On D. Hägele’s approach to the Bessis-Moussa-Villani conjecture

Peter S. Landweber* and Eugene R. Speer**
Department of Mathematics
Rutgers University
New Brunswick, New Jersey 08903 USA

Abstract. The reformulation of the Bessis-Moussa-Villani conjecture given by Lieb and Seiringer asserts that the coefficient $\alpha_{p,r}(A, B)$ of $t^r$ in the polynomial $\text{Tr}(A + tB)^p$, with $A, B$ positive semidefinite matrices, is nonnegative for all $p, r$. We propose a natural extension of a method of attack on this problem due to Hägele, and investigate for what values of $p, r$ the method is successful, obtaining a complete determination when either $p$ or $r$ is odd.

Key words and phrases. Bessis-Moussa-Villani (BMV) conjecture, positive definite matrices, trace inequalities.

2000 Mathematics Subject Classification. 15A90, 15A48, 15A45.

1. Introduction

In [3], Daniel Hägele gives an ingenious and simple proof that if $A$ and $B$ are $n \times n$ positive semidefinite matrices then for $p = 7$ all coefficients $\alpha_{p,r}(A, B)$ of $t$ in the polynomial

$$\text{Tr}(A + tB)^p \equiv \sum_{r=0}^{p} \alpha_{p,r}(A, B)t^r,$$

where $\text{Tr} M$ denotes the trace of the matrix $M$, are nonnegative. If this result could be proved for general $p$ it would imply [7] a conjecture of Bessis, Moussa, and Villani [1]. On the other hand, it was also shown in [3] that the same method does not suffice to prove the positivity of $\alpha_{6,3}$ (we will occasionally abbreviate “$\alpha_{p,r}(A, B) \geq 0$ for all positive semidefinite $A, B$” as “$\alpha_{p,r} \geq 0$”). Thus it is of interest to investigate for what values of $p$ and $r$ the method does or does not succeed in establishing $\alpha_{p,r} \geq 0$.

In this note we give several results, both negative and positive, in this direction. We must to some extent consider separately two possible cases, according to the parity of $p$ and $r$, and in each of these cases we define two related integers $k$ and $q$:

**Case 1:** $p$ and $r$ are odd. Then $p = 2k + 1$, $r = 2q + 1$;

**Case 2:** $p$ is even and $r$ is odd. Then $p = 2k + 2$, $r = 2q + 1$.

One further case,

**Case 3:** $p$ is odd and $r$ even,
is included implicitly; it is easy to verify that all our results for Case 1 imply corresponding results for Case 3, obtained by replacing $r$ with $p - r$. We will not consider in detail the
case in which both $p$ and $r$ are even; results in this case have been obtained by Klep and
Schweighofer \cite{5, 6} and by Burgdorf \cite{2}, as we discuss briefly in Section 4. In each of Cases 1
and 2 we define precisely a proof strategy which is the natural generalization of that of \cite{3}
and investigate its success. We are able to classify completely the pairs $(p, r)$ for which the
method succeeds; unfortunately, although these include one infinite class $(p, r$ odd with
$r = p - 4)$, the method does not succeed in enough cases to establish the BMV conjecture.

Results of this sort should be viewed in the light of an important theorem of Hillar
\cite{4}, which implies that if $\alpha_{p, r} \geq 0$ then also $\alpha_{p', r'} \geq 0$ if $p \geq p'$, $r \geq r'$, and $p - r \geq p' - r'$.
For example, it is pointed out in \cite{3} that although the proof method used there does not
apply directly when $p = 6$, $r = 3$, the nonnegativity of $\alpha_{6, 3}$ follows from the corresponding
result for $p = 7$, $r = 3$; similarly, our result that $\alpha_{p, p - 4} \geq 0$ for $p$ odd implies the positivity
of $\alpha_{p, r}$, for all $p$, when $r \leq 4$ or $r \geq p - 4$. Moreover, it follows that to establish the full
BMV conjecture it suffices to establish positivity of $\alpha_{p_n, r_n}$ for some sequences $p_n, r_n$ with
$p_n \to \infty$, $r_n \to \infty$, and $p_n - r_n \to \infty$ as $n \to \infty$. Our results leave open the possibility of
proving the BMV conjecture by successfully applying the method of \cite{3} to such a sequence
with $p_n, r_n$ even.

In order to describe the method more precisely we write $X_0 \equiv A$ and $X_1 \equiv B$. Let
$E_{p, r}$ be the set of binary strings of length $p$, $s = s_1 \cdots s_p$, containing exactly $r$ 1’s, and for
$s \in E_{p, r}$ write $Y_s = X_{s_1} \cdots X_{s_p}$. Then

$$\alpha_{p, r}(A, B) = \sum_{s \in E_{p, r}} \text{Tr}(Y_s).$$

(2)

Now for coefficients $c = (c_u)_{u \in E_{k, q}} \in \mathbb{C}^{E_{k, q}}$ define $Z(c) = \sum_{u \in E_{k, q}} c_u Y_u$. Then we will have
$\alpha_{p, r}(A, B) \geq 0$ if we can show that for some appropriately chosen $c^{(m)} = (c^{(m)}_u)_{u \in E_{k, q}},$
$1 \leq m \leq M$,

$$\alpha_{p, r}(A, B) = \begin{cases} \sum_m \text{Tr}(Z(c^{(m)})BZ(c^{(m)})^*), & \text{in case 1,} \\ \sum_m \text{Tr}(Z(c^{(m)})BZ(c^{(m)})^*A), & \text{in case 2.} \end{cases}$$

(3)

This follows from the fact that if $a$ and $b$ are the nonnegative square roots of $A$ and
$B$, respectively, then $\text{Tr}(Z(c)BZ(c)^*) = \text{Tr}[(Z(c)b)(Z(c)b)^*]$ and $\text{Tr}(Z(c)BZ(c)^*A) = \text{Tr}[(aZ(c)b)(aZ(c)b)^*]$.

To relate (2) with (3) we must make explicit the effect of the invariance of the trace
under cyclic permutations. Let $\tilde{E}_{p, r}$ be the set of equivalence classes of $E_{p, r}$ modulo cyclic
permutations, with $\pi : E_{p, r} \to \tilde{E}_{p, r}$ the canonical projection. Then (2) becomes

$$\alpha_{p, r}(A, B) = \sum_{\tilde{s} \in \tilde{E}_{p, r}} |\tilde{s}| \text{Tr}(Y_{s(\tilde{s})}),$$

(4)

where $|\tilde{s}|$ is the number of elements in $\tilde{s}$ and $s(\tilde{s})$ is some element of $\tilde{s}$. Similarly, if we
define $\sigma \equiv \sigma_{p, r} : E_{k, q} \to E_{p, r}$ by

$$\sigma_{p, r}(u, v) = \begin{cases} u_1 \cdots u_k 1 v_k \cdots v_1, & \text{in case 1,} \\ u_1 \cdots u_k 1 v_k \cdots v_1 0, & \text{in case 2,} \end{cases}$$

2
then the right hand side of (3) becomes

$$\sum_{\tilde{s} \in \tilde{E}_{p,r}} \sum_{(u,v) \in (\pi \sigma)^{-1}(\tilde{s})} \sum_{m} c^{(m)}_u c^{(m)}_v \text{Tr}(Y_{\sigma(u,v)}),$$

(5)

so that (3) will hold for all \(A, B\) if for all \(\tilde{s} \in \tilde{E}_{p,r}\),

$$|\tilde{s}| = \sum_{(u,v) \in (\pi \sigma)^{-1}(\tilde{s})} \sum_{m} c^{(m)}_u c^{(m)}_v.$$  

(6)

The generalization of the method of [3] referred to above is established.

**Condition H:** There exist \(M \geq 1\) and coefficients \(c^{(m)}, m = 1, \ldots, M\), such that (6) is satisfied for all \(\tilde{s} \in \tilde{E}_{p,r}\).

Before proceeding we verify a fact which is obviously necessary for the existence of such \(c^{(m)}\).

**Proposition 1:** For any \(p\) and \(r\) and any \(\tilde{s} \in \tilde{E}_{p,r}\) there exist \(u, v \in \tilde{E}_{k,q}\) such that \(\sigma_{p,r}(u, v) \in \tilde{s}\).

**Proof:** We give the proof in Case 1; Case 2 is similar. A useful geometric picture (the reader might draw a sketch) is obtained by letting \(C \subset \mathbb{C}\) denote the set of \(p\)th roots of unity and identifying an element \(s = s_1 \cdots s_p \in E_{p,r}\) with a map \(s : C \to \{0, 1\}\) labeling the elements of \(C\); the identification is via \(s(\exp 2j\pi i/p) = s_j, j = 1, \ldots, p\). Any \(\theta \in \mathbb{R}\) defines the line \(L_\theta\) in \(C\) through the origin and the point \(z_\theta = \exp i\theta\), oriented from the origin toward \(z_\theta\). Let \(N_1(\theta)\) be the number of points \(\omega \in C\) for which \(s(\omega) = 1\) and which lie to the right of \(L_\theta\), let \(N_2(\theta)\) be the number of such points which lie to the left of \(L_\theta\), and let \(N(\theta) = N_1(\theta) - N_2(\theta)\). \(N(\theta)\) is odd unless \(\pm z_\theta \in C\) with \(s(\pm z_\theta) = 1\), in which case it is even, and if \(N(\theta_0) = 0\) for some \(\theta_0\) then we can immediately read off the desired \(u, v\). But taking \(\theta\) with \(\pm z_\theta \notin C\), so that \(N(\theta)\) is odd, we observe that \(N(\theta + \pi) = -N(\theta)\) and so \(N(\theta_0) = 0\) for some intermediate \(\theta_0\). ■

2. Positive results

In this section we show that Condition H holds in the following cases:

**Case 1:** \(r = 1; r = p - 2; r = p - 4\); and \(p = 11, r = 3\). The cases \(r = 1\) and \(r = p - 2\) are easy (in each case one takes \(M = 1\) and \(c^{(1)}_u = 1\) for all \(u \in E_{k,q}\)); the remaining cases are covered in Theorems 2 and 3 below.

**Case 2:** \(r = 1\) and \(r = p - 1\). These follow the pattern of the two easy cases above; verification is left to the reader.

**Theorem 2:** Condition H holds if \(p = 11\) and \(r = 3\).
Proof: Defining

\[ Z_1 = Y_{00001} + Y_{00010} + Y_{00100} + Y_{01000} - Y_{10000}, \]
\[ Z_2 = \sqrt{2} (Y_{00100} - Y_{01000} - Y_{10000}), \]
\[ Z_3 = 2 (Y_{00100} - Y_{01000}), \]
\[ Z_4 = 2 Y_{01000}, \]

and using the fact that, since \( p = 11 \) is prime, \( |\tilde{s}| = 11 \) for all \( \tilde{s} \in \tilde{E}_{11,3} \), one finds easily that (compare \( (3) \), case 1)

\[ \alpha_{11,3}(A, B) = 11 \sum_{i=1}^{4} \text{Tr}(Z_i B Z_i^*). \]

We remark that both positive and negative coefficients occur among the \( c_u^{(m)} \) implicitly defined by \( (7) \). It can easily be shown that no solution in which all the coefficients are positive is possible; this is in contrast to the situation for the case \( p = 7, r = 3 \) discussed in [3] and for the cases treated in Theorem 3 below.

**Theorem 3:** Condition \( H \) holds if \( p \) is odd and \( r = p - 4 \).

Note that the case \( p = 7, r = 3 \) of this theorem appears in [3]; the case \( p = 9, r = 5 \) was obtained by Klep and Schweighofer (see [5]). After we had completed our work we learned that Theorem 3 was obtained independently by Burgdorf [2].

The theorem will follow almost immediately from the next lemma.

**Lemma 4:** Let \( p = 2k + 1 \geq 5 \) and let \( r = p - 4 = 2q + 1 \). Then \( E_{k,q} \) may be partitioned as \( E_{k,q} = \bigcup_{m=1}^{k-1} D_m \) in such a way that for every \( \tilde{s} \in \tilde{E}_{p,r} \) there exists a unique \( m, 1 \leq m \leq k - 1 \), and unique \( u, v \in D_m \), such that \( \sigma(u, v) \in \tilde{s} \).

**Proof of Theorem 3:** Set \( p = 2k + 1 \) and \( p - 4 = 2q + 1 \). We must find coefficients \( c^{(m)} = (c_u^{(m)})_{u \in E_{k,q}} \) satisfying \( (6) \); since \( p \) and \( p - 4 \) are relatively prime, \( |\tilde{s}| = p \) for every \( \tilde{s} \in \tilde{E}_{p,r} \) and so equivalently we must find \( c^{(m)} \) satisfying

\[ \sum_{m} \sum_{(u,v) \in (\pi\sigma)^{-1}(\tilde{s})} c_u^{(m)} c_v^{(m)} = 1. \]

But from Lemma 4, \( (8) \) holds if \( c^{(m)}, m = 1, \ldots, k - 1, \) is the characteristic function of \( D_m: c_u^{(m)} = 1 \) if \( u \in D_m, c_u^{(m)} = 0 \) otherwise. **■**

The next proof is somewhat complicated; it might help the reader to work through it in the case \( p = 9, r = 5 \) (this was the case that suggested the general result).

**Proof of Lemma 4:** Recalling that an element \( u \in E_{k,q} \) is a binary string \( u_1 u_2 \cdots u_k \), we define

\[ D_1 = \{ u \in E_{k,q} \mid u_1 = 0 \}, \]
\[ D_2 = \{ u \in E_{k,q} \mid u_1 = 1, u_k = 0 \}, \]
\[ D_3 = \{ u \in E_{k,q} \mid u_1 = u_k = 1, u_2 = 0 \}, \]
\[ D_4 = \{ u \in E_{k,q} \mid u_1 = u_k = u_2 = 1, u_{k-1} = 0 \}, \]

etc.,

4
and in general, for \( j \geq 0 \),

\[
D_{2j+1} = \{ u \in E_{k,q} \mid u_1 = u_2 = \cdots = u_j = u_k = u_{k-1} = \cdots = u_{k-j+1} = 1, \, u_{j+1} = 0 \},
\]

\[
D_{2j+2} = \{ u \in E_{k,q} \mid u_1 = u_2 = \cdots = u_{j+1} = u_k = u_{k-1} = \cdots = u_{k-j+1} = 1, \, u_{k-j} = 0 \}.
\]

It is clear that the \( D_m \) so defined form a partition of \( E_{k,q} \). We will write \( \tilde{D}_m = \sigma(D_m \times D_m) \), so that we must prove that for any \( \tilde{s} \in \tilde{E}_{p,r} \), \( |\tilde{s} \cap \bigcup_{m=1}^{k-1} \tilde{D}_m| = 1 \).

Note that a string \( u \in E_{k,q} \) contains exactly two zeros, and if \( u \in D_m \) then the position of one of these zeros is fixed and there are \( k - m \) possible positions for the remaining one; thus \( |D_m| = k - m \). Note also that \( u, v \in D_m \) if and only if the form of \( \sigma(u, v) \) is

\[
\begin{aligned}
1^j w 1^j 1^j 1 1^j x 0 1^j, & \quad \text{if } m = 2j + 1, \, j \geq 0, \\
1^{j+1} w 1^j 1^j x 0 1^j 0 1^{j+1}, & \quad \text{if } m = 2j + 2, \, j \geq 0
\end{aligned}
\]  

(9a) (9b)

where \( w, x \in E_{k-m,k-m-1} \) are arbitrary.

Now fix \( \tilde{s} \in \tilde{E}_{p,r} \). There are nonnegative integers \( n_0, \ldots, n_3 \), with \( n_0 + n_1 + n_2 + n_3 = 2k - 3 \), such that \( \tilde{s} \) consists of all cyclic permutations of the string

\[
01^{n_0}01^{n_1}01^{n_2}01^{n_3}.
\]

(10)

We must show that precisely one element of \( \tilde{s} \) has one of the forms (9).

Consider first (9a); the initial \( 1^j 0 \) and final \( 0 1^j \) there imply that if that string is put in the form (10) by a cyclic permutation then it will contain a substring \( 01^2j 0 \), i.e., that if an element in \( \tilde{s} \) has the form (9a) then one of the integers \( n_i \) must be even. Conversely, if \( n_i \) is even for some \( i \), with \( n_i = 2j_i \) \((j_i \geq 0)\), then the string \( s_i \in \tilde{s} \) defined by

\[
s_i = 1^{j_i} 0 1^{n_i+1} 0 1^{n_i+2} 0 1^{n_i+3} 0 1^{j_i}
\]  

(11a)

(here addition on the indices of the \( n_i \)'s is taken modulo 4) will lie in \( \tilde{D}_{n_i+1} \) if \( n_{i+1} \) and \( n_{i-1} \) satisfy certain additional constraints, which we discuss below. The discussion of (9b) is similar: if some \( n_i \) is odd, \( n_i = 2j_i + 1 \) \((j_i \geq 0)\), then the cyclic permutation of (10) in which the block \( 1^{n_i} \) is moved to the center is a candidate to lie in \( \tilde{D}_{n_i+1} \). If \( j_i + n_i - 1 + 2 \leq k \) and \( j_i + n_i + 1 + 2 \leq k \) (the only case that will be relevant, since (9b) has two zeros on each side of its center) then this string has the form

\[
s_i = 1^{k-(j_i+n_i+1+2)} 0 1^{n_i-1} 0 1^{j_i} 1 1^{j_i} 0 1^{n_i+1} 0 1^{k-(j_i+n_i+1+2)},
\]

(11b)

and will lie in \( \tilde{D}_{n_i+1} \) under further constraints on \( n_{i\pm1} \). We see that for each \( i, i = 0, 1, 2, 3 \), there is one possible element of \( \tilde{s} \) which could lie in \( D_{n_i+1} \), given by (11a) or (11b) as \( n_i \) is even or odd.

Now we ask what further conditions on \( n_{i\pm1} \) would imply that (11a) has the form (9a) or (11b) the form (9b). Consider first (11a), and recall that here \( n_i = 2j_i \). The second zero in (11a) is located at position \( j_i + n_{i+1} + 2 \), and for (11a) to have the form (9a) it is necessary that this zero lie to the left of a block \( 1^{j_i} 1 1^{j_i} \) at the center of the string, that is, to the left of position \( k - j_i + 1 \). Thus \( s_i \in \tilde{D}_{n_i+1} \) is possible only if \( j_i + n_{i+1} + 2 < k - j_i + 1 \),
i.e., only if \( n_i + n_i + 1 \leq k - 2 \). Combining this result with that of a similar analysis of the position of the third zero shows that

\[
s_i \in \bar{D}_{n_i+1} \text{ if and only if } n_i + n_i + 1 \leq k - 2 \text{ and } n_i + n_i - 1 \leq k - 2.
\]

(12)

The analysis of (11b), where \( n_i = 2j_i + 1 \), is similar: for this to have the form (9b), there must be at least \( j_i + 1 \) initial ones in the string, requiring that \( k - (j_i + n_i + 1 + 2) \geq j_i + 1 \); since there must also be \( j_i + 1 \) ones at the end of the string we are led again to the conclusion (12).

Finally we observe that the condition that \( \sum_{i=0}^{3} n_i = 2k - 3 \) implies that of any pair of inequalities \( n_i + n_i + 1 \leq k - 2 \) and \( n_{i+2} + n_{i+3} \leq k - 2 \) exactly one must be true. This implies that the condition of (12) will be satisfied for exactly one value of \( i \) (modulo 4), so that \( s_i \in \bar{D}_{n_i+1} \) (that is, \( s \cap \bar{D}_{n_i+1} = \{s_i\} \)) holds for precisely one value of \( i \). From (11a) or (11b) one can then read off the unique \( u, v \in D_{n_i+1} \) such that \( \sigma(u, v) = s_i \).

3. Negative results

In this section we show that Condition H does not hold in the following cases:

Case 1: \( 5 \leq r \leq p - 6; \ p \geq 13, \ r = 3; \) and \( p = 9, \ r = 3 \).

Case 2: \( 3 \leq r \leq p - 3 \).

The method of proof in all of these cases is similar to the argument of [3] establishing a negative result for \( p = 6, \ r = 3 \).

Throughout the rest of this section we assume that we are in case 1 or case 2, that is, \( r = 2q + 1 \) is odd, but to the extent possible we treat these two cases in a unified manner, so that for the moment either \( p = 2k + 1 \) or \( p = 2k + 2 \). If \( u, v \in E_{k,q} \) we write \( \bar{N}(u, v) = |\pi(\sigma(u, v))| \) and \( N(u, v) = |(\pi \sigma)^{-1}(\pi(\sigma(u, v)))| \); that is, \( \bar{N}(u, v) \) is the number of distinct strings obtained from \( \sigma(u, v) \) by cyclic permutation, and \( N(u, v) \) is the number of ordered pairs \( (w, x) \in E_{k,q} \times E_{k,q} \) such that \( \sigma(w, x) \) is obtained from \( \sigma(u, v) \) by a cyclic permutation. We will compute \( N(u, v) \) using the following simple remark. 

Remark 5: Let \( k' = p - k - 1 \) so that \( k' = k \) in case 1, \( k' = k + 1 \) in case 2. Then for any \( s \in E_{p,r} \) with \( |\pi(s)| = p, |(\pi \sigma)^{-1}(\pi(s))| \) is equal to the number of indices \( i, 1 \leq i \leq p, \) such that (i) \( s_i = 1 \) and the preceding (if \( i \geq k' + 1 \)) or succeeding (if \( i \leq p - k' \)) \( k' \) entries of \( s \)—that is \( s_{i-k'} \cdots s_{i-1} \) or \( s_{i+1} \cdots s_{i+k'} \), respectively—contain exactly \( q \) ones, and (ii) in case 2, if also \( i - k' = 0 \) or \( i + k' = 0 \), respectively. Of course if \( s = \sigma(u, v) \) then \( i = k + 1 \) satisfies this criterion. The application of this remark in any particular case is straightforward but tedious; we give a full discussion of one case in the proof of Lemma 6 and after that we are rather sketchy, leaving the details to the reader. It is probably most helpful to work out a simple example in each case.

We now define \( w = 0^{k-q} 1^q \in E_{k,q} \).

Lemma 6: Suppose that \( u \in E_{k,q} \). Then (a) \( \bar{N}(w, u) = p \), and (b) if \( u_1 = 0 \) or \( p \) is even (i.e., we are in case 2) then \( N(w, u) = 1 \). In particular, (c) \( \bar{N}(w, w) = p \) and \( N(w, w) = 1 \).
Proof: (a) The string $\sigma(w, u)$ contains a substring of at least $q + 1$ consecutive ones, and since there are a total of $2q + 1$ ones in the string, no nontrivial cyclic permutation of $\sigma(w, u)$ can coincide with it.

(b) Under either hypothesis, $s \equiv \sigma(w, u)$ has the form $s = 0^{k-q} 1^q 1 s_{k+2} \cdots s_{p-1} 0$; the key observation is that for $1 \leq j \leq q + 1$ the last $j$ entries of $s$ can contain at most $j - 1$ ones, and so entries $k + 2, \ldots, p - j$ must contain at least $q - j + 1$ ones. We show that no index $i$, $1 \leq i \leq p$, other than $i = k + 1$, can satisfy criterion (i) of Remark 5. Suppose then that $s_i = 1$ and $i \neq k + 1$. There are three possible cases: if $k - q + 1 \leq i \leq k$ then $s_{i+1} \cdots s_{i+k'} = 1^{k-i} 1 s_{k+2} \cdots s_{p-(k+1-i)}$ contains, by the observation above, at least $(k - i) + 1 + (q - k + i) = q + 1$ ones; if $k + 2 \leq i \leq p - q - 1$ then $s_{i-k'} \cdots s_{i-1}$ contains the substring $s_{k-q+1} \cdots s_{k+1} = 1^{q+1}$; and if $p - q \leq i \leq p$ then $s_{i-k'} \cdots s_{i-1} = 1^{p-i} 1 s_{k+2} \cdots s_{i-1}$ contains at least $(p - i) + 1 + (q - p + i) = q + 1$ ones.

(c) This is an immediate consequence of (a) and (b). ■

Lemma 7: Suppose there exist $x, y, z \in E_{k,q}$, all distinct from $w$ and with $x \neq y$ and $x \neq z$, such that

\begin{align}
N(w, x) = N(w, y) = N(x, x) &= 1, \\
\tilde{N}(w, x) = \tilde{N}(w, y) = \tilde{N}(x, x) &= p, \\
N(z, z) &= 3, \text{ with } \pi(\sigma(z, z)) = \{\sigma(z, z), \sigma(x, y), \sigma(y, x)\}. \tag{13c}
\end{align}

Then Condition H does not hold.

We remark that the requirement that all of $x, y, z$ and $w$ be distinct, except for the possibility that $y = z$, actually follows from (13) and Lemma 6.

Proof: We suppose that for some $M$ and $c^{(m)}$, (6) holds for all $\tilde{s}$, and derive a contradiction. From (6) applied to $\pi(\sigma(w, w))$, $\pi(\sigma(x, x))$, and $\pi(\sigma(w, x))$ we have, using Lemma 6(c) and (13a)–(13b),

\begin{equation}
p = \sum_m c_w^{(m)} c_w^{(m)} = \sum_m c_x^{(m)} c_x^{(m)} = \sum_m c_w^{(m)} c_x^{(m)}.
\end{equation}

These equations, together with the standard necessary condition for equality to hold in the Cauchy-Schwarz inequality, then imply that

\begin{equation}
c_w^{(m)} = c_x^{(m)}, \quad \text{for } m = 1, \ldots, M.
\end{equation}

But, first from (6) applied to $\pi(\sigma(w, y))$, and then from (15),

\begin{equation}
p = \sum_m c_w^{(m)} c_y^{(m)} = \sum_m c_x^{(m)} c_y^{(m)}.
\end{equation}

Finally, from (6) applied to $\pi(\sigma(z, z))$, (13c), and then (16),

\begin{equation}
\tilde{N}(z, z) = \sum_m c_z^{(m)} c_z^{(m)} + \sum_m c_x^{(m)} c_z^{(m)} + \sum_m c_y^{(m)} c_{z}^{(m)} = \sum_m c_z^{(m)} c_z^{(m)} + 2p \geq 2p,
\end{equation}

7
a contradiction, since $\tilde{N}(z, z)$ must divide $p$. ■

**Theorem 8:** If $r$ is odd and (a) $p$ is odd and $5 \leq r \leq p - 6$, (b) $p$ is odd, $p \geq 13$, and $r = 3$, or (c) $p$ is even and $3 \leq r \leq p - 3$, then Condition H does not hold.

**Proof:** (a) In this case we claim that the strings

$$
x = 010^{k-q-1}1^{q-1}, \quad y = 0^{k-q-2}1^q0^2, \quad \text{and} \quad z = 01^q0^{k-q-1},
$$

fulfill the conditions of Lemma 7. Since $2 \leq q \leq k - 3$ we have $x \neq y$ and $x \neq z$ (although $y = z$ if $q = k - 3$). The conditions

$$
N(w, x) = N(w, y) = 1, \quad \tilde{N}(w, x) = \tilde{N}(w, y) = p,
$$

follow from Lemma 6, since $x_1 = y_1 = 0$.

Consider now $\sigma(x, x) = 010^{k-q-1}1^{q-1}11^q10^{k-q-1}10$; this contains a unique string of $2q - 1 \geq 3$ consecutive ones and so can never coincide with a cyclic permutation of itself, so that indeed $\tilde{N}(x, x) = p$. A detailed analysis using Remark 5, as in the proof of Lemma 6(b) (but by symmetry it is necessary to consider only $i \leq k$), shows that $\tilde{N}(x, x) = 1$.

Finally consider $s \equiv \sigma(z, z) = 01^q0^{k-q-1}10^{k-q-1}10$. Again, consideration of the sizes of the three blocks of consecutive ones shows that $\tilde{N}(z, z) = p$. To find $\tilde{N}(z, z)$ we note that a cyclic permutation which brings the one at position $i = 2$ of $s$ to the center position $i = k + 1$ yields that string $0^{k-q-2}1^q0^211^q10^{k-q-1}10 = \sigma(y, x)$, and one obtains $\sigma(x, y)$ by a cyclic permutation bringing the one at $i = p - 1$ in $s$ to $i = k + 1$. However, if $3 \leq i \leq q + 1$ then $s_{i+1} \cdots s_{i+k}$ contains at most $q - 1$ ones, with a similar conclusion if $p - q - 1 \leq i \leq p - 2$, so that $\tilde{N}(z, z) = 3$ and (13c) holds.

(b) In this case the strings

$$
x = 0^{k-3}10^2, \quad y = z = 010^{k-2},
$$

fulfill the conditions of Lemma 7; the verification is similar to the above.

(c) If $5 \leq r \leq p - 3$ then the strings

$$
x = 10^{k-q}1^{q-1}, \quad y = 0^{k-q-1}1^q0, \quad z = 1^q0^{k-q},
$$

fulfill the conditions of Lemma 7; again the verification is similar to that of case (a). If $r = 3$ and $p \geq 8$ then the conclusion follows from the case $r = p - 3$ after the interchange of $A$ and $B$. Finally, the result for case $p = 6, r = 3$ was established in [3]. ■

The next result covers the one remaining negative result not included in Theorem 8. It is stated without proof in [5].

**Theorem 9:** If $p = 9$ and $r = 3$ then Condition H does not hold.

**Proof:** Again we suppose that there exist $c^{(m)}$, $m = 1, \ldots, M$, so that (6) holds for all $\tilde{s}$, and derive a contradiction by looking at a few specific choices of $\tilde{s}$, as given in Table 1; there we write $v_1 = 0001, v_2 = 0100$ (with $v_1, v_2 \in E_{4,1}$).
From (6) applied to $\tilde{s}_1$, $\tilde{s}_2$, and $\tilde{s}_3$, we have

\[ 9 = \sum_m c_{v_1}^{(m)} c_{v_1}^{(m)} = \sum_m c_{v_1}^{(m)} c_{v_2}^{(m)}; \quad 3 = \sum_m c_{v_2}^{(m)} c_{v_2}^{(m)}. \] (18)

These equations, however, are inconsistent with the Cauchy-Schwarz inequality. ■

4. Concluding remarks

In recent work [6] Klep and Schweighofer give a systematic algebraic language in which to discuss the method of [3]. They introduce the associative $\mathbb{R}$-algebra $\mathbb{R}\langle a, b \rangle$ with noncommuting generators $a$ and $b$ ($X$ and $Y$ in the notation of [6]), furnished with a natural involution $f \mapsto f^*$ obtained by reversing each word in the generators. They further define $\Sigma^2 \subset \mathbb{R}\langle a, b \rangle$ to be the cone of elements $f \in \mathbb{R}\langle a, b \rangle$ which may be written as sums of Hermitian squares, $f = \sum_i g_i^* g_i$, and $\Theta^2$ to be the cone of elements which are cyclically equivalent to elements of $\Sigma^2$, where two elements $f$ and $g$ are cyclically equivalent if their difference is a sum of commutators. It follows that if $f(a, b) \in \Theta^2$ and $a, b$ are nonnegative $n \times n$ matrices then $\text{Tr}(f(a, b)) \geq 0$, so in order to show that $a_{p,r} \geq 0$ it suffices to verify that $S_{p,r}(a^2, b^2) \in \Theta^2$, where $S_{p,r}(a^2, b^2) \in \mathbb{R}\langle a, b \rangle$ denotes the sum of all possible products of $r$ factors $b^2$ and $p-r$ factors $a^2$.

It is immediate that if hypothesis $H$ is satisfied for some $p, r$ falling under Case 1 or Case 2, or if $H$ is satisfied for $p, p-r$ with $p, r$ falling under Case 3, then $S_{p,r}(a^2, b^2) \in \Theta^2$; further, it follows from a result of [6] (Proposition 2.2) that the converse also holds. This means that the results of Sections 2 and 3 establish, for every $p, r$ with either $p$ or $r$ odd, whether or not $S_{p,r}(a^2, b^2) \in \Theta^2$. In particular, we can conclude that the approach of [3] (at least as formulated in [6]) when applied to such $p$ and $r$ cannot establish the BMV conjecture for any $p$ larger than 9.

Thus to make progress on the BMV conjecture using this approach one must consider cases in which both $p$ and $r$ are even. In this direction, Klep and Schweighofer show [6] that $S_{14,4}(a^2, b^2)$ and $S_{14,6}(a^2, b^2)$ belong to $\Theta^2$, which, together with results of [4] or by independent arguments given in [6], implies that the BMV conjecture is satisfied for $p = 13$ and indeed, by [4], for $p \leq 13$. Moreover, Burgdorf [2] has obtained a version of Theorem 3 strengthened to include $p, r$ even: she shows that $S_{p,4}(a^2, b^2) \in \Theta^2$ (and hence $S_{p,p-4}(a^2, b^2) \in \Theta^2$) for all $p \geq 4$.

| Name of $\tilde{s}$ | Typical $s \in \tilde{s}$ | $|\tilde{s}|$ | $(\pi\sigma)^{-1}(\tilde{s})$ |
|---------------------|--------------------------|----------|------------------|
| $\tilde{s}_1$       | 000111000                | 9        | $\{(v_1, v_1)\}$ |
| $\tilde{s}_2$       | 010010010                | 3        | $\{(v_2, v_2)\}$ |
| $\tilde{s}_3$       | 000110010                | 9        | $\{(v_1, v_2)\}$ |

Table 1
Acknowledgments

We thank I. Klep, M. Schweighofer, and S. Burgdorf for communicating their results to us in manuscript form, and C. Hillar and D. Hägele for helpful comments on a preliminary version of this paper.

References

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

[2] S. Burgdorf, private communication.

[3] D. Hägele, Proof of the cases $p \leq 7$ of the Lieb-Seiringer formulation of the Bessis-Moussa-Villani conjecture. *J. Stat. Phys.* **127**, 1167–1171 (2007).

[4] C. J. Hillar, Advances on the Bessis-Moussa-Villani trace conjecture. *Linear Alg. Appl.* **426**, 130–142 (2007).

[5] I. Klep (joint work with M. Schweighofer), Sums of Hermitian squares, Connes’ embedding problem and the BMV conjecture, *Mathematisches Forschungsinstitut Oberwolfach*, Report No. 14/2007: Reelle Algebraische Geometrie, pp. 27–30. March 11th–March 17th, 2007.

[6] I. Klep and M. Schweighofer, Sums of hermitian squares and the BMV conjecture. arXiv:0710.1074.

[7] E. H. Lieb and R. Seiringer, Equivalent forms of the Bessis-Moussa-Villani conjecture. *J. Stat. Phys.* **115**, 185–190 (2004).