., *Quantum Entropy and its Use, Texts and Monographs in Physics* (Springer, Berlin, Germany, 1993).
- ³⁸Eisert, J., “A note on lower bounds to variational problems with guarantees,” [arXiv:2301.06142](https://arxiv.org/abs/2301.06142) (2023).
- ³⁹More precisely, we use the fact that the Hamiltonian is invariant under the action of the unitaries $U \in \{\sigma_x, \sigma_y, \sigma_z, S\}$ where $S = \exp(i\pi(1 - \sigma_z)/4)$, i.e., $U^{\otimes 2}h(U^\dagger)^{\otimes 2} = h$. The set of states (29) is exactly the set left invariant by this action.