WEIGHTED COMPLEMENTARITY PROBLEMS - A NEW PARADIGM FOR COMPUTING EQUILIBRIA

FLORIAN A. POTRA*

Abstract. This paper introduces the notion of a weighted Complementarity Problem (wCP), which consists in finding a pair of vectors (x, s) belonging to the intersection of a manifold with a cone, such that their product in a certain algebra, $x \circ s$, equals a given weight vector w. When w is the zero vector, then wCP reduces to a Complementarity Problem (CP). The motivation for introducing the more general notion of a wCP lies in the fact that several equilibrium problems in economics can be formulated in a natural way as wCP. Moreover, those formulations lend themselves to the development of highly efficient algorithms for solving the corresponding equilibrium problems. For example, Fisher's competitive market equilibrium model can be formulated as a wCP that can be efficiently solved by interior-point methods. Moreover, it is shown that the Quadratic Programming and Weighted Centering problem, which generalizes the notion of a Linear Programming and Weighted Centering problem recently proposed by Anstreicher, can be formulated as a special linear monotone wCP. The main contribution of the paper is to introduce and analyze two interior-point methods for solving general monotone linear wCPs.

Key words. weighted complementarity, interior-point, path-following, Fisher equilibrium

AMS subject classifications. 90C51, 90C33

1. Introduction. The aim of this paper is to introduce the notion of a weighted Complementarity Problem (wCP), to study its theoretical properties, and to construct numerical methods for its numerical solution. This notion significantly extends the notion of a complementarity problem (CP). Generally speaking, wCP consists in finding a pair of vectors (x, s) belonging to the intersection of a manifold with a cone, such that their product in a certain algebra, $x \circ s$, equals a given weight vector w. When w is the zero vector, wCP reduces to a Complementarity Problem (CP). With nonzero weight vectors, the theory of wCP becomes more complicated than the theory of CP. However, many of the essential properties of CP extend to wCP. Also, many interior-point methods for CP can be extended to efficient algorithms for solving wCP.

We have been motivated to introduce the notion of wCP by the fact that wCP can be used for modeling a larger class of problems from science and engineering. Even when a problem can also be modeled by CP, the wCP model leads to a more efficient numerical solution method. For example, we will show that the Fisher market equilibrium problem, which can be modeled as a *nonlinear* CP, can also be modeled a *linear* wCP. The latter can be solved more efficiently than the former. Ye's algorithm [25] for solving the Fisher problem with linear utilities can be viewed as a numerical method for solving a particular instance of a linear wCP. It consists of two phases. First, the potential reduction method from [24, page 106] is used to construct a starting point satisfying a restrictive centering condition [25, (5)]. Then a modified primal-dual path-following algorithm is employed to compute an ε-approximate solution of the Fisher market equilibrium problem. We note that the path followed by Ye's algorithm is nonsmooth, and that the primal-dual path-following algorithm, which is a modification of the standard primal-dual path-following algorithm for linear complementarity problems of Kojima et al. [14], is a short-step algorithm that has optimal computational complexity but rather poor practical performance [22]. Ye

*Department of Mathematics and Statistics, University of Maryland Baltimore County, 1000 Hilltop Circle, Baltimore, MD 22150 (potra@umbc.edu).

[25, Corollary 1] mentions that the predictor-corrector method of Mizuno-Todd-Ye (MTY) can be used instead of the short-step algorithm, but gives no details about how MTY is to be modified for this problem.

Recently the computational complexity results for the Fisher problem from [25] have been improved by Anstreicher [2]. He proposes a generalization of the Eisenberg-Gale formulation of the Fisher problem [6], called Linear Programming and Weighted Centering (LPWC) problem and shows that it possesses a natural dual problem DPWC. He obtains a lower iteration complexity for DPWC by utilizing a combination of the volumetric [1, 23] and logarithmic [17] barriers. It turns out that LPWC generalizes both linear programming (LP), and the problem of finding the weighted analytic center of a polytope [3, 10]. In the present paper we consider a generalization of LPWC, called Quadratic Programming and Weighted Centering (QPWC) and we show that this problem and its dual lead to a monotone linear wCP.

The main contribution of our paper is to propose and analyze two interior-point methods for solving general monotone linear wCPs. The first method can be interpreted as an extension of the largest-step path-following method of McShane [15]. The algorithm requires one matrix factorization per iteration, has the same computational complexity as the short-step algorithm, but is much more efficient in practice. We also propose an extension of the MTY predictor-corrector method which requires two matrix factorizations at each iteration. Both algorithms follow a smooth central path. When applied to the Fisher problem, they can use the standard starting point described in [25] without having to use another method for obtaining a better centered starting point.

Although, for the sake of generality, we present the notion of a nonlinear wCP in a Jordan algebra setting, we only study the case of a linear wCP over the nonnegative orthant of $\mathbb{R}^n$. In this case all the components of the weight vector w are nonnegative, which is denoted by $w \geq 0$. When $w = 0$ this problem reduces to a mixed horizontal linear complementarity problem, while in the case when all components of w are strictly positive ($w > 0$) we obtain a weighted analytic centering problem. These two problems have been extensively studied and their properties are well understood. The more general case when some of the components of w are equal to zero and the other components are strictly positive ($w \geq 0$) appears to be more difficult to analyze, and generalization of algorithms developed for the cases $w = 0$ and $w > 0$ is nontrivial.

Conventions. We denote by $\mathbb{N}$ the set of all nonnegative integers. $\mathbb{R}$, $\mathbb{R}_+$, $\mathbb{R}_{++}$ denote the set of real, nonnegative real, and positive real numbers respectively. The symbol e represents the vector of all ones, with dimension given by the context. We denote by $\log t$ the natural logarithm of t.

We denote component-wise operations on vectors by the usual notations for real numbers. Thus, given two vectors u, v of the same dimension, uv, u/v, etc. will denote the vectors with components $u_i v_i$, u_i/v_i, etc. This notation is consistent as long as component-wise operations always have precedence in relation to matrix operations. Note that $Auv = A(uv) \neq (Au)v$. Also, if f is a scalar function and v is a vector, then $f(v)$ denotes the vector with components $f(v_i)$. For example if $v \in \mathbb{R}^n_+$, then $\sqrt{v}$ denotes the vector with components $\sqrt{v_i}$, and $1 - v$ denotes the vector with components $1 - v_i$. Traditionally the vector $1 - v$ is written as $e - v$, where e is the vector of all ones. Inequalities are to be understood in a similar fashion. For example if $v \in \mathbb{R}^n$, then $v \geq 3$ means that $v_i \geq 3,\ i = 1, \ldots, n$. Traditionally this is written as $v \geq 3e$. For a vector $v \in \mathbb{R}^n$ we denote $\max v = \max\{v_i \ :\ i = 1, \ldots, n\}$ and $\min v = \min\{v_i \ :\ i = 1, \ldots, n\}$. If $\|\,.\,\|$ is a vector norm on $\mathbb{R}^n$ and A is a matrix, then the operator norm induced by $\| \, . \, \|$ is defined by $\| A \| = \max\{\| Ax \| \; ; \; \| x \| = 1\}$. As a particular case we note that if U is the diagonal matrix defined by the vector u, then $\| U \|_2 = \| u \|_\infty$.

Throughout this paper we use the MATLAB-like notation $[\, u \, ; \, v \, ; \, w \,]$ to denote the column vector $[u^T v^T w^T]^T$. Given a matrix P, we denote by $\operatorname{Ran} P$ its range (or column space) and by $\operatorname{Ker} P$ its kernel (or null space).

2. The weighted complementarity problem. We first present the notion of a wCP in a general Jordan algebra setting [7], since this general framework reveals the essential geometric features of a wCP and will be used in subsequent papers. If $(\mathcal{J}, \circ)$ is an Euclidean Jordan algebra, we denote by $\mathcal{K} = \{x \circ x \; : \; x \in \mathcal{J}\}$ the cone formed by the squares of its elements. It turns out that the cone $\mathcal{K}$ is symmetric, in the sense that it is self-dual, $\mathcal{K} = \mathcal{K}^*$, and its automorphism group acts transitively on its interior. Symmetric cones are intimately related to Euclidean Jordan algebras, since it can be proved that a cone is symmetric if and only if it is the cone of squares of some Euclidean Jordan algebra. Interestingly, symmetric cones are also intimately related to self-scaled barriers, since a cone is symmetric if and only if it is self-scaled, i.e., it admits a self-scaled barrier. Interior-point methods for solving linear programming problems over self-scaled cones were studied in [18, 19]. The relation between the self-scaled cones and the symmetric cones (also called homogenous self-dual cones) was noted in [12]. Classical results in Euclidean Jordan algebras show that every symmetric cone can be decomposed into five irreducible symmetric cones (see also [13] and the literature cited therein). The most important symmetric cones used in mathematical programming are the positive orthant, the cone of positive semidefinite matrices, and the second-order (Lorentz) cone. The extension of interior-point algorithms to a Jordan algebra setting, was first detailed in [9, 8]. In [8] one considers the problem of minimizing a convex quadratic programming problem over the intersection of a linear manifold and a symmetric cone, and it is shown that the monotone linear complementarity problem over symmetric cones can be reduced to such a problem. In fact, the reverse is also true. The linear manifold considered in [8] was a subset of $\mathcal{J} \times \mathcal{J}$ that possessed a certain monotony property.

Since the Lagrange multipliers $y \in I\!R^m$ corresponding to equality constraints play a special role in equilibrium problems, we consider a (possibly nonlinear) manifold $\mathcal{M} \subset \mathcal{J} \times \mathcal{J} \times I\!R^m$. Given such a manifold and a vector $w \in \mathcal{K}$, we define a wCP as the problem of finding

$$(x, s, y) \in \mathcal{M} \cap \mathcal{K} \times \mathcal{K} \times I\!R^m, \quad \text{such that } x \circ s = w. \tag{2.1}$$

If $\mathcal{M}$ is defined as

$$\mathcal{M} = \{(x, s, y) \in \mathcal{J} \times \mathcal{J} \times I\!R^m \; : \; F(x, s, y) = 0\}, \tag{2.2}$$

where $F : \mathcal{J} \times \mathcal{J} \times I\!R^m \to \mathcal{J} \times I\!R^m$ is a given nonlinear map, then (2.1) can be written as the problem of finding a solution of the following nonlinear system

$$\begin{aligned} x \circ s &= w \\ F(x, s, y) &= 0 \\ x, s &\in \mathcal{K} \end{aligned} \tag{2.3}$$

With a zero weight vector $w = 0$, this reduces to a mixed nonlinear complementarity problems over the Euclidean Jordan algebra $\mathcal{J}$. Such problems were considered, for example, in [27]. There are no existence results for the solution of (2.3) in the general case. When $w = 0$ an existence result follows from [27, Theorem 3.10] (see also [27, Corollary 4.4]). The uniqueness of the solution is not guaranteed even for the particular case of a monotone linear complementarity problem.

2.1. The linear wCP over the nonnegative orthant. In this subsection we will analyze the case where $\mathcal{J} = \mathbb{R}^n$, $x \circ s = xs$, $\mathcal{K} = \mathbb{R}^n_+$, and F is an affine mapping. Then (2.3) can be written as

$$\begin{aligned} xs &= w \\ Px + Qs + Ry &= a \\ x, s &\geq 0 \end{aligned} \tag{2.4}$$

Here $P \in \mathbb{R}^{(n+m)\times n}, Q \in \mathbb{R}^{(n+m)\times n}, R \in \mathbb{R}^{(n+m)\times m}$ are given matrices, $a \in \mathbb{R}^{n+m}$ is a given vector, and $w \in \mathbb{R}^n_+$ is a given weight vector (the data of the problem). The matrix R is assumed to have full column rank. In the first equation above, xs denotes the vector having as components the product of the corresponding components of x and s. In other words, xs is the componentwise product of the vectors x and s. The notation $x \geq 0$ means that all components of the vector x are nonnegative, i.e., $x \in \mathbb{R}^n_+$. Similarly, $x > 0$ means that all components of the vector x are positive, i.e., $x \in \mathbb{R}^n_{++}$. wCP (2.4) is called monotone if

$$P\Delta x + Q\Delta s + R\Delta y = 0 \text{ implies } \Delta x^T \Delta s \geq 0\,, \tag{2.5}$$

and it is called skew-symmetric if

$$P\Delta x + Q\Delta s + R\Delta y = 0 \text{ implies } \Delta x^T \Delta s = 0\,. \tag{2.6}$$

If (2.4) is monotone and strictly feasible then it has a solution. This statement will be proved in a more general setting in [21].

In the next subsection we will show that the Fisher problem reduces to a skew-symmetric wCP.

2.2. The Fisher equilibrium problem as a wCP. We consider a market composed of $n_c \geq 2$ consumers and $n_p \geq 2$ producers. Consumer i has a budget $w_i > 0$ to spend on buying goods from the producers in such a way that an individual utility function is maximized. The price equilibrium is an assignment of prices to goods, so that when every consumer buys a maximal bundle of goods then the market clears, meaning that all the money is spent and all the goods are sold. Without loss of generality, we assume that producer j has one unit of some good to sell. Let the individual utility function of consumer i be of the form

$$u_i = \sum_{j=1}^{n_p} u_{ij} x_{ij}\,, \tag{2.7}$$

where u_{ij} is the utility coefficient of consumer i for the good produced by producer j, and x_{ij} represents the amount of good bought by consumer i from producer j. We assume that the following inequalities are satisfied for all i and j:

$$w_i > 0, \quad u_{ij} \geq 0, \quad \sum_{k=1}^{n_c} u_{kj} > 0, \quad \sum_{k=1}^{n_p} u_{ik} > 0\,. \tag{2.8}$$

Under these assumptions Eisenberg and Gale [6] proved that the market clearing prices are given by the optimal Lagrange multipliers for the last n_p equality constraints of the following convex optimization problem:

$$
\begin{aligned}
(2.9)\qquad & \underset{u_i,\, x_{ij}}{\text{maximize}} \quad \sum_{i=1}^{n_c} w_i \log u_i \\
& \text{subject to} \\
& \quad u_i - \sum_{j=1}^{n_p} u_{ij} x_{ij} = 0\,, \quad i = 1, \ldots, n_c \\
& \quad \sum_{i=1}^{n_c} x_{ij} = 1, \qquad j = 1, \ldots, n_p \\
& \quad u_i \geq 0,\ x_{ij} \geq 0\,, \quad i = 1, \ldots, n_c\,,\ j = 1, \ldots, n_p.
\end{aligned}
$$

This optimization problem can be written under the form

$$
\begin{aligned}
(2.10)\qquad & \underset{x}{\text{maximize}} \sum_{i=1}^{n} w_i \log x_i \\
& \text{subject to} \quad Ax = b \\
& \qquad\qquad\quad x \geq 0\,,
\end{aligned}
$$

where x is a n-dimensional vector, with $n = n_c(n_p + 1)$, having its first n_c coordinates formed by $u_1, ..., u_{n_c}$, and the remaining $n_c n_p$ coordinates consisting of the variables x_{ij},

$$
x = [u_1, ..., u_{n_c}, x_{1\,1}, \ldots, x_{1\,n_p}, x_{2\,1}, \ldots, x_{2\,n_p}, \ldots, x_{n_c\,1}, \ldots, x_{n_c\,n_p}]^T.
$$

The n-dimensional weight vector w has its first n_c coordinates equal to $w_1, \ldots, w_{n_c}$, and the remaining coordinates equal to zero. A is a full rank $m \times n$-matrix, with $m = n_c + n_p$, and b is an m-dimensional vector having its first n_c coordinates equal to zero, and the remaining n_p coordinates equal to 1, i.e., $b = [0; e]$. By introducing the vectors

$$
a^1 = -[u_{1\,1}, \ldots, u_{1\,n_p}]^T,\ a^2 = -[u_{2\,1}, \ldots, u_{2\,n_p}]^T,\ \ldots, a^{n_c} = -[u_{n_c\,1}, \ldots, u_{n_c\,n_p}]^T,
$$

we can write

$$
A = \begin{pmatrix}
1 & 0 & \cdots & 0 & a^{1\,T} & 0 & \cdots & 0 \\
0 & 1 & \cdots & 0 & 0 & a^{2\,T} & \cdots & 0 \\
. & . & \cdots & . & . & . & \cdots & . \\
0 & 0 & \cdots & 1 & 0 & 0 & \cdots & a^{n_c\,T} \\
0 & 0 & \cdots & 0 & e_1^T & e_1^T & \cdots & e_1^T \\
0 & 0 & \cdots & 0 & e_2^T & e_2^T & \cdots & e_2^T \\
. & . & \cdots & . & . & . & \cdots & . \\
0 & 0 & \cdots & 0 & e_{n_p}^T & e_{n_p}^T & \cdots & e_{n_p}^T
\end{pmatrix}
= \begin{pmatrix}
I_{n_c} & A_1 & A_2 & \cdots & A_{n_c} \\
0 & I_{n_p} & I_{n_p} & \cdots & I_{n_p}
\end{pmatrix}.
$$

Let us now consider a general optimization problem of the form (2.10), with arbitrary $w \in I\!R^n_+$, $A \in I\!R^{m \times n}$, $b \in I\!R^m$. We assume that A is full rank. When all components of the weight vector w are positive, i.e. $w > 0$, (2.10) is known as a Weighted Analytic

Center Problem [3, 10]. In the present paper we consider the more general case where some of the components of w may vanish. This makes the problem more difficult. We note that if x is the solution of (2.10), then we must have $x_i > 0$ whenever $w_i > 0$, so that in this case the product $w_i x_i^{-1}$ is well defined. If $w_i = 0$, then we take by definition $w_i x_i^{-1} = 0$ for any value of x_i. With this convention, the KKT conditions for (2.10) can be written as a nonlinear CP:

$$
\begin{aligned}
xv &= 0 \\
wx^{-1} + v - A^T y &= 0 \\
Ax &= b \\
x, v &\geq 0
\end{aligned}
\tag{2.11}
$$

We note that the second equation above is defined for any $v \in I\!R^n_+$, $y \in I\!R^m$, and any $x \in I\!R^n_+$ such that $x_i \neq 0$ whenever $w_i > 0$. By denoting $s = A^T y$, we have $s = wx^{-1} + v \geq 0$. Multiplying this equation by x we obtain the following linear wCP:

$$
\begin{aligned}
xs &= w \\
s - A^T y &= 0 \\
Ax &= b \\
x, s &\geq 0
\end{aligned}
\tag{2.12}
$$

This a particular case of wCP (2.4) with

$$
P = \begin{pmatrix} A \\ 0 \end{pmatrix}, \; Q = \begin{pmatrix} 0 \\ I \end{pmatrix}, \; R = \begin{pmatrix} 0 \\ -A^T \end{pmatrix}, \; a = \begin{pmatrix} b \\ 0 \end{pmatrix},
\tag{2.13}
$$

which is easily shown to be skew-symmetric. Indeed, if the right-hand-side of (2.6) is satisfied then $\Delta x \in \operatorname{Ker} A$ and $\Delta s \in \operatorname{Ran} A^T$, so that $\Delta x^T \Delta s = 0$.

We note that (2.12) was first obtained by Ye [25], who used the Eisenberg and Gale [6] formulation of the Fisher problem.

2.3. The Quadratic Programming and Weighted Centering problem. In this section we introduce a more general convex optimization problem that leads to a monotone linear wCP. Given an $n \times n$ symmetric positive semidefinite matrix M, a full rank $m \times n$ matrix A, with $m < n$, and vectors $f \in I\!R^n$, $w \in I\!R^n_+$, $b \in I\!R^m$, we consider the following optimization problems:

$$
\begin{aligned}
&\underset{x}{\text{minimize}} \quad \varphi(x) := \frac{1}{2} x^T M x + f^T x - \sum_{i=1}^n w_i \log x_i \\
&\text{subject to} \qquad Ax = b \\
&\qquad\qquad\qquad x \geq 0\,,
\end{aligned}
\tag{2.14}
$$

and

$$
\begin{aligned}
&\underset{u,s,y}{\text{maximize}} \quad \psi(u,s,y) := -\frac{1}{2} u^T M u + b^T y + \sum_{i=1}^n w_i \log s_i + \sum_{i=1}^n w_i (1 - \log w_i) \\
&\text{subject to} \qquad s = Mu - A^T y + f \\
&\qquad\qquad\qquad s \geq 0.
\end{aligned}
\tag{2.15}
$$

In (2.14) and (2.14) we have tacitly made the convention that if $w_i = 0$ then the corresponding terms $w_i \log x_i$, $w_i \log s_i$, $w_i(1 - \log w_i)$ are set to zero. By denoting

$$
\mathcal{I} = \{ i \in \{1, \ldots, n\} \;:\; w_i > 0 \}\,,
\tag{2.16}
$$

we have

$$\sum_{i=1}^{n} w_i \log x_i = \sum_{i\in\mathcal{I}} w_i \log x_i,\ \sum_{i=1}^{n} w_i \log s_i = \sum_{i\in\mathcal{I}} w_i \log s_i,\ \sum_{i=1}^{n} w_i \log w_i = \sum_{i\in\mathcal{I}} w_i \log w_i.$$

We call (2.14) a Quadratic Programming and Weighted Centering (QPWC) problem. For $M = 0$ it reduces to the recently introduced notion of a Linear Programming and Weighted Centering (LPWC) problem [2]. We say that x is strictly feasible (or an interior point) for (2.14) if $Ax = b$ and $x > 0$. Similarly (u, s, y) is called strictly feasible (or an interior point) for (2.15) if $s = Mu - A^T y + f > 0$. We note that in the special case $M = 0$, Anstreicher [2] considers a more general notion of strict feasibility requiring only that $x_i > 0,\ s_i > 0$ for all $i \in \mathcal{I}$. Since in this paper we are concerned with interior-point methods, we will only consider the more restrictive notion of strict feasibility defined above.

THEOREM 2.1.

1. *(weak duality) If x is feasible for (2.14) and (u, s, y) is feasible for (2.15) then $\varphi(x) \geq \psi(u, s, y)$;*

2. *(optimality conditions) x is an optimal solution for (2.14) and (u, s, y) is an optimal solution for (2.15) if and only if*

$$xs = w, \quad Mx = Mu. \tag{2.17}$$

Moreover, in this case (x, s, y) is also an optimal solution for (2.15), and we have $\varphi(x) = \psi(u, s, y) = \psi(x, s, y)$;

3. *(strong duality) If (2.14) and (2.15) are strictly feasible then they have optimal solutions x^*, (x^*, s^*, y^*) with $\varphi(x^*) = \psi(x^*, s^*, y^*)$.*

Proof. Let x be feasible for (2.14) and (u, s, y) be feasible for (2.15). If $x_i = 0$ for some $i \in \mathcal{I}$ then $\varphi(x) = +\infty$, and if $s_i = 0$ for some $i \in \mathcal{I}$ then $\psi(u, s, y) = -\infty$. Therefore in what follows, we can assume without loss of generality that $x_i > 0$ and $s_i > 0$ for all $i \in \mathcal{I}$. Using the feasibility assumption and the inequality

$$\nu - \sigma \log \nu \geq \sigma - \sigma \log \sigma, \tag{2.18}$$

we deduce that

$$\begin{aligned}
&\varphi(x) - \psi(u, s, y) \\
&= \frac{1}{2} x^T M x + \frac{1}{2} u^T M u + f^T x - b^T y - \sum_{i=1}^{n} w_i \log(x_i s_i) - \sum_{i=1}^{n} w_i (1 - \log w_i) \\
&\geq \frac{1}{2} x^T M x + \frac{1}{2} u^T M u + f^T x - b^T y - \sum_{i=1}^{n} x_i s_i \\
&= \frac{1}{2} x^T M x + \frac{1}{2} u^T M u + (s - Mu + A^T y)^T x - (Ax)^T y - x^T s \\
&= \frac{1}{2} (x - u)^T M (x - u) \geq 0.
\end{aligned}$$

Since (2.18) holds with equality if and only if $\sigma = \nu$, it follows that $\varphi(x) = \psi(u, s, y)$ if and only if (2.17) is satisfied. Obviously, in this case x is an optimal solution for (2.14), and (u, s, y), (x, s, y) are optimal solutions for (2.15). Assume now that x is an optimal solution for (2.14). Since φ is convex and the constraints are linear, the KKT conditions must hold. Therefore there are vectors $\overline{y}\,,\overline{v}$ such that

$$\overline{v} = Mx + f - \frac{w}{x} - A^T\overline{y} \geq 0.$$

We note that if x is optimal then $x_i > 0$ whenever $w_i > 0$, since otherwise the objective function is $+\infty$. Therefore the ith component of the vector $\frac{w}{x}$ is well defined whenever $w_i > 0$. All the other components of $\frac{w}{x}$ are set to 0 by definition. If we define

$$\overline{s} = Mx + f - A^T\overline{y} = \overline{v} + \frac{w}{x} \geq 0,$$

then $(x, \overline{s}\,, \overline{y}\,)$ is feasible for (2.15), and we have $x\overline{s} = w$. Hence $(x, \overline{s}\,, \overline{y}\,)$ is an optimal solution for (2.15), and (2.17) must hold for any other optimal solution (u, s, y) of (2.15). This finishes the proof of points 1) and 2) of our theorem. In order to complete the proof of 3), we have to show that under the assumption of strict feasibility (2.14) has an optimal solution. Let $\overline{x}$, (u, s, y) be strictly feasible for (2.14) and (2.15) respectively, i.e.,

$$A\overline{x} = b, \quad s = Mu - A^Ty + f, \quad \overline{x} > 0, \quad s > 0,$$

and consider the level set

$$\mathcal{L} = \{x \ : \ Ax = b, \ x \geq 0, \ \varphi(x) \leq \varphi(\overline{x}\,)\}.$$

For any $x \in \mathcal{L}$ we have $x_i > 0$, $\forall i \in \mathcal{I}$, and

$$\begin{aligned}
\varphi(\overline{x}\,) &\geq \frac{1}{2}x^TMx + f^Tx - \sum_{i=1}^n w_i \log x_i \\
&= \frac{1}{2}x^TMx + x^Ts + y^TAx - u^TMx - \sum_{i=1}^n w_i \log x_i \\
&= \frac{1}{2}x^TMx - u^TMx + b^Ty + x^Ts + \sum_{i=1}^n w_i \log x_i\,.
\end{aligned}$$

It follows that

$$\begin{aligned}
\overline{\xi} &:= \varphi(\overline{x}\,) - \frac{1}{2}u^TMu - b^Ty - \sum_{i=1}^n w_i \log s_i \\
&= \frac{1}{2}(x-u)^TM(x-u) + x^Ts - \sum_{i=1}^n w_i \log(x_is_i) \\
&\geq \sum_{i\notin\mathcal{I}} x_is_i + \sum_{i\in\mathcal{I}} x_is_i - w_i\log(x_is_i)\,. \qquad (2.19)
\end{aligned}$$

It is easy to show that for any $i \in \mathcal{I}$ there is $\zeta_i > 0$ such that $t - w_i \log t \geq .5t$, $\forall t \geq \zeta_i$. By denoting

$$\zeta = \max_i \zeta_i, \quad \mathcal{I}_1 = \{i \in \mathcal{I} \ : \ x_is_i \geq \zeta\}, \quad \mathcal{I}_2 = \{i \in \mathcal{I} \ : \ x_is_i < \zeta\}, \quad \overline{\sigma} = \min_i s_i\,,$$

and using (2.18), we deduce that

$$\begin{aligned}\bar{\xi} &\geq \sum_{i \notin \mathcal{I}} x_i s_i + \frac{1}{2} \sum_{i \in \mathcal{I}_1} x_i s_i + \sum_{i \in \mathcal{I}_2} x_i s_i - w_i \log(x_i s_i) \\ &\geq \bar{\sigma} \sum_{i \notin \mathcal{I}} x_i + \frac{\bar{\sigma}}{2} \sum_{i \in \mathcal{I}_1} x_i + \sum_{i \in \mathcal{I}_2} w_i - w_i \log(w_i).\end{aligned}$$

It follows that

$$x_j \leq \min\{\zeta,\ \frac{2}{\bar{\sigma}}(\bar{\xi} - \sum_{i \in \mathcal{I}_2} w_i - w_i \log(w_i))\}, \quad j = 1, \ldots, n\,.$$

Hence, the level set $\mathcal{L}$ is compact. Since φ is convex and bounded below on $\mathcal{L}$, it will attain its minimum on $\mathcal{L}$. ∎

The optimality conditions from the above theorem give rise to a linear wCP of the form (2.4) with

$$P = \begin{pmatrix} A \\ M \end{pmatrix},\ Q = \begin{pmatrix} 0 \\ -I \end{pmatrix},\ R = \begin{pmatrix} 0 \\ -A^T \end{pmatrix},\ a = \begin{pmatrix} b \\ -f \end{pmatrix}.$$

It can be shown that this wCP is monotone. In the particular case when $M = 0$, the wCP is skew-symmetric.

3. Ye's method for solving the Fisher equilibrium problem. In this section we review the interior-point method developed by Ye [25] for the Fisher equilibrium problem. In order to be able to better compare his method with the two methods proposed in this paper we will present Ye's method as a method for solving the linear wCP (2.4) for the case (2.13). For this problem Ye considered a starting point $z^0 = [x^0; s^0; y^0]$ where the first n_c coordinates of x^0 are given by

$$u_i^0 = \frac{1}{n_p} \sum_{k=1}^{n_p} u_{ik}, \quad i = 1, \ldots, n_c \tag{3.1}$$

and the remaining $n_c n_p$ coordinates are equal to $1/n_p$,

$$x_{ij}^0 = \frac{1}{n_p}, \quad i = 1, \ldots, n_c\,,\ j = 1, \ldots, n_p. \tag{3.2}$$

We have clearly $Ax^0 = b$ and $x^0 > 0$. The dual vector $y^0 = [q^0; p^0]$ has components

$$q_i^0 = \frac{\widehat{\beta}}{u_i^0}\,,\ i = 1, \ldots, n_c\,; \quad p_j^0 = 2n_p\widehat{\beta},\ j = 1, \ldots, n_p. \tag{3.3}$$

Finally, the components of the slack vector $s^0 = Ay^0$,

$$s^0 = [v_1^0, ..., v_{n_c}^0, s_{1\,1}^0, \ldots, s_{1\,n_p}^0, s_{2\,1}^0, \ldots, s_{2\,n_p}^0, \ldots, s_{n_c\,1}^0, \ldots, s_{n_c\,n_p}^0]^T,$$

are given by

$$v_i^0 = q_i^0\,,\ i = 1, \ldots, n_c\,; \quad s_{i\,j}^0 = p_j^0 - q_i^0 u_{i\,j}\,,\ i = 1, \ldots, n_c\,,\ j = 1, \ldots, n_p\,. \tag{3.4}$$

It follows that

$$
\begin{aligned}
u_i^0 v_i^0 &= \widehat{\beta}\,,\; i = 1,\ldots,n_c\,, \\
x_{ij}^0 s_{ij}^0 &= 2\widehat{\beta} - \frac{\widehat{\beta}\,u_{ij}}{\sum_{k=1}^{n_p} u_{ik}} \in [\,\widehat{\beta},\; 2\widehat{\beta}\,]\,,\; i = 1,\ldots,n_c\,,\; j = 1,\ldots,n_p\,, \\
\mu_0 &= \frac{x^{0\,T} s^0}{n_c + n_p n_c} = \frac{2\,n_p\,\widehat{\beta}}{n_p + 1}\,.
\end{aligned} \tag{3.5}
$$

Let us denote the feasible set of wCP (2.4) by

$$
\mathcal{F} = \{ z = [x; s; y] \in I\!R^{2n+m} \;:\; Px + Qs + Ry = a,\; x \geq 0,\; s \geq 0 \}\,. \tag{3.6}
$$

Its relative interior

$$
\mathcal{F}^0 = \{ z = [x; s; y] \in I\!R^{2n+m} \;:\; Px + Qs + Ry = a,\; x > 0,\; s > 0 \} \tag{3.7}
$$

is called the set of strictly feasible (or interior) points. The central path of $\mathcal{F}$ is the curve given by the set of all points $[t; z] = [t; x; s; y]$, with $t > 0$, satisfying

$$
\begin{array}{rl}
xs & = te \\
Px + Qs + Ry & = a \\
x > 0,\; s > 0 &
\end{array}\,. \tag{3.8}
$$

If $[t; z]$ is on the central path, then obviously

$$
\mu = \mu(z) = \frac{x^T s}{n} = t\,.
$$

Therefore, a good proximity measure of a point $z = [x; s; y] \in \mathcal{F}$ to the central path (3.8) is given by

$$
\delta_2(z) = \left\| \frac{xs}{\mu(z)} - e \right\|_2\,, \quad \mu(z) = \frac{x^T s}{n}\,. \tag{3.9}
$$

Using this proximity measure we can define the following neighborhood of the central path (3.8):

$$
\mathcal{N}_2(\alpha) = \{ z \in \mathcal{F}^0 \;:\; \delta_2(z) \leq \alpha \}. \tag{3.10}
$$

For the starting point $z^0 = [x^0; s^0; y^0]$ we have

$$
\mu_0^2 \delta_2(z^0)^2 = \widehat{\beta}^2 \left(\frac{(n_p - 3)n_c}{n_p + 1} + \sum_{i=1}^{n_c} \sum_{j=1}^{n_p} \left(\frac{u_{ij}}{\sum_{k=1}^{n_p} u_{ik}} \right)^2 \right) \leq \frac{2(n_p - 1)n_c \widehat{\beta}^2}{n_p + 1}\,.
$$

It follows that

$$
\delta_2(z^0) \leq \frac{\sqrt{(n_p^2 - 1)n_c}}{\sqrt{2}\,n_p} < \frac{\sqrt{n_c}}{\sqrt{2}}\,. \tag{3.11}
$$

In [25] Ye uses the starting point z^0 with

$$
\widehat{\beta} = \frac{n_p + 1}{2n_p}\,\|w\|_\infty\,, \tag{3.12}
$$

which implies

$$\mu_0 = \mu(z^0) = \|w\|_\infty . \tag{3.13}$$

Starting from this point, the potential reduction method from [24, page 106] is used to construct a point $\hat{z} \in \mathcal{N}_2(\alpha)$ for some $\alpha < 1$. This is achieved in $O(n_c)$ iterations. We note that at page 324 of [25] it is claimed that only $O(\log(n_c\, n_p))$ iterations are needed. This is clearly a missprint, as acknowledged by Ye in a private communication. Then, the so-called modified primal-dual path-following algorithm, starting from $\hat{z}$, is used to follow the modified central path

$$\left\{ \begin{array}{rcl} xs & = & \widehat{w}(t) \\ Px + Qs + Ry & = & a \end{array} \right. , \quad \widehat{w}(t) = \max\{te, w\}, \quad t > 0\,. \tag{3.14}$$

Here $x, s \in I\!R^n_{++}$, $w \in I\!R^n_{+}$, with $n = n_c(n_p + 1)$, but only n_c components of w are nonzero. If $t \geq \|w\|_\infty$ then (3.14) reduces to (3.8), and we have $t = \mu = \mu(z) = x^T s/n$. On the other hand for any $[t; z]$ satisfying (3.14) and $t \leq \min\{w_i \ : \ w_i > 0\}$ we have

$$\mu = \frac{e^T w + n_c n_p t}{n_c(n_p+1)} = t + \frac{e^T w - n_c t}{n_c(n_p+1)}\,,$$

so that $\mu \to e^T w/n$ as $t \to 0$.
The l_2-neighborhood of the modified central path (3.14) is defined by

$$\widehat{\mathcal{N}}_2(w,\alpha) = \{[t;z] \in I\!R_{++} \times \mathcal{F}^0 \ : \ \|xs - \widehat{w}(t)\|_2 \leq \alpha t\}\,.$$

The point $\hat{z}$ produced by the potential reduction method satisfies $[\hat{t}; \hat{z}] \in \widehat{\mathcal{N}}_2(w,\alpha)$, with $\hat{t} = \mu(\hat{z}) \geq \|w\|_\infty$ and $\alpha < 1$. At each iteration, the so-called modified primal-dual path-following algorithm from [25] starts with a point $[t; z] \in \widehat{\mathcal{N}}_2(w,\alpha)$ and produces a point $[t_+; z^+] \in \widehat{\mathcal{N}}_2(w,\alpha)$ with $t_+ = (1 - \alpha/\sqrt{n}\,)\, t$. Let us denote the set of ε-approximate solutions of wCP (2.4) by

$$\mathcal{S}_\varepsilon = \{z = [x; s; y] \in \mathcal{F} \ : \ \|xs - w\|_2 \leq \varepsilon\}\,. \tag{3.15}$$

If $[t; z] \in \widehat{\mathcal{N}}_2(w,\alpha)$, then

$$\|xs - w\|_2 \leq \|xs - \widehat{w}(t)\|_2 + \|\widehat{w}(t) - w\|_2 \leq \alpha t + \|te\|_2 \leq \left(\alpha + \sqrt{n}\,\right) t. \tag{3.16}$$

Therefore, the modified primal-dual path-following algorithm from [25] produces an ε-approximate solution in at most $O\left(\sqrt{n}\, \log\left(\|w\|_\infty \sqrt{n}/\varepsilon\right)\right)$ iterations. If we add the cost of the potential reduction method, it follows that an ε-approximate solution for the Fisher problem is obtained in at most

$$O(\sqrt{n_c n_p}\, \log\left((n_c + n_p)\, \|w\|_\infty/\varepsilon\right) + O(n_c) \tag{3.17}$$

iterations. Since at each iteration we have $t_+ = (1 - \alpha/\sqrt{n}\,)\, t$, the algorithm belongs to the class of short-step methods, and therefore its practical performance is close to the worst case bounds reflected in the above iteration complexity result. In what follows we show that it is not necessary to first find a point in $\mathcal{N}_2(\alpha)$ with $\alpha < 1$. Instead, we start with z^0, define a new (smooth) central path emanating from it, and consider a new neighborhood of this central path. We will show that our algorithms find an ε-approximate solution for the Fisher problem in at most $O(\sqrt{n_c n_p}\, \log\left((n_c + n_p)\, \|w\|_\infty/\varepsilon\right)$ iterations. Our algorithms are long step algorithms and therefore their practical performance are much better than indicated by this iteration complexity result. Moreover our algorithms work for any monotone linear wCP.

4. Two interior-point methods for solving monotone linear wCPs. In this section we present a long step path-following method and a predictor-corrector method for solving a general linear wCP of the form (2.4) that is monotone in the sense of (2.5). The long step method may be interpreted as a generalization of McShane's largest step algorithm [15]. The name "largest step algorithm" was given by Gonzaga [11] in the case of monotone complementarity problems (see also [4, 5]). The predictor-corrector method can be considered as a generalization of the Mizuno-Todd-Ye method [16], which was the first algorithm for solving linear programming problems having both polynomial complexity and superlinear convergence. The main difference between the two algorithms to be introduced in this section is that the long step path-following algorithm uses only one matrix factorization per iteration while the predictor-corrector method uses two factorizations.

4.1. A long step path-following method. Let us consider the notations from (2.4), (3.6) and (3.7). Given a strictly feasible starting point $z^0 = [x^0, s^0; y^0] \in \mathcal{F}^0$, we denote

$$t_0 = \mu(z^0),\ c = x^0 s^0,\ \gamma = \frac{\min\ c}{t_0},\ \ w(t) = (1 - t/t_0)w + (t/t_0)c,\ t \in (0, t_0], \tag{4.1}$$

where $\min\ c = \min\{c_j\ :\ j = 1 \ldots, n\}$. We define the central path of wCP (2.4) emanating from z^0 as the set of all points $[t; z] = [t; x; s; y]$, with $t \in (0, t_0]$, satisfying

$$\begin{array}{rl} xs & = w(t) \\ Px + Qs + Ry & = a \\ x > 0,\ s > 0 & \end{array} . \tag{4.2}$$

By construction $[t^0; z^0]$ belongs to this path. We note that for the Fisher problem we have $t_0 = \|\, w \,\|_\infty$ and $\gamma > 1/2$, as indicated by (3.5), (3.12) and (3.13).

Before describing our interior-point methods let us make some remarks about the modified central path (3.14) used by Ye [25] and the central path (4.2). First, as noted above the starting point belongs by construction to the central path (4.2). This is not the case with the modified central path (3.14). In fact, as mentioned in Section 3, a potential reduction method, which can be considered a "Phase I" algorithm, is used to produce a point in a certain neighborhood of the modified central path. The difference between the two central paths consists in the fact that right-hand side of the first equation in (3.14) is given by $\widehat{w}(t)$, while the right-hand side of the corresponding equation in (4.2) is $w(t)$. Since the wCP is monotone, strictly feasible, and $\widehat{w}(t) > 0,\quad w(t) > 0,\quad \forall t \in (0,\ t_0]$, it follows that both (3.14) and (4.2) have unique solutions for any $t \in (0,\ t_0]$. However, while $w(t)$ is smooth on $(0,\ t_0]$, $\widehat{w}(t)$ is not smooth at the points in the set $\{w_i\ :\ w_i > 0\}$.

Given a parameter α such that

$$0 \le \frac{\gamma}{3} \le \alpha \le \frac{2\gamma}{3}\,, \tag{4.3}$$

we define the following neighborhood of the above central path:

$$\mathcal{N}_2(w, c, \alpha) = \{[t; z] = [t; x; s; y] \in (0, t_0] \times \mathcal{F}^0\ :\ \|xs - w(t)\| \le \alpha t\}.$$

For our starting point we have $x^0 s^0 = c = w(t_0)$, so that $[t_0, z^0] \in \mathcal{N}_2(w, c, \alpha)$.

At a typical iteration of our algorithm we have a point $[t; z] \in \mathcal{N}_2(w, c, \alpha)$, for some $t \le t_0$. Since $c \ge \gamma t_0\, e$, it follows that

$$xs \ge w(t) - \alpha te \ge (t/t_0)c - \alpha te \ge (\gamma - \alpha)te = \beta te, \quad \beta = \gamma - \alpha \ge \frac{\gamma}{3} \ge \frac{\alpha}{2}\,\cdot \tag{4.4}$$

Let us denote

$$t(\theta) = (1-\theta)t,\ z(\theta) = [x(\theta); s(\theta; y(\theta))] = [x + u(\theta); s + v(\theta); y + d(\theta)]\,, \tag{4.5}$$

where $u(\theta), v(\theta), d(\theta)$ are the solutions of the following linear system

$$\left\{ \begin{array}{rcl} su(\theta) + xv(\theta) & = & w(t(\theta)) - xs \\ Pu(\theta) + Qv(\theta) + Rd(\theta) & = & 0 \end{array} \right. . \tag{4.6}$$

Using the stepsize

$$\theta^+ = \max\{\widehat{\theta} \in [0,1]\ :\ [t(\theta); z(\theta)] \in \mathcal{N}_2(w, a, \alpha),\ \forall \theta \in [0, \widehat{\theta}\,]\}, \tag{4.7}$$

we obtain the new point

$$[t_+; z^+] := [t(\theta^+); z(\theta^+)] \in \mathcal{N}_2(w, c, \alpha)\,, \tag{4.8}$$

and we can begin a new iteration.

In order to efficiently compute the stepsize defined in (4.7) we first solve the following two linear systems

$$\left\{ \begin{array}{rcl} s\check{u} + x\check{v} & = & w(t) - xs \\ P\check{u} + Q\check{v} + R\check{d} & = & 0 \end{array} \right. ,\quad \left\{ \begin{array}{rcl} s\hat{u} + x\hat{v} & = & w - xs \\ P\hat{u} + Q\hat{v} + R\hat{d} & = & 0 \end{array} \right. . \tag{4.9}$$

The solution of the linear system (4.6) can be written under the form

$$u(\theta) = (1-\theta)\check{u} + \theta\hat{u},\ v(\theta) = (1-\theta)\check{v} + \theta\hat{v},\ d(\theta) = (1-\theta)\check{d} + \theta\hat{d}. \tag{4.10}$$

From (4.5) and (4.6) we have

$$x(\theta)s(\theta) = w(t(\theta)) + u(\theta)s(\theta) = w(t) + \theta(t/t_0)(w - c) + u(\theta)s(\theta)\,. \tag{4.11}$$

By denoting

$$\psi(\theta) = \|x(\theta)s(\theta) - w(t(\theta))\|_2^2 - \alpha^2 t(\theta)^2,$$

we deduce that

$$[t(\theta); z(\theta)] \in \mathcal{N}_2(w, c, \alpha) \quad \text{if and only if} \quad \psi(\theta) \le 0.$$

From (4.5),(4.6),(4.9),(4.10),(4.11) we have:

$$\begin{aligned} \psi(\theta) &= \|u(\theta)v(\theta)\|_2^2 - (1-\theta)^2 t^2 \alpha^2 \\ &= \|(1-\theta)^2 \check{u}\check{v} + \theta(1-\theta)\left(\check{u}\hat{v} + \check{v}\hat{u}\right) + \theta^2 \hat{u}\hat{v}\|_2^2 - (1-\theta)^2 t^2 \alpha^2. \end{aligned}$$

Finally, by developing the square of the norm of the sum of three vectors we obtain an explicit form of $\psi(\theta)$ as a quartic in θ:

$$\begin{aligned} \psi(\theta) = {} & (1-\theta)^4 \parallel \check{u}\check{v} \parallel_2^2 + 2\theta(1-\theta)^3 e^T(\check{u}^2\check{v}\hat{v} + \check{u}\check{v}^2\hat{u}) \\ & + \theta^2(1-\theta)^2 \left(\parallel \check{u}\hat{v} + \check{v}\hat{u} \parallel_2^2 + 2e^T(\check{u}\check{v}\hat{u}\hat{v}) \right) \\ & + 2\theta^3(1-\theta)e^T(\check{u}\hat{u}^2\hat{v} + \check{v}\hat{v}\hat{u}^2) + \theta^4 \parallel \hat{u}\hat{v} \parallel_2^2 - (1-\theta)^2 t^2 \alpha^2. \end{aligned} \tag{4.12}$$

In what follows we will use the following technical result, which is well known in the interior-point literature (see for example [20, Lemma 3.1]).

LEMMA 4.1. *If (2.5) is satisfied then the linear system*

$$\begin{cases} su + xv &= g \\ Pu + Qv + Rd &= 0 \end{cases}$$

has a unique solution for any $x \in I\!R^n_{++}, s \in I\!R^n_{++}, g \in I\!R^n$*, and the following inequality holds*

$$\| uv \|_2 \leq \frac{1}{\sqrt{8}} \left\| (xs)^{-1/2} g \right\|_2^2 .$$

At the current iteration we have $[t; z] \in \mathcal{N}_2(w, c, \alpha)$, and by using (4.4) we obtain

$$\| \check{u}\check{v} \|_2 \leq \frac{1}{\beta t \sqrt{8}} \| xs - w(t) \|_2^2 \leq \frac{\alpha^2 t}{\beta \sqrt{8}} \leq \frac{2\alpha t}{\sqrt{8}} < \alpha t \,,$$

so that $\psi(0) = \| \check{u}\check{v} \|_2^2 - t^2\alpha^2 < 0$ and $\psi(1) = \| \hat{u}\hat{v} \|_2^2 \geq 0$. Therefore the quartic equation $\psi(\theta) = 0$ has at least one root in the interval $(0, 1]$ and our steplength can be computed as

(4.13) $\theta^+ =$ the smallest root of the quation $\psi(\theta) = 0$ from the interval $(0, 1]$.

Our algorithm can be therefore formally defined as

Algorithm 1 (*Largest Step*)
Given a starting point $z^0 = [x^0; s^0; y^0] \in \mathcal{F}^0$:
  Consider the notation from (4.1);
  Choose a parameter α satisfying (4.3);
  Set $k \leftarrow 0$;
  repeat
    Set $z = [x; s; y] \leftarrow z^k$, $t \leftarrow t_k$;
    Solve the linear systems (4.9);
    Compute steplength θ^+ from (4.12) and (4.13);
    Compute t_+, z^+ from (4.5), (4.8), and (4.10);
    Set $\theta_k \leftarrow \theta^+$, $z^{k+1} \leftarrow z^+$, $t_{k+1} \leftarrow t_+$;
    Set $k \leftarrow k + 1$.
  continue

We note that the linear systems from (4.9) have the same matrix, so that the algorithm requires only one matrix factorization per iteration.

In the remainder of this section we will establish the computational complexity of Algorithm 1. For convenience we introduce the following notation:

$$\rho = 1 + \frac{\|c - w\|_2}{t_0} \,. \tag{4.14}$$

THEOREM 4.2. *If wCP (2.4) is monotone, then Algorithm 1 is well defined and generates an iteration sequence satisfying the following properties:*

$$[t_k; z^k] \in \mathcal{N}_2(w, c, \alpha) \,;$$
$$t_{k+1} = (1 - \theta_k) t_k \,;$$
$$\theta_k \geq \frac{\alpha}{9\rho} \,.$$

Proof. The first two properties hold from construction. In order to prove the last property we first note that if $[t(\theta); z(\theta)]$ is given by (4.5) then according to (4.4), (4.11), and Lemma 4.1 we have

$$\|x(\theta)s(\theta) - w(t(\theta))\|_2 = \|u(\theta)s(\theta)\|_2 \le \frac{\|w(t(\theta)) - xs\|_2^2}{\sqrt{8}\min xs} \le \frac{\|w(t(\theta)) - xs\|_2^2}{\sqrt{2}\alpha t}\,.$$

In order to majorize the last term above we note that

$$\|w(t(\theta)) - xs\|_2 = \|w(t) + \theta(t/t_0)(w - c) - xs\|_2 < t(\alpha + \theta\rho).$$

If $0 < \theta \le \alpha/(9\rho) < 1/9$ then

$$\frac{\|x(\theta)s(\theta) - w(t(\theta))\|_2}{\alpha t(\theta)} \le \frac{(\alpha + \theta\rho)^2}{\sqrt{2}\alpha^2(1-\theta)} \le \frac{(\alpha + \alpha/9)^2}{\sqrt{2}\alpha^2(1 - 1/9)} < .99\,.$$

It follows that $[t(\theta); z(\theta)] \in \mathcal{N}_2(w, c, \alpha)$, $\forall\theta \in (0,\, \alpha/(9\rho)]$, which implies $\theta^+ \ge \alpha/(9\rho)$, where θ^+ is the stepsize (4.7) used by our algorithm. ∎

Corollary 4.3. *If wCP (2.4) is monotone, then Algorithm 1 finds an ε-approximate solution for this problem (i.e., a point $z \in \mathcal{S}_\varepsilon$, where $\mathcal{S}_\varepsilon$ is defined in (3.15)) in at most*

$$O\left(\frac{x^{0\,T}s^0/n + \|\, x^0s^0 - w \,\|_2}{\min x^0 s^0} \log \frac{x^{0\,T}s^0/n + \|\, x^0s^0 - w \,\|_2}{\varepsilon}\right)$$

iterations.

Proof. From Theorem 4.2, (4.1), (4.3), and (4.14) we have

$$\begin{aligned}\|\, x^k s^k - w \,\|_2 &\le \|\, x^k s^k - w(t_k) \,\|_2 + \|\, w(t_k) - w \,\|_2 \le t_k(\alpha + \|\, w - c \,\|_2 / t_0) \\ &< t_k\rho \le \left(1 - \frac{\alpha}{9\rho}\right)^k t_0\rho \le \left(1 - \frac{\gamma}{27\rho}\right)^k t_0\rho \le \left(1 - \frac{\min x^0s^0}{27 t_0\rho}\right)^k t_0\rho\,,\end{aligned}$$

and the complexity result follows from a standard argument. ∎

Corollary 4.4. *When applied to the wCP generated by the Fisher problem (see (2.13)) with starting point z^0 given by (3.1)-(3.4) and (3.12), Algorithm 1 finds an ε-approximate solution for this problem in at most*

$$O\left(\sqrt{n_c n_p}\, \log \frac{(n_c + n_p)\,\|w\|_\infty}{\varepsilon}\right)$$

iterations.

Proof. By writing

$$\|c - w\|_2^2 = \|w\|_2^2 + \|c\|_2^2 - 2w^Tc = \|w\|_2^2 + \|c\|_2^2 - \frac{n_p + 1}{n_p}\|w\|_\infty\|w\|_1\,,$$

and using the relation $\|w\|_\infty^2 \le \|w\|_2^2 \le \|w\|_\infty\|w\|_1$ we deduce that

$$\|c - w\|_2^2 \le \|c\|_2^2 - \frac{\|w\|_2^2}{n_p} \le \|c\|_2^2 - \frac{\|w\|_\infty^2}{n_p}\,.$$

On the other hand we have

$$\begin{aligned}\|c\|_2^2 &= n_c\widehat{\beta}^2 + \sum_{i,j}\left(s^0_{ij}u^0_{ij}\right)^2 = \widehat{\beta}^2\left(n_c + \sum_{i=1}^{n_c}\sum_{j=1}^{n_p}\left(2 - \frac{u_{ij}}{\sum_{k=1}^{n_p} u_{ik}}\right)^2\right)\\ &= \widehat{\beta}^2\left(n_c + \sum_{i=1}^{n_c}\sum_{j=1}^{n_p}\left(4 - 2\frac{u_{ij}}{\sum_{k=1}^{n_p} u_{ik}} + \frac{u_{ij}^2}{\left(\sum_{k=1}^{n_p} u_{ik}\right)^2}\right)\right)\\ &= \widehat{\beta}^2\left(4n_cn_p - n_c + \sum_{i=1}^{n_c}\frac{\sum_{j=1}^{n_p} u_{ij}^2}{\left(\sum_{k=1}^{n_p} u_{ik}\right)^2}\right)\\ &= \left(\frac{n_p+1}{2n_p}\right)^2\left(4n_cn_p - n_c + \sum_{i=1}^{n_c}\frac{\sum_{j=1}^{n_p} u_{ij}^2}{\left(\sum_{k=1}^{n_p} u_{ik}\right)^2}\right)\|\, w\,\|_\infty^2 \,.\end{aligned}$$

Since $\sum_{j=1}^{n_p} u_{ij}^2 \le \left(\sum_{j=1}^{n_p} u_{ij}\right)^2$ we deduce that for all $n_c, n_p \ge 1$ there holds

$$\|c-w\| \le \|c\| \le \frac{n_p+1}{n_p}\sqrt{n_cn_p}\,\|\, w\,\|_\infty \le 2\sqrt{n_cn_p}\,\|\, w\,\|_\infty \,.$$

Finally, from (3.5), (3.12), (3.13) we have

$$\min c \ge \|\, w\,\|_\infty/2\,, \quad t_0 = {x^0}^T s^0/n = \|\, w\,\|_\infty\,,$$

and the desired complexity result follows Corollary 4.3 . ∎

4.2. A predictor-corrector method. As mentioned in the previous subsection, Algorithm 1 requires only one matrix factorization per iteration. At a cost of two matrix factorizations per iteration, we can generalize the Mizuno-Todd-Ye predictor-corrector algorithm to our setting. The purpose of the predictor is to improve as much as possible the optimality measure t while not departing too much from the central path (4.2). The algorithm depends on two parameters α and $\overline{\alpha}$, such that

$$\frac{\gamma}{3} \le \alpha < \overline{\alpha} \le \frac{2\gamma}{3}\,, \quad \frac{4\alpha}{3} \le \overline{\alpha} \le \sqrt{2}\alpha\,. \tag{4.15}$$

The above relations are satisfied for example by $\alpha = \sqrt{2}\gamma/3$, and $\overline{\alpha} = 2\gamma/3$.

4.2.1. The predictor. At the beginning of the predictor step we are given a point $[t; z] \in \mathcal{N}(w, c, \alpha)$ and we compute the predictor direction $[u; v; d]$ as the solution of the linear system

$$\left\{\begin{array}{rcl} su + xv &=& w - xs\\ Pu + Qv + Rd &=& 0\end{array}\right. . \tag{4.16}$$

We define

$$\begin{gathered}x(\theta) = x + \theta u,\; s(\theta) = s + \theta v,\; y(\theta) = y + \theta d,\; t(\theta) = (1-\theta)t,\\ z(\theta) = [\,x(\theta)\,;\, s(\theta)\,;\, y(\theta)\,]\,.\end{gathered} \tag{4.17}$$

The stepsize along this direction is taken as

$$\overline{\theta} = \max\{\widehat{\theta} \in [0,1] \;:\; [\,t(\theta)\,;\, z(\theta)\,] \in \mathcal{N}_2(w, c, \overline{\alpha}\,),\ \forall \theta \in [0, \widehat{\theta}\,]\}, \tag{4.18}$$

We have

$$x(\theta)s(\theta) = (x+\theta u)(s+\theta v) = (1-\theta)xs + \theta w + \theta^2 uv\,,$$
$$w(t(\theta)) = w + (1-\theta)t(c-w)\,,$$
$$x(\theta)s(\theta) - w(t(\theta)) = (1-\theta)(xs - w(t)) + \theta^2 uv\,.$$

Therefore the inequality

$$\|x(\theta)s(\theta) - w(t(\theta))\|_2 \le \overline{\alpha}\, t(\theta)$$

can be written as

$$\beta_0(1-\theta)^2 + 2\beta_1(1-\theta)\theta^2 + \beta_2\theta^4 \le 0, \tag{4.19}$$

where

$$\beta_0 = \frac{\|xs - w(t)\|_2^2}{t^2} - \overline{\alpha}^{\,2},\ \beta_1 = \frac{(uv)^T(xs - w(t))}{t^2},\ \beta_2 = \frac{\|uv\|_2^2}{t^2}\,. \tag{4.20}$$

Since $[t; z] \in \mathcal{N}(w, c, \alpha)$ we have $\beta_0 \le \alpha^2 - \overline{\alpha}^{\,2} < 0$. If $uv = 0$ then $\beta_1 = \beta_2 = 0$, so that in this case we have $\overline{\theta} = 1$. Therefore, in what follows we assume $\beta_2 > 0$. By using the substitution $\phi = (1-\theta)/\theta^2$, (4.19) can be reduced to the following quadratic inequality

$$\beta_0 + 2\beta_1\phi + \beta_2\phi^2 \le 0. \tag{4.21}$$

The left-hand-side of the above inequality is strictly negative for $\phi = 0$ (since $\beta_0 < 0$), and strictly positive for ϕ sufficiently large (since $\beta_2 > 0$). Therefore the above inequality holds for all $\phi \le \overline{\phi}$, where

$$\overline{\phi} = \frac{-\beta_0}{\beta_1 + \sqrt{\beta_1^2 - \beta_0\beta_2}}\,. \tag{4.22}$$

It follows that stepsize θ^+ defined in (4.18) is explicitly given by

$$\overline{\theta} = \begin{cases} 1 & \text{if } uv = 0 \\[2ex] \dfrac{2\overline{\phi}}{\overline{\phi} + \sqrt{\overline{\phi} + \overline{\phi}^{\,2}}} & \text{if } uv \neq 0 \end{cases}\,. \tag{4.23}$$

Having computed this steplength we obtain the predicted point

$$[\,\bar{t}\,;\,\bar{z}\,] = \left[\,t(\overline{\theta}\,)\,;\,z(\overline{\theta}\,)\,\right] = [\,(1-\overline{\theta}\,)t\,;\,x(\overline{\theta}\,)\,;\,s(\overline{\theta}\,)\,;\,y(\overline{\theta}\,)\,] \in \mathcal{N}_2(w, c, \overline{\alpha}\,). \tag{4.24}$$

4.2.2. The corrector. A corrector steps usually follows a predictor step. It starts with a point $[\,t\,;\,z\,] \in \mathcal{N}(w, c, \overline{\alpha}\,)$ and produces a point $[\,t\,;\,z^+\,] \in \mathcal{N}(w, c, \alpha)$. Note that the measure of optimality t remains unchanged, but the measure of proximity to the path (4.2) is improved, wherefrom the name corrector. The direction of the corrector is computed as the solution of the linear system

$$\begin{cases} su + xv & = & w(t) - xs \\ Pu + Qv + Rd & = & 0 \end{cases}\,. \tag{4.25}$$

By taking a unit step along this direction we obtain the points

$$t^+ = t, \quad x^+ = x + u, \quad s^+ = s + v, \quad y^+ = y + d, \quad z^+ = [\, x^+\, ;\, s^+\, ;\, y^+\,]. \tag{4.26}$$

Proceeding as in (4.4) with $\overline{\alpha}$ instead of α and using (4.15) we deduce that $xs \geq (\gamma - \overline{\alpha}\,)te \geq \overline{\alpha}\,/2$. According to Lemma 4.1 and (4.15) we have

$$\left\| \, x^+ s^+ - w(t) \, \right\|_2 = \| \, uv \, \|_2 \leq \frac{\| \, w(t) - xs \, \|_2^2}{\min xs\sqrt{8}} \leq \frac{\overline{\alpha}^{\,2} t}{\overline{\alpha}\,\sqrt{2}} \leq \alpha t. \tag{4.27}$$

Hence $[\, t^+\, ;\, z^+\,] \in \mathcal{N}(w, c, \alpha)$.

Algorithm 2 (*Predictor-Corrector*)
Given a starting point $z^0 = [x^0; s^0; y^0] \in \mathcal{F}^0$:
Consider the notation from (4.1);
Choose parameters α and $\overline{\alpha}$ satisfying (4.15);
Set $k \leftarrow 0$;
repeat
Predictor
Set $z = [x; s; y] \leftarrow z^k$, $t \leftarrow t_k$;
Solve the linear system (4.16);
Compute steplength $\overline{\theta}$ from (4.20), (4.22), and (4.23);
Compute $\overline{t}\,, \overline{z}$ from (4.17) and (4.24);
Set $\overline{\theta}_k \leftarrow \overline{\theta}$, $\overline{z}^{\,k} \leftarrow \overline{z}$, $t_{k+1} \leftarrow \overline{t}$;
Corrector
Set $z = [x; s; y] \leftarrow \overline{z}^{\,k}$, $t \leftarrow t_{k+1}$;
Solve the linear systems (4.25);
Compute z^+ from (4.26);
Set $z^{k+1} \leftarrow z^+$;
Set $k \leftarrow k + 1$.
continue

THEOREM 4.5. *If wCP (2.12) is monotone, then Algorithm 2 is well defined and generates an iteration sequence satisfying the following properties*

$$[t_k; z^k] \in \mathcal{N}_2(w, c, \alpha)\ ,\ [t_{k+1}; \overline{z}^{\,k}] \in \mathcal{N}_2(w, c, \overline{\alpha}\,)\,;$$
$$t_{k+1} = (1 - \overline{\theta}_{\,k})t_k\,;$$
$$\overline{\theta}_{\,k} \geq \frac{2\alpha}{3\rho}\,.$$

Proof. The first two properties have already been proved. In order to find a lower bound for $\overline{\theta}$ we note that

$$|\beta_1| \leq \frac{\|xs - w(t)\|_2 \|uv\|_2}{t^2} = \sqrt{\beta_0 + \overline{\alpha}^{\,2}}\sqrt{\beta_2}, \quad 0 \leq \beta_0 + \overline{\alpha}^{\,2} \leq \alpha^2,$$

and therefore

$$\begin{aligned} \overline{\phi} &\geq \frac{-\beta_0}{|\beta_1| + \sqrt{\beta_1^2 - \beta_0\beta_2}} \geq \frac{-\beta_0}{\left(\sqrt{\beta_0 + \overline{\alpha}^{\,2}} + \overline{\alpha}\right)\sqrt{\beta_2}} \\ &\geq \frac{-\beta_0}{(\alpha + \overline{\alpha}\,)\sqrt{\beta_2}} \geq \frac{\overline{\alpha}\, - \alpha}{\sqrt{\beta_2}}\,. \end{aligned}$$

On the other hand we have

$$\|xs - w\|_2 \leq \|xs - w(t)\|_2 + \|w(t) - w\|_2 \leq (\alpha + \|c - w\|_2)t \leq \rho t\,,$$
$$\|uv\|_2 \leq \frac{\|xs - w\|_2^2}{\sqrt{8}\min xs} < \frac{\rho^2 t}{\alpha\sqrt{2}}\,,$$

which leads to the lower bound

$$\overline{\phi} \geq \frac{\alpha(\overline{\alpha} - \alpha)\sqrt{2}}{\rho^2} \geq \frac{\alpha^2\sqrt{2}}{3\rho^2} > \frac{4\alpha^2}{9\rho^2} =: \widehat{\phi}\,. \tag{4.28}$$

Since $\alpha < 2\gamma/3 \leq 2/3$, and $\rho > 1$, we have $\widehat{\phi} < 1/5$. Using (4.23) and the fact that

$$\frac{2\phi}{\phi + \sqrt{\phi + \phi^2}} \geq \sqrt{\phi}, \quad \forall \phi \in [0, .5],$$

we obtain

$$\overline{\theta} > \frac{2\widehat{\phi}}{\widehat{\phi} + \sqrt{\widehat{\phi} + \widehat{\phi}^2}} \geq \sqrt{\widehat{\phi}} = \frac{2\alpha}{3\rho}\,.$$

∎

The following corollaries are easily proved.

COROLLARY 4.6. *If wCP (2.4) is monotone, then Algorithm 2 finds an ε-approximate solution for this problem (i.e., a point $z \in \mathcal{S}_\varepsilon$, where $\mathcal{S}_\varepsilon$ is defined in (3.15)) in at most*

$$O\left(\frac{x^{0\,T} s^0/n + \left\| x^0 s^0 - w \right\|_2}{\min x^0 s^0} \log \frac{x^{0\,T} s^0/n + \left\| x^0 s^0 - w \right\|_2}{\varepsilon}\right)$$

iterations.

COROLLARY 4.7. *When applied to the wCP generated by the Fisher problem (see (2.13)) with starting point z^0 given by (3.1)-(3.4) and (3.12), Algorithm 2 finds an ε-approximate solution for this problem in at most*

$$O\left(\sqrt{n_c n_p}\, \log \frac{(n_c + n_p)\,\|w\|_\infty}{\varepsilon}\right)$$

iterations.

4.3. Comparison between the two interior-point methods. In what follows we make some remarks about the similarities and the differences between the long step method (Algorithm 1) and the predictor-corrector method (Algorithm 2).

First, we note that if the optimal stepsize θ^+ as defined by (4.7) was known, then the new point $z^+ = z(\theta^+)$ could be obtained by solving the linear system (4.6) with $\theta = \theta^+$. Definition (4.7), while motivating the name "largest step method" given to Algorithm 1, does not lend itself to a direct computation of θ^+. In Algorithm 1 we compute θ^+ as the smallest root of the quartic (4.12) in the interval $(0, 1]$, where $\check{u}, \check{v}, \hat{u}, \hat{v}$ are obtained by solving the linear systems (4.9). The first system in (4.9) gives the so-called centering direction (the Newton direction for the equations defining the central path (4.2)), while the second system in (4.9) gives the so-called affine scaling direction (the Newton direction for the equations defining the wCP (2.4)). Since both systems have the same matrix, only one matrix factorization is needed.

Algorithm 2 also uses the centering direction and the affine scaling direction but computed at different points. The predictor uses the affine scaling direction at the current point and obtains the predicted point by taking the largest stepsize on this direction that keeps the point in the larger neighborhood. The corrector uses the centering direction computed at the predicted point and obtains the new iterate by taking a unit stepsize on this direction. It is shown that the corrected point belongs to the original neighborhood. Since the affine scaling direction and the centering directions are computed at different points two matrix factorizations are needed at each iteration.

5. Conclusions. In this paper we have introduced the notion of a weighted complementarity problem (wCP) in a general setting. We have shown that the Fisher market equilibrium problem can be formulated as a skew-symmetric linear wCP. We have also shown that the notion of a Linear Programming and Weighted Centering (LPWC) problem recently introduced by Anstreicher [2] reduces to a skew-symmetric linear wCP. The more general notion of a Quadratic Programming and Weighted Centering (QPWC) problem, introduced in the present paper, reduces to a monotone linear wCP. We have proposed two interior-point methods for solving general monotone linear wCPs and have established their computational complexity. The first method generalizes the largest step interior-point method of McShane [15], while the second method generalizes the Mizuno-Todd-Ye predictor corrector method [16]. If the weight vector w is equal to 0, and if the corresponding problems have a strict complementarity solution, it is known that the first method is superlinearly convergent, while the second method is quadratically convergent. These asymptotic convergence results hold also for nonzero weight vectors. This can be shown by appropriately modifying the arguments from [26, 15, 4]. A rigorous proof, in a much more general setting, will be given in a subsequent paper [21]. When applied to the wCP generated by the Fisher equilibrium problem, both algorithms have the same iteration complexity as the one obtained by Ye [25]. However our algorithms work directly with the starting point proposed in [25], without having to use another method to obtain a better centered starting point. Moreover, the central path followed by our algorithm is smooth, while the central path proposed in [25] is not.

REFERENCES

[1] K. M. Anstreicher. Volumetric path following algorithms for linear programming. *Math. Programming*, 76(1, Ser. B):245–263, 1997. Interior point methods in theory and practice (Iowa City, IA, 1994).

[2] K.M. Anstreicher. Interior-point algorithms for a generalization of linear programming and weighted centering. Technical report, Dept. of Management Sciences, University of Iowa, February 2011. www.optimization-online.org/DB_HTML/2011/02/2937.html.

[3] D. S. Atkinson and P. M. Vaidya. A scaling technique for finding the weighted analytic center of a polytope. *Math. Programming*, 57(2, Ser. B):163–192, 1992.

[4] J. F. Bonnans and C. C. Gonzaga. Convergence of interior point algorithms for the monotone linear complementarity problem. *Mathematics of Operations Research*, 21:1–25, 1996.

[5] J. F. Bonnans and F. A. Potra. On the convergence of the iteration sequence of infeasible path following algorithms for linear complementarity problems. *Mathematics of Operations Research*, 22(2):378–407, 1997.

[6] E. Eisenberg and D. Gale. Consensus of subjective probabilities: the pari-mutuel method. *Ann. Math. Statist.*, 30:165–168, 1959.

[7] J. Faraut and A. Korányi. *Analysis on symmetric cones*. Oxford Mathematical Monographs. The Clarendon Press Oxford University Press, New York, 1994. , Oxford Science Publications.

[8] L. Faybusovich. Euclidean Jordan algebras and interior-point algorithms. *Positivity*, 1(4):331–357, 1997.

[9] L. Faybusovich. Linear systems in Jordan algebras and primal-dual interior-point algorithms. *J. Comput. Appl. Math.*, 86(1):149–175, 1997. Special issue dedicated to William B. Gragg (Monterey, CA, 1996).

[10] Robert M. Freund. Projective transformations for interior-point algorithms, and a superlinearly convergent algorithm for the w-center problem. *Math. Programming*, 58(3, Ser. A):385–414, 1993.

[11] C. C. Gonzaga. The largest step path following algorithm for monotone linear complementarity problems. *Math. Programming*, 76(2, Ser. A):309–332, 1997.

[12] O. Güler. Barrier functions in interior point methods. *Math. Oper. Res.*, 21(4):860–885, 1996.

[13] R. A. Hauser and O. Güler. Self-scaled barrier functions on symmetric cones and their classification. *Found. Comput. Math.*, 2(2):121–143, 2002.

[14] M. Kojima, S. Mizuno, and A. Yoshise. An $O(\sqrt{n}\,L)$ iteration potential reduction algorithm for linear complementarity problems. *Math. Programming*, 50(3, (Ser. A)):331–342, 1991.

[15] K. A. McShane. Superlinearly convergent $O(\sqrt{n}L)$-iteration interior-point algorithms for linear programming and the monotone linear complementarity problem. *SIAM Journal on Optimization*, 4(2):247–261, 1994.

[16] S. Mizuno, M. J. Todd, and Y. Ye. On adaptive-step primal-dual interior-point algorithms for linear programming. *Mathematics of Operations Research*, 18(4):964–981, 1993.

[17] Y. Nesterov and A. Nemirovsky. *Interior Point Polynomial Methods in Convex Programming*. SIAM Publications. SIAM, Philadelphia, 1994.

[18] Yu. E. Nesterov and M. J. Todd. Self-scaled barriers and interior-point methods for convex programming. *Math. Oper. Res.*, 22(1):1–42, 1997.

[19] Yu. E. Nesterov and M. J. Todd. Primal-dual interior-point methods for self-scaled cones. *SIAM J. Optim.*, 8(2):324–364 (electronic), 1998.

[20] F. A. Potra. The Mizuno-Todd-Ye algorithm in a larger neighborhood of the central path. *European Journal of Operational Research*, 143:257–267, 2002.

[21] F.A. Potra. Sufficient weigthed complementarity problems. Technical report, University of Maryland Baltimore County, in preparation.

[22] C. Roos, T. Terlaky, and J.-P. Vial. *Interior point methods for linear optimization*. Springer, New York, 2006.

[23] P. M. Vaidya. A new algorithm for minimizing convex functions over convex sets. *Math. Programming*, 73(3, Ser. A):291–341, 1996.

[24] Y. Ye. *Interior Point Algorithms : Theory and Analysis*. Wiley-Interscience Series in Discrete Mathematics and Optimization. John Wiley and Sons, 1997.

[25] Y. Ye. A path to the Arrow-Debreu competitive market equilibrium. *Math. Program.*, 111(1-2, Ser. B):315–348, 2008.

[26] Y. Ye and K. Anstreicher. On quadratic and $O(\sqrt{n}L)$ convergence of predictor-corrector algorithm for LCP. *Mathematical Programming*, 62(3):537–551, 1993.

[27] A. Yoshise. Interior point trajectories and a homogeneous model for nonlinear complementarity problems over symmetric cones. *SIAM Journal on Optimization*, 17(4):1129–1153, 2006.