49th IEEE Conference on Decision and Control
December 15-17, 2010
Hilton Atlanta Hotel, Atlanta, GA, USA

Distributed Iterative Regularization Algorithms for Monotone Nash games

Aswin Kannan and Uday V. Shanbhag

***Abstract*— In this paper, we consider the development of single-timescale schemes for the distributed computation of Nash equilibria. In general, equilibria associated with convex Nash games over continuous strategy sets are wholly captured by the solution set of a variational inequality. Our focus is on Nash games whose equilibrium conditions are given by monotone variational inequalities, a class referred to as *monotone Nash games*. Unless suitably strong assumptions (such as strong monotonicity) are imposed on the mapping corresponding to the variational inequality, distributed schemes for computing equilibria often require the solution of a sequence of regularized problems, each of which has a unique solution. Such schemes operate on two timescales and are generally harder to implement in online settings. Motivated by this shortcoming, this work focuses on the development of three single timescale iterative regularization schemes that require precisely *one projection* step at every iteration. The first is an iterative Tikhonov regularization scheme while the second is an analogously constructed iterative proximal-point method. Both schemes are characterized by the property that the regularization/centering parameter are updated after every iteration, rather than when one has an approximate solution to the regularized problem. Finally, a modified form of the proximal-point scheme is also presented where the weight on the proximal term is updated as well.**

I. Introduction

Consider an $N-$person Nash game in which the ith player solves

$\mathrm{Ag}(x_{-i})$	minimize	$f_i(x_i; x_{-i})$
	subject to	$x_i \in K_i$

for $i = 1, \ldots, N$ where $x_{-i} = (x_j)_{j \neq i}$, $K_i \subseteq \mathbb{R}^{n_i}$ is a closed and convex set and $\sum_{i=1}^N n_i = n$. Additionally, we assume that for $i = 1, \ldots, N$, the function f_i is a differentiable real-valued function given by $f_i : \mathbb{R}^n \to \mathbb{R}$ and convex in x_i for all $x_{-i} \in \prod_{j \neq i} K_j$. A Nash equilibrium of the aforementioned noncooperative game is given by a tuple $\{x_i^*\}_{i=1}^N$, where $x_i^* \in SOL(\mathrm{Ag}(x_{-i}^*))$ for $i = 1, \ldots, N$ and SOL(Ag) denotes the set of solutions to problem (Ag). Throughout this paper, we refer to this canonical Nash game by $\mathcal{G}$. The convexity of the objectives and the associated strategy sets allow us to claim that that the first-order equilibrium conditions are necessary and sufficient. In fact, these conditions can be shown to be equivalent to a scalar variational inequality $\mathrm{VI}(K, F)$ (see result from [5, Ch. 1]) where

$$F(x) \triangleq \begin{pmatrix} \nabla_{x_1} f_1(x) \\ \vdots \\ \nabla_{x_N} f_N(x) \end{pmatrix} \quad \text{and} \quad K \triangleq \prod_{i=1}^N K_i.$$

Recall that $\mathrm{VI}(K, F)$ is a problem requiring an $x \in K$ such that $F(x)^T(y - x) \geq 0, \forall y \in K$. We assume throughout this paper that the mapping $F : \mathbb{R}^n \to \mathbb{R}^n$ is a single-valued mapping possessing a monotonicity property over K namely: $(F(x) - F(y))^T(x - y) \geq 0$ for all $x, y \in K$. We refer to the resulting class of Nash games as *monotone* Nash games.

Much research has been carried out on the development of algorithms for monotone variational inequalities, amongst these being projection-based methods, interior-point algorithms, etc (see [5, Ch. 12]). Of these, projection-based schemes are natural candidates for employment within a distributed framework. Yet, by themselves, standard projection methods require a degree of well-posedness in order to claim global convergence. In particular, if F is either strongly monotone or co-coercive, we may develop a *single timescale* fixed steplength scheme under the caveat that the steplength is sufficiently small. A single timescale scheme refers to one where each iterate is obtained by a single gradient or projection step. However, if F is merely monotone, generally such an avenue is unavailable. In merely monotone cases, a set of classical techniques reside in the realm of regularization and proximal-point methods and requires solving a *sequence* of well-posed problems, each of which might require a distributed iterative process in itself. This is effectively a *two timescale* method where the regularization/proximal method makes changes at a slower timescale while solutions to the regularized problems in an exact or inexact form are found at a faster timescale. Our motivation lies in developing *distributed single timescale* extensions of standard regularization/proximal methods with appropriate convergence guarantees. In general, these schemes require updating the regularization parameters at every iteration, rather than when an acceptable subproblem solution is available, a property that leads to the schemes being referred to as *iterative regularization techniques.*

Our work is motivated by a host of settings where online distributed schemes assume relevance. Examples of these include communications [4], [10], bandwidth and spectrum allocation [2], [3] and optical networks [7], [8]. Of note is recent work of Pavel and her coauthors [7] where an extragradient scheme [5] in a deterministic regime, capable of accommodating monotone Nash games. Also of interest

Department of Industrial and Enterprise Systems Engineering, University of Illinois, Urbana IL 61801, Email: {akannan,udaybag}@illinois.edu. This work has been partially supported by NSF award CCF-0728863 and DOE award DE-SC0003879.

is recent work that examines best response schemes in the context of monotone Nash games [10].

The key contributions of this paper lie in the introduction and analysis of several iterative regularization techniques. The first of these extends a standard Tikhonov regularization method to the single-timescale regime by requiring that the regularization parameter be updated at every iteration. Correspondingly, in the context of proximal-point methods, we suggest that the centering parameter around which the prox-term is constructed is also updated at every iteration. A third scheme is presented where the weighting parameter in the prox-term is also updated at every iteration. In the instance of Tikhonov, the scheme accommodates merely monotone mappings while the proximal-point based methods currently require strict monotonicity.

This paper is divided into five sections. In section II, we provide a brief background to the methods of interest. Section III, presents Tikhonov regularization methods and Section IV presents a single timescale iterative proximal point method while section V introduces a modified iterative proximal scheme.

II. Preliminaries

Our analysis is restricted to Nash games that lead to monotone variational inequalities, a class of games that takes on the name *monotone Nash games*. Our goal, however, lies in developing distributed schemes implementable in networked settings. Since, strong monotonicity of the mapping allows us to directly construct precisely such a class of schemes, this forms our starting point. In the absence of strong monotonicity however, a direct application of direct projection schemes does not guarantee a contraction and thereby convergence [5]. A possible avenue for alleviating the challenge is through regularization methods, a set of techniques that address this ill-posedness by sequentially solving a set of strongly monotone problems. One such technique is an exact Tikhonov regularization method. Let

$$F^k \triangleq F(z) + \epsilon^k z, \tag{1}$$

where $z^{k+1} = SOL(K, F^k)$ and ϵ^k denotes the regularization at the kth iteration. Under suitable conditions (see [5, Ch.12]) the sequence $\{z^k\}$ converges to z^* as $\epsilon^k \to 0$. The subproblems are strongly monotone variational inequalities and can be solved by a host of iterative methods. Note that inexact solutions of such problems also leads to convergent schemes [5, Ch. 12].

An alternative scheme is one that maintains a regularization based on the change in consecutive iterates via a proximal term. Such methods are also referred to as *heavy-ball methods* [9]. Given a scalar $\theta > 0$, such a scheme employs an F^k be redefined as

$$F^k \triangleq F(z) + \theta(z - z^k), \tag{2}$$

where $z^{k+1} = SOL(K, F^k)$. This scheme is guaranteed to converge for monotone variational inequalities (see [1], [5]). A key shortcoming, however, is that both schemes require the solution of a sequence of problems, leading to a natively two timescale scheme.

Single timescale schemes have several advantages. First, they are far easier to implement since the complexity of the scheme is restricted to solving the projection problem. Second, online implementations that require coordination in a networked setting are far easier to manage, making them significantly attractive over schemes that require solving complex subproblems prior to updating their decisions. Third, these schemes are easily distributed providing an avenue for solving truly large-scale networked game-theoretic problems. With this being the major motivation, we provide a unified framework for stating three different single timescale schemes namely the iterative Tikhonov scheme, the iterative proximal scheme and the modified iterative proximal scheme.

Given a game $\mathcal{G}$, the general form of the single timescale scheme is as follows:

$$z_i^{k+1} = \Pi_{K_i}\big(z_i^k - \gamma_k \left(F(z_i^k, z_{-i}^k) + \theta_1^k z_i^k - \theta_2^k z_i^{k-1}\right)\big). \tag{3}$$

Based on the choices of the parameter sequences $\{\theta_1^k\}$ and $\{\theta_2^k\}$, we have the following schemes:

a) Case 1: If, $\theta_1^k = \epsilon^k \to 0$ and $\theta_2^k = 0$, then the scheme is referred to as the iterative Tikhonov regularization (ITR) scheme.

b) Case 2: If $\theta_1^k = \theta_2^k = \theta$ is fixed, then the scheme is referred to as the iterative proximal point (IPP) scheme.

c) Case 3: If $\theta_1^k = \theta_2^k = \frac{c}{\gamma_k}$ where $c \in (0,1)$, then the scheme is referred to as the modified iterative proximal point (MIPP) scheme.

III. Iterative Tikhonov schemes

In standard Tikhonov regularization schemes, one constructs a sequence of exact (or inexact) solutions to well-posed regularized problems and the regularization parameter is driven to zero at the slower timescale. In contrast, we consider a class of *iterative* Tikhonov schemes in which the steplength and regularization parameter are changed at the same rates. We proceed to show that this scheme is indeed convergent. The convergence statement for general monotone mappings also appears in [6] but does so *without a proof* and employs slightly different assumptions. Our proof is original and was presented in a shortened form in past work by the second author [11]. Here, we restate the proofs with some modifications with the intent of generalizing it to partially coordinated settings. Notably, the result for symmetric Nash games is inspired by a result for optimization problems stated in [9]. The following Lemmas from [9] are employed in developing our convergence theory.

Lemma 1: Let $u_{k+1} \le q_k u_k + \alpha_k$, $0 \le q_k < 1$, $\alpha_k \ge 0$ and

$$\sum_{k=0}^{\infty}(1 - q_k) = \infty, \ \frac{\alpha_k}{1 - q_k} \to 0, \ k \to \infty.$$

Then $\lim_{k\to\infty} u_k \le 0$ and if $u_k > 0$, then $\lim_{k\to\infty} u_k = 0$.

Lemma 2: Let $u_{k+1} \le (1 + v_k)u_k + p_k$, $u_k, v_k, p_k \ge 0$ and

$$\sum_{k=0}^{\infty} v_k < \infty, \ \sum_{k=0}^{\infty} p_k < \infty, \ k \to \infty.$$

Then $\lim_{k\to\infty} u_k = \bar{u} \geq 0$

Our proof of convergence relies on relating the iterates of the proposed ITR scheme to that of the original Tikhonov scheme. The following Lemma is reproduced from [11] provides a bound between consecutive iteratives of the standard Tikhonov scheme.

Lemma 3: Let the mapping ∇F be monotone and suppose $SOL(K, F)$ be nonempty and bounded. Consider the standard exact tikhonov scheme defined by (1). If $M := \|w^*\|_2$, then

$$\|y^k - y^{k-1}\| \leq \frac{M(\epsilon^{k-1} - \epsilon^k)}{\epsilon^k}.$$

Proof: Omitted (See [11]). ∎

The proof of convergence for a constant steplength scheme when Jacobian ∇F is assumed to be symmetric has been omitted and may be found in detail in [11]. When ∇F is not necessarily symmetric, we present a diminishing steplength scheme in Section III-A.

A. General monotone Nash games

When considering games with mappings that have general Jacobians, it is not possible to obtain a bound by means of the mean value theorem.Therefore, we follow a different methodology, namely a diminishing step size to ensure a convergent sequence. The regularization parameter $\{\epsilon^k\}$ and the steplength sequence $\{\gamma_k\}$ are assumed to satisfy the following assumptions.

Assumption 1: (A1) The mapping $F(\mathbf{x})$ is Lipschitz continuous with constant L. The regularization parameter ϵ^k and steplength γ_k satisfy $\sum_{k=1}^{\infty} \gamma^k \epsilon^k = \infty, \epsilon^{k+1} \leq \epsilon^k, \forall k, \gamma_{k+1} \leq \gamma_k \forall k, \lim_{k\to\infty} \gamma_k/\epsilon^k = 0, \sum_{k=0}^{\infty} \gamma_k^2 < \infty, \sum_{k=0}^{\infty} (\gamma_k \epsilon^k)^2 < \infty$ and

$$\lim_{k\to\infty} \frac{\epsilon^{k-1} - \epsilon^k}{\gamma_k (\epsilon^k)^2} = 0. \tag{4}$$

Theorem 1: Suppose the $\mathcal{G}$ has a nonempty bounded set of equilibria. Let K be convex and let F be monotone on K. Let assumption (A1) hold. Then $w^k \to w^*$ as $k \to \infty$, where w^k is obtained via the iterative Tikhonov scheme and w^* is the least-norm solution of $\mathrm{VI}(K, F)$.

Proof: By the triangle inequality, $\|w^{k+1} - w^*\|$ can be bounded by terms 1 and 2:

$$\|w^{k+1} - w^*\| \leq \underbrace{\|w^{k+1} - y^k\|}_{term\,1} + \underbrace{\|y^k - w^*\|}_{term\,2}.$$

Of these, term 2 converges to zero from the convergence statement of Tikhonov regularization methods. It suffices to show that term 1 converges to zero as $k \to \infty$ which follows as shown next. By using the non-expansivity of the Euclidean projector, $\|w^{k+1} - y^k\|^2$ is given by

$$\begin{aligned} \|w^{k+1} - y^k\|^2 &= \|\Pi_K \left(w^k - \gamma_k (F(w^k) + \epsilon^k w^k)\right) \\ &\quad - \Pi_K \left(y^k - \gamma_k (F(y^k) + \epsilon^k y^k)\right) \|^2 \\ &\leq \| \left(w^k - \gamma_k (F(w^k) + \epsilon^k w^k)\right) \\ &\quad - \left(y^k - \gamma_k (F(y^k) + \epsilon^k y^k)\right) \|^2. \end{aligned}$$

This expression can be further simplified as

$$\begin{aligned} &\|(1 - \gamma_k \epsilon^k)(w^k - y^k) - \gamma_k (F(w^k) - F(y^k))\|^2 \\ &= (1 - \gamma_k \epsilon^k)^2 \|w^k - y^k\|^2 + \gamma_k^2 \|F(w^k) - F(y^k)\|^2 \\ &- 2\gamma_k (1 - \gamma_k \epsilon_k)(w^k - y^k)^T (F(w^k) - F(y^k)) \\ &\leq (1 - 2\gamma_k \epsilon_k + \gamma_k^2 (L^2 + \epsilon_k^2)) \|w^k - y^k\|^2, \end{aligned}$$

where the last inequality follows from $\gamma_k \epsilon_k \leq 1$ and the monotonicity of $F(x)$ over K. If Lemma 1 can indeed be invoked then it follows that $\|w^{k+1} - y^k\| \to 0$ as $k \to \infty$. The remainder of the proof shows that the conditions for employing Lemma 1 do hold. It can be seen that

$$\begin{aligned} \|w^{k+1} - y^k\| &\leq q_k \|w^k - y^k\| \\ &\leq q_k \|w^k - y^{k-1}\| + q_k \|y^k - y^{k-1}\| \\ &\leq q_k \|w^k - y^{k-1}\| + q_k M \frac{(\epsilon^{k-1} - \epsilon^k)}{\epsilon^k}, \end{aligned}$$

where the second inequality is a consequence of the triangle inequality and the third inequality follows from Lemma 3. Suppose $q_k := \sqrt{(1 - 2\gamma_k \epsilon_k + \gamma_k^2 (L^2 + \epsilon_k^2)}$. Invoking Lemma 1 requires showing that

$$\sum_{k=0}^{\infty} (1 - q_k) = \infty \quad \text{and} \quad \lim_{k\to\infty} \frac{q_k}{1 - q_k} M \frac{(\epsilon^{k-1} - \epsilon^k)}{\epsilon^k} = 0.$$

It is easily seen that

$$\begin{aligned} \sum_{k=0}^{\infty} (1 - q_k) = \sum_{k=0}^{\infty} \frac{1 - q_k^2}{1 + q_k} &= \sum_{k=0}^{\infty} \left(\frac{2\gamma_k \epsilon_k - \gamma_k^2 (L^2 + \epsilon_k^2)}{1 + q_k} \right) \\ &> \sum_{k=0}^{\infty} (2\gamma_k \epsilon_k - \gamma_k^2 (L^2 + \epsilon_k^2)) = \infty, \end{aligned}$$

where the inequality follows from $q_k < 1$ and the final equality follows from $\sum_{k=0}^{\infty} \gamma_k \epsilon^k = \infty$ and the square summability of $\gamma_k \epsilon_k$ and γ^k. The second requirement follows by observing that

$$\begin{aligned} \frac{q_k}{1 - q_k} M \frac{(\epsilon^{k-1} - \epsilon^k)}{\epsilon^k} &= \frac{q_k (1 + q_k)}{1 - q_k^2} M \frac{(\epsilon^{k-1} - \epsilon^k)}{\epsilon^k} \\ &= \frac{q_k (1 + q_k)}{2\gamma_k \epsilon_k - \gamma_k^2 (L^2 + \epsilon_k^2)} M \frac{(\epsilon^{k-1} - \epsilon^k)}{\epsilon^k} \\ &= \underbrace{\frac{q_k (1 + q_k)}{2 - \frac{\gamma_k}{\epsilon^k} (L^2 + \epsilon_k^2)}}_{Term\,1} M \underbrace{\frac{(\epsilon^{k-1} - \epsilon^k)}{\gamma_k (\epsilon^k)^2}}_{Term\,2}. \end{aligned}$$

Since $\gamma_k, \epsilon^k \to 0$, it follows that $q_k \to 1$. Since $\gamma_k/\epsilon^k \to 0$, Term 1 tends to 1 as $k \to \infty$. By assumption, Term 2 tends to zero as $k \to \infty$. ∎

Again, the following result shows that a feasible choice of steplength and regularization parameter sequences do indeed exist for satisfying (A1), under a slightly different set of assumptions from [11].

Lemma 4: Consider assumption (A1). Then the update scheme $\epsilon^k = k^{-\alpha}$ and $\gamma^k = k^{-\beta}$ satisfies (4) in (A1) where $\frac{1}{2} < \alpha + \beta < 1$ and $\beta > \alpha$ for all k.

Proof: Omitted (see [11]). ∎

IV. Iterative proximal point scheme

In this section, we consider an alternate technique that uses a proximal term of the form $\theta(z - z^{k-1})$ rather than $\epsilon^k z$ in modifying the map. Consequently, when such a method is applied to a variational inequality VI(K, F), a sequence of iterates is constructed, each of which requires the solution of a modified strongly modified problem VI(K, F^k)), where $F^k \triangleq F(z) + \theta(z - z^k)$. The convergence of the proximal-point algorithm is shown to hold under an assumption of monotonicity of the original mapping and for a positive θ. Note that the solution of each subproblem, given by $z^{k+1} = (SOL(K, F^k))$, is given by the solution of the fixed-point problem

$$z^{k+1} = \Pi_K(z^{k+1} - \gamma\left(F(z^{k+1}) + \theta(z^{k+1} - z^k)\right).$$

Under assumptions of convexity of the set K and monotonicty of F on K, the convergence of the standard proximal-point algorithm has been established in [1], [5].

In the spirit of the iterative Tikhonov regularization scheme, we present a single timescale *iterative proximal point (IPP) method* alternative. In a game-theoretic generalization of this scheme, a projection step using the deviation between the k^{th} and $k-1^{th}$ iterates yields the $k+1^{th}$ iterate and is formally stated as

$$z_i^{k+1} = \Pi_{K_i}\left(z_i^k - \gamma_k(F_i(z^k) + \theta(z_i^k - z_i^{k-1}))\right), \qquad (5)$$

for $i = 1, \ldots, N$. Next, under an assumption of strict monotonicity of $F(x)$ and the boundedness of K, we establish the global convergence of the IPP scheme.

Theorem 2: Consider the game $\mathcal{G}$ and assume that K is a compact and convex set and F is a continuous and strictly monotone map on K. Let $\{z_k\}$ denote the set of iterates defined by the iterative proximal scheme (5). Let $\sum_{k=1}^\infty \gamma_k = \infty$ and $\sum_{k=1}^\infty \gamma_k^2 < \infty$. Then $\lim_{k\to\infty} z_k = w^*$.

Proof: We begin by expanding $\|z^{k+1} - w^*\|$ and by using the non-expansivity property of projection.

$$\begin{aligned}&\|z^{k+1} - w^*\|^2 = \|\Pi_K(z^k - \gamma_k(F(z^k) + \theta(z^k - z^{k-1})))\\ &- \Pi_K(w^* - \gamma_k F(w^*))\|^2\\ &\le \|(z^k - w^*) - \gamma_k(F(z^k) - F(w^*)) - \gamma_k\theta(z^k - z^{k-1})\|^2.\end{aligned}$$

Expanding the right hand side,

$$\begin{aligned}&\|z^{k+1} - w^*\|^2\\ &\le \|z^k - w^*\|^2 + (\gamma_k L)^2\|z^k - w^*\|^2 + (\gamma_k\theta)^2\|z^k - z^{k-1}\|^2\\ &- 2\gamma_k(z^k - w^*)^T(F(z^k) - F(w^*))\\ &- 2\gamma_k\theta(z^k - z^{k-1})^T\left(z^k - w^* - \gamma_k(F(z^k) - F(w^*))\right).\end{aligned}$$

Using Lipschitz and monotonicity properties of $F(x)$, we have

$$\begin{aligned}&\|z^{k+1} - w^*\|^2\\ &\le (1 + \gamma_k^2 L^2)\|z^k - w^*\|^2 + (\gamma_k\theta)^2\|z^k - z^{k-1}\|^2\\ &\underbrace{-2\gamma_k\theta(z^k - z^{k-1})^T\left((z^k - w^*) - \gamma_k(F(z^k) - F(w^*))\right)}_{\textbf{Term 1}}.\end{aligned}$$

Term 1 can bounded from above by the use of the Cauchy-Schwartz inequality, the boundedness of the iterates, namely $\|z^k - w^*\| \le C$, and the Lipschitz continuity of F, as shown next.

$$\begin{aligned}&\|z^{k+1} - w^*\|^2\\ &\le (1 + \gamma_k^2 L^2)\|z^k - w^*\|^2 + (\gamma_k\theta)^2\|z^k - z^{k-1}\|^2\\ &+ 2\gamma_k\theta\|z^k - z^{k-1}\|\left(\|z^k - w^*\| + \gamma_k\|F(z^k) - F(w^*)\|\right)\\ &\le (1 + \gamma_k^2 L^2)\|z^k - w^*\|^2 + (\gamma_k\theta)^2\|z^k - z^{k-1}\|^2\\ &+ 2\gamma_k\theta C\|(z^k - z^{k-1})\| + 2\gamma_k^2\theta LC\|(z^k - z^{k-1})\|.\end{aligned}$$

Next, we derive a bound on $\|z^k - z^{k-1}\|$ by leveraging the non-expansivity of the Euclidean projector.

$$\begin{aligned}&\|z^k - z^{k-1}\|\\ &= \|\Pi_K(z^{k-1} - \gamma_{k-1}(F(z_{k-1}) + \theta(z^{k-1} - z^{k-2}))))\\ &- \Pi_K(z^{k-1})\|\\ &\le \|(z^{k-1} - \gamma_{k-1}(F(z_{k-1}) + \theta(z^{k-1} - z^{k-2}))) - (z^{k-1})\|\\ &= \| - \gamma_{k-1}(F(z_{k-1}) + \theta(z^{k-1} - z^{k-2})))\|.\end{aligned}$$

It follows from the boundedness of K and the continuity of $F(z)$, that there exists a $\beta > 0$ such that $\|F(z)\| \le \beta$ for all $z \in K$, implying that $\|z^k - z^{k-1}\| \le \gamma_{k-1}(\beta + \theta C)$. The bound on $\|z^k - z^{k-1}\|$ allows us to derive an upper bound $\|z^{k+1} - w^*\|^2$:

$$\begin{aligned}&\|z^{k+1} - w^*\|^2 \le (1 + \gamma_k^2 L^2)\|z^k - w^*\|^2\\ &+ (\gamma_k\theta)^2\gamma_{k-1}^2(\beta + \theta C)^2 + 2\gamma_k\gamma_{k-1}\theta C(1 + \gamma_k L)(\beta + \theta C)\\ &\le (1 + \underbrace{\gamma_k^2 L^2}_{\triangleq v_k})\|z^k - w^*\|^2\\ &+ \underbrace{(\gamma_k\theta)^2\gamma_k^2(\beta + \theta C)^2 + 2\gamma_k^2\theta C(1 + \gamma_k L)(\beta + \theta C)}_{\triangleq p_k}.\end{aligned}$$

The above sequence can be compactly represented as the recursive sequence $u_{k+1} \le (1 + v_k)u_k + p_k$, where

$$\sum_{k=0}^\infty v_k = L^2\sum_{k=0}^\infty \gamma_k^2 < \infty, \quad \sum_{k=0}^\infty p_k < \infty,$$

the latter a consequence of the square summability of γ_k. It follows from Lemma 2 that $u_k \to \bar{u} \ge 0$. It remains to show that $\bar{u} = 0$.

Recall that $\|z^{k+1} - w^*\|^2$ is bounded as per the following expression:

$$\begin{aligned}\|z^{k+1} - w^*\|^2 &\le (1 + v_k)\|z^k - w^*\|^2 + p_k\\ &- 2\gamma_k(z^k - w^*)^T(F(z^k) - F(w^*))\end{aligned}$$

Suppose $\bar{u} > 0$. It follows that along every subsequence, we have that $\mu_k = (z_k - w^*)^T(F(z_k) - w^*) \ge \mu' > 0, \forall k$. This is a consequence of the strict monotonicity of F whereby if $(F(z^k) - F(w^*))^T(z^k - w^*) \to 0$ if $z^k \to w^*$. Since $\bar{u} > 0$, it follows that $\|z^k - w^*\|^2 \to \bar{u} > 0$.

Then by summing over all k, we obtain

$$\lim_{k\to\infty} \|z^{k+1} - w^*\|^2 \le \|z^0 - w^*\|^2 + \sum_{k=0}^{\infty} \gamma_k^2 L^2 \|z^k - w^*\|^2 - 2\sum_{k=0}^{\infty} \gamma_k \mu_k + \sum_{k=0}^{\infty} p_k.$$

Since v_k and p_k are summable and $\mu_k \ge \mu' > 0$ for all k, we have that

$$\begin{aligned}\lim_{k\to\infty} \|z^{k+1} - w^*\|^2 &\le \|z^0 - w^*\|^2 + \sum_{k=0}^{\infty} \gamma_k^2 L^2 \|z^k - w^*\|^2 \\ &- 2\sum_{k=0}^{\infty} \gamma_k \mu_k + \sum_{k=0}^{\infty} p_k \\ &\le \|z^0 - w^*\|^2 + \sum_{k=0}^{\infty} \gamma_k^2 L^2 \|z^k - w^*\|^2 \\ &- 2\mu' \sum_{k=0}^{\infty} \gamma_k + \sum_{k=0}^{\infty} p_k \le -\infty,\end{aligned}$$

where the latter follows from observing that $\sum_{k=0}^{\infty} \gamma_k^2 < \infty, \sum_{k=0}^{\infty} \gamma_k = \infty$ and $\|z_k - w^*\| \le C$. But this is a contradiction, implying that along some subsequence, we have that $\mu_k \to 0$ and $\liminf_{k\to\infty} \|z^k - w^*\|^2 = 0$. But we know that $\{z^k\}$ has a limit point and that the sequence z_k converges. Therefore, we have that $\lim_{k\to\infty} z_k = w^*$. ∎

V. Modified iterative proximal point scheme

In this section, we consider a modified form of the iterative proximal-point scheme. This modification is motivated by observing that the proximal term at the kth iterate, given by $\theta(z^k - z^{k-1})$, is $O(\gamma_{k-1})$, a consequence of the bound developed in the previous result. Therefore, could one develop an iterative proximal-point method where the proximal term was less dependent on the steplength?. Furthermore, it would be expected that this proximal term would converge to zero at a slower rate. Yet, would we find that the behavior of the trajectory was indeed smoother and this proximal term would provide better damping.

We consider a modified IPP scheme in which the ith player takes a step given by

$$z_i^{k+1} = \Pi_{K_i}\left(z_i^k - \gamma_k(F_i(z^k) + \theta_k(z_i^k - z_i^{k-1}))\right), \quad (6)$$

for $i = 1, \ldots, N$, where $\theta_k = c/\gamma_k$ and $c \in (0,1)$. It follows that this scheme can be effectively stated as

$$z_i^{k+1} = \Pi_{K_i}\left((1-c)z_i^k + cz_i^{k-1} - \gamma_k F_i(z^k))\right).$$

Essentially, a projection step is carried out using a convex combination of z_i^k and z_i^{k-1}. Note that even in the standard iterative proximal-point scheme, such a combination is used; however, in that setting, combination is of the form $(1 - \gamma_k\theta)z_i^k + \gamma_k\theta z_i^{k-1}$ and as one proceeds, $\gamma_k \to 0$ and one places less and less emphasis on the past. In this setting, we use a fixed convex combination, specified by the parameter $c \in (0,1)$. Prior to proving our main convergence statement, we prove an intermediate result.

Lemma 5: Suppose $u_k \le cu_{k-1} + \gamma_{k-1}\beta$ where $c \in (0,1)$, $\{\gamma_k\} > 0$ is a decreasing bounded sequence with $\sum_{k=0}^{\infty} \gamma_k < \infty$ and $0 \le u_k < \infty$ for all k. Then, we have that $\sum_{k=1}^{\infty} \gamma_k u_k < \infty$.

Proof: By definition, we have that $u_k \le cu_{k-1} + \gamma_{k-1}\beta$ for $k = 1, \ldots,$. Multiplying this expression by γ_{k-1} and summing over k, we obtain

$$\sum_{j=1}^{k} \gamma_{j-1} u_j \le c\sum_{j=0}^{k-1} \gamma_j u_j + \beta \sum_{j=0}^{k} \gamma_j^2.$$

But $\gamma_k \le \gamma_{k-1}$ for all k, implying that

$$\begin{aligned}\sum_{j=1}^{k} \gamma_j u_j \le \sum_{j=1}^{k} \gamma_{j-1} u_j &\le c\sum_{j=0}^{k-1} \gamma_j u_j + \beta \sum_{j=0}^{k-1} \gamma_j^2 \\ \implies (1-c)\sum_{j=1}^{k-1} \gamma_j u_j + \gamma_k u_k &\le c\gamma_0 u_0 + \beta \sum_{j=0}^{k-1} \gamma_j^2.\end{aligned}$$

By taking limits, it follows that

$$(1-c)\lim_{k\to\infty} \sum_{j=0}^{k-1} \gamma_j u_j + \underbrace{\lim_{k\to\infty} \gamma_k u_k}_{=0} \le c\gamma_0 u_0 + \beta \lim_{k\to\infty} \sum_{j=0}^{k} \gamma_j^2$$

$$\implies \lim_{k\to\infty} \sum_{j=0}^{k-1} \gamma_j u_j < \infty,$$

since $\sum_{j=0}^{\infty} \gamma_j^2 < \infty$, $\lim_{k\to\infty} \gamma_k = 0$ and u_k is bounded. ∎

Proposition 3: Consider a game $\mathcal{G}$ and and assume that K is a compact and convex set and F is a continuous and strictly monotone map on K. Let $\{z_k\}$ denote the set of iterates defined by the iterative proximal scheme (6). Let $\sum_{k=1}^{\infty} \gamma_k = \infty$ and $\sum_{k=1}^{\infty} \gamma_k^2 < \infty$. In addition let $\theta_k = \frac{c}{\gamma_k}$ where $c \in (0,1)$ for $k \ge 0$. Then $\lim_{k\to\infty} z_k = w^*$.

Proof: We begin by observing that $\|z^k - z^{k-1}\|$ can now be bounded as follows:

$$\begin{aligned}&\|z_k - z_{k-1}\| \\ &\le \|\Pi_K(z_{k-1} - \gamma_{k-1}(F(z_{k-1}) + \theta_k(z_{k-1} - z_{k-2}))) \\ &- \Pi_K(z_{k-1})\| \\ &\le \| - \gamma_{k-1}(F(z_{k-1}) + \theta_k(z_{k-1} - z_{k-2}))\| \\ &\le \gamma_{k-1}\|F(z_k)\| + \gamma_{k-1}\theta_{k-1}\|z_{k-1} - z_{k-2}\| \\ &= \gamma_{k-1}\|F(z_k)\| + c\|z_{k-1} - z_{k-2}\|.\end{aligned}$$

The above sequence is of the form $u^{k+1} = q_k u^k + \alpha_k$, where $q_k = \gamma_{k-1}\theta_{k-1} = c$ and $\alpha_k = \gamma_{k-1}\beta$ where $\|F(z^k)\| \le \beta$. Therefore, we have that

$$\sum_{k=1}^{\infty}(1-q_k) = \infty \quad \text{and} \quad \lim_{k\to\infty} \frac{\alpha_k}{1-q_k} = \lim_{k\to\infty} \frac{\alpha_k}{(1-c)} = 0.$$

Therefore from Lemma 1, the sequence $\{z^k\}$ converges and

$$\lim_{k\to\infty} \|z^k - z^{k-1}\| = 0.$$

Since w^* is bounded and fixed, it is clear that, $\|z^k - w^*\|$ converges to $\bar{u} \ge 0$. It suffices to show that $\bar{u} \equiv 0$.

We proceed by contradiction and assume that $\bar{u} > 0$. We begin by recalling the definition of iterates and leveraging properties of the projection operator, we have

$$\begin{aligned}
&\|z_{k+1}-w^*\|^2 \\
&\leq \|z_k-\gamma_k(F(z_k)+\theta_k(z_k-z_{k-1}))-(w^*-\gamma_k F(w^*))\|^2 \\
&= \|(1-\gamma_k\theta_k)(z_k-w^*)-\gamma_k(F(z_k)-F(w^*)) \\
&\quad +\gamma_k\theta_k(z_{k-1}-w^*)\|^2.
\end{aligned}$$

By expanding the expression on the right, we have

$$\begin{aligned}
&\underbrace{(1-\gamma_k\theta_k)^2\|(z_k-w^*)\|^2}_{\textbf{Term 1}} \\
&+\gamma_k^2\|(F(z_k)-F(w^*))\|^2+\underbrace{\gamma_k^2\theta_k^2\|(z_{k-1}-w^*)\|^2}_{\textbf{Term 2}} \\
&+\underbrace{2\gamma_k\theta_k(1-\gamma_k\theta_k)(z_k-w^*)^T(z_{k-1}-w^*)}_{\textbf{Term 3}} \\
&\underbrace{-2\gamma_k^2\theta_k(F(z_k)-F(w^*))^T(z_{k-1}-w^*)}_{\textbf{Term 4}} \\
&\underbrace{-2\gamma_k(1-\gamma_k\theta_k)(z_k-w^*)^T(F(z_k)-F(w^*))}_{\textbf{Term 5}}.
\end{aligned}$$

By using Cauchy-Schwartz on term 3, and subsequently combining with terms 1 and 2, leads to term 6 below. Additionally, terms 4 and 5 when added together, along with the application of Cauchy-Schwartz, lead to the corresponding terms below.

$$\begin{aligned}
&\|z_{k+1}-w^*\|^2 \leq \underbrace{((1-\gamma_k\theta_k)\|z_k-w^*\|+\gamma_k\theta_k\|z_{k-1}-w^*\|)^2}_{\textbf{Term 6}} \\
&-2\gamma_k(z_k-w^*)^T(F(z_k)-F(w^*)) \\
&+\underbrace{2\gamma_k^2\theta_k\|F(z_k)-F(w^*)\|\|z_k-z_{k-1}\|+\gamma_k^2\|F(z_k)-F(w^*)\|^2}_{\triangleq d_k}.
\end{aligned}$$

Consequently, $\|z^{k+1}-w^*\|^2$ can be expressed as

$$\begin{aligned}
\|z^{k+1}-w^*\|^2 &\leq ((1-\gamma_k\theta_k)\|z_k-w^*\|+\gamma_k\theta_k\|z_{k-1}-w^*\|)^2 \\
&\quad -2\gamma_k(z_k-w^*)^T(F(z_k)-F(w^*))+d_k.
\end{aligned}$$

It can be observed that $d_k \leq \gamma_k^2\beta^2+2c\beta\gamma_k\|z^k-z^{k-1}\|$. From Lemma 5, it is clear that $\sum_{k=1}^{\infty}\gamma_k\|z^k-z^{k-1}\|<\infty$. This allows us to claim that $\sum_{k=1}^{\infty}d_k<\infty$. By recalling that $\gamma_k\theta_k=c$, the error $\|z^{k+1}-w^*\|^2$ can be further bounded

$$\begin{aligned}
\|z^{k+1}-w^*\|^2 &\leq ((1-c)\|z_k-w^*\|+c\|z_{k-1}-w^*\|)^2+d_k \\
&\quad -2\gamma_k(z_k-w^*)^T(F(z_k)-F(w^*)).
\end{aligned}$$

By summing over k, we have

$$\begin{aligned}
&\sum_{k=1}^{K}\|z^{k+1}-w^*\|^2 \leq \sum_{k=1}^{K}((1-c)\|z_k-w^*\|+c\|z_{k-1}-w^*\|)^2 \\
&+\sum_{k=1}^{K}d_k-2\sum_{k=1}^{K}\gamma_k(z_k-w^*)^T(F(z_k)-F(w^*)).
\end{aligned}$$

In the expression above, we observe that the coefficient of $\|z^k-w^*\|^2$ for $2\leq k\leq K-2$ is given by $(1-(1-c)^2-c^2)=2c(1-c)$. It follows that the inequality above can be expressed as

$$\begin{aligned}
&\sum_{k=1}^{K}c(1-c)\left(\|z_{k-1}-w^*\|-\|z_k-w^*\|\right)^2-c\|z^0-w^*\|^2 \\
&-\|z^1-w^*\|^2+c\|z^K-w^*\|^2+\|z^{K+1}-w^*\|^2 \\
&\leq \sum_{k=1}^{K}d_k-2\sum_{k=1}^{K}\gamma_k(z_k-w^*)^T(F(z_k)-F(w^*)).
\end{aligned}$$

Taking limits, we have

$$\begin{aligned}
&\sum_{k=1}^{\infty}c(1-c)\left(\|z_k-w^*\|-\|z_{k-1}-w^*\|\right)^2-c\|z^0-w^*\|^2 \\
&-\|z^1-w^*\|^2+c\lim_{k\to\infty}\|z^k-w^*\|^2+\lim_{k\to\infty}\|z^{k+1}-w^*\|^2 \\
&\leq \sum_{k=1}^{\infty}(d_k-\gamma_k\mu_k).
\end{aligned}$$

Since $\bar{u}>0$, it follows that along every subsequence, we have that $\mu_k=2(z_k-w^*)^T(F(z_k)-w^*)\geq\mu'>0,\forall k$. This is a consequence of the strict monotonicity of F whereby if $(F(z^k)-F(w^*))^T(z^k-w^*)\to 0$ if $z^k\to w^*$. By noting that $\sum_{k=0}^{\infty}d_k<\infty$, $\sum_{k=0}^{\infty}\gamma_k\mu_k=\infty$ and the boundedness of z^k, it emerges that

$$\sum_{k=1}^{\infty}(c(1-c))^2((\|z_k-w^*\|-\|z_{k-1}-w^*\|)^2\leq-\infty,$$

a contradiction to the nonnegativity of the left-hand side. Therefore along some subsequence, we have that $\mu_k\to 0$ and $\liminf_{k\to\infty}\|z^k-w^*\|^2=0$. But we know that $\left\{z^k\right\}$ has a limit point and that the sequence z_k converges. Therefore, we have that $\lim_{k\to\infty}z_k=w^*$. ■

References

[1] E. Allevi, A. Gnudi, and I.V. Konnov, **The Proximal Point Method for Nonmonotone Variational Inequalities**, Mathematical Methods of Operations Research **63** (2004), no. 3, 553–565.

[2] T. Alpcan and T. Başar, **A game-theoretic framework for congestion control in general topology networks**, 41th IEEE Conf. Decision and Control (Las Vegas, NV), December 2002.

[3] T. Alpcan and T. Basar, **Distributed algorithms for nash equilibria of flow control games**, Annals of Dynamic Games **7** (2003).

[4] E. Altman, T. Basar, T. Jiménez, and N. Shimkin, **Competitive routing in networks with polynomial cost**, INFOCOM (3), 2000, pp. 1586–1593.

[5] F. Facchinei and J-S. Pang, **Finite dimensional variational inequalities and complementarity problems: Vols i and ii**, Springer-Verlag, NY, Inc., 2003.

[6] I. V. Konnov, **Equilibrium models and variational inequalities**, Elsevier, 2007.

[7] Y. Pan and L. Pavel, **Games with coupled propagated constraints in optical network with multi-link topologies**, Automatica **45** (2009), 871–880.

[8] L. Pavel, **A noncooperative game approach to OSNR optimization in optical networks**, IEEE Transactions on Automatic Control **51** (2006), no. 5, 848–852.

[9] B.T. Polyak, **Introduction to optimization**, Optimization Software Inc., New York, 1987.

[10] G. Scutari, F. Facchinei, J-S. Pang, and D.P. Palomar, **Monotone games in communications: Theory, algorithms, and models**, submitted to IEEE Transactions on Information Theory (April 2010).

[11] H. Yin, U.V. Shanbhag, and P.G. Mehta, **Nash equilibrium problems with congestion costs and shared constraints**, Conditionally accepted by IEEE Transactions on Automatic Control.