A paradox in bosonic energy computations via semidefinite programming relaxations

M Navascués1,6, A García-Sáez2, A Acín3, S Pironio4 and M B Plenio5

1 School of Physics, University of Bristol, BS8 1TL Bristol, UK
2 Departament d’Estructura i Constituents de la Matèria, Universitat de Barcelona, E-08028 Barcelona, Spain
3 ICFO—Institut de Ciències Fotòniques, E-08860 Castelldefels, Barcelona, Spain
4 Laboratoire d’Information Quantique, Université Libre de Bruxelles, B-1050 Bruxelles, Belgium
5 Institut für Theoretische Physik, Universität Ulm, D-89069 Ulm, Germany
E-mail: m.navascues@bristol.ac.uk

New Journal of Physics 15 (2013) 023026 (21pp)
Received 26 September 2012
Published 15 February 2013
Online at http://www.njp.org/
doi:10.1088/1367-2630/15/2/023026

Abstract. We show that the recent hierarchy of semidefinite programming relaxations based on non-commutative polynomial optimization and reduced density matrix variational methods exhibits an interesting paradox when applied to the bosonic case: even though it can be rigorously proven that the hierarchy collapses after the first step, numerical implementations of higher-order steps generate a sequence of improving lower bounds that converges to the optimal solution. We analyze this effect and compare it with a similar behavior observed in implementations of semidefinite programming relaxations for commutative polynomial minimization. We conclude that the method converges due to the rounding errors occurring during the execution of the numerical program, and show that convergence is lost as soon as computer precision is incremented. We support this conclusion by proving that for any element $p$ of a Weyl algebra which is non-negative in the Schrödinger representation there exists another element $\tilde{p}$ arbitrarily close to $p$ that admits a sum of squares decomposition.

6 Author to whom any correspondence should be addressed.

Content from this work may be used under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.
1. Introduction

Computing the energy spectrum of a finite set of indistinguishable particles subject to a given potential is a standard problem appearing in many branches of physics, e.g. in quantum chemistry, atomic physics or condensed matter physics. Although traditionally the main approaches to this problem have been variational [1], in the last decade it has been attacked with success by means of semidefinite programming (SDP) formulations [2–4] of the constraints on second-order reduced density matrices proposed in [5–7] (the so-called 2-reduced density matrix (2-RDM) method). These SDP methods can be viewed as particular instances of a more general non-commutative polynomial optimization approach [8, 9], which extends to the non-commutative setting the method developed by Lasserre [10] and Parrilo [11] for scalar polynomial optimization. Roughly speaking, a non-commutative optimization problem consists in finding the minimal eigenvalue of a Hermitian polynomial of non-commutative operators. To solve such problems, one can define a hierarchy of SDP relaxations, each of which corresponds to finding a sum of squares (SOS) decomposition of the polynomial to be minimized which provides a lower bound $p^k$ on the optimal solution $p^*$ of the original problem, with $p^1 \leq p^2 \leq \cdots \leq p^*$. This approach reduces to the 2-RDM method when applied to fermionic systems, but more generally it is also highly successful, e.g., in characterizing the set of quantum correlations in quantum information science [12]. In recent years, modifications of this algorithm exploiting translational invariance have been proposed independently by Hübener and Barthel [13] and Baumgratz and Plenio [14] as an alternative to variational techniques in condensed matter physics. In [13–15], it was suggested to apply such SDP methods to compute the ground-state energy of bosonic systems, i.e. to find the minimal eigenvalue of Weyl polynomials.

In this paper, we point out that any computer implementation of the SDP hierarchies [8, 9] to bosonic systems will exhibit the non-commutative analogue of an effect already observed in similar algorithms for commutative polynomial minimization [16, 17]. On the one hand, it
can be proven that any relaxation beyond the first one will not provide better lower bounds on $p^*$, i.e. $p^k = p^1$ for all $k \geq 1$. On the other hand, numerically it is observed that the bounds $\hat{p}^1$, $\hat{p}^2$, … output by the computer form an increasing sequence, with $\lim_{k \to \infty} \hat{p}^k = p^*$.

We will show that the resolution of this ‘mathematical paradox’ follows the same lines as the commutative one. Namely, even though there exist positive Weyl polynomials $p$ that do not admit a SOS decomposition, for any such polynomial there exists an arbitrarily small perturbation $p \to \tilde{p} = p + \epsilon g$ such that $\tilde{p}$ can be decomposed as an SOS of Weyl polynomials. The rounding errors introduced by the computer while executing the algorithm correspond to such a perturbation, and so numerical implementation of the SDP method converges to the correct answer of the problem. This is, therefore, an example of a numerical method that converges, not in spite of rounding errors, but because of them.

This paper is structured as follows. In section 2, we provide the basic definitions and facts about Weyl algebras that are used in the remainder of the text. In section 3, we will describe the SDP method and illustrate the paradox with a numerical example. The resolution of the paradox will be given in section 4, where we will prove that any positive Weyl polynomial can be perturbed to an SOS of polynomials. Finally, in section 5 we will present our conclusions.

2. Definitions and basic results on Weyl algebras

2.1. Weyl algebras and Weyl polynomials

A Weyl algebra $\mathcal{W}_n$ is a $\ast$-algebra with $2n$ generators $\bar{a} = (a_1, a_2, \ldots, a_n)$ and $\bar{a}^* = (a_1^*, a_2^*, \ldots, a_n^*)$ satisfying the canonical commutation relations (CCRs)

$$[a_i, a_j] = [a_i^*, a_j^*] = 0, \quad [a_i, a_j^*] = \delta_{ij}, \quad \text{for all } i, j \in \{1, \ldots, n\}. \quad (1)$$

An element $p$ of $\mathcal{W}_n$ is thus a linear combination (with complex coefficients) of words in the $2n$ letters $\bar{a}$ and $\bar{a}^*$. The words and elements of $\mathcal{W}_n$ can also be viewed as monomials and polynomials, respectively, in the $2n$ variables $\bar{a}$ and $\bar{a}^*$.

Using the CCRs, any element $p$ of a Weyl algebra can be brought to the normal form

$$p = \sum_{\bar{s}, \bar{i}} p_{\bar{s}, \bar{i}} (a^*)^{\bar{i}} \bar{a}^{\bar{s}}, \quad (2)$$

where $\bar{s} = (s_1, \ldots, s_n), \bar{i} = (t_1, \ldots, t_n) \in \mathbb{N}^n$, $a^i$ denotes the monomial $a^i = \prod_{l=1}^n a_l^{t_l}$ and similarly $(a^*)^\bar{i} = \prod_{l=1}^n a_l^{*t_l}$, and where $p_{\bar{s}, \bar{i}} \in \mathbb{C}$ are complex coefficients. For instance, the elements $a_1 a_1^*$ and $a_1 a_2 a_2^* a_1^*$, expressed in normal form, are $1 + a_1^* a_1$ and $1 + a_1^* a_1 + a_2^* a_2 + a_1^* a_2^* a_1 a_2$, respectively. We will later show that the decomposition (2) is unique for each $p \in \mathcal{W}_n$. To check whether two different polynomials $p, p'$ in the variables $\bar{a}, \bar{a}^*$ represent the same element $p = p' \in \mathcal{W}_n$ it is therefore enough to write $p, p'$ in normal form.

This last observation suggests a natural norm in $\mathcal{W}_n$: let $p \in \mathcal{W}_n$, and let (2) be its normal decomposition. Then, we define the $l_1$-norm of $p$ as

$$l_1(p) = \sum_{\bar{s}, \bar{i}} |p_{\bar{s}, \bar{i}}|. \quad (3)$$

It is also useful to distinguish the elements of $\mathcal{W}_n$ by resorting to the concept of degree. We say that the degree of an element $p$ of $\mathcal{W}_n$ is $\deg(p) = \max\{|\bar{s}| + |\bar{i}| : p_{\bar{s}, \bar{i}} \neq 0\}$ for $p$ expressed in normal form (2). It is easy to see that the degree of any monomial $s = \prod_{k=1}^d s_k$ with $s_k \in \{a_j, a_j^*\}_{j=1}^n$ is equal to $\deg(s) = |s| = d$.  

New Journal of Physics 15 (2013) 023026 (http://www.njp.org/)
Given a monomial $s$, we denote by $\frac{s}{2} \frac{1}{2}$ its anti-normal ordering, that is, the monomial $s'$ that results when we reorder the letters appearing in the expression of $s$ in such a way that all the letters $a_1, a_2, \ldots, a_n$ end up on the left and all the letters $(a_1^*, a_2^*, \ldots, a_n^*)$ end up on the right. For example, $\frac{s}{2} a_2^* (a_1^*)^2 a_3 = \frac{s}{2} a_1^* a_1^* a_2^* = a_1^* (a_1^*)^2 a_2^*$.

In this paper, we will be mainly concerned with Hermitian elements of $\mathcal{W}_n$, i.e. those polynomials $p$ such that $p = p^*$. If $p$ is decomposed as in (2), the hermiticity condition thus translates as $p_{i,j} = p^*_{j,i}$, where $p^*_{i,j}$ denotes the complex conjugate of the number $p_{i,j}$.

2.2. The Schrödinger representation

Let $\pi$ be a mapping $\pi : \mathcal{W}_n \to L(\mathcal{H})$, for some Hilbert space $\mathcal{H}$, where $L(\mathcal{H})$ denotes the space of linear (not necessarily bounded) operators of $\mathcal{H}$. We say that $\pi$ is a $\ast$-representation of $\mathcal{W}_n$ if and only if

1. $\pi(1) = \mathbb{1}_\mathcal{H}$.
2. $\pi(p q) = \pi(p) \pi(q), \pi(p + q) = \pi(p) + \pi(q)$ for all $p, q \in \mathcal{W}_n$.
3. $\pi(p^*) = \pi(p)^*$ for all $p \in \mathcal{W}_n$.

We now show that $\mathcal{W}_n$ admits a representation. For this, let $H$ be a separable Hilbert space, and let $\{|s\rangle, s \in \mathbb{N}\}$ be an orthonormal basis for $H$, which we call the number basis. If we denote by $\tilde{a} \in L(H)$ the linear operator defined by

$$\tilde{a} |s\rangle = \begin{cases} 0 & \text{for } s = 0, \\ \sqrt{s} |s - 1\rangle & \text{otherwise}, \end{cases}$$

then its adjoint $\tilde{a}^*$ satisfies

$$\tilde{a}^* |s\rangle = \sqrt{s + 1} |s + 1\rangle$$

and so it can be verified that

$$[\tilde{a}, \tilde{a}^*]|s\rangle = |s\rangle$$

for all $s \in \mathbb{N}$.

Defining $\mathcal{H} = H^{\otimes n}$, we can then build a representation $\pi_S : \mathcal{W}_n \to L(\mathcal{H})$ for the Weyl algebra $\mathcal{W}_n$ through

$$\pi_S(a_i) = \mathbb{1}^{\otimes k - 1} \otimes \tilde{a} \otimes \mathbb{1}^{\otimes n - k}.$$  

This representation of $\mathcal{W}_n$ is known as the Schrödinger (or Fock) representation. From now on, we always refer to this representation and write $\pi$ for $\pi_S$ for simplicity.

2.3. Weyl polynomial minimization

The Schrödinger representation admits a clear physical interpretation: given a set of $n$ one-dimensional particles, it associates with each particle $k \in \{1, \ldots, n\}$ a pair of creation and annihilation operators $\pi(a_k), \pi(a_k^*)$. The operators describing the position and momentum of particle $k$ along the real line are then given, respectively, by $\pi(x_k) \equiv \pi\left(\frac{a_k + a_k^*}{\sqrt{2}}\right)$ and $\pi(p_k) \equiv \pi\left(\frac{a_k - a_k^*}{\sqrt{2m_i}}\right)$. If these particles are subject to a potential $V(x_1, \ldots, x_n)$, the energy operator of the system, in non-relativistic approximation, will be given by

$$\pi(E) = \pi\left(\sum_{i=1}^{n} \frac{p_i^2}{2m_i} + V(x_1, \ldots, x_n)\right)$$

$\text{New Journal of Physics} \ 15 \ (2013) \ 023026 \ (\text{http://www.njp.org/})$
where $m_i \in \mathbb{R}^+$ is the mass of particle $i$. In particular, the minimum energy of the system will be given by

$$\lambda_{\text{inf}}(E) = \inf \{ \langle \phi | \pi(E) | \phi \rangle : |\phi\rangle \in \mathcal{S}, \langle \phi | \phi \rangle = 1 \},$$

(8)

where $\mathcal{S}$ is the Schwartz space, that is, the set of states $|\phi\rangle$ (i.e. vectors of $\mathcal{H}$) satisfying $\| \pi(p) |\phi\rangle \| < \infty$ for all $p \in \mathcal{W}_n$. Note that the minimization over $\langle \phi | E | \phi \rangle$ makes sense, because the energy is a Hermitian operator, $E = E^*$, and consequently, $\langle \phi | E | \phi \rangle \in \mathbb{R}$ is a real quantity for all $|\phi\rangle \in \mathcal{S}$.

The right-hand side is a polynomial in the complex variables $\bar{p}$ for some coefficients $a_j$. As an application of the coherent states, let us show that the decomposition (2) is unique, or, equivalently, that $0$ admits a unique representation. Suppose thus that

$$0 = \sum_{\tilde{s}, \tilde{t}} p_{\tilde{s}, \tilde{t}} a^{s \tilde{s}} a^{\tilde{t}}$$

(13)

for some coefficients $p_{\tilde{s}, \tilde{t}} \in \mathbb{C}$. Using the Schrödinger representation, we have that

$$0 = \langle \tilde{\alpha} | \pi(0) | \tilde{\alpha} \rangle = \sum_{\tilde{s}, \tilde{t}} p_{\tilde{s}, \tilde{t}} \alpha^{s \tilde{s}} \alpha^{\tilde{t}}$$

for all $\tilde{\alpha} \in \mathbb{C}^n$.

The right-hand side is a polynomial in the complex variables $\tilde{\alpha}, \tilde{\alpha}^*$, and it can only be equal to zero for all values of $\tilde{\alpha}$ if $p_{\tilde{s}, \tilde{t}} = 0$ for all $\tilde{s}, \tilde{t}$. 

---

New Journal of Physics 15 (2013) 023026 (http://www.njp.org/)
3. Sum-of-squares decompositions, the semidefinite programming hierarchy and the paradox

Given a Hermitian polynomial \( p \in \mathcal{W}_n \), we say that \( p \) admits an SOS decomposition if there exist polynomials \( f_i \in \mathcal{W}_n \) such that

\[
p = \sum_i f_i^* f_i.
\]  

(15)

We denote by \( \Sigma^2 \) the set of all such polynomials. It is clear that if \( p \in \Sigma^2 \), then \( \pi(p) \geq 0 \), since, for any \( |\phi\rangle \in \mathcal{S} \),

\[
\langle \phi | \pi(p) | \phi \rangle = \sum_i \langle \phi | \pi(f_i^*) \pi(f_i) | \phi \rangle \geq 0.
\]  

(16)

However, the opposite implication is not true, not even in \( \mathcal{W}_1 \). Indeed, as noted by Schmüdgen [18], the family of polynomials \( p_\epsilon = (a_1^* a_1 - 1)(a_1 a_1^* - 2) + \epsilon \) satisfies \( \pi(p_\epsilon) \geq 0 \), for \( \epsilon \geq 0 \), but nevertheless \( p_\epsilon \notin \Sigma^2 \), for \( \epsilon < \frac{1}{4} \).

Given a \( p \in \mathcal{W}_n \), a possible scheme for finding a lower bound \( \lambda^* \) on the solution \( \lambda_{inf}(p) \) of (10) is thus to solve the problem

\[
\lambda^* \equiv \sup \{ \lambda \in \mathbb{R} : p - \lambda \in \Sigma^2 \}.
\]  

(17)

This principle is the one behind the polynomial minimization algorithms developed by Lasserre [10] and Parrilo [11], and their non-commutative analogue [8, 9]. Such algorithms work by searching for SOS decompositions of \( p - \lambda \) with some degree constraint. Applied to the Weyl minimization problem, this results in the following sequence of programs:

\[
\lambda^k \equiv \sup \{ \lambda \in \mathbb{R} : p - \lambda \in \Sigma^2_k \}.
\]  

(18)

Here \( k \) is an integer such that \( 2k \geq \text{deg}(p) \) and indexing the successive programs in the sequence and \( \Sigma^2_k \) is the set of polynomials which admit a decomposition of the form (15) with \( \text{deg}(f_i) \leq k \), for all \( i \). Each of these problems is a semidefinite program, as one can check that

\[
\lambda^k = \max \{ \lambda \in \mathbb{R} : p - \lambda = (\bar{w}^k)^* Z \bar{w}^k, Z \geq 0 \},
\]  

(19)

where \( \bar{w}^k \) is a vector whose components are the normally ordered monomials of degree \( \leq k \). Clearly, \( \lambda^k \leq \lambda^{k+1} \) and \( \lim_{k \to \infty} \lambda^k = \lambda^* \). The programs (18) and (19) thus form a converging hierarchy of SDP relaxations for the problem (17). Supplemented with a boundedness condition (that is not satisfied in the present case of Weyl polynomials), it can further be shown that this hierarchy necessarily converges to the optimal solution of the problem (10) [8, 9], as problems (17) and (10) then turn out to be equivalent [19].

Let \( L \) denote an arbitrary functional on the Weyl algebra, i.e. \( L : \mathcal{W}_n \to \mathbb{C} \). Then, the dual of problems (18) and (19) can be shown to be

\[
\mu^k = \min \{ L(p) : L(1) = 1, L(qq^*) \geq 0 \text{ for all } q \in \mathcal{W}_n, \text{deg}(q) \leq k \}
\]  

(20)

or explicitly in SDP form

\[
\mu^k = \min \left\{ \sum_{\tilde{i}, \tilde{j}} p_{\tilde{i}, \tilde{j}} y_{\tilde{i}, \tilde{j}} : y_{0,0} = 1, M_k(y) \geq 0 \right\}.
\]  

(21)
In this last formulation, \( p_{i,j} \) are the coefficients of \( p \) in normal form \( 2 \), \( y = \{ y_{i,j} : \| \bar{s} + \bar{r} \|_1 \leq 2k \} \subset \mathbb{C} \) are the optimization variables, and \( M_k(y) \) is the moment matrix of order \( k \), a matrix whose rows and columns are indexed by pairs of vectors \( (\bar{s}, \bar{r}) \in \mathbb{N}^n \times \mathbb{N}^n \), with \( \| \bar{s} + \bar{r} \|_1 \leq k \), and with entries defined by
\[
[M_k(y)]_{(\bar{s}, \bar{r}), (\bar{a}, \bar{v})} = \sum_{\bar{q}, \bar{r}} c_{\bar{q}, \bar{r}} y_{\bar{q}, \bar{r}},
\]
(22)
where \( c_{\bar{q}, \bar{r}} \) are the normal form coefficients of the monomial \( (a^{*} \bar{s} \bar{a}^{*})^{\bar{a}} a^{\bar{v}} \), i.e.
\[
(a^{*} \bar{s} \bar{a}^{*})^{\bar{a}} a^{\bar{v}} = \sum_{\bar{q}, \bar{r}} c_{\bar{q}, \bar{r}} a^{\bar{q}} a^{\bar{r}}.
\]
(23)

The next lemma shows that problems \( (18) \)–\( (21) \) are, in fact, equivalent:

**Lemma 1.** If there exists a feasible point of program \( (19) \), then \( \mu^k = \lambda^k \).

**Proof.** By Sylvester’s criterion \( 20 \), it just suffices to show that problem \( (20) \) admits a strictly feasible point, i.e. that there exists a functional \( L \) such that \( L(1) = 1 \) and \( L(pp^*) > 0 \), for all \( p \neq 0 \). Now, consider the functional \( L(p) = \text{tr} (\Omega \pi(p)) \), with
\[
\Omega = \frac{1}{(2\pi)^n} \int d\bar{a} e^{-|\bar{a}|^2} |\bar{a} \rangle \langle \bar{a}|.
\]
(24)

This functional satisfies \( L(1) = 1 \). Also, for any non-zero polynomial \( p \neq 0 \in \mathcal{W}_n \) we have that
\[
L(p^* p) = \frac{1}{(2\pi)^n} \int d\bar{a} e^{-|\bar{a}|^2} \langle \bar{a} | \pi(p^* p) | \bar{a} \rangle \geq \frac{1}{(2\pi)^n} \int d\bar{a} e^{-|\bar{a}|^2} |\langle \bar{a} | \pi(p) | \bar{a} \rangle|^2
\]
\[
= \frac{1}{(2\pi)^n} \int d\bar{a} e^{-|\bar{a}|^2} |p(\bar{a}, \bar{a}^*)|^2 > 0.
\]
(25)

The last inequality comes from the fact that \( p(\bar{a}, \bar{a}^*) \neq 0 \) and that the integration takes place in all \( \mathbb{C}^n \).

Even though the above SDP hierarchy cannot be guaranteed to converge to the optimal value \( \lambda_{\text{inf}}(p) \) of \( (10) \), every SDP step provides a lower bound on \( \lambda_{\text{inf}}(p) \). Based on the successful applications of this SDP hierarchy to fermionic systems and quantum correlations, where in practice very good lower bounds are obtained after only a few SDP relaxations, one could expect a good overall performance also in the context of Weyl polynomials. However, as the next result shows, no improvement over the first lower bound can be obtained by considering higher steps in the hierarchy.

**Lemma 2.** Let \( p \in \Sigma^2 \). Then, \( p \in \Sigma^k \) for \( 2k \leq \deg(p) \).

**Proof.** By hypothesis, \( p \in \Sigma^2 \), and thus there exist polynomials \( f_i \) such that \( (3) \) holds. We will show that all such polynomials satisfy \( 2 \deg(f_i) \leq \deg(p) \).

---

Note that, if \( p_{i,j} \in \mathbb{R} \), \( \forall \bar{s}, \bar{r} \), one can assume \( \{ y_{i,j} : \| \bar{s} + \bar{r} \|_1 \leq 2k \} \subset \mathbb{R} \).

*New Journal of Physics* 15 (2013) 023026 (http://www.njp.org/)
Suppose, in contrast, that
\[ p = \sum f_i^* f_i - f^* f \]  
with \( 2 \deg(f) > \deg(p) \). Then
\[ p(\tilde{a}, \tilde{a}^*) = \langle \tilde{a} | \pi(p) | \tilde{a} \rangle \geq \langle \tilde{a} | \pi(f^*) \pi(f) | \tilde{a} \rangle \geq |\langle \tilde{a} | \pi(f) | \tilde{a} \rangle|^2 = |f(\tilde{a}, \tilde{a}^*)|^2. \]

Now, denote by \( LT(f) \) the leading terms of \( f \), i.e.
\[ LT(f)(\tilde{a}, \tilde{a}^*) = \sum_{|\tilde{\alpha} + \tilde{\beta}| = \deg(f)} f_{\tilde{\alpha}, \tilde{\beta}} (\tilde{\alpha}^*)^* \tilde{\beta}^* \]
and choose \( \tilde{\beta} \in \mathbb{C}^n \) such that \( LT(f)(\tilde{\beta}, \tilde{\beta}^*) = c \neq 0 \). Then it is straightforward that \( |f(r\tilde{\beta}, r\tilde{\beta}^*)| = O(r^{\deg(f)}) \). However, \( |p(r\tilde{\beta}, r\tilde{\beta}^*)| \leq O(r^{\deg(p)}) \). For (27) to hold, we must thus have that \( 2 \deg(f) \leq \deg(p) \).

\[ \square \]

What lemma 2 shows is that, for any polynomial \( p \), the sequence of values \( \lambda_{k_0}, \lambda_{k_0+1}, \ldots \), with \( k_0 = \lceil \deg(p)/2 \rceil \) is constant and equal to \( \lambda^* \). In other words, the first SDP relaxation of the problem already provides the best approximation to \( \lambda_{\min}(p) \) attainable with SOS decompositions.

How does such an approximation perform? Consider the uniparametric family of one-dimensional Hamiltonians \( \{ E_{m} : m \in \mathbb{R} \} \), with
\[ E_{m} = \frac{1}{2} \tilde{p}^2 + m\tilde{x}^2 + \tilde{x}^4 \]  
with \( \tilde{x} \equiv (\tilde{u} + \tilde{u}^*)/\sqrt{2}, \tilde{p} \equiv (\tilde{u} - \tilde{u}^*)/i\sqrt{2} \). Note that, for \( m < 0 \), \( E_{m} \) corresponds to the interesting double-well potential. Figure 1 shows a plot of \( \lambda^2 \) as a function of \( m \), together with an upper bound on \( \lambda_{\min}(E_{m}) \) obtained by an exact diagonalization of \( \pi(E_{m}) \) in the truncated number basis \( \{|s\rangle\}_{s=0}^{100} \). We used the solver SDPT3-4.0 [21] and the MATLAB package YALMIP [22] to carry out the SDP calculations. It is clear that, as soon as \( m < 0 \), the approximation given by \( \lambda^2 \) becomes worse.

From our discussions above, it follows that \( \lambda^2 = \lambda^3 = \lambda^4 = \cdots \), and so \( \lambda^2 \) represents the best lower bound on \( \lambda_{\min}(E_{m}) \) achievable with the SDP hierarchy. Figure 2 shows, however, that this is not the case. Indeed, we see that subsequent relaxations of the problem return lower bounds which are closer and closer to the variational upper bound, until, at \( \lambda^6 \), both bounds become practically indistinguishable. What is happening?

4. Resolution of the paradox

As mentioned in the introduction, the above paradox is not new in commutative polynomial optimization: indeed, Henrion and Lasserre [16] noted that the numerical implementation of their SDP algorithm for polynomial minimization returned the optimal value of the two-dimensional Motzkin polynomial, instead of its SOS value \( -\infty \). Lasserre successfully solved this paradox by proving that any commutative positive polynomial can be approximated arbitrarily well by an SOS decomposition [23]. The accepted resolution of the paradox was that the rounding errors occurring during the numerical computations perturbed the polynomial \( p \) to be minimized to another one of higher degree admitting an SOS decomposition [17].
Figure 1. Bounds on $\lambda_{inf}(E_m)$ as a function of $m$. The upper curve represents an upper bound obtained through variational techniques. The lower curve represents the lower bound $\lambda^2$ corresponding to the first relaxation of the problem.

Figure 2. Lower bounds on $\lambda_{inf}(E_m)$ as a function of $m$ corresponding to the second, third, fourth and fifth SDP relaxations, in ascending order. The black line represents the variational estimate, quite close to $\lambda^6$ (red line).

In this section, we will prove a non-commutative analogue of this result, namely that Weyl polynomials which are positive semidefinite in the Schrödinger representation can be perturbed to a higher degree polynomial in $\Sigma^2$. This is formally stated in the following theorem:

**Theorem 1.** Let $p$ be an element of $\mathcal{W}_n$ such that $\pi(p) \succeq 0$. Then, for any $\epsilon, \delta > 0$, there exists a polynomial $\tilde{p} \in \Sigma^2$ such that $l_1(p - \tilde{p}) < \epsilon$ and $|\lambda_{inf}(\tilde{p}) - \lambda_{inf}(p)| < \delta$. 

*New Journal of Physics* 15 (2013) 023026 (http://www.njp.org/)
The proof of this theorem follows straightforwardly from the next three lemmas. In these lemmas, the constant $c$ is arbitrary but fixed at $c > 2$.

**Lemma 3.** Let $p \in \mathcal{W}_n$ be a polynomial such that $\pi(p) \geq 0$. Let

$$g'_c = \sum_{|\vec{t}| \leq r} c^{||\vec{t}||_1} (n+||\vec{t}||_1-1)! a^\dagger (a^\dagger)^*,$$

where the sum runs over all vectors $\vec{t} \in \mathbb{N}^n$ of length $||\vec{t}||_1$ less that or equal to $r$. Then for any $\epsilon > 0$, there exists some $r_0$ such that for all $r > r_0$,

$$\tilde{p} = p + \epsilon g'_c \in \Sigma^2.$$  \hfill (30)

**Lemma 4.** Let $\tilde{p}'$ be defined as in lemma 3. Then

$$l_1(p - \tilde{p}') \leq \epsilon \cdot \frac{c}{c-2}$$

for any $r > 0$.

**Lemma 5.** Let $p$ be a Hermitian polynomial in $\mathcal{W}_n$ such that $\lambda_{\text{inf}}(p)$ exists, and let $g'_c$ be defined as in (30). Then, for any $\delta > 0$, there exists a number $q \in \mathbb{R}^+$ such that

$$\lambda_{\text{inf}}(p) \leq \lambda_{\text{inf}}(p + \epsilon g'_c) \leq \lambda_{\text{inf}}(p) + \delta + \epsilon q,$$

for all $r, \epsilon > 0$.

The proofs of lemmas 3 and 5 are given here below. The proof of lemma 4 follows from the results presented in appendix A. This lemma implies that $p$ and $\tilde{p}'$ can be made arbitrarily close if we take $\epsilon$ small enough. Moreover, if we express $p - \tilde{p}'$ in the anti-normal form, i.e. $p - \tilde{p}' = \sum_{i,j} p_{i,j}^\alpha (a_i^\dagger)^*$, the corresponding natural norm $l_1(p - \tilde{p}') = \sum_{i,j} |p_{i,j}^\alpha|$ is trivially bounded by $\epsilon \cdot e^{l/c}$. This implies that any computer implementation of program (19) where the polynomial to minimize is expressed in normal or anti-normal form will require a lot of precision in order to distinguish $\tilde{p}'$ from $p$ for low values of $\epsilon$.

**Proof of lemma 3.** The demonstration of this lemma will make use of two lemmas, proven in appendices B and C, respectively. \hfill \Box

**Lemma 6.** Let $s \in \mathcal{W}_n$ be a monomial. Then

$$\hat{x} s s^* \hat{x} - ss^* \in \Sigma^2.$$ \hfill (34)

**Lemma 7.** Let $L$ be a linear functional in $\mathcal{W}_n$. If $L(h^*h) \geq 0$ for any $h \in \mathcal{W}_n$ and there exist $c, d > 0, k \in \mathbb{N}$ such that for any monomial $s \in \mathcal{W}_n$, the relation

$$|L(s)| \leq dc^{|s|} \Gamma \left( \frac{|s| + k + 1}{2} \right)$$

holds, then there exists a non-normalized quantum state $\rho$ (i.e. a non-negative, trace-class operator) such that $L(h) = \text{tr}(\rho \pi(h))$, for any polynomial $h \in \mathcal{W}_n$. 

---

New Journal of Physics 15 (2013) 023026 (http://www.njp.org/)
Let us now proceed with the proof of lemma 3. Following Lasserre and Netzer [24], let \( \mathcal{W}_n(r) \) be the set of elements of \( \mathcal{W}_n \) with degree less than or equal to \( r \), and consider the semidefinite program

\[
\epsilon_r^* = \min \{ L(p) | L : \mathcal{W}_n(2r) \to \mathbb{R}, \text{L linear}, \ L(g_{r}^c) \leq 1, \ L(h^zh) \geq 0, \forall h \in \mathcal{W}_n(r) \}. \tag{36}
\]

Noting that \( L = 0 \) is an admissible linear functional, we have that the problem has feasible points and that \( \epsilon_r^* \leq 0 \). The condition \( L(g_{r}^c) \leq 1 \), together with lemma 6, implies that the diagonal entries of all feasible moment matrices \( M_r(y) \) are upper bounded, and so the absolute values of the rest of the entries, due to positive semidefiniteness. From these two observations, it follows that our problem admits a solution, i.e. \( \epsilon_r^* \neq -\infty \) is attainable for a feasible choice of \( L \).

The dual of (36) is

\[
\max \{ -\epsilon : \epsilon \geq 0, \ p + \epsilon g_{r}^c \in \Sigma^2 \}. \tag{37}
\]

That this problem has solutions for any \( p = p^* \) follows from the fact that, for all \( \mu \in \mathbb{C} \) and any pair of monomials \( s, t \in \mathcal{W}_n \),

\[
\mu s^*t + \mu^*t^*s + |\mu|[s^*s + t^*t] \in \Sigma^2. \tag{38}
\]

And thus, invoking lemma 6,

\[
\mu s^*t + \mu^*t^*s + |\mu|[\bar{s}^*s\bar{t}^*\bar{t}] \in \Sigma^2. \tag{39}
\]

By increasing the value of \( \epsilon \), at some point we will therefore have that \( p + \epsilon g_{r}^c \in \Sigma^2 \).

Moreover, in this particular case, there is no duality gap, i.e. the solutions of both the primal and dual problems coincide. Again, this can be established by invoking the quantum state (24): choosing \( \sigma > 0 \) such that \( \text{tr}\{\Omega_{\pi}(g_{r}^c)\} < 1/\sigma \), it follows that \( L(h) \equiv \sigma \text{tr}(\Omega_{\pi}(h)) \) is a strictly feasible point of (36) and, thus, the solutions of the dual and primal problems are the same [20].

This, together with the fact that \( g_{s}^c - g_{r}^c \) is an SOS for \( s \geq r \), implies that, for all \( \epsilon \geq -\epsilon_r^* \geq 0 \), the polynomial

\[
p + \epsilon g_{r}^c (s \geq r) \tag{40}
\]

is also an SOS.

The sequence \( (\epsilon_r^*)_r \) is, therefore, an increasing one. We will next prove that \( \lim_{r \to \infty} \epsilon_r^* = 0 \), so that the \( \epsilon > 0 \) appearing in the formulation of lemma 3 can be taken as arbitrarily small.

Consider the sequence \( L_r^* \) of functionals that attain the solutions \( \epsilon_r^* \) of the problem, and denote by \( M_r^* \) their corresponding moment matrices (the entries \( (M_r^*)_{(\bar{s},\bar{t}), (\bar{u},\bar{v})} \) where either \( \|\bar{s} + \bar{t}\|_1 > r \) or \( \|\bar{u} + \bar{v}\|_1 > r \) are assumed to be completed with zeros). By lemma 6, we have that

\[
(M_r^*)_{(\bar{s},\bar{t}), (\bar{u},\bar{v})} \leq (M_r^*)_{(\bar{s},\bar{t}, 0), (\bar{u},\bar{v}, 0)} \leq \frac{(n + \|\bar{s} + \bar{t}\|_1 - 1)!}{(n - 1)!} \epsilon_r^* \|\bar{s} + \bar{t}\|_1 =: d(\bar{s}, \bar{t}). \tag{41}
\]

Now, perform the transformation

\[
(M_r^*)_{(\bar{s},\bar{t}), (\bar{u},\bar{v})} \to (N_r^*)_{(\bar{s},\bar{t}), (\bar{u},\bar{v})} \equiv d(\bar{s}, \bar{t})^{1/2}(M_r^*)_{(\bar{s},\bar{t}), (\bar{u},\bar{v})} d(\bar{u}, \bar{v})^{1/2}. \tag{42}
\]

\( N_r^* \) is thus positive semidefinite and its diagonals are upper bounded by 1 for all \( r \); it follows that all the entries of the matrices \( N_r^* \) are in the interval \([1, -1]\). By the Banach–Alaoglu theorem, the sequence \( N_r^* \) admits a subsequence \( \{r_i\} \) that converges in the weak-* topology to a limit \( N^* \to \hat{N} \) when \( i \to \infty \) [25]. Undoing the previous change of coordinates, we are left with an
infinite-sized matrix $\hat{M}$ that defines a linear functional $\hat{L}(p) \equiv \sum_{i,j} P_{(i,j)} \hat{M}_{(0,i),(0,j)}$ on the Weyl algebra.

This functional satisfies $\hat{L}(h^* h) \geq 0$, for any polynomial $h$. Moreover, for any sequence $s = s_1 s_2$, with $|s_1| = \left\lceil \frac{|s|}{2} \right\rceil$, $|s_2| = \left\lfloor \frac{|s|}{2} \right\rfloor$,

$$|\hat{L}(s)| \leq (\hat{L}(s_1^* s_1) \hat{L}(s_2^* s_2))^{1/2} \leq \left( \frac{(n + |s_1| - 1)! (n + |s_2| - 1)!}{(n - 1)! (n - 1)!} \right)^{1/2} \sqrt{c^{|s_1| + |s_2|}} \leq \sqrt{c^{|s|}} \Gamma \left( \frac{2n + 2 + |s|}{2} \right).$$

By lemma 7, this last condition implies that there exists a non-normalized quantum state $\rho \in S_1(\mathcal{H})$ in the Schrödinger representation such that $\hat{L}(h) = \text{tr}(\rho \pi(h))$, for all $h$.

Now

$$\lim_{r \to \infty} \epsilon^*_r = \hat{L}(p) = \text{tr}(\rho \pi(p)) \geq 0,$$

where the last inequality follows from the non-negativity assumption on $p$. On the other hand, $\epsilon^*_r \leq 0 \forall r$, and therefore, we have that $\lim_{r \to \infty} \epsilon^*_r = 0$. 

**Proof of lemma 5.** Given a vector $\tilde{m} \in \mathbb{N}^n$, we will denote by $|\tilde{m}\rangle$ the number state $|m_1\rangle \otimes |m_2\rangle \otimes \cdots \otimes |m_n\rangle$. Now, if $\pi(p) \geq 0$, then $\lambda_{\text{inf}}(p)$ exists and can be written as the limit of a sequence of the form $\langle \phi_i | \pi(p) | \phi_i \rangle$, where $|\phi_i\rangle$ are normalized quantum states. Such states can be, in turn, approximated with arbitrary precision by finite linear combinations of number states. Choose, then, a number $M \in \mathbb{N}$ such that the normalized state $|\Phi\rangle \equiv \sum_{|m| \leq M} d_{\tilde{m}} |\tilde{m}\rangle$ satisfies $\lambda_{\text{inf}}(p) \leq \langle \Phi | \pi(p) | \Phi \rangle \leq \lambda_{\text{inf}}(p) + \delta$.

It can be verified that, for any pair of number states $|\tilde{m}\rangle$, $|\tilde{m}'\rangle$, with $\|\tilde{m}\|_\infty, \|\tilde{m}'\|_\infty \leq M$, and any monomial $s$ of the annihilation and creation operators, the inequality $|\langle \tilde{m} | ss^* | \tilde{m}' \rangle| \leq \frac{(s + M)!}{M!} \delta_{\tilde{m}, \tilde{m}'}$ holds. Therefore,

$$\langle \Phi | ss^* | \Phi \rangle \leq \frac{(|s| + M)!}{M!} \sum_{\tilde{m}} |d_{\tilde{m}}|^2 = \frac{(|s| + M)!}{M!}.$$

Finally, choose $c > 1$. It follows that

$$\langle \Phi | g^r_\epsilon | \Phi \rangle \leq \sum_{l=0}^{\infty} \frac{\sharp \{ \tilde{l} : |\tilde{l}| \leq l \} (l + M)!}{M!} \frac{1}{c^l} = \sum_{l=0}^{\infty} \frac{(l + M)!}{M!} \frac{1}{c^l} = \left( \frac{c}{c - 1} \right)^{M+1},$$

where $\sharp \{ \bullet \}$ denotes the cardinality of the set $\{ \bullet \}$. Here, in order to identify the second and third expressions we made use of proposition (A.2) in appendix A. Equating to $q$ the last result, we arrive at the promised lemma.

Note that lemma 7 alone provides an alternative explanation for the observed convergence to the optimal solution, this time from the point of view of the dual problem (20). The reason why $\mu^h$ does not necessarily converge in theory to $\lambda_{\text{inf}}(p)$ for the exact problem is because the values $L(s)$ in each program grow faster than $dc^{|s|} \Gamma \left( \frac{|s| + M}{2} \right)$. If, however, due to finite numerical
Figure 3. Plot of $\lambda^k - \lambda^2$ as a function of the parameter $\lambda_{\text{Star}}$ in the case $m = -1$.

precision our solvers limit the magnitude of such momenta, low-order relaxations $\mu^k$ should provide a better approximation to $\lambda_\text{inf}$. 8

Increasing the numerical precision of our programs should thus have a twofold effect: on the one hand, it should allow the computer to distinguish between $p$ and its perturbation $\tilde{p}$. On the other hand, it should extend the moment matrix search space to include matrices with entries of very different magnitude. A high-precision numerical computation should, therefore, make the curves in figure 2 collapse to the same line.

We used the SDP solver Semi-Definite Programming Algorithm GNU Multiple Precision (SDPA-GMP) [4, 26] to compute SDP approximations of $E_{-1}$, $E_1$ with a precision of 600 digits. Figures 3 and 4 show the outputs $\lambda^k - \lambda^2$ of both problems as a function of $\lambda_{\text{Star}}$, an internal parameter of SDPA-GMP that constrains the magnitudes of the entries of the moment matrix $^9[26]$. Note that, in agreement with the above interpretation, the difference between $\lambda^k$ and $\lambda^2 = \lambda^\star$ tends to zero as the constraints on the moment matrix disappear (right end of figures 3 and 4). Conversely, the solutions of higher order relaxations start to differ from $\lambda^2$ and become closer to the actual solution of the problem as we restrict the magnitude of the entries of the moment matrix (left end of figures 3 and 4).

The reader may wonder why the above effect is not observed in fermionic Hamiltonian minimization [2–4], where the goal is to minimize polynomials of creation and annihilation operators satisfying the anti-commutation relations $\{a_i, a_j^\dagger\} = 2\delta_{ij}$. The reason is twofold: firstly, taking $i = j$ in the previous expression it follows that these operators are bounded, and so the corresponding sequence of SDP bounds is guaranteed to converge to the ground state energy [8, 9]. Secondly, since these operators have unit norm, the effect of a perturbation $\varepsilon(a_i)^*a_i$ can only affect the value of the objective function by an amount of at most $\pm\varepsilon$. The SDP scheme is thus robust against rounding errors.

8 This does not apply to high-order relaxations corresponding to $k \gg 1$, as a constraint of the form $L(a_i^k(a_i^k)^*) < K$ becomes unimplementable as soon as $K > k!$.
9 More concretely, $(\omega_{\text{Star}}\lambda_{\text{Star}})^I - M_k(y) \geq 0$. In our calculations, we set $\omega_{\text{Star}} = 2$. 

New Journal of Physics 15 (2013) 023026 (http://www.njp.org/)
Figure 4. Plot of $\lambda^k - \lambda^2$ as a function of the parameter lambdaStar in the case $m = 1$.

5. Conclusion

We have identified a paradoxical behavior in the application to bosonic systems of the SDP methods widely used in quantum chemistry energy calculations and quantum information. Namely, we have pointed out that numerical implementations of the method seem to converge despite a simple theoretical argument showing that the first SDP relaxation should already provide the best lower bound to the problem at hand. This phenomenon is similar to an analogous behavior observed in commutative polynomial optimization [16, 17] and we suggested that the paradox arises from rounding errors introduced in the numerical commutation. We provided a theoretical basis for this assumption by proving that for any bosonic Hamiltonian to be minimized there exists a perturbation of it whose ground state energy can accurately be approximated by the SDP method. Furthermore, we showed that the effect disappears as soon as we increase the computer precision.

Our results suggest that the above problem could be avoided by constraining the values of the diagonal elements of the moment matrices in each program. Thanks to such constraints, computer implementations of the primal problem could return reliable solutions without having to resort to extremely high-precision numerical calculations. This approach will be explored in a forthcoming paper.

It is worth noting that Cimprič [27] proposed to use Schmüdgen’s positivstellensatz for Weyl algebras [18] to introduce a different SDP hierarchy than the one presented here in order to find rigorous lower bounds on the minimum value of arbitrary Weyl polynomials. The application of this method, however, requires high-precision SDP solvers.

Let us conclude with a problem for the non-commutative real algebraic geometry community. We have shown that the set of SOS polynomials is dense in the set of positive semidefinite elements of the Weyl algebra, i.e. we can approximate any polynomial which is positive semidefinite in the Schrödinger representation by an SOS. It would be interesting to know if this kind of results also holds in other algebras important for quantum chemistry. For
instance, if such an ‘approximation property’ were also true in algebras containing coulombian elements of the type $1/|\bar{x}_i - \bar{x}_j|$, then we would be able to estimate electronic molecular energies without the need of introducing orbital basis sets.

Acknowledgments

MN has been supported by the Templeton Foundation. We acknowledge Monique Laurent and Frank Vallentin for interesting discussions. SP acknowledges financial support from the Brussels-Capital Region through a BB2B grant.

Appendix A. Bounding the norm of $g'_c$

The goal of this section is to prove the next lemma, from which lemma 4 is a direct corollary.

Lemma A.1. Let $c > 2$. Then

$$l_1(g'_c) \leq \frac{c}{c - 2}. \quad (A.1)$$

For the proof of this lemma, we will rely on the next two propositions.

Proposition A.1.

$$a^k(a^k)^* = \sum_{m=0}^{k} \frac{k!^2}{m!^2(k-m)!} (a^m)^* a^m. \quad (A.2)$$

Proof. Clearly, $a^k(a^k)^* = \sum_{l,m} s_{l,m} (a^*)^l a^m$. Evaluating the mean value of the Schrödinger representation of both polynomials with respect to an arbitrary coherent state $|\alpha\rangle$, we have that

$$\sum_{l,m} s_{l,m} (a^*)^l a^m = e^{-|\alpha|^2} \sum_{j=0}^{\infty} \frac{(j+k)!}{j!^2} |\alpha|^{2j} =: g(|\alpha|^2). \quad (A.3)$$

It follows that

$$s_{l,m} = \delta_{l,m} \frac{1}{m!} \frac{d^m g(x)}{dx^m} \bigg|_{x=0}. \quad (A.4)$$

Now, it can be proven, by induction, that

$$\frac{d^m g(x)}{dx^m} = \frac{k!}{(k-m)!} e^{-x} \sum_{j=0}^{\infty} \frac{(j+k)!}{j!(j+m)!} x^j $$

if $m \leq k$, while $\frac{d^m g(x)}{dx^m} = 0$ if $m > k$. The statement of the proposition follows from these two relations. \qed

Proposition A.2. Let $\sharp \{\bar{t} \in \mathbb{N}^n : \|\bar{t}\|_1 = k\}$ denote the number of elements of $\mathbb{N}^n$ satisfying $\|\bar{t}\|_1 = k$. Then

$$\sharp \{\bar{t} \in \mathbb{N}^n : \|\bar{t}\|_1 = k\} = \frac{(n+k-1)!}{(n-1)!k!}. \quad (A.6)$$
Proof. Our aim is to compute the number of ways in which \( k \) identical balls can be contained in \( n \) different boxes. Clearly, any possible configuration can be represented uniquely by a sequence of \( k \) dots ‘.’ and \( n - 1 \) bars ‘|’. The number of balls \( n_1 \) in box 1 would then correspond to the number of dots on the left of the first bar; the number of balls \( n_j \) inside box \( j \), for \( 2 \leq j \leq n - 1 \), to the number of dots between the \((j - 1)\)th and the \( j \)th bars; the number of balls in box \( n \), to the number of dots on the right of the \((n - 1)\)th box. For instance, the configuration \( n_1 = 2, n_2 = 0, n_3 = 1 \) would be represented by ‘..|.’.

It is elementary that the number of permutations of \( n + k - 1 \) elements, out of which \( n - 1 \) and \( k \) are indistinguishable, is equal to \( \frac{(n+k-1)!}{(n-1)!k!} \).

\[ \square \]

**Proof of lemma A.1.** Proposition A.1 implies that

\[
I_1(a^k(a^k)^*) = \sum_{m=0}^{k} \frac{k!^2}{m!^2(k-m)!} \leq \sum_{m=0}^{k} \frac{k!}{m!(k-m)!} k! = 2^k k! \quad (A.7)
\]

The last expression is logarithmically superadditive, i.e. \( \prod l \leq k! \leq 2^k k! \), for all sets of natural numbers \( \{k_1, k_2, \ldots\} \) such that \( \sum_i k_i = k \). It follows that the bound given by equation (A.7) also holds for \( I_1(a^i(a^j)^*) \), with \( i \in \{0, 1, \ldots, r\} \) and \( \|\bar{t}\| = k \). We thus have that

\[
I_1(g_c^i) \leq \sum_{\|\bar{t}\| \leq r} \frac{(n-1)!}{(n + \|\bar{t}\| - 1)!} I_1(a^i(a^j)^*) \leq \sum_{k=0}^{r} \frac{\#\{\bar{t} : \|\bar{t}\| \leq k\} 2^k}{(n + k - 1) c^k}
\]

\[
= \sum_{k=0}^{r} \left( \frac{2}{c} \right)^k \leq \sum_{k=0}^{\infty} \left( \frac{2}{c} \right)^k = \frac{c}{c - 2},
\]

where in the third inequality, we have made use of proposition A.2.

\[ \square \]

**Appendix B. Anti-normal ordered monomials**

The following appendix establishes lemma 6.

**Proposition B.1.** For any \( k \in \mathbb{N} \),

\[
a^{k+l}(a^{k+l})^* - a^s a^k(a^k)^* a \in \Sigma^2. \quad (B.1)
\]

Proof. Using the CCRs, we have that

\[
a^* a^k(a^k)^* a = -k a^{k-1}(a^k)^* a - (k + 1) a^k(a^k)^* a + a^{k+1}(a^k)^{k+1}. \quad (B.2)
\]

Now, by induction, we have that, for any \( 0 \leq l \leq k \), \( a^l(a^k)^* a^k \in \Sigma^2 \). Indeed, for \( l = 0 \) the result is obvious. Suppose now that the result holds for \( l \). Then

\[
a^{k+l}(a^k)^* a^{k-l-1} = a^l a(a^k)^* a^{k-l-1} = ka^l(a^k)^* a^{k-l-1} + a^l(a^k)^* a^{k-l} \quad (B.3)
\]

and the last expression belongs to \( \Sigma^2 \) by hypothesis.

It follows that

\[
a^{k+l}(a^{k+l})^* - a^s a^k(a^k)^* a a = (k + 1) a^k(a^k)^* a + ka^{k+1}(a^k)^* a \in \Sigma^2. \quad (B.4)
\]

\[ \square \]
Proposition B.2. Let \( s \in \mathcal{W}_1 \) be an arbitrary monomial of length \( k \). Then
\[
a^k(a^k)^* - ss^* \in \Sigma^2. \tag{B.5}
\]

Proof. We will prove the proposition by induction. Suppose, thus, that the proposition holds for all monomials of length smaller than or equal to \( k \), and let \( s \) be an arbitrary monomial with \( |s| = k + 1 \). There are two possibilities:

1. \( s = a\tilde{s} \), with \( |\tilde{s}| = k \). Then we have that
\[
a^{k+1}(a^{k+1})^* - ss^* = a(a^k(a^k)^* - \tilde{s}\tilde{s}^*)a^* = \sum_i a f_i f_i^* a^* \in \Sigma^2. \tag{B.6}
\]

2. \( s = a^*\tilde{s} \), with \( |\tilde{s}| = k \). Then we have that
\[
a^{k+1}(a^{k+1})^* - ss^* = \{a^{k+1}(a^{k+1})^* - a^*a^k(a^k)^*a\} + \{a^*(a^k)^* - \tilde{s}\tilde{s}^*a\}. \tag{B.7}
\]

The first term in brackets is an SOS by proposition B.1; the second term is an SOS due to the induction hypothesis.

To complete the induction we also have to show that the proposition also holds for \( k = 1 \). But this is trivial, since, in that case, \( aa^* - ss^* \) equals 0 (1), for \( s = a \) (\( s = a^* \)).

\[\square\]

Lemma B.1. Let \( s \in \mathcal{W}_n \) be a monomial. Then
\[
\frac{\delta ss^*}{\delta^*} - ss^* \in \Sigma^2. \tag{B.8}
\]

Proof. Proposition B.2 already shows that the lemma holds for \( n = 1 \). Now, suppose that the lemma holds for \( n \), and let \( s = tu \in \mathcal{W}_{n+1} \), with \( t \) (\( u \)) being a word with the letters \( a_1, \ldots, a_n, a_1^*, \ldots, a_n^* \) (\( a_{n+1}, a_{n+1}^* \)). Let \( \tilde{t} \in \mathbb{N}^n \) be such that \( \frac{\delta tu}{\delta^*} = a^\tilde{t} (a^\tilde{t})^* \). Then
\[
\frac{\delta ss^*}{\delta^*} - ss^* = a^\tilde{t} a_{n+1}^* (a_{n+1})^*a^\tilde{t} a_{n+1}^* - a^\tilde{t} uu^* (a^\tilde{t})^* + uu^* \frac{\delta tu}{\delta^*} - uu^* \frac{\delta tu}{\delta^*} a^\tilde{t} a_{n+1}^* - uu^* a^\tilde{t} a_{n+1}^*. \tag{B.9}
\]

The first two terms on the right-hand side admit an SOS decomposition due to proposition B.2. The two remaining terms belong to \( \Sigma^2 \) because of the induction hypothesis.

\[\square\]

Appendix C. States in the Schrödinger representation

In this appendix, we demonstrate the following lemma:

Lemma C.1. Let \( L \) be a linear functional in \( \mathcal{W}_n \). If \( L(h^*h) \geq 0 \) for any \( h \in \mathcal{W}_n \) and there exist \( c, d > 0, k \in \mathbb{N} \) such that for any monomial \( s \in \mathcal{W}_n \), the relation
\[
|L(s)| \leq dc^{|s|} \Gamma\left(\frac{|s| + k + 1}{2}\right) \tag{C.1}
\]
holds, then there exists a non-normalized quantum state \( \rho \) (a non-negative trace class operator) such that
\[
L(h) = \text{tr}(\rho \pi(h)) \tag{C.2}
\]
for any polynomial \( h \in \mathcal{W}_n \).
Proof. Suppose that, indeed, such a functional exists and define its characteristic function \( \chi(\bar{\xi}) \) as
\[
\chi(\bar{\xi}) = \sum_{l=0}^{\infty} \frac{L((i\bar{\xi}^\sigma \bar{R})^l)}{l!} =: L(e^{i\bar{\xi}^\sigma \bar{R}}),
\]
where \( \bar{R} \in \mathcal{W}_n^{2n} \) is the vector of polynomials
\[
\bar{R} = \left( \begin{array}{c}
a_1 + a_1^* \sqrt{2} \\
a_1 - a_1^* i \sqrt{2} \\
a_n + a_n^* \sqrt{2} \\
a_n - a_n^* i \sqrt{2}
\end{array} \right)
\]
and \( \sigma \) denotes the symplectic form, i.e. \( \sigma = \bigoplus_{l=1}^{n} \left( \begin{array}{cc}
0 & 1 \\
-1 & 0
\end{array} \right) \).

Using (C.1), we have that
\[
|\chi(\bar{\xi})| \leq d \sum_{l=0}^{\infty} \frac{(2n\hat{c})^l \Gamma(l+1/2)}{l!} \Gamma(l+1/2)
\]
with \( \hat{c} = \frac{1}{\sqrt{2}} \max_k \{|\xi_{2k+1} + i\xi_{2k+2}|\} \). That the last series converges for any value of \( \hat{c} \) follows from the relation
\[
\int_{-\infty}^{\infty} |x|^k e^{-x^2} dx = \Gamma\left(\frac{k+1}{2}\right).
\]
This allows us to write
\[
\sum_{l=0}^{\infty} a^l \Gamma\left(\frac{l+1/2}{2}\right) = \int_{-\infty}^{\infty} e^{a|x|} |x|^k e^{-x^2} dx
\]
and the last integral converges for all \( a \in \mathbb{R} \).

Now, define the operator
\[
\rho \equiv \frac{1}{(2\pi)^n} \int d\bar{\xi} \chi(\bar{\xi}) W_{-\bar{\xi}}
\]
with \( W_{\bar{\xi}} = e^{i\bar{\xi}^\sigma \bar{R}} \) being the so-called Weyl operator [28]. We will next prove that, for any polynomial \( h \), \( \text{tr}(\rho \pi(h)) = L(h) \).

Because \( \text{tr}(W_{\bar{\xi}} W_{\bar{\eta}}) = (2\pi)^n \delta(\bar{\xi} + \bar{\eta}) \), it follows that \( \text{tr}(\rho W_{\bar{\xi}}) = \chi(\bar{\xi}) = L(W_{\bar{\xi}}) \). Moreover, from the Weyl relations
\[
W_{\bar{\xi}} W_{\bar{\eta}} = e^{-i\bar{\xi}^\sigma \bar{\eta}/2} W_{\bar{\xi} + \bar{\eta}},
\]
it is immediate that \( \text{tr}(\rho W_{\bar{\xi}} W_{\bar{\eta}}) = L(e^{-i\bar{\xi}^\sigma \bar{\eta}/2} W_{\bar{\xi} + \bar{\eta}}) \equiv f(\bar{\xi}, \bar{\eta}) \). We will now show that
\[
f(\bar{\xi}, \bar{\eta}) = \lim_{r \to \infty} f'(\bar{\xi}, \bar{\eta}),
\]
where \( f'(\bar{\xi}, \bar{\eta}) \equiv \sum_{l,m=0}^{r} L\left(\frac{(i\bar{\xi}^\sigma \bar{R})^l}{l!} \frac{(i\bar{\eta}^\sigma \bar{R})^m}{m!}\right) \).

\[\text{New Journal of Physics 15 (2013) 023026 (http://www.njp.org/)}\]
Note that $f$ is analytic, and that $f^r$ corresponds to a sort of truncated Taylor expansion of $f$ with respect to the variables $\tilde{\xi}, \tilde{\eta}$, where only those monomials $\tilde{\xi}^i \tilde{\eta}^j$ with $\| \tilde{\xi} \|_1, \| \tilde{\eta} \|_1 \leq r$ are present. Indeed, if we take any number of derivatives of $\tilde{\xi}_i, \tilde{\eta}_j$ (not more than $r$ of each) on both sides of (C.9) and then evaluate on the point $\tilde{\xi} = \tilde{\eta} = 0$, we will arrive at two polynomials $p_1, p_2 \in W_n$ on each side. Since this is a general relation between operators, both polynomials must be the same modulo the CCRs. It follows that

\[
\frac{\partial^m \partial^{m'}}{\prod_{k=1}^m \partial \tilde{\xi}_k \prod_{l=1}^{m'} \partial \tilde{\eta}_l} f(\tilde{\xi}, \tilde{\eta}) \big|_{\tilde{\xi} = \tilde{\eta} = 0} = L(p_1) = L(p_2) = \frac{\partial^m \partial^{m'}}{\prod_{k=1}^m \partial \tilde{\xi}_k \prod_{l=1}^{m'} \partial \tilde{\eta}_l} f^r(\tilde{\xi}, \tilde{\eta}) \big|_{\tilde{\xi} = \tilde{\eta} = 0}
\]

(C.11)

for $m, m' \leq r$. Since $f(\tilde{\xi}, \tilde{\eta})$ is analytic in $\tilde{\xi}, \tilde{\eta}$, the sum of the absolute values of the coefficients of its Taylor expansion must converge. This implies that $\lim_{r \to \infty} f(\tilde{\xi}, \tilde{\eta}) - f^r(\tilde{\xi}, \tilde{\eta}) \to 0$ for fixed $\tilde{\xi}, \tilde{\eta}$, and so we obtain equation (C.10).

Analogously, one can prove that

\[
\text{tr} \left( \rho \prod_{j=1}^k W_{i_j} \right) = \lim_{N \to \infty} L \left( \prod_{j=1}^k \sum_{l=0}^N (i \tilde{\xi}_j \cdot \sigma \tilde{R}^j)^l \right)
\]

(C.12)

for any $k$.

Now, let $p = \prod_{k=1}^m p_k \in W_n$, with $p_k \in \left( \frac{a_i + a_i^*}{2}, \frac{a_i - a_i^*}{2i} \right)_{i=1}^n$. Then

\[
L(p) = \frac{\partial^m}{\partial \epsilon_1, \ldots, \epsilon_k} L \left( \prod_{k=1}^m \sum_{l=0}^{\infty} (i \epsilon_k s_k^j)^l \right) \big|_{\epsilon = 0} = \frac{\partial^m}{\partial \epsilon_1, \ldots, \epsilon_k} \text{tr} \left( \rho \prod_{k=1}^m e^{i \epsilon_k p_k} \right) = \frac{\text{tr}(\rho \pi(p))}{i^m}.
\]

(C.13)

By taking linear combinations of the former expectation values, we thus have that $\text{tr}(\rho \pi(h)) = L(h)$ for any $h \in \mathcal{W}_n$.

It only remains to show that $\rho$ is a quantum state, i.e. it is trace-class and positive semidefinite.

From the quantum Bochner–Kinchin theorem [29], we know that $\chi(\tilde{\xi})$ is the characteristic function of a non-normalized quantum state if and only if

1. $\chi(\tilde{\xi})$ is continuous at the origin;
2. for any $r \in \mathbb{N}, \tilde{\xi}_1, \tilde{\xi}_2, \ldots, \tilde{\xi}_r \in \mathbb{R}^{2n}$, and $c_1, c_2, \ldots, c_r \in \mathbb{C}$ the relation

\[
\sum_{k,l=0}^r c_k c_l^* \chi(\tilde{\xi}_k - \tilde{\xi}_l) e^{i \tilde{\xi}_l \sigma \tilde{z}_l / 2} \geq 0
\]

(C.14)

holds.
From (C.5), we know that \( \chi(\bar{\xi}) \) is not only continuous everywhere but even analytic. On the other hand, from (C.10) we have that

\[
\sum_{k,l=0}^{r} c_k c_l^* \chi(\bar{\xi}_k - \bar{\xi}_l) e^{i\bar{\xi}_k \sigma \bar{\xi}_l / 2} = \lim_{m \to \infty} L(h_m h_m^*),
\]

where \( h_m = \sum_k c_k \sum_{l=0}^{m} (i\bar{\xi}_k \sigma \bar{\xi}_l) / l! \). Since \( h_m \in \mathcal{W}_n \), \( L(h_m h_m^*) \geq 0 \), and so the above limit is non-negative. \( \square \)

References

[1] Szabo A and Ostlund N S 1996 Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory (Mineola, NY: Dover)
[2] Nakata M, Nakatsuji H, Ehara M, Fukuda M, Nakata K and Fujisawa K 2001 J. Chem. Phys. 114 8282
[3] Mazziotti D A 2004 Phys. Rev. Lett. 93 213001
[4] Nakata M, Braams B J, Fujisawa K, Fukuda M, Percus J K, Yamashita M and Zhao Z 2008 J. Chem. Phys. 128 164113
[5] Coleman A J 1963 Rev. Mod. Phys. 35 668
[6] Garrod C and Percus J K 1964 J. Math. Phys. 5 1756
[7] Erdahl R M 1978 Int. J. Quantum Chem. 13 697
[8] Pironio S, Navascués M and Acín A 2010 SIAM J. Optim. 20 2157–80
[9] Navascués M, Pironio S and Acín A 2011 Noncommutative polynomial optimization Handbook on Semidefinite, Cone and Polynomial Optimization ed M F Anjos and J Lasserre (Berlin: Springer) pp 601–34
[10] Lasserre J B 2001 SIAM J. Optim. 11 796–817
[11] Parrilo P A 2003 Math. Program. Ser. B 96 293–320
[12] Navascués M and Pironio S A 2007 Acín, Phys. Rev. Lett. 98 010401
Navascués M and Pironio S A 2008 Acín New J. Phys. 10 073013
Doherty A C, Liang Y C, Toner B and Wehner S 2008 Proc. IEEE Conf. on Computational Complexity (College Park, MD: IEEE Computer Society)
[13] Hübenener R and Barthel T 2012 Phys. Rev. Lett. 108 200404
[14] Baumgratz T and Plenio M B 2012 New J. Phys. 14 023027
[15] Hammond J R and Mazziotti D A 2005 Phys. Rev. A 71 062503
[16] Henrion D and Lasserre J B 2005 Detecting global optimality and extracting solutions in GloptiPoly Positive Polynomials in Control (Lecture Notes on Control and Information Sciences vol 312) ed D Henrion and A Garulli (Berlin: Springer) pp 293–310
[17] Waki H, Nakata M and Muramatsu M 2012 Strange behaviors of interior-point methods for solving semidefinite programming problems in polynomial optimization Comput. Optim. Appl. 53 823–44
[18] Schmüdgen K 2005 Math. Ann. 331 779–94
[19] Helton J W and McCullough S A 2004 A Positivstellensatz for non-commutative polynomials Trans. Am. Math. Soc. 356 721–37
[20] Vandenberghe L and Boyd S 1996 Semidefinite programming SIAM Rev. 38 49
[21] Toh K C, Todd M J and Tutuncu R H 1999 SDPT3—a MATLAB software package for semidefinite programming Optim. Methods Softw. 11 545
Tutuncu R H, Toh K C and Todd M J 2003 Solving semidefinite-quadratic-linear programs using SDPT3 Math. Prog. Ser. B 95 189
[22] Lofberg J 2004 YALMIP: A Toolbox for Modeling and Optimization in MATLAB http://control.ee.ethz.ch/~joloef/yalmip.php

New Journal of Physics 15 (2013) 023026 (htp://www.njp.org/)
[23] Lasserre J B 2001 SIAM J. Optim. 11 751
[24] Lasserre J B and Netzer T 2007 Math. Z. 256 99
[25] Reed M and Simon B 1980 Functional Analysis (New York: Academic)
[26] Fujisawa K, Fukuda M, Kobayashi K, Kojima M, Nakata K, Nakata M and Yamashita M 2008 SDPA (Semidefinite Programming Algorithm) and User’s SDPA-GMP Manual Version 7.1.0; Res. Rep. Math. Comput. Sci. B 448 1–35
Nakata M 2010 A numerical evaluation of highly accurate multiple-precision arithmetic version of semidefinite programming solver: SDPA-GMP-QD-DD Proc. 2010 IEEE Multi-Conf. on Systems and Control pp 29–34
[27] Cimprič J 2010 J. Math. Anal. Appl. 369 443–52
[28] Petz D 1990 An Invitation to the Algebra of Canonical Commutation Relations (Leuven: Leuven University Press)
[29] Holevo A S 1982 Probabilistic and Statistical Aspects of Quantum Theory (Amsterdam: North-Holland)