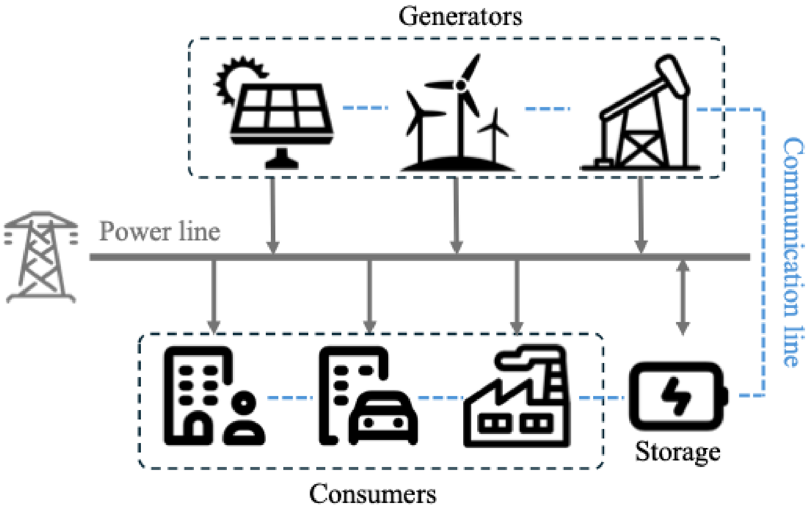
Incentive Design for Safe Nash Equilibrium Learning in Large Populations via Control Barrier Functions

Yifeng Xiao[1], Jair Cert´orio[2], Nuno C. Martins[2], Yasser Shoukry[3], Pierluigi Nuzzo[1]

[1]University of California, Berkeley
[2]University of Maryland, College Park
[3]University of California, Irvine

The Challenge: Safety-Aware Incentive Design for Population Games

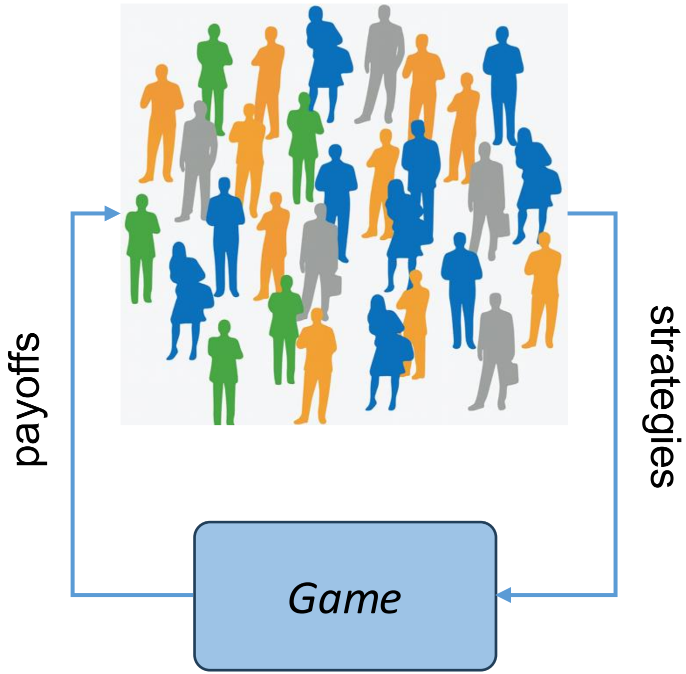


Systems whose behaviors emerge from the strategic decisions of a population of agents.

- **Evolutionary Dynamics Model (EDM):** governs how the population updates its strategies over time.
- **Exogenous System (ES) dynamics:** describe how the system's state evolves in response to the population's decisions.

*How to design incentives that **guide population behavior** while ensuring **safety and stability**?*

Population Game Dynamics



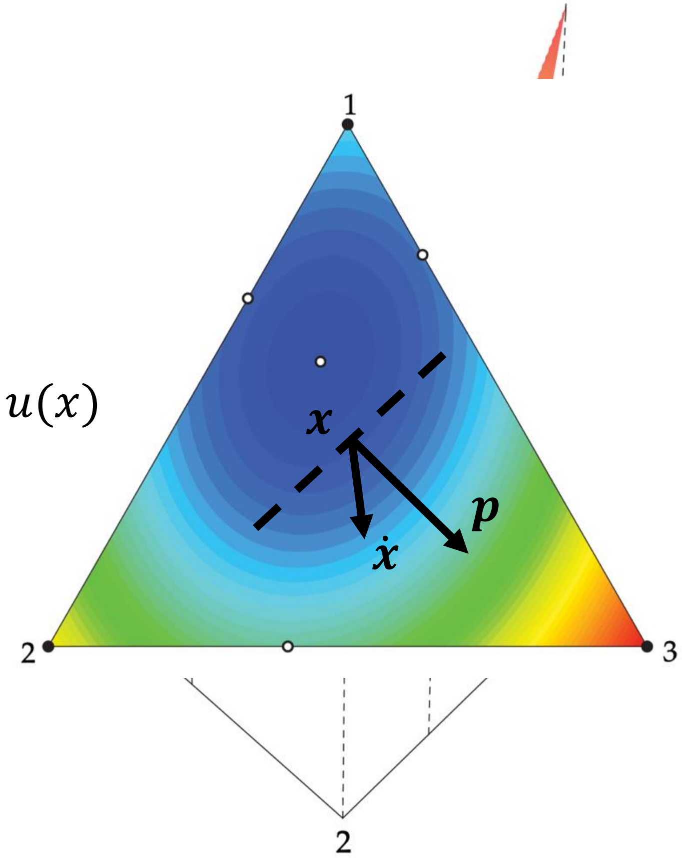
- *Strategy set*: $\{1, \ldots, n\}$, the set of strategies available to agents.
- *Population states*
  - x_i: the fraction of the population selecting strategy i
  - $x = \{x_1, \ldots, x_n\}$ is referred to as the population state
- *Payoff function*
  - p_i: payoff obtained by following the i-th strategy
  - $p = \{p_1, \ldots, p_n\}$ is the payoff vector
- *EDM* specifies how the population state evolves in response to given payoffs.

$$\dot{x}(t) = V(p(t), x(t))$$

$$V_i(p, x) := \sum_{j=1}^{n} (x_j \tau_{ji}(p, x) - x_i \tau_{ij}(p, x))$$

where $\tau\colon \mathbb{R}^n \times \mathbb{X} \to \mathbb{R}^{n \times n}$ is a learning rule, i.e., revision protocol

Positive Correlation and Potential Function



- A learning rule τ is ***positive correlated (PC)*** if for all $(x, p) \in \mathbb{X} \times \mathbb{R}^n$,

$$\dot{x} \neq 0 \leftrightarrow p^\top \dot{x} > 0$$

Consider a potential function $u(x)$ capturing information about payoffs,

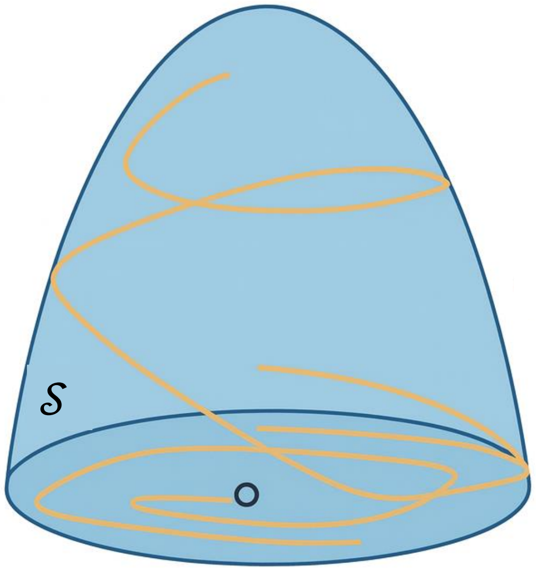
$$p(x) = \nabla_x u(x)$$

then the state trajectory always evolves in the direction that increases the potential if $\dot{x} \neq 0$.

- A learning rule τ is ***Nash stationary (NS)*** if, given the best response map $\mathcal{M}(p) \coloneqq argmax_x p^\top x$:

$$\dot{x} = 0 \leftrightarrow x \in \mathcal{M}(p)$$

Control Barrier Function (CBF)



Given a system of the standard form:

$$\dot{z} = f(z) + g(z)u,$$

with state $z \in \mathbb{D} \subset \mathbb{R}^n$, input $u \in \mathbb{U} \subset \mathbb{R}^m$, and functions f, g are Lipschitz continuous.

Consider $\mathcal{S}$ and $h(z)$:

$$\mathcal{S} = \{z \in \mathbb{D} : h(z) \geq 0\},$$
$$\partial\mathcal{S} = \{z \in \mathbb{D} : h(z) = 0\},$$
$$Int(\mathcal{S}) = \{z \in \mathbb{D} : h(z) > 0\}.$$

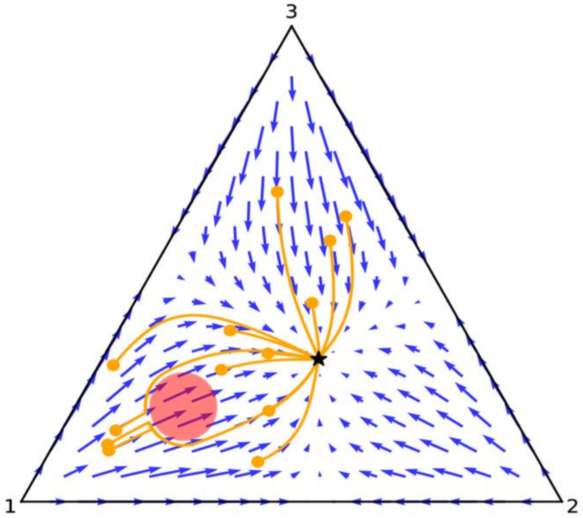
$h(z)$ is a CBF if $\nabla_z h(z) \neq 0$ for all $z \in \partial\mathcal{S}$ and there exists an extended class $\mathcal{K}$ function α, such that for all $z \in \mathcal{S}$, there is u such that

$$L_f h(z) + L_g h(z)u \geq -\alpha(h(z))$$

where L_f and L_g denote the Lie derivatives along f and g.

CBF-Based Population State Avoidance

Given an EDM $\dot{x}(t)$ satisfying the PC and NS conditions, and an unsafe region Ω_x along with CBF $h(x)$, the payoff $P(x)$ can be designed as:



$$P(x) = \nabla_x h(x), \quad x \in \partial\Omega_x.$$

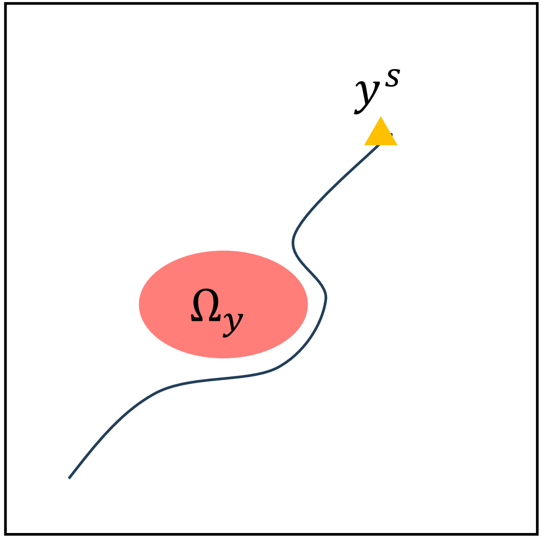
Nagumo's theorem:

$$\mathbb{X} \backslash \Omega_x \text{ is invariant} \iff \dot{h}(x) \geq 0, \quad x \in \partial\Omega_x.$$

PC condition:

$$\dot{x}^\top p = \dot{x}^\top \nabla_x h(x) \geq 0 \iff \dot{h}(x) \geq 0.$$

Exogenous System Dynamics



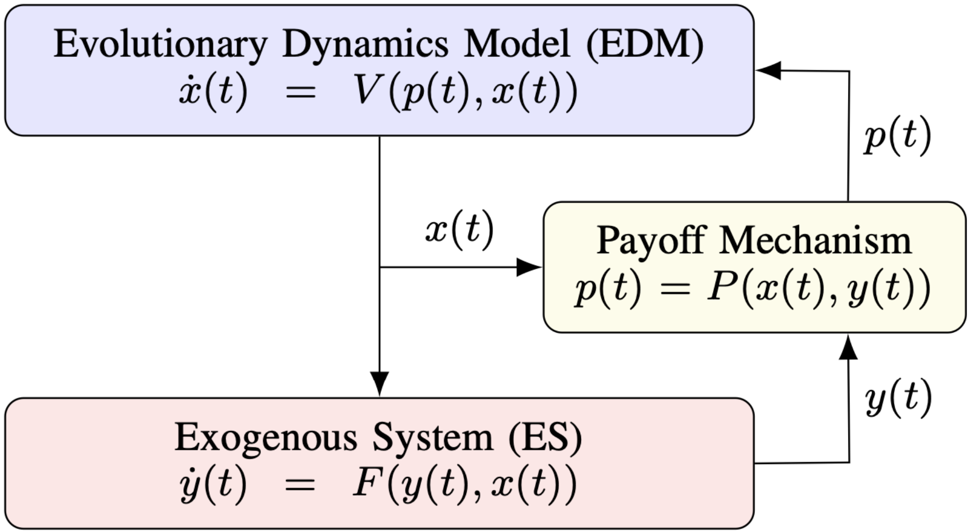
Consider an ES coupled with the EDM and model it with a general affine function:

$$\dot{y} = F\big(y(t), x(t)\big) = f\big(y(t)\big) + g\big(y(t)\big)^\top x(t),$$

where x is the population state, where $f\colon \mathbb{Y} \to \mathbb{R}^n$ and $g\colon \mathbb{Y} \to \mathbb{R}^{n\times m}$ are Lipschitz continuous functions. We assume that if $x(t) \equiv x$, the system has a unique equilibrium point $y^*(x)$.

How to design incentives that ***guide population behavior*** *while ensuring* ***safety and stability*** *in exogeneous system?*

System coupling with an EDM and an ES

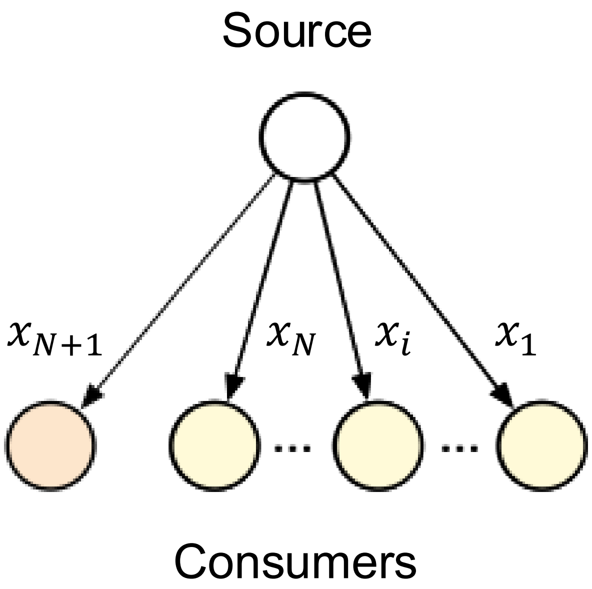


- ES evolves with respect to the population's decisions.
- Unsafe region Ω_y is defined over y.
- Target point y^s

Goal: Design a payoff function $P(x)$ to ensure

$$y(t) \notin \Omega_y, \quad t \geq 0,$$
$$y(t) \to y^s \ as\ t \to \infty.$$

Control Limits and Equilibrium Points



In coupled system, e.g., resource allocation, x_i denotes the proportion of control resources.

$$\sum_{i=1}^{N} x_i(t) = 1$$

Introduce an auxiliary strategy x_{N+1}:

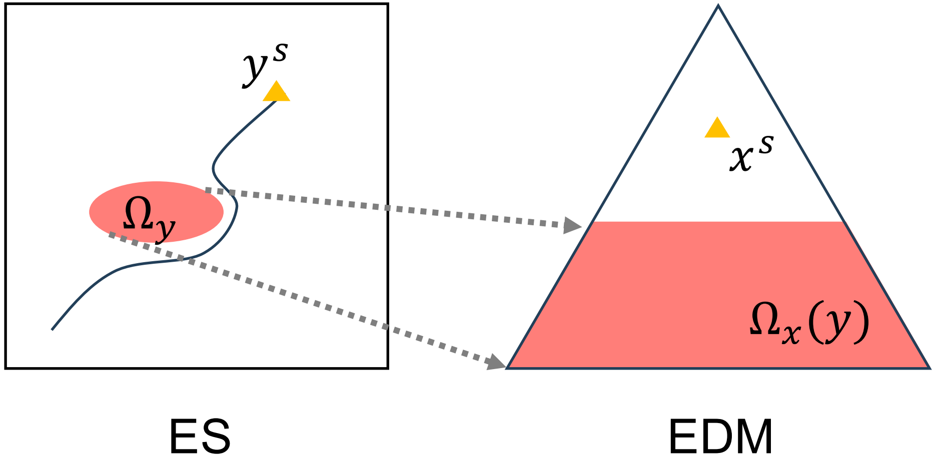
$$\sum_{i=1}^{N+1} x_i(t) = 1, \ \ \sum_{i=1}^{N} x_i(t) \leq 1$$

Equilibrium Point:

$$\dot{y} = 0 \implies x^s = \underline{-(g(y^s)^{-1}})^\top f(y^s), \quad x_{N+1}^s = 1 - \sum_{i=1}^{N} x_i^s.$$

invertible

CBF-Based Safe Payoff Design



Attractive potential:

$$u_a(x) := -\frac{K_a}{2}\|x - x^s\|_2^2 + \frac{K_a}{2},$$

Repulsive potential:

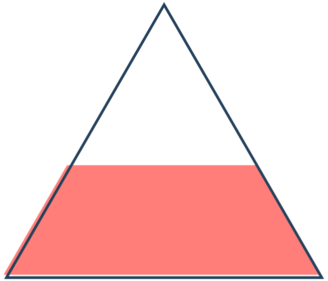
$$u_r(x, y) := \underline{K_r \phi(x, y)},$$

How to connect ϕ with $h(y)$?

Connect ϕ with $h(y)$

Based on CBF definition, the ES state y will always be in $\mathbb{Y} \backslash \Omega_y$ as long as the control action x is from

$$\mathcal{K}_{cbf}(y) = \{x \in \mathbb{X} : \phi(x, y) \geq 0\},$$
$$\phi(x, y) = L_f h(y) + L_g h(y) x + \alpha(h(y)).$$



$\Omega_x(y) = \mathbb{X} \backslash \mathcal{K}_{cbf}(y)$

$\phi(x, y)$ can be used in the potential function:

$$u_r(x, y) := K_r \phi(x, y)$$

Combined potential function:

Dominates when $\phi(x, y) > \delta$

$$u(x, y) = w(x, y) u_a(x) + (1 - w(x, y)) u_r(x, y)$$

Dominates when $\phi(x, y) \rightarrow \delta$, for a small δ

Stability Analysis and Safety Condition

Payoff function:

$$P(x,y) = \gamma \nabla_x u(x,y),$$

scaling factor

Lyapunov function:

$$U(x,y) = -u(x,y) + \frac{K_a}{2} > U(x^s, y^s) = 0,$$

for all $(x,y) \neq (x^s, y^s)$.

And we want to prove the time derivative of $U(x,y)$ is negative definite.

$$\dot{U}(x,y) < 0 \iff \dot{u}(x,y) > 0.$$

$$\dot{u}(x,y) > 0 \iff \dot{x}^\top \nabla_x u(x,y) + \dot{y}^\top \nabla_y u(x,y) > 0$$

?

$$\frac{1}{\gamma} \dot{x}^\top p > 0$$

Stability Analysis and Safety Condition

$$\dot{x}^\top \nabla_x u(x,y) + \dot{y}^\top \nabla_y u(x,y) > 0$$

$$\underset{x,y}{\text{minimize}} \quad F(y,x)^\top \nabla_y u(x,y),$$

$$\text{s.t.} \quad \delta \geq \phi(x,y) \geq 0, \quad \sum_{i=1}^{N} x_i \leq 1.$$

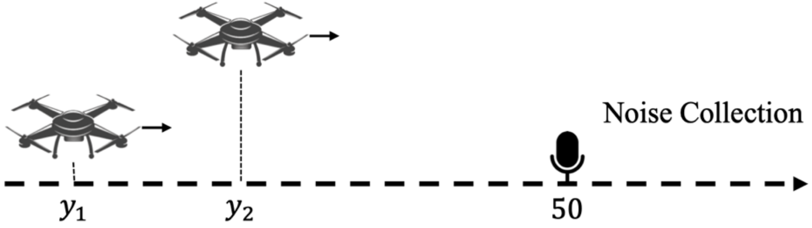
If $u_y^* > 0$, we get $\dot{u}(x,y) > 0$, otherwise

Theorem. Suppose the EDM satisfies the PC and NS conditions, and for any $\beta > 0$, there exists a factor γ such that $\beta V(p,x) \leq V(\gamma p, x)$. If the initial state (x_0, y_0) satisfies $\phi(x_0, y_0) \geq 0$ and the target state (x^s, y^s) satisfies $\phi(x^s, y^s) > \delta$, the payoff function is designed as

$$P(x,y) = \nabla_x u(x,y),$$

which ensures the safety and asymptotic stability of the coupled system.

Simplified Drone System with Noise Control



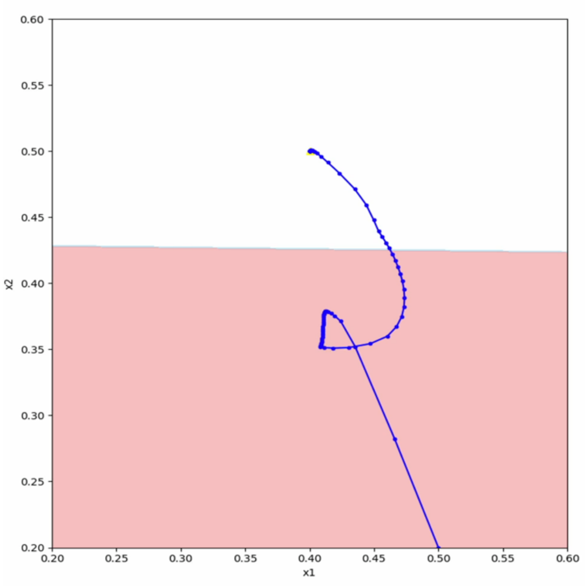
- Two active drones with total resource constraints.
- EDM - $x(t)$ - resource fractions; ES - $y(t)$ - drone position.
- The restricted region is defined by the combined drone-collector distance:

$$\Omega_y = \{y \in \mathbb{R}^2 : \left\| y - 50 \right\|_2 \leq 5\}$$

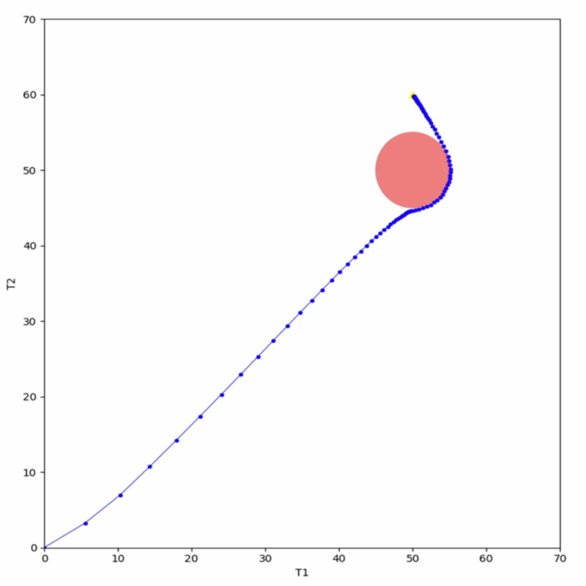
Two drones cannot be close to the collector simultaneously.

Goal: Allocate control resources to guide the drones to a target while satisfying the constraints.

Safe Control with CBF-Based Payoff Design

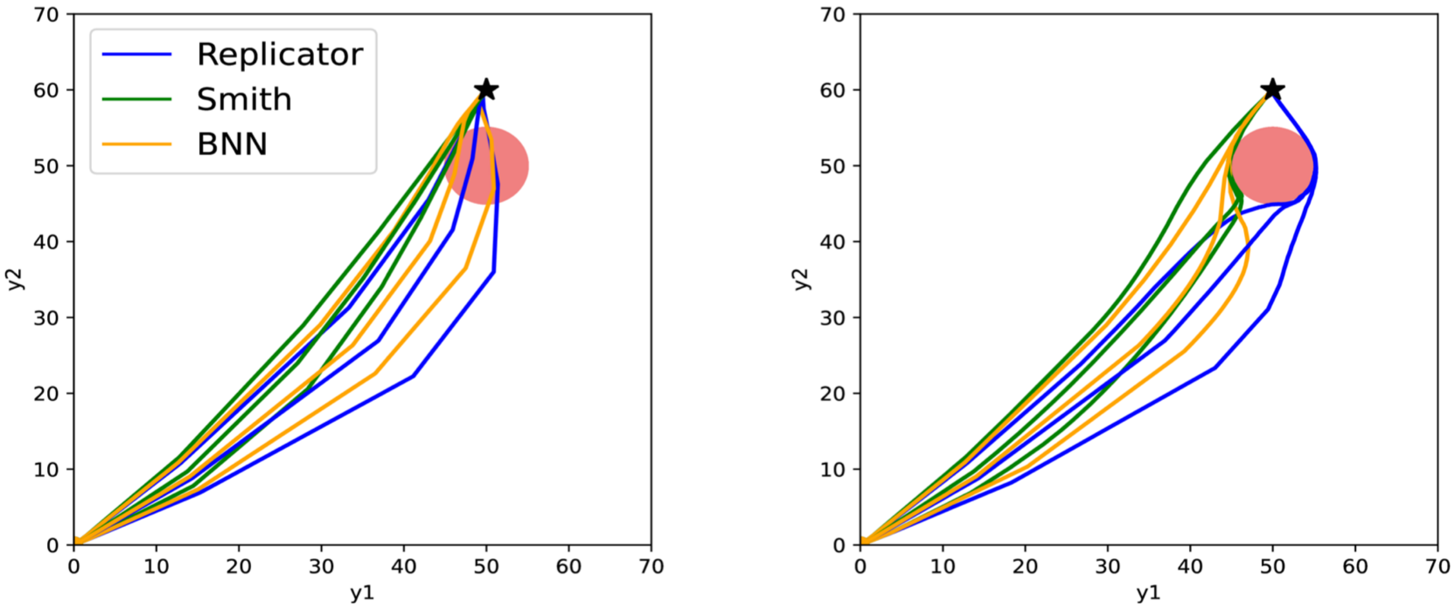


EDM State: Resource evolution



ES State: Drone position

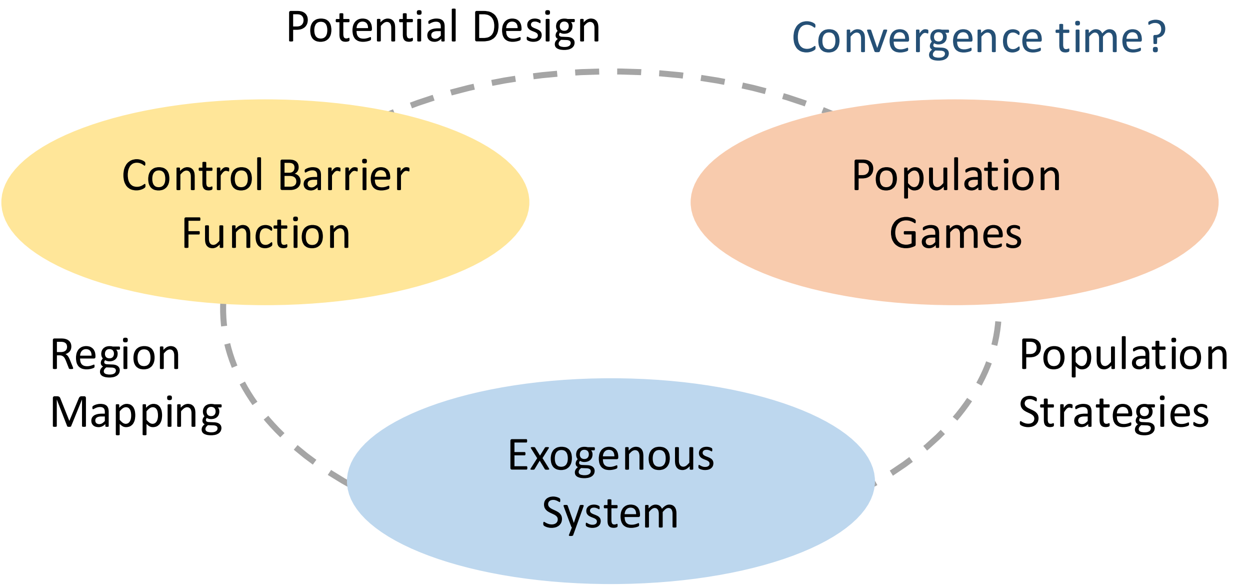
ES Trajectories with Different EDMs



Conclusions

Robustness?

Convergence time?



High-Order System?

Real-world Applications?