CPS:Medium:Safe Learning-Enabled Cyberphysical Systems, CNS-2038493

Co-PIs: M. Sznaier and O. Camps,
Robust Systems Lab Northeastern University


L. Von Engelbrechten


Z. Taylor-Liang


S. Wong


A. Comas


T. Dai


Y. Zhang


O. Camps

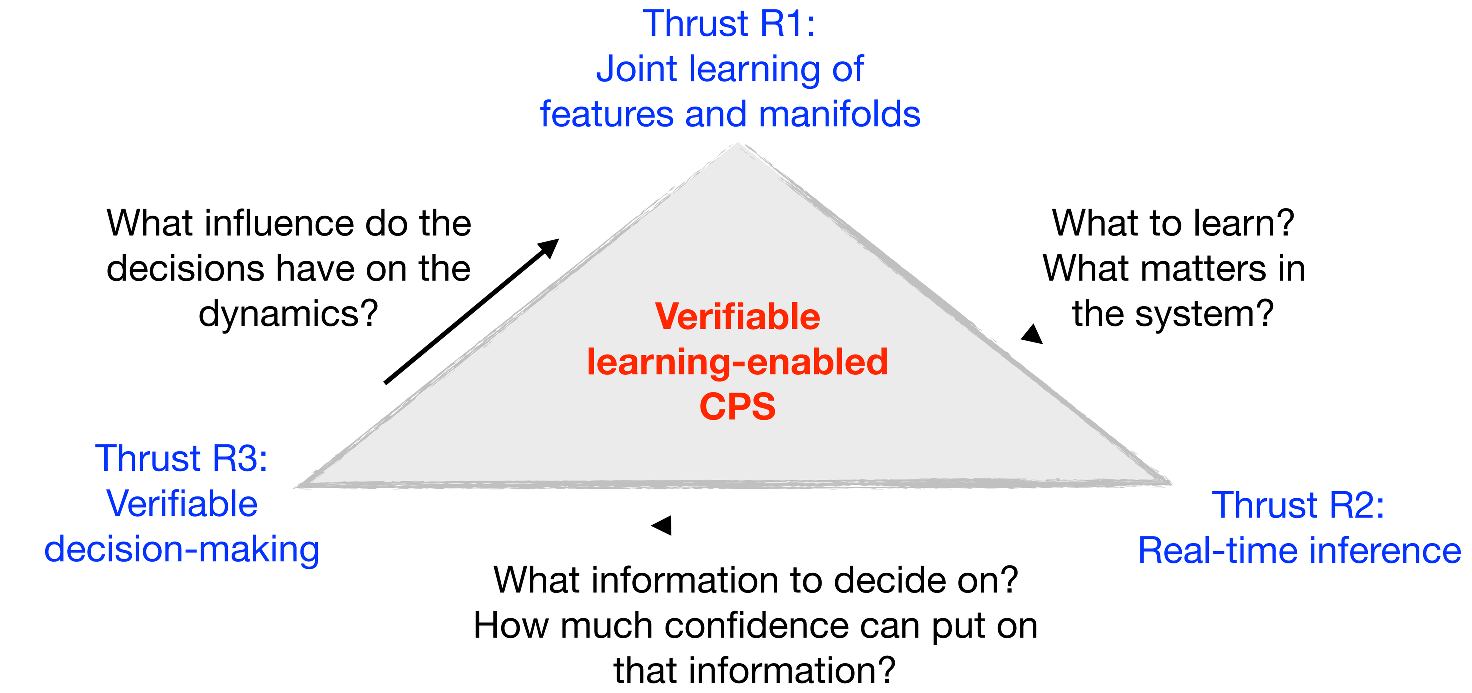

M. Sznaier

Motivation

Design autonomous CPS capable of safely operating in and adapting to previously unseen scenarios.
(Humans can do it!)

Challenges

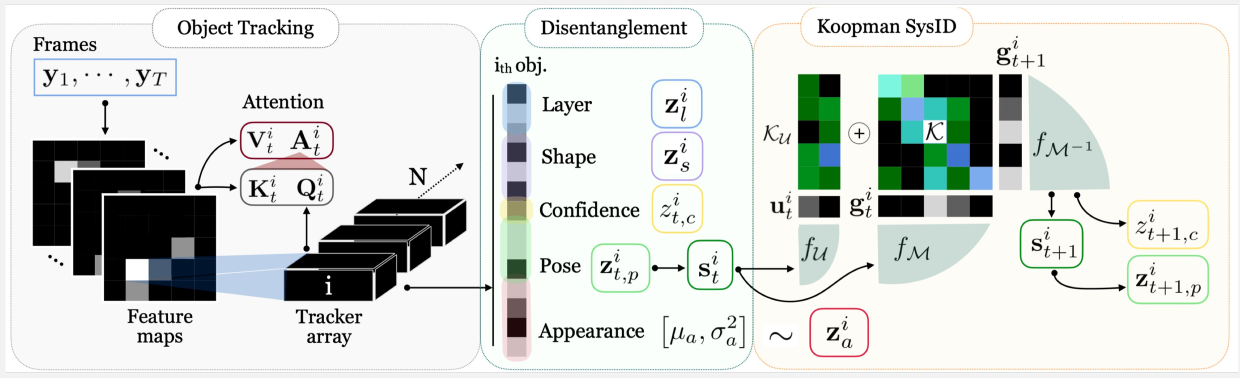
- Lack of training data (often single execution).
- Need to act while learning (no re-do!).
- Actionable information sparsely encoded in large data sets.



Joint learning of features and manifolds

- Goal: learn parsimonious dynamical representations.
- Main idea: search for manifold where the dynamics are linear (Koopman operators).
- Technical details:
  - Use delayed coordinates and attention based tracking
  - Enforce linear manifold dynamic
  - Interpretable, manipulable representations

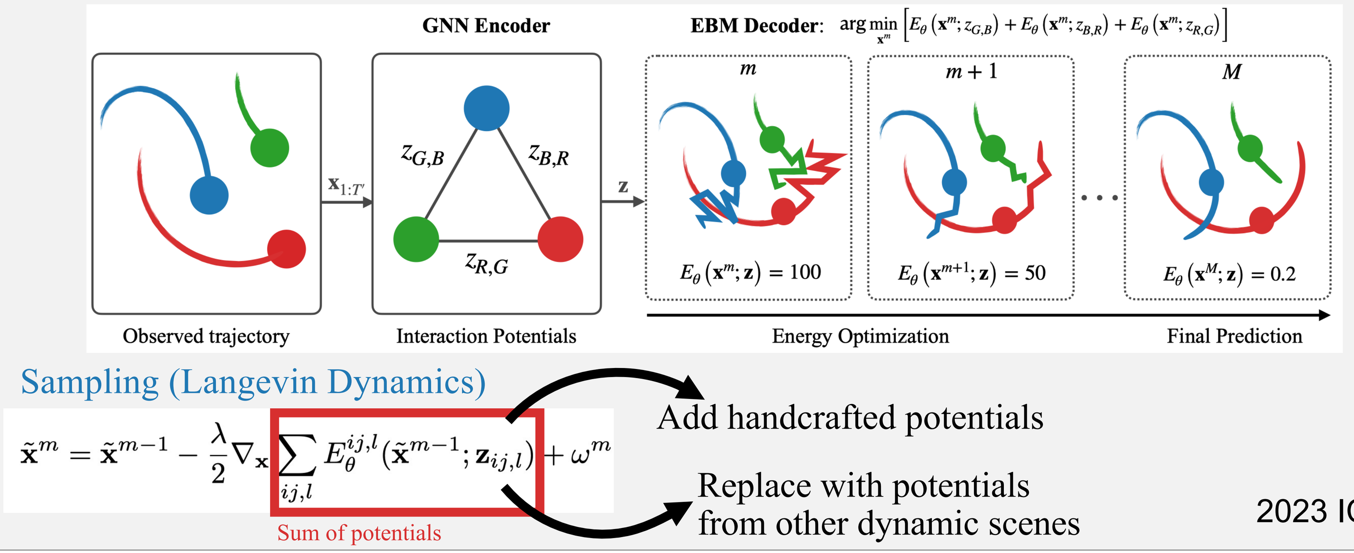
OKID: Object-centric Koopman-based Interpretable Decomposition



2023 L4DC

Learning dynamic interactions

- Goal: explain complex dynamics of interacting agents.
- Main idea: model interactions with a fully-connected graph and Energy-based Models (EBM).
- Technical details:
  - Trajectories are encoded as sets of potentials conditioning EBMs
  - At inference time, trajectories are predicted by satisfying these potentials



2023 ICML (Oral)

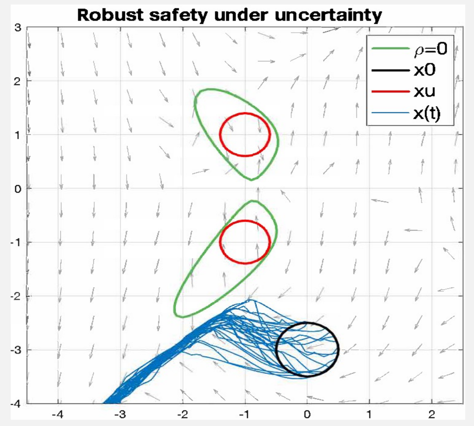
Verifiable safe data driven control

- Goal: data driven avoidance of an unsafe set.
- Main idea: enforce robust Nagumo-invariance .
- Technical details:
  - Prior: non linear dynamics of the form:
    $\dot{\mathbf{x}} = \mathbf{F}\phi(\mathbf{x}) + \mathbf{G}\gamma(\mathbf{x})\mathbf{u} + \eta \quad \mathbf{F}, \mathbf{G}$ unknown
  - Initial condition set: $\mathcal{X}_o$, bad set: $\mathcal{X}_u \doteq \{\mathbf{x} : h(\mathbf{x}) \geq 0\}$
  - Experimental data: $(\mathbf{u}(t_k), \mathbf{x}(t_k), \dot{\mathbf{x}}(t_k)),\ 0 \leq t_k \leq T$
  - Fact: bad set is avoided if there exist $\rho(x), \psi(x)$
    $\rho(\mathbf{x}) \geq 0,\ \forall \mathbf{x} \in \mathcal{X}_0,\ \rho(\mathbf{x}) < 0, \forall \mathbf{x} \in \mathcal{X}_u$ and

$$\nabla \cdot \left[\rho(\mathbf{x})\Big(\sum_i \alpha_i \mathbf{f}_i(\mathbf{x}(t)) + \mathbf{w}\Big) + \psi(\mathbf{x}) \sum_j \beta_j \mathbf{g}_j(\mathbf{x})\right] - \rho(\mathbf{x})h(\mathbf{x}) > 0$$

for all plants that could have generated the data

  - Fact: reduces to a convex SDP via duality
  - Fact: only imposes $d\rho/dt < 0$ when $\rho = 0$



2023 IEEE LCSS

Application

- Public space monitoring to mitigate unsafe situations.


Monitoring social distancing

Active shooter scenario (joint with M. Siami)

Scientific Impact

- Rapprochement of Systems Theory, ML, Viability.
- Efficient extraction of actionable information from large data sets.
- Frugal, explainable architectures for dynamics oriented learning.

Broader Impact and Outreach

- Certified safe learning enabled systems that can operate in close proximity to humans.
- Applications: health care, infrastructure monitoring, public space monitoring.
- Outreach through Northeastern's UPLIFT program.
- Internships at Adobe, Mathworks, Google.

Northeastern

Award CNS-2038493