Contemporary Mathematics
Volume **636**, 2015
http://dx.doi.org/10.1090/conm/636/12736

Nonlinear Equilibrium for Resource Allocation Problems

Roman A. Polyak

Abstract. We consider Nonlinear Equilibrium (NE) for the optimal allocation of limited resources. The NE is a generalization of Walras-Wald equilibrium, which is equivalent to J. Nash equilibrium in n-person concave games. Finding NE is equivalent to solving a variational inequality (VI) with a monotone and smooth operator on $\Omega = \mathbb{R}^n_+ \otimes \mathbb{R}^m_+$. 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 pseudo-gradient projection (PGP) and extra pseudo-gradient (EPG) methods require $O(n^2)$ operations per step, because in both cases the main operation per step is matrix by vector multiplication.

We prove convergence, establish global Q-linear rate and estimated computational complexity for both the PGP and EPG methods under various assumption on the input data.

Both methods can be viewed as pricing mechanisms for establishing economic equilibrium. On the other hand, they are primal-dual decomposition methods.

1. Introduction

For several decades, Linear Programming (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 a priori 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 a priori 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.

2010 *Mathematics Subject Classification.* Primary 90C33, 90C47, 91B50.

Key words and phrases. Nonlinear equilibrium, Duality, Walras-Wald equilibrium, Pseudo-gradient, Extra-pseudo-gradient, Linear programming.

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

We consider an alternative to the LP approach for optimal resource allocation, which is based on the Generalized Walras-Wald Equilibrium [**15**].

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 $b(\lambda) = (b_1(\lambda), \dots, b_m(\lambda))^T$.

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

$$x^* \in \operatorname{Argmax}\{(c(x^*), x) | Ax \le b(\lambda^*), x \in \mathbb{R}^n_+\}, \tag{1.1}$$

$$\lambda^* \in \operatorname{Argmin}\{(b(\lambda^*), \lambda) | A^T\lambda \ge c(x^*), \lambda \in \mathbb{R}^m_+\} \tag{1.2}$$

nonlinear equilibrium (NE).

The primal-dual LP solution which one obtains from (1.1)–(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 [**15**].

In this paper, we relax 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 the strong monotonicity at the equilibrium y^* or just monotonicity.

The projected gradient method for convex optimization was introduced in the 60's (see [**7**],[**13**]). Some variations of this method were used in [**3**] for solving VI. The projected gradient method for convex optimization has mainly theoretical value, because even in case of linear constraints it requires solving at each step a quadratic programming problem. In case of simple feasible sets, however, the projected gradient type methods can be very efficient. In this paper, we used projected pseudo-gradient type methods for solving VI, which is equivalent to (1.1)–(1.2). Projection on Ω is a very simple procedure; therefore, the main operation per step is computing the pseudo-gradient, which requires only matrix by vector multiplication.

We show that under local strong monotonicity and Lipschitz continuity of the operators b and c the PGP converges globally with Q-linear rate and the ratio depends only on the condition number of the VI operator. We establish complexity bound for the PGP 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 strong monotonicity at the equilibrium y^*, the convergence of the PGP becomes problematic. Therefore in the second part, of the paper we consider the extra pseudo-gradient (EPG) method.

The extragradient method was first introduced by G. Korpelevich in the 70's for finding saddle points [**11**]. Over the years, it became an important tool for solving VI (see [**1**],[**2**],[**4**]-[**6**],[**8**]-[**11**] 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.

We also show that the EPG method converges to the NE y^* if both the price c and resource b operators are just monotone and satisfy 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 PGP. This is our third contribution.

The paper is organized as follows. The basic assumptions are introduced in the following section. In Section 3, we recall the difference between the classical Walras-Wald equilibrium and NE and show the equivalence of finding NE to solving a particular VI. In Section 4, we establish convergence properties and the complexity bound of the PGP. In Section 5, we prove convergence of the EPG method under minimum assumptions on the input data. In Section 6, we establish global Q-linear convergence rate and the complexity bound for the EPG. In the Appendix, we estimate the Lipschitz constant for the VI operator, which plays an important role in both the PGP 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. The following three sets of data are 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 $1 \le i \le m$, which is required to produce one item of good $1 \le j \le n$;
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 $\lambda = (\lambda_1, \dots, \lambda_i, \dots, \lambda_m)$;
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 that 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 condition for b and c, the PGP generates a primal-dual sequence, which converges to y^* with Q-linear rate [**15**].

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

$$(b(\lambda) - b(\lambda^*), \lambda - \lambda^*) \ge \beta\|\lambda - \lambda^*\|^2,\ \beta > 0,\ \forall \lambda \in \mathbb{R}^m_+, \tag{2.1}$$

$$(c(x) - c(x^*), x - x^*) \le -\alpha\|x - x^*\|^2,\ \alpha > 0\ \forall x \in \mathbb{R}^n_+. \tag{2.2}$$

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

$$\|b(\lambda) - b(\lambda^*)\| \le L_b\|\lambda - \lambda^*\|,\ \forall \lambda \in \mathbb{R}^m_+ \tag{2.3}$$

and

$$\|c(x) - c(x^*\| \le L_c\|x - x^*\|,\ \forall x \in \mathbb{R}^n_+ \tag{2.4}$$

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

We will say that the price and resource operators are well-defined if (2.1)–(2.4) hold.

The assumption (2.1) 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 (2.2) 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.

In other words, at the equilibrium the resource availability has to be sensitive to the prices variation and the prices for a product has to be sensitive to production variation.

The Lipschitz conditions (2.3)–(2.4) assume that deviation from the NE can not lead to uncontrolled changes of prices for goods and resource availability.

3. Generalized Walras-Wald Equilibrium.

In this section, we recall 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 [**14**]. He received the Nobel Prize in Economics in 1994 for his discovery. 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 renowned work "Elements of Pure Economics". Moreover, it had 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 [**12**].

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

$$x^* \in \operatorname{Argmax}\{(c(x^*), x) | Ax \leq b, x \in \mathbb{R}^n_+\}, \tag{3.1}$$

$$\lambda^* \in \operatorname{Argmin}\{(b, \lambda) | A^T\lambda \geq c(x^*), \lambda \in \mathbb{R}^m_+\}. \tag{3.2}$$

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 (3.1) and LP Duality (1947) to show the existance of $\lambda^* \in \mathbb{R}^m_+$ in (3.2).

The equivalence of H. Kuhn's version of WW equilibrium and J. Nash equilibrium in a concave n-person game was established in [**17**].

One obtains WW equilibrium from NE by assuming $b(\lambda) = b$ in (1.1)–(1.2). So the NE (1.1)–(1.2) is a natural extension of the WW equilibrium, which makes it in a sense "symmetric".

Our next step is to recall that finding NE from (1.1)–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 (1.1)–1.2) exists.

THEOREM 1. *Finding $y^* = (x^*; \lambda^*)$ from (1.1)–1.2) is equivalent to solving the following VI*

$$(g(y^*), y - y^*) \leq 0, \forall y = (x; \lambda) \in \Omega, \tag{3.3}$$

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

PROOF. For a given $y \in \Omega$, we consider the following map

$$y \to \omega(y) = \text{Argmax}\{\Phi(y, Y) | Y \in \Omega\}, \tag{3.4}$$

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

In [**15**], it was shown that NE $y^* = (x^*, \lambda^*)$ is a fixed point of the map (3.4), i.e.,

$$y^* \in \text{Argmax}\{\Phi(y^*, Y) | Y \in \Omega\}. \tag{3.5}$$

Under a fixed $y \in \Omega$, the gradient of $\Phi(y; Y)$ in Y at $Y = y$ is called a pseudo-gradient

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

For the convex optimization problem (3.5), the fact that y^* is among its solutions means that for y^* the following optimality criteria holds

$$(g(y^*), y - y^*) \leq 0, \forall y \in \Omega. \tag{3.6}$$

In other words, NE $y^* = (x^*, \lambda^*)$ is a solution of VI (3.6).

On the other hand, for any solution $\bar{y} = (\bar{x}; \bar{\lambda})$ of VI (3.6), we have

$$(g(\bar{y}), y) \leq (g(\bar{y}), \bar{y}), \forall y \in \Omega. \tag{3.7}$$

Therefore $g(\bar{y}) \leq 0$ and $\bar{y} \geq 0$, so $(g(\bar{y}), \bar{y}) \leq 0$ or

$$c(\bar{x}) \leq A^T\bar{\lambda}, \bar{x} \geq 0 \ \textit{and} \ A\bar{x} \leq b(\bar{\lambda}), \bar{\lambda} \geq 0. \tag{3.8}$$

For $y = 0$ from (3.7) follows

$$(g(\bar{y}), \bar{y}) \geq 0, \tag{3.9}$$

hence

$$(g(\bar{y}), \bar{y}) = 0. \tag{3.10}$$

It follows from (3.8) that $\bar{x}$ is a feasible solution for the primal LP

$$\max\{(c(\bar{x}), x) | Ax \leq b(\bar{\lambda}), x \geq 0\} \tag{3.11}$$

and $\bar{\lambda}$ is a feasible solution for the dual LP

$$\min\{(b(\bar{\lambda}), \lambda) | A^T\lambda \geq c(\bar{x}), \lambda \geq 0\}. \tag{3.12}$$

It follows from (3.10) that for the primal feasible solution $\bar{x}$ and dual feasible solution $\bar{\lambda}$ the complementarity conditions

$$(A\bar{x} - b(\bar{\lambda}), \bar{\lambda}) = 0, (A^T\bar{\lambda} - c(\bar{x}), \bar{x}) = 0$$

are satisfied.

Therefore $\bar{x}$ solves (3.11) and $\bar{\lambda}$ solves (3.12), i.e., $\bar{x} = x^*$ and $\bar{\lambda} = \lambda^*$. □

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

$$x^* \in D : (c(x^*), x - x^*) \leq 0, \forall x \in D. \tag{3.13}$$

Solving (3.13), generally speaking, is more difficult than solving the corresponding primal-dual LP, i.e., (1.1)–1.2) when $b(\lambda) \equiv b$ and $c(x) \equiv c$. It may seem that finding NE $y^* = (x^*, \lambda^*)$ is more difficult than solving the VI (3.13).

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

The fundamental difference between NE (1.1)–1.2) and WW (3.13) follows from the geometry of their feasible sets $\Omega = \mathbb{R}^n_+ \otimes \mathbb{R}^m_+$ and D. The simplicity of Ω makes pseudo-gradient projection type methods particularly suitable for solving VI (3.3), because it requires matrix by vector multiplication as the main operation per step whereas pseudo-gradient projection methods for solving VI (3.13) require solving a quadratic programming problem at each step.

In the following sections, we will concentrate on a pseudo-gradient projection method for solving VI (3.3).

4. Pseudo-Gradient Projection Method

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

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

The vector v is called a projection of u on Q. Later we will need the following two well-known properties of the projection operator P_Q (see, for example, [**15**]).

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

$$\|P_Q(u_1) - P_Q(u_2)\| \leq \|u_1 - u_2\|, \forall u_1, u_2 \in \mathbb{R}^q \tag{4.1}$$

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

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

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

$$u^* = P_Q(u^* + tg(u^*)). \tag{4.2}$$

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

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

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

$$g(y) = (c(x) - A^T\lambda; Ax - b(\lambda)). \tag{4.3}$$

We are ready to describe the PGP method for solving the VI (3.3). 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 already been found. The PGP method finds the next approximation y_{s+1} by the formula

$$y_{s+1} = P_\Omega(y_s + tg(y_s)). \tag{4.4}$$

In other words, each step of the PGP method consists of updating the production vector x_s and the price vector λ_s by the following formulas:

$$x_{j,s+1} = [x_{j,s} + t(c(x_s) - A^T\lambda_s)_j]_+,\ j = 1, \dots, n, \tag{4.5}$$

$$\lambda_{i,s+1} = [\lambda_{i,s} + t(Ax_s - b(\lambda_s))_i]_+,\ i = 1, \dots, m. \tag{4.6}$$

The step length $t > 0$ will be specified later. The method (4.5)–4.6) 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 PGP method (4.5)–4.6) can be viewed as a pricing mechanism for finding equilibrium.

It follows from (4.5) 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 (4.6) 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 strongly monotone at λ^* and x^*, i.e., (2.1)–2.2) hold, then the operator g is strongly monotone at y^* and for $\gamma = \min\{\alpha, \beta\}$ the following inequality holds:*

$$(g(y) - g(y^*), y - y^*) \le -\gamma\|y - y^*\|^2,\ \forall y \in \Omega. \tag{4.7}$$

PROOF. We have

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

Using (2.1) and (2.2) for $\gamma = \min\{\alpha, \beta\}$, we obtain (4.7) □

LEMMA 2. *If b and c satisfy Lipschitz conditions (2.3)-(2.4) at λ^* and x^*, then the operator $g : \Omega \to \mathbb{R}^{n+m}$ given by (4.3) satisfies Lipschitz condition at y^*, i.e., there is an $L > 0$ such that*

$$\|g(y) - g(y^*)\| \le L\|y - y^*\|,\ \forall y \in \Omega. \tag{4.8}$$

For the proof of Lemma 2 and the upper bound for L see the Appendix.

REMARK 2. We will assume later that for a given $x \in \mathbb{R}^n$ finding c(x) does not require more than $O(n^2)$ operations, and for a given $\lambda \in \mathbb{R}^m$ finding $b(\lambda)$ does not require more than $O(m^2)$ operations. We also assume that $n \ge m$. From (4.5)-(4.6) follows that each step of the PGP method (4.4) does not require more than $O(n^2)$ operations.

EXAMPLE 3. Let $c(x) = \nabla(\frac{1}{2}x^TCx + c^Tx)$ and $b(\lambda) = \nabla(\frac{1}{2}\lambda^TB\lambda + b^T\lambda)$, 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 PGP method (4.4) 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 PGP method (4.4).

THEOREM 4. *If operators b and c are well-defined i.e., (2.1)-(2.4) hold then:*

1) *for any $0 < t < 2\gamma L^{-2}$ the PGP method (4.4) globally converges to NE $y^* = (x^*; \lambda^*)$ with Q-linear rate and the ratio $0 < q(t) = (1 - 2t\gamma + t^2L^2)^{1/2} < 1$, i.e.,*

$$\|y_{s+1} - y^*\| \le q(t)\|y_s - y^*\|; \tag{4.9}$$

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

$$\|y_{s+1} - y^*\| \le (1 - \varkappa^2)^{1/2}\|y_s - y^*\|; \tag{4.10}$$

3) *for the PGP complexity, we have the following bound*

$$\mathrm{Comp}(PGP) = O(n^2\varkappa^{-2}\ln\varepsilon^{-1}), \tag{4.11}$$

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

PROOF. 1) From (4.4), non-expansive property of operator P_Ω (4.1) and optimality criteria (4.2) follows

$$\begin{aligned}\|y_{s+1} - y^*\|^2 &= \|P_\Omega(y_s + tg(y_s)) - P_\Omega(y^* + tg(y^*))\|^2\\ &\le \|y_s + tg(y_s) - y^* - tg(y^*)\|^2\\ &= (y_s - y^* + t(g(y_s) - g(y^*)), y_s - y^* + t(g(y_s) - g(y^*)))\\ &= \|y - y^*\|^2 + 2t(y_s - y^*, g(y_s) - g(y^*))\\ &\quad + t^2\|g(y_s) - g(y^*)\|^2. \end{aligned} \tag{4.12}$$

For well-defined b and c from (4.7) (4.8) and (4.12), we obtain

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

Hence for $0 < t < 2\gamma L^{-2}$, we have $0 < q(t) = (1 - 2t\gamma + t^2L^2)^{\frac{1}{2}} < 1$. In other words, the projection operator (4.4) is contractive, which means that for any given $t \in (0, 2\gamma L^{-2})$ the PGP method globally converges with Q-linear rate, i.e., (4.9) holds.

2) For $t = \gamma L^{-2} = \mathrm{argmin}\{q(t)|t > 0\}$, we have

$$q = q(\gamma L^{-2}) = (1 - (\gamma L^{-1})^2)^{\frac{1}{2}} = (1 - \varkappa^2)^{\frac{1}{2}},$$

i.e., (4.10) holds.

3) Let $0 < \varepsilon \ll 1$ be the required accuracy, then in view of (4.10) it takes $O((\ln q)^{-1}\ln\varepsilon)$ steps to find an ε-approximation for the NE $y^* = (x^*, \lambda^*)$. It follows from Remark 2 that each PGP step (4.4) does not require more than $O(n^2)$ operations. Therefore, finding the ε-approximation to NE $y^* = (x^*, \lambda^*)$ requires

$$N = O\left(n^2\frac{\ln\varepsilon}{\ln q}\right) = O\left(n^2\frac{\ln\varepsilon^{-1}}{\ln q^{-1}}\right)$$

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 PGP method we obtain (4.11). □

If $\gamma = \min\{\alpha, \beta\} = 0$, then pseudo-gradient $g : \Omega \to \mathbb{R}^{m+n}$ defined by (4.3) is not even locally strongly monotone, therefore (4.9) cannot guarantee convergence of the PGP method (4.4). In the following section, we consider the extra pseudo-gradient method (EPG) for finding NE (1.1)-(1.2) in the absence of local strong monotonicity of both operators b and c.

The extragradient method was first introduced by G. Korpelevich ([**11**]) in the 70s for finding saddle points. Lately, it became a popular tool for solving VI (see [**1**],[**2**],[**4**]-[**6**],[**8**]-[**11**] 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.,

$$\|g(y_1) - g(y_2)\| \leq L\|y_1 - y_2\|, \ \forall y_1, y_2 \in \Omega. \tag{4.13}$$

5. Extra Pseudo-Gradient Method for finding NE

The application of G. Korpelevich's extragradient method [**11**] for solving VI (3.3) leads to the following extra pseudo-gradient (EPG) method for finding NE $y^* = (x^*; \lambda^*)$. Each step of the EPG method consists in two phases: the predictor phase and the corrector phase. We start with initial approximation $y_0 = (x_0; \lambda_0)) \in \mathbb{R}^n_{++} \otimes \mathbb{R}^m_{++}$. Let assume that the vector $y_s = (x_s; \lambda_s)$ has been found already.

The predictor phase consists of finding

$$\hat{y}_s = P_\Omega(y_s + tg(y_s)) = [y_s + tg(y_s)]_+. \tag{5.1}$$

The corrector phase finds the new approximation

$$y_{s+1} = P_\Omega(y_s + tg(\hat{y}_s)) = [y_s + tg(\hat{y}_s)]_+. \tag{5.2}$$

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

In other words, the first phase predicts the new production vector

$$\hat{x}_s = [x_s + t(c(x_s) - A^T\lambda_s)]_+ \tag{5.3}$$

and a new price vector

$$\hat{\lambda}_s = [\lambda_s + t(Ax_s - b(\lambda_s))]_+. \tag{5.4}$$

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 second phase corrects the production vector

$$x_{s+1} = [x_s + t(c(\hat{x}_s) - A^T\hat{\lambda}_s)]_+ \tag{5.5}$$

and the price vector

$$\lambda_{s+1} = [\lambda_s + t(A\hat{x}_s - b(\hat{\lambda}_s))]_+. \tag{5.6}$$

The meaning of the formulas (5.3)-(5.4) and (5.5)-(5.6) is similar to the meaning of the formulas (4.5)-(4.6).

The formulas (5.1)-(5.2) can be viewed as a pricing mechanism for finding the NE $y^* = (x^*; \lambda^*)$.

THEOREM 5. *If c and b are monotone operators and Lipschitz condition (4.13) is satisfied, then for any $t \in (0, (\sqrt{2}L)^{-1})$ the EPG method (5.1)-(5.2) generates a convergent sequence $\{y_s\}_{s=1}^{\infty}$ and $\lim_{s\to\infty} y_s = y^*$.*

PROOF. Let us consider vector $h_s = y_s + tg(y_s) - \hat{y}_s$, then from (5.1) we have

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

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

$$(tg(y_s) + (y_s - \hat{y}_s), y - \hat{y}_s) \leq 0. \tag{5.7}$$

For $h_{s+1} = y_s + tg(\hat{y}_s) - y_{s+1}$ from (5.2) follows $(h_{s+1}, y - y_{s+1}) \leq 0$, $\forall y \in \Omega$.

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

$$(tg(\hat{y}_s) + (y_s - y_{s+1}), y - y_{s+1}) \leq 0. \tag{5.8}$$

From (5.1), (5.2) and non-expansive property of the operator P_Ω, which is defined by (4.1), as well as Lipschitz condition (4.13), we obtain

$$\begin{aligned} ||y_{s+1} - \hat{y}_s|| &= ||P_\Omega(y_s + tg(\hat{y}_s)) - P_\Omega(y_s + tg(y_s))|| \\ &\leq t\,||g(\hat{y}_s) - g(y_s)|| \\ &\leq tL\,||\hat{y}_s - y_s||\,. \end{aligned} \tag{5.9}$$

From (5.8) for $y = y^*$ we have

$$(y_s - y_{s+1} + tg(\hat{y}_s), y^* - y_{s+1}) \leq 0. \tag{5.10}$$

By taking $y = y_{s+1}$ in (5.7), 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

$$(y_s - \hat{y}_s, y_{s+1} - \hat{y}_s) + t(g(\hat{y}_s), y_{s+1} - \hat{y}_s) - t(g(\hat{y}_s) - g(y_s), y_{s+1} - \hat{y}_s) \leq 0. \tag{5.11}$$

Then using (5.9), 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 (5.11) we have

$$(y_s - \hat{y}_s, y_{s+1} - \hat{y}_s) + t(g(\hat{y}_s), y_{s+1} - \hat{y}_s) - (tL)^2\,||\hat{y}_s - y_s||^2 \leq 0. \tag{5.12}$$

By adding (5.10) and (5.12), we obtain

$$\begin{aligned} (y_s - y_{s+1}, y^* - y_{s+1}) + t(g(\hat{y}_s), y^* - y_{s+1}) + (y_s - \hat{y}_s, y_{s+1} - \hat{y}_s) \\ + t(g(\hat{y}_s), y_{s+1} - \hat{y}_s) - (tL)^2\,||\hat{y}_s - y_s||^2 \\ = (y_s - y_{s+1}, y^* - y_{s+1}) + t(g(\hat{y}_s), y^* - \hat{y}_s) \\ + (y_s - \hat{y}_s, y_{s+1} - \hat{y}_s) - (tL)^2\,||y_s - \hat{y}_s||^2 \leq 0. \end{aligned} \tag{5.13}$$

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 (5.13) and using the monotonicity inequality

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

from (5.13), we obtain

$$2(y_s - y_{s+1}, y^* - y_{s+1}) + 2(y_s - \hat{y}_s, y_{s+1} - \hat{y}_s) - 2(tL)^2 ||\hat{y}_s - y_s||^2 \leq 0. \tag{5.14}$$

Using identity

$$2(u - v, w - v) = ||u - v||^2 + ||v - w||^2 - ||u - w||^2 \tag{5.15}$$

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 (5.14) as follows:

$$||y_{s+1} - y^*||^2 + (1 - 2(tL)^2) ||y_s - \hat{y}_s||^2 + ||\hat{y}_s - y_{s+1}||^2 \leq ||y_s - y^*||^2 . \tag{5.16}$$

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

$$||y_{N+1} - y^*||^2 + (1 - 2(tL)^2) \sum_{s=0}^{N} ||y_s - \hat{y}_s||^2 + \sum_{s=0}^{N} ||\hat{y}_s - y_{s+1}||^2 \leq ||y_0 - y^*||^2 ,$$

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

$$\sum_{s=0}^{N} ||y_s - \hat{y}_s||^2 < \infty \quad , \quad \sum_{s=0}^{N} ||\hat{y}_s - y_{s+1}||^2 < \infty.$$

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 (5.16) 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 $\{y_{s_i}\}_{s_i \geq 1}^{\infty}$, i.e., $\lim\limits_{s_i \to \infty} y_{s_i} = \bar{y}$. Due to (a), we have $\lim\limits_{s_i \to \infty} \hat{y}_{s_i} = \bar{y}$, and due to (b) we have $\lim\limits_{s_i \to \infty} y_{s_i+1} = \bar{y}$. 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 (4.2) follows $\bar{y} = y^*$, which together with $||y_{s+1} - y^*|| < ||y_s - y^*||$ for $s \geq 1$ leads to $\lim\limits_{s \to \infty} y_s = y^*$. The proof of Theorem 2 is completed. □

REMARK 6. From (5.16) for any $0 < t < (\sqrt{2}L)^{-1}$, we have

$$||y_{s+1} - y^*||^2 + (1 - 2(tL)^2)(||y_s - \hat{y}_s||^2 + ||\hat{y}_s - y_{s+1}||^2) \leq ||y_s - y^*||^2 . \tag{5.17}$$

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 (5.17), we obtain

$$||y_{s+1} - y^*||^2 \leq ||y_s - y^*||^2 - \mu(t) ||y_s - y^* - (y_{s+1} - y^*)||^2 . \tag{5.18}$$

Using the triangle inequality

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

we can rewrite (5.17) 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 we can rewrite the last inequality as follows:

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

which leads to

$$\sup_{s \geq 1} ||y_{s+1} - y^*|| \, (||y_s - y^*||)^{-1} = q \leq 1. \tag{5.19}$$

In the following section, we show that under local strong monotonicity (2.1)–(2.2) and Lipschitz condition (4.13), the EPG method (5.1)–(5.2) converges globally with Q-linear rate, i.e., (5.19) 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 (2.1), (2.2) and Lemma 1 that for $\gamma = \min\{\alpha, \beta\}$, we have

$$(g(y) - g(y^*), y - y^*) \leq -\gamma \, ||y - y^*||^2 \quad , \forall y \in \Omega \tag{6.1}$$

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 (6.1), we obtain

$$(g(y), y - y^*) \leq -\gamma \, ||y - y^*||^2 \quad , \forall y \in \Omega. \tag{6.2}$$

THEOREM 7. *If (2.1) and (2.2) are satisfied and the Lipschitz condition (4.13) holds, 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*

$$q\left(\frac{1}{2L}\right) = \frac{1 + \varkappa}{1 + 2\varkappa};$$

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

$$||y_{s+1} - y^*|| \leq \sqrt{1 - 0.5\varkappa} \, ||y_s - y^*|| \, ; \tag{6.3}$$

4)

$$\text{Comp}(EPG) \leq O(n^2 \varkappa^{-1} \ln \varepsilon^{-1}). \tag{6.4}$$

PROOF. 1) It follows from (5.1)–(5.2), the non-expansive property of the projection operator P_Ω and Lipschitz condition (4.13) 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 in the proof of Theorem 5, we obtain

$$\begin{aligned} &(y_s - y_{s+1}, y^* - y_{s+1}) + (y_s - \hat{y}_s, y_{s+1} - \hat{y}_s) \\ &+t(g(\hat{y}_s), y^* - \hat{y}_s) - (tL)^2 \, ||\hat{y}_s - y_s||^2 \leq 0. \end{aligned} \tag{6.5}$$

From (6.2) 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 (6.5) as follows:

$$\begin{aligned} (6.6) \qquad & 2(y_s - y_{s+1}, y^* - y_{s+1}) + 2(y_s - \hat{y}_s, y_{s+1} - \hat{y}_s) \\ & +2\gamma t \, ||\hat{y}_s - y^*||^2 - 2(tL)^2 \, ||\hat{y}_s - y_s||^2 \leq 0. \end{aligned}$$

Applying identity (5.15) to the scalar products in (6.6), 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

$$\begin{aligned} (6.7) \qquad & ||y_{s+1} - y^*||^2 + ||\hat{y}_s - y_{s+1}||^2 + (1 - 2(tL)^2) \, ||y_s - \hat{y}_s||^2 \\ & +2\gamma t \, ||\hat{y}_s - y^*||^2 \leq ||y_s - y^*||^2 . \end{aligned}$$

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 (6.7) as follows:

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

or

$$\begin{aligned} & ||y_{s+1} - y^*||^2 + ||\hat{y}_s - y_{s+1}||^2 + (1 + 2\gamma t - 2(tL)^2) \, ||\hat{y}_s - y_s||^2 \\ (6.8) \qquad & +4\gamma t(\hat{y}_s - y_s, y_s - y^*) \leq (1 - 2\gamma t) \, ||y_s - y^*||^2 . \end{aligned}$$

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:

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

Therefore from (6.8) we have

$$\begin{aligned} & ||y_{s+1} - y^*||^2 + ||\hat{y}_s - y_{s+1}||^2 \\ & + \left\| \sqrt{\nu(t)}(\hat{y}_s - y_s) + 2(y_s - y^*) \frac{\gamma t}{\sqrt{\nu(t)}} \right\|^2 \leq \left(1 - 2\gamma t + \frac{4(\gamma t)^2}{\nu(t)} \right) ||y_s - y^*||^2 . \end{aligned}$$

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

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

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

$$(6.9) \qquad q\left(\frac{1}{2L}\right) = 1 - \varkappa + \frac{\varkappa^2}{0.5 + \varkappa} = \frac{1 + \varkappa}{1 + 2\varkappa}.$$

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 (6.9) 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 (6.3) holds.

4) It follows from (6.3) that for a given accuracy $0 < \epsilon \ll 1$ and $q = \sqrt{1 - 0.5\varkappa}$, the EPG method requires that

$$s = O\left(\frac{\ln \epsilon^{-1}}{\ln q^{-1}}\right)$$

steps to get $y_s : \|y_s - y*\| \leq \epsilon$.

It follows from (5.1)–(5.2) and Remark 2 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

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

i.e., the bound (6.4) holds true. □

REMARK 8. For small $\varkappa > 0$, the complexity bound (6.4) is much better than the PGP bound (4.11). On the other hand, the EPG requires two projections at each step instead of one, as in the case of PGP, 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 $1 > \varkappa > 0.5$ and n is large enough, then the PGP could be more efficient.

7. Appendix

The important part of both the PGP and EPG methods is the Lipschitz constant $L > 0$ in (4.13).

Let us find an upper bound for $L > 0$.

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

$$||A||_I = \max_{1\leq j\leq n}\sum_{i=1}^{m}|a_{ij}| \leq 1 \text{ and } ||A||_{II} = \max_{1\leq i\leq m}\sum_{j=1}^{n}|a_{ij}| \leq 1. \tag{7.1}$$

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

$$|c_j(x_1) - c_j(x_2)| \leq L_{c,j}\,||x_1 - x_2||\,, \quad \forall (x_1, x_2) \in \mathbb{R}^n_+ \otimes \mathbb{R}^n_+ \tag{7.2}$$

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

$$|b_i(\lambda_1) - b_i(\lambda_2)| \leq L_{b,i}\,||\lambda_1 - \lambda_2||\,, \quad \forall (\lambda_1, \lambda_2) \in \mathbb{R}^m_+ \otimes \mathbb{R}^m_+. \tag{7.3}$$

Using (7.2), we obtain

$$||c(x_1) - c(x_2)|| = \sqrt{\sum_{j=1}^{n}(c_j(x_1) - c_j(x_2))^2} \leq \sqrt{\sum_{j=1}^{n}L^2_{c,j}\,||x_1 - x_2||^2}$$
$$\leq L_c\sqrt{n\|x_1 - x_2\|^2} = L_c\sqrt{n}\,||x_1 - x_2||$$

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

Using (7.3), we obtain

$$||b(\lambda_1) - b(\lambda_2)|| = \sqrt{\sum_{i=1}^{m}(b_i(\lambda_1) - b_i(\lambda_2))^2} \le \sqrt{\sum_{i=1}^{m} L_{b,i}^2 \|\lambda_1 - \lambda_2\|^2}$$
$$\le L_b\sqrt{m\|\lambda_1 - \lambda_2\|^2} = L_b\sqrt{m}\,||\lambda_1 - \lambda_2||$$

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

(7.4)

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

For $||A|| = \sqrt{\lambda_{\max}(A^TA)}$ and $\left|\left|A^T\right|\right| = \sqrt{\lambda_{\max}(AA^T)}$, in view of (7.1), 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 (7.4) 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

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

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 (3.1)–(3.2) was achieved by replacing the fixed resource vector b by the resource operator $b : \mathbb{R}^m_+ \to \mathbb{R}^m_+$ (see [**15**]). This is not only justifiable from the market standpoint but it leads to new methods, which are based on projection type techniques for solving VI. 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, one can view both PGP and EPG as primal-dual decomposition methods.

The complexity bounds (4.11) and (6.4) show that in a number of instances finding NE by PGP or EPG can be cheaper than solving a correspondent LP by interior point methods.

Both PGP and EPG can be used for very large scale resources allocation problems when simplex or interior point methods for solving LP are difficult to use due to the necessity of solving large linear systems of equations at each step.

The "symmetrization" also helps to avoid the combinatorial nature of LP. On the other hand, finding NE drastically reduces the complexity as compared with

using PGP or EPG for finding Walras-Wald Equilibrium, which requires at each step solving one or two quadratic programming problem:

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

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

Both the PGP and the EPG can be viewed as pricing mechanisms for finding NE, which make the prices $c(x^*)$ consistent with the output x^* and the resource availability $b(\lambda^*)$ consistent with the resource prices λ^*.

Moreover, we have

$$(c(x^*) - A^T\lambda^*)_j < 0 \quad \Rightarrow \quad x_j^* = 0 \tag{8.1}$$

$$x_j^* > 0 \quad \Rightarrow \quad (c(x^*) - A^T\lambda^*)_j = 0 \tag{8.2}$$

$$(Ax^* - b(\lambda^*))_i < 0 \quad \Rightarrow \quad \lambda_i^* = 0 \tag{8.3}$$

$$\lambda_i^* > 0 \quad \Rightarrow \quad (Ax^* - b(\lambda^*))_i = 0 \tag{8.4}$$

It follows from (8.1) that at the equilibrium the market is cleared from goods, the prices for which can not cover their production expenses. It follows from (8.3) that a resource has no value if the supply is greater than its demand. It follows from (8.2) that at the equilibrium for each product on the market the price is equal to its production expenses. It follows from (8.4) that for every resource in demand the supply is equal to the demand.

Finally, at the equilibrium 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 complexity bounds (4.11) and (6.4), as well as the numerical results obtained, show that in a number of instances finding NE by the EPG method can be cheaper than solving a correspondent LP by interior point methods.

We have some encouraging numerical results, but the new technology for solving RAP needs much more numerical work and economic analysis before it will become a practical tool.

Acknowledgement. The author is grateful to the reviewers for valuable comments.

References

[1] A. Antipin, *The gradient and exstragradient approaches in bilinear equilibrium programming*, A. Dorodnizin Computing Center RAS (in Russian), 2002.

[2] A. Auslender and M. Teboulle, *Interior projection-like methods for monotone variational inequalities*, Math. Program. **104** (2005), no. 1, Ser. A, 39–68, DOI 10.1007/s10107-004-0568-x. MR2167425 (2006d:90232)

[3] A.B. Bakushinskij and B.T. Polyak, *On the solution of variational inequalities*, Sov. Math Doklady **14** (1974), 1705-1710.

[4] Y. Censor, A. Gibali, and S. Reich, *Extensions of Korpelevich's extragradient method for the variational inequality problem in Euclidean space*, Optimization **61** (2012), no. 9, 1119–1132, DOI 10.1080/02331934.2010.539689. MR2966121

[5] Y. Censor, A. Gibali, and S. Reich, *The subgradient extragradient method for solving variational inequalities in Hilbert space*, J. Optim. Theory Appl. **148** (2011), no. 2, 318–335, DOI 10.1007/s10957-010-9757-3. MR2780566 (2011k:49014)

[6] Y. Censor, A. Gibali, and S. Reich, *Strong convergence of subgradient extragradient methods for the variational inequality problem in Hilbert space*, Optim. Methods Softw. **26** (2011), no. 4-5, 827–845, DOI 10.1080/10556788.2010.551536. MR2837800 (2012g:49012)

[7] A. A. Goldstein, *Convex programming in Hilbert space*, Bull. Amer. Math. Soc. **70** (1964), 709–710. MR0165982 (29 #3262)
[8] A. N. Iusem and B. F. Svaiter, *A variant of Korpelevich's method for variational inequalities with a new search strategy*, Optimization **42** (1997), no. 4, 309–321, DOI 10.1080/02331939708844365. MR1609571 (98m:49020)
[9] E. N. Khobotov, *A modification of the extragradient method for solving variational inequalities and some optimization problems* (Russian), Zh. Vychisl. Mat. i Mat. Fiz. **27** (1987), no. 10, 1462–1473, 1597. MR918546 (88m:49005)
[10] I. Konnov, *Combined relaxation methods for variational inequalities*, Lecture Notes in Economics and Mathematical Systems, vol. 495, Springer-Verlag, Berlin, 2001. MR1795730 (2001i:49001)
[11] G. M. Korpelevič, *An extragradient method for finding saddle points and for other problems* (Russian), Èkonom. i Mat. Metody **12** (1976), no. 4, 747–756. MR0451121 (56 #9408)
[12] H. W. Kuhn, *On a theorem of Wald*, Linear inequalities and related systems, Annals of Mathematics Studies, no. 38, Princeton University Press, Princeton, N. J., 1956, pp. 265–273. MR0082913 (18,629d)
[13] E. Levitin and B. Polyak, *Constrained minimization methods*, Zeh. Vychisl. Mat. i Met Fiz. **6** (1966), 787-823.
[14] J. Nash, *Non-cooperative games*, Ann. of Math. (2) **54** (1951), 286–295. MR0043432 (13,261g)
[15] R. A. Polyak, *Finding generalized Walras-Wald equilibrium*, Methods Funct. Anal. Topology **14** (2008), no. 3, 242–254. MR2458489 (2009j:91129)
[16] J. B. Rosen, *Existence and uniqueness of equilibrium points for concave n-person games*, Econometrica **33** (1965), 520–534. MR0194210 (33 #2423)
[17] S. Zuchovitsky, R. Polyak and M. Primak, *N-person concave game and one production model*, Soviet Math. Dokl. **11** (1970), no. 2, 523-526.

Department of Mathematics, The Technion - Israel Institute of Technology, 32000 Haifa, Israel
E-mail address: rpolyak@techunix.technion.ac.il