SUMS OF HERMITIAN SQUARES AND THE BMV CONJECTURE

IGOR KLEP AND MARKUS SCHWEIGHOFER

Abstract. We show that all the coefficients of the polynomial

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

are nonnegative whenever $m \leq 13$ is a nonnegative integer and A and B are positive semidefinite matrices of the same size. This has previously been known only for $m \leq 7$. The validity of the statement for arbitrary m has recently been shown to be equivalent to the Bessis-Moussa-Villani conjecture from theoretical physics. In our proof, we establish a connection to sums of hermitian squares of polynomials in noncommuting variables and to semidefinite programming. As a by-product we obtain an example of a real polynomial in two noncommuting variables having nonnegative trace on all symmetric matrices of the same size, yet not being a sum of hermitian squares and commutators.

1. Introduction

While attempting to simplify the calculation of partition functions in quantum statistical mechanics, Bessis, Moussa and Villani (BMV) conjectured in 1975 [BMV] that for any hermitian $n \times n$ matrices A and B with B positive semidefinite, the function

$$\varphi^{A,B} : \mathbb{R} \to \mathbb{R}, \quad t \mapsto \operatorname{tr}\left(e^{A-tB}\right)$$

is the Laplace transform of a positive measure $\mu^{A,B}$ on $\mathbb{R}_{\geq 0}$. That is,

$$\varphi^{A,B}(t) = \int_0^\infty e^{-tx}\, d\mu^{A,B}(x)$$

for all $t \in \mathbb{R}$. By Bernstein's theorem, this is equivalent to $\varphi^{A,B}$ being completely monotone, i.e.,

$$(-1)^s \frac{d^s}{dt^s}\varphi^{A,B}(t) \geq 0$$

for all $s \in \mathbb{N}_0$ and $t \in \mathbb{R}_{\geq 0}$.

Due to its importance (cf. [BMV, LiSe]) there is an extensive literature on this conjecture. Nevertheless it has resisted all attempts at proving it. For an overview of all the approaches before 1998 leading to partial results, we refer the reader to Moussa's survey [Mou].

Date: July 31, 2008.

2000 *Mathematics Subject Classification.* Primary 11E25, 13J30, 15A90; Secondary 15A45, 08B20, 90C22.

Key words and phrases. Bessis-Moussa-Villani (BMV) conjecture, sum of hermitian squares, trace inequality, semidefinite programming.

The first author acknowledges the financial support from the state budget by the Slovenian Research Agency (project No. Z1-9570-0101-06).

Supported by the DFG grant "Barrieren".

In 2004, Lieb and Seiringer [LiSe] achieved a breakthrough paving the way to a series of new attempts at proving the BMV conjecture. They succeeded in restating the conjecture in the following purely algebraic form:

Conjecture 1.1 (BMV, algebraic form)**.** *The polynomial*

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

has only nonnegative coefficients whenever A and B are $n\times n$ positive semidefinite matrices.

The coefficient of t^k in p 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). It is easy to see that these coefficients are real for hermitian A, B.

Suppose A, B are positive semidefinite $n\times n$ matrices. For $k \le 2$ or $m-k \le 2$, each word appearing in $S_{m,k}(A,B)$ has nonnegative trace as is easily seen. This proves the conjecture for $m \le 5$. For $n \le 2$, A can (as always) be assumed to be diagonal and after a diagonal change of basis also B has only nonnegative entries. Hence the conjecture is trivial for $n \le 2$. The first nontrivial case $(m,k,n) = (6,3,3)$ was verified by Hillar and Johnson [HJ] with the help of a computer algebra system by considering entries of both 3×3 matrices, A and B, as scalar and therefore *commuting* variables. Hägele [Häg] shifted the focus from scalars to symbolic computation with matrices (regardless of their size) and gave a surprisingly simple argument settling the case $(m,k) = (7,3)$ and thus also $(m,k) = (7,4)$ by symmetry. Combined with the easy observations from above, this proves Conjecture 1.1 for $m = 7$.

Hägele then deduced the case $m = 6$, which he could not solve directly with his technique, by appealing to the following seminal result due to Hillar [Hi1]: If Conjecture 1.1 is true for m, then it is also true for all $m' < m$ [Hi1, Corollary 1.8]. A strengthening [Hi1, Theorem 1.7] of this result (see Section 4 for a precise statement) is crucial for our main contribution:

Theorem 1.2. *The* BMV *Conjecture* 1.1 *holds for* $m \le 13$.

We exploit semidefinite programming to find certain certificates for nonnegativity of $\operatorname{tr}(S_{m,k}(A,B))$ which are dimensionless (i.e., valid for all n). These certificates are algebraic identities in the ring of polynomials in two *noncommuting* variables involving sums of hermitian squares. The found identities are exact though obtained with the help of numerical computations. But they exist only for certain pairs (m,k) and we have to rely on Hillar's work to deduce Theorem 1.2. For instance, such a sum of hermitian squares certificate does not exist for $(m,k) = (6,3)$, see Example 3.5.

With the benefit of hindsight, Hägele's argument can be read as such a certificate for the case $(m,k) = (7,3)$. However, the certificates we give for $(m,k) = (14,4)$ and $(m,k) = (14,6)$ are much more involved and seem to be impossible to find by hand.

This paper is organized as follows. Section 2 develops the appropriate algebraic framework needed for the desired nonnegativity certificates. In Section 3 the existence of such a certificate is transformed into a linear matrix inequality (LMI) enabling us to search for these certificates using semidefinite programming (SDP). Section 4 explains the overall argument for the proof of Theorem 1.2. The proof itself is presented in full detail in Section 5. A synopsis of our results and other recent developments is given in Section 6, where we also relate the BMV conjecture to another just as old open problem of Connes on II_1 factors. Finally, in the appendix we streamline the proof of the mentioned crucial result of Hillar and give an alternative argument to prove the BMV conjecture for $m = 13$ avoiding Hillar's theorem.

2. From matrices to symbols

The gist of our method is to model the matrices as *noncommuting* variables instead of disaggregating them into scalar entries modeled by *commuting* variables. To this end we introduce the ring of polynomials in two noncommuting variables.

Remark 2.1. It is easy to see [KS2, Lemma 3.15] that the nonnegativity of $\operatorname{tr}(S_{m,k}(A,B))$ for all positive semidefinite *complex* A and B of all sizes need only be checked for all positive semidefinite (in particular symmetric) *real* A and B of all sizes (by identifying $n \times n$ complex matrices with $2n \times 2n$ real matrices). We therefore work over the real numbers.

We write $\langle X, Y \rangle$ for the monoid freely generated by X and Y, i.e., $\langle X, Y \rangle$ consists of *words* in two letters (including the empty word denoted by 1). Let $\mathbb{R}\langle X, Y \rangle$ denote the associative $\mathbb{R}$-algebra freely generated by X and Y. The elements of $\mathbb{R}\langle X, Y \rangle$ are polynomials in the noncommuting variables X and Y with coefficients in $\mathbb{R}$. An element of the form aw where $0 \neq a \in \mathbb{R}$ and $w \in \langle X, Y \rangle$ is called a *monomial* and a its *coefficient*. Hence words are monomials whose coefficient is 1. We endow $\mathbb{R}\langle X, Y \rangle$ with the involution $p \mapsto p^*$ fixing $\mathbb{R} \cup \{X, Y\}$ pointwise. Recall that an *involution* has the properties $(p+q)^* = p^* + q^*$, $(pq)^* = q^* p^*$ and $p^{**} = p$ for all $p, q \in \mathbb{R}\langle X, Y \rangle$. In particular, for each word $w \in \langle X, Y \rangle$, w^* is its reverse.

Definition 2.2. Two polynomials $f, g \in \mathbb{R}\langle X, Y \rangle$ are called *cyclically equivalent* ($f \overset{\text{cyc}}{\sim} g$) if $f - g$ is a sum of commutators in $\mathbb{R}\langle X, Y \rangle$. Here elements of the form $pq - qp$ are called *commutators* ($p, q \in \mathbb{R}\langle X, Y \rangle$).

This definition reflects the fact that $\operatorname{tr}(AB) = \operatorname{tr}(BA)$ for square matrices A and B of the same size. The following proposition shows that cyclic equivalence can easily be checked and will be used tacitly in the sequel. Part (c) is a special case of [KS2, Theorem 2.1] motivating the definition of cyclic equivalence.

Proposition 2.3. (a) *For $v, w \in \langle X, Y \rangle$, we have $v \overset{\text{cyc}}{\sim} w$ if and only if there are $v_1, v_2 \in \langle X, Y \rangle$ such that $v = v_1 v_2$ and $w = v_2 v_1$.*

(b) *Two polynomials $f = \sum_{w \in \langle X,Y \rangle} a_w w$ and $g = \sum_{w \in \langle X,Y \rangle} b_w w$ ($a_w, b_w \in \mathbb{R}$) are cyclically equivalent if and only if for each $v \in \langle X, Y \rangle$,*

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

(c) *Suppose $f \in \mathbb{R}\langle X, Y \rangle$ and $f^* = f$. Then $f \overset{\text{cyc}}{\sim} 0$ if and only if $\operatorname{tr}(f(A,B)) = 0$ for all real symmetric matrices A and B of the same size.*

Definition 2.4. For each subset $S \subseteq \mathbb{R}\langle X, Y \rangle$, we introduce the set

$$\operatorname{Sym} S := \{g \in S \mid g^* = g\}$$

of its *symmetric elements*. Elements of the form g^*g ($g \in \mathbb{R}\langle X,Y\rangle$) are called *hermitian squares*. We denote by

$$\Sigma^2 := \{\sum_i g_i^* g_i \mid g_i \in \mathbb{R}\langle X,Y\rangle\} \subseteq \operatorname{Sym}\mathbb{R}\langle X,Y\rangle$$

the convex cone of all sums of hermitian squares and by

$$\begin{aligned}\Theta^2 &:= \{f \in \mathbb{R}\langle X,Y\rangle \mid \exists g \in \Sigma^2 : f \overset{\text{cyc}}{\sim} g\}\\ &= \Sigma^2 + \{\sum_i (g_i h_i - h_i g_i) \mid g_i, h_i \in \mathbb{R}\langle X,Y\rangle\} \subseteq \mathbb{R}\langle X,Y\rangle\end{aligned}$$

the convex cone of all polynomials that are cyclically equivalent to a sum of hermitian squares.

The following theorem proved in [Hel] also holds for several variables and motivates the use of sums of hermitian squares (see [HP] for a survey of recent developments). We will only use the easy implication from (i) to (ii).

Theorem 2.5 (Helton). *The following are equivalent for $f \in \operatorname{Sym}\mathbb{R}\langle X,Y\rangle$:*

(i) $f \in \Sigma^2$*;*
(ii) $f(A,B)$ *is positive semidefinite for all* $n \in \mathbb{N}$ *and* $A, B \in \operatorname{Sym}\mathbb{R}^{n\times n}$.

To obtain the desired type of certificates we try to merge Proposition 2.3(c) with Theorem 2.5. However, such certificates do not always exist.

Remark 2.6. Consider the following conditions for $f \in \mathbb{R}\langle X,Y\rangle$:

(i) $f \in \Theta^2$;
(ii) $\operatorname{tr}(f(A,B)) \geq 0$ for all $n \in \mathbb{N}$ and $A, B \in \operatorname{Sym}\mathbb{R}^{n\times n}$.

Then (i) implies (ii) but not vice versa. For instance,

$$YX^4Y + XY^4X - 3XY^2X + 1 \in \operatorname{Sym}\mathbb{R}\langle X,Y\rangle$$

satisfies (ii) but not (i) (see [KS2, Example 4.4] for details). Later on we will see further such examples.

3. From symbols to matrices

To search systematically for the certificates just introduced, we develop a *noncommutative* version of the Gram matrix method. The corresponding theory for polynomials in *commuting* variables is well-known and has been studied and used extensively, see e.g. [CLR, PS].

Checking whether a polynomial in noncommuting variables is an element of Σ^2 or Θ^2, respectively, is most efficiently done via the so-called *Gram matrix method*. Given a symmetric $f \in \mathbb{R}\langle X,Y\rangle$ of degree $\leq 2d$ and a vector $\bar{v}$ containing all words in X, Y of degree $\leq d$, there is a real symmetric matrix G with $f = \bar{v}^* G \bar{v}$. (Here $\bar{v}^*$ arises from $\bar{v}$ by applying the involution entrywise to the transposed vector $\bar{v}^t$.) Every such matrix G is called a *Gram matrix* for f. Obviously, the set of all Gram matrices for f is an affine subspace.

Example 3.1. Consider the polynomial

$$h := X^4 + 2XYX + 2X^2 + Y^2 + 2Y + 1 \in \operatorname{Sym}\mathbb{R}\langle X,Y\rangle.$$

Since h has degree four, we choose

$$\bar{v} := [1, X, Y, X^2, XY, YX, Y^2]^t.$$

Then every Gram matrix for h has the form

$$G = \begin{bmatrix} 1 & 0 & 1 & a & 0 & 0 & b \\ 0 & 2-2a & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1-2b & 0 & 0 & 0 & 0 \\ a & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ b & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \in \operatorname{Sym}\mathbb{R}^{7\times 7}.$$

We will revisit this example below.

From Cholesky's decomposition we deduce that $f \in \operatorname{Sym}\mathbb{R}\langle X,Y\rangle$ is a sum of hermitian squares if and only if it has a positive semidefinite Gram matrix. Indeed, if $G = C^*C$ is a positive semidefinite Gram matrix for f, then $f = \bar{v}^*C^*C\bar{v} = (C\bar{v})^*(C\bar{v}) = \sum_i g_i^*g_i \in \Sigma^2$ where $g_i \in \mathbb{R}\langle X,Y\rangle$ is the i-th entry of the vector $C\bar{v}$. The converse follows the same line of reasoning.

Example 3.1 continued. There is no positive semidefinite Gram matrix G for h since the determinant of the submatrix

$$\begin{bmatrix} G_{22} & G_{26} \\ G_{62} & G_{66} \end{bmatrix} = \begin{bmatrix} 2-2a & 1 \\ 1 & 0 \end{bmatrix}$$

is always negative. Hence $h \notin \Sigma^2$.

The existence of a sum of hermitian squares decomposition of $f \in \operatorname{Sym}\mathbb{R}\langle X,Y\rangle$ is equivalent to an LMI feasibility problem. As such it can be decided by solving the SDP

$$\text{minimize } \operatorname{tr}(G) \quad \text{subject to} \quad \bar{v}^*G\bar{v} = f,\ G \text{ positive semidefinite.}$$

Note that $\bar{v}^*G\bar{v} = f$ are just linear constraints on the entries of G as one sees by comparing coefficients. The objective function $G \mapsto \operatorname{tr}(G)$ is often a good choice for finding nice low rank matrices G but can be replaced by any other function linear in the entries of G. If the polynomial is dense (no sparsity), the dimension of the LMI is equal to $(2^{d+1}-1)\times(2^{d+1}-1)$. For more on SDP, we refer the reader to the survey [Tod].

Likewise, checking whether $f \in \Theta^2$ can be done by solving the SDP

$$\text{minimize } \operatorname{tr}(G) \quad \text{subject to} \quad \bar{v}^*G\bar{v} \overset{\text{cyc}}{\sim} f,\ G \text{ positive semidefinite.}$$

By Proposition 2.3(b), $\bar{v}^*G\bar{v} \overset{\text{cyc}}{\sim} f$ are again linear constraints on the entries of G.

For the sake of convenience, from now on a real symmetric matrix G will be called a *Gram matrix* for $f \in \mathbb{R}\langle X,Y\rangle$ (with respect to a vector of words $\bar{v}$) if $f \overset{\text{cyc}}{\sim} \bar{v}^*G\bar{v}$.

Example 3.1 continued. Every Gram matrix (in the new sense) for h has the form

$$\begin{bmatrix} 1 & 0 & 1 & 1-\frac{1}{2}a_1 & -a_2-a_3 & a_2 & \frac{1}{2}-\frac{1}{2}a_4 \\ 0 & a_1 & a_3 & 0 & -a_6-a_7+1 & a_6 & -a_8-a_9 \\ 1 & a_3 & a_4 & a_7 & a_8 & a_9 & 0 \\ 1-\frac{1}{2}a_1 & 0 & a_7 & 1 & -a_{10} & a_{10} & -\frac{1}{2}a_{11}-\frac{1}{2}a_{12} \\ -a_2-a_3 & -a_6-a_7+1 & a_8 & -a_{10} & a_{11} & 0 & -a_5 \\ a_2 & a_6 & a_9 & a_{10} & 0 & a_{12} & a_5 \\ \frac{1}{2}-\frac{1}{2}a_4 & -a_8-a_9 & 0 & -\frac{1}{2}a_{11}-\frac{1}{2}a_{12} & -a_5 & a_5 & 0 \end{bmatrix}.$$

Setting $a_4 = a_7 = 1$ and all other a_i to zero, we get the positive semidefinite matrix $G = \begin{bmatrix}1 & 0 & 1 & 1 & 0 & 0 & 0\end{bmatrix}^* \begin{bmatrix}1 & 0 & 1 & 1 & 0 & 0 & 0\end{bmatrix}$ with corresponding representation $h \overset{\text{cyc}}{\sim} (X^2+Y+1)^2 \in \Sigma^2$, i.e., $h \in \Theta^2$.

In the proof of our main result we will use the Gram matrix method to show that certain $S_{m,k}(X^2,Y^2) \in \Theta^2$. We start by dramatically reducing the sizes of corresponding SDPs with a monomial reduction. For this, we need a technical lemma.

Lemma 3.2. *Let $p_i \in \mathbb{R}\langle X,Y\rangle$.*

(a) *If for $A,B \in \operatorname{Sym}\mathbb{R}^{n\times n}$, $\operatorname{tr}\left(\sum_i (p_i^*p_i)(A,B)\right) = 0$, then $p_i(A,B) = 0$ for all i.*

(b) *If $\sum_i p_i^*p_i \overset{\text{cyc}}{\sim} 0$, then $p_i = 0$ for all i.*

Proof. (a) Denote by e_j the canonical basis vectors of $\mathbb{R}^n$. Then

$$0 = \operatorname{tr}\Big(\sum_i (p_i^*p_i)(A,B)\Big) = \sum_{i,j} \langle (p_i^*p_i)(A,B)e_j, e_j\rangle = \sum_{i,j} \langle p_i(A,B)e_j, p_i(A,B)e_j\rangle.$$

Hence $p_i(A,B)e_j = 0$ for all i,j and thus $p_i(A,B) = 0$ for all i.

(b) If $\sum_i p_i^*p_i \overset{\text{cyc}}{\sim} 0$, then $\operatorname{tr}(\sum_i p_i(A,B)^*p_i(A,B)) = 0$, and by the above, $p_i(A,B) = 0$ for all symmetric A and B of all sizes n. This implies $p_i = 0$ for all i (see e.g. [KS1, Proposition 2.3]). ☐

Not only do we drastically reduce the number of words needed in the Gram method for $S_{m,k}(X^2,Y^2)$ but we also impose a block structure on the Gram matrix G with blocks G_i. This is done in the following proposition. We use self-explanatory notation like $\{X^2,Y^2\}^\ell$ for the set of all words that are concatenations of ℓ copies of X^2 and Y^2.

Proposition 3.3. *Fix $m,k \in \mathbb{N}$.*

(a) *If m and k are even, set*

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

(b) *If m is odd and k is even, set*

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

(c) *If m and k are odd, set*

$$V_1 := \left\{ v \in Y\{X^2, Y^2\}^{\frac{m-1}{2}} \mid \deg_X v = m-k,\, \deg_Y v = k \right\},$$
$$V_2 := \left\{ v \in \{X^2, Y^2\}^{\frac{m-1}{2}} Y \mid \deg_X v = m-k,\, \deg_Y v = k \right\}.$$

(d) *If m is even and k is odd, set*

$$V_1 := \left\{ v \in X\{X^2, Y^2\}^{\frac{m}{2}-1} Y \mid \deg_X v = m-k,\, \deg_Y v = k \right\},$$
$$V_2 := \left\{ v \in Y\{X^2, Y^2\}^{\frac{m}{2}-1} X \mid \deg_X v = m-k,\, \deg_Y v = k \right\}.$$

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 $G_i \in \operatorname{Sym} \mathbb{R}^{V_i \times V_i}$ such that

$$S_{m,k}(X^2, Y^2) \overset{\text{cyc}}{\sim} \sum_i \bar{v}_i^* G_i \bar{v}_i. \tag{1}$$

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

$$S_{m,k}(X^2, Y^2) \overset{\text{cyc}}{\sim} \sum_{i,j} p_{i,j}^* p_{i,j}. \tag{2}$$

Proof. The second statement is clear since

$$\sum_i \bar{v}_i^* G_i \bar{v}_i = \sum_i \bar{v}_i^* C_i^* C_i \bar{v}_i = \sum_i (C_i \bar{v}_i)^* C_i \bar{v}_i = \sum_{i,j} p_{i,j}^* p_{i,j}.$$

We assume without loss of generality that $1 \le k \le m-1$. Suppose that $S_{m,k}(X^2, Y^2) \in \Theta^2$, i.e.,

$$S_{m,k}(X^2, Y^2) \overset{\text{cyc}}{\sim} \sum_j p_j^* p_j \tag{3}$$

for finitely many $0 \ne p_j \in \mathbb{R}\langle X, Y\rangle$. Set $d := \max_j \deg_Y p_j$ and let P_j be the sum of all monomials of degree d with respect to Y appearing in p_j.

Fix real symmetric matrices A and B of the same size. For any real λ, we have $\lambda^{2k} \operatorname{tr}(S_{m,k}(A^2, B^2)) = \operatorname{tr}(\sum_j p_j(A, \lambda B)^* p_j(A, \lambda B))$. We consider this as an equality of real polynomials in λ.

If we assume $d > k$, then $\operatorname{tr}(\sum_j P_j(A, B)^* P_j(A, B)) = 0$ since the degree of the right hand side polynomial cannot exceed the degree of the left hand side polynomial. By (a) of Lemma 3.2, we get $P_j(A, B) = 0$ for all j. Since A and B were arbitrary, this implies $P_j = 0$ by Lemma 3.2(b), contradicting the choice of d. Therefore all monomials appearing in p_j have degree $\le k$ in Y. By similar arguments, one shows that all p_j are actually homogeneous of degree $m-k$ in X and homogeneous of degree k in Y, i.e., $p_j \in \operatorname{span}_{\mathbb{R}} W$ where W is the set of all words of length m with the letter X appearing $m-k$ times and the letter Y appearing k times.

Claim. Suppose we are in one of the cases (a)–(d) and $v_i \in V_i$ for each i. Then $v_i^* v_j \overset{\text{cyc}}{\sim} u$ for some $u \in \{X^2, Y^2\}^m$ if and only if $i = j$.

Proof of claim. The "if" part is immediate. To show the "only if" part, we assume that $i \ne j$ and show that $v_i^* v_j$ contains $YX^\ell Y$ or $XY^\ell X$ as a subword for some odd ℓ. Then the claim follows by Proposition 2.3(a).

The existence of such a subword must be checked case by case. As an example, consider (a). By symmetry arguments, it suffices to look at $v_1^* v_2$ and $v_2^* v_3$. In the former case, the letter at position $m+1$ in $v_1^* v_2$ is an X which is followed to the left and right hand side by finitely many X^2. This block of X's has odd length and is embraced at both ends by a Y since we have assumed $k \geq 1$. In the latter case, there is an X at the m-th and a Y at the $(m+1)$-st position in $v_2^* v_3$. This Y is followed to the right hand side by finitely many Y^2 giving a block of Y's of odd length surrounded by X's.

The other cases (b)–(d) are essentially the same, proving the claim.

Write each p_j as $p_j = \sum_i p_{i,j} + q_j$ where $p_{i,j} \in \operatorname{span}_{\mathbb{R}} V_i$ and $q_j \in \operatorname{span}_{\mathbb{R}} U$ with $U := W \setminus \bigcup_i V_i$. By the claim, $p_j^* p_j = \sum_i p_{i,j}^* p_{i,j} + r_j$ where $\sum_i p_{i,j}^* p_{i,j}$ is a linear combination of words that are cyclically equivalent to a word in $\{X^2, Y^2\}^m$ and r_j is in the linear span of words not cyclically equivalent to a word in $\{X^2, Y^2\}^m$. By part (b) of Proposition 2.3, it follows that (3) can be split into

$$S_{m,k}(X^2, Y^2) \overset{\text{cyc}}{\sim} \sum_{i,j} p_{i,j}^* p_{i,j} \qquad \text{and} \qquad 0 \overset{\text{cyc}}{\sim} \sum_j r_j.$$

Now let J be the index set consisting of all j and define matrices $C_i \in \mathbb{R}^{J \times V_i}$ by $[p_{i,j}]_{j \in J} = C_i \bar{v}_i$. Then the matrices $G_i := C_i^* C_i$ are positive semidefinite and satisfy (1). $\square$

We illustrate the proposition by two examples.

Example 3.4. We have $S_{8,4}(X^2, Y^2) \in \Theta^2$. For instance, with

$$\begin{aligned}
\bar{v}_1 &= [Y^2X^2Y^2X^2,\ Y^4X^4,\ X^2Y^4X^2,\ Y^2X^4Y^2,\ X^4Y^4,\ X^2Y^2X^2Y^2]^t, \\
\bar{v}_2 &= [XY^4X^3,\ XY^2X^2Y^2X,\ X^3Y^4X]^t, \\
\bar{v}_3 &= [Y^3X^4Y,\ YX^2Y^2X^2Y,\ YX^4Y^3]^t
\end{aligned}$$

and

$$G_1 = \begin{bmatrix} 4 & 4 & 0 & 3 & 1 & 1 \\ 4 & 4 & 0 & 3 & 1 & 1 \\ 0 & 0 & 3 & 0 & 3 & 3 \\ 3 & 3 & 0 & 3 & 0 & 0 \\ 1 & 1 & 3 & 0 & 4 & 4 \\ 1 & 1 & 3 & 0 & 4 & 4 \end{bmatrix}, \quad G_2 = G_3 = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix},$$

$S_{8,4}(X^2, Y^2) \overset{\text{cyc}}{\sim} \sum_{i=1}^3 \bar{v}_i^* G_i \bar{v}_i$. The matrices G_i which we found using SDP are positive semidefinite as can be seen from their characteristic polynomials

$$\begin{aligned}
p_{G_1} &= -108t^3 + 129t^4 - 22t^5 + t^6 \in \mathbb{R}[t], \\
p_{G_2} = p_{G_3} &= 2t^2 - t^3 \in \mathbb{R}[t].
\end{aligned}$$

Alternatively, we can use the Cholesky decompositions $G_i = C_i^* C_i$ for

$$C_1 = \frac{1}{2}\begin{bmatrix} 4 & 4 & 0 & 3 & 1 & 1 \\ 0 & 0 & 2\sqrt{3} & 0 & 2\sqrt{3} & 2\sqrt{3} \\ 0 & 0 & 0 & \sqrt{3} & -\sqrt{3} & -\sqrt{3} \end{bmatrix}, \quad C_2 = C_3 = \begin{bmatrix} 1 & 0 & -1 \end{bmatrix}.$$

A first nontrivial nonnegativity certificate of this type was found in an ad hoc fashion by Hägele [Häg], namely

$$\begin{aligned}
S_{7,3}(X^2, Y^2) \overset{\text{cyc}}{\sim}\ & 7(Y^2X^4Y)^*(Y^2X^4Y) + \\
& 7(X^2Y^2X^2Y + X^4Y^3)^*(X^2Y^2X^2Y + X^4Y^3) \in \Sigma^2.
\end{aligned} \tag{4}$$

This proves Conjecture 1.1 for $m = 7$ (since the cases $k \le 2$ and $m - k \le 2$ are trivial and $S_{7,4}(X^2, Y^2) = S_{7,3}(Y^2, X^2) \in \Theta^2$). Note that the representation (4) uses only words from V_1 of Proposition 3.3(c). Hägele also showed that there is no such representation for $S_{6,3}(X^2, Y^2)$ using only words from V_1 of Proposition 3.3(d). However, he speculated that admitting more words might lead to such a representation meaning in our setup that $S_{6,3}(X^2, Y^2) \in \Theta^2$. Our next example proves that this is not the case.

Example 3.5. We show that $S_{6,3}(X^2, Y^2) \notin \Theta^2$. Suppose, by way of contradiction, that $S_{6,3}(X^2, Y^2) \in \Theta^2$. Then by Proposition 3.3(d), with the basis

$$V = \{Y^3X^3,\ YX^2Y^2X,\ XY^2X^2Y,\ X^3Y^3\}$$

we can find a positive semidefinite Gram matrix for $S_{6,3}(X^2, Y^2)$ that is block diagonal of the form

$$G_{6,3} = \begin{bmatrix} a_{11} & a_{12} & 0 & 0 \\ a_{12} & a_{22} & 0 & 0 \\ 0 & 0 & b_{11} & b_{12} \\ 0 & 0 & b_{12} & b_{22} \end{bmatrix} \in \mathbb{R}^{4\times 4}.$$

With $\bar{v} = [v]_{v\in V}$, it follows from $S_{6,3}(X^2, Y^2) \overset{\text{cyc}}{\sim} \bar{v}^* G_{6,3} \bar{v}$ that

$$G_{6,3} = \begin{bmatrix} a_{11} & a_{12} & 0 & 0 \\ a_{12} & a_{22} & 0 & 0 \\ 0 & 0 & 2 - a_{22} & 6 - a_{12} \\ 0 & 0 & 6 - a_{12} & 6 - a_{11} \end{bmatrix}.$$

For a positive semidefinite matrix of this form, $0 \le a_{11} \le 6$, $0 \le a_{22} \le 2$,

$$a_{12}^2 \le a_{11}a_{22}, \tag{5}$$

$$(6 - a_{12})^2 \le (6 - a_{11})(2 - a_{22}). \tag{6}$$

By adding (5) and (6), we obtain

$$36 - 12a_{12} + 2a_{12}^2 \le 12 - 2a_{11} - 6a_{22} + 2a_{11}a_{22}.$$

As $-2a_{11} - 6a_{22} + 2a_{11}a_{22} = a_{22}(a_{11} - 6) + a_{11}(a_{22} - 2) \le 0$, this implies

$$0 \ge a_{12}^2 - 6a_{12} + 12 = (a_{12} - 3)^2 + 3,$$

a contradiction. Hence $S_{6,3}(X^2, Y^2) \notin \Theta^2$.

4. Strategy of the proof

An important ingredient in the proof of Theorem 1.2 will be the following descent result of Hillar [Hi1, Theorem 1.7]:

Theorem 4.1 (Hillar)**.** *The failure of Conjecture* 1.1 *for a certain* (m, k) *implies failure for all* (m', k') *with* $m' - k' \ge m - k$ *and* $k' \ge k$*.*

In view of this theorem it suffices to prove Conjecture 1.1 for $(m, k) = (14, 4)$ and $(m, k) = (14, 6)$. To do this we apply our Gram matrix method to prove that $S_{14,4}(X^2, Y^2) \in \Theta^2$ and $S_{14,6}(X^2, Y^2) \in \Theta^2$.

Since the search for positive semidefinite Gram matrices is done by SDP, the entries of the found matrices are only floating point numbers and do not provide a sound proof for the existence of a certificate of nonnegativity. However, in our case, there happen to exist such Gram matrices with *rational* entries and we have employed several strategies and heuristics to find them.

First, we have detected symmetries and patterns in the numerical solutions and imposed them as additional constraints in subsequent SDPs. Second, we have worked with different objective functions in order to find solutions with some "nice" rational entries that could be fixed. Finally, we have employed rounding techniques involving heuristics to guess the prime factors appearing in the denominators of the presumably rational entries. All too often, we have however lost numerical stability and had to backtrack in this manually guided refinement process.

For a systematic treatment of finding exact rational sum of squares certificates for polynomials in *commuting* variables we refer the reader to [PP], see also [Hi2] and the references therein.

5. Proof of Theorem 1.2

As mentioned above, it suffices to show that $S_{14,4}(X^2,Y^2), S_{14,6}(X^2,Y^2) \in \Theta^2$ (cf. the table on page 15 below). Let

$$\begin{aligned}\bar{v}_{14,4} =& [Y^2X^{10}Y^2,\ X^4Y^2X^2Y^2X^4,\ X^6Y^4X^4,\ X^2Y^2X^6Y^2X^2,\ X^4Y^2X^4Y^2X^2,\\ & X^8Y^4X^2+X^6Y^2X^2Y^2X^2,\ X^4Y^4X^6Y^2+X^2Y^2X^8Y^2,\\ & X^{10}Y^4+X^8Y^2X^2Y^2+X^6Y^2X^4Y^2]^t\end{aligned}$$

and

$$G_{14,4} = \begin{bmatrix} 7 & 0 & 0 & 0 & 0 & 0 & 7 & 7 \\ 0 & 7 & 7 & 0 & 7 & 7 & 0 & 0 \\ 0 & 7 & 14 & 0 & 7 & 7 & 0 & 0 \\ 0 & 0 & 0 & 7 & 7 & 7 & 7 & 7 \\ 0 & 7 & 7 & 7 & 14 & 14 & 7 & 7 \\ 0 & 7 & 7 & 7 & 14 & 14 & 7 & 7 \\ 7 & 0 & 0 & 7 & 7 & 7 & 14 & 14 \\ 7 & 0 & 0 & 7 & 7 & 7 & 14 & 14 \end{bmatrix}.$$

Then $S_{14,4}(X^2,Y^2) \overset{\text{cyc}}{\sim} \bar{v}_{14,4}^* G_{14,4} \bar{v}_{14,4}$. The matrix $G_{14,4}$ is positive semidefinite with Cholesky decomposition $G_{14,4} = L_{14,4}^* L_{14,4}$, where

$$L_{14,4} = \sqrt{7} \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}.$$

We now consider $S_{14,6}(X^2,Y^2)$. Let $A_{14,6}$ be the symmetric 15×15 matrix from page 12 and

$$\begin{aligned}\bar{u}_{14,6} =& [Y^3X^6Y^2X^2Y,\ YX^2Y^2X^2Y^2X^4Y,\ Y^3X^4Y^2X^4Y,\ YX^2Y^4X^6Y,\\ & Y^3X^2Y^2X^6Y,\ Y^5X^8Y,\ YX^4Y^4X^4Y,\ YX^2Y^2X^4Y^2X^2Y,\ Y^3X^8Y^3,\\ & YX^8Y^5,\ YX^6Y^2X^2Y^3,\ YX^6Y^4X^2Y,\ YX^4Y^2X^4Y^3,\\ & YX^4Y^2X^2Y^2X^2Y,\ YX^2Y^2X^6Y^3]^t.\end{aligned}$$

From the matrices on pages 13 and 14 we form a symmetric 35×35 matrix $B_{14,6}$ as follows: The top left 18×19 block is given by the matrix on page 13, the bottom left 17×19 block is given on page 14 and the other entries are obtained from

$$[B_{14,6}]_{i,j} = [B_{14,6}]_{36-j,36-i} \quad \text{for} \quad i,j > 19.$$

Let

$$\begin{aligned}\bar{w}_{14,6} =[&Y^2X^2Y^2X^6Y^2,\ Y^4X^8Y^2,\ Y^2X^6Y^4X^2,\ Y^2X^4Y^2X^2Y^2X^2,\ X^2Y^4X^4Y^2X^2,\\ &Y^2X^2Y^2X^4Y^2X^2,\ Y^4X^6Y^2X^2,\ X^2Y^2X^2Y^4X^4,\ Y^2X^4Y^4X^4,\\ &X^2Y^4X^2Y^2X^4,\ Y^2X^2Y^2X^2Y^2X^4,\ Y^4X^4Y^2X^4,\ X^2Y^6X^6,\ Y^2X^2Y^4X^6,\\ &Y^4X^2Y^2X^6,\ Y^6X^8,\ X^4Y^6X^4,\ X^2Y^2X^2Y^2X^2Y^2X^2,\ Y^2X^4Y^2X^4Y^2,\\ &X^8Y^6,\ X^6Y^2X^2Y^4,\ X^6Y^4X^2Y^2,\ X^6Y^6X^2,\ X^4Y^2X^4Y^4,\\ &X^4Y^2X^2Y^2X^2Y^2,\ X^4Y^2X^2Y^4X^2,\ X^4Y^4X^4Y^2,\ X^4Y^4X^2Y^2X^2,\\ &X^2Y^2X^6Y^4,\ X^2Y^2X^4Y^2X^2Y^2,\ X^2Y^2X^4Y^4X^2,\ X^2Y^2X^2Y^2X^4Y^2,\\ &X^2Y^4X^6Y^2,\ Y^2X^8Y^4,\ Y^2X^6Y^2X^2Y^2]^t\end{aligned}$$

Then

$$S_{14,6}(X^2,Y^2) \overset{\text{cyc}}{\sim} \bar{u}_{14,6}^* A_{14,6}\bar{u}_{14,6} + \bar{w}_{14,6}^* B_{14,6}\bar{w}_{14,6}. \tag{7}$$

Both matrices $A_{14,6}$ and $B_{14,6}$ are positive semidefinite as is easily checked by looking at the corresponding characteristic polynomials using symbolic computation. Hence $S_{14,6}(X^2,Y^2) \in \Theta^2$. By Theorem 4.1, this proves the BMV conjecture for $m \leq 13$.

Remark 5.1. The word vectors $\bar{u}_{14,6}$ and $\bar{w}_{14,6}$ as well as the matrices on pages 12, 13 and 14 can be found in the Mathematica notebook that is available with the electronic version of the source of this article:

`http://arxiv.org/abs/0710.1074`

In the same file we also provide code that verifies the nonnegativity certificate (7) when executed.

$$
\begin{bmatrix}
\frac{737}{45} & -\frac{13}{10} & \frac{247}{90} & \frac{497}{90} & \frac{497}{90} & \frac{497}{90} & -\frac{8}{5} & \frac{7}{2} & \frac{122}{35} & \frac{3}{19} & \frac{3}{19} & \frac{3}{19} & \frac{746}{243} & \frac{413}{180} & \frac{199}{32} \\
-\frac{13}{10} & \frac{7}{5} & -\frac{7}{15} & -\frac{19}{30} & -\frac{19}{30} & -\frac{19}{30} & \frac{221}{162} & \frac{17}{10} & -\frac{17}{20} & \frac{5}{4} & \frac{5}{4} & \frac{5}{4} & -\frac{43}{81} & -\frac{48}{49} & \frac{413}{180} \\
\frac{247}{90} & -\frac{7}{15} & 3 & 0 & 0 & 0 & -\frac{31}{6} & -\frac{392}{81} & \frac{5377}{1215} & \frac{175}{972} & \frac{175}{972} & \frac{175}{972} & \frac{1437}{500} & -\frac{43}{81} & \frac{746}{243} \\
\frac{497}{90} & -\frac{19}{30} & 0 & \frac{7}{3} & \frac{7}{3} & \frac{7}{3} & \frac{235}{247} & \frac{227}{90} & -\frac{13}{90} & -\frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & \frac{175}{972} & \frac{5}{4} & \frac{3}{19} \\
\frac{497}{90} & -\frac{19}{30} & 0 & \frac{7}{3} & \frac{7}{3} & \frac{7}{3} & \frac{235}{247} & \frac{227}{90} & -\frac{13}{90} & -\frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & \frac{175}{972} & \frac{5}{4} & \frac{3}{19} \\
\frac{497}{90} & -\frac{19}{30} & 0 & \frac{7}{3} & \frac{7}{3} & \frac{7}{3} & \frac{235}{247} & \frac{227}{90} & -\frac{13}{90} & -\frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & \frac{175}{972} & \frac{5}{4} & \frac{3}{19} \\
-\frac{8}{5} & \frac{221}{162} & -\frac{31}{6} & \frac{235}{247} & \frac{235}{247} & \frac{235}{247} & \frac{2251}{200} & \frac{2251}{200} & -\frac{18211}{2240} & \frac{235}{247} & \frac{235}{247} & \frac{235}{247} & -\frac{31}{6} & \frac{221}{162} & -\frac{8}{5} \\
\frac{7}{2} & \frac{17}{10} & -\frac{392}{81} & \frac{227}{90} & \frac{227}{90} & \frac{227}{90} & \frac{2251}{200} & \frac{3902}{225} & -\frac{373}{45} & \frac{227}{90} & \frac{227}{90} & \frac{227}{90} & -\frac{392}{81} & \frac{17}{10} & \frac{7}{2} \\
\frac{122}{35} & -\frac{17}{20} & \frac{5377}{1215} & -\frac{13}{90} & -\frac{13}{90} & -\frac{13}{90} & -\frac{18211}{2240} & -\frac{373}{45} & \frac{712}{105} & -\frac{13}{90} & -\frac{13}{90} & -\frac{13}{90} & \frac{5377}{1215} & -\frac{17}{20} & \frac{122}{35} \\
\frac{3}{19} & \frac{5}{4} & \frac{175}{972} & -\frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & \frac{235}{247} & \frac{227}{90} & -\frac{13}{90} & \frac{7}{3} & \frac{7}{3} & \frac{7}{3} & 0 & -\frac{19}{30} & \frac{497}{90} \\
\frac{3}{19} & \frac{5}{4} & \frac{175}{972} & -\frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & \frac{235}{247} & \frac{227}{90} & -\frac{13}{90} & \frac{7}{3} & \frac{7}{3} & \frac{7}{3} & 0 & -\frac{19}{30} & \frac{497}{90} \\
\frac{3}{19} & \frac{5}{4} & \frac{175}{972} & -\frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & \frac{235}{247} & \frac{227}{90} & -\frac{13}{90} & \frac{7}{3} & \frac{7}{3} & \frac{7}{3} & 0 & -\frac{19}{30} & \frac{497}{90} \\
\frac{746}{243} & -\frac{43}{81} & \frac{1437}{500} & \frac{175}{972} & \frac{175}{972} & \frac{175}{972} & -\frac{31}{6} & -\frac{392}{81} & \frac{5377}{1215} & 0 & 0 & 0 & 3 & -\frac{7}{15} & \frac{247}{90} \\
\frac{413}{180} & -\frac{48}{49} & -\frac{43}{81} & \frac{5}{4} & \frac{5}{4} & \frac{5}{4} & \frac{221}{162} & \frac{17}{10} & -\frac{17}{20} & -\frac{19}{30} & -\frac{19}{30} & -\frac{19}{30} & -\frac{7}{15} & \frac{7}{5} & -\frac{13}{10} \\
\frac{199}{32} & \frac{413}{180} & \frac{746}{243} & \frac{3}{19} & \frac{3}{19} & \frac{3}{19} & -\frac{8}{5} & \frac{7}{2} & \frac{122}{35} & \frac{497}{90} & \frac{497}{90} & \frac{497}{90} & \frac{247}{90} & -\frac{13}{10} & \frac{737}{45}
\end{bmatrix}
$$

$$\left[\begin{array}{ccccccccccccccccccc}
9 & 9 & 5 & 7 & 0 & \frac{7}{3} & \frac{7}{3} & 0 & \frac{11}{3} & 0 & \frac{7}{3} & \frac{7}{3} & 0 & \frac{7}{3} & \frac{7}{3} & \frac{7}{3} & 0 & 0 & \frac{5}{2} \\
9 & 9 & 5 & 7 & 0 & \frac{7}{3} & \frac{7}{3} & 0 & \frac{11}{3} & 0 & \frac{7}{3} & \frac{7}{3} & 0 & \frac{7}{3} & \frac{7}{3} & \frac{7}{3} & 0 & 0 & \frac{5}{2} \\
5 & 5 & 28 & \frac{7}{2} & 5 & 7 & 7 & \frac{16}{19} & -\frac{21}{2} & \frac{16}{19} & 7 & 7 & \frac{16}{19} & 7 & 7 & 7 & 2 & 4 & \frac{13}{3} \\
7 & 7 & \frac{7}{2} & \frac{6349}{200} & 14 & 7 & 7 & 8 & 11 & \frac{373}{90} & 7 & 7 & \frac{373}{90} & 7 & 7 & 7 & \frac{22}{9} & 2 & \frac{23}{2} \\
0 & 0 & 5 & 14 & 25 & \frac{7}{2} & \frac{7}{2} & \frac{1066}{81} & \frac{7}{2} & \frac{3494}{741} & \frac{7}{2} & \frac{7}{2} & \frac{3494}{741} & \frac{7}{2} & \frac{7}{2} & \frac{7}{2} & \frac{85}{27} & 7 & 0 \\
\frac{7}{3} & \frac{7}{3} & 7 & 7 & \frac{7}{2} & 7 & 7 & \frac{7}{2} & 7 & \frac{7}{2} & 7 & 7 & \frac{7}{2} & 7 & 7 & 7 & 1 & \frac{7}{2} & \frac{14}{3} \\
\frac{7}{3} & \frac{7}{3} & 7 & 7 & \frac{7}{2} & 7 & 7 & \frac{7}{2} & 7 & \frac{7}{2} & 7 & 7 & \frac{7}{2} & 7 & 7 & 7 & 1 & \frac{7}{2} & \frac{14}{3} \\
0 & 0 & \frac{16}{19} & 8 & \frac{1066}{81} & \frac{7}{2} & \frac{7}{2} & 21 & 10 & 8 & \frac{7}{2} & \frac{7}{2} & 8 & \frac{7}{2} & \frac{7}{2} & \frac{7}{2} & 18 & \frac{3}{2} & -\frac{5}{2} \\
\frac{11}{3} & \frac{11}{3} & -\frac{21}{2} & 11 & \frac{7}{2} & 7 & 7 & 10 & 28 & 6 & 7 & 7 & 6 & 7 & 7 & 7 & 6 & 4 & -\frac{1}{2} \\
0 & 0 & \frac{16}{19} & \frac{373}{90} & \frac{3494}{741} & \frac{7}{2} & \frac{7}{2} & 8 & 6 & 5 & \frac{7}{2} & \frac{7}{2} & 5 & \frac{7}{2} & \frac{7}{2} & \frac{7}{2} & \frac{11}{2} & -\frac{1}{4} & 0 \\
\frac{7}{3} & \frac{7}{3} & 7 & 7 & \frac{7}{2} & 7 & 7 & \frac{7}{2} & 7 & \frac{7}{2} & 7 & 7 & \frac{7}{2} & 7 & 7 & 7 & 1 & \frac{7}{2} & \frac{14}{3} \\
\frac{7}{3} & \frac{7}{3} & 7 & 7 & \frac{7}{2} & 7 & 7 & \frac{7}{2} & 7 & \frac{7}{2} & 7 & 7 & \frac{7}{2} & 7 & 7 & 7 & 1 & \frac{7}{2} & \frac{14}{3} \\
0 & 0 & \frac{16}{19} & \frac{373}{90} & \frac{3494}{741} & \frac{7}{2} & \frac{7}{2} & 8 & 6 & 5 & \frac{7}{2} & \frac{7}{2} & 5 & \frac{7}{2} & \frac{7}{2} & \frac{7}{2} & \frac{11}{2} & -\frac{1}{4} & 0 \\
\frac{7}{3} & \frac{7}{3} & 7 & 7 & \frac{7}{2} & 7 & 7 & \frac{7}{2} & 7 & \frac{7}{2} & 7 & 7 & \frac{7}{2} & 7 & 7 & 7 & 1 & \frac{7}{2} & \frac{14}{3} \\
\frac{7}{3} & \frac{7}{3} & 7 & 7 & \frac{7}{2} & 7 & 7 & \frac{7}{2} & 7 & \frac{7}{2} & 7 & 7 & \frac{7}{2} & 7 & 7 & 7 & 1 & \frac{7}{2} & \frac{14}{3} \\
\frac{7}{3} & \frac{7}{3} & 7 & 7 & \frac{7}{2} & 7 & 7 & \frac{7}{2} & 7 & \frac{7}{2} & 7 & 7 & \frac{7}{2} & 7 & 7 & 7 & 1 & \frac{7}{2} & \frac{14}{3} \\
0 & 0 & 2 & \frac{22}{9} & \frac{85}{27} & 1 & 1 & 18 & 6 & \frac{11}{2} & 1 & 1 & \frac{11}{2} & 1 & 1 & 1 & \frac{7396}{315} & -\frac{11}{3} & -\frac{16}{3} \\
0 & 0 & 4 & 2 & 7 & \frac{7}{2} & \frac{7}{2} & \frac{3}{2} & 4 & -\frac{1}{4} & \frac{7}{2} & \frac{7}{2} & -\frac{1}{4} & \frac{7}{2} & \frac{7}{2} & \frac{7}{2} & -\frac{11}{3} & \frac{52}{3} & -5
\end{array}\right]$$

$$
\begin{bmatrix}
\frac{5}{2} & \frac{5}{2} & \frac{13}{3} & \frac{23}{2} & 0 & \frac{14}{3} & \frac{14}{3} & -\frac{5}{2} & -\frac{1}{2} & 0 & \frac{14}{3} & \frac{14}{3} & 0 & \frac{14}{3} & \frac{14}{3} & \frac{14}{3} & -\frac{16}{3} & -5 & 28 \\
-\frac{77}{90} & -\frac{77}{90} & -\frac{7}{6} & \frac{9}{5} & 1 & -2 & -2 & 1 & -\frac{10}{3} & -2 & -2 & -2 & -2 & -2 & -2 & -2 & 1 & \frac{7}{2} & \frac{14}{3} \\
-\frac{77}{90} & -\frac{77}{90} & -\frac{7}{6} & \frac{9}{5} & 1 & -2 & -2 & 1 & -\frac{10}{3} & -2 & -2 & -2 & -2 & -2 & -2 & -2 & 1 & \frac{7}{2} & \frac{14}{3} \\
-\frac{77}{90} & -\frac{77}{90} & -\frac{7}{6} & \frac{9}{5} & 1 & -2 & -2 & 1 & -\frac{10}{3} & -2 & -2 & -2 & -2 & -2 & -2 & -2 & 1 & \frac{7}{2} & \frac{14}{3} \\
0 & 0 & -1 & -1 & -\frac{13}{4} & -2 & -2 & 1 & -\frac{31}{27} & -2 & -2 & -2 & -2 & -2 & -2 & -2 & \frac{11}{2} & -\frac{1}{4} & 0 \\
-\frac{77}{90} & -\frac{77}{90} & -\frac{7}{6} & \frac{9}{5} & 1 & -2 & -2 & 1 & -\frac{10}{3} & -2 & -2 & -2 & -2 & -2 & -2 & -2 & 1 & \frac{7}{2} & \frac{14}{3} \\
-\frac{77}{90} & -\frac{77}{90} & -\frac{7}{6} & \frac{9}{5} & 1 & -2 & -2 & 1 & -\frac{10}{3} & -2 & -2 & -2 & -2 & -2 & -2 & -2 & 1 & \frac{7}{2} & \frac{14}{3} \\
0 & 0 & -1 & -1 & -\frac{13}{4} & -2 & -2 & 1 & -\frac{31}{27} & -2 & -2 & -2 & -2 & -2 & -2 & -2 & \frac{11}{2} & -\frac{1}{4} & 0 \\
-\frac{28829}{4480} & -\frac{28829}{4480} & -\frac{55591}{20007} & -8 & 0 & -\frac{10}{3} & -\frac{10}{3} & \frac{7}{2} & -\frac{757}{81} & -\frac{31}{27} & -\frac{10}{3} & -\frac{10}{3} & -\frac{31}{27} & -\frac{10}{3} & -\frac{10}{3} & -\frac{10}{3} & 6 & 4 & -\frac{1}{2} \\
0 & 0 & \frac{9}{2} & -\frac{229}{81} & -\frac{1327}{972} & 1 & 1 & \frac{109987}{10080} & \frac{7}{2} & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 18 & \frac{3}{2} & -\frac{5}{2} \\
-\frac{77}{90} & -\frac{77}{90} & -\frac{7}{6} & \frac{9}{5} & 1 & -2 & -2 & 1 & -\frac{10}{3} & -2 & -2 & -2 & -2 & -2 & -2 & -2 & 1 & \frac{7}{2} & \frac{14}{3} \\
-\frac{77}{90} & -\frac{77}{90} & -\frac{7}{6} & \frac{9}{5} & 1 & -2 & -2 & 1 & -\frac{10}{3} & -2 & -2 & -2 & -2 & -2 & -2 & -2 & 1 & \frac{7}{2} & \frac{14}{3} \\
0 & 0 & \frac{99031}{13440} & -\frac{44}{3} & -\frac{1240243}{162000} & 1 & 1 & -\frac{1327}{972} & 0 & -\frac{13}{4} & 1 & 1 & -\frac{13}{4} & 1 & 1 & 1 & \frac{85}{27} & 7 & 0 \\
-\frac{413}{180} & -\frac{413}{180} & \frac{1369}{180} & -\frac{195323}{22050} & -\frac{44}{3} & \frac{9}{5} & \frac{9}{5} & -\frac{229}{81} & -8 & -1 & \frac{9}{5} & \frac{9}{5} & -1 & \frac{9}{5} & \frac{9}{5} & \frac{9}{5} & \frac{22}{9} & 2 & \frac{23}{2} \\
1 & 1 & 6 & \frac{1369}{180} & \frac{99031}{13440} & -\frac{7}{6} & -\frac{7}{6} & \frac{9}{2} & -\frac{55591}{20007} & -1 & -\frac{7}{6} & -\frac{7}{6} & -1 & -\frac{7}{6} & -\frac{7}{6} & -\frac{7}{6} & 2 & 4 & \frac{13}{3} \\
-\frac{2246}{315} & -\frac{2246}{315} & 1 & -\frac{413}{180} & 0 & -\frac{77}{90} & -\frac{77}{90} & 0 & -\frac{28829}{4480} & 0 & -\frac{77}{90} & -\frac{77}{90} & 0 & -\frac{77}{90} & -\frac{77}{90} & -\frac{77}{90} & 0 & 0 & \frac{5}{2} \\
-\frac{2246}{315} & -\frac{2246}{315} & 1 & -\frac{413}{180} & 0 & -\frac{77}{90} & -\frac{77}{90} & 0 & -\frac{28829}{4480} & 0 & -\frac{77}{90} & -\frac{77}{90} & 0 & -\frac{77}{90} & -\frac{77}{90} & -\frac{77}{90} & 0 & 0 & \frac{5}{2}
\end{bmatrix}
$$

6. Concluding remarks

6.1. **Current state of the BMV conjecture.** The following table shows the examples we have computed on an ordinary PC running Mathematica with the NCAlgebra package [HMS], Yalmip [Löf] and the SDP solver SeDuMi [Stu]. Most of the computations took a few seconds, some of them a few minutes.

m	$k=0$, $k=1$, $k=2$, …, $k=m-1$, $k=m$
0	+
1	+ +
2	+ + +
3	+ + + +
4	+ + + + +
5	+ + + + + +
6	+ + + ⊖ + + +
7	+ + + ⊕ ⊕ + + +
8	+ + + − ⊕ − + + +
9	+ + + − ⊕ ⊕ − + + +
10	+ + + − ⊕ − ⊕ − + + +
11	+ + + **+** **+** − − **+** **+** + + +
12	+ + + − **+** − − − **+** − + + +
13	+ + + − **+** − − − − **+** − + + +
14	+ + + − ⊕ − ⊕ − ⊕ − ⊕ − + + +
15	+ + + − **+** − − − − − − **+** − + + +
16	+ + + − **+** − − − − − − − **+** − + + +
17	+ + + − **+** − − − − − − − − **+** − + + +
18	+ + + − **+** − − − ? ? ? − − − **+** − + + +
19	+ + + − **+** − − − ? ? ? ? − − − **+** − + + +

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

symbol	meaning
+	$S_{m,k}$ is in Θ^2 for trivial reasons
⊕	$S_{m,k}$ is in Θ^2 (with proof)
+	$S_{m,k}$ is in Θ^2 (numerical evidence)
⊖	$S_{m,k}$ is not in Θ^2 (with proof)
−	$S_{m,k}$ is not in Θ^2 (numerical evidence)

Legend

While finishing our paper, Landweber and Speer sent us a closely related preprint [LaSp] where they prove for example that $S_{m,4}(X^2, Y^2) \in \Theta^2$ for odd m and that $S_{11,3}(X^2, Y^2) \in \Theta^2$. Their certificates only use words from V_1 of Proposition 3.3.

They also give results on the negative side, which imply by Proposition 6.1 below that $S_{m,k}(X^2, Y^2) \notin \Theta^2$ in the following cases:

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

The compatibility between our setup and the setup of Landweber and Speer [LaSp] is provided by the following proposition communicated to us by Eugene Speer. We thank him for letting us include this result.

Proposition 6.1. *Retain the notation from Proposition* 3.3 *and assume that m or k is odd. Then $S_{m,k}(X^2, Y^2) \in \Theta^2$ if and only if $S_{m,k}(X^2, Y^2) \overset{\text{cyc}}{\sim} \bar{v}_1^* G_1 \bar{v}_1$ for some positive semidefinite G_1 (or equivalently, if and only if $S_{m,k}(X^2, Y^2) \overset{\text{cyc}}{\sim} \bar{v}_2^* G_2 \bar{v}_2$ for some positive semidefinite G_2).*

Proof. One direction is trivial and for the converse suppose that $S_{m,k}(X^2, Y^2) \in \Theta^2$. Then by Proposition 3.3, $S_{m,k}(X^2, Y^2) \overset{\text{cyc}}{\sim} \sum_{i=1}^2 \bar{v}_i^* G_i \bar{v}_i$ for some positive semidefinite G_1, G_2. Note that $w \in V_1$ if and only if $w^* \in V_2$. Hence,

$$\bar{v}_1^* G_1 \bar{v}_1 = \sum_{v,u \in V_1} v^* (G_1)_{vu} u = \sum_{w,z \in V_2} w (G_1')_{wz} z^* \overset{\text{cyc}}{\sim} \sum_{w,z \in V_2} z^* (G_1')_{wz} w = \bar{v}_2^* G_1' \bar{v}_2,$$

where G_1' is a positive semidefinite matrix obtained from G_1 by a relabelling of rows and columns. Thus

$$S_{m,k}(X^2, Y^2) \overset{\text{cyc}}{\sim} \sum_{i=1}^2 \bar{v}_i^* G_i \bar{v}_i \overset{\text{cyc}}{\sim} \bar{v}_2^* (G_1' + G_2) \bar{v}_2$$

and similarly $S_{m,k}(X^2, Y^2) \overset{\text{cyc}}{\sim} \bar{v}_1^* (G_1 + G_2') \bar{v}_1$. $\square$

Independently of the work of Landweber and Speer, the doctoral student Burgdorf [Bur], initially guided by further numerical experiments, found a combinatorial proof of $S_{m,4}(X^2, Y^2) \in \Theta^2$ for all m.

To summarize, the table on page 15 can be updated as follows:

8	$+\ +\ +\ \ominus\ \oplus\ \ominus\ +\ +\ +$
9	$+\ +\ +\ \ominus\ \oplus\ \oplus\ \ominus\ +\ +\ +$
10	$+\ +\ +\ \ominus\ \oplus\ \ominus\ \oplus\ \ominus\ +\ +\ +$
11	$+\ +\ +\ \oplus\ \oplus\ \ominus\ \ominus\ \oplus\ \oplus\ +\ +\ +$
12	$+\ +\ +\ \ominus\ \oplus\ \ominus\ -\ \ominus\ \oplus\ \ominus\ +\ +\ +$
13	$+\ +\ +\ \ominus\ \oplus\ \ominus\ \ominus\ \ominus\ \ominus\ \oplus\ \ominus\ +\ +\ +$
14	$+\ +\ +\ \ominus\ \oplus\ \ominus\ \oplus\ \ominus\ \oplus\ \ominus\ \oplus\ \ominus\ +\ +\ +$
15	$+\ +\ +\ \ominus\ \oplus\ \ominus\ \ominus\ \ominus\ \ominus\ \ominus\ \ominus\ \oplus\ \ominus\ +\ +\ +$
16	$+\ +\ +\ \ominus\ \oplus\ \ominus\ -\ \ominus\ -\ \ominus\ -\ \ominus\ \oplus\ \ominus\ +\ +\ +$
17	$+\ +\ +\ \ominus\ \oplus\ \ominus\ \ominus\ \ominus\ \ominus\ \ominus\ \ominus\ \ominus\ \ominus\ \oplus\ \ominus\ +\ +\ +$
18	$+\ +\ +\ \ominus\ \oplus\ \ominus\ -\ \ominus\ ?\ \ominus\ ?\ \ominus\ -\ \ominus\ \oplus\ \ominus\ +\ +\ +$
19	$+\ +\ +\ \ominus\ \oplus\ \ominus\ \ominus\ \ominus\ \ominus\ \ominus\ \ominus\ \ominus\ \ominus\ \ominus\ \ominus\ \oplus\ \ominus\ +\ +\ +$
20	$+\ +\ +\ \ominus\ \oplus\ \ominus\ ?\ \ominus\ ?\ \ominus\ ?\ \ominus\ ?\ \ominus\ ?\ \ominus\ \oplus\ \ominus\ +\ +\ +$
21	$+\ +\ +\ \ominus\ \oplus\ \ominus\ \ominus\ \ominus\ \ominus\ \ominus\ \ominus\ \ominus\ \ominus\ \ominus\ \ominus\ \ominus\ \ominus\ \ominus\ \oplus\ \ominus\ +\ +\ +$
22	$+\ +\ +\ \ominus\ \oplus\ \ominus\ ?\ \ominus\ ?\ \ominus\ ?\ \ominus\ ?\ \ominus\ ?\ \ominus\ ?\ \ominus\ \oplus\ \ominus\ +\ +\ +$

Moreover, the table continues like one would expect from looking at the lines $m = 19, 20, 21, 22$. Hillar's descent Theorem 4.1 together with positive results for $k = 4$ (by Landweber and Speer and, independently, by Burgdorf) establishes Conjecture 1.1 for $k \leq 4$ and $m - k \leq 4$. Also, there is still the possibility of proving the BMV conjecture in the same manner by replacing a suitable sequence of ?, which only occur for even m and k, by $\oplus$.

Very recently, using analytical methods, Fleischhack [Fle] and, independently, Friedland [Fri] have shown the following: For *fixed* positive semidefinite A, B and $k \in \mathbb{N}$ there is an $m' \geq k$, such that $\operatorname{tr} S_{m,k}(A, B) \geq 0$ for all $m \geq m'$. If m' could be chosen *independently* of A, B, then Conjecture 1.1 would follow by Hillar's descent theorem.

6.2. **Relation to Connes' embedding conjecture.** In [KS2] we studied the following conditions for real symmetric polynomials f in noncommuting variables $\bar{X} := (X_1, \ldots, X_r)$:

(i) $\operatorname{tr}(f(A_1, \ldots, A_r)) \geq 0$ for all $n \in \mathbb{N}$ and all $A_i \in \operatorname{Sym} \mathbb{R}^{n \times n}$ with $\|A_i\| \leq 1$;
(ii) $\tau(f(a_1, \ldots, a_r)) \geq 0$ for all II_1-factors $\mathcal{F}$ and all $a_i \in \operatorname{Sym} \mathcal{F}$ with $\|a_i\| \leq 1$;
(iii) $\forall \varepsilon \in \mathbb{R}_{>0}\ \exists g \in \mathbb{R}\langle \bar{X} \rangle$:

$$f + \varepsilon \overset{\text{cyc}}{\sim} g \in M := \{\sum_i g_i^* g_i + \sum_{i,j} h_{ij}^* (1 - X_i^2) h_{ij} \mid g_i, h_{i,j} \in \mathbb{R}\langle \bar{X} \rangle\}.$$

We proved that (ii) and (iii) are equivalent and imply (i). Moreover, we showed that the converse implication (i) $\Rightarrow$ (ii) is equivalent to an old conjecture of Connes about type II_1-factors.

In Example 3.5 we have seen that $S_{6,3}(X^2, Y^2) \notin \Theta^2$, hence the tracial version of Helton's sum of hermitian squares theorem [Hel] fails (cf. also Remark 2.6). By homogeneity, even $S_{6,3}(X^2, Y^2) + \varepsilon \notin \Theta^2$ for all $\varepsilon \in \mathbb{R}$. Similarly, there is no $g \in M$ with $S_{6,3}(X^2, Y^2) \overset{\text{cyc}}{\sim} g$ although $S_{6,3}(X^2, Y^2)$ satisfies (i). However, it is unknown whether $S_{6,3}(X^2, Y^2)$ satisfies (ii) (or equivalently, (iii)). If it does not, then Connes' embedding conjecture fails.

Acknowledgments. We would like to thank Christopher Hillar for introducing the second author to the BMV conjecture at an IMA workshop in Minneapolis. The main part of the work was done at the Universität Konstanz, the former host institution of the second author, during a stay of the first author financed by the DFG. Part of the work was also done during the Real Algebraic Geometry workshop in Oberwolfach in March 2007. A preliminary report of this work appeared in the Oberwolfach reports [KS3]. We would like to thank Peter Landweber and Eugene Speer for the careful reading of a previous version of the manuscript. They provided us with a detailed list of comments and corrections that improved the exposition as well as some of the results. Also, Pierre Moussa and Sabine Burgdorf contributed some valuable remarks. Finally, we would like to thank two anonymous referees for their suggestions which greatly contributed to the overall presentation.

Appendix A. Euler-Lagrange equations

Hillar's proof of the descent Theorem 4.1 relies on [Hi1, Corollary 3.6]. In this section we prove a similar statement, Lemma A.1, which can alternatively be used to prove the descent theorem by a simple inspection of Hillar's proof.

Our proof of Lemma A.1 uses only Lagrange multipliers and is shorter and simpler than Hillar's variational proof of [Hi1, Corollary 3.6]. However, the two results are not entirely reconcilable.

For a variational approach to the original form of the BMV conjecture, we refer the reader to [LeC], see also [Mou].

Lemma A.1. *Given $n \in \mathbb{N}$, suppose that (A, B) minimizes $\operatorname{tr}(S_{m,k}(A^2, B^2))$ among all symmetric $A, B \in \mathbb{R}^{n\times n}$ of Hilbert-Schmidt norm 1. Suppose further that A and B are positive semidefinite. Then*

$$AS_{m-1,k}(A^2, B^2) = \frac{m-k}{m} \operatorname{tr}(S_{m,k}(A^2, B^2))A \qquad \textit{and} \tag{8}$$

$$BS_{m-1,k-1}(A^2, B^2) = \frac{k}{m} \operatorname{tr}(S_{m,k}(A^2, B^2))B. \tag{9}$$

Proof. We actually prove more. We fix an arbitrary $B \in \operatorname{Sym}\mathbb{R}^{n\times n}$ and show that (8) holds when a positive semidefinite matrix A minimizes $\operatorname{tr}(S_{m,k}(A^2, B^2))$ among all $A \in \operatorname{Sym}\mathbb{R}^{n\times n}$ with $\|A\|_{\mathrm{HS}} = 1$. Then a corresponding statement will hold for (9) by symmetry. Recall that the Hilbert-Schmidt norm on $\operatorname{Sym}\mathbb{R}^{n\times n}$ is induced by the scalar product given by $\langle A, B\rangle_{\mathrm{HS}} := \operatorname{tr}(AB) = \sum_{i,j} A_{i,j}B_{i,j}$. We use the method of Lagrange multipliers and therefore compute the first derivatives of the functions $f, g : \operatorname{Sym}\mathbb{R}^{n\times n} \to \mathbb{R}$ given by

$$f : A \mapsto \operatorname{tr}(S_{m,k}(A^2, B^2)) \qquad \text{and} \qquad g : A \mapsto \operatorname{tr}(A^2) = \|A\|^2_{\mathrm{HS}}.$$

The derivatives $Df(A)[H]$ and $Dg(A)[H]$ at $A \in \operatorname{Sym}\mathbb{R}^{n\times n}$ along the direction $H \in \operatorname{Sym}\mathbb{R}^{n\times n}$ are the coefficients of the linear terms of $f(A+\lambda H)$ and $g(A+\lambda H)$ considered as polynomials in λ, respectively. Since

$$g(A+\lambda H) = \operatorname{tr}((A+\lambda H)(A+\lambda H)) = \operatorname{tr}(A^2) + \lambda(\operatorname{tr}(AH) + \operatorname{tr}(HA)) + \lambda^2 \operatorname{tr}(H^2),$$

we get $Dg(A)[H] = \operatorname{tr}(AH) + \operatorname{tr}(HA) = \operatorname{tr}(2AH) = \langle 2A, H\rangle$, i.e., the gradient of g in A is $\nabla g(A) = 2A$.

The calculation of $Df(A)[H]$ is more complicated, but follows the same scheme, namely that one occurrence of A^2 at a time can be replaced by AH or HA. The idea is the same as in the proof of [Hi1, Lemma 2.1]. We have

$$\begin{aligned}
0 &= \operatorname{tr}\left(\sum_{i=1}^{m}(A^2+tB^2)^{i-1}((AH+HA)-(AH+HA))(A^2+tB^2)^{m-i}\right)\\
&= \operatorname{tr}\left(m(AH+HA)(A^2+tB^2)^{m-1}\right) -\\
&\quad \operatorname{tr}\left(\sum_{i=1}^{m}(A^2+tB^2)^{i-1}(AH+HA)(A^2+tB^2)^{m-i}\right)
\end{aligned}$$

and the coefficient of t^k in the last expression is

$$\operatorname{tr}(m(AH+HA)S_{m-1,k}(A^2, B^2)) - Df(A)[H].$$

This implies

$$Df(A)[H] = \langle m(AS_{m-1,k}(A^2, B^2) + S_{m-1,k}(A^2, B^2)A), H\rangle$$

and therefore $\nabla f(A) = m(AS_{m-1,k}(A^2, B^2) + S_{m-1,k}(A^2, B^2)A)$.

If A is now a minimizer as stated, then we obtain a Lagrange multiplier $\mu \in \mathbb{R}$ such that $\nabla f(A) = \mu\nabla g(A)$ (since $\nabla g(A) = 2A \neq 0$), i.e.,

$$AS_{m-1,k}(A^2, B^2) + S_{m-1,k}(A^2, B^2)A = \mu A. \tag{10}$$

We now subtract the two equations that can be obtained from (10) by multiplication with A from the left and right, respectively, and see that A^2 commutes with $S_{m-1,k}(A^2,B^2)$. If A is in addition positive semidefinite, then also A commutes with $S_{m-1,k}(A^2,B^2)$. Therefore (10) becomes $AS_{m-1,k}(A^2,B^2)=\frac{\mu}{2}A$. Moreover,

$$\frac{\mu}{2}=\operatorname{tr}(\frac{\mu}{2}A^2)=\operatorname{tr}(A^2S_{m-1,k}(A^2,B^2))=\frac{m-k}{m}\operatorname{tr}(S_{m,k}(A^2,B^2))$$

by [Hi1, Lemma 2.1]. □

Appendix B. Self-contained proof of Conjecture 1.1 for $m=13$

Instead of Hillar's descent Theorem 4.1 one can use special features of the found nonnegativity certificates for $S_{14,4}(X^2,Y^2)$ and $S_{14,6}(X^2,Y^2)$ to deduce Conjecture 1.1 for m *equal to* 13. We include this since the ideas might be helpful in future algebraic approaches to the BMV conjecture.

Retain the notation from Section 5. From the Cholesky decomposition of $G_{14,4}$ we deduce that

$$S_{14,4}(X^2,Y^2)\overset{\text{cyc}}{\sim}\sum_{i=1}^{4}g_i^*g_i$$

for

$$\begin{aligned}
g_1&=\sqrt{7}(Y^2X^{10}Y^2+X^4Y^4X^6Y^2+X^2Y^2X^8Y^2+X^{10}Y^4+X^8Y^2X^2Y^2+\\
&\qquad X^6Y^2X^4Y^2),\\
g_2&=\sqrt{7}(X^4Y^2X^2Y^2X^4+X^6Y^4X^4+X^4Y^2X^4Y^2X^2+X^8Y^4X^2+\\
&\qquad X^6Y^2X^2Y^2X^2),\\
g_3&=\sqrt{7}X^6Y^4X^4,\\
g_4&=\sqrt{7}(X^2Y^2X^6Y^2X^2+X^4Y^2X^4Y^2X^2+X^8Y^4X^2+X^6Y^2X^2Y^2X^2+\\
&\qquad X^4Y^4X^6Y^2+X^2Y^2X^8Y^2+X^{10}Y^4+X^8Y^2X^2Y^2+X^6Y^2X^4Y^2).
\end{aligned}$$

We now turn to $S_{14,6}(X^2,Y^2)$. Let $[1]_{35\times 35}$ be the 35×35 matrix with all entries equal to 1. Then $B_{14,6}-\lambda[1]_{35\times 35}$ is positive semidefinite whenever

$$\lambda\le\frac{588889450102066403443857277324727 1387}{6345100314096416989598091089889990510969779}\approx 9.281\cdot 10^{-7}.$$

As $\bar{w}_{14,6}^*[1]_{35\times 35}\bar{w}_{14,6}=S_{7,3}(X^2,Y^2)^2$, this implies that for some $h_i\in\mathbb{R}\langle\bar{X}\rangle$,

$$S_{14,6}(X^2,Y^2)\overset{\text{cyc}}{\sim}10^{-7}S_{7,3}(X^2,Y^2)^2+\sum_i h_i^*h_i. \tag{11}$$

We are now ready to prove Conjecture 1.1 for $m=13$. It is easy to see that $S_{13,k}(X^2,Y^2)\in\Theta^2$ for $k\in\{0,1,2,11,12,13\}$. Let us consider $S_{13,3}(A^2,B^2)$ for positive semidefinite $A,B\in\mathbb{R}^{n\times n}$. Suppose there are such A,B with

$$\operatorname{tr}(S_{13,3}(A^2,B^2))<0. \tag{12}$$

By Lemma A.1, we may without loss of generality assume that A and B satisfy (8) and (9) (with $m=13$ and $k=3$). Then $AS_{12,3}(A^2,B^2)$ and $BS_{12,2}(A^2,B^2)$ are negative semidefinite, A commutes with $S_{12,3}(A^2,B^2)$ and B commutes with $S_{12,2}(A^2,B^2)$. Hence

$$S_{13,3}(A^2,B^2)=A^2S_{12,3}(A^2,B^2)+B^2S_{12,2}(A^2,B^2)$$

is negative semidefinite and so is $BS_{13,3}(A^2,B^2)B$. By the above, $S_{14,4}(X^2,Y^2)\in\Theta^2$, so

$$0\leq \operatorname{tr}(S_{14,4}(A^2,B^2))=\frac{14}{10}\operatorname{tr}(B^2S_{13,3}(A^2,B^2))=\frac{14}{10}\operatorname{tr}(BS_{13,3}(A^2,B^2)B)\leq 0.$$

(For the first equality see e.g. [Hi1, Lemma 2.1].) As $S_{14,4}(X^2,Y^2)\overset{\text{cyc}}{\sim}\sum_{i=1}^4 g_i^*g_i$ with $g_3=\sqrt{7}X^6Y^4X^4$ and $\operatorname{tr}(S_{14,4}(A^2,B^2))=0$, $A^6B^4A^4=0$ by Lemma 3.2. In particular, $\operatorname{tr}((B^2A^5)^*(B^2A^5))=0$, hence $B^2A^5=0$. Repeating this we obtain $BA^{5/2}=A^{5/2}B=0$. But then $S_{13,3}(A^2,B^2)=0$, contradicting (12). This proves the BMV conjecture for $(m,k)\in\{(13,3),(13,10)\}$. Similarly, the cases $(m,k)=(13,4)$ and $(m,k)=(13,9)$ can be handled.

Let us now consider $S_{13,5}(A^2,B^2)$ for positive semidefinite $A,B\in\mathbb{R}^{n\times n}$. Suppose there are such A,B with

$$\operatorname{tr}(S_{13,5}(A^2,B^2))<0. \tag{13}$$

As before, we can deduce that $\operatorname{tr}(S_{14,6}(A^2,B^2))=0$. From (11) it follows that $S_{7,3}(A^2,B^2)=0$. By (4), this implies $B^2A^4B=0$, thus $B^{3/2}A^2=A^2B^{3/2}=0$. Therefore $S_{13,5}(A^2,B^2)=0$, contradicting (13). This settles Conjecture 1.1 for $(m,k)\in\{(13,5),(13,8)\}$. To conclude the proof we note that the two remaining cases $(m,k)=(13,6)$ and $(m,k)=(13,7)$ can be handled similarly.

References

[BMV] D. Bessis, P. Moussa, M. Villani: Monotonic converging variational approximations to the functional integrals in quantum statistical mechanics, J. Math. Phys. **16**, 2318–2325 (1975)

[Bur] S. Burgdorf: Sums of Hermitian Squares as an Approach to the BMV Conjecture, preprint `http://arxiv.org/abs/0802.1153`

[CLR] M.D. Choi, T.Y. Lam, B. Reznick: Sums of squares of real polynomials, *Proc. Sympos. Pure Math.*, **58**, Part 2, 103–126 (1995)

[Fle] C. Fleischhack: Asymptotic Positivity of Hurwitz Product Traces, preprint `http://arxiv.org/abs/0804.3665`

[Fri] S. Friedland: Remarks on BMV conjecture, preprint `http://arxiv.org/abs/0804.3948`

[Häg] D. Hägele: Proof of the cases $p\leq 7$ of the Lieb-Seiringer formulation of the Bessis-Moussa-Villani conjecture, J. Statist. Phys. **127**, no. 6, 1167–1171 (2007)

[Hel] J.W. Helton: "Positive" noncommutative polynomials are sums of squares, Ann. of Math. (2) **156**, no. 2, 675–694 (2002)

[HMS] J.W. Helton, R.L. Miller, M. Stankus: NCAlgebra: A Mathematica Package for Doing Non Commuting Algebra, available from `http://www.math.ucsd.edu/~ncalg/`

[HP] J.W. Helton, M. Putinar: Positive polynomials, the spectral theorem and optimization, to appear in the IMA volume on Applications of Algebraic Geometry `http://arxiv.org/abs/math/0612103`

[Hi1] C.J. Hillar: Advances on the Bessis-Moussa-Villani Trace Conjecture, Linear Algebra Appl. **426**, no. 1, 130–142 (2007)

[Hi2] C.J. Hillar: Sums of polynomial squares over totally real fields are rational sums of squares, Proc. Amer. Math. Soc., to appear `http://arxiv.org/abs/0704.2824`

[HJ] C.J. Hillar, C.R. Johnson: On the positivity of the coefficients of a certain polynomial defined by two positive definite matrices, J. Stat. Phys. **118**, no. 3–4, 781–789 (2005)

[KS1] I. Klep, M. Schweighofer: A Nichtnegativstellensatz for polynomials in noncommuting variables, Israel J. Math. **161**, no. 1, 17–27 (2007)

[KS2] I. Klep, M. Schweighofer: Connes' embedding conjecture and sums of hermitian squares, Adv. Math. **217**, no. 4, 1816–1837 (2008)

[KS3] I. Klep, M. Schweighofer: Sums of hermitian squares, Connes' embedding problem and the BMV Conjecture, Oberwolfach Reports, vol. 4, no. 1, 779–782 (2007)

[LeC] K.J. Le Couteur: Representation of the function $\mathrm{Tr}(\exp(A - \lambda B))$ as a Laplace transform with positive weight and some matrix inequalities, J. Phys. A **13**, no. 10, 3147–3159 (1980)

[Löf] J. Löfberg: YALMIP: A Toolbox for Modeling and Optimization in MATLAB, Proceedings of the CACSD Conference, Taipei, Taiwan, 284–289 (2004) `http://control.ee.ethz.ch/~joloef/yalmip.php`

[LaSp] P.S. Landweber, E.R. Speer: On D. Hägele's approach to the Bessis-Moussa-Villani conjecture, preprint `http://arxiv.org/abs/0711.0672`

[LiSe] E.H. Lieb, R. Seiringer: Equivalent forms of the Bessis-Moussa-Villani conjecture, J. Statist. Phys. **115**, no. 1–2, 185–190 (2004)

[Mou] P. Moussa: On the representation of $\mathrm{Tr}(e^{(A-\lambda B)})$ as a Laplace transform, Rev. Math. Phys. **12**, no. 4, 621–655 (2000)

[PS] P.A. Parrilo, B. Sturmfels: Minimizing polynomial functions, Ser. Discrete Math. Theor. Comput. Sci. **60**, 83–99 (2003)

[PP] H. Peyrl, P.A. Parrilo: A Macaulay 2 package for computing sum of squares decompositions of polynomials with rational coefficients, Proceedings of the 2007 international workshop on symbolic-numeric computation, London, Ontario, Canada, 207–208 (2007)

[Stu] J. Sturm: Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones, Optim. Methods Softw. **11/12**, no. 1–4, 625–653 (1999)

[Tod] M.J. Todd: Semidefinite optimization, Acta Numerica **10**, 515–560 (2001)

Igor Klep, Univerza v Ljubljani, Oddelek za matematiko Inštituta za matematiko, fiziko in mehaniko, Jadranska 19, 1111 Ljubljana, Slovénie

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

Markus Schweighofer, Université de Rennes 1, Laboratoire de Mathématiques, Campus de Beaulieu, 35042 Rennes cedex, France

E-mail address: `markus.schweighofer@univ-rennes1.fr`