A simple and efficient algorithm for computing approximate Nash equilibria in two-person zero-sum sequential games with imcomplete information

Anonymous Author(s)
Affiliation
Address
email

Abstract

We present a simple primal-dual algorithm for computing approximate Nash equilibria in two-person zero-sum sequential games with imcomplete information and perfect recall (like Texas Hold'em Poker) based on the notion of enlarged subgradients. Our algorithm only performs basic iterations (i.e iterations involving no calls to external first-order oracles, etc.) and is applicable to a broad class of two-person zero-sum games including simultaneous games and sequential games with imcomplete information and perfect recall. The applicability to the latter kind of games is thanks to the sequence-form representation [10] which allows us to encode any such game as a matrix game with polyhedral strategy profiles, the size of the representation being linear in ... The number of iterations needed to produce a Nash equilibrium with a given precision is inversely proportional to the desired precision. As proof-of-concept, we apply our algorithm to solve Kuhn 3-card Poker.

keywords: sequential game; imcomplete information; perfect recall; approximate Nash equilibrium; primal-dual algoithm; convex-optimization

1 Introduction

A game-theoretic approach to playing games strategically optimally consists in computing Nash-equilibria (infact, approximations thereof) offline, and playing one's part of the equilibrium online. This technique is the driving-force behind solution concepts like CFR [18, 13], CFR$^+$ **TODO: add ref!!!**, and other variants, which have recently had profound success in Poker. However, solving games for equilibria remains a mathematical and computational challenge, especially in sequential games with imperfect information. This paper proposes a simple and fast algorithm for solving for such equilibria approximately, in a sense which will be made clear shortly.

1.1 Notation, terminology, and some aspects of modern convex analysis

General notions. Given a set X, 2^X denotes the *powerset* of X, i.e the set of all subsets of X, or equivalently the set of all binary functions on X. Let m and n be positive integers. Given two vectors $z, w \in \mathbb{R}^n$, their inner product will be denoted $\langle z, w\rangle := \sum_j z_j w_j$. The components of z will be denoted z_0, z_1, ..., z_{n-1} (indexing begins from 0, not 1). The notation "$z \geq 0$" means that all the components of z are nonnegative. $\mathbb{R}^n_+ := \{z \in \mathbb{R}^n \mid z \geq 0\}$ is the nonnegative n-dimensional *orthant*. $\|z\|$ denotes the 2-*norm* of z defined by $\|z\| := \sqrt{\langle z, z\rangle}$. $(z)_+ := \max(0, z) \in \mathbb{R}^n_+$ is the point-wise maximum of z with 0. For example, $((-2, \pi))_+ = (max(-2, 0), max(\pi, 0)) = (0, \pi)$. The operator $(.)_+$ is the well-known (multi-dimensional) *ramp* function. The n-simplex denoted

Δ_n, is defined by $\Delta_n := \{z \in \mathbb{R}^n_+ | z_0 + z_1 + ... z_{n-1} = 1\}$. Given a matrix $A \in \mathbb{R}^{m \times n}$, its *spectral norm*, denoted $\|A\|$, is defined to be the largest *singular value* of A, i.e the largest *eigenvalue* of $A^T A$ (or equivalently, of AA^T).

Modern convex analysis. Given a subset C of $\mathbb{R}^n$, i_C denotes its *indicator function* defined by

$$i_C(x) = 0 \text{ if } x \in C \text{ and } +\infty \text{ otherwise.} \tag{1}$$

At times, we will write $i_{x \in C}$ for $i_C(x)$ (to ease notation, etc.). Let $f : \mathbb{R}^n \to (-\infty, +\infty]$ be a convex function. The *effective domain* of f, denoted $dom(f)$, is defined as

$$dom(f) := \{x \in \mathbb{R}^n | f(x) < +\infty\}. \tag{2}$$

If $dom(f) \neq \emptyset$ then we say f is *proper*. The *subgradient* of f is the set-valued function

$$\partial f : \mathbb{R}^n \to 2^{\mathbb{R}^n},\ \ x \mapsto \{s \in \mathbb{R}^n | f(z) \geq f(x) + \langle s, z - x\rangle, \forall z \in \mathbb{R}^n\}. \tag{3}$$

Of course $\partial f(x)$ reduces to the singleton $\{\nabla f(x)\}$ in case f is differentiable at x. If f is convex, its *proximal operator* is the function $\mathrm{prox}_f : \mathbb{R}^n \to \mathbb{R}^n$ defined by

$$\mathrm{prox}_f(x) := \min_{z \in \mathbb{R}^n} \frac{1}{2}\|z - x\|^2 + f(z). \tag{4}$$

For example, if C is a closed convex subset of $\mathbb{R}^n$, then $\mathrm{prox}_C = \mathrm{proj}_C$, the orthogonal projector onto C. Thus proximal operators generalize orthogonal projectors. For example $\mathrm{prox}_{i_{\mathbb{R}^n_+}}(z) \equiv \mathrm{proj}_{\mathbb{R}^n_+}(z) \equiv (z)_+$. One also has the useful characterization

$$p = \mathrm{prox}_f(x) \text{ iff } x - p \in \partial f(p). \tag{5}$$

The interested reader should refer to [5] for a more elaborate exposition of proximal operators and their use in modern convex-optimization.

1.2 Nash-equilibrium concepts and statement of the problem

1.2.1 Preliminaries

The sequence-form for two-person zero-sum games was introduced in [10], and the theory was further developed in [11, 17, 16], where it was established that for sequential two-person zero-sum games with imcomplete information and perfect recall, there exist sparse matrices $A \in \mathbb{R}^{n_1 \times n_2}$, $E_1 \in \mathbb{R}^{l_1 \times n_1}$, $E_2 \in \mathbb{R}^{l_2 \times n_2}$, and vectors $e_1 \in \mathbb{R}^{l_1}, e_2 \in \mathbb{R}^{l_2}$ such that n_1, n_2, l_1, and l_2 are all linear in the size of the game tree (number of states in the game) and such that Nash equilibria correspond to pairs (x, y) of *realization plans* which solve the primal LCP

$$\text{minimize } \langle e_1, p\rangle, \quad \text{subject to: } (y, p) \in \mathbb{R}^{n_2} \times \mathbb{R}^{l_1}, y \geq 0, E_2 y = e_2, -Ay + E_1^T p \geq 0. \tag{6}$$

and the dual LCP

$$\text{maximize} - \langle e_2, q\rangle, \ \text{subject to: } (x, q) \in \mathbb{R}^{n_1} \times \mathbb{R}^{l_2}, x \geq 0, E_1 x = e_1, A^T x + E_2^T q \geq 0. \tag{7}$$

The vectors $p = (p_0, p_1, ..., p_{l_2 - 1}) \in \mathbb{R}^{l_2}$ and $q = (q_0, q_1, ..., q_{l_1 - 1}) \in \mathbb{R}^{l_1}$ are dual variables. A is the *payoff matrix* and each E_k is a matrix whose entries are -1, 0 or 1, and each e_k is a vector of the form $(1, 0, ..., 0)$. In this so-called *sequence-form* representation, the strategy profile of player k is the convex polyhedron

$$Q_k := \{z \in \mathbb{R}^{n_k}_+ |\ E_k z = e_k\}. \tag{8}$$

Note that the LCPs above have the equivalent saddle point formulation

$$\underset{y \in Q_2}{\text{minimize}}\ \underset{x \in Q_1}{\text{maximize}}\ \langle x, Ay\rangle \tag{9}$$

At a feasible point (y, p, x, q), the *primal-dual gap* $\tilde{G}(y, p, x, q)$ of these primal-dual pair of LCPs is given by[1]

$$\begin{aligned} 0 \leq \tilde{G}(y, p, x, q) &:= \langle e_1, p\rangle - (\langle e_2, q\rangle) = \langle e_1, p\rangle + \langle e_2, q\rangle \\ &= G(x, y) := \max\{\langle u, Ay\rangle - \langle x, Av\rangle | (u, v) \in Q_1 \times Q_2\}. \end{aligned} \tag{10}$$

[1] The inequality being due to *weak duality*.

It was shown (see Theorem 3.14 of [16]) that a pair $(x, y) \in Q_1 \times Q_2$ of realization plans is a solution to the LCPs (6) and (7) (i.e is a Nash equilibrium for the game) if and only if there exist vectors p and q such that

$$-Ay + E_1^T p \geq 0, \;\; A^T x + E_2^T q \geq 0, \; \langle x, -Ay + E_1^T p\rangle = 0, \; \langle y, A^T x + E_2^T q\rangle = 0. \tag{11}$$

Moreover, at equilibria, *strong duality* holds and the value of the game equals $p_0 = -q_0$, i.e the primal-dual gap $\tilde{G}(y, p, x, q)$ defined in (10) vanishes at equilibria.

1.2.2 A note about matrix games on simplexes.

It should be noted that any matrix $A \in \mathbb{R}^{n_1 \times n_2}$ specifies a matrix game with payoff matrix A, for which each player's strategy profile is a simplex; this simplex can be written in the form (8) by taking $E_k := (1, 1, ..., 1) \in \mathbb{R}^{1 \times n_k}$ and $e_k = 1 \in \mathbb{R}^1$. Thus every matrix game on simplexes can be seen as a sequential game. Thus the results presented in this manuscript can be trivially applied such games in particular. Here, the polyhedral strategy profiles Q_k defined in (8) reduce to simplexes Δ_{n_k}, and the primal-dual gap function $G(x, y)$ writes

$$G(x, y) = \max\{\langle u, Ay\rangle - \langle x, Av\rangle | (u, v) \in \Delta_{n_1} \times \Delta_{n_2}\} = \max_{0 \leq j < n_1} (Ay)_j - \min_{0 \leq i < n_2} (A^T x)_i. \tag{12}$$

1.2.3 Approximate feasibility and optimality

Solving the LCPs (6) and (7) exactly is impossible in practice (indeed, this system of problems is NP-hard [10]) and such a precision doesn't have any fundamental practical advantage. Instead, it is customary compute "approximate" Nash equilibria.

Definition 1 (**Nash ϵ-equilibria**)**.** *Given $\epsilon > 0$, a Nash ϵ-equilibrium is a pair (x^*, y^*) of realization plans such that there exists dual vectors p^* and q^* for problems* (6) *and* (7) *such that the primal-dual gap at (y^*, p^*, x^*, q^*) doesn't exceed ϵ. That is,*

$$0 \leq \tilde{G}(y^*, p^*, x^*, q^*) \leq \epsilon. \tag{13}$$

Let us now, introduce the notion of ϵ-subgradients [8].

Definition 2 (ϵ-subgradient)**.** *Given a scalar $\epsilon > 0$ and a function $f : \mathbb{R}^n \to [-\infty, +\infty]$, the ϵ-enlarged subgradient (or ϵ-subgradient, for short) of f is the set-valued function*

$$\partial_\epsilon f : \mathbb{R}^n \to 2^{\mathbb{R}^n}, \;\; x \mapsto \{s \in \mathbb{R}^n | f(z) \geq f(x) + \langle s, z - x\rangle - \epsilon, \forall z \in \mathbb{R}^n\}. \tag{14}$$

The idea behind ϵ-subgradients is the following. Say we wish to minimize the function f. Replace the usual necessary condition "$0 \in \partial f(x)$" for the optimality of x with the weaker condition "$\partial_\epsilon f(x)$ contains a sufficiently small vector v". In fact, it is easy to see that, for each point $x \in \mathbb{R}^n$, we have $\lim_{\epsilon \to 0^+} \partial_\epsilon f(x) = \partial f(x)$. This approximation concept for subgradients yields the following concept of approximate Nash equlibria (adapted from [8]).

Definition 3 (**Nash (ϵ_1, ϵ_2)-equilibria**)**.** *Given tolerance levels $\epsilon_1, \epsilon_2 > 0$, a Nash (ϵ_1, ϵ_2)-equilibrium for the GSP* (21) *is any quadruplet (x^*, y^*, x^*, q^*) for which there exists a perturbation vector v^* such that*

$$\|v^*\| \leq \epsilon_1 \textit{ and } \; v^* \in \partial_{\epsilon_2}[\hat{\Psi}_1(.,., x^*, q^*) + \hat{\Psi}_2(y^*, p^*, .,.)](y^*, p^*, x^*, q^*). \tag{15}$$

Such a vector v^* is called a Nash ϵ-residual at (x^*, y^*, x^*, q^*).

The above definition is a generalization of the notion of Nash equilibria since:

- Exact Nash equilibria correspond to Nash $(0, 0)$-equilibria.
- Nash ϵ-equilibria (in the sense of Definition 1) correspond to Nash $(0, \epsilon)$-equilibria.

1.3 Our contribution

We develop a $\mathcal{O}(1/\epsilon)$ primal-dual algorithm (Algorithm 1) for computing Nash $(\epsilon, 0)$-equilibria in sequential two-person zero-sum games with imcomplete information and perfect recall.

2 Related work

We present a selection of algorithms that is representative of the efforts that have been made in the literature to compute Nash ϵ-equilibria for two-person zero-sum games with imcomplete information like Texas Hold'em Poker, etc.

In [9], a nested iterative procedure using the Excessive Gap Technique (EGT) [14] was used to to solve the equilibrium problem (9). The authors reported a $\mathcal{O}(1/\epsilon)$ convergence rate (which derives from the general EGT theory) for the outer-most iteration loop.

[7] proposed a modified verson of the technqiues in [9] and proved a $\mathcal{O}\left(\left(\|A\|/\delta\right) ln\left(1/\epsilon\right)\right)$ convergence rate in terms of the number of calls made to a first-order oracle. Here $\delta = \delta(A, E_1, E_2, e_1, e_2) > 0$ is a certain *condition number* for the game. The crux of their technique was to observe that (9) can further be written a the minimization of the primal-dual gap function $G(x, y)$ (defined in (10)) for the game[2], viz

$$\text{minimize}\{G(x, y) | (x, y) \in Q_1 \times Q_2\}, \tag{16}$$

and then show there exists a scalar $\delta > 0$ such that for any pair of realization plans $(x, y) \in Q_1 \times Q_2$,

$$\text{"distance between } (x, y) \text{ and the set of equilibria"} \le G(x, y)/\delta. \tag{17}$$

Their algorithm is then derived by iteratively applying Nesterov smoothing [15] with a geometrically decreasing sequence of tolerance levels $\epsilon_{n+1} = \epsilon_n/\gamma$ (with $\gamma > 1$) G. It should be noted however that

- The constant $\delta > 0$ can be arbitrarily small, and so the factor $\|A\|/\delta$ in the $\mathcal{O}\left(\left(\|A\|/\delta\right) ln\left(1/\epsilon\right)\right)$ convergence rate can be arbitrarily large for ill-conditioned games.
- The reported linear convergence rate is not in terms of basic operations (addition, multiplication, matvec, clipping, etc.), but in terms of the number of calls to a first-order oracle. For example the complicated projections proj_{Q_k} are applied at each iteration.

The primal-dual algorithm first developed in [2] was proposed [3] in to solve matrix games on simplexes. It should be stressed that such matrix games are considerably simpler than the games considered here. Indeed, the authors in [3] use the fact that computing the euclidean projection of a point unto a simplex can be done in linear time as in [6]. In contrast, no such efficient algorithm is known nor is likely to exist for the polyhedra Q_k defined in (8), the strategy profiles for players in the games considered here. It should however be noted that such this projection can still be done iteratively using, for example, the algorithm in proposition 4.2 of [4]. Unfortunately, as with any nested iterative scheme, one would have to solve this sub-problem with finer and finer precision.

Sampling techniques like the CFR (CounterFactual Regret minimization) and its many variants [**?**, 13, 1] have also become state-of-the-art.

Finally, let us note that for the class of games considered here (sequential games with imcomplete information), the LCPs (6) and (7) are exceedingly larger than what state-of-the-art LCP and interior-point solvers can handle. See for example [9].

3 Reformulation as a Generalized Saddle-point Problem (GSP)

One cannot directly attack the LCPs (6) and (7) via a traditional primal-dual algorithm (for example [2, 3]) because computing the euclidean projections proj_{Q_k} is very difficult (in fact, such subproblems would have to be solved iteratively[3]). Also, the primal-dual gap might explode even at "nearly

[2]The minimizers of G are precisely the equilibria of the game.

[3]An "exception to the rule" is the case where the $Q_k's$ are simplexes, so that the projections can be computed exactly using [6].

feasible" points, leaving a primal-dual algorithm with no clear on whether progress is being made or not. So need to way to

- avoid having to compute the projections proj_{Q_k},
- have control over how far we are from the set of equilibria and avoid infinite (and thus non-informative) primal-dual gaps.

In the next theorem, we show that the said LCP problems can be conveniently written as a *Generalized Saddle-point Problem (GSP)* in the sense of [8].

Theorem 4. *Define two proper closed convex functions*

$$\left.\begin{aligned} g_1 &: \mathbb{R}^{n_2} \times \mathbb{R}^{l_1} \to (-\infty, +\infty], \quad g_1(y,p) := i_{y\geq 0} + \langle e_1, p\rangle \\ g_2 &: \mathbb{R}^{n_1} \times \mathbb{R}^{l_2} \to (-\infty, +\infty], \quad g_2(x,q) := i_{x\geq 0} + \langle e_2, q\rangle \end{aligned}\right\} \tag{18}$$

Also define two bilinear forms Ψ_1, $\Psi_2 : \mathbb{R}^{n_2} \times \mathbb{R}^{l_1} \times \mathbb{R}^{n_1} \times \mathbb{R}^{l_2} \to \mathbb{R}$ with $\Psi_2 = -\Psi_1$ by letting

$$K := \begin{bmatrix} A & -E_1^T \\ E_2 & 0 \end{bmatrix}, \quad \Psi_1(y,p,x,q) := \left\langle \begin{bmatrix} x \\ q \end{bmatrix}, K \begin{bmatrix} y \\ p \end{bmatrix} \right\rangle = \langle x, Ay\rangle - \langle x, E_1^T p\rangle + \langle q, E_2 y\rangle, \tag{19}$$

and, define the functions $\hat{\Psi}_1$, $\hat{\Psi}_2 : \mathbb{R}^{n_2} \times \mathbb{R}^{l_1} \times \mathbb{R}^{n_1} \times \mathbb{R}^{l_2} \to (-\infty, +\infty]$ by

$$\left.\begin{aligned} &\hat{\Psi}_1(y,p,x,q) := \Psi_1(y,p,x,q) + g_1(y,p) \text{ if } y \geq 0, \hat{\Psi}_1(y,p,x,q) := \infty \text{ otherwise} \\ &\hat{\Psi}_2(y,p,x,q) := \Psi_2(y,p,x,q) + g_2(x,q) \text{ if } x \geq 0, \hat{\Psi}_2(y,p,x,q) := \infty \text{ otherwise.} \end{aligned}\right\} \tag{20}$$

Finally, define the sets $S_1 := \mathbb{R}_+^{n_2} \times \mathbb{R}^{l_1}$ and $S_2 := \mathbb{R}_+^{n_1} \times \mathbb{R}^{l_2}$ and consider the GSP(Ψ_1, Ψ_2, g_1, g_2): Find a quadruplet $(y^*, p^*, x^*, q^*) \in S_1 \times S_2$ such that

$$\hat{\Psi}_1(y^*, p^*, x^*, q^*) \leq \hat{\Psi}_1(y, p, x^*, q^*) \text{ and } \hat{\Psi}_2(y^*, p^*, x^*, q^*) \leq \hat{\Psi}_2(y^*, p^*, x, q) \forall (y,p,x,q) \in S_1 \times S_2. \tag{21}$$

Then a quadruplet $(y^*, p^*, x^*, q^*) \in \mathbb{R}^{n_2} \times \mathbb{R}^{l_1} \times \mathbb{R}^{n_1} \times \mathbb{R}^{l_2}$ solves the Nash equilibrium LCPs (6) *and* (7) *iff it solves GSP(Ψ_1, Ψ_2, g_1, g_2). In order words, GSP(Ψ_1, Ψ_2, g_1, g_2) is equivalent to the Nash equilibrium LCPs* (6) *and* (7)*.*

Proof. Note that $dom(g_1) = S_1$ and $dom(g_2) = S_2$ and observe that since the equilibria of the LCPs (6) and (7), and the equilibria of GSP(Ψ_1, Ψ_2, g_1, g_2) are contained in $S_1 \times S_2$, it suffices to show that at any point $(y,p,x,q) \in S_1 \times S_2$, the primal-dual gap between the primal LCP (6) and the dual LCP (7) equals the primal-dual gap of GSP(Ψ_1, Ψ_2, g_1, g_2). Indeed, the unconstrained objective in (6), say $a(x,y)$, can be computed as

$$\begin{aligned} a(y,p) &= \langle e_1, p\rangle + i_{y\geq 0} + i_{-Ay+E_1^T p\geq 0} + i_{E_2 y = e_2} \\ &= g_1(y,p) + \max_{x'\geq 0} \langle x', Ay - E_1^T p\rangle + \max_{q'} \langle q', E_2 y - e_2\rangle \\ &= g_1(y,p) + \max_{x',q'} \langle x', Ay\rangle - \langle x', E_1^T p\rangle + \langle q', E_2 y\rangle - (i_{x'\geq 0} + \langle e_2, q\rangle) \\ &= g_1(y,p) - \min_{x',q'} \Psi_2(y,p,x',q') + g_2(x',q') = g_1(y,p) - \underbrace{\min_{x',q'} \hat{\Psi}_2(y,p,x',q')}_{\phi_2(y,p)} \\ &= g_1(y,p) - \phi_2(y,p). \end{aligned}$$

Similarly, the unconstrained objective, say $b(x,q)$, in the dual LCP (7) writes

$$\begin{aligned} b(x,q) &= -\langle q, e_2\rangle - i_{x\geq 0} - i_{A^T x + E_2^T q \geq 0} - i_{E_1 x = e_1} \\ &= -g_2(x,q) + \min_{y'\geq 0} \langle y', A^T x + E_2^T q\rangle + \min_{p'} \langle p', e_1 - E_1 x\rangle \\ &= -g_2(x,q) + \min_{y',p'} \Psi_1(y',p',x,q) + g_1(y',p') = -g_2(x,q) + \underbrace{\min_{y',p'} \hat{\Psi}_1(y',p',x,q)}_{\phi_1(x,q)} \\ &= -g_2(x,q) + \phi_1(x,q). \end{aligned}$$

Thus, noting that $-\infty \le \phi_1(x,q), \phi_2(y,p), < +\infty$ (so that all the operations below are valid), one computes the primal-dual gap between the primal LCP (6) and dual the LCP (7) at (y,p,x,q) as

$$\begin{aligned} a(y,p) - b(x,q) &= g_1(y,p) - \phi_2(y,p) + g_2(x,q) - \phi_1(x,q) \\ &= \Psi_1(y,p,x,q) + g_1(y,p) - \phi_2(y,p) + \Psi_2(y,p,x,q) + g_2(x,q) - \phi_1(x,q) \\ &= \hat{\Psi}_1(y,p,x,q) + \hat{\Psi}_2(y,p,x,q) - \phi_1(x,q) - \phi_2(y,p) \\ &= \text{primal-dual gap of GSP}(\Psi_1, \Psi_2, g_1, g_2) \text{ at } (y,p,x,q), \end{aligned}$$

where the second equality follows from the zero-sum condition $\Psi_1 + \Psi_2 := 0$. □

By Theorem 4, solviing for a Nash equilibrium for the game is equivalent to solving the GSP (21), which as it turns out, is simpler conceptually. The rest of the paper will be devoted to developing an algorithm for solving the latter.

4 The proposed primal-dual algorithm

We now derive the algorithm derive the algorithm which is the main object of this manuscript.

4.1 Algorithm and convergence rate

Algorithm 1 Primal-dual algorithm for computing approximate Nash Equilbria in two-person zero-sum games with imcomplete information and perfect recall

Require: $\epsilon > 0$; $(y^{(0)}, p^{(0)}, x^{(0)}, q^{(0)}) \in \mathbb{R}^{n_2} \times \mathbb{R}^{l_1} \times \mathbb{R}^{n_1} \times \mathbb{R}^{l_2}$.
Ensure: A Nash (ρ, ϵ)-equilibrium $(y^*, p^*, x^*, q^*) \in S_1 \times S_2$ for the GSP (21).

1: $\lambda \leftarrow 1/\|K\|$, $v_a^{(0)} \leftarrow 0$, $k \leftarrow 0$
2: **while** $\|v_a^{(k)}\|/k \ge \epsilon$ **do**
3: $\quad y^{(k+1)} \leftarrow (y^{(k)} - \lambda(A^T x^{(k)} + E_2^T q^{(k)}))_+$, $\;p^{(k+1)} \leftarrow p^{(k)} - \lambda(e_1 - E_1 x^{(k)})$
4: $\quad x^{(k+1)} \leftarrow (x^{(k)} + \lambda(Ay^{(k+1)} - E_1^T p^{(k+1)}))_+$, $\;\Delta x^{(k+1)} \leftarrow x^{(k+1)} - x^{(k)}$
5: $\quad \Delta q^{(k+1)} \leftarrow \lambda(E_2 y - e_2)$, $\;q^{(k+1)} \leftarrow q^{(k)} + \Delta q^{(k+1)}$
6: $\quad y^{(k+1)} \leftarrow y^{(k+1)} - \lambda(A^T \Delta x^{(k+1)} + E_2^T \Delta q^{(k+1)})$, $\;\Delta y^{(k+1)} \leftarrow y^{(k+1)} - y^{(k)}$
7: $\quad p^{(k+1)} \leftarrow p^{(k+1)} + \lambda E_1 \Delta x^{(k+1)}$, $\;\Delta p^{(k+1)} \leftarrow p^{(k+1)} - p^{(k)}$
8: $\quad v_a^{(k+1)} \leftarrow v_a^{(k)} + \frac{1}{\lambda}(\Delta y^{(k+1)}, \Delta p^{(k+1)}, \Delta x^{(k+1)}, \Delta q^{(k+1)})$
9: $\quad k \leftarrow k+1$
10: **end while**

Theorem 5 (Ergodic $\mathcal{O}(1/\epsilon)$ convergence)**.** *Let d_0 be the euclidean distance between the starting point $(y^{(0)}, p^{(0)}, x^{(0)}, q^{(0)})$ of Algorithm 1 and the set of equilibria for the GSP (21). Then given any $\epsilon > 0$, there exists an index $k_0 \le \frac{2d_0\|K\|}{\epsilon}$ such that after k_0 iterations the algorithm produces the following a quadruplet $(y^{k_0}, p^{k_0}, x^{k_0}, q^{k_0})$ and a vector v^{k_0} such that*

$$\|v^{k_0}\| \le \epsilon \textit{ and } \; v^{k_0} \in \partial[\hat{\Psi}_1(.,.,x^{k_0}, q^{k_0}) + \hat{\Psi}_2(y^{k_0}, p^{k_0}, .,.)](y^{k_0}, p^{k_0}, x^{k_0}, q^{k_0}). \tag{22}$$

Thus Algorithm 1 outputs an $(\epsilon, 0)$-Nash equilibrium for the GSP (21) in at most $\frac{2d_0\|K\|}{\epsilon}$ iterations.

Proof. It is clear to see that the quadruplet $(\Psi_1, \Psi_2, g_1, g_2)$ satisfies assumptions B.1, B.2, B.3, B.5, and B.6 of [8] with $L_{xx} = L_{yy} = 0$ and $L_{xy} = L_{yx} = \|K\|$. Now, one easily computes the proximal operator of g_j in closed-form as $\text{prox}_{\lambda g_j}(z,v) \equiv ((z)_+, v - \lambda e_j)$. With all these ingredients in place, Algorithm 1 is then obtained from [8, Algorithm T-BD], applied on the GSP (21) with the choice of parameters: $\sigma = 1 \in (0,1]$, $\sigma_x = \sigma_y = 0 \in [0,\sigma)$, $\lambda_{xy} := \frac{1}{\sigma L_{xy}}\sqrt{(\sigma^2 - \sigma_x^2)(\sigma^2 - \sigma_y^2)} = \sigma/\|K\| = 1/\|K\|$, and $\lambda = \lambda_{xy} \in (0, \lambda_{xy}]$. The convergence result follows immediately from [8, Theorem 4.2]. □

4.2 Practical considerations

Efficient computation of Ay and A^Tx. In Algorithms 1 most of the time is spent pre-multiplying vectors by A and A^T. For *flop*-type Poker games like *Texas Hold'em* and *Rhode Island Hold'em*, A (and thus A^T too) is very big (up 10^{14} rows and columns!) but has a rich block-diagnoal structure which can be carefully exploited, as was done in [9].

Computing $\|K\|$. Power iteration (Perron-Frobenius)...

5 Experimantal results

To access the practical quality of the proposed algorithm, we experimented it on real and simulated matrix games.

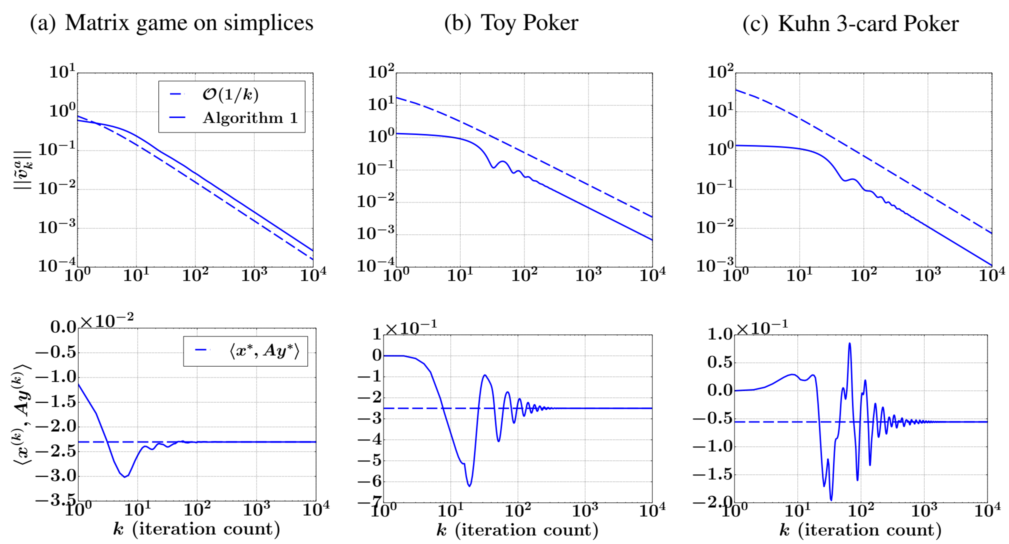

Figure 1: Convergence curves of Algorithm 1. **Top row**: Evolution of ergodic primal-dual gap. **Bottom row**: Evolution of value of game.

Matrix game on simplexes. As in [15, 3], we generate a 1000×1000 random matrix whose entries are uniformly identically distributed in the closed interval $[-1, 1]$.

Toy Poker game. The sequence-form representation for this Toy Poker game is given by (showing only nonzero entries):

$E_1 \in \mathbb{R}^{3\times 5}$ with $E_1(0,0) = E_1(1,1) = E_1(1,2) = E_1(2,3) = E_1(2,4) = 1$, $E_1(1,0) = E_1(2,0) = -1$;

$E_2 \in \mathbb{R}^{3\times 5}$ with $E_2(0,0) = E_2(1,1) = E_2(1,2) = E_2(2,3) = E_2(2,4) = 1$, $E_2(1,0) = E_2(2,0) = -1$; and

$A \in \mathbb{R}^{5\times 5}$ with $A(2,0) = A(4,0) = -0.5$, $A(1,3) = 1$, $A(3,1) = -1$, $A(1,2) = A(1,4) = A(3,2) = A(3,4) = 0.25$.

Kuhn 3-card Poker. The game has $n_1 = n_2 = 13$ and $l_1 = l_2 = 7$, and its sequence-form is given by:

$E_1 \in \mathbb{R}^{7\times 13}$ with $E_1(0,0) = E_1(1,9) = E_1(1,12) = E_1(2,1) = E_1(2,4) = E_1(3,5) = E_1(3,8) = E_1(4,2) = E_1(4,3) = E_1(5,6) = E_1(5,7) = E_1(6,10) = E_1(6,11) = 1$, $E_1(1,0) = E_1(2,0) = E_1(3,0) = E_1(4,1) = E_1(5,5) = E_1(6,9) = -1$;

$E_2 \in \mathbb{R}^{7\times 13}$ with $E_2(0,0) = E_2(1,7) = E_2(1,8) = E_2(2,9) = E_2(2,10) = E_2(3,5) = E_2(3,6) = E_2(4,11) = E_2(4,12) = E_2(5,1) = E_2(5,2) = E_2(6,3) = E_2(6,4) = 1$, $E_2(1,0) = E_2(2,0) = E_2(3,0) = E_2(4,0) = E_2(5,0) = E_2(6,0) = -1$; and

$A \in \mathbb{R}^{13\times 13}$ with $A(3,8) = A(3,12) = A(4,6) = A(4,10) = A(7,12) = A(8,10) = -0.333333$, $A(1,7) = A(1,11) = A(2,8) = A(2,12) = A(5,11) = A(6,4) = A(6,12) = A(10,4) = A(10,8) = -0.166667$, $A(7,4) = A(8,2) = A(11,4) = A(11,8) = A(12,2) = A(12,6) = 0.333333$, $A(4,5) = A(4,9) = A(5,3) = A(8,1) = A(8,9) = A(9,3) = A(9,7) = A(12,1) = A(12,5) = 0.166667$.

The pair $(x^*, y^*) \in \mathbb{R}^{13} \times \mathbb{R}^{13}$ of realization plans given by

$$
\begin{aligned}
x^* &= [1, 0.759, 0.759, 0, 0.241, 1, 0.425, 0.575, 0, 0.275, 0, 0.275, 0.725]^T,\\
y^* &= [1, 1, 0, 0.667, 0.333, 0.667, 0.333, 1, 0, 0, 1, 0, 1]^T
\end{aligned}
$$

is a Nash 10^{-4}-equlibrium computed in 1500 iterations of Algorithm 1. The convergence curves are shown in Fig 1. One easy checks that this equilibrium is feasible. Indeed,

$$
\begin{aligned}
E_1x^* - e_1 = [&4.76 \times 10^{-5}, -1.91 \times 10^{-5}, 5.67 \times 10^{-5}, 8.23 \times 10^{-6}, 2.90 \times 10^{-5},\\
&- 8.62 \times 10^{-7}, -1.96 \times 10^{-5}]^T
\end{aligned}
$$

and

$$
\begin{aligned}
E_2y^* - e_2 = [&- 7.04 \times 10^{-7}, 2.27 \times 10^{-6}, -3.29 \times 10^{-6}, -1.50 \times 10^{-6},\\
&2.92 \times 10^{-6}, -4.97 \times 10^{-7}, -5.85 \times 10^{-7}]^T
\end{aligned}
$$

Finally, one checks that

$$
x^{*T}Ay^* = -0.05555,
$$

which agrees to 5 d.p with the value of $-1/18$ computed analytically by H. W. Kuhn in his 1950 paper [12]. The evolution of the dual gap and the expected value of the game across iterations are shown in Figure 1.

6 Concluding remarks

Making use of the sequence-form representation [10, 17, 16], we have deviced a primal-dual algorithm for computing Nash equilibria in two-person zero-sum sequential games with imcomplete information (like Texas Hold'em, etc.). Our algorithm is simple to implement, with a very low constant cost per iteration, and enjoys a rigorous convergence theory with a proven $\mathcal{O}(1/min(\rho,\epsilon))$ convergence in terms of basic operations (matvec products, clipping, etc.), to a Nash (ρ,ϵ)-equilibrium of the game. Equilibrium problems are saddle-point convex-concave problems, and as such a natural choice for algorithms for solving them would be in the family of primal-dual algorithms. The author believes such primal-dual schemes will receive more attention in the algorithmic game theory community in future.

References

[1] Bowling, M., Burch, N., Johanson, M., Tammelin, O.: Heads-up limit holdem poker is solved. Science 347(6218), 145–149 (2015)

[2] Chambolle, A., Pock, T.: A first-order primal-dual algorithm for convex problems with applications to imaging. J. Math. Imaging Vis. 40, 120 (2011)

[3] Chambolle, A., Pock, T.: On the ergodic convergence rates of a first-order primal-dual algorithm (2014)

[4] Combettes, P.L., Dũng, D., Vũ, B.C.: Dualization of signal recovery problems. Set-Valued and Variational Analysis 18(3-4), 373–404 (2010)

[5] Combettes, P.L., Pesquet, J.C.: Proximal splitting methods in signal processing. In: Fixed-point algorithms for inverse problems in science and engineering, pp. 185–212. Springer New York (2011)

[6] Duchi, J., Shalev-Shwartz, S., Singer, Y., Chandra, T.: Efficient projections onto the l 1-ball for learning in high dimensions. In: Proceedings of the 25th international conference on Machine learning. pp. 272–279. ACM (2008)

[7] Gilpin, A., Pea, J., Sandholm, T.: First-order algorithm with $\mathcal{O}(\ln(1/\epsilon))$ convergence for ϵ-equilibrium in two-person zero-sum games. Mathematical Programming 133(1-2), 279–298 (2012)

[8] He, Y., Monteiro, R.D.: Accelerating block-decomposition first-order methods for solving generalized saddle-point and nash equilibrium problems. Optimization-online preprint (2013)

[9] Hoda, S., Gilpin, A., Pena, J., Sandholm, T.: Smoothing techniques for computing nash equilibria of sequential games. Mathematics of Operations Research 35(2), 494–512 (2010)

[10] Koller, D., Megiddo, N.: The complexity of two-person zero-sum games in extensive form. Games and economic behavior 4(4), 528–552 (1992)

[11] Koller, D., Megiddo, N., Von Stengel, B.: Fast algorithms for finding randomized strategies in game trees. In: Proceedings of the twenty-sixth annual ACM symposium on Theory of computing. pp. 750–759. ACM (1994)

[12] Kuhn, H.W.: A simplified two-person poker. Annals of Math (1950)

[13] Lanctot, M., Waugh, K., Zinkevich, M., Bowling, M.: Monte carlo sampling for regret minimization in extensive games. In: Advances in Neural Information Processing Systems. pp. 1078–1086 (2009)

[14] Nesterov, Y.: Excessive gap technique in nonsmooth convex minimization. SIAM Journal on Optimization 16(1), 235–249 (2005)

[15] Nesterov, Y.: Smooth minimization of non-smooth functions. Math. Program. Ser. A 103, 127152 (2005)

[16] von Stengel, B.: Equilibrium Computation for Two-player Games in Strategic and Extensive Form. N. Nisan, T. Roughgarden, E. Tardos, and V. Vazirani (eds.), Algorithmic Game Theory. Cambridge Univ. Press, Cambridge UK

[17] von Stengel, B.: Efficient computation of behavior strategies. Games and Economic Behavior 14(2), 220–246 (1996b)

[18] Zinkevich, M., Johanson, M., Bowling, M., Piccione, C.: Regret minimization in games with incomplete information. Advances in Neural Information Processing Systems 20, 1729–1736 (2008)