Strategy Synthesis for Partially-known Switched Stochastic Systems

John Jackson[1], Luca Laurenti[2], Eric Frew[1], Morteza Lahijanian[1]
[1]Smead Aerospace Engineering, University of Colorado Boulder
[2]Computer Science, TU Delft

john.m.jackson@colorado.edu

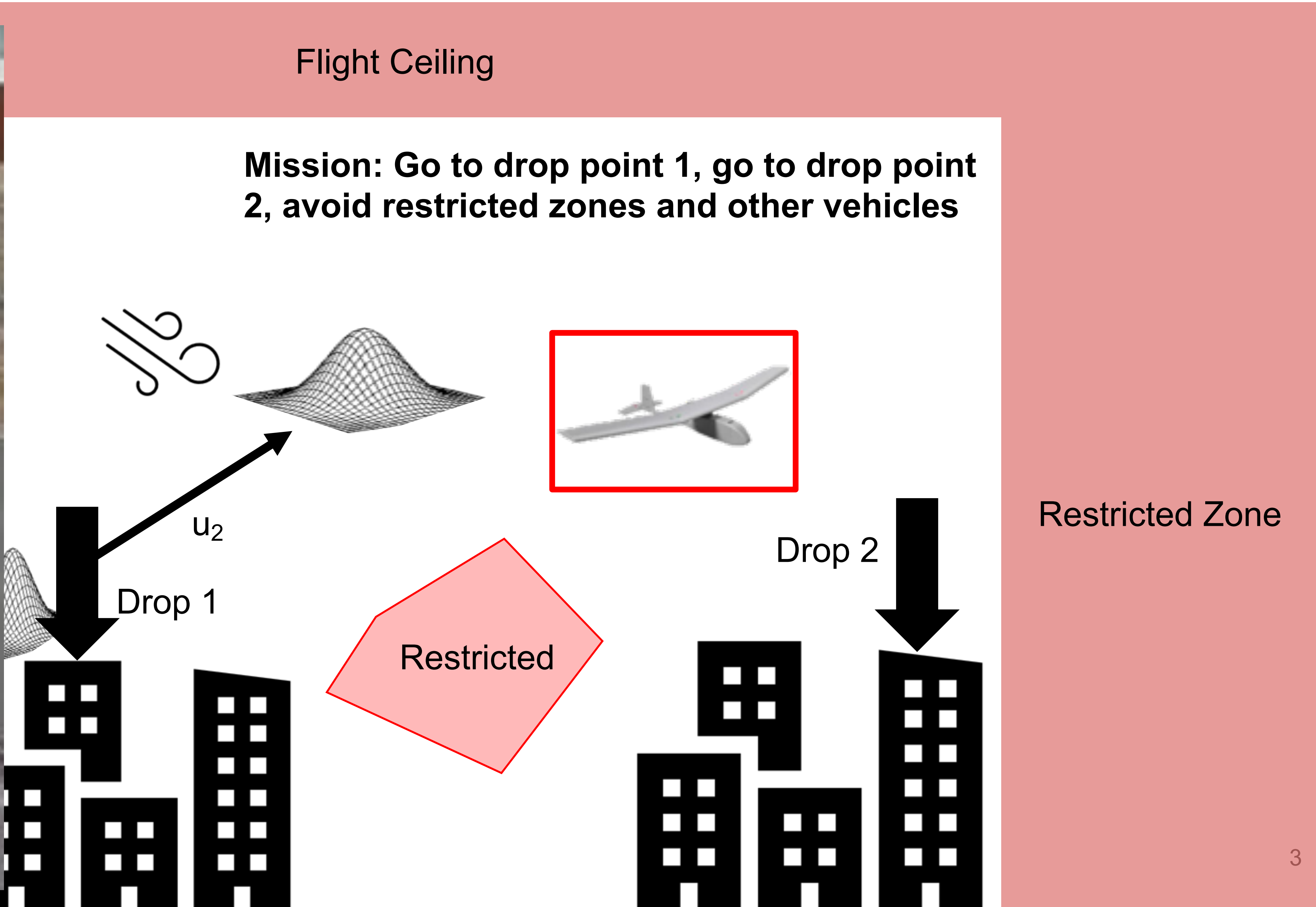
HSCC 2021

Research and Engineering Center for Unmanned Vehicles
UNIVERSITY OF COLORADO **BOULDER**

ARIA Systems
Assured Reliable Interactive Autonomous

Safety-Critical Urban Flying

Mission: Go to drop point 1, go to drop point 2, avoid restricted zones and other vehicles



Problem Overview

$$\mathbf{x}(k+1) = \overbrace{f_{\mathbf{u}(k)}(\mathbf{x}(k))}^{\textbf{Known}} + \overbrace{g_{\mathbf{u}(k)}(\mathbf{x}(k))}^{\textbf{Unknown}} + \underbrace{\mathbf{w}(k)}_{\text{sub-Gaussian noise}}$$

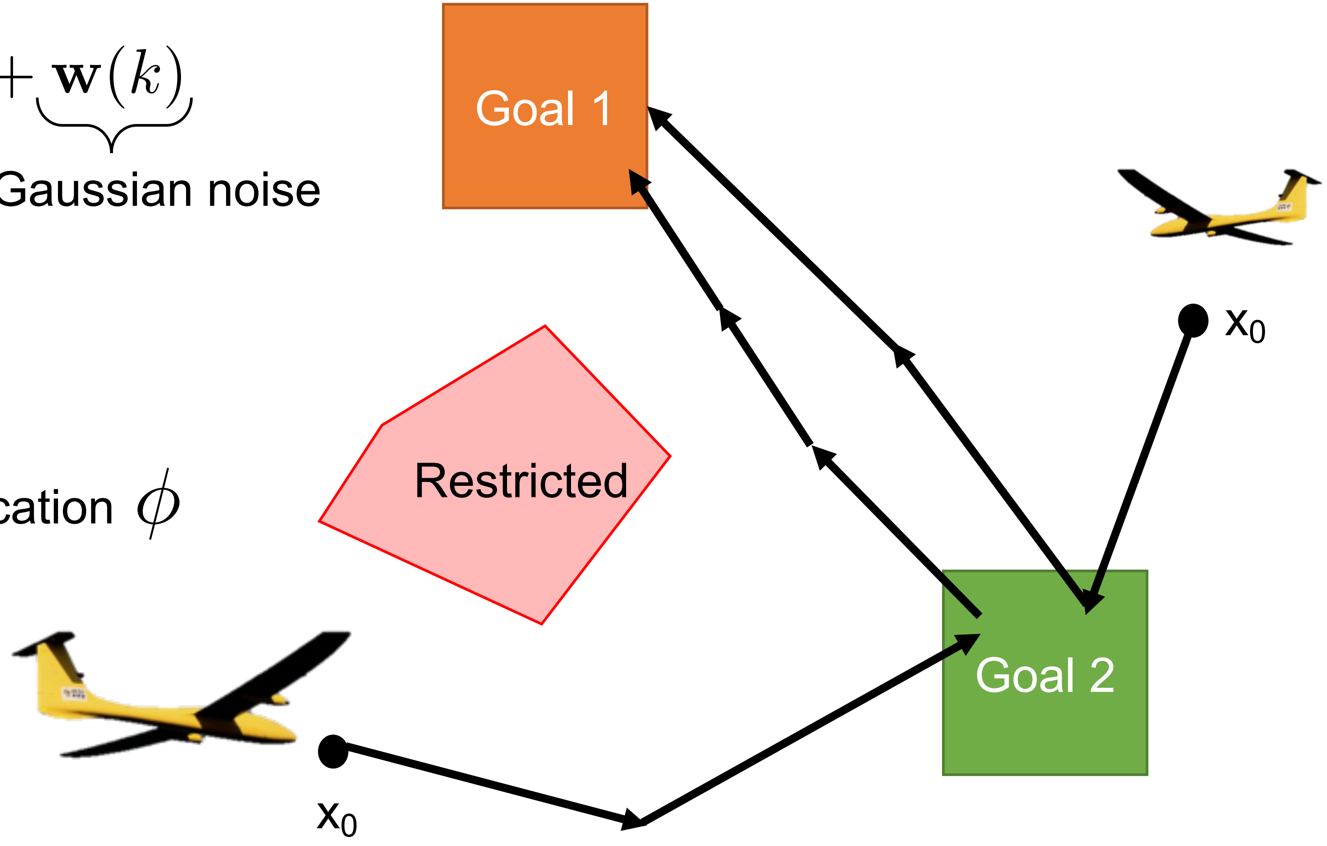
$$\mathbf{u}(k) \in \{u_1, \ldots, u_M\}$$

Given: a dataset generated by system, continuity assumptions on unknown dynamics.

Mission: Go to Goal 1, Go to Goal 2, always avoid restricted airspace. $\Leftrightarrow$ LTLf Specification ϕ

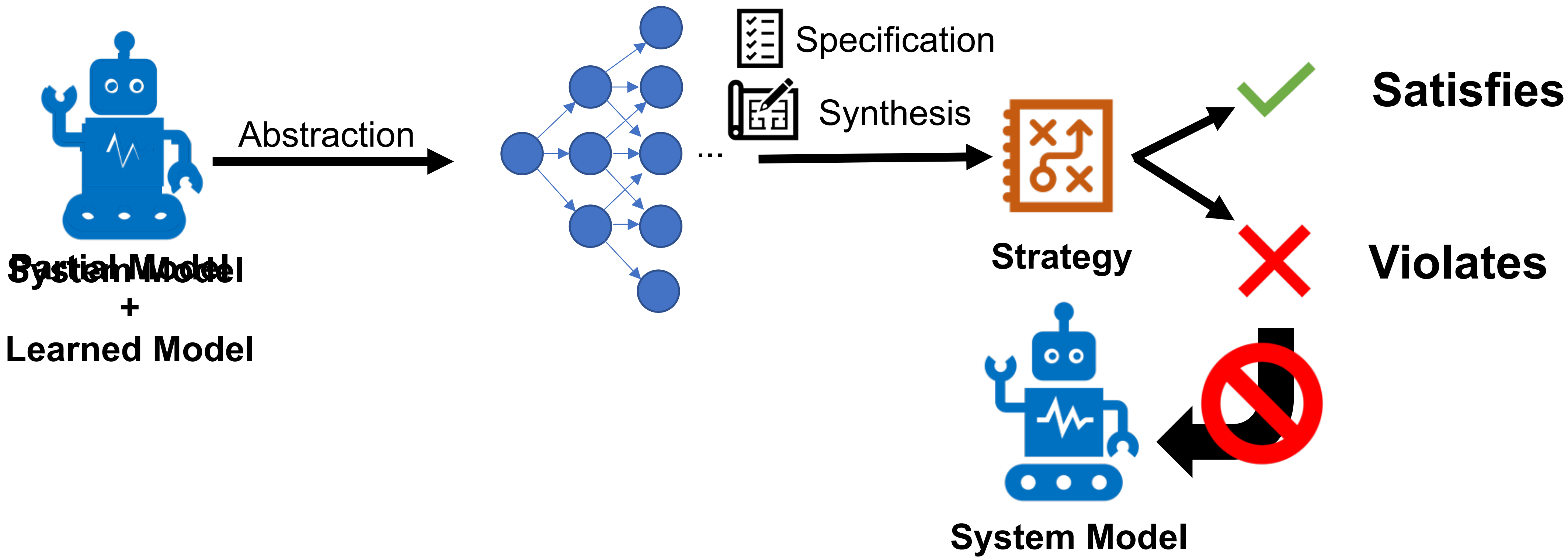
Strategy: π maps from path to control

$$\Pr(x_0 \models \phi \mid \pi^*) = p^*$$



Problem: Synthesize a switching strategy for the partially known system as close to the true optimal strategy as possible.

Abstraction-Based Synthesis



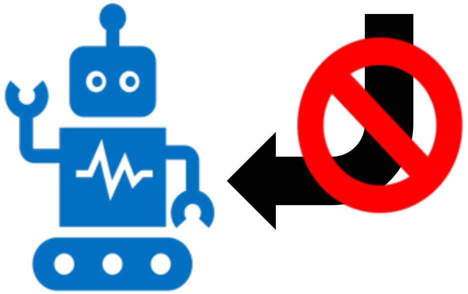
Challenges

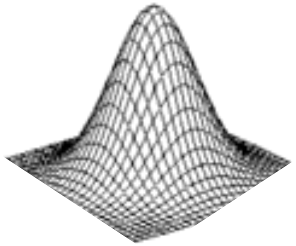
How to construct an abstraction for a partially-known model?

How to generate a strategy with guarantees for the latent system?

Data-Driven Approaches

$$\mathbf{x}(k+1) = A_i \mathbf{x}(k)$$





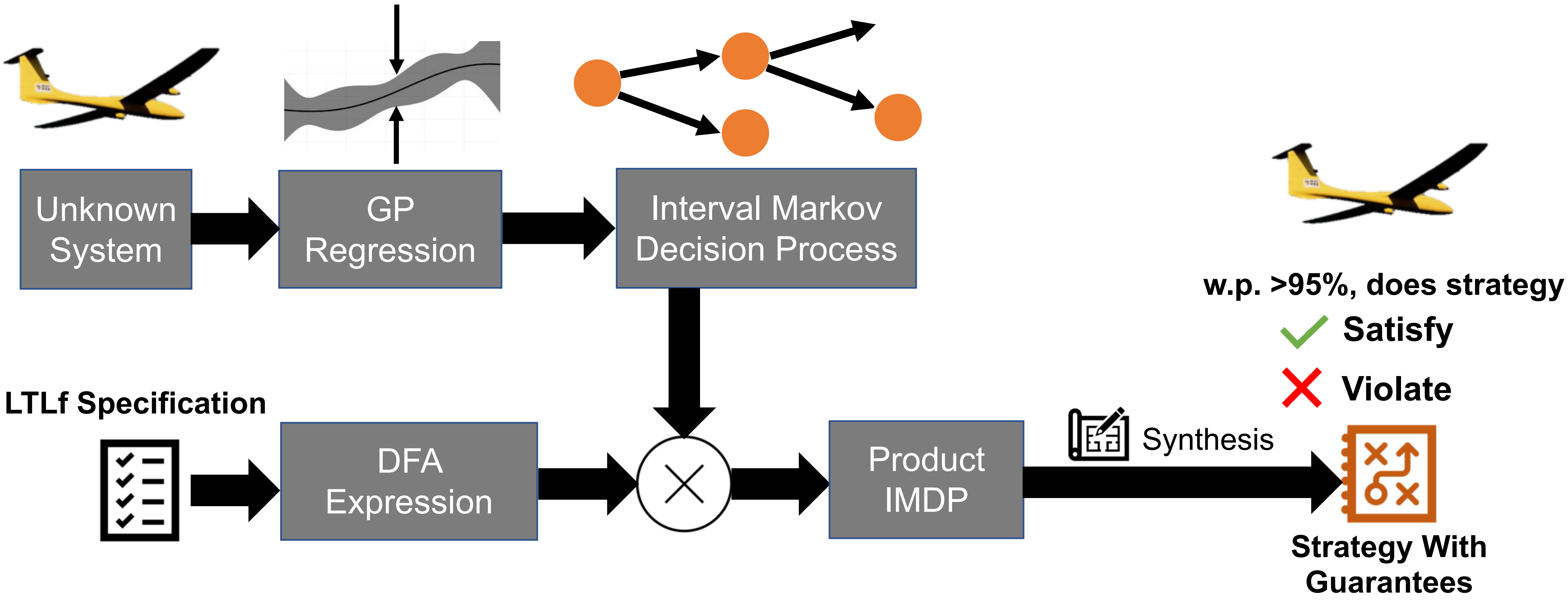
Specification?

Completeness?

- Parameterized linear systems learned from data e.g. [1], [2]
- Polynomial approximations e.g. [3]
- Gaussian process regression and synthesis for
  - Safe reinforcement learning e.g. [4], [5]
  - Control barrier certificates e.g. [6]

Solution: Data-driven abstraction-based strategy synthesis for a **switched stochastic** system with **guarantees that extend to the latent system**.

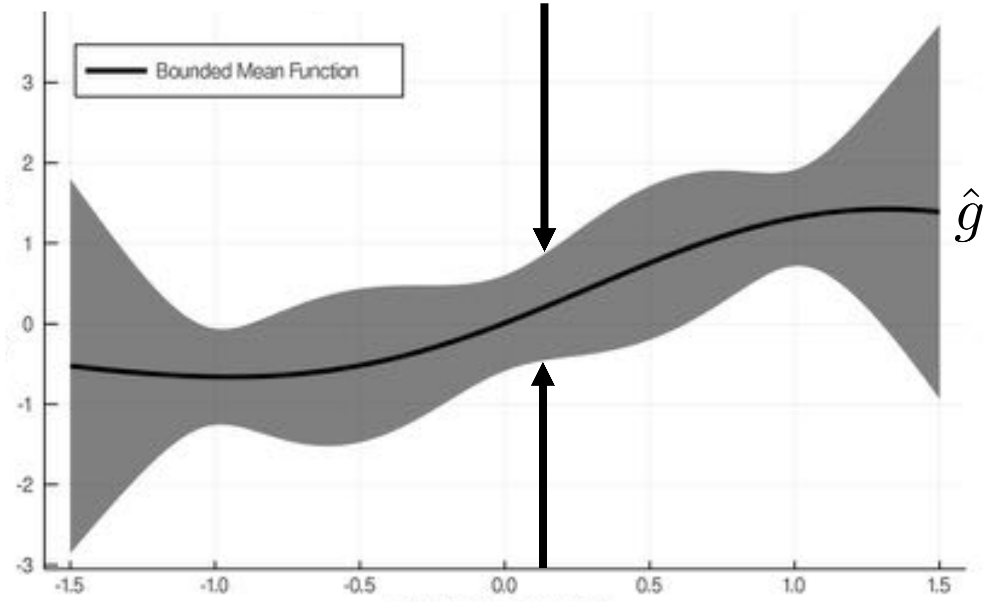
Framework Overview



Learning with Gaussian Processes

$$y_{k+1} = x_{k+1} - f_{u_k}(x_k) = g_{u_k}(x_k) + w_k$$

$\epsilon(\delta)$



$\hat{g}(x)$

Given: prior mean and covariance functions, dataset

Procedure: Joint-Gaussian assumption and Bayesian conditioning

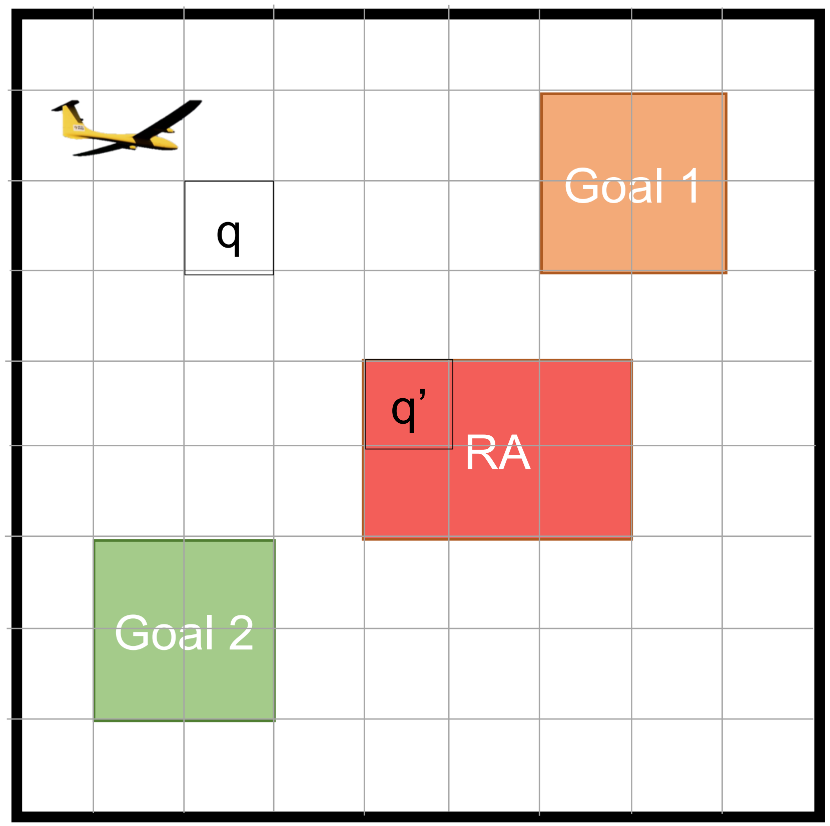
Result: MLE estimate of function, posterior covariance

RKHS Assumption [8, Chowdhury 2017]

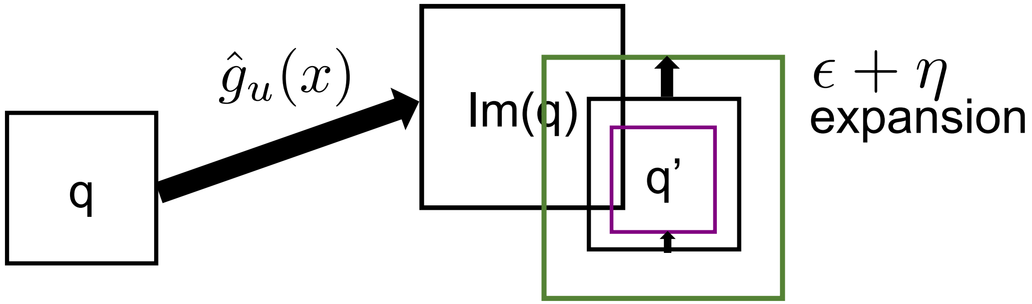
- Lifts Gaussian assumption
- Unknown function lies in the span of the prior covariance function
- General probabilistic error bounds using GP regression

Probabilistic Error: $\Pr(|g(x) - \hat{g}(x)| \leq \epsilon(\delta)) \geq 1 - \delta \quad \forall x \in X$

Interval MDP Abstraction



"Unsafe"
"Out of Bounds"



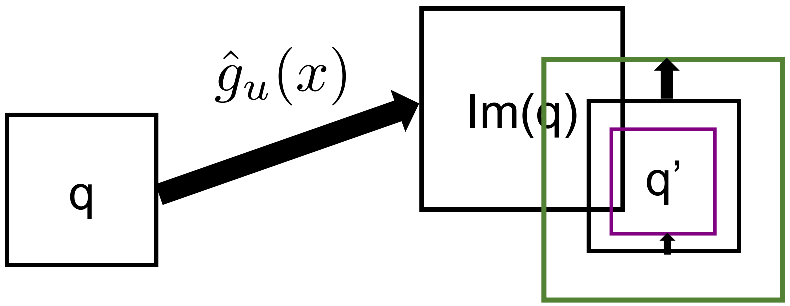
$$\Pr(q \to q')$$

Uncertainties

1. Process Noise (Known) η, P_η
2. Learning Error – RKHS Analysis ϵ, P_ϵ
3. Discretization Error

An IMDP defines a **space** of MDPs using **transition probability intervals** under each action.

Interval MDP Abstraction



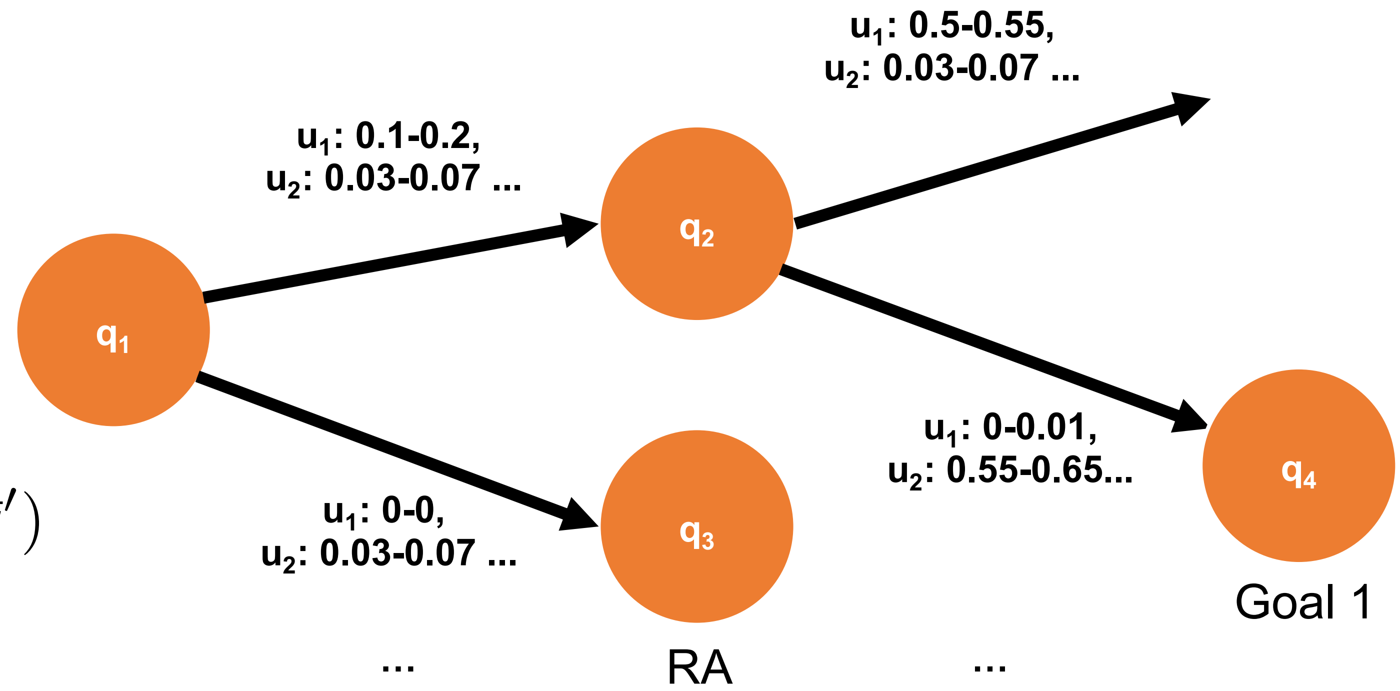
Uncertainties

1. Process Noise (Known) η, P_η
2. Learning Error – RKHS Analysis ϵ, P_ϵ
3. Discretization Error

$$\underline{\Pr}(q \to \underline{q}') \leq \Pr(q \to q') \leq \overline{\Pr}(q \to \overline{q}')$$

$$\overline{\Pr} = \mathbf{1}^{\overline{q}}_{Im(q)} P_\epsilon P_\eta (1 - P_\epsilon)$$

$$\underline{\Pr} = \mathbf{1}^{\underline{q}}_{Im(q)} P_\epsilon P_\eta$$



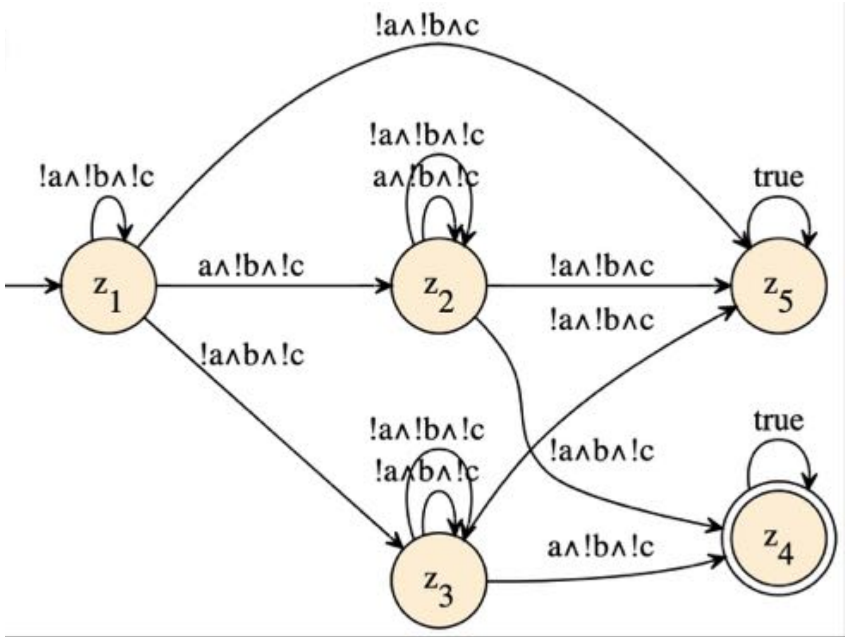
An IMDP defines a **space** of MDPs using **transition probability intervals** under each action.

LTLf and DFA

- Linear temporal logic over finite traces (LTLf) specifies behavior for finite time
- All regular languages have Deterministic Finite Automata (DFA) associations
  - Finite-state object to track progress towards satisfying ϕ

$$\phi = \mathcal{G}\neg c \wedge \mathcal{F}a \wedge \mathcal{F}b$$

both accept the same language!

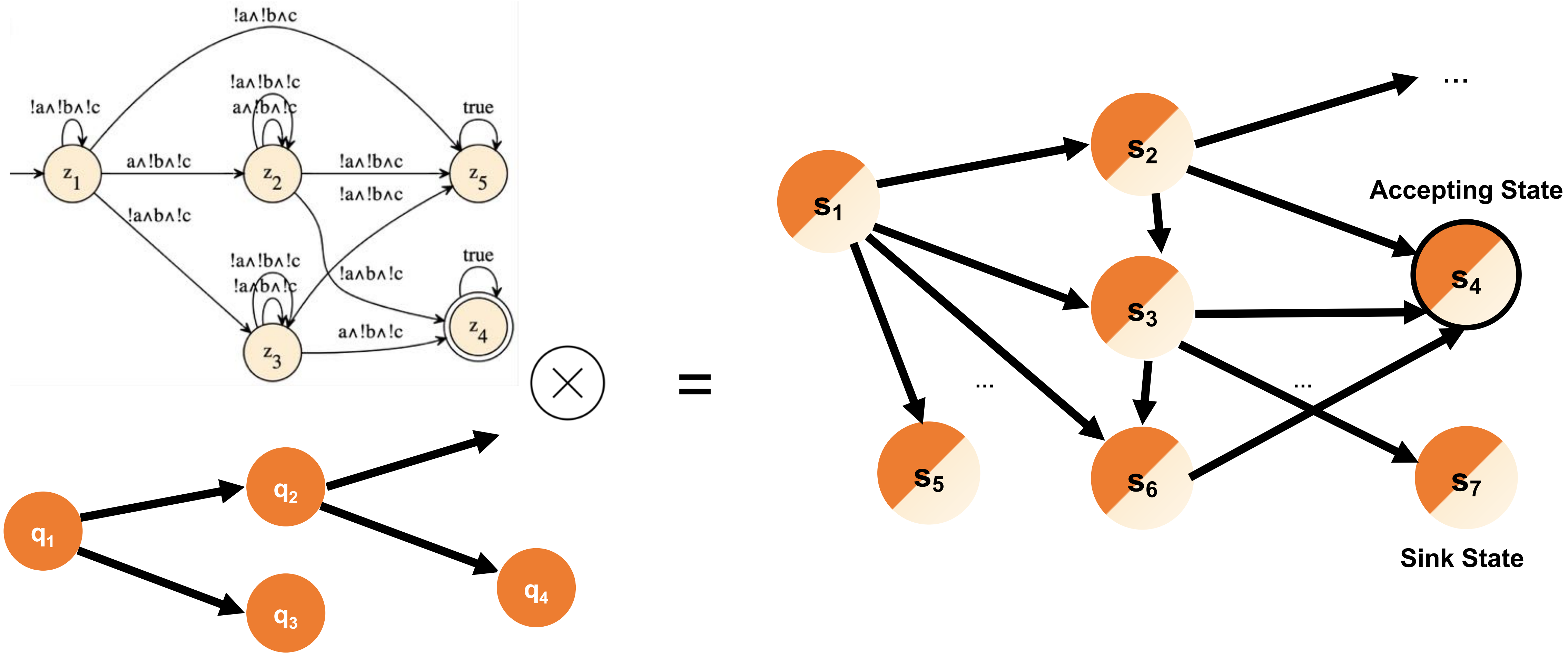


Sink State

Accepting State

[De Giacomo IJCAI'13]

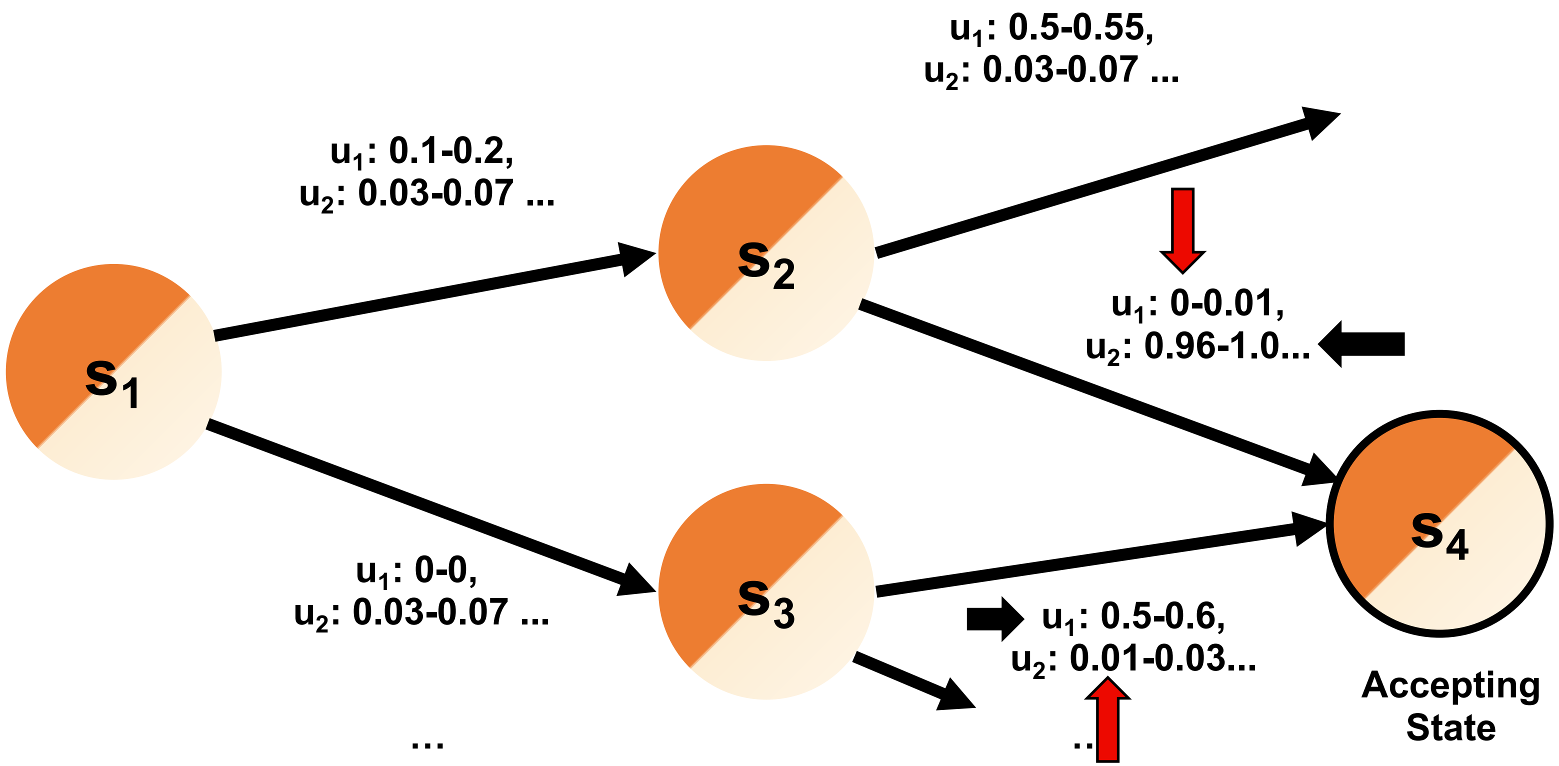
Product of DFA and IMDP



- Cartesian product of IMDP and DFA states, feasible transition probability intervals
- Paths on the product embed the history of DFA
- **The product of the system IMDP and DFA is also an IMDP**

Synthesis on an IMDP

- Synthesis on IMDP is a two-player game
- Efficient value iteration over strategies and transition distributions [Lahijanian TAC'15]



Player 1 chooses actions

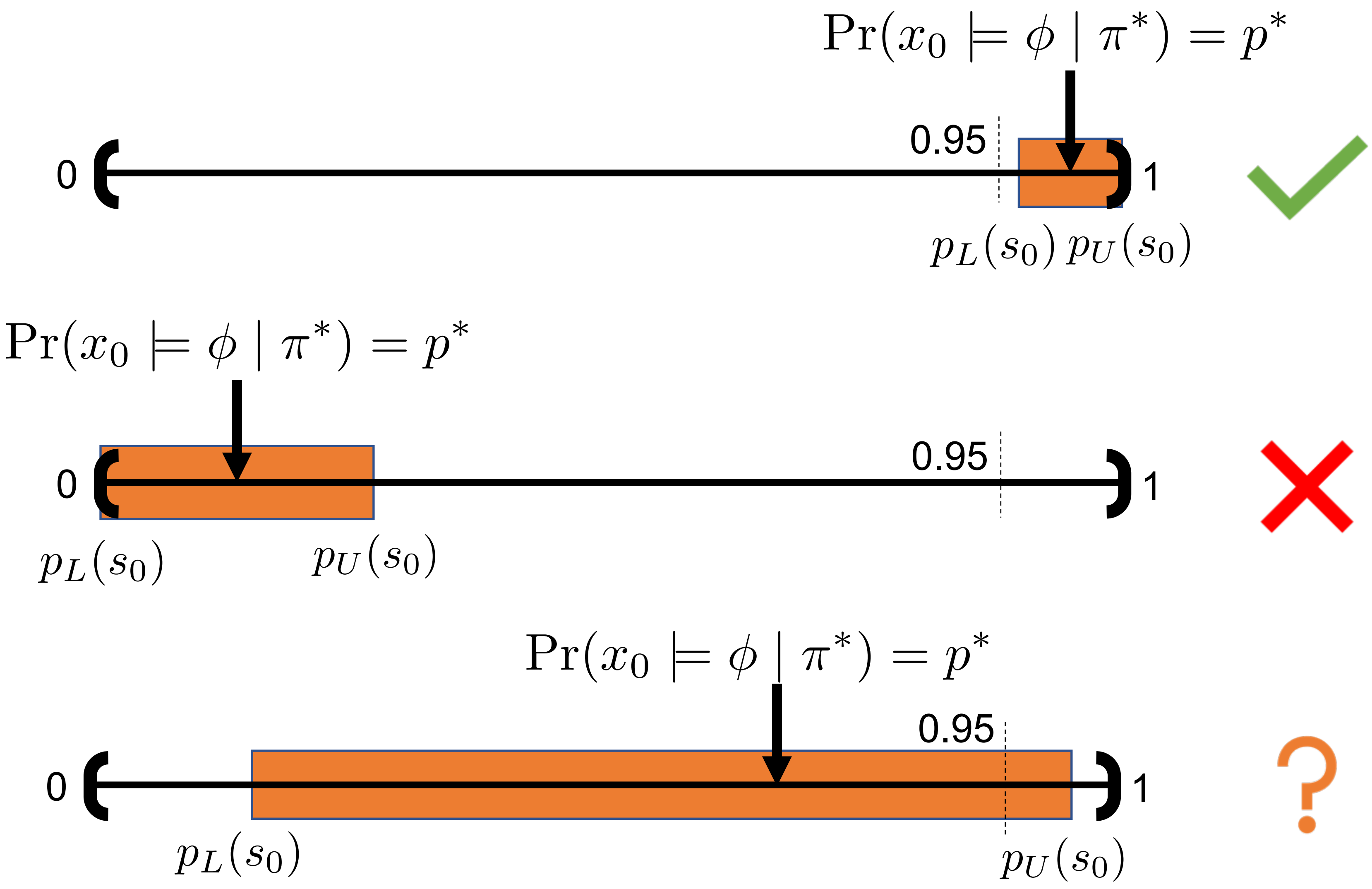
Player 2 chooses transition probability

$$p_L(s_0) = \max_{\pi} \min_{\gamma} p(s_0 \models \phi \mid \pi, \gamma)$$

$$p_U(s_0) = \max_{\pi} \max_{\gamma} p(s_0 \models \phi \mid \pi, \gamma)$$

Optimal Robust Strategy: Choose actions with the highest lower-bound of transition to the accepting state (accounting for player 2's worst actions)

Policy Optimality

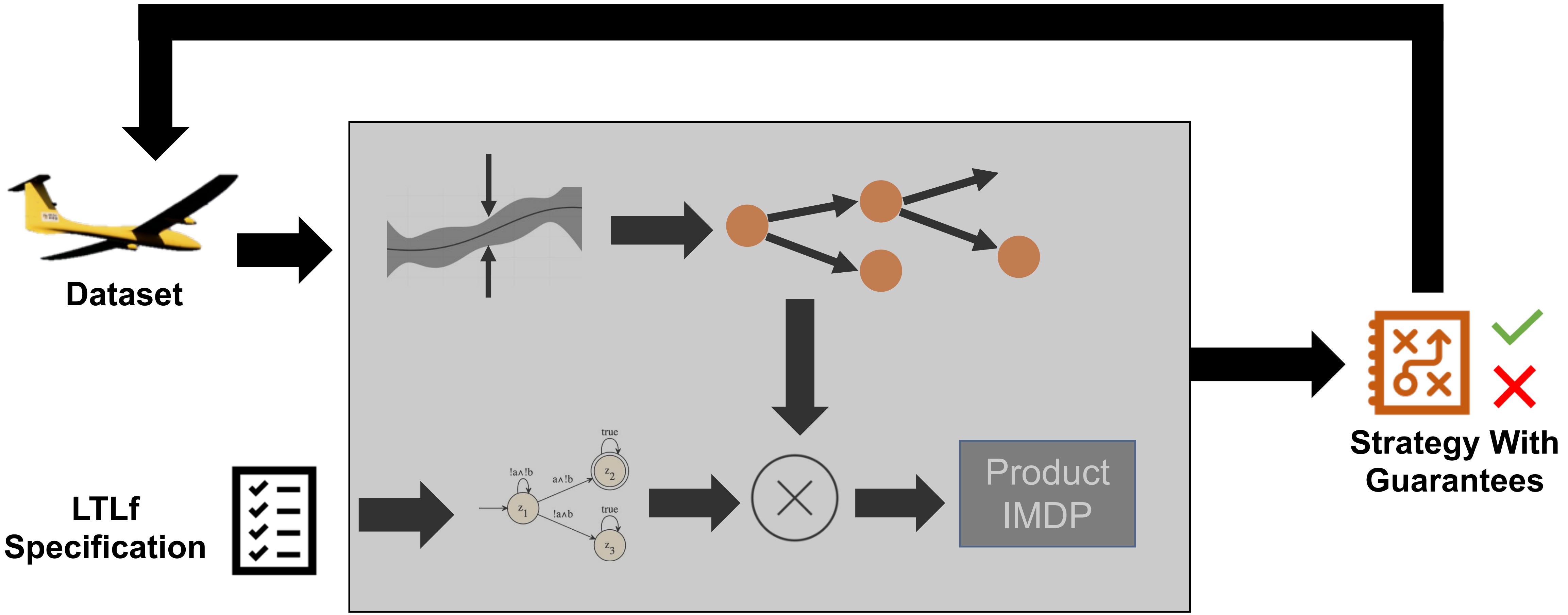


$$p_L(s_0) \geq 0.95$$

$$p_U(s_0) < 0.95$$

To Get Better Bounds
- More Data
- Finer Discretization

Framework Overview



End-to-End Learning and Strategy Synthesis

Linear Switched System

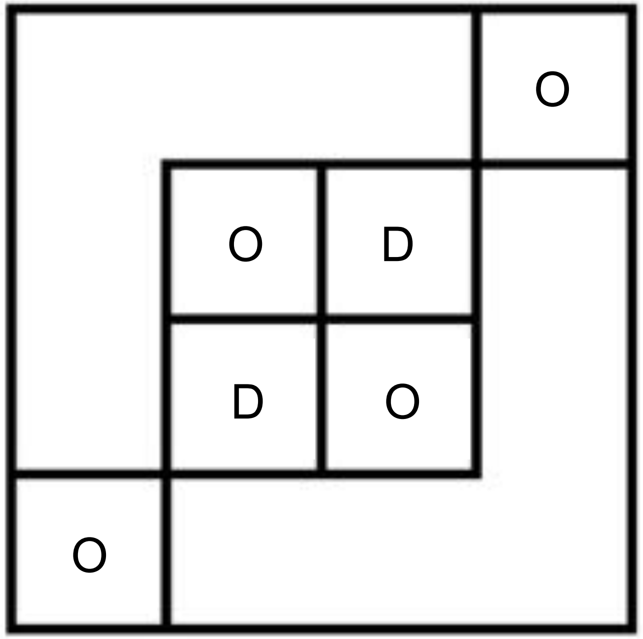
$$\mathbf{x}(k+1) = A_i \mathbf{x}(k) + \mathbf{w}(k)$$

$$A_1 = \begin{bmatrix} 0.4 & 0.1 \\ 0 & 0.5 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 0.4 & 0.5 \\ 0 & 0.5 \end{bmatrix},$$
$$A_3 = \begin{bmatrix} 0.4 & 0 \\ 0.5 & 0.5 \end{bmatrix},$$

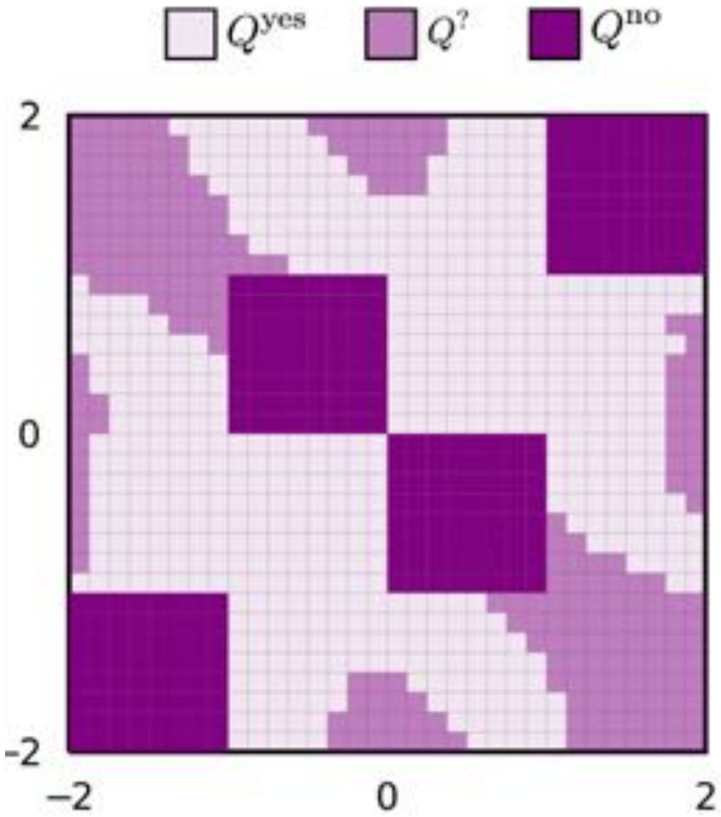
→ 200 i.i.d. datapoints per mode
→ noise variance: 0.01^2

$$\phi = \neg O \ \mathcal{U} \ D$$

State Space Labels



Is probability of success > 0.95?



Example from [Lahijanian et. al. TAC'15]

Nonlinear System Example

$$\mathbf{x}(k+1) = \mathbf{x}(k) + g_u(\mathbf{x}(k)) + \mathbf{w}(k)$$

→ 300 i.i.d. datapoints per mode

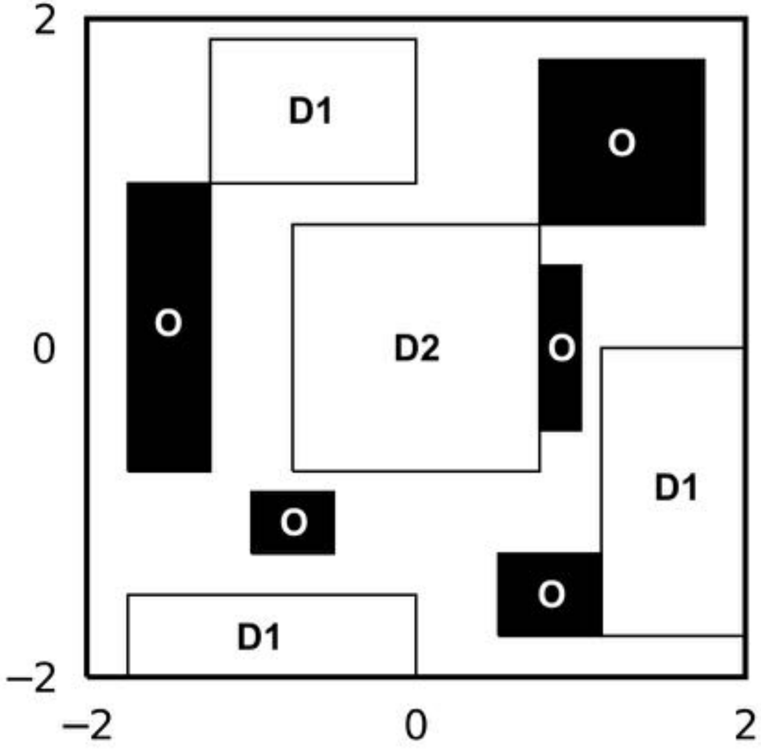
→ noise variance: 0.01^2

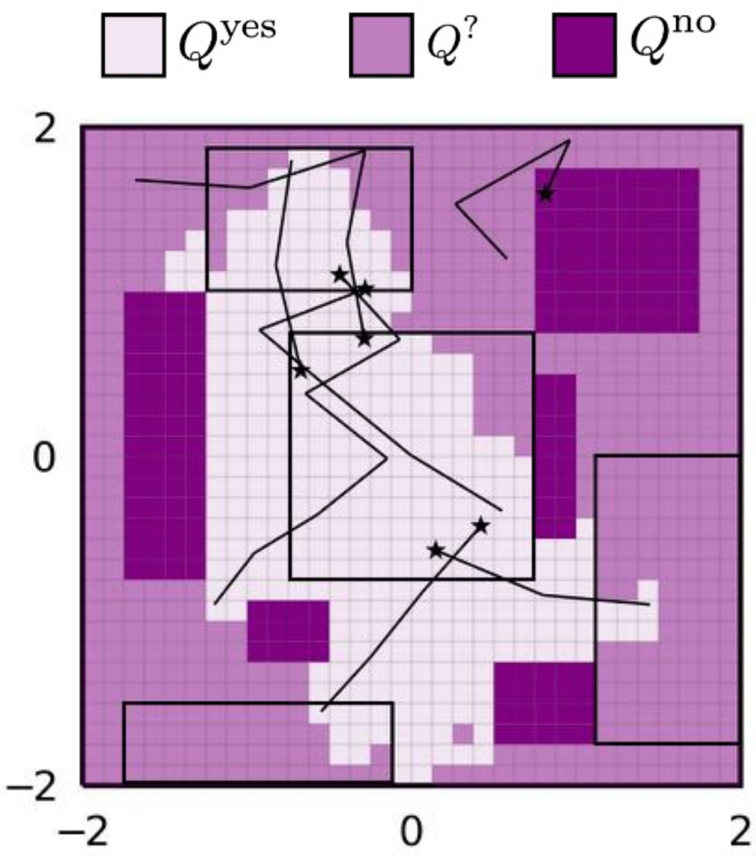
$$g_u(x) = \begin{cases} 0.5 + 0.2\sin(x(2)), 0.4\cos(x(1)) & \text{if } u = 1 \\ -0.5 + 0.2\sin(x(2)), 0.4\cos(x(1)) & \text{if } u = 2 \\ 0.4\cos(x(2)), 0.5 + 0.2\sin(x(1)) & \text{if } u = 3 \\ 0.4\cos(x(2)), -0.5 + 0.2\sin(x(1)) & \text{if } u = 4 \end{cases}$$

cardinal movements with nonlinear terms

$$\phi \equiv \mathcal{G}(\neg O) \wedge \mathcal{F}(D1) \wedge \mathcal{F}(D2)$$

State Space Labels

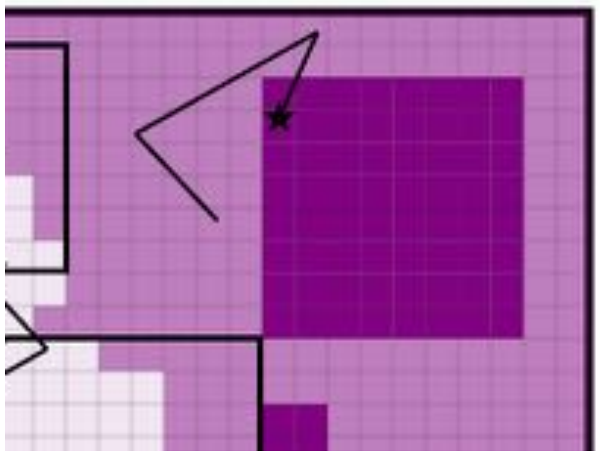




Conclusion

Data-driven abstraction-based strategy synthesis for a **switched stochastic** system with **guarantees that extend to the latent system**.

- Optimal choices for discretization and parameter bounds, RKHS constants
- Extension to online strategy refinement and synthesis
- Working on a Julia toolbox for the community 🌞



Thank you for listening!

John Jackson

john.m.jackson@colorado.edu

NSF

C-UAS

References

[1] S. Haesaert, P. M. Van den Hof, and A. Abate, "Data-driven and model-based verification via bayesian identification and reachability analysis," Automatica, vol. 79, pp. 115–126, 2017.

[2] J. Kenanian, A. Balkan, R. M. Jungers, and P. Tabuada, "Data driven stability analysis of black-box switched linear systems," Automatica, vol. 109, p. 108533, 2019.

[3] M. Ahmadi, A. Israel, and U. Topcu, "Safety assessemt based on physically-viable data-driven models," in 2017 IEEE 56th Annual Conference on Decision and Control (CDC), pp. 6409–6414, IEEE, 2017.

[4] Berkenkamp, F., Turchetta, M., Schoellig, A. P. & Krause, A. Safe Model-based Reinforcement Learning with Stability Guarantees. Arxiv (2017).

[5] Akametalu, A. K. et al. Reachability-based safe learning with Gaussian processes. 53rd Ieee Conf Decis Control 1424–1431 (2014) doi:10.1109/cdc.2014.7039601.

[6] Jagtap, P., Pappas, G. J. & Zamani, M. Control Barrier Functions for Unknown Nonlinear Systems using Gaussian Processes*. 2020 59th IEEE Conf Decis Control Cdc 00, 3699–3704 (2020).

[7] Wells, A. M., Lahijanian, M., Kavraki, L. E. & Vardi, M. Y. LTLf Synthesis on Probabilistic Systems. Electron Proc Theor Comput Sci 326, 166–181 (2020).

[8] S.R. Chowdhury and A.Gopalan,"On kernelized multi-armed bandits," in *Proceedings of the 34th International Conference on Machine Learning-Volume 70*, pp. 844–853, JMLR. org, 2017.

[9] De Giacomo, G., & Vardi, M. Y. (2013). Linear temporal logic and linear dynamic logic on finite traces. In IJCAI'13 Proceedings of the Twenty-Third international joint conference on Artificial Intelligence (pp. 854-860). Association for Computing Machinery.

[10] M. Lahijanian, S. B. Andersson, and C. Belta, "Formal verification and synthesis for discrete-time stochastic systems," *IEEE Transactions on Automatic Control*, vol. 60, pp. 2031–2045, Aug. 2015.

Parameter Variations

Minimum Probability of Satisfaction

Theorem 1 Parameter Uniformly Increasing