Nonlinear Equilibrium for optimal resource allocation

Roman A Polyak (rpolyak@gmu.edu)
George Mason University
SEOR & Mathematical Sciences Departments
George Mason University
Fairfax, VA 22030, USA

Abstract

We consider Nonlinear Equilibrium (NE) for optimal allocation of limited resources. The NE is a generalization of the Walras-Wald equilibrium, which is equivalent to J. Nash equilibrium in an n-person concave game. Finding NE is equivalent to solving a variational inequality (VI) with a monotone and smooth operator on $\Omega = \mathbb{R}^n_+ \otimes \mathbb{R}^m_+$. The projection on Ω is a very simple procedure, therefore our main focus is two methods for which the projection on Ω is the main operation. Both projected pseudo-gradient (PPG) and extra pseudo-gradient (EPG) methods require $O(n^2)$ operations per step. We proved convergence, established global Q-linear rate and estimated computational complexity for both PPG and EPG methods. The methods can be viewed as pricing mechanisms for establishing economic equilibrium.

Keywords: Nonlinear Equilibrium, Duality, Walras-Wald equilibrium, Pseudo-gradient, Extra-pseudo-gradient, Linear Programming.

1 Introduction

For several decades LP has been widely used for optimal resource allocation. In 1975 L. V. Kantorovich and T. C. Koopmans shared the Nobel Prize in Economics "for their contributions to the theory of optimum allocation of limited resources."

The LP approach uses two fundamental assumptions:

a) The price vector $c = (c_1, \dots, c_n)^T$ for goods is fixed, given priory and independent of the production output vector $x = (x_1, \dots, x_n)^T$.

b) The resource vector $b = (b_1, \dots, b_m)^T$ is also fixed, given priory and the resource availability is independent of the resource price vector $\lambda = (\lambda_1, \dots, \lambda_n)^T$.

Unfortunately, such assumptions do not reflect the basic market law of supply and demand. Therefore, the LP models might lead to solutions which are not always practical. Also, a small change of at least one component of the price vector c might lead to a drastic change of the primal solution. Similarly, a small variation of the resource vector b might lead to a dramatic change of the dual solution.

We consider an alternative to the LP approach for optimal resource allocation, which is based on the Generalized Walras-Wald Equilibrium [10].

The fixed price vector $c = (c_1, \dots, c_n)^T$ is replaced by a price operator $c : \mathbb{R}^n_+ \to \mathbb{R}^n_+$, which maps the production output vector $x = (x_1, \dots, x_n)^T$ into the price vector $c(x) = (c_1(x), \dots, c_n(x))^T$.

Similarly, the fixed resource vector $b = (b_1, \dots, b_m)^T$ is replaced by the resource operator $b : \mathbb{R}^m_+ \to \mathbb{R}^m_+$, which maps the resource price vector $\lambda = (\lambda_1, \dots, \lambda_n)^T$ into the resource availability vector $b(\lambda) = (b_1(\lambda), \dots, b_m(\lambda))^T$.

The pair of vectors $y^* = (x^*, \lambda^*) \in \Omega :$

$$(c(x^*), x^*) = \max\{(c(x^*), x) | Ax \leq b(\lambda^*), x \in \mathbb{R}^n_+\} \tag{1}$$

$$(b(\lambda^*), \lambda^*) = \min\{(b(\lambda^*), \lambda) | A^T\lambda \geq c(x^*), \lambda \in \mathbb{R}^m_+\} \tag{2}$$

we call nonlinear equilibrium (NE).

The primal-dual LP solution which one obtains from (1)-(2) when $c(x) \equiv c$ and $b(\lambda) \equiv b$ can be viewed as linear equilibrium (LE).

The strong monotonicity assumptions for both the price operator $c : \mathbb{R}^n_+ \to \mathbb{R}^n_+$ and the resource operator $b : \mathbb{R}^m_+ \to \mathbb{R}^m_+$ guarantee the existence and uniqueness of the NE [10].

In this paper we reduce the strong monotonicity assumptions for both operators $c : \mathbb{R}^n_+ \to \mathbb{R}^n_+$ and $b : \mathbb{R}^m_+ \to \mathbb{R}^m_+$ to local strong monotonicity at the equilibrium y^* or just monotonicity.

We show that under local strong monotonicity and Lipschitz continuity the PPG converges globally with Q-linear rate and the ratio depends only on the condition number of the VI operator. We established the computational complexity of the PPG method in terms of the condition number, the size of the problem and the required accuracy. This is our first contribution.

In the absence of even local strong monotonicity the convergence of the PPG becomes problematic. Therefore in the second part of the paper we consider the extra pseudo-gradient (EPG) method for finding y^*.

The extra gradient method was first introduced by G. Korpelevich in the 70's for finding saddle points[7]. Over the years, it became an important tool for solving VI (see [1], [2],[3],[4],[5],[6] and references therein).

Application of the EPG for finding NE leads to a two stage algorithm. At the first stage the EPG predicts both the production and the price vector. At the second stage it corrects them dependent on the prices for the predicted output and resource availability for the predicted resource prices. It requires projecting the primal-dual vector on Ω twice, which is still only $O(n^2)$ operations.

It was shown that the EPG method converges to the NE y^* if both the price c and resource b operators are just monotone and satisfy a Lipschitz condition. This is our second contribution.

Under local strong monotonicity the EPG method globally converges with Q-linear rate, and the ratio is defined by the condition number of the VI operator. For a small condition number, the EPG has a better ratio and a much better complexity bound than the PPG. This is our third contribution.

The paper is organized as follows. The basic assumptions are introduced in the following section. In section three we remind the difference between the classical Walras-Wald equilibrium and NE and show the equivalence of finding NE to solving a particular VI. In section four we establish convergence properties and computational complexity of the PPG. In section five we prove convergence of the EPG method under minimum assumptions on the input data. In section six we establish global Q-linear convergence rate and computational complexity of the EPG. In the Appendix we estimate the Lipschitz constant for the VI operator which plays an important role in both PPG and EPG methods. We conclude the paper by discussing important properties of the NE and the fundamental differences between NE and LE.

2 Basic Assumptions.

We consider an economy which produces n goods by consuming m resources. There are three sets of data required for problem formulation.

1) The technological matrix $A : \mathbb{R}^n_+ \to \mathbb{R}^m_+$ which "transforms" resources into goods, *i.e.* a_{ij} defines the amount of factor $1 \leq i \leq m$ which is required to produce one item of good $1 \leq j \leq n$.

2) The resource operator $b : \mathbb{R}^m_+ \to \mathbb{R}^m_+$, where $b_i(\lambda)$ is the availability of the resource i under the resource price vector $\lambda = (\lambda_1, \ldots, \lambda_i, \ldots, \lambda_m)$.

3) The price operator $c : \mathbb{R}^n_+ \to \mathbb{R}^n_+$, where $c_j(x)$ is the price for one item of good j under the production output $x = (x_1, \dots, x_j, \dots, x_n)$.

We assume the matrix A does not have zero rows or columns, which means that each resource is used for the production of at least one of the goods and each good requires at least one of the resources.

Under strong monotonicity of b on $\mathbb{R}^m_+$ and c on $\mathbb{R}^n_+$ the NE y^* exists and is unique. Finding y^* is equivalent to solving a VI with a strongly monotone operator on Ω. Therefore under Lipschitz continuity of b and c the PPG generates a primal-dual sequence, which converges to y^* with Q-linear rate [10].

In this paper we replace the global strong monotonicity of b and c with corresponding properties only at the NE y^*:

$$(b(\lambda) - b(\lambda^*), \lambda - \lambda^*) \geq \beta \|\lambda - \lambda^*\|^2, \; \beta > 0, \; \forall \lambda \in \mathbb{R}^m_+ \tag{3}$$

$$(c(x) - c(x^*), x - x^*) \leq -\alpha \|x - x^*\|^2, \; \alpha > 0 \, \forall x \in \mathbb{R}^n_+ \tag{4}$$

In the first part we also replace the global Lipschitz continuity of b and c by corresponding local assumptions:

$$\|b(\lambda) - b(\lambda^*)\| \leq L_b \|\lambda - \lambda^*\|, \; \forall \lambda \in \mathbb{R}^m_+ \tag{5}$$

and

$$\|c(x) - c(x^*\| \leq L_c \|x - x^*\|, \; \forall x \in \mathbb{R}^n_+ \tag{6}$$

where $\|\cdot\|$ is the Euclidean norm.

We will say that the price and resource operators are well defined if (3)-(6) hold.

The assumption (3) implies that an increase of the price λ_i for any resource $1 \leq i \leq m$ when the rest is fixed at the equilibrium level leads to an increase of the resource availability $b_i(\lambda)$ and the margin for the resource increase has a positive lower bound. Conversely, it follows from (4) that any increase of production x_j, $1 \leq i \leq n$ when the rest is fixed at the equilibrium level leads to a decrease of the price $c_j(x)$ per item for good j. Moreover, the margin of the price decrease has a negative upper bound. The Lipschitz conditions (5)-(6) assume that deviation from the NE y^* can't lead to uncontrolled changes of prices for goods and resource availability.

3 Generalized Walras-Wald Equilibrium.

In this section we remind that NE is a generalization of Walras-Wald (WW) equilibrium, which is equivalent to a particular VI.

The notion of equilibrium in a concave n-person game was introduced by J. Nash in 1950 [9]. He received the Nobel Prize in Economics for his discovery in 1994. For many years it was not clear whether J. Nash equilibrium had anything to do with economic equilibrium introduced as early as 1874 by Leon Walras in his most admired work "Elements of Pure Economics". Moreover, it has not been clear for a long time whether Walras equations have a solution.

The first substantial contribution was due to Abraham Wald, who in the mid 1930's proved the existence of Walras equilibrium under some special assumptions on the price vector-fuction $c(x)$. These assumptions, unfortunately, were hard to justify from an economic standpoint [8].

In the mid 1950's, Harold Kuhn modified the WW model. H. Kuhn's version of WW equilibrium consists of finding $y^* = (x^*; \lambda^*)$:

$$(c(x^*), x^*) = \max\{(c(x^*), X) | AX \leq b, X \in \mathbb{R}^n_+\} \tag{7}$$

$$(b, \lambda^*) = \min\{(b, \Lambda) | A^T\Lambda \geq c(x^*), \Lambda \in \mathbb{R}^m_+\}. \tag{8}$$

He proved the existence of the WW equilibrium under minimum assumptions on the input data, using two basic tools: Kakutani's fixed point Theorem (1941) to show the existance of $x^* \in \mathbb{R}^n_+$ in (7) and LP Duality (1947) to show the existance of $\lambda^* \in \mathbb{R}^m_+$ in (8).

The equivalence of H. Kuhn's version of WW equilibrium and J. Nash equilibrium in a concave n-person game was established in [12].

One obtains WW equilibrium from NE by assuming $b(\lambda) = b$ in (1)-(2). So the NE (1)-(2) is a natural extension of the WW equilibrium, which makes it in a sense "symmetric".

Our next step is to show that finding NE from (1)-(2) is equivalent to solving a particular variational inequality (VI).

We assume that NE $y^* = (x^*; \lambda^*) \in \Omega = \mathbb{R}^n_+ \otimes \mathbb{R}^m_+$ defined by (1)-(2) exists.
Theorem 1. *Finding $y^* = (x^*; \lambda^*)$ from (1)-(2) is equivalent to solving the following VI*

$$(g(y^*), Y - y^*) \leq 0, \forall Y = (X; \Lambda) \in \Omega \tag{9}$$

where $g(y) = g(x, \lambda) = \big(c(x) - A^T\lambda; Ax - b(\lambda)\big)$.

Proof. Let $y^* = (x^*; \lambda^*)$ be defined by (1)-(2).

Solving the dual LP pair (1)-(2) is equivalent to finding a saddle point of the corresponding Lagrangian

$$\mathbb{L}(x^*, \lambda^*; X, \Lambda) = (c(x^*), X) - (\Lambda, AX - b(\lambda^*)), \tag{10}$$

i.e. finding

$$\max_{X \in \mathbb{R}^n_+} \min_{\Lambda \in \mathbb{R}^m_+} \mathbb{L}(x^*, \lambda^*; X, \Lambda) = \min_{\Lambda \in \mathbb{R}^m_+} \max_{X \in \mathbb{R}^n_+} \mathbb{L}(x^*, \lambda^*; X, \Lambda). \tag{11}$$

The problem (11) is equivalent to finding a J. Nash equilibrium of a concave two person game with the following payoff functions:

$$\varphi_1(x,\lambda;X,\lambda) = (c(x),X) - (\lambda, AX - b(\lambda)) = (c(x) - A^T\lambda, X) + (\lambda, b(\lambda))$$
$$\varphi_2(x,\lambda;x,\Lambda) = (Ax - b(\lambda), \Lambda)$$

where $X \in \mathbb{R}^n_+$ is the strategy of the first player and $\Lambda \in \mathbb{R}^m_+$ is the strategy of the second player.

Let $y = (x;\lambda)$ and $Y = (X;\Lambda) \in \Omega$. The corresponding normalized payoff function (see [11]) is defined as follows

$$\Phi(y,Y) = (c(x) - A^T\lambda, X) + (Ax - b(\lambda), \Lambda) + (\lambda, b(\lambda))$$

Therefore, finding a saddle point (11) is equivalent to finding a normalized J. Nash equilibrium of a two person concave game, i.e. to finding such $y^* \in \Omega$ that

$$\Phi(y^*,y^*) = \max\{\Phi(y^*,Y)|Y \in \Omega\}. \quad (12)$$

Under fixed $y = y^*$ the function $\Phi(y^*,Y)$ is concave in $Y \in \Omega$, therefore the problem (12) is a convex programming problem. Let's consider the pseudo-gradient of $\Phi(y,Y)$ at y^*, i.e.

$$g(y^*) = \nabla_Y \Phi(y^*,Y)_{|Y=y^*} = (c(x^*) - A^T\lambda^*; Ax^* - b(\lambda^*))$$

The fact that y^* is the solution of the problem (12) means that

$$(g(y^*), Y - y^*) \leq 0, \ \forall Y \in \Omega, \quad (13)$$

i.e. y^* is the solution of the VI (9).

On the other hand, let's assume that $\bar{y} \in \Omega$ is the solution of the following VI

$$(g(\bar{y}, Y - \bar{y}) \leq 0, \ \forall Y \in \Omega, \quad (14)$$

then

$$\max\{(g(\bar{y}), Y - \bar{y})|Y \in \Omega\} = (g(\bar{y}), \bar{y} - \bar{y}) = 0. \quad (15)$$

From (15) we obtain $g(\bar{y}) \leq 0$ because assuming that at least one component of the vector $g(\bar{y})$ is positive, we obtain

$$\max\{(g(\bar{y}), Y - \bar{y})|Y \in \Omega\} = \infty$$

Therefore, if $\bar{y} = (\bar{x};\bar{\lambda}) \geq 0$ solves the variational inequality (14), then

$$c(\bar{x}) - A^T\bar{\lambda} \leq 0 \text{ and } A\bar{x} - b(\bar{\lambda}) \leq 0.$$

Moreover, for $1 \leq j \leq n$ we have either

$$\begin{array}{ccc} (c(\bar{x}) - A^T\bar{\lambda})_j < 0 & \Rightarrow & \bar{x}_j = 0, \\ & \text{or} & \\ \bar{x}_j > 0 & \Rightarrow (c(\bar{x}) - A^T\bar{\lambda}) = 0 & \end{array} \tag{16}$$

and for $1 \leq i \leq m$ we have either

$$\begin{array}{ccc} (A\bar{x} - b(\bar{\lambda}))_i < 0 & \Rightarrow & \bar{\lambda}_i = 0, \\ & \text{or} & \\ \bar{\lambda}_i > 0 & \Rightarrow & (A\bar{x} - b(\bar{\lambda}))_i = 0. \end{array} \tag{17}$$

Hence, $\bar{y} = (\bar{x}; \bar{\lambda}) \in \Omega$ is a primal-dual feasible solution, which satisfied the complementarity condition (17)-(18). Therefore, the vector $Y = \bar{y}$ is the solution of the following primal-dual LP pair

$$\max\left\{(c(\bar{x}), X) | AX \leq b(\bar{\lambda}), X \in \mathbb{R}^n_+\right\} = (c(\bar{x}), \bar{x})$$
$$\max\left\{(b(\bar{\lambda}, \Lambda) | A^T\Lambda \geq c(\bar{x}), \Lambda \in \mathbb{R}^m_+\right\} = (b(\bar{\lambda}), \bar{\lambda}),$$

i.e. $\bar{y} = y^*$.

□

Let $D = \{X : AX \leq b, X \in \mathbb{R}^n_+\}$, then the classical WW equilibrium is equivalent to the following VI

$$x^* \in D : (c(x^*), X - x^*) \leq 0, \forall X \in D. \tag{18}$$

Solving (18), generally speaking, is more difficult than solving the corresponding primal-dual LP, i.e. (1)-(2) when $b(\lambda) \equiv b$ and $c(x) \equiv c$. It looks like finding NE (x^*, λ^*) from (1)-(2) is more difficult than solving the VI (18).

In fact, as we will see later, finding NE $y^* = (x^*; \lambda^*)$ from (1)-(2) in a number of instances can be even easier than solving the corresponding LP.

The fundamental difference between NE (1)-(2) and WW (18) follows from the geometry of their feasible sets $\Omega = \mathbb{R}^n_+ \otimes \mathbb{R}^m_+$ and D. The simplicity of Ω makes projection type methods particularly suitable for solving VI (9), whereas projection methods for solving VI (18) require solving a quadratic programming problem at each step.

In the following sections we will concentrate on a projected pseudo-gradient method for solving VI (9).

4 Projected Pseudo-Gradient Method

Let Q be a closed convex set in $\mathbb{R}^n$, then for each $u \in \mathbb{R}^q$ there is

$$v = P_Q(u) = \text{argmin}\{\|w-u\| | w \in Q\}$$

The vector v is called projection of u on Q. Later we will need the following two well known facts about the projection operator P_Q.

First, operator $P_Q : u \in \mathbb{R}^q \to v \in \Omega$ is non-expansive i.e.

$$\|P_Q(u_1) - P_Q(u_2)\| \leq \|u_1 - u_2\|, \forall u_1, u_2 \in \mathbb{R}^q \tag{19}$$

Second, vector $u^* \in \mathbb{R}^q$ is a solution of the VI

$$(g(u^*), u - u^*) \leq 0, \forall u \in Q$$

if and only if for any $t > 0$ the vector u^* is a fixed point of the map $P_Q(I + tg) : \Omega \to \Omega$, i.e.

$$u^* = P_\Omega(u^* + tg(u^*)). \tag{20}$$

For a vector $u \in \mathbb{R}^q$ the projection on $\mathbb{R}^q_+$ is given by the formula

$$v = P_{\mathbb{R}^q_+}(u) = [u]_+ = ([u_1]_+, \dots, [u_q]_+)^T,$$

where for $1 \leq i \leq q$ we have

$$[u_i]_+ = \begin{cases} u_i, & u_i \geq 0 \\ 0, & u_i < 0. \end{cases}$$

Therefore the projection $P_\Omega(y)$ of $y = (x, \lambda) \in \mathbb{R}^n \otimes \mathbb{R}^m$ on $\Omega = \mathbb{R}^n_+ \otimes \mathbb{R}^m_+$ is defined by the following formula

$$P_\Omega(y) = [y]_+ = ([x]_+; [\lambda]_+).$$

We remind that the VI operator $g : \Omega \to \mathbb{R}^{n+m}$ is defined by the formula

$$g(y) = (c(x) - A^T\lambda; Ax - b(\lambda)) \tag{21}$$

We are ready to describe the PPG method for solving the VI (9). Let $y^0 = (x^0; \lambda^0) \in \mathbb{R}^n_{++} \otimes \mathbb{R}^m_{++}$ be a starting point and $(x^s; \lambda^s)$ has been found already. The PPG method finds the next approximation y^{s+1} by the formula

$$y^{s+1} = P_\Omega(y^s + tg(y^s)). \tag{22}$$

In other words, each step of the PPG method consists of updating goods and prices by the following formulas

$$x_j^{s+1} = [x_j^s + t(c(x^s) - A^T\lambda^s)_j]_+,\ j = 1,\dots,n \tag{23}$$

$$\lambda_i^{s+1} = [\lambda_i^s + t(Ax^s - b(\lambda^s))_i]_+,\ i = 1,\dots,m. \tag{24}$$

The step length $t > 0$ we will specify later. The method (23)-(24) can be viewed as a projected explicit Euler method for solving the following system of differential equations

$$\begin{aligned}\frac{dx}{dt} &= c(x) - A^T\lambda \\ \frac{d\lambda}{dt} &= Ax - b(\lambda).\end{aligned}$$

On the other hand, the PPG method (23)-(24) can be viewed as a pricing mechanism for finding equilibrium.

It follows from (23) that if the current price $c_j(x^s)$ for an item of good j exceeds the expenses $(A^T\lambda^s)_j$ required to produce this item, then the production of good j has to be increased. On the other hand, if the current price $c_j(x^s)$ is less than the current expenses $(A^T\lambda^s)_j$, then the production of good j has to be reduced.

It follows from (24) that if the current consumption $(Ax^s)_i$ of resource i exceeds the current availability $b_i(\lambda^s)$, then the price for the resource has to be increased.If the availability $b_i(\lambda^s)$ of resource i exceeds consumption $(Ax^s)_i$, then the price for an item of the resource has to be reduced.

Lemma 1. *If the operators b and c are locally strongly monotone, i.e. (3)-(4) hold, then the operator g is locally strongly monotone and for $\gamma = \min\{\alpha, \beta\}$ the following inequality holds*

$$(g(y) - g(y^*), y - y^*) \le -\gamma\|y - y^*\|^2,\ \forall y \in \Omega. \tag{25}$$

Proof. We have

$$\begin{aligned}(g(y) - g(y^*), y - y^*) &= (c(x) - A^T\lambda - c(x^*) + A^T\lambda^*, x - x^*) \\ &+ (Ax - b(\lambda) - Ax^* + b(\lambda^*, \lambda - \lambda^*) \\ &= (c(x) - c(x^*), x - x^*) - (A^T(\lambda - \lambda^*), x - x^*) \\ &+ (A(x - x^*), \lambda - \lambda^*) - (b(\lambda) - b(\lambda^*), \lambda - \lambda^*).\end{aligned}$$

Using (3) and (4) for $\gamma = \min\{\alpha, \beta\}$ we obtain (25) □

Lemma 2. *If b and c satisfy the local Lipschitz conditions (5)-(6) then the operator $g : \Omega \to \mathbb{R}^{n+m}$ given by (21) satisfies the local Lipschitz condition, i.e. there is an $L > 0$ such that*

$$\|g(y) - g(y^*)\| \le L\|y - y^*\|,\ \forall y \in \Omega \tag{26}$$

For the proof and the upper bound for L see Appendix.

Remark 1. *We will assume later that for a given* $x \in \mathbb{R}^n$ *finding c(x) does not require more then* $O(n^2)$ *operations and for a given* $\lambda \in \mathbb{R}^m$ *finding* $b(\lambda)$ *does not require more then* $O(m^2)$ *operations. We also assume that* $n \geq m$. *From (23)-(24) follows that each step of the PPG method (22) does not require more than* $O(n^2)$ *operations.*

Example 1. *Let* $c(x) = \nabla(\frac{1}{2}x^T Cx + c^T x)$ *and* $b(\lambda) = \nabla(\frac{1}{2}\lambda^T B\lambda + b^T\lambda)$, *where* $C : \mathbb{R}^n \to \mathbb{R}^n$ *is symmetric negative semidefinite and* $B : \mathbb{R}^m \to \mathbb{R}^m$ *symmetric positive semidefinite, then each step of PPG method (22) requires* $O(n^2)$ *operations.*

Let $\varkappa = \gamma L^{-1}$ be the condition number of the VI operator g. The following Theorem establishes the global Q-linear convergence rate and complexity of the PPG method (22).

Theorem 2. *If operators b and c are well defined i.e. (3)-(6) hold then:*

1) *for any* $0 < t < 2\gamma L^{-2}$ *the PPG method (22) globally converges to NE* $y^* = (x^*; \lambda^*)$ *with Q-linear rate and the ratio* $0 < q(t) = (1 - 2t\gamma + t^2 L^2)^{1/2} < 1$, *i.e.*

$$\|y^{s+1} - y^*\| \leq q(t)\|y^s - y^*\|; \tag{27}$$

2) *for* $t = \gamma L^{-2} = \min\{q(t)|t > 0\}$ *the following bound holds*

$$\|y^{s+1} - y^*\| \leq (1 - \varkappa^2)^{1/2}\|y^s - y^*\|; \tag{28}$$

3) *for the PPG complexity we have the following bound*

$$Comp(PPG) = O(n^2 \varkappa^{-2} \ln \varepsilon^{-1}), \tag{29}$$

where $\varepsilon > 0$ *is the required accuracy.*

Proof. 1) From (22), semicontractibility (19) and optimality criteria (20) follows

$$\begin{aligned}
\|y^{s+1} - y^*\|^2 &= \|P_\Omega(y^s + tg(y^s)) - P_\Omega(y^* + tg(y^*))\|^2 \\
&\leq \|y^s + tg(y^s) - y^* - tg(y^*)\|^2 \\
&= (y^s - y^* + t(g(y^s) - g(y^*)), y^s - y^* + t(g(y^s) - g(y^*))) \\
&= \|y - y^*\|^2 + 2t(y^s - y^*, g(y^s) - g(y^*)) \\
&+ t^2\|g(y^s) - g(y^*)\|^2
\end{aligned} \tag{30}$$

For well defined b and c from (25), (26) and (30) we obtain

$$\|y^{s+1} - y^*\|^2 \leq \|y^s - y^*\|^2(1 - 2t\gamma + t^2 L^2)$$

Hence for $0 < t < 2\gamma L^{-2}$ we have $0 < q(t) = 1 - 2t\gamma + t^2 L^2 < 1$. In other words the projection operator (22) is contractive, which means that for any given $t \in (0, 2\gamma L^{-2})$ the PPG method globally converges with Q-linear rate, i.e. (27) holds.

2) For $t = \gamma L^{-2} = \text{argmin}\{q(t)|t > 0\}$ we have $q = q(\gamma L^{-2}) = (1-(\gamma L^{-1})^2)) = (1-\varkappa^2)$ i.e. (28) holds.

3) Let $0 < \varepsilon \ll 1$ be the required accuracy, then in view of (28) it takes $O((\ln q)^{-1} \ln \varepsilon)$ steps to find an ε-approximation for the NE $y^* = (x^*, \lambda^*)$. It follows from Remark 1 that each PPG step (22) does not require more than $O(n^2)$ operations. Therefore, finding the ε-approximation to NE $y^* = (x^*, \lambda^*)$ requires $N = O(n^2 \ln \varepsilon^{-1} (\ln q^{-1})^{-1})$ operations. In view of $(\ln q^{-1})^{-1} = (-\frac{1}{2}\ln(1-\varkappa^2))^{-1}$ and keeping in mind $\ln(1+x) \leq x$, $\forall x > -1$ we have $\ln(1-\varkappa^2) \leq -\varkappa^2$ i.e. $-\frac{1}{2}\ln(1-\varkappa^2) \geq \frac{1}{2}\varkappa^2$ or $(\ln q^{-1})^{-1} = (-\frac{1}{2}\ln(1-\varkappa^2))^{-1} \leq 2\varkappa^{-2}$, so for the overall complexity of the PPG method we obtain (29).

□

If $\gamma = \min\{\alpha, \beta\} = 0$, then pseudo-gradient $g : \Omega \to \mathbb{R}^{m+n}$ defined by (21) is not even locally strongly monotone, therefore (27) cannot guarantee convergence of the PPG method (22). In the following section we consider the extra pseudo-gradient method (EPG) for finding NE (1)-(2) in the absence of local strong monotonicity of both operators b and c.

The extra gradient method was first introduced by G. Korpelevich ([7]) in the 70s for finding saddle points. Lately, it became a popular tool for solving VI problems (see [1], [2],[3],[4],[5],[6] and references therein).

First we show that EPG converges to the NE for any monotone operators b and c which satisfy a Lipschitz condition on $\Omega = \mathbb{R}^n_+ \otimes \mathbb{R}^m_+$, i.e.

$$\|g(y_1) - g(y_2)\| \leq L\|y_1 - y_2\|, \ \forall y_1, y_2 \in \Omega \tag{31}$$

5 Extra Pseudo-Gradient Method for finding NE

The application of G. Korpelevich extra gradient method [7] for solving VI (9) leads to the following two phase extra pseudo-gradient (EPG) method for finding NE (1)-(2).

At each step the predictor phase consists of finding

$$\hat{y}_s = P_\Omega(y_s + tg(y_s)) = [y_s + tg(y_s)]_+, \tag{32}$$

while the corrector phase finds the new approximation

$$y_{s+1} = P_\Omega(y_s + tg(\hat{y}_s)) = [y_s + tg(\hat{y}_s)]_+. \tag{33}$$

The step length $t > 0$ will be specified later.

We start with initial approximation $y_0 = (x_0; \lambda_0)) \in \mathbb{R}^n_{++} \otimes \mathbb{R}^m_{++}$. The formulas (32)-(33) can be viewed as a pricing mechanism for finding the NE $y^* = (x^*; \lambda^*)$.

The first phase predicts the new production vector

$$\hat{x}_s = [x_s + t(c(x_s) - A^T \lambda_s)]_+ \tag{34}$$

and a new price vector

$$\hat{\lambda}_s = [\lambda_s + t(Ax_s - b(\lambda_s))]_+. \tag{35}$$

The pair $(\hat{x}_s; \hat{\lambda}_s)$ in turn predicts the price vector $c(\hat{x}_s) = (c_1(\hat{x}_s), \ldots, c_n(\hat{x}_s))$ and the resource availibility vector $b(\hat{\lambda}_s) = (b_1(\hat{\lambda}_s), \ldots, b_m(\hat{\lambda}_s))$.

The corrector phase establishes the new production level

$$x_{s+1} = [x_s + t(c(\hat{x}_s) - A^T \hat{\lambda}_s)]_+ \tag{36}$$

and the new price level

$$\lambda_{s+1} = [\lambda_s + t(A\hat{x}_s - b(\hat{\lambda}_s))]_+ \tag{37}$$

for the resources.

The meaning of the formulas (34)-(35) and (36)-(37) is similar to the meaning of the formulas (23)-(24).

Theorem 3. *If c and b are monotone operators and Lipschitz condition (31) is satisfied, then for any $t \in (0, (\sqrt{2}L)^{-1})$ the EPG method (32)-(33) generates a sequence $\{y_s\}_{s=1}^{\infty}$ such that $\lim_{s\to\infty} y_s = y^*$.*

Proof. Let's consider vector $h_s = y_s + tg(y_s) - \hat{y}_s$, then from (32) we have

$$(h_s, y - \hat{y}_s) \le 0, \quad \forall y \in \Omega = \mathbb{R}^n_+ \otimes \mathbb{R}^m_+,$$

i.e. for a given $t > 0$ and $\forall y \in \Omega$ we have

$$(g(y_s) + t^{-1}(y_s - \hat{y}_s), y - \hat{y}_s) \le 0. \tag{38}$$

For $h_{s+1} = y_s + tg(\hat{y}_s - y_{s+1}$ from (33) follows $(h_{s+1}, y - y_{s+1}) \le 0,\ \forall y \in \Omega$.

Therefore, for a given $t > 0$ and $\forall y \in \Omega$ we have

$$(g(\hat{y}_s) + t^{-1}(y_s - y_{s+1}), y - y_{s+1}) \le 0. \tag{39}$$

From (32), (33) and the semicontractability of the projection operator (19) as well as Lipschitz condition (31) we obtain

$$\begin{aligned} ||y_{s+1} - \hat{y}_s|| &= ||P_\Omega(y_s + tg(\hat{y}_s)) - P_\Omega(y_s + tg(y_s)|| \\ &\leq t\, ||g(\hat{y}_s) - g(y_s)|| \\ &\leq tL\, ||\hat{y}_s - y_s|| \,. \end{aligned} \tag{40}$$

From (39) for $y = y^*$ we have

$$(y_s - y_{s+1} + tg(\hat{y}_s), y^* - y_{s+1}) \leq 0. \tag{41}$$

By taking $y = y_{s+1}$ in (38) we obtain

$$(y_s - \hat{y}_s, y_{s+1} - \hat{y}_s) + t(g(y_s), y_{s+1} - \hat{y}_s) \leq 0,$$

or

$$(y_s - \hat{y}_s, y_{s+1} - \hat{y}_s) + t(g(\hat{y}_s), y_{s+1} - \hat{y}_s) - t(g(\hat{y}_s) - g(y_s), y_{s+1} - \hat{y}_s) \leq 0. \tag{42}$$

Then using (40) we obtain

$$\begin{aligned} (g(\hat{y}_s) - g(y_s), y_{s+1} - \hat{y}_s) &\leq ||g(\hat{y}_s) - g(y_s)||\, ||y_{s+1} - \hat{y}_s|| \\ &\leq tL^2\, ||\hat{y}_s - y_s||^2 \,. \end{aligned}$$

Therefore, from (42) we have

$$(y_s - \hat{y}_s, y_{s+1} - \hat{y}_s) + t(g(\hat{y}_s), y_{s+1} - \hat{y}_s) - (tL)^2\, ||\hat{y}_s - y_s||^2 \leq 0. \tag{43}$$

By adding (41) and (43) we obtain

$$\begin{aligned} (y_s - y_{s+1}, y^* - y_{s+1}) + t(g(\hat{y}_s), y^* - y_{s+1}) &+ (y_s - \hat{y}_s, y_{s+1} - \hat{y}_s) \\ + t(g(\hat{y}_s), y_{s+1} - \hat{y}_s) &- (tL)^2\, ||\hat{y}_s - y_s||^2 \\ = (y_s - y_{s+1}, y^* - y_{s+1}) &+ t(g(\hat{y}_s), y^* - \hat{y}_s) \\ &+ (y_s - \hat{y}_s, y_{s+1} - \hat{y}_s) - (tL)^2\, ||y_s - \hat{y}_s||^2 \leq 0. \end{aligned} \tag{44}$$

From $(g(y^*), y - y^*) \leq 0$, $\forall y \in \Omega$ we obtain $(g(y^*), \hat{y}_s - y^*) \leq 0$ or $t(-g(y^*), y^* - \hat{y}_s) \leq 0$. Adding the last inequality to the left hand side of (44) and using the monotonicity inequality

$$(g(\hat{y}_s) - g(y^*), y^* - \hat{y}_s) \geq 0$$

from (44) we obtain

$$2(y_s - y_{s+1}, y^* - y_{s+1}) + 2(y_s - \hat{y}_s, y_{s+1} - \hat{y}_s) - 2(tL)^2\, ||\hat{y}_s - y_s||^2 \leq 0. \tag{45}$$

Using identity

$$2(u-v,w-v)=||u-v||^2+||v-w||^2-||u-w||^2 \tag{46}$$

with $u=y_s$, $v=y_{s+1}$, and $w=y^*$ we obtain

$$2(y_s-y_{s+1},y^*-y_{s+1})=||y_s-y_{s+1}||^2+||y_{s+1}-y^*||^2-||y_s-y^*||^2 .$$

Using the same identity with $u=y_s$, $v=\hat{y}_s$, and $w=y_{s+1}$ we obtain

$$2(y_s-\hat{y}_s,y_{s+1}-\hat{y}_s)=||y_s-\hat{y}_s||^2+||\hat{y}_s-y_{s+1}||^2-||y_s-y_{s+1}||^2 .$$

Therefore, we can rewrite (45) as follows

$$||y_{s+1}-y^*||^2+(1-2(tL)^2)\,||y_s-\hat{y}_s||^2+||\hat{y}_s-y_{s+1}||^2\le||y_s-y^*||^2 . \tag{47}$$

By adding up the last inequality from $s=0$ to $s=N$ we obtain

$$||y_{N+1}-y^*||^2+(1-2(tL)^2)\sum_{s=1}^{N}||y_s-\hat{y}_s||^2+\sum_{s=1}^{N}||\hat{y}_s-y_{s+1}||^2\le||y_0-y^*||^2 ,$$

which means that for $0<t<\frac{1}{\sqrt{2}L}$ we obtain

$$\sum_{s=1}^{N}||y_s-\hat{y}_s||^2<\infty \quad , \quad \sum_{s=1}^{N}||\hat{y}_s-y_{s+1}||^2<\infty.$$

In other words we have

(a) $||y_s-\hat{y}_s||\to 0$ and (b)$||\hat{y}_s-y_{s+1}||\to 0$.

It follows from (47) that $\{||y_s-y^*||\}_{s=1}^{\infty}$ is a monotone decreasing sequence, hence the sequence $\{y_s\}_{s=0}^{\infty}$ is bounded. Therefore there exists a convergent subsequence $\{y_{s_i}\}_{s_i\ge1}^{\infty}$, *i.e.* $\lim_{s_i\to\infty}y_{s_i}=\bar{y}$. Due to (a) we have $\lim_{s_i\to\infty}\hat{y}_{s_i}=\bar{y}$ and due to (b) we have $\lim_{s_i\to\infty}y_{s_i+1}=\bar{y}$. Keeping in mind the continuity of the operator g we obtain

$$\begin{aligned}\bar{y}=\lim_{s_i\to\infty}y_{s_i+1}&=\lim_{s_i\to\infty}[y_{s_i}+tg(\hat{y}_{s_i})]_+\\&=[\bar{y}+tg(\bar{y})]_+,\end{aligned}$$

i.e. $\bar{y}=P_\Omega(\bar{y}+tg(\bar{y}))$ for $t>0$. Therefore from (20) follows $\bar{y}=y^*$, which together with $||y_{s+1}-y^*||<||y_s-y^*||$ for $s\ge1$ leads to $\lim_{s\to\infty}y_s=y^*$. The proof of Theorem 2 is completed. □

Remark 2. *From (47) for any* $0 < t < (\sqrt{2}L)^{-1}$ *we have*

$$||y_{s+1} - y^*||^2 + (1 - 2(tL)^2)(||y_s - \hat{y}_s||^2 + ||\hat{y}_s - y_{s+1}||^2) \leq ||y_s - y^*|| \tag{48}$$

Using $||a - b||^2 \leq 2(||a - c||^2 + ||c - b||^2)$ *with* $a = y_s$, $b = y_{s+1}$, $c = \hat{y}_s$ *and* $\mu(t) = 0.5(1 - 2(tL)^2)$ *from (48) we obtain*

$$||y_{s+1} - y^*||^2 \leq ||y_s - y^*||^2 - \mu(t)\, ||y_s - y^* - (y_{s+1} - y^*)||^2 . \tag{49}$$

Using the triangle inequality

$$||y_s - y_{s+1}|| \geq ||y_s - y^*|| - ||y_{s+1} - y^*||$$

we can rewrite (48) as follows

$$||y_{s+1} - y^*||^2 \leq ||y_s - y^*||^2 - \mu(t)(||y_s - y^*|| - ||y_{s+1} - y^*||)^2$$

Let $r = ||y_{s+1} - y^*||\, ||y_s - y^*||^{-1}$ *then the last inequality we can rewrite as follows*

$$(1 + \mu(t))r^2 - 2\mu(t)r + (\mu(t) - 1) \leq 0,$$

which leads to

$$\sup_{s \geq 1} ||y_{s+1} - y^*||\, (||y_s - y^*||)^{-1} = q \leq 1. \tag{50}$$

In the following section we show that under local strong monotonicity (3)-(4) and Lipschitz condition (31) the EPG method (32)-(33) converges globally with Q-linear rate, i.e. (50) takes place with $0 < q < 1$.

Moreover the EPG has a better ratio and in a number of instances much better complexity bound than the PPG.

6 Convergence rate of the EPG method

It follows from (3) and (4) and Lemma 1 that for $\gamma = \min\{\alpha, \beta\}$ we have

$$(g(y) - g(y^*), y - y^*) \leq -\gamma\, ||y - y^*||^2 \quad , \forall y \in \Omega \tag{51}$$

or

$$(g(y), y - y^*) - (g(y^*), y - y^*) \leq -\gamma\, ||y - y^*||^2 , \forall y \in \Omega.$$

Keeping in mind that $(g(y^*), y - y^*) \leq 0$, $\forall y \in \Omega$ from (51) we obtain

$$(g(y), y - y^*) \leq -\gamma\, ||y - y^*||^2 \quad \forall y \in \Omega. \tag{52}$$

Theorem 4. *If (3) and (4) are satisfied and Lipschitz condition (31) hold, then for* $\nu(t) = 1 + 2\gamma t - 2(tL)^2$ *and the ratio* $q(t) = 1 - 2\gamma t + 4(\gamma t)^2(\nu(t))^{-1}$ *the following bounds hold*

1) $||y_{s+1} - y^*||^2 \leq q(t)\,||y_s - y^*||^2$, $0 < q(t) < 1$, $\forall t \in (0, (\sqrt{2}L)^{-1})$,

2) *for* $t = \frac{1}{2L}$ *we have*

$$q\left(\frac{1}{2L}\right) \leq \frac{1+\varkappa}{1+2\varkappa},$$

3) *for any* $\varkappa \in [0, 0.5]$ *we have*

$$||y_{s+1} - y^*|| \leq \sqrt{1 - 0.5\varkappa}\,||y_s - y^*||\,, \tag{53}$$

4)

$$Comp(EPG) \leq O(n^2\varkappa^{-1}\ln\varepsilon^{-1}). \tag{54}$$

Proof. 1) It follows from (32)-(33), the semi-contractability of the projection operator P_Ω and Lipschitz condition (31) that

$$\begin{aligned} ||\hat{y}_s - y_{s+1}|| &= ||P_\Omega(y_s + tg(y_s)) - P_\Omega(y_s + tg(\hat{y}_s))|| \\ &\leq t\,||g(y_s) - g(\hat{y}_s)|| \\ &\leq tL\,||y_s - \hat{y}_s||\,. \end{aligned}$$

Using arguments similar to those in Theorem 3 we obtain

$$\begin{aligned} (y_s - y_{s+1}, y^* - y_{s+1}) + (y_s - \hat{y}_s, y_{s+1} - \hat{y}_s) + \\ t(g(\hat{y}_s), y^* - \hat{y}_s) - (tL)^2\,||\hat{y}_s - y_s||^2 \leq 0. \end{aligned} \tag{55}$$

From (52) with $y = \hat{y}_s$ we obtain

$$(g(\hat{y}_s), y^* - \hat{y}_s) \geq \gamma\,||\hat{y}_s - y^*||^2\,.$$

Therefore we can rewrite (55) as follows

$$\begin{aligned} &2(y_s - y_{s+1}, y^* - y_{s+1}) + 2(y_s - \hat{y}_s, y_{s+1} - \hat{y}_s) + \\ &2\gamma t\,||\hat{y}_s - y^*||^2 - 2(tL)^2\,||\hat{y}_s - y_s||^2 \leq 0. \end{aligned} \tag{56}$$

Applying identity (46) to the scalar products in (56) we obtain

$$\begin{aligned} ||y_s - y_{s+1}||^2 + ||y_{s+1} - y^*||^2 - ||y_s - y^*||^2 + ||y_s - \hat{y}_s||^2 + \\ ||\hat{y}_s - y_{s+1}||^2 - ||y_s - y_{s+1}||^2 + 2\gamma t\,||\hat{y}_s - y^*|| - 2(tL)^2\,||y_s - \hat{y}s||^2 \leq 0, \end{aligned}$$

or

$$||y_{s+1} - y^*||^2 + ||\hat{y}_s - y_{s+1}||^2 + (1 - 2(tL)^2)\,||y_s - \hat{y}_s||^2 + \tag{57}$$
$$2\gamma t\,||\hat{y}_s - y^*||^2 \le ||y_s - y^*||^2 .$$

Using

$$\begin{aligned}||\hat{y}_s - y^*||^2 &= (\hat{y}_s - y_s + y_s - y^*, \hat{y}_s - y_s + y_s - y^*)\\ &= ||\hat{y}_s - y_s||^2 + 2(\hat{y}_s - y_s, y_s - y^*) + ||y_s - y^*||^2\end{aligned}$$

we can rewrite (57) as follows

$$||y_{s+1} - y^*||^2 + ||\hat{y}_s - y_{s+1}||^2 + (1 - 2(tL)^2)\,||\hat{y}_s - y_s||^2 +$$
$$2\gamma t\,||\hat{y}_s - y_s||^2 + 4\gamma t(\hat{y}_s - y_s, y_s - y^*) + 2\gamma t\,||y_s - y^*||^2 \le ||y_s - y^*||^2 ,$$

or

$$||y_{s+1} - y^*||^2 + ||\hat{y}_s - y_{s+1}||^2 + (1 + 2\gamma t - 2(tL)^2)\,||\hat{y}_s - y_s||^2 +$$
$$4\gamma t(\hat{y}_s - y_s, y_s - y^*) \le (1 - 2\gamma t)\,||y_s - y^*||^2 . \tag{58}$$

By introducing $\nu(t) = 1 + 2\gamma t - 2(tL)^2$ we can rewrite the third and fourth term of the left hand side as follows

$$\left\|\sqrt{\nu(t)}(\hat{y}_s - y_s) + (y_s - y^*)\frac{\gamma t}{\sqrt{\nu(t)}}\right\|^2 - \frac{4(\gamma t)^2\,||y_s - y^*||^2}{\nu(t)}.$$

Therefore from (58) we have

$$||y_{s+1} - y^*||^2 + ||\hat{y}_s - y_{s+1}||^2 +$$
$$\left\|\sqrt{\nu(t)}(\hat{y}_s - y_s) + (y_s - y^*)\frac{\gamma t}{\sqrt{\nu(t)}}\right\|^2 \le \left(1 - 2\gamma t + \frac{4(\gamma t)^2}{\nu(t)}\right)||y_s - y^*||^2$$

Hence, for $q(t) = 1 - 2\gamma t + 4(\gamma t)^2(\nu(t))^{-1}$, we obtain

$$||y_{s+1} - y^*||^2 \le q(t)\,||y_s - y^*||^2 .$$

2) For $t = \frac{1}{2L}$ and $\varkappa = \gamma L^{-1}$ we have

$$q\left(\frac{1}{2L}\right) = 1 - \varkappa + \frac{\varkappa^2}{0.5 + \varkappa} = \frac{1 + \varkappa}{1 + 2\varkappa} \tag{59}$$

It is easy to see that for every $t \in (0, (\sqrt{2}L)^{-1})$ we have $0 < q(t) < 1$.

3) It follows from (59) that for any $0 \leq \varkappa \leq 0.5$ we have

$$q\left(\frac{1}{2L}\right) \leq 1 - 0.5\varkappa$$

Therefore the bound (53) holds.

4) It follows from (53) that for a given accuracy $0 < \epsilon \ll 1$ and $q = \sqrt{1-0.5\varkappa}$ the EPG method requires $s = \ln \epsilon^{-1}(\ln q^{-1})^{-1}$ step to get $y_s : \|y_s - y*\| \leq \epsilon$. It follows from (32)-(33) and Remark 1 that each step of EPG requires $O(n^2)$ operations per step, therefore the overall complexity of the EPG method is bounded by $O(n^2 \ln \epsilon^{-1}(\ln q^{-1})^{-1})$.

   Then $(\ln q^{-1})^{-1} = (-\frac{1}{2}\ln(1-0.5\varkappa))^{-1}$. Due to $\ln(1+x) \leq x, \ \forall x > -1$ we obtain $\ln(1-0.5\varkappa) \leq -0.5\varkappa$, hence $-\frac{1}{2}\ln(1-0.5\varkappa) \geq 0.25\varkappa$ and $(\ln q^{-1})^{-1} \leq 4\varkappa^{-1}$.

   Therefore the overall EPG complexity is

$$Comp(EPG) \leq O(n^2\varkappa^{-1}\ln \epsilon^{-1}),$$

i.e. the bound (54) holds true. The proof of Theorem 4 is completed. □

Remark 3. *For small $\varkappa > 0$ the complexity bound (54) is much better than the PPG bound (29). On the other hand the EPG requires two projections at each step instead of one as in the case of PPG, but keeping in mind the relatively low cost to project on Ω one can still expect the EPG to be more efficient. However, in the case when $\varkappa < 1$ is close to 1 and n is large enough, then the PPG could be more efficient.*

7 Appendix

The important part of both PPG and EPG methods is the Lipschitz constant $L > 0$ in (31).

Let's find the upper bound for $L > 0$.

To simplify our considerations we assume that the matrix A is rescaled, so

$$||A||_I = \max_{1\leq j\leq n}\sum_{i=1}^{m}|a_{ij}| \leq 1 \text{ and } ||A||_{II} = \max_{1\leq i\leq m}\sum_{j=1}^{n}|a_{ij}| \leq 1. \tag{60}$$

We assume as always that the components of vector functions $c(x)$ and $b(\lambda)$ satisfy Lipschitz condition, *i.e.* for any $1 \leq j \leq n$ there is $L_{c,j}$ that

$$|c_j(x_1) - c_j(x_2)| \leq L_{c,j}\,||x_1 - x_2||\,, \quad \forall (x_1, x_2) \in \mathbb{R}^n_+ \otimes \mathbb{R}^n_+ \tag{61}$$

and for any $1 \le i \le m$ there is $L_{b,i}$ that

$$|b_i(\lambda_1) - b_i(\lambda_2)| \le L_{b,i} \left|\left|\lambda_1 - \lambda_2\right|\right|, \quad \forall(\lambda_1, \lambda_2) \in \mathbb{R}^m_+ \otimes \mathbb{R}^m_+ \tag{62}$$

Using (61) we obtain

$$\begin{aligned} ||c(x_1) - c(x_2)|| &= \sqrt{\sum_{j=1}^{n}(c_j(x_1) - c_j(x_2))^2} \le \sqrt{\sum_{j=1}^{n} L^2_{c,j} \left|\left|x_1 - x_2\right|\right|^2} \\ &\le L_c\sqrt{n\|x_1 - x_2\|^2} = L_c\sqrt{n} \left|\left|x_1 - x_2\right|\right| \end{aligned}$$

where $L_c = \max_{1 \le j \le n} L_{c,j}$.

Using (62) we obtain

$$\begin{aligned} ||b(\lambda_1) - b(\lambda_2)|| &= \sqrt{\sum_{i=1}^{m}(b_i(\lambda_1) - b_i(\lambda_2))^2} \le \sqrt{\sum_{i=1}^{m} L^2_{b,i}\|\lambda_1 - \lambda_2\|^2} \\ &\le L_b\sqrt{m\|\lambda_1 - \lambda_2\|^2} = L_b\sqrt{m} \left|\left|\lambda_1 - \lambda_2\right|\right| \end{aligned}$$

where $L_b = \max_{1 \le i \le m} L_{b,i}$. Therefore,

$$\begin{aligned} ||g(y_1) - g(y_2)|| &\le \left|\left|c(x_1) - A^T\lambda_1 - c(x_2) + A^T\lambda_2\right|\right| + ||Ax_1 - b(\lambda_1) - Ax_2 + b(\lambda_2)|| \\ &\le ||c(x_1) - c(x_2)|| + \left|\left|A^T\right|\right| ||\lambda_1 - \lambda_2|| + ||A|| \, ||x_1 - x_2|| + ||b(\lambda_1) - b(\lambda_2)|| \\ &\le L_c\sqrt{n} \, ||x_1 - x_2|| + \left|\left|A^T\right|\right| ||\lambda_1 - \lambda_2|| + ||A|| \, ||x_1 - x_2|| + L_b\sqrt{m} \, ||\lambda_1 - \lambda_2|| \\ &= (L_c\sqrt{n} + ||A||) \, ||x_1 - x_2|| + (L_b\sqrt{m} + \left|\left|A^T\right|\right|) \, ||\lambda_1 - \lambda_2|| \end{aligned} \tag{63}$$

For $||A|| = \sqrt{\lambda_{\max}(A^TA)}$ and $\left|\left|A^T\right|\right| = \sqrt{\lambda_{\max}(AA^T)}$, in view of (60) we have

$$||A|| \le \sqrt{n} \, ||A||_I \le \sqrt{n}$$

and

$$\left|\left|A^T\right|\right| \le \sqrt{m} \left|\left|A^T\right|\right|_I \le \sqrt{m}$$

Hence, from (63) follows

$$||g(y_1) - g(y_2)|| \le \sqrt{n}(L_c + 1) \, ||x_1 - x_2|| + \sqrt{m}(L_b + 1) \, ||\lambda_1 - \lambda_2||$$

Assuming $n > m$ and taking $\hat{L} = \max\{L_c, L_b\}$ we obtain

$$\begin{aligned} ||g(y_1) - g(y_2)|| &\le \hat{L}(\sqrt{n} + 1) \left[||x_1 - x_2|| + ||\lambda_1 - \lambda_2||\right] \\ &\le \sqrt{2}\hat{L}(\sqrt{n} + 1) \, ||y_1 - y_2|| \end{aligned}$$

In other words, $L \le \sqrt{2}\hat{L}(\sqrt{n} + 1) = O(\sqrt{n})$.

8 Concluding Remarks

The "symmetrization" of the classical Walras-Wald Equilibrium (7)-(8) was achieved by replacing the fixed resource vector b by the resource operator $b : \mathbb{R}^m_+ \to \mathbb{R}^m_+$ (see [10]). This lead to NE $y^* = (x^*; \lambda^*)$ (1)-(2), which is not only justifiable from the market standpoint, but opens new opportunities for finding efficient numerical methods for optimal allocation limited resources.

The PPG method (22) and the EPG method (32)-(33) can be viewed as two pricing mechanisms for finding NE $y^* = (x^*; \lambda^*)$, which make the prices $c(x^*)$ consistent with the output x^* and make the resource availability $b(\lambda^*)$ consistent with the resource prices λ^*.

Moreover we have

$$(c(x^*) - A^T\lambda^*)_j < 0 \quad \Rightarrow \quad x_j^* = 0 \tag{64}$$

$$x_j^* > 0 \quad \Rightarrow \quad (c(x^*) - A^T\lambda^*)_j = 0 \tag{65}$$

$$(Ax^* - b(\lambda^*))_i < 0 \quad \Rightarrow \quad \lambda_i^* = 0 \tag{66}$$

$$\lambda_i^* > 0 \quad \Rightarrow \quad (Ax^* - b(\lambda^*))_i = 0 \tag{67}$$

It follows from (64) that at the equilibrium the market is cleared from goods, the prices for which can't cover their production expenses. It follows from (66) that a resource has no value if the supply is greater than its demand. It follows from (65) that at the equilibrium for each product on the market the price is equal to its production expenses. It follows from (67) that for every resource in demand the supply is equal to the demand. Finally, the total cost of the goods on the market is equal to the total production cost, i.e.

$$(c(x^*), x^*) = (b(\lambda^*), \lambda^*).$$

The "symmetrization" helps to avoid the combinatorial nature of LP and on the other hand NE drastically simplifies the complexity as compared with Projected Pseudo-Gradient method for the Classical Walras-Wald Equilibrium, which requires at each step solving a quadratic programming problem:

$$P_\Omega(x + tg(x)) = \text{argmin}\{||y - (x + tg(x))|| \, |y \in \Omega\},$$

where $\Omega = \{x : Ax \leq b, x \geq 0\}$.

The complexity bounds (29) and (54) as well as the numerical results obtained show that in a number of instances finding NE by EPG method can be cheaper than solving a correspondent LP by interior point methods.

At each step the production vector x_s and the price vector λ_s are updated by simple formulas and it can be done in parallel. In other words both the PPG and EPG one can view as decomposition methods, which decompose the entire problem on simple subproblems at each step.

Both PPG and EPG can be used for very large scale resources allocation problems when both simplex and interior point methods for solving LP is difficult to use due to the necessity to solve large linear systems at each step.

Both PPG and EPG methods were tested on large scale randomly generated NE.

The numerical results obtained corroborate the theory and provide strong evidence of the feasibility of NE technique for optimal resource allocation.

Acknowledgement. The research was partially supported by NSF Grant CCF-0836338.

References

[1] Antipin, A. (2002). The Gradient and Exstragradient Approaches in Bilinear Equilibrium Programming. A. *Dorodnizin Computing Center RAS (in Russian)*

[2] Auslender, A. and Teboulle, M. (2005). Interior projection-like methods for monotone variational inequalities. *Mathematical programming*, 104(1):39-68.

[3] Censor, Y., Gibali, A., and Reich, S. (2010). Extensions of Korpelevichs extragradient method for the variational inequality problem in Euclidean space. *Optimization*, (1):1-14.

[4] Iusem, A. and Svaiter, B. (1997). A variant of Korpelevichs method for variational inequalities with a new search strategy. *Optimization*, 42(4):309-321.

[5] Khobotov, E. (1987). Modification of the extra-gradient method for solving variational inequalities and certain optimization problems. *USSR Computational Mathematics and Mathematical Physics*, 27(5):120-127.

[6] Konnov, I. (2001). *Combined relaxation methods for variational inequalities.*Springer Verlag.

[7] Korpelevich, G. (1976). Extragradient method for finding saddle points and other problems. *Matecon*, 12(4):747-756.

[8] Kuhn, H. (1956). On a theorem of Wald. *Linear Inequalities and Related Systems*, pages 256-73.

[9] Nash, J. (1951). Non-cooperative games. *The Annals of Mathematics*, 54(2):286-295.

[10] Polyak, R. (2008). Finding generalized Walras-Wald equilibrium. *Methods of Functional Analysis and Topology*, 14(3):242-254.

[11] Rosen, J. (1965). Existence and uniqueness of equilibrium point for concave n-person games. *Econometrica*, 33(3):520-534.

[12] Zuchovitsky, S., Polyak, R., and Primak, M. (1970). N-person concave game and one production model. In *Soviet Math. Dokl*, volume 11.