DIMACS Technical Report 98-38
July 1998

ON THE ARITHMETIC-GEOMETRIC MEAN INEQUALITY AND ITS RELATIONSHIP TO LINEAR PROGRAMMING, MATRIX SCALING, AND GORDAN'S THEOREM

by

Bahman Kalantari [1]

[1] Department of Computer Science, Rutgers University, New Brunswick, NJ 08903 (kalantar@cs.rutgers.edu)

DIMACS is a partnership of Rutgers University, Princeton University, AT&T Labs-Research, Bell Labs, Bellcore and NEC Research Institute.

DIMACS is an NSF Science and Technology Center, funded under contract STC–91–19999; and also receives support from the New Jersey Commission on Science and Technology.

ABSTRACT

It is a classical inequality that the minimum of the ratio of the (weighted) arithmetic mean to the geometric mean of a set of positive variables is equal to one, and is attained at the center of the positivity cone. While there are numerous proofs of this fundamental homogeneous inequality, in the presence of an arbitrary subspace, and/or the replacement of the arithmetic mean with an arbitrary linear form, the new minimization is a nontrivial problem. We prove a generalization of this inequality, also relating it to linear programming, to the diagonal matrix scaling problem, as well as to Gordan's theorem. Linear programming is equivalent to the search for a nontrivial zero of a linear or positive semidefinite quadratic form over the nonnegative points of a given subspace. The goal of this paper is to present these intricate, surprising, and significant relationships, called *scaling dualities*, and via an elementary proof. Also, to introduce two conceptually simple polynomial-time algorithms that are based on the scaling dualities, significant bounds, as well as Nesterov and Nemirovskii's machinery of self-concordance. The algorithms are simultaneously applicable to linear programming, to the computation of a separating hyperplane, to the diagonal matrix scaling problem, and to the minimization of the arithmetic-geometric mean ratio over the positive points of an arbitrary subspace. The scaling dualities, the bounds, and the algorithms are special cases of a much more general theory on convex programming, developed by the author. For instance, via the scaling dualities semidefinite programming is a problem dual to a generalization of the classical trace-determinant ratio minimization over the positive definite points of a given subspace of the Hilbert space of symmetric matrices.

Key words. Arithmetic-geometric mean inequality, Linear programming, Gordan's theorem, Diagonal matrix scaling, Convexity.

1. Introduction. The classical (weighted) arithmetic-geometric mean inequality can be viewed as a statement on the minimization of the ratio of the two (weighted) means over the positive orthant: the minimum is attained at the center of this cone, any positive scalar multiple of the vector of ones. While there are numerous proofs of this fundamental homogeneous inequality, see e.g. Hardy, Littlewood, and Pólya [5], Bullen, Mitrinović, and Vasić [3], Mitrinović, Pečarić, and Fink [18], and Alzer [1]; in the presence of a proper subspace not containing the center, the new minimization is a nontrivial problem. A more general version of the above problem is the minimization of the ratio of an arbitrary linear function to the geometric mean, over the positive points of an arbitrary subspace. Still an even more general version is when the ratio is replaced with that of an arbitrary positive semidefinite quadratic form, to the square of the geometric mean.

On the other hand linear programming can be formulated as the problem of computing a nontrivial zero of a linear function over the intersection of the nonnegative orthant and a given subspace of the Euclidean space, or proving that such zero does not exist. This problem is in fact Karmarkar's canonical linear programming [15]. A more general, more interesting, and more important problem than Karmarkar's canonical LP problem is when the linear form is replaced with a positive semidefinite quadratic form, considered in Kalantari [7]. Indeed, as proved in [7], by removing the positive semidefiniteness assumption, the zero-finding problem becomes NP-complete. Linear programming is also equivalent to the problem of testing if the convex-hull of a given set of points in the Euclidean space contains the origin. A more extended version of this problem is the following: determine if the convex-hull of a given set of points contains the origin; otherwise, compute a separating hyperplane. In fact the latter problem calls for an algorithmic proof of Gordan's theorem.

Another problem of interest is the diagonal matrix scaling problem. Given a positive semidefinite symmetric matrix, this is the problem of computing a positive definite diagonal matrix such that pre and post multiplication of the given matrix by the computed diagonal matrix results in a matrix whose row sums (hence column sums) equal prescribed positive numbers; or proving that such diagonal matrix does not exist. The case of diagonal scaling of matrices with nonnegative entries, not necessarily positive semidefinite, has been a problem of interest for a very long time. The matrix scaling problem can also be stated in the presence of a subspace. In this more general case of the problem, the diagonal matrix is replaced with the product of a diagonally scaled orthogonal projection matrix, and a positive definite diagonal matrix induced by a positive point of the underlying subspace. A diagonal scaling problem can also be defined for the case of linear forms over a given subspace. This turns out to be a problem dual to Karmarkar's canonical LP.

The goal of this paper is to present the intricate, surprising, and significant relationships between the above stated problems, in the form dualities, called *scaling dualities*, and via an elementary proof. Also, to introduce two conceptually simple polynomial-time algorithms that are based on the scaling dualities, several significant bounds, as well as Nesterov and Nemirovskii's machinery of self-concordance. The algorithms are simultaneously applicable to linear programming, to the computation of a separating hyperplane, to the diagonal matrix scaling problem, and to the minimization of the arithmetic-geometric mean ratio over the positive points of a given subspace. The scaling dualities, the bounds, and the algorithms are special cases of a much more general theory on convex programming, developed by the author (see Kalantari [13]). For instance, via the scaling dualities semidefinite programming is a problem dual to a generalization of the classical trace-determinant ratio minimization over the positive definite points of a given subspace of the Hilbert space of symmetric matrices. Indeed, all the theorems to be stated in this paper, as well as the two algorithms can be shown to hold for analogous problems defined with respect to the cone of positive semidefinite symmetric matrices. Although the proof of general scaling dualities is nontrivial, the proof of those presented in this paper are established via elementary means. The simplicity of the proofs is due to the very special nature of the underlying problems, i.e., linearity or quadraticity of the homogeneous objective function, as well as symmetric properties of the underlying cone, the nonnegative orthant.

In § 2, we preset the main theorems. In § 3 and § 4, we present their proofs. In § 5, we describe the two algorithms, state their complexities, the main ingredients that result in the algorithms, and their application in the derivation of the claimed complexity results. One of the major ingredient is the scaling dualities. Other ingredients are significant bounding theorems that will be stated without proof. The proof of these bounds is given in [13] and in much more generality. Another ingredient is a theorem that combines two fundamental but most basic properties from Nesterov and Nemirovskii's theory on self-concordance, [19]. In our concluding remarks, § 6, we briefly describe generalization of the theorems and the algorithms, in particular with respect to corresponding problems defined over the cone of positive semidefinite matrices.

2. The arithmetic-geometric mean inequality, linear programming, Gordan's separation theorem, and matrix scaling. Let $K = \{x \in \Re^n : x \geq 0\}$, the nonnegative orthant, and $K^\circ = \{x \in \Re^n : x > 0\}$, the positive orthant. Let $\lambda \in K^\circ$ be an arbitrary vector of *weights*, and $\theta = \sum_{i=1}^n \lambda_i$. The classical (weighted) arithmetic-geometric mean inequality is the following

$$\frac{\frac{1}{\theta}\lambda^T x}{\left(\prod_{i=1}^n x_i^{\lambda_i}\right)^{1/\theta}} \geq 1, \quad \forall x \in K^\circ. \tag{2.1}$$

Moreover, the minimum is attained at any positive scalar multiple of $e = (1, \ldots, 1)^T$.

Let W be a subspace of $\Re^n$. If W is a proper subspace, we will assume that $W = \{x \in \Re^n : Ax = 0\}$, where A is a given $m \times n$ matrix of rank m. Let $S = \{x \in \Re^n : \|x\| = 1\}$. We will assume that $d^0 \in W \cap K^\circ \cap S$ is available. Let $\phi(x) = \frac{1}{2}x^T Qx$, where Q is an $n \times n$ symmetric matrix, also assumed to be positive semidefinite, i.e., $x^T Qx \geq 0$, for all $x \in \Re^n$. Let

$$F_\lambda(x) = -\sum_{i=1}^n \lambda_i \ln x_i. \tag{2.2}$$

Linear programming is equivalent to the problem of testing if ϕ has a nontrivial zero over $W \cap K$, see Kalantari [7]. Indeed, the latter problem is more general and more important than Karmarkar's canonical linear programming problem [15], which is the following problem: given $c \in \Re^n$, determine if $c^T x = 0$, for some $x \in W \cap K$, $x \neq 0$. It is easy to see that ϕ has a desired zero if and only if $\mu = 0$, where

$$\mu = \min\{\phi(x) : x \in W \cap K \cap S\}. \tag{2.3}$$

We shall refer to the problem of testing if $\mu = 0$ as *homogeneous programming* (HP). Let the *logarithmic potential*, and the *homogeneous potential* be defined as

$$\psi_\lambda(x) = \phi(x) + F_\lambda(x), \quad X_\lambda(x) = \frac{\phi(x)}{exp(-\frac{2}{\theta}F_\lambda(x))} = \frac{\phi(x)}{\prod_{i=1}^n x_i^{2\lambda_i/\theta}}, \tag{2.4}$$

respectively. Let

$$\psi_\lambda^* = \inf\{\psi_\lambda(x) : x \in W \cap K^\circ\}, \quad X_\lambda^* = \inf\{X_\lambda(x) : x \in W \cap K^\circ\}. \tag{2.5}$$

The *scaling problem* (SP) is to determine if $\psi_\lambda^* = -\infty$, and if ψ_λ^* is finite, to compute its value, together with a corresponding minimizer d_λ^*. The *Homogeneous scaling problem* (HSP) is to determine if $X_\lambda^* = -\infty$, and if X_λ^* is finite, to compute its value, together with a corresponding minimizer $\bar{d}_\lambda^*$. Since $\phi(x)$ is convex and $F_\lambda(x)$ is strictly convex, $\psi_\lambda(x)$ is strictly convex.

For a given $d \in K^\circ$, let

$$D_\lambda = \nabla^2 F_\lambda(d)^{-1/2} = \mathrm{diag}(\frac{d_1}{\sqrt{\lambda_1}}, \cdots, \frac{d_n}{\sqrt{\lambda_n}}), \tag{2.6}$$

$$e_\lambda = D_\lambda^{-1} d = (\sqrt{\lambda_1}, \cdots, \sqrt{\lambda_n})^T, \tag{2.7}$$

$$\phi_{\lambda,d}(x) = \phi(D_\lambda x) = \frac{1}{2} x^T D_\lambda Q D_\lambda x, \tag{2.8}$$

$$\psi_{\lambda,d}(x) = \psi(D_\lambda x) = \phi_{\lambda,d}(x) + F_\lambda(x) + F_\lambda(d). \tag{2.9}$$

Let $P_{\lambda,d}$ be the orthogonal projection operator onto the subspace $W_{\lambda,d} = D_\lambda^{-1} W = \{x \in \Re^n : AD_\lambda x = 0\}$. Thus, if $W = \Re^n$, then $P_{\lambda,d} = I$, the identity matrix. Otherwise,

$$P_{\lambda,d} = I - D_\lambda A^T (AD_\lambda^2 A^T)^{-1} AD_\lambda. \tag{2.10}$$

The *algebraic scaling problem* (ASP) is to test the solvability of the *scaling equation* (SE) :

$$P_{\lambda,d} \nabla \phi_{\lambda,d}(e_\lambda) = P_{\lambda,d} D_\lambda Q D_\lambda e_\lambda = e_\lambda, \quad d \in W \cap K^\circ. \tag{2.11}$$

THEOREM 2.1. *The following statements are equivalent:*

(1) : $\mu > 0$.

(2) : $\exists\ d \in W \cap K^\circ$ *such that* $\phi(d) > 0$, *and* $X_\lambda(x) \geq X_\lambda(d)$, $\forall\ x \in W \cap K^\circ$.

(3) : $\exists\ d \in W \cap K^\circ$ *such that* $\psi_\lambda(x) \geq \psi_\lambda(d)$, $\forall\ x \in W \cap K^\circ$.

(4) : $\exists\ d \in W \cap K^\circ$ *such that* $P_{e,d} D_e Q D_e e = P_{e,d} \lambda$.

(5) : $\exists\ d \in W \cap K^\circ$ *such that* $P_{e,d} D_e Q D_e e > 0$.

(6) : $\exists\ d \in W \cap K^\circ$ *such that* $P_{\lambda,d} D_\lambda Q D_\lambda e_\lambda = e_\lambda$.

(7) : $\exists\ d \in W \cap K^\circ$ *such that* $\|P_{\lambda,d} D_\lambda Q D_\lambda e_\lambda - e_\lambda\| < \lambda_{min} = \min\{\lambda_1, \ldots, \lambda_n\}$.

(8) : $\exists\ d \in W \cap K^\circ$ *such that* $P_{\lambda,d} D_\lambda Q D_\lambda e_\lambda > 0$.

(9) : $\exists\ d \in W \cap K^\circ$ *such that* $P_{\lambda^2,d} D_{\lambda^2} Q D_{\lambda^2} \lambda = \lambda$, *where* $\lambda^2 = (\lambda_1^2, \cdots, \lambda_n^2)^T$.

Moreover, (3), (4), *and* (6) *have a unique and common solution* d, *and* d *also satisfies* (2).

The following is an obvious but important corollary of Theorem 2.1.

COROLLARY 2.2. (Scaling dualities for positive semidefinite quadratic forms) *Either* $\mu = 0$, *or condition* (i) *is true,* $i = 2, \ldots, 9$; *but not both.* □

The equivalence of the statements of Theorem 2.1 justifies the naming of the corresponding dualities of Corollary 2.2 as *scaling dualities.* From the algorithmic point of view it turns out that the dualities that concern diagonal scaling are the more important dualities.

Indeed, Theorem 2.1 can be stated in more generality, where $\phi(x)$ can be an arbitrary convex homogeneous function. The corresponding proof however becomes more involved than the simple proofs presented in this paper. For the proof of some of the corresponding parts of Theorem 2.1, as derived for general convex homogeneous functions, see Kalantari [11]. We mention that linearly constrained convex quadratic programming can be formulated as the problem of computing a nontrivial root of the homogeneous function $\frac{1}{2}x^TQx/e^Tx + c^Tx$ over $W \cap K$, for some positive semidefinite Q, and $c \in \Re^n$, see Kalantari [13]. In fact, most of the scaling dualities implied by Theorem 2.1 can be stated in much more generality, where the underlying cone is an arbitrary closed convex pointed cone, and ϕ an arbitrary convex form defined over this cone (see § 5). Our study of what we now call scaling dualities, and their algorithmic significance began in Kalantari [6], and was continued in [7]-[11]. We emphasize that the goal of the present paper is to present these significant dualities for the very special problems considered, and via an elementary proof. Also, to introduce two conceptually simple, but very capable polynomial-time algorithms that are motivated by these dualities. Indeed, the linear programming/matrix scaling algorithm of Khachiyan and Kalantari [16] is also based on a scaling duality, implied by Gordan's theorem (see [13]).

Remark 1. When $W = \Re^n$, Corollary 2.2 implies that either there exists $x \geq 0, x \neq 0$, such that $Qx = 0$, or there exists $d > 0$ such that $Qd > 0$; but not both. It is easy to show that this duality is equivalent to Gordan's theorem, see Kalantari [11]. Gordan's theorem (see Dantzig [2], Schrijver [20]) is the following: given a real $m \times n$ matrix B, either there exists $x \geq 0, x \neq 0$, such that $Bx = 0$, or there exists y such that $B^Ty > 0$; but not both. If the latter condition occurs, the vector y induces a hyperplane separating the column vectors of B from the origin. If we let $Q = B^TB$, and $Qd > 0$, then Bd can be taken to be y. We will discuss the complexity of computing a desired zero, or a separating hyperplane in § 5.

Remark 2. When $W = \Re^n$, Corollary 2.2 implies that either there exists $x \geq 0, x \neq 0$, such that $Qx = 0$, or given any $\lambda > 0$, there exists $d > 0$ such that $D_eQD_ee = \lambda$; but not both. When $\lambda = e$, this is the quasi doubly-stochastic scaling problem, a problem which has been of considerable interest for matrices of nonnegative entries, see e.g. Marshal and Olkin [17], Kalantari [11]. Another statement implied by this corollary is that $\mu > 0$ if and only if given any $\lambda > 0$, there exists $d > 0$, such that $D_\lambda QD_\lambda e_\lambda = e_\lambda$. In other words, there exists a diagonal scaling of Q such that e_λ is a fixed-point. Moreover, the latter result is true if and only if there exits a diagonal scaling of Q such that $D_{\lambda^2}QD_{\lambda^2}\lambda = \lambda$, i.e., λ itself is a fixed point. The case of $D_eQD_ee = e$ scaling was shown to be polynomially solvable in [16] via a simple analysis, see also [13]. But in fact all other cases considered in this paper, including matrix scaling over a given subspace, are polynomially solvable. However, the polynomial-time solvability of the subspace constrained case is considerably demanding, requiring several important and nontrivial auxiliary results (see § 5).

The following theorem will be shown to be a consequence of Theorem 2.1. It in particular implies a generalization of the familiar (weighted) arithmetic-geometric mean inequality. Let $c \in \Re^n$ be arbitrary, and define

$$f_\lambda(x) = \frac{c^Tx}{\prod_{i=1}^n x_i^{\lambda_i/\theta}}, \quad g_\lambda(x) = c^Tx - \sum_{i=1}^n \lambda_i \ln x_i \tag{2.12}$$

THEOREM 2.3. *The following statements are equivalent:*

$(1')$: $c^Tx > 0, \ \forall\, x \in W \cap K, \ x \neq 0.$

$(2')$: $\exists\, d \in W \cap K^\circ$ *such that* $c^Td > 0$, *and* $f_\lambda(x) \geq f_\lambda(d), \ \forall x \in W \cap K^\circ$.

$(3')$: $\exists\, d \in W \cap K^\circ$ *such that* $g_\lambda(x) \geq g_\lambda(d), \ \forall x \in W \cap K^\circ$.

$(4')$: $\exists\, d \in W \cap K^\circ$ *such that* $P_{e,d}D_ec = P_{e,d}\lambda$.

$(5')$: $\exists\, d \in W \cap K^\circ$ *such that* $P_{e,d}D_ec > 0$.

$(6')$: $\exists\, d \in W \cap K^\circ$ *such that* $P_{\lambda,d}D_\lambda c = e_\lambda$.

$(7')$: $\exists\ d \in W \cap K^{\circ}$ *such that* $\|P_{\lambda,d}D_{\lambda}c - e_{\lambda}\| < \lambda_{min} = \min\{\lambda_1, \ldots, \lambda_n\}$.

$(8')$: $\exists\ d \in W \cap K^{\circ}$ *such that* $P_{\lambda,d}D_{\lambda}c > 0$.

$(9')$: $\exists\ d \in W \cap K^{\circ}$ *such that* $P_{\lambda^2,d}D_{\lambda^2}c = \lambda$, *where* $\lambda^2 = (\lambda_1^2, \cdots, \lambda_n^2)^T$.

Moreover, $(3')$, $(4')$, *and* $(6')$ *have a unique and common solution* d, *and* d *also satisfies* $(2')$.

The following is an obvious but important corollary of Theorem 2.3.

COROLLARY 2.4. (Scaling dualities for linear forms) *Either* $c^T x = 0$, *for some* $x \in W \cap K$, $x \neq 0$, *or condition* (i) *is true,* $i = 2', \ldots, 9'$; *but not both.* □

Remark 3. For $i = 2'$, Corollary 2.4 implies that either Karmarkar's canonical LP is solvable, or the potential function $f_{\lambda}(x)$ is unbounded from below; but not both. In fact Karmarkar's algorithm, [15], can be viewed as an algorithmic proof of this fact, see also Kalantari [6], [9]. Another duality implied by this corollary is that either Karmarkar's canonical LP is solvable, or there there exists $d \in W \cap K^{\circ}$ such that $P_{e,d}D_e c > 0$, but not both. This duality which implies Gordan's theorem gives rise to a simple variation of Karmarkar's algorithm (see Kalantari [10]), a polynomial-time algorithm that actually establishes this duality.

3. Proof of Theorem 2.1.

Proof. We will prove the theorem by proving the following circle of implications

$$(1) \Longrightarrow (2) \Longrightarrow (3) \Longrightarrow (4) \Longrightarrow (6) \Longrightarrow (7) \Longrightarrow (8) \Longrightarrow (1),$$

and

$$(1) \iff (5), \quad (6) \iff (9).$$

$(1) \Longrightarrow (2)$: Suppose $\mu > 0$. Consider the set $W \cap K^{\circ} \cap S$, $S = \{x \in E : \|x\| = 1\}$. Since $X_{\lambda}(x)$ approaches infinity as x approaches a boundary point of $W \cap K^{\circ} \cap S$, its infimum is attained at some point, say, x^*. We claim that x^* must be the minimum of $X_{\lambda}(x)$ over $W \cap K^{\circ}$. Otherwise, there exists $\bar{x} \in W \cap K^{\circ}$ such that $X_{\lambda}(\bar{x}) < X_{\lambda}(x^*)$. But as X_{λ} is homogeneous of degree zero, we get $X_{\lambda}(\bar{x}/\|\bar{x}\|) = X_{\lambda}(\bar{x}) < X_{\lambda}(x^*)$, a contradiction.

$(2) \Longrightarrow (3)$: Suppose there exists $d \in W \cap K^{\circ}$ such that $\phi(d) > 0$, and $X_{\lambda}(x) \geq X_{\lambda}(d)$, for all $x \in W \cap K^{\circ}$. From the first order optimality condition, d must be a stationary point of X_{λ}, i.e.,

$$P\nabla X_{\lambda}(d) = 0, \tag{3.1}$$

where $P = P_{e,e} = I - A^T(AA^T)^{-1}A$. Differentiating $X_{\lambda}(x)$, we get

$$\nabla X_{\lambda}(x) = \left(\nabla\phi(x) + \frac{2}{\theta}\phi(x)\nabla F_{\lambda}(x)\right) exp(\frac{2}{\theta}F_{\lambda}(x)). \tag{3.2}$$

For any positive real α, we have

$$\nabla\phi(\alpha x) = \alpha\nabla\phi(x), \quad \nabla F_{\lambda}(\alpha x) = \frac{1}{\alpha}\nabla F_{\lambda}(x). \tag{3.3}$$

From (3.1), (3.2), and (3.3) we conclude that for any positive real α, we have

$$P\left(\alpha\nabla\phi(d) + \frac{2}{\theta}(\alpha^2\phi(d))(\frac{1}{\alpha}\nabla F_{\lambda}(d))\right) = 0. \tag{3.4}$$

In particular, since $\phi(d) > 0$, we can let $\alpha = \sqrt{\theta / 2\phi(d)}$. Then, setting $d^* = \alpha d$, we have

$$P \nabla \psi_\lambda(d^*) = 0. \tag{3.5}$$

Equivalently, this implies

$$\nabla \psi_\lambda(d^*) = A^T v, \tag{3.6}$$

for some vector v of Lagrange multipliers. Since ϕ is convex, for all $x \in W \cap K^\circ$ we have

$$\psi_\lambda(x) \geq \psi_\lambda(d^*) + \nabla \psi_\lambda(d^*)^T (x - d^*) = \psi_\lambda(d^*) + v^T A(x - d^*) = \psi_\lambda(d^*), \tag{3.7}$$

hence (3) is satisfied.

(3) $\Longrightarrow$ (4) : Suppose d satisfies (3). Then $P \nabla \psi_\lambda(d) = 0$. Equivalently,

$$\nabla \psi_\lambda(d) = Q D_e e - D_e^{-1} \lambda = A^T v, \tag{3.8}$$

for some v. Multiplying the latter equation by D_e, we get

$$D_e Q D_e e - \lambda = D_e A^T v. \tag{3.9}$$

Multiply the above by AD_e, we can solve for v to get

$$v = (A^T D_e^2 A)^{-1} A D_e (D_e Q D_e - \lambda). \tag{3.10}$$

Substituting the above in the previous equation, we get

$$P_{e,d}(D_e Q D_e - \lambda) = 0. \tag{3.11}$$

Thus, (4) is valid.

(4) $\Longrightarrow$ (6) : If d satisfies (4), then it follows that $\nabla \psi_\lambda(d) = A^T v$, for some v. Equivalently,

$$\nabla \psi_\lambda(d) = Q D_\lambda e_\lambda - D_\lambda^{-1} e_\lambda = A^T v. \tag{3.12}$$

Multiplying the latter equation by D_λ, we get

$$D_\lambda Q D_\lambda e_\lambda - e_\lambda = D_\lambda A^T v. \tag{3.13}$$

Multiply the above by AD_λ, we can solve for v to get

$$v = (A^T D_\lambda^2 A)^{-1} A D_\lambda (D_\lambda Q D_\lambda e_\lambda - e_\lambda). \tag{3.14}$$

Substituting the above in the previous equation, we get

$$P_{\lambda,d}(D_\lambda Q D_\lambda e_\lambda - e_\lambda) = 0. \tag{3.15}$$

Since $AD_\lambda e_\lambda = Ad = 0$, $P_{\lambda,d} e_\lambda = e_\lambda$. But this implies (6) is valid.

(6) $\Longrightarrow$ (7) : This is immediate.

(7) $\Longrightarrow$ (8) : Suppose that d satisfies (7). Assume that a given $y \in \Re^n$ satisfies the inequality $\|y - \lambda\| < \lambda_{min}$. Then, $|y_i - \lambda_i| < \lambda_{min}$, for all $i = 1, \ldots, n$. Thus, it follows that $y > 0$. Hence (8) must hold.

$(8) \Longrightarrow (1)$: Suppose that d satisfies (8), and $\mu = 0$. Since Q is positive semidefinite, there exist $x \in W \cap K$, $x \neq 0$, such that

$$Qx = 0, Ax = 0. \tag{3.16}$$

Let $w = D_\lambda^{-1}x$. Note that $w \geq 0$, $w \neq 0$. Since $P_{\lambda,d}w = w$, and $D_\lambda e_\lambda = d$, we have

$$w^T P_{\lambda,d} D_\lambda Q D_\lambda e_\lambda = w^T D_\lambda Q D_\lambda e_\lambda = x^T Q d. \tag{3.17}$$

On the one hand, since $Qx = 0$, we must have $x^T Qd = 0$. On the other hand, since $P_{\lambda,d}D_\lambda Q D_\lambda e_\lambda > 0$, and $w \geq 0$, $w \neq 0$, we get $x^T Qd > 0$, a contradiction. Thus, if (8) holds, then so does (1).

$(1) \Longrightarrow (5)$: Since (1) implies (4), given any $\lambda \in K^\circ$, we can select $\lambda = e$. Then, there exists $d \in W \cap K^\circ$ such that $P_{e,d}D_e Q D_e e = P_{e,d}e = e$. Thus, (5) holds.

$(5) \Longrightarrow (1)$: Suppose that d satisfies (5). Then, d satisfies (8) for $\lambda = e$. Hence, $\mu > 0$.

$(6) \Longrightarrow (9)$: This implication and and its converse are trivial.

To prove the remaining claims of the theorem, we need to observe that since ψ_λ is strictly convex, the minimizer d of ψ_λ, if it exists, must be unique. We have already seen that if d satisfies (4) or (6), then it must satisfy $P\nabla\psi_\lambda(d) = 0$. Thus, d is the minimizer of ψ. Hence it satisfies (3) and is unique. Next we show that d also satisfies (2). Suppose that d' satisfies (2). Then, $P\nabla X_\lambda(d') = 0$. Since $\phi(d') > 0$, some positive scalar multiple of d', say $\alpha d'$ satisfies $P\nabla\psi_\lambda(\alpha d') = 0$. But then we must have $\alpha d' = d$. Since $X_\lambda(\alpha x) = X(x)$, for all $x \in W \cap K^\circ$, d also satisfies (2). □

4. Proof of Theorem 2.3.

Proof. Assume that $(1')$ holds. Let $Q = cc^T$. Then, the corresponding μ is positive. From Theorem 2.1 we conclude that for each $i = 2, \ldots, 9$, condition (i) is satisfied.

$(2) \Longrightarrow (2')$: Since $X_\lambda(x) = f_\lambda^2(x)$, and $c^T x > 0$, for $x \in W \cap K$, $x \neq 0$, $(2')$ follows.

$(3) \Longrightarrow (3')$: Suppose that d satisfies (3). Then $cc^T d - D_e^{-1}\lambda = A^T v$, for some vector of lagrange multipliers, v. Since $\alpha = c^T d > 0$, from the above equation we get $c - \bar{D}_e^{-1}\lambda = A^T \bar{v}$, where $\bar{d} = d/\alpha$, $\bar{v} = v/\alpha$. But this implies that $\bar{d}$ is the minimizer of $g_\lambda(x) = c^T x + F_\lambda(x)$.

$(4) \Longrightarrow (4')$: Suppose d satisfies (4), i.e., $P_{e,d}D_e cc^T D_e e = P_{e,d}\lambda$. Since $\alpha = c^T d > 0$, $\hat{d} = \alpha d$ lies in $W \cap K^\circ$. Since $P_{e,d} = P_{e,\hat{d}}$, we get $P_{e,\hat{d}}\hat{D}_e c = P_{e,\hat{d}}\lambda$.

Analogously to the above, it follows that for $i = 5 - 8$, if d satisfies (i), then $\hat{d} = (c^T d)d$ satisfies (i').

$(8') \Longrightarrow (1')$: Follows from similar argument given for proof of the corresponding part of Theorem 2.1.

$(6') \Longrightarrow (9')$: This and its converse are trivial.

The proof of the remaining parts follow in a similar fashion as in Theorem 2.1. □

5. Two polynomial-time algorithms for HP, SP, HSP, and ASP. Here we will describe two polynomial-time algorithms for solving ϵ-approximate version of these problems. These algorithms are conceptually easy. However, their analysis requires the use of scaling dualities, significant bounds, as well as some results from Nesterov and Nemirovskii's machinery of self-concordance. First, we define ϵ-approximate version of the four problems. Next, we describe the two algorithms. We will then state without proof the main ingredients, other than the scaling dualities already proved. Finally, we will justify how the main ingredients result in the stated complexity results. For the proof of the stated theorems and in much more generality, see Kalantari [13].

Given $\epsilon \in (0,1]$, ϵ-HP, ϵ-SP, ϵ-HSP, and ϵ-ASP are defined as follows. ϵ-HP is to compute $d \in W \cap K \cap S$ such that $\phi(d) \leq \epsilon$, or proving that such point does not exist. Given $\lambda \in K^\circ$, ϵ-ASP is to test the solvability of $\|P_{\lambda,d} D_\lambda Q D_\lambda e_\lambda - e_\lambda\| < \epsilon$. ϵ-SP is to compute, if it exists, a point $d \in W \cap K^\circ$ such that $\psi_\lambda(d) - \psi_\lambda^* \leq \epsilon$. ϵ-HSP is to compute, if it exists, a point $d \in W \cap K^\circ$ such that $X_\lambda(d)/X_\lambda^* \leq exp(\epsilon)$. In particular, if $\epsilon = \lambda_{min}$, then from Theorem 2.1 the solution of ϵ-ASP will result in a point $d \in W \cap K^\circ$ such that $P_{\lambda,d} D_\lambda Q D_\lambda e_\lambda > 0$. For the convex-hull problem (see Remarks 1 and 3) such a point gives rise to a separating hyperplane.

Let $u = -P\nabla\psi_\lambda(d^0)$, where d^0 is a given point in $W \cap K^\circ \cap S$. For each $t \in (0, \infty)$, define

$$f_\lambda^{(t)}(x) = t\phi(x) + tu^T x + F_\lambda(x).$$

Given $d \in W \cap K^\circ$, let

$$f_{\lambda,d}^{(t)}(x) = f_\lambda^{(t)}(D_\lambda x) = t\phi_{\lambda,d}(x) + tu^T Dx + F_\lambda(x) + F_\lambda(d).$$

Also, given $d \in W \cap K^\circ$, let y be the Newton direction with respect to ψ_λ, i.e., the solution to the equation $P\nabla^2\psi_\lambda(d)y = -P\nabla\psi_\lambda(d)$, where P is the orthogonal projection operator onto W. The Newton decrement is defined as $\delta(d) = [y^T \nabla^2 \psi_\lambda(d) y]^{1/2}$. Let $\bar{\delta} \in [\delta_*, 1)$, where $\delta_* = 2 - \sqrt{3}$. The Newton iterate is

$$d' = NEW(\psi_\lambda, d) = d + \sigma(d)y, \qquad \sigma(d) = \begin{cases} \frac{1}{1+\delta(d)}, & \text{if } \delta(d) > \bar{\delta}; \\ \frac{1-\delta(d)}{\delta(d)(3-\delta(d))}, & \text{if } \delta(d) \in [\delta_*, \bar{\delta}]; \\ 1, & \text{if } \delta(d) < \delta_*. \end{cases}$$

Analogously, corresponding to $f_\lambda^{(t)}$, one defines y_t, $\delta_t(d)$, and $d'_t = NEW(f_\lambda^{(t)}, d)$.

Consider the following Potential-Reduction algorithm:

Potential-Reduction:

Initialization. Let $d = d^0$.

Iterative Step. Replace d with $d' = NEW(\psi_\lambda, d)$ and repeat.

THEOREM 5.1. (Potential-Reduction Complexity Theorem, Kalantari [13]) *Assume that $\lambda \geq e$. Let ϵ be in $(0,1)$. Consider the Potential-Reduction algorithm, and let $\sigma = exp[\frac{2}{\theta}\psi_\lambda(d^0) - 1 + \ln\frac{\theta}{2}]$, and $R = \sup\{exp(-\frac{2}{\theta}F_\lambda(x)) : x \in W \cap K^\circ, \|x\| = 1\}$.*

If $\mu = 0$, the number of iterations to solve ϵ-HP is $O(\theta \ln \frac{R\sigma}{\epsilon})$. If $\mu > 0$, the number of iterations to solve ϵ-SP, ϵ-HSP, or ϵ-ASP is $O(\theta \ln \frac{R\sigma}{\mu} + \ln\ln \frac{\|Q\|}{\epsilon})$. □

Consider the following Path-Following algorithm, where $t_* \in (0,1)$ is an appropriately selected number:

Path-Following:

Initialization. Let $t = 1$, $d = d^0$.

Phase I. While $t > t_*$, replace (d, t) with (d'_t, t'), where $d'_t = NEW(f_\lambda^{(t)}, d)$, $t' = t\, exp(-c_*/\sqrt{\theta})$, c_* an absolute positive constant (can be taken equal to $1/9$).

Phase II. Replace d with $d' = NEW(f^{(t)}, d)$ and repeat.

THEOREM 5.2. (Path-Following Complexity Theorem, Kalantari [13]) *Assume $\lambda \geq e$. Let ϵ be in $(0,1)$. Consider the Path-Following algorithm.*

If $\mu = 0$, the number of iterations to solve ϵ-HP is $O(\sqrt{\theta}\ln\frac{\theta\|u\|}{\epsilon} + \ln\ln||Q||)$. If $\mu > 0$, the number of iterations to solve ϵ-ASP is $O(\sqrt{\theta}\ln\frac{\theta\|u\|}{\mu\epsilon} + \ln\ln||Q||)$, and the number of iterations to solve ϵ-SP or ϵ-HSP is $O(\sqrt{\theta}\ln\frac{\theta\|u\|}{\mu} + \ln\ln\frac{\|Q\|}{\epsilon})$. $\square$

Remark 4. The above algorithms also apply to the case of an arbitrary $\lambda > 0$, not necessarily satisfying $\lambda \geq e$. Essentially all is needed is to replace λ with $\lambda' = \lambda/\lambda_{min}$. These algorithm can be used to solve linear programs, to compute separating hyperplanes, to diagonally scale positive semidefinite symmetric matrices over an arbitrary subspace, or to compute the minimum of the ratio of an arbitrary linear (positive semidefinite form) to the geometric mean (square of the geometric mean) over the positive points of an arbitrary subspace. The implementation of the above two algorithms for ϵ-SP, ϵ-HSP, and ϵ-ASP requires the computation of a lower bounds on μ, assuming that it is positive. For the computation of such a lower bound for the case where Q has rational or algebraic entries, see [12]. These algorithms can also be applied to solve Karmarkar's canonical LP. All is needed is to replace the linear form c^Tx with $(c^Tx)^2$. In particular, taking $\lambda = e$, the Path-Following algorithm gives an $O(\sqrt{n}\ln\frac{n\|u\|}{\epsilon})$ iteration complexity for solving ϵ-approximate version of Karmarkar's problem. As usual, over the rational inputs ϵ can be taken to be $O(2^{-L})$, where L is the *size* of LP (see [12] for different notions of size, over the rationals or the algebraic numbers). The Path-Following algorithm not only results in the best-known iteration complexity for solving LP (see e.g. [4]), while being one of the simplest of such algorithms, but also is an algorithm that is capable of solving more general diagonal matrix scaling problems than that considered in [16], as well as having the capability of solving SP and HSP.

We now explain the main ingredients, other than the scaling dualities already proved, that result in the above two complexity theorems. The following theorem characterizes basic properties of Newton's method (see [19], Theorem 2.2.2), and main properties of the parametric family $f_\lambda^{(t)}$ (see [19], Theorem 3.1.1):

THEOREM 5.3. (Nesterov and Nemirovskii [19]) *Given $d \in W \cap K^\circ$, $d' = NEW(\psi_\lambda, d) \in W \cap K^\circ$, and*

$$\begin{cases} \psi_\lambda(d') - \psi_\lambda(d) \leq -(\delta(d) - \ln(1+\delta(d))), & \text{if } \delta(d) > \bar{\delta}; \\ \delta(d') \leq \frac{1}{4}(6\delta(d) - \delta^2(d) - 1), & \text{if } \delta(d) \in [\delta_*, \bar{\delta}]; \\ \delta(d') \leq \frac{\delta^2(d)}{(1-\delta(d))^2}, & \text{if } \delta(d) < \delta_*. \end{cases}$$

If $\delta(d) < \frac{1}{3}$, then $\psi_\lambda^* = \inf\{\psi_\lambda(x) : x \in W \cap K^\circ\} > -\infty$ and

$$\psi_\lambda(d) - \psi_\lambda^* \leq \frac{\omega^2(\delta(d))(1+\omega(\delta(d))}{2(1-\omega(\delta(d))}, \quad \omega(\delta(d)) = 1 - (1 - 3\delta(d))^{1/3}.$$

Given $t \in (0, \infty)$, all the above apply to $f_\lambda^{(t)}$. Moreover, suppose that $\delta_t(d) \leq 1/4$, then $\delta_{t'}(d) \leq 1/4$, where $t' = t\ exp(\frac{-c_*}{\sqrt{\theta}})$. $\square$

Now we state some other significant results.

THEOREM 5.4. (Kalantari [13], § 8) *Assume that $\lambda \geq e$. Suppose that $d \in W \cap K^\circ$ satisfies $\delta(d) < \delta_*/2$. Then,*

$$\delta(d) \leq ||P_{\lambda,d}\nabla\psi_{\lambda,d}(e_\lambda)|| = O\left(\frac{\theta||Q||}{\mu}\delta(d)\right).$$

Suppose that $\mu > 0$. Given any $t \in (0, 1]$, suppose that $d \in W \cap K^\circ$ satisfies $\delta_t(d) < \delta_*$. Then,

$$||P_{\lambda,d}\nabla f_{\lambda,d}^{(t)}(e_\lambda)|| = O\left(\frac{\theta||Q||\ ||u||}{\mu}\delta_t(d)\right).$$

In either case ($\mu > 0$, or $\mu \le 0$), given any $t \in (0,1]$, suppose that $d \in W \cap K^\circ$ is a point obtained via Phase I of the Path-Following algorithm, and $\delta_t(d) < \delta_*/2$. Then,

$$\|P_{\lambda,d}\nabla f^{(t)}_{\lambda,d}(e_\lambda)\| = O\left(\|Q\|\delta_t(d)(\frac{1}{t})^{\frac{\sqrt{\theta}}{c_*}}\right). \quad \square$$

THEOREM 5.5. (Kalantari [13], § 9) *Let the k-th iterate of the Potential-Reduction algorithm be denoted by d^k. If $\delta(d^k) \ge \delta_*$, then*

$$\phi\left(\frac{d^k}{\|d^k\|}\right) \le R\sigma \ exp\left(-\frac{2k\xi_*}{\theta}\right), \qquad \xi_* = \delta_* - \ln(1+\delta_*). \quad \square$$

THEOREM 5.6. (Kalantari [13], § 9) *Suppose that $\mu > 0$. Then, for all $x \in W \cap K^\circ$ we have*

$$\ln\left(\frac{X_\lambda(x)}{X^*_\lambda}\right)^{\frac{\theta}{2}} \le \psi_\lambda(x) - \psi^*_\lambda. \quad \square$$

The following theorem is the final ingredient, a significant consequence of the scaling dualities:

THEOREM 5.7. (Kalantari [13], § 10) *Let γ be a number in $(0,1]$. Given $t \in (0,1]$, suppose that $d \in W \cap K^\circ$ satisfies $\|P_{\lambda,d}\nabla f^{(t)}_{\lambda,d}(e_\lambda)\| \le \frac{1}{2}\gamma$. Let $\hat{d} = \sqrt{t}d$. If $\|P_{\lambda,\hat{d}}\nabla \psi_{\lambda,\hat{d}}(e_\lambda)\| \ge \gamma$, then $\phi(\frac{d}{\|d\|}) \le C(\gamma)t$, where*

$$C(\gamma) = \frac{1}{\gamma^2}[2\theta + \gamma(\sqrt{\theta}+1)]\|u\|^2. \quad \square$$

Now we describe how Theorem 2.1, and Theorems 5.3-5.7 imply the desired complexity theorems, Theorem 5.1, and Theorem 5.2. We will first describe this for Theorem 5.1.

From Theorem 2.1, HP is solvable (i.e. $\mu = 0$) if and only if the other three problems are not. If $\mu = 0$, Theorem 5.5 gives the desired complexity bound to solve ϵ-HP. If $\mu > 0$, since $\phi(d^k/\|d^k\|) \ge \mu$, Theorem 5.5 implies that in $O(\theta \ln \frac{R\sigma}{\mu})$ iterations we obtain a point $x \in W \cap K^\circ$ such that $\delta(x) < \delta_*$. Let x^k be the subsequent iterates of the Potential-Reduction algorithm, where $x^0 = x$. From Theorem 5.3 it follows that

$$\delta(x^k) \le r_*^{2^k}, \quad r_* = \frac{\delta_*}{(1-\delta_*)^2}. \tag{5.1}$$

The above together with the first bound in Theorem 5.4 implies that the number of iterations, k, to get a point x^k satisfying ϵ-ASP, is $O(\ln\ln \frac{\theta\|Q\|}{\mu\epsilon})$. Hence the claimed complexity for solving ϵ-ASP. From Theorem 5.3 it also follows that (see [13]) once we have obtained a point x^k such that $\delta(x^k) < \min\{\delta_*, \frac{\epsilon}{3}(1-\omega(\delta_*))\}$, we have

$$\psi_\lambda(x^k) - \psi^*_\lambda \le \epsilon. \tag{5.2}$$

This together with the bound in Theorem 5.6 gives the desired complexity for solving ϵ-SP, or ϵ-HSP.

We will next derive the complexity result of Theorem 5.2. From Theorem 5.3 it follows that given any $t \in (0,1]$, since $\delta_1(d^0) = 0$, the number of iterations, k_t, to obtain $d \in W \cap K^\circ$ such that $\delta_t(d) < \delta_*$, satisfies

$$O\left(\sqrt{\theta}\ln(\frac{1}{t})\right). \tag{5.3}$$

Suppose that $\mu = 0$, then from Theorem 2.1 we must have $||P_{\lambda,\hat{d}} \nabla \psi_{\lambda,\hat{d}}(e_\lambda)|| \geq \gamma_{min}$. From this, and Theorem 5.7 it suffices to compute $d \in W \cap K^\circ$ such that

$$(5.4) \qquad ||P_{\lambda,d} \nabla f_d^{(t_*)}(e_d)|| \leq \frac{1}{2}\gamma_{min},$$

where t_* satisfies

$$(5.5) \qquad \frac{1}{t_*} = \frac{C(\gamma_{min})}{\epsilon}.$$

Thus, the number of iterations of Phase I will equal $O(\sqrt{\theta} \ln \frac{\theta ||u||}{\epsilon})$, and at the termination of this phase we have a point $x \in W \cap K^\circ$ such that $\delta_{t_*}(x) < \delta_*$. Let $x^0 = x, x^1, \ldots, x^k$ be the sequence of iterates of Phase II. From Theorem 5.3, it follows that $\delta_{t_*}(x^k) \leq r_*^{2^k}$. From this and the third upper bound given in Theorem 5.4, it follows that the number of iterations of Phase II to solve ϵ-HP is $O(\ln \ln ||Q|| \epsilon^{-\sqrt{\theta}/c_*})$. Hence the claimed combined complexity of the two phases.

Suppose that $\mu > 0$. To solve ϵ-ASP, from Theorem 5.7 it suffices to compute $d \in W \cap K^\circ$ such that

$$(5.6) \qquad ||P_{\lambda,d} \nabla f_d^{(t_*)}(e_d)|| \leq \frac{\epsilon}{2},$$

where t_* satisfies

$$(5.7) \qquad \frac{1}{t_*} = \frac{2C(\epsilon)}{\mu}.$$

Thus, the number of iterations of Phase I is $O(\sqrt{\theta} \ln \frac{\theta ||u||}{\mu \epsilon})$. At the termination of this phase we have a point $x \in W \cap K^\circ$ such that $\delta_{t_*}(x) < \delta_*$. Let $x^0 = x, x^1, \ldots, x^k$ be the sequence of iterates of Phase II. From the second bound of Theorem 5.4, and since $\delta_{t_*}(x^k) \leq r_*^{2^k}$, the number of iterations of Phase II is $O(\ln \ln \frac{\theta ||Q|| \, ||u||}{\mu \epsilon})$. Hence the claimed combined complexity.

To solve ϵ-SP, or ϵ-HSP, we first solve δ_*-ASP, i.e., compute a point $d > 0$ such that $||P_{\lambda,d} D_\lambda Q D_\lambda e_\lambda - e_\lambda|| < \delta_*$, in $O(\sqrt{\theta} \ln \frac{\theta ||u||}{\mu})$ iterations. We then replace d with $NEW(\psi_\lambda, d)$, and repeat this step. The number of latter iterations can be estimated as it was done for the Potential-Reduction algorithm.

6. Concluding remarks. In this paper we have proved several dualities relating a generalization of the classical arithmetic geometric mean inequality to linear programming, to Gordan's theorem, and to several diagonal matrix scaling problems. These dualities, called scaling dualities, although were derived via elementary means are remarkable and significant from the theoretical and algorithmic point of view. Many of the scaling dualities can be stated with respect to much more general convex programming problems, relating their equivalent HP formulation to problems that can be viewed as analogues of SP, HSP, and ASP, considered in this paper. The latter three problems are genuinely dual to the HP formulation of convex programming problems, i.e., HP is solvable if and only if the other three are not. Indeed, as shown in [13], given an arbitrary closed convex pointed cone K in a finite dimensional space, a logarithmically homogeneous barrier $F(x)$ for the interior of K, a subspace W, and a homogeneous function ϕ with homogeneous degree $p > 0$, the scaling equation can be defined as $P_d \nabla \phi_d(e_d) = e_d$, or $P_d \nabla^2 \phi_d(e_d) = (p-1)^{-1} e_d$, if $p \neq 1$; where $d \in W \cap K^\circ$, $D = \nabla^2 F(d)^{-1/2}$, $\phi_d(x) = \phi(Dx)$, $e_d = D^{-1}d$, and P_d the orthogonal projection onto $W_d = D^{-1}W$. For instance, semidefinite programming can be formulated as the problem of testing the existence of a nontrivial zero of $tr(cx)$, over an arbitrary subspace of the space of symmetric matrices, and its intersection with the cone of positive semidefinite symmetric matrices, $K = S_n^+$, where c is a given symmetric matrix, and $tr(\cdot)$ is the trace function. Now suppose that given $\lambda \in K^\circ$, we take $F_\lambda(x) = -\ln det(x^\lambda)$, where x^λ is defined as follows. Firstly, given a matrix $x \in K^\circ$, the matrix

$exp(x) = \sum_{j=1}^{\infty} x^j/j!$ is well-defined and lies in K°. If x is a diagonal positive definite matrix, define $\ln x$ as the matrix $\mathrm{diag}(\ln x_1, \ldots, \ln x_n)$. If x is an arbitrary positive definite matrix, define $\ln x = U_x \ln \Lambda_x U_x^T$, where $x = U_x \Lambda_x U_x^T$, with U_x unitary, and Λ_x the diagonal matrix of eigenvalues of x. Define x^λ as the matrix $exp(\lambda^{1/2} \ln x \lambda^{1/2})$. Now if for a given $d \in K^\circ$ we define $D_\lambda = \nabla^2 F_\lambda(d)^{-1/2}$, and $e = I$ the identity matrix, and D_e the operator that maps $x \in S_n^+$ to $d^{1/2} x d^{1/2}$, then all the results proved or presented in this paper apply, verbatim. These include all the complexity theorems, and the auxiliary theorems that imply them. The quantity λ_{min} will stand for the minimum eigenvalue of λ. These results are the subject of a forthcoming paper, [14]. In particular, in [14] we prove the following inequality which is a generalization of the arithmetic-geometric mean inequality as well as the well-known trace-determinant inequality: let λ be an arbitrary matrix in K°, the cone of positive definite symmetric matrices, and $\theta = tr(\lambda)$. Then,

$$\frac{\frac{1}{\theta} tr(\lambda x)}{\left(\det(x^\lambda)\right)^{1/\theta}} \geq 1, \quad \forall x \in K^\circ.$$

REFERENCES

[1] H. Alzer, A proof of arithmetic-geometric mean inequality, *Amer. Math. Monthly*, 103 (1996) 585.

[2] G.B. Dantzig, *Linear Programming and Extensions* (Princeton University Press, Princeton, New Jersey, 1963).

[3] P.S. Bullen, D.S. Mitrinović, and P.M. Vasić, *Means and Their Inequalities* (Reidel, Dordrecht, 1988).

[4] C.C. Gonzaga, Path-following methods for linear programming, *SIAM Review*, 34 (1992) 167-224.

[5] G.H. Hardy, J.E. Littlewood, and G. Pólya, *Inequalities* (Cambridge University Press, Cambridge, 1952).

[6] B. Kalantari, Karmarkar's algorithm with improved steps, *Math. Programming*, 46 (1990) 73-78.

[7] B. Kalantari, Canonical problems for quadratic programming and projective methods for their solution, *Contemporary Mathematics*, 114 (1990) 243-263.

[8] B. Kalantari, Derivation of a generalized and strengthened Gordan theorem from generalized Karmarkar potential and logarithmic barrier functions, Technical Report LCSR-TR-121, Department of Computer Science, Rutgers University, New Brunswick, NJ, 1989.

[9] B. Kalantari, Generalization of Karmarkar's algorithm to convex homogeneous functions, *Oper. Res. Lett.*, 11 (1992) 93-98.

[10] B. Kalantari, A simple polynomial time algorithm for a convex hull problem equivalent to linear programming, *Combinatorics Advances*, Kluwer Academic Publishers (1995) 207-216.

[11] B. Kalantari, A Theorem of the alternative for multihomogeneous functions and its relationship to diagonal scaling of matrices, *Linear Algebra and its Applications*, 236 (1996) 1-24.

[12] B. Kalantari and M.R. Emamy-K, On linear programming and matrix scaling over the algebraic numbers, *Linear Algebra and its Applications*, 262 (1997) 283- 306.

[13] B. Kalantari, Scaling dualities and self-concordant homogeneous programming in finite dimensional spaces, Technical Report DCS-TR 359, Department of Computer Science, Rutgers University, New Brunswick, NJ, 1998.

[14] B. Kalantari, On the trace-determinant inequality, semidefinite programming, and matrix scaling over the cone matrices, forthcoming.

[15] N. Karmarkar, A new polynomial time algorithm for linear programming, *Combinatorica*, 4 (1984) 373-395.

[16] L. Khachiyan and B. Kalantari, Diagonal matrix scaling and linear programming, *SIAM J. Optim.*, 4 (1992) 668-672.

[17] A.W. Marshall and I. Olkin, Scaling of matrices to achieve specified row and column sums, *Numer. Math.*, 12 (1968) 83-90.

[18] D.S. Mitrinović, J.E. Pečarić, and A.M. Fink, *Classical and New Inequalities in Analysis* (Kluwer Academic Publishers, Dordrecht, 1993).

[19] Y. Nesterov and A.S. Nemirovskii, *Interior-Point Polynomial Algorithms in Convex Programming* (SIAM, Philadelphia, PA, 1994).

[20] A. Schrijver, *Theory of Linear and Integer Programming* (John Wiley and Sons, New York, 1986).