EXTENDING THE HÖLDER TYPE INEQUALITY OF BLAKLEY AND ROY TO NON-SYMMETRIC NON-SQUARE MATRICES

THOMAS H. PATE

ABSTRACT. Suppose m, n, and k are positive integers, and let $\langle \cdot, \cdot \rangle$ denote standard inner product on the spaces $\mathbb{R}^p$, $p > 0$. We show that if D is an $m \times n$ non-negative real matrix, and u and v are non-negative unit vectors in $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively, then

(1) $$\langle (DD^t)^k Du, v \rangle \geq \langle Du, v \rangle^{2k+1},$$

with equality if and only if $\langle (DD^t)^k Du, v \rangle = 0$, or there exists $\alpha > 0$ such that $Du = \alpha v$ and $D^t v = \alpha u$. This inequality extends to non-symmetric non-square matrices a 1965 result of Blakley and Roy which asserts that if D is a non-negative $n \times n$ symmetric matrix, and $u \in \mathbb{R}^n$ is a non-negative unit vector, then

(2) $$\langle D^k u, u \rangle \geq \langle Du, u \rangle^k,$$

with equality, when $k \geq 2$, if and only if $\langle D^k u, u \rangle = 0$, or there exists $\alpha > 0$ such that $Du = \alpha u$. The generality of the inequality (1) derives not only from the fact that D is not assumed to be symmetric or square, but from the fact that we admit two unit vectors u and v instead of the single unit vector u appearing in inequality (2) of Blakley and Roy. We apply our result to verify the conjecture of A. Sidorenko (1993) in the non-symmetric case provided that the underlying graph is a path.

1. INTRODUCTION AND MAIN RESULT

Our main result is a new inequality involving non-negative matrices. If $m, n \in \mathbb{N}$, the set of positive integers, then $\mathbb{R}_+^{m \times n}$ denotes the set of all $m \times n$ non-negative real matrices. We let $\mathbb{R}_+^n$, essentially $\mathbb{R}_+^{n \times 1}$, denote the set of all non-negative real vectors of length n. The notation $\mathbb{R}^n$ and $\mathbb{C}^n$ will have their usual meanings, and we equip these vector spaces with standard inner product, denoted by $\langle \cdot, \cdot \rangle$. Thus, if $x = (x_1, x_2, \ldots, x_n)^t$ and $y = (y_1, y_2, \ldots, y_n)^t$ are in $\mathbb{R}^n$, then $\langle x, y \rangle = \sum_{i=1}^n x_i y_i$, while if x and y are in $\mathbb{C}^n$, then $\langle x, y \rangle = \sum_{i=1}^n x_i \bar{y}_i$, where $A \mapsto A^t$ is the transpose map. We let $\mathbb{R}^{m \times n}$ and $\mathbb{C}^{m \times n}$ denote, respectively, the set of all $m \times n$ real matrices, and the set of all $m \times n$ complex matrices. The norm associated with $\langle \cdot, \cdot \rangle$, the 2-norm, is denoted by $\| \cdot \|$.

In the paper titled "A Hölder Type Inequality for Symmetric Matrices with Non-negative Entries", Blakley and Roy [1] proved the following.

Theorem 1. *Suppose n and ℓ are positive integers. If D is a symmetric member of $\mathbb{R}_+^{n\times n}$, then*

$$(3) \qquad \langle D^\ell u, u\rangle \;\geq\; \langle Du, u\rangle^\ell, \quad \forall u \in \mathbb{R}_+^n \quad such \; that \quad \|u\| = 1,$$

with equality, when $\ell \geq 2$, if and only if $\langle D^\ell u, u\rangle = 0$, or there exists $\alpha > 0$ such that $Du = \alpha u$.

We first extend Theorem 1 when ℓ is even, obtaining Theorem 2 below. This extension, which is by far the easier, depends upon Lemma 1, which is elementary, and probably well known. A proof is included for the sake of completeness. The extension in case ℓ is odd, which we consider to be our main result, and is presented in Theorem 4. In both the even and odd cases we eliminate the assumption that D is symmetric. Moreover, in the odd case we obtain an extension that is considerably stronger than would likely be expected. Necessary and sufficient conditions for equality are presented in both cases, and an application to the conjecture of Sidorenko [4] is given.

If $x = (x_1, x_2, \ldots, x_n)$ and $q = (q_1, q_2, \ldots, q_n)$ are non-negative sequences, $\sum_{j=1}^n q_j = 1$, and $\beta > 0$, then $\left(\sum_{i=1}^n q_i x_i^\beta\right)^{1/\beta}$, which we denote by $\mathfrak{M}_\beta(x, q)$, is known as a weighted power mean. According to [2, Theorem 96], the function $t \to \mathfrak{M}_t(x, q)$ is non-decreasing on $[1, \infty)$. In particular, $\mathfrak{M}_\beta(x, q) \geq \mathfrak{M}_1(x, q)$ when $\beta > 1$, with equality if and only if there exists a number c such that $x_i = c$ whenever $q_i \neq 0$.

Lemma 1. *Suppose $A \in \mathbb{C}^{n\times n}$, and A is both Hermitian and positive semidefinite. If x is a unit vector in $\mathbb{C}^n$, and $\alpha > 1$, then $\langle A^\alpha x, x\rangle \geq \langle Ax, x\rangle^\alpha$, with equality if and only if x is an eigenvector of A.*

Proof. Since A is Hermitian and positive semidefinite, there exists a unitary $U \in \mathbb{C}^{n\times n}$, and a non-negative diagonal matrix Λ, with diagonal sequence $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_n)$, such that $A = U^*\Lambda U$, where U^* denotes the conjugate transpose of U. Letting $y = (y_1, y_2, \ldots, y_n)^t$ denote Ux, and letting $p = (p_1, p_2, \ldots, p_n)$, where $p_i = |y_i|^2$ for each i, we have $\sum_{i=1}^n p_i = 1$. Hence, for any $\beta > 0$,

$$(4) \qquad \langle A^\beta x, x\rangle = \langle U^*\Lambda^\beta U x, x\rangle = \langle \Lambda^\beta y, y\rangle = \sum_{i=1}^n \lambda_i^\beta |y_i|^2 = \left(\mathfrak{M}_\beta(\lambda, p)\right)^\beta.$$

But $\alpha > 1$, so $\mathfrak{M}_\alpha(\lambda, p) \geq \mathfrak{M}_1(\lambda, p)$. Combining this with (4) we obtain

$$(5) \qquad \langle A^\alpha x, x\rangle = \left(\mathfrak{M}_\alpha(\lambda, p)\right)^\alpha \geq \left(\mathfrak{M}_1(\lambda, p)\right)^\alpha = \langle Ax, x\rangle^\alpha,$$

which is the desired inequality.

If x is a unit eigenvector of A, then $Ax = \gamma x$ for some $\gamma \geq 0$; hence, $\langle A^\alpha x, x\rangle = \gamma^\alpha = \langle Ax, x\rangle^\alpha$. Conversely, if $\langle A^\alpha x, x\rangle = \langle Ax, x\rangle^\alpha$, then $\mathfrak{M}_\alpha(\lambda, p) = \mathfrak{M}_1(\lambda, p)$ by (5), so there exists a number c such that $\lambda_i = c$ whenever $|y_i|^2 \neq 0$. But this is equivalent to the statement that y is an eigenvector of Λ, which is in turn equivalent to the statement that x is an eigenvector of A. The condition for equality is therefore correct as stated. $\qquad\square$

We make frequent use of the Cauchy-Schwartz inequality, particularly its associated necessary and sufficient condition for equality. If V is a complex vector space with inner product $\langle \cdot, \cdot\rangle$, then for all $x, y \in V$ we have $|\langle x, y\rangle|^2 \leq \langle x, x\rangle\langle y, y\rangle$, with equality if and only if there exists a number c such that $y = cx$ or $x = cy$. If equality results, and $x \neq 0$, then there exists c such that $y = cx$, and if $y \neq 0$,

then there exists c such that $x = cy$. Alternately, equality results if and only if the subspace $\text{Span}\{x, y\}$ has dimension less than 2.

Theorem 2. *Suppose $m, n \in \mathbb{N}$, and $A \in \mathbb{C}^{m \times n}$. If $u \in \mathbb{C}^n$, $v \in \mathbb{C}^m$, and $\|u\| = \|v\| = 1$, then*

$$\text{(6)} \qquad \langle (A^*A)^\alpha u, u \rangle \geq |\langle Au, v \rangle|^{2\alpha}, \quad \forall \alpha > 1,$$

with equality if and only if $Au = 0$, or there exists a non-zero $\zeta \in \mathbb{C}$ such that $Au = \zeta v$ and $A^*v = \bar{\zeta}u$.

Proof. Assuming that $\alpha > 1$, and that $u \in \mathbb{C}^n$ and $v \in \mathbb{C}^m$ are a unit, we apply Lemma 1 to the positive semidefinite Hermitian matrix A^*A to obtain

$$\text{(7)} \qquad \langle (A^*A)^\alpha u, u \rangle \geq \langle A^*Au, u \rangle^\alpha = \left(\langle v, v \rangle \langle Au, Au \rangle \right)^\alpha \geq |\langle Au, v \rangle|^{2\alpha},$$

where the second inequality in (7) is an instance of Cauchy-Schwartz. The inequality (6) is therefore true. In order for (6) to reduce to equality, both of the inequalities in (7) must be equalities. By Lemma 1 the first inequality reduces to equality if and only if there exists $\gamma \in \mathbb{R}$ such that $A^*Au = \gamma u$. The second inequality reduces to equality if and only if there exists $\zeta \in \mathbb{C}$ such that $Au = \zeta v$. If $\zeta = 0$, then $Au = 0$. If $\zeta \neq 0$, then $\gamma u = A^*Au = \zeta A^*v$; so $A^*v = (\gamma/\zeta)u$. But,

$$\zeta = \zeta \langle v, v \rangle = \langle Au, v \rangle = \langle u, A^*v \rangle = (\gamma/\bar{\zeta})\langle u, u \rangle = \gamma/\bar{\zeta},$$

so $\gamma = |\zeta|^2$ and $\gamma/\zeta = \bar{\zeta}$. Therefore, $Au = \zeta v$ and $A^*v = \bar{\zeta}u$. If $Au = 0$, then both sides of (6) reduce to zero, while if $Au = \zeta v$ and $A^*v = \bar{\zeta}u$ for some $\zeta \in \mathbb{C}$, then both sides of (6) reduce to $|\zeta|^{2\alpha}$. The stated conditions are therefore both necessary and sufficient for equality. $\qquad \square$

Note that if u and v are non-zero vectors in $\mathbb{C}^n$ and $\mathbb{C}^m$, respectively, then we may apply Theorem 2 to the vectors $\|u\|^{-1}u$ and $\|v\|^{-1}v$ to obtain the inequality

$$\text{(8)} \qquad \langle u, u \rangle^{\alpha-1} \langle v, v \rangle^\alpha \langle (A^*A)^\alpha u, u \rangle \geq |\langle Au, v \rangle|^{2\alpha},$$

where inequality results if and only if $Au = 0$ or if there exists $\zeta \in \mathbb{C}$ such that $Au = (\zeta\rho)v$ and $A^*v = (\zeta\rho^{-1})u$, where $\rho = \|u\|/\|v\|$. It is this inequality and the corresponding inequality that follows from Theorem 4 that we will apply to the conjecture of Sidorenko.

By specializing Theorem 2 to the real case we obtain the following extension of Theorem 1.

Theorem 3. *Suppose $m, n \in \mathbb{N}$, and $D \in \mathbb{R}^{m \times n}$. If $u \in \mathbb{R}^n$, $v \in \mathbb{R}^m$, and $\|u\| = \|v\| = 1$, then*

$$\text{(9)} \qquad \langle (D^tD)^\alpha u, u \rangle \geq (\langle Du, v \rangle)^{2\alpha}, \quad \forall \alpha > 1,$$

with equality if and only if $Du = 0$, or there exists $\zeta \in \mathbb{R}$ such that $Du = \zeta v$ and $D^tv = \zeta u$.

Note that Theorem 3 improves upon Blakley and Roy in several ways. Not only do we remove the assumption that D is symmetric and that u is non-negative, but there is the additional unit vector v, and the exponent α which need not in this case be a positive integer. One could say that the inequality of Blakley and Roy is more or less expected in the case when ℓ is even. It is the case when ℓ is odd that is surprising and significant.

The natural analogue to Theorem 4 when ℓ is even would be the inequality

$$(10) \qquad \langle (D^t D)^k u, v \rangle \geq (\langle Du, v \rangle)^{2k},$$

which should hold if $D \in \mathbb{R}_+^{n \times n}$, $u \in \mathbb{R}_+^n$ and $v \in \mathbb{R}_+^n$ are a unit, and $k \in \mathbb{N}$. Unfortunately, this inequality is false, as can be demonstrated with a simple example. Let

$$(11) \qquad D = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}, \quad u = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \text{and} \quad v = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

Then, u and v are non-negative unit vectors in $\mathbb{R}^2$, and $D \in \mathbb{R}_+^{2 \times 2}$. Moreover, it is easy to see that

$$(12) \qquad (D^t D)^k = \begin{bmatrix} 0 & 0 \\ 0 & 2^k \end{bmatrix}, \quad \text{so} \ \langle (D^t D)^k u, v \rangle = 0, \ \text{while} \ \langle Du, v \rangle = \frac{1}{\sqrt{2}}.$$

The inequality (10) therefore fails for all $k \in \mathbb{N}$.

Theorem 4. *If m, n, and k are in $\mathbb{N}$, and D is a member of $\mathbb{R}^{m \times n}$ with non-negative entries, then*

$$(13) \qquad \langle (DD^t)^k Du, v \rangle \geq \langle Du, v \rangle^{2k+1},$$

for all $u \in \mathbb{R}_+^n$ and $v \in \mathbb{R}_+^m$ such that $\langle u, u \rangle = \langle v, v \rangle = 1$, with equality if and only if $\langle (DD^t)^k Du, v \rangle = 0$, or there exists $\alpha > 0$ such that $Du = \alpha v$ and $D^t v = \alpha u$.

If u and v are arbitrary non-zero members of $\mathbb{R}_+^n$ and $\mathbb{R}_+^m$, respectively, then we may apply Theorem 4 to the unit vectors $\|u\|^{-1}u$ and $\|v\|^{-1}v$ to obtain the equivalent inequality

$$(14) \qquad \langle u, u \rangle^k \langle v, v \rangle^k \langle (DD^t)^k Du, v \rangle \geq \langle Du, v \rangle^{2k+1},$$

with equality if and only if $\langle (DD^t)^k Du, v \rangle = 0$, or there exists $\alpha > 0$ such that $Du = (\alpha \rho) v$ and $D^t v = (\alpha/\rho) u$, where $\rho = \|u\|/\|v\|$. The proof of Theorem 4 is included in Section 3.

Given $q \in \mathbb{N}$, we let $\mathbf{1}_q$ denote the vector of length q, each of whose entries is 1. Setting $u = \mathbf{1}_n$ and $v = \mathbf{1}_m$ in (14) and then in (8), we obtain the twin inequalities

$$(15) \qquad \begin{aligned} m^k n^k \langle (DD^t)^k D \mathbf{1}_n, \mathbf{1}_m \rangle &\geq \langle D\mathbf{1}_n, \mathbf{1}_m \rangle^{2k+1} \quad \text{and} \\ m^k n^{k-1} \langle (D^t D)^k \mathbf{1}_n, \mathbf{1}_n \rangle &\geq \langle D\mathbf{1}_n, \mathbf{1}_m \rangle^{2k}, \end{aligned}$$

where in both cases we assume that $D \in \mathbb{R}_+^{m \times n}$. According to Theorem 4 the condition for equality in the first of these inequalities is that $\langle (DD^t)^k D \mathbf{1}_n, \mathbf{1}_m \rangle = 0$, or there exists $\alpha > 0$ such that $D\mathbf{1}_n = \alpha \sqrt{n/m} \, \mathbf{1}_m$, and $D^t \mathbf{1}_m = \alpha \sqrt{m/n} \, \mathbf{1}_n$. But, if $D \neq 0$, then the former condition cannot be satisfied, for if $\langle (DD^t)^k D \mathbf{1}_n, \mathbf{1}_m \rangle$ is 0, then $(DD^t)^k D = 0$. So $(DD^t)^k DD^t$, which is the same as $(DD^t)^{k+1}$, must also be zero. But DD^t is positive semidefinite and symmetric, so $(DD^t)^{k+1} = 0$ implies that $DD^t = 0$, which implies that $D = 0$. Therefore the necessary and sufficient condition for equality in either of the inequalities in (16) is that $D = 0$, or there exists $\alpha > 0$ such that $D\mathbf{1}_n = \alpha \sqrt{n/m} \, \mathbf{1}_m$, and $D^t \mathbf{1}_m = \alpha \sqrt{m/n} \, \mathbf{1}_n$. Letting $\mathcal{S}(A)$ denote the sum of the entries of matrix A, we may restate (15) and its condition for equality as follows.

Theorem 5. *Suppose m, n, and k are positive integers. If $D \in \mathbb{R}_+^{m \times n}$, then*

$$(16) \qquad m^k n^k \mathcal{S}\big((DD^t)^k D\big) \geq \big(\mathcal{S}(D)\big)^{2k+1} \ \text{and} \ m^k n^{k-1} \mathcal{S}\big((D^t D)^k\big) \geq \big(\mathcal{S}(D)\big)^{2k},$$

with equality in either inequality if and only if each row sum of D is $\mathcal{S}(D)/m$, and each column sum of D is $\mathcal{S}(D)/n$.

Of course, it follows from Theorem 5 that if D is a non-negative $n \times n$ matrix whose entries sum to n, then equality results in (16) if and only if D is doubly stochastic. Thus, we have

Theorem 6. *Suppose n and k are positive integers and $D \in \mathbb{R}_+^{n \times n}$. If the sum of the entries of D is n, then*

$$(17) \qquad \mathcal{S}\big((DD^t)^k D\big) \geq n \quad \text{and} \quad \mathcal{S}\big((D^tD)^k\big) \geq n,$$

with equality in either inequality if and only if D is doubly stochastic.

2. The conjecture of Sidorenko

We now show how Theorem 4 applies to Sidorenko's Conjecture 1 [4]. Suppose $h(x,y)$ is bounded, non-negative, and measurable on $[0,1]^2$ with respect to Lebesgue measure which we denote by μ, and let G be a bipartite graph with edge set E whose first part has p vertices, and whose second part has q vertices. Sidorenko's Conjecture 1 asserts that

$$(18) \qquad \int \prod_{(i,j) \in E} h(x_i, y_j) \, d\mu^{p+q} \geq \left(\int h(x,y) d\mu^2 \right)^{|E|}.$$

Inequality (18), particularly the integral on the left side, is not easy to understand, so we present a simple example. Suppose the graph G has vertex set $\{1,2,3,4,5\}$ and edge set $\{(1,2),(3,2),(3,4),(5,4)\}$. If we let $\mathcal{A}$ denote $\{1,3,5\}$ and let $\mathcal{B}$ denote $\{2,4\}$, then G is bipartite with first part $\mathcal{A}$ and second part $\mathcal{B}$. In this case $p = 3$, $q = 2$, and $|E| = 4$; moreover, G is a simple path that zig-zags back and forth from $\mathcal{A}$ to $\mathcal{B}$. If we assume that h is continuous and non-negative on $[0,1]^2$, then the integral (18) transforms into

$$(19) \qquad \begin{aligned} \int_0^1 \int_0^1 & \int_0^1 \int_0^1 \int_0^1 h(x_1, x_2) \, h(x_3, x_2) \, h(x_3, x_4) \, h(x_5, x_4) \, dx_1 \, dx_2 \, dx_3 \, dx_4 \, dx_5 \\ & \geq \left(\int_0^1 \int_0^1 h(x,y) \, dx \, dy \right)^4, \end{aligned}$$

which, as we shall demonstrate, is a consequence of Theorem 4, our main result.

Theorem 1 implies the more general inequality (18) whenever h is symmetric on $[0,1]^2$ and G is a path. We will show that (18) is true if G is a path even when h is not symmetric. The inequality (19) of the above example is then a consequence of our results. Additional notation is required. If $p \in \mathbb{N}$, then I_p denotes $\{1,2,\dots,p\}$, and if $\mathcal{T}$ is a set, then $I(\mathcal{T},p)$ denotes the set of all functions from $\mathcal{T}$ to I_p. Let G be a bipartite graph with first part $\mathcal{A}$ of cardinality r, second part $\mathcal{B}$ of cardinality s, and edge set E, where $(i,j) \in E$ not only means that there is an edge from vertex i to vertex j, but that $i \in \mathcal{A}$ and $j \in \mathcal{B}$. Suppose $m,n \in \mathbb{N}$, and let h be a function from $I_m \times I_n$ to $[0,\infty)$; that is, $h \in \mathbb{R}_+^{m \times n}$. An implication of Sidorenko's conjecture is that

$$(20) \qquad K_{m,n}(G) \sum_{f \in I(\mathcal{A},m)} \sum_{g \in I(\mathcal{B},n)} \prod_{(p,q) \in E} h_{f(p),g(q)} \geq \left(\sum_{i=1}^m \sum_{j=1}^n h_{ij} \right)^{|E|},$$

where $K_{m,n}(G)$ is a coefficient that depends upon m, n, the graph G, and the manner in which Sidoreko's conjecture is converted to a discrete statement. In the cases that we consider, namely that G is a path and h not necessarily symmetric, the value of $K_{m,n}(G)$ will be apparent.

The direct application of Theorem 4 involves the assumption that $|E|$ is odd, so we set $|E| = 2k+1$. The case $|E|$ even is considered separately. By a path we mean what is sometimes called a simple path, so each vertex is visited exactly once with first vertex in $\mathcal{A}$, and, since $|E|$ is odd, last vertex in $\mathcal{B}$. This implies that $\mathcal{A}$ and $\mathcal{B}$ both have exactly $k+1$ elements. We index our vertices with the integers $1, 2, \ldots, 2k+2$, letting $\mathcal{A} = \{1, 3, 5, \ldots, 2k+1\}$ and $\mathcal{B} = \{2, 4, \ldots, 2k+2\}$. The graph G consists of the ordered pair $(1, 2)$ and all ordered pairs of the form $(i, i\pm1)$ where $i \in \{3, 5, \ldots, 2k+1\}$. A discretization of Conjecture 1 of [4] appropriate to this case is

$$(21) \qquad \sum_{f \in I(\mathcal{A},m)} \sum_{g \in I(\mathcal{B},n)} \prod_{(p,q) \in E} h_{f(p),g(q)} \frac{1}{m^{k+1}} \frac{1}{n^{k+1}} \geq \left(\sum_{i=1}^{m} \sum_{j=1}^{n} h_{ij} \frac{1}{m} \frac{1}{n} \right)^{|E|},$$

where the summations above approximate the integrals of Sidorenko. The factors $m^{-k-1}n^{-k-1}$ and $m^{-1}n^{-1}$ correspond to the expressions $d\mu^{2k+2}$ and $d\mu^2$. But, $|E| = 2k+1$, so (21) is the same as (20), provided that $K_{m,n}(G) = m^k n^k$. Imagining the odd integers written in a vertical column adjacent to a similar column containing the even integers, we see the graph G as a simple zig-zag pattern going back and forth from $\mathcal{A}$ to $\mathcal{B}$. Consider the term $\prod_{(p,q) \in E} h_{f(p),g(q)}$ appearing in (20). Since the edges of G are $(1, 2), (3, 2), (3, 4), (5, 4), (5, 6), \ldots, (2k+1, 2k), (2k+1, 2k+2)$, this product appears as

$$(22) \quad h_{f(1),g(2)} h_{f(3),g(2)} h_{f(3),g(4)} h_{f(5),g(4)} h_{f(5),g(6)} h_{f(7),g(6)} h_{f(7),g(8)}$$
$$\cdots h_{f(2k+1),g(2k)} h_{f(2k+1),g(2k+2)}.$$

If we substitute h^t for h in alternate positions in (22) we obtain

$$(23) \quad h_{f(1),g(2)} h^t_{g(2),f(3)} h_{f(3),g(4)} h^t_{g(4),f(5)} h_{f(5),g(6)} h^t_{g(6),f(7)} h_{f(7),g(8)}$$
$$\cdots h^t_{g(2k),f(2k+1)} h_{f(2k+1),g(2k+2)}.$$

The left side of (20), minus the factor $K_{m,n}(G)$ which is currently $m^k n^k$, is therefore equal to

$$(24) \quad \sum_{f(1)=1}^{m} \sum_{f(3)=1}^{m} \cdots \sum_{f(2k+1)=1}^{m} \sum_{g(2)=1}^{n} \sum_{g(4)=1}^{n}$$
$$\cdots \sum_{g(2k+2)=1}^{n} h_{f(1),g(2)} \prod_{i=1}^{k} h^t_{g(2i),f(2i+1)} h_{f(2i+1),g(2i+2)},$$

which is easily recognized to be the sum of the entries of the matrix $h(h^t h)^k = (hh^t)^k h$. Thus, if we define $\mathbf{1}_j$ as above for each j and make the substitution $h = D$, then (20) transforms into

$$(25) \qquad m^k n^k \langle (DD^t)^k D \, \mathbf{1}_n, \mathbf{1}_m \rangle \geq \langle D \, \mathbf{1}_n, \mathbf{1}_m \rangle^{2k+1},$$

which, since $\langle \mathbf{1}_q, \mathbf{1}_q \rangle = q$ for all $q \in \mathbb{N}$, is an instance of the first part of (15), and therefore a consequence of Theorem 4. The discrete version of the conjecture of Sidorenko is therefore true when G is a path of odd length.

If $|E| = 2k$, that is, $|E|$ even, then we will apply the second part of (15). In this case we let $\mathcal{A} = \{1, 3, \ldots, 2k+1\}$ as before, and $\mathcal{B} = \{2, 4, \ldots, 2k\}$. The edge set E is $\{(2i \pm 1, 2i) : i \in I_k\}$, so there are $2k+1$ vertices indexed by the members of I_{2k+1}, and an analysis similar to that presented in the odd case indicates that $K_{m,n}(G) = m^{k-1} n^k$. The term corresponding to (23) is therefore

$$(26) \quad h_{f(1),g(2)} h_{f(3),g(2)} h_{f(3),g(4)} h_{f(5),g(4)} h_{f(5),g(6)} h_{f(7),g(6)} h_{f(7),g(8)}$$
$$\cdots h_{f(2k-1),g(2k)} h_{f(2k+1),g(2k)},$$

which may be rewritten as

$$(27) \quad h_{f(1),g(2)} h^t_{g(2),f(3)} h_{f(3),g(4)} h^t_{g(4),f(5)} h_{f(5),g(6)} h^t_{g(6),f(7)} h_{f(7),g(8)}$$
$$\cdots h_{f(2k-1),g(2k)} h^t_{g(2k),f(2k+1)}.$$

Mindful of (27), we see that the left side of (20), if we omit the factor $m^{k-1} n^k$, transforms into

$$(28) \quad \sum_{f(1)=1}^{m} \sum_{f(3)=1}^{m} \cdots \sum_{f(2k+1)=1}^{m} \sum_{g(2)=1}^{n} \sum_{g(4)=1}^{n} \cdots \sum_{g(2k)=1}^{n} \prod_{i=1}^{k} h_{f(2i-1),g(2i)} h^t_{g(2i),f(2i+1)},$$

which is easily seen to be the sum of the entries of $(hh^t)^k$. That is, if we set $D = h$, then (20) is

$$(29) \quad m^{k-1} n^k \langle (DD^t)^k \mathbf{1}_m, \mathbf{1}_m \rangle \geq \langle D\mathbf{1}_n, \mathbf{1}_m \rangle^{2k},$$

which is also an instance of (15) with the roles of m and n reversed. Conjecture 1 of [4] is therefore true in its discrete form, provided that G is a path, even if h is not symmetric. The corresponding continuous version now follows via the application of appropriate limiting theorems.

According to [4, pg. 2] the inequality (18) is true for the disjoint union of graphs $G_1, G_2, \ldots, G_p$ if it is true for each of the graphs separately. This in conjunction with Theorem 4 leads to additional matrix inequalities that are, however, not as easy to state as (13).

3. Proof of Theorem 4

We begin with some notation and a technical lemma that will simplify the proof. By $\Omega_{m,n}$ we shall mean the set of all triples (D, u, v) such that $D \in \mathbb{R}^{m \times n}_+$, $u \in \mathbb{R}^n_+$, $v \in \mathbb{R}^m_+$, and $\langle u, u \rangle = \langle v, v \rangle = 1$. Fix $k \in \mathbb{N}$ and define the function $F_{m,n}$ on $\Omega_{m,n}$ by

$$(30) \quad F_{m,n}(D, u, v) = \langle (DD^t)^k Du, v \rangle - \langle Du, v \rangle^{2k+1}.$$

The problem is to show that $F_{m,n}(D, u, v) \geq 0$ for all $(D, u, v) \in \Omega_{m,n}$, and to verify that the stated conditions are necessary and sufficient for equality.

Lemma 2. *If m, n, and k are in $\mathbb{N}$, and $\phi : \Omega_{m,n} \to \Omega_{n,m}$ is defined by $\phi(D, u, v) = (D^t, v, u)$, then ϕ is a bijection satisfying $F_{n,m} \circ \phi = F_{m,n}$. Moreover, if $P_1 \in \mathbb{R}^{n \times n}$, and $P_2 \in \mathbb{R}^{m \times m}$ are permutation matrices, then $F_{m,n}(P_2 D P_1^t, P_1 u, P_2 v) = F_{m,n}(D, u, v)$ for all $(D, u, v) \in \Omega_{m,n}$.*

Proof. It is obvious that ϕ is a bijection between $\Omega_{m,n}$ and $\Omega_{n,m}$. Since $(DD^t)^t = DD^t$ we have

$$(31) \quad \langle (DD^t)^k Du, v \rangle = \langle u, D^t (DD^t)^k v \rangle = \langle u, (D^t D)^k D^t v \rangle = \langle (D^t D)^k D^t v, u \rangle.$$

Moreover, $\langle Du, v \rangle = \langle D^t v, u \rangle$. Therefore,

$$
\begin{aligned}
F_{m,n}(D, u, v) &= \langle (D^t D)^k D^t v, u \rangle - \langle D^t v, u \rangle^{2k+1} \\
&= F_{n,m}(D^t, v, u) = (F_{n,m} \circ \phi)(D, u, v),
\end{aligned}
\tag{32}
$$

so $F_{n,m} \circ \phi = F_{m,n}$. Let $\tilde{D} = P_2 D P_1^t$. Note that $\tilde{D}^t = P_1 D^t P_2^t$ and that $\tilde{D}\tilde{D}^t = P_2 D D^t P_2^t$, so $(\tilde{D}\tilde{D}^t)^k = P_2 (DD^t)^k P_2^t$, and $(\tilde{D}\tilde{D}^t)^k \tilde{D} = P_2 (DD^t)^k D P_1^t$. Thus,

$$
\begin{aligned}
F_{m,n}(\tilde{D}, P_1 u, P_2 v) &= \langle (\tilde{D}\tilde{D}^t)^k \tilde{D} P_1 u, P_2 v \rangle - \langle \tilde{D} P_1 u, P_2 v \rangle^{2k+1} \\
&= \langle P_2 (DD^t)^k D P_1^t P_1 u, P_2 v \rangle - \langle P_2 D P_1^t P_1 u, P_2 v \rangle^{2k+1} \\
&= \langle (DD^t)^k Du, v \rangle - \langle Du, v \rangle^{2k+1} = F_{m,n}(D, u, v),
\end{aligned}
\tag{33}
$$

so the second part of Lemma 2 is also true. $\qquad\square$

Lemma 3. *Suppose $m, n \in \mathbb{N}$, $D \in \mathbb{R}_+^{m \times n}$, $u \in \mathbb{R}^n$, $v \in \mathbb{R}^m$, and $\langle u, u \rangle = \langle v, v \rangle = 1$. If there exists $\gamma_1, \gamma_2 \in \mathbb{R}$ such that $Du = \gamma_1 v$ and $D^t v = \gamma_2 u$, then $\gamma_1 = \gamma_2$, and $F_{m,n}(D, u, v) = 0$. If we assume further that $u > 0$ and $v > 0$, then $\gamma_1 = \gamma_2 = \alpha$, where $\alpha \geq 0$, and α^2 is the spectral radius of $D^t D$.*

Proof. Suppose $D \in \mathbb{R}_+^{m \times n}$. If $\langle u, u \rangle = \langle v, v \rangle = 1$, where $u \in \mathbb{R}^n$ and $v \in \mathbb{R}^m$, and there exist real γ_1 and γ_2 such that $Du = \gamma_1 v$ and $D^t v = \gamma_2 u$, then $\langle Du, v \rangle = \gamma_1 \langle v, v \rangle = \gamma_1$, while $\langle Du, v \rangle = \langle u, D^t v \rangle = \gamma_2 \langle u, u \rangle = \gamma_2$; thus, $\gamma_1 = \gamma_2$. Moreover, letting γ denote γ_1 and γ_2, we have $D^t Du = \gamma^2 u$, so $(D^t D)^k u = \gamma^{2k} u$ and $(DD^t)^k Du = D(D^t D)^k u = \gamma^{2k} Du = \gamma^{2k+1} v$. Thus, $\langle (DD^t)^k Du, v \rangle = \gamma^{2k+1}$. But $\langle Du, v \rangle = \gamma$, so $\langle Du, v \rangle^{2k+1}$ is also γ^{2k+1}. Therefore, $F_{m,n}(D, u, v) = 0$.

Assume that $u > 0$ and $v > 0$ entrywise. If $D = 0$, then $\alpha = 0$, and $\alpha = 0 = \gamma$ in accordance with the lemma, so assume that $D \neq 0$. Since D is non-negative and non-zero, $D^t D$, which is positive semidefinite and symmetric, is also non-negative and non-zero. By the Perron-Frobenius theory [3, Theorem 8.3.1], and the Rayleigh-Ritz characterization of eigenvalues of Hermitian matrices [3, Theorem 4.2.2], we have

$$
\alpha^2 = \sup_{x \neq 0} \frac{\langle D^t Dx, x \rangle}{\langle x, x \rangle}.
\tag{34}
$$

Since $\langle D^t Dx, x \rangle \geq 0$ for all $x \in \mathbb{R}^n$, if α^2 were 0, then we would have $\langle D^t Dx, x \rangle = \langle Dx, Dx \rangle = 0$ for all $x \in \mathbb{R}^n$. But this immediately implies that $D = 0$. Therefore, $D \neq 0$ implies $\alpha > 0$.

Letting $\gamma = \gamma_1 = \gamma_2$ as above, we have $D^t Du = \gamma D^t v = \gamma^2 u$. Since $u \neq 0$, γ^2 is therefore an eigenvalue of the symmetric matrix $D^t D$, as is α^2. Let x be a non-zero member of $\mathbb{R}^n$ such that $D^t Dx = \alpha^2 x$. Since $D^t D$ has non-negative entries, we know that we may assume that x has non-negative entries. If $\gamma^2 \neq \alpha^2$, then, because $D^t D$ is symmetric, the eigenvectors u and x must be orthogonal with respect to $\langle \cdot, \cdot \rangle$. But, x is non-negative and non-zero, and u is strictly positive, so this is impossible. Therefore, $\gamma^2 = \alpha^2$, and $\gamma = \alpha$. $\qquad\square$

Lemma 4. *Suppose $m, n \in \mathbb{N}$, and $(D, u, v) \in \Omega_{m,n}$. If there exists $\rho > 0$ such that $Du = \rho v$, then either $D^t v = \rho u$ and $F_{m,n}(D, u, v) = 0$, or $D^t v \neq \rho u$ and $F_{m,n}(D, u, v) > 0$.*

Proof. If $Du = \rho v$ and $D^t v = \rho u$, then $F_{m,n}(D, u, v) = 0$ by Lemma 3. Assume that $Du = \rho v$, but $D^t v \neq \rho u$. Then, by Lemma 3, we cannot have $D^t v = \beta u$ for

any β. Thus $\mathrm{Span}\{D^t v, u\}$ must have dimension 2, and the following instance of the Cauchy-Schwartz inequality must in fact be strict inequality; that is, we have

$$(35) \quad \rho^2 = \langle Du, v \rangle^2 = \langle u, D^t v \rangle^2 < \langle u, u \rangle \langle D^t v, D^t v \rangle = \langle D^t v, D^t v \rangle = \langle DD^t v, v \rangle.$$

Therefore,

$$(36) \quad \langle Du, v \rangle^{2k+1} = \rho^{2k+1} < \rho \langle DD^t v, v \rangle^k \le \rho \langle (DD^t)^k v, v \rangle = \langle (DD^t)^k Du, v \rangle,$$

where we have obtained the third relation in (36) by applying Lemma 1 to DD^t. Thus, if $D^t v \ne \rho u$, then we must have $\langle Du, v \rangle^{2k+1} < \langle (DD^t)^k Du, v \rangle$.   □

As a special case of Theorem 4, suppose $u \in \mathbb{R}_+^n$, and $\|u\| = 1$. If $Du = 0$, then $F_{m,n}(D, u, v) = 0$ for any $v \in \mathbb{R}_+^m$. So, assume that $Du \ne 0$, and let $v = \rho^{-1} Du$, where $\rho = \|Du\|$. Then, $v \in \mathbb{R}_+^m$, $\|v\| = 1$, and $Du = \rho v$, so Lemma 4 tells us that either $D^t v = \rho u$, and $F_{m,n}(D, u, v) = 0$, or $D^t v \ne \rho u$, and $F_{m,n}(D, u, v) > 0$. Thus, given any unit vector $u \in \mathbb{R}_+^n$ there is always at least one unit vector $v \in \mathbb{R}_+^m$ such that (13) is true. Moreover, if $Du \ne 0$, and $v = \|Du\|^{-1} Du$, then we have strict inequality unless $D^t v = \|Du\| u$.

Our proof of Theorem 4 is inductive. As noted previously, k remains fixed; moreover, we assume that $D \ne 0$. If $D = 0$ there is nothing to prove. Let $\mathcal{Z}$ denote the set of all pairs $(m, n) \in \mathbb{N} \times \mathbb{N}$ such that Theorem 4 is true. We will show that $\mathcal{Z} = \mathbb{N} \times \mathbb{N}$. Lemma 2 implies that

$$(37) \qquad\qquad (m, n) \in \mathcal{Z} \quad \text{if and only if} \quad (n, m) \in \mathcal{Z}.$$

We begin by showing that $(m, n) \in \mathcal{Z}$ if either m or n is 1. Due to (37), it is sufficient to consider the case $n = 1$. When $n = 1$, D has a single non-negative non-zero column which we denote by x. Since $u \in \mathbb{R}_+^1$, and $\langle u, u \rangle = 1$, we must have $u = 1$. Moreover, $Du = x$, $D^t v = \langle x, v \rangle$, and $D^t Du = \langle x, x \rangle u$, so $(D^t D)^k u = \langle x, x \rangle^k u$, $D(D^t D)^k u = (DD^t)^k Du = \langle x, x \rangle^k x$, and $\langle (DD^t)^k Du, v \rangle = \langle x, v \rangle (\langle x, x \rangle)^k$. Since, $\langle Du, v \rangle = \langle x, v \rangle$, we have

$$(38) \qquad \langle (DD^t)^k Du, v \rangle - \langle Du, v \rangle^{2k+1} = \langle x, v \rangle \big(\langle x, x \rangle^k - \langle x, v \rangle^{2k} \big).$$

By the Cauchy-Schwartz inequality, $\langle x, v \rangle^2 \le \langle x, x \rangle \langle v, v \rangle = \langle x, x \rangle$; thus, $\langle x, x \rangle^k - \langle x, v \rangle^{2k} \ge 0$. Since x and v are non-negative, $\langle x, v \rangle$ is also non-negative; therefore, $\langle x, v \rangle \big(\langle x, x \rangle^k - \langle x, v \rangle^{2k} \big) \ge 0$. This proves that the inequality (13) is true when $n = 1$. It is clear from (38) that $\langle (DD^t)^k Du, v \rangle = \langle Du, v \rangle^{2k+1}$ if and only if $\langle x, v \rangle = 0$ or $\langle x, x \rangle^k - \langle x, v \rangle^{2k} = 0$. We noted above that $\langle (DD^t)^k Du, v \rangle = \langle x, v \rangle (\langle x, x \rangle)^k$; thus, $\langle x, v \rangle = 0$ if and only if $\langle (DD^t)^k Du, v \rangle = 0$. Moreover, $\langle x, x \rangle^k - \langle x, v \rangle^{2k} = 0$ if and only if $\langle x, x \rangle - \langle x, v \rangle^2 = 0$, which, due to the condition for equality in Cauchy-Schwartz, can only happen if $v = \beta x$ for some number β. Since $\langle v, v \rangle = 1$, we then have $1 = \langle v, v \rangle = \beta^2 \langle x, x \rangle$, which implies that $\beta = (\langle x, x \rangle)^{-1/2}$. Therefore, $D^t v = \langle x, v \rangle = \beta \langle x, x \rangle = (\langle x, x \rangle)^{1/2} u$, and $Du = x = \beta^{-1} v = (\langle x, x \rangle)^{1/2} v$. This proves that equality can occur in (38) only if $Du = \alpha v$ and $D^t v = \alpha u$, where $\alpha = (\langle x, x \rangle)^{1/2}$, or $\langle (DD^t)^k Du, v \rangle = 0$. Since it is obvious that equality happens under either of these two conditions, Theorem 4 is true when $n = 1$. At this point we know that $\mathcal{Z}$ contains all pairs (m, n), where $m = 1$ or $n = 1$.

We now assume that $(p, q) \in \mathcal{Z}$ if $p < m$ or if $p = m$ and $q \le n - 1$, and we show that $(m, n) \in \mathbb{Z}$. An important point is that for the pair (m, n) under consideration it is true that both $(m - 1, n)$ and $(m, n - 1)$ are in $\mathcal{Z}$. For each $p \in \mathbb{N}$ let $\mathcal{S}^p$ denote the unit sphere in $\mathbb{R}^n$; that is, $\mathcal{S}^p$ is the set of all $x \in \mathbb{R}^p$ such that $\|x\| = 1$, and let $\mathcal{W}_p$ denote $\mathcal{S}^p \cap \mathbb{R}_+^p$. We denote $\mathcal{W}_n \times \mathcal{W}_m$ by $\mathcal{W}_{m,n}$, and partition $\mathcal{W}_{m,n}$

THOMAS H. PATE

into $\mathcal{W}_{m,n,+}$, the set of all pairs $(x, y) \in \mathcal{W}_{m,n}$ such that both x and y have strictly positive entries, and $\mathcal{W}_{m,n,0}$, which is the set of all $(x, y) \in \mathcal{W}_{m,n}$ such that at least one of x and y has a zero entry. When there is no danger of confusion we will write $\mathcal{W}_+$ instead of $\mathcal{W}_{m,n,+}$, and $\mathcal{W}_0$ instead of $\mathcal{W}_{m,n,0}$.

Suppose $(u, v) \in \mathcal{W}_{m,n,0}$. This means that at least one of u and v has a zero entry. We choose permutation matrices $P_1 \in \mathbb{R}^{n \times n}$ and $P_2 \in \mathbb{R}^{m \times m}$ such that the zero entries of u are grouped together at the bottom of $P_1 u$, and the zero entries of v are grouped together at the bottom of $P_2 v$. Due to Lemma 2 we know that $F_{m,n}(D, u, v) = F_{m,n}(P_2 D P_1^t, P_1 u, P_2 v)$, so we replace D with $P_2 D P_1^t$, u with $P_1 u$, v with $P_2 v$, and rename our variables D, u, and v. At this point we know that either the last entry of u or the last entry of v is zero. We will assume that the last entry of u is 0. If instead each entry of u is positive, then the last entry of v is zero. In this case we apply Lemma 2, obtaining that $F_{m,n}(D, u, v) = F_{n,m}(D^t, v, u)$, and proceed with (D^t, v, u) and the pair (n, m), instead of (D, u, v) and the pair (m, n). We have noted previously that both $(m-1, n)$ and $(m, n-1)$ are in $\mathcal{Z}$. From this it follows from (37) that $(n, m-1) \in \mathcal{Z}$. If we are considering (D^t, v, u) our argument would use the fact that $(n, m-1) \in \mathcal{Z}$ to show that $(n, m) \in \mathcal{Z}$. We then employ (37) once more to conclude that $(m, n) \in \mathcal{Z}$. It was therefore permissible to assume that the last entry of u is 0. Let $u = (u_1, u_2, \ldots, u_n)^t$, where $u_n = 0$, and let $\hat{u}$ be the member of $\mathbb{R}_+^{n-1}$ obtained by deleting the last entry of u. Let $\hat{D}$ denote the $m \times (n-1)$-matrix obtained from D by deleting its last column, and let $\tilde{D}$ denote the matrix obtained from D by replacing each of the entries of the last column of D with 0. Then, $\hat{D} \in \mathbb{R}_+^{m \times (n-1)}$, and $\hat{u} \in \mathcal{W}_{n-1}$. Since $(m, n-1) \in \mathcal{Z}$, and $\langle Du, v \rangle = \langle \hat{D}\hat{u}, v \rangle$, we therefore have

$$(39) \qquad \langle Du, v \rangle^{2k+1} = \langle \hat{D}\hat{u}, v \rangle^{2k+1} \leq \langle (\hat{D}\hat{D}^t)^k \hat{D}\hat{u}, v \rangle.$$

Moreover, $0 \leq \tilde{D} \leq D$ entrywise, so $(\tilde{D}\tilde{D}^t)^k \tilde{D} \leq (DD^t)^k D$ entrywise. Therefore

$$(40) \qquad \langle (\hat{D}\hat{D}^t)^k \hat{D}\hat{u}, v \rangle = \langle (\tilde{D}\tilde{D}^t)^k \tilde{D}u, v \rangle \leq \langle (DD^t)^k Du, v \rangle,$$

where the equality follows via a consideration of block multiplication of partitioned matrices. Combining (39) with (40), we obtain that

$$(41) \qquad 0 \leq \langle Du, v \rangle^{2k+1} = \langle \hat{D}\hat{u}, v \rangle^{2k+1} \leq \langle (\hat{D}\hat{D}^t)^k \hat{D}\hat{u}, v \rangle \leq \langle (DD^t)^k Du, v \rangle.$$

The inequality (13) is therefore true in this case. We now consider the case for equality. If $\langle (DD^t)^k Du, v \rangle = 0$, then by (41) we also have $\langle Du, v \rangle^{2k+1} = 0$, which in turn implies that $F_{m,n}(D, u, v) = 0$. Similarly, if there exists $\alpha > 0$ such that $Du = \alpha v$ and $D^t v = \alpha u$, then Lemma 3 guarantees us that $F_{m,n}(D, u, v) = 0$. The stated condition is therefore sufficient for equality on $\mathcal{W}_0$. Now, suppose conversely that $\langle (DD^t)^k Du, v \rangle = \langle Du, v \rangle^{2k+1}$; that is, $F_{m,n}(D, u, v) = 0$. Assume also that $\langle (DD^t)^k Du, v \rangle \neq 0$. This means that we have equality throughout (41), so

$$(42) \qquad \langle (\hat{D}\hat{D}^t)^k \hat{D}\hat{u}, v \rangle \neq 0 \quad \text{and} \quad \langle \hat{D}\hat{u}, v \rangle^{2k+1} = \langle (\hat{D}\hat{D}^t)^k \hat{D}\hat{u}, v \rangle.$$

Inductive hypothesis now implies that there exists $\alpha > 0$ such that $\hat{D}\hat{u} = \alpha v$ and $\hat{D}^t v = \alpha \hat{u}$. But $\hat{D}\hat{u} = Du$, so we have $Du = \alpha v$. Under these conditions Lemma 4 tells us that

$$(43) \qquad D^t v = \alpha u \quad \text{or} \quad \langle (DD^t)^k Du, v \rangle > \langle Du, v \rangle^{2k+1}.$$

Since the inequality in (43) is excluded, we must have both $Du = \alpha v$ and $D^t v = \alpha u$. This proves that the inequality holds on $\mathcal{W}_0$, and reduces to equality if and only if the stated conditions hold.

Since $\mathcal{W}_{m,n}$ is compact as a subset of $\mathbb{R}^m \times \mathbb{R}^n$, the function $F_{m,n}(D, \cdot, \cdot)$, which we now denote by F, is bounded on $\mathcal{W}_{m,n}$, and we must therefore assume its absolute minimum value at some point of $\mathcal{W}_{m,n}$. If this point is not in $\mathcal{W}_0$, then it must be in $\mathcal{W}_+$, and must therefore be a critical point of F restricted to $\mathcal{W}_+$. We will complete the proof of Theorem 4 by showing that $F(u, v) \geq 0$ for each critical point $(u, v) \in \mathcal{W}_+$, with equality if and only if there exists $\alpha > 0$ such that $Du = \alpha v$ and $D^t v = \alpha u$, or if $\langle (DD^t)^k Du, v \rangle = 0$.

By extension $F(\cdot, \cdot)$ is a function from $\mathbb{R}^n \times \mathbb{R}^m$ to $\mathbb{R}$. We think of derivatives in the sense of Fréchet; thus, the partial derivative of F with respect to its first variable at the point (u, v), which we denote by $F_1(u, v)$, is a linear functional on $\mathbb{R}^n$. Similarly, $F_2(u, v)$, the partial derivative of F with respect to its second variable, is a linear functional on $\mathbb{R}^m$. For example, if we define the function $G : \mathbb{R}^p \times \mathbb{R}^p \to \mathbb{R}$ by $G(x, y) = \langle x, y \rangle^2$, then $G_1(x, y)(z) = 2\langle x, y \rangle \langle z, y \rangle$ for all $z \in \mathbb{R}^p$. Similarly, $G_2(x, y)(z) = 2\langle x, y \rangle \langle x, z \rangle$ for all $z \in \mathbb{R}^p$.

We are concerned with the restriction of F to $\mathcal{W}_n \times \mathcal{W}_m$, which is $(\mathcal{S}^n \cap \mathbb{R}^n_+) \times (\mathcal{S}^m \cap \mathbb{R}^m_+)$. If (u, v) is interior to this set, then $u > 0$ and $v > 0$ entrywise, and $\langle u, u \rangle = \langle v, v \rangle = 1$. To find the critical points of F restricted in this way we consider the function

$$H(u, v, \lambda_1, \lambda_2) = F(u, v) + (\lambda_1/2)(1 - \langle u, u \rangle) + (\lambda_2/2)(1 - \langle v, v \rangle)$$

defined for all $(u, v) \in \mathbb{R}^n_+ \times \mathbb{R}^m_+$, and $(\lambda_1, \lambda_2) \in \mathbb{R} \times \mathbb{R}$. In order that (u, v) be critical for F restricted as noted, it is necessary that $(u, v, \lambda_1, \lambda_2)$ be critical for H. Thus, each of the four partial derivatives H_i, $1 \leq i \leq 4$, must be zero at $(u, v, \lambda_1, \lambda_2)$. Of course, that $H_3(u, v, \lambda_1, \lambda_2) = H_4(u, v, \lambda_1, \lambda_2) = 0$ merely conveys that the constraints $\|u\| = 1$ and $\|v\| = 1$ are satisfied. We focus upon H_1 and H_2. If $(u, v, \lambda_1, \lambda_2)$ is critical with respect to H, then we must have

(44)
$$H_1(u, v, \lambda_1, \lambda_2)(z_1) = \langle (DD^t)^k Dz_1, v \rangle - (2k+1)\langle Du, v \rangle^{2k} \langle Dz_1, v \rangle - \lambda_1 \langle z_1, u \rangle = 0,$$
$$\forall z_1 \in \mathbb{R}^n,$$

and

(45)
$$H_2(u, v, \lambda_1, \lambda_2)(z_2) = \langle (DD^t)^k Du, z_2 \rangle - (2k+1)\langle Du, v \rangle^{2k} \langle Du, z_2 \rangle - \lambda_2 \langle z_2, v \rangle = 0,$$
$$\forall z_2 \in \mathbb{R}^m.$$

We will use equations (44) and (45) to obtain information about $(u, v, \lambda_1, \lambda_2)$. Setting $z_1 = u$ in (44) and $z_2 = v$ in (45) we obtain that

(46)
$$\langle (DD^t)^k Du, v \rangle - (2k+1)\langle Du, v \rangle^{2k+1} - \lambda_1 = 0$$

and

(47)
$$\langle (DD^t)^k Du, v \rangle - (2k+1)\langle Du, v \rangle^{2k+1} - \lambda_2 = 0.$$

Subtracting (47) from (46), we obtain that $\lambda_1 = \lambda_2$. Letting λ denote λ_1, equations (44) and (45) become

(48)
$$\langle (DD^t)^k Dz_1, v \rangle - (2k+1)\langle Du, v \rangle^{2k} \langle Dz_1, v \rangle = \lambda \langle z_1, u \rangle, \quad \forall z_1 \in \mathbb{R}^n,$$

and

$$\langle (DD^t)^k Du, z_2 \rangle - (2k+1)\langle Du, v \rangle^{2k} \langle Du, z_2 \rangle = \lambda \langle z_2, v \rangle, \quad \forall z_2 \in \mathbb{R}^m. \tag{49}$$

But (48) and (49) hold for all $z_1 \in \mathbb{R}^n$ and $z_2 \in \mathbb{R}^m$, respectively. Hence, the vector equations

$$(D^tD)^k D^t v - (2k+1)\langle Du, v \rangle^{2k} D^t v = \lambda u \tag{50}$$

and

$$(DD^t)^k Du - (2k+1)\langle Du, v \rangle^{2k} Du = \lambda v \tag{51}$$

must also hold for some λ if $(u, v) \in \mathcal{W}_+$ is critical. Let $\mathcal{C}$ denote the set of all $(x, y) \in \mathcal{W}_+$ such that (50) and (51) are satisfied when $u = x$ and $v = y$. Since $F(x, y) \geq 0$ for all $(x, y) \in \mathcal{W}_0$, if there exists $(x, y) \in \mathcal{W}_{m,n}$ such that $F(x, y) < 0$, then there must exist $(x, y) \in \mathcal{C}$ such that $F(x, y) < 0$. We will show that $F(x, y) \geq 0$ for all $(x, y) \in \mathcal{C}$. We will first consider the case $\lambda = 0$, but first we need some definitions.

Since $D \neq 0$ we know that $D^tD \in \mathbb{R}_+^{n \times n}$ and that its spectral radius α^2, where $\alpha > 0$, is a positive eigenvalue of D^tD having an associated non-negative unit eigenvector which we denote by η_1. Thus, $\|\eta_1\| = 1$ and $D^tD\eta_1 = \alpha^2 \eta_1$. Let $\eta_2 = \alpha^{-1} D\eta_1$ so that $D\eta_1 = \alpha \eta_2$. Then, $\langle \eta_2, \eta_2 \rangle = \alpha^{-2} \langle D\eta_1, D\eta_1 \rangle = \alpha^{-2} \langle D^tD\eta_1, \eta_1 \rangle = \langle \eta_1, \eta_1 \rangle = 1$ and $D^t\eta_2 = \alpha^{-1} D^tD\eta_1 = \alpha \eta_1$. Thus, we have

$$D\eta_1 = \alpha \eta_2, \quad D^t \eta_2 = \alpha \eta_1, \quad \eta_1 \geq 0, \quad \eta_2 \geq 0, \quad \text{and} \quad \|\eta_1\| = \|\eta_2\| = 1, \tag{52}$$

and $F(\eta_1, \eta_2) = 0$. Setting $\lambda = 0$ in (50) and (51) we obtain the equations

$$(D^tD)^k D^t v = (2k+1)\langle Du, v \rangle^{2k} D^t v \tag{53}$$

and

$$(DD^t)^k Du = (2k+1)\langle Du, v \rangle^{2k} Du. \tag{54}$$

Equation (54) implies that $\langle (DD^t)^k Du, \eta_2 \rangle - (2k+1)\langle Du, v \rangle^{2k} \langle Du, \eta_2 \rangle = 0$. But, $\langle Du, \eta_2 \rangle = \langle u, D^t\eta_2 \rangle = \alpha \langle u, \eta_1 \rangle$. Since $\alpha > 0$, $\|u\| = \|\eta_1\| = 1$, $u > 0$, and $\eta_1 \geq 0$, we have $\alpha \langle u, \eta_1 \rangle > 0$. Moreover, $DD^t\eta_2 = \alpha^2 \eta_2$, so we have $(DD^t)^k \eta_2 = \alpha^{2k} \eta_2$ and $D^t(DD^t)^k \eta_2 = (D^tD)^k D^t\eta_2 = \alpha^{2k+1} \eta_1$. Thus, $\langle (DD^t)^k Du, \eta_2 \rangle = \langle u, (D^tD)^k D^t\eta_2 \rangle = \alpha^{2k+1} \langle u, \eta_1 \rangle$. Combining these results we obtain the equation

$$\langle (DD^t)^k Du, \eta_2 \rangle - (2k+1)\langle Du, v \rangle^{2k} \langle Du, \eta_2 \rangle = \alpha \langle u, \eta_1 \rangle \big(\alpha^{2k} - (2k+1)\langle Du, v \rangle^{2k} \big) = 0. \tag{55}$$

Since $\alpha \langle u, \eta_1 \rangle > 0$, (55) implies that

$$\alpha^{2k} = (2k+1)\langle Du, v \rangle^{2k}. \tag{56}$$

Substituting (56) into (53) and (54) we obtain the equations

$$(D^tD)^k D^t v = \alpha^{2k} D^t v \quad \text{and} \quad (DD^t)^k Du = \alpha^{2k} Du. \tag{57}$$

Now, substituting the second of the equations in (57) into F, we obtain that

$$F(u, v) = \langle (DD^t)^k Du, v \rangle - \langle Du, v \rangle^{2k+1} = \langle Du, v \rangle \big(\alpha^{2k} - \langle Du, v \rangle^{2k} \big). \tag{58}$$

But $D \neq 0$, and both u and v are strictly positive, therefore $\langle Du, v \rangle > 0$. Moreover, the Cauchy-Schwartz inequality says that $\langle Du, v \rangle^2 \leq \langle Du, Du \rangle \langle v, v \rangle = \langle D^tDu, u \rangle \leq \alpha^2$, since α^2 is the largest eigenvalue of the symmetric matrix D^tD. Therefore, $\alpha^{2k} \geq \langle Du, v \rangle^{2k}$ and $F(u, v) \geq 0$. This proves that $F(u, v) \geq 0$ for any $(u, v) \in \mathcal{C}$ associated with $\lambda = 0$. Now, for equality to happen it would have to be true that

$\langle Du, v \rangle^2 = \alpha^2$, which can only happen if there exists β such that $Du = \beta v$. But then $\beta > 0$, since $u > 0$ and $v > 0$, so Lemma 4 says that we must have either $D^t v = \beta u$ or $\langle (DD^t)^k Du, v \rangle > \langle Du, v \rangle^{2k+1}$. Since the latter possibility is excluded, we have $Du = \beta v$ and $D^t v = \beta u$. Thus, if $(u, v) \in \mathcal{C}$ via association with $\lambda = 0$, then either $F(u, v) > 0$ or else $F(u, v) = 0$, and there exists $\beta > 0$ such that $Du = \beta v$ and $D^t v = \beta u$. In fact, $\beta = \alpha$.

Henceforth, we assume that $\lambda \neq 0$. Note that since u and η_1 are unit non-negative vectors, if we had $\beta_1 \eta_1 = u$ for some β_1, then β_1 would be 1; that is, $u = \eta_1$. In this case $(DD^t)^k Du = \alpha^{2k+1} \eta_2$, so equation (51) transforms into

$$(59) \qquad \alpha^{2k+1} \big(1 - (2k+1) \langle \eta_2, v \rangle^{2k} \big) \eta_2 = \lambda v.$$

Since $\lambda \neq 0$, equation (59) says that $v = \beta_2 \eta_2$ for some β_2. But η_2 and v are non-negative unit vectors, so $\beta_2 = 1$; that is, $v = \eta_2$. Thus, the assumption that $u = \beta_1 \eta_1$ implies that $u = \eta_1$ and $v = \eta_2$, in which case we have $F(u, v) = 0$, and $Du = \alpha v$ and $D^t v = \alpha u$. Similarly, if we start with the assumption that $v = \beta_2 \eta_2$, then we can prove in the same way that $v = \eta_2$ and $u = \eta_1$, so our theorem is true in either of these cases. Henceforth, we assume that u is not a multiple of η_1 and v is not a multiple of η_2. Thus, letting $Z_u = \mathrm{Span}\{u, \eta_1\}$ and $Z_v = \mathrm{Span}\{v, \eta_2\}$, we know that $\dim(Z_u) = \dim(Z_v) = 2$. Let Z_1 denote the orthogonal complement of the subspace of $\mathbb{R}^n$ generated by η_1; thus, $Z_1 = \big(\mathrm{Span}\{\eta_1\} \big)^{\perp}$. Similarly, let $Z_2 = \big(\mathrm{Span}\{\eta_2\} \big)^{\perp}$. Since $\dim(Z_1) = n-1$, $\dim(Z_u) = 2$, and $u \notin Z_1$, the dimension of $Z_u \cap Z_1$ must be 1. Let ξ_1 be a unit vector in $Z_u \cap Z_1$ such that $\langle u, \xi_1 \rangle > 0$. Clearly there must exist a unit vector $\xi_1 \in Z_u \cap Z_1$. If $\langle \xi_1, u \rangle = 0$, then, since $\{\xi_1, \eta_1\}$ is an orthonormal basis for Z_u, u would be a multiple of η_1, which is a case that has already been excluded, and if $\langle \xi_1, u \rangle < 0$, then we replace ξ_1 with $-\xi_1$. Similarly, $\dim(Z_v \cap Z_2) = 1$, and there exists a unit vector $\xi_2 \in Z_v \cap Z_2$ such that $\langle \xi_2, v \rangle > 0$. The set $\{\eta_2, \xi_2\}$ is an orthonormal basis for Z_v just as $\{\eta_1, \xi_1\}$ is an orthonormal basis for Z_u. Thus we have $u = a\eta_1 + b\xi_1$ and $v = c\eta_2 + d\xi_2$, where $a = \langle u, \eta_1 \rangle > 0$, $b = \langle u, \xi_1 \rangle > 0$, $c = \langle v, \eta_2 \rangle > 0$, and $d = \langle v, \xi_2 \rangle > 0$. Due to (52) we have $D\eta_1 = \alpha \eta_2$ and $D^t \eta_2 = \alpha \eta_1$; thus, it is easily seen that $(DD^t)^k D\eta_1 = \alpha^{2k+1} \eta_2$ and $(D^t D)^k D^t \eta_2 = \alpha^{2k+1} \eta_1$. Letting K denote $(2k+1) \langle Du, v \rangle^{2k}$, we employ (51) to obtain that

$$(60) \quad \begin{aligned} \lambda \langle v, \eta_2 \rangle &= \langle (DD^t)^k Du, \eta_2 \rangle - K \langle Du, \eta_2 \rangle = \langle u, (D^t D)^k D^t \eta_2 \rangle - K \langle u, D^t \eta_2 \rangle \\ &= \alpha^{2k+1} \langle u, \eta_1 \rangle - \alpha K \langle u, \eta_1 \rangle = \alpha \langle u, \eta_1 \rangle \big(\alpha^{2k} - K \big). \end{aligned}$$

Taking the inner product of both sides of (50) with η_1, using the fact that $(DD^t)^k D\eta_1 = \alpha^{2k+1} \eta_2$, and proceeding as in (60) we obtain the equation

$$(61) \qquad \lambda \langle u, \eta_1 \rangle = \alpha \langle v, \eta_2 \rangle \big(\alpha^{2k} - K \big).$$

Since $\langle v, \eta_2 \rangle = c$ and $\langle u, \eta_1 \rangle = a$, equations (60) and (61) imply that

$$(62) \qquad \lambda c = \alpha a \big(\alpha^{2k} - K \big) \quad \text{and} \quad \lambda a = \alpha c \big(\alpha^{2k} - K \big).$$

Since $\lambda \neq 0$ and $a > 0$, the term $\alpha^{2k} - K \neq 0$. We may therefore divide the first equation in (62) by the second to obtain that $c/a = a/c$ or $a^2 = c^2$. But $a > 0$ and $c > 0$, so $a = c$; that is, $\langle u, \eta_1 \rangle = \langle v, \eta_2 \rangle > 0$. But $a^2 + b^2 = c^2 + d^2 = 1$, $b > 0$, and $d > 0$; therefore $a = c$ implies that $b = d$. Thus, $u = a\eta_1 + b\xi_1$ and $v = a\eta_2 + b\xi_2$.

Consider the functions $x : [0, 1] \to \mathbb{R}^n$ and $y : [0, 1] \to \mathbb{R}^m$ defined by $x(t) = \sqrt{1 - t^2}\eta_1 + t\xi_1$ and $y(t) = \sqrt{1 - t^2}\eta_2 + t\xi_2$. Clearly, $\|x(t)\| = \|y(t)\| = 1$ for all

$t \in [0,1]$; moreover, we have $x(0) = \eta_1$, $x(1) = \xi_1$, $y(0) = \eta_2$, and $y(1) = \xi_2$. That $u = a\eta_1 + b\xi_1$ and $v = a\eta_2 + b\xi_2$ implies that if we let t_0 denote b, then $0 < t_0 < 1$, $x(t_0) = u$, and $y(t_0) = v$. Although u is in the interior of $\mathbb{R}_+^n$ and v is in the interior of $\mathbb{R}_+^m$, neither ξ_1 nor ξ_2 can be an interior point. Otherwise we could not have $\langle \eta_1, \xi_1 \rangle = 0$ and $\langle \eta_2, \xi_2 \rangle = 0$. There is therefore a number t_1 such that $t_1 > t_0$, $x(t_1)$ is on the boundary of $\mathbb{R}_+^n$, but $x(t)$ is in the interior of $\mathbb{R}_+^n$ for all t such that $t_0 \le t < t_1$. Similarly, there is a number t_2 such that $t_2 > t_0$, $y(t_2)$ is on the boundary of $\mathbb{R}_+^m$, but $y(t)$ is in the interior of $\mathbb{R}_+^m$ for all t such that $t_0 \le t < t_2$. Let $\bar{t}$ denote the minimum of t_1 and t_2. Then, for all t such that $t_0 \le t < \bar{t}$, both $x(t)$ and $y(t)$ are in the interiors of $\mathbb{R}_+^n$ and $\mathbb{R}_+^m$, respectively, while either $x(\bar{t})$ or $y(\bar{t})$ is a boundary point. Define $\zeta : [0, \bar{t}] \to \mathbb{R}$ by

$$(63) \qquad \zeta(t) = F_{m,n}(D, x(t), y(t)) = \langle (DD^t)^k Dx(t), y(t) \rangle - \langle Dx(t), y(t) \rangle^{2k+1}.$$

Then, $\zeta(t_0) = F_{m,n}(D, u, v)$. We will use a convexity argument to show that $\zeta(t_0) > 0$, thereby completing the proof. Note that $\zeta(0) = F_{m,n}(D, \eta_1, \eta_2) = 0$, and, since either $x(\bar{t})$ is a boundary point of $\mathbb{R}_+^n$ or $y(\bar{t})$ is a boundary point of $\mathbb{R}_+^m$, we must have $\zeta(\bar{t}) \ge 0$. We require a simpler expression for $\zeta(t)$. We have $Dx(t) = \sqrt{1-t^2} D\eta_1 + tD\xi_1 = \alpha\sqrt{1-t^2}\eta_2 + tD\xi_1$, so

$$(64)$$
$$\langle Dx(t), y(t) \rangle = \alpha(1-t^2)\langle \eta_2, \eta_2 \rangle + \alpha t\sqrt{1-t^2}\langle \eta_2, \xi_2 \rangle + t\sqrt{1-t^2}\langle D\xi_1, \eta_2 \rangle + t^2\langle D\xi_1, \xi_2 \rangle.$$

But $\langle \eta_2, \eta_2 \rangle = 1$, $\langle \eta_2, \xi_2 \rangle = 0$, and $\langle D\xi_1, \eta_2 \rangle = \langle \xi_1, D^t\eta_2 \rangle = \alpha\langle \xi_1, \eta_1 \rangle = 0$, so

$$(65) \qquad \langle Dx(t), y(t) \rangle = \alpha(1-t^2) + t^2\langle D\xi_1, \xi_2 \rangle = \alpha - t^2 K_0,$$

where $K_0 = \alpha - \langle D\xi_1, \xi_2 \rangle$. Similarly, $(D^tD)^k(\eta_1) = \alpha^{2k}\eta_1$, so we have

$$(66) \qquad \begin{aligned} (D^tD)^k x(t) &= \alpha^{2k}\sqrt{1-t^2}\,\eta_1 + t(D^tD)^k\xi_1 \text{ and} \\ (DD^t)^k Dx(t) &= \alpha^{2k+1}\sqrt{1-t^2}\,\eta_2 + t(DD^t)^k D\xi_1. \end{aligned}$$

Therefore,

$$(67) \qquad \langle (DD^t)^k Dx(t), y(t) \rangle = \alpha^{2k+1}(1-t^2) + t^2\langle (DD^t)^k D\xi_1, \xi_2 \rangle$$

because the mixed inner product terms involving $\langle \eta_2, \xi_2 \rangle$ and $\langle (DD^t)^k D\xi_1, \eta_2 \rangle$ are both 0. Letting K_1 denote $\alpha^{2k+1} - \langle (DD^t)^k D\xi_1, \xi_2 \rangle$, we therefore have

$$(68)$$
$$\langle (D^tD)^k x(t), y(t) \rangle = \alpha^{2k+1} - t^2 K_1 \quad \text{and} \quad \zeta(t) = \alpha^{2k+1} - t^2 K_1 - (\alpha - t^2 K_0)^{2k+1}$$

for all $t \in [0, \bar{t}]$. We claim that $K_0 \ge 0$ and that $K_1 \ge 0$; the cases $K_0 = 0$ and $K_0 > 0$ are then considered separately. By the Cauchy-Schwartz inequality we have $(\langle D\xi_1, \xi_2 \rangle)^2 \le \langle D\xi_1, D\xi_1 \rangle\langle \xi_2, \xi_2 \rangle = \langle D^tD\xi_1, \xi_1 \rangle \le \alpha^2$, since α^2 is the largest eigenvalue, and spectral radius, of D^tD; thus, $K_0 \ge 0$. Consider the case $\alpha = \langle D\xi_1, \xi_2 \rangle$; that is, $K_0 = 0$. This causes the above instance of Cauchy-Schwartz to reduce to equality, so there must exist $\beta \in \mathbb{R}$ such that $D\xi_1 = \beta\xi_2$. But then $\langle D\xi_1, \xi_2 \rangle = \beta\langle \xi_2, \xi_2 \rangle = \beta$, so $\beta = \alpha$. Similarly,

$$(69) \quad \alpha^2 = (\langle D\xi_1, \xi_2 \rangle)^2 = (\langle \xi_1, D^t\xi_2 \rangle)^2 \le \langle D^t\xi_2, D^t\xi_2 \rangle\langle \xi_1, \xi_1 \rangle = \langle DD^t\xi_2, \xi_2 \rangle \le \alpha^2,$$

since DD^t and D^tD have the same spectral radius. So, one more instance of Cauchy-Schwartz reduces to equality. Consequently, we have $(\langle \xi_1, D^t\xi_2 \rangle)^2 = \langle D^t\xi_2, D^t\xi_2 \rangle$, which can only happen if there exists $\gamma \in \mathbb{R}$ such that $D^t\xi_2 = \gamma\xi_1$. But then $\langle \xi_1, D^t\xi_2 \rangle = \langle D\xi_1, \xi_2 \rangle = \gamma$, so $\gamma = \alpha$. At this point we have $D\xi_1 = \alpha\xi_2$ and $D^t\xi_2 = \alpha\xi_1$. Since $u = a\eta_1 + b\xi_1$ and $v = a\eta_2 + b\xi_2$, we then have $Du = \alpha v$ and

$D^t v = \alpha u$, and $F(u, v) = 0$. Our theorem is therefore true in this case because the asserted inequality reduces to equality and the conditions for equality are in force.

Henceforth we assume that $K_0 > 0$. Note that

$$
(70) \quad
\begin{aligned}
(\langle (DD^t)^k D\xi_1, \xi_2 \rangle)^2 &\le \langle (DD^t)^k D\xi_1, (DD^t)^k D\xi_1 \rangle \langle \xi_2, \xi_2 \rangle \\
&= \langle D(D^t D)^k \xi_1, D(D^t D)^k \xi_1 \rangle = \langle (D^t D)^{2k+1} \xi_1, \xi_1 \rangle \le \alpha^{4k+2},
\end{aligned}
$$

since the largest eigenvalue of $(D^t D)^{2k+1}$ is α^{4k+2}. Consequently, $\langle (DD^t)^k D\xi_1, \xi_2 \rangle \le \alpha^{2k+1}$ and $K_1 \ge 0$. Noting that ζ is a function of t^2 we define the function $\omega : [0, \bar{t}^2] \to \mathbb{R}$ by

$$
(71) \qquad \omega(z) = \zeta(\sqrt{z}) = \alpha^{2k+1} - zK_1 - (\alpha - zK_0)^{2k+1}.
$$

Computing the second derivative of ω with respect to z we obtain

$$
(72) \qquad \omega''(z) = -(2k+1)(2k)K_0^2(\alpha - zK_0)^{2k-1}.
$$

Due to (65) we know that $\langle Dx(\bar{t}), y(\bar{t}) \rangle = \alpha - \bar{t}^2 K_0 \ge 0$; thus, $K_0 > 0$ implies that $\alpha - t^2 K_0 > 0$ for all $t \in [0, \bar{t})$. This implies that $\alpha - zK_0 > 0$ for all $z \in [0, \bar{t}^2)$, which in turn implies that $\omega''(z) < 0$ for all $z \in (0, \bar{t}^2)$. At this point we know that $\omega(0) = 0$, $\omega(\bar{t}^2) \ge 0$, and ω is strictly concave down on $(0, \bar{t}^2)$. This implies ([2, 75]) that $\omega(z) > 0$ for all $z \in (0, \bar{t}^2)$. In particular, we must have $0 < \omega(t_0^2) = \zeta(t_0) = F_{m,n}(D, u, v)$. This completes the proof.

References

[1] G.R. Blakley, and Prabir Roy, A Hölder type inequality for symmetric matrices with non-negative entries, *Proceedings of the American Mathematical Society*, (6)**16**(1965), 1244-1245. MR0184950 (32:2421)

[2] G. Hardy, J.E. Littlewood, and G. Pólya, *Inequalities*, Cambridge University Press, Cambridge, 1999.

[3] Roger A. Horn and Charles R. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge, 1985. MR832183 (87e:15001)

[4] A. Sidorenko, A correlation inequality for bipartite graphs, *Graphs and Combinatorics*, **9**(1993), 201-204. MR1225933 (94b:05177)

[5] A. Sidorenko, Inequalities for functionals generated by bipartite graphs (in Russian), *Discrete Math and Applications* (in Russian), (3)**3**(1991), 50-65. MR1138091 (92h:05136)

Department of Mathematics, Auburn University, Auburn, Alabama 36849