ADVANCES ON THE BESSIS-MOUSSA-VILLANI TRACE CONJECTURE

CHRISTOPHER J. HILLAR

ABSTRACT. A long-standing conjecture asserts that the polynomial

$$p(t) = \mathrm{Tr}[(A + tB)^m]$$

has nonnegative coefficients whenever m is a positive integer and A and B are any two $n \times n$ positive semidefinite Hermitian matrices. The conjecture arises from a question raised by Bessis, Moussa, and Villani (1975) in connection with a problem in theoretical physics. Their conjecture, as shown recently by Lieb and Seiringer, is equivalent to the trace positivity statement above. In this paper, we derive a fundamental set of equations satisfied by A and B that minimize or maximize a coefficient of $p(t)$. Applied to the Bessis-Moussa-Villani (BMV) conjecture, these equations provide several reductions. In particular, we prove that it is enough to show that (1) it is true for infinitely many m, (2) a nonzero (matrix) coefficient of $(A + tB)^m$ always has at least one positive eigenvalue, or (3) the result holds for singular positive semidefinite matrices. Moreover, we prove that if the conjecture is false for some m, then it is false for all larger m. Finally, we outline a general program to settle the BMV conjecture that has had some recent success.

1. INTRODUCTION

In 1975, while studying partition functions of quantum mechanical systems, Bessis, Moussa, and Villania formulated a conjecture regarding a positivity property of traces of matrices [1]. If this property holds, explicit error bounds in a sequence of Padé approximants follow. Let A and B be $n \times n$ Hermitian matrices with B positive semidefinite, and let

$$\phi^{A,B}(t) = \mathrm{Tr}[\exp(A - tB)].$$

The original formulation of the conjecture asserts that the function $\phi^{A,B}$ is completely monotone; in other words, $\phi^{A,B}$ is the Laplace transform of a positive measure $\mu^{A,B}$ supported on $[0, \infty)$:

$$\mathrm{Tr}[\exp(A - tB)] = \int_0^\infty \exp(-tx) d\mu^{A,B}(x).$$

Equivalently, the derivatives of the function $f(t) = \phi^{A,B}(t)$ alternate signs:

$$(-1)^m f^{(m)}(t) \geq 0, \quad t > 0, \; m = 0, 1, 2, \ldots.$$

1991 *Mathematics Subject Classification.* 15A24, 15A45, 15A90, 33Cxx, 44A10, 47A50, 47N50, 49J40.

Key words and phrases. Bessis-Moussa-Villani (BMV) conjecture, positive definite matrices, trace inequality, Euler-Lagrange equations, words in two matrices.

Supported under a National Science Foundation Postdoctoral Research Fellowship.

Since its introduction in [1], many partial results and substantial computational experimentation have been given [2, 3, 4, 5, 7, 8, 9, 14, 15], all in favor of the conjecture's validity. However, despite much work, very little is known about the problem, and it has remained unresolved except in very special cases. Recently, Lieb and Seiringer in [13], and as previously communicated to us [8], have reformulated the conjecture of [1] as a question about the traces of certain sums of words in two positive definite matrices. In what follows, we shall use the standard convention that a *positive definite* (resp. *positive semidefinite*) matrix is one that is complex Hermitian and has positive eigenvalues (resp. nonnegative eigenvalues).

Conjecture 1.1 (Bessis-Moussa-Villani)**.** *The polynomial* $p(t) = \mathrm{Tr}\left[(A+tB)^m\right]$ *has all nonnegative coefficients whenever* A *and* B *are* $n \times n$ *positive semidefinite matrices.*

Remark 1.2. Although not immediately obvious, the polynomial $p(t)$ has all real coefficients (see Corollary 2.4).

The coefficient of t^k in $p(t)$ is the trace of $S_{m,k}(A,B)$, the sum of all words of length m in A and B, in which k B's appear (it has been called the k-th Hurwitz product of A and B). In [8], among other things, it was noted that, for $m < 6$, each constituent word in $S_{m,k}(A,B)$ has nonnegative trace. Thus, the above conjecture is valid for $m < 6$ and arbitrary positive integers n. It was also noted in [8] (see also [1]) that the conjecture is valid for arbitrary m and $n < 3$. Thus, the first case in which prior methods did not apply and the conjecture was in doubt, is $m = 6$ and $n = 3$. Even in this case, all coefficients, except $\mathrm{Tr}[S_{6,3}(A,B)]$, were known to be nonnegative (also as shown in [8]). It was only recently [9], using heavy computation, that this remaining coefficient was shown to be nonnegative.

Much of the subtlety of Conjecture 1.1 can be seen by the fact that $S_{m,k}(A,B)$ need not have all nonnegative eigenvalues, and in addition that some words within the $S_{m,k}(A,B)$ expression can have negative trace (see [8], where it is shown that $\mathrm{Tr}[ABABBA]$ can be negative).

Our advancement is the introduction of a fundamental pair of matrix equations satisfied by A and B that minimize or maximize a coefficient of $p(t)$. In what follows, we will be using the natural Euclidean norm on the set of complex $n \times n$ matrices:

$$\|A\| = \mathrm{Tr}[AA^*]^{1/2}.$$

(Here, C^* denotes the *conjugate transpose* of a complex matrix C). The precise statement of our main result is the following.

Theorem 1.3. *Let* $m > k > 0$ *be positive integers, and let* A *and* B *be positive semidefinite matrices of norm* 1 *that minimize (resp. maximize) the quantity* $\mathrm{Tr}[S_{m,k}(A,B)]$ *over all positive semidefinite matrices of norm* 1*. Then,* A *and* B *satisfy the following pair of equations:*

$$\begin{cases} AS_{m-1,k}(A,B) & = \; A^2\mathrm{Tr}[AS_{m-1,k}(A,B)] \\ BS_{m-1,k-1}(A,B) & = \; B^2\mathrm{Tr}[BS_{m-1,k-1}(A,B)]. \end{cases} \tag{1.1}$$

We call (1.1) the *Euler-Lagrange equations* for Conjecture 1.1. The name comes from the resemblance of our techniques to those of computing the first variation in the calculus of variations. We should remark that there have been other variational approaches to this problem [2, 3]; a review can be found in [15]. Although we are motivated by Conjecture 1.1, we discovered that these equations are also satisfied by a minimization (resp. maximization) over Hermitian matrices A and B of norm 1 (see Corollary 3.7), and it is natural to consider this more general situation. In this regard, we present the following application of the Euler-Lagrange equations.

Theorem 1.4. *If A and B are Hermitian matrices of norm 1 and $m > 1$, then*

$$|\mathrm{Tr}[S_{m,k}(A,B)]| \leq \binom{m}{k}.$$

Moreover, if $m > k > 0$, then equality holds only when $A = \pm B$, and if in addition $m > 2$, then A has precisely one nonzero eigenvalue.

Remark 1.5. When $m = 1$, this theorem fails to hold. For example, let A be the $n \times n$ diagonal matrix $A = \mathrm{diag}(n^{-1/2}, \ldots, n^{-1/2})$. Then $\|A\| = 1$, but $\mathrm{Tr}[S_{1,0}(A,B)] = \mathrm{Tr}[A] = n^{1/2} > 1$ for $n > 1$.

It is easy to see that this maximum is at least $\binom{m}{k}$, and using elementary considerations involving the Cauchy-Schwartz inequality, one can show that

$$|\mathrm{Tr}[S_{m,k}(A,B)]| \leq \|S_{m,k}(A,B)\| n^{1/2} \leq \binom{m}{k} \|A\|^{m-k} \|B\|^k n^{1/2} = \binom{m}{k} n^{1/2}.$$

However, we do not know if a dependency on the size of the matrices involved can be removed without appealing to equations (1.1).

As a strategy to prove Conjecture 1.1, we offer the following.

Conjecture 1.6. *Let $m > k > 0$ be positive integers. Positive semidefinite (resp. Hermitian) matrices A and B of norm 1 that satisfy the Euler-Lagrange equations commute.*

From this result, Conjecture 1.1 would be immediate. Of course, Theorem 1.4 implies that Conjecture 1.6 holds for the case of Hermitian maximizers and minimizers. We next list some of the major consequences of the equations found in Theorem 1.3. The first one implies that counterexamples to Conjecture 1.1 are closed upwards. The precise statement is given by the following.

Theorem 1.7. *Suppose that there exist integers M, K and $n \times n$ positive definite matrices A and B such that $\mathrm{Tr}[S_{M,K}(A,B)] < 0$. Then, for any $m \geq M$ and $k \geq K$ such that $m - k \geq M - K$, there exist $n \times n$ positive definite A and B making $\mathrm{Tr}[S_{m,k}(A,B)]$ negative.*

Corollary 1.8. *If the Bessis-Moussa-Villani conjecture is true for some m_0, then it is also true for all $m < m_0$.*

Corollary 1.8 also reduces the BMV conjecture to its "asymptotic" formulation.

Corollary 1.9. *If the Bessis-Moussa-Villani conjecture is true for infinitely many m, then it is true for all m.*

Corollary 1.9 motivates a general program to solve the BMV conjecture, and there is evidence that this approach is more than a theoretical possibility. For instance, Hägele [6] has used this approach and Corollary 1.8 to prove the conjecture for all $m \leq 7$ (and all n). Inspired by Hägele's ideas, Klep and Schweighofer [12] used semidefinite programming techniques to prove the conjecture for all $m \leq 9$. It should be noted that these techniques provably fail for the difficult $m = 6$ case, making the appeal to Corollary 1.8 fundamental.

A next result characterizes the BMV conjecture in terms of the eigenvalues of the matrix $S_{m,k}(A,B)$.

Theorem 1.10. *Fix positive integers $m \geq k$ and n. Then,* $\mathrm{Tr}[S_{m,k}(A,B)] \geq 0$ *for all $n \times n$ positive semidefinite A and B if and only if whenever $S_{m,k}(A,B) \neq 0$, it has at least one positive eigenvalue.*

Remark 1.11. This theorem can be viewed as a transfer principle for the BMV conjecture: instead of proving positivity for the sum of all the eigenvalues, we need only show it for at least one of them. Thus, our original conjecture can be made to resemble a variant of Perron's Theorem for nonnegative matrices.

Conjecture 1.12. *Fix positive integers $m \geq k$ and n, and positive semidefinite $n \times n$ matrices A and B. Then $S_{m,k}(A,B)$ either has a positive eigenvalue or is the zero matrix.*

Our final result generalizes a fact first discovered in [9] (there only the real case was considered), and it implies that it is enough to prove the Bessis-Moussa-Villani conjecture for singular A and B.

Theorem 1.13. *Let m, n be positive integers, and suppose that* $\mathrm{Tr}\left[(A+tB)^{m-1}\right]$ *has nonnegative coefficients for each pair of $n \times n$ positive semidefinite matrices A and B. If $p(t) = \mathrm{Tr}\left[(A+tB)^{m}\right]$ has nonnegative coefficients whenever A, B are singular $n \times n$ positive semidefinite matrices, then $p(t)$ has nonnegative coefficients whenever A and B are arbitrary $n \times n$ positive semidefinite matrices.*

The organization of this paper is as follows. In Section 2, we recall some facts about Hurwitz products, and in Section 3 we derive the two equations found in Theorem 1.3. Finally, in Section 4, we use these equations to prove our main Theorems 1.4, 1.7, 1.10, and 1.13.

2. Preliminaries

We begin with a review of some basic facts involving Hurwitz products; some of this material can be found in [9]. The coefficients $S_{m,k}(A,B)$ may be generated via the recurrence:

$$S_{m,k}(A,B) = AS_{m-1,k}(A,B) + BS_{m-1,k-1}(A,B). \tag{2.1}$$

The following lemmas will be useful for computing the traces of the $S_{m,k}$.

Lemma 2.1. *Fix integers $m > k \geq 0$. For any two $n \times n$ matrices A and B, we have*

$$\mathrm{Tr}\left[S_{m,k}(A,B)\right] = \frac{m}{m-k}\mathrm{Tr}\left[AS_{m-1,k}(A,B)\right].$$

Proof. Consider the following chain of equalities:

$$\begin{aligned}
0 &= \mathrm{Tr}\left[\sum_{i=1}^{m} (A+tB)^{i-1}(A-A)(A+tB)^{m-i}\right] \\
&= \mathrm{Tr}\left[mA(A+tB)^{m-1}\right] - \mathrm{Tr}\left[\sum_{i=1}^{m} (A+tB)^{i-1} A (A+tB)^{m-i}\right] \\
&= \mathrm{Tr}\left[mA(A+tB)^{m-1}\right] - \mathrm{Tr}\left[\frac{d}{dy}(Ay+tB)^{m}\right]\Bigg|_{y=1} \\
&= \mathrm{Tr}\left[mA(A+tB)^{m-1}\right] - \frac{d}{dy}\left[\mathrm{Tr}\,(Ay+tB)^{m}\right]\Bigg|_{y=1}.
\end{aligned}$$

Since $S_{m,k}(Ay, B) = y^{m-k} S_{m,k}(A, B)$, it follows that the coefficient of t^k in the last expression above is just

$$m\mathrm{Tr}[AS_{m-1,k}(A,B)] - (m-k)\mathrm{Tr}[S_{m,k}(A,B)].$$

□

Lemma 2.2. *Fix integers $m \geq k > 0$. For any two $n \times n$ matrices A and B, we have*

$$\mathrm{Tr}\left[S_{m,k}(A,B)\right] = \frac{m}{k}\mathrm{Tr}\left[BS_{m-1,k-1}(A,B)\right].$$

Proof. Follows from Lemma 2.1 by taking the trace of both sides of equation (2.1). □

Let A and B be $n \times n$ Hermitian matrices. Since $S_{m,k}(A,B)$ is the sum of all words of length m in A and B with k B's, it follows that the conjugate transpose of $S_{m,k}(A,B)$ simply permutes its constituent summands. This verifies the following fact.

Lemma 2.3. *If A and B are $n \times n$ Hermitian matrices, then the matrix $S_{m,k}(A,B)$ is Hermitian.*

Corollary 2.4. *The polynomial $p(t) = \mathrm{Tr}\left[(A+tB)^m\right]$ has all real coefficients whenever A and B are $n \times n$ Hermitian matrices.*

Although $S_{m,k}(A,B)$ is Hermitian for Hermitian A and B, it need not be positive definite even when A and B are $n \times n$ positive definite matrices, $n > 2$. Examples are easily generated, and computational experiments suggest that it is usually not positive definite.

Finally, we record a useful fact about positive definite congruence.

Lemma 2.5. *Let C be any complex $n \times n$ matrix and let A be an $n \times n$ positive semidefinite matrix. Then CAC^* is positive semidefinite.*

Proof. See [10, p. 399]. □

3. Derivation of the Euler-Lagrange Equations

The arguments for our main theorems are based on a variational observation. It says that an expression $\mathrm{Tr}[S_{m,k}(A,B)]$ is minimized or maximized when A and B satisfy the Euler-Lagrange equations (see Corollary 3.6). Before presenting a proof of this fact, we give a series of technical preliminaries.

Proposition 3.1. *Let $m > k > 0$ be positive integers. Let B be any given Hermitian $n \times n$ matrix, and suppose that A is a positive semidefinite matrix of norm 1 that minimizes (resp. maximizes)*

$$\mathrm{Tr}[S_{m,k}(A, B)]$$

over all positive semidefinite matrices of norm 1. Let $\varepsilon > 0$, and let $C := C(x) = (c_{rs}(x))$ be an $n \times n$ matrix with entries $c_{rs}(x) = u_{rs}(x) + iv_{rs}(x)$ in which u_{rs} and v_{rs} are differentiable functions $u_{rs}, v_{rs} : [-\varepsilon, \varepsilon] \to \mathbb{R}$. Moreover, suppose that $C(0) = I$ and $CAC^* \neq 0$ for all $x \in [-\varepsilon, \varepsilon]$. Then

$$\mathrm{Tr}\left[\frac{d}{dx}\left(\frac{CAC^*}{\|CAC^*\|}\right) S_{m-1,k}(A, B)\right]\Bigg|_{x=0} = 0.$$

Proof. Let A, B, and C be as in the statement of the theorem. Keeping in mind Corollary 2.4, we may consider the differentiable function $f : [-\varepsilon, \varepsilon] \to \mathbb{R}$ given by

$$f(x) = \mathrm{Tr}\left[S_{m,k}\left(\frac{CAC^*}{\|CAC^*\|}, B\right)\right].$$

By hypothesis, the minimum (resp. maximum) of f is achieved at $x = 0$. Consequently, it follows that

$$\frac{df(x)}{dx}\bigg|_{x=0} = 0. \tag{3.1}$$

Next, notice that,

$$\frac{d}{dx}\mathrm{Tr}\left[\left(\frac{CAC^*}{\|CAC^*\|} + tB\right)^m\right] = \mathrm{Tr}\left[\frac{d}{dx}\left(\frac{CAC^*}{\|CAC^*\|} + tB\right)^m\right]$$
$$= \mathrm{Tr}\left[\sum_{i=1}^{m}\left(\frac{CAC^*}{\|CAC^*\|} + tB\right)^{i-1} \frac{d}{dx}\left(\frac{CAC^*}{\|CAC^*\|} + tB\right)\left(\frac{CAC^*}{\|CAC^*\|} + tB\right)^{m-i}\right].$$

When $x = 0$, the above expression evaluates to

$$\mathrm{Tr}\left[\sum_{i=1}^{m}(A + tB)^{i-1} \frac{d}{dx}\left(\frac{CAC^*}{\|CAC^*\|}\right)(A + tB)^{m-i}\right]\Bigg|_{x=0}$$
$$= \mathrm{Tr}\left[m\frac{d}{dx}\left(\frac{CAC^*}{\|CAC^*\|}\right)(A + tB)^{m-1}\right]\Bigg|_{x=0}$$

It follows, therefore, from (3.1) that

$$\mathrm{Tr}\left[\frac{d}{dx}\left(\frac{CAC^*}{\|CAC^*\|}\right) S_{m-1,k}(A, B)\right]\Bigg|_{x=0} = 0.$$

□

A corresponding statement can be made by fixing A and minimizing (resp. maximizing) over B.

Proposition 3.2. *Let $m > k > 0$ be positive integers. Let A be any given Hermitian $n \times n$ matrix, and let B be a positive semidefinite matrix of norm 1 that minimizes (resp. maximizes)*

$$\mathrm{Tr}[S_{m,k}(A, B)]$$

over all positive semidefinite matrices of norm 1. Let $\varepsilon > 0$, and let $C := C(x) = (c_{rs}(x))$ be an $n \times n$ matrix with entries $c_{rs}(x) = u_{rs}(x) + iv_{rs}(x)$ in which u_{rs}

and v_{rs} are differentiable functions $u_{rs}, v_{rs} : [-\varepsilon, \varepsilon] \to \mathbb{R}$. Moreover, suppose that $C(0) = I$ and $CBC^* \neq 0$ for all $x \in [-\varepsilon, \varepsilon]$. Then

$$\mathrm{Tr}\left[\frac{d}{dx}\left(\frac{CBC^*}{\|CBC^*\|}\right) S_{m-1,k-1}(A,B)\right]\bigg|_{x=0} = 0.$$

Proof. The proof is similar to that of Proposition 3.1, so we omit it. □

In our next lemma, we compute the derivative found in Propositions 3.1 and 3.2. For notational simplicity, the entry-wise derivative of the matrix C evaluated at the point $x = 0$ will be denoted by C'.

Lemma 3.3. *With the hypotheses as in Proposition 3.1, we have*

$$\frac{d}{dx}\left(\frac{CAC^*}{\|CAC^*\|}\right)\bigg|_{x=0} = C'A + AC'^* - \mathrm{Tr}[C'A^2]A - \overline{\mathrm{Tr}[C'A^2]}A. \tag{3.2}$$

Proof. A straightforward application of the product rule [11, p. 490] for (matrix) differentiation gives

$$\frac{d}{dx}\left(\frac{CAC^*}{\|CAC^*\|}\right) = \frac{d}{dx}\left(\frac{1}{\|CAC^*\|}\right) CAC^* + \frac{1}{\|CAC^*\|}\left(\frac{dC}{dx}AC^* + CA\frac{dC^*}{dx}\right).$$

Next, we compute

$$\begin{aligned}
\frac{d}{dx}\|CAC^*\|^{-1} &= -\|CAC^*\|^{-2}\frac{d}{dx}\|CAC^*\| \\
&= -\|CAC^*\|^{-2}\frac{d}{dx}\left(\mathrm{Tr}[CAC^*CAC^*]^{1/2}\right) \\
&= -(1/2)\|CAC^*\|^{-2}\,\mathrm{Tr}[CAC^*CAC^*]^{-1/2}\frac{d}{dx}\mathrm{Tr}[CAC^*CAC^*].
\end{aligned}$$

The product expansion of $\frac{d}{dx}\mathrm{Tr}[CAC^*CAC^*]$ occurring in this last line is:

$$\begin{aligned}
&\mathrm{Tr}\left[\frac{dC}{dx}AC^*CAC^* + CA\frac{dC^*}{dx}CAC^* + CAC^*\frac{dC}{dx}AC^* + CAC^*CA\frac{dC^*}{dx}\right] \\
&= 2\mathrm{Tr}\left[\frac{dC}{dx}AC^*CAC^*\right] + \overline{2\mathrm{Tr}\left[\frac{dC}{dx}AC^*CAC^*\right]}.
\end{aligned}$$

Finally, setting $x = 0$ and using the assumptions that $\|A\| = 1$ and $C(0) = I$, equation (3.2) follows. □

We now have enough to prove the main results of this section.

Theorem 3.4. *Let $m > k > 0$ be positive integers. Let B be any given Hermitian $n \times n$ matrix, and let A be a positive semidefinite matrix of norm 1 that minimizes (resp. maximizes)*

$$\mathrm{Tr}[S_{m,k}(A,B)]$$

over all positive semidefinite matrices of norm 1. Then

$$AS_{m-1,k}(A,B) = A^2\mathrm{Tr}[AS_{m-1,k}(A,B)]. \tag{3.3}$$

Proof. Let A and B be as in the hypotheses of the theorem. By using different matrices C in the statement of Proposition 3.1, we will produce a set of equations satisfied by the entries of $AS_{m-1,k}(A,B)$ that combine to make the single matrix equation (3.3). For ease of presentation, we introduce the following notation. For integers r, s, let E_{rs} denote the $n \times n$ matrix with all zero entries except for a 1 in the (r, s) entry.

Fix integers $1 \leq r, s \leq n$ and take $C = I + xE_{rs}$. Since C is invertible for all $x \in [-1/2, 1/2]$, it follows that $CAC^* \neq 0$ for all such x. Therefore, the hypotheses of Lemma 3.3 are satisfied. The formula there is

$$\frac{d}{dx}\left(\frac{CAC^*}{\|CAC^*\|}\right)\bigg|_{x=0} = C'A + AC'^* - \mathrm{Tr}[C'A^2]A - \overline{\mathrm{Tr}[C'A^2]}A,$$

in which C' is the entry-wise derivative of the matrix C evaluated at the point $x = 0$. Additionally, Proposition 3.1, along with a trace manipulation, tells us that

$$\begin{aligned}
\Big(\mathrm{Tr}[C'A^2] + \overline{\mathrm{Tr}[C'A^2]}\Big)&\mathrm{Tr}\left[AS_{m-1,k}(A,B)\right] \\
&= \mathrm{Tr}\left[C'AS_{m-1,k}(A,B) + AC'^*S_{m-1,k}(A,B)\right] \\
&= \mathrm{Tr}\left[C'AS_{m-1,k}(A,B)\right] + \mathrm{Tr}\left[S_{m-1,k}(A,B)AC'^*\right] \\
&= \mathrm{Tr}\left[C'AS_{m-1,k}(A,B)\right] + \overline{\mathrm{Tr}\left[C'AS_{m-1,k}(A,B)\right]}.
\end{aligned} \tag{3.4}$$

Since $C' = E_{rs}$, a computation shows that for any matrix N, the trace of $C'N$ is just the (s, r) entry of N. In particular, it follows from (3.4) that the (s, r) entries of $AS_{m-1,k}(A, B) + \overline{AS_{m-1,k}(A, B)}$ and $(A^2 + \overline{A^2})\mathrm{Tr}[AS_{m-1,k}(A, B)]$ coincide. We have therefore proved that

$$AS_{m-1,k}(A,B) + \overline{AS_{m-1,k}(A,B)} = (A^2 + \overline{A^2})\mathrm{Tr}[AS_{m-1,k}(A,B)]. \tag{3.5}$$

We next perform a similar examination using the matrices $C = I + ixE_{rs}$ to arrive at a second matrix identity. Combining equation (3.2) and Proposition 3.1 as before, we find that

$$AS_{m-1,k}(A,B) - \overline{AS_{m-1,k}(A,B)} = (A^2 - \overline{A^2})\mathrm{Tr}[AS_{m-1,k}(A,B)]. \tag{3.6}$$

The theorem now follows by adding these two equations and dividing both sides of the result by 2. □

Similar arguments using Proposition 3.2 in place of Proposition 3.1 produce the following results.

Theorem 3.5. *Let $m > k > 0$ be positive integers. Let A be any given Hermitian $n \times n$ matrix, and let B be a positive semidefinite matrix of norm 1 that minimizes (resp. maximizes)*

$$\mathrm{Tr}[S_{m,k}(A,B)]$$

over all positive semidefinite matrices of norm 1. Then

$$BS_{m-1,k-1}(A,B) = B^2\mathrm{Tr}[AS_{m-1,k-1}(A,B)].$$

Combining the statements of this section, we have finally derived the Euler-Lagrange equations (1.1) for Conjecture 1.1.

Corollary 3.6. *Let $m > k > 0$ be positive integers, and let A and B be positive semidefinite matrices of norm 1 that minimize (resp. maximize) the quantity $\mathrm{Tr}[S_{m,k}(A, B)]$ over all positive semidefinite matrices of norm 1. Then A and B satisfy the following pair of equations:*

$$\begin{cases} AS_{m-1,k}(A,B) &= A^2\mathrm{Tr}[AS_{m-1,k}(A,B)] \\ BS_{m-1,k-1}(A,B) &= B^2\mathrm{Tr}[BS_{m-1,k-1}(A,B)]. \end{cases}$$

Our proof generalizes to show that the same equations hold for Hermitian minimizers (resp. maximizers), or more generally, for classes of unit norm matrices with the same inertia. This result is the main ingredient in our proof of Theorem 1.4 concerning the maximum of $\mathrm{Tr}[S_{m,k}(A,B)]$.

Corollary 3.7. *Let $m > k > 0$ be positive integers, and let A and B be Hermitian matrices of norm 1 that minimize (resp. maximize) $\mathrm{Tr}[S_{m,k}(A,B)]$ over all Hermitian matrices of norm 1. Then A and B must satisfy the following pair of equations:*

$$\begin{cases} AS_{m-1,k}(A,B) & = \; A^2\mathrm{Tr}[AS_{m-1,k}(A,B)] \\ BS_{m-1,k-1}(A,B) & = \; B^2\mathrm{Tr}[BS_{m-1,k-1}(A,B)]. \end{cases}$$

In general, we conjecture that trace minimizers commute (Conjecture 1.6), a claim that would imply Conjecture 1.1. We close this section with one more application of the Euler-Lagrange equations.

Corollary 3.8. *Suppose that the minimum of $\mathrm{Tr}[S_{m,k}(A,B)]$ over the set of positive semidefinite matrices is zero, and let A and B be positive semidefinite matrices that achieve this minimum. Then, $S_{m,k}(A,B) = 0$.*

Proof. When $k = m$ or $k = 0$, the claim is clear. Therefore, suppose that $m > k > 0$. Let A and B be positive semidefinite matrices with $\mathrm{Tr}[S_{m,k}(A,B)] = 0$. If either A or B is zero then the corollary is trivial. Otherwise, consider

$$0 = \frac{\mathrm{Tr}[S_{m,k}(A,B)]}{\|A\|^{m-k}\|B\|^k} = \mathrm{Tr}[S_{m,k}(\widetilde{A},\widetilde{B})],$$

in which $\widetilde{A} = A/\|A\|$ and $\widetilde{B} = B/\|B\|$. Combining equations (1.1) with the assumptions, it follows that $\widetilde{A}S_{m-1,k}(\widetilde{A},\widetilde{B}) = 0$ and $\widetilde{B}S_{m-1,k-1}(\widetilde{A},\widetilde{B}) = 0$. Moreover, equation (2.1) implies that

$$S_{m,k}(\widetilde{A},\widetilde{B}) = \widetilde{A}S_{m-1,k}(\widetilde{A},\widetilde{B}) + \widetilde{B}S_{m-1,k-1}(\widetilde{A},\widetilde{B}) = 0.$$

Multiplying both sides of this identity by $\|A\|^{m-k}\|B\|^k$ completes the proof. □

4. Proofs of the Main Theorems

We first use the Euler-Lagrange equations to prove Theorem 1.4.

Proof of Theorem 1.4. Let $m > 1$ and n be positive integers. Since our arguments are the same in both cases, we consider determining the maximum of $\mathrm{Tr}[S_{m,k}(A,B)]$. Let M be the compact set of Hermitian matrices with norm 1 and choose $(A,B) \in M \times M$ that maximizes $\mathrm{Tr}[S_{m,k}(A,B)]$. If $k = 0$, then the desired inequality is of the form

$$\mathrm{Tr}[A^m] \leq \sum_{i=1}^{n} |\lambda_i|^m \leq \sum_{i=1}^{n} \lambda_i^2 = \|A\| = 1,$$

in which $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of A. A similar argument holds for $m = k$. Therefore, we assume below that $m > k > 0$.

The Euler-Lagrange equations from Corollary 3.7 imply that

$$AS_{m-1,k}(A,B) = A^2\mathrm{Tr}[AS_{m-1}(A,B)]. \tag{4.1}$$

Performing a uniform, unitary similarity, we may assume that A is diagonal of the form $A = \mathrm{diag}(\lambda_1, \ldots, \lambda_r, 0, \ldots, 0)$, in which $\lambda_1, \ldots, \lambda_r$ are nonzero. Let

$\widetilde{A} = \mathrm{diag}(\lambda_1^{-1}, \ldots, \lambda_r^{-1}, 0, \ldots, 0)$ be the pseudo-inverse of A, and set $D = \widetilde{A}A$. Multiplying both sides of (4.1) by $\widetilde{A}$, it follows that

$$DS_{m-1,k}(A,B) = A\mathrm{Tr}[AS_{m-1,k}(A,B)].$$

Taking the norm of both sides of this expression and applying Lemma 2.1, we have

$$\frac{m-k}{m}\mathrm{Tr}[S_{m,k}(A,B)] = \|DS_{m-1,k}(A,B)\| \leq \|S_{m-1,k}(A,B)\| \leq \binom{m-1}{k}. \tag{4.2}$$

It follows that $\binom{m}{k} \leq \mathrm{Tr}[S_{m,k}(A,B)] \leq \frac{m}{m-k}\binom{m-1}{k} = \binom{m}{k}$ as desired.

We next verify the final assertions in the statement of the theorem. From above, every inequality in the chain (4.2) is an equality. Thus, each term occurring in

$$\|S_{m-1,k}(A,B)\| = \sum_W \|W(A,B)\|,$$

a sum over length $m-1$ words W with k B's, takes the value 1. In particular, we have that $1 = \|A^{m-k-1}B^k\|$. When $m-1 > k > 1$, an application of Lemma 4.1 below completes the proof of the theorem. The remaining cases $k = 1$ or $m = k+1$ are dealt with as follows.

Without loss of generality, we may suppose that $k = 1$ (interchange the roles of the matrices A and B). Applying the Cauchy-Schwartz inequality, we obtain the following chain of inequalities:

$$1 = \mathrm{Tr}[A^{m-1}B]^2 = \left(\sum_{i=1}^n \lambda_i^{m-1} b_{ii}\right)^2 \leq \sum_{i=1}^n (\lambda_i^2)^{m-1} \sum_{i=1}^n b_{ii}^2 \leq \|A\|\|B\| = 1. \tag{4.3}$$

It follows that each inequality in (4.3) is an equality. In particular, the second-to-last identity says that B is diagonal. Moreover, equality in Cauchy-Schwartz implies that $\lambda_i^{m-1} = \delta b_{ii}$ for some real number δ and all i. Since $1 = \left|\sum_{i=1}^n \lambda_i^{m-1} b_{ii}\right| = \delta^2$, it follows that $A = \pm B$. If, in addition, $m > 2$ and A has more than 1 nonzero eigenvalue, then

$$1 = \sum_{i=1}^n \lambda_i^2 > \sum_{i=1}^n (\lambda_i^2)^{m-1} = \sum_{i=1}^n b_{ii}^2 = 1,$$

a contradiction. Therefore, the conclusions of the theorem hold for $k = 1$. □

Lemma 4.1. *Suppose that A and B are Hermitian matrices of norm 1 and $r > 0$ and $s > 1$ are integers such that $\|A^rB^s\| = 1$. Then, $A = \pm B$ has only 1 nonzero eigenvalue.*

Proof. Performing a uniform, unitary similarity, we may suppose that B is a diagonal matrix with entries less than or equal to 1 in absolute value. From the hypotheses, we have

$$1 = \|A^rB^s\| \leq \|A^r\|\|B^s\| \leq \|A\|^r\|B\|^s = 1.$$

Therefore, $\|B^s\| = 1 = \|B\|$, and since $s > 1$, this implies that B has a single nonzero eigenvalue. It follows that $\|A^rB^s\| = \|A^rB\| = 1$ is equal to the absolute value of the $(1,1)$ entry of A^r. Finally, since $\|A^r\| = 1$, the matrix A^r has only one nonzero entry, and therefore, A has only one nonzero eigenvalue. Thus, $A^r = \pm A$ and since $A^r = \pm B$, it follows that $A = \pm B$. □

The argument for our next result uses the following well-known fact; we provide a proof for completeness (see also Theorem 7.6.3 and Problem 9, p. 468 in [10]).

Lemma 4.2. *If P and Q are positive semidefinite matrices, then PQ has all nonnegative eigenvalues.*

Proof. Suppose first that P is positive definite. Then PQ is similar to

$$P^{-1/2}PQP^{1/2} = P^{1/2}QP^{1/2}.$$

In particular, PQ is similar to a positive semidefinite matrix by Lemma 2.5. Therefore, in this case PQ has all nonnegative eigenvalues. The general version of the claim now follows from continuity. $\square$

We are now prepared to present a proof that counterexamples to Conjecture 1.1 are closed upward. Theorem 1.10 closely follows.

Proof of Theorem 1.7. Suppose that Conjecture 1.1 is false for some m and k and let A and B be real positive semidefinite matrices of unit norm that minimize

$$\mathrm{Tr}[S_{m,k}(A,B)] = \frac{m}{m-k}\mathrm{Tr}[AS_{m-1,k}(A,B)] = \frac{m}{k}\mathrm{Tr}[BS_{m-1,k-1}(A,B)].$$

We show that for these same matrices A and B, we have $\mathrm{Tr}[S_{m+1,k}(A,B)] < 0$ and $\mathrm{Tr}[S_{m+1,k+1}(A,B)] < 0$.

Combining equation (2.1) and the identities (1.1) from Corollary 3.6, it follows that

$$S_{m,k}(A,B) = A^2\mathrm{Tr}[AS_{m-1,k}(A,B)] + B^2\mathrm{Tr}[BS_{m-1,k-1}(A,B)]. \tag{4.4}$$

This matrix is negative semidefinite since it is the sum of two such matrices. Hence, the product $AS_{m,k}(A,B)$ has all non-positive eigenvalues by Lemma 4.2. Thus, Lemma 2.1 implies that

$$\mathrm{Tr}[S_{m+1,k}(A,B)] = \frac{m}{m+1-k}\mathrm{Tr}[AS_{m,k}(A,B)] \leq 0. \tag{4.5}$$

In the case of equality, multiplying equation (4.4) on the left by A and taking the trace of both sides, it follows that

$$0 = \mathrm{Tr}[A^3]\mathrm{Tr}[AS_{m-1,k}(A,B)] + \mathrm{Tr}[AB^2]\mathrm{Tr}[BS_{m-1,k-1}(A,B)].$$

However, $\mathrm{Tr}[AB^2] \geq 0$ by Lemma 4.2 and since A is nonzero, we must have $\mathrm{Tr}[A^3] > 0$. This gives a contradiction to equality in (4.5). It follows that $\mathrm{Tr}[S_{m+1,k}(A,B)] < 0$ as desired.

In the same manner, we can also prove that

$$\mathrm{Tr}[S_{m+1,k+1}(A,B)] = \frac{m+1}{k+1}\mathrm{Tr}[BS_{m,k}(A,B)]$$

is negative. The conclusions of the theorem now follow immediately. $\square$

Proof of Theorem 1.10. We first prove the direction ($\Leftarrow$) using the contrapositive. Suppose that $\mathrm{Tr}[S_{m,k}(A,B)]$ can be made negative. The proof of Theorem 1.7 shows that there exist positive semidefinite matrices A and B such that $S_{m,k}(A,B)$ is negative semidefinite and $\mathrm{Tr}[S_{m,k}(A,B)] < 0$ (so that $S_{m,k}(A,B)$ is nonzero). It follows that the second implication in the statement of the theorem is false. The converse is clear. $\square$

Finally, we work out the proof of Theorem 1.13; the argument is similar in spirit to the proof of Theorem 1.4.

Proof of Theorem 1.13. Suppose we know that Conjecture 1.1 is true for the power $m-1$ and also suppose that for some k there exist $n \times n$ positive definite matrices A and B such that $\mathrm{Tr}[S_{m,k}(A,B)]$ is negative. Clearly, we must have $m > k > 0$. By homogeneity, there are positive definite A and B with norm 1 such that $\mathrm{Tr}[S_{m,k}(A,B)]$ is negative. Let M be the set of positive semidefinite matrices with norm 1 and choose $(A,B) \in M \times M$ that minimizes $\mathrm{Tr}[S_{m,k}(A,B)]$; our goal is to show that A and B must both be singular.

Suppose by way of contradiction that A is invertible. The Euler-Lagrange equations say that

$$AS_{m-1,k}(A,B) = A^2 \mathrm{Tr}[AS_{m-1,k}(A,B)].$$

Multiplying both sides of this equation by A^{-1} and taking the trace, it follows that

$$\mathrm{Tr}[S_{m-1,k}(A,B)] = \mathrm{Tr}[A]\mathrm{Tr}[AS_{m-1,k}(A,B)].$$

By hypothesis, $\mathrm{Tr}[S_{m-1,k}(A,B)]$ is nonnegative. Therefore, using Lemma 2.1, we have

$$\begin{aligned} \frac{m-k}{m} \mathrm{Tr}\left[S_{m,k}(A,B)\right] &= \mathrm{Tr}\left[AS_{m-1,k}(A,B)\right] \\ &= \frac{\mathrm{Tr}[S_{m-1,k}(A,B)]}{\mathrm{Tr}[A]} \\ &\geq 0, \end{aligned}$$

a contradiction ($\mathrm{Tr}[A]$ is nonzero since A is nonzero). It follows that A must be singular. A similar examination with B also shows that it must be singular.

Thus, if Conjecture 1.1 is true for singular A and B, it must be true for invertible A and B as well. This completes the proof of the theorem. □

5. Acknowledgments

We would like to thank Scott Armstrong for several interesting and useful remarks concerning a preliminary version of this manuscript.

References

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

[2] 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: Math. Gen. **13** (1980), 3147–3159.

[3] K. J. Le Couteur, *Some problems of statistical mechanics and exponential operators*, pp 209–235 in Proceedings of the International Conference and Winter School of Frontiers of Theoretical Physics, eds. F. C. Auluck, L.S. Kothari, V.S. Nanda, Indian National Academy, New Dehli, 1977, Published by the Mac Millan Company of India, 1978.

[4] M. Drmota, W. Schachermayer and J. Teichmann, *A hyper-geometric approach to the BMV-conjecture*, Monatshefte fur Mathematik **146** (2005), 179–201.

[5] M. Fannes and D. Petz, *Perturbation of Wigner matrices and a conjecture*, Proc. Amer. Math. Soc. **131** (2003), 1981–1988.

[6] D. Haegele, *Proof of the cases* $p \leq 7$ *of the Lieb-Seiringer formulation of the Bessis-Moussa-Villani conjecture*, math.FA/0702217.

[7] F. Hansen, *Trace functions as Laplace transforms*, J. Math. Phys., **47** 043504 (2006).

[8] C. Hillar and C. R. Johnson, *Eigenvalues of words in two positive definite letters*, SIAM J. Matrix Anal. Appl., **23** (2002), 916–928.

[9] C. Hillar and C. R. Johnson, *On the positivity of the coefficients of a certain polynomial defined by two positive definite matrices*, J. Stat. Phys., **118** (2005), 781–789.

[10] R. Horn and C. R. Johnson, *Matrix analysis*, Cambridge University Press, New York, 1985.

[11] R. Horn and C. R. Johnson, *Topics in matrix analysis*, Cambridge University Press, New York, 1991.

[12] I. Klep and M. Schweighofer, private communication, 2007.

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

[14] Nathan Miller, 3×3 *cases of the Bessis-Moussa-Villani conjecture*, Princeton University Senior Thesis, 2004.

[15] P. Moussa, *On the representation of* $\mathrm{Tr}\left(e^{A-\lambda B}\right)$ *as a Laplace transform*, Rev. Math. Phys. **12**, 621–655 (2000).

Department of Mathematics, Texas A&M University, College Station, TX 77843

E-mail address: chillar@math.tamu.edu