Asymmetric Perturbation in Solving Bilinear Saddle-Point Optimization

ICML 2026 · Oral Presentation

CyberAgent **AI Lab**

Kenshi Abe · Mitsuki Sakamoto · Kaito Ariu (CyberAgent) · Atsushi Iwasaki (UEC)

Bilinear saddle-point optimization

We solve a two-player zero-sum bilinear game:

$$\min_{x \in \mathcal{X}} \max_{y \in \mathcal{Y}} x^\top A y$$

- A — the $m \times n$ game (payoff) matrix
- $x \in \mathcal{X}$, $y \in \mathcal{Y}$ — mixed strategies on polytopes

EXAMPLES

Two-player zero-sum normal-form games

Two-player zero-sum extensive-form games — under sequence-form strategy representation

SOLUTION CONCEPT

Nash equilibrium (x^*, y^*) — neither player can profitably deviate from its strategy:

$$\forall (x, y) \in \mathcal{X} \times \mathcal{Y}, \quad (x^*)^\top A y \le (x^*)^\top A y^* \le x^\top A y^*$$

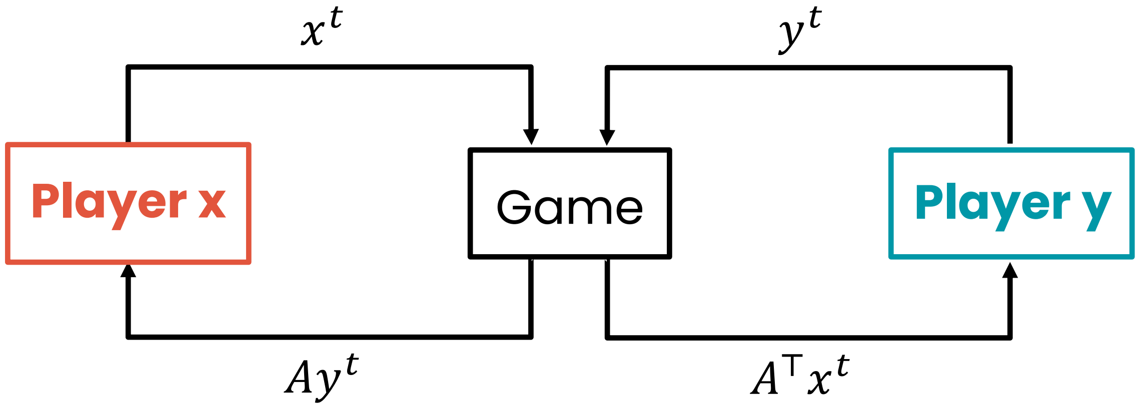
- x^* — minimax strategy
- y^* — maximin strategy

Online learning in games

We consider computing equilibrium via online learning.

In the online learning setting, at each iteration $t \in [T]$:

1. Players submit their current strategies x^t (or y^t) to the game.
2. Each player observes the gradient vector Ay^t (or $A^\top x^t$) of its payoff.
3. Players update their strategies x^{t+1} (or y^{t+1}) based on the observed gradient.



Gradient Descent-Ascent (GDA)

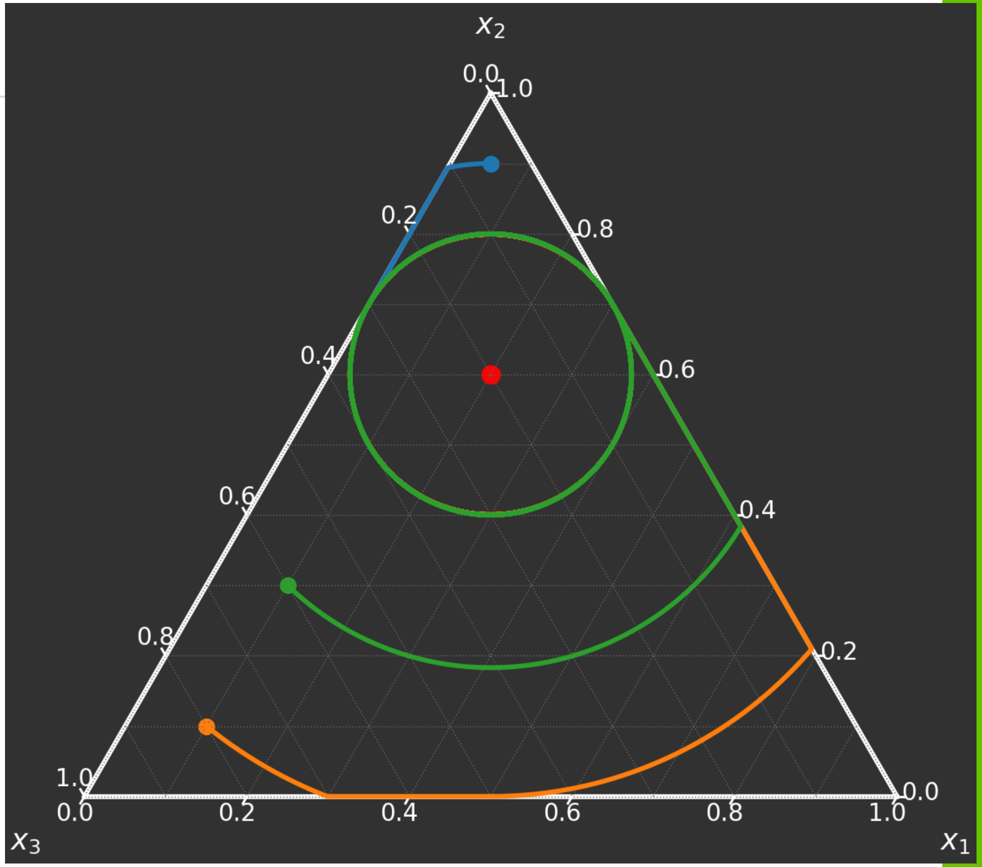
Each player takes a projected gradient step:

$$x^{t+1} = \Pi_{\mathcal{X}}(x^t - \eta A y^t)$$
$$y^{t+1} = \Pi_{\mathcal{Y}}(y^t + \eta A^\top x^{t+1})$$

PROBLEM

The strategy sequence may fail to converge.

[Mertikopoulos et al., 2018]



GOAL

Develop a learning algorithm whose strategies directly converge to a minimax (resp. maximin) strategy x^* (resp. y^*).

Payoff perturbation

In payoff perturbation, each player adds a strongly-convex penalty to its own payoff:

[Facchinei & Pang, 2003]

$$x^{t+1} = \Pi_{\mathcal{X}}\left(x^t - \eta(Ay^t + \mu x^t)\right)$$
$$y^{t+1} = \Pi_{\mathcal{Y}}\left(y^t + \eta(A^\top x^{t+1} - \mu y^t)\right)$$

Equivalently, solving a perturbed game:

$$\min_{x \in \mathcal{X}} \max_{y \in \mathcal{Y}} \left\{ x^\top A y + \frac{\mu}{2} ||x||^2 - \frac{\mu}{2} ||y||^2 \right\}$$

μ: the strength of the penalty

KEY PROPERTY

Strongly-convex–strongly-concave $\Rightarrow$ **linear convergence** to (x^μ, y^μ).

(x^μ, y^μ): the unique equilibrium of the perturbed game.

The limitation of symmetric perturbation

The perturbed solution **only approximates the true equilibrium**, and the error grows with μ:

$$\mathrm{dist}(x^{\mu}, \mathcal{X}^*) = \mathcal{O}(\mu)$$

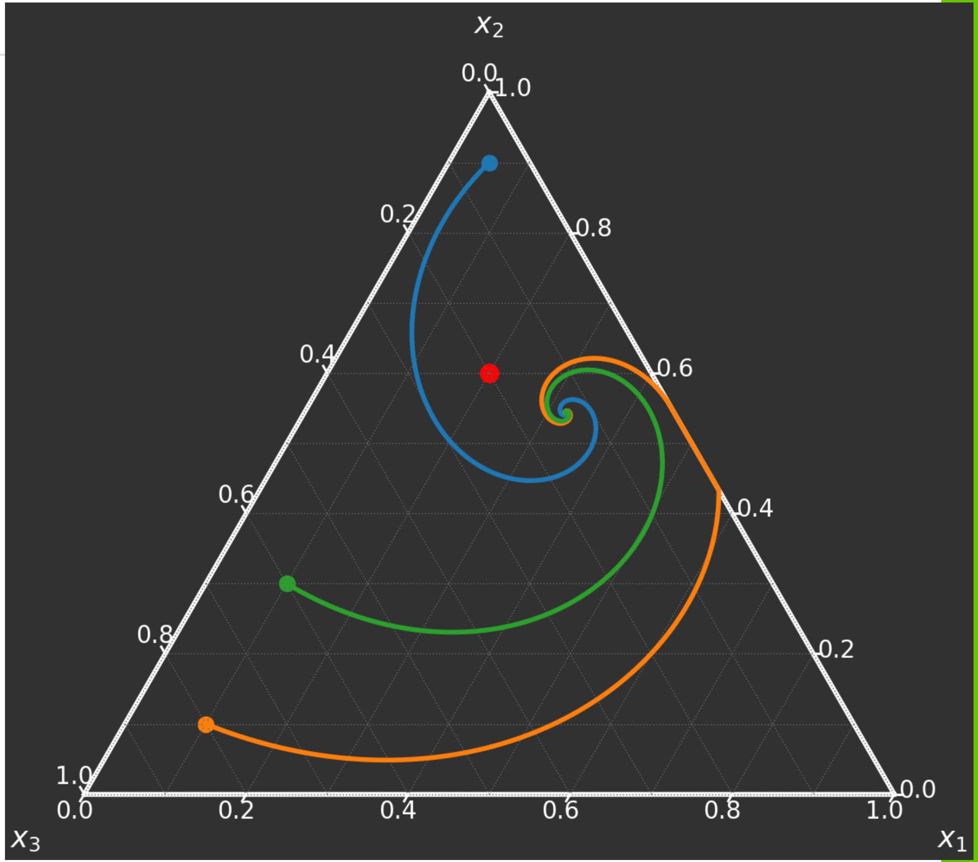
[Liu et al., 2023; Abe et al., 2024]

$\mathcal{X}^*$: the set of minimax strategies

CONSEQUENCE

μ must go to zero to remove the error (e.g., $\mu = \mathcal{O}(1/t)$).

→ **no faster than $\tilde{\mathcal{O}}(1/t)$ rates.**



QUESTION

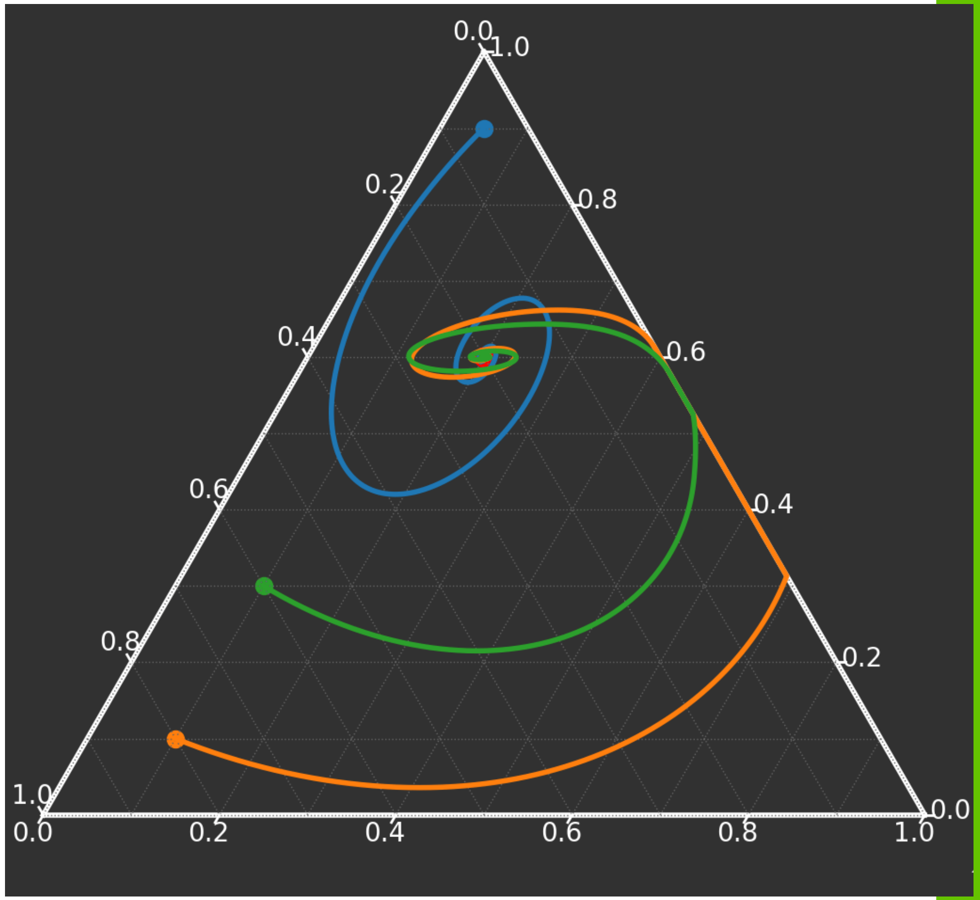
Can we recover an original-game equilibrium strategy without $\mu \to 0$?

Our idea: asymmetric perturbation

Perturb **only player x's payoff**;
player y's payoff remains unperturbed:

ASYMMETRIC PERTURBATION (OURS)

$$\min_{x\in\mathcal{X}} \max_{y\in\mathcal{Y}} \left\{x^\top A y + \frac{\mu}{2}||x||^2 \cancel{- \frac{\mu}{2}||y||^2}\right\}$$



Equilibrium invariance property

Under **asymmetric perturbation**:

THEOREM (INFORMAL)

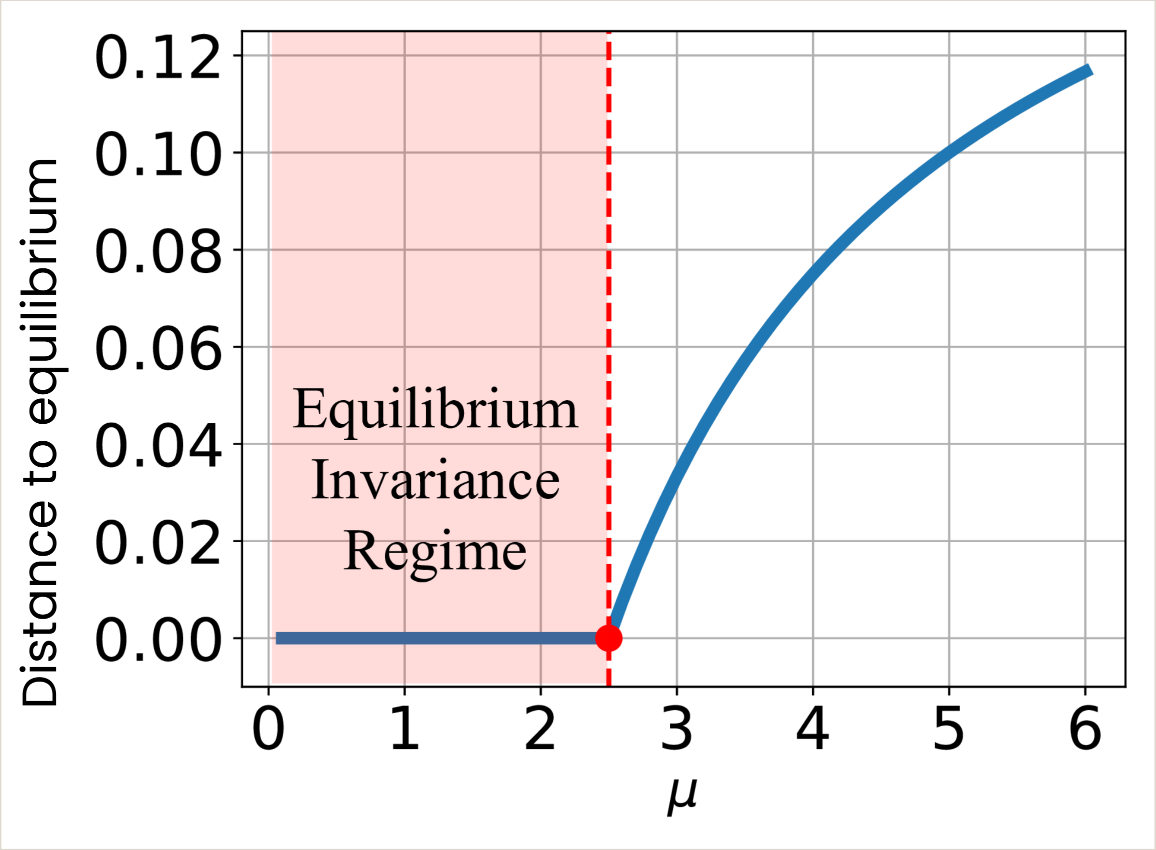
μ below a game-dependent threshold

$$\Rightarrow \ \mathrm{dist}(x^{\mu}, \mathcal{X}^{*}) = 0$$

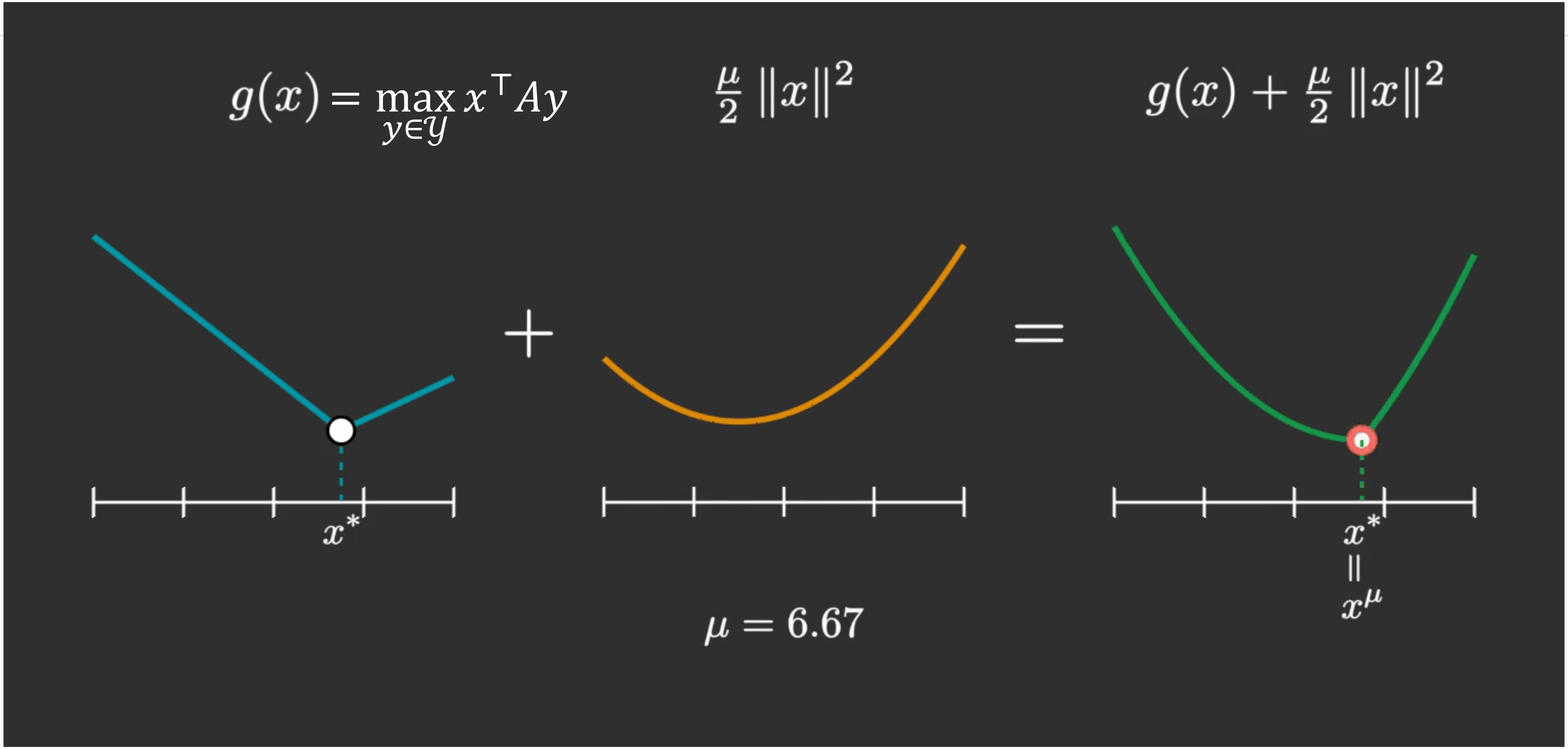
$\mathcal{X}^{*}$: the set of minimax strategies of the original game

CONSEQUENCE

No need to drive μ to zero;
a sufficiently small μ is enough.



Why it works?: near-linear growth



Asymmetrically Perturbed GDA (AsymP-GDA)

A GDA-style learning algorithm for solving asymmetrically perturbed games:

ASYMP-GDA (OURS)

$$x^{t+1} = \Pi_{\mathcal{X}}\left(x^t - \eta(Ay^t + \mu x^t)\right)$$

$$y^{t+1} = \Pi_{\mathcal{Y}}\left(y^t + \eta(A^\top x^{t+1} \cancel{- \mu y^t})\right)$$

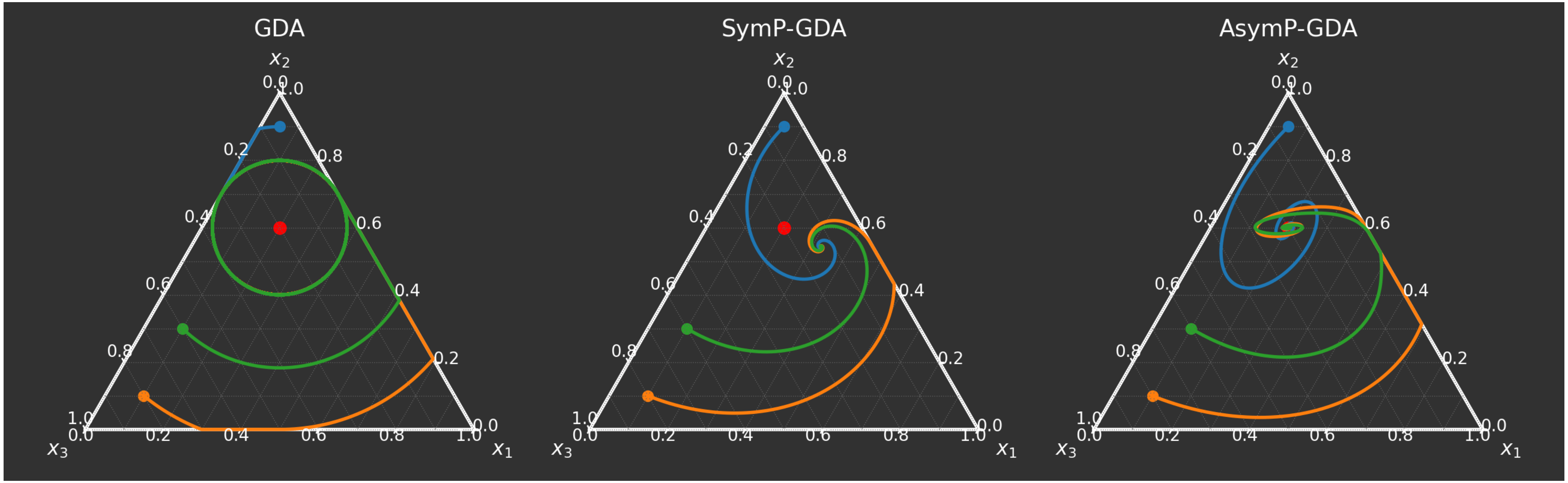
Convergence guarantee

THEOREM (INFORMAL)

μ below a game-dependent threshold

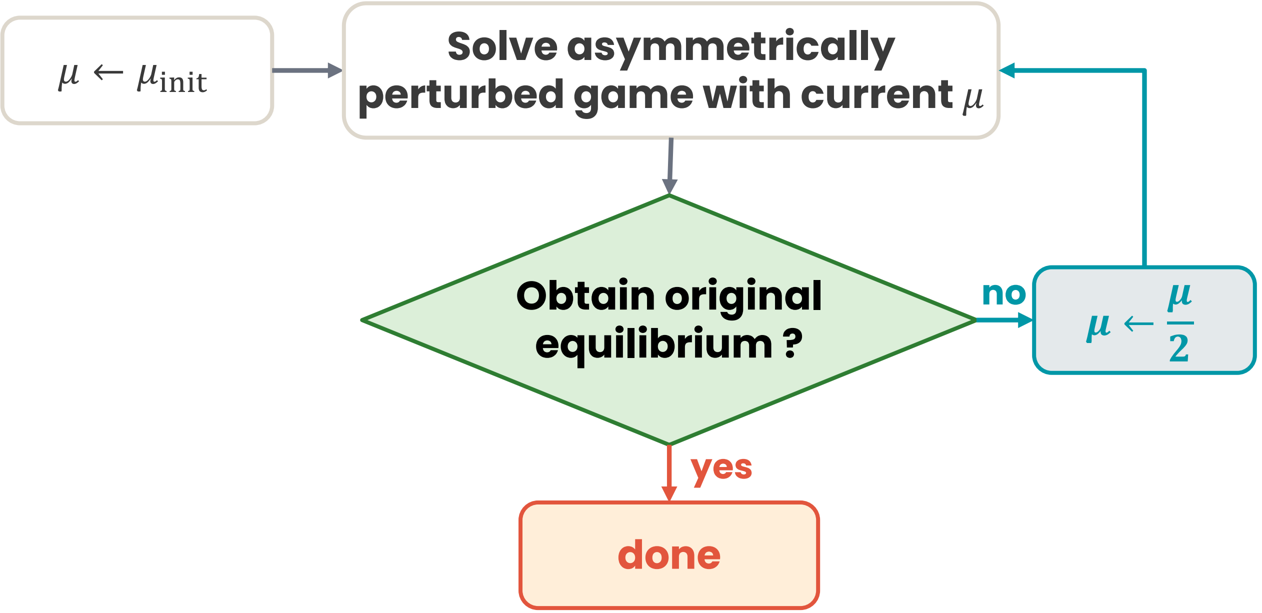
$\Rightarrow$ $\mathrm{dist}(x^t, \mathcal{X}^*) = \exp\big(-\Omega(t)\big)$ for the perturbed player x!

Provably faster *than the $\tilde{O}(1/t)$ rate of symmetric perturbation in the same bilinear setting.*



Parameter-free variant

Equilibrium invariance needs μ **below an unknown threshold**.
⇒ We remove this requirement **while keeping the linear rate!**



$\mu \leftarrow \mu_{\text{init}}$

Solve asymmetrically perturbed game with current μ

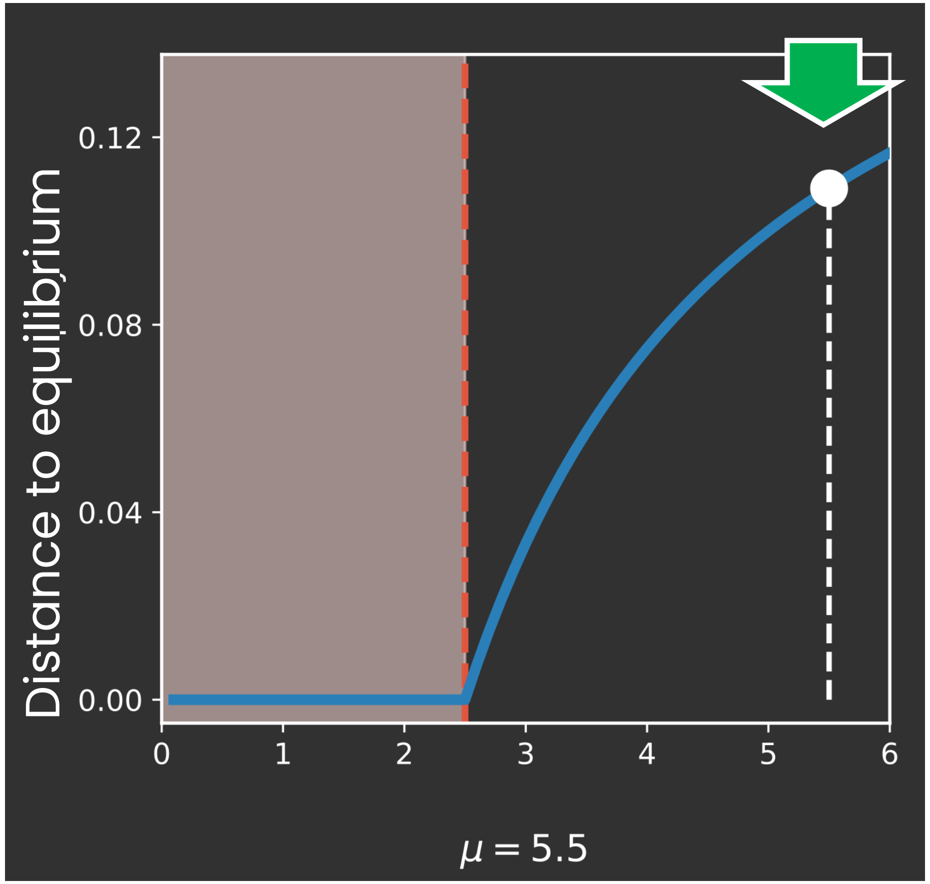
Obtain original equilibrium ?

no

$\mu \leftarrow \dfrac{\mu}{2}$

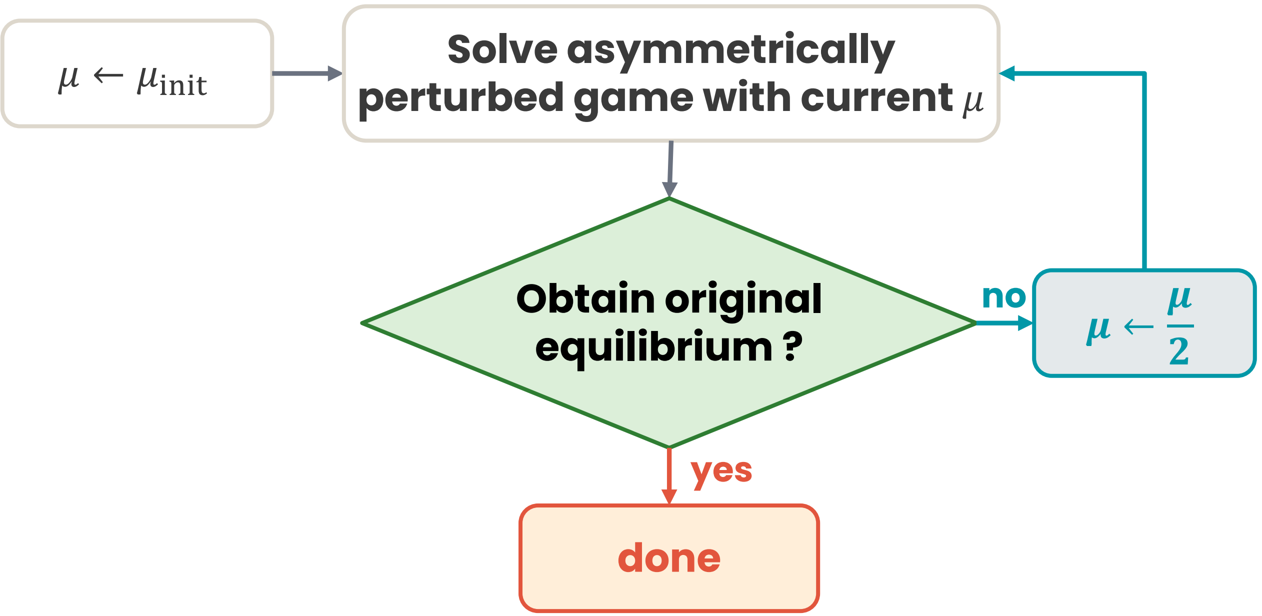
yes

done



Parameter-free variant

Equilibrium invariance needs μ **below an unknown threshold**.
⇒ We remove this requirement **while keeping the linear rate!**



$\mu \leftarrow \mu_{\text{init}}$

Solve asymmetrically perturbed game with current μ

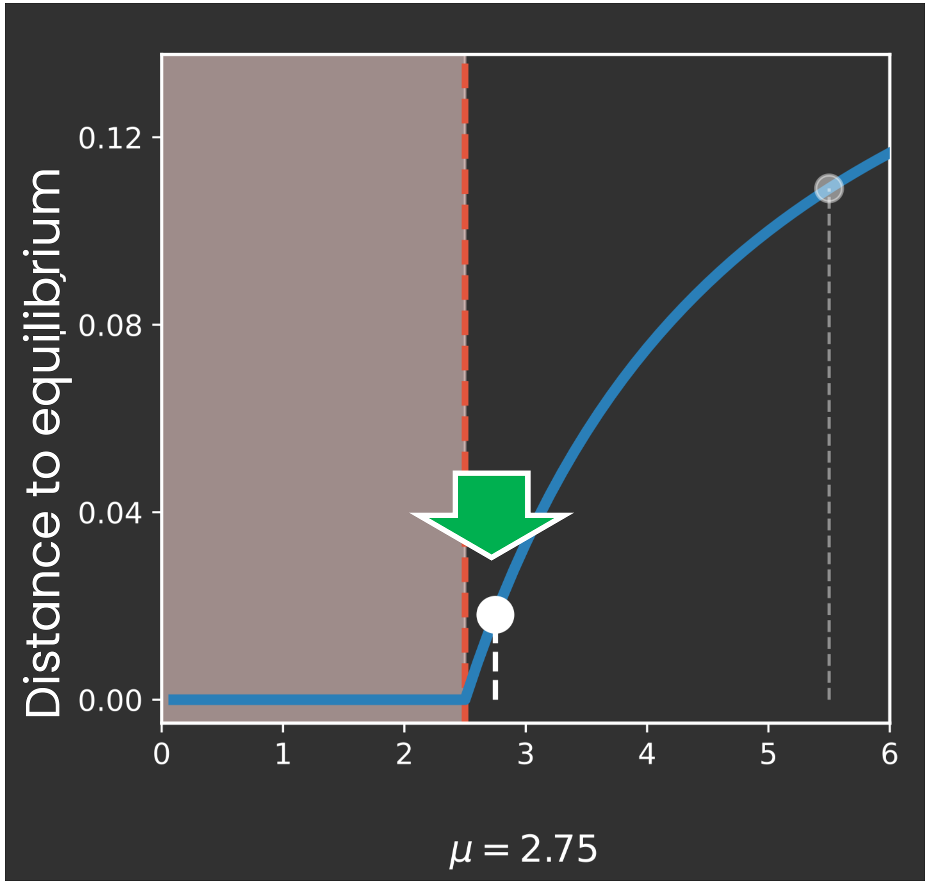
Obtain original equilibrium ?

no

$\mu \leftarrow \frac{\mu}{2}$

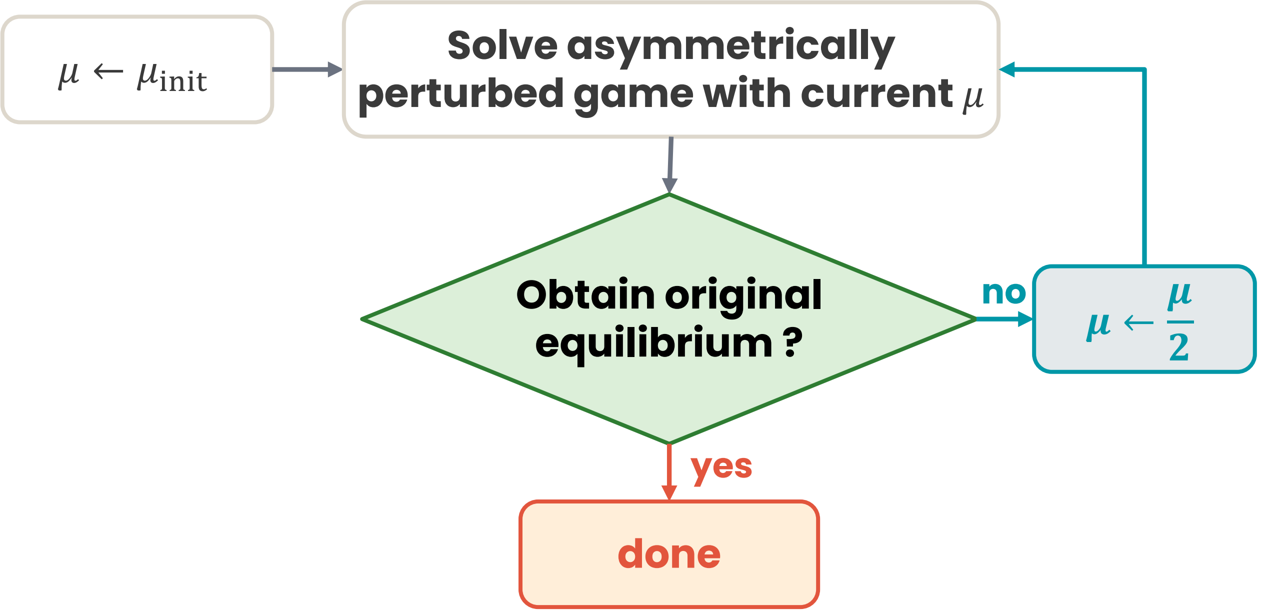
yes

done



Parameter-free variant

Equilibrium invariance needs μ **below an unknown threshold**.
⇒ We remove this requirement **while keeping the linear rate!**



$\mu \leftarrow \mu_{\text{init}}$

Solve asymmetrically perturbed game with current μ

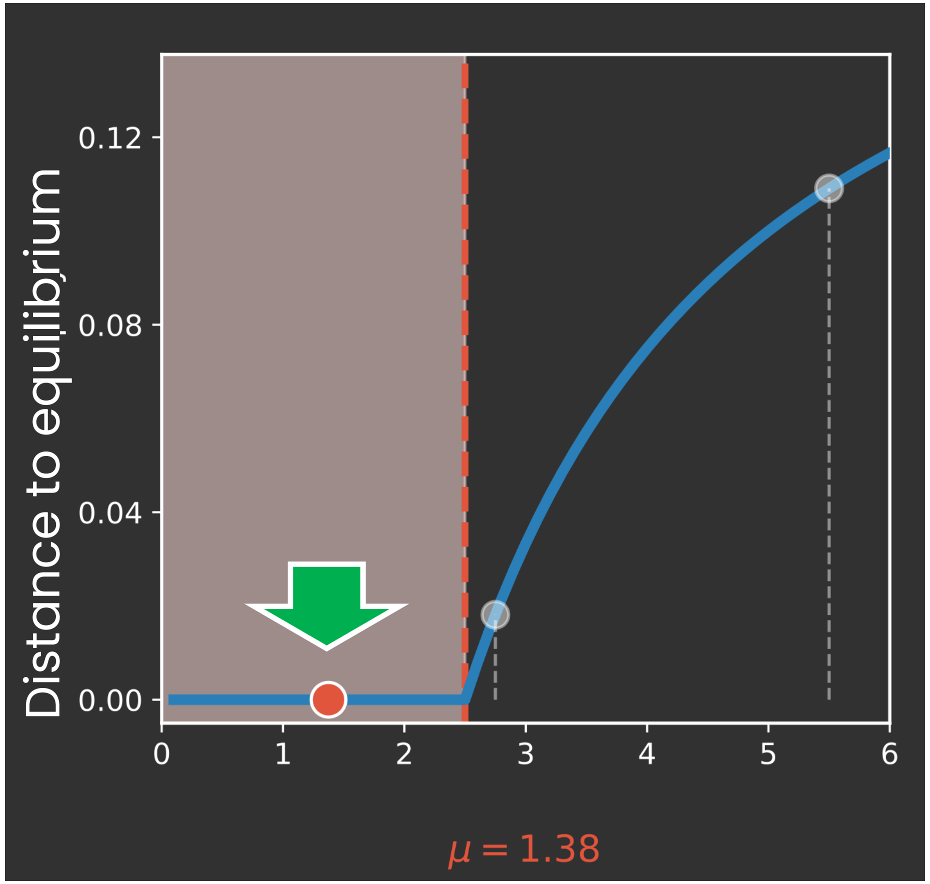
Obtain original equilibrium ?

no

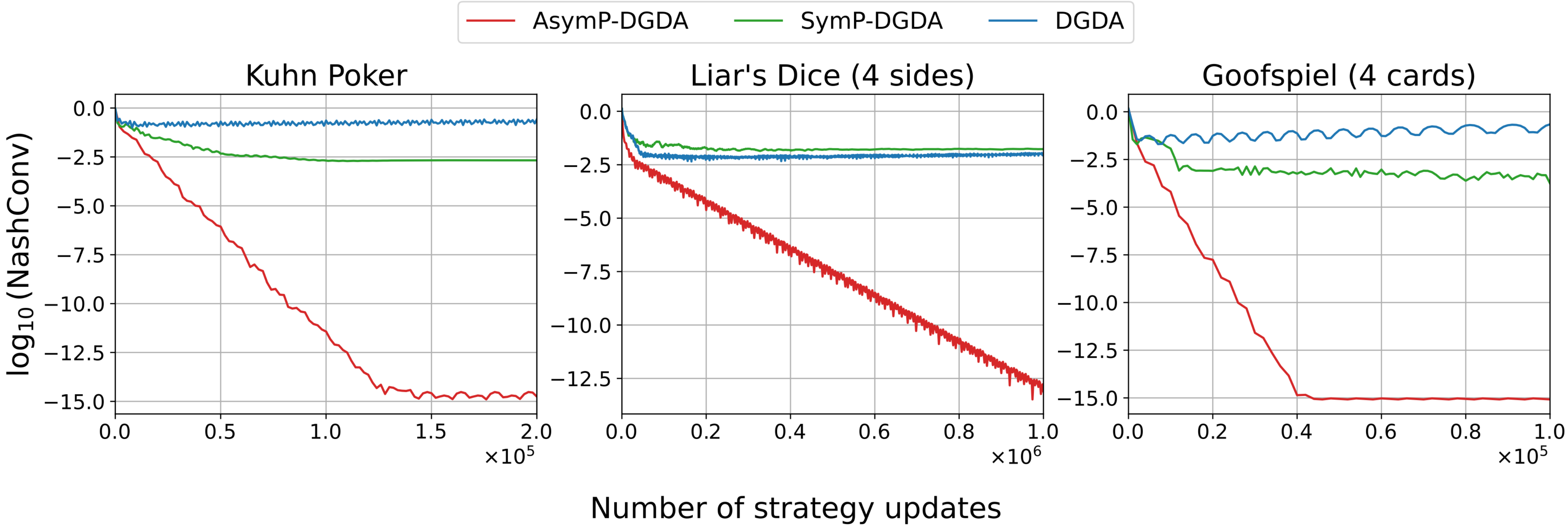
$\mu \leftarrow \frac{\mu}{2}$

yes

done



Experiments in extensive-form games



Asymmetric perturbation beats symmetric perturbation in extensive-form games!

TAKEAWAYS

Visit our poster!

Hall A #4617 · Jul 7 · 2:00–3:45 PM

Equilibrium invariance	Perturbing one player makes the perturbed equilibrium coincide exactly with the original — no need to drive μ to zero.
Linear rate	AsymP-GDA gives linear convergence, beating the $\tilde{O}(1/t)$ of symmetric perturbation.
Parameter-free	A halving procedure retains a linear rate, without knowledge of game-dependent constants.

Code: github.com/CyberAgentAILab/asymmetrically-perturbed-gda

References

- **[Abe et al., 2024]** Abe, K., Ariu, K., Sakamoto, M., and Iwasaki, A. “Adaptively perturbed mirror descent for learning in games.” In ICML, 2024.
- **[Facchinei & Pang, 2003]** Facchinei, F. and Pang, J.-S. “Finite-dimensional variational inequalities and complementarity problems.” Springer, 2003.
- **[Liu et al., 2023]** Liu, M., Ozdaglar, A., Yu, T., and Zhang, K. “The power of regularization in solving extensive-form games.” In ICLR, 2023.
- **[Mertikopoulos et al., 2018]** Mertikopoulos, P., Papadimitriou, C., and Piliouras, G. “Cycles in adversarial regularized learning.” In SODA, 2018.