

A NOTE ON THE NONEXISTENCE OF SUM OF SQUARES CERTIFICATES FOR THE BMV CONJECTURE

KRISTIJAN CAFUTA, IGOR KLEP¹, AND JANEZ POVH²

ABSTRACT. The algebraic reformulation of the BMV conjecture is equivalent to a family of dimension-free tracial inequalities involving positive semidefinite matrices. Sufficient conditions for these to hold in the form of algebraic identities involving polynomials in non-commuting variables have been given by Markus Schweighofer and the second author. Later the existence of these certificates has been settled for all but one case, which is resolved in this note.

1. INTRODUCTION

In an attempt to simplify the calculation of partition functions of quantum mechanical systems Bessis, Moussa and Villani [BMV75] conjectured in 1975 that for any two symmetric matrices A, B , where B is positive semidefinite, the function $t \mapsto \operatorname{tr}(e^{A-tB})$ is the Laplace transform of a positive Borel measure with real support. This would permit the calculation of explicit upper and lower bounds of energy levels in multiple particle systems. For an overview of mostly analytical approaches before 1998 we refer the reader to Moussa's survey [Mou00].

In 2004, Lieb and Seiringer [LS04] restated the conjecture in the following purely algebraic form: all the coefficients of the polynomial

$$p_m = \operatorname{tr}((A + tB)^m) \in \mathbb{R}[t]$$

are nonnegative whenever $m \in \mathbb{N}$ and A and B are positive semidefinite matrices of the same size. The coefficient of t^k in p_m is the trace of $S_{m,k}(A, B) :=$ the sum of all words of length m in A and B in which B appears exactly k times (and therefore A exactly $m - k$ times).

In his ingenious 2007 paper [Häg07], Hägele found a dimension-free algebraic certificate proving $\operatorname{tr}(S_{7,3}(A, B)) \geq 0$ for all positive semidefinite A, B , and then used Hillar's important descent theorem [Hil07] to deduce the same property for $S_{6,3}(A, B)$. Motivated by this, Schweighofer and the second author [KS08b] established an approach to the BMV conjecture using sums of hermitian squares of polynomials in noncommuting variables combined with Hillar's descent theorem and proved the conjecture for $m \leq 13$. To describe the method in detail we introduce some notation.

Date: June 8, 2010.

2000 Mathematics Subject Classification. Primary 11E25, 90C22; Secondary 08B20, 13J30, 90C90.

Key words and phrases. Bessis-Moussa-Villani (BMV) conjecture, noncommutative polynomial, sum of hermitian squares, commutator, cyclic equivalence, semidefinite programming, free positivity.

¹Partially supported by the Slovenian Research Agency (programs no. P1-0222 and P1-0288).

²Supported by the Slovenian Research Agency (project no. 1000-08-210518).

1.1. Notation. The main feature of this method is to model the matrices as *noncommuting variables* in a noncommutative polynomial ring. Let $\langle X, Y \rangle$ be the monoid freely generated by $\{X, Y\}$, i.e., $\langle X, Y \rangle$ consists of **words** in the two noncommuting letters X, Y (including the empty word denoted by 1). We consider the free algebra $\mathbb{R}\langle X, Y \rangle$ on $\langle X, Y \rangle$, i.e., the ring of polynomials in the noncommuting variables X, Y with coefficients from \mathbb{R} . The elements of $\mathbb{R}\langle X, Y \rangle$ are linear combinations of words from $\langle X, Y \rangle$ and are called **NC polynomials**. The length of the longest word in an NC polynomial $f \in \mathbb{R}\langle X, Y \rangle$ is the **degree** of f and is denoted by $\deg f$. Likewise we consider the X -degree $\deg_X f$ and the Y -degree $\deg_Y f$.

Definition 1.1. Two polynomials $f, g \in \mathbb{R}\langle X, Y \rangle$ are called **cyclically equivalent** ($f \stackrel{\text{cyc}}{\sim} g$) if $f - g$ is a sum of commutators in $\mathbb{R}\langle X, Y \rangle$, i.e., there are $p_i, q_i \in \mathbb{R}\langle X, Y \rangle$ with $f - g = \sum(p_i q_i - q_i p_i) = \sum[p_i, q_i]$.

This definition reflects the fact that $\text{tr}(AB) = \text{tr}(BA)$ for square matrices A and B of the same size. Cyclic equivalence can easily be checked. One readily verifies that words $v, w \in \langle X, Y \rangle$ are cyclically equivalent if and only if there are $v_1, v_2 \in \langle X, Y \rangle$ such that $v = v_1 v_2$ and $w = v_2 v_1$. Two polynomials $f = \sum_{w \in \langle X, Y \rangle} a_w w$ and $g = \sum_{w \in \langle X, Y \rangle} b_w w$ from $\mathbb{R}\langle X, Y \rangle$ (here, only finitely many of the $a_w, b_w \in \mathbb{R}$ are nonzero) are cyclically equivalent if and only if for each $v \in \langle X, Y \rangle$,

$$\sum_{\substack{w \in \langle X, Y \rangle \\ w \stackrel{\text{cyc}}{\sim} v}} a_w = \sum_{\substack{w \in \langle X, Y \rangle \\ w \stackrel{\text{cyc}}{\sim} v}} b_w.$$

We equip $\mathbb{R}\langle X, Y \rangle$ with the **involution** $*$ that fixes $\mathbb{R} \cup \{X, Y\}$ pointwise and thus reverses words, e.g., if $p = (X^2 - XY^3)$, then

$$p^* = (X^2 - XY^3)^* = X^2 - Y^3 X.$$

So $\mathbb{R}\langle X, Y \rangle$ is the $*$ -algebra freely generated by two symmetric letters. Let $\text{Sym } \mathbb{R}\langle X, Y \rangle = \{f \in \mathbb{R}\langle X, Y \rangle \mid f^* = f\}$ denote the set of all **symmetric elements**. The involution $*$ extends naturally to matrices (in particular, to vectors) over $\mathbb{R}\langle X, Y \rangle$. For instance, if $V = (v_i)$ is a (column) vector of NC polynomials $v_i \in \mathbb{R}\langle X, Y \rangle$, then V^* is the row vector with components v_i^* . We shall also use V^t to denote the row vector with components v_i .

Given an NC polynomial $f \in \mathbb{R}\langle X, Y \rangle$ it is natural to substitute symmetric matrices A, B of the same size for the variables X and Y yielding a matrix $f(A, B)$ of the same size. The involution $*$ is compatible with matrix transposition in the sense that $f(A, B)^t = f^*(A, B)$. For instance, for the polynomial p defined above, we have

$$p(A, B) = A^2 - AB^3.$$

A special case of [KS08a, Theorem 2.1] (the main motivation for the definition of cyclic equivalence) says that a symmetric $f \in \mathbb{R}\langle X, Y \rangle$ is a sum of commutators (i.e., $f \stackrel{\text{cyc}}{\sim} 0$) if and only if $\text{tr}(f(A, B)) = 0$ for all real symmetric matrices A and B of the same size.

We now turn to notions related to positivity. Recall that the *Löwner ordering* \succeq on symmetric matrices is defined as $A \succeq B$ if and only if $A - B$ is positive semidefinite.

Definition 1.2. We denote by

$$(1) \quad \Sigma^2 = \left\{ \sum_{i=1}^m g_i^* g_i \mid m \in \mathbb{N}, g_i \in \mathbb{R}\langle X, Y \rangle \right\} \subseteq \text{Sym } \mathbb{R}\langle X, Y \rangle$$

the convex cone of all **sums of hermitian squares** and by

$$(2) \quad \Theta^2 = \{f \in \mathbb{R}\langle X, Y \rangle \mid \exists g \in \Sigma^2 : f \stackrel{\text{cyc}}{\approx} g\} \subseteq \mathbb{R}\langle X, Y \rangle$$

the convex cone of all polynomials that are **cyclically equivalent to a sum of hermitian squares**. That is, Θ^2 consists of all polynomials that can be written as a sum of hermitian squares and commutators.

The importance of these sets for us is given by the following elementary observations:

Proposition 1.3. *Let $f \in \mathbb{R}\langle X, Y \rangle$.*

- (1) *If $f \in \Sigma^2$, then $f(A, B) \succeq 0$ for all symmetric matrices A and B of the same size.*
- (2) *If $f \in \Theta^2$, then $\text{tr}(f(A, B)) \geq 0$ for all symmetric matrices A and B of the same size.*

By Helton's theorem [Hel02], the converse of (1) holds: if for $f \in \mathbb{R}\langle X, Y \rangle$, $f(A, B)$ is positive semidefinite for all symmetric matrices A and B of the same size, then $f \in \Sigma^2$. On the other hand, the converse of (2) fails in general, that is, there are examples of NC polynomials f satisfying the trace positivity condition of part (2) of Proposition 1.3 yet $f \notin \Theta^2$, cf. [KS08a, Example 4.4] or [KS08b, Example 3.5]. Nevertheless, this part of Proposition 1.3 yields a useful sufficient condition for tracial positivity and was exploited by Schweighofer and the second author [KS08b] to prove the BMV conjecture for $m \leq 13$.

To model *positive semidefiniteness* with the aid of *symmetric* noncommuting variables, we consider polynomials in X^2, Y^2 : if $S_{m,k}(X^2, Y^2) \in \Theta^2$ for some m, k , then the t^k coefficient of p_m is nonnegative for all positive semidefinite matrices A, B of all sizes.

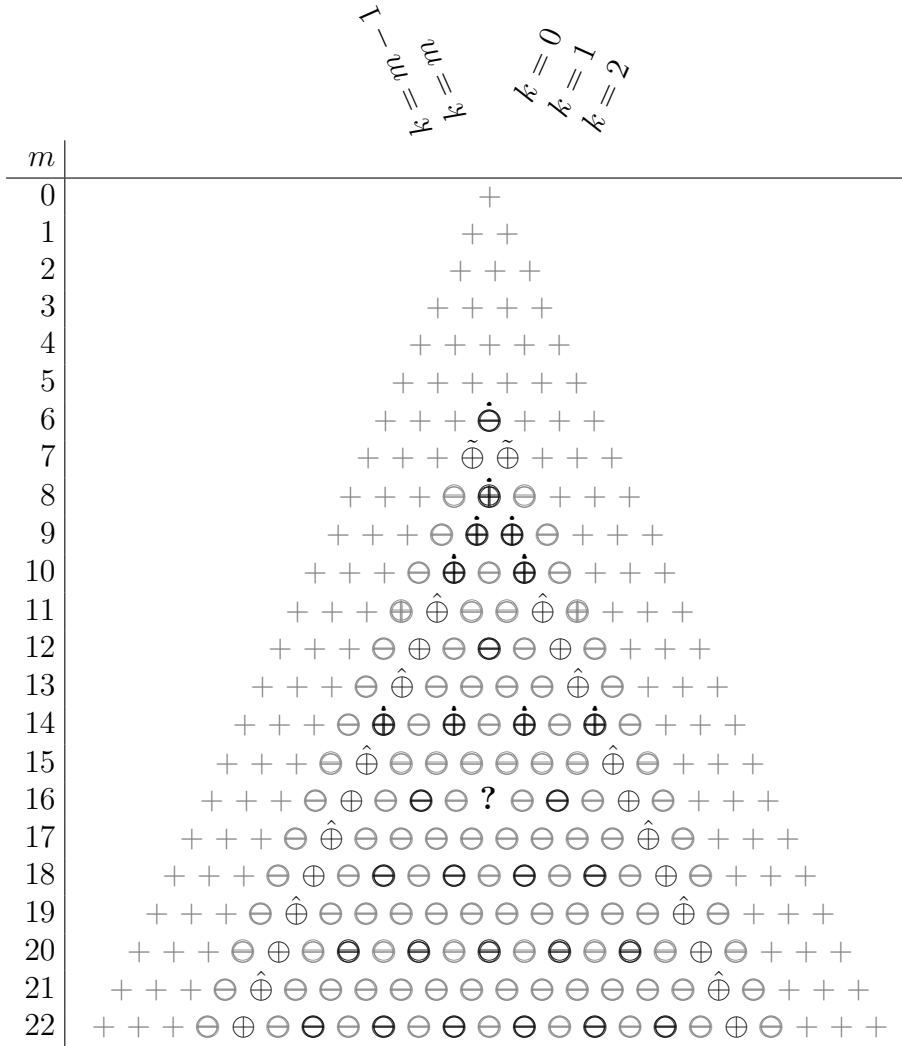
1.2. An xmas tree. Much work has been done in determining whether $S_{m,k}(X^2, Y^2) \in \Theta^2$ for a given pair (m, k) . It is easy to see $S_{m,k}(X^2, Y^2) \in \Theta^2$ for $k \leq 2$ or $m - k \leq 2$. In our terminology, the first (nontrivial) certificate can be extracted from Hägele [Häg07] to show $S_{7,3}(X^2, Y^2) \in \Theta^2$. This was followed upon in [KS08b] where among the main results were $S_{6,3}(X^2, Y^2) \notin \Theta^2$, $S_{14,4}(X^2, Y^2) \in \Theta^2$ and $S_{14,6}(X^2, Y^2) \in \Theta^2$. The latter results combined with Hillar's descent theorem [Hil07] imply that the BMV conjecture holds for $m \leq 13$. Hillar's theorem implies that the BMV conjecture holds if and only if it holds for an infinite number of m 's. This is a crucial ingredient for the sum of squares approach to the conjecture as it is clear that not all $S_{m,k}(X^2, Y^2)$ are members of Θ^2 .

Here is a brief overview of the latest developments. Landweber and Speer [LS09] proved for example that $S_{m,4}(X^2, Y^2) \in \Theta^2$ for odd m and that $S_{11,3}(X^2, Y^2) \in \Theta^2$. They also give results on the negative side, implying that $S_{m,k}(X^2, Y^2) \notin \Theta^2$ in the following cases:

- (1) m is odd and $5 \leq k \leq m - 5$; (2) $m \geq 13$ is odd and $k = 3$;
- (3) m is even, k is odd and $3 \leq k \leq m - 3$; (4) $(m, k) = (9, 3)$.

Independently of the work of Landweber and Speer, Burgdorf [Bur] found a combinatorial proof of $S_{m,4}(X^2, Y^2) \in \Theta^2$ for all m . Together with Hillar's descent theorem this implies that the BMV conjecture holds for all pairs (m, k) with $k \leq 4$ or $m - k \leq 4$. The last contribution of negative results is given by Collins, Dykema and Torres-Ayala [CDTA]: $S_{12,6}(X^2, Y^2) \notin \Theta^2$ and for even m, k with $6 \leq k \leq m - 10$, $S_{m,k}(X^2, Y^2) \notin \Theta^2$.

We present the state-of-the-art knowledge conveniently in the form of a table:



Is $S_{m,k}(X^2, Y^2) \in \Theta^2$?

(The tree continues following the pattern in rows 20, 21 and 22.)

	authors	color
	Hägele [Häg07]	$\tilde{\oplus}$
	Klep and Schweighofer [KS08b]	$\dot{\oplus}$ $\dot{\ominus}$
	Burgdorf [Bur], Landweber and Speer [LS09]	$\hat{\oplus}$
	Burgdorf [Bur]	\oplus
	Landweber and Speer [LS09]	\oplus \ominus
	Collins, Dykema and Torres-Ayala [CDTA]	\ominus

symbol	meaning
+	$S_{m,k}$ is in Θ^2 for trivial reasons
\oplus	$S_{m,k}$ is in Θ^2 (with proof)
\ominus	$S_{m,k}$ is not in Θ^2 (with proof)

Legend

The aim of this article is to settle the remaining case, i.e., we prove (what was conjectured in [KS08b, pg. 754] based on numerical evidence) $S_{16,8}(X^2, Y^2) \notin \Theta^2$.

2. GRAM MATRIX METHOD AND SEMIDEFINITE PROGRAMMING

In this section we explain how a desired nonmembership certificate can be obtained. The main idea is to construct a linear map $L : \mathbb{R}\langle X, Y \rangle \rightarrow \mathbb{R}$ satisfying

$$(3) \quad L(\Theta^2 \mathcal{P}) \subseteq \mathbb{R}_{\geq 0} \quad L(S_{16,8}(X^2, Y^2)) < 0.$$

2.1. Gram matrix method. Checking whether a polynomial in noncommuting variables is an element of Σ^2 and Θ^2 , respectively, is most efficiently done via the so-called Gram matrix method [KS08b, KP10], well-known in the commutative setting [CLR95, PS03].

Theorem 2.1 (Klep, Schweighofer [KS08b, Proposition 3.3]). *Suppose m, k are even and set*

$$(4) \quad \begin{aligned} V_1 &:= \{v \in \{X^2, Y^2\}^{\frac{m}{2}} \mid \deg_X v = m - k, \deg_Y v = k\}, \\ V_2 &:= \{v \in X\{X^2, Y^2\}^{\frac{m}{2}-1}X \mid \deg_X v = m - k, \deg_Y v = k\}, \\ V_3 &:= \{v \in Y\{X^2, Y^2\}^{\frac{m}{2}-1}Y \mid \deg_X v = m - k, \deg_Y v = k\}. \end{aligned}$$

Let \bar{v}_i denote the vector $[v]_{v \in V_i}$. Then $S_{m,k}(X^2, Y^2) \in \Theta^2$ if and only if there exist positive semidefinite matrices $G_i \in \text{Sym } \mathbb{R}^{V_i \times V_i}$ such that

$$(5) \quad S_{m,k}(X^2, Y^2) \stackrel{\text{cyc}}{\approx} \sum_i \bar{v}_i^* G_i \bar{v}_i.$$

If $G_i = H_i^* H_i$ and $H_i \in \mathbb{R}^{J_i \times V_i}$ (J_i some index set), then with $[p_{i,j}]_{j \in J_i} := H_i \bar{v}_i$ we have

$$(6) \quad S_{m,k}(X^2, Y^2) \stackrel{\text{cyc}}{\approx} \sum_{i,j} p_{i,j}^* p_{i,j}.$$

Any symmetric block matrix $G = \begin{bmatrix} G_1 & & \\ & G_2 & \\ & & G_3 \end{bmatrix}$ satisfying $f \stackrel{\text{cyc}}{\approx} \sum_i \bar{v}_i^* G_i \bar{v}_i$ for some $f \in \mathbb{R}\langle X, Y \rangle$, is called a **Gram matrix** for f . If $f = \sum_i \bar{v}_i^* G_i \bar{v}_i$, then we call G an **exact Gram matrix**. (We emphasize this is not the standard definition.)

2.2. The certificate. Let us now return to the question whether $S_{16,8}(X^2, Y^2) \in \Theta^2$. Following Theorem 2.1 we must determine whether there exists a positive semidefinite matrix G such that $S_{16,8}(X^2, Y^2) \stackrel{\text{cyc}}{\approx} W^* G W$, where W is the vector $[\bar{v}_1^t \quad \bar{v}_2^t \quad \bar{v}_3^t]^t$ obtained in Theorem 2.1. Here, \bar{v}_1 has length 70, while \bar{v}_2 and \bar{v}_3 have length 35.

Therefore we obtain a semidefinite feasibility problem [WSV00] in the matrix variable G of order 140, where the linear constraints on G express that for each product of words $w \in \{p^*q \mid p, q \in W\}$ we have

$$(7) \quad \sum_{\substack{p, q \in W \\ p^*q \stackrel{\text{cyc}}{\approx} w}} G_{p,q} = \sum_{u \stackrel{\text{cyc}}{\approx} w} a_u,$$

where a_u is the coefficient of u in $S_{16,8}(X^2, Y^2)$. There are 4485 equivalence classes (with respect to cyclic equivalence) of words in $\{p^*q \mid p, q \in W\}$, yielding 4485 linear constraints in the semidefinite program.

By Theorem 2.1, we may restrict ourselves to words $w = p^*q$ with $p, q \in V_i$. Moreover, since $S_{16,8}(X^2, Y^2)$ is symmetric we can merge for each pair w and w^* , where $w \in \{p^*q \mid p, q \in V_i \text{ for some } i = 1, 2, 3\}$, the constraints (7) into a single constraint

$$(8) \quad \sum_{\substack{p, q \in W \\ p^*q \stackrel{\text{cyc}}{\sim} w \vee p^*q \stackrel{\text{cyc}}{\sim} w^*}} G_{p,q} = \sum_{u \stackrel{\text{cyc}}{\sim} w \vee u \stackrel{\text{cyc}}{\sim} w^*} a_u.$$

Thus we obtain a semidefinite program in a block diagonal matrix variable with blocks of orders 70, 35 and 35, with 440 linearly independent linear constraints instead of the initial 4485 constraints. Hence we reduced the number of constraints by about 90 %.

To prove this problem is *infeasible*, we find a separating hyperplane with the help of semidefinite programming. Fix $m = 16$, $k = 8$ and let \mathcal{V} denote the vector space of all block diagonal symmetric matrices as in Theorem 2.1. To each $0 \neq G \in \mathcal{V}$ we can associate the NC polynomial $W^*GW \in \mathbb{R}\langle X, Y \rangle$ of degree 32. Let \mathcal{P} denote the vector space of all such polynomials. Each $f \in \mathcal{P}$ has an exact Gram matrix. It is even *unique* since f is *homogeneous* [KP10, Proposition 2.3]. Let $\Theta^2\mathcal{P}$ denote the set of NC polynomials in \mathcal{P} with a positive semidefinite Gram matrix.

Lemma 2.2. $\Theta^2\mathcal{P} = \Theta^2 \cap \mathcal{P}$.

Proof. This is a straightforward extension of the proof of [KS08b, Proposition 3.3]. ■

Every linear map $L : \mathcal{P} \rightarrow \mathbb{R}$ can be presented as

$$(9) \quad f \mapsto \langle B_1, G_1 \rangle + \langle B_2, G_2 \rangle + \langle B_3, G_3 \rangle = \text{tr}(B_1 G_1) + \text{tr}(B_2 G_2) + \text{tr}(B_3 G_3)$$

for some (symmetric) block matrix $B_L = \begin{bmatrix} B_1 & & \\ & B_2 & \\ & & B_3 \end{bmatrix}$, where $\begin{bmatrix} G_1 & & \\ & G_2 & \\ & & G_3 \end{bmatrix}$ is an exact Gram matrix for f . Conversely, equation (9) can be used to define $L : \mathcal{P} \rightarrow \mathbb{R}$ due to the uniqueness of the exact Gram matrix for polynomials in \mathcal{P} .

Let $\{C_j \mid j \in J\}$ denote a basis of $\{A_{p^*q} \mid i \in \{1, 2, 3\}, (p, q) \in V_i \times V_i\}^\perp \subseteq \mathcal{V}$, where A_w are the constraint matrices from our original feasibility SDP (8).

We are now in a position to present the desired SDP constructing a separating hyperplane.

Proposition 2.3. *Let G_0 denote any Gram matrix for $S_{16,8}(X^2, Y^2)$. Consider the semidefinite feasibility problem*

$$(10) \quad \begin{aligned} B = \begin{bmatrix} B_1 & & \\ & B_2 & \\ & & B_3 \end{bmatrix} &\succeq 0, \\ \langle B, G_0 \rangle &= -100, \\ \langle B, C_j \rangle &= 0 \quad \text{for all } j \in J. \end{aligned}$$

Then (10) is feasible if and only if $S_{16,8}(X^2, Y^2) \notin \Theta^2$.

Proof. Suppose first $L : \mathcal{P} \rightarrow \mathbb{R}$ is linear. By the self-duality of the cone of positive semidefinite matrices, $L(\Theta^2\mathcal{P}) \subseteq \mathbb{R}_{\geq 0}$ if and only if $B_L \succeq 0$. If this holds, then $L(f) = 0$ for all $f \in \mathcal{P}$ with $f \stackrel{\text{cyc}}{\sim} 0$ by Lemma 2.2. Hence

$$(11) \quad \langle B_L, H \rangle = 0$$

for all $H \in \mathcal{V}$ satisfying $W^*HW \stackrel{\text{cyc}}{\approx} 0$. The later condition can be rephrased as $\langle A_w, H \rangle = 0$ for all A_w . So H is in the span of the C_j . In particular, (11) can be equivalently written as

$$(12) \quad \langle B_L, C_j \rangle = 0 \quad \text{for all } j \in J.$$

The above shows that every B feasible for (10) gives rise to a linear functional $L : \mathcal{P} \rightarrow \mathbb{R}$ as in (9) with the following properties:

$$L(\Theta^2\mathcal{P}) \subseteq \mathbb{R}_{\geq 0} \quad \text{and} \quad L(S_{16,8}(X^2, Y^2)) = -100.$$

Hence $S_{16,8}(X^2, Y^2) \notin \Theta^2$ by Theorem 2.1.

Conversely, assume $S_{16,8}(X^2, Y^2) \notin \Theta^2$. As the convex cone Θ^2 is closed (cf. [BK, Lemma 4.5]), there exists a separating linear functional $L : \mathbb{R}\langle X, Y \rangle \rightarrow \mathbb{R}$ satisfying $L(\Theta^2) \subseteq \mathbb{R}_{\geq 0}$ and $L(S_{16,8}(X^2, Y^2)) < 0$. The restriction of L to \mathcal{P} is of the form (9) and hence (after scaling) yields a feasible point B for (10). ■

Remark 2.4. We chose -100 in (10) for numerical reasons, since the trace in $\langle B, G_0 \rangle = \text{tr}(BG_0)$ is not normalized. This results in a larger smallest eigenvalue of B , making it easier to round and eventually prove its positive semidefiniteness.

Theorem 2.5. (10) is feasible.

2.3. Proof of Theorem 2.5. We explain how this was verified using a computer.

A general SDP solver (such as SDPT3, SDPA or SeDuMi; see Mittelman’s website [Mit] for a benchmark of state-of-the-art solvers) will produce a *floating point* feasible solution for (10). However, finding a *symbolic* (e.g. *rational*) feasible point requires additional work. We proceed as follows: run (10) as an SDP with *trivial* objective function, since under a strict feasibility assumption the interior point methods yield solutions in the relative interior of the optimal face, which in our case is the whole feasibility set. If strict complementarity is additionally provided, the interior point methods lead to the *analytic center* of the feasibility set [GS98, HdKR02].

In our example this produces a *nonsingular* matrix B' with smallest eigenvalue approximately $\varepsilon = 0.41$ and distance to the affine subspace generated by the linear constraints of (10) being approximately $\delta = 7.1 \cdot 10^{-8}$. Taking a very close rational approximation B'' of B' (e.g. $\tau = \|B'' - B'\|$ satisfies $\tau^2 + \delta^2 < \varepsilon^2$) and then projecting B'' onto the affine subspace yields a *rational* matrix B feasible for (10); see [PP08, Proposition 8 and Fig. 1 on pg. 276] for details and proof of correctness.

We also *explicitly* computed a rational (even integer) feasible point for a small modification of (10), where $\langle B, G_0 \rangle$ is negative but not necessarily -100 . All the data needed to verify the correctness is available from our NCSOSTools [CKP] website

<http://ncsostools.fis.unm.si/>

See also the Appendix for a fuller explanation. ■

Corollary 2.6. $S_{16,8}(X^2, Y^2) \notin \Theta^2$.

2.4. More on rational certificates for SDP. Consider a feasibility SDP in primal form

$$(FSDP) \quad \begin{array}{l} X \succeq 0 \\ \text{s. t. } \langle A_i, X \rangle = b_i, \quad i = 1, \dots, m \end{array}$$

and assume the input data A_i, b is rational. If the problem is feasible, does there exist a *rational* solution? If so, can one use a combination of numerical and symbolic computation to produce one?

Example 2.7. Some caution is necessary, as a feasible SDP of the form (FSDP) need not admit a rational solution. For a concrete example, note that

$$\begin{bmatrix} 2 & x \\ x & 1 \end{bmatrix} \oplus \begin{bmatrix} x & 1 & 0 \\ 1 & x & 1 \\ 0 & 1 & x \end{bmatrix} \succeq 0 \quad \Leftrightarrow \quad x = \sqrt{2}.$$

On the other hand, if (FSDP) does admit a feasible *positive definite* solution, then it admits a (positive definite) *rational* solution [PP08]; this was exploited in the proof of our main theorem above.

The existence of rational feasible points is thus an important issue in polynomial optimization as it provides exact sum of squares certificates for non-negativity or for lower bounds of polynomials. Building upon the pioneering work of Peyrl and Parrilo, Kalfoten, Li, Yang and Zhi [KLYZ08] provide an extension to sums of squares of rational functions that also touches upon possible singularities in floating point feasible solutions. We also mention [EDZ], where the authors propose an algorithm for detecting and returning a rational point in a convex basic closed semialgebraic set defined by rational polynomials. As a special case, they obtain a procedure deciding whether a multivariate polynomial $f \in \mathbb{Q}[\bar{X}]$ (in commuting variables \bar{X}) is a sum of squares in $\mathbb{Q}[\bar{X}]$. The latter problem is closely related to an open problem of Sturmfels which is at the same time a particular instance of the rationality problem for (FSDP): if $f \in \mathbb{Q}[\bar{X}]$ is a sum of squares in $\mathbb{R}[\bar{X}]$, is f a sum of squares in $\mathbb{Q}[\bar{X}]$?

This problem is still open and the main contribution is given by Hillar in [Hil09]: if f is assumed to be a sum of squares of polynomials with coefficients in a *totally real* field, then the answer to Sturmfels' question is affirmative.

In general, given an SDP with rational input, each coordinate of its optimal solution is an algebraic number and the degrees of minimal polynomials of these algebraic numbers are studied in [NRS10].

Acknowledgments. The authors thank Markus Schweighofer for insightful comments and discussions. We thank the referee for useful suggestions that helped improve the presentation of the paper.

APPENDIX: THE MATLAB VERIFICATION

The data package contains the following data:

- S168 ... the BMV polynomial $S_{\{16,8\}}(X^2, Y^2)$
- V ... the vector of all words of order 16, which can appear in an SOHS polynomial cyclically equivalent to S168.
Note that $V=[V1;V2;V3]$, where V_i is a vector as in Theorem 2.1
- G0 = a (block diagonal) Gram matrix for S168 - the one we used in (10).

- A ... a matrix of order 19600 x 440 ... each column of A corresponds to equations as in (7) or (8)

- C ... a matrix of order 19600 x 3305 ... columns of C are pairwise orthogonal and also orthogonal to columns of A ... matrix reformulations of columns of C are exactly matrices C_i from (10).

Note: we kept in A and C only columns which corresponds to the diagonal blocks, as described in Lemma 2.3, hence we have in A and C altogether $70 \cdot 71/2 + 2 \cdot 35 \cdot 36/2 = 3745$ columns.

- B ... a solution of (10). Note that $B = \text{blockDiag}(B_1, B_2, B_3)$ - a PsD matrix of order 140x140.

Instructions (some of this requires NCSOSTools):

1. To reproduce the polynomial $S_{\{16,8\}}(X^2, Y^2)$, run

```
S168=BMVq(16,8);
```

2. To check that G_0 is a Gram matrix for S168 call

```
Snew=V'*G0*V; NCisCycEq(S168,Snew)
```

(Caution: the last command must give answer 1 and takes quite some time to evaluate.)

3. To check that A contains the equations (7) compute

```
trace(reshape(A(:,i),140,140)*G0)
```

which must be the number of all words in S168, which are cyclically equivalent to the words w or w^* , underlying the i -th equation

4. To check that B is feasible for the linear constraints in (10) run

```
norm(C'*B(:))==0, trace(B*G0)<0
```

Note that $\text{trace}(B \cdot G_0) = -8$ and not -100 as stated in (10) - the reason is that B is integer matrix obtained by rounding and projecting and rescaling.

5. B is an integer matrix. To see that is it PsD, compute

```
min(eig(B))>0
```

Alternatively, for a symbolic verification, please use our Mathematica notebook

`bmv_16_8-ldlt.nb`

available from <http://ncsostools.fis.unm.si> where the LDU factorization is given.

REFERENCES

- [BK] S. Burgdorf and I. Klep. The truncated tracial moment problem. To appear in *J. Operator Theory*, <http://arxiv.org/abs/1001.3679>.
- [BMV75] D. Bessis, P. Moussa, and M. Villani. Monotonic converging variational approximations to the functional integrals in quantum statistical mechanics. *J. Mathematical Phys.*, 16(11):2318–2325, 1975.
- [Bur] S. Burgdorf. Sums of hermitian squares as an approach to the BMV conjecture. To appear in *Linear Multilinear Algebra*, <http://arxiv.org/abs/0802.1153>.
- [CDTA] B. Collins, K.J. Dykema, and F. Torres-Ayala. Sum-of-squares results for polynomials related to the Bessis-Moussa-Villani conjecture. Preprint, <http://arxiv.org/abs/0905.0420>.
- [CKP] K. Cafuta, I. Klep, and J. Povh. NCSOSTools: a computer algebra system for symbolic and numerical computation with noncommutative polynomials, available from <http://ncsostools.fis.unm.si>.
- [CLR95] M.D. Choi, T.Y. Lam, and B. Reznick. Sums of squares of real polynomials. In *K-theory and algebraic geometry: connections with quadratic forms and division algebras (Santa Barbara, CA, 1992)*, volume 58 of *Proc. Sympos. Pure Math.*, pages 103–126. Amer. Math. Soc., Providence, RI, 1995.
- [EDZ] M.S. El Din and L. Zhi. Computing rational points in convex semi-algebraic sets and sos decompositions. Preprint, <http://arxiv.org/abs/0910.2973>.
- [GS98] D. Goldfarb and K. Scheinberg. Interior point trajectories in semidefinite programming. *SIAM J. Optim.*, 8(4):871–886, 1998.
- [Häg07] D. Hägele. Proof of the cases $p \leq 7$ of the Lieb-Seiringer formulation of the Bessis-Moussa-Villani conjecture. *J. Stat. Phys.*, 127(6):1167–1171, 2007.
- [HdKR02] M. Halická, E. de Klerk, and C. Roos. On the convergence of the central path in semidefinite optimization. *SIAM J. Optim.*, 12(4):1090–1099, 2002.
- [Hel02] J.W. Helton. “Positive” noncommutative polynomials are sums of squares. *Ann. of Math. (2)*, 156(2):675–694, 2002.
- [Hil07] C.J. Hillar. Advances on the Bessis-Moussa-Villani trace conjecture. *Linear Algebra Appl.*, 426(1):130–142, 2007.
- [Hil09] C.J. Hillar. Sums of squares over totally real fields are rational sums of squares. *Proc. Amer. Math. Soc.*, 137(3):921–930, 2009.
- [KLYZ08] E. Kaltofen, B. Li, Z. Yang, and L. Zhi. Exact certification of global optimality of approximate factorizations via rationalizing sums-of-squares with floating point scalars. In *ISSAC 2008*, pages 155–163. ACM, New York, 2008.
- [KP10] I. Klep and J. Povh. Semidefinite programming and sums of hermitian squares of noncommutative polynomials. *J. Pure Appl. Algebra*, 214:740–749, 2010.
- [KS08a] I. Klep and M. Schweighofer. Connes’ embedding conjecture and sums of Hermitian squares. *Adv. Math.*, 217(4):1816–1837, 2008.
- [KS08b] I. Klep and M. Schweighofer. Sums of Hermitian squares and the BMV conjecture. *J. Stat. Phys.*, 133(4):739–760, 2008.
- [LS04] E.H. Lieb and R. Seiringer. Equivalent forms of the Bessis-Moussa-Villani conjecture. *J. Stat. Phys.*, 115(1-2):185–190, 2004.
- [LS09] P.S. Landweber and E.R. Speer. On D. Hägele’s approach to the Bessis-Moussa-Villani conjecture. *Linear Algebra Appl.*, 431(8):1317–1324, 2009.

- [Mit] H. Mittelman. <http://plato.asu.edu/sub/pns.html>.
- [Mou00] P. Moussa. On the representation of $\text{Tr}(e^{(A-\lambda B)})$ as a Laplace transform. *Rev. Math. Phys.*, 12(4):621–655, 2000.
- [NRS10] J. Nie, K. Ranestad, and B. Sturmfels. The algebraic degree of semidefinite programming. *Math. Program.*, 122(2, Ser. A):379–405, 2010.
- [PP08] H. Peyrl and P.A. Parrilo. Computing sum of squares decompositions with rational coefficients. *Theoretical Computer Science*, 409(2):269–281, 2008.
- [PS03] P.A. Parrilo and B. Sturmfels. Minimizing polynomial functions. In *Algorithmic and quantitative real algebraic geometry (Piscataway, NJ, 2001)*, volume 60 of *DIMACS Ser. Discrete Math. Theoret. Comput. Sci.*, pages 83–99. Amer. Math. Soc., Providence, RI, 2003.
- [WSV00] H. Wolkowicz, R. Saigal, and L. Vandenberghe. *Handbook of Semidefinite Programming*. Kluwer, 2000.

KRISTIJAN CAFUTA, UNIVERZA V LJUBLJANI, FAKULTETA ZA ELEKTROTEHNIKO, LABORATORIJ ZA UPORABNO MATEMATIKO, TRŽAŠKA 25, 1000 LJUBLJANA, SLOVENIA

E-mail address: kristijan.cafuta@fe.uni-lj.si

IGOR KLEP, UNIVERZA V MARIBORU, FAKULTETA ZA NARAVOSLOVJE IN MATEMATIKO, KOROŠKA 160, 2000 MARIBOR, AND UNIVERZA V LJUBLJANI, FAKULTETA ZA MATEMATIKO IN FIZIKO, JADRANSKA 19, 1111 LJUBLJANA, SLOVENIA

E-mail address: igor.klep@fmf.uni-lj.si

JANEZ POVH, FAKULTETA ZA INFORMACIJSKE ŠTUDIJE V NOVEM MESTU, NOVI TRG 5, 8000 NOVO MESTO, SLOVENIA

E-mail address: janez.povh@fis.unm.si

NOT FOR PUBLICATION

CONTENTS

1. Introduction	1
1.1. Notation	2
1.2. An xmas tree	3
2. Gram matrix method and semidefinite programming	5
2.1. Gram matrix method	5
2.2. The certificate	5
2.3. Proof of Theorem 2.5	7
2.4. More on rational certificates for SDP	7
Acknowledgments	8
Appendix: The Matlab verification	8
References	10
Index	12