Stochastic Optimal Control of Linear-Quadratic Altruistic Systems with Perfect State Information

Amit Enbal — PH.D. Candidate

Supervisor: Yaakov Oshman — Prof. Em.

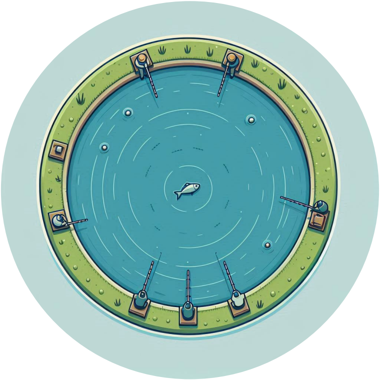
July 14th, 2025

Introduction

Motivation

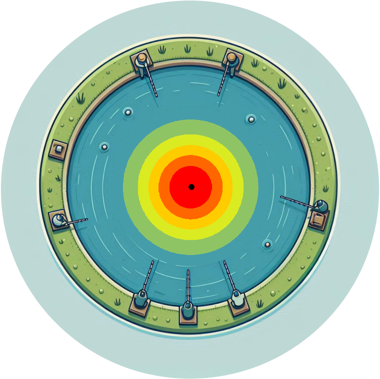
- Consider a group of fishermen $\{x_n\}_{n=1}^{N}$ throwing baits to catch a fish.
- Fish's position X is uncertain, with a known distribution.
- Problem: place $\{x_n\}_{n=1}^{N}$ to minimize the expected miss

$$\mathbb{E}\left[\min_n \|x_n - X\|^2\right].$$



Motivation (ctd')

- Observations:
  1. Problem posses inherent redundancy.
  2. Agent-wise optimal strategy is clearly not group-optimal.
- Conclusion: optimal group performance force fishermen to sacrifice individual performance.
- Example of an **Altruistic system**.



Outline

1 Introduction

2 Static Altruistic Systems

3 The Linear Quadratic Altruistic Problem

4 The Predictive LQA Problem

5 Application to Cooperative Interception

6 Conclusions

Static Altruistic Systems

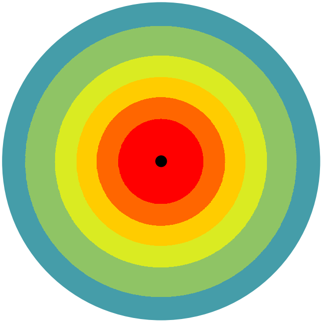
Individual Performance

- $\mathcal{X} = \{x_n\}_{n=1}^{N} \subset \mathbb{R}^s$ is a system of static agents.
- $X : \Omega \to \mathbb{R}^s$ is a random vector with an absolutely continuous measure $\mathbb{P}$.
- The individual cost of the n-th agent is

$$\mathcal{J}_n(x_n) = \mathbb{E}[g(x_n, X)] \,.$$

- The egoistical law of the n-th agent is

$$\mathring{x}_n := \operatorname{argmin} \mathcal{J}_n(x_n).$$



$$\mathcal{J}_n(x_n) = \mathbb{E}\left[\|x_n - X\|^2\right],$$
$$X \sim \mathcal{N}(0, I)$$

Collective Performance

- $\mathcal{X}$ is called altruistic if its collective performance is measured by

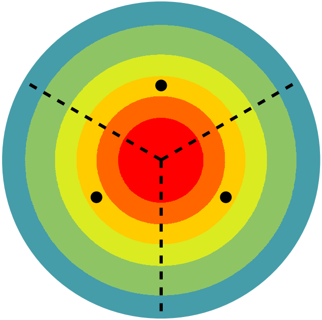
$$\mathcal{J}(x_1,\ldots,x_N) = \mathbb{E}\left[\min_n g(x_n, X)\right].$$

- The altruistic law is $\{x_n^*\}_{n=1}^N := \operatorname{argmin} \mathcal{J}$.
- For the quadratic case $\mathcal{J}$ is known as the N-centers function

$$\phi_{N,X}(x_1,\ldots,x_N) := \mathbb{E}\left[\min_n \|x_n - X\|^2\right],$$

and its altruistic law satisfies

$$x_n^* = C_X(V_{x_n^*|\mathcal{X}}).$$



$\mathcal{J} = \phi_{3,X},\ X \sim \mathcal{N}(0, I)$
(Flury 1990)

The Linear Quadratic Altruistic Problem

Problem Formulation

- Consider a cooperative system of N homogeneous agents with the linear dynamics

$$x_{n,k+1} = A_k x_{n,k} + B_k u_{n,k} - w_k,\; n = 1, \dots, N.$$

- Find the system control law $u_{n,k}(x_1, \dots, x_N)$ to minimize

$$\mathcal{J} = \mathbb{E}\left[\min_n \left\{\|x_{n,K}\|_Q^2\right\} + \sum_{n=1}^{N}\sum_{k=0}^{K-1} \|u_{n,k}\|_{R_k}^2\right].$$

- Assumptions:
  1. $\{x_{n,k}\}_{n=1}^{N}$ are perfectly known at the beginning of each stage k .
  2. $\{w_k\}_{k=0}^{K-1}$ is a zero-mean white noise with finite second moments and an absolutely continuous probability measure.
  3. $w_k \perp\!\!\!\perp x_{n,k}, u_{n,k}\ \forall n, k$.

Terminal Dynamics

- Preforming a terminal projection transformation

$$z_{n,k} := \Psi(K,k)x_{n,k}, \quad \Psi(K,k) := Q^{1/2}\prod_{s=k}^{K-1} A_s,$$

- the problem's cost can be rewritten as

$$\mathcal{J} = \mathbb{E}\left[\min_{n=1,\dots,N}\left\{\|z_{n,K}\|^2\right\} + \sum_{n=1}^{N}\sum_{k=0}^{K-1}\|u_{n,k}\|_{R_k}^2\right],$$

- with the terminal dynamics

$$z_{n,k+1} = z_{n,k} + \bar{B}_k u_{n,k} - \bar{w}_k, \quad n = 1,\dots,N,$$
$$\bar{B}_k := \Psi(K,k+1)B_k, \quad \bar{w}_k := \Psi(K,k+1)w_k.$$

The Dynamic Programming Algorithm

- Define the concatenated vectors

$$z_{:,k} := \begin{bmatrix} z_{1,k}^T & \dots & z_{N,k}^T \end{bmatrix}^T, u_{:,k} = \begin{bmatrix} u_{1,k}^T & \dots & u_{N,k}^T \end{bmatrix}^T.$$

- The problem's Bellman equations are

$$J_K(z_{:,K}) = \min_n \left\| z_{n,K} \right\|^2,$$

$$J_k(z_{:,k}) = \min_{u_{:,k}} \mathop{\mathbb{E}}_{\bar{w}_k} \left[\sum_{n=1}^{N} \left\| u_{n,k} \right\|_{R_k}^2 + J_{k+1}(z_{:,k+1}) \mid z_{:,k} \right].$$

- For $k = K - 1$

$$J_{K-1}(z_{:,K-1}) = \min_{u_{:,K-1}} \left\{ \sum_{n=1}^{N} \left\| u_{n,K-1} \right\|_{R_{K-1}}^2 + \mathop{\mathbb{E}}_{\bar{w}_{K-1}} \left[\min_n \left\| z_{n,K-1} + \bar{B}_{K-1} u_{n,K-1} - \bar{w}_{K-1} \right\|^2 \mid z_{:,K-1} \right] \right\}.$$

Direct Solution

Theorem (Optimal Last-Step Controls)

The optimal controls at stage $k = K-1$ satisfy

$$u^*_{n,K-1} = L_{n,K-1}\left(z_{n,K-1} - C_{\bar{w}_{K-1}}(V_{\hat{z}_{n,K}|\hat{\mathcal{Z}}_K})\right),$$

$$L_{n,K-1} \coloneqq -P_{n,K}\left(R_{K-1} + P_{n,K}\bar{B}^T_{K-1}\bar{B}_{K-1}\right)^{-1}\bar{B}^T_{K-1},$$

$$P_{n,K} \coloneqq \mathbb{P}_{\bar{w}_{K-1}}(V_{\hat{z}_{n,K}|\hat{\mathcal{Z}}_K}),$$

where $\hat{\mathcal{Z}}_K = \{\hat{z}_{n,K}\}_{n=1}^N$ are solutions of equation system

$$\hat{z}_{n,K} = z_{n,K-1} + \bar{B}_{K-1}L_{n,K-1}\left(z_{n,K-1} - C_{\bar{w}_{K-1}}(V_{\hat{z}_{n,K}|\hat{\mathcal{Z}}_K})\right), \quad n = 1,\ldots,N.$$

Direct Solution (ctd')

Theorem (Last-Step Cost-to-Go)

The cost-to-go function at stage $K-1$ is given by

$$J_{K-1}(z_{:,K-1}) = \sum_{n=1}^{N} \Bigg\{ P_{n,k} \mathbb{E}\left[\left\| \bar{w}_{K-1} - C_{\bar{w}_{K-1}}(V_{\hat{z}_{n,K}|\hat{\mathcal{Z}}_K}) \right\|^2 \mid \bar{w}_{K-1} \in V_{\hat{z}_{n,K}|\hat{\mathcal{Z}}_K} \right] + \left\| z_{n,K-1} - C_{\bar{w}_{K-1}}(V_{\hat{z}_{n,K}|\hat{\mathcal{Z}}_K}) \right\|^2_{P_{n,K-1}} \Bigg\},$$

$$P_{n,K-1} := P_{n,k} \left(I + \bar{B}_{K-1} L_{n,K-1} \right).$$

The Predictive LQA Problem

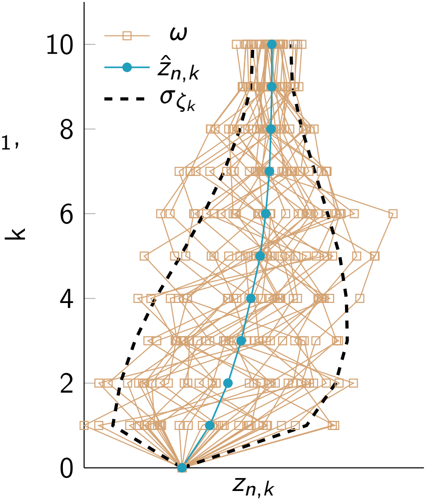
Deterministic and Stochastic (Pseudo) Systems

- Agents follow their egoistical law

$$\mathring{u}_k(z_{n,k}) = \mathring{L}_k z_{n,k},$$
$$\mathring{L}_k = -\left(\bar{B}_k^T \mathring{P}_{k+1} \bar{B}_k + R_k\right)^{-1} \bar{B}_k^T \mathring{P}_{k+1},$$
$$\mathring{P}_k = \mathring{P}_{k+1}\left(I + \bar{B}_k \mathring{L}_k\right), \quad \mathring{P}_K = 1.$$

- Dynamics can be decomposed to

$$\hat{z}_{n,k+1} = \hat{z}_{n,k} + \bar{B}_k \mathring{u}_k(\hat{z}_{n,k}),$$
$$\hat{z}_{n,k} := \mathbb{E}[z_{n,k} \mid z_{n,0}],$$
$$\zeta_{k+1} = \zeta_k + \bar{B}_k \mathring{u}_k(\zeta_k) + \bar{w}_k,$$
$$\zeta_k := \hat{z}_{n,k} - z_{n,k}.$$



The Predictive Problem Formulation

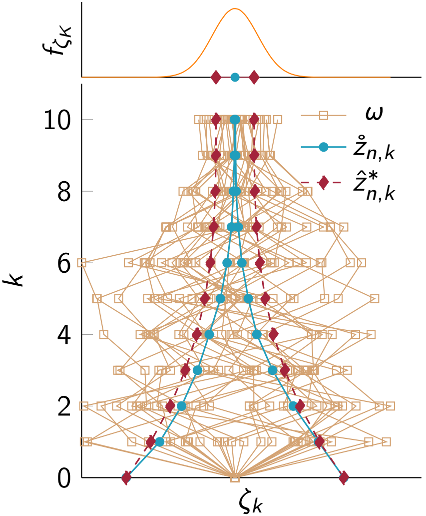
- The kth stage subproblem is

$$\mathcal{J}_k = \mathbb{E}\Bigg[\min_n \left\{ \left\| \hat{z}_{n,K} - \zeta_K \right\|^2 \right\} + \sum_{n=1}^{N} \sum_{s=k}^{K-1} \left\| \mathring{u}_s(\hat{z}_{n,s}) + \tilde{u}_{n,s} \right\|^2_{R_s} \mid z_{:,k} \Bigg]$$

$$\hat{z}_{n,s+1} = \hat{z}_{n,s} + \bar{B}_s \left(\mathring{u}_s(\hat{z}_{n,s}) + \tilde{u}_{n,s} \right),$$

$$\zeta_{s+1} = \zeta_s + \bar{B}_s \mathring{u}_s(\zeta_s) + \bar{w}_s.$$

- $u_{n,s} = \mathring{u}_s(\hat{z}_{n,s}) + \tilde{u}_{n,s}$ where $\tilde{u}_{n,s}$ is limited to OL laws for tractability.



The Predictive Problem Formulation (ctd')

- By the definition of the Voronoi partition of $\hat{\mathcal{Z}}_K := \{\hat{z}_{n,K}\}_{n=1}^N$

$$\min_n \|\hat{z}_{n,K} - \zeta_K\|^2 = \sum_{n=1}^N \mathbb{I}_{V_{\hat{z}_{n,K}|\hat{\mathcal{Z}}_K}}(\zeta_K)\|\hat{z}_{n,K} - \zeta_K\|^2$$

$$\leq \sum_{n=1}^N \mathbb{I}_{S_{n,k}}(\zeta_K)\|\hat{z}_{n,K} - \zeta_K\|^2 \quad \forall \mathcal{S}_k = \{S_{n,k}\}_{n=1}^N .$$

Lemma

$$\mathcal{J}_k(z_{:,k}, \{\tilde{u}_{:,s}\}_{s=k}^{K-1}) = \min_{\mathcal{S}_k=\{S_{n,k}\}_{n=1}^N} \hat{\mathcal{J}}_k(z_{:,k}, \mathcal{S}_k, \{\tilde{u}_{:,s}\}_{s=k}^{K-1}) :=$$

$$\min_{\mathcal{S}_k=\{S_{n,k}\}_{n=1}^N} \mathbb{E}\left[\sum_{n=1}^N \left\{ \mathbb{I}_{S_{n,k}}(\zeta_K)\|\hat{z}_{n,K} - \zeta_K\|^2 + \sum_{s=k}^{K-1} \|\mathring{u}_s(\hat{z}_{n,s}) + \tilde{u}_{n,s}\|_{R_s}^2 \right\} \mid z_{:,k}\right]$$

Necessary Conditions for the Optimal OL Perturbations

Lemma (Optimal Partition for Given Perturbations)

The optimal partition to minimizes $\hat{\mathcal{J}}_k(z_{:,k}, \mathcal{S}_k, \{\tilde{u}_{n,s}\}_{s=k}^{K-1})$ *is* $\mathcal{V}_{\hat{\mathcal{Z}}_K} = \left\{V_{\hat{z}_{1,K}|\hat{\mathcal{Z}}_K}, \ldots, V_{\hat{z}_{N,K}|\hat{\mathcal{Z}}_K}\right\}$.

Lemma (Optimal Perturbations for a Given Partition)

The optimal OL perturbations to minimizes $\hat{\mathcal{J}}_k(z_{:,k}, \mathcal{S}_k, \{\tilde{u}_{n,s}\}_{s=k}^{K-1})$ *is*

$$\tilde{u}^*_{n,s}(\hat{z}_{n,s}, S_{n,k}) = -\mathring{L}_s \hat{z}_{n,s} + \tilde{L}_{n,s} \left(\hat{z}_{n,s} - C_{\zeta_K}(S_{n,k})\right),$$

$$\tilde{L}_{n,s} = -\left(R_s + \bar{B}_s^T \tilde{P}_{n,s+1} \bar{B}_s\right)^{-1} \bar{B}_s^T \tilde{P}_{n,s+1},$$

$$\tilde{P}_{n,s} = \tilde{P}_{n,s+1}\left(I + \bar{B}_s \tilde{L}_{n,s}\right), \ \tilde{P}_{n,K} = \mathbb{P}_{\zeta_K}(S_{n,k}),$$

with the optimal value $\hat{J}_k(z_{:,k}, \mathcal{S}_k) := \hat{\mathcal{J}}_k(z_{:,k}, \mathcal{S}_k, \{\tilde{u}^*_{n,s}\}_{s=k}^{K-1})$.

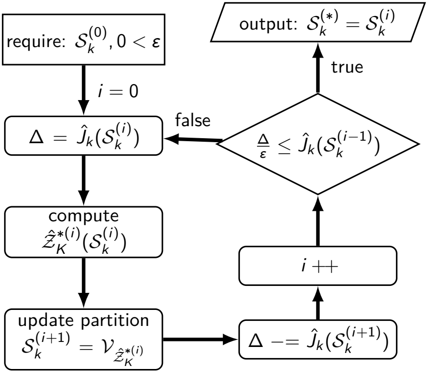
The Partition Improvement Algorithm

Properties:

1. Descent algorithm
2. If $u^*_{K-1,n}$ is the optimal action to minimize $\mathcal{J}$ at stage $K-1$ then

$$u^*_{n,K-1} = \mathring{u}_{K-1}(z_{n,K-1}) + \tilde{u}^*_{n,K-1}(z_{N,K-1}, \mathcal{S}^*_{K-1}),$$

where $\mathcal{S}^*_{K-1}$ is a stationary point of the algorithm.



The Partition Improvement Algorithm

The PLQA Law

Definition (The PLQA Law)

The PLQA law is the open-loop feedback law

$$u_{n,k}^{\mathrm{PLQA}}(z_{n,k}, S_{n,k}) = \mathring{u}_{n,k} + \tilde{u}_{n,k}^{*} = \tilde{L}_{n,k}\left(z_{n,k} - C_{\zeta_K}(S_{n,k})\right).$$

Theorem (Superiority over the Egoistical Law)

For all $k = 0, \ldots, K-1$ there exists a partition $\mathcal{S}_k = \{S_{n,k}\}_{n=1}^{N}$ for which the PLQA law achieves an equal or a lower cost $\mathcal{J}_k$ than as if each agent followed its egoistical law.

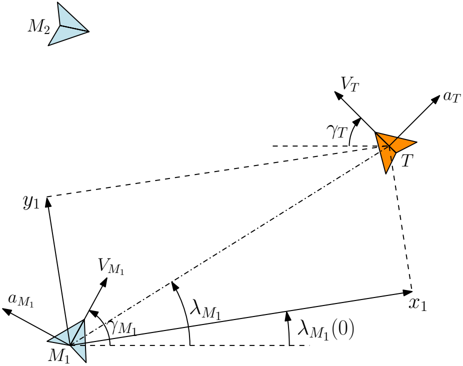
Application to Cooperative Interception

Cooperative Interception

- Consider the head-on 2-on-1 planar engagement

$$\dot{x}_n = -V_T\cos(\gamma_T) - V_{M_n}\cos(\gamma_{M_n}),$$
$$\dot{y}_n = V_T\sin(\gamma_T) - V_{M_n}\sin(\gamma_{M_n}),$$
$$\dot{\gamma}_{M_n} = u_n/V_{M_n}, \quad \dot{\gamma}_T = w/V_T.$$

- Interceptors' commands u_n were applied at a fixed rate $\frac{1}{\Delta t}$.
- Target's commands w modeled as a continuous-time zero-mean white Gaussian noise of intensity σ^2.



Cooperative Interception (ctd')

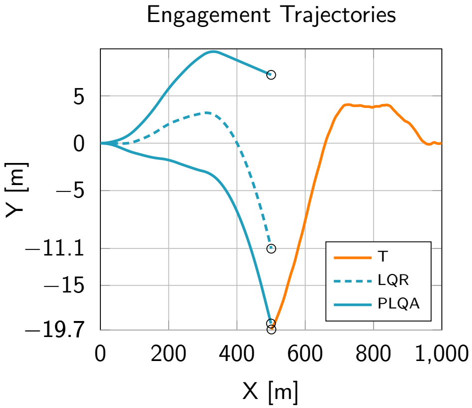
- Missile termination times $t_{f_1} = t_{f_2} := t_f$ were assumed to be equal, allowing the discrete-time cost

$$\mathcal{J} = \mathbb{E}\left[\min_{n=1,\dots,N}\left\{y_{n,K}^2\right\} + \frac{1}{a_{M_{\max}}^2}\sum_{n=1}^{N}\sum_{k=0}^{K-1} u_{n,k}^2\right], \quad K = \left\lceil \frac{t_f}{\Delta t} \right\rceil .$$

	Simulation Parameters				
V_{M_n}	500	[m/s]	V_T	500	[m/s]
R_0	1	[km]	Δt	0.05	[s]
σ^2	2500	$\left[\mathrm{m}^2/\mathrm{s}^3\right]$	$a_{M_{\max}}$	20	g

Simulation Results — Single Run Analysis

- Overlaid 2V1 engagement trajectories of LQR and PLQA guidance laws.
- All agents start with identical initial conditions.
- LQR agent trajectories coincide, incurring a miss of 8.6[m].
- PLQA agents spread, incurring a substantial lower miss of 0.628[m].



Simulation Results — Monte-Carlo Simulation

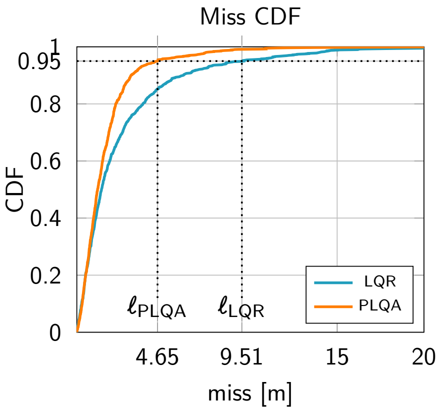
- CDF evaluated based on 1000 target realizations.
- Define a missile's *lethal radius*

$$\text{miss} := \min_n \sqrt{x_{n,K}^2 + y_{n,K}^2},$$

$$\ell : \mathbb{P}(\text{miss} \leq \ell) = 0.95.$$

- PLQA offers
  - 51% reduction in lethal radius
  - 35% reduction in average miss

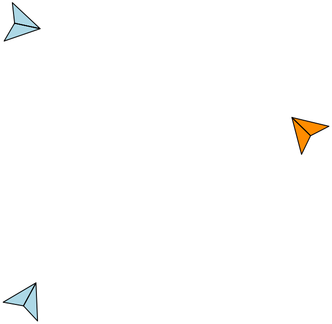
compared to the LQR.



Conclusions

Conclusions

- Introduced the concept of systems altruism.
- Derived a novel cooperative law, guaranteed to outperform its egoistical counterpart.
- Demonstrated the results for a cooperative interception scenario, showcasing the superior performance of the altruistic law.



Thank you for your attention!