



Universiteit
Leiden
The Netherlands

Certificates of quantum many-body properties assisted by machine learning

Requena, B.; Munoz-Gil, G.; Lewenstein, M.; Dunjko, V.; Tura, J.

Citation

Requena, B., Munoz-Gil, G., Lewenstein, M., Dunjko, V., & Tura, J. (2023). Certificates of quantum many-body properties assisted by machine learning. *Physical Review Research*, 5(1). doi:10.1103/PhysRevResearch.5.013097

Version: Publisher's Version

License: [Creative Commons CC BY 4.0 license](https://creativecommons.org/licenses/by/4.0/)

Downloaded from: <https://hdl.handle.net/1887/3718623>

Note: To cite this publication please use the final published version (if applicable).

Certificates of quantum many-body properties assisted by machine learning

Borja Requena ¹, Gorka Muñoz-Gil ^{1,2}, Maciej Lewenstein ^{1,3}, Vedran Dunjko ⁴, and Jordi Tura ^{5,6,*}

¹*ICFO - Institut de Ciències Fòniques, The Barcelona Institute of Science and Technology, Av. Carl Friedrich Gauss 3, 08860 Castelldefels (Barcelona), Spain*


²*Institute for Theoretical Physics, University of Innsbruck, Technikerstr. 21a, A-6020 Innsbruck, Austria*

³*ICREA, Pg. Lluís Companys 23, 08010 Barcelona, Spain*

⁴*LIACS, Leiden University, Niels Bohrweg 1, 2333 CA Leiden, The Netherlands*

⁵*Max-Planck-Institut für Quantenoptik, Hans-Kopfermann-Str. 1, 85748 Garching, Germany*

⁶*Instituut-Lorentz, Universiteit Leiden, P.O. Box 9506, 2300 RA Leiden, The Netherlands*

 (Received 23 April 2021; revised 13 October 2022; accepted 19 December 2022; published 10 February 2023)

Computationally intractable tasks are often encountered in physics and optimization. They usually comprise a cost function to be optimized over a so-called feasible set, which is specified by a set of constraints. This may yield, in general, to difficult and nonconvex optimization tasks. A number of standard methods are used to tackle such problems: variational approaches focus on parametrizing a subclass of solutions within the feasible set. In contrast, relaxation techniques have been proposed to approximate it from outside, thus complementing the variational approach to provide ultimate bounds to the global optimal solution. In this paper, we propose a novel approach combining the power of relaxation techniques with deep reinforcement learning in order to find the best possible bounds within a limited computational budget. We illustrate the viability of the method in two paradigmatic problems in quantum physics and quantum information processing: finding the ground state energy of many-body quantum systems, and building energy-based entanglement witnesses of quantum local Hamiltonians. We benchmark our approach against other classical optimization algorithms such as breadth-first search or Monte Carlo, and we characterize the effect of transfer learning. We find the latter may be indicative of phase transitions with a completely autonomous approach. Finally, we provide tools to tackle other common applications in the field of quantum information processing with our method.

DOI: [10.1103/PhysRevResearch.5.013097](https://doi.org/10.1103/PhysRevResearch.5.013097)

I. INTRODUCTION

In physics, we often encounter optimization tasks that are computationally intractable. These are often related to the characterization of complex systems that entail poorly scaling problems. For instance, in quantum physics, the complexity of characterizing many-body systems scales exponentially with the system size, rapidly becoming an intractable task. In such cases, we often rely on methods that provide an accurate approximation to the actual solution. There exist two main paradigmatic approaches: the variational ansatz and relaxation methods (see Fig. 1).

On the one hand, the variational ansatz consists in parametrizing a family of solutions with the hope that it contains a good approximation to the optimal one. In the context of quantum sciences, it has found tremendous success in areas so diverse as quantum chemistry [1–7], condensed matter [8–13], and quantum machine learning [14,15], and it is the main pillar upon which many modern quantum algorithms rest [13,16–25]. However, variational approaches are

suboptimal by construction and only yield an upper bound (in a minimization problem) to the optimal solution. Increasing the complexity of the ansatz may result in a better approximation at the price of increasing the overall computational cost. Furthermore, the distance between the obtained solution and the optimal one is unknown, in general. Even when we obtain it, we need additional methods to prove it is indeed the case.

On the other hand, we find relaxation methods. Any optimization task is characterized by the problem's constraints. All the points that fulfill them define the feasible set, which may be hard to optimize over in many cases. Therefore, to ease the optimization process, we may consider relaxing some of the constraints, obtaining a relaxed larger set. However, the minimum over a larger set can only be less or equal than the original one, thus resulting in a lower bound to the problem. Hence, the combination of the two approaches, variational and relaxations, yields an upper and lower bound that conform an uncertainty interval around the optimal solution, as illustrated in Fig. 1(b).

Relaxation techniques have been widely used in quantum information processing since its dawn. In this field, semidefinite programming (SDP) has been a successful and recurrent tool to build relaxations [26–28]. Perhaps, the most paradigmatic example in the context of entanglement theory is the Peres criterion, which is a relaxation from the set of separable states to the set of states that are positive under partial transposition (PPT) [29]. The membership problem in the separable

*Corresponding author: tura@lorentz.leidenuniv.nl

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/) license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

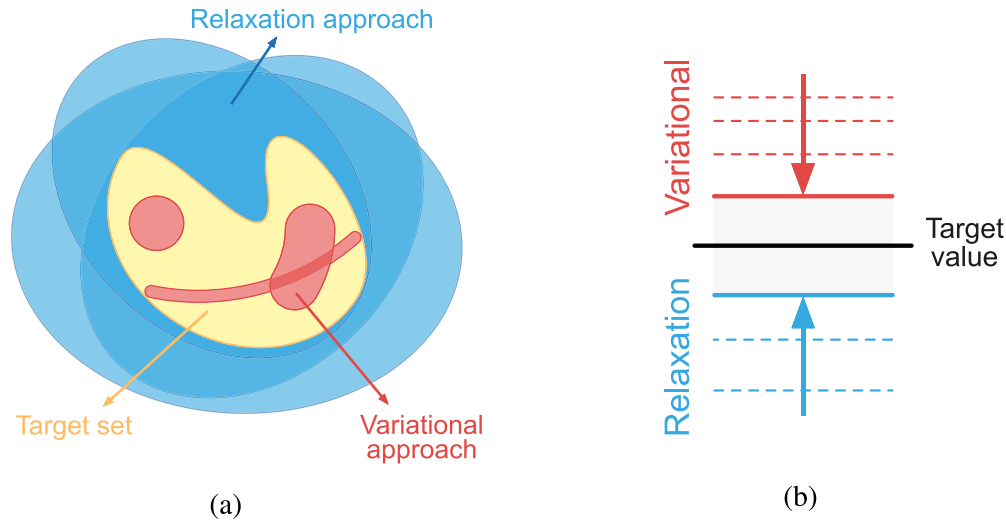


FIG. 1. Interpretation of exact solutions and bounds obtained through variational and relaxation methods. (a) Schematic representation of an optimization task. The goal is to optimize a function over a hard to characterize set (yellow set). The variational approach allows us to parametrize subsets within the one of interest (different red sets). Different parametrizations yield different subsets that are more or less convenient depending on the task. Relaxation techniques can efficiently represent larger sets than the one of interest (different blue sets) exploiting, for instance, convexity or linearity. Neither different variational approaches nor different relaxations need to be contained into one another, so the sets they represent are, in general, incomparable; (b) Values of the objective function. In black, the optimal unknown value. In red, the different minima obtained by variational methods. The smaller the value, the better the bound. In blue, different minima obtained by relaxation techniques. The greater the value, the more accurate their associated certificate. In grey, the uncertainty region where the optimal solution lies, given by the best variational and the best certificate obtained so far.

set is NP-hard [30], whereas checking the PPT criterion is very simple. Relaxation techniques also play a major role in the device-independent version of quantum information processing [31]. For instance, in cryptographic security proofs, we need to consider all possible quantum attacks, which are hard to characterize, motivating research for supraquantum theories that are analytically tractable [32,33]. In the quest to characterize the set of quantum correlations [34], several operationally simple, outer approximations have been proposed [35–42], as well as systematic relaxations via SdP [27,43,44]. Additionally, relaxation methods have found a wide range of applications in quantum physics and chemistry [45–54]. A recurrent theme is to find solutions that are simple enough to be understandable and computationally tractable, while being as accurate as possible.

Strongly relaxed problems may be easier to solve, although they may also yield looser bounds. At the same time, some relaxations may be more elegant/smarter than others, yielding better bounds while using similar computational resources. These generally exploit useful properties of the system, such as the existence of symmetries. However, without any prior knowledge about the problem at hand, or whenever it lacks such appealing properties, efficient relaxations may be highly elusive. Hence, it is crucial to devise methods to find, among all possible relaxations of the original problem, the best trade-off between accuracy and simplicity. Finding such optimal relaxation is a complex combinatorial optimization problem whose solution can reveal relevant properties of the underlying system.

To find it, we propose to harness the power of machine learning (ML). In the past decade, there has been an outstanding rise of the ML field, achieving remarkable feats [55–57]

with deep implications in industry and academia. In physics, all kinds of ML techniques have been used to tackle multiple challenging problems in the field [58,59]. For instance, in condensed matter, different models have been used to identify quantum phases of matter [60–62] and have been combined with existing variational methods to characterize complex Hamiltonians [63–65]. Closer to the problem we tackle in this paper, ML techniques have shown great success at dealing with combinatorial problems [66]. There have been proposed various approaches, from supervised learning of neural networks [67] to unsupervised methods over graphs [68] and, in particular, reinforcement learning (RL) [69]. While traditional algorithms rely on heuristics and specific insight about the problem, ML approaches are able to solve many of them faster and without any prior knowledge or assumption.

RL techniques [70] are responsible for some of the biggest breakthroughs in the ML field. Many of these relate to the efficient exploration of vast state spaces, leading to superhuman skills in games [55,71] or major advances in protein-structure prediction [72]. In physics, RL has been widely used for similar tasks, such as the design and control of quantum experiments [73,74], the preparation of quantum states [75], or the optimization of quantum error correction codes [76], which all feature broad combinatorial spaces akin to the task presented here.

Therefore, in this paper, we combine RL techniques with SdP-based relaxations to systematically search for optimal relaxations within a given computational budget. In the proposed scheme, an RL agent has access to a black box that computes the relaxation of the problem by solving an SdP [see Fig. 3(a)]. The agent can increase or decrease the relaxation level and observe an output that depends on both the

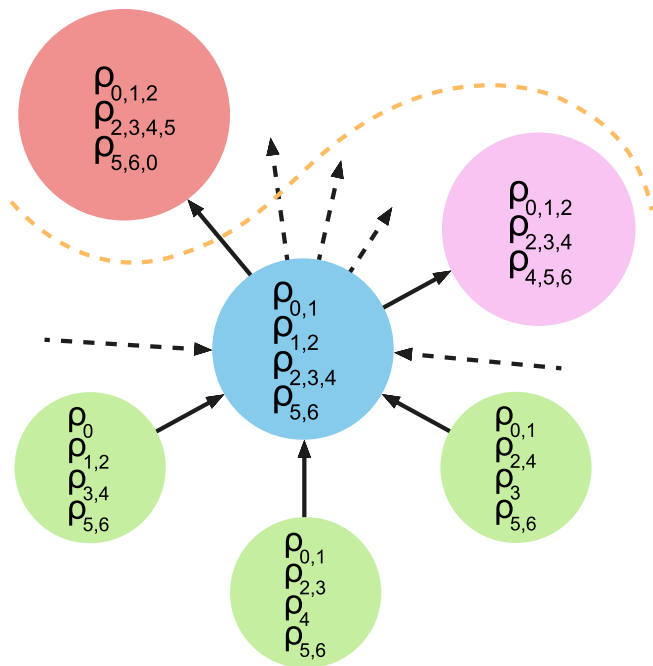


FIG. 2. Poset structure of the constraint space. The different circles represent Ξ_C for different $C \subseteq \mathcal{P}([n])$. The arrows represent the partial order relation \preceq , so that $\Xi_C \preceq \Xi_{C'}$ is represented by an arrow from Ξ_C to $\Xi_{C'}$. Only the arrows relative to the central node are drawn. Dashed arrows indicate that there exist many more $\Xi_{C'}$ arriving/departing from the central node that are simply not drawn. The orange-dashed line separates those Ξ_C that fall into the allowed computational budget (green, blue, and pink nodes) from those that are too costly (red). Moving vertically up into the diagram provides better certificates, but at a higher cost. Since \preceq is a partial order relation, some nodes (e.g., the three at the bottom) are incomparable.

associated computational cost and the quality of the obtained bound. We illustrate the procedure in tackling two paradigmatic problems in quantum physics and quantum information: finding the ground-state energy of local many-body Hamiltonians and building energy-based entanglement witnesses. We first show our results on the first problem in which, even for very simple scenarios, we find counterintuitive optimal relaxations. Then, we compare our RL approach to other classical optimization algorithms and we show how to use transfer learning to explore the phase diagram of the Hamiltonians in an autonomous way. Finally, we present our results in the second problem tackling two different examples. Applying RL to obtain useful relaxations can be seen as a meta-algorithm with a wide range of applicability. Here, we present it in two different cases of study without hindering its more general flavor. In Sec. VII we discuss how the same principles apply to diverse areas of quantum information processing.

The paper is structured as follows: In Sec. II, we describe the mathematical formalism to build relaxations for our cases of study. In Sec. III, we introduce the constraint space over which the RL agent operates, and we introduce the optimization framework in Sec. IV. We present our main results in Sec. V and Sec. VI. Then, we discuss how our framework naturally applies to various relevant problems in quantum information in Sec. VII. Finally, we conclude in Sec. VIII.

II. BUILDING RELAXATIONS

In this section, we introduce the methods to systematically build relaxations, which we consider throughout the paper. These relaxations are based on a semidefinite program (SdP), whose optimal solution constitutes a certificate, as we detail in the section right below. We incorporate this methodology in the RL framework, introduced in Sec. IV, as a black box module with which an RL agent can interact. The agent can define the set of constraints to consider in the SdP, although it is agnostic to the calculation of the certificate and it only receives a reward at the end. Then, the agent can decide to tighten or loosen the relaxation by modifying the constraints. Here, we thoroughly present the mathematical formalism of the different steps involved in the process.

In the interest of simplicity, we introduce the main concepts of our methodology applied to the optimization tasks that we tackle throughout this paper: finding the ground-state energy of quantum local Hamiltonians and building energy-based entanglement witnesses. While these do not restrict the applicability of our paper to other areas in quantum information (see Sec. VII), it shall certainly ease the exposition. The mathematical formalism of the second task is an extension of the first one. Hence, we show here how to relax the ground state energy problem and, then, we extend it to the entanglement witnessing in Sec. VIA.

A. Finding the ground-state energy

Consider the optimization task to find the ground-state energy E_0 of a quantum local Hamiltonian

$$H = \sum_{i=1}^m H_i. \tag{1}$$

The Hamiltonian H acts on n qubits, and it is a sum of terms H_i , each of which acts on at most $k = O(1)$ qubits. The sum Eq. (1) has therefore $m = O(\text{poly}(n))$ terms. The support of H_i , denoted $\text{supp}(H_i)$ is the set of qubits where H_i acts non-trivially. The supports of the different H_i may overlap; *i.e.*, $\text{supp}(H_i) \cap \text{supp}(H_j)$ may not be empty.

To find E_0 , a possibility is to directly construct a quantum state that has E_0 energy with respect to H . Therefore, a first possible approach is to parametrize a family of quantum states $|\psi(\theta)\rangle$ exploiting some known properties of H . We can safely assume the parametrization yields a valid (*i.e.*, normalized) quantum state for any value of the parameters θ . Additionally, by construction, $\langle \psi(\theta) | H | \psi(\theta) \rangle \geq E_0$ for all θ . Let us denote

$$\gamma = \min_{\theta} \langle \psi(\theta) | H | \psi(\theta) \rangle, \tag{2}$$

which satisfies $\gamma \geq E_0$ by construction. An example of such a parametrization would be to describe $|\psi(\theta)\rangle$ as a tensor network contraction, which exploits the locality properties of H , limiting the entanglement present in its ground state [12,77–80].

Complexity theory results (in particular, QMA-hardness) strongly suggest that finding, or even approximating, the ground state energy of a local Hamiltonian is a hard task, even for a quantum computer [81–83]. Furthermore, this hardness persists in physically relevant instances [84]. Notice that, even

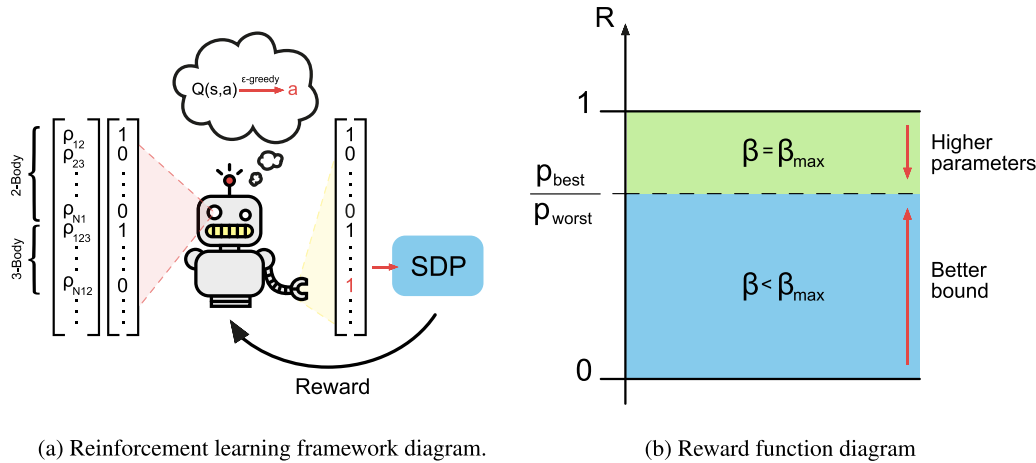


FIG. 3. (a) Schematic representation of the reinforcement learning framework. First, the agent observes the state: a one-hot encoding of the active constraints. Given the observation, it estimates the Q-values associated to the possible actions with a deep Q-network. Then, it decides which action to take according to an ϵ -greedy policy, bringing the agent to a new state. Finally, the black box solves the SdP associated to the new state, providing the agent with a reward, whose parts from Eq. (9) are illustrated in (b).

if we found the actual solution $|\psi(\theta)\rangle$, we cannot prove, solely from that, that it is the global minimum [85].

It is therefore highly desirable to obtain a bound from the other side; *i.e.*, a value β for which one can prove $E_0 \geq \beta$. This would guarantee $E_0 \in [\beta, \gamma]$ and, thus, help determine whether it is worth to refine the search depending on $|\gamma - \beta| < \epsilon$. However, for a proof of the type $E_0 \geq \beta$, constructing an example $|\psi(\theta)\rangle$ is not good enough. We need a proof that is satisfied by all valid quantum states and, possibly, a larger set, as long as it makes the proof simpler. Such a proof is referred to as a certificate, and it is typically obtained by numerical means. SdP is a natural tool to obtain such certificates upon which we capitalize in our paper. The optimality of the SdP solution or, at least, a valid bound for a certificate follows from strong or weak duality properties, respectively (see Appendix A).

B. Building a trivial relaxation

A common technique to build a relaxation for the local Hamiltonian problem is via the triangle inequality [45,86–88],

$$\min_{\rho} \text{Tr}[\rho H] \geq \sum_i \min_{\rho_i} \text{Tr}[\rho_i \hat{H}_i], \tag{3}$$

where ρ and ρ_i are density matrices acting on the support of H and H_i respectively. Note that i refers to a Hamiltonian term and it has nothing to do with the i -th party. Furthermore, in Eq. (3), the \hat{H}_i are sums of some local terms H_j of Eq. (1), grouped so that $\text{supp}(\hat{H}_i)$ is as large as possible while still allowing for computation of their minimal eigenvalue. This size obviously depends on the available computational resources.

Let us observe that the right-hand side in Eq. (3) is a sum of minima, where each minimization is carried out independently. Due to this independence, in general, it is not the case that different ρ_i are mutually compatible; *i.e.*, that there exists a global state ρ such that each ρ_i is the corresponding partial trace of ρ . The converse is true; however, every valid quantum state ρ has an associated set of partial traces ρ_i , but given a

set of ρ_i , a global ρ may not exist. This is what proves the inequality Eq. (3).

The minimization of the right-hand side of Eq. (3) is equivalent to solving the following SdP (cf. Appendix A):

$$\beta_{\theta} := \min_{\{\rho_i\}} \sum_i \text{Tr}[\rho_i \hat{H}_i] \tag{4}$$

$$\text{s.t.} \quad \rho_i \geq 0$$

$$\text{Tr}[\rho_i] = 1.$$

Since there is no mutual compatibility enforced among the ρ_i , and each is treated independently, the triangle inequality Eq. (3) constitutes a trivial relaxation. A natural way to strengthen the relaxation is to impose further restrictions on the collection of possible ρ_i , in such a way that any quantum state would also satisfy them. The strongest restriction possible is to directly ask that $\{\rho_i\}$ come from a global quantum state. Unfortunately, this would be equivalent to finding the value of E_0 , which is QMA-complete. Furthermore, it is strongly connected to solving the so-called quantum marginal problem (QMP), which is also QMA-complete [81–83]. The QMP has been solved completely in very rare instances, such as the global state being symmetric [89] or for the case of one-body marginals [90–92]. Nevertheless, the SdP based formulation Eq. (4) motivates a hierarchy of relaxations based on solving the QMP up to some degree of compatibility.

C. Building tighter relaxations

In order to build relaxations that yield a tighter bound than that of the triangle inequality, our first observation is that the set $\{\rho_i\}$ does not need to fulfill any mutual compatibility constraint. It would be natural to expect that, at least, the partial traces on different supports’ intersection match. This will reduce the space of solutions, provided that $\{\rho_i\}$ must fulfill additional conditions. Therefore, since the minimization is over a smaller set, its result can only be a tighter bound.

Hence, the first level of compatibility we might want to ask for is that ρ_i and ρ_j yield the same reduced density matrix (RDM) on their common support, which we shall denote $\rho_{i \wedge j}$,

$$\text{Tr}_{\text{supp}(\rho_j)^c}[\rho_i] = \text{Tr}_{\text{supp}(\rho_i)^c}[\rho_j] \equiv \rho_{i \wedge j}. \tag{5}$$

Here, the partial trace $\text{Tr}_S(\cdot)$ denotes that we eliminate subsystem S and the superindex c indicates the complementary set. Thus, Tr_{S^c} produces the RDM acting on the subsystem S . Note that the partial trace condition is linear in ρ_i . Therefore, it can be naturally imported into Eq. (4) and still be formulated in terms of a SdP,

$$\beta_1 := \min_{\{\rho_i\}} \sum_i \text{Tr}[\rho_i \hat{H}_i] \quad (6)$$

s.t. $\rho_i \geq 0$
 $\text{Tr}[\rho_i] = 1$
 $\text{Tr}_{\text{supp}(\rho_i)^c}[\rho_i] = \rho_{i \wedge j}$.

Given that the sets of $\{\rho_i\}$ that satisfy the constraints of Eq. (6) also satisfy the constraints of Eq. (4), we have $\beta_\emptyset \leq \beta_1 \leq E_0$, by construction.

The certificates obtained from Eq. (6) can be further strengthened by adding virtual RDMs. For instance, even if H is 2-local, we might want to ask, e.g., that the two-body RDMs acting on *Alice* – *Bob* and *Bob* – *Charlie* are such that they both come from a virtual three-body density matrix acting on *Alice* – *Bob* – *Charlie*. The latter is not strictly necessary in order to compute the energy, for 2-body density matrices suffice, but this compatibility condition further restricts the set $\{\rho_i\}$, therefore improving the bound. In mathematical jargon, this method is known as representing the feasible set as a projected spectrahedra [93]. Hence, instead of solely asking that ρ_i and ρ_j yield the same RDM on their intersection, now we might impose a stronger constraint, which is that ρ_i and ρ_j come from a valid density matrix $\rho_{i \vee j}$ defined on the union of their supports,

$$\beta_2 := \min_{\{\rho_{i \vee j}\}} \sum_i \text{Tr}[\rho_i \hat{H}_i] \quad (7)$$

s.t. $\rho_{i \vee j} \geq 0$
 $\text{Tr}[\rho_{i \vee j}] = 1$
 $\text{Tr}_{\text{supp}(\rho_i)^c}[\rho_{i \vee j}] = \rho_i$.

We observe that the constraints imposed in Eq. (7) automatically imply those of Eq. (6), so we have omitted their writing, as they became redundant.

We also observe that, although now we have $\beta_\emptyset \leq \beta_1 \leq \beta_2 \leq E_0$, the cost of solving Eq. (7) is substantially higher than that of Eq. (6), because the SdP variables $\rho_{i \vee j}$ act on more qubits than ρ_i and the cost of representing them grows expo-

nentially in the number of qubits. Similarly, the relaxations from Eq. (7) can be further strengthened by considering compatibility with more regions, yielding a chain of inequalities $\beta_\emptyset \leq \beta_1 \leq \beta_2 \leq \dots \leq E_0$.

In Eq. (7) the compatibility constraints are enforced on all possible pairs (i, j) . However, not all the constraints are equally useful. In an extreme case, when $\text{supp}(\rho_i) \cap \text{supp}(\rho_j) = \emptyset$, adding the variable $\rho_{i \vee j}$ with its respective constraints makes no difference. Indeed, since $\text{Tr}[\rho_i \hat{H}_i + \rho_j \hat{H}_j] = \text{Tr}[(\rho_i \otimes \rho_j)(\hat{H}_i \otimes \mathbb{1}_j + \mathbb{1}_i \otimes \hat{H}_j)]$, the choice $\rho_{i \vee j} = \rho_i \otimes \rho_j$ is always possible, as it satisfies the rest of constraints, therefore not changing β_2 . We remark this tensor product choice is possible because the supports do not intersect. However, if we define $\rho_{i \vee j}$ as a variable in Eq. (7), we increase its computational complexity without improving the bound, thus yielding a worse certificate.

In Appendix A we give details on the basics of SdP and how to obtain mathematical proofs from their solutions.

III. THE CONSTRAINT SPACE

Following the methodology presented in the previous section, we combine it with the RL formalism in order to create a suitable environment with well defined actions, states, rewards, and their respective relationships, which constitute the fundamental elements of any RL problem. With this goal, we propose to use the space of constraints, which induces an underlying structure for the action and state spaces that the RL agent will explore, as we explain in Sec. IV. In this section, we present the constraint space and study its structure.

Continuing with our example, let us consider a set of n qubits, labeled from 0 to $n-1$, and denote $[n] = \{0, \dots, n-1\}$. Let $\mathcal{P}([n]) = \{\emptyset, \{0\}, \{1\}, \dots, \{n-1\}, \{0, 1\}, \{0, 2\}, \dots, [n]\}$ denote the parts of $[n]$; i.e., the set of all subsets of $[n]$, thus containing 2^n elements.

Notice that we can associate a certificate to every subset $C \subseteq \mathcal{P}([n])$ in the following way: for each element $S \in C$, corresponding to a subset of $[n]$, we consider the RDM acting on the qubits labeled by the elements in S , which we denote ρ_S . Let us denote $\Xi_C := \{\rho_S\}_{S \in C}$ the collection of RDMs associated to C . By enforcing compatibility on their overlapping supports, we can define the SdP

$$\beta_C := \min_{\Xi_C} \sum_i \langle H_i \rangle \quad (8)$$

s.t. $\rho_S \geq 0 \quad \forall S \in C$
 $\text{Tr}[\rho_S] = 1$
 $\text{Tr}_{R^c}[\rho_S] = \text{Tr}_{R^c}[\rho_{S'}] \quad \forall R \subseteq S \cap S', \quad S, S' \in C,$

where the partial trace over the whole system is set to one by convention $\text{Tr}_{[n]}[\rho] = 1$. We have written the objective function as $\sum_i \langle H_i \rangle$ for the following reasons: first, C could be small enough so that there is no $S \in C$ such that $\text{supp}(H_i) \subseteq S$. If this is the case, then we substitute $\langle H_i \rangle$ by the minimal eigenvalue of H_i , in the same spirit as the trivial relaxation from Eq. (4). Hence, if $C = \emptyset$, the cost function of Eq. (8) amounts to the sum of the minimal eigenvalue of each H_i . Otherwise, if Ξ_C contains a density matrix ρ_i whose support contains the support of H_i , we simply compute

$\langle H_i \rangle = \text{Tr}[\rho_i H_i]$. Note that, in case that multiple density matrices from Ξ_C could be used to compute $\langle H_i \rangle$, the last constraint of Eq. (8) guarantees the result is well defined; i.e., independent of the choice $\rho_i \in \Xi_C$. In practice, the last constraint of Eq. (8) rarely needs to be imposed over all the subsets of the intersection, and it is enough to take $R = S \cap S'$ for all pairs $S, S' \in C$. In Appendix B we discuss these implications in a detailed way. Regardless of the constraint implementation of Eq. (8), the SdP yields a valid lower bound.

Furthermore, given a set of constraints $C \subseteq \mathcal{P}([n])$, it is not necessary to define Eq. (8) over all the variables contained in Ξ_C . If some $S \in C$ is contained in another $S' \in C$, such that $S \subseteq S'$, we can simply use $\rho_{S'}$, as it contains all the information on ρ_S . This choice is well-defined due to the constraints in Eq. (8) and it naturally defines a simplification function $s : \Xi_C \mapsto s(\Xi_C)$, which allows to simplify the SdP by removing redundant variables.

One of the main motivations of this paper is to optimize the quality of the lower bound within a limited computational budget. The asymptotic complexity of an SdP with m variables of matrix size n depends on the method that it is used to solve it. A rough estimate is $O(m^2n^2)$, but iteration costs of the algorithm are not factored in [94]. There exist interior-point methods, which are faster than the ellipsoid method [95], e.g., Alizadeh's algorithm runs in $\tilde{O}(\sqrt{m}(m+n^3)L)$ time, where L is an input parameter and the \tilde{O} notation is used to suppress $\text{polylog}(mn/\varepsilon)$ terms, where ε is the required precision [96,97]. In our case, we use the self-dual minimization method SeDuMi [98], which has a complexity $\tilde{O}(m^2n^{5/2} + n^{7/2})$ for large-scale instances, although there are algorithms of $\tilde{O}(nm^3)$, suitable for small matrix sizes [99]. Interestingly, quantum algorithms have been proposed to solve SdP [100], and ML methods have been studied to aid the SdP solver [101].

In light of the whole zoo of algorithms for SdP and their various complexities, it is clear the time complexity of an SdP instance is highly dependent on the solver used. Nevertheless, for our case study it is important that a given computational budget will determine a set of maximal (m, n) that are allowed, which we estimate by effectively limiting the size and contents of Ξ (cf. Sec. V B).

The space of constraints forms a partially ordered set (poset) with respect to the following partial order relation. Given $C, C' \in \mathcal{P}([n])$, we say $C \preceq C'$ if, and only if, for each $S \in C$ there exists a $S' \in C'$ such that $S \subseteq S'$. The motivation of the partial order relation \preceq is that $C \preceq C'$ implies $\beta_C \leq \beta_{C'}$ by construction: every density matrix in Ξ_C can be obtained by tracing out some elements of another density matrix in $\Xi_{C'}$, and the constraints in Eq. (8) enforce mutual compatibility among all the elements in Ξ_C and $\Xi_{C'}$. In Fig. 2 we illustrate such structure, which motivates the definition of the RL formalism in Sec. IV.

IV. CONSTRAINT OPTIMIZATION WITH REINFORCEMENT LEARNING

In this section, we propose a method to obtain the best possible certificate within a certain computational cost by exploring the constraint space described in Sec. III. Due to the high amount of structure in this extensive combinatorial space, we propose to use reinforcement learning (RL) [70] with function approximation. In our proposed framework, it naturally favors lower cost solutions and can optimize the exploration strategy based on previous experiences. In such spaces, experience in one region may be useful in others, e.g., in periodic systems, actions in one domain should be identical to actions in another, which further allows for easy transfer of learning without explicit analysis of the model parameters (see Sec. V C).

To this end, we frame the optimization problem as a Markov decision process (MDP). The MDP is defined through a state space, an action space, a transition function between states given an action, and a reward function, which associates a value to each state-action-state tuple. We detail all these elements below. A learning agent, as the learning program is called in RL terminology, explores the constraint space with the goal to find the set of constraints $C^* \subseteq \mathcal{P}([n])$ that provides the best possible certificate within a limited computational budget, while using the least amount of resources. In algorithmic terms, we distinguish two main independent parts:

(i) An environment that hosts the constraint space restricted by the computational budget, as in Fig. 2. It also contains a black box that takes a set of constraints C as input, computes β_C by solving the associated SdP [Eq. (8)] and outputs a reward.

(ii) A learning agent that navigates the environment's constraint space (i). At every point in the space, it inputs the corresponding set of constraints C into the black box. The agent can choose to strengthen or loosen the constraints, effectively exploring the constraint space with its actions. In doing so, the agent obtains different rewards that guide it towards finding the optimal relaxation. Note that the agent is completely agnostic about the actual physical problem at hand.

We aim to understand up to which extent such a fully automated approach may help in studying physical systems. In the following, we connect the MDP components to our running example. See Fig. 3(a) for a schematic depiction.

State space. The state space corresponds to the constraint space introduced in Sec. III, in which each state is a set of constraints $C \subseteq \mathcal{P}([n])$ and it is bound by the computational budget, as illustrated in Fig. 2. We represent the states by one-hot encoding of the active constraints $S \in C$: considering a set of 2^n -dimensional canonical vectors with only a nonzero unit element, each representing an element $S \in \mathcal{P}([n])$, a state vector is the sum of the vectors that encode the components $S \in C$. Equivalently, it identifies the set $\Xi_C = \{\rho_S\}_{S \in C}$ of RDMS that enter as variables in Eq. (8). As shown in the leftmost part of Fig. 3(a), the RDMS ρ_S are ordered according to their dimension in the state vector. Out of the 2^n possible variables, we need only consider $\text{poly}(n)$ of them, effectively reducing the state vector size: we can ignore the 1-body constraints as well as those ρ_S whose sole contribution to the cost of solving the associated SdP would exceed the computational budget. With a computational budget B , this leaves $n^{O(\log(B))}$ available RDMS to construct the certificate. If no $S \in C$ is such that $i \in S$, the 1-body constraint corresponding to $\rho_{\{i\}}$ is added by default. Therefore, the smallest set of constraints that we allow for is $C = \{\{0\}, \dots, \{n-1\}\}$, represented by a state vector of zeros. We take it as the initial state of the MDP unless stated otherwise.

Actions. An action a consists of either adding or removing a constraint, driving the agent from one state to another. In practice, actions flip bits in the state vector corresponding to the encoded constraints. The agent is free to add a constraint of any size, as long as the cost associated to the resulting set is within the computational budget. For instance, the agent can start by adding a 4-body constraint, e.g., ρ_{0123} , to the

initial state. In contrast, removing a constraint has a different effect. In order to keep the state space exploration consistent, removing a constraint splits it into its most immediate components of a lower degree. For instance, in 1D, removing ρ_{0123} would result in ρ_{012} and ρ_{123} . Note that a valid action always corresponds to an arrow (in both directions) in the poset depicted in Fig. 2, remaining within the boundary (not crossing the orange dashed line).

Transition function. The transition function is a simple deterministic function implicitly defined above: $T(C|a, C')$ is a Kronecker delta, attaining unit value if the constraint configuration C is reached by adding or removing the constraint specified by the action a from the set of constraints C' .

Reward. We define the reward function to match the overall optimization goal, provided that the learning agent aims to maximize the obtained reward. The reward associated to a state C depends on: (1) the energy bound β_C , obtained solving its associated SdP [Eq. (8)], and (2) its computational cost. In practice, we take the amount of free parameters in the SdP, which we denote by p , as a representation of the computational cost. Note that, in general, given an optimization problem, we have no prior knowledge about the optimal β^* and its cost p^* . Therefore, in order to compute the reward, we rely on a set of references that are updated as the agent explores. More precisely, we keep track of the best and worst bounds obtained, β_{\max} and β_{\min} respectively, and the best and worst set of parameters with which the best bound β_{\max} has been observed, denoted p_{best} and p_{worst} respectively. We compute the reward associated to a state by comparing its corresponding β and p to the reference values,

$$R(\beta, p) = \frac{p_{\text{best}}}{p_{\text{worst}}} \cdot \begin{cases} \frac{p_{\text{worst}}}{p} & \text{if } \beta = \beta_{\max} \\ \left(\frac{\beta - \beta_{\min}}{\beta_{\max} - \beta_{\min}} \right)^d & \text{otherwise,} \end{cases} \quad (9)$$

where d is a fixed exponent that controls the shape of the line $(\beta - \beta_{\min})/(\beta_{\max} - \beta_{\min})$. We introduce this exponent to provide better discrimination among the higher bounds, typically $d = 5$. Notice that $p_{\text{worst}} \geq p_{\text{best}}$ and, therefore, $p_{\text{worst}}/p \geq 1$. Thus, the prefactor $p_{\text{best}}/p_{\text{worst}} \leq 1$ ensures that $R(\beta, p) \in [0, 1]$, $\forall \beta, p$. Figure 3(b) shows a schematic of the reward function. In summary, the reward function mainly focuses on the resulting bound β , unless various states provide the maximum possible bound β_{\max} . In this case, those with higher computational costs are penalized.

The agent. Within the proposed framework, we can perform the constraint optimization with various methods. As mentioned before, we propose to use RL with function approximation. The learning program or agent specifies the policy by which actions are taken with the ultimate goal of maximizing the obtained reward. More precisely, we use double deep Q learning [102–104] with an ϵ -greedy policy π . At each state C , the agent estimates the Q values $Q^\pi(C, a)$ of each possible action a , a measure of the expected rewards associated with taking each action and then following the policy π . The ϵ -greedy policy dictates that actions are taken according to

$$\pi(C) = \begin{cases} \arg \max_a Q^\pi(C, a), & \text{with probability } (1 - \epsilon) \\ \text{uniform random } a, & \text{with probability } \epsilon. \end{cases} \quad (10)$$

Figure 3(a) shows a schematic representation of the whole process. In Sec. V we show that such an approach leads to the optimal relaxation faster compared to other classical optimization methods and, sometimes, it is able to find it even where the other methods fail.

V. GROUND-STATE ENERGY BOUNDS FOR THE HEISENBERG XX MODEL

We apply the method described in the previous sections to find lower bounds to the ground-state energy of quantum local Hamiltonians. To illustrate the process, we focus on a paradigmatic condensed matter model: the anti-ferromagnetic 1D quantum Heisenberg XX model [105], described by the Hamiltonian

$$H = \sum_{i=0}^{n-1} J_i (\sigma_i^x \sigma_{i+1}^x + \sigma_i^y \sigma_{i+1}^y) + \sum_{i=0}^{n-1} B_i \sigma_i^z, \quad (11)$$

where σ^α , $\alpha = x, y, z$, are the Pauli matrices, J_i is the antiferromagnetic exchange interaction between spins, and B_i is the strength of the external magnetic field. We consider periodic boundary conditions, such that $\sigma_n^\alpha = \sigma_0^\alpha$. In the homogeneous case, i.e., $J_i = J, B_i = B \forall i$, the model presents a quantum phase transition at $B = 2J$ [106] between an antiferromagnetic and a paramagnetic phase, in which the entanglement vanishes [107–109], see also [110]. We will hence refer to these phases as entangled and unentangled, respectively. Although the 1D XX model Eq. (11) is efficiently solvable via the Jordan-Wigner transformation [111], corresponding to a quadratic fermionic Hamiltonian [112–114], the agent is oblivious to such information. We emphasize that the points in the search space have no semantics to the agent, which, moreover, is not provided with any information about the Hamiltonian in any explicit way. This guarantees that our approach is as generally applicable as possible.

A. Optimal relaxations

Here, we present the results applying the RL method to the homogeneous version of the aforementioned Hamiltonian. For this case, we consider a computational budget that allows for the allocation of up to half of all the possible 3-body constraints. With these conditions, we find the best approximation to the ground state across the whole phase diagram of the Hamiltonian.

Unentangled phase ($B/J \geq 2$). In the unentangled phase, the ground state can be perfectly described by the set of independent 1-body RDMs. Therefore, we would expect the optimal set of constraints to be the minimum that the agent can consider $C = \{\{0\}, \dots, \{n-1\}\}$. Nevertheless, this is only true in the extreme case of $J = 0$. In a general scenario, with $0 < 2J \leq B$, the optimal set of constraints is made out of 2-body RDMs, as shown in Fig. 4(d). This is to provide support for the 2-body terms of the local Hamiltonian. Recall that, in our implementation, whenever a term H_i of the Hamiltonian is not supported by the set of RDMs $\Xi_C = \{\rho_S\}_{S \in C}$, we take $\langle H_i \rangle$ to be its minimal eigenvalue $\min(\sigma(H_i)) = -J$. With 2-body constraints, the resulting RDMs are rank-1 projectors, which correspond to pure states such that $\langle H_i \rangle = 0$ for the 2-body terms, thus yielding a better energy bound. Increasing the size

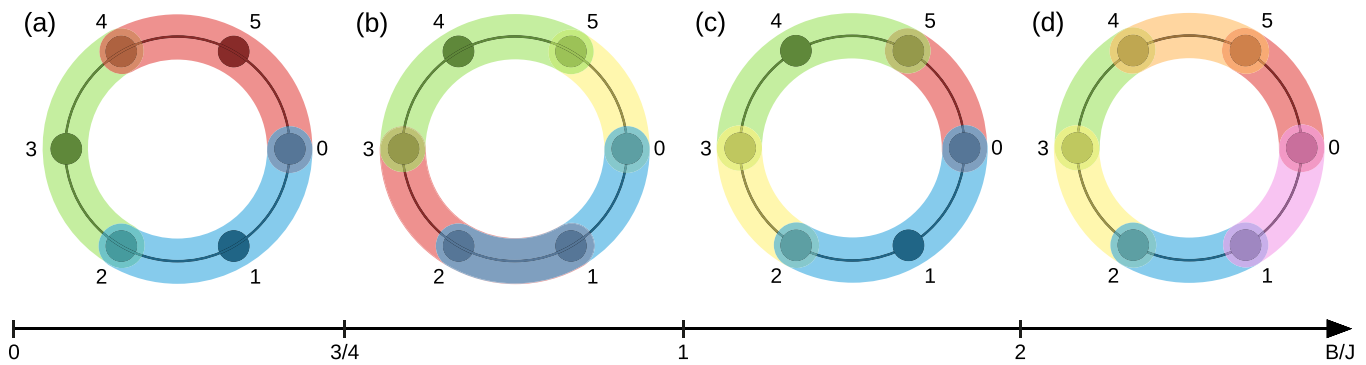


FIG. 4. Illustrative representation of the optimal constraints C^* that yield the best ground-state energy bound with minimal cost for the homogeneous Heisenberg XX model in 1D [Eq. (11)]. Each color represents the support of an RDM ($S \in C^*$) and we impose compatibility constraints over the overlapping areas. The results are obtained with a budget that allows for the allocation of up to half of the 3-body RDMs and $n = 6$ spins (black circles). For this case, the RL algorithm finds four different optimal solutions across the phase diagram. Interestingly, the entangled phase ($B/J < 2$) shows two intermediate solutions: (c) and (b), before the expected one (a) at $B/J < 3/4$.

of the constraints any further does not improve the energy bound at all.

Entangled phase ($B/J < 2$). In the case of the entangled ground state, its exact energy can only be obtained by considering the system as a whole, corresponding to $C = \{[n]\}$. Therefore, the agent can only provide the best possible approximation to the exact energy within the allowed computational budget. Just like in the previous case, it may seem reasonable to expect the optimal set of constraints to be unique for the whole phase. Nevertheless, the agent finds three separate regimes as depicted in Fig. 4:

(i) Close to the phase transition, the optimal relaxation is obtained by alternating 2-body and 3-body constraints, as shown in Fig. 4(c). The resulting certificate has a lower computational cost than (a) and (b), but its energy bound is higher than (a) and the same as (b).

(ii) In an intermediate regime, as shown in Fig. 4(b), the best possible certificate is obtained combining the overlap of some of the largest possible constraints with the inclusion of smaller constraints. It has the same computational cost as (a) but it provides a higher energy bound.

(iii) Deep into the phase, as shown in Fig. 4(a), the best possible certificate is obtained by evenly distributing all the largest possible constraints throughout the system. A priori, we would expect this to be the optimal solution throughout the whole phase.

In the entangled phase, the two intermediate optimal relaxations (b) and (c) provide better bounds than the set of constraints (a) in Fig. 4, even with (c) being a significantly stronger relaxation than the others. In fact, (c) yields the exact same energy bound as (b) in the $1 \leq B/J < 2$ regime, as we show in Appendix C, and its therefore optimal due to its significantly lower computational cost. Similarly, in the entangled phase, they all yield the same bound and thus the optimal is the simplest one, (d). This simple scenario shows that evaluating the quality of a relaxation beforehand is not a trivial task and it becomes even less straightforward when considering arbitrary Hamiltonians and computational budgets. Additionally, the given budget may also allow the allocation of a few 4-body RDMs, which we take into account in the optimization. However, the agent finds that it is better to

combine several 3-body and 2-body RDMs rather than using a limited amount of 4-body ones.

Already in such a simple scenario, the agent is able to find a rich set of intermediate solutions, which may, at first glance, seem counterintuitive. The solutions are, nevertheless, closely related to the actual entanglement structure of the ground state of the system [108]. This shows that the agent is able to capture physical properties of the system, even when various possible solutions are very close in terms of cost and quality.

In Fig. 4 we show the solution of a small system of $n = 6$ sites for illustrative purposes. In larger systems, we observe that the same optimal patterns remain consistent, suggesting that the qualitative solutions obtained in small systems can be used for larger ones with similar properties. In Appendix C, we provide further details about the quality of the obtained certificates throughout the phase space and show that the optimal sets of constraints do remain optimal across different system sizes. Furthermore, in Appendix D we compare the resulting energy bounds with other state-of-the-art techniques to lower bound quantum Hamiltonians. We show that the proposed SdP formulation [Eq. (8)] provides a generic approach which includes some of the existing methods as particular cases and, therefore, achieving equal or better results by construction. Moreover, the main principles of other works, such as enforcing symmetries [54], can also be incorporated into our framework introducing by additional constraints. Nonetheless, the presented approach does not rely on such symmetric constraint, hence allowing us to tackle Hamiltonians of arbitrary properties.

As a final remark, notice that the ground state of the unentangled phase is a product state, meaning that the exact solution lies within the budget with which the agent is provided. In contrast, in the entangled one, the ground state can only be exactly described by its full density matrix, meaning that the exact solution falls outside of the computational budget. With the framework we here present, when the agent is far from using the whole budget, it may be seen as a strong indication that the provided result is exact. In Appendix E, we discuss some of these particular cases of interest.

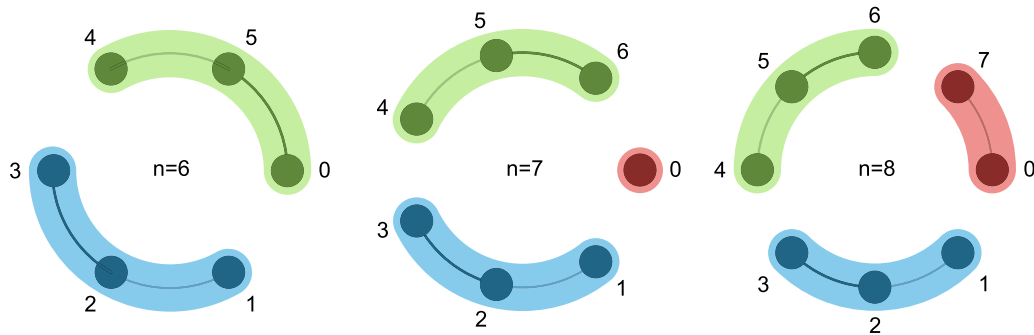


FIG. 5. Schematic representation of the inhomogeneous XX Hamiltonian (11) with its optimal set of constraints for the cases of $n = 6, 7, 8$ spins. The external magnetic field is fixed, $B_i = 1 \forall i$, and the interaction strength between spins, $J_i = i \bmod 3$, is represented by the intensity of the black line: solid ($J_i = 2$), transparent ($J_i = 1$) and no line ($J_i = 0$). The support of the RDMs that constitute the optimal relaxation are depicted in different colours, as in Fig. 4

B. Comparison with other optimization methods

As we briefly mention at the end of Sec. IV, the proposed framework allows for the straightforward application of several optimization algorithms besides RL. In this section, we evaluate the quality of the RL results in comparison to two informative points of reference: breadth first search (BFS) [115] and Monte Carlo (MC) optimization [116].

For the comparison, we consider an inhomogeneous version of the XX Heisenberg model from Eq. (11) in which we keep a constant magnetic field $B_i = 1 \forall i$ and tune the interaction strength $J_i = i \bmod 3$. This provides us with isolated groups of three interacting sites. Note that, depending on the system size, there may be exclusively triplets, triplets and an isolated site or triplets and a pair, as we show in Fig. 5 for the cases of $n = 6, 7$, and 8 , respectively.

Such model allows us to determine the optimal set of constraints beforehand while posing one of the hardest optimization instances, as it is a unique point in the constraint space. This way, we can compare the performance of the different algorithms with respect to the optimal solution. As performance metric, we compute the rewards, as in Eq. (9), with full knowledge of $\beta_{\max}, \beta_{\min}, p_{\text{best}}, p_{\text{worst}}$. This provides a measure of distance/similarity to the optimal configuration, obtaining reward 1 when reaching it.

Note that the algorithms have different ways to explore the state space. Hence, in order to compare them in the fairest way, we do not take into account repeated visits to the states. Contrary to the BFS, both the RL and the MC agents can go back and forth revisiting the same states. Given that the main computational cost comes from solving the SdPs, we keep a memory of the solutions obtained during the exploration, so that revisiting a state implies a negligible computational cost.

Consequently, we evaluate the overall performance tracking the best obtained reward for every new visited state. In Fig. 6, we depict the amount of new states visited by fifty agents before they reach a reward of 0.95 on average. We repeat the process for several system sizes, with which the constraint space increases exponentially. We tune the hyperparameters for the RL and MC optimizations at a system size of $n = 10$ and we keep them throughout the whole process (see Appendix F).

First, we compare the agent performance with a budget that allows the agents to allocate half of all the available 3-body constraints. We show the results in Fig. 6(a). For small systems, there are no substantial differences in performance, given that the state space is reduced. Already at $n = 11$, the BFS is not able to find the optimal bound within a reasonable time. While the MC optimization provides better results for small systems, it is outperformed by the RL agent at $n = 16$. We hypothesize that, at this size, the overhead of learning is overcome by the increasing complexity of the state space.

In order to test this hypothesis, we conduct the same experiment with a larger computational budget that allows the agents to allocate all the 3-body constraints. With this, for the same system sizes, the agents encounter significantly larger constraint spaces (see Appendix F 3). We show the results in Fig. 6(b). In this case, the differences between the MC and RL optimizations are relatively smaller for smaller systems and the RL agents outperform the MC optimization earlier on. This means that, for large state spaces, the learning cost involved in the RL optimization pays off, making it better than following a simple MC heuristic. In addition, unlike the RL, the MC shows a strong dependency on a proper hyperparametrization, e.g., choosing an appropriate inverse temperature, provided that, as soon as the parameters are not optimised for the specific problem, the performance is dramatically affected. Proper parameter tuning is, in itself, a computationally costly task, given the constraint-space size. The RL scheme, being quite resilient to its hyperparametrization, provides a significant advantage in this sense, allowing us to tune it in reduced systems.

C. Analysis with transfer learning across the phase space

An interesting feature of the proposed framework is that none of its parts require prior information about the actual problem. This suggests the possibility of exploring a given constraint optimization problem and its underlying system in a completely autonomous way. One way to take advantage of this feature is by performing transfer learning (TL) [117]. In order to do so, we start by training an agent to solve a system under the action of a Hamiltonian. Then, we leverage

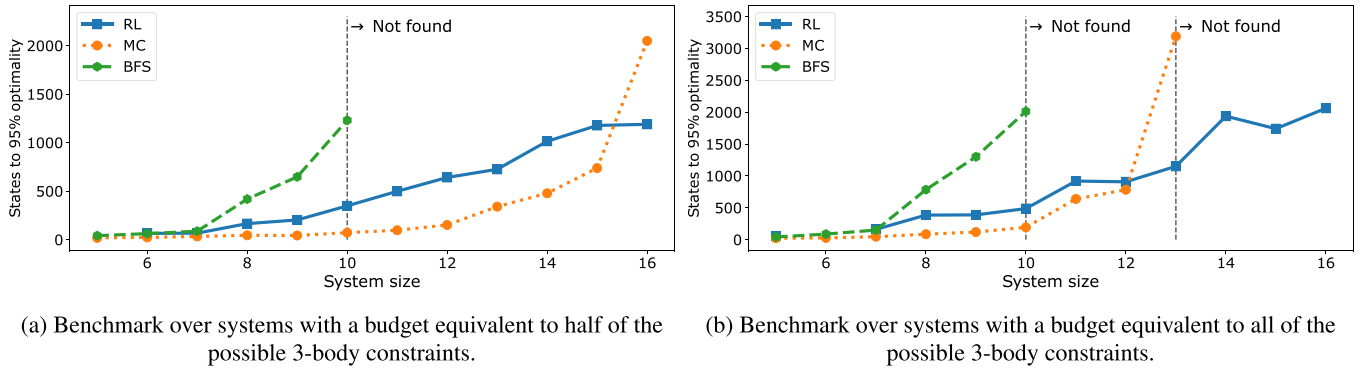


FIG. 6. Benchmark of the performance of the three optimization algorithms: Breadth first search (BFS), Monte Carlo (MC), and deep reinforcement learning (RL). The algorithms are evaluated in two scenarios: allowing up to (a) half of all the 3-body constraints and (b) all the 3-body constraints. The dashed vertical lines indicate the system size beyond which the overlapping algorithm was unable to find the optimal state in less than 4000 visited states.

the experience obtained by the agent in the initial task using it as the initial condition to solve a new problem with a similar Hamiltonian.

For this task, we consider an homogeneous version of the Heisenberg XX model [Eq. (11)]. As we have seen, while this Hamiltonian presents a unique quantum phase transition at $B/J = 2$, it has four different optimal relaxations across the phase diagram (see Sec. V A). We train an agent to find the optimal constraints deep in one phase, with $B/J = 5$. Then, we use the resulting trained agent to find the optimal relaxations for the rest of the phase space. In Fig. 7 we show the ratio between the time it takes the algorithm to converge with TL t_{TL} and the time it takes with a cold start t_0 , i.e., starting from scratch. Hence, with $t_{TL}/t_0 < 1$ there is favorable TL and with $t_{TL}/t_0 > 1$ there is negative transfer. We obtain the convergence time averaging the training results of fifty independent agents, shown on the right panel of Fig. 7 (see also [118]).

We observe that TL in the same phase is significantly favorable. Indeed, for this particular problem, the optimal set of constraints is the same across the whole phase, including the critical point [cases (d) and (c) in Fig. 7, respectively]. When applied across phases, the advantage of TL dimin-

ishes sharply. Close to the phase transition [case (b)], there appears a local minimum in which some agents get stuck and, under the given conditions, it takes them hundreds of training episodes to correct it. In this regime, the TL still provides an advantage regarding convergence, although it does not help avoiding the sub-optimal relaxation. Deep into the opposite phase [case (a)], even though TL barely affects the performance, as $t_{TL}/t_0 \simeq 1$, it has a slightly negative impact.

The vertical lines of Fig. 7 show the phase transition (solid) and the intermediate points in which the optimal set of constraints changes (dashed). The results suggest that losing the convergence advantage from TL can be indicative of changes in the ground state of the system, such as phase transitions. Hence, this approach can be used to infer properties of the physical system in a completely autonomous way by exploiting the failure of the method, such as in [119,120].

VI. ENTANGLEMENT WITNESSES FROM THE HEISENBERG XX MODEL

Here, we show how to apply the presented framework in the context of entanglement detection [121]. In particular,

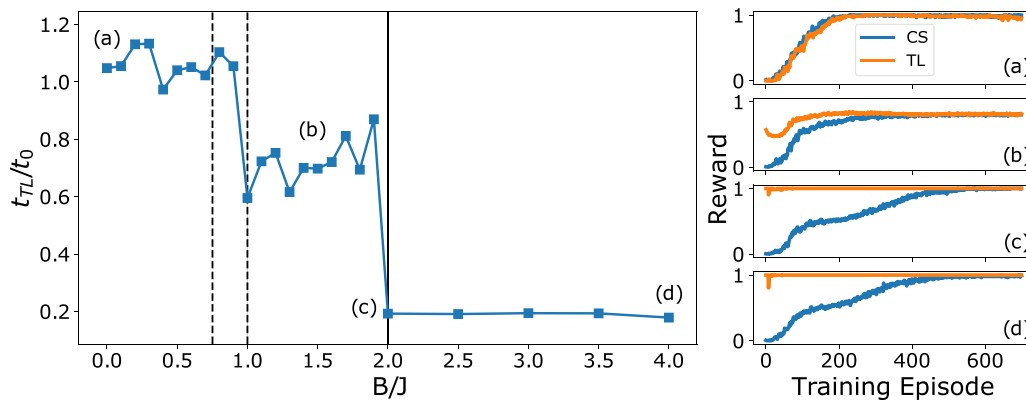
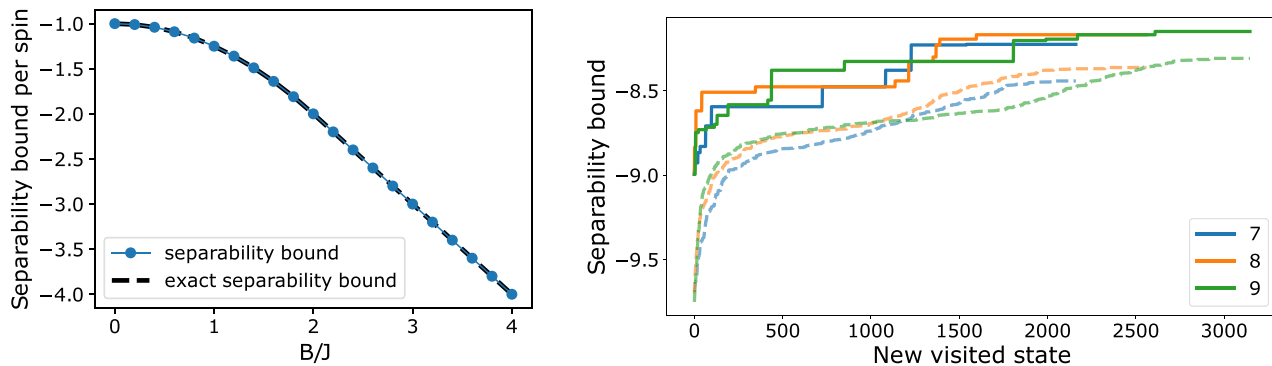


FIG. 7. Transfer learning results for an ensemble of fifty independent agents. (Left) Convergence time ratio between transfer learning and cold start as function of the parameter B/J [recall Eq. (11)]. The pretrained models are taken from $B/J = 5$ and they are used as starting point in the optimization for different values of the parameter. The vertical lines indicate qualitative changes in the optimal solution, with the solid line corresponding to the phase transition. (Right) Average reward obtained at the final state of an evaluation episode, performed after each training episode, with transfer learning (TL) and cold start (CS).



(a) Resulting bounds for the homogeneous Hamiltonian in 1D compared to the exact ones across the phase space.

(b) Bounds obtained by fifty independent agents with a random Hamiltonian and three different budgets: allowing 7, 8 and 9 3-body RDMs. The solid lines show the best obtained separability bound among all agents and the dashed lines show their mean.

FIG. 8. Separability bounds obtained for the Heisenberg XX model.

we address the task of building energy-based entanglement witnesses from local Hamiltonians [122,123], a paradigmatic problem in quantum information processing. We illustrate how our constraint optimization framework benefits from being agnostic to the problem it is solving, as we only need to adapt the black box module that handles the SdP optimization from Eq. (8). Once we have that, we can readily apply the entire pipeline to the new problem.

A. Energy-based entanglement witnesses from local Hamiltonians

Entanglement is a fundamental property of quantum mechanics and it plays a key role in quantum information processing applications. However, characterizing or detecting entanglement in experimental applications can be hard due to the limited available information about the quantum state. A way to detect entanglement is through entanglement witnesses, which are observables whose expectation value can certify whether the state measured is entangled.

We can construct entanglement witnesses by choosing the observable to be the Hamiltonian itself. Let $\Delta E = \langle H \rangle - E_{\text{sep}}$ be our witness, where $\langle H \rangle$ is the expected energy of our state

and E_{sep} is the minimum energy of the Hamiltonian in the set of separable states. This way, if $\Delta E < 0$, the quantum state lies outside of the separable set and, thus, it is entangled. Note, however, that if $\Delta E \geq 0$, the witness does not decide.

To find E_{sep} , we need to solve a global optimization task of $\langle H \rangle$ over the set of separable states. The search can be restricted to pure product states via a convex roof argument, although that does not simplify the complexity of the optimization. Therefore, we need to enforce that the global quantum state ρ is fully separable in Eq. (8). Nonetheless, even though the RDMs of a fully separable ρ are also separable, deciding membership to the set of separable quantum states is NP hard [30], even in simpler instances [124–126].

We can relax this problem by considering the set of states that are positive under partial transposition (PPT), which contains the set of separable states. These states are easy to characterize, as membership in the PPT set can be checked efficiently. However, they include some entangled states, thus being a relaxation of the set of separable states. Hence, we can modify Eq. (8), including the PPT constraints, to yield a lower bound on the separability bound of a local Hamiltonian H ,

$$\begin{aligned}
 \beta_C^{\text{full-sep}} := & \min_{\rho_S} \sum_i \langle H_i \rangle \\
 \text{s.t.} & \rho_S \geq 0 & \forall S \in C \\
 & \rho_S^{\Gamma_A} \geq 0 & \forall A \subseteq S \\
 & \text{Tr}[\rho_S] = 1 \\
 & \text{Tr}_{R^c}[\rho_S] = \text{Tr}_{R^c}[\rho_{S'}] & \forall R \subseteq S \cap S', \quad S, S' \in C,
 \end{aligned} \tag{12}$$

where the superscript Γ_A indicates that the partial transposition ($\mathbb{1}_{A^c} \otimes T_A$) is applied to the elements of $S = A \cup A^c$. Note that this is a linear operation since it simply permutes elements of ρ_S . Thus, any quantum state satisfying $\text{Tr}[\rho H] < \beta_C^{\text{full-sep}}$ must contain some entanglement.

We can further tighten the optimization in Eq. (12) in several directions. For instance, we can consider symmetric

extensions [26] to improve the approximation of the PPT set to the separable set. In some cases, we can ask that the bound detects a higher degree of entanglement, yielding a k -producibility bound. In this direction, there have been proposed device-independent witnesses of entanglement depth [127,128], based on relaxations of the quantum marginal problem via an SdP. In these cases, we can tighten the

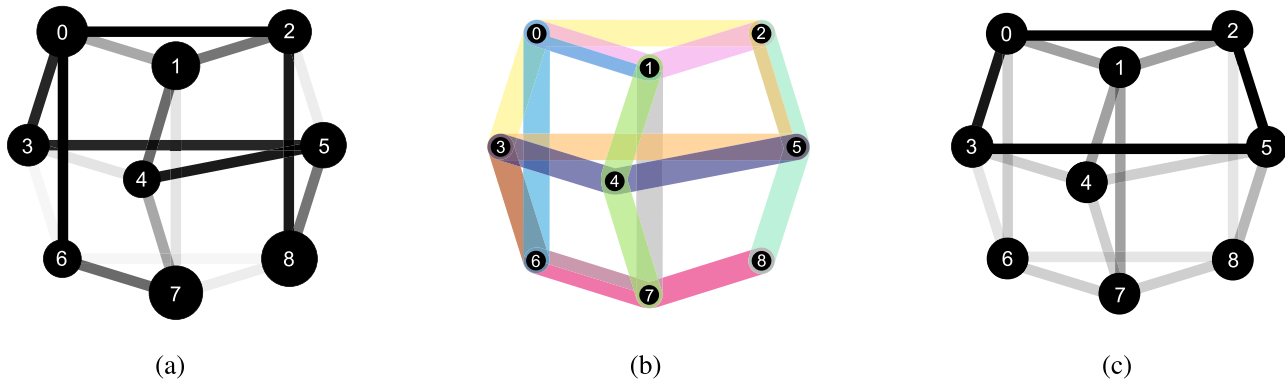


FIG. 9. Pictorial representation of the results obtained with a random Hamiltonian. (a) Representation of the random Hamiltonian parameters. The line transparency and the circle radius respectively indicate the relative values of J_i and B_i . (b) Representation of the best constraints found by the agents. Each color represent the support of an RDM considered in the relaxation, as in Fig. 4. (c) Histogram of the terms contained in the best constraints of the top ten performing agents. The line transparency represents the term frequency normalized by the maximum.

relaxation by imposing compatibility with larger quantum states.

B. Optimal separability bounds

We start by considering an homogeneous version of the Heisenberg XX model in 1D, as in Sec. V A. In this case, there is an analytical expression for the separability bound, which corresponds to the mean-field ground-state energy for negative J in cubic lattices [129]. We use this case as reference to validate our results. With the RL method, we find that we can recover the exact separability bound using exclusively 2-body RDMs, as we show in Fig. 8(a). Notice that the separability bound in the unentangled phase ($B/J \geq 2$) coincides with the exact ground-state energy that we show in Appendix C and Appendix D, proving that the ground state is, indeed, separable.

Then, we consider an heterogeneous version of the Hamiltonian in a graph with random parameters, $J_i, B_i \sim \text{uniform}[0, 1)$. In Fig. 9(a), we show a schematic representation of the system. Unlike in the homogeneous case, here, there are no apparent properties of the system that we can exploit to obtain an expression for the separability bound. Furthermore, we cannot even devise a strategy to build an efficient relaxation, as we do in Sec. V B.

We search for the optimal relaxation with fifty independent RL agents with three different budgets: allowing seven, eight, and nine 3-body constraints. We show the results in Fig. 8(b), where we see that we reach a higher bounds with every additional 3-body RDM. However, as the constraint space grows, it takes longer to reach the best configuration.

We depict the best obtained relaxation in Fig. 9(b), where we can see that the agent has significantly favoured some regions in detriment of others, e.g., using a 2-body RDM in sites 3 and 7 despite having enough budget to span the whole system with 3-body RDMs. To see this more clearly, in Fig. 9(c), we represent the frequency with which each connection in the graph appears in the best constraints found by the top ten performing agents. This result highlights a few clear key elements to obtain better bounds, even though some

of them may not seem relevant a priori given the Hamiltonian, such as the $\{2, 5\}$ term, which has a weak interaction with strong fields on both sites. Looking further into these agents we found that the 3-body terms $\{0, 1, 2\}$, $\{2, 5, 8\}$, and $\{3, 4, 5\}$ are significantly overrepresented with respect to the rest, as they span the regions with the strongest interactions in the system.

VII. APPLICATION TO OTHER PROBLEMS

The framework we presented here applies to the meta-problem of obtaining the best certificates given a computational budget by finding the most suitable convex relaxation. Although we have focused on two applications centered in lower-bounding energies of local Hamiltonians, our methodology can be directly applied to many other tasks. The only requirement is to adapt the black box routine from Eq. (8) to the new tasks, as we do in Sec. VI A, and appropriately map the constraint space to the new problem. Once it is done, its implementation to the new tasks is straightforward, provided that the optimization framework is entirely agnostic to the actual problem.

Convex sets arise naturally in quantum information in many flavors. An efficient way to characterize them is through linear witnesses. Among those, witnesses that can be easily measured are clearly preferred. This property means, in practice, that they consist of an $O(\text{poly}(n))$ number of terms. An important subclass of them is that in which these terms are local; i.e., acting on $O(1)$ parties at most. In Appendix G we thoroughly discuss how to perform the SdP formalization of some relevant problems in quantum information. In Appendix G 1, we discuss how our approach can be used to optimize outer approximations to the set of quantum correlations, in Appendix G 2, we consider the more general problem of finding better sum-of-squares representations of multivariate polynomials and, in Appendix G 3, we discuss how our method can be applied in problems that are amenable to linear programming, such as finding outer approximations to projections of the set of correlations that satisfy the no-signalling principle.

VIII. CONCLUSIONS

In this paper, we have introduced a novel approach to construct optimal relaxations to obtain certificates of quantum many-body properties. We have proposed a machine learning approach, based on deep reinforcement learning, to find such certificates given a finite computational budget. Throughout this paper, we have showcased its properties in the context of approximating the ground state energy and the separability bounds of quantum local Hamiltonians.

With the proposed framework, an RL agent is able to find the best possible certificate with the lowest computational cost within the computational budget. We have studied the validity of the method in the well-known Heisenberg XX model, showing that the agent is able to correctly characterize the ground state across the phase diagram without any kind of information about the physical system at hand. Indeed, we have shown how the certificates found by the agent change according to the ground state, as the structure of the constraints that suffice for a good approximation correlates with the system's phase and the entanglement properties of the ground state. Nonetheless, unravelling their precise relation is a matter deserving future investigation.

Already in small systems, we observe that the agent can capture the complexity of the system of study and go beyond more trivial relaxations, even when these are close in terms of the objective function. We have also shown that the agent is able to solve the opposite case, in which simpler certificates provide better bounds than more complex ones. Besides, we have shown that the qualitative relaxations obtained in reduced systems can be generalized to larger ones, as these remain consistent for any size. Hence, the constraint optimization can be performed in a reduced version of the original problem in order to minimize the computational workload.

Additionally, we have shown that the reinforcement learning approach handles large optimization spaces rather successfully, strongly outperforming other classical optimization algorithms. Furthermore, we have shown how to take advantage of transfer learning, positively impacting scalability. Moreover, we have characterized its behavior to find that it may be indicative of variations in the nature of the ground state of the system of study, some of which due to phase transitions. Hence, constituting an autonomous method to explore the system's phase diagram.

Finally, we have applied our framework in the context of entanglement detection. We have shown that we can efficiently recover the analytical solutions for simple cases and, as a final result, we apply our method to the case of a random Hamiltonian, to which there is no easy solution. Then, we provide the tools to generalize the framework to other common tasks such as optimizing outer approximations to the set of quantum correlations. Actually, the presented framework can be readily extended to other tasks in quantum information that are based on finding good outer approximations of convex sets that are hard to describe.

As future work, it remains open the question of which properties of the Hamiltonian have led to better bounds with simpler certificates. Furthermore, transfer learning can be used to analyze common patterns between different Hamiltonians. Besides, the architecture of the reinforcement learning

agent can be adapted to allow for the transfer learning between problems of different sizes. As an additional step, it would be interesting to study how introducing explicit information about the Hamiltonian may affect the optimization process. For instance, whether a RL agent can help in designing better adiabatic schedules [130] or whether better certificates can be built by combining RL following an adiabatic path.

IX. CODE AVAILABILITY

We provide the source code for the method proposed in this paper as a Python library in Ref. [131], with tools and extended explanations to reproduce the presented results and to use the method in various scenarios beyond the ones we consider throughout this paper.

ACKNOWLEDGMENTS

We thank A. Acín, F. Alet, F. Baccari, M. Lubasch, and N. Pancotti for enlightening discussions. We acknowledge the contribution of Aina Guirao to the design of the figures. B.R., G.M.-G., and M.L. acknowledge support from ERC AdG NOQIA, Agencia Estatal de Investigación (PGC2018-097027-B-I00/10.13039/501100011033, CEX2019-000910-S/10.13039/501100011033, Plan National FIDEUA PID2019-106901GB-I00, FPI, QUANTERA MAQS PCI2019-111828-2, QUANTERA DYNAMITE PCI2022-132919, Proyectos de I+D+I “Retos Colaboración” QUSPIN RTC2019-007196-7), MICIIN with funding from European Union NextGenerationEU(PRTR-C17.I1), and by Generalitat de Catalunya, Fundació Privada Cellex, Fundació Mir-Puig, Generalitat de Catalunya (European Social Fund FEDER and CERCA program, AGAUR Grant No. 2021 SGR 01452, QuantumCAT U16-011424, co-funded by ERDF Operational Program of Catalonia 2014-2020), Barcelona Supercomputing Center MareNostrum (FI-2022-1-0042), EU Horizon 2020 FET-OPEN OPTOLogic (Grant No. 899794), EU Horizon Europe Program (Grant Agreement 101080086 — NeQST), National Science Centre, Poland (Symfonia Grant No. 2016/20/W/ST4/00314), ICFO Internal “QuantumGaudi” project; European Union's Horizon 2020 research and innovation program under the Marie-Skłodowska-Curie Grant Agreement No. 101029393 (STREDCH) and No. 847648 (“La Caixa” Junior Leaders fellowships ID100010434: LCF/BQ/PI19/11690013, LCF/BQ/PI20/11760031, LCF/BQ/PR20/11770012, LCF/BQ/PR21/11840013). G.M.-G. acknowledges support from the Austrian Science Fund (FWF) through SFB BeyondC F7102 and from Fundació Obra Social la Caixa (LCF-ICFO grant). J.T. thanks the Alexander von Humboldt foundation for support. This project has received funding from the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - Project No. 414325145 in the framework of the Austrian Science Fund (FWF): SFB F7104. This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement No. 899354. This work has received support from the European Union's Horizon Europe program through the ERC StG FINE-TEA-SQUAD (Grant No. 101040729). The authors also acknowledge support from

the Quantum Delta program. This publication is part of the ‘Quantum Inspire – the Dutch Quantum Computer in the Cloud’ project (with Project No. NWA.1292.19.194) of the NWA research program ‘Research on Routes by Consortia (ORC)’, which is funded by the Netherlands Organization for Scientific Research (NWO). This work was supported by the Dutch Research Council (NWO/OCW), as part of the Quantum Software Consortium programme (Project No. 024.003.037). Views and opinions expressed in this work are, however, those of the author(s) only and do not necessarily reflect those of the European Union, European Climate, Infrastructure and Environment Executive Agency (CINEA), nor any other granting authority. Neither the European Union nor any granting authority can be held responsible for them.

APPENDIX A: SDP-BASED CERTIFICATES

In this section we discuss the details on how a proof is obtained via a semidefinite program (SdP). To this end, let us recall the (primal) form of a SdP in canonical form

$$\begin{aligned} \min_X \quad & \langle C, X \rangle \\ \text{s.t.} \quad & \langle A_i, X \rangle = b_i \\ & X \geq 0. \end{aligned} \tag{A1}$$

To each primal SdP one can associate a dual SdP, which is the following optimization problem:

$$\max_{\mathbf{y}} \quad C - \sum_i y_i A_i \geq 0. \tag{A2}$$

$$\begin{aligned} \beta_C := \min_{\mathcal{C}} \quad & \sum_i \langle H_i \rangle \\ \text{s.t.} \quad & \rho_S \geq 0 \\ & \text{Tr}[\rho_S] = 1 \\ & \text{Tr}_R[\rho_S] = \text{Tr}_R[\rho_{S'}], \quad \text{where } R = S \cap S', \quad \forall S, S' \in \mathcal{C}. \end{aligned} \tag{B1}$$

In other words, Eq. (B1) enforces that reduced states are equal in every pairwise intersection of constraints. Both Eq. (8) and Eq. (B1) yield valid certificates, but Eq. (B1) might not implement all the compatibility conditions that one would naively expect, in some pathological cases. Here we discuss an example (see Fig. 10). Consider a system of 4 qubits in a 1 – D geometry on a ring, *i.e.*, with periodic boundary conditions. Consider furthermore that our set of constraints is $C = \{\{0, 1, 2\}, \{1, 2, 3\}, \{2, 3, 0\}, \{3, 0, 1\}\}$. Equation (B1) would require that the SdP takes into account the variables $\rho_{0,1}, \rho_{0,2}, \rho_{0,3}, \rho_{1,2}, \rho_{1,3},$ and $\rho_{2,3}$. These come from the 6 ways to choose 2 elements from a 4-element set, like C . Note,

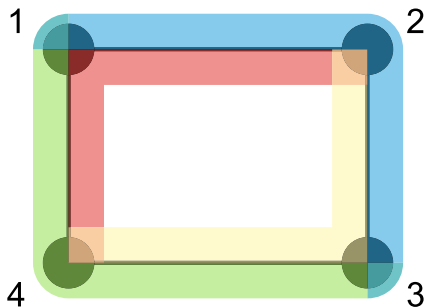


FIG. 10. A pathological case for Eq. (B1).

Although the primal SdP Eq. (A1) and the dual SdP Eq. (A2) are different optimization problems any two X and \mathbf{y} that satisfy their respective problem constraints obey the weak duality relation

$$\langle C, X \rangle \geq \mathbf{y}'\mathbf{b}. \tag{A3}$$

This means that a feasible solution X of Eq. (A1) upper bounds the value of the objective function of Eq. (A2) (in particular, its maximum value). Conversely, any feasible solution \mathbf{y} of Eq. (A2) lower bounds the value of the objective function of Eq. (A1) (in particular, its minimum value). Therefore, one can construct a mathematical proof for a bound $E_0 \geq \beta$ by transforming the SdP into primal canonical form Eq. (A1), constructing its dual problem and finding a dual feasible solution of the latter Eq. (A2). In practice, this transformation is taken care of automatically with SdP parsers such as cvx [132,133] or yalmip [134] and an SdP solver (e.g., [98,135]) numerically finds the values of the optimal X and \mathbf{y} of both problems. Furthermore, Eq. (A3) typically becomes tight at optimality of both Eq. (A1) and Eq. (A2) under reasonable regularity conditions, such as strict feasibility [93].

APPENDIX B: PATHOLOGICAL CASES

It might appear that a more natural way to define Eq. (8) would be as follows:

however, that the SdP does not enforce conditions that one would naturally expect, such as

$$\rho_3 \equiv \text{Tr}_2[\rho_{23}] = \text{Tr}_0[\rho_{03}] = \text{Tr}_1[\rho_{13}].$$

In other words, the two-body reduced density matrices stemming from Eq. (B1) do not need, *a priori*, to have compatible supports in their intersections. This caveat is resolved in the formulation of Eq. (8). In practice, however, pathological cases such as the one depicted in Fig. 10 are quite rare. For instance, the same scheme with a three-body constraint centered at each site, but for a number of parties larger than 4 would automatically generate all the single-body terms.

APPENDIX C: OPTIMAL CONSTRAINTS ACROSS SYSTEM SIZES

In Sec. V A we presented the result of applying the proposed method to the XX Heisenberg model Eq. (11) with a computational budget that allowed the agent to allocate up to half of the 3-body constraints. Here, we explore in further detail how the different sets depicted in Fig. 4 vary in energy throughout the phase space. Moreover, we show that even for bigger system sizes the same qualitative relaxations remain optimal across the different sizes. In particular, we will study systems of sizes $n = 6, 12, 24, 36$.

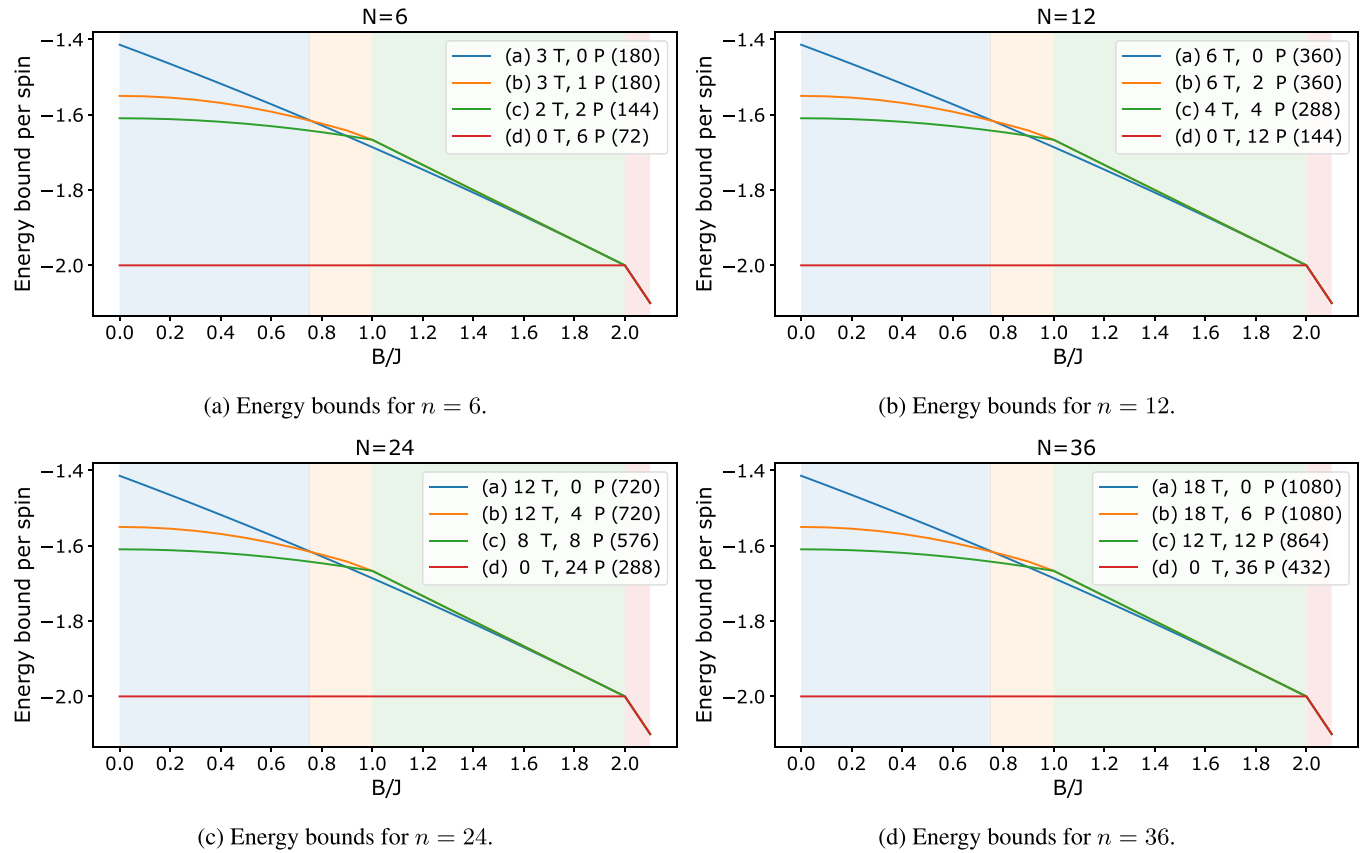


FIG. 11. Energy bounds obtained for system sizes of $n = 6, 12, 24, 36$ with the sets of constraints that are optimal in some region of the phase space. The labels (a), (b), (c), (d) refer to the qualitative relaxations of Fig. 4, followed by the number of triplets (T) and pairs (P) constituting the RDMs and, in brackets, the cost associated to solving the associated SdP in terms of the number of free variables in the SdP. The shaded background indicates the color of the set of constraints that is optimal within the range.

Let us first address what we mean by qualitative relaxations and how these are generalized to different sizes. The optimal sets of constraints shown in Fig. 4 can be seen as patterns of reduced density matrices (RDMs) that span the system and can therefore be reproduced at any size. This way, sets of constraints made out of the same RDM pattern may constitute the same qualitative relaxation. Let us describe these patterns and provide some examples:

(a) Span the system with evenly distributed 3-body RDMs.

- (i) $n = 6$: $C = \{\{0, 1, 2\}, \{2, 3, 4\}, \{4, 5, 0\}\},$
 $C = \{\{1, 2, 3\}, \{3, 4, 5\}, \{5, 0, 1\}\},$
- (ii) $n = 12$: $C = \{\{0, 1, 2\}, \{2, 3, 4\}, \{4, 5, 6\},$
 $\{6, 7, 8\}, \{8, 9, 10\}, \{10, 11, 0\}\}.$

(b) Span the system with 3-body RDMs, including an additional 2-body RDM by overlapping two of the 3-body ones. This is the least straightforward pattern to generalize, provided that it can either be interpreted as having only one extra 2-body RDM, or including some additional 2-body RDMs every few sites. We have found that, for the considered system sizes, the optimal set of constraints is found by including these 2-body RDMs every 6 sites, i.e., spanning the system by repetition of the 6-body pattern.

- (i) $n = 6$: $C = \{\{0, 1, 2\}, \{2, 3, 4\}, \{3, 4, 5\},$
 $\{5, 0\}\}, C = \{\{0, 1, 2\}, \{2, 3\}, \{3, 4, 5\}, \{4, 5, 0\}\},$

- (ii) $n = 12$: $C = \{\{0, 1, 2\}, \{2, 3, 4\}, \{3, 4, 5\},$
 $\{5, 6\}, \{6, 7, 8\}, \{8, 9, 10\}, \{9, 10, 11\}, \{11, 0\}\}.$
- (c) Span the system alternating 3-body and 2-body RDMs.
 - (i) $n = 6$: $C = \{\{0, 1, 2\}, \{2, 3\}, \{3, 4, 5\}, \{5, 0\}\},$
 $C = \{\{0, 1\}, \{1, 2, 3\}, \{3, 4\}, \{4, 5, 0\}\},$
 - (ii) $n = 12$: $C = \{\{0, 1, 2\}, \{2, 3\}, \{3, 4, 5\},$
 $\{5, 6\}, \{6, 7, 8\}, \{8, 9\}, \{9, 10, 11\}, \{11, 0\}\}.$

(d) Span the system with 2-body RDMs.

- (i) $n = 6$: $C = \{\{0, 1\}, \{1, 2\}, \{2, 3\}, \{3, 4\}, \{4, 5\},$
 $\{5, 0\}\}, C = \{\{1, 2\}, \{2, 3\}, \{3, 4\}, \{4, 5\}, \{5, 0\}\},$
- (ii) $n = 12$: $C = \{\{0, 1\}, \{1, 2\}, \{2, 3\}, \{3, 4\}, \{4, 5\},$
 $\{5, 6\}, \{6, 7\}, \{7, 8\}, \{8, 9\}, \{9, 10\}, \{10, 11\},$
 $\{11, 0\}\}.$

In Fig. 11 we show the energy bounds obtained by all the sets of constraints that, at some point along the phase diagram, are optimal. Indeed, we find that the optimal sets of constraints at different system sizes represent the same qualitative relaxations and the regimes under which these are optimal are all the same. This suggests that a reduced version of the original problem can be used in order to find the optimal set of constraints, significantly reducing the computational cost of the optimization.

Additionally, we see that the relaxation (c), in some regions of the phase space, provides the same energy bound as (b) and, even more, yields a better bound than (a), while being a much simpler certificate that involves 20% less SdP variables

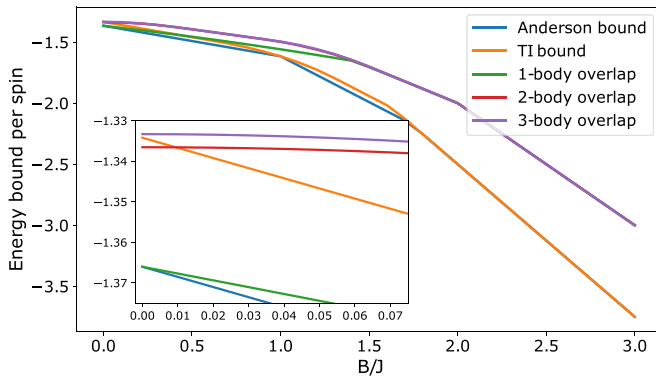


FIG. 12. Comparison between the Anderson bound [45], the method from [54] imposing TI symmetry and our SdP formulation Eq. (8) to lower bound the ground state energy of the homogeneous XX Heisenberg model Eq. (11). We consider 5-body RDMs for all methods and three instances of our method with increasing overlap between them. In the extreme case of $B/J = 0$, the Anderson bound is equivalent to considering 1-body overlaps between the RDMs in our method. On the other hand, the symmetry-based method yields a bound between what we obtain with a 2-body and a 3-body overlap. With increasing B/J , both methods are quickly outperformed by our approach, even with the least possible overlap.

than (a) and (b). Overall, the relative behavior of each set of constraints is the same across system sizes and they all converge to the same value at $B/J = 2$, where the phase transition happens.

APPENDIX D: LOWER BOUND METHODS

In this Appendix, we provide an overview of known techniques specific for our case of study: obtaining lower energy bounds to quantum Hamiltonians. Additionally, we compare the performance of some of them with our method based on semidefinite programming (SdP). Since the early 1950s, with the introduction of Anderson’s method [45], several other methods to lower bound the ground state energy of quantum many-body Hamiltonians have been developed. Anderson’s method corresponds to the triangle inequality Eq. (3): one splits the local Hamiltonian into a sum of operators that act on as many qubits as computational resources allow for exact diagonalization. Then, one independently combines their ground-state energies without imposing any kind of compatibility amongst the different ground states. Below, we briefly survey some recent papers that improve upon this idea by imposing restrictions onto the feasible set, e.g., in the form of a symmetry [54], and compare their performance with our proposed SdP formulation Eq. (8). The comparison results are showcased in Fig. 12.

(i) Reference [46] proposes a slight generalization of Anderson’s method, based on the following eigenvalue inequality: if f_0 denotes the minimal eigenvalue of a linear operator f acting on a finite-dimensional Hilbert space, then $(f + g)_0 \geq f_0 + g_0$ for all f and g acting on a finite-dimensional Hilbert space. By applying this principle to a Hamiltonian H with translationally invariant symmetry, one obtains $H_0 \geq mh_0$, where H is expressed as a sum of m terms of the form h_0 acting on different qubits. The h_0 terms can be

viewed as fundamental cells of a tiling of a 1D or 2D lattice. Hence, this procedure is equivalent to applying Anderson’s method m times, with translational invariant (TI) symmetry as an essential ingredient. Moreover, it can only produce solutions with the same TI symmetry, by construction. In general, optimal solutions of symmetric problems may not possess the same symmetry as the original problem, but rather there will be an orbit of solutions according to that symmetry, as we have shown in Fig. 4 for a TI system.

(ii) Reference [50] proposes a method specifically developed for 1D quantum spin systems. It is based on the Golden-Thompson inequality $\text{Tr}[e^{A+B}] \leq \text{Tr}[e^A e^B]$, with which one divides the Hamiltonian H into two parts $H = A + B$, for which the matrix exponentials can be easily calculated. In order to fulfill such condition, both the decomposition in Eq. (6) from [50] and the following derivation rely on the matrix-product-state-like structure given by the 1D geometry. One may either group spins into larger-dimensional sites, analogously to taking larger reduced density matrices (RDMs) in our method, or group them differently in order to obtain lower minimal eigenvalues locally, as the agent would do by exploring the constraint space. Importantly, the contraction of tensors necessary for the method in [50] can only be done efficiently in 1D geometries or, more generally, in constant treewidth graphs. Contrarily, the method we propose in Eq. (8) is not bound to such restrictions.

(iii) Reference [136] proposes a moment method, developed in the same spirit to those of the NPA-hierarchy discussed in Appendix G 1. One starts by defining a family of Hermitian operators $\{O_k\}_{1 \leq k \leq L}$ and considers any linear combination of them $O_\alpha := \sum_k \alpha_k O_k$. Then, by construction, $\text{Tr}[O_\alpha^\dagger O_\alpha \rho] \geq 0$ for any density matrix ρ and for any $\alpha \in \mathbb{C}^L$, but the converse is not necessarily true. One can construct a moment matrix X with entries $X_{kl} := \langle O_k^\dagger O_l \rangle$, which is positive semidefinite by construction. Indeed, for any $\alpha \in \mathbb{C}^L$ we have that $\alpha^\dagger X \alpha = \text{Tr}[O_\alpha^\dagger O_\alpha \rho] \geq 0$. But that alone does not guarantee that ρ is a valid density matrix. Therefore, the following inequality holds:

$$\min_{\rho \geq 0, \text{Tr}[\rho]=1} \langle H, \rho \rangle \geq \min_{X \geq 0, \langle A_i, X \rangle = b_i} \langle H, \rho \rangle, \quad (\text{D1})$$

where the objective function in the right-hand side can be expressed as a linear combination of the terms X_{kl} if the set $\{O_k\}_k$ is sufficiently rich. Additionally, equality constraints between the X_{kl} terms can be imposed as part of a SdP of variable size, e.g., normalization. Here, the quality of the solution highly depends on identifying meaningful operators O_k (cf. Appendix G 1) and, as acknowledged by the authors in [136], *the choice of the set $\{O_k\}_k$ is not unique and the optimal choice is not obvious*, which further motivates our RL-based approach.

(iv) Reference [54] considers a general symmetry onto the system and its solution, which translates into a system of equations involving the RDMs used to compute the energy. Positivity is imposed by the reparametrization of the density matrices $\rho = \tau^\dagger \tau / \text{Tr}[\tau^\dagger \tau]$ and symmetry on the clusters’ RDM is imposed manually. Then one must carry on a numerical minimization on the resulting parameter space, which is nonlinear due to the squaring of the trace. Moreover, one must consider that the local minima found is a global

minimum in order to obtain a certificate. In our case, the certificate always follows from SdP’s weak duality (see Appendix A). The local minima problems, positivity constraints and symmetry conditions in the RDMs challenges in [54] could be easily circumvented via a SdP formulation in the RDMs, although one would still be looking for a solution that satisfies a given symmetry. Furthermore, symmetries can be incorporated in SdP programs in natural ways, usually through block-diagonalizations in the SdP variables induced by the symmetries [137–139].

As we have discussed, most existing methods focus on cases that intensively exploit properties such as TI symmetries and/or 1D geometries. Remarkably, it is not clear how to extend existing methods to cases that do not have these restrictions. The method proposed in Eq. (8) does not require either. In order to provide a benchmark against existing methods, we have focused on those that rely explicitly on the collection of RDMs [45,54] rather than on moment methods [136], for which we would need to change the agent’s black box in order to give a fair comparison. Despite our method could be further improved in efficiency by exploiting symmetries in the problem [139], we have kept the general case and used a symmetric problem for the benchmark.

We perform the comparison using the homogeneous version of the XX Heisenberg model Eq. (11), with $J_i = 1, B_i = B, \forall i$ and we vary B . We take 5-body RDMs in all methods and, in our approach, we control the compatibility constraints with the amount of overlap between RDMs (more overlap means more constraints). This way, we can choose from overlapping a single spin in consecutive RDMs up to four spins. The resulting bounds are shown in Fig. 12, where we observe that, in the extreme case of no external field $B = 0$, the Anderson bound is equivalent to a 1-body overlap relaxation and the symmetry-based method lies between the 2-body and 3-body overlap ones. As B/J increases, our method quickly outperforms the others, even with the simplest possible relaxation. Notice that we could still obtain an even higher bound with 4-body overlaps. In agreement with what we see in Fig. 11 from Appendix C, the bounds provided by our method converge to the same value at $B/J = 2$, corresponding to the phase transition, beyond which the result is the exact ground state energy. Similarly, the symmetry-based method and the Anderson bound eventually converge to the same energy values, but well below the exact energy.

These results show that our SdP formulation Eq. (8) can outperform the other approaches without making any assumption about the symmetries of the system nor any *ad hoc* parametrization, which makes it applicable to any problem. Naturally, we can decrease our computational cost by imposing symmetries that are known beforehand as additional constraints in the SdP besides the compatibility ones, which allows for tighter bounds within the given computational budget.

APPENDIX E: PARTICULAR CASES

Here we analyze some cases where the local Hamiltonian Eq. (1) enjoys desirable properties that make a certificate easier to obtain.

Frustration-free. If H is a frustration-free Hamiltonian, its lowest energy eigenstate coincides with a lowest energy state of each of the individual terms H_i . In other words, global ground states correspond to local ground states. In this case, let $|\psi\rangle$ be the ground state of H . It is also a ground state of every H_i and it defines a set of RDMs $\rho_i = \text{Tr}_{\text{supp}(H_i)^c} |\psi\rangle\langle\psi|$. Note that frustration-freeness guarantees that the contribution of each term equals its algebraic minimum $\text{Tr}[\rho_i H_i] = \min \sigma(H_i)$. Hence, the minimal set of constraints C_\emptyset [cf. Eq. (4)] already reproduces the ground-state energy. On the one hand, given a term H_i , there is no ρ_i that yields a smaller value than $\min \sigma(H_i)$. On the other hand, the set of RDMs that correspond to the actual ground state satisfy this condition. This implies that strengthening the constraints in Eq. (8) to any $C \supseteq C_\emptyset$ will be of no effect in increasing β_C .

A couple of comments are in order:

(i) Obtaining a set of constraints Ξ_C , which recovers an exact lower bound $\beta_C = E_0$ does not automatically imply that we can recover the ground state configuration, even if the problem is fully classical. For instance, even if H corresponds to a classical 3-SAT problem: H can be written in the computational basis as a sum of projectors Π_i that act nontrivially on 3 variables x_{i_1}, x_{i_2} and x_{i_3} . Since $\Pi_i \geq 0$ and there exists a satisfiable instance, we obtain $\beta_C = 0$ for any relaxation. By inspecting the values of the ρ_i that the SdP Eq. (8) outputs, it does not need to be the case that ρ_i is a rank-1 projector onto the solution state $|x_{i_1} x_{i_2} x_{i_3}\rangle$ and thus directly interpretable as part of the solution to 3-SAT.

(ii) Frustration-free Hamiltonians constitute an important class of models. The ground state of short-range, gapped Hamiltonians can be well approximated by that of a frustration-free Hamiltonian by increasing their locality to be $O(\log(n))$ [140]. Frustration-free Hamiltonians comprise notable models, both commuting and anticommuting: On the one hand, frustration-free, commuting models include the toric code [141,142], Levin-Wen models [143] and quantum error correcting codes [144]. Importantly, graph states [145] or, more generally, stabilizer states such as the cluster state [146] are included in this class. Graph states can be approximated as ground states of two-body Hamiltonians [147], although it has been shown for spin-1/2 that this approximation cannot be made exact (ground states of frustration-free 2-local qubit Hamiltonians are unentangled [148,149]), even if we drop the frustration-freeness condition [150,151]. On the other hand, frustration-free, noncommuting models include the Affleck-Kennedy-Lieb-Tasaki (AKLT) [152], Rokhsar-Kivelson models [153,154] and parent Hamiltonians that are defined from injective projected entangled-pair states (PEPS) [78,155–158]. Sufficient conditions on when a Hamiltonian must be frustration-free have been studied in [159].

Mutually commuting terms. If H is a sum of mutually commuting terms, its eigenstates correspond to eigenstates of each of the H_i . Note, however, that the order of the eigenenergies in H needs not correspond to the order of the eigenenergies in H_i . For instance, changing H_i to $-H_i$ reverses the order of the eigenstates, but leaves commutativity untouched. The simplest example of a commuting, nonfrustration-free Hamiltonian is to consider $H = \sum_{\langle i,j \rangle \in E} \sigma_z^{(i)} \otimes \sigma_z^{(j)}$, where E are the edges of a triangle. In this case, tightening the constraints

in Eq. (8) helps in better capturing the frustration in the model, thus improving β_C , as a larger number of sites is considered.

APPENDIX F: OPTIMIZATION DETAILS

In this section, we provide a brief description of the optimization methods with details about the specific parameters we have used to obtain the results shown in Fig. 6. The implementation details can also be found in the open repository in Ref. [131] hosting the source code with extended explanations. Furthermore, we show the scaling of the combinatorial space over which the optimizations are carried over.

1. Reinforcement learning parameters

The RL optimization has several hyper-parameters that dictate both the deep Q-network architecture [103] and the learning procedure of the RL agent. Given that the size of the state vectors depends on the size of the actual problem (see Sec. IV), the RL agent must be adapted to each system size. This way, we define many of the parameters as a function of the system size n , the state vector size s , and the number of possible actions $a = s + 1$, as the agent is allowed to remain in the same state.

The agent architecture has three fully connected hidden layers with a rectified linear unit (ReLU) activation function. The input layer has size s , the first layer has size $3s$, the second layer has size $2a$, the third layer has size $2a$, and the output layer has size a . This network is copied as a target network for double Q learning [104].

The learning procedure is structured in learning episodes in which the agent performs a trajectory through the state space, always starting from the same initial state, as described at the beginning of Sec. IV. Throughout the learning episodes, the agent gathers experience in the form of *State-Action-New State-Reward* tuples that are stored in a memory. At the end of each episode, the agent replays a batch of steps from the memory to learn.

We set an episode length of the order of the system size n , modified according to the computational budget. For a low budget, such as in Fig. 6(a) with half of the 3-body constraints available, the episode length can be slightly lower than n , e.g., $0.7n$. In contrast, for a high budget, as in Fig. 6(b) with all the 3-body constraints available, the episode length needs to be slightly higher than n , e.g., $1.2n$. This way, we guarantee that the agent has enough time to allocate all the possible constraints with some margin for errors. We set the batch size for the experience replay to 20 episodes and the agent starts learning once it has visited as many states as a fifth of the batch size. We use a learning rate of 5×10^{-3} and we update the target network every 5 episodes.

Finally, in order to enforce the agent to explore, we vary the value of ϵ in the ϵ – greedy policy throughout the training process. We start with $\epsilon_0 = 0.9$ and we make it decay exponentially after every training episode, labeled by e , such that $\epsilon_e = \max\{0.1, \delta^e \epsilon_0\}$ with $\delta \in (0, 1)$. In the benchmarking from Fig. 6, we have taken $\delta = 0.5$ for systems $n \leq 7$ and $\delta = 0.95$ for systems $n > 7$. In small systems, the constraint space is reduced and the exploration is not needed, while a proper exploration is critical in larger problems.

2. Baseline optimization methods

Here we describe the methods used to benchmark our results: breadth first search (BFS) and Monte Carlo (MC) optimization.

BFS does not have any hyperparameter. Starting from the initial state, it builds a queue of states to visit by recursively expanding each state. Expanding a state consists of appending, at the end of the queue, all the possible states that can be reached from it through valid actions. We have taken randomized orders in the state expansion and we do not consider states that have already been visited or that are already in the queue.

The MC optimization has only one hyperparameter, that is, the effective temperature T . The algorithm consists of proposing random valid actions to go from one state to another. Then, the movement is accepted or rejected depending on the reward associated to the old and new states, $R_{\text{old}}, R_{\text{new}}$, with acceptance probability $p(R_{\text{new}}, R_{\text{old}}) = \min\{1, e^{(R_{\text{new}} - R_{\text{old}})/T}\}$. We tune the effective temperature to obtain a 50% acceptance ratio in a long, well converged, optimization. The results from Fig. 6(a) are obtained with $T = 0.084$ and the results from Fig. 6(b) with $T = 0.097$.

3. State space scaling

As we mention in the main text, we deal with a vast combinatorial constraint space. Here, we study its scaling with respect to the system size n in the 1D case considered throughout Sec. V. In 1D, we have n constraints of each size and we mainly consider two budgets that either allow to allocate up to half of the 3-body constraints or all of them. As we explain in Sec. IV, the 1-body constraints are implicitly represented by the lack of active constraints in the state vector, so we do not need to account for them. Additionally, even though these budgets also allow the agent to allocate a few 4-body constraints, these add a few additional combinations of, at most, $O(n)$. Thus, we will not count them for the sake of simplicity.

Hence, we effectively consider n 2-body constraints and n 3-body constraints. For example, in the case of $n = 4$, we consider all the possible combinations of the constraint set $\{\{0, 1\}, \{1, 2\}, \{2, 3\}, \{0, 3\}, \{0, 1, 2\}, \{1, 2, 3\}, \{0, 2, 3\}, \{0, 1, 3\}\}$. The agent can activate any arbitrary number of 2-body constraints yielding a total of $\xi_2 = \sum_{j=0}^n \binom{n}{j}$ possible combinations among them. Similarly, we can count the possible combinations of 3-body constraints ξ_3 , although it depends on the budget

$$\xi_3 = \begin{cases} \sum_{k=0}^{n//2} \binom{n}{k} & \text{budget for half of the 3-body constraints} \\ \sum_{k=0}^n \binom{n}{k} & \text{budget for all of the 3-body constraints,} \end{cases} \quad (\text{F1})$$

where $n//2$ denotes the integer half, e.g., $7//2 = 3$.

Then, we must consider the possible combinations between 2-body and 3-body constraints. Depending on the overlap, every additional 3-body constraint can include between one or two 2-body ones. For instance $\{\{0, 1, 2\}, \{1, 2, 3\}\}$

contain $\{\{0, 1\}, \{1, 2\}, \{2, 3\}\}$, while $\{\{0, 1, 2\}, \{0, 2, 3\}\}$ contain $\{\{0, 1\}, \{1, 2\}, \{2, 3\}, \{0, 3\}\}$. Therefore, the amount of possible combinations of 2- and 3-body constraints strongly depends on the overlap between the 3-body ones. Hence, we can build an upper and lower bound to the size of the constraint space by considering that all states have maximum and minimum overlaps.

$$|C| = \begin{cases} \sum_{k=0}^{n/2} \binom{n}{k} \sum_{j=0}^{n-2k} \binom{n-2k}{j} & \text{budget for half of the 3-body constraints} \\ \sum_{k=0}^{n/2} \binom{n}{k} \sum_{j=0}^{n-2k} \binom{n-2k}{j} + \sum_{k=n//2+1}^n \binom{n}{k} & \text{budget for all of the 3-body constraints.} \end{cases} \quad (F2)$$

For every possible combination of 3-body constraints ξ_3 , we consider all the possible 2-body combinations ξ_{23} considering minimum overlap between the 3-body terms. The first term of the sum ($k = 0$) corresponds to ξ_2 . The additional term in the case of the larger budget accounts for the possible combinations of 3-body constraints that do not leave room to fit any 2-body term.

Similarly, in the case with maximum overlap, we have $\xi_{23} = \sum_{j=0}^{n-k-1} \binom{n-k-1}{j}$ possible 2-body combinations for $k > 0$ 3-body constraints. Thus, we can obtain an upper bound to the size of the constraint space,

$$|C| = \begin{cases} \sum_{j=0}^n \binom{n}{j} + \sum_{k=1}^{n/2} \binom{n}{k} \sum_{j=0}^{n-k-1} \binom{n-k-1}{j} & \text{budget for half of the 3-body constraints} \\ 1 + \sum_{j=0}^n \binom{n}{j} + \sum_{k=1}^{n-1} \binom{n}{k} \sum_{j=0}^{n-k-1} & \text{budget for all of the 3-body constraints.} \end{cases} \quad (F3)$$

The first sum in the first line and second lines correspond to x_{i_2} for the case of $k = 0$ 3-body constraints. In the second line, the one is the state with all the possible constraints active $k = n$.

In Sec. VB, we consider a problem with a unique optimal relaxation. The agents need to find them in state spaces of the order of 10^2 for $n = 5$ up to 10^7 for $n = 16$, posing an increasingly harder problem. In Sec. VIB, the constraint space we consider in the 2D case is of the order of up to 10^{11} for the highest budget.

APPENDIX G: GENERALIZATIONS

In Sec. VII, we briefly mention the generalization of the presented method to other common tasks in the field of quantum information processing, beyond the two examples of lower bounding the ground-state energy and the separability bound of local Hamiltonians. We present, here, a nonexhaustive set of examples explicitly showing how to implement the black box routine from Eq. (8) to such tasks. This allows the straightforward implementation of the RL framework, provided that it is entirely agnostic to the actual problem.

1. Outer approximations to the set of quantum correlations

The set of correlations that are produced by quantum mechanics is also a convex set [160]. A whole program aiming at its characterization has obtained several operationally-motivated characterizations of it [35–42]. Systematic methods also yield relaxations, which can be made arbitrarily accurate at a higher computational cost [27,43,44]. In this case, we note we also have a poset structure that can be exploited to build a similar constraint space.

Let us recall that the so-called Navascués-Pironio-Acín (NPA) hierarchy [27] chooses a set of operators $S = \{\mathbb{1}, A_0, B_0, \dots\}$ from which it builds a moment matrix $\Gamma = S^\dagger S$. Nontrivial relationships among the entries of Γ are

Let us first consider only the cases in which the overlap between 3-body constraints is minimum, which is the most likely scenario in the lower budget and greatly simplifies the calculation. For a case with k 3-body constraints, we have $\xi_{23} = \sum_{j=0}^{n-2k} \binom{n-2k}{j}$ possible 2-body constraint combinations. This way, we obtain a lower bound for the size of the constraint space,

imposed by the algebra generated by the elements of S : commutation relations or identities such as $A_i^\dagger A_i = \mathbb{1}$ impose linear constraints among the entries of Γ . Then, given a Bell inequality that can be formally represented as $I = \text{Tr}[C\Gamma]$, one can find a lower bound to its value over the quantum set by solving

$$\beta_Q^S := \begin{array}{ll} \min_{\Gamma} & \text{Tr}[C\Gamma] \\ \text{s.t.} & \Gamma \succeq 0 \\ & \Gamma_{00} = 1 \\ & \text{Tr}[C_i \Gamma] = 0 \end{array} \quad , \quad (G1)$$

where C and C_i are real matrices, thus obtaining a quantum Bell inequality of the form $I \geq \beta_Q^S$. Note that C picks the coefficients of I and the C_i enforces the conditions arising from the operator algebra. For instance, if the Bell scenario is such that the outcomes of the measurements are ± 1 then $A_k^2 - \mathbb{1} = 0$. Then C_i picks the entries A_k^2 and $\mathbb{1}$ in Γ with the appropriate coefficients, imposing the equality constraint. Similarly, commutation relations such as $[A_k, B_l] = 0$ are enforced in the same way.

In this case, the poset structure lies in the definition of the set of operators S . The partial order relation \preceq corresponds to the inclusion order relation \subseteq between two different sets of operators and the agent can perform actions in a similar way, by adding and removing operators.

In analogy to Appendix E, some witnesses for the quantum set admit a proof for a ver low operator degree in S . The paradigmatic example is the CHSH Bell inequality [161], which can be shown to be bounded by $2\sqrt{2}$,

$$\begin{aligned} & 2\sqrt{2}\mathbb{1} - (A_0 B_0 + A_0 B_1 + A_1 B_0 - A_1 B_1) \\ &= \frac{1}{\sqrt{2}} \sum_{i=0}^1 \left(A_i - \frac{B_0 + (-1)^i B_1}{\sqrt{2}} \right)^\dagger \\ & \quad \times \left(A_i - \frac{B_0 + (-1)^i B_1}{\sqrt{2}} \right) \succeq 0. \end{aligned} \quad (G2)$$

On the other hand, inequalities such as the so-called I_{3322} inequality [162] do not seem to admit a tight proof for their quantum bound, even in the case that S contains operators up to degree 5 [137,163]. Having simple certificates such as those of the form of Eq. (G2) turns out to be extremely convenient for proofs in device-independent quantum information processing protocols, such as self-testing: When β_Q^S is tight, it means that the quantum state and measurements yield exactly zero expectation value on all the sos terms [cf. Eq. (G2)]. These equations then impose conditions that allow to characterize the states and/or measurements performed to some extent, solely from their statistics [164–171].

2. Improving sum-of-squares representations of non-negative polynomials

Semidefinite programming optimization is essentially equivalent to finding sum-of-squares decompositions [93]. The latter arises naturally when trying to answer the following question: Given a real polynomial in d variables, does it take non-negative values for all points in \mathbb{R}^d ? This is precisely Hilbert’s 17th problem [172]. On the one hand, it is a trivial observation that every polynomial that admits a sum-of-squares representation is non-negative by construction, and the latter can be efficiently found via a SdP. Unfortunately, not every non-negative polynomial admits a sum-of-squares decomposition in terms of polynomials [173]. In fact, although Hilbert’s problem was solved by Artin in 1927 [174], who showed that every non-negative polynomial admits a sum-of-squares representation in terms of rational functions, the problem remains NP-hard. By controlling the degree of the denominator in the rational function sos, one also obtains a hierarchy.

Interestingly, non-negative polynomials also appear naturally in physics and optimization. For instance, imagine one wants to find the minimal energy of a classical local Hamiltonian. This task appears naturally in the verification of quantum optimizers [175] or in the context of finding the classical bound of a Bell inequality with few-body correlators [114]. In these cases, there are some geometric properties imprinted in the cost function, which one would like that they persist in the sos decomposition. However, such decompositions are not unique in general. When the underlying graph that connects variables that interact directly is chordal, the sparsity in the objective function percolates to a sparse sos decomposition [176–179]. However, in the case that the underlying graph has a complicated chordal extension, it is significantly harder to obtain good sos decompositions, since there is no systematic method in this case, making the situation amenable to a RL agent. It would be interesting to see to which extent a RL agent recovers a perfect elimination ordering stemming from a chordal graph and whether it can find effective strategies when the graph is approximately chordal.

3. Optimization of nonlocality depth witnesses from few-body Bell inequalities

Here we consider the following multipartite Bell scenario, where n parties labeled from $[n]$ are space-like separated and each of them can perform m measurements each yielding

d possible outcomes. At the end of the experiment, parties have collected enough statistics to estimate the conditional probability distribution $p(\mathbf{a}|\mathbf{x})$, where \mathbf{x} is an n -dimensional vector denoting a collective choice of measurements and \mathbf{a} is also an n -dimensional vector labeling the corresponding outcomes. Studying Bell nonlocality in such a multipartite scenario easily turns into a highly complex task, even from the point of view of designing or finding relevant Bell inequalities [180–184]. Furthermore, in the multipartite scenario, analogously to entanglement [121,123,127,128,185], Bell nonlocality can come in many flavors, from fully local models to bilocal models that are only falsified by genuinely nonlocal correlations. In addition, the multipartite Bell scenario poses the extra challenge of a consistent time ordering in defining a partially local model, otherwise it could be self-contradicting [186,187]. This caveat can be avoided by defining a so-called k -local model, which is a mixture of models of the form

$$p(\mathbf{a}|\mathbf{x}) = \sum_{\lambda} p(\lambda) \prod_{i=1}^L p(\mathbf{a}_{S_i}|\mathbf{x}_{S_i}, \lambda), \tag{G3}$$

where $\{S_i\}_{i=1}^L$ form a partition of $[n]$ with $|S_i| \leq k$, the so-called response functions $p(\mathbf{a}_{S_i}|\mathbf{x}_{S_i}, \lambda)$ satisfy the no-signalling principle and $\mathbf{a}_S, \mathbf{x}_S$ indicate that we select from \mathbf{a} or \mathbf{x} , respectively, only those components whose index belongs to $S \subseteq [n]$. By mixing models of the form Eq. (G3), one constructs k -local models, in this case, under no-signalling constraints.

In [188,189] a way to optimize Bell inequalities for k -nonlocality depth was proposed for large system sizes, leveraging on two factors that simplify the problem: designing Bell inequalities that are (i) permutationally invariant and (ii) composed of two-body correlators only. As can be inferred from Eq. (G3), in order to construct these inequalities or to find their k -local bound given one, one needs to know a characterization, in terms of extremal points, of the projected no-signalling polytope for $|S_i|$ parties in the relevant Bell scenario. This way, one can construct the k -local polytope [182,188,190]. Unfortunately, the polytope of nonsignalling correlations admits an easy description only in terms of inequalities. In terms of vertices, the so-called PR-boxes [35], it has been shown that finding all PR-boxes is equivalent to finding all Bell inequalities [191]. Therefore, it seems that such a daunting task [192,193] could benefit from a relaxation approach, which we here describe: By allowing for a simpler characterization of the projected no-signalling polytope, one can hope to construct k -local models with less extremal points to be considered.

Let us recall that the no-signalling principle states that the probability distributions $p(\mathbf{a}|\mathbf{x})$, apart from satisfying the relations $\sum_{\mathbf{a}} p(\mathbf{a}|\mathbf{x}) = 1$ for all \mathbf{x} and $p(\mathbf{a}|\mathbf{x}) \geq 0$ for every \mathbf{a}, \mathbf{x} , as does any mathematically sound probability distribution, they also satisfy the so-called no-signalling principle, which reads

$$p(\mathbf{a}_S|\mathbf{x}_S \cup \mathbf{x}_{S^c}) = p(\mathbf{a}_S|\mathbf{x}_S \cup \mathbf{x}'_{S^c}), \forall \mathbf{a}_S, \mathbf{x}_S, \mathbf{x}_{S^c}, \mathbf{x}'_{S^c}, S \subseteq [n] \tag{G4}$$

where we have split \mathbf{x} into those components labeled in S, \mathbf{x}_S and those in its complementary set, and $p(\mathbf{a}_S|\mathbf{x})$ is defined as

the marginal probability distribution

$$p(\mathbf{a}_S|\mathbf{x}) = \sum_{\mathbf{a}_{S^c}} p(\mathbf{a}|\mathbf{x}). \quad (\text{G5})$$

Note that, operationally, the NS principle imposes that the marginal probability distribution that a subset S of parties observe does not depend on the inputs \mathbf{x}_{S^c} received by the rest of the parties during the experiment. Hence, the rest of the parties cannot signal information to the parties in S by choosing a particular set of inputs. Furthermore, the no-signalling principle tells us that the quantities $p(\mathbf{a}_S|\mathbf{x}_S)$ are well defined.

We can now build the relaxation as follows: The projected no-signalling polytope, in terms of 2-body correlations,

is given in terms of the marginals $p(\mathbf{a}_S|\mathbf{x}_S)$, where $|S| = 2$ (we can take particular linear combinations of them to build symmetric correlators). Each of the $p(\mathbf{a}_S|\mathbf{x}_S)$ stemmed from a common $p(\mathbf{a}|\mathbf{x})$, but at a first relaxation level, this assumption can be dropped. The relaxation hierarchy is then built by imposing compatibility at larger and larger levels: for instance, given $S, W \subseteq [n]$, $|S| = |W| = 2$, $|S \cap W| = 1$, we can impose that there exists a no-signalling three-partite $p(\mathbf{a}_{SUW}|\mathbf{x}_{SUW})$ with appropriate marginals. From the set inclusion relation one recovers the same poset structure in the constraint space. Since now the problem is linear, one can build outer approximations to the projected no-signalling polytope by means of a linear programming black box or, equivalently, a diagonal SdP.

-
- [1] A. Kandala, A. Mezzacapo, K. Temme, M. Takita, M. Brink, J. M. Chow, and J. M. Gambetta, Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets, *Nature (London)* **549**, 242 (2017).
- [2] A. Kandala, K. Temme, A. D. Córcoles, A. Mezzacapo, J. M. Chow, and J. M. Gambetta, Error mitigation extends the computational reach of a noisy quantum processor, *Nature (London)* **567**, 491 (2019).
- [3] A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. O'Brien, A variational eigenvalue solver on a photonic quantum processor, *Nat. Commun.* **5**, 4213(2014).
- [4] B. P. Lanyon, J. D. Whitfield, G. G. Gillett, M. E. Goggin, M. P. Almeida, I. Kassal, J. D. Biamonte, M. Mohseni, B. J. Powell, M. Barbieri, A. Aspuru-Guzik, and A. G. White, Towards quantum chemistry on a quantum computer, *Nat. Chem.* **2**, 106 (2010).
- [5] C. Hempel, C. Maier, J. Romero, J. McClean, T. Monz, H. Shen, P. Jurcevic, B. P. Lanyon, P. Love, R. Babbush, A. Aspuru-Guzik, R. Blatt, and C. F. Roos, Quantum Chemistry Calculations on a Trapped-Ion Quantum Simulator, *Phys. Rev. X* **8**, 031022(2018).
- [6] P. O'Malley, R. Babbush, I. Kivlichan, J. Romero, J. McClean, R. Barends, J. Kelly, P. Roushan, A. Tranter, N. Ding, B. Campbell, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth, A. Fowler, E. Jeffrey, E. Lucero, A. Megrant, J. Mutus, *et al.*, Scalable quantum simulation of molecular energies, *Phys. Rev. X* **6**, 031007 (2016).
- [7] T. E. O'Brien, B. Senjean, R. Sagastizabal, X. Bonet-Monroig, A. Dutkiewicz, F. Buda, L. DiCarlo, and L. Visscher, Calculating energy derivatives for quantum chemistry on a quantum computer, *npj Quantum Inf.* **5**, 113 (2019).
- [8] S. R. White, Density Matrix Formulation for Quantum Renormalization Groups, *Phys. Rev. Lett.* **69**, 2863 (1992).
- [9] S. R. White, Density-matrix algorithms for quantum renormalization groups, *Phys. Rev. B* **48**, 10345 (1993).
- [10] F. Verstraete, D. Porras, and J. I. Cirac, Density Matrix Renormalization Group and Periodic Boundary Conditions: A Quantum Information Perspective, *Phys. Rev. Lett.* **93**, 227205 (2004).
- [11] A. J. Daley, C. Kollath, U. Schollwöck, and G. Vidal, Time-dependent density-matrix renormalization-group using adaptive effective hilbert spaces, *J. Stat. Mech.: Theory Exp.* (2004) P04005.
- [12] R. Orús, A practical introduction to tensor networks: Matrix product states and projected entangled pair states, *Ann. Phys.* **349**, 117 (2014).
- [13] C. Bravo-Prieto, J. Lumbrales-Zarapico, L. Tagliacozzo, and J. I. Latorre, Scaling of variational quantum circuit depth for condensed matter systems, *Quantum* **4**, 272 (2020).
- [14] J. Biamonte, P. Wittek, N. Pancotti, P. Rebentrost, N. Wiebe, and S. Lloyd, Quantum machine learning, *Nature (London)* **549**, 195 (2017).
- [15] V. Dunjko and H. J. Briegel, Machine learning & artificial intelligence in the quantum domain: a review of recent progress, *Rep. Prog. Phys.* **81**, 074001 (2018).
- [16] E. Farhi, J. Goldstone, and S. Gutmann, A quantum approximate optimization algorithm, *arXiv:1411.4028*.
- [17] C. Kokail, C. Maier, R. van Bijnen, T. Brydges, M. K. Joshi, P. Jurcevic, C. A. Muschik, P. Silvi, R. Blatt, C. F. Roos, and P. Zoller, Self-verifying variational quantum simulation of lattice models, *Nature (London)* **569**, 355 (2019).
- [18] G. E. Crooks, Performance of the quantum approximate optimization algorithm on the maximum cut problem, *arXiv:1811.08419*.
- [19] L. Zhou, S.-T. Wang, S. Choi, H. Pichler, and M. D. Lukin, Quantum Approximate Optimization Algorithm: Performance, Mechanism, and Implementation on Near-Term Devices, *Phys. Rev. X* **10**, 021067 (2020).
- [20] M. P. Harrigan, K. J. Sung, M. Neeley, K. J. Satzinger, F. Arute, K. Arya, J. Atalaya, J. C. Bardin, R. Barends, S. Boixo *et al.*, Quantum approximate optimization of non-planar graph problems on a planar superconducting processor, *Nat. Phys.* **17**, 332 (2021).
- [21] Y. Herasymenko and T. O'Brien, A diagrammatic approach to variational quantum ansatz construction, *Quantum* **5**, 596 (2021).
- [22] A. Garcia-Saez and J. I. Latorre, Addressing hard classical problems with adiabatically assisted variational quantum eigensolvers *arXiv:1806.02287*.
- [23] R. Sagastizabal, X. Bonet-Monroig, M. Singh, M. A. Rol, C. C. Bultink, X. Fu, C. H. Price, V. P. Ostroukh, N. Muthusubramanian, A. Bruno, M. Beekman, N. Haider, T. E. O'Brien, and L. DiCarlo, Experimental error mitigation via

- symmetry verification in a variational quantum eigensolver, *Phys. Rev. A* **100**, 010302(2019).
- [24] J. Tura, Quantum algorithms for near-term devices, in *Quantum Software Engineering & Programming*, edited by M. Piattini, G. Peterssen, R. Perez-Castillo, J. L. Hevia, and M. A. Serrano, CEURWorkshop Proceedings No. 2561 (Aachen, 2020), pp. 38–50.
- [25] M. Benedetti, D. Garcia-Pintos, O. Perdomo, V. Leyton-Ortega, Y. Nam, and A. Perdomo-Ortiz, A generative modeling approach for benchmarking and training shallow quantum circuits, *npj Quantum Inf.* **5**, 45(2019).
- [26] A. C. Doherty, P. A. Parrilo, and F. M. Spedalieri, Complete family of separability criteria, *Phys. Rev. A* **69**, 022308 (2004).
- [27] M. Navascués, S. Pironio, and A. Acín, Bounding the Set of Quantum Correlations, *Phys. Rev. Lett.* **98**, 010401 (2007).
- [28] F. Baccari, D. Cavalcanti, P. Wittek, and A. Acín, Efficient Device-Independent Entanglement Detection for Multipartite Systems, *Phys. Rev. X* **7**, 021042 (2017).
- [29] A. Peres, Separability Criterion for Density Matrices, *Phys. Rev. Lett.* **77**, 1413 (1996).
- [30] L. Gurvits, Classical deterministic complexity of Edmonds' problem and quantum entanglement, in *Proceedings of the Thirty-Fifth ACM Symposium on Theory of Computing - STOC'03* (ACM Press, New York, 2003).
- [31] A. Acín, N. Brunner, N. Gisin, S. Massar, S. Pironio, and V. Scarani, Device-Independent Security of Quantum Cryptography against Collective Attacks, *Phys. Rev. Lett.* **98**, 230501 (2007).
- [32] R. Gallego, L. Masanes, G. D. L. Torre, C. Dhara, L. Aolita, and A. Acín, Full randomness from arbitrarily deterministic events, *Nat. Commun.* **4**, 2654(2013).
- [33] R. Augusiak, M. Demianowicz, M. Pawłowski, J. Tura, and A. Acín, Elemental and tight monogamy relations in nonsignaling theories, *Phys. Rev. A* **90**, 052323(2014).
- [34] W. Slofstra, The set of quantum correlations is not closed, *Forum Math Pi* **7**, e1 (2019).
- [35] S. Popescu and D. Rohrlich, Quantum nonlocality as an axiom, *Found. Phys.* **24**, 379 (1994).
- [36] G. Brassard, H. Buhrman, N. Linden, A. A. Méthot, A. Tapp, and F. Unger, Limit on Nonlocality in Any World in Which Communication Complexity Is Not Trivial, *Phys. Rev. Lett.* **96**, 250401 (2006).
- [37] N. Linden, S. Popescu, A. J. Short, and A. Winter, Quantum Nonlocality and Beyond: Limits from Nonlocal Computation, *Phys. Rev. Lett.* **99**, 180502 (2007).
- [38] M. Navascués and H. Wunderlich, A glance beyond the quantum model, *Proc. R. Soc. London A* **466**, 881 (2010).
- [39] M. Pawłowski, T. Paterek, D. Kaszlikowski, V. Scarani, A. Winter, and M. Zukowski, Information causality as a physical principle, *Nature (London)* **461**, 1101 EP (2009).
- [40] T. Fritz, A. B. Sainz, R. Augusiak, J. B. Brask, R. Chaves, A. Leverrier, and A. Acín, Local orthogonality as a multipartite principle for quantum correlations, *Nat. Commun.* **4**, 2263 (2013), article.
- [41] R. Gallego, L. E. Würflinger, A. Acín, and M. Navascués, Quantum Correlations Require Multipartite Information Principles, *Phys. Rev. Lett.* **107**, 210403 (2011).
- [42] M. Navascués, Y. Guryanova, M. J. Hoban, and A. Acín, Almost quantum correlations, *Nat. Commun.* **6**, 6288 (2015), article.
- [43] M. Navascués, S. Pironio, and A. Acín, A convergent hierarchy of semidefinite programs characterizing the set of quantum correlations, *New J. Phys.* **10**, 073013 (2008).
- [44] S. Pironio, M. Navascués, and A. Acín, Convergent relaxations of polynomial optimization problems with noncommuting variables, *SIAM J. Optim.* **20**, 2157 (2010).
- [45] P. W. Anderson, Limits on the energy of the antiferromagnetic ground state, *Phys. Rev.* **83**, 1260 (1951).
- [46] L. Epele, H. Fanchiotti, C. García-Canal, and A. Pacheco, Lower bound for the ground energy of spin systems, *Physica A* **173**, 500 (1991).
- [47] D. A. Mazziotti and R. M. Erdahl, Uncertainty relations and reduced density matrices: Mapping many-body quantum mechanics onto four particles, *Phys. Rev. A* **63**, 042113 (2001).
- [48] G. Gidofalvi and D. A. Mazziotti, Boson correlation energies via variational minimization with the two-particle reduced density matrix: Exact N-representability conditions for harmonic interactions, *Phys. Rev. A* **69**, 042511(2004).
- [49] D. A. Mazziotti, Realization of Quantum Chemistry without Wave Functions through First-Order Semidefinite Programming, *Phys. Rev. Lett.* **93**, 213001 (2004).
- [50] T. J. Osborne, Lower bound for the ground-state energy density of a 1D quantum spin system, [arXiv:cond-mat/0508428](https://arxiv.org/abs/cond-mat/0508428).
- [51] D. A. Mazziotti, Variational reduced-density-matrix method using three-particle n -representability conditions with application to many-electron molecules, *Phys. Rev. A* **74**, 032501 (2006).
- [52] D. A. Mazziotti, Structure of fermionic density matrices: Complete N-Representability Conditions, *Phys. Rev. Lett.* **108**, 263002(2012).
- [53] D. A. Mazziotti, Enhanced Constraints for Accurate Lower Bounds on Many-Electron Quantum Energies from Variational Two-Electron Reduced Density Matrix Theory, *Phys. Rev. Lett.* **117**, 153001 (2016).
- [54] F. Uskov and O. Lychkovskiy, A variational lower bound on the ground state of a many-body system and the squaring parametrization of density matrices, *J. Phys.: Conf. Ser.* **1163**, 012057 (2019).
- [55] O. Vinyals, I. Babuschkin, W. M. Czarnecki, M. Mathieu, A. Dudzik, J. Chung, D. H. Choi, R. Powell, T. Ewalds, P. Georgiev *et al.*, Grandmaster level in starcraft ii using multi-agent reinforcement learning, *Nature (London)* **575**, 350 (2019).
- [56] A. Ramesh, M. Pavlov, G. Goh, S. Gray, C. Voss, A. Radford, M. Chen, and I. Sutskever, [Zero-shot text-to-image generation](https://arxiv.org/abs/2205.06175), in *Proceedings of the 38th International Conference on Machine Learning*, edited by M. Meila and T. Zhang, Proceedings of Machine Learning Research, Vol. 139 (PMLR, 2021), pp. 8821–8831.
- [57] S. Reed, K. Zolna, E. Parisotto, S. G. Colmenarejo, A. Novikov, G. Barth-Maron, M. Gimenez, Y. Sulsky, J. Kay, J. T. Springenberg *et al.*, A generalist agent, [arXiv:2205.06175](https://arxiv.org/abs/2205.06175).
- [58] G. Carleo, I. Cirac, K. Cranmer, L. Daudet, M. Schuld, N. Tishby, L. Vogt-Maranto, and L. Zdeborová, Machine learning and the physical sciences, *Rev. Mod. Phys.* **91**, 045002(2019).
- [59] A. Dawid, J. Arnold, B. Requena, A. Gresch, M. Płodzień, K. Donatella, K. A. Nicoli, P. Stornati, R. Koch, M. Büttner *et al.*,

- Modern applications of machine learning in quantum sciences, [arXiv:2204.04198](#).
- [60] J. Carrasquilla and R. G. Melko, Machine learning phases of matter, *Nat. Phys.* **13**, 431 (2017).
- [61] P. Huembeli, A. Dauphin, and P. Wittek, Identifying quantum phase transitions with adversarial neural networks, *Phys. Rev. B* **97**, 134109(2018).
- [62] B. S. Rem, N. Käming, M. Tarnowski, L. Asteria, N. Fläschner, C. Becker, K. Sengstock, and C. Weitenberg, Identifying quantum phase transitions using artificial neural networks on experimental data, *Nat. Phys.* **15**, 917 (2019).
- [63] G. Carleo and M. Troyer, Solving the quantum many-body problem with artificial neural networks, *Science* **355**, 602 (2017).
- [64] Y. Nomura, A. S. Darmawan, Y. Yamaji, and M. Imada, Restricted boltzmann machine learning for solving strongly correlated quantum systems, *Phys. Rev. B* **96**, 205152 (2017).
- [65] M. Hibat-Allah, M. Ganahl, L. E. Hayward, R. G. Melko, and J. Carrasquilla, Recurrent neural network wave functions, *Phys. Rev. Res.* **2**, 023358(2020).
- [66] Y. Bengio, A. Lodi, and A. Prouvost, Machine learning for combinatorial optimization: A methodological tour d’horizon, *Eur. J. Oper. Res.* **290**, 405 (2021).
- [67] O. Vinyals, M. Fortunato, and N. Jaitly, Pointer networks, in *Advances in Neural Information Processing Systems*, edited by C. Cortes, N. Lawrence, D. Lee, M. Sugiyama, and R. Garnett, Vol. 28 (Curran Associates, Inc., New York, 2015).
- [68] N. Karalias and A. Loukas, Erdos goes neural: An unsupervised learning framework for combinatorial optimization on graphs, in *Advances in Neural Information Processing Systems*, edited by H. Larochelle, M. Ranzato, R. Hadsell, M. Balcan, and H. Lin, Vol. 33 (Curran Associates, Inc., New York, 2020) pp. 6659–6672.
- [69] N. Mazyavkina, S. Sviridov, S. Ivanov, and E. Burnaev, Reinforcement learning for combinatorial optimization: A survey, *Comput. Oper. Res.* **134**, 105400 (2021).
- [70] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction* (MIT press, Cambridge, 2018).
- [71] D. Silver, T. Hubert, J. Schrittwieser, I. Antonoglou, M. Lai, A. Guez, M. Lanctot, L. Sifre, D. Kumaran, T. Graepel *et al.*, A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play, *Science* **362**, 1140 (2018).
- [72] A. W. Senior, R. Evans, J. Jumper, J. Kirkpatrick, L. Sifre, T. Green, C. Qin, A. Židek, A. W. Nelson, A. Bridgland *et al.*, Improved protein structure prediction using potentials from deep learning, *Nature (London)* **577**, 706 (2020).
- [73] A. A. Melnikov, H. P. Nautrup, M. Krenn, V. Dunjko, M. Tiersch, A. Zeilinger, and H. J. Briegel, Active learning machine learns to create new quantum experiments, *Proc. Natl. Acad. Sci. U.S.A.* **115**, 1221 (2018).
- [74] M. Bukov, A. G. Day, D. Sels, P. Weinberg, A. Polkovnikov, and P. Mehta, Reinforcement Learning in Different Phases of Quantum Control, *Phys. Rev. X* **8**, 031086 (2018).
- [75] X.-M. Zhang, Z. Wei, R. Asad, X.-C. Yang, and X. Wang, When does reinforcement learning stand out in quantum control? A comparative study on state preparation, *npj Quantum Inf.* **5**, 1 (2019).
- [76] H. P. Nautrup, N. Delfosse, V. Dunjko, H. J. Briegel, and N. Friis, Optimizing quantum error correction codes with reinforcement learning, *Quantum* **3**, 215 (2019).
- [77] F. Verstraete, V. Murg, and J. Cirac, Matrix product states, projected entangled pair states, and variational renormalization group methods for quantum spin systems, *Adv. Phys.* **57**, 143 (2008).
- [78] N. Schuch, I. Cirac, and D. Pérez-García, PEPS as ground states: Degeneracy and topology, *Ann. Phys.* **325**, 2153 (2010).
- [79] N. Schuch and J. I. Cirac, Matrix product state and mean-field solutions for one-dimensional systems can be found efficiently, *Phys. Rev. A* **82**, 012314(2010).
- [80] Y. Zhou, E. M. Stoudenmire, and X. Waintal, What Limits the Simulation of Quantum Computers? *Phys. Rev. X* **10**, 041038 (2020).
- [81] J. Kempe and O. Regev, 3-local Hamiltonian is QMA-complete, *Quantum Info. Comput.* **3**, 258 (2003).
- [82] J. Kempe, A. Kitaev, and O. Regev, The complexity of the local Hamiltonian problem, *SIAM J. Comput.* **35**, 1070 (2006).
- [83] D. Aharonov, D. Gottesman, S. Irani, and J. Kempe, The power of quantum systems on a line, *Commun. Math. Phys.* **287**, 41 (2009).
- [84] N. Schuch and F. Verstraete, Computational complexity of interacting electrons and fundamental limitations of density functional theory, *Nat. Phys.* **5**, 732 (2009).
- [85] J. Czartowski, K. Szymański, B. Gardas, Y. V. Fyodorov, and K. Życzkowski, Separability gap and large-deviation entanglement criterion, *Phys. Rev. A* **100**, 042326 (2019).
- [86] R. Tarrach and R. Valent, Exact lower bounds to the ground-state energy of spin systems: The two-dimensional $S=1/2$ antiferromagnetic Heisenberg model, *Phys. Rev. B* **41**, 9611 (1990).
- [87] A. Chandran, D. Kaszlikowski, A. Sen(De), U. Sen, and V. Vedral, Regional Versus Global Entanglement in Resonating-Valence-Bond States, *Phys. Rev. Lett.* **99**, 170502 (2007).
- [88] F. Alet, D. Braun, and G. Misguich, Comment on “Regional Versus Global Entanglement in Resonating-Valence-Bond States”, *Phys. Rev. Lett.* **101**, 248901(2008).
- [89] A. Aloy, M. Fadel, and J. Tura, The quantum marginal problem for symmetric states: Applications to variational optimization, nonlocality and self-testing, *New J. Phys.* **23**, 033026 (2021).
- [90] M. Walter, B. Doran, D. Gross, and M. Christandl, Entanglement polytopes: Multiparticle entanglement from single-particle information, *Science* **340**, 1205 (2013).
- [91] A. Sawicki, M. Oszmaniec, and M. Kuś, Critical sets of the total variance can detect all stochastic local operations and classical communication classes of multiparticle entanglement, *Phys. Rev. A* **86**, 040304 (2012).
- [92] A. Sawicki, M. Oszmaniec, and M. Kuś, Convexity of momentum map, morse index, and quantum entanglement, *Rev. Math. Phys.* **26**, 1450004 (2014).
- [93] R. T. Grigoriy Blekherman, Pablo A. Parrilo, ed., *Semidefinite Optimization and Convex Algebraic Geometry* (SIAM, 2013).
- [94] M. Navascués, M. Owari, and M. B. Plenio, Power of symmetric extensions for entanglement detection, *Phys. Rev. A* **80**, 052306 (2009).
- [95] M. Grötschel, L. Lovász, and A. Schrijver, *Geometric Algorithms and Combinatorial Optimization* (Springer, Berlin, 1993).
- [96] F. Alizadeh, Interior point methods in semidefinite programming with applications to combinatorial optimization, *SIAM J. Optim.* **5**, 13 (1995).

- [97] S. Arora, E. Hazan, and S. Kale, Fast algorithms for approximate semidefinite programming using the multiplicative weights update method, in *46th Annual IEEE Symposium on Foundations of Computer Science (FOCS'05)* (IEEE, Piscataway, NJ, 2005).
- [98] J. F. Sturm, Using sedumi 1.02, a matlab toolbox for optimization over symmetric cones, *Optim. Methods Softw.* **11**, 625 (1999).
- [99] D. Peaucelle, D. Henrion, L. Yann, and K. Taitz, User's guide for sedumi interface 1.04 (LAAS-CNRS, Toulouse, 2002).
- [100] F. G. Brandao and K. M. Svore, Quantum speed-ups for solving semidefinite programs, in *2017 IEEE 58th Annual Symposium on Foundations of Computer Science (FOCS)* (IEEE, Piscataway, NJ, 2017).
- [101] T. Kriváchy, Y. Cai, J. Bowles, D. Cavalcanti, and N. Brunner, High-speed batch processing of semidefinite programs with feedforward neural networks, *New J. Phys.* **23**, 103034 (2021).
- [102] C. J. C. H. Watkins, Learning from delayed rewards, Ph.D. thesis, King's College, Cambridge, 1989.
- [103] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski *et al.*, Human-level control through deep reinforcement learning, *Nature (London)* **518**, 529 (2015).
- [104] H. v. Hasselt, A. Guez, and D. Silver, Deep reinforcement learning with double q-learning, in *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence, AAAI'16* (AAAI Press, Palo Alto, CA, 2016), pp. 2094-2100.
- [105] E. Lieb, T. Schultz, and D. Mattis, Two soluble models of an antiferromagnetic chain, *Ann. Phys.* **16**, 407 (1961).
- [106] S. Sachdev, *Quantum Phase Transitions* (Cambridge University Press, Cambridge, 2009).
- [107] X. Wang, Entanglement in the quantum Heisenberg XY model, *Phys. Rev. A* **64**, 012313 (2001).
- [108] X. Wang, Thermal and ground-state entanglement in Heisenberg XX qubit rings, *Phys. Rev. A* **66**, 034302 (2002).
- [109] A. D. Pasquale, G. Costantini, P. Facchi, G. Florio, S. Pascazio, and K. Yuasa, XX model on the circle, *Eur. Phys. J.: Spec. Top.* **160**, 127 (2008).
- [110] H. Nishimori and Y. Ozeki, Ground-state long-range order in the two-dimensional xxz model, *J. Phys. Soc. Jpn.* **58**, 1027 (1989).
- [111] P. Jordan and E. Wigner, Über das paulische Äquivalenzverbot, *Z. Phys.* **47**, 631 (1928).
- [112] M. A. Nielsen, The Fermionic Canonical Commutation Relations and the Jordan-Wigner Transform, School of Physical Sciences The University of Queensland, Vol. 59, 2005.
- [113] M.-C. Bañuls, J. I. Cirac, and M. M. Wolf, Entanglement in fermionic systems, *Phys. Rev. A* **76**, 022311 (2007).
- [114] J. Tura, G. De las Cuevas, R. Augusiak, M. Lewenstein, A. Acín, and J. I. Cirac, Energy as a Detector of Nonlocality of Many-Body Spin Systems, *Phys. Rev. X* **7**, 021005 (2017).
- [115] T. H. D. C. Cormen, C. E. M. Leiserson, R. L. M. Rivest, and C. C. U. Stein, *Introduction to Algorithms* (MIT Press, Cambridge, 2009).
- [116] S. Kirkpatrick, C. D. Gelatt Jr., and M. P. Vecchi, Optimization by simulated annealing, *Science* **220**, 671 (1983).
- [117] M. E. Taylor and P. Stone, Transfer learning for reinforcement learning domains: A survey, *J. Mach. Learn. Res.* **10**, 1633 (2009).
- [118] A. Dawid, P. Huembeli, M. Tomza, M. Lewenstein, and A. Dauphin, Phase detection with neural networks: Interpreting the black box, *New J. Phys.* **22**, 115001 (2020).
- [119] E. P. L. van Nieuwenburg, Y.-H. Liu, and S. D. Huber, Learning phase transitions by confusion, *Nat. Phys.* **13**, 435 (2017).
- [120] K. Kottmann, P. Huembeli, M. Lewenstein, and A. Acín, Unsupervised Phase Discovery with Deep Anomaly Detection, *Phys. Rev. Lett.* **125**, 170603(2020).
- [121] O. Gühne and G. Tóth, Entanglement detection, *Phys. Rep.* **474**, 1 (2009).
- [122] G. Tóth and O. Gühne, Detection of multipartite entanglement with two-body correlations, *Appl. Phys. B* **82**, 237 (2006).
- [123] O. Gühne, G. Tóth, and H. J. Briegel, Multipartite entanglement in spin chains, *New J. Phys.* **7**, 229 (2005).
- [124] N. Yu, Separability of a mixture of dicke states, *Phys. Rev. A* **94**, 060101(R) (2016).
- [125] J. Tura, A. Aloy, R. Quesada, M. Lewenstein, and A. Sanpera, Separability of diagonal symmetric states: A quadratic conic optimization problem, *Quantum* **2**, 45 (2018).
- [126] C. Marconi, A. Aloy, J. Tura, and A. Sanpera, Entangled symmetric states and copositive matrices, *Quantum* **5**, 561 (2021).
- [127] A. Aloy, J. Tura, F. Baccari, A. Acín, M. Lewenstein, and R. Augusiak, Device-Independent Witnesses of Entanglement Depth from Two-Body Correlators, *Phys. Rev. Lett.* **123**, 100507(2019).
- [128] J. Tura, A. Aloy, F. Baccari, A. Acín, M. Lewenstein, and R. Augusiak, Optimization of device-independent witnesses of entanglement depth from two-body correlators, *Phys. Rev. A* **100**, 032307(2019).
- [129] G. Tóth, Entanglement witnesses in spin models, *Phys. Rev. A* **71**, 010301 (2005).
- [130] B. F. Schiffer, J. Tura, and J. I. Cirac, Adiabatic spectroscopy and a variational quantum adiabatic algorithm, *PRX Quantum* **3**, 020347 (2022).
- [131] B. Requena, G. Muñoz-Gil, and J. Tura, Borjarequena/bounce, <https://doi.org/10.5281/zenodo.4585623>.
- [132] M. Grant and S. Boyd, CVX: Matlab software for disciplined convex programming, version 2.1, <http://cvxr.com/cvx>.
- [133] M. Grant and S. Boyd, Graph implementations for nonsmooth convex programs, in *Recent Advances in Learning and Control*, Lecture Notes in Control and Information Sciences, edited by V. Blondel, S. Boyd, and H. Kimura (Springer-Verlag, Berlin, 2008), pp. 95–110, http://stanford.edu/~boyd/graph_dcp.html.
- [134] J. Lofberg, YALMIP : a toolbox for modeling and optimization in MATLAB, in *2004 IEEE International Conference on Robotics and Automation (IEEE Cat. No.04CH37508)* (IEEE, Piscataway, NJ, 2004).
- [135] K. C. Toh, M. Todd, R. Tütüncü, and R. H. Tutuncu, SDPT3 - a matlab software package for semidefinite programming, *Optim. Methods Softw.* **11**, 545 (1999).
- [136] T. Baumgratz and M. B. Plenio, Lower bounds for ground states of condensed matter systems, *New J. Phys.* **14**, 023027 (2012).
- [137] D. Rosset, Characterization of correlations in quantum networks, Ph.D. thesis, University of Geneva, 2015.
- [138] A. Tavakoli, D. Rosset, and M.-O. Renou, Enabling Computation of Correlation Bounds for Finite-Dimensional Quantum Systems via Symmetrization, *Phys. Rev. Lett.* **122**, 070501(2019).

- [139] K. Gatermann and P. A. Parrilo, Symmetry groups, semidefinite programs, and sums of squares, *J. Pure Appl. Algebra* **192**, 95 (2004).
- [140] M. B. Hastings, Solving gapped Hamiltonians locally, *Phys. Rev. B* **73**, 085115(2006).
- [141] A. Kitaev, Fault-tolerant quantum computation by anyons, *Ann. Phys.* **303**, 2 (2003).
- [142] A. Kitaev, Anyons in an exactly solved model and beyond, *Ann. Phys.* **321**, 2 (2006).
- [143] M. A. Levin and X.-G. Wen, String-net condensation: A physical mechanism for topological phases, *Phys. Rev. B* **71**, 045110(2005).
- [144] D. Gottesman, An introduction to quantum error correction and fault-tolerant quantum computation, in *Quantum Information Science and its Contributions to Mathematics, Proceedings of Symposia in Applied Mathematics*, Vol. 68 (2010), pp. 13–58.
- [145] M. Hein, W. Dür, J. Eisert, R. Raussendorf, M. Van den Nest, and H. Briegel, Entanglement in graph states and its applications, in *International School of Physics Enrico Fermi*, Vol. 162 (2006).
- [146] H. J. Briegel and R. Raussendorf, Persistent Entanglement in Arrays of Interacting Particles, *Phys. Rev. Lett.* **86**, 910 (2001).
- [147] A. S. Darmawan and S. D. Bartlett, Graph states as ground states of two-body frustration-free Hamiltonians, *New J. Phys.* **16**, 073013 (2014).
- [148] S. Bravyi, Efficient algorithm for a quantum analogue of 2-sat, *Contemp. Math.* **536**, 33 (2011).
- [149] J. Chen, X. Chen, R. Duan, Z. Ji, and B. Zeng, No-go theorem for one-way quantum computing on naturally occurring two-level systems, *Phys. Rev. A* **83**, 050301(R) (2011).
- [150] M. A. Nielsen, Cluster-state quantum computation, *Rep. Math. Phys.* **57**, 147 (2006).
- [151] M. V. den Nest, K. Luttmmer, W. Dür, and H. J. Briegel, Graph states as ground states of many-body spin-1/2 Hamiltonians, *Phys. Rev. A* **77**, 012301 (2008).
- [152] I. Affleck, T. Kennedy, E. H. Lieb, and H. Tasaki, Rigorous Results on Valence-Bond Ground States in Antiferromagnets, *Phys. Rev. Lett.* **59**, 799 (1987).
- [153] D. S. Rokhsar and S. A. Kivelson, Superconductivity and the Quantum Hard-Core Dimer Gas, *Phys. Rev. Lett.* **61**, 2376 (1988).
- [154] C. Castelnovo, C. Chamon, C. Mudry, and P. Pujol, From quantum mechanics to classical statistical physics: Generalized Rokhsar-Kivelson Hamiltonians and the stochastic matrix form decomposition, *Ann. Phys.* **318**, 316 (2005).
- [155] D. Perez-Garcia, F. Verstraete, M. Wolf, and J. Cirac, Matrix product state representations, *Quantum Inf. Comput.* **7**, 401 (2007).
- [156] D. Pérez-García, F. Verstraete, M. Wolf, and J. Cirac, Peps as unique ground states of local Hamiltonians, *Quantum Inf. Comput.* **8**, 650 (2008).
- [157] J. I. Cirac, J. Garre-Rubio, and D. Pérez-García, Mathematical open problems in projected entangled pair states, *Rev. Mat. Complut.* **32**, 579 (2019).
- [158] E. Cruz-Rico, Efficient Preparation of a Family of Ground States of 2D Local Hamiltonians and their Verification Properties, Master's thesis, Ludwig-Maximilians-Universität München, 2020.
- [159] O. Sattath, S. C. Morampudi, C. R. Laumann, and R. Moessner, When a local Hamiltonian must be frustration-free, *Proc. Natl. Acad. Sci. USA* **113**, 6433 (2016).
- [160] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner, Bell nonlocality, *Rev. Mod. Phys.* **86**, 419 (2014).
- [161] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, Proposed Experiment to Test Local Hidden-Variable Theories, *Phys. Rev. Lett.* **23**, 880 (1969).
- [162] M. Froissart, Constructive generalization of Bell's inequalities, *Il Nuovo Cimento B* **64**, 241 (1981).
- [163] K. F. Pál and T. Vértesi, Maximal violation of a bipartite three-setting, two-outcome Bell inequality using infinite-dimensional quantum systems, *Phys. Rev. A* **82**, 022116 (2010).
- [164] I. Šupić and J. Bowles, Self-testing of quantum systems: A review, *Quantum* **4**, 337 (2020).
- [165] A. Acín, S. Massar, and S. Pironio, Randomness Versus Nonlocality and Entanglement, *Phys. Rev. Lett.* **108**, 100402 (2012).
- [166] T. H. Yang and M. Navascués, Robust self-testing of unknown quantum systems into any entangled two-qubit states, *Phys. Rev. A* **87**, 050102(R) (2013).
- [167] C. Bamps and S. Pironio, Sum-of-squares decompositions for a family of Clauser-Horne-Shimony-Holt-like inequalities and their application to self-testing, *Phys. Rev. A* **91**, 052111(2015).
- [168] A. Coladangelo, K. T. Goh, and V. Scarani, All pure bipartite entangled states can be self-tested, *Nat. Commun.* **8**, 15485 EP (2017), article.
- [169] I. Šupić, A. Coladangelo, R. Augusiak, and A. Acín, Self-testing multipartite entangled states through projections onto two systems, *New J. Phys.* **20**, 083041 (2018).
- [170] J. Kaniewski, I. Šupić, J. Tura, F. Baccari, A. Salavrakos, and R. Augusiak, Maximal nonlocality from maximal entanglement and mutually unbiased bases, and self-testing of two-qutrit quantum systems, *Quantum* **3**, 198 (2019).
- [171] F. Baccari, R. Augusiak, I. Šupić, J. Tura, and A. Acín, Scalable Bell Inequalities for Qubit Graph States and Robust Self-Testing, *Phys. Rev. Lett.* **124**, 020402(2020).
- [172] D. Hilbert, Mathematical problems, *Bull. Am. Math.* **8**, 437 (1902).
- [173] T. S. Motzkin, The arithmetic-geometric inequality, in *Inequalities (Proc. Sympos. Wright-Patterson Air Force Base, Ohio, 1965)* (Academic Press, New York, 1967), pp. 205–224.
- [174] E. Artin, Über die zerlegung definiter funktionen in quadrate, *Abh. Math. Semin. Univ. Hambg.* **5**, 100 (1927).
- [175] F. Baccari, C. Gogolin, P. Wittek, and A. Acín, Verifying the output of quantum optimizers with ground-state energy lower bounds, *Phys. Rev. Res.* **2**, 043163(2020).
- [176] L. Vandenberghe and M. S. Andersen, Chordal graphs and semidefinite optimization, *Found. Trend. Optim.* **1**, 241 (2015).
- [177] D. Cifuentes and P. A. Parrilo, Exploiting chordal structure in polynomial ideals: A Röbner bases approach, *SIAM J. Discrete Math.* **30**, 1534 (2016).
- [178] D. Cifuentes and P. A. Parrilo, Chordal networks of polynomial ideals, *SIAM J. Appl. Algebra Geometry* **1**, 73 (2017).
- [179] Y. Zheng, G. Fantuzzi, A. Papachristodoulou, P. Goulart, and A. Wynn, Chordal decomposition in operator-splitting

- methods for sparse semidefinite programs, *Math. Program.* **180**, 489 (2020).
- [180] R. F. Werner and M. M. Wolf, All-multipartite bell-correlation inequalities for two dichotomic observables per site, *Phys. Rev. A* **64**, 032112(2001).
- [181] M. Żukowski and C. Brukner, Bell's Theorem for General N -Qubit States, *Phys. Rev. Lett.* **88**, 210401 (2002).
- [182] C. Śliwa, Symmetries of the Bell correlation inequalities, *Phys. Lett. A* **317**, 165 (2003).
- [183] J.-D. Bancal, N. Gisin, and S. Pironio, Looking for symmetric Bell inequalities, *J. Phys. A: Math. Theor.* **43**, 385303 (2010).
- [184] J. Tura, R. Augusiak, A. B. Sainz, T. Vértesi, M. Lewenstein, and A. Acín, Detecting nonlocality in many-body quantum states, *Science* **344**, 1256 (2014).
- [185] B. Lücke, J. Peise, G. Vitagliano, J. Arlt, L. Santos, G. Tóth, and C. Klempt, Detecting Multiparticle Entanglement of Dicke States, *Phys. Rev. Lett.* **112**, 155304(2014).
- [186] R. Gallego, L. E. Würflinger, A. Acín, and M. Navascués, Operational Framework for Nonlocality, *Phys. Rev. Lett.* **109**, 070401(2012).
- [187] J.-D. Bancal, J. Barrett, N. Gisin, and S. Pironio, Definitions of multipartite nonlocality, *Phys. Rev. A* **88**, 014102(2013).
- [188] F. Baccari, J. Tura, M. Fadel, A. Aloy, J.-D. Bancal, N. Sangouard, M. Lewenstein, A. Acín, and R. Augusiak, Bell correlation depth in many-body systems, *Phys. Rev. A* **100**, 022121(2019).
- [189] V. J. L. Pastor, Exploring nonlocal correlations in many-body systems, Master's thesis, Universitat Politècnica de Catalunya, 2018.
- [190] J. Tura, R. Augusiak, A. Sainz, B. Lücke, C. Klempt, M. Lewenstein, and A. Acín, Nonlocality in many-body quantum systems detected with two-body correlators, *Ann. Phys.* **362**, 370 (2015).
- [191] T. Fritz, Polyhedral duality in bell scenarios with two binary observables, *J. Math. Phys.* **53**, 072202 (2012).
- [192] L. Babai, L. Fortnow, and C. Lund, Non-deterministic exponential time has two-prover interactive protocols, *Comput. Comple.* **1**, 3 (1991).
- [193] B. Chazelle, An optimal convex hull algorithm in any fixed dimension, *Discrete Comput. Geom.* **10**, 377 (1993).