

Certified bounds on optimization problems in quantum theory

Younes Naceur^{1,2}, Jie Wang³, Victor Magron^{2,4}, and Antonio Acín^{1,5}

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

²LAAS-CNRS - Laboratoire d'Analyse et d'Architecture des Systèmes, Centre National de la Recherche Scientifique, Toulouse, France

³State Key Laboratory of Mathematical Sciences, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing, China

⁴Institute of Mathematics, Toulouse, France

⁵ICREA - Institució Catalana de Recerca i Estudis Avançats, 08010 Barcelona, Spain

Semidefinite relaxations of polynomial optimization have become a central tool for addressing the non-convex optimization problems over non-commutative operators that are ubiquitous in quantum information theory and, more in general, quantum physics. Yet, as these global relaxation methods are solved numerically, the bounds issued by the semidefinite solver can - and often do - exceed the global optimum, undermining their certifiability. To counter this issue, we introduce a framework for extracting rigorous certified bounds on non-commutative optimization problems from inexact numerical data, and apply it to several paradigmatic problems in quantum information theory. An extension to sparsity and symmetry-adapted semidefinite relaxations is also provided and compared to the general dense scheme. Our results establish rational post-processing as a practical route to reliable certification, pushing semidefinite optimization toward a certifiable standard for quantum information science.

1 Introduction

Non-commutative polynomial optimization (NPO) investigates the task of bounding the spectrum of a polynomial in non-commuting operator-valued variables, where said underlying operators are subject to polynomial inequality and equality constraints. As this matches the

natural setting of optimization problems in quantum theory, many applications of NPO have been studied therein. Such problem classes comprise, not exclusively, the computation of the maximal quantum violation of a given Bell expression [1], the ground-state problem of many-body quantum systems [2], as well as fundamental problems in causality [3], and contextuality [4]. Generically, polynomial optimization problems (in commuting or non-commuting variables) are hard to solve exactly, which is why often *relaxations* of these problems are the only tools available. By virtue of relaxations, optimization is performed over a *bigger*, often convex feasible set. Note that relaxation to a *convex* problem is often crucial for its computational tractability, since for these problems a manifold of efficient numerical methods exist. Moreover, they are free from local optima, yielding rigorous lower and upper bounds to minimization and maximization problems, respectively.

The *Navascués-Pironio-Acín* (NPA) hierarchy [5, 6] constitutes a very general relaxation framework based on available convex computational methods that allows one to obtain a monotone series of outer bounds $\{\lambda_d\}_{d=d_{\min}}^{\infty}$ (upper bounds to maximization problems) for a given NPO. The bounds are obtained numerically by solving a series of *Semidefinite Programs* (SDP) of increasing sizes, where under mild conditions this series converges to the global optimum of the problem, $\lambda_{\infty} = \lambda_{\max}$.

This hierarchy has been extensively studied and applied to several emblematic problems in quantum theory. In recent years, several efforts have been made to extend the efficiency,

Younes Naceur: Younes.Naceur@icfo.eu

scalability, and applicability of non-commutative semidefinite relaxations. As in the commuting case, the exploitation of the problem’s sparsity structure has been shown to significantly improve the scalability of the hierarchy [7, 8]. To complement sparsity exploitation, one can further reduce the size of the underlying optimization problem by exploiting the physical symmetries of the underlying problems [9]. To obtain tighter bounds at a given relaxation order, the use of first-order optimality conditions [10, 11] has been explored. Various extensions of the hierarchy, increasing its applicability, include the trace optimization of non-commutative polynomials [12], as well as optimization over polynomials in state expectation values [13]. Recently, the framework was also extended to deal with NPO’s subject to differential constraints [14].

While the computational efficiency of semidefinite relaxations is one of their principal advantages, it also entails a significant limitation. In practice, the underlying semidefinite programs are solved using floating-point arithmetic, and the resulting feasible points satisfy the positivity and equality constraints of the problem only approximately. Consequently, the outer bounds obtained from hierarchies such as NPA, which are often described in the literature as *certifying* that a given objective value cannot exceed the numerical bound λ_d , do *not* constitute formal certificates. As will be illustrated in later sections, the unavoidable rounding errors inherent to numerical solvers can invalidate the positivity or feasibility of the solution. This absence of numerical exactness constitutes an unresolved bottleneck for using semidefinite relaxations as proof systems in quantum theory.

Although obtaining exact, rather than approximate, bounds is an evident step toward mathematical rigor, there are additional, more practical arguments for their necessity: Many optimization protocols, —such as branch-and-bound [15], or constraint-generation schemes [16]— rely on iterative schemes, reusing previously computed SDP bounds as input for subsequent computations. These methods typically rely on the exactness of the input bounds; however, if the input bounds are only approximate, the inference chains often lose validity, undermining their claimed guarantees. Similarly, in witness-based scenarios, such as randomness certification or entanglement de-

tection, a numerical upper bound that is even slightly too low can lead to wrong physical conclusions. Establishing exact, algebraically verifiable SDP bounds is therefore a prerequisite for turning semidefinite relaxations into rigorous tools for certification in quantum information science, bridging the gap between efficient numerical computation and mathematical proof.

To address this issue, we present a three-step post-processing procedure that turns numerical SDP outputs into rigorous, exact certificates for non-commutative polynomial optimization. The method applies to constrained and unconstrained problems, and is extended to sparsity- and symmetry-adapted scenarios. Starting from a numerical Gram-matrix certificate, we (1) round the floating-point solution to nearby rational data, (2) project it (Frobenius-optimally) onto the affine subspace enforcing the certificate equalities, and (3) lift the resulting matrix back into the positive-semidefinite (PSD) cone to restore feasibility. After the rounding step, the subsequent algebraic operations are performed using exact rational arithmetic, thereby avoiding any further round-off error. Additionally, when the projected matrix lies strictly inside the PSD cone, our method can be used to further tighten the numerical bound.

The manuscript is organized as follows: Section 2 introduces the theory of semidefinite relaxations for non-commutative polynomial optimization problems. In Section 3 we present our algorithmic contribution, that is a rigorous method of how to obtain certifiable bounds from numerically inaccurate certificates. Section 4 extends the certification scheme to the sparsity- and symmetry-adapted version of the hierarchy. Sections 5 and 6 are dedicated to our numerical tests. In Section 5, we apply our method to the maximal violation of a family of bipartite Bell inequalities, as well as a variant of the tilted CHSH inequality, and investigate its performance. Section 6 deals with the application of our scheme to certify bounds on quantum many-body ground-state observables. In Section 7 we conclude our work and outline directions for further research.

2 Semidefinite relaxations for non-commutative optimization problems

We start by presenting the polynomial optimization problems analyzed in this work. Given a separable Hilbert space $\mathcal{H}$, let $\mathcal{B}(\mathcal{H})$ be the set of bounded operators acting on $\mathcal{H}$. We consider the problem of finding the maximal eigenvalue $\lambda_{\max}$ of a polynomial $f(\underline{X})$ with rational coefficients in tuples of bounded, self-adjoint operator-valued variables $\underline{X} = \{X_1, \dots, X_n\}$ over a feasible set K given by

$$\{\underline{X} \in \mathcal{B}(\mathcal{H}) \mid h_i(\underline{X}) = 0, g_i(\underline{X}) \succeq 0; h_i \in H, g_i \in G\}, \quad (1)$$

which is characterized by a set of polynomial inequality and equality constraints $G = \{g_i(\underline{X})\}_{i=1, \dots, m_G}$ and $H = \{h_i(\underline{X})\}_{i=1, \dots, m_H}$, respectively. The dimension of the variables $\underline{X}$ is unrestricted, allowing for feasible solutions acting on a Hilbert space $\mathcal{H}$ of arbitrary dimension — a key feature of this framework, enabling the treatment of device-independent scenarios in quantum information science.

Without loss of generality, only maximization problems are considered in this section. The generic optimization problem then reads as

$$\begin{aligned} \sup_{|\psi\rangle, \mathcal{H}, \underline{X} \in \mathcal{B}(\mathcal{H})} & \quad \langle \psi | f(\underline{X}) | \psi \rangle \\ \text{s.t.} & \quad h_i(\underline{X}) = 0, \quad i = 1, \dots, m_H, \\ & \quad g_i(\underline{X}) \succeq 0, \quad i = 1, \dots, m_G. \end{aligned} \quad (2)$$

The optimization can be interpreted as follows: We maximize over all possible (finite and infinite dimensional) Hilbert spaces $\mathcal{H}$, the normalized states therein, $|\psi\rangle \in \mathcal{H}$, as well as the tuples of bounded operators $\underline{X} \in \mathcal{B}(\mathcal{H})$ acting upon it. This is equivalent to optimizing for the maximal eigenvalue of the polynomial $f(\underline{X})$ s.t. $\underline{X} \in K$. Generically, this problem is non-convex and computationally intractable [17], which is why studying a relaxed form of the problem is often necessary.

2.1 Relaxations of polynomial optimization

One starts by recasting the optimization as a problem of non-negativity

$$\begin{aligned} \inf_{\lambda \in \mathbb{R}} & \quad \lambda \\ \text{s.t.} & \quad \lambda 1 - f(\underline{X}) \succeq 0 \quad \forall \underline{X} \in K \end{aligned} \quad (3)$$

which is equivalent to (2) so far: one rephrases the maximization problem (2) as a minimization problem over valid upper bounds, where the optima of the two problems coincide. Above, we denoted the identity operator by 1. In what follows, by a slight abuse of notation, the same symbol will also be used to denote the constant polynomial 1 and the identity matrix. Now, as the task of characterizing the set of non-negative polynomials $\mathcal{P}_K(\underline{X})$ on a basic closed semialgebraic set K is again a computationally intractable task, a version of a central result by Helton and McCullough [18] is used. To do so, we rephrase Problem (3) in the framework of operator theory.

Let $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ denote the underlying field of real or complex numbers, respectively. We then consider all polynomial objects as members of the $*$ -algebra $\mathbb{K}(\underline{X})$ spanned by $\underline{X}$, and equipped with the involution $*$, acting on $\mathbb{K}(\underline{X})$ by reversing the word order of monomials and conjugating their coefficients. Given $r \in \mathbb{K}(\underline{X})$, the polynomial r^*r is called a *Hermitian square*. Sums of the form $\sum_i r_i^* r_i$ are called *Sums-of-Hermitian-Squares* (SOHS), and *Sums-of-Squares* (SOS) when $\mathbb{K} = \mathbb{R}$.

In addition, we define the quadratic module $\mathcal{K}$ associated to the feasible set K as

$$\mathcal{K} = \{p \in \mathbb{K}(\underline{X}) \mid p = \sum_i r_i^* r_i + \sum_{i,j} u_{ij}^* g_i u_{ij} + \sum_{i,j} v_{ij}^* h_i w_{ij}; \quad r_i, u_{ij}, v_{ij}, w_{ij} \in \mathbb{K}(\underline{X})\}. \quad (4)$$

A quadratic module is said to fulfill the *Archimedean* property if and only if there exists some $C > 0$ s.t. $C - \sum_{i=1}^n X_i^2 \in \mathcal{K}$. This property guarantees boundedness of the feasible operator tuples, a crucial aspect for the asymptotic convergence of the hierarchy.

Theorem 1 (Helton–McCullough Positivstellensatz [18]). *Let K be a basic closed semialgebraic set with associated Archimedean quadratic module $\mathcal{K}$. If $p \in \mathbb{K}(\underline{X})$ is positive definite on K , then $p \in \mathcal{K}$.*

Using this result, we rephrase the non-negativity condition as

$$\begin{aligned} \inf_{\lambda \in \mathbb{R}} & \quad \lambda \\ \text{s.t.} & \quad \lambda 1 - f \in \mathcal{K}, \end{aligned} \quad (5)$$

where we passed from explicit operator evaluations $f(X)$ over explicit Hilbert spaces and operator tuples to abstract polynomials f in an algebra, which is why the optimization variables $\mathcal{H}$ and $\underline{X}$ are omitted. The next step consists of relaxing this problem by not considering all polynomials in $\mathcal{K}$, but only those up to a certain degree. To do so, we restrict the summands of $p \in \mathcal{K}$ to have maximal degree $2d$, which defines $\mathcal{K}_{2d}$, the *truncated* quadratic module of degree $2d$. This allows one to define the d -th level of the non-commutative SOHS hierarchy, which is dual to the NPA hierarchy:

$$\lambda_d = \inf_{\lambda \in \mathbb{R}} \lambda \quad \text{s.t.} \quad \lambda 1 - f \in \mathcal{K}_{2d}, \quad (6)$$

where the minimal level for which the relaxation can be run is given by

$$d_{\min} = \lceil \max\{\deg(f), \deg(h_i), \deg(g_i)\}/2 \rceil. \quad (7)$$

Now, obviously, $\mathcal{K}_{2d} \subseteq \mathcal{K}_{2d+2} \subseteq \dots \subseteq \mathcal{P}_{\mathcal{K}}(\underline{X})$, which implies $\lambda_{d_{\min}} \geq \lambda_{d_{\min}+1} \geq \dots \geq \lambda_{\max}$, yielding a series of monotonically decreasing upper bounds λ_d to the original problem.

Most conveniently, each membership certificate of $\mathcal{K}_{2d}$ can be written as a matrix positivity condition. Such optimization problems take the form of an SDP, enabling the use of a well developed toolbox for convex optimization. We note here that, in contrast, checking membership of $\mathcal{K}$ is computationally intractable, as no efficient numerical methods are known for this task.

It has been shown in [5], see also [6, 19], that under the Archimedean condition this series of upper bounds $\{\lambda_d\}_{d_{\min}}^{\infty}$ converges asymptotically, i.e., $\lambda_{\infty} = \lambda_{\max}$. However, there are instances where the relaxations converge at a finite order, i.e., $\exists d \in \mathbb{N} : \lambda_d = \lambda_{\max}$, where the occurrence is tied to a so-called *flatness* or *rank loop* condition on the moment matrix. For a more detailed discussion of the theory of non-commutative semidefinite relaxations, the reader is directed to [20].

Now, many problems naturally occurring in quantum theory exhibit polynomial equality constraints. One can make use of this fact to significantly reduce the size of the underlying SDP by working over the associated quotient space, e.g. comparing equivalence classes up to the constraining equality relations: Let

$\mathcal{I} = \{\sum_i v_i h_i w_i \mid v_i, w_i \in \mathbb{K}(\underline{X}); h_i \in H\}$ denote the polynomial *two-sided ideal* generated by H . Given a monomial order, any polynomial can be reduced to its unique equivalence class modulo the generators of the ideal. The resulting polynomial is called the *normal form* $\mathcal{N}(\cdot)$ of the original one, where the set of normal forms of monomials in $\mathbb{K}(\underline{X})$ forms a canonical basis of the quotient space $\mathbb{K}(\underline{X})/\mathcal{I}$. It is then sufficient to optimize only over said equivalence classes in the degree- d truncation of the quotient algebra, $\mathbb{K}(\underline{X})_d/\mathcal{I}$:

$$\begin{aligned} \lambda_d = \inf_{\lambda \in \mathbb{R}, f_i, u_i \in \mathbb{K}(\underline{X})/\mathcal{I}} \lambda \quad (8) \\ \text{s.t.} \quad \lambda 1 - f = \sum_i f_i^* f_i + \sum_{i,j} u_{ij}^* g_i u_{ij} \mod \mathcal{I}, \\ \deg(f_i) \leq d, \deg(u_{ij}) \leq \left\lfloor \frac{d - \deg(g_i)}{2} \right\rfloor := d_i. \end{aligned}$$

As many problems native to quantum theory exclusively feature equality constraints, we are going to restrict the following illustration to the case of no inequalities, $G = \emptyset$, for now.

To make the SDP underlying (8) more apparent, define a monomial basis v_d of $\mathbb{K}(\underline{X})_d/\mathcal{I}$, with cardinality denoted by $s_{n,d}^{\mathcal{I}}$. The problem can then be phrased as an optimization over PSD *Gram matrices* $G_0 \in \mathbb{K}^{s_{n,d}^{\mathcal{I}} \times s_{n,d}^{\mathcal{I}}}$:

$$\begin{aligned} \lambda_d = \inf_{\lambda \in \mathbb{R}, G_0 \in \mathbb{K}^{s_{n,d}^{\mathcal{I}} \times s_{n,d}^{\mathcal{I}}}} \lambda \quad (9) \\ \text{s.t.} \quad \lambda 1 - f = v_d^* G_0 v_d \mod \mathcal{I}, \\ G_0 \succeq 0, \end{aligned}$$

where the unconstrained case is recovered by working over the trivial ideal $\mathcal{I} = \emptyset$. The optimization can be interpreted as follows: Some of the monomial coefficients of f predetermine entries of G_0 , while others are undetermined. When running the numerical SDP solver, it attempts to find the best pair (λ, G_0) in the intersection of the PSD cone with the affine subspace spanned by the coefficients of $\lambda - f(\underline{X})$.

The size of PSD matrices in the hierarchy grows exponentially in d , where in the unconstrained case it reads as

$$s_{n,d} = \dim(\mathbb{K}(\underline{X})_d) = \frac{n^{d+1}-1}{n-1}, \quad (10)$$

which in general leads to only lower relaxation levels being numerically addressable. For constrained problems, working in the quotient algebra of dimension $\dim(\mathbb{K}(\underline{X})_d/\mathcal{I}) = s_{n,d}^{\mathcal{I}} < s_{n,d}$

often reduces the matrix sizes drastically. As mentioned earlier, improving the scalability of the NPA hierarchy, and SDPs in general is an active field of research, where the interested reader is directed to [21] for a survey on scalability improvements by sparsity exploitation.

It is also possible to study this problem on its dual side, optimizing over the set of *quantum moments* rather than non-negative polynomials [5, 6]. In that setting, the *moment matrix* replaces the Gram matrix as the central PSD variable. Although the dual viewpoint is often used in the literature, we restrict our discussion to the primal side, as it is the more natural setting for algebraic certification. Note also that the works [5, 6] follow the opposite convention when defining the primal and dual problem.

While the above framework can be used to address generic non-commutative optimization problems of the form (2), polynomial optimization does not constitute the exclusive use of SDPs in quantum theory. SDPs naturally occur, not exhaustively, in applications such as entanglement detection through the Doherty-Parrilo-Spedalieri hierarchy [22], the quantum marginal problem [23], tomography problems [24], or quantum network correlations [25]. In many cases, the SDP structure of the problem arises from the PSD nature of the underlying quantum states or measurements.

3 Extracting certified bounds

As anticipated in Section 1, SDP solvers aiming to certify non-negativity of $\lambda - f$ via the numerical tuple (λ, G_0) might issue inaccurate solutions due to several reasons, such as inexact termination, failure of strict feasibility, or poor conditioning of the problem. Although in theory SDPs are exactly solvable by quantifier elimination [26], this method has doubly exponential complexity in the number of variables [27]. In contrast, if only approximate solutions are required, SDPs can be solved in polynomial time using convex optimization methods such as the ellipsoid or interior-point algorithms [28]. Since scalability is crucial for many problems in quantum theory, this work focuses on post-processing of the numerical data rather than aiming for exact solutions of the underlying SDPs.

Various works discuss the appropriate handling

of obtaining exact Sum-of-Squares (SOS) decompositions from inexact data. For univariate polynomials with real or complex coefficients, there are dedicated algorithms approximating complex roots of perturbed positive polynomials [29, 30], or solving norm equations together with using Euler's identity [31]. In the multivariate case, one can either rely on the so-called *rounding and projection* [32] (see also [33]) or perturbation-compensation algorithms [34], implemented in the RealCertify package [35]. For constrained compact problems, one can apply rigorous interval arithmetic techniques to obtain a certified upper bound [36]. Similarly to that setting, tailored methods to obtain certified bounds to SDP relaxations to the optimal power-flow problems have been presented in [37]. Related computer-assisted proof techniques have also been used in quantum information to certify finite-dimensional SDP bounds for specific problem instances, for example in channel-discrimination problems, by projecting numerical SDP solutions onto affine matrix constraints and restoring positivity [38]. In contrast, our method aims to extract certified solutions to a wider range of problems, encompassing any problem following the general NPO formulation (2). Other algorithms have been designed for specific problems with finitely many minimizers [39, 40]. Such methods provide exact, rational SOS decompositions, together with a certified rational upper bound of the problem solution. If one is interested in certifying global optimality, one can rely on the recently developed high-precision solver [41] that provides exact SOS decompositions, possibly with non-rational coefficients. On the other hand, if one primarily seeks highly accurate numerical bounds rather than exact certificates, one may use arbitrary-precision SDP solvers (e.g., SDPA-GMP [42]) at the expense of a larger computational cost. However, the output remains a numerical approximation and does not constitute a proof unless further validated by an adequate certification method. The rounding and projection algorithm also has been extended to unconstrained non-commutative problems in [43] to obtain exact rational SOHS decompositions. Our main algorithmic contribution is based on this rounding and projection scheme, in order to handle constraints for maximal eigenvalue problems, as well as sparse/symmetric input data.

Our method offers a complementary guarantee in situations where rigorous bounds to a given problem are of primary interest, independently of the convergence of the original problem.

We now present our constructive procedure that converts an approximate numerical SDP solution of Problem (9) into a formally exact rational certificate. Starting with the case of an unconstrained NPO ($G = H = \mathcal{I} = \emptyset$) let (λ_d, G_0) be the numerical tuple of upper bound $\lambda_d \in \mathbb{R}$ and Gram matrix $G_0 \in \mathbb{K}^{s_{n,d} \times s_{n,d}}$, as issued by the SDP solver of one's choice. The left hand side (LHS) L_d of the numerical certificate at order d , suggesting $\lambda_d > \lambda_{\max}$ then reads

$$L_d := \lambda_d 1 - f \approx v_d^* G_0 v_d. \quad (11)$$

The method to obtain an exact certificate from (11) consists of three algebraic steps: (1) rounding of the floating point data to rationally represented data in $\mathbb{Q}$ with controlled denominator, (2) projection, and (3) lifting, each preserving or restoring feasibility within the truncated quadratic module $\mathcal{K}_{2d}$. After the rounding step, the algebraic operations are performed using exact rational arithmetic, so that no additional round-off error is introduced. A geometric picture of the procedure is displayed in Figure 1. After the first two steps, rounding and projection, the solver's tuple (λ_d, G_0) has been mapped onto the affine subspace $\mathcal{L}$ consistent with the coefficients of $\lambda_d 1 - f$, thereby enforcing exact equality in Eq. (11). We call the obtained tuple $(\mathcal{P}(\tilde{G}_0), \tilde{\lambda}_d)$ a *pre-certificate*. In the case that $\mathcal{P}(\tilde{G}_0) \succeq 0$, we established a proper SOHS certificate, proving $\tilde{\lambda}_d$ as an exact upper bound. If the resulting matrix $\mathcal{P}(\tilde{G}_0)$ is not PSD, in step (3) we lift it back into the PSD cone by adding a rational correction justified by the constraint structure. This lifted matrix $\mathcal{P}_+(\tilde{G}_0)$ then certifies a *weaker* bound $\lambda_d^{\text{rat}} > \tilde{\lambda}_d$, which holds exactly. That is, step (3) is only necessary if $\mathcal{P}(\tilde{G}_0)$ is not PSD. However, in some situations, one can still apply step (3) when $\mathcal{P}(\tilde{G}_0) \succeq 0$ to further tighten the bound $\tilde{\lambda}_d$, as it is shown later.

In the rest of the section, we provide a rigorous formulation of the three steps in the process, both for problems with and without equalities.

3.1 Round and Project

To obtain a nearby rational tuple $(\mathcal{P}(\tilde{G}_0), \tilde{\lambda}_d)$ from their floating-point counterparts in the un-

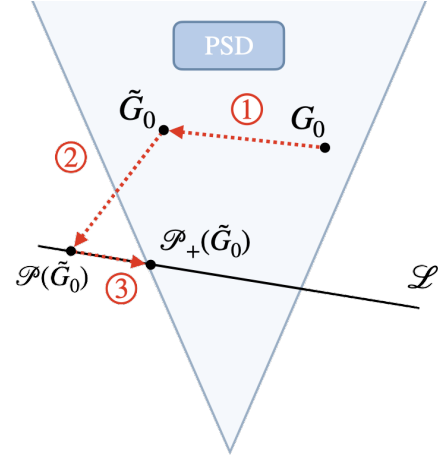


Figure 1: The geometric picture of our procedure. (1) G_0 in floating-point representation is rounded to a rational $\tilde{G}_0$. (2) $\tilde{G}_0$ is projected onto $\mathcal{L}$ in exact arithmetic, possibly leaving the PSD cone. (3) $\mathcal{P}(\tilde{G}_0)$ is lifted back into the PSD cone, certifying λ_d^{rat} as an upper bound.

constrained case, one can make use of the *Round and Project* procedure from [43]: Fix a given rounding precision η and start by rounding the floating-point data (λ_d, G_0) entry-wise to rational representations $(\tilde{\lambda}_d, \tilde{G}_0)$, where established methods such as *continued fractions* can be used. This rounding step should be understood as a reconstruction of adequate rational representations of the floating-point numbers: every floating-point number already represents an exact rational number, and one could in principle use this exact representation directly in all subsequent exact computations. We instead reconstruct a nearby rational number of controlled precision in order to suppress insignificant numerical noise and to avoid unnecessarily large numerators and denominators in the subsequent exact rational arithmetic operations.

Once the reconstructed data $(\tilde{\lambda}_d, \tilde{G}_0)$ have been chosen, the subsequent algebraic projection and lifting operations are performed exactly over the rational coefficient field to not introduce any additional round-off error. Using exact arithmetic instead of floating-point arithmetic is hence crucial for the algebraic manipulation to preserve the exactness of the certificate in intermediate manipulation steps. In the rest of the work, $\tilde{A}$ represents the entrywise rationalized form of the object A , where rationalization on polynomials is understood as acting on their coefficient vector in the monomial basis. We denote by $\mathbb{Q}$ the rational subfield of $\mathbb{K}$ and by $\mathbb{Q}\langle X \rangle$ the associated rational

polynomial algebra (real or complex).

Now, as mentioned earlier, since f is assumed to have rational coefficients, it constrains the entries of G_0 corresponding to its support. To recover this relationship after rounding, one has to project the entries of $\tilde{G}_0$ back into the affine space $\mathcal{L}$ defined through the coefficients of $\tilde{L}_d := \tilde{\lambda}_d 1 - f$ via [43]:

$$\mathcal{P}(\tilde{G}_0)_{\alpha,\beta} = (\tilde{G}_0)_{\alpha,\beta} - \frac{1}{n(\alpha,\beta)} \sum_{\gamma^*\delta \sim \alpha^*\beta} (\tilde{G}_0)_{\gamma,\delta} - (\tilde{L}_d)_{\alpha^*\beta}, \quad (12)$$

which can be expressed compactly with a rational correction matrix Δ as $\mathcal{P}(\tilde{G}_0) = \tilde{G}_0 + \Delta$. The tuple (α, β) refers to the indexing with respect to monomials $\alpha, \beta \in \mathbb{K}\langle X \rangle_d$, and $p_{\alpha^*\beta}$ refers to the monomial coefficient of $\alpha^*\beta$ in p . Intuitively, this projection enforces exact algebraic consistency between the rounded Gram matrix and the target polynomial by averaging over equivalent monomial pairs. Here, $n(\alpha, \beta)$ accounts for the number of such pairs:

$$n(\alpha, \beta) = \{(\gamma, \delta) \mid \gamma^*\delta = \alpha^*\beta\}. \quad (13)$$

When considering problems with equality constraints, this projection can be adapted by working over the quotient algebra $\mathbb{K}\langle X \rangle/\mathcal{I}$ by enforcing coefficient equality of *normal forms* $\mathcal{N}(\alpha^*\beta)$ of monomials $\alpha^*\beta$, i.e., their equivalence classes in the quotient algebra. The goal is again to find a correction matrix Δ such that the corrected certificate holds exactly up to normal form reductions

$$\mathcal{N}(\tilde{L}_d) = \mathcal{N}\left(v_d^*(\tilde{G}_0 + \Delta)v_d\right). \quad (14)$$

As several entries (α, β) map to the same normal form, there are in general infinitely many ways of constructing Δ . The most natural choice is the correction with the least Frobenius deviation from the original rationalized Gram matrix: As the numerical Gram matrices are PSD up to the solver precision, deviating the least possible from them will minimize the distance to the PSD cone after projection. The next Lemma provides the formal construction of this projection.

Lemma 2 (Frobenius optimal Gram projection). *For a given optimization problem (9), let $\alpha^*\beta$ be the monomials corresponding to the basis elements (α, β) of the Gram matrix G_0 . By construction, they form a basis of $\mathbb{K}\langle X \rangle_{2d}$. Let t index the basis of equivalence classes in $\mathbb{K}\langle X \rangle_{2d}/\mathcal{I}$.*

The associated normal form operation can then be expressed as

$$\mathcal{N} : \mathbb{K}\langle X \rangle_{2d} \rightarrow \mathbb{K}\langle X \rangle_{2d}/\mathcal{I} \quad (15)$$

$$\alpha^*\beta \mapsto \sum_t n_{\alpha,\beta}^t t$$

where $n_{\alpha,\beta}^t$ is the coefficient with which the word $\alpha^*\beta$ reduces to t under $\mathcal{N}$. Define the rational numerical certificate residual coefficients r_t per reduced word t as

$$r_t = \mathcal{N}(\tilde{L}_d)_t - \mathcal{N}\left(\sum_{\alpha,\beta} v_{d,\alpha}^*(\tilde{G}_0)_{\alpha,\beta} v_{d,\beta}\right)_t$$

$$= \mathcal{N}(\tilde{L}_d)_t - \sum_{\alpha,\beta} n_{\alpha,\beta}^t (\tilde{G}_0)_{\alpha,\beta}. \quad (16)$$

Let Ξ be the matrix with entries

$$(\Xi)_{t,s} := \sum_{\alpha,\beta} n_{\alpha,\beta}^t n_{\alpha,\beta}^s. \quad (17)$$

If Ξ is non-singular, the following projection $\mathcal{P}(\tilde{G})$ yields exact equality between the left hand side (LHS) and right hand side (RHS) of the SOHS certificate:

$$\mathcal{P}(\tilde{G}_0)_{\alpha,\beta} = (\tilde{G}_0)_{\alpha,\beta} + \Delta_{\alpha,\beta}$$

$$= (\tilde{G}_0)_{\alpha,\beta} + \sum_t (\Xi^{-1}r)_t n_{\alpha,\beta}^t. \quad (18)$$

Furthermore, it is the optimal projection $\mathcal{P}$ in terms of Frobenius deviation

$$\|\mathcal{P}(\tilde{G}_0) - \tilde{G}_0\|_F^2 = \sum_{\alpha,\beta} |(\mathcal{P}(\tilde{G}_0) - \tilde{G}_0)_{\alpha,\beta}|^2 \quad (19)$$

from the rounded numerical Gram matrix $\tilde{G}_0$.

Proof of Lemma 2. Let the assumptions in Lemma 2 hold. We phrase the problem of finding the correction Δ with least Frobenius norm as a convex quadratic optimization problem:

$$\inf_{\Delta \in \mathbb{Q}^{s_{n,d}^{\mathcal{I}} \times s_{n,d}^{\mathcal{I}}}} \frac{1}{2} \|\Delta\|_F^2 \quad (20)$$

$$\text{s.t.} \quad r_t = \sum_{\alpha,\beta} n_{\alpha,\beta}^t \Delta_{\alpha,\beta}, \quad \forall t,$$

which can be solved using Lagrangian multipliers λ_t . Stationarity of the Lagrangian yields

$$\Delta_{\alpha,\beta} = \sum_t \lambda_t n_{\alpha,\beta}^t. \quad (21)$$

Plugging back into the constraints yields

$$r_t = \Xi \lambda_t. \quad (22)$$

If Ξ is non-singular, using (21) and (22) one finds the global optimizer

$$\Delta_{\alpha,\beta} = \sum_t (\Xi^{-1} r)_t n_{\alpha,\beta}^t. \quad (23)$$

□

Now, in many settings, such as Bell inequalities or ground states of quantum many body systems, the equality constraints defining the problem relate two monomials to each other. Hence, the associated normal forms map monomials onto monomials instead of polynomials. We call such normal forms and the associated ideals *binomial*. In such cases, one finds a diagonal $\Xi \succ 0$ and the Frobenius optimal solution always exists:

Corollary 3. *If $\mathcal{N}$ is binomial, the Frobenius ideal projection reduces to*

$$\mathcal{P}(\tilde{G})_{\alpha,\beta} = \tilde{G}_{\alpha,\beta} + \frac{r_{\mathcal{N}(\alpha^*\beta)}}{n_{\mathcal{I}}(\alpha,\beta)} \quad (24)$$

with

$$n_{\mathcal{I}}(\alpha,\beta) = \#\{(\gamma,\delta) : \mathcal{N}(\alpha^*\beta) = \mathcal{N}(\gamma^*\delta)\}. \quad (25)$$

Proof of Corollary 3. Binomial normal forms can be expressed as

$$\mathcal{N} : \alpha^*\beta \mapsto \delta_{\mathcal{N}(\alpha^*\beta),t} t \quad (26)$$

which allows us to rewrite the residuals r_t as

$$r_t = \mathcal{N}(\tilde{L}_d)_t - \sum_{\alpha,\beta:\mathcal{N}(\alpha^*\beta)=t} \tilde{G}_{\alpha,\beta}. \quad (27)$$

Applying the method of Lagrangian multipliers yields

$$\Delta_{\alpha,\beta} = \lambda_t \delta_{\mathcal{N}(\alpha^*\beta),t}. \quad (28)$$

Plugging this into the constraints yields

$$r_t = \sum_{\alpha,\beta:\mathcal{N}(\alpha^*\beta)=t} \lambda_t \Rightarrow \lambda_t = \frac{r_t}{n_{\mathcal{I}}(\alpha,\beta)}, \quad (29)$$

which, when used with 28, implies the claim. □

As (12) is implied by (24), in what follows we are going to restrict our considerations to the quotient case without loss of generality.

The projection formula is also readily adapted to problems featuring inequality constraints. Recall that in that case, the numerical certificate reads

$$\lambda 1 - f = \sum_i f_i^* f_i + \sum_{i,j} u_{ij}^* g_i u_{ij} \mod \mathcal{I}, \quad (30)$$

which, when expressed through Gram matrices, features one matrix G_0 corresponding to the SOHS $\sum_i f_i^* f_i$, and m_G matrices $\{G_i\}_{i=1}^{m_G}$ of sizes $s_{n,d_i}^{\mathcal{I}}$ corresponding to the so-called localizing terms $\sum_{i,j} u_{ij}^* g_i u_{ij}$. To deal with such constraints, one adapts the LHS from Equation (11) to feature the inequality terms, and projects the rounded principal Gram matrix $\tilde{G}_0$ onto the rationalized LHS as before to match coefficients.

To do so, start by constructing the rational LHS $\tilde{L}_d$ featuring the inequality terms:

$$\tilde{L}_d := \tilde{\lambda}_d 1 - f - \sum_i \sum_j \tilde{u}_{ij}^* g_i \tilde{u}_{i,j}. \quad (31)$$

To obtain the rational polynomials $\tilde{u}_{ij}$ in the decomposition, one starts by rounding the floating-point localizing Gram matrices $\{G_i\}_{i=1}^{m_G}$ to rationals $\{\tilde{G}_i\}_{i=1}^{m_G}$. Afterwards, to obtain a nearby rational PSD Gram matrix, we numerically project $\tilde{G}_i$ onto the PSD cone via an eigendecomposition and rationally reconstruct the resulting factors:

$$\mathcal{P}(\tilde{G}_i) = \tilde{U}_i \tilde{U}_i^T; \quad \tilde{U}_i = \tilde{V}_i \text{diag} \left(\max(\sqrt{\tilde{\lambda}_{ij}}, 0) \right) \quad (32)$$

with $\tilde{V}$ being the matrix of rationalized eigenvectors of $\tilde{G}_i$ and $\{\lambda_{ij}\}_{j=1,\dots,s_{n,d_i}}$ the corresponding numerical eigenvalues. From these, the polynomials $\tilde{u}_{ij}$ are obtained as

$$\tilde{u}_{ij} = (\tilde{U}_i^T v_{di})_j, \quad (33)$$

where v_{di} denotes the monomial basis appropriate for the localizing matrix of g_i . Using Lemma 2 to project $\tilde{G}_0$ onto $\tilde{L}_d$ from Equation (31) yields a pre-certificate $\tilde{L}_d = v_d^* \mathcal{P}(\tilde{G}_0) v_d$, and with $\mathcal{P}(\tilde{G}_0) \succeq 0$ it would certify membership of $\tilde{\lambda}_d 1 - f$ in $\mathcal{K}_{2d}$ and hence $\tilde{\lambda}_d$ as an upper bound to f . This procedure also applies to moment inequality constraints of the form $\langle m_i \rangle \geq 0$, where the associated multiplier polynomials $\tilde{u}_{ij}$ are scalar valued (degree 0).

Performing the Round and Project procedure, in the unconstrained or constrained case, results in an exact pre-certificate

$$\tilde{L}_d = v_d^* \mathcal{P}(\tilde{G}_0) v_d \mod \mathcal{I}, \quad (34)$$

where this expression constitutes an exact algebraic identity, ensuring that any remaining infeasibility stems solely from the non-positivity of $\mathcal{P}(\tilde{G}_0)$.

As anticipated, in the best case, the resulting projected Gram matrix $\mathcal{P}(\tilde{G}_0)$ is already PSD. This can only be the case if the rounded bound $\tilde{\lambda}_d$ is a valid upper bound to the global optimum, $\tilde{\lambda}_d \geq \lambda_{\max}$, and hence leaves us with the certified bound $\tilde{\lambda}_d$.

3.2 Lifting positivity

The remaining question is how to repair a pre-certificate (34) whose Gram matrix was projected out of the PSD cone, rendering it vacuous. This particularly happens if $\tilde{\lambda}_d \not\geq \lambda_{\max}$. The next Theorem provides, under reasonably mild assumptions, an explicit, constraint-dependent lifting that guarantees a valid upper bound.

Theorem 4 (Certified bounds to NPOs). *Let $\mathcal{P}(\tilde{G}_0)$ be the rounded and projected Gram matrix obtained from Lemma 2 with $\mu_{\min} \in \mathbb{Q}$ a lower bound to its spectrum, and let $s_{n,d}^{\mathcal{I}}$ denote the number of monomials in n variables up to degree d after normal form reduction. Assume that the underlying variables $\underline{X}$ are subject to one of the following polynomial constraints:*

- *Ball constraints:* $1 \succeq \sum_i X_i^2$,
- *Box constraints:* $1 - X_i^2 \succeq 0 \ \forall i$,
- *Unipotence:* $1 - X_i^2 = 0 \ \forall i$,
- *Projection constraints:* $X_i^2 = X_i \ \forall i$,

then, the rational bound $\lambda_d^{\text{rat}} \in \mathbb{Q}$ with

$$\lambda_d^{\text{rat}} := \tilde{\lambda}_d + \delta_d; \quad \delta_d := -\min\{\mu_{\min}, 0\} \cdot s_{n,d}^{\mathcal{I}} \quad (35)$$

constitutes a certified upper bound to the optimization problem, i.e., $\lambda_d^{\text{rat}} \geq \lambda_{\max}$.

Importantly, the condition on the existence of constraints is rather weak: In the non-locality scenario, Bell type expressions can be defined in terms of projectors, while in the quantum many-body setting, the underlying local spin operators naturally fulfill unipotence.

The geometric intuition behind the proof of Theorem 4 is that the non-PSD Gram matrix $\mathcal{P}(\tilde{G}_0)$ is lifted by adding a multiple of the identity to the certificate, leading to a weaker bound

λ^{rat} . The allowed correction magnitude is guaranteed to be feasible within the quadratic module due to the following Lemma.

Lemma 5 (Constant SOHS decomposition). *Let $\varepsilon > 0$. If the variables $\underline{X}$ are subject to constraints as in Theorem 4, one can find the following exact, constant SOHS decomposition*

$$\varepsilon s_{n,d}^{\mathcal{I}} 1 = \varepsilon v_d^* v_d + \sum_k r_k^* r_k + \sum_i q_{ki}^* c_i q_{ki} \quad \text{mod } \mathcal{I} \quad (36)$$

for some $r_k, q_{ki} \in \mathbb{Q}\langle \underline{X} \rangle_d$ and v_d the monomial basis of $\mathbb{K}\langle \underline{X} \rangle_d / \mathcal{I}$.

Lemma 5 can be proven by induction, as shown in Appendix A. Using this, we can prove Theorem 4:

Proof of Theorem 4. Let the assumptions in Theorem 4 hold. We are going to show that for λ^{rat} there exists some PSD gram matrix $\mathcal{P}_+(\tilde{G}_0)$ and polynomials p_{ij} such that

$$\begin{aligned} \lambda^{\text{rat}} 1 - f &= v_d^* \left(\mathcal{P}_+(\tilde{G}_0) \right) v_d + \sum_{i,j} p_{ij}^* g_i p_{ij} \quad \text{mod } \mathcal{I} \\ &\Leftrightarrow \lambda^{\text{rat}} 1 - f \in \mathcal{K}_{2d}, \end{aligned} \quad (37)$$

which implies the claim.

Consider the exact pre-certificate (34). If $\mu_{\min} < 0$ then adding (36) with $\varepsilon = -\mu_{\min}$ to both sides of the certificate lifts $\mathcal{P}(\tilde{G}_0)$ back into the PSD cone as follows:

$$\begin{aligned} \tilde{\lambda}_d 1 - \mu_{\min} s_{n,d}^{\mathcal{I}} 1 - f &= \sum_{i,j} \tilde{u}_{ij}^* g_i \tilde{u}_{ij} \quad (38) \\ &= v_d^* \left(\mathcal{P}(\tilde{G}_0) - \mu_{\min} 1 \right) v_d + \sum_k r_k^* r_k + \sum_{i,k} q_{ki}^* c_i q_{ki}. \end{aligned}$$

Intuitively, this addition corresponds to lifting the spectrum of $\mathcal{P}(\tilde{G}_0)$ by $-\mu_{\min} > 0$ while staying within the quadratic module, thereby restoring positivity without changing the algebraic identity of the certificate.

Expressing the r_k contributions with a PSD matrix R_0 of the same size, and joining terms corresponding to constraints c_i yields

$$\begin{aligned} \tilde{\lambda}_d 1 - \mu_{\min} s_{n,d}^{\mathcal{I}} 1 - f &= v_d^* \left(\mathcal{P}(\tilde{G}_0) - \mu_{\min} 1 + R_0 \right) v_d \\ &\quad + \sum_{i,j} p_{ij}^* g_i p_{ij}, \end{aligned} \quad (39)$$

which implies the claim with

$$\mathcal{P}_+(\tilde{G}_0) = \left(\mathcal{P}(\tilde{G}_0) - \mu_{\min} 1 + R_0 \right) \quad (40)$$

depending on the constraints involved. $\square$

It should be noted that a numerical bound may have been correct even if the corresponding Gram matrix G_0 is projected outside the PSD cone. Consequently, our method allows one to (a) recover valid bounds from possibly incorrect ones, and (b) validate correct bounds, but it does not provide a necessary certificate of invalidity for numerical bounds.

The above argument only holds if $\mu_{\min}$ actually constitutes a rigorous lower bound to the spectrum of $\mathcal{P}(\tilde{G}_0)$. Unfortunately, the Abel-Ruffini theorem implies no closed form solution for the eigenvalues of generic square matrices of dimension larger than four, as they are defined as the roots of the associated characteristic polynomial. A rigorous lower bound on the smallest eigenvalue can instead be obtained using certified interval arithmetic. In our implementation, `ArbLib.jl` [44] performs arbitrary-precision floating-point ball arithmetic, yielding an interval guaranteed to contain a given eigenvalue. The certified lower endpoint of this enclosure constitutes a certified lower bound to the eigenspectrum. Since this endpoint is a finite binary floating-point number, we use its exact rational representation in the subsequent algebraic computations. As the floating-point bound is already certified and holds exactly, its exact rational representation can be used while maintaining validity throughout. We emphasize that the working precision of `ArbLib` controls the tightness of the eigenvalue enclosure, but not the validity of the certification. Throughout this work, all eigenvalue bounds entering the certification procedure are understood in this rigorous sense.

Theorem 4 establishes that any numerical non-negativity certificate returned by an SDP solver can be transformed into an exact rational bound by a finite algebraic correction depending on the minimal eigenvalue defect and the problem size. An important observation is that the bound correction term $\delta_d = -\min\{\mu_{\min}, 0\} \cdot s_{n,d}^T$ is affected by two factors, (i) the distance from the projected $\mathcal{P}(\tilde{G}_0)$ to the PSD cone, and (ii) the size of the problem in terms of variables and relaxation order. While working with high precision solvers and rounding arithmetic can lead to smaller distances to the PSD cone, the term $s_{n,d}^T$ ultimately dominates the bound correction, as it will be illustrated in later sections.

3.3 Bound tightening of interior pre-certificates

Now, in the case of the projected Gram matrix being positive-definite (PD), e.g. $\mathcal{P}(\tilde{G}_0) \succ 0$, it is not necessary to perform the lifting step to restore feasibility. However, it is possible to leverage the positive eigenspectrum to obtain a *tighter* bound $\lambda_d^{\text{rat}} < \tilde{\lambda}_d$ by shifting the PSD matrix closer to the boundary of its cone. This can be done in two ways, (1) in the presence of unipotency constraints $X_i^2 = 1$, one can use a modification of Lemma 5 with $\varepsilon < 0$, while in the other cases one can make use of a rank-1 PSD downgrade, as explained in the next Lemma.

Theorem 6 (Certified tightening for interior pre-certificates). *Assume that the Frobenius optimal projection 2 leads to $\mathcal{P}(\tilde{G}_0) \succ 0$ lying in the interior of the PSD cone. Let $\mu_{\min} > 0$ be a certified lower bound to the spectrum, and let*

$$\tau = \frac{1}{e_0^* \mathcal{P}(\tilde{G}_0)^{-1} e_0} \quad (41)$$

where $e_0 \in \mathbb{K}^{s_{n,d}^T}$ denotes the coordinate vector selecting the constant monomial s.t. $e_0^* v_d = 1$. Then, a tighter bound $\lambda_d^{\text{rat}} < \tilde{\lambda}_d$ can be obtained depending on the constraints involved:

- (i) *If the variables are subject to unipotency constraints, $\{1 - X_i^2\} \subseteq H$, then $\lambda_d^{\text{rat}} = \tilde{\lambda}_d - \mu_{\min} \cdot s_{n,d}^T$ constitutes a rigorous upper bound to the optimization problem*
- (ii) *Independently of the constraints, $\lambda_d^{\text{rat}} = \tilde{\lambda}_d - \tau$ constitutes a rigorous upper bound to the optimization problem, where $\tau \in [\mu_{\min}, \mu_{\max}]$ with $\mu_{\max}$ an upper bound to the spectrum of $\mathcal{P}(\tilde{G}_0)$.*

The proof for this Theorem can be found in Appendix B. Applying this Theorem in the case of $\mathcal{P}(\tilde{G}_0) \succ 0$ lets one further tighten the numerical rounded bounds while certifying them. Note that as the numerical optima of the SDP lie on the boundary of the PSD cone, in the generic case the eigenspectrum of $\mathcal{P}(\tilde{G}_0)$ is not positive as the Frobenius distance between $\mathcal{P}(\tilde{G}_0)$ and $\tilde{G}_0$ is minimized. However, in numerically delicate instances where PSD solvers terminate in the interior of the cone, the above scenario does occur, as it is observed and interpreted in Section 6. Also note that in the case of unipotency constraints

one can apply either of the two tightening strategies, while procedure (i) carries an extra factor $s_{n,d}^T$ which, for large problems, will typically result in a tighter bound than procedure (ii).

Concludingly, we established a rigorous method to obtain mathematical non-negativity proof from the solver's data and presented an adapted method for interior solutions where the algorithm provides certified bounds tighter than their numerical counterpart.

The same above reasoning of rounding, projecting, and lifting, can be extended to structured relaxations such as sparsity or symmetry-adapted hierarchies, as presented in the next section.

4 Exact certificates under sparsity and symmetry

There exist problem for which it is necessary to go to higher orders of the hierarchy to get tight bounds to the solution of the optimization problem. In those cases, the exponential growth of the PSD matrices with the relaxation order often makes higher intractable in practice, especially if a large number of variables is involved. This bottleneck of semidefinite relaxations is of particular importance, for instance, in quantum many-body physics, as one is often interested in the behavior of systems of large sizes. One way to be able to address such problems, is to employ a sparse version of the hierarchy, which relies on a sparse analogue of Theorem 1. Applying the sparse instead of the standard, dense hierarchy has led to improvements in scalability by orders of magnitude, if the problem admits certain structure [2, 8]. In other cases, the symmetry of the problem can be exploited to achieve a block-diagonal form of the Gram matrix and thus reducing computational complexity. In this section, we adapt the procedure described above to get rigorous bounds for systems that are sparse or symmetric. First, we will focus on the sparse case.

4.1 Hierarchies for sparse systems

We start by giving a brief introduction to correlative sparsity and the associated SOHS hierarchy. The idea is that one can exploit a sparse *correlation pattern* of the variables, to then construct several, smaller Gram (or moment) matrices, indexed by *cliques* of variables that correlate

among each other. This is in stark contrast to the dense hierarchy, where a single PSD block is indexed by monomials between all variables in the problem. Employing the sparse hierarchy leads to more PSD conditions of smaller sizes instead of one larger one. As the dominant computational cost for SDP solvers arises from the largest block in the PSD conditions, employing methods to reduce its dimension is crucial to improve the efficiency of SDP hierarchies.

To start, build the correlation graph $C = \{V, E\}$ with vertices $V = \underline{X}$ given by the operators in the problem, and edges E between them whenever they occur together multiplicatively in the support of the objective f , or additively in the constraints in G , and H .

The next step to obtain the hierarchy consists in preprocessing the correlation graph C : A *chordal graph* is a graph in which every cycle of length $k \geq 4$ contains a chord, i.e., an edge joining two non-adjacent vertices of the cycle. An important fact is that every non-chordal graph can be extended to be chordal by adding edges to E , which is called a *chordal extension*. Note that there is often more than one way to render a given correlation graph chordal.

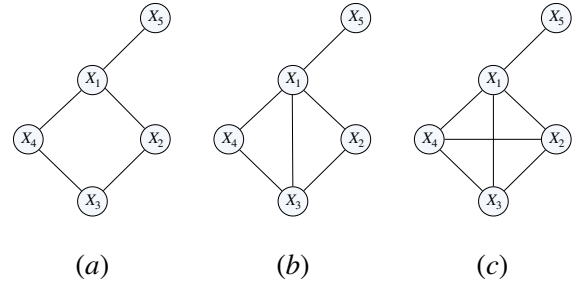


Figure 2: (a) The correlation graph C for $f(\underline{X}) = X_1X_5 + X_1X_2 + X_2X_3 + X_3X_4 + X_4X_1$. (b) A minimal chordal extension of C . (c) Another chordal extension of C .

After performing the chordal extension, one identifies the so-called *maximal cliques* in the graph: A clique is a subset $C_k \subset V$ such that every pair $X_i, X_j \in C_k$ satisfies $(i, j) \in E$, which is called maximal if no other clique C'_k with $C_k \subsetneq C'_k$ exists. As an example, in Figure 2 (b), the maximal cliques can be identified as $C_1 = \{X_1, X_2, X_3\}$, $C_2 = \{X_1, X_3, X_4\}$, and $C_3 = \{X_1, X_5\}$. When opting for the chordal extension (c) instead, one can identify the two cliques $C_1 = \{X_1, X_2, X_3, X_4\}$ and $C_2 = \{X_1, X_5\}$. We

note here that, by construction of the correlation graph, each constraint features variables from a single clique. Hence, we define $J_k \subseteq [m_G]$ as the set of indices i of inequality constraints contained in C_k . Furthermore, let N_C denote the number of maximal cliques, $n_k = |C_k|$ the cardinality of the clique, and $\mathbb{K}\langle C_k \rangle$ the $*$ -algebra spanned by C_k .

Having constructed the chordal extension and identified the maximal cliques, one can define the d -th relaxation level of the sparse SOHS hierarchy:

$$\begin{aligned} \lambda_d^{\text{sp}} = & \inf_{\lambda \in \mathbb{R}, f_{k,\ell}, u_{ij}^{(k)} \in \mathbb{K}\langle C_k \rangle / \mathcal{I}} \lambda \\ \text{s.t. } & \lambda \mathbf{1} - f = \\ & \sum_{k=1}^{N_C} \left(\sum_{\ell} f_{k,\ell}^* f_{k,\ell} + \sum_{i \in J_k} \sum_j (u_{ij}^{(k)})^* g_i u_{ij}^{(k)} \right) \bmod \mathcal{I} \\ & \deg(f_{k,\ell}) \leq d, \deg(u_{ij}^{(k)}) \leq \left\lfloor \frac{d - \deg(g_i)}{2} \right\rfloor. \end{aligned} \quad (42)$$

Similarly to the dense hierarchy 8, the first condition characterizes membership in the associated *sparse truncated quadratic module*, which according to the sparse version [7] of Theorem 1 certifies positivity on K . The main difference to the dense case is that the SOHS splits over contributions per clique, i.e., summands $f_{k,\ell}$ and $u_{ij}^{(k)}$ stemming from the clique-wise algebras $\mathbb{K}\langle C_k \rangle / \mathcal{I}$. Hence, when expressing the SOHS in terms of PSD Gram matrices, one is left with N_C principal Gram matrices $G_0^{(k)}$ indexed by the variables in C_k , as well as m_G localizing-type Gram matrices associated to the $u_{ij}^{(k)}$, equivalently indexed by variables from C_k . This manifests the anticipated reduction in the maximal PSD block size by employing the sparse hierarchy.

As mentioned, for a given correlations graph, there are different ways of constructing chordal extensions, which eventually introduces a trade-off between computational complexity (smaller cliques) and tightness of the relaxation (larger cliques). Depending on the chosen chordal extension of the correlation graph, one obtains PSD constraints on larger or smaller blocks, where the dense hierarchy 8 is recovered in the case of a single clique $C_1 = \underline{X}$. This introduces a second dimension to the hierarchy, i.e., the degree of the chordal extension, where more complete graphs are associated to larger computational complexity while providing tighter bounds. Intuitively,

allowing for SOHS summands from a bigger space might let one find more elaborate expressions to certify a tighter bound. It has been shown in [7] that in the case of a chordal extension, one recovers the dense asymptotic convergence results, $\lambda_\infty^{\text{sp}} = \lambda_{\max}$. However, at each finite relaxation order, the dense bounds lower bound the sparse bounds, $\lambda_d \leq \lambda_d^{\text{sp}}$.

4.2 Exact certificates for sparse systems

Now that we have introduced the numerical sparse SOHS certificates, we are going to establish a rigorous way on how to extract a certified upper bound from the corresponding numerical data. The rounding–projection–lifting strategy developed for the dense hierarchy extends naturally to such relaxations which exploit correlative sparsity. Let the sparse numerical certificate be denoted by the tuple $(\lambda_d^{\text{sp}}, \{G_0^{(k)}\}_{k=1}^{N_C}, \{G_i^{(k)}\}_{i=1}^{m_G})$.

The numerical certificate (possibly over a quotient space) then reads as

$$L_d \approx \sum_k \left(v_d^{(k)} \right)^* G_0^{(k)} v_d^{(k)} \bmod \mathcal{I}, \quad (43)$$

where $v_d^{(k)}$ denotes the monomial basis vector of $\mathbb{K}\langle C_k \rangle_d / \mathcal{I}$, and the LHS L_d includes the inequality localizing contributions $G_i^{(k)}$ exactly as in the dense case. The normal form $\mathcal{N}$ can then be expressed on the monomial basis of each clique

$$\mathcal{N} : (\alpha, \beta)_k \mapsto \sum_t n_{\alpha\beta k}^t t, \quad (44)$$

where reductions occur in the global quotient $\mathbb{K}\langle \underline{X} \rangle / \mathcal{I}$, so reduced words t can receive contributions from multiple cliques. We search again for rational correction matrices $\Delta^{(k)}$ such that the certificate holds exactly for all reduced word coefficients t :

$$\mathcal{N}(L_d)_t = \mathcal{N} \left(\sum_k \left(v_d^{(k)} \right)^* (G_0^{(k)} + \Delta^{(k)}) v_d^{(k)} \right)_t \quad (45)$$

For that, one again defines the residuals per reduced words, now involving Gram matrices associated to all cliques:

$$r_t = \mathcal{N}(L_d)_t - \sum_{\alpha, \beta, k} n_{\alpha\beta k}^t \left(G_0^{(k)} \right)_{\alpha, \beta}, \quad (46)$$

where (α, β, k) runs over monomial pairs (α, β) belonging to clique k . Following exactly the same

argument as in Theorem 2 but summing over tuples involving the clique index (α, β, k) let us derive the Frobenius optimal projection in the sparse case.

Lemma 7 (Frobenius optimal sparse Gram projection). *Let $(\alpha^* \beta)_k \in \mathbb{K}\langle X \rangle_{2d}$ be the word corresponding to indices (α, β) of Gram matrix $G_0^{(k)}$. Let*

$$(\Xi)_{t,s} = \sum_{\alpha, \beta, k} n_{\alpha\beta k}^t n_{\alpha\beta k}^s. \quad (47)$$

If Ξ is invertible, the Frobenius optimal correction reads

$$\Delta_{\alpha, \beta}^{(k)} = \sum_t (\Xi^{-1} r)_t n_{\alpha\beta k}^t, \quad (48)$$

where for binomial $\mathcal{N}$ the Frobenius optimal correction reads

$$\Delta_{\alpha, \beta}^{(k)} = \frac{r_{\mathcal{N}(\alpha^* \beta)}}{n_{\mathcal{I}}(\alpha, \beta, k)} \quad (49)$$

with

$$n_{\mathcal{I}}(\alpha, \beta, k) = \#\{(\gamma, \delta, l) : \mathcal{N}((\alpha^* \beta)_k) = \mathcal{N}((\gamma^* \delta)_l)\}. \quad (50)$$

The proof follows exactly the same argument as the proof of Lemma 2 and is hence not presented here.

Intuitively, this result tells us how to update the entries of Gram matrices associated to different cliques, such that the projected Gram matrices deviate the least from their numerical counterpart. Again, this is a natural choice because we expect the least deviation of the eigenvalues of the numerical Gram matrices, hence minimizing the bound loss associated to our scheme.

In the case of the projected Gram matrices leaving the PSD cone, we perform a similar lifting of λ as in the dense case, where we again denote the rounded, projected, and lifted matrices with $\mathcal{P}_+(\tilde{G}_0^{(k)})$.

Theorem 8 (Sparse certified bounds). *Assume an inaccurate numerical certificate $(\lambda_d^{\text{sp}}, \{G_0^{(k)}\}_{k=1}^{N_C}, \{G_i^{(k)}\}_{i=1}^{m_G})$ with the Gram matrices $\{\mathcal{P}(G_0^{(k)})\}$ admitting lower spectral bounds $\mu_{\min}^k$ after rounding and projecting. Furthermore, assume that all of the variables are subject to one of the constraint sets $\{c_i\}$ from Theorem 4. Then,*

$$\lambda_d^{\text{sp, rat}} = \tilde{\lambda}_d^{\text{sp}} + \delta_d^{\text{sp}} := \tilde{\lambda}_d^{\text{sp}} - \sum_k \min\{\mu_{\min}^k, 0\} \cdot s_{n_k, d}^{\mathcal{I}} \quad (51)$$

where $s_{n_k, d}^{\mathcal{I}} = \dim(\mathbb{K}\langle C_k \rangle_d / \mathcal{I})$, constitutes an exact upper bound to the maximization problem, i.e., $\lambda^{\text{sp, rat}} \geq \lambda_{\max}$.

Proof of Theorem 8. Let $\{C_k\}_{k=1}^{N_C}$ be the maximal cliques of a chordal extension of C . For each k , let $v_d^{(k)}$ be the monomial basis of $\mathbb{K}\langle C_k \rangle_d / \mathcal{I}$ and J_k the indices of constraints supported on C_k .

Assume an exact pre-certificate

$$L_d = \sum_k \left(v_d^{(k)}\right)^* \mathcal{P}(\tilde{G}_0^{(k)}) v_d^{(k)} \mod \mathcal{I}. \quad (52)$$

For blocks $\mathcal{P}(\tilde{G}_0^{(k)}) \not\geq 0$, Lemma 5 with $\varepsilon = -\mu_{\min}^k$ on variables of C_k gives

$$-\mu_{\min}^k s_{n_k, d}^{\mathcal{I}} 1 = -\mu_{\min}^k (v_d^{(k)})^* v_d^{(k)} + \sum_i r_i^{(k)*} r_i^{(k)} + \quad (53)$$

$$\sum_{i,j} q_{ij}^{(k)*} c_i q_{ij}^{(k)} \mod \mathcal{I}.$$

Fixing k and adding one of those expressions to 52 yields a PSD update as in the dense case

$$\mathcal{P}_+(\tilde{G}_0^{(k)}) = \mathcal{P}(\tilde{G}_0^{(k)}) - \mu_{\min}^k 1 + R_0^{(k)} \succeq 0, \quad (54)$$

and raises $\tilde{\lambda}_d$ on the LHS by $-\mu_{\min}^k s_{n_k, d}^{\mathcal{I}}$.

Aggregating over all k with $\mu_{\min}^k < 0$ gives

$$\lambda_d^{\text{sp, rat}} - f = \sum_{k=1}^{N_C} (v_d^{(k)})^* \mathcal{P}_+(\tilde{G}_0^{(k)}) v_d^{(k)} + \sum_k \sum_{i \in J_k} \sum_j (p_{ij}^{(k)})^* g_i p_{ij}^{(k)} \in \mathcal{K}_{2d}^{\text{sp}}, \quad (55)$$

which implies the claim. $\square$

We note that if $\exists k : \mathcal{P}(\tilde{G}_0^{(k)}) \succ 0$, the bounds can also be tightened clique-wise as in Theorem 6, which can partly compensate negative eigenvalues from other cliques.

4.3 Exact certificates for symmetric systems

The structure of the numerical certificate (43) is not exclusive to sparse relaxations, but also occurs in the case of block diagonal Gram matrices G_0 . As block-diagonalizable Gram matrices correspond to associated symmetries, one can adapt Theorem 8 to also be able to obtain certified bounds on symmetry-reduced problems of the corresponding form.

Lemma 9 (Frobenius optimal Gram projection over polynomial bases). *Assume that due to the symmetry of a problem, the underlying Gram matrix G_0 can be block-diagonalized by a transformation U into K blocks $G_0^{(k)}$ of maximal size N . Depending on U , each block will be indexed by a non-monomial basis, e.g., polynomials $\{b_i^{(k)}\}_{i=1}^N$, which can be expressed in the monomial basis of the full space $\{v_\alpha\}_{\alpha=1}^{KN}$ as*

$$b_i^{(k)} = \sum_{\alpha} b_{i,\alpha}^{(k)} v_{\alpha}, \quad (56)$$

where $b_{i,\alpha}^{(k)}$ are the matrix elements of U . Assuming an ideal $\mathcal{I}$ generating a normal form expressed on the monomial basis as in (8), the corresponding symmetry-adapted non-negativity certificate reads as

$$\begin{aligned} \mathcal{N}(f - \lambda 1) &= \mathcal{N} \left(\sum_{i,j,k} (b_i^{(k)})^* G_{ij}^{(k)} b_j^{(k)} \right) \\ &= \sum_{i,j,k,\alpha,\beta} (b_{i,\alpha}^{(k)})^* G_{ij}^{(k)} b_{j,\beta}^{(k)} \mathcal{N}(v_{\alpha}^* v_{\beta}), \\ \mathcal{N}(f - \lambda 1)_t &= \sum_{i,j,k,\alpha,\beta} (b_{i,\alpha}^{(k)})^* G_{ij}^{(k)} b_{j,\beta}^{(k)} n_{\alpha\beta}^t := \sum_{i,j,k} \tilde{n}_{ijk}^t G_{ij}^{(k)}, \end{aligned} \quad (57)$$

which has the same form as in Theorem 8. Defining the matrix Ξ accordingly, the Frobenius optimal projection for the Gram blocks indexed by polynomial bases is given by

$$\Delta_{i,j}^{(k)} = \sum_t (\Xi^{-1} r)_t \tilde{n}_{ijk}^t. \quad (58)$$

The proof follows the same arguments as the proof of Lemmas 2 and 7.

Note that to establish a certified bound in the case of $\mathcal{P}(\tilde{G}_0^{(k)}) \not\equiv 0$, one cannot simply apply Lemma 5 and lift the Gram blocks independently as in the sparse case, as they are not indexed by the monomial basis in which constraints $\{c_i\}$ are expressed. Instead, one has to lift the full block-diagonal ambient Gram matrix $\mathcal{P}(\tilde{G}_0) := \text{diag}(\mathcal{P}(\tilde{G}_0^{(1)}), \dots, \mathcal{P}(\tilde{G}_0^{(K)}))$, where eigenvalue computations can still be performed block-wise due to the unitary invariance of the matrices' spectra.

Theorem 10 (Certified bounds in symmetry-adapted bases). *Assume an inaccurate numerical certificate $(\lambda_d, \{G_0^{(k)}\}_{k=1}^K, \{G_i^{(k)}\}_{i=1}^{m_G})$ to an NPO where $G_0^{(k)}$ are blocks indexed by*

polynomials $\{b_i^{(k)}\}_{i=1}^N$ that stem from block-diagonalization of a Gram matrix in the monomial basis, and assume variables are subject to one of the constraint sets $\{c_i\}$ from Theorem 4. Let

$$\mathcal{P}(\tilde{G}_0^{(k)}) = \tilde{G}_0^{(k)} + \Delta^{(k)} \quad k \in [K] \quad (59)$$

be the set of blocks after rounding and projecting according to Lemma 9 with certified bounds to minimal eigenvalues $\mu_{\min}^k$. Then,

$$\lambda_d^{\text{rat}} = \tilde{\lambda}_d + \delta_d = \tilde{\lambda}_d - \min_k(\{\mu_{\min}^k, 0\}) \cdot s_{n,d}^{\mathcal{I}} \quad (60)$$

constitutes a certified upper bound to the maximization problem, i.e., $\lambda_d^{\text{rat}} \geq \lambda_{\max}$.

The proof follows from the invariance of the eigenspectrum under unitary transformations and applying Lemma 5 to $\mathcal{P}(\tilde{G}_0)$ in the monomial basis. Again, if all blocks fulfill $\mathcal{P}(\tilde{G}_0^{(k)}) \succ 0$, the same argument as in Theorem 6 holds, which lets one further tighten the numerical symmetry-adapted bound.

One has to note here that even though the projection scheme in the sparse and symmetric case exhibit a similar structure, they also strongly differ in terms of computational complexity. In the sparse (and trivially, dense) case, one can show that the rationalization scheme carries *less* computational complexity with respect to the matrix sizes than solving the underlying SDP. However, in the symmetry-adapted case, computing the Frobenius optimal projection may exceed the cost of SDP solving, depending on the block partitioning of the initial Gram matrix G_0 . A discussion of the complexity in both sparsity and symmetry-adapted cases is given in Appendix C.

As a second remark we note that in the case of sparse cliques with sufficiently small overlap, one finds

$$s_{n,d}^{\mathcal{I}} > \sum_k s_{n_{C_k},d}^{\mathcal{I}} \quad (61)$$

in which case the loss in bounds through rationalization in the dense case, δ_d , often exceeds the sparse loss, δ_d^{sp} . This fact can lead to the counterintuitive phenomenon that rationalized sparse bounds become tighter than the corresponding rationalized dense bounds, $\lambda_d^{\text{sp,rat}} < \lambda_d^{\text{rat}}$, as it is illustrated in the next section.

5 Maximal quantum violation of Bell inequalities

As a first benchmark for our certification method, as well to demonstrate its necessity, we apply it to the problem of determining the maximal quantum violation of Bell inequalities. All of the numerical experiments in this work were modeled in Julia using the libraries NCTSSOS [8] and QMBCertify [2], and the resulting SDPs were solved with the commercial solver MOSEK [45], using 10^{-8} tolerances for primal feasibility, dual feasibility, and the relative duality gap. The solver choice is critical for numerical reliability, where MOSEK was selected for its state-of-the-art interior-point implementation and its widespread adoption in semidefinite optimization. Certified bounds on minimal eigenvalues $\mu_{\min}$ of the rounded and projected Gram matrices (as explained in Section 3.2) have been computed using `Arblib.jl` [44], where exact rational representations of the certified enclosures returned by `Arblib` are used for subsequent exact arithmetic operations necessary for the algebraic lifting. The code to obtain certified bounds and reproduce our results is integrated in the Julia libraries NCTSSOS¹ and QMBCertify² packages for Bell and many-body benchmarks respectively.

As an illustrative example, consider the famous Clauser-Horne-Shimony-Holt (CHSH) inequality [46] $\langle A_0 B_0 \rangle + \langle A_0 B_1 \rangle + \langle A_1 B_0 \rangle - \langle A_1 B_1 \rangle \leq 2$. Here, $\langle A_i B_j \rangle$ denote expectation values of products of dichotomic observables A_i and B_j . Intuitively one can think of a game involving two parties, Alice and Bob, that perform either of the two measurements randomly, trying to maximize the correlations constituting the inequality. Such inequalities characterize the boundaries between *local behaviors*, admitting a description in terms of local hidden variables, and *quantum behaviors*, e.g., correlations achievable by measurements on a joint bipartite quantum system. If a given strategy exceeds the local bound, equal to 2 in the case of the CHSH inequality, one can conclude that no local hidden-variable model can reproduce the corresponding correlations, and hence that non-local (quantum or supraquantum) resources were

used. A central question along these lines is *how* different the emergent correlations can be for quantum states with respect to local strategies, which is captured in the *maximal quantum violation* of a given Bell expression, known to be equal to $2\sqrt{2}$ in the CHSH case.

Finding maximal violations of Bell expressions to bound the set of quantum correlations constitutes the first family of problems for which non-commutative SDP relaxations were exploited [1]. Such scenarios can be extended to feature an arbitrary number of parties, measurements per parties, and outcomes per measurement, depending on the setting of interest. In this section we focus on two-party Bell expressions featuring m_A and m_B dichotomic measurements on Alice's and Bob's side, respectively. In general, an optimization problem of this class reads

$$\begin{aligned} \lambda_{\max} = & \sup_{\mathcal{H}, |\psi\rangle \in \mathcal{H}, A_i, B_j \in \mathcal{B}(\mathcal{H})} \sum_{i=1}^{m_A} \sum_{j=1}^{m_B} c_{ij} \langle A_i B_j \rangle \\ & + \sum_{i=1}^{m_A} a_i \langle A_i \rangle + \sum_{j=1}^{m_B} b_j \langle B_j \rangle \quad (62) \\ \text{s.t.} \quad & A_i^2 = 1, \quad i = 1, \dots, m_A, \\ & B_j^2 = 1, \quad j = 1, \dots, m_B, \\ & [A_i, B_j] = 0, \end{aligned}$$

where the coefficients $\{c_{ij}, a_i, b_j\}$ encrypt the type of Bell inequality, and commutation of A_i and B_j is enforced to model the bipartite Hilbert space structure. This problem exhibits the same structure as (2), allowing us to make use of the NPA hierarchy and our certification scheme.

The need for certification already becomes clear in the simplest case of the CHSH inequality: Running hierarchy level $d = 2$ for CHSH yields the following (wrong) upper bound:

$$\lambda_2 \approx 2.828427122879 \not\geq 2\sqrt{2} = \lambda_{\max}, \quad (63)$$

smaller than the known maximal quantum violation. However, the certified bound fulfills $\lambda_2^{\text{rat}} > \lambda_{\max}$, as illustrated by the following truncated decimal expansion of the exact rational representation:

$$\lambda_2^{\text{rat}} \approx 2.828427125402 > 2\sqrt{2} = \lambda_{\max}. \quad (64)$$

We emphasize here that while the numerical bound λ_2 *does* hold when truncated at the 8th digit, corresponding to the SDP tolerance of the

¹<https://github.com/wangjie212/NCTSSOS/tree/master/examples>

²<https://github.com/wangjie212/QMBCertify/tree/main/examples>

solver, the main conceptual point of this work is to actually *certify* the soundness of a given bound via an exact algebraic certificate, as it cannot be guaranteed by merely truncating numerical bounds at the expected numerical precision.

5.1 Certified bounds to maximum quantum violations of Bell inequalities $A_2 - A_{89}$

To further investigate this behavior, we ran our certification methods on a catalog of 88 inequivalent two-party, dichotomic Bell inequalities, called $A_2 - A_{89}$, first introduced in [47], for relaxation orders $d \in \{1, 2, 3\}$. These problems constitute adequate benchmark problems for our method as accurate numerical reference values have been established for 79/88 inequalities in the benchmarking set [48], allowing us to judge the quality of both the numerical bounds as well as their certified counterparts. Besides this, the catalog carries a manifold of different scenarios and hence Gram matrix sizes, allowing us to test our method on a wide spread of settings. We start with the analysis of the dense certification scheme.

As a first question it is interesting to see how the certification loss

$$\delta_d := \lambda_d^{\text{rat}} - \lambda_d = -\min\{\mu_{\min}, 0\} \cdot s_{n,d}^{\mathcal{I}} \quad (65)$$

is correlated with the quality of the initial numerical certificate: we expect the loss in tightness to be worse if the initial numerical certificate had a large numerical error. To do so, as a measure of inaccuracy of the numerical certificate at relaxation order d , we introduce the 1-norm $\mathcal{D}_d$ of the difference between the numerical left hand side L_d

$$L_d = \lambda_d 1 - f, \quad (66)$$

and the numerical right hand side R_d of the certificate

$$R_d = v_d^* G_0 v_d, \quad (67)$$

in terms of reduced basis words indexed by t :

$$\mathcal{D}_d := \sum_t |L_d - R_d|_t, \quad (68)$$

which lets one investigate how the certification loss δ_d depends on the inaccuracy of the numerical certificate.

Figure 3a shows the behavior of the bound loss δ_d depending on the numerical certificate error

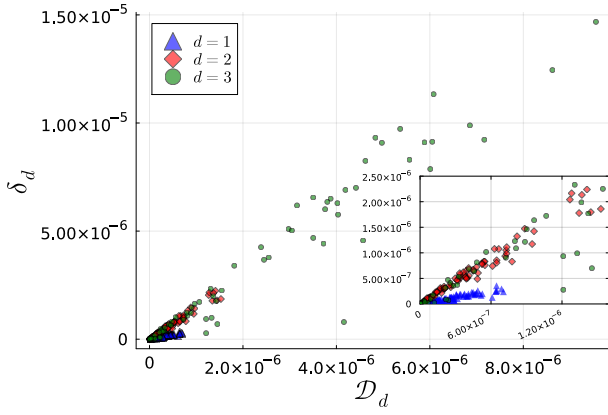
$\mathcal{D}_d$. One can see empirically that large certificate errors $\mathcal{D}_d$ lead to higher bound losses, which is what we intuitively expect: as established in Section 3, the bound loss depends on two variables, the Gram matrix size $s_{n,d}^{\mathcal{I}}$, and the minimal eigenvalue of the projected matrix $\mu_{\min}$, which are both correlated with the certificate error $\mathcal{D}_d$. We expect both $\mu_{\min}$ and $s_{n,d}^{\mathcal{I}}$ to independently affect the bound loss δ_d , which is further examined in the following figures.

Figure 3b shows the trend that larger numerical certificate errors produce larger distances of projected matrices $\mathcal{P}(\tilde{G}_0)$ to the PSD cone. This is expected, as large certificate errors lead to large Frobenius distances between $\mathcal{P}(\tilde{G}_0)$ and G_0 , possibly projecting the matrices further out of the PSD cone to achieve coefficient equalities. Hence, we expect the minimal, negative, projected eigenvalue bound $\mu_{\min}$ to grow in magnitude as the numerical certificate error $\mathcal{D}_d$ worsens, yielding larger certification losses δ_d .

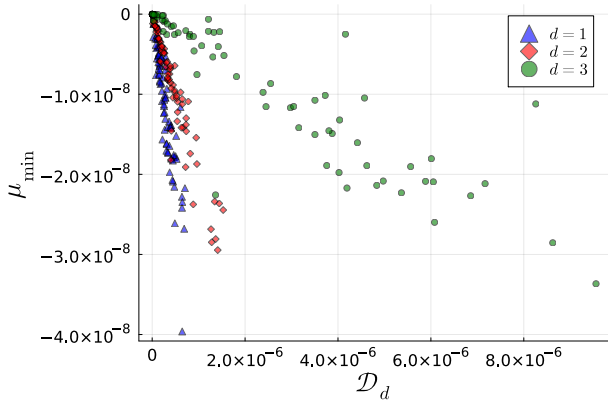
Figure 3c shows the behavior of the bound loss δ_d with respect to the Gram matrix size $s_{n,d}^{\mathcal{I}}$. As the two are multiplicatively correlated, we see that larger $s_{n,d}^{\mathcal{I}}$ can result in larger δ_d . However, one can also see that at a fixed size $s_{n,d}^{\mathcal{I}}$, the bound loss has a high variance, which is reflected by the variance in certificate errors $\mathcal{D}_d$ and hence minimal eigenvalues $\mu_{\min}$. This indicates that the quality of the numerical certificate is at least as important as the combinatorial growth for the tightness of the certified bounds.

To summarize, our first benchmarks indicate that a large numerical certificate error $\mathcal{D}_d$ often leads to a large bound loss δ_d , firstly due to the correlation of $\mathcal{D}_d$ and δ_d with the size $s_{n,d}^{\mathcal{I}}$ of the Gram matrix, and secondly through the larger Frobenius distance $\|\mathcal{P}(\tilde{G}_0) - \tilde{G}_0\|_F$ induced through projection, enabling larger negative eigenvalues $\mu_{\min}$.

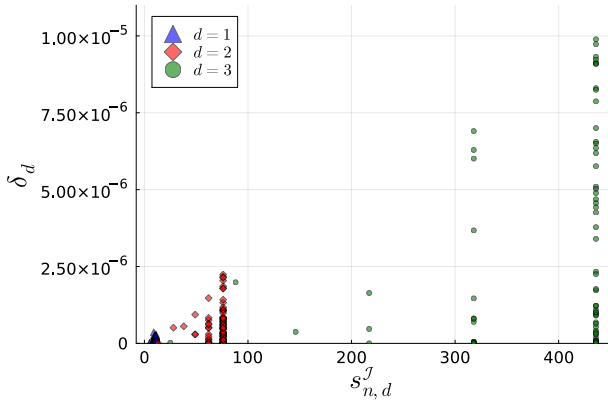
In [48], the authors established that relaxations for the majority of Bell inequalities $A_2 - A_{89}$ match constructive lower bounds up to the numerical accuracy of 10^{-8} at latest at order $d = 3$, where the obtained reference maximal violations λ_{ref} are computed in [49, 50]. Splitting the inequalities into three sets, $\mathcal{B}_i, i \in \{1, 2, 3\}$ according to the hierarchy level at which the lower and upper bounds numerically coincide (sorting the remaining inequalities to $\mathcal{B}_3$), allows us to investigate the behavior of the obtained numerical



(a) Numerical certificate error $\mathcal{D}_d$ vs. bound loss δ_d . The worse the accuracy of the initial numerical certificate measured by its 1-norm error, the worse the resulting bound loss δ_d .



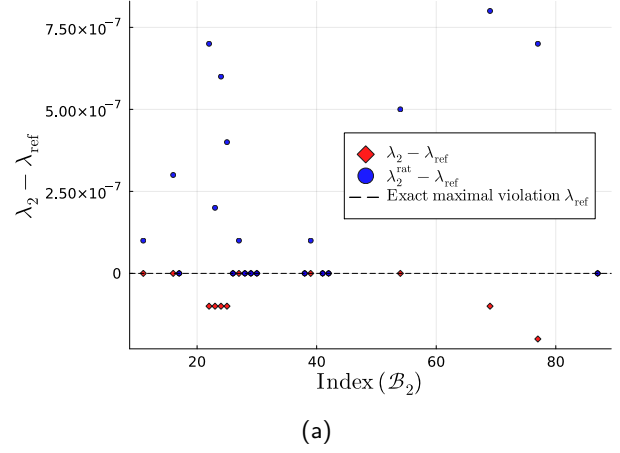
(b) Numerical certificate error $\mathcal{D}_d$ vs. minimal eigenvalue after projection $\mu_{\min}$. A large certificate error $\mathcal{D}_d$ leads to large Frobenius distances $\|\tilde{G}_0 - \mathcal{P}(\tilde{G}_0)\|_F$ trying to enforce coefficient equality in the certificate, leading to possibly large distances to the PSD cone after projection.



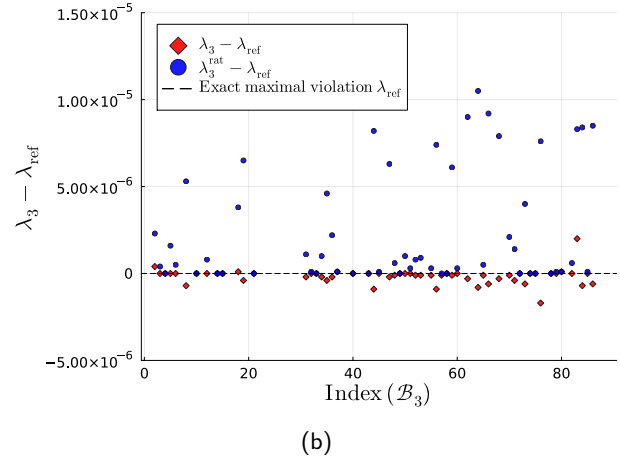
(c) Gram-matrix size $s_{n,d}^T$ vs. bound loss δ_d . The larger the size of the Gram matrix $s_{n,d}^T$, the larger the possible bound loss δ_d . However, at fixed matrix size, δ_d is determined by the minimal eigenvalue $\mu_{\min}$.

Figure 3: Numerical results for the certification scheme on Bell inequalities A2-A89

bounds, as well as the certified bounds, with respect to these reference maximal violations. As $|\mathcal{B}_1| = 2$, we restrict the analysis to the cases of $\mathcal{B}_2$ and $\mathcal{B}_3$.



(a)



(b)

Figure 4: Numerical bounds λ_d and rationalized bounds λ_d^{rat} relative to reference violations λ_{ref} for $\mathcal{B}_2$ (a) and $\mathcal{B}_3$ (b).

Figure 4 shows how the numerical and rational bounds relate to the numerical reference maximal violations³. Note that a point in the positive halfspace denotes a bound which is consistent with the reference values, while points in the negative halfspace are necessarily invalid. We see that while a significant fraction of the numerical bounds λ_d at both relaxation orders are invalid, all produced rational bounds are proper upper bounds to λ_{ref} . In fact, we find that 6/20 numerical bounds in $\mathcal{B}_2$ are incorrect, while 26/66 bounds in $\mathcal{B}_3$ are incorrect. This demonstration

³Bell inequalities with indices $i \in \{7, 9, 10, 15, 28, 42, 44, 48, 66\}$ are not displayed because their numerical bounds hold by a significant margin.

highlights the necessity of certified bounds as reliable alternatives to their numerical counterparts.

5.2 Comparison between the dense and sparse scheme

While the minimal projected eigenvalue $\mu_{\min}$ inherently depends on the certificate error $\mathcal{D}_d$, and hence on the numerical precision of the solver, one can mitigate the dependence on the Gram matrix size by making use of the sparse hierarchy and associated certification scheme. To do so, we applied the sparse scheme to the same set of Bell inequalities and compared the sparse certification bound loss δ_d^{sp} to the dense one.

Figure 5 shows the difference between the dense and the sparse bound loss, $\delta_d - \delta_d^{\text{sp}}$ for different relaxation orders $d \in \{1, 2, 3\}$. Points in the positive half space hence correspond to experiments where $\delta_d > \delta_d^{\text{sp}}$, i.e., where the sparse certification scheme lead to a smaller bound correction after projection.

As one can see, often times the dense loss δ_d exceeds the sparse loss δ_d^{sp} , which, when neglecting the distribution of minimal eigenvalues, can be explained by the often smaller total problem size in the sparse case with respect to the dense case. As discussed in Section 4, by exploiting sparsity, some monomials, i.e., those involving variables from different cliques, do not appear in any of the sparse gram matrices, while they do in the dense description. On the other hand, a large variable overlap of cliques could cause the sum of sparse block sizes ($\sum_k s_{n_k,d}^{\mathcal{I}}$) to become larger than the dense block size ($s_{n,d}^{\mathcal{I}}$). However, if the overlap is small enough, the total size will be smaller than in the dense case, which is reflected in the figures.

This behavior can lead to the counterintuitive phenomenon that the certified sparse bounds become *tighter* than their rational dense counterparts, in stark contrast to results from the numerical relaxations. This happens exactly when the excess in δ_d compensates for the numerical difference $\lambda_d < \lambda_d^{\text{sp}}$. In our experiments we find that at $d = 1$, 70/89 rational sparse bounds were tighter than the dense counterparts. At $d = 3$ we find that 6/88 sparse bounds were tighter, while at $d = 2$ all dense bounds were tighter. This dependency on the relaxation order can be explained by two competing mechanisms, assuming cliques with sufficiently small overlap: On the one hand, the numerical bound difference $\lambda_d - \lambda_d^{\text{sp}}$ is ex-

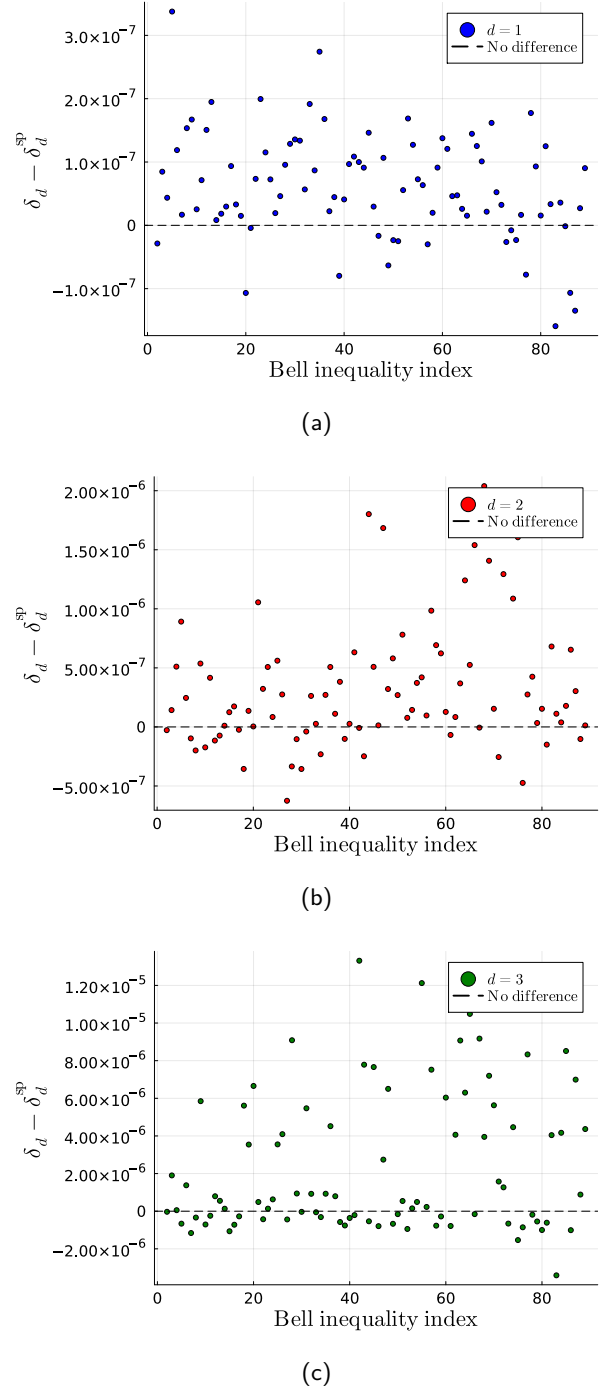


Figure 5: Difference in bound losses $\delta_d - \delta_d^{\text{sp}}$ between dense and sparse certification schemes for Bell inequalities A_2 – A_{89} at relaxation orders $d = 1$ (a), $d = 2$ (b), and $d = 3$ (c).

pected to increase with increasing d , as the sparse hierarchy fails to capture the increasing amount of clique overlap monomials. On the other hand, the dense Gram matrix size grows more rapidly with d , leading to a larger multiplicative factor when computing the certified bounds λ_d^{rat} .

5.3 Tilted CHSH inequality

As a last benchmark in the non-locality scenario we investigated the performance of our scheme to estimate the maximal quantum violation of a variant of the so called *tilted* CHSH inequality [51], introduced in [52]:

$$\begin{aligned} \mathcal{B}_t(\theta_A, \theta_B) &= \sum_{x,y} (-1)^{xy} \langle A_x B_y \rangle + \theta_A \langle A_0 \rangle + \theta_B \langle B_0 \rangle \\ &\leq 2 + \theta_A + \theta_B. \end{aligned} \quad (69)$$

The study of said inequality poses a computationally relevant problem for our framework: On the one hand, this tilted CHSH inequality is self-testing, and the authors provide an analytical solution for its maximal quantum violation. On the other hand, the authors show that this tilted CHSH inequality is *intractable* using the NPA hierarchy as $\theta_A + \theta_B \rightarrow 2$, as increasingly high relaxation levels are needed for convergence. In fact, the authors showed, using the arbitrary-precision solver SDPA-GMP [42], that for parameters $\theta_A = \theta_B = 0.999$, even NPA level $d = 10$ does not achieve the analytical optimal violation. This provides an interesting application for the certification scheme, as it allows us (a) to investigate its performance at high relaxation orders without finite convergence and (b) compare computed bounds to the exact maximum quantum violations.

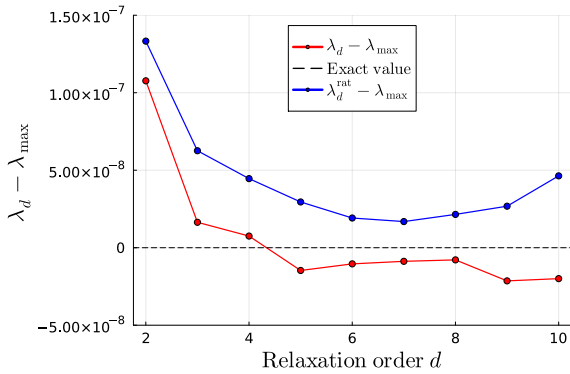


Figure 6: Numerical bounds λ_d and certified bounds λ_d^{rat} with respect to the analytical maximal violation $\lambda_{\max}$ of the tilted CHSH inequality with parameters $\theta_A = \theta_B = 0.999$

As one can see in Figure 6, even though high-precision calculations showed that the hierarchy does not converge at the displayed relaxation orders, the numerical upper bounds λ_d computed

with MOSEK are below the analytical maximum for relaxation orders $d > 4$, resulting in wrong bounds, which furthermore do not monotonically decrease. On the other hand, our certification scheme produces valid bounds at all relaxation orders, with the tightest bound achieved at order $d = 7$. In this scenario where numerical bounds are stagnating, the certified bounds start to worsen due to the rapid growth of $s_{n,d}^{\mathcal{I}}$ with respect to d .

6 Ground-state observables in quantum many-body systems

A primary objective of quantum many-body physics is to understand the properties of the state of minimal energy, or ground state of a physical system. A conceptually simple method to do so is *exact diagonalization*, where one represents the Hamiltonian describing a system as a matrix in a chosen basis and computes its full eigenspectrum numerically [53]. While this method provides numerically exact solutions, the Hilbert space dimension typically grows exponentially with the particle number N (for instance, $\dim \mathcal{H} = 2^N$ for N spin-1/2 particles), rendering exact diagonalization infeasible rapidly. To mitigate this, one often restricts the variational ground-state search to physically motivated manifolds [54]. Notably, tensor network methods [55], e.g. Matrix-Product States or Projected Entangled Pair States, have been shown to significantly scale up the feasible problem sizes. However, a variational Ansatz, in general, does not provide any guarantee for how well the Ansatz space captures the actual ground state of the problem, and can just provide upper bounds to the true ground-state energy

$$E_0^{\text{var}} \geq E_0. \quad (70)$$

For any other observable, variational methods just offer an estimation, with no guarantee of how close it is to the true value or whether it bounds it from above or below. The NPA hierarchy constitutes a counterpart to these variational Ansätze by virtue of relaxation, hence allowing one to (numerically) certify an interval for E_0 :

$$E_0^{\text{var}} \geq E_0 \geq E_0^{\text{SDP}}. \quad (71)$$

Now, in [2] the authors make use of such inequality constraints on the ground-state energy

to obtain bounds on ground-state observables beyond the energy (e.g. magnetization and local correlators). This constitutes a particularly interesting application for our method, as it constitutes one of the aforementioned iterative algorithms, reusing SDP bounds of former runs in consecutive programs. If wrong numerical lower bounds are issued, $E_0^{\text{SDP}} \not\leq E_0$, imposing such a wrong inequality in the computation of other ground state observables might cause an accumulation of errors. Our certification method precisely eliminates this failure mode by turning the SDP output into a rigorously valid lower bound.

We start by applying our certification scheme to the computation of ground state energies of the translationally invariant $J_1 - J_2$ Heisenberg chain with nearest- and second-nearest neighbor interactions and periodic boundary conditions with Hamiltonian

$$H = \frac{1}{4} \sum_{i=1}^N \sum_{a \in \{x,y,z\}} (\sigma_i^a \sigma_{i+1}^a + J_2 \sigma_i^a \sigma_{i+2}^a), \quad (72)$$

where σ_i^a denotes the Pauli operator of type a acting on site i . We adapt the convention $J_1 = 1$ from the underlying work.

In [2, 56], the authors exploit sparsity and various symmetries to significantly improve the scalability of the hierarchy, allowing them to study large system sizes. Exploiting the sign symmetry of the model lets one block-diagonalize the dense Gram matrix into four smaller blocks $G^{(j)}, j \in \{1, 2, 3, 4\}$. Due to the model's symmetry under permutation of the Pauli operators $\{\sigma_i^x, \sigma_i^y, \sigma_i^z\}$, one can identify three of the blocks with each other, leaving one with two independent PSD blocks $G^{(j)}, j \in \{1, 2\}$ which can be further block diagonalized to blocks $G_i^{(j)}, i \in \{1, \dots, N/2 + 1\}$ due to the translational symmetry of the model. The associated transformation is the discrete Fourier transform $P \in \mathbb{C}^{N \times N}$:

$$P_{i,j} = \frac{1}{\sqrt{N}} e^{-2\pi i(i-1)(j-1)/N}, \quad i, j = 1, \dots, N. \quad (73)$$

For details on these symmetry exploitations the reader is directed to [56].

After exploiting the symmetries, the numerical SOHS certificate for lower bounds λ_d to said

ground state problems can be obtained as

$$\begin{aligned} \frac{H}{N} - \lambda_d &\simeq \left[1, (1_{s_1} \otimes P_1) \vec{b}_1\right]^* G_1^{(1)} \left[1, (1_{s_1} \otimes P_1) \vec{b}_1\right] \\ &+ \sum_{i=2}^{\frac{N}{2}+1} \left((1_{s_1} \otimes P_i) \vec{b}_1\right)^* G_i^{(1)} \left((1_{s_1} \otimes P_i) \vec{b}_1\right) \\ &+ \sum_{i=1}^{\frac{N}{2}+1} \left((1_{s_2} \otimes P_i) \vec{b}_2\right)^* G_i^{(2)} \left((1_{s_2} \otimes P_i) \vec{b}_2\right), \end{aligned} \quad (74)$$

where $\vec{b}_j$ with $|\vec{b}_j| = N s_j$ are monomial bases of the two sign symmetry sectors, 1_{s_j} denotes the identity matrix of size s_j , and P_i the i -th row of P . Note that one can interpret the discrete Fourier transformation as acting on the bases or on the Gram matrix blocks. Hence, one can either perform the projection on the polynomially indexed blocks $G_i^{(j)}$ according to Lemma 9, or inflate them back to the monomial basis and perform the projection according to Lemma 7. As certification in the monomial basis becomes significantly more expensive for large N (see Appendix C), all computations have been performed in the symmetry-adapted basis. In this scenario, the larger blocks $G^{(j)}$ are indexed by independent monomial bases $\vec{b}_j$, while their subblocks $G_i^{(j)}$ have monomial overlap in their indexing polynomials $(1_{s_j} \otimes P_i) \vec{b}_j$. This hence represents an instance featuring both sparsity and symmetry, where an intermediate lifting between Theorems 8 and 10 can be applied:

$$\lambda_d^{\text{rat}} = \tilde{\lambda}_d + \min_i \{\mu_{\min}^{i1}\} \cdot N s_1 + \min_i \{\mu_{\min}^{i2}\} \cdot N s_2, \quad (75)$$

where $\{\mu_{\min}^{ij}\}$ denotes the minimal eigenvalue of block $G_i^{(j)}$. We denote the arithmetic mean of the minima involved by $\overline{\mu_{\min}} = \sum_j \min_i \{\mu_{\min}^{ij}\} / 2$.

Figure 7 shows the results for applying our certification scheme to a $J_1 - J_2$ Heisenberg chain up to length $N = 50$. One can see from Figure 7a that the certified bounds are consistent with the upper bounds computed by DMRG, recovering the results from [2]. Figure 7b shows the bound shift by going from numerical to rational bounds depending on N . What is remarkable here is that in this scenario the bound shift *decreases* with an increasing problem size. This can be explained by the negative minimal $\overline{\mu_{\min}}$ of the projected blocks $\mathcal{P}(\tilde{G}_i^{(j)})$ decreasing in magnitude with increasing N (see Figure 7c), which indicates that the numerical certificates carry a smaller error.

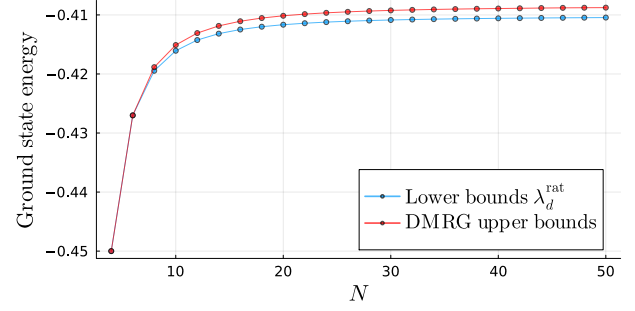
This is most likely due to the fact that at larger N more matrix elements $(G_i^{(j)})_{\alpha,\beta}$ are assigned to the same equivalence class under constraints and symmetries, possibly leading to a more effective cancellation of errors. The second factor preventing a rapid increase in δ_d is the total size of all Gram matrices $s_{n,d}^T = N(s_1 + s_2)$, growing linearly with N (Figure 7d) due to the translational invariance of the model. In conclusion, this first experiment suggests that the numerical results of the certification method seem to improve when symmetries are exploited. However, they also come with an increased computational cost, as it is shown in Appendix C.

One can extend the rationalization scheme to observables beyond the ground state energy. The numerical SOHS certificate then involves scalar Gram multipliers $\kappa_1, \kappa_2 > 0$ that correspond to the inequalities in E_0 . The corresponding numerical certificate for lower (or upper) bounding an observable O under previously established energy constraints reads

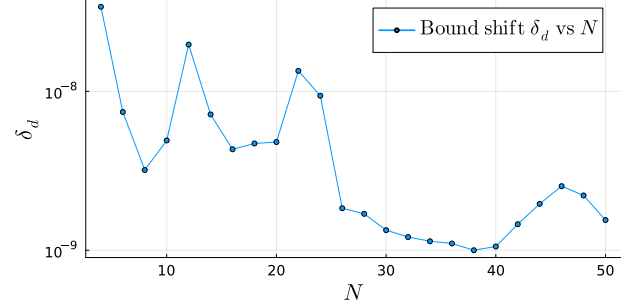
$$\begin{aligned} \frac{\pm O}{N} \mp \lambda_d &\simeq \left[1, (1_{s_1} \otimes P_1) \vec{b}_1\right]^* G_1^{(1)} \left[1, (1_{s_1} \otimes P_1) \vec{b}_1\right] \\ &+ \sum_{i=2}^{\frac{N}{2}+1} \left((1_{s_1} \otimes P_i) \vec{b}_1\right)^* G_i^{(1)} \left((1_{s_1} \otimes P_i) \vec{b}_1\right) \\ &+ \sum_{i=1}^{\frac{N}{2}+1} \left((1_{s_2} \otimes P_i) \vec{b}_2\right)^* G_i^{(2)} \left((1_{s_2} \otimes P(i)) \vec{b}_2\right) \\ &+ \kappa_1(H - E_0^{\text{SDP}}) + \kappa_2(E_0^{\text{var}} - H). \end{aligned} \quad (76)$$

To obtain a rational certificate from the numerical one, one first has to establish rational, certified bounds on E_0 for the multiplier terms. For the SDP bound we use the previously established certified lower bounds on H . As establishing certified variational bounds is not the scope of this work, we use $E_0^{\text{var}, \text{rat}} = (\widetilde{E_0^{\text{var}}} + \delta E_0)$ as a rational upper bound, where δE_0 is chosen larger than the DMRG precision, and the rounding precision η of lower magnitude than δE_0 to ensure a valid bound. Next, one proceeds to construct a rational LHS featuring the rational inequality terms and associated rationalized multipliers as described in Section 3. As the aforementioned symmetry considerations also hold in this scenario, one can then proceed to perform the projection according to Lemma 9 and obtain certified bounds as in (75).

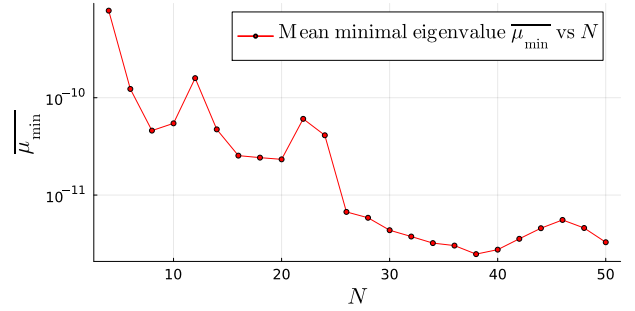
We start by investigating nearest-neighbor correlations $\langle \sigma_i^a \sigma_{i+1}^a \rangle$ in the $J_1 - J_2$ Heisenberg chain



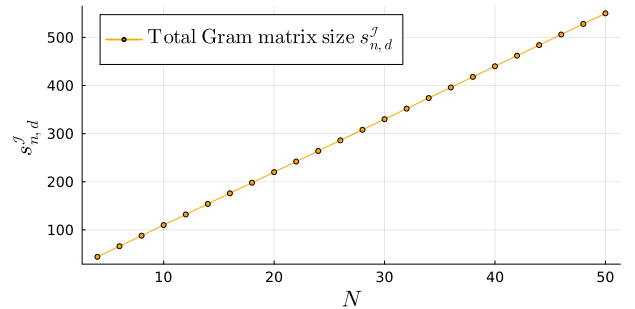
(a) Certified lower bounds λ_d^{rat} and DMRG upper bounds to E_0



(b) Rationalization bound shift $\delta_d = \lambda_d - \lambda_d^{\text{rat}}$



(c) Mean minimal eigenvalue $\overline{\mu_{\min}}$ of the projected Gram blocks $\mathcal{P}(\tilde{G}_i^{(j)})$



(d) Total size of Gram matrices $\mathcal{P}(\tilde{G}_i^{(j)})$

Figure 7: Numerical results for obtaining certified bounds to the $J_1 - J_2$ Heisenberg chain ground state energy depending on the problem size N . Computations were performed at an intermediate relaxation order between $d = 2$ and $d = 3$.

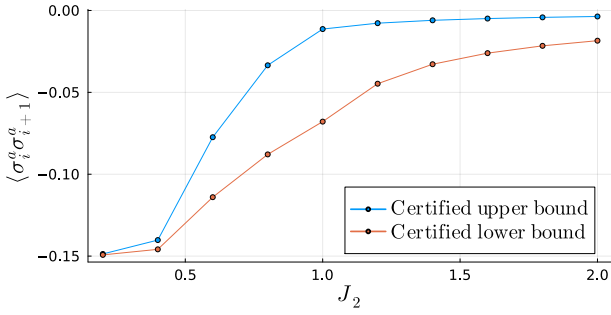


Figure 8: Certified bounds to nearest-neighbor correlations $\langle \sigma_i^a \sigma_{i+1}^a \rangle$ in a $J_1 - J_2$ Heisenberg chain of length $N = 12$. Computations were performed at an intermediate relaxation order between $d = 3$ and $d = 4$.

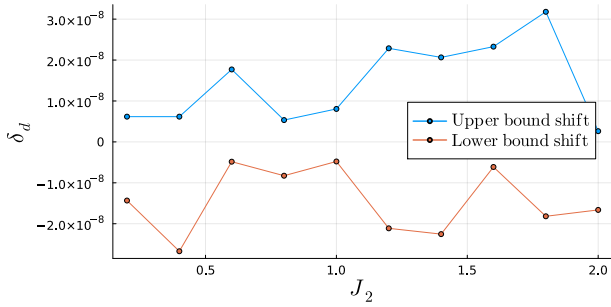


Figure 9: Bound shift δ_d for certified bounds on $\langle \sigma_i^a \sigma_{i+1}^a \rangle$ from Figure 8

of length $N = 12$ dependent on the second-nearest neighbor couplings J_2 . The results are shown in Figure 8, where we are able to certify antiferromagnetic correlations $\langle \sigma_i^a \sigma_{i+1}^a \rangle < 0$ for all J_2 , recovering the results from [2]. Figure 9 shows the corresponding bound shifts δ_d , where we find no obvious correlations between the shifts δ_d and the coupling J_2 . As the problem size $s_{n,d}^{\mathcal{I}}$ is fixed for all J_2 , the differences in δ_d stem solely from differences in minimal eigenvalues $\mu_{\min}^{ij}$ of the projected Gram blocks. This indicates that the numerical problem is equally well conditioned for all J_2 , in stark contrast to the following application.

Figure 10 shows certified bounds to second neighbor spin correlations $\langle \sigma_i^a \sigma_{i+2}^a \rangle$ in the $J_1 - J_2$ Heisenberg chain of length $N = 12$. We are able to certify ferromagnetic correlations $\langle \sigma_i^a \sigma_{i+2}^a \rangle > 0$ for couplings $J_2 \leq J_2^+ < 0.5$, and antiferromagnetic correlations $\langle \sigma_i^a \sigma_{i+2}^a \rangle < 0$ for $J_2 \geq J_2^- > 0.5$. Note that in the analysis of [2], the transition point was exactly at $J_2 = 0.5$. At this value, the so called *Majumdar-Ghosh point*, two exactly degenerate ground states $|\psi_1\rangle$ and $|\psi_2\rangle$ emerge [57], leading to a distinct behavior in the

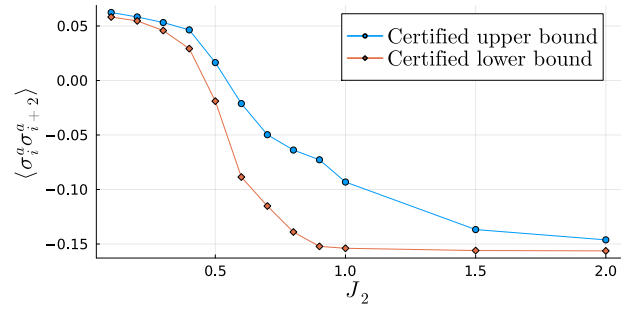


Figure 10: Certified bounds to second neighbor spin correlations $\langle \sigma_i^a \sigma_{i+2}^a \rangle$ in the $J_1 - J_2$ Heisenberg chain of length $N = 12$ for different values of J_2 . An intermediate relaxation order between $d = 3$ and $d = 4$ was used for computations.

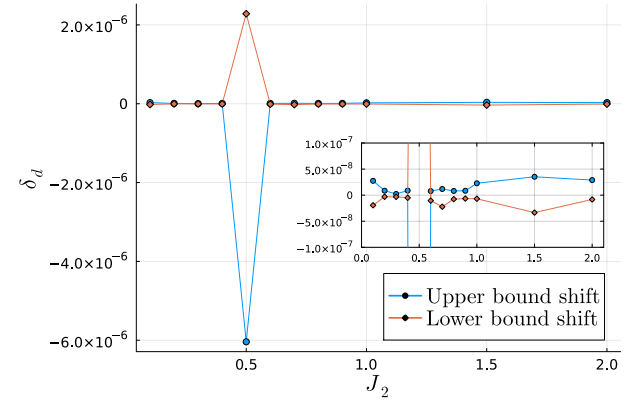


Figure 11: Bound shifts δ_d for certified bounds on $\langle \sigma_i^a \sigma_{i+2}^a \rangle$ from Figure 10.

observed bound shift, as displayed in Figure 11: Remarkably, while for $J_2 \neq 0.5$ projected Gram matrices $\mathcal{P}(\tilde{G}_i^{(j)})$ with negative minimal eigenvalues $\mu_{\min}^{ij} < 0$ emerge, for $J_2 = 0.5$ the numerical, as well as the projected Gram blocks, lie on the *interior* of the PSD cone, carrying positive minimal eigenvalues $\mu_{\min}^{ij} > 0$ of relatively large magnitude. This is in accordance with the observation of the SDP solver not being able to provide an optimal solution at $J_2 = 0.5$, but issuing a feasible point: As the objective function is linear and the feasible set is the intersection of an affine space $\mathcal{L}$ and the PSD cone, optimal numerical solutions have to lie on the boundary of the PSD cone and approximately fulfill $\mu_{\min}^{ij} \approx 0$. In the case of $\mu_{\min}^{ij} > 0$, we are able to tighten the bounds to our problem by applying Theorem 6 (i) to shift the projected blocks from the interior of the PSD cone to its boundary.

This apparent ill-conditioning can be explained by the two-fold degeneracy of the ground state

at $J_2 = 0.5$: As $|\psi_1\rangle$ and $|\psi_2\rangle$ share exactly the same ground state energy E_0 , any convex combination will do so too, leading to an increase in dimensionality of the optimal subspace. Hence, the energy constraints employed to compute bounds on $\langle \sigma_i^a \sigma_{i+2}^a \rangle$ do *not* restrict the search space in the same manner as for problems with non-degenerate ground states, as it is the case for the other investigated values of J_2 for a finite chain.

This phenomenon suggests that, while in general our certification procedure produces weakened bounds, it is possible to construct tighter rational bounds λ_d^{rat} with respect to their numerical counterpart λ_d if SDP solutions are issued in the interior of the PSD cone. Conclusively, this suggests a second possible application of our method to obtain certified *and* tighter bounds to problems which are intractable to SDP solvers.

7 Conclusion and Outlook

In this work we presented a constructive method to obtain certified outer bounds to non-commutative optimization problems by postprocessing numerical data obtained from the NPA hierarchy. We do so by extracting an exact SOHS certificate from an SDP solver's Gram data, which consists of (1) rounding the data to rationals (2) performing a Frobenius optimal projection back onto the affine constraint subspace, and (3) a constraint dependent lifting of the projected Gram matrices into the PSD cone. We further extended this procedure to sparsity and symmetry-adapted scenarios, and analyzed the method's complexity relative to the underlying SDP. We also provided a tightening procedure which for PD projected Gram matrices leads to tighter, certified bounds.

We then applied our method to obtain certified bounds on (a) the maximal violation of a catalog of bipartite Bell inequalities, and (b) quantum many-body ground-state observables, where we found moderate losses in tightness that reflect a justified cost for obtaining rigorous guarantees. In the sparse case we observed that sparse rational bounds can become tighter than their dense counterparts, while in the symmetry-adapted scenario we identified a computational bottleneck that we suspect to be common to any SDP-based computer-assisted proof in arbitrary symmetry-adapted bases. In addition, we also find that

in regimes of inexact SDP termination, the rational bounds can become tighter than their numerical, uncertified counterparts, which is of independent interest when dealing with numerical instabilities. In summary, we provided analytical constructions and numerical evidence that promote numerical semidefinite relaxations toward computer-assisted proofs in quantum theory.

Our method is readily extendable to modified non-commutative hierarchies such as trace- or state-optimization, as they differ from the presented hierarchy solely up to moment equality constraints. Following the same argument, it is also possible to extend this method to differentially constrained non-commutative polynomial optimization problems, allowing one to certify non-polynomial quantities in quantum theory. We also emphasize that if our post-processing framework proves applicable to non-commutative optimization problems arising in other application domains, this would further enhance its significance.

There are two main venues for further investigation, namely (i) optimality of the bound-lifting step in the dense case, and (ii) scalability in the symmetry-adapted setting: Since our bound-lifting strategy scales with the size of the Gram blocks, which grows rapidly in the dense hierarchy, we observe increasing bound losses at higher relaxation orders. A refined bound-lifting procedure featuring a more involved spectral analysis of the Gram matrices, or a proof of optimality of the presented method hence deserve further investigation. Secondly, as we showed in Appendix C, when Gram blocks are indexed by polynomial bases, even establishing the numerical non-negativity certificate might exceed the cost of solving the associated SDP. Understanding how structure exploitation of the unitary basis change might reduce the certificate complexity is hence another open question. In addition, it would be interesting to systematically quantify the dependence of the certification loss on solver accuracy: Using higher-precision solvers should yield more accurate numerical certificates and hence smaller bound differences δ_d , at the expense of computational efficiency. Regarding the last point, it would be also interesting to obtain *lower* bounds on the relaxed maxima, providing the counterpart to our certified bounds and yielding certified enclosures for the initial numerical upper bound,

allowing one to further judge its quality, e.g., as demonstrated in [38]. To do so, one has to apply the projection on the dual (pseudo-moment) side of the certificate, where the exact projection onto the pseudo-moment cone poses the main technical challenge.

Beyond these directions, it would also be worthwhile to further study the certification scheme for numerically delicate instances in which the solver returns solutions strictly inside the PSD cone, and to systematically leverage such cases to obtain tighter bounds.

Acknowledgements

We thank Nando Leijenhorst, Jonas Britz, Erik Altelaarrea Ferre, and Andreas Leitherer for useful discussions. This work has been supported by European Union's HORIZON-MSCA-2023-DN-JD programme under the Horizon Europe (HORIZON) Marie Skłodowska-Curie Actions, grant agreement 101120296 (TENORS), the European Union Quanteria project Veriqtas, the European Union Quanteria project COMPUTE, the Government of Spain (Severo Ochoa CEX2019-000910-S, FUNQIP and NextGeneration EU PRTR-C17.I1) and European Union (QSNP, 101114043 and PASQuanS2.1, 101113690), Fundació Cellex, Fundació Mir-Puig, Generalitat de Catalunya (CERCA program), the ERC AdG CERQUTE, the AXA Chair in Quantum Information Science, the National Key R&D Program of China under grant No. 2023YFA1009401, the Natural Science Foundation of China under grant No. 12571333, and the International Partnership Program of Chinese Academy of Sciences under grant No. 167GJHZ2023001FN. This work was granted access to the HPC resources of CALMIP supercomputing center under the allocation 2016-23035.

Author contribution

The authors acknowledge the use of Large Language Models for mathematical writing and code development.

A Proof of Lemma 5

Proof of Lemma 5. We deal with the four different constraint families separately.

Box and unipotency constraints. Assume $\{c_i\}_{i=1,\dots,n} = \{1 - X_i^2\}_{i=1,\dots,n} \subseteq G$, e.g. Box constraints. We show that for $\varepsilon > 0$ there exists the following rational decomposition

$$\varepsilon s_{n,d} = \varepsilon v_d^* v_d + \sum_{i,k} q_{ki}^* (1 - X_i^2) q_{ki}. \quad (77)$$

with polynomials $q_{ki} \in \mathbb{Q}_{d-1}\langle X \rangle$. To do so, we first show that

$$1 = w^* w + \sum_{i,k} q_{ki}^* (1 - X_i^2) q_{ki}. \quad (78)$$

for all monomials w with $|w| \leq d$. Proceed by induction on d . The variables X_j yield the base case for words of length $|w| = 1$:

$$1 = X_j^2 + (1 - X_j^2). \quad (79)$$

Assume the induction hypothesis (78) holds for w with $|w| \leq d-1$. Multiplying by X_j from both sides gives

$$X_j^2 = u^* u + \sum_{i,k} X_j q_{ki}^* (1 - X_i^2) q_{ki} X_j. \quad (80)$$

with $u = X_j w$. Plugging this expression into (79) yields

$$1 = u^* u + \sum_{i,k} X_j q_{ki}^* (1 - X_i^2) q_{ki} X_j + (1 - X_j^2)$$

which is of the desired form. Summing over all u with $|u| \leq d$ and multiplying by ε proves the claim (77). The full argument also holds for unipotency constraints.

Ball constraints. Assume $c_1 = (1 - \sum_i X_i^2) \in G$. We want to show the existence of

$$\varepsilon s_{n,d} 1 = \varepsilon v_d^* v_d + \sum_k r_k^* r_k + \sum_k q_k^* \left(1 - \sum_i X_i^2\right) q_k. \quad (81)$$

for $\varepsilon > 0$, $r_k \in \mathbb{Q}_d\langle X \rangle$ and $q_k \in \mathbb{Q}_{d-1}\langle X \rangle$. We again proceed by induction on $|w|$, where we want to show that for all w with $|w| \leq d$

$$1 = w^* w + \sum_k r_k^* r_k + \sum_k q_k^* \left(1 - \sum_i X_i^2\right) q_k. \quad (82)$$

The base case is again given by the variables X_j :

$$1 = X_j^2 + \sum_{i \neq j} X_i^2 + \left(1 - \sum_i X_i^2\right). \quad (83)$$

Assuming (82) holds for $|w| \leq d-1$, multiplying the expression from both sides by X_j gives

$$X_j^2 = u^*u + \sum_k X_j r_k^* r_k X_j + \sum_k X_j q_k^* \left(1 - \sum_i X_i^2\right) q_k X_j. \quad (84)$$

Substitute into (83)

$$1 = u^*u + \sum_{i \neq j} X_i^2 + \sum_k X_j r_k^* r_k X_j + \sum_k X_j q_k^* \left(1 - \sum_i X_i^2\right) q_k X_j + \left(1 - \sum_i X_i^2\right), \quad (85)$$

which again proves the claim (81) by summing over all $|u| \leq d$ and scaling by ε .

Projection constraints. Assume $\{c_i\}_{i=1}^n = \{X_i - X_i^2\}_{i=1}^n \subseteq H$. We want to show

$$\varepsilon s_{n,d} 1 = \varepsilon v_d^* v_d + \sum_k r_k^* r_k + \sum_{i,k} q_{ki}^* (X_i - X_i^2) q_{ki}. \quad (86)$$

Which, as in the other cases, is ensured if the following holds for all w with $|w| \leq d$:

$$1 = w^*w + \sum_k r_k^* r_k + \sum_{i,k} q_{ki}^* (X_i - X_i^2) q_{ki}, \quad (87)$$

which again is proven by induction. The following holds for all j

$$1 = X_j^2 + (X_j - 1)^2 + 2(X_j - X_j^2). \quad (88)$$

Assuming (87) holds for $|w| = d-1$, multiplying it by X_j^2 gives

$$X_j^2 = u^*u + \sum_k X_j r_k^* r_k X_j + \sum_{i,k} X_j q_{ki}^* (X_i - X_i^2) q_{ki} X_j. \quad (89)$$

Plugging into (88) yields

$$1 = u^*u + \sum_k X_j r_k^* r_k X_j + \sum_{i,k} X_j q_{ki}^* (X_i - X_i^2) q_{ki} X_j + (X_j - 1)^2 + 2(X_j - X_j^2) \quad (90)$$

which is of the target SOHS form, yielding (86) by summing over all words and multiplying by ε . We note that the same argument also holds for $\{c_i\}_{i=1}^n \subseteq G$.

In all three cases, we obtain the desired SOHS decomposition for $f s_{n,d}$. The summing argument holds when working over a quotient algebra by replacing $s_{n,d} \rightarrow s_{n,d}^{\mathcal{I}}$, as well as when restricted to specific variable subsets in the sparse case, as in Theorem 8. $\square$

B Proof of Theorem 6

Proof of Theorem 6. We start by proving the tightening scheme (i) from the Theorem.

Let $(\tilde{\lambda}_d, \mathcal{P}(\tilde{G}_0))$ be an exact pre-certificate obtained after rounding and projection

$$\tilde{\lambda}_d 1 - f = v_d^* \mathcal{P}(\tilde{G}_0) v_d \mod \mathcal{I}, \quad (91)$$

and let $\mathcal{P}(\tilde{G}_0) \succ 0$ admit a positive lower bound $\mu_{\min} > 0$ to its eigenspectrum. Recall the following constant SOHS decomposition from Equation 77 in Appendix A:

$$s_{n,d} 1 = v_d^* v_d + \sum_{i,k} q_{ki}^* (1 - X_i^2) q_{ki} \quad (92)$$

Multiplying by any $\varepsilon \in \mathbb{R}$ and working modulo the underlying ideal then yields the following relation

$$\varepsilon s_{n,d}^{\mathcal{I}} 1 = \varepsilon v_d^* v_d \mod \mathcal{I} \quad (93)$$

where this holds particularly for $\varepsilon < 0$ and hence for $\varepsilon = -\mu_{\min} < 0$. This relation follows from the vector space property of the ideal. Adding Equations 91 and 93 then yields the following certificate in the quotient space

$$\begin{aligned} (\tilde{\lambda}_d 1 - \mu_{\min} s_{n,d}^{\mathcal{I}}) - f &= \\ v_d^* (\mathcal{P}(\tilde{G}_0) - \mu_{\min} 1) v_d &\mod \mathcal{I}, \end{aligned} \quad (94)$$

which certifies $\lambda_d^{\text{rat}} = (\tilde{\lambda}_d - \mu_{\min} s_{n,d}^{\mathcal{I}}) < \tilde{\lambda}_d$ as an upper bound to the optimization problem.

We will now prove the second tightening scheme (ii) of the Theorem. Again, assume an exact precertificate as in Equation 91, where no further assumptions on the constraints of the problem are taken. Our goal is to find the maximum $\tau > 0$ such that the following rank-1 downdate is still PSD:

$$\mathcal{P}(\tilde{G}_0) - \tau e_0 e_0^* \succeq 0 \quad (95)$$

considering such downgraded matrix then leads to the following SOHS decomposition

$$(\tilde{\lambda}_d - \tau) 1 - f = v_d^* (\mathcal{P}(\tilde{G}_0) - \tau e_0 e_0^*) v_d \mod \mathcal{I}, \quad (96)$$

Start by considering the following block hermitian matrix

$$S := \begin{pmatrix} \mathcal{P}(\tilde{G}_0) & e_0 \\ e_0^* & \tau^{-1} \end{pmatrix} \quad (97)$$

Now, the Schur complement of τ^{-1} lets one characterize the PSD condition 95:

$$(\mathcal{P}(\tilde{G}_0) - \tau e_0 e_0^*) \succeq 0 \Leftrightarrow S \succeq 0 \quad (98)$$

On the other hand, the Schur complement of $\mathcal{P}(\tilde{G}_0)$ yields

$$\tau^{-1} - e_0^* \mathcal{P}(\tilde{G}_0)^{-1} e_0 \geq 0 \Leftrightarrow S \succeq 0. \quad (99)$$

Putting both arguments together yields

$$(\mathcal{P}(\tilde{G}_0) - \tau e_0 e_0^*) \succeq 0 \Leftrightarrow \tau \leq \frac{1}{e_0^* \mathcal{P}(\tilde{G}_0)^{-1} e_0} \quad (100)$$

Hence, for $\tau = \frac{1}{e_0^* \mathcal{P}(\tilde{G}_0)^{-1} e_0}$, Equation 96 is a valid SOHS certificate, certifying $\lambda_d^{\text{rat}} = \tilde{\lambda}_d - \tau$ as a valid upper bound to the optimization problem. One finds $\tau \in [\mu_{\min}, \mu_{\max}]$ due to the relation of the eigenspectra of $\mathcal{P}(\tilde{G}_0)$ and its inverse. We also note that as $\mathcal{P}(\tilde{G}_0) \in \mathbb{Q}^{s_{n,d}^{\mathcal{I}} \times s_{n,d}^{\mathcal{I}}}$ we have $\tau \in \mathbb{Q}$, which is why no further rationalization is needed if the exact inverse of $\mathcal{P}(\tilde{G}_0)$ is computed. $\square$

C Complexity considerations

In this section we address the computational complexity of the projection schemes in Lemmas 7 and 9 (dense, sparse, and symmetry-adapted), as well as the complexity to compute the associated bound shift δ_d .

First we show that the projection scheme (7), dense or sparse, over a binomial ideal is computationally less complex than solving the underlying SDP. We give the argument in the sparse case as it trivially includes the dense case. We adopt the standard interior-point complexity model, in which an SDP with S blocks of maximal size N requires $\mathcal{O}(SN^3)$ arithmetic operations per iteration, and $\mathcal{O}(\sqrt{SN})$ iterations [58]. As the scaling with respect to the number of constraints involved is not made explicit, this constitutes an agnostic lower bound on the overall complexity.

According to Theorem 8, the Frobenius optimal projection over a quotient space generated by a binomial ideal is given by

$$\Delta_{\alpha,\beta}^{(k)} = \frac{r_{\mathcal{N}(\alpha^*\beta)}}{n_{\mathcal{I}}(\alpha,\beta,k)} \quad (101)$$

Now, for S cliques and associated Gram blocks with largest blocksize N , this update has asymptotic complexity $\mathcal{O}(SN^2)$. To compute the maps

$n_{\mathcal{I}}(\alpha,\beta,k)$, one has to iterate over all blocks S of maximal size N^2 , also yielding complexity $\mathcal{O}(SN^2)$. Now, once the maps $n_{\mathcal{I}}(\alpha,\beta,k)$ are computed, the residuals $r_{\mathcal{N}(\alpha^*\beta)}$ are also obtained in $\mathcal{O}(SN^2)$ according to Equation (46), leading to a total complexity of $\mathcal{O}(SN^2)$, which is dominated by the SDP cost per iteration of $\mathcal{O}(SN^3)$.

We are now going to show that in the case of Gram matrices indexed by non-monomial bases, as it is the case when a Gram matrix is block diagonalized due to symmetry, the computational complexity of the rationalization scheme can become dominant depending on the problem parameters.

Assume a given Gram matrix can be block diagonalized due to symmetries of the problem, as it is the case in [2]. Assume that after block-diagonalization we are again left with S blocks of maximal size N . Each Gram block k is indexed by a polynomial basis

$$b_i^{(k)} = \sum_{\alpha \in SN} b_{\alpha,i}^{(k)} v_{\alpha} \quad (102)$$

where $(v_{\alpha})_{\alpha=1}^{SN}$ indexes the Gram matrix G_0 , e.g., constitutes the monomial indexing basis before block diagonalization. The naive projection would be to transform the large block diagonal matrix back to the monomial basis, losing the block diagonal structure and ending up with complexity $\mathcal{O}(S^2 N^2)$ according to the previous dense argument. In that case, the algorithm is quadratic in S while the SDP solving cost is linear in S . Thus, in the case of many small blocks, performing the projection in the monomial basis might become more costly than solving the SDP.

Secondly, one can perform the projection directly in the polynomial basis as in Lemma 9. Defining all variables accordingly, the numerical certificate reads

$$\mathcal{N}(f - \lambda 1)_t \approx \sum_{i,j,k} \tilde{n}_{ijk}^t G_{ij}^{(k)} \quad (103)$$

Now, in case of the underlying ideal $\mathcal{I}$ being binomial, the map n_{ijk}^t reads

$$\tilde{n}_{ijk}^t = \sum_{\alpha,\beta}^{SN} (b_{i,\alpha}^{(k)})^* b_{j,\beta}^{(k)} \delta_{\mathcal{N}(\alpha^*\beta),t} \quad (104)$$

In the most general case, each indexing polynomial $b_i^{(k)}$ can have the whole monomial basis of

size SN as its support. This leads to a complexity of building $\tilde{n}_{ijk}^t$ in a polynomial basis of $\mathcal{O}(S^2N^2)$, which for large S dominates the linear cost in S per SDP iteration.

The complexity argument above also holds for the establishment of numerical SOHS certificates involving symmetries from SDP data and is hence independent of our rationalization scheme. This implies that any SOHS-based proof for non-negativity in symmetry-adapted bases might become computationally more complex than solving the underlying SDP, which constitutes a general bottleneck for SDP assisted proofs for non-negativity in symmetry-adapted scenarios.

Note that besides the asymptotic argument, there is an associated cost per arithmetic operation performed in rational numbers. Depending on the rounding accuracy, the rationalized Gram matrix entries will have numerators and denominators of large bit size. Depending on the size of the rational numbers, rational arithmetic can become orders of magnitude more expensive than the associated operations using floating point numbers. This introduces a degree of freedom, where one can trade rounding accuracy and hence quality of the result for computational complexity.

In addition to performing the projection, to obtain a certified bound λ_d^{rat} , one needs to compute the minimal eigenvalues of the projected Gram matrices, where in the sparse case this results in complexity $\mathcal{O}(SN^3)$, which is dominated by the overall SDP complexity. In the symmetry-adapted scenario, while the projection complexity is independent of the choice of basis, the eigenvalue computation complexity does heavily depend on the choice: Performing the eigenvalue computation in the inflated monomial basis results in a complexity of $\mathcal{O}(S^3N^3)$ while computing the eigenvalues in the block-diagonal basis has complexity $\mathcal{O}(SN^3)$. Hence, for large S the eigenvalue computation in the monomial basis might become more costly than solving the underlying SDP, while the computation in the symmetry-adapted basis is always cheaper. Hence, from a computational point of view, it is preferable to perform the certification scheme in the symmetry-adapted basis than translating back to the monomial basis.

These asymptotic considerations are consistent with our numerical observations: In the Bell ex-

ample, projection and eigenvalue costs remain negligible relative to the SDP solving, while in the many-body scenario the combination of high rounding precision and a large number of small blocks lead to the projection eventually becoming the bottleneck of the computation.

References

- [1] Miguel Navascués, Stefano Pironio, and Antonio Acín. “Bounding the set of quantum correlations”. *Phys. Rev. Lett.* **98**, 010401 (2007).
- [2] Jie Wang, Jacopo Surace, Irénée Frérot, Benoît Legat, Marc-Olivier Renou, Victor Magron, and Antonio Acín. “Certifying ground-state properties of many-body systems”. *Physical Review X* **14**, 031006 (2024).
- [3] Mateus Araújo, Cyril Branciard, Fabio Costa, Adrien Feix, Christina Giarmatzi, and Časlav Brukner. “Witnessing causal nonseparability”. *New Journal of Physics* **17**, 102001 (2015).
- [4] Armin Tavakoli, Emmanuel Zambrini Cruzeiro, Roope Uola, and Alastair A. Abbott. “Bounding and Simulating Contextual Correlations in Quantum Theory”. *PRX Quantum* **2**, 020334 (2021).
- [5] Miguel Navascués, Stefano Pironio, and Antonio Acín. “A convergent hierarchy of semidefinite programs characterizing the set of quantum correlations”. *New Journal of Physics* **10**, 073013 (2008).
- [6] Stefano Pironio, Miguel Navascués, and Antonio Acín. “Convergent relaxations of polynomial optimization problems with noncommuting variables”. *SIAM Journal on Optimization* **20**, 2157–2180 (2010).
- [7] Igor Klep, Victor Magron, and Janez Povh. “Sparse noncommutative polynomial optimization”. *Mathematical Programming* **193**, 789–829 (2022).
- [8] Jie Wang and Victor Magron. “Exploiting term sparsity in noncommutative polynomial optimization”. *Computational Optimization and Applications* **80**, 483–521 (2021).
- [9] Cordian Riener, Thorsten Theobald, Lina Jansson Andrén, and Jean B. Lasserre. “Exploiting symmetries in SDP-relaxations

- for polynomial optimization”. *Mathematics of Operations Research* **38**, 122–141 (2013).
- [10] Hamza Fawzi, Omar Fawzi, and Samuel O. Scalet. “Certified algorithms for equilibrium states of local quantum Hamiltonians”. *Nature Communications* **15**, 7394 (2024).
- [11] Mateus Araújo, Igor Klep, Andrew J. P. Garner, Tamás Vértesi, and Miguel Navascués. “First-order optimality conditions for non-commutative optimization problems”. *Foundations of Computational Mathematics* (2026). [arXiv:2311.18707](https://arxiv.org/abs/2311.18707).
- [12] Sabine Burgdorf, Kristijan Cafuta, Igor Klep, and Janez Povh. “The tracial moment problem and trace-optimization of polynomials”. *Mathematical Programming* **137**, 557–578 (2013).
- [13] Igor Klep, Victor Magron, Jurij Volčič, and Jie Wang. “State polynomials: positivity, optimization and nonlinear Bell inequalities”. *Mathematical Programming* **207**, 645–691 (2024).
- [14] Mateus Araújo, Andrew J. P. Garner, and Miguel Navascués. “Non-commutative optimization problems with differential constraints”. *Quantum* **10**, 2169 (2026). [arXiv:2408.02572](https://arxiv.org/abs/2408.02572).
- [15] Flavio Baccari, Christian Gogolin, Peter Wittek, and Antonio Acín. “Verifying the output of quantum optimizers with ground-state energy lower bounds”. *Phys. Rev. Res.* **2**, 043163 (2020).
- [16] Bissan Ghaddar, Juan C. Vera, and Miguel F. Anjos. “A dynamic inequality generation scheme for polynomial programming”. *Mathematical Programming* **156**, 21–57 (2016).
- [17] Monique Laurent. “Sums of squares, moment matrices and optimization over polynomials”. In Mihai Putinar and Seth Sullivant, editors, *Emerging applications of algebraic geometry. Volume 149 of The IMA Volumes in Mathematics and its Applications*, pages 157–270. Springer, New York, NY (2009).
- [18] J. W. Helton and Scott McCullough. “A Positivstellensatz for non-commutative polynomials”. *Transactions of the American Mathematical Society* **356**, 3721–3737 (2004).
- [19] Andrew C. Doherty, Yeong-Cherng Liang, Ben Toner, and Stephanie Wehner. “The quantum moment problem and bounds on entangled multi-prover games”. In *Proceedings of the 2008 IEEE 23rd Annual Conference on Computational Complexity. Pages 199–210. CCC ’08USA* (2008). IEEE Computer Society.
- [20] Sabine Burgdorf, Igor Klep, and Janez Povh. “Optimization of Polynomials in Non-Commuting Variables”. *SpringerBriefs in Mathematics*. Springer International Publishing, Cham (2016).
- [21] Victor Magron and Jie Wang. “Sparse polynomial optimization: theory and practice”. *World Scientific*. (2023).
- [22] Andrew C. Doherty, Pablo A. Parrilo, and Federico M. Spedalieri. “Complete family of separability criteria”. *Phys. Rev. A* **69**, 022308 (2004).
- [23] Xiao-Dong Yu, Timo Simnacher, Nikolai Wyderka, H. Chau Nguyen, and Otfried Gühne. “A complete hierarchy for the pure state marginal problem in quantum mechanics”. *Nature Communications* **12**, 1012 (2021).
- [24] Marco Cattaneo, Matteo A. C. Rossi, Keijo Korhonen, Elsi-Mari Borrelli, Guillermo García-Pérez, Zoltán Zimborás, and Daniel Cavalcanti. “Self-consistent quantum measurement tomography based on semidefinite programming”. *Phys. Rev. Res.* **5**, 033154 (2023).
- [25] Alejandro Pozas-Kerstjens, Rafael Rabelo, Łukasz Rudnicki, Rafael Chaves, Daniel Cavalcanti, Miguel Navascués, and Antonio Acín. “Bounding the sets of classical and quantum correlations in networks”. *Phys. Rev. Lett.* **123**, 140503 (2019).
- [26] Hirokazu Anai. “On solving semidefinite programming by quantifier elimination”. *RIMS Kôkyûroku* **1038**, 154–162 (1998). [url: https://www.kurims.kyoto-u.ac.jp/~kyodo/kokyuroku/contents/pdf/1038-21.pdf](https://www.kurims.kyoto-u.ac.jp/~kyodo/kokyuroku/contents/pdf/1038-21.pdf).
- [27] Christopher W. Brown and James H. Davenport. “The complexity of quantifier elimination and cylindrical algebraic decomposition”. In *Proceedings of the 2007 international symposium on Symbolic and algebraic computation. Pages 54–60. Waterloo Ontario Canada* (2007). ACM.
- [28] Yurii Nesterov and Arkadii Nemirovskii. “Interior-point polynomial algorithms in con-

- vex programming”. *Number 13 in SIAM Studies in Applied and Numerical Mathematics*. SIAM. Philadelphia, Pa (1994).
- [29] Victor Magron, Mohab Safey El Din, and Markus Schweighofer. “Algorithms for weighted sum of squares decomposition of non-negative univariate polynomials”. *Journal of Symbolic Computation* **93**, 200–220 (2019).
 - [30] Victor Magron, Mohab Safey El Din, Markus Schweighofer, and Trung Hieu Vu. “Exact SOHS decompositions of trigonometric univariate polynomials with Gaussian coefficients”. In *Proceedings of the 2022 International Symposium on Symbolic and Algebraic Computation*. Pages 325–332. (2022).
 - [31] Przemysław Koprowski, Victor Magron, and Tristan Vaccon. “Pourchet’s theorem in action: decomposing univariate nonnegative polynomials as sums of five squares”. In *Proceedings of the 2023 International Symposium on Symbolic and Algebraic Computation*. Pages 425–433. (2023).
 - [32] Helfried Peyrl and Pablo A Parrilo. “Computing sum of squares decompositions with rational coefficients”. *Theoretical Computer Science* **409**, 269–281 (2008).
 - [33] Erich L. Kaltofen, Bin Li, Zhengfeng Yang, and Lihong Zhi. “Exact certification in global polynomial optimization via sums-of-squares of rational functions with rational coefficients”. *Journal of Symbolic Computation* **47**, 1–15 (2012).
 - [34] Victor Magron and Mohab Safey El Din. “On exact Reznick, Hilbert-Artin and Putinar’s representations”. *Journal of Symbolic Computation* **107**, 221–250 (2021).
 - [35] Victor Magron and Mohab Safey El Din. “RealCertify: a Maple package for certifying non-negativity”. *ACM Communications in Computer Algebra* **52**, 34–37 (2018).
 - [36] V. Magron, X. Allamigeon, S. Gaubert, and B. Werner. “Formal proofs for Nonlinear Optimization”. *Journal of Formalized Reasoning* **8**, 1–24 (2015). [arXiv:1404.7282](https://arxiv.org/abs/1404.7282).
 - [37] Antoine Oustry, Claudia d’Ambrosio, Leo Liberti, and Manuel Ruiz. “Certified and accurate SDP bounds for the ACOPF problem”. *Electric Power Systems Research* **212**, 108278 (2022).
 - [38] Jessica Bavaresco, Mio Murao, and Marco Túlio Quintino. “Strict hierarchy between parallel, sequential, and indefinite-causal-order strategies for channel discrimination”. *Phys. Rev. Lett.* **127**, 200504 (2021).
 - [39] Victor Magron, Mohab Safey El Din, and Trung-Hieu Vu. “Sum of squares decompositions of polynomials over their gradient ideals with rational coefficients”. *SIAM Journal on Optimization* **33**, 63–88 (2023).
 - [40] Lorenzo Baldi, Teresa Krick, and Bernard Mourrain. “An effective positivstellensatz over the rational numbers for finite semi-algebraic sets”. *Mathematics of Computation* (2025).
 - [41] Nando Leijenhorst and David de Laat. “Solving clustered low-rank semidefinite programs arising from polynomial optimization”. *Mathematical Programming Computation* **16**, 503–534 (2024).
 - [42] Maho Nakata. “A numerical evaluation of highly accurate multiple-precision arithmetic version of semidefinite programming solver: SDPA-GMP, -QD and -DD”. In *Proceedings of the IEEE International Symposium on Computer-Aided Control System Design*. Pages 29–34. (2010).
 - [43] Kristijan Cafuta, Igor Klep, and Janez Povh. “Rational sums of hermitian squares of free noncommutative polynomials”. *Ars Mathematica Contemporanea* **9**, 243–259 (2015).
 - [44] Fredrik Johansson. “Arb: efficient arbitrary-precision midpoint-radius interval arithmetic”. *IEEE Transactions on Computers* **66**, 1281–1292 (2017).
 - [45] MOSEK ApS. “MOSEK optimizer api for Julia 11.0.30”. (2025). url: <https://docs.mosek.com/11.0/juliaapi/index.html>.
 - [46] John F. Clauser, Michael A. Horne, Abner Shimony, and Richard A. Holt. “Proposed experiment to test local hidden-variable theories”. *Phys. Rev. Lett.* **23**, 880–884 (1969).
 - [47] Tsuyoshi Ito, Hiroshi Imai, and David Avis. “Bell inequalities stronger than the clauser-horne-shimony-holt inequality for three-level isotropic states”. *Phys. Rev. A* **73**, 042109 (2006).
 - [48] Timotej Hrga, Igor Klep, and Janez Povh. “Certifying Optimality of Bell Inequality Violations: Noncommutative Polynomial Optimization through Semidefinite Programming

- and Local Optimization”. *SIAM Journal on Optimization* **34**, 1341–1373 (2024).
- [49] Károly F Pál and Tamás Vértesi. “Quantum Bounds on Bell inequalities”. *Physical Review A* **79**, 022120 (2009).
- [50] Károly F. Pál and Tamás Vértesi. “Maximal violation of a bipartite three-setting, two-outcome Bell inequality using infinite-dimensional quantum systems”. *Phys. Rev. A* **82**, 022116 (2010).
- [51] Antonio Acín, Serge Massar, and Stefano Pironio. “Randomness versus nonlocality and entanglement”. *Phys. Rev. Lett.* **108**, 100402 (2012).
- [52] Nicolas Gigena, Ekta Panwar, Giovanni Scala, Mateus Araújo, Máté Farkas, and Anubhav Chaturvedi. “Self-testing tilted strategies for maximal loophole-free non-locality”. *npj Quantum Information* **11**, 82 (2025).
- [53] Alexander Weiße and Holger Fehske. “Exact Diagonalization Techniques”. In *Computational Many-Particle Physics*. Pages 529–544. Springer Berlin Heidelberg (2008).
- [54] Dian Wu, Riccardo Rossi, Filippo Vicentini, Nikita Astrakhantsev, Federico Becca, Xiaodong Cao, Juan Carrasquilla, Francesco Ferrari, Antoine Georges, Mohamed Hibat-Allah, Masatoshi Imada, Andreas M. Läuchli, Guglielmo Mazzola, Antonio Mezzacapo, Andrew Millis, Javier Robledo Moreno, Titus Neupert, Yusuke Nomura, Jannes Nys, Olivier Parcollet, Rico Pohle, Imelda Romero, Michael Schmid, J. Maxwell Silvester, Sandro Sorella, Luca F. Tocchio, Lei Wang, Steven R. White, Alexander Wietek, Qi Yang, Yiqi Yang, Shiwei Zhang, and Giuseppe Carleo. “Variational benchmarks for quantum many-body problems”. *Science* **386**, 296–301 (2024).
- [55] J. Ignacio Cirac, David Pérez-García, Norbert Schuch, and Frank Verstraete. “Matrix product states and projected entangled pair states: Concepts, symmetries, theorems”. *Rev. Mod. Phys.* **93**, 045003 (2021).
- [56] Jie Wang, David Jansen, Irénée Frérot, Marc-Olivier Renou, Victor Magron, and Antonio Acín. “Scalable ground-state certification of quantum spin systems via structured noncommutative polynomial optimization”. *SciPost Physics* **21**, 074 (2026). [arXiv:2604.01555](https://arxiv.org/abs/2604.01555).
- [57] Chanchal K. Majumdar and Dipan K. Ghosh. “On Next-Nearest-Neighbor Interaction in Linear Chain. I”. *J. Math. Phys.* **10**, 1388–1398 (1969).
- [58] Aharon Ben-Tal and Arkadi Nemirovski. “Lectures on modern convex optimization: Analysis, algorithms, and engineering applications”. *Society for Industrial and Applied Mathematics*. (2001).