Monotone Near-Zero-Sum Games

Ruichen Luo[1] Sebastian U. Stich[2] Krishnendu Chatterjee[1]

[1]Institute of Science and Technology Austria (ISTA)
3400 Klosterneuburg, Austria

[2]CISPA Helmholtz Center for Information Security
66386 St. Ingbert, Germany

Oct 10, 2025 in St. Ingbert, Germany

Outline for section 1

NEP and VIP

Two-Person Nash Equilibrium Problem (NEP):

- Player 1: $\max_{\mathbf{x}\in X} u_1(\mathbf{x},\mathbf{y})$; Player 2: $\max_{\mathbf{y}\in Y} u_2(\mathbf{x},\mathbf{y})$.
- X and Y are compact and convex; u_1 and u_2 are L-smooth on $X \times Y$.
- Nash equilibrium $(\mathbf{x}^*,\mathbf{y}^*) \in X \times Y$ s.t.

$$u_1(\mathbf{x}^*,\mathbf{y}^*) \geq u_1(\mathbf{x},\mathbf{y}^*), \text{ for all } \mathbf{x} \in X,$$
$$u_2(\mathbf{x}^*,\mathbf{y}^*) \geq u_2(\mathbf{x}^*,\mathbf{y}), \text{ for all } \mathbf{y} \in Y.$$

Variational Inequality Problem (VIP):

- Operator $\mathcal{F}: Z \to Z$.
- (Weak) Solution $\mathbf{z}^* \in Z$:

$$\langle \mathcal{F}(\mathbf{z}^*), \mathbf{z}-\mathbf{z}^* \rangle \geq 0, \text{ for all } \mathbf{z} \in Z.$$

- Minimization of smooth convex function g: $\mathcal{F} = \nabla g$.
- NEP of concave games where u_1 (u_2) is concave in $\mathbf{x}$ ($\mathbf{y}$):

$$\mathcal{F} = -[\nabla_{\mathbf{x}} u_1, \nabla_{\mathbf{y}} u_2].$$

NEP and VIP

Two-Person Nash Equilibrium Problem (NEP):

- Player 1: $\max_{\mathbf{x}\in X} u_1(\mathbf{x},\mathbf{y})$; Player 2: $\max_{\mathbf{y}\in Y} u_2(\mathbf{x},\mathbf{y})$.
- X and Y are compact and convex; u_1 and u_2 are L-smooth on $X \times Y$.
- Nash equilibrium $(\mathbf{x}^*,\mathbf{y}^*) \in X \times Y$ s.t.

$$u_1(\mathbf{x}^*,\mathbf{y}^*) \geq u_1(\mathbf{x},\mathbf{y}^*), \text{ for all } \mathbf{x} \in X,$$
$$u_2(\mathbf{x}^*,\mathbf{y}^*) \geq u_2(\mathbf{x}^*,\mathbf{y}), \text{ for all } \mathbf{y} \in Y.$$

Variational Inequality Problem (VIP):

- Operator $\mathcal{F}: Z \to Z$.
- (Weak) Solution $\mathbf{z}^* \in Z$:

$$\langle \mathcal{F}(\mathbf{z}^*), \mathbf{z} - \mathbf{z}^* \rangle \geq 0, \text{ for all } \mathbf{z} \in Z.$$

- Minimization of smooth convex function g: $\mathcal{F} = \nabla g$.
- NEP of concave games where u_1 (u_2) is concave in $\mathbf{x}$ ($\mathbf{y}$):

$$\mathcal{F} = -[\nabla_{\mathbf{x}} u_1, \nabla_{\mathbf{y}} u_2].$$

NEP and VIP

Two-Person Nash Equilibrium Problem (NEP):

- Player 1: $\max_{\mathbf{x}\in X} u_1(\mathbf{x},\mathbf{y})$; Player 2: $\max_{\mathbf{y}\in Y} u_2(\mathbf{x},\mathbf{y})$.
- X and Y are compact and convex; u_1 and u_2 are L-smooth on $X \times Y$.
- Nash equilibrium $(\mathbf{x}^*,\mathbf{y}^*) \in X \times Y$ s.t.

$$u_1(\mathbf{x}^*,\mathbf{y}^*) \geq u_1(\mathbf{x},\mathbf{y}^*), \text{ for all } \mathbf{x} \in X,$$
$$u_2(\mathbf{x}^*,\mathbf{y}^*) \geq u_2(\mathbf{x}^*,\mathbf{y}), \text{ for all } \mathbf{y} \in Y.$$

Variational Inequality Problem (VIP):

- Operator $\mathcal{F}: Z \to Z$.
- (Weak) Solution $\mathbf{z}^* \in Z$:

$$\langle \mathcal{F}(\mathbf{z}^*), \mathbf{z} - \mathbf{z}^* \rangle \geq 0, \text{ for all } \mathbf{z} \in Z.$$

- Minimization of smooth convex function g: $\mathcal{F} = \nabla g$.
- NEP of concave games where u_1 (u_2) is concave in $\mathbf{x}$ ($\mathbf{y}$):

$$\mathcal{F} = -[\nabla_{\mathbf{x}} u_1, \nabla_{\mathbf{y}} u_2].$$

NEP and VIP

Two-Person Nash Equilibrium Problem (NEP):

- Player 1: $\max_{\mathbf{x} \in X} u_1(\mathbf{x}, \mathbf{y})$; Player 2: $\max_{\mathbf{y} \in Y} u_2(\mathbf{x}, \mathbf{y})$.
- X and Y are compact and convex; u_1 and u_2 are L-smooth on $X \times Y$.
- Nash equilibrium $(\mathbf{x}^*, \mathbf{y}^*) \in X \times Y$ s.t.

$$u_1(\mathbf{x}^*, \mathbf{y}^*) \geq u_1(\mathbf{x}, \mathbf{y}^*), \text{ for all } \mathbf{x} \in X,$$
$$u_2(\mathbf{x}^*, \mathbf{y}^*) \geq u_2(\mathbf{x}^*, \mathbf{y}), \text{ for all } \mathbf{y} \in Y.$$

Variational Inequality Problem (VIP):

- Operator $\mathcal{F}: Z \to Z$.
- (Weak) Solution $\mathbf{z}^* \in Z$:

$$\langle \mathcal{F}(\mathbf{z}^*), \mathbf{z} - \mathbf{z}^* \rangle \geq 0, \text{ for all } \mathbf{z} \in Z.$$

- Minimization of smooth convex function g: $\mathcal{F} = \nabla g$.
- NEP of concave games where u_1 (u_2) is concave in $\mathbf{x}$ ($\mathbf{y}$):

$$\mathcal{F} = -[\nabla_{\mathbf{x}} u_1, \nabla_{\mathbf{y}} u_2].$$

NEP and VIP

Two-Person Nash Equilibrium Problem (NEP):

- Player 1: $\max_{\mathbf{x} \in X} u_1(\mathbf{x}, \mathbf{y})$; Player 2: $\max_{\mathbf{y} \in Y} u_2(\mathbf{x}, \mathbf{y})$.
- X and Y are compact and convex; u_1 and u_2 are L-smooth on $X \times Y$.
- Nash equilibrium $(\mathbf{x}^*, \mathbf{y}^*) \in X \times Y$ s.t.

$$u_1(\mathbf{x}^*, \mathbf{y}^*) \geq u_1(\mathbf{x}, \mathbf{y}^*), \text{ for all } \mathbf{x} \in X,$$
$$u_2(\mathbf{x}^*, \mathbf{y}^*) \geq u_2(\mathbf{x}^*, \mathbf{y}), \text{ for all } \mathbf{y} \in Y.$$

Variational Inequality Problem (VIP):

- Operator $\mathcal{F}\colon Z \to Z$.
- (Weak) Solution $\mathbf{z}^* \in Z$:

$$\langle \mathcal{F}(\mathbf{z}^*), \mathbf{z} - \mathbf{z}^* \rangle \geq 0, \text{ for all } \mathbf{z} \in Z.$$

- Minimization of smooth convex function g: $\mathcal{F} = \nabla g$.
- NEP of concave games where u_1 (u_2) is concave in $\mathbf{x}$ $(\mathbf{y})$:

$$\mathcal{F} = -\left[\nabla_{\mathbf{x}} u_1, \nabla_{\mathbf{y}} u_2\right].$$

NEP and VIP

Two-Person Nash Equilibrium Problem (NEP):

- Player 1: $\max_{\mathbf{x}\in X} u_1(\mathbf{x},\mathbf{y})$; Player 2: $\max_{\mathbf{y}\in Y} u_2(\mathbf{x},\mathbf{y})$.
- X and Y are compact and convex; u_1 and u_2 are L-smooth on $X \times Y$.
- Nash equilibrium $(\mathbf{x}^*,\mathbf{y}^*) \in X \times Y$ s.t.

$$u_1(\mathbf{x}^*,\mathbf{y}^*) \geq u_1(\mathbf{x},\mathbf{y}^*), \text{ for all } \mathbf{x} \in X,$$
$$u_2(\mathbf{x}^*,\mathbf{y}^*) \geq u_2(\mathbf{x}^*,\mathbf{y}), \text{ for all } \mathbf{y} \in Y.$$

Variational Inequality Problem (VIP):

- Operator $\mathcal{F}\colon Z \to Z$.
- (Weak) Solution $\mathbf{z}^* \in Z$:

$$\langle \mathcal{F}(\mathbf{z}^*), \mathbf{z} - \mathbf{z}^* \rangle \geq 0, \text{ for all } \mathbf{z} \in Z.$$

- Minimization of smooth convex function g: $\mathcal{F} = \nabla g$.
- NEP of concave games where u_1 (u_2) is concave in $\mathbf{x}$ $(\mathbf{y})$:

$$\mathcal{F} = -\left[\nabla_{\mathbf{x}} u_1, \nabla_{\mathbf{y}} u_2\right].$$

NEP and VIP

Two-Person Nash Equilibrium Problem (NEP):

- Player 1: $\max_{\mathbf{x}\in X} u_1(\mathbf{x},\mathbf{y})$; Player 2: $\max_{\mathbf{y}\in Y} u_2(\mathbf{x},\mathbf{y})$.
- X and Y are compact and convex; u_1 and u_2 are L-smooth on $X \times Y$.
- Nash equilibrium $(\mathbf{x}^*,\mathbf{y}^*) \in X \times Y$ s.t.

$$u_1(\mathbf{x}^*,\mathbf{y}^*) \geq u_1(\mathbf{x},\mathbf{y}^*), \text{ for all } \mathbf{x} \in X,$$
$$u_2(\mathbf{x}^*,\mathbf{y}^*) \geq u_2(\mathbf{x}^*,\mathbf{y}), \text{ for all } \mathbf{y} \in Y.$$

Variational Inequality Problem (VIP):

- Operator $\mathcal{F}\colon Z \to Z$.
- (Weak) Solution $\mathbf{z}^* \in Z$:

$$\langle \mathcal{F}(\mathbf{z}^*), \mathbf{z} - \mathbf{z}^* \rangle \geq 0, \text{ for all } \mathbf{z} \in Z.$$

- Minimization of smooth convex function g: $\mathcal{F} = \nabla g$.
- NEP of concave games where u_1 (u_2) is concave in $\mathbf{x}$ $(\mathbf{y})$:

$$\mathcal{F} = -\left[\nabla_{\mathbf{x}} u_1, \nabla_{\mathbf{y}} u_2\right].$$

NEP and VIP

Two-Person Nash Equilibrium Problem (NEP):

- Player 1: $\max_{\mathbf{x} \in X} u_1(\mathbf{x}, \mathbf{y})$; Player 2: $\max_{\mathbf{y} \in Y} u_2(\mathbf{x}, \mathbf{y})$.
- X and Y are compact and convex; u_1 and u_2 are L-smooth on $X \times Y$.
- Nash equilibrium $(\mathbf{x}^*, \mathbf{y}^*) \in X \times Y$ s.t.

$$u_1(\mathbf{x}^*, \mathbf{y}^*) \geq u_1(\mathbf{x}, \mathbf{y}^*), \text{ for all } \mathbf{x} \in X,$$
$$u_2(\mathbf{x}^*, \mathbf{y}^*) \geq u_2(\mathbf{x}^*, \mathbf{y}), \text{ for all } \mathbf{y} \in Y.$$

Variational Inequality Problem (VIP):

- Operator $\mathcal{F}\colon Z \to Z$.
- (Weak) Solution $\mathbf{z}^* \in Z$:

$$\langle \mathcal{F}(\mathbf{z}^*), \mathbf{z} - \mathbf{z}^* \rangle \geq 0, \text{ for all } \mathbf{z} \in Z.$$

- Minimization of smooth convex function g: $\mathcal{F} = \nabla g$.
- NEP of concave games where u_1 (u_2) is concave in $\mathbf{x}$ ($\mathbf{y}$):

$$\mathcal{F} = -\left[\nabla_{\mathbf{x}} u_1, \nabla_{\mathbf{y}} u_2\right].$$

Decomposition of two-person games

Reformulation:

- $u_1 = -g - h$; $u_2 = -g + h$.
- Coupling: $g = (-u_1 - u_2)/2$;
- Zero-sum: $h = (-u_1 + u_2)/2$.

VIP of

$$\begin{aligned}\mathcal{F} = -[\nabla_{\mathbf{x}} u_1, \nabla_{\mathbf{y}} u_2] &= \nabla g + [\nabla_{\mathbf{x}} h, -\nabla_{\mathbf{y}} h] \\ &\triangleq \nabla g + \mathcal{H}.\end{aligned}$$

Zero-sum games (strictly competitive games):

- $g = 0$, or
- $\mathcal{F} = [\nabla_{\mathbf{x}} h, -\nabla_{\mathbf{y}} h] = \mathcal{H}$.
- Minimax optimization:

$$\min_{\mathbf{x} \in X} \max_{\mathbf{y} \in Y} h(\mathbf{x}, \mathbf{y}).$$

Decomposition of two-person games

Reformulation:

- $u_1 = -g - h$; $u_2 = -g + h$.
- Coupling: $g = (-u_1 - u_2)/2$;
- Zero-sum: $h = (-u_1 + u_2)/2$.

VIP of

$$\mathcal{F} = -[\nabla_{\mathbf{x}} u_1, \nabla_{\mathbf{y}} u_2] = \nabla g + [\nabla_{\mathbf{x}} h, -\nabla_{\mathbf{y}} h] \triangleq \nabla g + \mathcal{H}.$$

Zero-sum games (strictly competitive games):

- $g = 0$, or
- $\mathcal{F} = [\nabla_{\mathbf{x}} h, -\nabla_{\mathbf{y}} h] = \mathcal{H}$.
- Minimax optimization:

$$\min_{\mathbf{x} \in X} \max_{\mathbf{y} \in Y} h(\mathbf{x}, \mathbf{y}).$$

Decomposition of two-person games

Reformulation:

- $u_1 = -g - h$; $u_2 = -g + h$.
- Coupling: $g = (-u_1 - u_2)/2$;
- Zero-sum: $h = (-u_1 + u_2)/2$.

VIP of

$$\begin{aligned}\mathcal{F} = -\left[\nabla_{\mathbf{x}} u_1, \nabla_{\mathbf{y}} u_2\right] &= \nabla g + \left[\nabla_{\mathbf{x}} h, -\nabla_{\mathbf{y}} h\right] \\ &\triangleq \nabla g + \mathcal{H}.\end{aligned}$$

Zero-sum games (strictly competitive games):

- $g = 0$, or
- $\mathcal{F} = [\nabla_{\mathbf{x}} h, -\nabla_{\mathbf{y}} h] = \mathcal{H}$.
- Minimax optimization:

$$\min_{\mathbf{x} \in X} \max_{\mathbf{y} \in Y} h(\mathbf{x}, \mathbf{y}).$$

Decomposition of two-person games

Reformulation:

- $u_1 = -g - h$; $u_2 = -g + h$.
- Coupling: $g = (-u_1 - u_2)/2$;
- Zero-sum: $h = (-u_1 + u_2)/2$.

VIP of

$$\begin{aligned}\mathcal{F} = -[\nabla_{\mathbf{x}} u_1, \nabla_{\mathbf{y}} u_2] &= \nabla g + [\nabla_{\mathbf{x}} h, -\nabla_{\mathbf{y}} h] \\ &\triangleq \nabla g + \mathcal{H}.\end{aligned}$$

Zero-sum games (strictly competitive games):

- $g = 0$, or
- $\mathcal{F} = [\nabla_{\mathbf{x}} h, -\nabla_{\mathbf{y}} h] = \mathcal{H}$.
- Minimax optimization:

$$\min_{\mathbf{x} \in X} \max_{\mathbf{y} \in Y} h(\mathbf{x}, \mathbf{y}).$$

Decomposition of two-person games

Reformulation:

- $u_1 = -g - h$; $u_2 = -g + h$.
- Coupling: $g = (-u_1 - u_2)/2$;
- Zero-sum: $h = (-u_1 + u_2)/2$.

VIP of

$$\mathcal{F} = -\left[\nabla_{\mathbf{x}} u_1, \nabla_{\mathbf{y}} u_2\right] = \nabla g + \left[\nabla_{\mathbf{x}} h, -\nabla_{\mathbf{y}} h\right] \triangleq \nabla g + \mathcal{H}.$$

Zero-sum games (strictly competitive games):

- $g = 0$, or
- $\mathcal{F} = [\nabla_{\mathbf{x}} h, -\nabla_{\mathbf{y}} h] = \mathcal{H}$.
- Minimax optimization:

$$\min_{\mathbf{x} \in X} \max_{\mathbf{y} \in Y} h(\mathbf{x}, \mathbf{y}).$$

Decomposition of two-person games

Reformulation:

- $u_1 = -g - h$; $u_2 = -g + h$.
- Coupling: $g = (-u_1 - u_2)/2$;
- Zero-sum: $h = (-u_1 + u_2)/2$.

VIP of

$$\begin{aligned}\mathcal{F} = -\left[\nabla_{\mathbf{x}} u_1, \nabla_{\mathbf{y}} u_2\right] &= \nabla g + [\nabla_{\mathbf{x}} h, -\nabla_{\mathbf{y}} h] \\ &\triangleq \nabla g + \mathcal{H}.\end{aligned}$$

Zero-sum games (strictly competitive games):

- $g = 0$, or
- $\mathcal{F} = [\nabla_{\mathbf{x}} h, -\nabla_{\mathbf{y}} h] = \mathcal{H}$.
- Minimax optimization:

$$\min_{\mathbf{x} \in X} \max_{\mathbf{y} \in Y} h(\mathbf{x}, \mathbf{y}).$$

Decomposition of two-person games

Reformulation:

- $u_1 = -g - h$; $u_2 = -g + h$.
- Coupling: $g = (-u_1 - u_2)/2$;
- Zero-sum: $h = (-u_1 + u_2)/2$.

VIP of

$$\begin{aligned}\mathcal{F} = -\left[\nabla_{\mathbf{x}} u_1, \nabla_{\mathbf{y}} u_2\right] &= \nabla g + [\nabla_{\mathbf{x}} h, -\nabla_{\mathbf{y}} h] \\ &\triangleq \nabla g + \mathcal{H}.\end{aligned}$$

Zero-sum games (strictly competitive games):

- $g = 0$, or
- $\mathcal{F} = [\nabla_{\mathbf{x}} h, -\nabla_{\mathbf{y}} h] = \mathcal{H}$.
- Minimax optimization:

$$\min_{\mathbf{x} \in X} \max_{\mathbf{y} \in Y} h(\mathbf{x}, \mathbf{y}).$$

Decomposition of two-person games

Reformulation:

- $u_1 = -g - h$; $u_2 = -g + h$.
- Coupling: $g = (-u_1 - u_2)/2$;
- Zero-sum: $h = (-u_1 + u_2)/2$.

VIP of

$$\begin{aligned}\mathcal{F} = -\left[\nabla_{\mathbf{x}} u_1, \nabla_{\mathbf{y}} u_2\right] &= \nabla g + \left[\nabla_{\mathbf{x}} h, -\nabla_{\mathbf{y}} h\right] \\ &\triangleq \nabla g + \mathcal{H}.\end{aligned}$$

Zero-sum games (strictly competitive games):

- $g = 0$, or
- $\mathcal{F} = [\nabla_{\mathbf{x}} h, -\nabla_{\mathbf{y}} h] = \mathcal{H}$.
- Minimax optimization:

$$\min_{\mathbf{x} \in X} \max_{\mathbf{y} \in Y} h(\mathbf{x}, \mathbf{y}).$$

Decomposition of two-person games

Reformulation:

- $u_1 = -g - h$; $u_2 = -g + h$.
- Coupling: $g = (-u_1 - u_2)/2$;
- Zero-sum: $h = (-u_1 + u_2)/2$.

VIP of

$$\begin{aligned}\mathcal{F} = -\left[\nabla_{\mathbf{x}} u_1, \nabla_{\mathbf{y}} u_2\right] &= \nabla g + \left[\nabla_{\mathbf{x}} h, -\nabla_{\mathbf{y}} h\right] \\ &\triangleq \nabla g + \mathcal{H}.\end{aligned}$$

Zero-sum games (strictly competitive games):

- $g = 0$, or
- $\mathcal{F} = [\nabla_{\mathbf{x}} h, -\nabla_{\mathbf{y}} h] = \mathcal{H}$.
- Minimax optimization:

$$\min_{\mathbf{x} \in X} \max_{\mathbf{y} \in Y} h(\mathbf{x}, \mathbf{y}).$$

Decomposition of two-person games

Reformulation:

- $u_1 = -g - h$; $u_2 = -g + h$.
- Coupling: $g = (-u_1 - u_2)/2$;
- Zero-sum: $h = (-u_1 + u_2)/2$.

VIP of

$$\begin{aligned}\mathcal{F} = -\left[\nabla_{\mathbf{x}} u_1, \nabla_{\mathbf{y}} u_2\right] &= \nabla g + \left[\nabla_{\mathbf{x}} h, -\nabla_{\mathbf{y}} h\right] \\ &\triangleq \nabla g + \mathcal{H}.\end{aligned}$$

Zero-sum games (strictly competitive games):

- $g = 0$, or
- $\mathcal{F} = [\nabla_{\mathbf{x}} h, -\nabla_{\mathbf{y}} h] = \mathcal{H}$.
- Minimax optimization:

$$\min_{\mathbf{x} \in X} \max_{\mathbf{y} \in Y} h(\mathbf{x}, \mathbf{y}).$$

Monotone games

Assumption 1 (convex-concave zero-sum part)

The function $h(\cdot,\cdot)$ is μ-**strongly convex**-ν-**strongly concave**.

Assumption 2 (jointly convex coupling part)

The function $g(\cdot,\cdot)$ is **jointly convex**.

Proposition 1

Under Assumptions 1 and 2, $\mathcal{F} = \nabla g + \mathcal{H}$ is $\min\{\mu,\nu\}$-strongly monotone.

Proof.

∇g is monotone and $\mathcal{H}$ is $\min\{\mu,\nu\}$-strongly monotone. $\square$

Monotone games: Assumptions 1 and 2.

- Monotone (general-sum) games: g is L-smooth;
- Monotone zero-sum games: g is 0-smooth.

Monotone games

Assumption 1 (convex-concave zero-sum part)

The function $h(\cdot,\cdot)$ is μ-**strongly convex**-ν-**strongly concave**.

Assumption 2 (jointly convex coupling part)

The function $g(\cdot,\cdot)$ is **jointly convex**.

Proposition 1

Under Assumptions 1 and 2, $\mathcal{F} = \nabla g + \mathcal{H}$ is $\min\{\mu,\nu\}$-strongly monotone.

Proof.

∇g is monotone and $\mathcal{H}$ is $\min\{\mu,\nu\}$-strongly monotone. □

Monotone games: Assumptions 1 and 2.

- Monotone (general-sum) games: g is L-smooth;
- Monotone zero-sum games: g is 0-smooth.

Monotone games

Assumption 1 (convex-concave zero-sum part)

The function $h(\cdot, \cdot)$ is μ**-strongly convex-**ν**-strongly concave**.

Assumption 2 (jointly convex coupling part)

The function $g(\cdot, \cdot)$ is **jointly convex**.

Proposition 1

Under Assumptions 1 and 2, $\mathcal{F} = \nabla g + \mathcal{H}$ is $\min\{\mu, \nu\}$-strongly monotone.

Proof.

∇g is monotone and $\mathcal{H}$ is $\min\{\mu, \nu\}$-strongly monotone. □

Monotone games: Assumptions 1 and 2.

- Monotone (general-sum) games: g is L-smooth;
- Monotone zero-sum games: g is 0-smooth.

Monotone games

Assumption 1 (convex-concave zero-sum part)

The function $h(\cdot,\cdot)$ is μ-**strongly convex**-ν-**strongly concave**.

Assumption 2 (jointly convex coupling part)

The function $g(\cdot,\cdot)$ is **jointly convex**.

Proposition 1

Under Assumptions 1 and 2, $\mathcal{F} = \nabla g + \mathcal{H}$ is $\min\{\mu, \nu\}$-strongly monotone.

Proof.

∇g is monotone and $\mathcal{H}$ is $\min\{\mu, \nu\}$-strongly monotone. $\square$

Monotone games: Assumptions 1 and 2.

- Monotone (general-sum) games: g is L-smooth;
- Monotone zero-sum games: g is 0-smooth.

Monotone games

Assumption 1 (convex-concave zero-sum part)

The function $h(\cdot,\cdot)$ is μ-**strongly convex**-ν-**strongly concave**.

Assumption 2 (jointly convex coupling part)

The function $g(\cdot,\cdot)$ is **jointly convex**.

Proposition 1

Under Assumptions 1 and 2, $\mathcal{F} = \nabla g + \mathcal{H}$ is $\min\{\mu,\nu\}$-strongly monotone.

Proof.

∇g is monotone and $\mathcal{H}$ is $\min\{\mu,\nu\}$-strongly monotone. □

Monotone games: Assumptions 1 and 2.

- Monotone (general-sum) games: g is L-smooth;
- Monotone zero-sum games: g is 0-smooth.

Monotone games

Assumption 1 (convex-concave zero-sum part)

The function $h(\cdot,\cdot)$ is μ-**strongly convex**-ν-**strongly concave**.

Assumption 2 (jointly convex coupling part)

The function $g(\cdot,\cdot)$ is **jointly convex**.

Proposition 1

Under Assumptions 1 and 2, $\mathcal{F} = \nabla g + \mathcal{H}$ is $\min\{\mu,\nu\}$*-strongly monotone.*

Proof.

∇g is monotone and $\mathcal{H}$ is $\min\{\mu,\nu\}$-strongly monotone. $\square$

Monotone games: Assumptions 1 and 2.

- Monotone (general-sum) games: g is L-smooth;
- Monotone zero-sum games: g is 0-smooth.

Monotone games

Assumption 1 (convex-concave zero-sum part)

The function $h(\cdot, \cdot)$ is μ-**strongly convex**-ν-**strongly concave**.

Assumption 2 (jointly convex coupling part)

The function $g(\cdot, \cdot)$ is **jointly convex**.

Proposition 1

Under Assumptions 1 and 2, $\mathcal{F} = \nabla g + \mathcal{H}$ *is* $\min\{\mu, \nu\}$*-strongly monotone.*

Proof.

∇g is monotone and $\mathcal{H}$ is $\min\{\mu, \nu\}$-strongly monotone. $\square$

Monotone games: Assumptions 1 and 2.

- Monotone (general-sum) games: g is L-smooth;
- Monotone zero-sum games: g is 0-smooth.

A historical overview

- Equal conditioning ($\mu = \nu$):
  - Bilinear coupling, zero sum[1]
  - General coupling, general sum[2] (multi-player)
- General conditioning ($\mu \neq \nu$):
  - Bilinear coupling, zero sum[3]
  - General coupling, zero sum[4]
  - General coupling, general sum: ?

---

[1] Yurii Nesterov. "Smooth minimization of non-smooth functions". In: *Mathematical programming* 103 (2005), pp. 127–152.

[2] Arkadi Nemirovski. "Prox-method with rate of convergence O (1/t) for variational inequalities with Lipschitz continuous monotone operators and smooth convex-concave saddle point problems". In: *SIAM Journal on Optimization* 15.1 (2004), pp. 229–251.

[3] Antonin Chambolle and Thomas Pock. "A first-order primal-dual algorithm for convex problems with applications to imaging". In: *Journal of mathematical imaging and vision* 40.1 (2011), pp. 120–145; Yunmei Chen, Guanghui Lan, and Yuyuan Ouyang. "Optimal primal-dual methods for a class of saddle point problems". In: *SIAM Journal on Optimization* 24.4 (2014), pp. 1779–1814; Kiran K. Thekumparampil, Niao He, and Sewoong Oh. "Lifted Primal-Dual Method for Bilinearly Coupled Smooth Minimax Optimization". In: *Proceedings of The 25th International Conference on Artificial Intelligence and Statistics*. Ed. by Gustau Camps-Valls, Francisco J. R. Ruiz, and

A historical overview

- Equal conditioning ($\mu = \nu$):
  - Bilinear coupling, zero sum[1]
  - General coupling, general sum[2] (multi-player)
- General conditioning ($\mu \neq \nu$):
  - Bilinear coupling, zero sum[3]
  - General coupling, zero sum[4]
  - General coupling, general sum: ?

[1] Nesterov, "Smooth minimization of non-smooth functions".

[2] Nemirovski, "Prox-method with rate of convergence O (1/t) for variational inequalities with Lipschitz continuous monotone operators and smooth convex-concave saddle point problems".

[3] Chambolle and Pock, "A first-order primal-dual algorithm for convex problems with applications to imaging"; Chen, Lan, and Ouyang, "Optimal primal-dual methods for a class of saddle point problems"; Thekumparampil, He, and Oh, "Lifted Primal-Dual Method for Bilinearly Coupled Smooth Minimax Optimization".

[4] Lin, Jin, and Jordan, "Near-optimal algorithms for minimax optimization"; Kovalev and Gasnikov, "The first optimal algorithm for smooth and strongly-convex-strongly-concave minimax optimization"; Lan and Li, "A Novel Catalyst Scheme for Stochastic Minimax Optimization".

A historical overview

- Equal conditioning ($\mu = \nu$):
  - Bilinear coupling, zero sum[1]
  - General coupling, general sum[2] (multi-player)
- General conditioning ($\mu \neq \nu$):
  - Bilinear coupling, zero sum[3]
  - General coupling, zero sum[4]
  - General coupling, general sum: ?

[1] Nesterov, "Smooth minimization of non-smooth functions".

[2] Nemirovski, "Prox-method with rate of convergence O (1/t) for variational inequalities with Lipschitz continuous monotone operators and smooth convex-concave saddle point problems".

[3] Chambolle and Pock, "A first-order primal-dual algorithm for convex problems with applications to imaging"; Chen, Lan, and Ouyang, "Optimal primal-dual methods for a class of saddle point problems"; Thekumparampil, He, and Oh, "Lifted Primal-Dual Method for Bilinearly Coupled Smooth Minimax Optimization".

[4] Lin, Jin, and Jordan, "Near-optimal algorithms for minimax optimization"; Kovalev and Gasnikov, "The first optimal algorithm for smooth and strongly-convex-strongly-concave minimax optimization"; Lan and Li, "A Novel Catalyst Scheme for Stochastic Minimax Optimization".

A historical overview

- Equal conditioning ($\mu = \nu$):
  - Bilinear coupling, zero sum[1]
  - General coupling, general sum[2] (multi-player)
- General conditioning ($\mu \neq \nu$):
  - Bilinear coupling, zero sum[3]
  - General coupling, zero sum[4]
  - General coupling, general sum: ?

[1] Nesterov, "Smooth minimization of non-smooth functions".

[2] Nemirovski, "Prox-method with rate of convergence O (1/t) for variational inequalities with Lipschitz continuous monotone operators and smooth convex-concave saddle point problems".

[3] Chambolle and Pock, "A first-order primal-dual algorithm for convex problems with applications to imaging"; Chen, Lan, and Ouyang, "Optimal primal-dual methods for a class of saddle point problems"; Thekumparampil, He, and Oh, "Lifted Primal-Dual Method for Bilinearly Coupled Smooth Minimax Optimization".

[4] Lin, Jin, and Jordan, "Near-optimal algorithms for minimax optimization"; Kovalev and Gasnikov, "The first optimal algorithm for smooth and strongly-convex-strongly-concave minimax optimization"; Lan and Li, "A Novel Catalyst Scheme for Stochastic Minimax Optimization".

A historical overview

- Equal conditioning ($\mu = \nu$):
  - Bilinear coupling, zero sum[1]
  - General coupling, general sum[2] (multi-player)
- General conditioning ($\mu \neq \nu$):
  - Bilinear coupling, zero sum[3]
  - General coupling, zero sum[4]
  - General coupling, general sum: ?

[1] Nesterov, "Smooth minimization of non-smooth functions".

[2] Nemirovski, "Prox-method with rate of convergence O (1/t) for variational inequalities with Lipschitz continuous monotone operators and smooth convex-concave saddle point problems".

[3] Chambolle and Pock, "A first-order primal-dual algorithm for convex problems with applications to imaging"; Chen, Lan, and Ouyang, "Optimal primal-dual methods for a class of saddle point problems"; Thekumparampil, He, and Oh, "Lifted Primal-Dual Method for Bilinearly Coupled Smooth Minimax Optimization".

[4] Lin, Jin, and Jordan, "Near-optimal algorithms for minimax optimization"; Kovalev and Gasnikov, "The first optimal algorithm for smooth and strongly-convex-strongly-concave minimax optimization"; Lan and Li, "A Novel Catalyst Scheme for Stochastic Minimax Optimization".

A historical overview

- Equal conditioning ($\mu = \nu$):
  - Bilinear coupling, zero sum[1]
  - General coupling, general sum[2] (multi-player)
- General conditioning ($\mu \neq \nu$):
  - Bilinear coupling, zero sum[3]
  - General coupling, zero sum[4]
  - General coupling, general sum: ?

[1] Nesterov, "Smooth minimization of non-smooth functions".

[2] Nemirovski, "Prox-method with rate of convergence O $(1/t)$ for variational inequalities with Lipschitz continuous monotone operators and smooth convex-concave saddle point problems".

[3] Chambolle and Pock, "A first-order primal-dual algorithm for convex problems with applications to imaging"; Chen, Lan, and Ouyang, "Optimal primal-dual methods for a class of saddle point problems"; Thekumparampil, He, and Oh, "Lifted Primal-Dual Method for Bilinearly Coupled Smooth Minimax Optimization".

[4] Lin, Jin, and Jordan, "Near-optimal algorithms for minimax optimization"; Kovalev and Gasnikov, "The first optimal algorithm for smooth and strongly-convex-strongly-concave minimax optimization"; Lan and Li, "A Novel Catalyst Scheme for Stochastic Minimax Optimization".

A historical overview

- Equal conditioning ($\mu = \nu$):
  - Bilinear coupling, zero sum[1]
  - General coupling, general sum[2] (multi-player)
- General conditioning ($\mu \neq \nu$):
  - Bilinear coupling, zero sum[3]
  - General coupling, zero sum[4]
  - General coupling, general sum: ?

---

[1] Nesterov, "Smooth minimization of non-smooth functions".

[2] Nemirovski, "Prox-method with rate of convergence O (1/t) for variational inequalities with Lipschitz continuous monotone operators and smooth convex-concave saddle point problems".

[3] Chambolle and Pock, "A first-order primal-dual algorithm for convex problems with applications to imaging"; Chen, Lan, and Ouyang, "Optimal primal-dual methods for a class of saddle point problems"; Thekumparampil, He, and Oh, "Lifted Primal-Dual Method for Bilinearly Coupled Smooth Minimax Optimization".

[4] Lin, Jin, and Jordan, "Near-optimal algorithms for minimax optimization"; Kovalev and Gasnikov, "The first optimal algorithm for smooth and strongly-convex-strongly-concave minimax optimization"; Lan and Li, "A Novel Catalyst Scheme for Stochastic Minimax Optimization".

Gradient complexity for ε-Nash equilibrium

Proposition 2 (Monotone (general-sum) games[5])

For monotone games, the gradient complexity is $\widetilde{\mathcal{O}}\left(\frac{L}{\min\{\mu,\nu\}} \cdot \log\left(\frac{1}{\varepsilon}\right)\right)$.

Proposition 3 (Monotone zero-sum games (minimax optimization)[6])

For monotone zero-sum games, the gradient complexity is

$$\widetilde{\mathcal{O}}\left(\frac{L}{\sqrt{\mu\nu}} \cdot \log\left(\frac{1}{\varepsilon}\right)\right).$$

[5] Paul Tseng. "On linear convergence of iterative methods for the variational inequality problem". In: *Journal of Computational and Applied Mathematics* 60.1-2 (1995), pp. 237–252.

[6] Kovalev and Gasnikov, "The first optimal algorithm for smooth and strongly-convex-strongly-concave minimax optimization"; Lan and Li, "A Novel Catalyst Scheme for Stochastic Minimax Optimization".

Gradient complexity for ε-Nash equilibrium

Proposition 2 (Monotone (general-sum) games[5])

For monotone games, the gradient complexity is $\widetilde{\mathcal{O}}\left(\frac{L}{\min\{\mu,\nu\}} \cdot \log\left(\frac{1}{\varepsilon}\right)\right)$.

Proposition 3 (Monotone zero-sum games (minimax optimization)[6])

For monotone zero-sum games, the gradient complexity is

$$\widetilde{\mathcal{O}}\left(\frac{L}{\sqrt{\mu\nu}} \cdot \log\left(\frac{1}{\varepsilon}\right)\right).$$

[5] Tseng, "On linear convergence of iterative methods for the variational inequality problem".

[6] Kovalev and Gasnikov, "The first optimal algorithm for smooth and strongly-convex-strongly-concave minimax optimization"; Lan and Li, "A Novel Catalyst Scheme for Stochastic Minimax Optimization".

Outline for section 2

New problem class

Assumption 3 (near-zero-sum)

There exists $\delta \in [0, L]$ such that the function $g(\cdot, \cdot)$ is δ-**smooth**.

Definition 1 (Monotone Near-Zero-Sum Games)

A two-person general-sum game is a **monotone δ-near-zero-sum game** if it satisfies Assumptions 1 to 3.

We study the NEP of monotone near-zero-sum games, or the VIP of

$$\mathcal{F}(\mathbf{x}, \mathbf{y}) = \nabla g(\mathbf{x}, \mathbf{y}) + \underbrace{\mathcal{H}(\mathbf{x}, \mathbf{y})}_{[\nabla_{\mathbf{x}} h(\mathbf{x},\mathbf{y}), -\nabla_{\mathbf{y}} h(\mathbf{x},\mathbf{y})]}, \quad \mathbf{x} \in X, \ \mathbf{y} \in Y, \qquad (1)$$

where X and Y are compact and convex, and

1. ∇g is δ-Lipschitz continuous; $\mathcal{H}$ is L-Lipschitz continuous ($\delta \leq L$);
2. $g(\mathbf{x}, \mathbf{y})$ is jointly convex; h is μ-strongly convex-ν-strongly concave.

New problem class

Assumption 3 (near-zero-sum)

There exists $\delta \in [0, L]$ such that the function $g(\cdot, \cdot)$ is δ-**smooth**.

Definition 1 (Monotone Near-Zero-Sum Games)

A two-person general-sum game is a **monotone δ-near-zero-sum game** if it satisfies Assumptions 1 to 3.

We study the NEP of monotone near-zero-sum games, or the VIP of

$$\mathcal{F}(\mathbf{x}, \mathbf{y}) = \nabla g(\mathbf{x}, \mathbf{y}) + \underbrace{\mathcal{H}(\mathbf{x}, \mathbf{y})}_{[\nabla_{\mathbf{x}} h(\mathbf{x},\mathbf{y}), -\nabla_{\mathbf{y}} h(\mathbf{x},\mathbf{y})]}, \quad \mathbf{x} \in X,\ \mathbf{y} \in Y, \tag{1}$$

where X and Y are compact and convex, and

1. ∇g is δ-Lipschitz continuous; $\mathcal{H}$ is L-Lipschitz continuous ($\delta \leq L$);
2. $g(\mathbf{x}, \mathbf{y})$ is jointly convex; h is μ-strongly convex-ν-strongly concave.

New problem class

Assumption 3 (near-zero-sum)

There exists $\delta \in [0, L]$ such that the function $g(\cdot, \cdot)$ is δ-**smooth**.

Definition 1 (Monotone Near-Zero-Sum Games)

A two-person general-sum game is a **monotone δ-near-zero-sum game** if it satisfies Assumptions 1 to 3.

We study the NEP of monotone near-zero-sum games, or the VIP of

$$\mathcal{F}(\mathbf{x}, \mathbf{y}) = \nabla g(\mathbf{x}, \mathbf{y}) + \underbrace{\mathcal{H}(\mathbf{x}, \mathbf{y})}_{[\nabla_{\mathbf{x}} h(\mathbf{x},\mathbf{y}), -\nabla_{\mathbf{y}} h(\mathbf{x},\mathbf{y})]}, \quad \mathbf{x} \in X, \ \mathbf{y} \in Y, \qquad (1)$$

where X and Y are compact and convex, and

1. ∇g is δ-Lipschitz continuous; $\mathcal{H}$ is L-Lipschitz continuous ($\delta \leq L$);
2. $g(\mathbf{x}, \mathbf{y})$ is jointly convex; h is μ-strongly convex-ν-strongly concave.

New problem class

Assumption 3 (near-zero-sum)

There exists $\delta \in [0, L]$ such that the function $g(\cdot, \cdot)$ is δ-**smooth**.

Definition 1 (Monotone Near-Zero-Sum Games)

A two-person general-sum game is a **monotone δ-near-zero-sum game** if it satisfies Assumptions 1 to 3.

We study the NEP of monotone near-zero-sum games, or the VIP of

$$\mathcal{F}(\mathbf{x}, \mathbf{y}) = \nabla g(\mathbf{x}, \mathbf{y}) + \underbrace{\mathcal{H}(\mathbf{x}, \mathbf{y})}_{[\nabla_{\mathbf{x}} h(\mathbf{x},\mathbf{y}), -\nabla_{\mathbf{y}} h(\mathbf{x},\mathbf{y})]}, \quad \mathbf{x} \in X, \ \mathbf{y} \in Y, \tag{1}$$

where X and Y are compact and convex, and

1. ∇g is δ-Lipschitz continuous; $\mathcal{H}$ is L-Lipschitz continuous ($\delta \le L$);
2. $g(\mathbf{x}, \mathbf{y})$ is jointly convex; h is μ-strongly convex-ν-strongly concave.

New problem class

Assumption 3 (near-zero-sum)

There exists $\delta \in [0, L]$ such that the function $g(\cdot, \cdot)$ is δ-**smooth**.

Definition 1 (MONOTONE NEAR-ZERO-SUM GAMES)

A two-person general-sum game is a **monotone δ-near-zero-sum game** if it satisfies Assumptions 1 to 3.

We study the NEP of monotone near-zero-sum games, or the VIP of

$$\mathcal{F}(\mathbf{x}, \mathbf{y}) = \nabla g(\mathbf{x}, \mathbf{y}) + \underbrace{\mathcal{H}(\mathbf{x}, \mathbf{y})}_{[\nabla_{\mathbf{x}} h(\mathbf{x},\mathbf{y}), -\nabla_{\mathbf{y}} h(\mathbf{x},\mathbf{y})]}, \quad \mathbf{x} \in X, \ \mathbf{y} \in Y, \qquad (1)$$

where X and Y are compact and convex, and

1. ∇g is δ-Lipschitz continuous; $\mathcal{H}$ is L-Lipschitz continuous ($\delta \leq L$);
2. $g(\mathbf{x}, \mathbf{y})$ is jointly convex; h is μ-strongly convex-ν-strongly concave.

Outline for section 3

An Attempt: Nesterov's Smoothing

Monotone zero-sum games ($u_1 + u_2 = 0$):

- $h_x(\mathbf{x}) \triangleq -u_1(\mathbf{x}, \mathbf{y}(\mathbf{x}))$ is μ-strongly-convex,
- where $\mathbf{y}(\mathbf{x}) \triangleq \arg\max_{\mathbf{y} \in Y} u_2(\mathbf{x}, \mathbf{y})$.

Nesterov's smoothing: (assuming $\mu \leq \nu$)

$$\mathbf{x}_{t+1} \approx \arg\min_{\mathbf{x} \in X} \left[h_x(\mathbf{x}) + \frac{\nu}{2} \|\mathbf{x} - \mathbf{x}_t\|^2 \right] .$$

- **Total gradient complexity:** $\widetilde{\mathcal{O}}\left(\frac{L}{\sqrt{\mu\nu}} \cdot \log^2\left(\frac{1}{\varepsilon}\right)\right)$.
- Outer loop: $\widetilde{\mathcal{O}}\left(\sqrt{\nu/\mu} \cdot \log\left(\frac{1}{\varepsilon}\right)\right)$ iterations;[7]
- Inner loop: $\widetilde{\mathcal{O}}\left(L/\nu \cdot \log\left(\frac{1}{\varepsilon}\right)\right)$ gradient queries.[8]

---

[7] Nesterov, "Smooth minimization of non-smooth functions".

[8] Tseng, "On linear convergence of iterative methods for the variational inequality problem".

An Attempt: Nesterov's Smoothing

Monotone zero-sum games ($u_1 + u_2 = 0$):

- $h_x(\mathbf{x}) \triangleq -u_1(\mathbf{x}, \mathbf{y}(\mathbf{x}))$ is μ-strongly-convex,
- where $\mathbf{y}(\mathbf{x}) \triangleq \arg\max_{\mathbf{y} \in Y} u_2(\mathbf{x}, \mathbf{y})$.

Nesterov's smoothing: (assuming $\mu \leq \nu$)

$$\mathbf{x}_{t+1} \approx \arg\min_{\mathbf{x} \in X} \left[h_x(\mathbf{x}) + \frac{\nu}{2} \|\mathbf{x} - \mathbf{x}_t\|^2 \right].$$

- **Total gradient complexity:** $\widetilde{\mathcal{O}}\left(\frac{L}{\sqrt{\mu\nu}} \cdot \log^2\left(\frac{1}{\varepsilon}\right)\right)$.
- Outer loop: $\widetilde{\mathcal{O}}\left(\sqrt{\nu/\mu} \cdot \log\left(\frac{1}{\varepsilon}\right)\right)$ iterations;[7]
- Inner loop: $\widetilde{\mathcal{O}}\left(L/\nu \cdot \log\left(\frac{1}{\varepsilon}\right)\right)$ gradient queries.[8]

[7] Nesterov, "Smooth minimization of non-smooth functions".

[8] Tseng, "On linear convergence of iterative methods for the variational inequality problem".

An Attempt: Nesterov's Smoothing

Monotone zero-sum games ($u_1 + u_2 = 0$):

- $h_x(\mathbf{x}) \triangleq -u_1(\mathbf{x}, \mathbf{y}(\mathbf{x}))$ is μ-strongly-convex,
- where $\mathbf{y}(\mathbf{x}) \triangleq \arg\max_{\mathbf{y} \in Y} u_2(\mathbf{x}, \mathbf{y})$.

Nesterov's smoothing: (assuming $\mu \leq \nu$)

$$\mathbf{x}_{t+1} \approx \arg\min_{\mathbf{x} \in X} \left[h_x(\mathbf{x}) + \frac{\nu}{2} \|\mathbf{x} - \mathbf{x}_t\|^2 \right] .$$

- **Total gradient complexity:** $\widetilde{\mathcal{O}}\left(\frac{L}{\sqrt{\mu\nu}} \cdot \log^2\left(\frac{1}{\varepsilon}\right)\right)$.
- Outer loop: $\widetilde{\mathcal{O}}\left(\sqrt{\nu/\mu} \cdot \log\left(\frac{1}{\varepsilon}\right)\right)$ iterations;[7]
- Inner loop: $\widetilde{\mathcal{O}}\left(L/\nu \cdot \log\left(\frac{1}{\varepsilon}\right)\right)$ gradient queries.[8]

---

[7] Nesterov, "Smooth minimization of non-smooth functions".

[8] Tseng, "On linear convergence of iterative methods for the variational inequality problem".

An Attempt: Nesterov's Smoothing

Monotone zero-sum games ($u_1 + u_2 = 0$):

- $h_x(\mathbf{x}) \triangleq -u_1(\mathbf{x}, \mathbf{y}(\mathbf{x}))$ is μ-strongly-convex,
- where $\mathbf{y}(\mathbf{x}) \triangleq \arg\max_{\mathbf{y} \in Y} u_2(\mathbf{x}, \mathbf{y})$.

Nesterov's smoothing: (assuming $\mu \leq \nu$)

$$\mathbf{x}_{t+1} \approx \arg\min_{\mathbf{x} \in X} \left[h_x(\mathbf{x}) + \frac{\nu}{2} \left\| \mathbf{x} - \mathbf{x}_t \right\|^2 \right] .$$

- **Total gradient complexity:** $\widetilde{\mathcal{O}}\left(\frac{L}{\sqrt{\mu\nu}} \cdot \log^2\left(\frac{1}{\varepsilon}\right)\right)$.
- Outer loop: $\widetilde{\mathcal{O}}\left(\sqrt{\nu/\mu} \cdot \log\left(\frac{1}{\varepsilon}\right)\right)$ iterations;[7]
- Inner loop: $\widetilde{\mathcal{O}}\left(L/\nu \cdot \log\left(\frac{1}{\varepsilon}\right)\right)$ gradient queries.[8]

[7] Nesterov, "Smooth minimization of non-smooth functions".

[8] Tseng, "On linear convergence of iterative methods for the variational inequality problem".

An Attempt: Nesterov's Smoothing

Monotone zero-sum games ($u_1 + u_2 = 0$):

- $h_x(\mathbf{x}) \triangleq -u_1(\mathbf{x}, \mathbf{y}(\mathbf{x}))$ is μ-strongly-convex,
- where $\mathbf{y}(\mathbf{x}) \triangleq \arg\max_{\mathbf{y} \in Y} u_2(\mathbf{x}, \mathbf{y})$.

Nesterov's smoothing: (assuming $\mu \leq \nu$)

$$\mathbf{x}_{t+1} \approx \arg\min_{\mathbf{x} \in X} \left[h_x(\mathbf{x}) + \frac{\nu}{2} \|\mathbf{x} - \mathbf{x}_t\|^2 \right].$$

- **Total gradient complexity:** $\widetilde{\mathcal{O}}\left(\frac{L}{\sqrt{\mu\nu}} \cdot \log^2\left(\frac{1}{\varepsilon}\right)\right)$.
- Outer loop: $\widetilde{\mathcal{O}}\left(\sqrt{\nu/\mu} \cdot \log\left(\frac{1}{\varepsilon}\right)\right)$ iterations;[7]
- Inner loop: $\widetilde{\mathcal{O}}\left(L/\nu \cdot \log\left(\frac{1}{\varepsilon}\right)\right)$ gradient queries.[8]

[7] Nesterov, "Smooth minimization of non-smooth functions".

[8] Tseng, "On linear convergence of iterative methods for the variational inequality problem".

An Attempt: Nesterov's Smoothing

Monotone zero-sum games ($u_1 + u_2 = 0$):

- $h_x(\mathbf{x}) \triangleq -u_1(\mathbf{x}, \mathbf{y}(\mathbf{x}))$ is μ-strongly-convex,
- where $\mathbf{y}(\mathbf{x}) \triangleq \arg\max_{\mathbf{y} \in Y} u_2(\mathbf{x}, \mathbf{y})$.

Nesterov's smoothing: (assuming $\mu \leq \nu$)

$$\mathbf{x}_{t+1} \approx \arg\min_{\mathbf{x} \in X} \left[h_x(\mathbf{x}) + \frac{\nu}{2} \left\| \mathbf{x} - \mathbf{x}_t \right\|^2 \right] .$$

- **Total gradient complexity:** $\widetilde{\mathcal{O}}\left(\frac{L}{\sqrt{\mu\nu}} \cdot \log^2\left(\frac{1}{\varepsilon}\right)\right)$.
- Outer loop: $\widetilde{\mathcal{O}}\left(\sqrt{\nu/\mu} \cdot \log\left(\frac{1}{\varepsilon}\right)\right)$ iterations;[7]
- Inner loop: $\widetilde{\mathcal{O}}\left(L/\nu \cdot \log\left(\frac{1}{\varepsilon}\right)\right)$ gradient queries.[8]

[7] Nesterov, "Smooth minimization of non-smooth functions".

[8] Tseng, "On linear convergence of iterative methods for the variational inequality problem".

An Attempt: Nesterov's Smoothing

Monotone zero-sum games ($u_1 + u_2 = 0$):

- $h_x(\mathbf{x}) \triangleq -u_1(\mathbf{x}, \mathbf{y}(\mathbf{x}))$ is μ-strongly-convex,
- where $\mathbf{y}(\mathbf{x}) \triangleq \arg\max_{\mathbf{y} \in Y} u_2(\mathbf{x}, \mathbf{y})$.

Nesterov's smoothing: (assuming $\mu \leq \nu$)

$$\mathbf{x}_{t+1} \approx \arg\min_{\mathbf{x} \in X} \left[h_x(\mathbf{x}) + \frac{\nu}{2} \|\mathbf{x} - \mathbf{x}_t\|^2 \right].$$

- **Total gradient complexity:** $\widetilde{\mathcal{O}}\left(\frac{L}{\sqrt{\mu\nu}} \cdot \log^2\left(\frac{1}{\varepsilon}\right)\right)$.
- Outer loop: $\widetilde{\mathcal{O}}\left(\sqrt{\nu/\mu} \cdot \log\left(\frac{1}{\varepsilon}\right)\right)$ iterations;[7]
- Inner loop: $\widetilde{\mathcal{O}}\left(L/\nu \cdot \log\left(\frac{1}{\varepsilon}\right)\right)$ gradient queries.[8]

[7] Nesterov, "Smooth minimization of non-smooth functions".

[8] Tseng, "On linear convergence of iterative methods for the variational inequality problem".

An Attempt: Nesterov's Smoothing

Monotone zero-sum games ($u_1 + u_2 = 0$):

- $h_x(\mathbf{x}) \triangleq -u_1(\mathbf{x}, \mathbf{y}(\mathbf{x}))$ is μ-strongly-convex,
- where $\mathbf{y}(\mathbf{x}) \triangleq \arg\max_{\mathbf{y} \in Y} u_2(\mathbf{x}, \mathbf{y})$.

Nesterov's smoothing: (assuming $\mu \leq \nu$)

$$\mathbf{x}_{t+1} \approx \arg\min_{\mathbf{x} \in X} \left[h_x(\mathbf{x}) + \frac{\nu}{2} \|\mathbf{x} - \mathbf{x}_t\|^2 \right].$$

- **Total gradient complexity:** $\widetilde{\mathcal{O}}\left(\frac{L}{\sqrt{\mu\nu}} \cdot \log^2\left(\frac{1}{\varepsilon}\right)\right)$.
- Outer loop: $\widetilde{\mathcal{O}}\left(\sqrt{\nu/\mu} \cdot \log\left(\frac{1}{\varepsilon}\right)\right)$ iterations;[7]
- Inner loop: $\widetilde{\mathcal{O}}\left(L/\nu \cdot \log\left(\frac{1}{\varepsilon}\right)\right)$ gradient queries.[8]

[7] Nesterov, "Smooth minimization of non-smooth functions".

[8] Tseng, "On linear convergence of iterative methods for the variational inequality problem".

Cannot be directly applied here!

Counterexample for non-zero-sum games:

- $X = [0,1] \times [1,2] \subseteq \mathbb{R}^2$ and $Y = [-1,0] \subseteq \mathbb{R}$;
- $u_1 = -\frac{1}{2}(x_1 - 1)^2 - \frac{1}{2}(x_2 - 1)^2 + \frac{1}{2}x_1 y$ and $u_2 = \frac{1}{2}x_2 y - (y+1)^2$.
- Converges to the **Stackelberg solution**: $\left(\mathbf{x} = \left(\frac{40}{63}, \frac{68}{63}\right),\ y = -\frac{46}{63}\right)$,
- NOT Nash equilibrium: $\left(\mathbf{x} = \left(\frac{5}{8}, 1\right),\ y = -\frac{3}{4}\right)$.

Cannot be directly applied here!

Counterexample for non-zero-sum games:

- $X = [0,1] \times [1,2] \subseteq \mathbb{R}^2$ and $Y = [-1,0] \subseteq \mathbb{R}$;
- $u_1 = -\frac{1}{2}(x_1 - 1)^2 - \frac{1}{2}(x_2 - 1)^2 + \frac{1}{2}x_1 y$ and $u_2 = \frac{1}{2}x_2 y - (y+1)^2$.
- Converges to the **Stackelberg solution**: $\left(\mathbf{x} = \left(\frac{40}{63}, \frac{68}{63}\right),\ y = -\frac{46}{63}\right)$,
- NOT Nash equilibrium: $\left(\mathbf{x} = \left(\frac{5}{8}, 1\right),\ y = -\frac{3}{4}\right)$.

Cannot be directly applied here!

Counterexample for non-zero-sum games:

- $X = [0, 1] \times [1, 2] \subseteq \mathbb{R}^2$ and $Y = [-1, 0] \subseteq \mathbb{R}$;
- $u_1 = -\frac{1}{2}(x_1 - 1)^2 - \frac{1}{2}(x_2 - 1)^2 + \frac{1}{2}x_1 y$ and $u_2 = \frac{1}{2}x_2 y - (y + 1)^2$.
- Converges to the **Stackelberg solution**: $(\mathbf{x} = \left(\frac{40}{63}, \frac{68}{63}\right),\ y = -\frac{46}{63})$,
- NOT Nash equilibrium: $(\mathbf{x} = \left(\frac{5}{8}, 1\right),\ y = -\frac{3}{4})$.

Cannot be directly applied here!

Counterexample for non-zero-sum games:

- $X = [0,1] \times [1,2] \subseteq \mathbb{R}^2$ and $Y = [-1,0] \subseteq \mathbb{R}$;
- $u_1 = -\frac{1}{2}(x_1 - 1)^2 - \frac{1}{2}(x_2 - 1)^2 + \frac{1}{2}x_1 y$ and $u_2 = \frac{1}{2}x_2 y - (y+1)^2$.
- Converges to the **Stackelberg solution**: $\left(\mathbf{x} = \left(\frac{40}{63}, \frac{68}{63}\right),\ y = -\frac{46}{63}\right)$,
- NOT Nash equilibrium: $\left(\mathbf{x} = \left(\frac{5}{8}, 1\right),\ y = -\frac{3}{4}\right)$.

Our idea

Potential Function: For all $\mathbf{z} = (\mathbf{x}, \mathbf{y}) \in X \times Y$,

$$\Delta(\mathbf{z}) \stackrel{\text{def}}{=} \max_{\widetilde{\mathbf{z}}=(\widetilde{\mathbf{x}},\widetilde{\mathbf{y}})\in X\times Y} \underbrace{g(\mathbf{z}) - g(\widetilde{\mathbf{z}})}_{\text{jointly convex coupling}} + \underbrace{h(\mathbf{x}, \widetilde{\mathbf{y}}) - h(\widetilde{\mathbf{x}}, \mathbf{y})}_{\text{convex-concave zero-sum}} .$$

Proposition 4

For all $\mathbf{z} = (\mathbf{x}, \mathbf{y}) \in X \times Y$, we have $\Delta(\mathbf{z}) \geq 0$ and

$$2\Delta(\mathbf{z}) \geq \max_{\widetilde{\mathbf{z}}=(\widetilde{\mathbf{x}},\widetilde{\mathbf{y}})\in X\times Y} u_1(\widetilde{\mathbf{x}}, \mathbf{y}) - u_1(\mathbf{x}, \mathbf{y}) + u_2(\mathbf{x}, \widetilde{\mathbf{y}}) - u_2(\mathbf{x}, \mathbf{y}) .$$

Proposition 5

Let $\mathbf{z}^* \in X \times Y$. Then, $\mathbf{z}^*$ is the Nash equilibrium if and only if $\Delta(\mathbf{z}^*) = 0$.

Our idea

Potential Function: For all $\mathbf{z} = (\mathbf{x}, \mathbf{y}) \in X \times Y$,

$$\Delta(\mathbf{z}) \stackrel{\text{def}}{=} \max_{\widetilde{\mathbf{z}}=(\widetilde{\mathbf{x}},\widetilde{\mathbf{y}})\in X\times Y} \underbrace{g(\mathbf{z}) - g(\widetilde{\mathbf{z}})}_{\text{jointly convex coupling}} + \underbrace{h(\mathbf{x}, \widetilde{\mathbf{y}}) - h(\widetilde{\mathbf{x}}, \mathbf{y})}_{\text{convex-concave zero-sum}} .$$

Proposition 4

For all $\mathbf{z} = (\mathbf{x}, \mathbf{y}) \in X \times Y$, *we have* $\Delta(\mathbf{z}) \geq 0$ *and*

$$2\Delta(\mathbf{z}) \geq \max_{\widetilde{\mathbf{z}}=(\widetilde{\mathbf{x}},\widetilde{\mathbf{y}})\in X\times Y} u_1(\widetilde{\mathbf{x}}, \mathbf{y}) - u_1(\mathbf{x}, \mathbf{y}) + u_2(\mathbf{x}, \widetilde{\mathbf{y}}) - u_2(\mathbf{x}, \mathbf{y}) .$$

Proposition 5

Let $\mathbf{z}^* \in X \times Y$. *Then,* $\mathbf{z}^*$ *is the Nash equilibrium if and only if* $\Delta(\mathbf{z}^*) = 0$.

Our idea

Potential Function: For all $\mathbf{z} = (\mathbf{x}, \mathbf{y}) \in X \times Y$,

$$\color{blue}{\Delta}\color{black}(\mathbf{z}) \stackrel{\text{def}}{=} \max_{\widetilde{\mathbf{z}}=(\widetilde{\mathbf{x}},\widetilde{\mathbf{y}})\in X\times Y} \quad \underbrace{g(\mathbf{z}) - g(\widetilde{\mathbf{z}})}_{\text{jointly convex coupling}} \quad + \quad \underbrace{h(\mathbf{x}, \widetilde{\mathbf{y}}) - h(\widetilde{\mathbf{x}}, \mathbf{y})}_{\text{convex-concave zero-sum}} \quad .$$

Proposition 4

For all $\mathbf{z} = (\mathbf{x}, \mathbf{y}) \in X \times Y$*, we have* $\color{blue}{\Delta}\color{black}(\mathbf{z}) \geq 0$ *and*

$$2\color{blue}{\Delta}\color{black}(\mathbf{z}) \geq \max_{\widetilde{\mathbf{z}}=(\widetilde{\mathbf{x}},\widetilde{\mathbf{y}})\in X\times Y} \quad u_1(\widetilde{\mathbf{x}}, \mathbf{y}) - u_1(\mathbf{x}, \mathbf{y}) + u_2(\mathbf{x}, \widetilde{\mathbf{y}}) - u_2(\mathbf{x}, \mathbf{y})\,.$$

Proposition 5

Let $\mathbf{z}^* \in X \times Y$*. Then,* $\mathbf{z}^*$ *is the Nash equilibrium if and only if* $\color{blue}{\Delta}\color{black}(\mathbf{z}^*) = 0$*.*

Our algorithm

Algorithm 1 Iterative Coupling Linearization (ICL)

Require: $\mathbf{x}_0 \in X$, $\mathbf{y}_0 \in Y$.

1: **for** $t = 0, 1, \cdots, T-1$ **do**

2: Let

$$\varphi_t(\mathbf{x}, \mathbf{y}) \stackrel{\text{def}}{=} \langle \nabla_{\mathbf{x}} g(\mathbf{x}_t, \mathbf{y}_t), \mathbf{x} \rangle + \frac{1}{2\eta_t} \|\mathbf{x}_t - \mathbf{x}\|^2 + h(\mathbf{x}, \mathbf{y}) \\ - \langle \nabla_{\mathbf{y}} g(\mathbf{x}_t, \mathbf{y}_t), \mathbf{y} \rangle - \frac{1}{2\eta_t} \|\mathbf{y}_t - \mathbf{y}\|^2 .$$

3: Find an **inexact saddle point** $\mathbf{z}_{t+1} \in X \times Y$ of φ_t s.t.

$$\langle \nabla_{\mathbf{x}} \varphi_t(\mathbf{z}_{t+1}), \mathbf{x}_{t+1} - \mathbf{x} \rangle - \langle \nabla_{\mathbf{y}} \varphi_t(\mathbf{z}_{t+1}), \mathbf{y}_{t+1} - \mathbf{y} \rangle \leq \varepsilon_t,$$

for all $\mathbf{x} \in X$ and $\mathbf{y} \in Y$.

4: **end for**

Our algorithm

Algorithm 2 Iterative Coupling Linearization (ICL)

Require: $\mathbf{x}_0 \in X$, $\mathbf{y}_0 \in Y$.

1: **for** $t = 0, 1, \cdots, T-1$ **do**

2: Let

$$\varphi_t(\mathbf{x}, \mathbf{y}) \stackrel{\text{def}}{=} \langle \nabla_{\mathbf{x}} g(\mathbf{x}_t, \mathbf{y}_t), \mathbf{x} \rangle + \frac{1}{2\eta_t} \|\mathbf{x}_t - \mathbf{x}\|^2 + h(\mathbf{x}, \mathbf{y}) \\ - \langle \nabla_{\mathbf{y}} g(\mathbf{x}_t, \mathbf{y}_t), \mathbf{y} \rangle - \frac{1}{2\eta_t} \|\mathbf{y}_t - \mathbf{y}\|^2 .$$

3: Find an **inexact saddle point** $\mathbf{z}_{t+1} \in X \times Y$ of φ_t s.t.

$$\langle \nabla_{\mathbf{x}} \varphi_t(\mathbf{z}_{t+1}), \mathbf{x}_{t+1} - \mathbf{x} \rangle - \langle \nabla_{\mathbf{y}} \varphi_t(\mathbf{z}_{t+1}), \mathbf{y}_{t+1} - \mathbf{y} \rangle \le \varepsilon_t,$$

for all $\mathbf{x} \in X$ and $\mathbf{y} \in Y$.

4: **end for**

Convergence analysis - descent lemma

Lemma 1 (Descent lemma)

In monotone δ-near-zero-sum games, for $\eta_t \leq \frac{1}{\delta}$, we have

$$\left(\frac{1}{2\eta_t} + \frac{\min\{\mu, \nu\}}{2}\right)\left\|\mathbf{z}_{t+1} - \mathbf{z}^*\right\|^2 \leq \frac{1}{2\eta_t}\left\|\mathbf{z}_t - \mathbf{z}^*\right\|^2 + \varepsilon_t \, .$$

Proof Sketch.

$$\begin{aligned}
0 = -\Delta(\mathbf{z}^*) &\leq g(\mathbf{z}_{t+1}) - g(\mathbf{z}^*) + h(\mathbf{x}_{t+1}, \mathbf{y}^*) - h(\mathbf{x}^*, \mathbf{y}_{t+1}) \\
&\leq \left\langle \nabla g(\mathbf{z}_t) + \mathcal{H}(\mathbf{z}_{t+1}) + \frac{1}{\eta_t}(\mathbf{z}_{t+1} - \mathbf{z}_t), \mathbf{z}_{t+1} - \mathbf{z}^* \right\rangle - \frac{1}{\eta_t}\langle \mathbf{z}_{t+1} - \mathbf{z}_t, \mathbf{z}_{t+1} - \mathbf{z}^* \rangle \\
&\quad - \frac{\mu}{2}\left\|\mathbf{x}_{t+1} - \mathbf{x}^*\right\|^2 - \frac{\nu}{2}\left\|\mathbf{y}_{t+1} - \mathbf{y}^*\right\|^2 + \frac{\delta}{2}\left\|\mathbf{z}_{t+1} - \mathbf{z}_t\right\|^2 \\
&\leq \varepsilon_t + \frac{1}{2\eta_t}\left\|\mathbf{x}_t - \mathbf{x}^*\right\|^2 - \left(\frac{1}{2\eta_t} + \frac{\mu}{2}\right)\left\|\mathbf{x}_{t+1} - \mathbf{x}^*\right\|^2 \\
&\quad + \frac{1}{2\eta_t}\left\|\mathbf{y}_t - \mathbf{y}^*\right\|^2 - \left(\frac{1}{2\eta_t} + \frac{\nu}{2}\right)\left\|\mathbf{y}_{t+1} - \mathbf{y}^*\right\|^2 - \left(\frac{1}{2\eta_t} - \frac{\delta}{2}\right)\left\|\mathbf{z}_{t+1} - \mathbf{z}_t\right\|^2 \, .
\end{aligned}$$

Convergence analysis - descent lemma

Lemma 1 (Descent lemma)

In monotone δ-near-zero-sum games, for $\eta_t \leq \frac{1}{\delta}$, we have

$$\left(\frac{1}{2\eta_t} + \frac{\min\{\mu, \nu\}}{2}\right) \|\mathbf{z}_{t+1} - \mathbf{z}^*\|^2 \leq \frac{1}{2\eta_t} \|\mathbf{z}_t - \mathbf{z}^*\|^2 + \varepsilon_t \,.$$

Proof Sketch.

$$\begin{aligned}
0 = -\Delta(\mathbf{z}^*) &\leq g(\mathbf{z}_{t+1}) - g(\mathbf{z}^*) + h(\mathbf{x}_{t+1}, \mathbf{y}^*) - h(\mathbf{x}^*, \mathbf{y}_{t+1}) \\
&\leq \left\langle \nabla g(\mathbf{z}_t) + \mathcal{H}(\mathbf{z}_{t+1}) + \frac{1}{\eta_t}(\mathbf{z}_{t+1} - \mathbf{z}_t), \mathbf{z}_{t+1} - \mathbf{z}^* \right\rangle - \frac{1}{\eta_t} \langle \mathbf{z}_{t+1} - \mathbf{z}_t, \mathbf{z}_{t+1} - \mathbf{z}^* \rangle \\
&\quad - \frac{\mu}{2} \|\mathbf{x}_{t+1} - \mathbf{x}^*\|^2 - \frac{\nu}{2} \|\mathbf{y}_{t+1} - \mathbf{y}^*\|^2 + \frac{\delta}{2} \|\mathbf{z}_{t+1} - \mathbf{z}_t\|^2 \\
&\leq \varepsilon_t + \frac{1}{2\eta_t} \|\mathbf{x}_t - \mathbf{x}^*\|^2 - \left(\frac{1}{2\eta_t} + \frac{\mu}{2}\right) \|\mathbf{x}_{t+1} - \mathbf{x}^*\|^2 \\
&\quad + \frac{1}{2\eta_t} \|\mathbf{y}_t - \mathbf{y}^*\|^2 - \left(\frac{1}{2\eta_t} + \frac{\nu}{2}\right) \|\mathbf{y}_{t+1} - \mathbf{y}^*\|^2 - \left(\frac{1}{2\eta_t} - \frac{\delta}{2}\right) \|\mathbf{z}_{t+1} - \mathbf{z}_t\|^2 \,.
\end{aligned}$$

Convergence analysis - outer and inner loops

Lemma 2 (Outer loop)

Let $\eta_t = \eta \in (0, \frac{1}{\delta}]$, *for all* $t \in [0, T-1] \cap \mathbb{Z}$. *Denote* $\theta = \frac{\min\{\mu,\nu\}}{\eta^{-1}+\min\{\mu,\nu\}}$. *Suppose* $\varepsilon_t \leq \frac{\theta\varepsilon}{4\eta}$, *for all* $t \in [0, T-1] \cap \mathbb{Z}$. *In monotone* δ*-near-zero-sum games, if* $T \geq \frac{1}{\theta} \log \frac{2(D_X^2+D_Y^2)}{\varepsilon}$, *then* $\|\mathbf{z}_T - \mathbf{z}^*\|^2 \leq \varepsilon$.

Lemma 3 (Inner loop[9])

Under Assumption 1, at each iteration $t \in [0, T-1] \cap \mathbb{Z}$, *for* $\eta_t \geq \frac{1}{L}$, *the inexact solution* $(\mathbf{x}_{t+1}, \mathbf{y}_{t+1})$ *in Algorithm 1 can be found with a gradient complexity of*

$$\mathcal{O}\left(\frac{L}{\sqrt{\left(\eta_t^{-1}+\mu\right)\left(\eta_t^{-1}+\nu\right)}} \cdot \log\left(\frac{L\left(D_X^2+D_Y^2\right)}{\varepsilon_t}\right)\right).$$

[9] Kovalev and Gasnikov, "The first optimal algorithm for smooth and strongly-convex-strongly-concave minimax optimization"; Lan and Li, "A Novel Catalyst Scheme for Stochastic Minimax Optimization".

Convergence analysis - outer and inner loops

Lemma 2 (Outer loop)

Let $\eta_t = \eta \in (0, \frac{1}{\delta}]$, for all $t \in [0, T-1] \cap \mathbb{Z}$. Denote $\theta = \frac{\min\{\mu,\nu\}}{\eta^{-1}+\min\{\mu,\nu\}}$. Suppose $\varepsilon_t \leq \frac{\theta\varepsilon}{4\eta}$, for all $t \in [0, T-1] \cap \mathbb{Z}$. In monotone δ-near-zero-sum games, if $T \geq \frac{1}{\theta} \log \frac{2(D_X^2+D_Y^2)}{\varepsilon}$, then $\|\mathbf{z}_T - \mathbf{z}^*\|^2 \leq \varepsilon$.

Lemma 3 (Inner loop[9])

Under Assumption 1, at each iteration $t \in [0, T-1] \cap \mathbb{Z}$, for $\eta_t \geq \frac{1}{L}$, the inexact solution $(\mathbf{x}_{t+1}, \mathbf{y}_{t+1})$ in Algorithm 1 can be found with a gradient complexity of

$$\mathcal{O}\left(\frac{L}{\sqrt{\left(\eta_t^{-1}+\mu\right)\left(\eta_t^{-1}+\nu\right)}} \cdot \log\left(\frac{L\left(D_X^2+D_Y^2\right)}{\varepsilon_t}\right)\right).$$

[9] Kovalev and Gasnikov, "The first optimal algorithm for smooth and strongly-convex-strongly-concave minimax optimization"; Lan and Li, "A Novel Catalyst Scheme for Stochastic Minimax Optimization".

Convergence analysis - total gradient complexity

Theorem 1 (Main result)

Denote $\eta = \min\left\{\frac{1}{\delta}, \frac{1}{\min\{\mu,\nu\}}\right\}$ *and* $\theta = \frac{\min\{\mu,\nu\}}{\eta^{-1}+\min\{\mu,\nu\}}$. *Let* $\eta_t = \eta$ *and* $\varepsilon_t = \frac{\theta\varepsilon}{4\eta}$, *for all* $t \in [0, T-1] \cap \mathbb{Z}$. *In monotone* δ*-near-zero-sum games, Algorithm 1 obtains an* ε*-accurate Nash equilibrium with a gradient complexity of*

$$\widetilde{\mathcal{O}}\left(\left(\frac{L}{\sqrt{\mu\nu}} + \frac{L}{\min\{\mu,\nu\}} \cdot \min\left\{1, \sqrt{\frac{\delta}{\mu+\nu}}\right\}\right) \cdot \log^2\left(\frac{1}{\varepsilon}\right)\right).$$

Proof.

Multiply outer loop iterations and inner loop gradient complexity. □

Remark (Acceleration conditioning)

$$\min\{\mu,\nu\} + \delta \ll \mu + \nu\,.$$

Convergence analysis - total gradient complexity

Theorem 1 (Main result)

Denote $\eta = \min\left\{\frac{1}{\delta}, \frac{1}{\min\{\mu,\nu\}}\right\}$ *and* $\theta = \frac{\min\{\mu,\nu\}}{\eta^{-1}+\min\{\mu,\nu\}}$. *Let* $\eta_t = \eta$ *and* $\varepsilon_t = \frac{\theta\varepsilon}{4\eta}$, *for all* $t \in [0, T-1] \cap \mathbb{Z}$. *In monotone* δ*-near-zero-sum games, Algorithm 1 obtains an* ε*-accurate Nash equilibrium with a gradient complexity of*

$$\widetilde{\mathcal{O}}\left(\left(\frac{L}{\sqrt{\mu\nu}} + \frac{L}{\min\{\mu,\nu\}} \cdot \min\left\{1, \sqrt{\frac{\delta}{\mu+\nu}}\right\}\right) \cdot \log^2\left(\frac{1}{\varepsilon}\right)\right).$$

Proof.

Multiply outer loop iterations and inner loop gradient complexity. □

Remark (Acceleration conditioning)

$$\min\{\mu,\nu\} + \delta \ll \mu + \nu\,.$$

Convergence analysis - total gradient complexity

Theorem 1 (Main result)

Denote $\eta = \min\left\{\frac{1}{\delta}, \frac{1}{\min\{\mu,\nu\}}\right\}$ *and* $\theta = \frac{\min\{\mu,\nu\}}{\eta^{-1}+\min\{\mu,\nu\}}$. *Let* $\eta_t = \eta$ *and* $\varepsilon_t = \frac{\theta\varepsilon}{4\eta}$, *for all* $t \in [0, T-1] \cap \mathbb{Z}$. *In monotone* δ*-near-zero-sum games, Algorithm 1 obtains an* ε*-accurate Nash equilibrium with a gradient complexity of*

$$\widetilde{\mathcal{O}}\left(\left(\frac{L}{\sqrt{\mu\nu}} + \frac{L}{\min\{\mu,\nu\}} \cdot \min\left\{1, \sqrt{\frac{\delta}{\mu+\nu}}\right\}\right) \cdot \log^2\left(\frac{1}{\varepsilon}\right)\right).$$

Proof.

Multiply outer loop iterations and inner loop gradient complexity. □

Remark (Acceleration conditioning)

$$\min\{\mu,\nu\} + \delta \ll \mu + \nu\,.$$

Outline for section 4

Application 1: Regularized matrix games

- Player 1 maximizes $u_1 = \langle \mathbf{Ax}, \mathbf{y} \rangle + \mathcal{R}(\mathbf{x}, \mathbf{y})$ over $\mathbf{x} \in X$;
  Player 2 maximizes $u_2 = \langle \mathbf{Bx}, \mathbf{y} \rangle - \mathcal{R}(\mathbf{x}, \mathbf{y})$ over $\mathbf{y} \in Y$.
- $\|\mathbf{A}\| \le L$, $\|\mathbf{B}\| \le L$; $\left\| \frac{\mathbf{A}+\mathbf{B}}{2} \right\| \le \beta$;
- $\mathcal{R}$ is L-smooth and μ-strongly concave-ν-strongly convex.

- **Examples:** transaction fee, tax rates.

- **Assume:** $\beta \le \frac{1}{2}\sqrt{\mu\nu}$.
- The game is $\min\{\frac{\mu}{2}, \frac{\nu}{2}\}$-strongly monotone.
- Variational inequality methods: $\widetilde{\mathcal{O}}\left(\frac{L}{\min\{\mu,\nu\}} \cdot \log\left(\frac{1}{\varepsilon}\right)\right)$.[10]

- Can we directly apply our ICL here?
  No! $g = -\left\langle \frac{\mathbf{A}+\mathbf{B}}{2}\mathbf{x}, \mathbf{y} \right\rangle$ is non-convex.

[10] Tseng, "On linear convergence of iterative methods for the variational inequality problem".

Application 1: Regularized matrix games

- Player 1 maximizes $u_1 = \langle \mathbf{A}\mathbf{x}, \mathbf{y} \rangle + \mathcal{R}(\mathbf{x}, \mathbf{y})$ over $\mathbf{x} \in X$;
  Player 2 maximizes $u_2 = \langle \mathbf{B}\mathbf{x}, \mathbf{y} \rangle - \mathcal{R}(\mathbf{x}, \mathbf{y})$ over $\mathbf{y} \in Y$.
- $\|\mathbf{A}\| \leq L$, $\|\mathbf{B}\| \leq L$; $\left\| \frac{\mathbf{A}+\mathbf{B}}{2} \right\| \leq \beta$;
- $\mathcal{R}$ is L-smooth and μ-strongly concave-ν-strongly convex.
- **Examples:** transaction fee, tax rates.
- **Assume:** $\beta \leq \frac{1}{2}\sqrt{\mu\nu}$.
- The game is $\min\{\frac{\mu}{2}, \frac{\nu}{2}\}$-strongly monotone.
- Variational inequality methods: $\widetilde{\mathcal{O}}\left(\frac{L}{\min\{\mu,\nu\}} \cdot \log\left(\frac{1}{\varepsilon}\right)\right)$.[10]
- Can we directly apply our ICL here?
  No! $g = -\left\langle \frac{\mathbf{A}+\mathbf{B}}{2}\mathbf{x}, \mathbf{y} \right\rangle$ is non-convex.

[10] Tseng, "On linear convergence of iterative methods for the variational inequality problem".

Application 1: Regularized matrix games

- Player 1 maximizes $u_1 = \langle \mathbf{Ax}, \mathbf{y} \rangle + \mathcal{R}(\mathbf{x}, \mathbf{y})$ over $\mathbf{x} \in X$;
  Player 2 maximizes $u_2 = \langle \mathbf{Bx}, \mathbf{y} \rangle - \mathcal{R}(\mathbf{x}, \mathbf{y})$ over $\mathbf{y} \in Y$.
- $\|\mathbf{A}\| \leq L$, $\|\mathbf{B}\| \leq L$; $\left\| \frac{\mathbf{A}+\mathbf{B}}{2} \right\| \leq \beta$;
- $\mathcal{R}$ is L-smooth and μ-strongly concave-ν-strongly convex.

- **Examples:** transaction fee, tax rates.

- **Assume:** $\beta \leq \frac{1}{2}\sqrt{\mu\nu}$.
- The game is $\min\{\frac{\mu}{2}, \frac{\nu}{2}\}$-strongly monotone.
- Variational inequality methods: $\widetilde{\mathcal{O}}\left(\frac{L}{\min\{\mu,\nu\}} \cdot \log\left(\frac{1}{\varepsilon}\right)\right)$.[10]

- Can we directly apply our ICL here?
  No! $g = -\left\langle \frac{\mathbf{A}+\mathbf{B}}{2}\mathbf{x}, \mathbf{y} \right\rangle$ is non-convex.

[10] Tseng, "On linear convergence of iterative methods for the variational inequality problem".

Application 1: Regularized matrix games

- Player 1 maximizes $u_1 = \langle \mathbf{A}\mathbf{x}, \mathbf{y} \rangle + \mathcal{R}(\mathbf{x}, \mathbf{y})$ over $\mathbf{x} \in X$;
  Player 2 maximizes $u_2 = \langle \mathbf{B}\mathbf{x}, \mathbf{y} \rangle - \mathcal{R}(\mathbf{x}, \mathbf{y})$ over $\mathbf{y} \in Y$.
- $\|\mathbf{A}\| \leq L$, $\|\mathbf{B}\| \leq L$; $\left\| \frac{\mathbf{A}+\mathbf{B}}{2} \right\| \leq \beta$;
- $\mathcal{R}$ is L-smooth and μ-strongly concave-ν-strongly convex.

- **Examples:** transaction fee, tax rates.

- **Assume:** $\beta \leq \frac{1}{2}\sqrt{\mu\nu}$.
- The game is $\min\{\frac{\mu}{2}, \frac{\nu}{2}\}$-strongly monotone.
- Variational inequality methods: $\widetilde{\mathcal{O}}\left(\frac{L}{\min\{\mu,\nu\}} \cdot \log\left(\frac{1}{\varepsilon}\right)\right)$.[10]

- Can we directly apply our ICL here?
  No! $g = -\left\langle \frac{\mathbf{A}+\mathbf{B}}{2}\mathbf{x}, \mathbf{y} \right\rangle$ is non-convex.

[10] Tseng, "On linear convergence of iterative methods for the variational inequality problem".

Application 1: Regularized matrix games

- Player 1 maximizes $u_1 = \langle \mathbf{A}\mathbf{x}, \mathbf{y} \rangle + \mathcal{R}(\mathbf{x}, \mathbf{y})$ over $\mathbf{x} \in X$;
  Player 2 maximizes $u_2 = \langle \mathbf{B}\mathbf{x}, \mathbf{y} \rangle - \mathcal{R}(\mathbf{x}, \mathbf{y})$ over $\mathbf{y} \in Y$.
- $\|\mathbf{A}\| \le L$, $\|\mathbf{B}\| \le L$; $\left\| \frac{\mathbf{A}+\mathbf{B}}{2} \right\| \le \beta$;
- $\mathcal{R}$ is L-smooth and μ-strongly concave-ν-strongly convex.

- **Examples:** transaction fee, tax rates.

- **Assume:** $\beta \le \frac{1}{2}\sqrt{\mu\nu}$.
- The game is $\min\{\frac{\mu}{2}, \frac{\nu}{2}\}$-strongly monotone.
- Variational inequality methods: $\widetilde{\mathcal{O}}\left(\frac{L}{\min\{\mu,\nu\}} \cdot \log\left(\frac{1}{\varepsilon}\right)\right)$.[10]

- Can we directly apply our ICL here?
  No! $g = -\left\langle \frac{\mathbf{A}+\mathbf{B}}{2}\mathbf{x}, \mathbf{y} \right\rangle$ is non-convex.

[10] Tseng, "On linear convergence of iterative methods for the variational inequality problem".

Application 1: Regularized matrix games

- Player 1 maximizes $u_1 = \langle \mathbf{Ax}, \mathbf{y} \rangle + \mathcal{R}(\mathbf{x}, \mathbf{y})$ over $\mathbf{x} \in X$;
  Player 2 maximizes $u_2 = \langle \mathbf{Bx}, \mathbf{y} \rangle - \mathcal{R}(\mathbf{x}, \mathbf{y})$ over $\mathbf{y} \in Y$.
- $\|\mathbf{A}\| \le L$, $\|\mathbf{B}\| \le L$; $\left\| \frac{\mathbf{A}+\mathbf{B}}{2} \right\| \le \beta$;
- $\mathcal{R}$ is L-smooth and μ-strongly concave-ν-strongly convex.

- **Examples:** transaction fee, tax rates.

- **Assume:** $\beta \le \frac{1}{2}\sqrt{\mu\nu}$.
- The game is $\min\{\frac{\mu}{2}, \frac{\nu}{2}\}$-strongly monotone.
- Variational inequality methods: $\widetilde{\mathcal{O}}\left(\frac{L}{\min\{\mu,\nu\}} \cdot \log\left(\frac{1}{\varepsilon}\right)\right)$.[10]

- Can we directly apply our ICL here?
  No! $g = -\left\langle \frac{\mathbf{A}+\mathbf{B}}{2}\mathbf{x}, \mathbf{y} \right\rangle$ is non-convex.

[10] Tseng, "On linear convergence of iterative methods for the variational inequality problem".

Application 1: Regularized matrix games

- Player 1 maximizes $u_1 = \langle \mathbf{Ax}, \mathbf{y} \rangle + \mathcal{R}(\mathbf{x}, \mathbf{y})$ over $\mathbf{x} \in X$;
  Player 2 maximizes $u_2 = \langle \mathbf{Bx}, \mathbf{y} \rangle - \mathcal{R}(\mathbf{x}, \mathbf{y})$ over $\mathbf{y} \in Y$.
- $\|\mathbf{A}\| \le L$, $\|\mathbf{B}\| \le L$; $\left\| \frac{\mathbf{A}+\mathbf{B}}{2} \right\| \le \beta$;
- $\mathcal{R}$ is L-smooth and μ-strongly concave-ν-strongly convex.

- **Examples:** transaction fee, tax rates.

- **Assume:** $\beta \le \frac{1}{2}\sqrt{\mu\nu}$.
- The game is $\min\{\frac{\mu}{2}, \frac{\nu}{2}\}$-strongly monotone.
- Variational inequality methods: $\widetilde{\mathcal{O}}\left(\frac{L}{\min\{\mu,\nu\}} \cdot \log\left(\frac{1}{\varepsilon}\right)\right)$.[10]

- Can we directly apply our ICL here?
  No! $g = -\left\langle \frac{\mathbf{A}+\mathbf{B}}{2}\mathbf{x}, \mathbf{y} \right\rangle$ is non-convex.

[10] Tseng, "On linear convergence of iterative methods for the variational inequality problem".

Application 1: Regularized matrix games

- Player 1 maximizes $u_1 = \langle \mathbf{A}\mathbf{x}, \mathbf{y} \rangle + \mathcal{R}(\mathbf{x}, \mathbf{y})$ over $\mathbf{x} \in X$;
  Player 2 maximizes $u_2 = \langle \mathbf{B}\mathbf{x}, \mathbf{y} \rangle - \mathcal{R}(\mathbf{x}, \mathbf{y})$ over $\mathbf{y} \in Y$.
- $\|\mathbf{A}\| \leq L$, $\|\mathbf{B}\| \leq L$; $\left\| \frac{\mathbf{A}+\mathbf{B}}{2} \right\| \leq \beta$;
- $\mathcal{R}$ is L-smooth and μ-strongly concave-ν-strongly convex.

- **Examples:** transaction fee, tax rates.

- **Assume:** $\beta \leq \frac{1}{2}\sqrt{\mu\nu}$.
- The game is $\min\{\frac{\mu}{2}, \frac{\nu}{2}\}$-strongly monotone.
- Variational inequality methods: $\widetilde{\mathcal{O}}\left(\frac{L}{\min\{\mu,\nu\}} \cdot \log\left(\frac{1}{\varepsilon}\right)\right)$.[10]

- Can we directly apply our ICL here?
  No! $g = -\left\langle \frac{\mathbf{A}+\mathbf{B}}{2}\mathbf{x}, \mathbf{y} \right\rangle$ is non-convex.

[10] Tseng, "On linear convergence of iterative methods for the variational inequality problem".

Application 1: Regularized matrix games

- Player 1 maximizes $u_1 = \langle \mathbf{A}\mathbf{x}, \mathbf{y} \rangle + \mathcal{R}(\mathbf{x}, \mathbf{y})$ over $\mathbf{x} \in X$;
  Player 2 maximizes $u_2 = \langle \mathbf{B}\mathbf{x}, \mathbf{y} \rangle - \mathcal{R}(\mathbf{x}, \mathbf{y})$ over $\mathbf{y} \in Y$.
- $\|\mathbf{A}\| \leq L$, $\|\mathbf{B}\| \leq L$; $\left\| \frac{\mathbf{A}+\mathbf{B}}{2} \right\| \leq \beta$;
- $\mathcal{R}$ is L-smooth and μ-strongly concave-ν-strongly convex.

- **Examples:** transaction fee, tax rates.

- **Assume:** $\beta \leq \frac{1}{2}\sqrt{\mu\nu}$.
- The game is $\min\{\frac{\mu}{2}, \frac{\nu}{2}\}$-strongly monotone.
- Variational inequality methods: $\widetilde{\mathcal{O}}\left(\frac{L}{\min\{\mu,\nu\}} \cdot \log\left(\frac{1}{\varepsilon}\right)\right)$.[10]

- Can we directly apply our ICL here?
  No! $g = -\left\langle \frac{\mathbf{A}+\mathbf{B}}{2}\mathbf{x}, \mathbf{y} \right\rangle$ is non-convex.

[10] Tseng, "On linear convergence of iterative methods for the variational inequality problem".

Application 1: Regularized matrix games

- Player 1 maximizes $u_1 = \langle \mathbf{A}\mathbf{x}, \mathbf{y} \rangle + \mathcal{R}(\mathbf{x}, \mathbf{y})$ over $\mathbf{x} \in X$;
  Player 2 maximizes $u_2 = \langle \mathbf{B}\mathbf{x}, \mathbf{y} \rangle - \mathcal{R}(\mathbf{x}, \mathbf{y})$ over $\mathbf{y} \in Y$.
- $\|\mathbf{A}\| \le L$, $\|\mathbf{B}\| \le L$; $\left\| \frac{\mathbf{A}+\mathbf{B}}{2} \right\| \le \beta$;
- $\mathcal{R}$ is L-smooth and μ-strongly concave-ν-strongly convex.

- **Examples:** transaction fee, tax rates.

- **Assume:** $\beta \le \frac{1}{2}\sqrt{\mu\nu}$.
- The game is $\min\{\frac{\mu}{2}, \frac{\nu}{2}\}$-strongly monotone.
- Variational inequality methods: $\widetilde{\mathcal{O}}\left(\frac{L}{\min\{\mu,\nu\}} \cdot \log\left(\frac{1}{\varepsilon}\right)\right)$.[10]

- Can we directly apply our ICL here?
  No! $g = -\left\langle \frac{\mathbf{A}+\mathbf{B}}{2}\mathbf{x}, \mathbf{y} \right\rangle$ is non-convex.

[10] Tseng, "On linear convergence of iterative methods for the variational inequality problem".

- **Convex Reformulation**:
  Player 1: $\max_{\mathbf{x}\in X} \widetilde{u}_1(\mathbf{x},\mathbf{y}) = u_1(\mathbf{x},\mathbf{y}) - \beta_2 \|\mathbf{y}\|^2$;
  Player 2: $\max_{\mathbf{y}\in Y} \widetilde{u}_2(\mathbf{x},\mathbf{y}) = u_2(\mathbf{x},\mathbf{y}) - \beta_1 \|\mathbf{x}\|^2$.
  $\widetilde{g}(\mathbf{x},\mathbf{y}) = \frac{\beta_1}{2}\|\mathbf{x}\|^2 - \left\langle \left(\frac{\mathbf{A}+\mathbf{B}}{2}\right)\mathbf{x}, \mathbf{y}\right\rangle + \frac{\beta_2}{2}\|\mathbf{y}\|^2$.
- **Choices of β_1, β_2:**
  If $\beta \le \min\{\frac{\mu}{2}, \frac{\nu}{2}\}$, let $\beta_1 = \beta_2 = \beta$;
  If $\frac{\mu}{2} \le \beta \le \frac{\nu}{2}$, let $\beta_1 = \frac{\mu}{2}$ and $\beta_2 = \frac{2\beta^2}{\mu}$;
  If $\frac{\nu}{2} \le \beta \le \frac{\mu}{2}$, let $\beta_1 = \frac{2\beta^2}{\nu}$ and $\beta_2 = \frac{\nu}{2}$.
- We have $\beta_1 \le \frac{\mu}{2}$, $\beta_2 \le \frac{\nu}{2}$, and $\sqrt{\beta_1\beta_2} = \beta$.
  Therefore, $\widetilde{g}$ is convex and $(\beta_1+\beta+\beta_2)$-smooth.
- **Our method:**

$$\widetilde{\mathcal{O}}\left(\left(\frac{L}{\sqrt{\mu\nu}} + \frac{L}{\min\{\mu,\nu\}}\cdot\frac{\beta}{\sqrt{\mu\nu}}\right)\cdot\log^2\left(\frac{1}{\varepsilon}\right)\right).$$

Acceleration conditioning: $\min\{\mu,\nu\} + \beta \ll \sqrt{\mu\nu}$.

- **Convex Reformulation**:
  Player 1: $\max_{\mathbf{x}\in X} \widetilde{u}_1(\mathbf{x},\mathbf{y}) = u_1(\mathbf{x},\mathbf{y}) - \beta_2 \|\mathbf{y}\|^2$;
  Player 2: $\max_{\mathbf{y}\in Y} \widetilde{u}_2(\mathbf{x},\mathbf{y}) = u_2(\mathbf{x},\mathbf{y}) - \beta_1 \|\mathbf{x}\|^2$.
  $\widetilde{g}(\mathbf{x},\mathbf{y}) = \frac{\beta_1}{2}\|\mathbf{x}\|^2 - \left\langle \left(\frac{\mathbf{A}+\mathbf{B}}{2}\right)\mathbf{x},\mathbf{y}\right\rangle + \frac{\beta_2}{2}\|\mathbf{y}\|^2$.
- **Choices of β_1, β_2:**
  If $\beta \le \min\{\frac{\mu}{2},\frac{\nu}{2}\}$, let $\beta_1 = \beta_2 = \beta$;
  If $\frac{\mu}{2} \le \beta \le \frac{\nu}{2}$, let $\beta_1 = \frac{\mu}{2}$ and $\beta_2 = \frac{2\beta^2}{\mu}$;
  If $\frac{\nu}{2} \le \beta \le \frac{\mu}{2}$, let $\beta_1 = \frac{2\beta^2}{\nu}$ and $\beta_2 = \frac{\nu}{2}$.
- We have $\beta_1 \le \frac{\mu}{2}$, $\beta_2 \le \frac{\nu}{2}$, and $\sqrt{\beta_1\beta_2} = \beta$.
  Therefore, $\widetilde{g}$ is convex and $(\beta_1+\beta+\beta_2)$-smooth.
- **Our method:**

$$\widetilde{\mathcal{O}}\left(\left(\frac{L}{\sqrt{\mu\nu}} + \frac{L}{\min\{\mu,\nu\}}\cdot\frac{\beta}{\sqrt{\mu\nu}}\right)\cdot\log^2\left(\frac{1}{\varepsilon}\right)\right).$$

Acceleration conditioning: $\min\{\mu,\nu\} + \beta \ll \sqrt{\mu\nu}$.

- **Convex Reformulation**:
  Player 1: $\max_{\mathbf{x}\in X} \widetilde{u}_1(\mathbf{x},\mathbf{y}) = u_1(\mathbf{x},\mathbf{y}) - \beta_2 \|\mathbf{y}\|^2$;
  Player 2: $\max_{\mathbf{y}\in Y} \widetilde{u}_2(\mathbf{x},\mathbf{y}) = u_2(\mathbf{x},\mathbf{y}) - \beta_1 \|\mathbf{x}\|^2$.
  $\widetilde{g}(\mathbf{x},\mathbf{y}) = \frac{\beta_1}{2}\|\mathbf{x}\|^2 - \left\langle \left(\frac{\mathbf{A}+\mathbf{B}}{2}\right)\mathbf{x}, \mathbf{y}\right\rangle + \frac{\beta_2}{2}\|\mathbf{y}\|^2$.
- **Choices of β_1, β_2:**
  If $\beta \le \min\{\frac{\mu}{2}, \frac{\nu}{2}\}$, let $\beta_1 = \beta_2 = \beta$;
  If $\frac{\mu}{2} \le \beta \le \frac{\nu}{2}$, let $\beta_1 = \frac{\mu}{2}$ and $\beta_2 = \frac{2\beta^2}{\mu}$;
  If $\frac{\nu}{2} \le \beta \le \frac{\mu}{2}$, let $\beta_1 = \frac{2\beta^2}{\nu}$ and $\beta_2 = \frac{\nu}{2}$.
- We have $\beta_1 \le \frac{\mu}{2}$, $\beta_2 \le \frac{\nu}{2}$, and $\sqrt{\beta_1\beta_2} = \beta$.
  Therefore, $\widetilde{g}$ is convex and $(\beta_1 + \beta + \beta_2)$-smooth.
- **Our method:**

$$\widetilde{\mathcal{O}}\left(\left(\frac{L}{\sqrt{\mu\nu}} + \frac{L}{\min\{\mu,\nu\}} \cdot \frac{\beta}{\sqrt{\mu\nu}}\right) \cdot \log^2\left(\frac{1}{\varepsilon}\right)\right).$$

Acceleration conditioning: $\min\{\mu,\nu\} + \beta \ll \sqrt{\mu\nu}$.

- **Convex Reformulation**:
  Player 1: $\max_{\mathbf{x}\in X} \widetilde{u}_1(\mathbf{x},\mathbf{y}) = u_1(\mathbf{x},\mathbf{y}) - \beta_2 \|\mathbf{y}\|^2$;
  Player 2: $\max_{\mathbf{y}\in Y} \widetilde{u}_2(\mathbf{x},\mathbf{y}) = u_2(\mathbf{x},\mathbf{y}) - \beta_1 \|\mathbf{x}\|^2$.
  $\widetilde{g}(\mathbf{x},\mathbf{y}) = \frac{\beta_1}{2}\|\mathbf{x}\|^2 - \left\langle \left(\frac{\mathbf{A}+\mathbf{B}}{2}\right)\mathbf{x}, \mathbf{y}\right\rangle + \frac{\beta_2}{2}\|\mathbf{y}\|^2$.
- **Choices of β_1, β_2:**
  If $\beta \le \min\{\frac{\mu}{2}, \frac{\nu}{2}\}$, let $\beta_1 = \beta_2 = \beta$;
  If $\frac{\mu}{2} \le \beta \le \frac{\nu}{2}$, let $\beta_1 = \frac{\mu}{2}$ and $\beta_2 = \frac{2\beta^2}{\mu}$;
  If $\frac{\nu}{2} \le \beta \le \frac{\mu}{2}$, let $\beta_1 = \frac{2\beta^2}{\nu}$ and $\beta_2 = \frac{\nu}{2}$.
- We have $\beta_1 \le \frac{\mu}{2}$, $\beta_2 \le \frac{\nu}{2}$, and $\sqrt{\beta_1\beta_2} = \beta$.
  Therefore, $\widetilde{g}$ is convex and $(\beta_1 + \beta + \beta_2)$-smooth.
- **Our method:**

$$\widetilde{\mathcal{O}}\left(\left(\frac{L}{\sqrt{\mu\nu}} + \frac{L}{\min\{\mu,\nu\}} \cdot \frac{\beta}{\sqrt{\mu\nu}}\right) \cdot \log^2\left(\frac{1}{\varepsilon}\right)\right).$$

Acceleration conditioning: $\min\{\mu,\nu\} + \beta \ll \sqrt{\mu\nu}$.

- **Convex Reformulation**:
  Player 1: $\max_{\mathbf{x}\in X} \widetilde{u}_1(\mathbf{x},\mathbf{y}) = u_1(\mathbf{x},\mathbf{y}) - \beta_2 \|\mathbf{y}\|^2$;
  Player 2: $\max_{\mathbf{y}\in Y} \widetilde{u}_2(\mathbf{x},\mathbf{y}) = u_2(\mathbf{x},\mathbf{y}) - \beta_1 \|\mathbf{x}\|^2$.
  $\widetilde{g}(\mathbf{x},\mathbf{y}) = \frac{\beta_1}{2}\|\mathbf{x}\|^2 - \left\langle \left(\frac{\mathbf{A}+\mathbf{B}}{2}\right)\mathbf{x}, \mathbf{y} \right\rangle + \frac{\beta_2}{2}\|\mathbf{y}\|^2$.
- **Choices of β_1, β_2:**
  If $\beta \le \min\{\frac{\mu}{2}, \frac{\nu}{2}\}$, let $\beta_1 = \beta_2 = \beta$;
  If $\frac{\mu}{2} \le \beta \le \frac{\nu}{2}$, let $\beta_1 = \frac{\mu}{2}$ and $\beta_2 = \frac{2\beta^2}{\mu}$;
  If $\frac{\nu}{2} \le \beta \le \frac{\mu}{2}$, let $\beta_1 = \frac{2\beta^2}{\nu}$ and $\beta_2 = \frac{\nu}{2}$.
- We have $\beta_1 \le \frac{\mu}{2}$, $\beta_2 \le \frac{\nu}{2}$, and $\sqrt{\beta_1\beta_2} = \beta$.
  Therefore, $\widetilde{g}$ is convex and $(\beta_1+\beta+\beta_2)$-smooth.
- **Our method:**

$$\widetilde{\mathcal{O}}\left(\left(\frac{L}{\sqrt{\mu\nu}} + \frac{L}{\min\{\mu,\nu\}}\cdot\frac{\beta}{\sqrt{\mu\nu}}\right)\cdot\log^2\left(\frac{1}{\varepsilon}\right)\right).$$

Acceleration conditioning: $\min\{\mu,\nu\} + \beta \ll \sqrt{\mu\nu}$.

- **Convex Reformulation**:
  Player 1: $\max_{\mathbf{x}\in X} \widetilde{u}_1(\mathbf{x},\mathbf{y}) = u_1(\mathbf{x},\mathbf{y}) - \beta_2 \|\mathbf{y}\|^2$;
  Player 2: $\max_{\mathbf{y}\in Y} \widetilde{u}_2(\mathbf{x},\mathbf{y}) = u_2(\mathbf{x},\mathbf{y}) - \beta_1 \|\mathbf{x}\|^2$.
  $\widetilde{g}(\mathbf{x},\mathbf{y}) = \frac{\beta_1}{2}\|\mathbf{x}\|^2 - \left\langle \left(\frac{\mathbf{A}+\mathbf{B}}{2}\right)\mathbf{x}, \mathbf{y} \right\rangle + \frac{\beta_2}{2}\|\mathbf{y}\|^2$.
- **Choices of β_1, β_2:**
  If $\beta \le \min\{\frac{\mu}{2}, \frac{\nu}{2}\}$, let $\beta_1 = \beta_2 = \beta$;
  If $\frac{\mu}{2} \le \beta \le \frac{\nu}{2}$, let $\beta_1 = \frac{\mu}{2}$ and $\beta_2 = \frac{2\beta^2}{\mu}$;
  If $\frac{\nu}{2} \le \beta \le \frac{\mu}{2}$, let $\beta_1 = \frac{2\beta^2}{\nu}$ and $\beta_2 = \frac{\nu}{2}$.
- We have $\beta_1 \le \frac{\mu}{2}$, $\beta_2 \le \frac{\nu}{2}$, and $\sqrt{\beta_1\beta_2} = \beta$.
  Therefore, $\widetilde{g}$ is convex and $(\beta_1 + \beta + \beta_2)$-smooth.
- **Our method:**

$$\widetilde{\mathcal{O}}\left(\left(\frac{L}{\sqrt{\mu\nu}} + \frac{L}{\min\{\mu,\nu\}} \cdot \frac{\beta}{\sqrt{\mu\nu}}\right) \cdot \log^2\left(\frac{1}{\varepsilon}\right)\right).$$

Acceleration conditioning: $\min\{\mu,\nu\} + \beta \ll \sqrt{\mu\nu}$.

- **Convex Reformulation**:
  Player 1: $\max_{\mathbf{x}\in X} \widetilde{u}_1(\mathbf{x},\mathbf{y}) = u_1(\mathbf{x},\mathbf{y}) - \beta_2 \|\mathbf{y}\|^2$;
  Player 2: $\max_{\mathbf{y}\in Y} \widetilde{u}_2(\mathbf{x},\mathbf{y}) = u_2(\mathbf{x},\mathbf{y}) - \beta_1 \|\mathbf{x}\|^2$.
  $\widetilde{g}(\mathbf{x},\mathbf{y}) = \frac{\beta_1}{2}\|\mathbf{x}\|^2 - \left\langle \left(\frac{\mathbf{A}+\mathbf{B}}{2}\right)\mathbf{x}, \mathbf{y} \right\rangle + \frac{\beta_2}{2}\|\mathbf{y}\|^2$.
- **Choices of β_1, β_2:**
  If $\beta \le \min\{\frac{\mu}{2}, \frac{\nu}{2}\}$, let $\beta_1 = \beta_2 = \beta$;
  If $\frac{\mu}{2} \le \beta \le \frac{\nu}{2}$, let $\beta_1 = \frac{\mu}{2}$ and $\beta_2 = \frac{2\beta^2}{\mu}$;
  If $\frac{\nu}{2} \le \beta \le \frac{\mu}{2}$, let $\beta_1 = \frac{2\beta^2}{\nu}$ and $\beta_2 = \frac{\nu}{2}$.
- We have $\beta_1 \le \frac{\mu}{2}$, $\beta_2 \le \frac{\nu}{2}$, and $\sqrt{\beta_1\beta_2} = \beta$.
  Therefore, $\widetilde{g}$ is convex and $(\beta_1 + \beta + \beta_2)$-smooth.

- **Our method:**

$$\widetilde{\mathcal{O}}\left(\left(\frac{L}{\sqrt{\mu\nu}} + \frac{L}{\min\{\mu,\nu\}} \cdot \frac{\beta}{\sqrt{\mu\nu}}\right) \cdot \log^2\left(\frac{1}{\varepsilon}\right)\right).$$

Acceleration conditioning: $\min\{\mu,\nu\} + \beta \ll \sqrt{\mu\nu}$.

- **Convex Reformulation**:
  Player 1: $\max_{\mathbf{x}\in X} \widetilde{u}_1(\mathbf{x},\mathbf{y}) = u_1(\mathbf{x},\mathbf{y}) - \beta_2 \|\mathbf{y}\|^2$;
  Player 2: $\max_{\mathbf{y}\in Y} \widetilde{u}_2(\mathbf{x},\mathbf{y}) = u_2(\mathbf{x},\mathbf{y}) - \beta_1 \|\mathbf{x}\|^2$.
  $\widetilde{g}(\mathbf{x},\mathbf{y}) = \frac{\beta_1}{2}\|\mathbf{x}\|^2 - \left\langle \left(\frac{\mathbf{A}+\mathbf{B}}{2}\right)\mathbf{x},\mathbf{y}\right\rangle + \frac{\beta_2}{2}\|\mathbf{y}\|^2$.
- **Choices of β_1, β_2:**
  If $\beta \le \min\{\frac{\mu}{2},\frac{\nu}{2}\}$, let $\beta_1 = \beta_2 = \beta$;
  If $\frac{\mu}{2} \le \beta \le \frac{\nu}{2}$, let $\beta_1 = \frac{\mu}{2}$ and $\beta_2 = \frac{2\beta^2}{\mu}$;
  If $\frac{\nu}{2} \le \beta \le \frac{\mu}{2}$, let $\beta_1 = \frac{2\beta^2}{\nu}$ and $\beta_2 = \frac{\nu}{2}$.
- We have $\beta_1 \le \frac{\mu}{2}$, $\beta_2 \le \frac{\nu}{2}$, and $\sqrt{\beta_1\beta_2} = \beta$.
  Therefore, $\widetilde{g}$ is convex and $(\beta_1+\beta+\beta_2)$-smooth.

- **Our method:**

$$\widetilde{\mathcal{O}}\left(\left(\frac{L}{\sqrt{\mu\nu}} + \frac{L}{\min\{\mu,\nu\}}\cdot\frac{\beta}{\sqrt{\mu\nu}}\right)\cdot\log^2\left(\frac{1}{\varepsilon}\right)\right).$$

Acceleration conditioning: $\min\{\mu,\nu\} + \beta \ll \sqrt{\mu\nu}$.

Application 2: Strictly competitive games with additional incentives

- Player 1 maximizes $u_1 = -g(\mathbf{x}, \mathbf{y}) - h(\mathbf{x}, \mathbf{y})$ over $\mathbf{x} \in X$;
  Player 2 maximizes $u_2 = -g(\mathbf{x}, \mathbf{y}) + h(\mathbf{x}, \mathbf{y})$ over $\mathbf{y} \in Y$.
- h is the **competition payoff**, which is L-smooth and μ-strongly convex-ν-strongly concave;
- g is the **additional incentive**, which is β-smooth.
- **Example:** co-opetition[11].
- **Assume:** g is convex and $\beta = o\left(\max\{\frac{\mu}{2}, \frac{\nu}{2}\}\right)$;
- **Or assume:** $\beta = o\left(\min\{\frac{\mu}{2}, \frac{\nu}{2}\}\right)$.
- Variational inequality methods: $\widetilde{\mathcal{O}}\left(\frac{L}{\min\{\mu,\nu\}} \cdot \log\left(\frac{1}{\varepsilon}\right)\right)$.[12]

[11] Adam M Brandenburger and Barry J Nalebuff. *Co-opetition.* Crown Currency, 2011.
[12] Tseng, "On linear convergence of iterative methods for the variational inequality problem".

Application 2: Strictly competitive games with additional incentives

- Player 1 maximizes $u_1 = -g(\mathbf{x}, \mathbf{y}) - h(\mathbf{x}, \mathbf{y})$ over $\mathbf{x} \in X$;
  Player 2 maximizes $u_2 = -g(\mathbf{x}, \mathbf{y}) + h(\mathbf{x}, \mathbf{y})$ over $\mathbf{y} \in Y$.
- h is the **competition payoff**, which is L-smooth and μ-strongly convex-ν-strongly concave;
- g is the **additional incentive**, which is β-smooth.
- **Example:** co-opetition[11].
- **Assume:** g is convex and $\beta = o\left(\max\{\frac{\mu}{2}, \frac{\nu}{2}\}\right)$;
- **Or assume:** $\beta = o\left(\min\{\frac{\mu}{2}, \frac{\nu}{2}\}\right)$.
- Variational inequality methods: $\widetilde{\mathcal{O}}\left(\frac{L}{\min\{\mu,\nu\}} \cdot \log\left(\frac{1}{\varepsilon}\right)\right)$.[12]

[11] Brandenburger and Nalebuff, *Co-opetition*.

[12] Tseng, "On linear convergence of iterative methods for the variational inequality problem".

Application 2: Strictly competitive games with additional incentives

- Player 1 maximizes $u_1 = -g(\mathbf{x},\mathbf{y}) - h(\mathbf{x},\mathbf{y})$ over $\mathbf{x} \in X$;
  Player 2 maximizes $u_2 = -g(\mathbf{x},\mathbf{y}) + h(\mathbf{x},\mathbf{y})$ over $\mathbf{y} \in Y$.
- h is the **competition payoff**, which is L-smooth and μ-strongly convex-ν-strongly concave;
- g is the **additional incentive**, which is β-smooth.

- **Example:** co-opetition[11].

- **Assume:** g is convex and $\beta = o\left(\max\{\frac{\mu}{2}, \frac{\nu}{2}\}\right)$;
- **Or assume:** $\beta = o\left(\min\{\frac{\mu}{2}, \frac{\nu}{2}\}\right)$.
- Variational inequality methods: $\widetilde{\mathcal{O}}\left(\frac{L}{\min\{\mu,\nu\}} \cdot \log\left(\frac{1}{\varepsilon}\right)\right)$.[12]

[11] Brandenburger and Nalebuff, *Co-opetition*.

[12] Tseng, "On linear convergence of iterative methods for the variational inequality problem".

Application 2: Strictly competitive games with additional incentives

- Player 1 maximizes $u_1 = -g(\mathbf{x}, \mathbf{y}) - h(\mathbf{x}, \mathbf{y})$ over $\mathbf{x} \in X$;
  Player 2 maximizes $u_2 = -g(\mathbf{x}, \mathbf{y}) + h(\mathbf{x}, \mathbf{y})$ over $\mathbf{y} \in Y$.
- h is the **competition payoff**, which is L-smooth and μ-strongly convex-ν-strongly concave;
- g is the **additional incentive**, which is β-smooth.

- **Example:** co-opetition[11].

- **Assume:** g is convex and $\beta = o\left(\max\{\frac{\mu}{2}, \frac{\nu}{2}\}\right)$;
- **Or assume:** $\beta = o\left(\min\{\frac{\mu}{2}, \frac{\nu}{2}\}\right)$.
- Variational inequality methods: $\widetilde{\mathcal{O}}\left(\frac{L}{\min\{\mu,\nu\}} \cdot \log\left(\frac{1}{\varepsilon}\right)\right)$.[12]

[11] Brandenburger and Nalebuff, *Co-opetition*.

[12] Tseng, "On linear convergence of iterative methods for the variational inequality problem".

Application 2:
Strictly competitive games with additional incentives

- Player 1 maximizes $u_1 = -g(\mathbf{x}, \mathbf{y}) - h(\mathbf{x}, \mathbf{y})$ over $\mathbf{x} \in X$;
  Player 2 maximizes $u_2 = -g(\mathbf{x}, \mathbf{y}) + h(\mathbf{x}, \mathbf{y})$ over $\mathbf{y} \in Y$.
- h is the **competition payoff**, which is L-smooth and μ-strongly convex-ν-strongly concave;
- g is the **additional incentive**, which is β-smooth.

- **Example:** co-opetition[11].

- **Assume:** g is convex and $\beta = o\left(\max\{\frac{\mu}{2}, \frac{\nu}{2}\}\right)$;
- **Or assume:** $\beta = o\left(\min\{\frac{\mu}{2}, \frac{\nu}{2}\}\right)$.
- Variational inequality methods: $\widetilde{\mathcal{O}}\left(\frac{L}{\min\{\mu,\nu\}} \cdot \log\left(\frac{1}{\varepsilon}\right)\right)$.[12]

[11] Brandenburger and Nalebuff, *Co-opetition*.
[12] Tseng, "On linear convergence of iterative methods for the variational inequality problem".

Application 2: Strictly competitive games with additional incentives

- Player 1 maximizes $u_1 = -g(\mathbf{x}, \mathbf{y}) - h(\mathbf{x}, \mathbf{y})$ over $\mathbf{x} \in X$;
  Player 2 maximizes $u_2 = -g(\mathbf{x}, \mathbf{y}) + h(\mathbf{x}, \mathbf{y})$ over $\mathbf{y} \in Y$.
- h is the **competition payoff**, which is L-smooth and μ-strongly convex-ν-strongly concave;
- g is the **additional incentive**, which is β-smooth.

- **Example:** co-opetition[11].

- **Assume:** g is convex and $\beta = o\left(\max\{\frac{\mu}{2}, \frac{\nu}{2}\}\right)$;
- **Or assume:** $\beta = o\left(\min\{\frac{\mu}{2}, \frac{\nu}{2}\}\right)$.
- Variational inequality methods: $\widetilde{\mathcal{O}}\left(\frac{L}{\min\{\mu,\nu\}} \cdot \log\left(\frac{1}{\varepsilon}\right)\right)$.[12]

[11] Brandenburger and Nalebuff, *Co-opetition*.

[12] Tseng, "On linear convergence of iterative methods for the variational inequality problem".

Application 2: Strictly competitive games with additional incentives

- Player 1 maximizes $u_1 = -g(\mathbf{x}, \mathbf{y}) - h(\mathbf{x}, \mathbf{y})$ over $\mathbf{x} \in X$;
  Player 2 maximizes $u_2 = -g(\mathbf{x}, \mathbf{y}) + h(\mathbf{x}, \mathbf{y})$ over $\mathbf{y} \in Y$.
- h is the **competition payoff**, which is L-smooth and μ-strongly convex-ν-strongly concave;
- g is the **additional incentive**, which is β-smooth.

- **Example:** co-opetition[11].

- **Assume:** g is convex and $\beta = o\left(\max\{\frac{\mu}{2}, \frac{\nu}{2}\}\right)$;
- **Or assume:** $\beta = o\left(\min\{\frac{\mu}{2}, \frac{\nu}{2}\}\right)$.
- Variational inequality methods: $\widetilde{\mathcal{O}}\left(\frac{L}{\min\{\mu,\nu\}} \cdot \log\left(\frac{1}{\varepsilon}\right)\right)$.[12]

[11] Brandenburger and Nalebuff, *Co-opetition*.

[12] Tseng, "On linear convergence of iterative methods for the variational inequality problem".

1. **Consider the case** where g is convex and $\beta = o\left(\max\{\frac{\mu}{2}, \frac{\nu}{2}\}\right)$.
   Our method:

$$\widetilde{\mathcal{O}}\left(\left(\frac{L}{\sqrt{\mu\nu}} + \frac{L}{\min\{\mu,\nu\}} \cdot \sqrt{\frac{\beta}{\mu+\nu}}\right) \cdot \log^2\left(\frac{1}{\varepsilon}\right)\right).$$

2. **Consider the case** where $\beta = o\left(\min\{\frac{\mu}{2}, \frac{\nu}{2}\}\right)$.
   Convex Reformulation:

$$\widetilde{u}_1(\mathbf{x},\mathbf{y}) = u_1(\mathbf{x},\mathbf{y}) - \beta\|\mathbf{y}\|^2; \quad \widetilde{u}_2(\mathbf{x},\mathbf{y}) = u_2(\mathbf{x},\mathbf{y}) - \beta\|\mathbf{x}\|^2.$$

   Let $\widetilde{g} = -\frac{1}{2}(\widetilde{u}_1 + \widetilde{u}_2)$ and $\widetilde{h} = \frac{1}{2}(-\widetilde{u}_1 + \widetilde{u}_2)$.
   Then $\widetilde{g}(\cdot)$ is convex and 2β-smooth.
   Our method:

$$\widetilde{\mathcal{O}}\left(\frac{L}{\sqrt{\mu\nu}} \cdot \log^2\left(\frac{1}{\varepsilon}\right)\right).$$

NB. This is in sharp contrast to $\beta = o\left(\frac{1}{2}\sqrt{\mu\nu}\right)$ in matrix games.

1. **Consider the case** where g is convex and $\beta = o\left(\max\{\frac{\mu}{2}, \frac{\nu}{2}\}\right)$.
   Our method:

$$\widetilde{\mathcal{O}}\left(\left(\frac{L}{\sqrt{\mu\nu}} + \frac{L}{\min\{\mu,\nu\}} \cdot \sqrt{\frac{\beta}{\mu+\nu}}\right) \cdot \log^2\left(\frac{1}{\varepsilon}\right)\right).$$

2. **Consider the case** where $\beta = o\left(\min\{\frac{\mu}{2}, \frac{\nu}{2}\}\right)$.
   Convex Reformulation:

$$\widetilde{u}_1(\mathbf{x},\mathbf{y}) = u_1(\mathbf{x},\mathbf{y}) - \beta\|\mathbf{y}\|^2; \quad \widetilde{u}_2(\mathbf{x},\mathbf{y}) = u_2(\mathbf{x},\mathbf{y}) - \beta\|\mathbf{x}\|^2.$$

   Let $\widetilde{g} = -\frac{1}{2}(\widetilde{u}_1 + \widetilde{u}_2)$ and $\widetilde{h} = \frac{1}{2}(-\widetilde{u}_1 + \widetilde{u}_2)$.
   Then $\widetilde{g}(\cdot)$ is convex and 2β-smooth.
   Our method:

$$\widetilde{\mathcal{O}}\left(\frac{L}{\sqrt{\mu\nu}} \cdot \log^2\left(\frac{1}{\varepsilon}\right)\right).$$

   NB. This is in sharp contrast to $\beta = o\left(\frac{1}{2}\sqrt{\mu\nu}\right)$ in matrix games.

1. **Consider the case** where g is convex and $\beta = o\left(\max\{\frac{\mu}{2}, \frac{\nu}{2}\}\right)$.
   Our method:

$$\widetilde{\mathcal{O}}\left(\left(\frac{L}{\sqrt{\mu\nu}} + \frac{L}{\min\{\mu, \nu\}} \cdot \sqrt{\frac{\beta}{\mu+\nu}}\right) \cdot \log^2\left(\frac{1}{\varepsilon}\right)\right) .$$

2. **Consider the case** where $\beta = o\left(\min\{\frac{\mu}{2}, \frac{\nu}{2}\}\right)$.
   Convex Reformulation:

$$\widetilde{u}_1(\mathbf{x}, \mathbf{y}) = u_1(\mathbf{x}, \mathbf{y}) - \beta \|\mathbf{y}\|^2 ; \quad \widetilde{u}_2(\mathbf{x}, \mathbf{y}) = u_2(\mathbf{x}, \mathbf{y}) - \beta \|\mathbf{x}\|^2 .$$

   Let $\widetilde{g} = -\frac{1}{2}(\widetilde{u}_1 + \widetilde{u}_2)$ and $\widetilde{h} = \frac{1}{2}(-\widetilde{u}_1 + \widetilde{u}_2)$.
   Then $\widetilde{g}(\cdot)$ is convex and 2β-smooth.
   Our method:

$$\widetilde{\mathcal{O}}\left(\frac{L}{\sqrt{\mu\nu}} \cdot \log^2\left(\frac{1}{\varepsilon}\right)\right) .$$

   NB. This is in sharp contrast to $\beta = o\left(\frac{1}{2}\sqrt{\mu\nu}\right)$ in matrix games.

1. **Consider the case** where g is convex and $\beta = o\left(\max\{\frac{\mu}{2}, \frac{\nu}{2}\}\right)$.
   Our method:

$$\widetilde{\mathcal{O}}\left(\left(\frac{L}{\sqrt{\mu\nu}} + \frac{L}{\min\{\mu,\nu\}} \cdot \sqrt{\frac{\beta}{\mu+\nu}}\right) \cdot \log^2\left(\frac{1}{\varepsilon}\right)\right).$$

2. **Consider the case** where $\beta = o\left(\min\{\frac{\mu}{2}, \frac{\nu}{2}\}\right)$.
   Convex Reformulation:

$$\widetilde{u}_1(\mathbf{x},\mathbf{y}) = u_1(\mathbf{x},\mathbf{y}) - \beta\|\mathbf{y}\|^2; \quad \widetilde{u}_2(\mathbf{x},\mathbf{y}) = u_2(\mathbf{x},\mathbf{y}) - \beta\|\mathbf{x}\|^2.$$

   Let $\widetilde{g} = -\frac{1}{2}(\widetilde{u}_1 + \widetilde{u}_2)$ and $\widetilde{h} = \frac{1}{2}(-\widetilde{u}_1 + \widetilde{u}_2)$.
   Then $\widetilde{g}(\cdot)$ is convex and 2β-smooth.
   Our method:

$$\widetilde{\mathcal{O}}\left(\frac{L}{\sqrt{\mu\nu}} \cdot \log^2\left(\frac{1}{\varepsilon}\right)\right).$$

   NB. This is in sharp contrast to $\beta = o\left(\frac{1}{2}\sqrt{\mu\nu}\right)$ in matrix games.

1. **Consider the case** where g is convex and $\beta = o\left(\max\{\frac{\mu}{2}, \frac{\nu}{2}\}\right)$.
   Our method:

$$\widetilde{\mathcal{O}}\left(\left(\frac{L}{\sqrt{\mu\nu}} + \frac{L}{\min\{\mu,\nu\}} \cdot \sqrt{\frac{\beta}{\mu+\nu}}\right) \cdot \log^2\left(\frac{1}{\varepsilon}\right)\right).$$

2. **Consider the case** where $\beta = o\left(\min\{\frac{\mu}{2}, \frac{\nu}{2}\}\right)$.
   Convex Reformulation:

$$\widetilde{u}_1(\mathbf{x},\mathbf{y}) = u_1(\mathbf{x},\mathbf{y}) - \beta\|\mathbf{y}\|^2; \quad \widetilde{u}_2(\mathbf{x},\mathbf{y}) = u_2(\mathbf{x},\mathbf{y}) - \beta\|\mathbf{x}\|^2.$$

   Let $\widetilde{g} = -\frac{1}{2}(\widetilde{u}_1 + \widetilde{u}_2)$ and $\widetilde{h} = \frac{1}{2}(-\widetilde{u}_1 + \widetilde{u}_2)$.
   Then $\widetilde{g}(\cdot)$ is convex and 2β-smooth.
   Our method:

$$\widetilde{\mathcal{O}}\left(\frac{L}{\sqrt{\mu\nu}} \cdot \log^2\left(\frac{1}{\varepsilon}\right)\right).$$

   NB. This is in sharp contrast to $\beta = o\left(\frac{1}{2}\sqrt{\mu\nu}\right)$ in matrix games.

1. **Consider the case** where g is convex and $\beta = o\left(\max\{\frac{\mu}{2}, \frac{\nu}{2}\}\right)$.
Our method:

$$\widetilde{\mathcal{O}}\left(\left(\frac{L}{\sqrt{\mu\nu}} + \frac{L}{\min\{\mu,\nu\}} \cdot \sqrt{\frac{\beta}{\mu+\nu}}\right) \cdot \log^2\left(\frac{1}{\varepsilon}\right)\right) .$$

2. **Consider the case** where $\beta = o\left(\min\{\frac{\mu}{2}, \frac{\nu}{2}\}\right)$.
Convex Reformulation:

$$\widetilde{u}_1(\mathbf{x},\mathbf{y}) = u_1(\mathbf{x},\mathbf{y}) - \beta\|\mathbf{y}\|^2 ; \quad \widetilde{u}_2(\mathbf{x},\mathbf{y}) = u_2(\mathbf{x},\mathbf{y}) - \beta\|\mathbf{x}\|^2 .$$

Let $\widetilde{g} = -\frac{1}{2}(\widetilde{u}_1 + \widetilde{u}_2)$ and $\widetilde{h} = \frac{1}{2}(-\widetilde{u}_1 + \widetilde{u}_2)$.
Then $\widetilde{g}(\cdot)$ is convex and 2β-smooth.
Our method:

$$\widetilde{\mathcal{O}}\left(\frac{L}{\sqrt{\mu\nu}} \cdot \log^2\left(\frac{1}{\varepsilon}\right)\right) .$$

NB. This is in sharp contrast to $\beta = o\left(\frac{1}{2}\sqrt{\mu\nu}\right)$ in matrix games.

1. **Consider the case** where g is convex and $\beta = o\left(\max\{\frac{\mu}{2}, \frac{\nu}{2}\}\right)$.
   Our method:

$$\widetilde{\mathcal{O}}\left(\left(\frac{L}{\sqrt{\mu\nu}} + \frac{L}{\min\{\mu,\nu\}} \cdot \sqrt{\frac{\beta}{\mu+\nu}}\right) \cdot \log^2\left(\frac{1}{\varepsilon}\right)\right) .$$

2. **Consider the case** where $\beta = o\left(\min\{\frac{\mu}{2}, \frac{\nu}{2}\}\right)$.
   Convex Reformulation:

$$\widetilde{u}_1(\mathbf{x},\mathbf{y}) = u_1(\mathbf{x},\mathbf{y}) - \beta\left\|\mathbf{y}\right\|^2 ; \quad \widetilde{u}_2(\mathbf{x},\mathbf{y}) = u_2(\mathbf{x},\mathbf{y}) - \beta\left\|\mathbf{x}\right\|^2 .$$

Let $\widetilde{g} = -\frac{1}{2}(\widetilde{u}_1 + \widetilde{u}_2)$ and $\widetilde{h} = \frac{1}{2}(-\widetilde{u}_1 + \widetilde{u}_2)$.
Then $\widetilde{g}(\cdot)$ is convex and 2β-smooth.
Our method:

$$\widetilde{\mathcal{O}}\left(\frac{L}{\sqrt{\mu\nu}} \cdot \log^2\left(\frac{1}{\varepsilon}\right)\right) .$$

NB. This is in sharp contrast to $\beta = o\left(\frac{1}{2}\sqrt{\mu\nu}\right)$ in matrix games.

1. **Consider the case** where g is convex and $\beta = o\left(\max\{\frac{\mu}{2}, \frac{\nu}{2}\}\right)$.
   Our method:

$$\widetilde{\mathcal{O}}\left(\left(\frac{L}{\sqrt{\mu\nu}} + \frac{L}{\min\{\mu,\nu\}} \cdot \sqrt{\frac{\beta}{\mu+\nu}}\right) \cdot \log^2\left(\frac{1}{\varepsilon}\right)\right).$$

2. **Consider the case** where $\beta = o\left(\min\{\frac{\mu}{2}, \frac{\nu}{2}\}\right)$.
   Convex Reformulation:

$$\widetilde{u}_1(\mathbf{x}, \mathbf{y}) = u_1(\mathbf{x}, \mathbf{y}) - \beta \|\mathbf{y}\|^2; \quad \widetilde{u}_2(\mathbf{x}, \mathbf{y}) = u_2(\mathbf{x}, \mathbf{y}) - \beta \|\mathbf{x}\|^2.$$

Let $\widetilde{g} = -\frac{1}{2}(\widetilde{u}_1 + \widetilde{u}_2)$ and $\widetilde{h} = \frac{1}{2}(-\widetilde{u}_1 + \widetilde{u}_2)$.
Then $\widetilde{g}(\cdot)$ is convex and 2β-smooth.
Our method:

$$\widetilde{\mathcal{O}}\left(\frac{L}{\sqrt{\mu\nu}} \cdot \log^2\left(\frac{1}{\varepsilon}\right)\right).$$

NB. This is in sharp contrast to $\beta = o\left(\frac{1}{2}\sqrt{\mu\nu}\right)$ in matrix games.

1. **Consider the case** where g is convex and $\beta = o\left(\max\{\frac{\mu}{2}, \frac{\nu}{2}\}\right)$.
Our method:

$$\widetilde{\mathcal{O}}\left(\left(\frac{L}{\sqrt{\mu\nu}} + \frac{L}{\min\{\mu,\nu\}} \cdot \sqrt{\frac{\beta}{\mu+\nu}}\right) \cdot \log^2\left(\frac{1}{\varepsilon}\right)\right).$$

2. **Consider the case** where $\beta = o\left(\min\{\frac{\mu}{2}, \frac{\nu}{2}\}\right)$.
Convex Reformulation:

$$\widetilde{u}_1(\mathbf{x},\mathbf{y}) = u_1(\mathbf{x},\mathbf{y}) - \beta\|\mathbf{y}\|^2; \quad \widetilde{u}_2(\mathbf{x},\mathbf{y}) = u_2(\mathbf{x},\mathbf{y}) - \beta\|\mathbf{x}\|^2.$$

Let $\widetilde{g} = -\frac{1}{2}(\widetilde{u}_1 + \widetilde{u}_2)$ and $\widetilde{h} = \frac{1}{2}(-\widetilde{u}_1 + \widetilde{u}_2)$.
Then $\widetilde{g}(\cdot)$ is convex and 2β-smooth.
Our method:

$$\widetilde{\mathcal{O}}\left(\frac{L}{\sqrt{\mu\nu}} \cdot \log^2\left(\frac{1}{\varepsilon}\right)\right).$$

NB. This is in sharp contrast to $\beta = o\left(\frac{1}{2}\sqrt{\mu\nu}\right)$ in matrix games.

1. **Consider the case** where g is convex and $\beta = o\left(\max\{\frac{\mu}{2}, \frac{\nu}{2}\}\right)$.
   Our method:

$$\widetilde{\mathcal{O}}\left(\left(\frac{L}{\sqrt{\mu\nu}} + \frac{L}{\min\{\mu,\nu\}} \cdot \sqrt{\frac{\beta}{\mu+\nu}}\right) \cdot \log^2\left(\frac{1}{\varepsilon}\right)\right).$$

2. **Consider the case** where $\beta = o\left(\min\{\frac{\mu}{2}, \frac{\nu}{2}\}\right)$.
   Convex Reformulation:

$$\widetilde{u}_1(\mathbf{x},\mathbf{y}) = u_1(\mathbf{x},\mathbf{y}) - \beta\|\mathbf{y}\|^2; \quad \widetilde{u}_2(\mathbf{x},\mathbf{y}) = u_2(\mathbf{x},\mathbf{y}) - \beta\|\mathbf{x}\|^2.$$

   Let $\widetilde{g} = -\frac{1}{2}(\widetilde{u}_1 + \widetilde{u}_2)$ and $\widetilde{h} = \frac{1}{2}(-\widetilde{u}_1 + \widetilde{u}_2)$.
   Then $\widetilde{g}(\cdot)$ is convex and 2β-smooth.
   Our method:

$$\widetilde{\mathcal{O}}\left(\frac{L}{\sqrt{\mu\nu}} \cdot \log^2\left(\frac{1}{\varepsilon}\right)\right).$$

   NB. This is in sharp contrast to $\beta = o\left(\frac{1}{2}\sqrt{\mu\nu}\right)$ in matrix games.

Outline for section 5

Preliminary experiments: matrix games with transaction fee

Let $\mathbf{M} \in \mathbb{R}^{m\times n}$ be the payoff matrix of Player 1, and then $-\mathbf{M}$ be the payoff matrix of Player 2, both *without transaction fee*.
A **transaction fee** of $\rho \in [0,1]$ is imposed on every payment.
$\mathbf{M}_+ = \frac{1}{2}(\mathbf{M} + |\mathbf{M}|), \quad \mathbf{M}_- = \frac{1}{2}(-\mathbf{M} + |\mathbf{M}|)\,.$
The payoff matrices of Player 1 and Player 2 *with transaction fee* are

$$\mathbf{A} = (1-\rho)\mathbf{M}_+ - \mathbf{M}_-, \quad \mathbf{B} = -\mathbf{M}_+ + (1-\rho)\mathbf{M}_-\,.$$

For instance:

Table 1: Payoff matrices without/with the transaction fee $\rho = 1\%$.

300/-300	-200/200
-100/100	400/-400

297/-300	-200/198
-100/99	396/-400

Preliminary experiments: matrix games with transaction fee

Let $\mathbf{M} \in \mathbb{R}^{m \times n}$ be the payoff matrix of Player 1, and
then $-\mathbf{M}$ be the payoff matrix of Player 2, both *without transaction fee*.
A **transaction fee** of $\rho \in [0,1]$ is imposed on every payment.
$\mathbf{M}_+ = \frac{1}{2}(\mathbf{M} + |\mathbf{M}|), \quad \mathbf{M}_- = \frac{1}{2}(-\mathbf{M} + |\mathbf{M}|)$.
The payoff matrices of Player 1 and Player 2 *with transaction fee* are

$$\mathbf{A} = (1-\rho)\mathbf{M}_+ - \mathbf{M}_-, \quad \mathbf{B} = -\mathbf{M}_+ + (1-\rho)\mathbf{M}_- .$$

For instance:

Table 1: Payoff matrices without/with the transaction fee $\rho = 1\%$.

300/-300	-200/200
-100/100	400/-400

297/-300	-200/198
-100/99	396/-400

Preliminary experiments: matrix games with transaction fee

Let $\mathbf{M} \in \mathbb{R}^{m\times n}$ be the payoff matrix of Player 1, and
then $-\mathbf{M}$ be the payoff matrix of Player 2, both *without transaction fee*.
A **transaction fee** of $\rho \in [0, 1]$ is imposed on every payment.
$\mathbf{M}_+ = \frac{1}{2}(\mathbf{M} + |\mathbf{M}|), \quad \mathbf{M}_- = \frac{1}{2}(-\mathbf{M} + |\mathbf{M}|)$.
The payoff matrices of Player 1 and Player 2 *with transaction fee* are

$$\mathbf{A} = (1-\rho)\mathbf{M}_+ - \mathbf{M}_-, \quad \mathbf{B} = -\mathbf{M}_+ + (1-\rho)\mathbf{M}_- .$$

For instance:

Table 1: Payoff matrices without/with the transaction fee $\rho = 1\%$.

300/-300	-200/200
-100/100	400/-400

297/-300	-200/198
-100/99	396/-400

Preliminary experiments: matrix games with transaction fee

Let $\mathbf{M} \in \mathbb{R}^{m \times n}$ be the payoff matrix of Player 1, and then $-\mathbf{M}$ be the payoff matrix of Player 2, both *without transaction fee*.
A **transaction fee** of $\rho \in [0, 1]$ is imposed on every payment.
$\mathbf{M}_+ = \frac{1}{2}(\mathbf{M} + |\mathbf{M}|), \quad \mathbf{M}_- = \frac{1}{2}(-\mathbf{M} + |\mathbf{M}|)\,.$
The payoff matrices of Player 1 and Player 2 *with transaction fee* are

$$\mathbf{A} = (1-\rho)\mathbf{M}_+ - \mathbf{M}_-, \quad \mathbf{B} = -\mathbf{M}_+ + (1-\rho)\mathbf{M}_-\,.$$

For instance:

Table 1: Payoff matrices without/with the transaction fee $\rho = 1\%$.

300/-300	-200/200
-100/100	400/-400

297/-300	-200/198
-100/99	396/-400

Preliminary experiments: matrix games with transaction fee

Let $\mathbf{M} \in \mathbb{R}^{m\times n}$ be the payoff matrix of Player 1, and
then $-\mathbf{M}$ be the payoff matrix of Player 2, both *without transaction fee*.
A **transaction fee** of $\rho \in [0,1]$ is imposed on every payment.
$\mathbf{M}_+ = \frac{1}{2}(\mathbf{M} + |\mathbf{M}|), \quad \mathbf{M}_- = \frac{1}{2}(-\mathbf{M} + |\mathbf{M}|)\,.$
The payoff matrices of Player 1 and Player 2 *with transaction fee* are

$$\mathbf{A} = (1-\rho)\mathbf{M}_+ - \mathbf{M}_-, \quad \mathbf{B} = -\mathbf{M}_+ + (1-\rho)\mathbf{M}_-\,.$$

For instance:

Table 1: Payoff matrices without/with the transaction fee $\rho = 1\%$.

300/-300	-200/200
-100/100	400/-400

297/-300	-200/198
-100/99	396/-400

Preliminary experiments: matrix games with transaction fee

Let $\mathbf{M} \in \mathbb{R}^{m\times n}$ be the payoff matrix of Player 1, and
then $-\mathbf{M}$ be the payoff matrix of Player 2, both *without transaction fee*.
A **transaction fee** of $\rho \in [0, 1]$ is imposed on every payment.
$\mathbf{M}_+ = \frac{1}{2}(\mathbf{M} + |\mathbf{M}|), \quad \mathbf{M}_- = \frac{1}{2}(-\mathbf{M} + |\mathbf{M}|)\,.$
The payoff matrices of Player 1 and Player 2 *with transaction fee* are

$$\mathbf{A} = (1-\rho)\mathbf{M}_+ - \mathbf{M}_-, \quad \mathbf{B} = -\mathbf{M}_+ + (1-\rho)\mathbf{M}_-\,.$$

For instance:

Table 1: Payoff matrices without/with the transaction fee $\rho = 1\%$.

300/-300	-200/200
-100/100	400/-400

297/-300	-200/198
-100/99	396/-400

Matrix games with transaction fee - experimental setup

Let $n = m = 10000$, $\mu = 10^{-4}$, $\nu = 1$, and $\varepsilon = 10^{-7}$.
Let $\mathcal{R}_1 = \frac{\mu}{2}\|\cdot\|^2$ and $\mathcal{R}_2 = \frac{\nu}{2}\|\cdot\|^2$.
We generate a sparse, random matrix $\mathbf{M} \in \mathbb{R}^{m\times n}$ s.t. $\|\mathbf{M}\| = 1$.
We choose the **transaction fee** ρ from $\{0.00\%, 0.03\%, \cdots, 0.18\%\}$.
Player 1: $\max_{\mathbf{x}\in\Delta_n} u_1(\mathbf{x},\mathbf{y}) = \mathcal{R}_1(\mathbf{x}) + \langle \mathbf{A}\mathbf{x}, \mathbf{y}\rangle - \mathcal{R}_2(\mathbf{y})$;
Player 2: $\max_{\mathbf{y}\in\Delta_m} u_2(\mathbf{x},\mathbf{y}) = -\mathcal{R}_1(\mathbf{x}) + \langle \mathbf{B}\mathbf{x}, \mathbf{y}\rangle + \mathcal{R}_2(\mathbf{y})$.

Our theory:

$$\widetilde{\mathcal{O}}\left(\left(\frac{L}{\sqrt{\mu\nu}} + \frac{L}{\min\{\mu,\nu\}}\cdot\frac{\rho\,|||\mathbf{M}|||}{\sqrt{\mu\nu}}\right)\cdot\log^2\left(\frac{1}{\varepsilon}\right)\right).$$

Matrix games with transaction fee - experimental setup

Let $n = m = 10000$, $\mu = 10^{-4}$, $\nu = 1$, and $\varepsilon = 10^{-7}$.
Let $\mathcal{R}_1 = \frac{\mu}{2}\|\cdot\|^2$ and $\mathcal{R}_2 = \frac{\nu}{2}\|\cdot\|^2$.
We generate a sparse, random matrix $\mathbf{M} \in \mathbb{R}^{m\times n}$ s.t. $\|\mathbf{M}\| = 1$.
We choose the **transaction fee** ρ from $\{0.00\%, 0.03\%, \cdots, 0.18\%\}$.
Player 1: $\max_{\mathbf{x}\in\Delta_n} u_1(\mathbf{x},\mathbf{y}) = \mathcal{R}_1(\mathbf{x}) + \langle \mathbf{A}\mathbf{x}, \mathbf{y}\rangle - \mathcal{R}_2(\mathbf{y})$;
Player 2: $\max_{\mathbf{y}\in\Delta_m} u_2(\mathbf{x},\mathbf{y}) = -\mathcal{R}_1(\mathbf{x}) + \langle \mathbf{B}\mathbf{x}, \mathbf{y}\rangle + \mathcal{R}_2(\mathbf{y})$.

Our theory:

$$\widetilde{\mathcal{O}}\left(\left(\frac{L}{\sqrt{\mu\nu}} + \frac{L}{\min\{\mu,\nu\}}\cdot\frac{\rho\,\|\|\mathbf{M}\|\|}{\sqrt{\mu\nu}}\right)\cdot\log^2\left(\frac{1}{\varepsilon}\right)\right).$$

Matrix games with transaction fee - experimental setup

Let $n = m = 10000$, $\mu = 10^{-4}$, $\nu = 1$, and $\varepsilon = 10^{-7}$.
Let $\mathcal{R}_1 = \frac{\mu}{2} \|\cdot\|^2$ and $\mathcal{R}_2 = \frac{\nu}{2} \|\cdot\|^2$.
We generate a sparse, random matrix $\mathbf{M} \in \mathbb{R}^{m \times n}$ s.t. $\|\mathbf{M}\| = 1$.
We choose the **transaction fee** ρ from $\{0.00\%, 0.03\%, \cdots, 0.18\%\}$.
Player 1: $\max_{\mathbf{x} \in \Delta_n} u_1(\mathbf{x}, \mathbf{y}) = \mathcal{R}_1(\mathbf{x}) + \langle \mathbf{A}\mathbf{x}, \mathbf{y} \rangle - \mathcal{R}_2(\mathbf{y})$;
Player 2: $\max_{\mathbf{y} \in \Delta_m} u_2(\mathbf{x}, \mathbf{y}) = -\mathcal{R}_1(\mathbf{x}) + \langle \mathbf{B}\mathbf{x}, \mathbf{y} \rangle + \mathcal{R}_2(\mathbf{y})$.

Our theory:

$$\widetilde{\mathcal{O}}\left(\left(\frac{L}{\sqrt{\mu\nu}} + \frac{L}{\min\{\mu, \nu\}} \cdot \frac{\rho \, |||\mathbf{M}|||}{\sqrt{\mu\nu}}\right) \cdot \log^2\left(\frac{1}{\varepsilon}\right)\right).$$

Matrix games with transaction fee - experimental setup

Let $n = m = 10000$, $\mu = 10^{-4}$, $\nu = 1$, and $\varepsilon = 10^{-7}$.
Let $\mathcal{R}_1 = \frac{\mu}{2}\|\cdot\|^2$ and $\mathcal{R}_2 = \frac{\nu}{2}\|\cdot\|^2$.
We generate a sparse, random matrix $\mathbf{M} \in \mathbb{R}^{m\times n}$ s.t. $\|\mathbf{M}\| = 1$.
We choose the **transaction fee** ρ from $\{0.00\%, 0.03\%, \cdots, 0.18\%\}$.
Player 1: $\max_{\mathbf{x}\in\Delta_n} u_1(\mathbf{x},\mathbf{y}) = \mathcal{R}_1(\mathbf{x}) + \langle \mathbf{Ax}, \mathbf{y}\rangle - \mathcal{R}_2(\mathbf{y})$;
Player 2: $\max_{\mathbf{y}\in\Delta_m} u_2(\mathbf{x},\mathbf{y}) = -\mathcal{R}_1(\mathbf{x}) + \langle \mathbf{Bx}, \mathbf{y}\rangle + \mathcal{R}_2(\mathbf{y})$.

Our theory:

$$\widetilde{\mathcal{O}}\left(\left(\frac{L}{\sqrt{\mu\nu}} + \frac{L}{\min\{\mu,\nu\}}\cdot\frac{\rho\,\|\|\mathbf{M}\|\|}{\sqrt{\mu\nu}}\right)\cdot\log^2\left(\frac{1}{\varepsilon}\right)\right).$$

Matrix games with transaction fee - experimental setup

Let $n = m = 10000$, $\mu = 10^{-4}$, $\nu = 1$, and $\varepsilon = 10^{-7}$.
Let $\mathcal{R}_1 = \frac{\mu}{2}\|\cdot\|^2$ and $\mathcal{R}_2 = \frac{\nu}{2}\|\cdot\|^2$.
We generate a sparse, random matrix $\mathbf{M} \in \mathbb{R}^{m\times n}$ s.t. $\|\mathbf{M}\| = 1$.
We choose the **transaction fee** ρ from $\{0.00\%, 0.03\%, \cdots, 0.18\%\}$.
Player 1: $\max_{\mathbf{x}\in\Delta_n} u_1(\mathbf{x},\mathbf{y}) = \mathcal{R}_1(\mathbf{x}) + \langle \mathbf{A}\mathbf{x}, \mathbf{y}\rangle - \mathcal{R}_2(\mathbf{y})\,;$
Player 2: $\max_{\mathbf{y}\in\Delta_m} u_2(\mathbf{x},\mathbf{y}) = -\mathcal{R}_1(\mathbf{x}) + \langle \mathbf{B}\mathbf{x}, \mathbf{y}\rangle + \mathcal{R}_2(\mathbf{y})\,.$

Our theory:

$$\widetilde{\mathcal{O}}\left(\left(\frac{L}{\sqrt{\mu\nu}} + \frac{L}{\min\{\mu,\nu\}}\cdot\frac{\rho\,|||\mathbf{M}|||}{\sqrt{\mu\nu}}\right)\cdot\log^2\left(\frac{1}{\varepsilon}\right)\right)\,.$$

Matrix games with transaction fee - numerical results

Table 2: Gradient queries (in thousands) for ε-Nash under various transaction fees.

Transaction Fee ρ / Methods	0.00%	0.03%	0.06%	0.09%	0.12%	0.15%	0.18%
ICL (Algorithm 1)	$\mathbf{9.1 \pm 0.0}$	$\mathbf{22.6 \pm 0.4}$	$\mathbf{42.2 \pm 0.3}$	$\mathbf{65.0 \pm 0.3}$	$\mathbf{75.7 \pm 0.3}$	113.7 ± 0.7	123.8 ± 0.6
OGDA[13]	93.9 ± 0.5	93.9 ± 0.5	93.9 ± 0.5	93.9 ± 0.5	93.9 ± 0.5	$\mathbf{94.0 \pm 0.6}$	$\mathbf{94.0 \pm 0.6}$
EG[14]	132.9 ± 0.8	132.9 ± 0.8	132.9 ± 0.8	132.9 ± 0.8	132.9 ± 0.8	132.9 ± 0.8	132.9 ± 0.8

Remark (Acceleration conditioning)

In experiments: $\rho \leq 0.12\%$;
In theory: $\rho \ll \sqrt{\mu\nu} = 1\%$.

[13] Leonid Denisovich Popov. "A modification of the Arrow-Hurwitz method of search for saddle points". In: *Mat. Zametki* 28.5 (1980), pp. 777–784.
[14] Galina M Korpelevich. "The extragradient method for finding saddle points and other problems". In: *Matecon* 12 (1976), pp. 747–756.

Matrix games with transaction fee - numerical results

Table 2: Gradient queries (in thousands) for ε-Nash under various transaction fees.

Methods \ Transaction Fee ρ	0.00%	0.03%	0.06%	0.09%	0.12%	0.15%	0.18%
ICL (Algorithm 1)	$\mathbf{9.1 \pm 0.0}$	$\mathbf{22.6 \pm 0.4}$	$\mathbf{42.2 \pm 0.3}$	$\mathbf{65.0 \pm 0.3}$	$\mathbf{75.7 \pm 0.3}$	113.7 ± 0.7	123.8 ± 0.6
OGDA[13]	93.9 ± 0.5	93.9 ± 0.5	93.9 ± 0.5	93.9 ± 0.5	93.9 ± 0.5	$\mathbf{94.0 \pm 0.6}$	$\mathbf{94.0 \pm 0.6}$
EG[14]	132.9 ± 0.8	132.9 ± 0.8	132.9 ± 0.8	132.9 ± 0.8	132.9 ± 0.8	132.9 ± 0.8	132.9 ± 0.8

Remark (Acceleration conditioning)

In experiments: $\rho \leq 0.12\%$;
In theory: $\rho \ll \sqrt{\mu\nu} = 1\%$.

[13] Popov, "A modification of the Arrow-Hurwitz method of search for saddle points".
[14] Korpelevich, "The extragradient method for finding saddle points and other problems".

Convergence plot

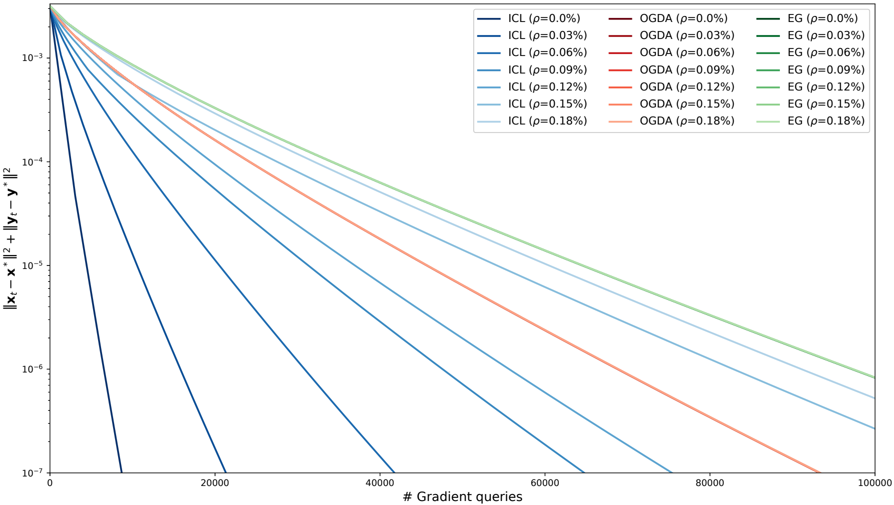

Figure 1: Comparisons of the convergence of the ICL, OGDA, and EG methods.

Conclusion

In this work:

- We define a new, intermediate class of *monotone near-zero-sum games*;
- We propose Iterative Coupling Linearization (ICL), which is faster when the game is *near-zero-sum and with imbalanced conditioning*;
- We apply our method to *regularized matrix games* and *competitive games with small additional incentives*.

Future work:

- Lower complexity bounds;
- Removal of the double logarithm;
- Non-Euclidean spaces (Mirror Prox[15]);
- Applications in other practical settings.

[15] Nemirovski, "Prox-method with rate of convergence O (1/t) for variational inequalities with Lipschitz continuous monotone operators and smooth convex-concave saddle point problems".

Conclusion

In this work:

- We define a new, intermediate class of *monotone near-zero-sum games*;
- We propose Iterative Coupling Linearization (ICL), which is faster when the game is *near-zero-sum and with imbalanced conditioning*;
- We apply our method to *regularized matrix games* and *competitive games with small additional incentives*.

Future work:

- Lower complexity bounds;
- Removal of the double logarithm;
- Non-Euclidean spaces (Mirror Prox[15]);
- Applications in other practical settings.

[15] Nemirovski, "Prox-method with rate of convergence O (1/t) for variational inequalities with Lipschitz continuous monotone operators and smooth convex-concave saddle point problems".

Conclusion

In this work:

- We define a new, intermediate class of *monotone near-zero-sum games*;
- We propose Iterative Coupling Linearization (ICL), which is faster when the game is *near-zero-sum and with imbalanced conditioning*;
- We apply our method to *regularized matrix games* and *competitive games with small additional incentives*.

Future work:

- Lower complexity bounds;
- Removal of the double logarithm;
- Non-Euclidean spaces (Mirror Prox[15]);
- Applications in other practical settings.

[15] Nemirovski, "Prox-method with rate of convergence O (1/t) for variational inequalities with Lipschitz continuous monotone operators and smooth convex-concave saddle point problems".

Conclusion

In this work:

- We define a new, intermediate class of *monotone near-zero-sum games*;
- We propose Iterative Coupling Linearization (ICL), which is faster when the game is *near-zero-sum and with imbalanced conditioning*;
- We apply our method to *regularized matrix games* and *competitive games with small additional incentives*.

Future work:

- Lower complexity bounds;
- Removal of the double logarithm;
- Non-Euclidean spaces (Mirror Prox[15]);
- Applications in other practical settings.

[15] Nemirovski, "Prox-method with rate of convergence O (1/t) for variational inequalities with Lipschitz continuous monotone operators and smooth convex-concave saddle point problems".

Conclusion

In this work:

- We define a new, intermediate class of *monotone near-zero-sum games*;
- We propose Iterative Coupling Linearization (ICL), which is faster when the game is *near-zero-sum and with imbalanced conditioning*;
- We apply our method to *regularized matrix games* and *competitive games with small additional incentives*.

Future work:

- Lower complexity bounds;
- Removal of the double logarithm;
- Non-Euclidean spaces (Mirror Prox[15]);
- Applications in other practical settings.

[15] Nemirovski, "Prox-method with rate of convergence $O(1/t)$ for variational inequalities with Lipschitz continuous monotone operators and smooth convex-concave saddle point problems".

Conclusion

In this work:

- We define a new, intermediate class of *monotone near-zero-sum games*;
- We propose Iterative Coupling Linearization (ICL), which is faster when the game is *near-zero-sum and with imbalanced conditioning*;
- We apply our method to *regularized matrix games* and *competitive games with small additional incentives*.

Future work:

- Lower complexity bounds;
- Removal of the double logarithm;
- Non-Euclidean spaces (Mirror Prox[15]);
- Applications in other practical settings.

[15] Nemirovski, "Prox-method with rate of convergence O $(1/t)$ for variational inequalities with Lipschitz continuous monotone operators and smooth convex-concave saddle point problems".

Conclusion

In this work:

- We define a new, intermediate class of *monotone near-zero-sum games*;
- We propose Iterative Coupling Linearization (ICL), which is faster when the game is *near-zero-sum and with imbalanced conditioning*;
- We apply our method to *regularized matrix games* and *competitive games with small additional incentives*.

Future work:

- Lower complexity bounds;
- Removal of the double logarithm;
- Non-Euclidean spaces (Mirror Prox[15]);
- Applications in other practical settings.

[15] Nemirovski, "Prox-method with rate of convergence O $(1/t)$ for variational inequalities with Lipschitz continuous monotone operators and smooth convex-concave saddle point problems".

Conclusion

In this work:

- We define a new, intermediate class of *monotone near-zero-sum games*;
- We propose Iterative Coupling Linearization (ICL), which is faster when the game is *near-zero-sum and with imbalanced conditioning*;
- We apply our method to *regularized matrix games* and *competitive games with small additional incentives*.

Future work:

- Lower complexity bounds;
- Removal of the double logarithm;
- Non-Euclidean spaces (Mirror Prox[15]);
- Applications in other practical settings.

---

[15] Nemirovski, "Prox-method with rate of convergence $O(1/t)$ for variational inequalities with Lipschitz continuous monotone operators and smooth convex-concave saddle point problems".