

Flam, Sjur; Ruszczynski, A.

Working Paper

Computing Normalized Equilibria in Convex-Concave Games

Working Paper, No. 2006:9

Provided in Cooperation with:

Department of Economics, School of Economics and Management, Lund University

Suggested Citation: Flam, Sjur; Ruszczynski, A. (2006) : Computing Normalized Equilibria in Convex-Concave Games, Working Paper, No. 2006:9, Lund University, School of Economics and Management, Department of Economics, Lund

This Version is available at:

<https://hdl.handle.net/10419/259930>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Computing Normalized Equilibria in Convex-Concave Games

S. D. FLÅM* AND A. RUSZCZYŃSKI†

April 26, 2006

ABSTRACT. This paper considers a fairly large class of noncooperative games in which strategies are jointly constrained. When what is called the Ky Fan or Nikaidô-Isoda function is convex-concave, selected Nash equilibria correspond to diagonal saddle points of that function. This feature is exploited to design computational algorithms for finding such equilibria.

To comply with some freedom of individual choice the algorithms developed here are fairly decentralized. However, since coupling constraints must be enforced, repeated coordination is needed while underway towards equilibrium.

Particular instances include zero-sum, two-person games - or minimax problems - that are convex-concave and involve convex coupling constraints.

Key words: Noncooperative games, Nash equilibrium, joint constraints, quasi-variational inequalities, exact penalty, subgradient projection, proximal point algorithm, partial regularization, saddle points, Ky Fan or Nikaidô-Isoda functions.

Math. Subj. Classification: 90C25, 91A10; *JEL Classification:* C63, C70.

1. INTRODUCTION

Noncooperative games are most often modeled without coupling constraints between various parties. Strategic interaction then affects preferences but leaves feasible domains invariant. An important exception came with Debreu's (1952) paper on existence of social equilibrium. There, in principle, a player's strategy may restrict other agents' freedom of choice. Later studies of such generalized settings include Rosen (1965), Bensoussan (1974), Harker (1991), Robinson (1993), Uryasev and Rubinstein (1994), Haurie and Krawczyk (2002), Pang and Fukushima (2005).

Ours is also a setting of mutually restricted choice. It is construed as a strategic-form noncooperative game, featuring a finite set I of players. Individual $i \in I$ seeks, with no collaboration, to minimize his private cost or loss $L_i(x) = L_i(x_i, x_{-i})$ with respect to own strategy x_i . As customary, $x_{-i} := (x_j)_{j \in I \setminus i}$ denotes the strategy profile taken by player i 's "adversaries."

In general, two types of constraints affect player i . For one, he must choose x_i from a fixed closed subset X_i of a finite-dimensional Euclidean space \mathbb{X}_i . For the other, his

*Corresponding author, Economics Department, Bergen University 5007 Norway; sjur.flaam@econ.uib.no. Thanks for support is due Ruhrgas and STINT.

†Rutgers University, Department of Management Science and Information Systems; rusz@business.rutgers.edu.

choice is regulated by a prescribed point-to-set correspondence $x_{-i} \mapsto C_i(x_{-i}) \subseteq \mathbb{X}_i$ in that $x_i \in C_i(x_{-i})$. Thus, any player's effective strategy set depends on his rivals' choices.

The problem addressed below is that of computing *generalized Nash equilibria*. By definition, any such equilibrium $x^* = (x_i^*)$ must satisfy, for each $i \in I$, the optimality condition that x_i^* minimizes $L_i(x_i, x_{-i}^*)$ subject to $x_i \in X_i \cap C_i(x_{-i}^*)$. To find strategy profiles $i \mapsto x_i^*$ of that sort is generally quite hard. Much simplification may obtain though, when - as assumed here - the product set

$$X := \{x = (x_i) : x_i \in X_i \cap C_i(x_{-i}) \text{ for all } i\} \quad (1)$$

is convex. This important and frequent situation was first studied by Rosen (1965). In the main part he required that the scaled marginal cost profile

$$M(r, x) := \left[r_i \frac{\partial L_i(x_i, x_{-i})}{\partial x_i} \right]_{i \in I}$$

be well defined and strictly monotone for some parameters $r_i > 0, i \in I$. By contrast, we shall accommodate nonsmooth data and rather assume, also for suitable $r = (r_i) > 0$, that

$$L(r, x, y) := \sum_{i \in I} r_i [L_i(x) - L_i(y_i, x_{-i})]$$

be convex in $x = (x_i)$ and concave in $y = (y_i)$. That assumption fits the setting of [6], [20], [21] and [33], but differs in three respects: first, data can be nonsmooth; second, we dispense with strict monotonicity of gradients, but third, we require convexity-concavity instead of merely *weak* versions of such curvature. The class at hand is larger than might first be believed.

Section 2 collects preliminaries. Thereafter we proceed to find equilibria. Throughout our enterprise the following disclaimer applies: While customary Nash equilibrium is self-enforcing, a generalized version, even when unique, need not share that desirable feature. Broadly, the reason is that a player may "hijack" the game. For instance, if some capacity is jointly constrained, a "quick" player could exhaust it by moving fast. Our concern is however, with computation, not enforcement. Reflecting on this, Section 3 brings out two new algorithms, both using partial regularizations, relaxed subgradient projections and averages of proposed solutions. These algorithms are specialized versions of general saddle-point methods developed in [19] and [30]. Section 4 proves convergence, and Section 5 displays example games.

2. PRELIMINARIES

Recent research on generalized Nash equilibrium has studied existence by means of quasi-variational inequalities [5], [15]. To solve such inequalities is typically hard. Accordingly, numerical methods are fairly few - and their practicality so far not clear [23], [24]. It deserves notice though, that exact penalty methods, when applicable,

may have much to offer. For a precise statement denote by $\text{dist}(x_i, C_i(x_{-i}))$ the distance from x_i to $C_i(x_{-i})$. Also, write $X_{-i} := \prod_{j \in I \setminus i} X_j$.

Proposition 1. (On exact penalty)

• Let $L_i(x_i, x_{-i})$ be Lipschitz continuous in $x_i \in X_i$ with modulus $< \lambda_i(x_{-i})$. Suppose $X_i \cap C_i(x_{-i})$ is nonempty closed whenever $x_{-i} \in X_{-i}$. Also suppose there is a penalty $\mathcal{P}_i(x_i, x_{-i})$ that vanishes when $x_i \in X_i \cap C_i(x_{-i})$ and satisfies

$$\mathcal{P}_i(x_i, x_{-i}) \geq \lambda_i(x_{-i}) \text{dist}(x_i, C_i(x_{-i})) \quad \text{whenever } x_i \notin C_i(x_{-i}).$$

Consider the game that has modified cost functions $i \mapsto L_i(x_i, x_{-i}) + \mathcal{P}_i(x_i, x_{-i})$ and no coupling constraints. Each Nash equilibrium of the latter game solves the original one.

• In particular, let X_i be compact convex, $L_i(x_i, x_{-i})$ jointly continuous - and convex Lipschitz in x_i with modulus smaller than a continuous $\lambda_i(x_{-i})$. If $x_{-i} \mapsto C_{-i}(x_{-i})$ is continuous with $X_i \cap C_i(x_{-i})$ nonempty convex for each $x_{-i} \in X_{-i}$, then the game with modified cost functions

$$i \mapsto L_i(x_i, x_{-i}) + \lambda_i(x_{-i}) \text{dist}(x_i, C_i(x_{-i})) \quad (2)$$

has a Nash equilibrium that solves the original constrained game.

Proof. Fix any $i \in I$ and $x_{-i} \in X_{-i}$. If $x_i \in X_i \setminus C_i(x_{-i})$, let \bar{x}_i be any closest approximation in $X_i \cap C_i(x_{-i})$. Then

$$\begin{aligned} L_i(\bar{x}_i, x_{-i}) + \mathcal{P}_i(\bar{x}_i, x_{-i}) &= L_i(\bar{x}_i, x_{-i}) < L_i(x_i, x_{-i}) + \lambda_i(x_{-i}) \|\bar{x}_i - x_i\| \\ &\leq L_i(x_i, x_{-i}) + \mathcal{P}_i(x_i, x_{-i}). \end{aligned}$$

Thus, in the modified game the best response of player i always belongs to $X_i \cap C_i(x_{-i})$. This takes care of the first bullet. For the second, note that $L_i(x_i, x_{-i})$, being convex in x_i near X_i , becomes indeed Lipschitz in that variable there. Further, objective (2) is convex in x_i and jointly continuous. Finally, because each X_i is nonempty compact convex, the assertion in the last bullet follows from standard existence results [22]. \square

Our chief concern is with computation of equilibria, not their existence. So henceforth we take existence for granted. Also, we shall deal with tractable instances that need neither quasi-variational inequalities nor exact penalty methods. As indicated above, tractability is had here as follows. While it's commonplace to demand that all images $X_i \cap C_i(x_{-i})$ are convex, we rather require that the product set X , as defined in (1), is convex. To see the bite of this assumption, suppose, quite generally, that

$$X_i \cap C_i(x_{-i}) = \{x_i : c_i(x_i, x_{-i}) \leq 0\}$$

for some real-valued function c_i . Then it suffices for convexity of $X_i \cap C_i(x_{-i})$ to have $c_i(x_i, x_{-i})$ convex (or merely quasi-convex) in x_i . By contrast, to ensure convexity of X one would typically require that each $c_i(x_i, x_{-i})$ be jointly convex.

Given the coupling constraint $x \in X$ and positive numbers $r_i, i \in I$, Rosen (1965) called the strategy profile $x^* = (x_i)$ a *normalized equilibrium* provided it

$$\text{minimize } \sum_{i \in I} r_i L_i(x_i, x_{-i}^*) \text{ subject to } x \in X.$$

Obviously, normalized equilibria are Nash equilibria, but, as illustrated in Example 7 below, the converse is not true in general - unless, of course, X_i always, and for each i , is contained in $C_i(x_{-i})$. Normalized equilibria are available under reasonable conditions. Following verbatim the proof of Theorem 3.1 in Nikaidô-Isoda (1955) we can strengthen Rosen's (1965) existence result slightly:

Proposition 2. (Existence of normalized equilibrium) *Suppose X as defined in (1) is nonempty compact convex. If $L_i(x_i, x_{-i})$ is convex in x_i , continuous in x_{-i} , and $\sum_{i \in I} L_i(x)$ is continuous, then there exists a normalized equilibrium. \square*

Proposition 2 motivates a **standing assumption**: X is nonempty compact convex, and each $L_i(x_i, x_{-i})$ is convex in x_i and finite-valued near X . Equilibrium is then fully characterized by essential marginal costs, that is, by partial subdifferentials $M_i(x) := \partial_{x_i} L_i(x)$ and normal cones. To state this, let $N(x)$ denote the normal cone to X at its member x , P the orthogonal projection onto X , and $M(r, x) := [r_i M_i(x)]_{i \in I}$ the vector of scaled subdifferentials. Then, standard optimality conditions of convex programming [27] yield:

Proposition 3. (Equilibria occur where essential marginal costs are zero) *The following three statements are necessary and sufficient for $x^* \in X$ to be a normalized equilibrium with strictly positive parameter vector $r = (r_i)$:*

- $\exists g^* \in M(r, x^*)$ such that $\langle g^*, x - x^* \rangle \geq 0$ for all $x \in X$;
- $0 \in M(r, x^*) + N(x^*)$;
- $x^* \in P[x^* - sM(r, x^*)]$ for all $s > 0$. \square

These bullets beg use of established computational techniques. In particular, because the first is a variational inequality, a plethora of corresponding algorithms may come into play [11]. Likewise, the second bullet directs attention to proximal point procedures [13], [14], [28], and especially, to splitting methods [8]. Finally, the last bullet indicates that subgradient projections might offer a good avenue [3], [10].

In any event, to make progress along any of these lines, it is desirable that the scaled marginal cost correspondence $x \rightarrow M(r, x)$ be monotone - or a fortiori strictly monotone [29]. However, even then each of the said approaches may meet significant difficulties. To wit, proximal point procedures, including those using splitting techniques, although yielding good convergence, are often difficult to implement. They typically require iterative solutions of similar perturbed games, each being almost as difficult to handle as the original one. Subgradient projection, with dwindling step-sizes, has opposite properties: implementation comes rather easily, but the method

often produces exceedingly slow convergence.

These observations lead us specialize on the data of the game, and to approach computation along different lines. For simpler notations, incorporate the parameter r_i into L_i ; that is, make the substitution $L_i \leftarrow r_i L_i$ - or alternatively, if possible, set $r_i = 1$. Correspondingly, introduce what we call the *Ky Fan* [1] or *Nikaidô-Isoda* [22] *function*,¹

$$L(x, y) := \sum_{i \in I} [L_i(x) - L_i(y_i, x_{-i})]. \quad (3)$$

Clearly, $x^* \in X$ is a normalized equilibrium iff

$$\sup_{y \in X} L(x^*, y) \leq 0.$$

Now, when solving this inequality system for x^* , it largely helps that $L(x, y)$ be convex in x . These observations motivate the inquiry below. They make us focus on games, declared *convex-concave*, that have Ky Fan function $L(x, y)$ convex-concave in x . By the standing assumption $L(x, y)$ is already concave in y . It turns out that convex-concave games admit Nash equilibria that not only are minimax saddle points of L , but they also lie on the diagonal.

To begin with, we notice that any saddle point (x^*, y^*) of L furnishes a normalized equilibrium x^* . This feature makes us inquire whether a normalized equilibrium x^* can be duplicated to constitute a diagonal saddle point (x^*, x^*) . As brought out in the next proposition, the answer is positive. For a main argument there we shall use the following result of independent interest.

Lemma 1. (Antisymmetry of partial derivatives) *Assume $L(x, y)$ is convex-concave when x, y are near the convex set X . Also assume $L(x, x) = 0$. Then*

$$\partial_x L(x, x) = -\partial_y L(x, x).$$

Proof. Define $h = x' - x$ with x' in a small neighborhood of x . By convexity of L with respect to x , for every $\alpha \in (0, 1)$,

$$\alpha L(x + h, x + \alpha h) + (1 - \alpha)L(x, x + \alpha h) \geq L(x + \alpha h, x + \alpha h) = 0.$$

Dividing by α and passing to the limit with $\alpha \downarrow 0$ we obtain

$$L(x + h, x) + \lim_{\alpha \downarrow 0} [\alpha^{-1} L(x, x + \alpha h)] \geq 0.$$

¹Nikaidô-Isoda (1955) focused on $\varphi(x, x') := -\sum_{i \in I} L_i(x_i, x'_{-i})$ and observed that equilibrium obtains iff $\min_y \max_x [\varphi(x, y) - \varphi(y, y)] = 0$; that is, iff $\max_y \min_x L(x, y) = 0$. The *Ky Fan inequality*

$$\sup_y L(x^*, y) \leq \sup_x L(x, x)$$

is central in Aubin's presentation of game theory (1993).

Clearly, for every $g \in \partial_y L(x, x)$,

$$\lim_{\alpha \downarrow 0} [\alpha^{-1} L(x, x + \alpha h)] \leq \langle g, h \rangle.$$

Thus

$$L(x + h, x) \geq \langle -g, h \rangle = \langle -g, h \rangle + L(x, x).$$

Since a feasible $x + h$ can be arbitrary in a sufficiently small neighborhood of x (such that all function values in the analysis above are finite), $-g \in \partial_x L(x, x)$. Consequently,

$$\partial_x L(x, x) \supseteq -\partial_y L(x, x).$$

In a symmetric way we can prove the converse inclusion. \square

We can now state a first main result.

Proposition 4. (On normalized equilibria) *If the game is convex-concave, then the following statements are equivalent:*

- (a) x^* is a normalized equilibrium;
- (b) $\sup_{y \in X} L(x^*, y) = 0$;
- (c) $\inf_{x \in X} L(x, x^*) = 0$;
- (d) (x^*, x^*) is a saddle point of L on $X \times X$.

Proof. (a) \Leftrightarrow (b). The equivalence follows directly from the definition of a normalized equilibrium.

(b) \Leftrightarrow (c). From (b) it follows that there is $g \in \partial_y L(x^*, x^*)$ such that $\langle g, x - x^* \rangle \leq 0$ for all $x \in X$. By Lemma 1, $-g \in \partial_x L(x^*, x^*)$, so

$$L(x, x^*) \geq \langle -g, x - x^* \rangle \geq 0 = L(x^*, x^*)$$

for every $x \in X$. The converse implication can be proved analogously.

((b) \wedge (c)) \Leftrightarrow (d). The equivalence is obvious, because $L(x^*, x^*) = 0$. \square

Proposition 4 allows us to address a related issue, namely: when is normalized equilibrium unique?

Proposition 5. (Uniqueness of normalized equilibrium) *In a convex-concave game suppose $L(x, y)$ is strictly convex in x or strictly concave in y . Then normalized equilibrium is unique.*

Proof. Suppose there are different normalized equilibria $x, \bar{x} \in X$. Then (\bar{x}, x) and (x, \bar{x}) are both saddle points of L . If $L(\cdot, x)$ is strictly convex, the inequality $L(\frac{1}{2}x + \frac{1}{2}\bar{x}, x) < \frac{1}{2}L(x, x) + \frac{1}{2}L(\bar{x}, x) = L(x, x)$ contradicts the minimality of $L(\cdot, x)$ at x . Similarly, in case $L(x, \cdot)$ is strictly concave, the inequality $L(x, \frac{1}{2}x + \frac{1}{2}\bar{x}) > \frac{1}{2}L(x, x) + \frac{1}{2}L(x, \bar{x}) = L(x, x)$ contradicts the maximality of $L(x, \cdot)$ at x . \square

3. PARTIAL REGULARIZATION METHODS

In ordinary Nash equilibrium every party, quite on his own, perfectly *predicts* the rivals' actions and *optimizes* his proper response. Here, given coupling constraints, some *coordination* is also called for. Reflecting on this, our purpose is to find Nash equilibrium using only iterative, single-agent programming albeit subject to necessary coordination. In this endeavour, while seeking diagonal saddle points of L , we shall adapt ideas developed for general minimax problems in [30]. Broadly, the procedure can be advertized as follows.

Besides the individuals $i \in I$, introduce a fictitious player concerned only with coordination. Suppose he recently suggested that the strategy profile $x \in X$ be used. Upon revising his suggestion x this particular agent predicts that individual $i \in I$ will respond with strategy

$$y_i^+ \in \arg \min L_i(\cdot, x_{-i}),$$

so as to fetch a reduction $L_i(x) - L_i(y_i^+, x_{-i})$ in own cost. Presumably the coordinating agent wants the overall cost reduction

$$L(x, y^+) = \sum_{i \in I} [L_i(x) - L_i(y_i^+, x_{-i})].$$

to be small. So, if possible, he might prudently change x in a "descent" direction

$$d_x \in -\partial_x L(x, y^+).$$

Similarly, individual $i \in I$, who recently opted for strategy y_i , predicts that the coordinating agent next will propose a profile x^+ such that $L(x^+, y) \leq 0$ or, a fortiori, one that satisfies

$$x^+ \in \arg \min_{x \in X} L(x, y). \quad (4)$$

In either case, such beliefs induce a change of his own response y_i , if possible, along a "descent" direction

$$d_{y_i} \in -\partial_{y_i} L_i(y_i, x_{-i}^+).$$

These loose ideas were intended to motivate and advertize the subsequent two algorithms. The broad outline, given above, must however, be refined on *four* accounts: *First*, some stability or inertia is needed in the predictions. For that purpose we shall introduce regularizing penalties of quadratic nature [30].

Second, the descent directions must be feasible. To that end we shall rely on projections, designed to enforce global, non-decomposable constraints [12].

Third, when updating x and y along proposed directions, appropriate step sizes are needed. At this juncture some techniques from subgradient projection methods will serve us well [25].

Fourth and finally, equality of the coordinating profile and the pattern of strategy responses is ensured by *compromising* the proposed updates.

All these matters are accounted for and incorporated in the following two algorithms:

Algorithm 1 (*Partial regularization in individual strategies*)

Initialization: Select an arbitrary starting point $x^0 \in X$ and set $\nu := 0$.

Predict individual strategies: Compute

$$y^{\nu+} := \arg \min \left\{ \sum_{i \in I} L_i(y_i, x_{-i}^{\nu}) + \frac{\rho}{2} \|y - x^{\nu}\|^2 : y \in X \right\}. \quad (5)$$

Test for termination: If $y^{\nu+} = x^{\nu}$, then stop: x^{ν} solves the problem.

Predict a coordinating strategy: Find $x^{\nu+} \in X$ such that

$$L(x^{\nu+}, x^{\nu}) \leq 0$$

and $\|x^{\nu+} - x^{\nu}\| \leq \kappa$ for some constant κ . In particular, $x^{\nu+} = x^{\nu}$ is one option.

Find direction of improvement: Select subgradients $g_x^{\nu} \in \partial_x L(x^{\nu}, y^{\nu+})$ and $g_{y_i}^{\nu} \in \partial_{x_i} L_i(x_i^{\nu}, x_{-i}^{\nu+})$, $i \in I$, and define a direction $d^{\nu} := (d_x^{\nu}, d_y^{\nu})$ with

$$d_x^{\nu} := P_x^{\nu} [-g_x^{\nu}], \quad d_y^{\nu} := P_y^{\nu} [-g_y^{\nu}],$$

where P_x^{ν} , P_y^{ν} denote orthogonal projections onto closed convex cones T_x^{ν} , T_y^{ν} containing the tangent cone $T(x^{\nu})$ of X at the current x^{ν} .

Calculate the step size: Let

$$\tau_{\nu} = \frac{\gamma_{\nu} [L(x^{\nu}, y^{\nu+}) - L(x^{\nu+}, x^{\nu})]}{\|d^{\nu}\|^2},$$

with $0 < \gamma_{\min} \leq \gamma_{\nu} \leq \gamma_{\max} < 2$.

Make a step: Update by the rules

$$\begin{aligned} x^{\nu++} &:= P [x^{\nu} + \tau_{\nu} d_x^{\nu}], \quad \text{and} \\ y^{\nu++} &:= P [x^{\nu} + \tau_{\nu} d_y^{\nu}], \end{aligned}$$

where P is the orthogonal projection onto X .

Strike a compromise: Let

$$x^{\nu+1} = \frac{1}{2} (x^{\nu++} + y^{\nu++}).$$

Increase the counter ν by 1 and continue to **Predict individual strategies**. \square

The second algorithm is symmetric to the first one in reversing the manner of prediction.

Algorithm 2. (*Partial regularization in the coordinating variable*) The method

proceeds as Algorithm 1, the only difference being in the prediction steps. Those are replaced by the following ones.

Predict individual strategies: Find $y^{\nu+} \in X$ such that

$$\sum_{i \in I} L_i(y_i^{\nu+}, x_{-i}^{\nu}) \leq \sum_{i \in I} L_i(x_i^{\nu}, x_{-i}^{\nu})$$

and $\|y^{\nu+} - x^{\nu}\| \leq \kappa$ for some constant κ . In particular, $y^{\nu+} = y^{\nu}$ is an easy and acceptable choice.

Predict the coordinating strategy: Compute

$$x^{\nu+} \in \arg \min \left\{ L(x, x^{\nu}) + \frac{\rho}{2} \|x - x^{\nu}\|^2 : x \in X \right\}. \quad \square \quad (6)$$

Some **remarks** are in order:

- Plainly, the approximating cones - and projections onto these - can be omitted. Indeed, simply take either cone to equal the entire space. If however, $X = \{x : c_k(x) \leq 0, k \in K\}$ for a finite set of differentiable convex functions $c_k, k \in K$, it is a tractable problem to project onto the cone generated by the gradients $\nabla c_k(x)$ of the active constraints; see [29].
- In the absence of coupling constraints, with $X = \prod_{i \in I} X_i$, prediction (5) of individual strategies decomposes into separate subproblems, one for each player.
- To execute (6) is generally more difficult than (5), given that $L(x, y)$ typically is less separable in x than in y .
- When compromising updates one need not use the constant, equal weight 1/2. Stage-varying choices $\alpha_x^{\nu} \geq 0$, $\alpha_y^{\nu} \geq 0$, $\alpha_x^{\nu} + \alpha_y^{\nu} = 1$ are applicable provided the weight be bounded away from 0 on the variable for which direction-finding was more elaborate (with minimization in the prediction step for the other variable)
- Both algorithms lend themselves to asynchronous implementations.
- The proximal parameter $\rho > 0$ may vary, provided it is bounded away from 0 and ∞ .
- Instead of quadratic terms in the prediction steps one can use more general mappings with similar properties; see [19].
- Procedures (4), (5), (6) invite duality methods. Then, if X equals $\{x : c(x) \leq 0\}$ and is properly qualified, Lagrange multipliers may guide good design of taxation or penalty schemes aimed at enforcement of equilibrium play; see [21], [29].

4. CONVERGENCE

Our convergence analysis follows the general lines of [19] and [30] with modifications that account for the special properties of our problem.

It simplifies the exposition to single out a key observation; namely, that our algorithmic step constitutes a Fejér mapping; see [9] and [25].

Lemma 2. (Fejér property) *Assume that the game is convex-concave and has a*

normalized equilibrium x^* . Define

$$W_\nu := \|x^\nu - x^*\|^2,$$

where $\{x^\nu\}$ is the sequence generated by any of the two algorithms defined in the previous section. Then for all ν

$$W_{\nu+1} \leq W_\nu - \frac{1}{2}\gamma_\nu(2 - \gamma_\nu)[L(x^\nu, y^{\nu+}) - L(x^{\nu+}, y^\nu)]^2 / \|d^\nu\|^2.$$

Proof. Invoking the non-expansiveness of projection, we have

$$\begin{aligned} \|x^{\nu++} - x^*\|^2 &= \|P[x^\nu + \tau_\nu d_x^\nu] - P[x^*]\|^2 \\ &\leq \|x^\nu + \tau_\nu d_x^\nu - x^*\|^2 \\ &= \|x^\nu - x^*\|^2 + 2\tau_\nu \langle d_x^\nu, x^\nu - x^* \rangle + \tau_\nu^2 \|d_x^\nu\|^2. \end{aligned}$$

Use now the orthogonal decomposition $-g_x^\nu = d_x^\nu + n_x^\nu$, n_x^ν being in the negative polar cone of T_x^ν , and observe that $x^* - x^\nu \in T_x^\nu$, to obtain

$$\langle d_x^\nu, x^\nu - x^* \rangle \leq \langle g_x^\nu, x^* - x^\nu \rangle \leq L(x^*, y^{\nu+}) - L(x^\nu, y^{\nu+}).$$

Whence,

$$\|x^{\nu++} - x^*\|^2 \leq \|x^\nu - x^*\|^2 - 2\tau_\nu [L(x^\nu, y^{\nu+}) - L(x^*, y^{\nu+})] + \tau_\nu^2 \|d_x^\nu\|^2.$$

Similarly,

$$\|y^{\nu++} - x^*\|^2 \leq \|x^\nu - x^*\|^2 + 2\tau_\nu [L(x^{\nu+}, x^\nu) - L(x^{\nu+}, x^*)] + \tau_\nu^2 \|d_y^\nu\|^2.$$

By convexity of the squared norm,

$$\|x^{\nu+1} - x^*\|^2 \leq \frac{1}{2} (\|x^{\nu++} - x^*\|^2 + \|y^{\nu++} - x^*\|^2).$$

Combining the last three inequalities we have

$$W_{\nu+1} \leq W_\nu - \tau_\nu [L(x^\nu, y^{\nu+}) - L(x^*, y^{\nu+}) - L(x^{\nu+}, x^\nu) + L(x^{\nu+}, x^*)] + \frac{1}{2}\tau_\nu^2 \|d^\nu\|^2.$$

Since, by Proposition 4, (x^*, x^*) is a saddle point of L , it follows that $L(x^*, y^{\nu+}) \leq L(x^{\nu+}, x^*)$. Therefore

$$W_{\nu+1} \leq W_\nu - \tau_\nu [L(x^\nu, y^{\nu+}) - L(x^{\nu+}, x^\nu)] + \frac{1}{2}\tau_\nu^2 \|d^\nu\|^2.$$

Here apply the stepsize rule to arrive at the required result. \square

The first convergence result can now be stated forthwith.

Theorem 1. (Convergence with regularized individual strategies) *Assume that the game is convex-concave and has a normalized equilibrium x^* . Then the sequence $\{x^\nu\}$ generated by Algorithm 1 converges to a normalized equilibrium.*

Proof. Since $L(x^{\nu+}, x^\nu) \leq L(x^\nu, x^\nu) = 0$ and $L(x^\nu, y^{\nu+}) \geq L(x^\nu, x^\nu) = 0$, from Lemma 1 we obtain,

$$W_{\nu+1} \leq W_\nu - \frac{1}{2}\gamma_\nu(2 - \gamma_\nu)[L(x^\nu, y^{\nu+})]^2 / \|d^\nu\|^2.$$

Evidently $\{W_\nu\}$ is non-increasing, hence bounded. The sequence $\{d^\nu\}$ is bounded, so $L(x^\nu, y^{\nu+}) \rightarrow 0$. By the definition of $y^{\nu+}$, there exists a subgradient $g \in \partial_y L(x^\nu, y^{\nu+})$ such that

$$\langle g - \rho(y^{\nu+} - x^\nu), h \rangle \leq 0,$$

for every feasible direction h at $y^{\nu+}$. Thus, with $h = x^\nu - y^{\nu+}$, one has

$$L(x^\nu, x^\nu) - L(x^\nu, y^{\nu+}) \leq \langle g, x^\nu - y^{\nu+} \rangle \leq -\rho\|y^{\nu+} - x^\nu\|^2,$$

so

$$L(x^\nu, y^{\nu+}) \geq \rho\|y^{\nu+} - x^\nu\|^2.$$

Consequently,

$$\lim_{\nu \rightarrow \infty} \|y^{\nu+} - x^\nu\|^2 = 0.$$

Let \hat{x} be an accumulation point of $\{x^\nu\}$ and y^+ be the associated accumulation point of $\{y^{\nu+}\}$. Then $y^+ = \hat{x}$, i.e.,

$$\hat{x} = \arg \min \left\{ \sum_{i \in I} L_i(y_i, \hat{x}_{-i}) + \frac{\rho}{2} \|y - \hat{x}\|^2 : y \in X \right\}.$$

This is necessary and sufficient for \hat{x} to be a normalized equilibrium. Substituting it for x^* in the definition of W_ν we conclude that the distance to \hat{x} is non-increasing. Consequently, \hat{x} is the only accumulation point of the sequence $\{x^\nu\}$. \square

Theorem 2. (Convergence under coordinated regularization) *Assume that the game is convex-concave and has a normalized equilibrium x^* . Then the sequence $\{x^\nu\}$ generated by Algorithm 2 is convergent to a normalized equilibrium.*

Proof. Proceeding analogously to the proof of Theorem 1 we arrive at the relation:

$$\lim_{\nu \rightarrow \infty} \|x^{\nu+} - x^\nu\|^2 = 0.$$

Let \hat{x} be an accumulation point of $\{x^\nu\}$ and x^+ be the associated accumulation point of $\{x^{\nu+}\}$. Then $x^+ = \hat{x}$, i.e.,

$$\hat{x} = \arg \min \left\{ L(x_i, \hat{x}_{-i}) + \frac{\rho}{2} \|x - \hat{x}\|^2 : x \in X \right\}.$$

By Proposition 4, this is necessary and sufficient for \hat{x} to be a normalized equilibrium. Substituting it for x^* in the definition of W_ν , we conclude that \hat{x} is the limit of the sequence $\{x^\nu\}$. \square

5. EXAMPLES OF CONVEX-CONCAVE GAMES

Convex-concave games may serve as standard models in their own right or as approximations to more complex data. This section concludes by indicating that the class at hand is more rich than might first be imagined. For all instances below the prime concern is with $L(x, y)$ being convex in x . In most cases the set X is left unspecified.

Example 1: (Two-person zero-sum games) *Any two-person zero-sum game with convex-concave cost $L_1(x_1, x_2)$ of player 1, is convex-concave.*

Proof. Since $L(x, y) = L_1(x_1, y_2) - L_1(y_1, x_2)$, the conclusion is immediate. \square

Example 2: (Games with affine interaction) *Let each $L_i(x_i, x_{-i})$ be jointly convex in (x_i, x_{-i}) and separately affine in x_{-i} . Then the game is convex-concave.* \square

Example 3: (Games with separable cost) *Let each $L_i(x_i, x_{-i}) = L_i(x_i) + L_{-i}(x_{-i})$ be separable and convex in x_i . Then the game is convex-concave.* \square

Example 4: (Games with bilinear interaction) *Suppose each cost function $L_i(x)$ is linear-quadratic in the sense that*

$$L_i(x) := \sum_{j \in I} [b_{ij}^T + x_i^T C_{ij}] x_j \quad (7)$$

for specified vectors $b_{ij} \in \mathbb{X}_j$ and matrices C_{ij} of appropriate dimension. If the corresponding $I \times I$ block matrix - featuring block C_{ij} in off-diagonal entry ij and $2C_{ii}$ on diagonal entry ii - is positive semidefinite, then the game is convex-concave.

Proof. Since $L_i(x_i, x_{-i})$ is affine in x_{-i} , it suffices to show that $\sum_{i \in I} L_i(x)$ is convex. Further, because linear terms can be ignored, it's enough to verify that $\sum_{i \in I} \sum_{j \in I} x_i^T C_{ij} x_j$ is convex. The Hessian of this double sum equals the described block matrix, and the conclusion follows. \square

Example 5: (Multi-person, finite-strategy matrix games) Suppose each player $i \in I$ has a finite set S_i of pure strategies. If he plays $s_i \in S_i$ against each rival $j \neq i$, the latter using strategy $s_j \in S_j$, then the cost incurred by the former agent equals

$$l_i(s_i, s_{-i}) := \sum_{j \in I} \{b_{ij}(s_j) + C_{ij}(s_i, s_j)\}$$

Here $C_{ij}(s_i, s_j)$ denotes the (s_i, s_j) entry of a prescribed $S_i \times S_j$ cost matrix C_{ij} . Pass now to mixed strategies $x_i \in X_i :=$ the probability simplex over S_i . Then format

(7) emerges again. Standard versions have $b_{ij} = 0$ and $C_{ii} = 0$. Zero-sum means the $\sum_{i \in I} \sum_{j \in I} x_i^T C_{ij} x_j = 0$, and evidently such games are convex. \square

Example 6: (*Cournot oligopoly*) A classical noncooperative game is the Cournot (1838) oligopoly model, still a workhorse in modern theories of industrial organization [32]. Generalizing it to comprise a finite set G of different goods, the model goes as follows: Firm $i \in I$ produces the commodity bundle $x_i \in \mathbb{R}^G$, thus incurring convex production cost $c_i(x_i)$ and gaining market revenues $p(\sum_{j \in I} x_j) \cdot x_i$. Here $p(\sum_{j \in I} x_j)$ is the price vector at which total demand equals the aggregate supply $\sum_{j \in I} x_j$. Suppose this inverse demand curve is affine and "slopes downwards" in the sense that $p(Q) = a - CQ$ where $a \in \mathbb{R}^G$ and C is a $G \times G$ positive semidefinite matrix. Then

$$L_i(x) = c_i(x_i) - (a - C \sum_{j \in I} x_j) \cdot x_i,$$

and the resulting Cournot oligopoly is convex-concave. A structurally similar model of river pollution has been studied, subject to linear coupling constraints, in [17], [21]. See also [31]. \square

Example 7: (*A game of location* [3]) Player $i = 1, 2$ lives in the Euclidean plane at the address $e_1 = (1, 0)$, $e_2 = (0, 1)$, respectively. While controlling $x_i \in X_i := (-\infty, 0]$ he wants to minimize the squared distance between his address and

$$X = \{(x_1, x_2) \in \mathbb{R}_-^2 : x_1 + x_2 \leq -1\}.$$

Each point on the line segment $[-e_1, -e_2] = [(-1, 0), (0, -1)]$ is a Nash equilibrium. However, only the midpoint $(-\frac{1}{2}, -\frac{1}{2})$ is normalized. \square

Acknowledgement: The paper greatly benefitted from comments given by two referees.

REFERENCES

- [1] J.-P. Aubin, *Optima and Equilibria*, Springer-Verlag, Berlin (1993).
- [2] A. Bensoussan, Points de Nash dans le cas de fonctionnelles quadratiques et jeux différentiels linéaires a N personnes, *SIAM J Control* 12, 460-499 (1974).
- [3] E. Cavazzuti and S. D. Flåm, Evolution to selected Nash equilibria, in F. Giannessi (ed.), *Nonsmooth Optimization Methods and Applications*, Gordon and Breach, London, 30-41 (1992)
- [4] A. Cournot, *Recherches sur les Principes Mathématiques de la Théorie des Richesses*, Paris. English translation: *Researches into the mathematical principles of the theory of wealth* (ed. N. Bacon) Macmillan, New York (1897).

- [5] D. Chan and J. S. Pang, The generalized quasi-variational inequality problem, *Mathematics of Operations Research* 7, 211-222 (1982).
- [6] J. Contreras, M. Klusch and J. B. Krawczyk, Numerical solution to Nash-Cournot equilibria in coupled constraint electricity markets, *IEEE Trans. on Power Systems* 19, 1, 195-206 (2004).
- [7] G. Debreu, A social equilibrium existence theorem, *Proceedings of the National Academy of Sciences* 38 (10) 886-893 (1952).
- [8] J. Eckstein and D. P. Bertsekas, On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators, *Mathematical Programming* 55 (3) 293-318 (1992)
- [9] I. I. Eremin and N. N. Astafiev, *Introduction to the Theory of Linear and Convex Programming* Nauka, Moscow (1976).
- [10] Yu. M. Ermoliev and S. Uryasev, On seeking a Nash equilibrium in n-person games, (in Russian) *Kibernetika* 3, 85-88 (1982).
- [11] F. Facchinei and J. S. Pang, *Finite-Dimensional Variational Inequalities and Complementarity Problems*, Springer-Verlag, Berlin (2003).
- [12] S. D. Flåm, On finite convergence and constraint identification of subgradient projection methods, *Math. Programming* 57 (3 A) 427-437 (1992).
- [13] S. D. Flåm and A. S. Antipin, Equilibrium programming using proximal-like algorithms, *Mathematical Programming* 78, 29-41 (1997).
- [14] C. D. Ha, A generalization of the proximal point algorithm, *SIAM J. Control and Optimization* 28(3) 5003-512 (1990).
- [15] P. T. Harker and J. S. Pang, Finite-dimensional variational inequalities and nonlinear complementarity problems: A survey of theory, algorithms and applications, *Mathematical Programming* 48, 2, 161-220 (1990).
- [16] P. T. Harker, Generalized Nash games and quasivariational inequalities, *European Journal of Operations Research* 54, 81-94 (1991).
- [17] A. Haurie and J. B. Krawczyk, Optimal charges on river effluent from lumped and distributed resources, *Environmental Modeling and Assessment* 2, 93-106 (1997)..
- [18] A. Haurie and J. B. Krawczyk, *An Introduction to Dynamic Games*, Internet textbook, URL: <http://ecolu-info.unige.ch/haurie/fame/textbook.pdf>.
- [19] M. Kallio and A. Ruszczyński, Perturbation methods for saddle point computation, Working Paper WP-94-38, IIASA (1994).

- [20] J. B. Krawczyk and S. Uryasev, Relaxation algorithms to find Nash equilibria with economic applications, *Environmental Modeling and Assessment* 5, 63-73 (2000).
- [21] J. B. Krawczyk, Coupled constraint Nash equilibria in environmental games, *Resources and Energy Economics* 27, 157-181 (2005).
- [22] H. Nikaidô and K. Isoda, Note on noncooperative convex games, *Pacific Journal of Mathematics*, 5, 1, 807-815 (1955).
- [23] J. S. Pang and M. Fukushima, Quasi-variational inequalities, generalized Nash equilibria, and multi-leader-follower games, *Computational Management Science* 1, 21-56 (2005).
- [24] J. S. Pang, Computing generalized Nash equilibria, Manuscript (2002).
- [25] B. T. Polyak, Minimization of nonsmooth functionals, *Zhurnal Vychislitelnoi Matematiki i Matematicheskoi Fiziki* 9, 509-521 (1969).
- [26] S. M. Robinson, Shadow prices for measures of effectiveness II, General model, *Operations Research* 41, 536-548 (1993).
- [27] R. T. Rockafellar, *Convex analysis*, Princeton University Press, Princeton (1970).
- [28] R. T. Rockafellar, Monotone operators and the proximal point algorithm, *SIAM J. Control and Optimization* 14 (5), 877-898 (1976).
- [29] J. B. Rosen, Existence and uniqueness of equilibrium points for concave n-person games, *Econometrica* 33 (3), 520-534 (1965).
- [30] A. Ruszczyński, A partial regularization method for saddle point seeking, Working Paper WP-94-20, IIASA (1994).
- [31] M. Tidball and G. Zaccour, An environmental game with coupling constraints, *Environmental Modeling and Assessment* 10, 153 - 158 (2005).
- [32] J. Tirole, *The Theory of Industrial Organization*, MIT Press, Cambridge (1988).
- [33] S. Uryasev and R. Y. Rubinstein, On relaxation algorithms in computation of noncooperative equilibria, *IEEE Transactions on Automatic Control* 39, 6, 1263-1267 (1994).