Synergistic Offline-Online Control Synthesis via Local Gaussian Process Regression

John Jackson[1], Luca Laurenti[2], Eric Frew[1], Morteza Lahijanian[1]

[1]University of Colorado Boulder, [2]TU Delft | Contact: john.m.jackson@colorado.edu

IEEE Conference on Decision and Control 2021

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

ARIA Systems
Assured Reliable Interactive Autonomous

Safety Critical Autonomous Systems

$$\underbrace{x_{k+1}}_{} = \overbrace{f(x_k, a_k)}^{\text{Known}} + \overbrace{g(x_k, a_k)}^{\text{Unknown!}} + \underbrace{w_k}_{\text{Sub-Gaussian Noise}}$$

Mission $\phi =$ Pick up the package and safely deliver it, avoid other vehicles and respect restrictions

- Linear Temporal Logic over finite traces (LTLf)

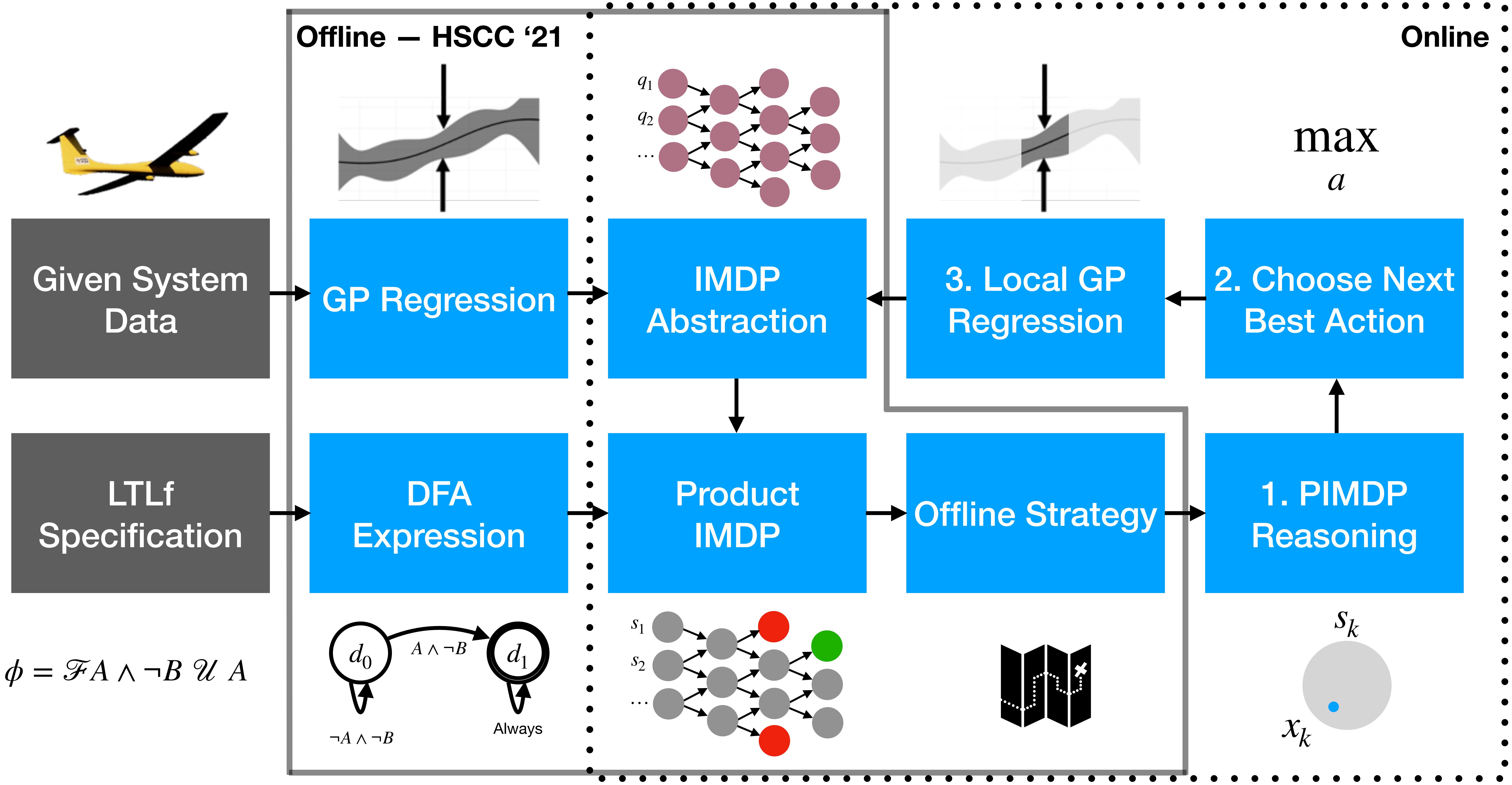
Want to find **strategy** π (maps from the vehicle path to actions) that has **guarantees** on satisfying ϕ

Given: a dataset generated by the system and the ability to collect more data.

How to use offline and online observations of g to synthesize a strategy π that maximizes the chance of satisfying ϕ?

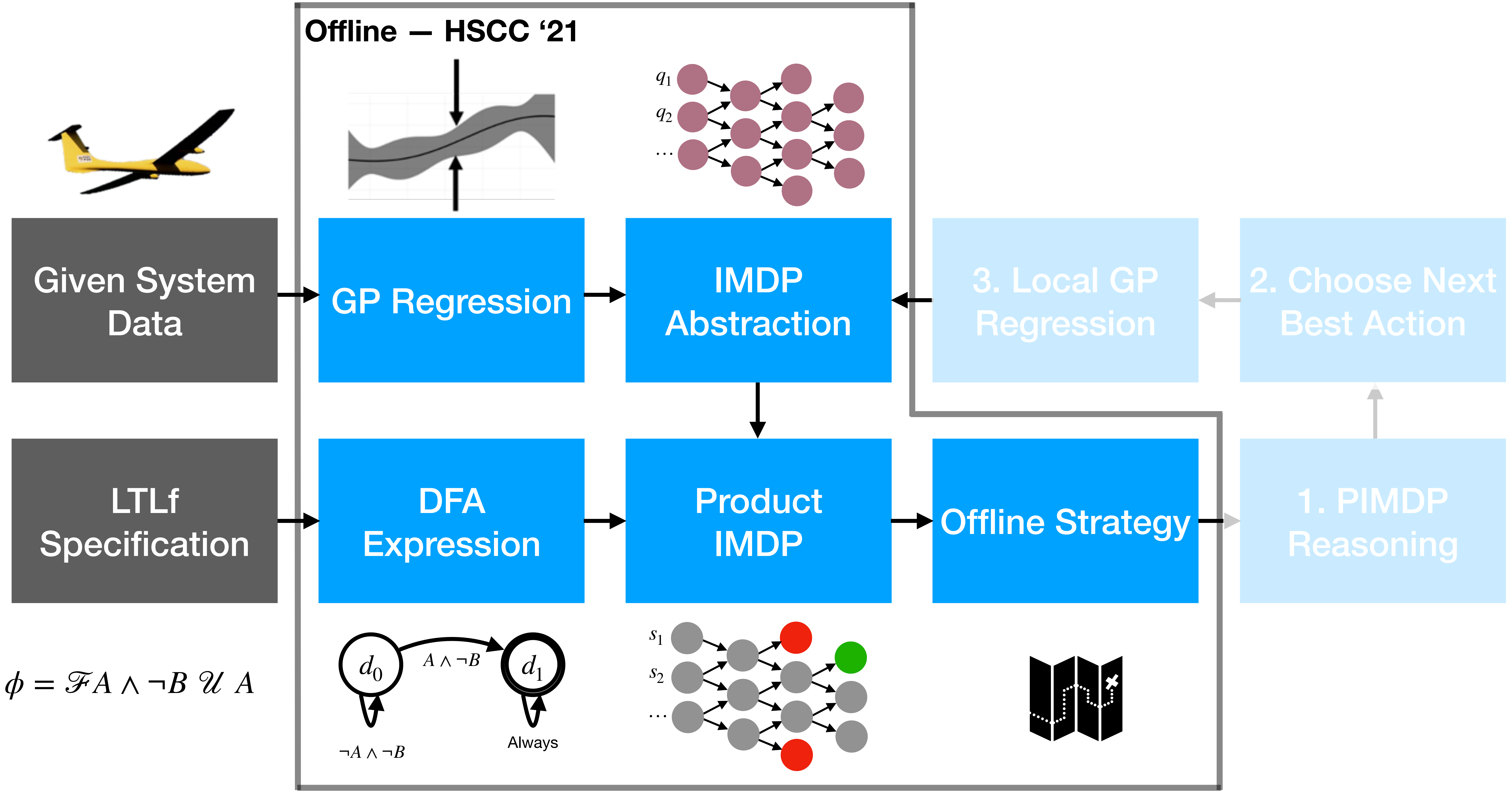
Current Framework Outline

Formal Control Synthesis + Machine Learning Regression



Jackson, John, et al. "Strategy synthesis for partially-known switched stochastic systems." Proceedings of the 24th International Conference on Hybrid Systems: Computation and Control. 2021.

Offline Components

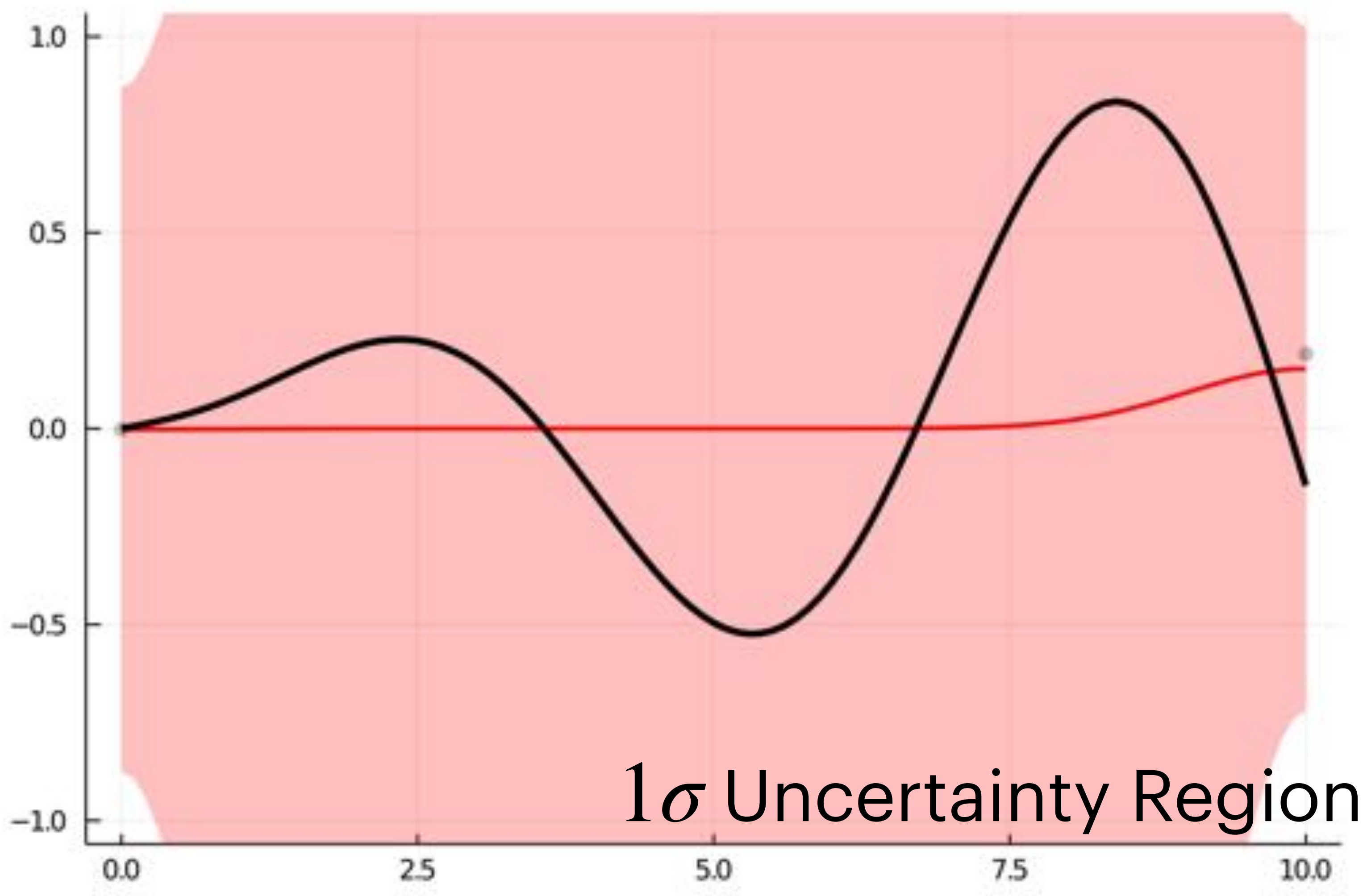


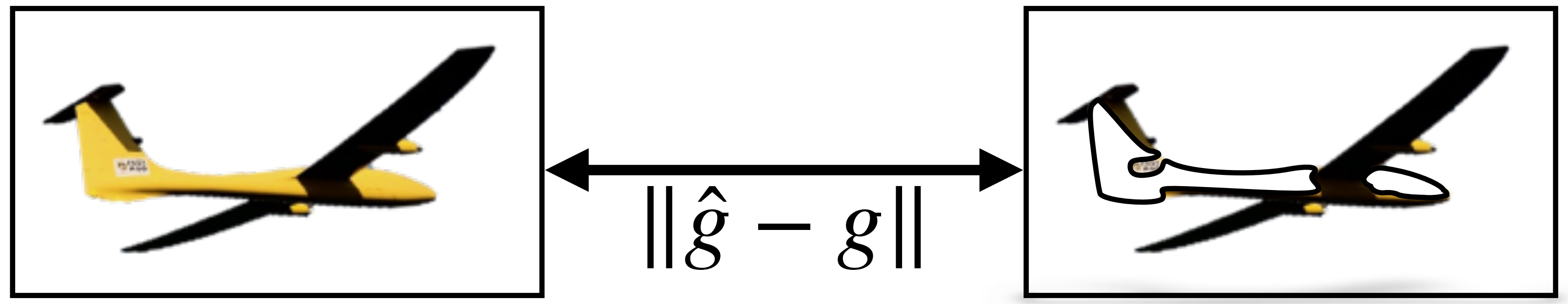
GP Regression and Abstraction

Gaussian Process (GP) Regression

True g (black)

Est. $\hat{g}$ (red)





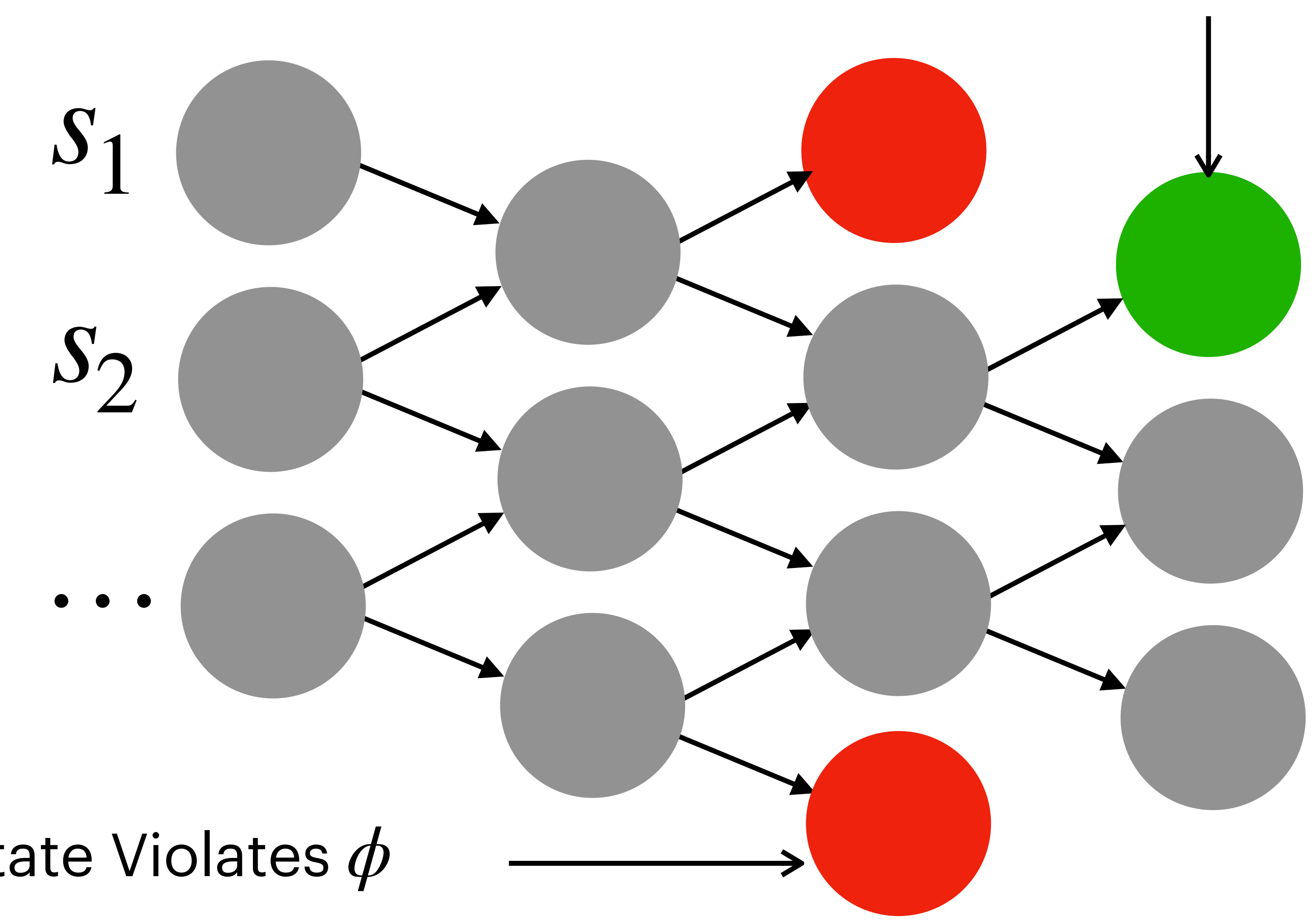
w_k may be non-Gaussian!

Reproducing Kernel Hilbert Space = Span of the Kernel Function

$$\Pr(\|\hat{g} - g\| \leq \epsilon) = 1 - \delta$$

Product IMDP

Interval MDP X DFA

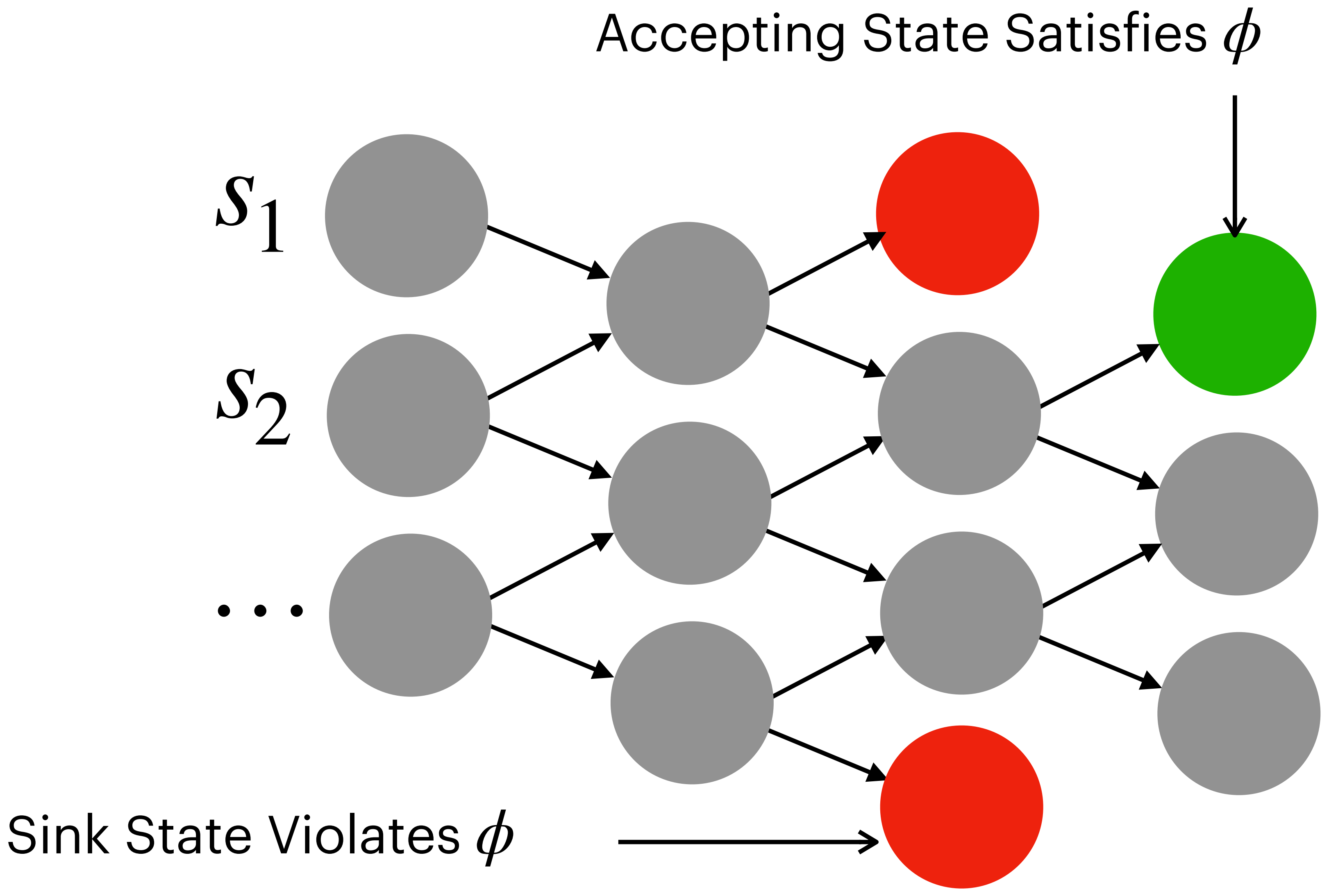


$$\Pr(s \to s' \mid a) \in [\underline{p}, \overline{p}]$$

- Abstracts systems behavior w.r.t ϕ
- Interval transitions account for
  - learning uncertainty
  - stochasticity
  - discretization error

PIMDP Abstraction and Synthesis

Interval MDP
X
DFA
→ **Product IMDP**



- Offline strategy $\pi : S \rightarrow A$ chooses actions to maximize **lower bound (LB) value function:**

$$\check{p}(s) = \max_{a} \min_{\gamma} \sum_{s' \in S} \gamma_s^a(s') \check{p}(s')$$

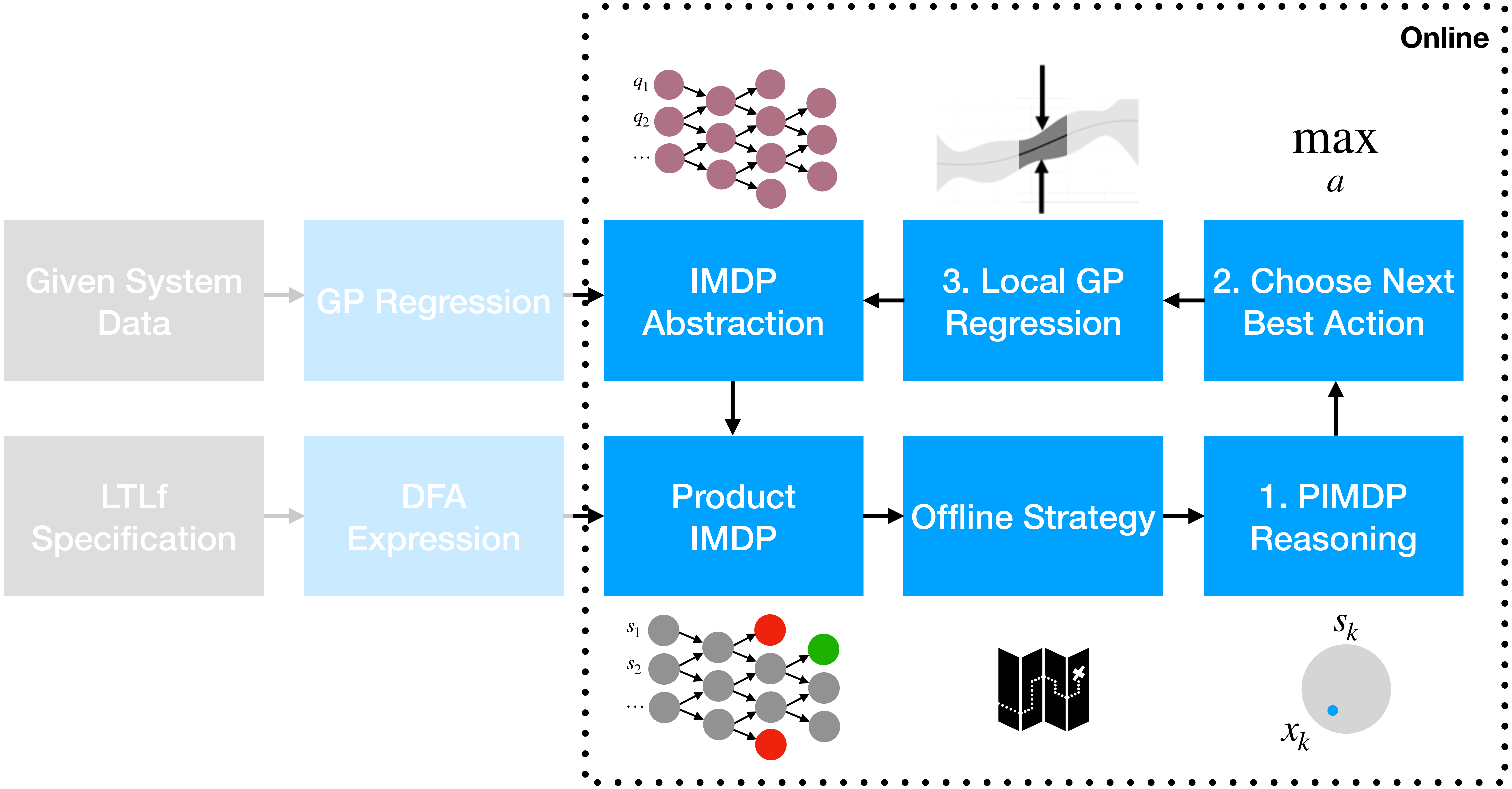
(a.k.a. minimum probability of satisfying ϕ from s)

- Strategy synthesis is a 2.5 player game:
  - Adversary γ that drives us to sink states
  - Actions (us) a that drives us to accepting states
- **Strategy and the guarantees map back to the system!**

When uncertainty is high (e.g. low data), the LB can be 0.

How can we refine the offline strategy using online data to increase the chance of satisfying ϕ, i.e. completing the mission?

Online Extension



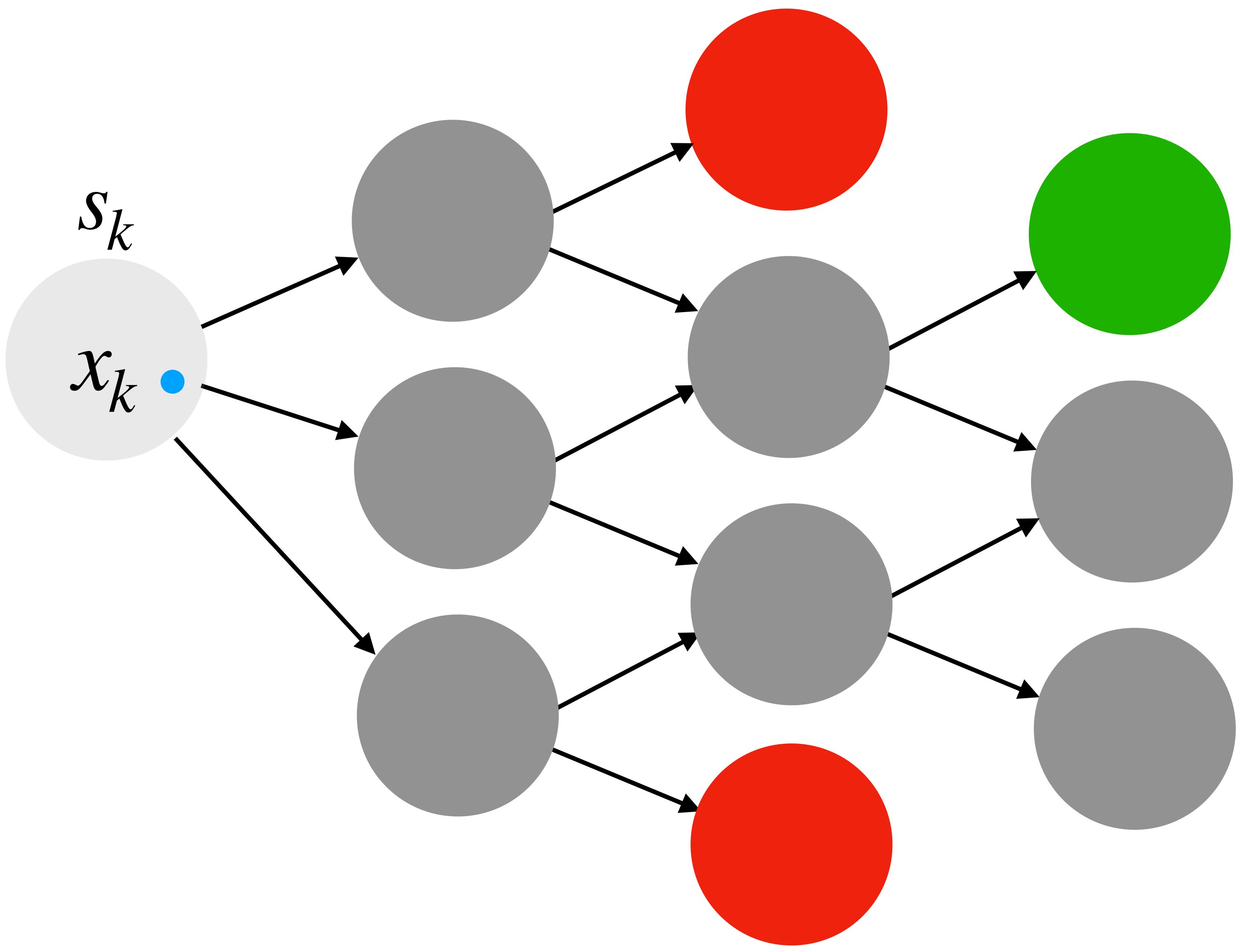
1. PIMDP Reasoning

Get Current State

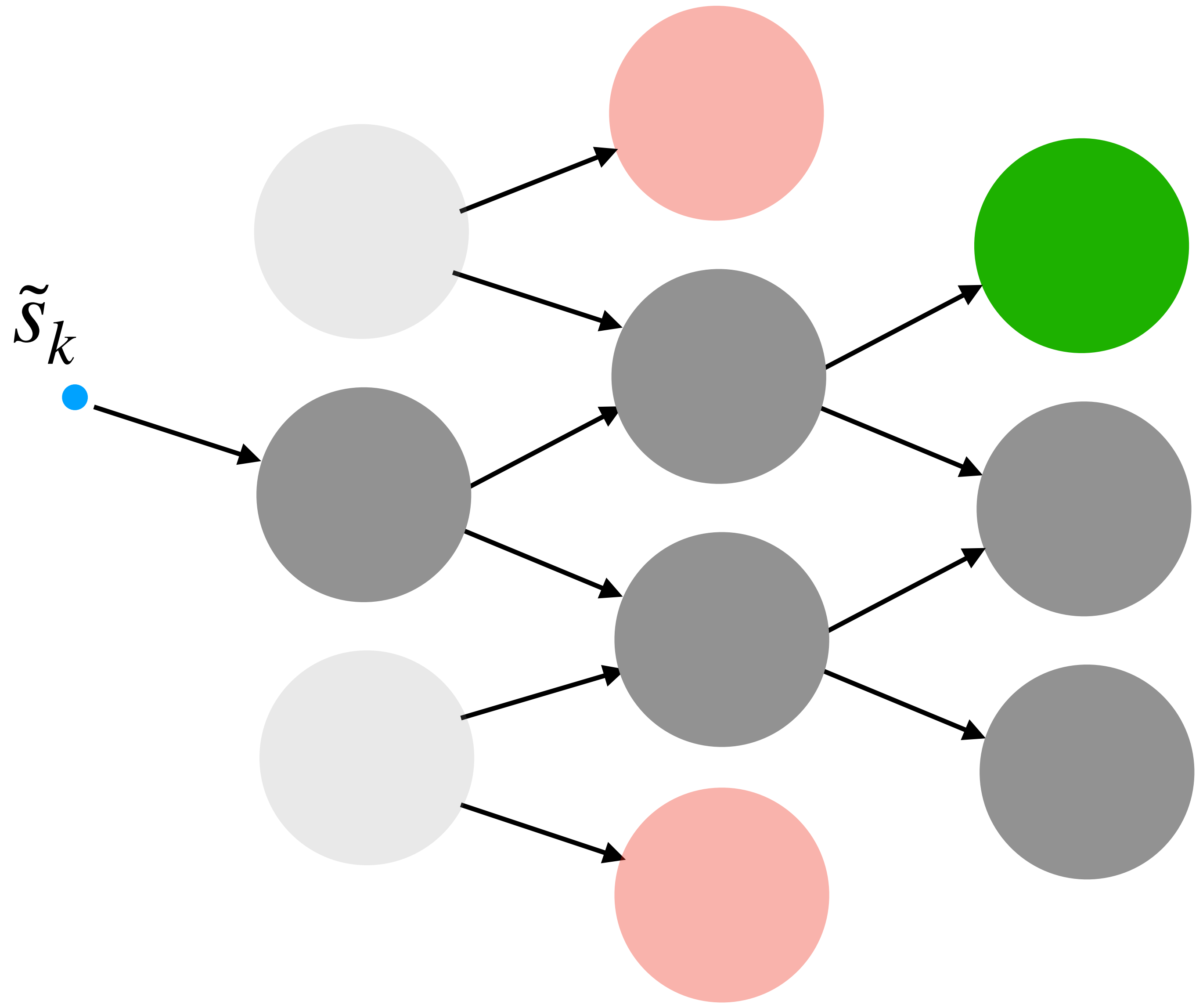
$$x_k = f(x_{k-1}, a_{k-1}) + g(x_{k-1}, a_{k-1}) + w_k$$

$$x_k \in s_k$$

Identify Location in PIMDP Abstraction



Prune Edges from New Augmented PIMDP State



Calculate Lower Bound Value Function

$$\check{p}(\tilde{s}_k) = \max_a \min_\gamma \sum_{s' \in S} \gamma^a_{\tilde{s}_k}(s')\check{p}(s')$$

LB Values are unchanged for all original PIMDP States.

Monotonic Increase in the lower bound value function!

2. Choose Next Best Action

Maximize LB Value

Calculate the LB probability of satisfying ϕ.

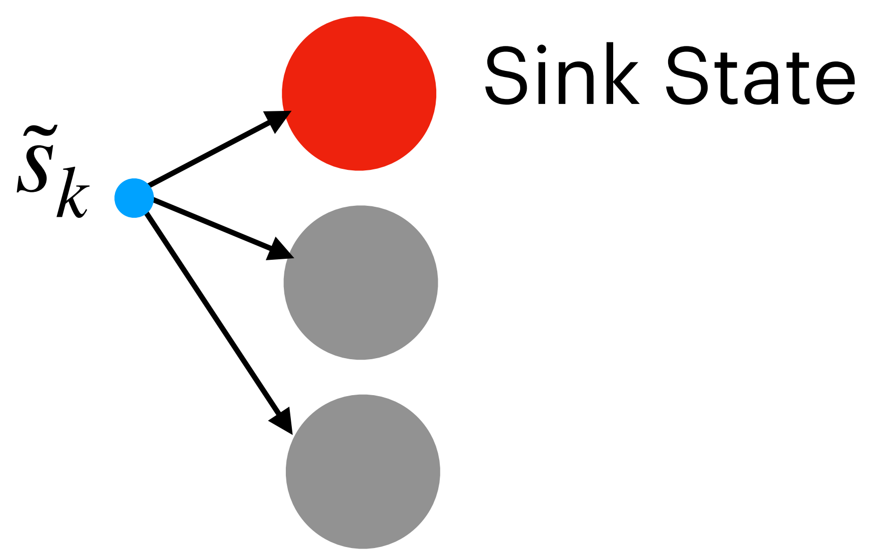
Choose a to maximize:

$$\check{p}(\tilde{s}_k) = \max_a \min_\gamma \sum_{s' \in S} \gamma^a_{\tilde{s}_k}(s')\check{p}(s')$$

- If LBV $\rightarrow$1, we are sure to satisfy ϕ
- Often get ties, e.g. LBV = 0 for all actions
- Other metrics to break ties?

+ Avoid Sink States

Determine the # of edges to the sink state over the next step.

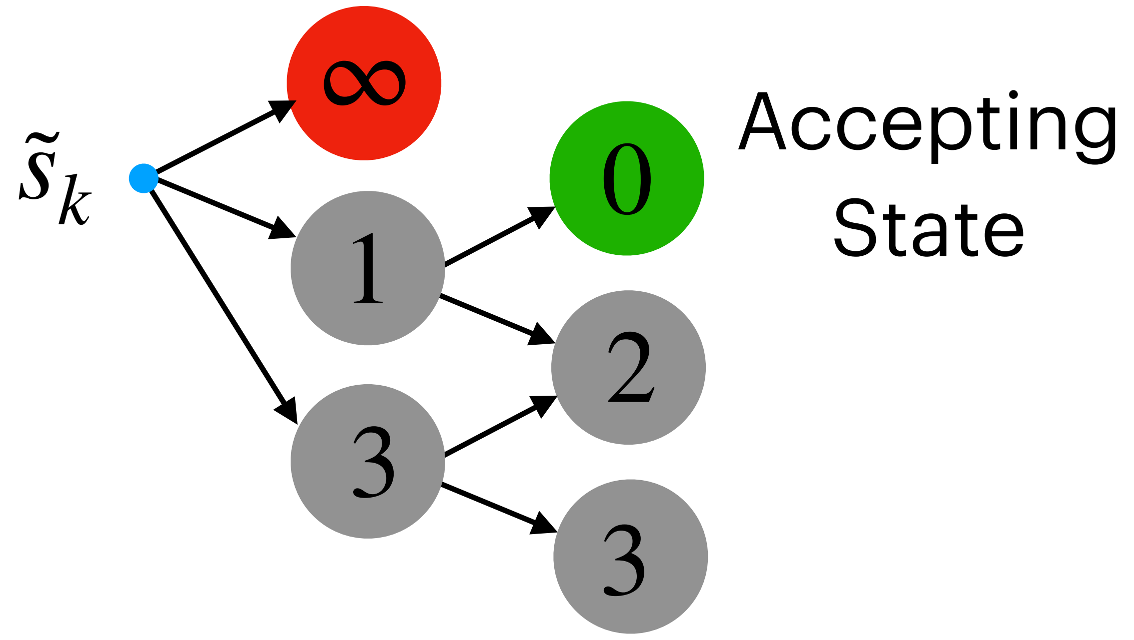


Choose a to minimize the chance of reaching a sink state in the PIMDP.

- Can lead to preferring safety over making progress

+ Progress on PIMDP

Calculate the # of edges to the accepting state.



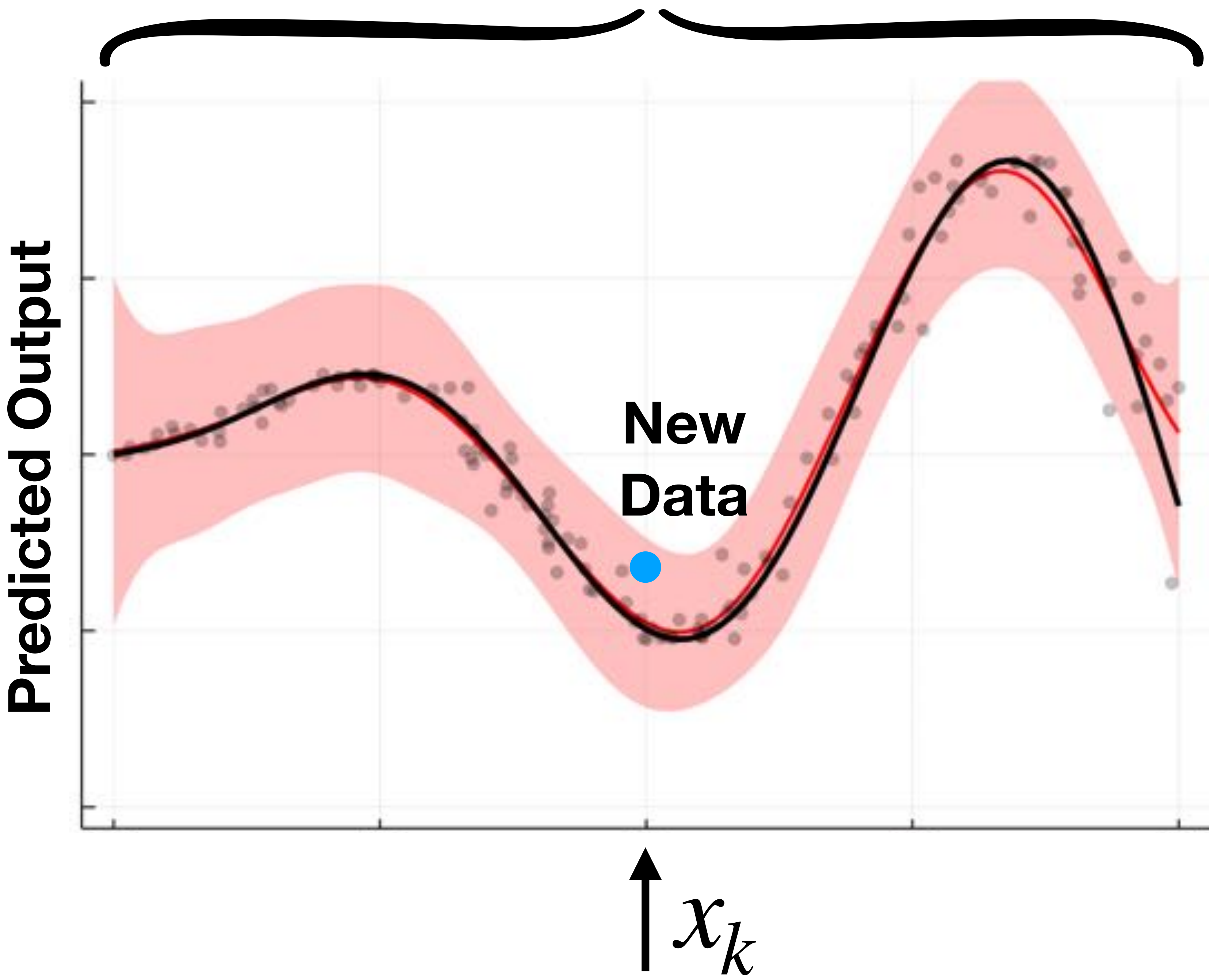
Choose a to minimize expected # of edges to the accepting state.

- Combine with sink metric

3. Local GP Regression and Update

Global GP

Use All Data to Predict $g(x_k)$

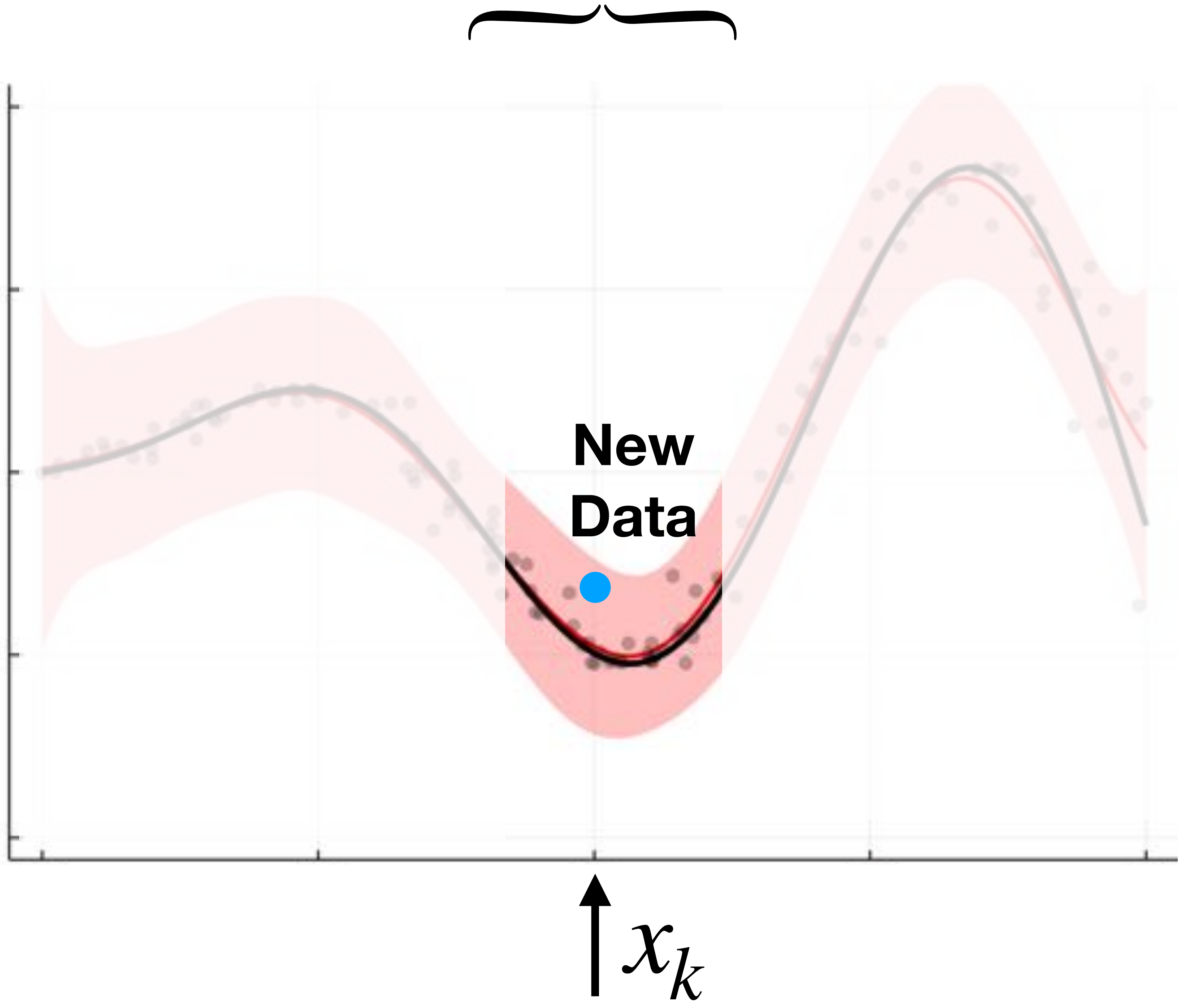


$+$ **smallest predictive variance (uncertainty) using all data**

$-$ $\mathcal{O}(n^3)$ **makes it expensive to update with new data**

Local GP

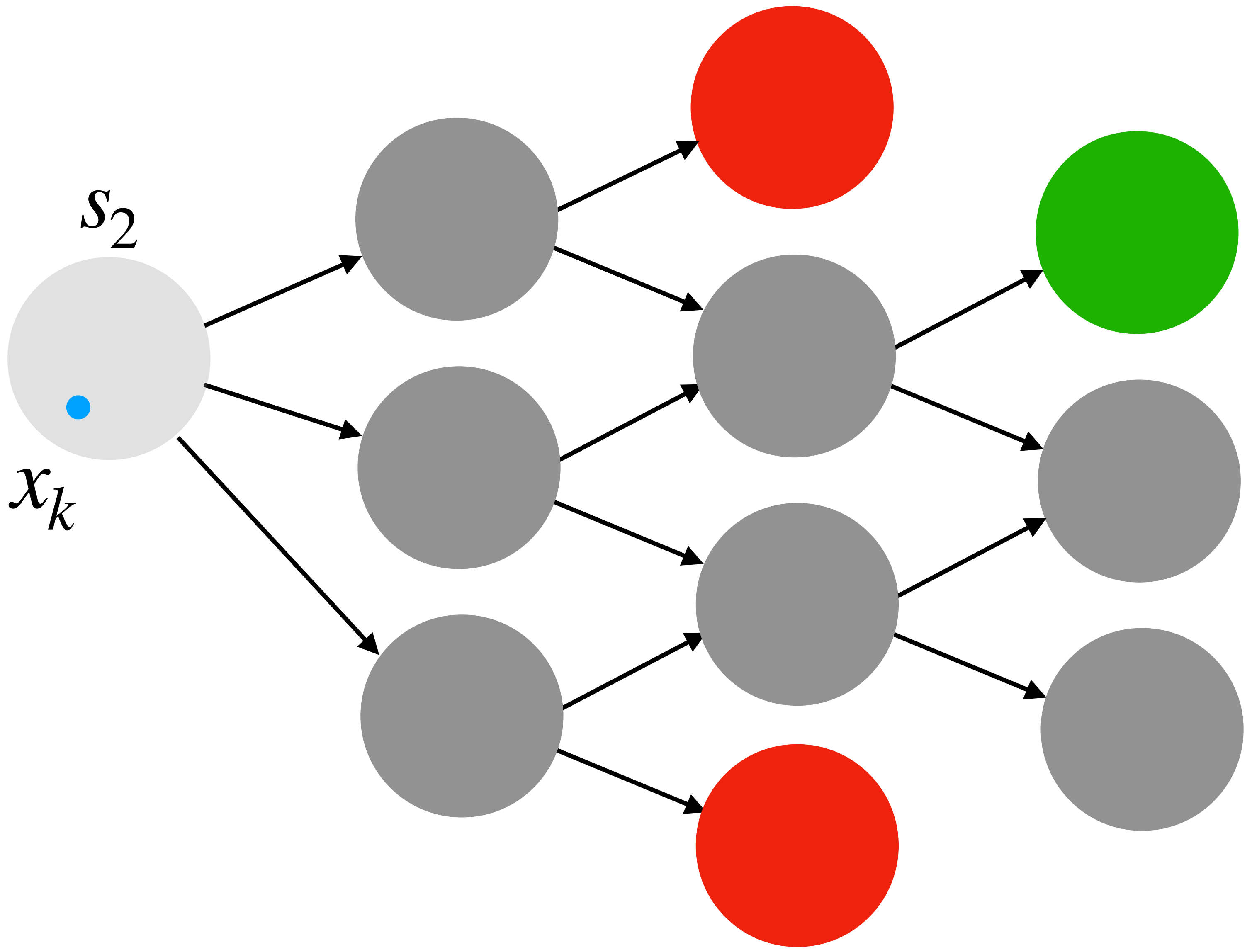
Use N-Nearest Data to Predict $g(x_k)$



$+$ **choosing $m < n$ data for faster updates and reasoning**

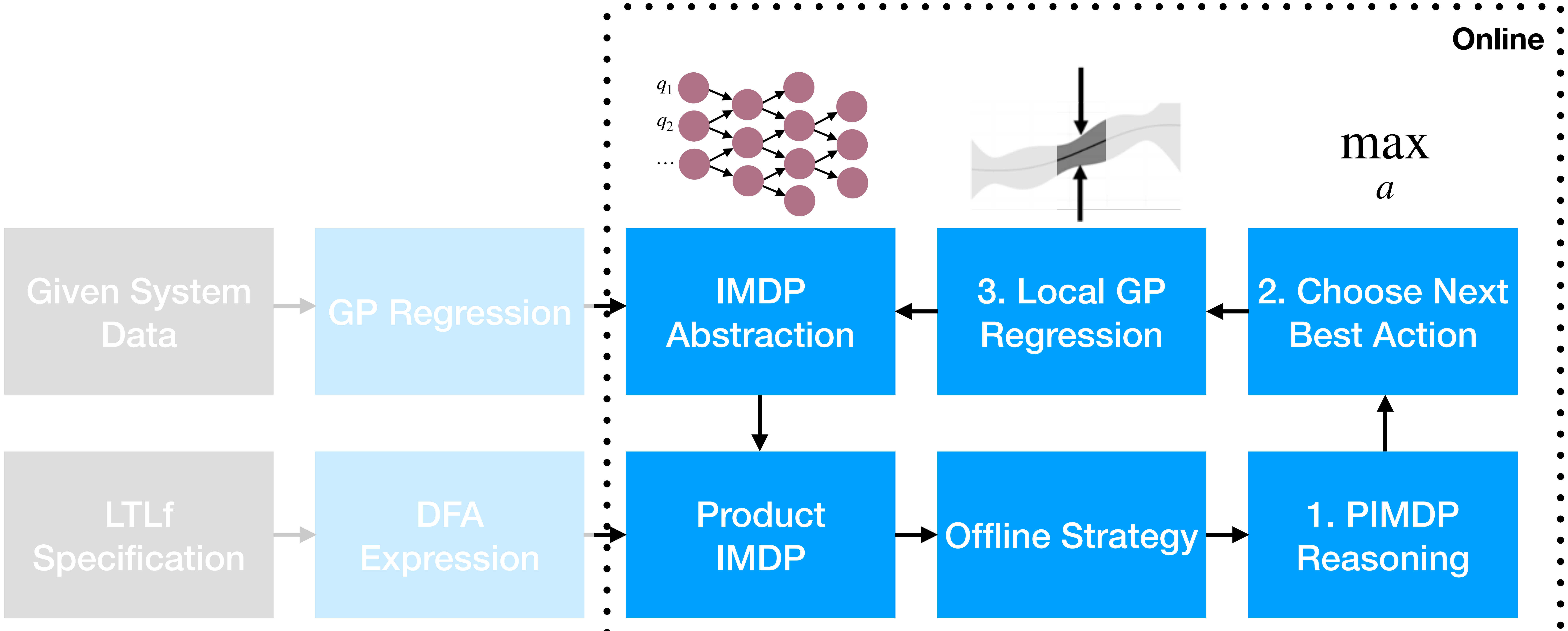
$-$ **increases predictive variance (uncertainty)**

Abstraction Update



- **update PIMDP states near the new data**
- **intervals shrink and become more certain as more data is collected**

Online Extension



Guaranteed to at least as well the offline strategy —
empirically, we observe improvements!

Example System

$$x_{k+1} = x_k + w_k + \begin{cases} [0.25 + 0.05 \sin x_k^{(2)}, 0.1 \cos x_k^{(1)}]^T & \text{if } u_k = u_1 \\ [-0.25 + 0.05 \sin x_k^{(2)}, 0.1 \cos x_k^{(1)}]^T & \text{if } u_k = u_2 \\ [0.1 \cos x_k^{(2)}, 0.25 + 0.05 \sin x_k^{(1)}]^T & \text{if } u_k = u_3 \\ [0.1 \cos x_k^{(2)}, -0.25 + 0.05 \sin x_k^{(1)}]^T & \text{if } u_k = u_4 \end{cases}$$

Unknown *a priori*

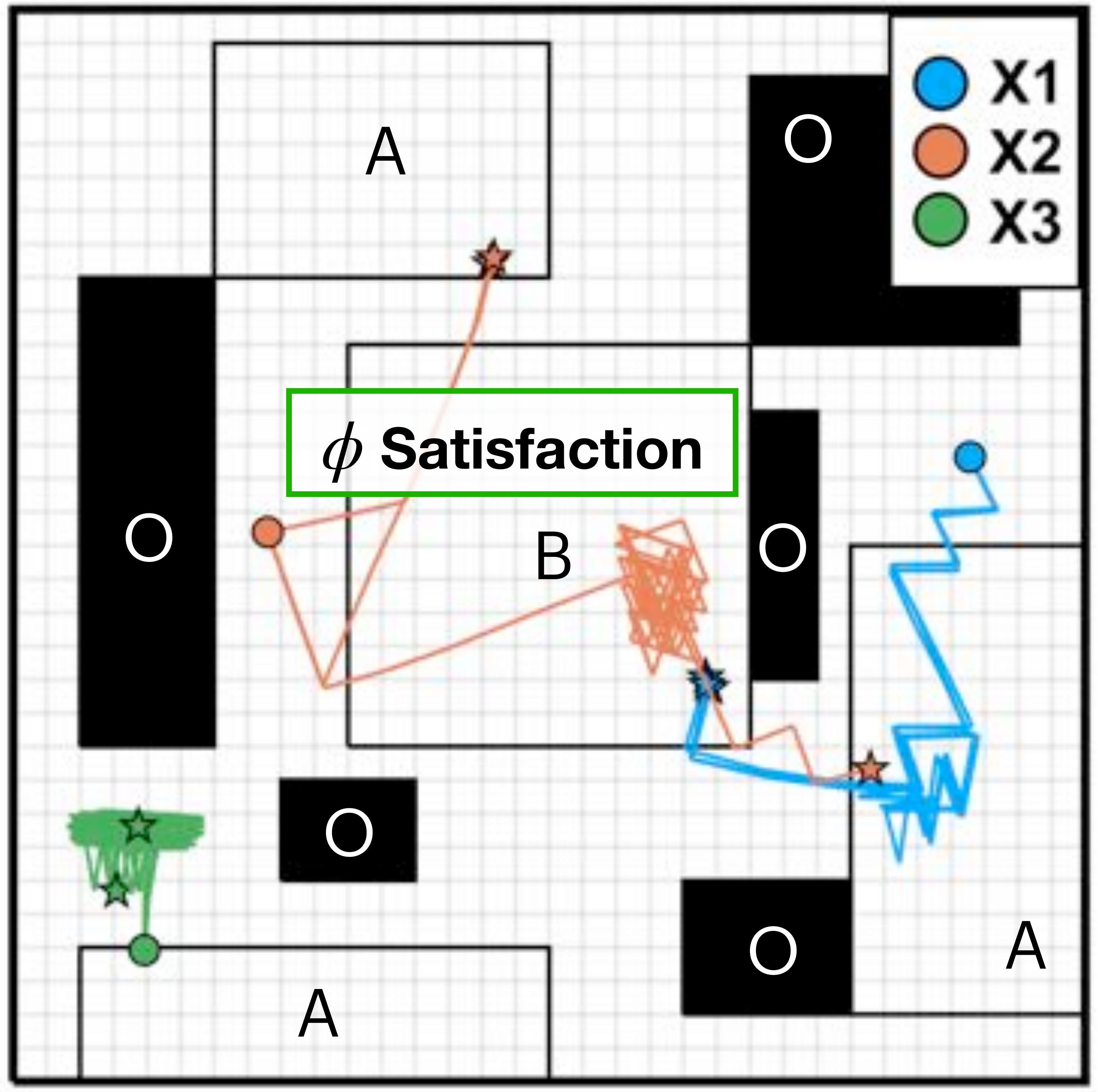
$\phi = \mathcal{F} A \wedge \mathcal{F} B \wedge \mathcal{G} \neg O$ (LTLf formula)

$=$ go to A and B in any order, always avoid O

Offline Abstraction and Synthesis with 200 datapoints

$N = 75$ Data Points for Local GPs

✓ Satisfies ϕ ✗ Violates ϕ ? Indeterminate



No Guarantees on Satisfaction!

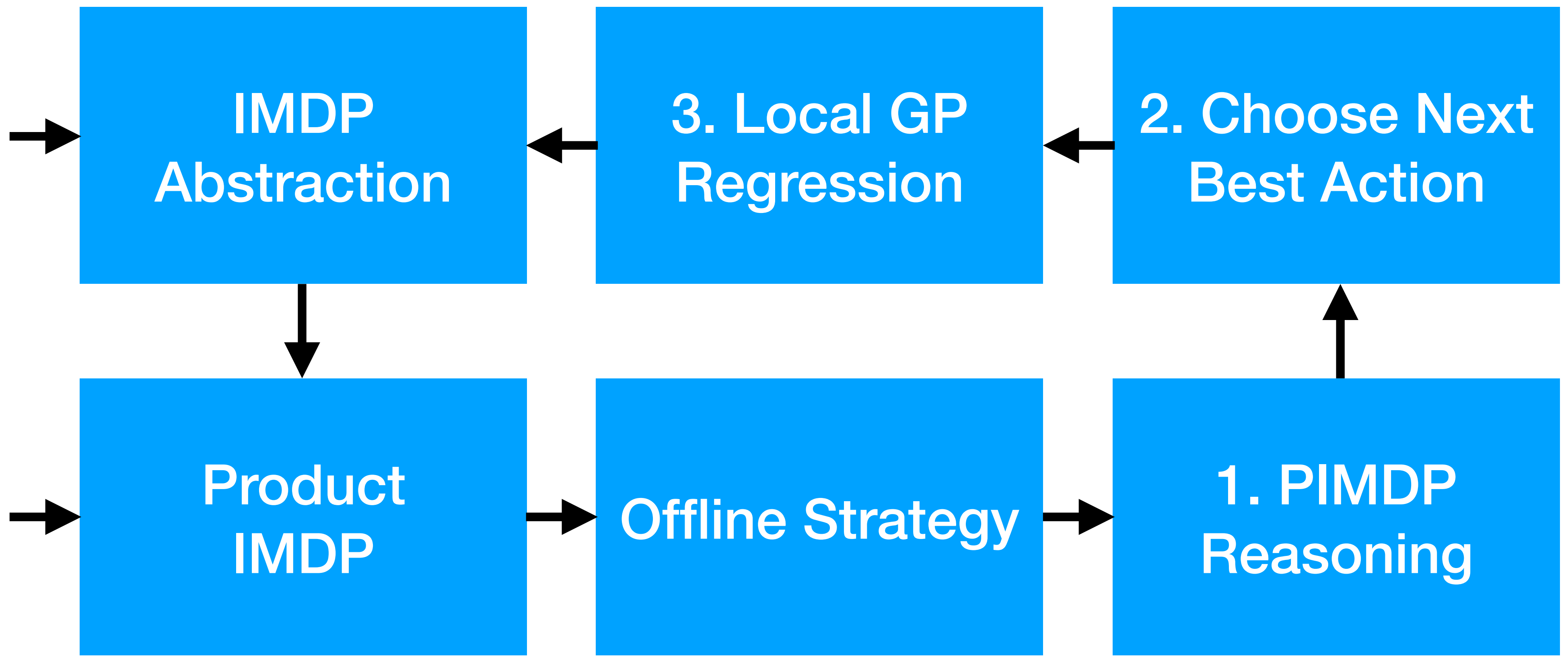
Sink + Progression Heuristic Simulations at X2

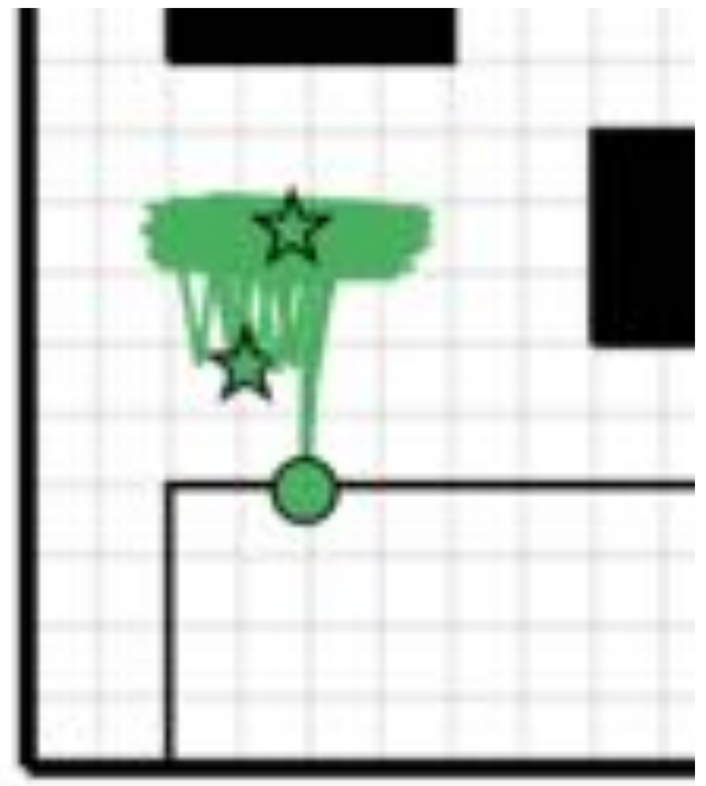
500 sims @ X2	Offline	All Data (Static)	Local (Static)	Local (Updates)
✓ % Satisfy	0%	76%	80.8%	**99.6%**
✗ % Violate	100%	0%	0%	**0%**
? % Indet.	0%	24%	19.2%	**0.4%**
IMDP Updates	-	-	-	**6258**

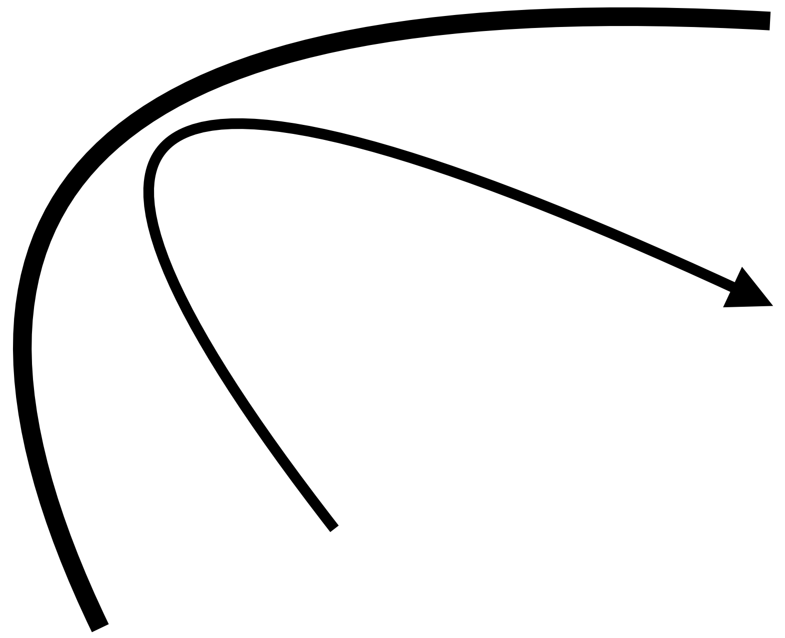
Sink + Progression Heuristic Simulations at X3

500 sims @ X3	Offline	All Data (Static)	Local (Static)	Local (Updates)
✓ % Satisfy	65.2%	0%	0%	**86.4%**
✗ % Violate	34.8%	9.8%	7.4%	**10.2%**
? % Indet.	0%	90.2%	92.6%	**3.4%**
IMDP Updates	-	-	-	**1814**

Conclusion & Future Work







- Online extension of GP, IMDP-based synthesis
- Extending theoretical guarantees
- Augmentation with abstraction-free approaches
- Identifying and testing autonomous systems applications

Thank you!

john.m.jackson@colorado.edu