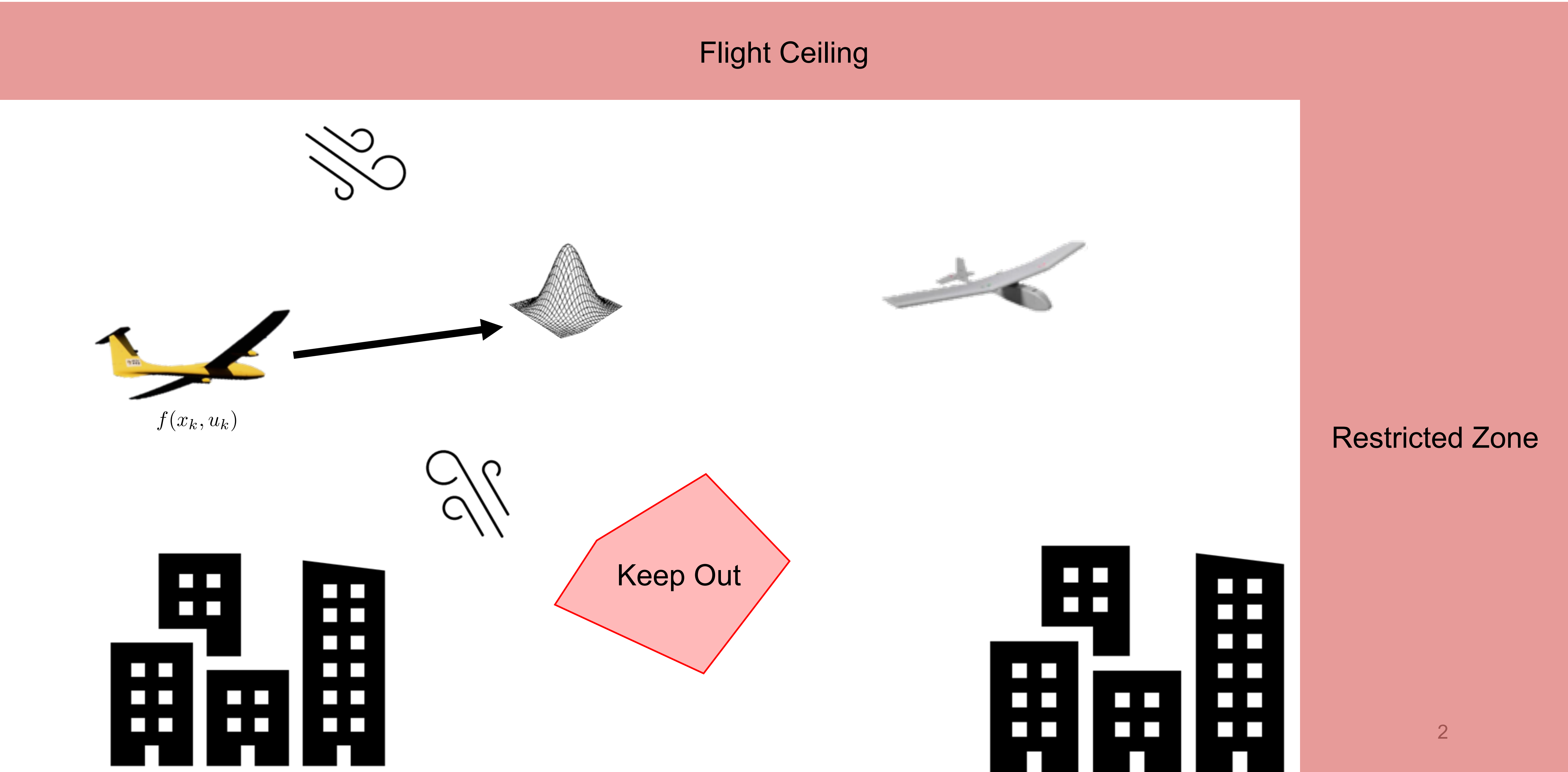
Safety Verification of Unknown Dynamical Systems via Gaussian Process Regression

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

john.m.jackson@colorado.edu

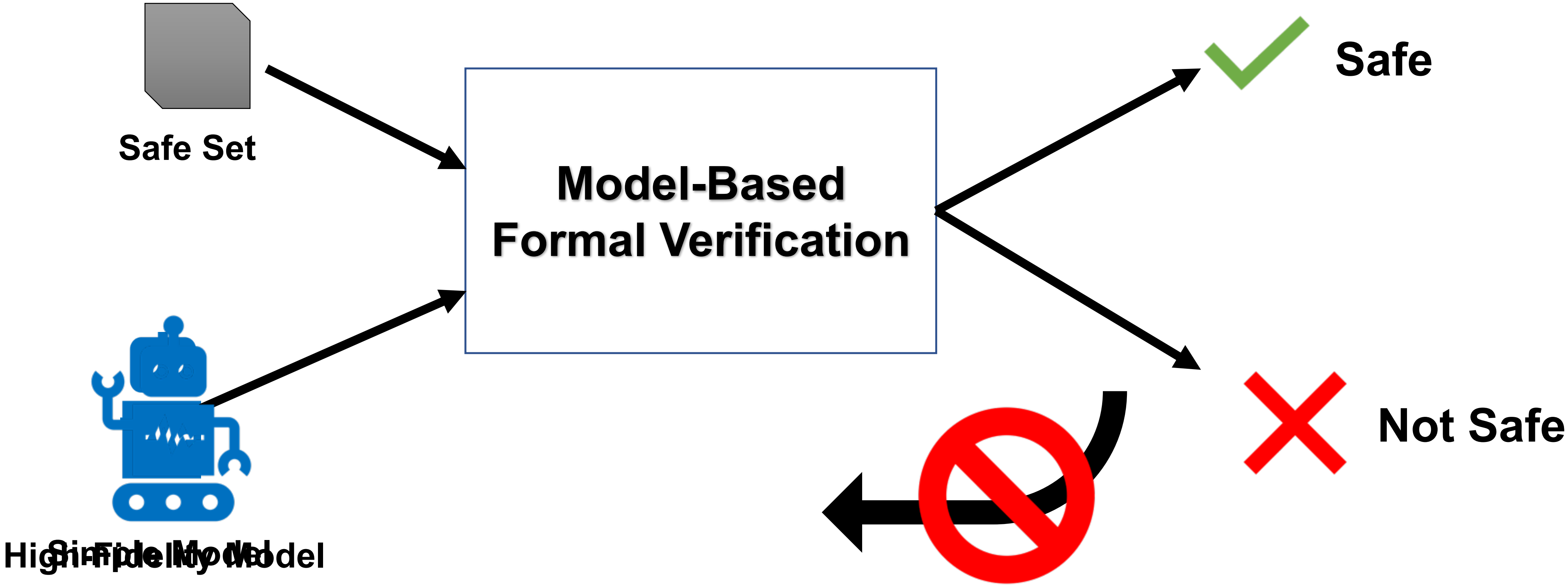
59th IEEE Conference on Decision and Control (2020)

Safety-Critical Urban Flying



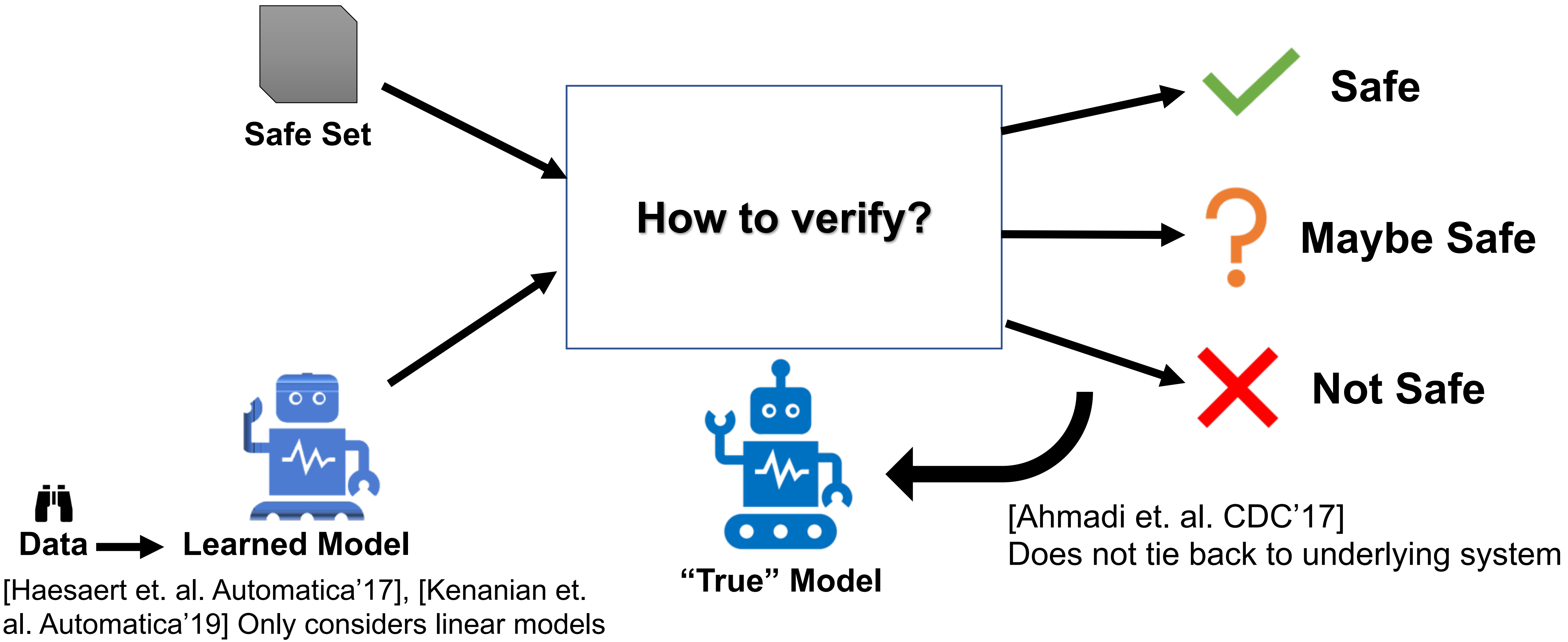
Model-Based Verification

Safety: Does the system leave the safe set?



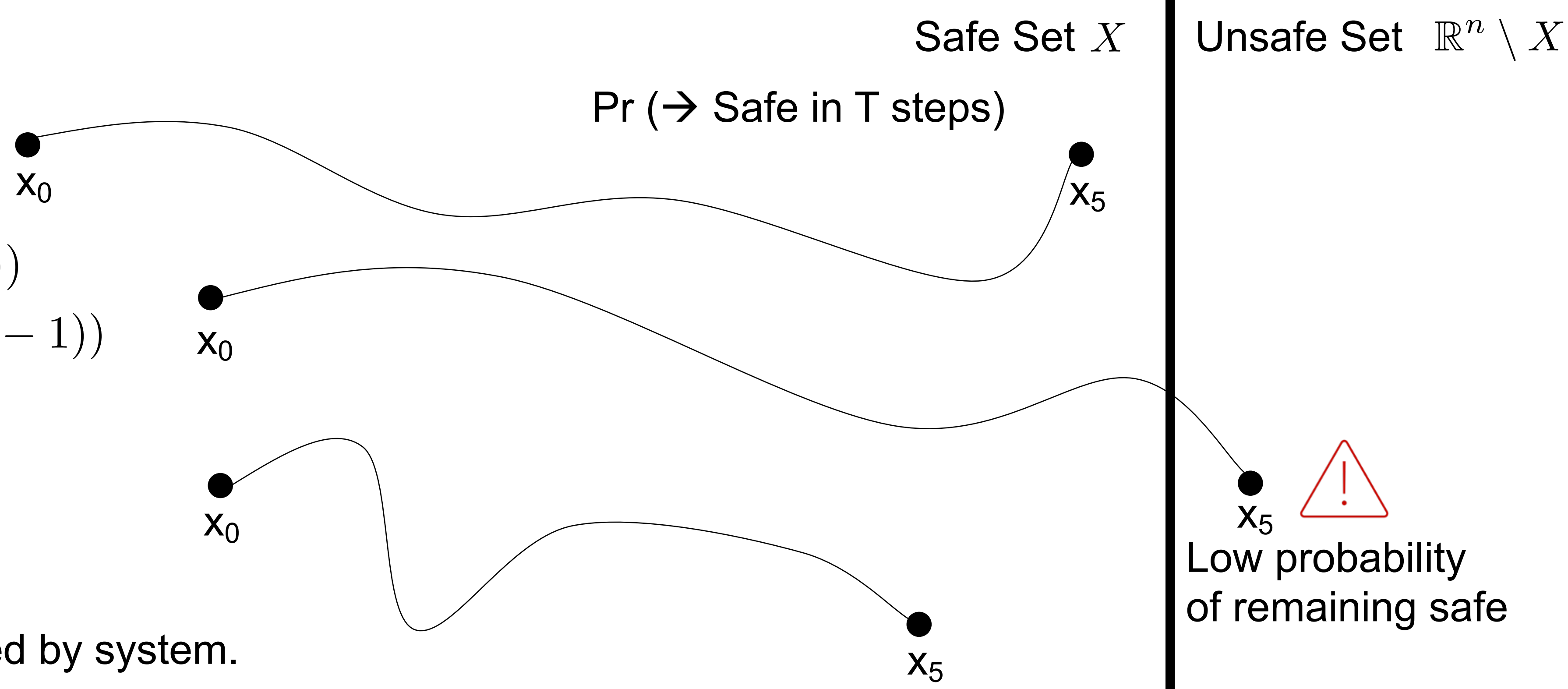
Data-Driven Verification

Safety: Does the system leave the safe set?



Challenge: Guarantees for potentially nonlinear systems.

Problem Overview



$$\mathbf{x}(k+1) = f(\mathbf{x}(k), \mathbf{u}(k))$$

$$\mathbf{y}(k) = \mathbf{x}(k) + \mathbf{v}(k, \mathbf{u}(k-1))$$

f is unknown

$\mathbf{x}(k) \in \mathbb{R}^n$

$\mathbf{u}(k) \in \mathcal{U}$

$\mathbf{v}$ is i.i.d. bdd. noise

Given: a dataset generated by system.

Problem: Compute bounds on the probability of remaining within the safe set over a horizon (i.e. for any control policy)

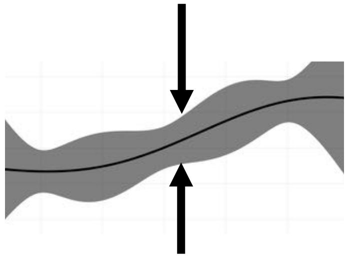
Solution Approach

Safe

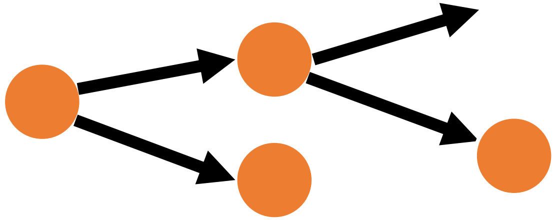
Maybe Safe

Not Safe

Unknown System



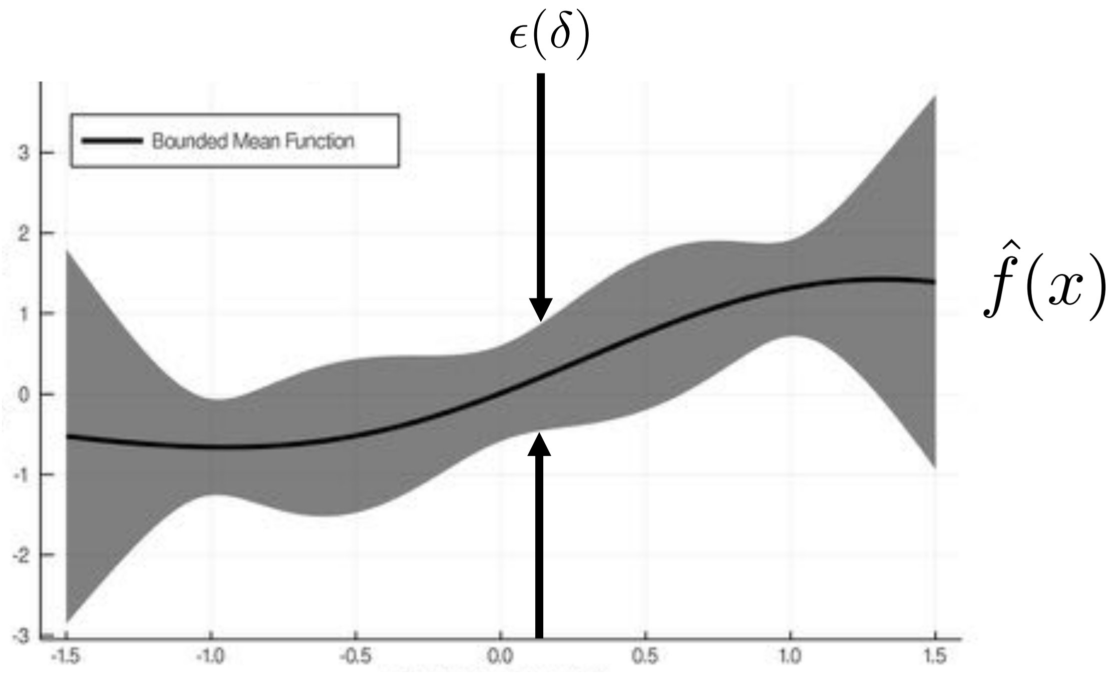
GP Regression



Interval Markov Decision Process

Probabilistic Bounds on Safety

Learning with Gaussian Processes



Given: prior mean and covariance functions, dataset

Procedure: Joint-Gaussian assumption and Bayesian conditioning

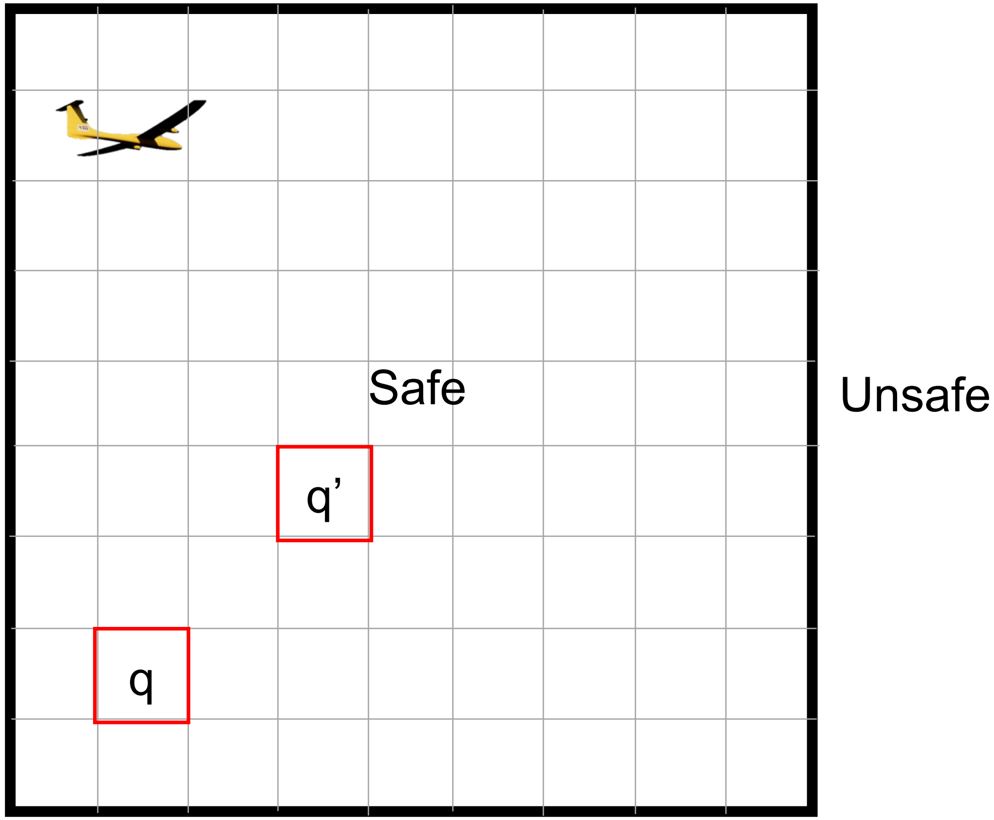
Result: MLE estimate of function, posterior covariance

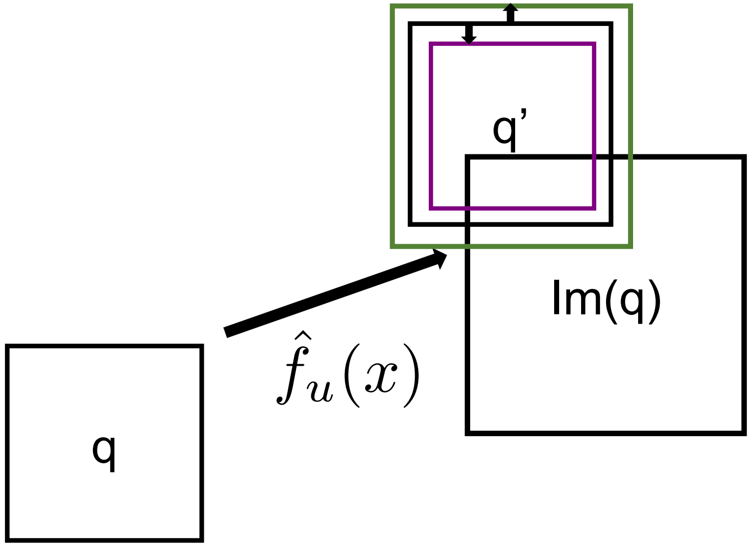
RKHS Assumption [Chowdhury et. al. JMLR’17**]**

- Unknown function lies in the span of the prior covariance function (continuity)
- General probabilistic error bounds using GP regression

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

Interval MDP Construction





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

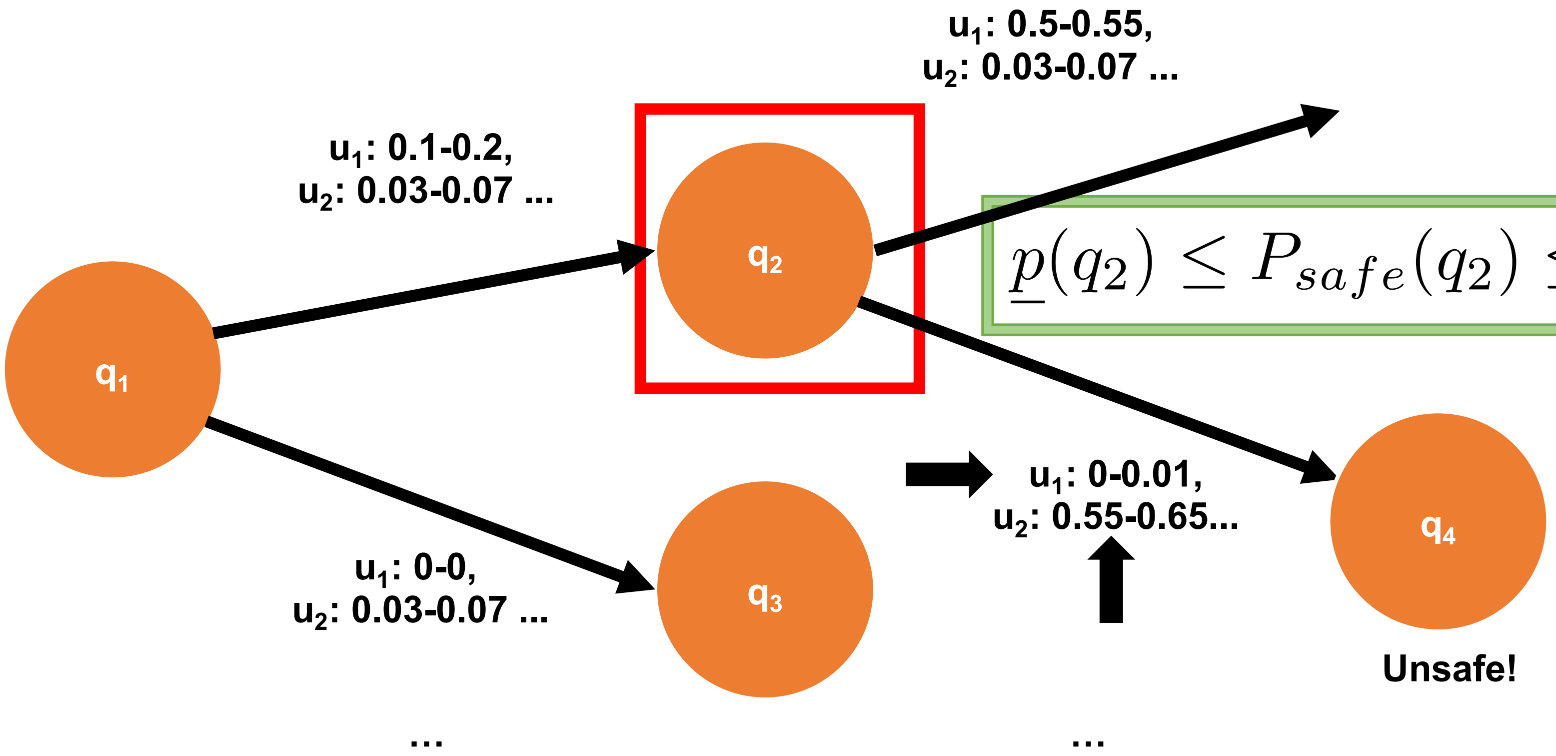
$$\overline{\Pr} = 1 - \Pr(|f(x) - \hat{f}(x)| \leq \epsilon)(1 - \mathbf{1}_{Im(q),q'})$$

$$\underline{\Pr} = \mathbf{1}_{Im(q),q'} \Pr(|f(x) - \hat{f}(x)| \leq \epsilon)$$

Accounts for **discretization** and **regression** error.

Verification on IMDP

→ Value Iteration over policies and adversaries (transition distributions) [Lahijanian et. al. TAC'15]



$$\boxed{\underline{p}(q_2) \leq P_{safe}(q_2) \leq \bar{p}(q_2)}$$

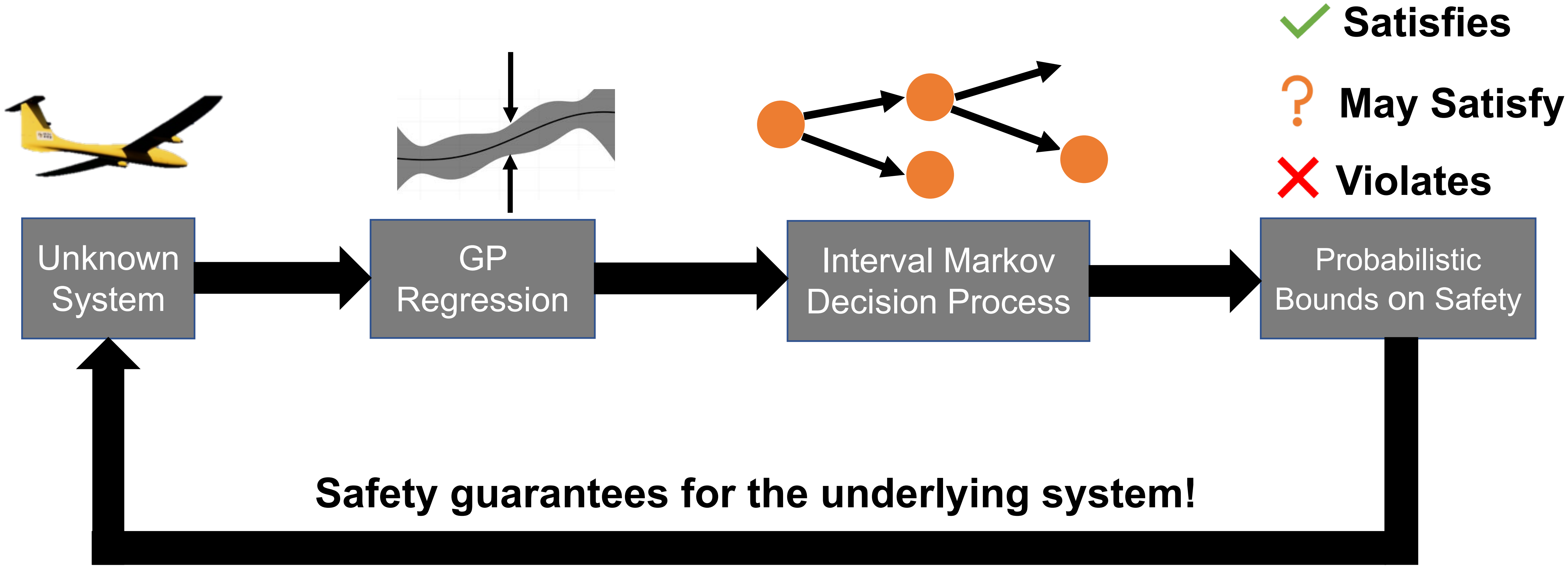
✓ $\underline{p}(q_2) \geq 0.95$

✗ $\bar{p}(q_2) \leq 0.95$

? otherwise

$$P_{safe}(q_2) = 1 - P_{unsafe}(q_2)$$

Approach

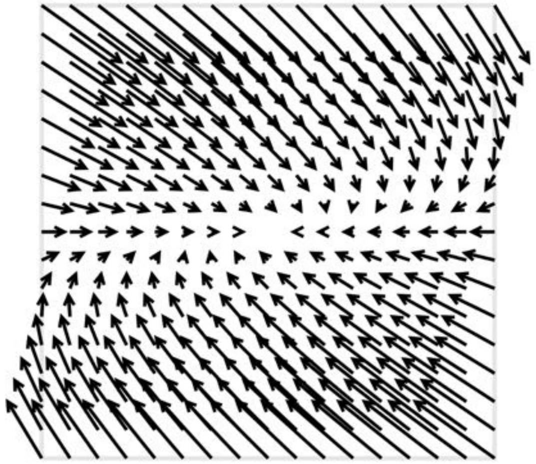

Satisfies
May Satisfy
Violates
Unknown System
GP Regression
Interval Markov Decision Process
Probabilistic Bounds on Safety
Safety guarantees for the underlying system!

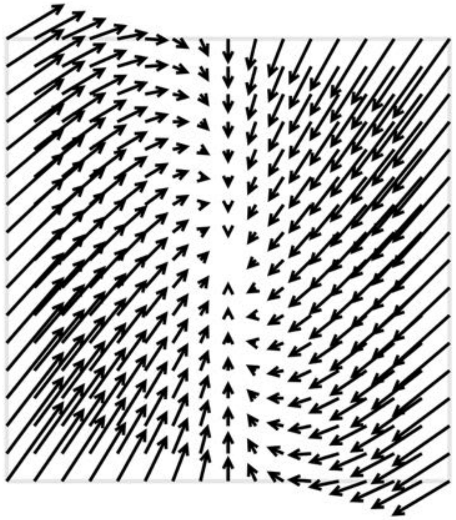
Switched System Verification

→ GPs trained w/ 1000 data points
→ noise variance: 0.01^2

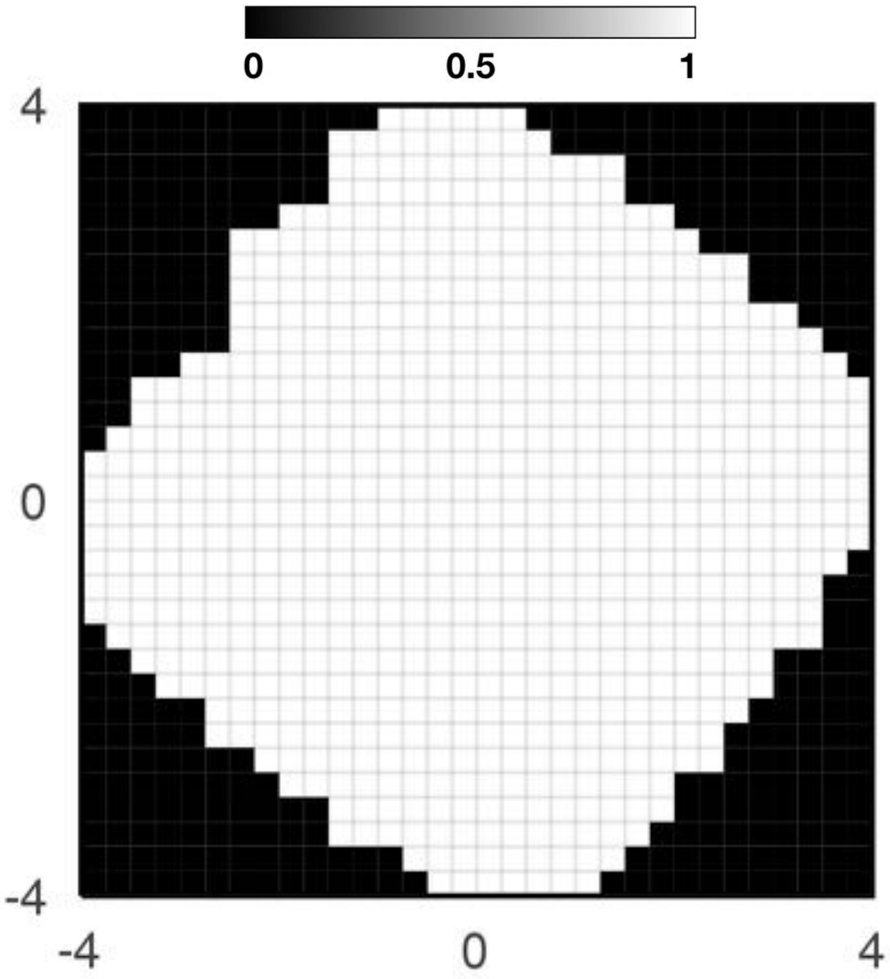
$$f(\mathbf{x}(k)) = A_i\,\mathbf{x}(k)$$

$$A_{\text{upper}} = \begin{bmatrix} 0.8 & 0.5 \\ 0 & 0.5 \end{bmatrix} \qquad A_{\text{lower}} = \begin{bmatrix} 0.5 & 0 \\ -0.5 & 0.8 \end{bmatrix}$$

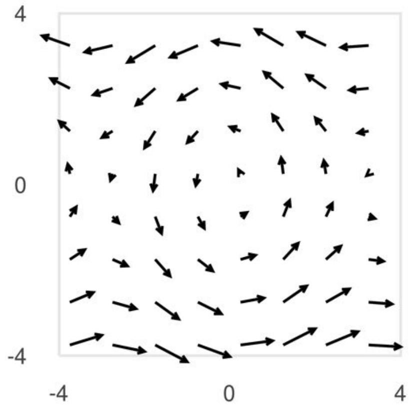




Minimum Probability of Safety (1000 steps)

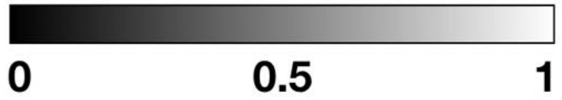


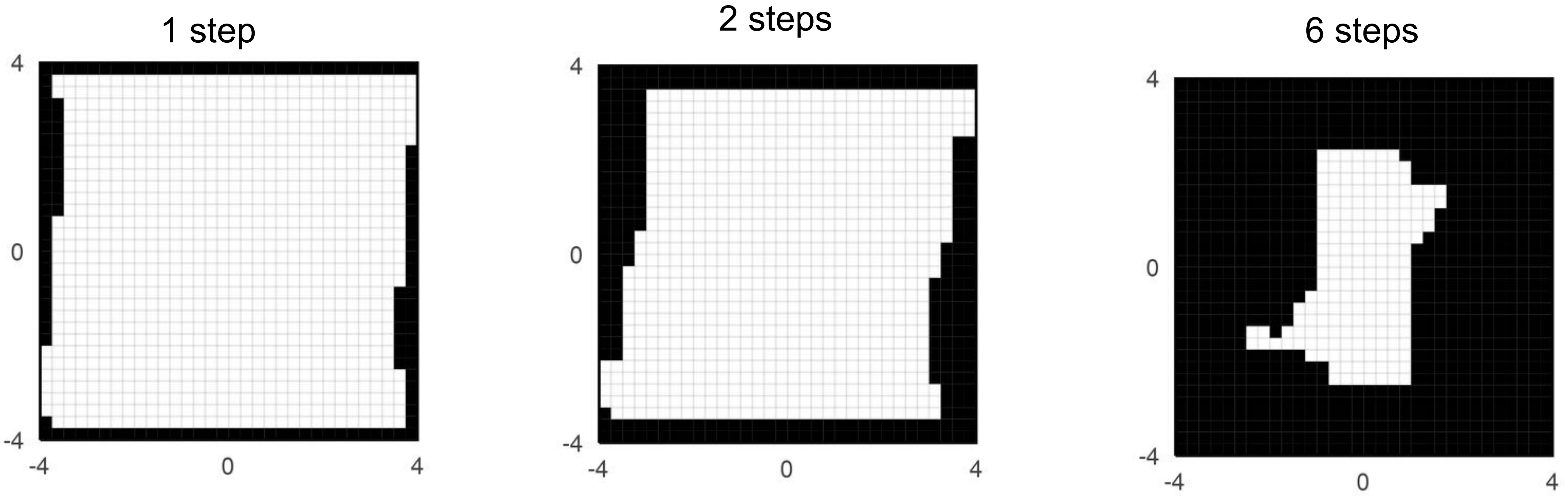
Nonlinear System Verification



$$f(\mathbf{x}(k)) = [\,\mathbf{x}_1(k) - 0.05\,\mathbf{x}_2(k),\ \ \mathbf{x}_2(k) + 0.1\sin(\mathbf{x}_1(k))\,]^T$$

Minimum Probability of Safety

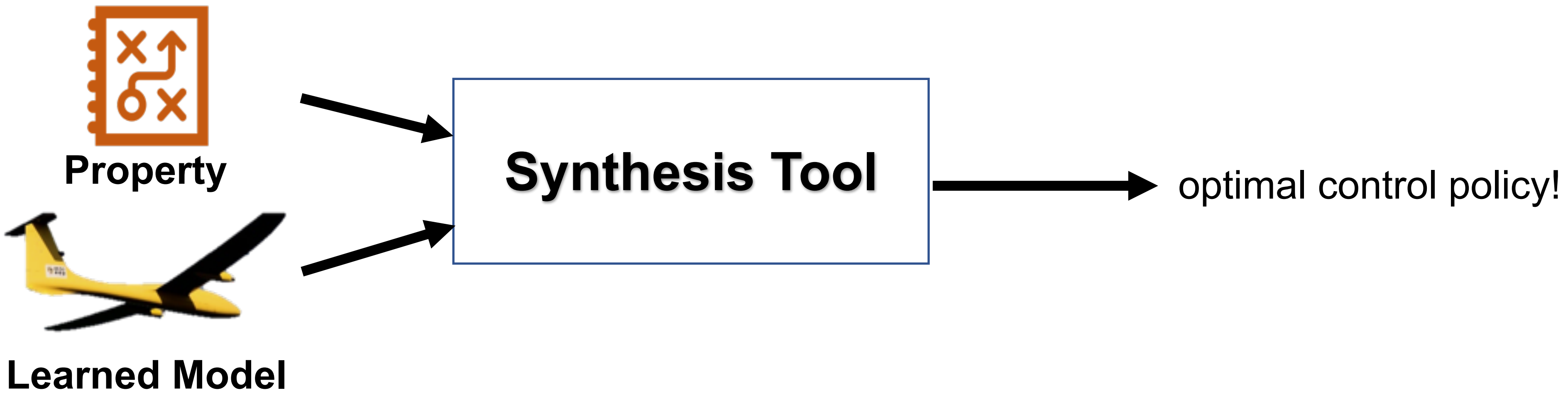




Future Work

- Optimal choices of discretization, transition bound parameters
- Bounding the cryptic RKHS constants in an easy-to-apply way
- Extension to online verification, synthesis problem

$\varphi = \Diamond\, target \wedge \Box \neg\, unsafe$



Thank you for listening!

John Jackson

john.m.jackson@colorado.edu

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] 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.

[5] 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.