Extragradient with player sampling for faster Nash equilibrium finding

Samy Jelassi [*1] **Carles Domingo-Enrich** [*2] **Damien Scieur** [3] **Arthur Mensch** [4 2] **Joan Bruna** [2]

Abstract

Data-driven modeling increasingly requires to find a Nash equilibrium in multi-player games, e.g. when training GANs. In this paper, we analyse a new extra-gradient method for Nash equilibrium finding, that performs gradient extrapolations and updates on a random subset of players at each iteration. This approach provably exhibits a better rate of convergence than full extra-gradient for non-smooth convex games with noisy gradient oracle. We propose an additional variance reduction mechanism to obtain speed-ups in smooth convex games. Our approach makes extrapolation amenable to massive multiplayer settings, and brings empirical speed-ups, in particular when using a heuristic cyclic sampling scheme. Most importantly, it allows to train faster and better GANs and mixtures of GANs.

A growing number of models in machine learning require to optimize over multiple interacting objectives. This is the case of generative adversarial networks (Goodfellow et al., 2014), imaginative agents (Racanière et al., 2017), hierarchical reinforcement learning (Wayne & Abbott, 2014) and multi-agent reinforcement learning (Bu et al., 2008). Solving saddle-point problems (see e.g., Rockafellar, 1970), that is key in robust learning (Kim et al., 2006) and image reconstruction (Chambolle & Pock, 2011), also falls in this category. These examples can be cast as games where players are parametrized modules that compete or cooperate to minimize their own objective functions.

To define a principled solution to a multi-objective optimization problem, we may rely on the notion of Nash equilibrium (Nash, 1951). At a Nash equilibrium, no player can improve its objective by unilaterally changing its strategy. The theoretical section of this paper considers the class of *convex n-player games*, for which Nash equilibria exist (Rosen, 1965). Finding a Nash equilibrium in this setting is equivalent to solving a variational inequality problem (VI) with a monotone operator (Rosen, 1965; Harker & Pang, 1990). This VI can be solved using first-order methods, that are prevalent in single-objective optimization for machine learning. Stochastic gradient descent (the simplest first-order method) is indeed known to converge to local minima under mild conditions met by ML problems (Bottou & Bousquet, 2008). Yet, while gradient descent can be applied simultaneously to different objectives, it may fail in finding a Nash equilibrium in very simple settings (see e.g., Letcher et al., 2019; Gidel et al., 2019). Two alternative modifications of gradient descent are necessary to solve the VI (hence Nash) problem: *averaging* (Magnanti & Perakis, 1997; Nedić & Ozdaglar, 2009) or *extrapolation* with averaging. The later was introduced as the *extra-gradient* (EG) method by Korpelevich (1976)); it is faster (Nemirovski, 2004) and can handle noisy gradients (Juditsky et al., 2011). Extrapolation corresponds to an *opponent shaping* step: each player anticipates its opponents' next moves to update its strategy.

In n-player games, extra-gradient computes $2n$ single player gradients before performing a parameter update. Whether in massive or simple two-players games, this may be an inefficient update strategy: early gradient information, computed at the beginning of each iteration, could be used to perform eager updates or extrapolations, similar to how alternated update of each player would behave. Therefore, we introduce and analyse new extra-gradient algorithms that extrapolate and update random or carefully selected subsets of players at each iteration (Fig. 1).

- We review the extra-gradient algorithm for differentiable games and outline its shortcomings (§3.1). We propose a doubly-stochastic extra-gradient (DSEG) algorithm (§3.2) that updates the strategies of a subset of players, thus performing *player sampling*. DSEG performs faster but noisier updates than the original full extra-gradient method (full EG, (Juditsky et al., 2011)), that uses a (once) stochastic gradient oracle. We introduce a variance reduction method to attenuate the noise added by player sampling in smooth games.
- We derive convergence rates for DSEG in the convex setting (§4), as summarized in Table 1. Proofs strongly relies on the specific structure of the noise introduced by

[*]Equal contribution [1]Princeton University, USA [2]NYU CIMS, New York, USA [3]Samsung SAIT AI Lab, Montreal, Canada [4]ENS, DMA, Paris, France. Correspondence to: Samy Jelassi <sjelassi@princeton.edu>.

Proceedings of the 37^{th} International Conference on Machine Learning, Vienna, Austria, PMLR 119, 2020.

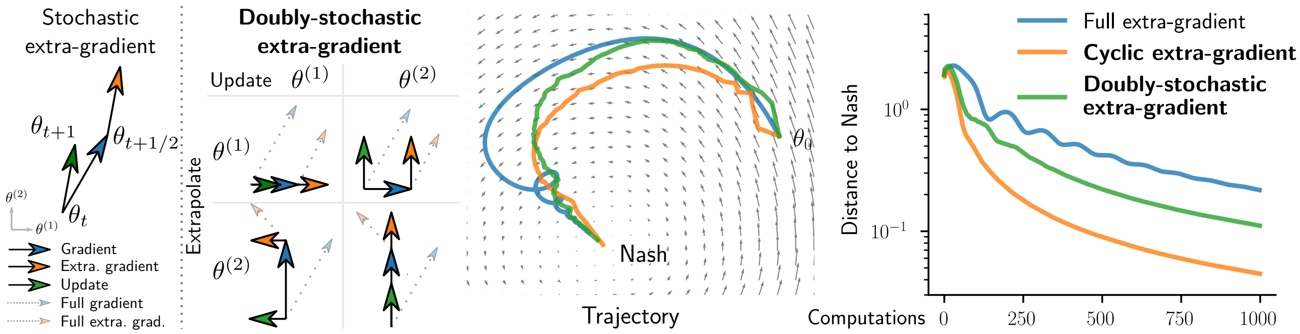


Figure 1. Left: We compute masked gradient during the extrapolation and update steps of the extra-gradient algorithm, to perform faster updates. *Right:* Optimization trajectories for doubly stochastic extra-gradient and full-update extra-gradient, on a convex single-parameter two-player convex game. Player sampling improves the expected rate of convergence toward the Nash equilibrium $(0, 0)$.

Table 1. New and existing (Juditsky et al., 2011) convergence rates for convex games, w.r.t. the number of gradient computations k. Doubly-stochastic extra-gradient (DSEG) *multiplies the noise contribution by a factor* $\alpha \triangleq \sqrt{b/n}$, where b is the number of sampled players among n. G bounds the gradient norm. L: Lip. constant of losses' gradient. σ^2 bounds the gradient estimation noise. Ω: diameter of the param. space. K: function of G, L, σ^2.

$\alpha \triangleq \sqrt{b/n}$	Non-smooth	Smooth
DSEG	$\mathcal{O}\left(n\sqrt{\frac{\Omega}{k}(G^2+\alpha^2\sigma^2)}\right)$	$\mathcal{O}\left(\alpha n \sigma \sqrt{\frac{\Omega}{k}} + \frac{K\Omega^{3/2}n^2}{\alpha^3 k^{3/2}}\right)$
Full EG	$\mathcal{O}\left(n\sqrt{\frac{\Omega}{k}(G^2+\sigma^2)}\right)$	$\mathcal{O}\left(n\sigma\sqrt{\frac{\Omega}{k}} + \frac{L\Omega n^{3/2}}{k}\right)$

player sampling. Our rates exhibit a better dependency on gradient noise compared to stochastic extra-gradient, and are thus interesting in the high-noise regime common in machine learning.

– Empirically, we first validate that DSEG is faster in massive differentiable convex games with noisy gradient oracles. We further show that non-random player selection improves convergence speed, and provide explanations for this phenomenon. In practical non-convex settings, we find that cyclic player sampling improves the speed and performance of GAN training (CIFAR10, ResNet architecture). The positive effects of extrapolation and alternation combine: DSEG should be used to train GANs, and even more to train *mixtures* of GANs.

2. Related work

Extra-gradient method. In this paper, we focus on finding the Nash equilibrium in convex n-player games, or equivalently the Variational Inequality problem (Harker & Pang, 1990; Nemirovski et al., 2010). This can be done using extrapolated gradient (Korpelevich, 1976), a "cautious" gradient descent approach that was promoted by Nemirovski (2004) and Nesterov (2007), under the name *mirror-prox*—we review this work in §3.1. Juditsky et al. (2011) propose a stochastic variant of mirror-prox, that assumes access to a noisy gradient oracle. In the convex setting, their results guarantees the convergence of the algorithm we propose, albeit with very slack rates. Our theoretical analysis refines these rates to show the usefulness of player sampling. Recently, Bach & Levy (2019) described a smoothness-adaptive variant of this algorithm similar to AdaGrad (Duchi et al., 2011), an approach that can be combined with ours. Yousefian et al. (2018) consider multi-agent games on networks and analyze a stochastic variant of extra-gradient that consists in randomly extrapolating and updating a single player. Compared to them, we analyse more general player sampling strategies. Moreover, our analysis holds for non-smooth losses, and provides better rates for smooth losses, through variance reduction. We also analyse precisely the reasons why player sampling is useful (see discussion in §4), an original endeavor.

Extra-gradient in non-convex settings. Extra-gradient has been applied in non-convex settings. Mertikopoulos et al. (2019) proves asymptotic convergence results for extra-gradient without averaging in a slightly non-convex case. Gidel et al. (2019) demonstrate the effectiveness of extra-gradient for GANs. They argue that it allows to escape the potentially chaotic behavior of simultaneous gradient updates (examplified by e.g. Cheung & Piliouras (2019)). Earlier work on GANs propose to replace simultaneous updates with alternated updates, with a comparable improvement (Gulrajani et al., 2017). In §5, we show that alternating player updates while performing opponent extrapolation improves the training speed and quality of GANs.

Opponent shaping and gradient adjustment. Extra-gradient can also be understood as an *opponent shaping* method: in the extrapolation step, the player looks one step in the future and anticipates the next moves of his opponents. Several recent works proposed algorithms that make use of the opponents' information to converge to an equilibrium

(Zhang & Lesser, 2010; Foerster et al., 2018; Letcher et al., 2019). In particular, the "Learning with opponent-learning awareness" (LOLA) algorithm is known for encouraging cooperation in cooperative games (Foerster et al., 2018). Lastly, some recent works proposed algorithms to modify the dynamics of simultaneous gradient descent by adding an adjustment term in order to converge to the Nash equilibrium (Mazumdar et al., 2019) and avoid oscillations (Balduzzi et al., 2018; Mescheder et al., 2017). One caveat of these works is that they need to estimate the Jacobian of the simultaneous gradient, which may be expensive in large-scale systems or even impossible when dealing with non-smooth losses as we consider in our setting. This is orthogonal to our approach that finds solutions of the original VI problem (4).

3. Solving convex games with partial first-order information

We review the framework of Cartesian convex games and the extra-gradient method in §3.1. Building on these, we propose to augment extra-gradient with player sampling and variance reduction in §3.2.

3.1. Solving convex games with gradients

In a game, each player observes a loss that depends on the independent parameters of all other players.

Definition 1. *A standard n-player game is given by a set of n players with parameters $\theta = (\theta^1, \dots, \theta^n) \in \Theta \subset \mathbb{R}^d$ where Θ decomposes into a Cartesian product $\prod_{i=1}^n \Theta^i$. Each player's parameter θ^i lives in $\Theta^i \subset \mathbb{R}^{d_i}$. Each player is given a loss function $\ell_i \colon \Theta \to \mathbb{R}$.*

For example, generative adversarial network (GAN) training is a standard game between a generator and discriminator that do not share parameters. We make the following assumption over the geometry of losses and constraints, that is the counterpart of the convexity assumption in single-objective optimization.

Assumption 1. *The parameter spaces $\Theta_1, \dots, \Theta_n$ are compact, convex and non-empty. Each player's loss $\ell_i(\theta^i, \theta^{-i})$ is convex in its parameter θ^i and concave in θ^{-i}, where θ^{-i} contains all other players' parameters. Moreover, $\sum_{i=1}^n \ell_i(\theta)$ is convex in θ.*

Ass. 1 implies that Θ has a diameter $\Omega \triangleq \max_{u,z\in\Theta} \|u - z\|_2$. Note that the losses may be non-differentiable. A simple example of Cartesian convex games satisfying Ass. 1, that we will empirically study in §5, are matrix games (e.g., rock-paper-scissors) defined by a positive payoff matrix $A \in \mathbb{R}^{d\times d}$, with parameters θ corresponding to n mixed strategies θ_i lying in the probability simplex $\triangle^{d_i}$.

Nash equilibria. Joint solutions to minimizing losses $(\ell_i)_i$ are naturally defined as the set of *Nash equilibria* (Nash, 1951) of the game. In this setting, we look for equilibria $\theta_\star \in \Theta$ such that

$$\forall\, i \in [n], \quad \ell_i(\theta^i_\star, \theta^{-i}_\star) = \min_{\theta^i \in \Theta^i} \ell_i(\theta^i, \theta^{-i}_\star). \tag{1}$$

A Nash equilibrium is a point where no player can benefit by changing his strategy while the other players keep theirs unchanged. Ass. 1 implies the existence of a Nash equilibrium (Rosen, 1965). We quantify the inaccuracy of a solution θ by the *functional Nash error*, also known as the Nikaidô & Isoda (1955) function:

$$\mathrm{Err}_N(\theta) \triangleq \sum_{i=1}^n \Big[\ell_i(\theta) - \min_{z\in\Theta_i} \ell_i(z, \theta^{-i})\Big]. \tag{2}$$

This error, computable through convex optimization, quantifies the gain that each player can obtain when deviating alone from the current strategy. In particular, $\mathrm{Err}_N(\theta) = 0$ if and only if θ is a Nash equilibrium; thus $\mathrm{Err}_N(\theta)$ constitutes a propose indication of convergence for sequence of iterates seeking a Nash equilibrium. We bound this value in our convergence analysis (see §4).

First-order methods and extrapolation. In convex games, as the losses ℓ_i are (sub)differentiable, we may solve (1) using first-order methods. We assume access to the *simultaneous gradient* of the game

$$F \triangleq (\nabla_1 \ell_1, \dots, \nabla_n \ell_n)^\top \in \mathbb{R}^d,$$

where we write $\nabla_i \ell_i \triangleq \nabla_{\theta^i} \ell_i$. It corresponds to the concatenation of the gradients of each player's loss with respect to its own parameters, and may be noisy. The losses ℓ_i may be non-smooth, in which case the gradients $\nabla_i \ell_i$ can be replaced by any subgradients. Simultaneous gradient descent, that explicitly discretizes the flow of the simultaneous gradient may converge slowly—e.g., in matrix games with skew-symmetric payoff and noiseless gradient oracle, convergence of the average iterate demands *decreasing* step-sizes. The extra-gradient method (Korpelevich, 1976) provides better guarantees (Nemirovski, 2004; Juditsky et al., 2011)—e.g., in the previous example, the step-size can remain constant. We build upon this method.

Extra-gradient consists in two steps: first, take a gradient step to go to an extrapolated point. Then use the gradient at the extrapolated point to perform a gradient step from the original point: at iteration τ,

$$\begin{aligned} &\text{(extrapolation)} \quad \theta_{\tau+1/2} = p_\Theta[\theta_\tau - \gamma_\tau F(\theta_\tau)], \\ &\text{(update)} \quad \theta_{\tau+1} = p_\Theta[\theta_\tau - \gamma_\tau F(\theta_{\tau+1/2})], \end{aligned} \tag{3}$$

where $p_\Theta[\cdot]$ is the Euclidean projection onto the constraint set Θ, i.e. $p_\Theta[z] = \mathrm{argmin}_{\theta\in\Theta} \|\theta - z\|_2^2$. This "cautious" approach allows to escape cycling orbits of the simultaneous gradient flow, that may arise around equilibrium points with skew-symmetric Hessians (see Fig. 1). The generalization of extra-gradient to general Banach spaces equipped by a Bregman divergence was introduced as the *mirror-prox* algorithm (Nemirovski, 2004). The new convergence results of §4 extend to the *mirror* setting (see §A.1). As recalled in Table 1, Juditsky et al. (2011) provide rates of convergence for the average iterate $\hat{\theta}_t = \frac{1}{t}\sum_{\tau=1}^t \theta_\tau$. Those rates are introduced for the equivalent variational inequality (VI) problem, finding

$$\theta_\star \in \Theta \text{ such that } F(\theta_\star)^\top(\theta - \theta_\star) \geqslant 0 \;\forall\, \theta \in \Theta, \tag{4}$$

where Ass. 1 ensures that the simultaneous gradient F is a monotone operator (see §A.2 for a review).

3.2. DSEG: Partial extrapolation and update for extra-gradient

The proposed algorithms are theoretically analyzed in the convex setting §4, and empirically validated in convex and non-convex setting in §5.

Caveats of extra-gradient. In systems with large number of players, an extra-gradient step may be computationally expensive due to the high number of backward passes necessary for gradient computations. Namely, at each iteration, we are required to compute $2n$ gradients before performing a first update. This is likely to be inefficient, as we could use the first computed gradients to perform a first extrapolation or update. This remains true for games down to two players. In a different setting, stochastic gradient descent (Robbins & Monro, 1951) updates model parameters before observing the whole data, assuming that partial observation is sufficient for progress in the optimization loop. Similarly, in our setting, partial gradient observation should be sufficient to perform extrapolation and updates toward the Nash equilibrium.

Player sampling. While standard extra-gradient performs at each iteration two passes of player's gradient computation, we therefore compute *doubly-stochastic simultaneous gradient estimates*, where only the gradients of a random subset of players are evaluated. This corresponds to evaluating a simultaneous gradient that is affected by *two* sources of noise. We sample a mini-batch $\mathcal{P}$ of players of size $b \leqslant n$, and compute the gradients for this mini-batch only. Furthermore, we assume that the gradients are noisy estimates, e.g., with noise coming from data sampling. We then compute a doubly-stochastic simultaneous gradient estimate $\tilde{F}$ as $\tilde{F} \triangleq (\tilde{F}^{(1)}, \dots, \tilde{F}^{(n)})^\top \in \mathbb{R}^d$ where

$$\tilde{F}^{(i)}(\theta, \mathcal{P}) \triangleq \begin{cases} \frac{n}{b} \cdot g_i(\theta) & \text{if } i \in \mathcal{P} \\ 0_{d_i} & \text{otherwise} \end{cases}, \tag{5}$$

Algorithm 1 Doubly-stochastic extra-gradient.

Input: initial point $\theta_0 \in \mathbb{R}^d$, stepsizes $(\gamma_\tau)_{\tau\in[t]}$, mini-batch size over the players $b \in [n]$.
With variance reduction (VR), $R \leftarrow \tilde{F}(\theta_0, [1, n])$ as in (5), i.e. the full simultaneous gradient.
for $\tau = 0, \dots, t$ **do**
  Sample mini-batches of players $\mathcal{P}$, $\mathcal{P}'$.
  Compute $\tilde{F}_{\tau+\frac{1}{2}} = \tilde{F}(\theta_\tau, \mathcal{P})$ using (5) or VR (Alg. 2).
  Extrapolation step: $\theta_{\tau+\frac{1}{2}} \leftarrow p_\Theta[\theta_\tau - \gamma_\tau \tilde{F}_{\tau+\frac{1}{2}}]$.
  Compute $\tilde{F}_{\tau+1} = \tilde{F}(\theta_{\tau+\frac{1}{2}}, \mathcal{P}')$ using (5) or VR
  Gradient step: $\theta_{\tau+1} \leftarrow p_\Theta[\theta_\tau - \gamma_\tau \tilde{F}_{\tau+1}]$.
Return $\hat{\theta}_t = [\sum_{\tau=0}^t \gamma_\tau]^{-1} \sum_{\tau=0}^t \gamma_\tau \theta_\tau$.

Algorithm 2 Variance reduced estimate of the simultaneous gradient with doubly-stochastic sampling

Input: point $\theta \in \mathbb{R}^d$, mini-batch $\mathcal{P}$, table of previous gradient estimates $R \in \mathbb{R}^d$.
Compute $\tilde{F}(\theta, \mathcal{P})$ as specified in equation (5).
for $i \in \mathcal{P}$ **do**
  Compute $\bar{F}^{(i)} \leftarrow \tilde{F}^{(i)}(\theta) + (1 - \frac{n}{b}) R^{(i)}$
  Update $R^{(i)} \leftarrow \frac{b}{n} \tilde{F}^{(i)}(\theta) = g_i(\theta)$
For $i \notin \mathcal{P}$, set $\bar{F}^{(i)} \leftarrow R^{(i)}$.
Return estimate $\bar{F} = (\bar{F}^{(1)}, \dots, \bar{F}^{(n)})$, table R.

and $g_i(\theta)$ is a noisy unbiased estimate of $\nabla_i \ell_i(\theta)$. The factor n/b in (5) ensures that the doubly-stochastic simultaneous gradient estimate is an unbiased estimator of the simultaneous gradient. Doubly-stochastic extra-gradient (DSEG) replaces the full gradients in the update (3) by the oracle (5), as detailed in Alg. 1.

Variance reduction for player noise. To obtain faster rates in convex games with smooth losses, we propose to compute a variance-reduced estimate of the gradient oracle (5). This mitigates the noise due to player sampling. Variance reduction is a technique known to accelerate convergence under smoothness assumptions in similar settings. While Palaniappan & Bach (2016); Iusem et al. (2017); Chavdarova et al. (2019) apply variance reduction on the noise coming from the gradient estimates, we apply it to the noise coming from the sampling over the players. We implement this idea in Alg. 2. We keep an estimate of $\nabla_i \ell_i$ for each player in a table R, which we use to compute *unbiased* gradient estimates with lower variance, akin to the approach of SAGA (Defazio et al., 2014) to reduce the variance of data noise.

Player sampling strategies. For convergence guarantees to hold, each player must have an equal probability of being sampled (*equiprobable player sampling condition*). Sampling uniformly over b-subsets of $[n]$ is a reasonable way to fulfill this condition as all players have probability $p = b/n$ of being chosen.

As a strategy to accelerate convergence, we propose to cycle over the $n(n-1)$ pairs of different players (with $b = 1$). At each iteration, we extrapolate the first player of the pair and update the second one. We shuffle the order of pairs once the block has been entirely seen. This scheme bridges extrapolation and alternated gradient descent: for GANs, it corresponds to extrapolate the generator before updating the discriminator, and vice-versa, cyclically. Although its convergence is not guaranteed, cyclic sampling over players is powerful for convex quadratic games (§5.1) and GANs (§5.2).

4. Convergence for convex games

We derive new rates for DSEG with random player sampling, improving the analysis of Juditsky et al. (2011). Player sampling can be seen as an extra source of noise in the gradient oracle. Hence the results of Juditsky et al. on stochastic extra-gradient guarantees the convergence of DSEG, as we detail in Corollary 1. Unfortunately, the convergence rates in this corollary do not predict any improvement of DSEG over full extra-gradient. Our main theoretical contribution is therefore a refinement of these rates for player-sampling noise. Improvements are obtained both for non-smooth and smooth losses, the latter using the proposed variance reduction approach. Our results predict better performance for DSEG in the high-noise regime. Results are stated here in Euclidean spaces for simplicity; they are proven in the more general mirror setting in App. B. In the analysis, we separately consider the two following assumptions on the losses.

Assumption 2a (Non-smoothness)**.** *For each $i \in [n]$, the loss ℓ_i has a bounded subgradient, namely $\max_{h \in \partial_i \ell_i(\theta)} \|h\|_2 \leqslant G_i$ for all $\theta \in \Theta$. In this case, we also define the quantity $G = \sqrt{\sum_{i=1}^n G_i^2/n}$.*

Assumption 2b (Smoothness)**.** *For each $i \in [n]$, the loss ℓ_i is once-differentiable and L-smooth, i.e. $\|\nabla_i \ell_i(\theta) - \nabla_i \ell_i(\theta')\|_2 \leqslant L\|\theta - \theta'\|_2$, for $\theta, \theta' \in \Theta$.*

Similar to Juditsky et al. (2011); Robbins & Monro (1951), we assume unbiasedness of the gradient estimate and boundedness of the variance.

Assumption 3. *For each player i, the noisy gradient g_i is unbiased and has bounded variance:*

$$\begin{aligned} \forall\, \theta \in \Theta, \quad & \mathbb{E}[g_i(\theta)] = \nabla_i \ell_i(\theta), \\ & \mathbb{E}[\|g_i(\theta) - \nabla_i \ell_i(\theta)\|_2^2] \leqslant \sigma^2. \end{aligned} \tag{6}$$

To compare DSEG to simple stochastic EG, we must take into account the cost of a single iteration, that we assume proportional to the number b of gradients to estimate at each step. We therefore set $k \triangleq 2\,b\,t$ to be the number of gradients estimates computed up to iteration t, and re-index the sequence of iterate $(\hat{\theta}_t)_{t \in \mathbb{N}}$ as $(\hat{\theta}_k))_{k \in 2b\mathbb{N}}$. We give rates with respect to k in the following propositions.

4.1. Slack rates derived from Juditsky et al.

Let us first recall the rates obtained by Juditsky et al. (2011) with noisy gradients but no player sampling.

Theorem 1 (Adapted from Juditsky et al. (2011))**.** *We consider a convex n-player game where Ass. 2a and Ass. 3 hold. We run Alg. 1 for t iterations without player sampling, thus performing $k = 2\,n\,t$ gradient evaluations. With optimal constant stepsize, the expected Nash error verifies*

$$\mathbb{E}\left[Err_N(\hat{\theta}_k)\right] \leqslant 14n\sqrt{\frac{\Omega}{3k}\,(G^2 + 2\sigma^2)}. \tag{7}$$

Assuming smoothness (Ass. 2b) and optimal stepsize,

$$\mathbb{E}\left[Err_N(\hat{\theta}_k)\right] \leqslant \max\left\{\frac{7\Omega L n^{3/2}}{k}, 14n\sqrt{\frac{2\Omega\sigma^2}{3k}}\right\}. \tag{8}$$

Player sampling fits within the framework of noisy gradient oracle (6), replacing the gradient estimates $(g_i)_{i \in [n]}$ with the estimates $(\tilde{F}^{(i)})_{i \in [n]}$ from (5), and updating the variance σ^2 accordingly. We thus derive the following corollary.

Corollary 1. *We consider a convex n-player game where Ass. 2a and Ass. 3 hold. We run Alg. 1 for t iterations with equiprobable player sampling, thus performing $k = 2\,b\,t$ gradient evaluations. With optimal constant stepsize, the expected Nash error verifies*

$$\mathbb{E}\left[\mathrm{Err}_N(\hat{\theta}_k)\right] \leqslant \mathcal{O}\left(n\sqrt{\frac{\Omega}{k}\left(\frac{n}{b}G^2 + \sigma^2\right)}\right).$$

Assuming smoothness (Ass. 2b) and optimal stepsize,

$$\mathbb{E}\left[Err_N(\hat{\theta}_k)\right] \leqslant \mathcal{O}\left(\frac{\Omega L n^{3/2}}{k} + n\sqrt{\frac{\Omega}{k}(\frac{n}{b}L^2\Omega^2 + \sigma^2)}\right).$$

The proof is in §B.1. The notation $\mathcal{O}(\cdot)$ hides numerical constants. Whether in the smooth or non-smooth case, the upper-bounds from Corollary 1 do not predict any improvement due to player sampling, as the factor before the gradient size G or $L\Omega$ is increased, and the factor before the noise variance σ remains constant.

4.2. Tighter rates using noise structure

Fortunately, a more cautious analysis allows to improve these bounds, by taking into account the noise structure induced by sampling in (5). We provide a new result in the non-smooth case, proven in §B.3.

Theorem 2. *We consider a convex n-player game where Ass. 2a and Ass. 3 hold. We run Alg. 1 for t iterations with equiprobable player sampling, thus performing $k = 2\,b\,t$ gradient evaluations. With optimal constant stepsize, the expected Nash error verifies*

$$\mathbb{E}\left[Err_N(\hat{\theta}_k)\right] \leqslant \mathcal{O}\left(n\sqrt{\frac{\Omega}{k}\left(G^2 + \frac{b}{n}\sigma^2\right)}\right). \quad (9)$$

Compared to Corollary 1, we obtain a factor $\sqrt{\frac{b}{n}}$ in front of the noise term $\frac{\sigma}{\sqrt{k}}$, without changing the constant before the gradient size G. We can thus expect faster convergence with noisy gradients. (9) is tightest when sampling a single player, i.e. when $b = 1$.

A similar improvement can be obtained with smooth losses thanks to the variance reduction technique proposed in Alg. 2. This is made clear in the following result, proven in §B.4.

Theorem 3. *We consider a convex n-player game where Ass. 2a, Ass. 2b and Ass. 3 hold. We run Alg. 1 for t iterations with equiprobable player sampling (Alg. 2), thus performing $k = 2\,b\,t$ gradient evaluations. With optimal constant stepsize, the expected Nash $\mathbb{E}\left[Err_N(\hat{\theta}_k)\right]$ error is upper-bounded by*

$$\mathcal{O}\left(\sqrt{\frac{b}{n}}n\sqrt{\frac{\Omega\sigma^2}{k}} + L^2(G^2+\sigma^2)n^2\left(\frac{n\Omega}{b\sigma^2 k}\right)^{3/2}\right) \quad (10)$$

The upper-bound (10) should be compared with the bound of full extra-gradient (8). With player sampling, the constant before the noise term (first term) is smaller by a factor $\sqrt{\frac{n}{b}}$. On the other hand, the gradient-size term scales as factor $(\frac{n}{b})^{3/2}$. Since the noise term is $\mathcal{O}(\frac{1}{\sqrt{k}})$ and the gradient size term is $\mathcal{O}(\frac{1}{k^{3/2}})$, for large k choosing low b is better. For $k \to \infty$, the bound (10) is once again tightest by sampling a random single player.

To sum up, doubly-stochastic extra-gradient convergence is controlled with a better rate than stochastic extra-gradient (EG) with non-smooth losses; with smooth losses, DSEG exhibits a better dependency on the noise but worse dependency on the gradient size. In the high noise regime, asymptotically DSEG brings the same improvement of a factor $\sqrt{\frac{b}{n}}$ before the constant $\frac{\sigma}{\sqrt{k}}$, for both smooth and non-smooth problems.

Step-sizes. The stepsizes of the previous propositions are assumed to be constant and are optimized knowing the geometry of the problem. They are explicit in App. B. As in full extra-gradient, convergence can be guaranteed without such knowledge using decreasing step-sizes. In experiments, we perform a grid-search over stepsizes to obtain the best results given a computational budget k.

5. Convex and non-convex applications

We show the performance of doubly-stochastic extra-gradient in the setting of quadratic games, comparing different sampling schemes. We assess the speed and final performance of DSEG in the practical context of GAN training. A *PyTorch/Numpy* package is attached.

5.1. Random convex quadratic games

We consider a game where n players can play d actions, with payoffs provided by a matrix $A \in \mathbb{R}^{nd\times nd}$, an horizontal stack of matrices $A_i \in \mathbb{R}^{(d\times nd)}$ (one for each player). The loss function ℓ_i of each player is defined as its expected payoff given the n mixed strategies $(\theta^1, \dots, \theta^n)$, i.e. $\forall\, i \in [n],\ \forall\, \theta \in \Theta = \triangle^{d_1} \times \dots \times \triangle^{d_n}$,

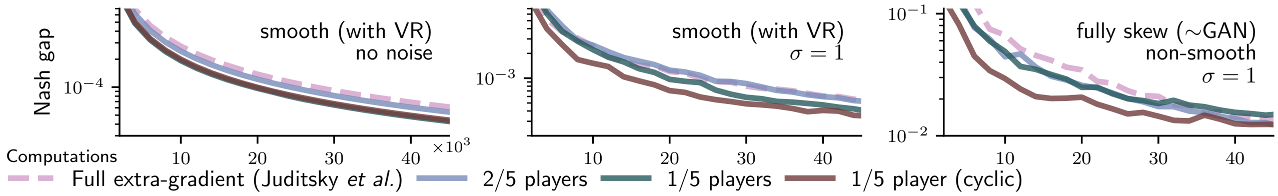
$$\ell_i(\theta^i, \theta_{-i}) = {\theta^i}^\top A_i \theta + \lambda \|\theta^i - \frac{1}{d_i}\|_1,$$

where λ is a regularization parameter that introduces non-smoothness and pushes strategies to snap to the simplex center. The positivity of A, i.e. $\theta^\top A\theta \geqslant 0$ for all $\theta \in \Theta$, is equivalent to the convexity of the game.

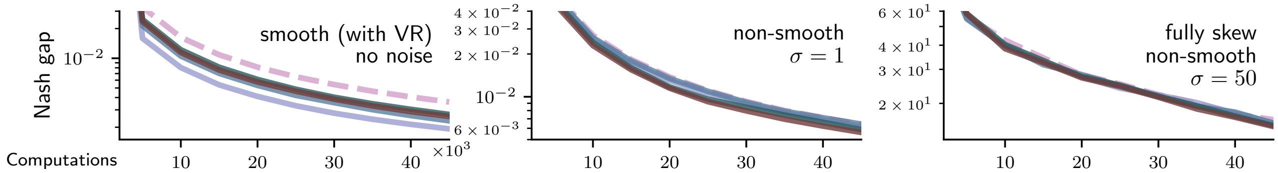
Experiments. We sample A as the weighted sum of a random symmetric positive definite matrix and a skew matrix. We compare the convergence speeds of extra-gradient algorithms, with or without player sampling. We vary three parameters: the variance σ of the noise in the gradient oracle (we add a Gaussian noise on each gradient coordinate), the non-smoothness λ of the loss, and the skewness of the matrix. We consider small games and large games ($n \in \{5, 50\}$). We use the (simplex-adapted) mirror variant of doubly-stochastic extra-gradient, and a constant stepsize, selected among a grid (see App. D). We use variance reduction when $\lambda = 0$ (smooth case). We also consider cyclic sampling in our benchmarks, as described in §3.2.

Results. Fig. 2 compares the convergence speed of player-sampled extra-gradient for the various settings and sampling schemes. As predicted by Theorem 2 and 3, the regime of convergence in $1/\sqrt{k}$ in the presence of noise is unchanged with player sampling. DSEG always brings a benefit in the convergence constants (Fig. 2a-b), in particular for smooth noisy problems (Fig. 2a center, Fig. 2b left). Most interestingly, cyclic player selection improves upon random sampling for small number of players (Fig. 2a).

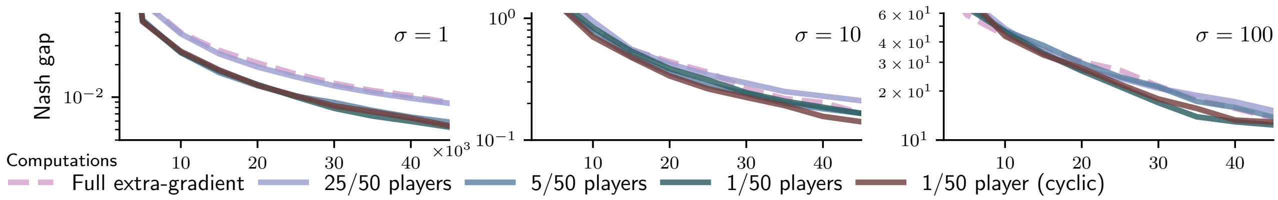
Fig. 2c highlights the trade-offs in Theorem 3: as the noise increase, the size of player batches should be reduced. Not that for skew-games with many players (Fig. 2b col. 3), our approach only becomes beneficial in the high-noise regime.



(a) 5-player games with increasingly hard configurations.



(b) 50-player games with increasingly hard configurations.



(c) 50-player smooth game with increasing noise (sampling with variance reduction).

Figure 2. Player sampled extra-gradient outperform vanilla extra-gradient for small noisy/non-noisy smooth/non-smooth games. Cyclic sampling performs better than random sampling, especially for 5 players (a). Higher sampling ratio is beneficial in high noise regime (c), Curves averaged over 5 games and 5 runs.

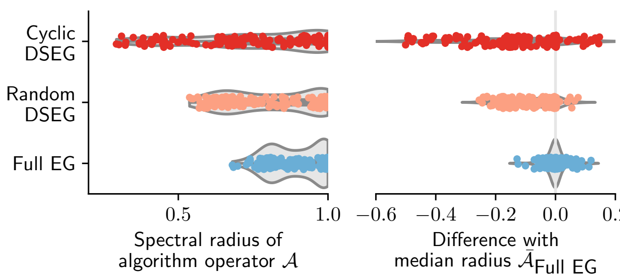


Figure 3. Left: Spectral radii of operators for random 2-player matrix games. *Right:* each radius is compared to the median radius obtained for full extra-gradient, within each category of skewness and conditioning of random payoff matrices. Cyclic sampling lowers spectral radii and improve convergence rates.

As predicted in §4, full EG should be favored with noiseless oracles (see App. D).

Spectral study of sampling schemes. The benefit of cyclic sampling can be explained for simple quadratic games. We consider a two-player quadratic game where $\ell_i(\theta) = {\theta^i}^\top A \theta$ for $i = 1, 2$, $\theta = (\theta^1, \theta^2)$ is an unconstrained vector of $\mathbb{R}^{2 \times d}$, and gradients are noiseless. In this setting, full EG and DSEG expected iterates follows a linear recursion $\mathbb{E}[\theta_{k+4}] = \mathcal{A}(\mathbb{E}[\theta_k])$, where k is the number of gradient evaluation and $\mathcal{A}$ is a linear "algorithm operator", computable in closed form. A lower spectral radius for $\mathcal{A}$ yields a better convergence rate for $(\mathbb{E}[\theta_k])_k$, in light of Gelfand (1941) formula—we compare spectral radii across methods.

We sample random payoff matrices A of varying skewness and condition number, and compare the spectral radius $\mathcal{A}$ associated to full EG, and DSEG with cyclic and random player selection. As summarized in Fig. 3, player sampling reduces the spectral radius of $\mathcal{A}$ on average; most interestingly, the reduction is more important using cyclic sampling. Spectral radii are not always in the same order across methods, hinting that sampling can be harmful in the worst cases. Yet cyclic sampling will perform best on average in this (simple) setting. We report details and further figures in App. C.

5.2. Generative adversarial networks (GANs)

We evaluate the performance of the player sampling approach to train a generative model on CIFAR10 (Krizhevsky & Hinton, 2009). We use the WGAN-GP loss (Gulrajani et al., 2017), that defines a non-convex two-player game. Our theoretical analysis indeed shows a $1/\sqrt{2}$ speed-up for noisy monotonous 2-player games—the following suggests that speed-up also arises in a non-convex setting. We compare the full stochastic extra-gradient (SEG) approach advocated by Gidel et al. (2019) to the cyclic sampling scheme proposed in §3.2 (i.e. *extra. D, upd. G, extra. G, upd. D*). We use the ResNet (He et al., 2016) architecture from Gidel

Figure 4. Training curves and samples using doubly-stochastic extragradient on CIFAR10 with WGAN-GP losses, for the best learning rates. Doubly-stochastic extrapolation allows faster and better training, most notably in term of Fréchet Inception Distance (10k). Curves averaged over 5 runs.

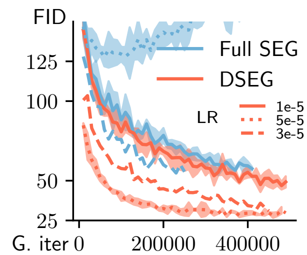
et al. (2019), and select the best performing stepsizes among a grid (see App. D). We use the Adam (Kingma & Ba, 2015) refinement of extra-gradient (Gidel et al., 2019) for both the baseline and proposed methods. The notion of functional Nash error does not exist in the non-convex setting. We estimate the convergence speed toward an equilibrium by measuring a quality criterion for the generator. We therefore evaluate the Inception Score (Salimans et al., 2016) and Fréchet Inception Distance (FID, Heusel et al. (2017) along training, and report their final values.

Results. We report training curves versus wall-clock time in Fig. 4. Cyclic sampling allows faster and better training, especially with respect to FID, which is more correlated to human appreciation (Heusel et al., 2017). Fig. 5 (right) compares our result to full extra-gradient with uniform averaging. It shows substantial improvements in FID, with results less sensitive to randomness. SEG itself slightly outperforms optimistic mirror descent (Gidel et al., 2019; Mertikopoulos et al., 2019).

Interpretation. Without extrapolation, alternated training is known to perform better than simultaneous updates in WGAN-GP (Gulrajani et al., 2017). Full extrapolation has been shown to perform similarly to alternated updates (Gidel et al., 2019). Our approach combine extrapolation with an alternated schedule. It thus performs better than extrapolating with simultaneous updates. It remains true across every learning rate we tested. Echoing our findings of §5.1, deterministic sampling is crucial for performance, as random player selection performs poorly (score 6.2 IS).

5.3. Mixtures of GANs

Finally, we consider a simple multi-player GAN setting, akin to Ghosh et al. (2018), where n different generators $(g_{\theta_i})_i$ seeks to fool m different discriminators $(f_{\varphi_j})_j$. We minimize $\sum_j \ell(g_{\theta_i}, f_{\varphi_j})$ for all i, and maximize $\sum_i \ell(g_{\theta_i}, f_{\varphi_j})$ for all j. Fake data is then sampled from mixture $\sum_{i=1}^n \delta_{i=J} g_{\theta_i}(\varepsilon)$, where J is sampled uniformly in $[n]$ and $\varepsilon \sim \mathcal{N}(0, I)$. We compare two methods: (i) SEG extrapolates and updates all $(g_{\theta_i})_i$, $(f_{\varphi_j})_j$ at the same time; (ii) DSEG extrapolates and updates successive pairs $(g_{\theta_j}, f_{\varphi_j})$ alternating the 4-step updates from §5.2.



Method	FID (50k)
SEG	19.69 ± 1.53
DSEG	$\mathbf{17.10 \pm 1.07}$
	IS
SEG	$8.26 \pm .16$
DSEG	$\mathbf{8.38 \pm .06}$

Figure 5. Left: Player sampling allows faster training of mixtures of GANs. *Right*: Player sampling trains better ResNet WGAN-GP. FID and IS computed on 50k samples, averaged over 5 runs.

Results. We compare the training curves of both SEG and DSEG in Fig. 5, for a range of learning rates. DSEG outperform SEG for all learning rates; more importantly, higher learning rates can be used for DSEG, allowing for faster training. DSEG is thus appealing for mixtures of GANs, that are useful to mitigate mode collapse in generative modeling. We report generated images in Appendix D.

6. Conclusion

We propose and analyse a doubly-stochastic extra-gradient approach for finding Nash equilibria. According to our convergence results, updating and extrapolating random sets of players in extra-gradient brings speed-up in noisy and non-smooth convex problems. Numerically, doubly-stochastic extra-gradient indeed brings speed-ups in convex settings, especially with noisy gradients. It brings speed-ups and *improve solutions* when training non-convex GANs and mixtures of GANs, thus combining the benefits of alternation and extrapolation in adversarial training. Numerical experiments show the importance of *sampling schemes*. We take a first step towards understanding the good behavior of *cyclic* player sampling through spectral analysis. We foresee interesting developments using player sampling in reinforcement learning: the policy gradients obtained using multi-agent actor critic methods (Lowe et al., 2017) are noisy estimates, a setting in which it is beneficial.

References

Bach, F. and Levy, K. A universal algorithm for variational inequalities adaptive to smoothness and noise. In *Proceedings of the Conference on Learning Theory*, 2019.

Balduzzi, D., Racanière, S., Martens, J., Foerster, J., Tuyls, K., and Graepel, T. The mechanics of n-player differentiable games. In *Proceedings of the International Conference on Machine Learning*, 2018.

Bottou, L. and Bousquet, O. The tradeoffs of large scale learning. In *Advances in Neural Information Processing Systems*, pp. 161–168, 2008.

Bu, L., Babu, R., De Schutter, B., et al. A comprehensive survey of multi-agent reinforcement learning. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 38(2):156–172, 2008.

Bubeck, S. Convex optimization: Algorithms and complexity. *Foundations and Trends in Machine Learning*, 8(3-4): 231–357, 2015.

Chambolle, A. and Pock, T. A first-order primal-dual algorithm for convex problems with applications to imaging. *Journal of Mathematical Imaging and Vision*, 40(1):120–145, 2011.

Chavdarova, T., Gidel, G., Fleuret, F., Foo, C.-S., and Lacoste-Julien, S. Reducing noise in GAN training with variance reduced extragradient. In *Advances in Neural Information Processing Systems*, 2019.

Cheung, Y. K. and Piliouras, G. Vortices Instead of Equilibria in MinMax Optimization: Chaos and Butterfly Effects of Online Learning in Zero-Sum Games. In *Proceedings of the Conference on Learning Theory*, 2019.

Defazio, A., Bach, F., and Lacoste-Julien, S. SAGA: A fast incremental gradient method with support for non-strongly convex composite objectives. In *Advances in Neural Information Processing Systems*, pp. 1646–1654, 2014.

Duchi, J., Hazan, E., and Singer, Y. Adaptive subgradient methods for online learning and stochastic optimization. *Journal of Machine Learning Research*, 12:2121–2159, 2011.

Foerster, J., Chen, R. Y., Al-Shedivat, M., Whiteson, S., Abbeel, P., and Mordatch, I. Learning with opponent-learning awareness. In *Proceedings of the International Conference on Autonomous Agents and Multi-Agent Systems*, 2018.

Gelfand, I. Normierte ringe. *Matematicheskii Sbornik*, 9(1): 3–24, 1941.

Ghosh, A., Kulharia, V., Namboodiri, V., Torr, P. H. S., and Dokania, P. K. Multi-agent diverse generative adversarial networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018.

Gidel, G., Berard, H., Vignoud, G., Vincent, P., and Lacoste-Julien, S. A variational inequality perspective on generative adversarial networks. In *International Conference on Learning Representations*, 2019.

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., and Bengio, Y. Generative adversarial nets. In *Advances in Neural Information Processing Systems*, pp. 2672–2680, 2014.

Gulrajani, I., Ahmed, F., Arjovsky, M., Dumoulin, V., and Courville, A. C. Improved training of Wasserstein GANs. In *Advances in Neural Information Processing Systems*, pp. 5767–5777, 2017.

Harker, P. T. and Pang, J.-S. Finite-dimensional variational inequality and nonlinear complementarity problems: a survey of theory, algorithms and applications. *Mathematical Programming*, 48(1-3):161–220, 1990.

He, K., Zhang, X., Ren, S., and Sun, J. Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770–778, 2016.

Heusel, M., Ramsauer, H., Unterthiner, T., Nessler, B., and Hochreiter, S. GANs trained by a two time-scale update rule converge to a local Nash equilibrium. In *Advances in Neural Information Processing Systems*, pp. 6626–6637, 2017.

Iusem, A., Jofré, A., Oliveira, R. I., and Thompson, P. Extragradient method with variance reduction for stochastic variational inequalities. *SIAM Journal on Optimization*, 27(2):686–724, 2017.

Juditsky, A., Nemirovski, A., and Tauvel, C. Solving variational inequalities with stochastic mirror-prox algorithm. *Stochastic Systems*, 1(1):17–58, 2011.

Kim, S.-J., Magnani, A., and Boyd, S. Robust Fisher discriminant analysis. In *Advances in Neural Information Processing Systems*, 2006.

Kingma, D. P. and Ba, J. Adam: A Method for Stochastic Optimization. In *International Conference on Learning Representations*, 2015.

Korpelevich, G. The extragradient method for finding saddle points and other problems. *Matecon*, 12:747–756, 1976.

Krizhevsky, A. and Hinton, G. Learning multiple layers of features from tiny images. Technical report, Citeseer, 2009.

Letcher, A., Foerster, J., Balduzzi, D., Rocktäschel, T., and Whiteson, S. Stable opponent shaping in differentiable games. In *International Conference on Learning Representations*, 2019.

Lowe, R., Wu, Y., Tamar, A., Harb, J., Abbeel, O. P., and Mordatch, I. Multi-agent actor-critic for mixed cooperative-competitive environments. In *Advances in Neural Information Processing Systems*, pp. 6379–6390, 2017.

Magnanti, T. L. and Perakis, G. Averaging schemes for variational inequalities and systems of equations. *Mathematics of Operations Research*, 22(3):568–587, 1997.

Mazumdar, E. V., Jordan, M. I., and Sastry, S. S. On finding local Nash equilibria (and only local Nash equilibria) in zero-sum games. *arXiv:1901.00838*, 2019.

Mertikopoulos, P., Lecouat, B., Zenati, H., Foo, C.-S., Chandrasekhar, V., and Piliouras, G. Optimistic mirror descent in saddle-point problems: Going the extra (gradient) mile. In *International Conference on Learning Representations*, 2019.

Mescheder, L., Nowozin, S., and Geiger, A. The numerics of GANs. In *Advances in Neural Information Processing Systems*, pp. 1825–1835, 2017.

Nash, J. Non-cooperative games. *Annals of Mathematics*, pp. 286–295, 1951.

Nedić, A. and Ozdaglar, A. Subgradient methods for saddle-point problems. *Journal of Optimization Theory and Applications*, 142(1):205–228, 2009.

Nemirovski, A. Prox-method with rate of convergence $o(1/t)$ for variational inequalities with lipschitz continuous monotone operators and smooth convex-concave saddle point problems. *SIAM Journal on Optimization*, 15(1):229–251, 2004.

Nemirovski, A., Onn, S., and Rothblum, U. G. Accuracy certificates for computational problems with convex structure. *Mathematics of Operations Research*, 35(1):52–78, 2010.

Nemirovsky, A. S. and Yudin, D. B. *Problem complexity and method efficiency in optimization.* Wiley, 1983.

Nesterov, Y. Dual extrapolation and its applications to solving variational inequalities and related problems. *Mathematical Programming*, 109(2-3):319–344, 2007.

Nikaidô, H. and Isoda, K. Note on non-cooperative convex games. *Pacific Journal of Mathematics*, 5(Suppl. 1):807–815, 1955.

Palaniappan, B. and Bach, F. Stochastic variance reduction methods for saddle-point problems. In *Advances in Neural Information Processing Systems*, pp. 1416–1424, 2016.

Racanière, S., Weber, T., Reichert, D., Buesing, L., Guez, A., Rezende, D. J., Badia, A. P., Vinyals, O., Heess, N., Li, Y., et al. Imagination-augmented agents for deep reinforcement learning. In *Advances in Neural Information Processing Systems*, pp. 5690–5701, 2017.

Robbins, H. and Monro, S. A stochastic approximation method. *The Annals of Mathematical Statistics*, 22(3): 400–407, 1951.

Rockafellar, R. T. Monotone operators associated with saddle-functions and minimax problems. In *Proceedings of Symposia in Pure Mathematics*, volume 18.1, pp. 241–250, 1970.

Rosen, J. B. Existence and uniqueness of equilibrium points for concave n-person games. *Econometrica*, 33(3):520–534, 1965.

Salimans, T., Goodfellow, I., Zaremba, W., Cheung, V., Radford, A., Chen, X., and Chen, X. Improved techniques for training GANs. In *Advances in Neural Information Processing Systems*, pp. 2234–2242, 2016.

Wayne, G. and Abbott, L. Hierarchical control using networks trained with higher-level forward models. *Neural Computation*, 26(10):2163–2193, 2014.

Yousefian, F., Nedić, A., and Shanbhag, U. V. On stochastic mirror-prox algorithms for stochastic cartesian variational inequalities: Randomized block coordinate and optimal averaging schemes. *Set-Valued and Variational Analysis*, 26(4):789–819, 2018.

Zhang, C. and Lesser, V. Multi-agent learning with policy prediction. In *Proceedings of the AAAI Conference on Artificial Intelligence*, 2010.