Institute for
Data
Driven
Dynamical
Design
ID$_4$

# Probabilistic Prediction of Material Stability: Integrating Convex Hulls into Active Learning

Andrew Novick*[1], Diana Cai*[2], Quan Nguyen[3], Vladan Stevanovic[1], Roman Garnett[3], Ryan Adams[2], Eric Toberer[1]
[1] Department of Physics, Colorado School of Mines
[2] Department of Computer Science, Princeton University
[3] Department of Computer Science and Engineering, Washington University in St. Louis

**Stability prediction is accelerated by treating the convex hull as a probabilistic object, allowing for an efficient active learning process that minimizes the number of thermodynamic calculations necessary to define the convex hull.**

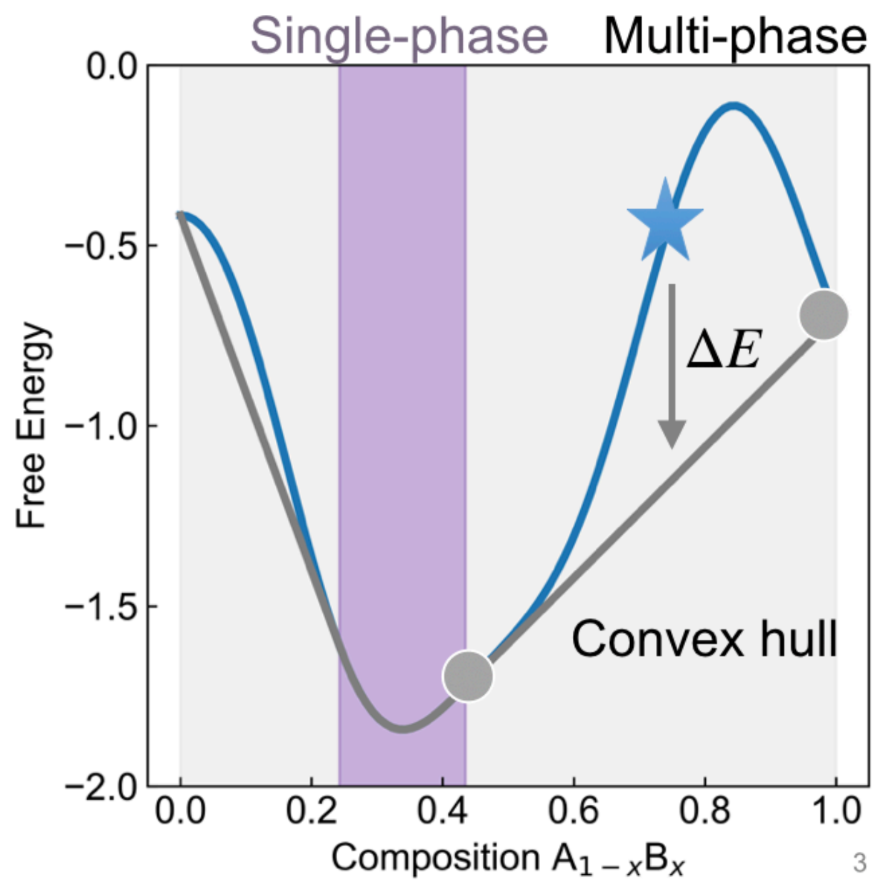
## Motivation

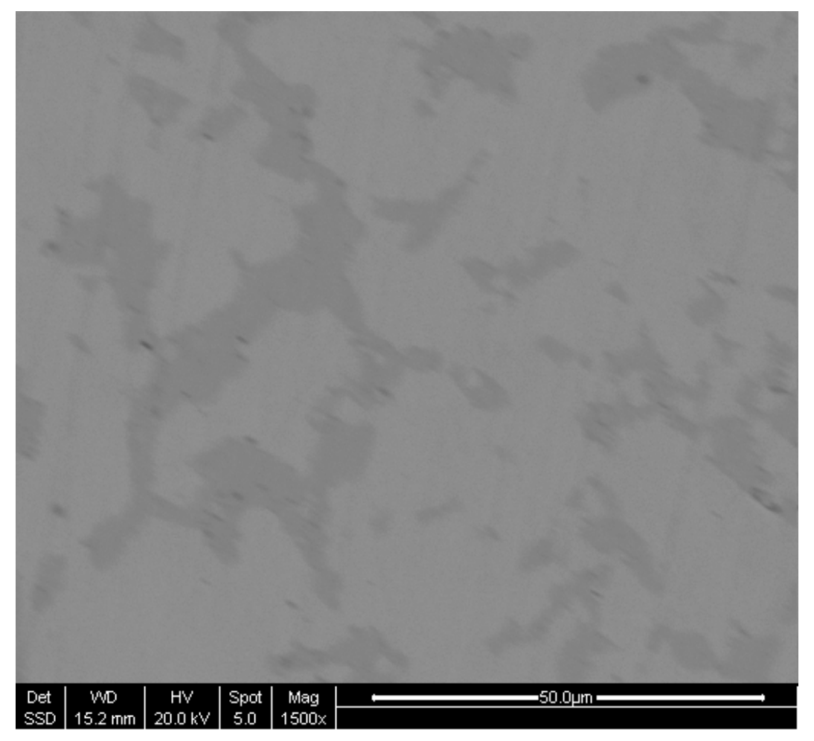
- Understanding thermodynamic stability is foundational to chemical and materials design.
- High-fidelity surrogate models not always available for complex materials (e.g., high-entropy alloys, highly correlated materials, liquids)

## Research Goal

- Accelerate stability predictions by treating convex hull as probabilistic object
- Minimize the number of thermodynamic calculation necessary to define convex hull

## Phase Diagram Construction from Convex Hulls

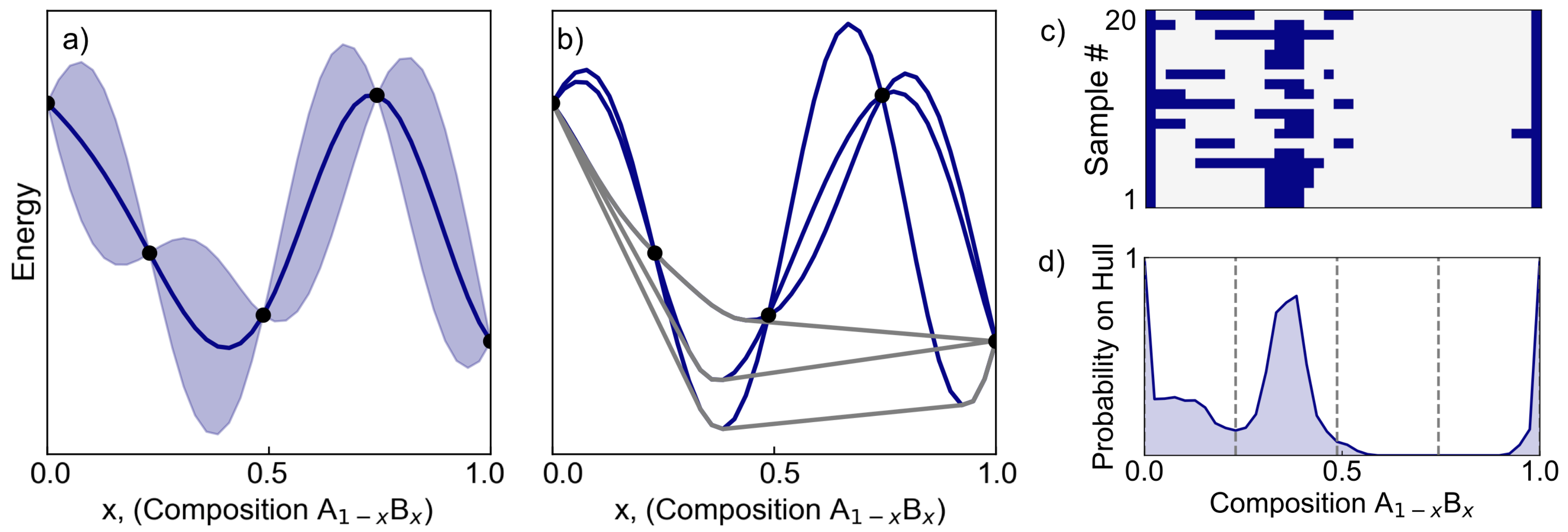




- Materials have option to separate into multiple phases to lower energy
- Competition between being a single phase or multiple phases—captured with convex hull
- Stability prediction requires information from all other competing compositions

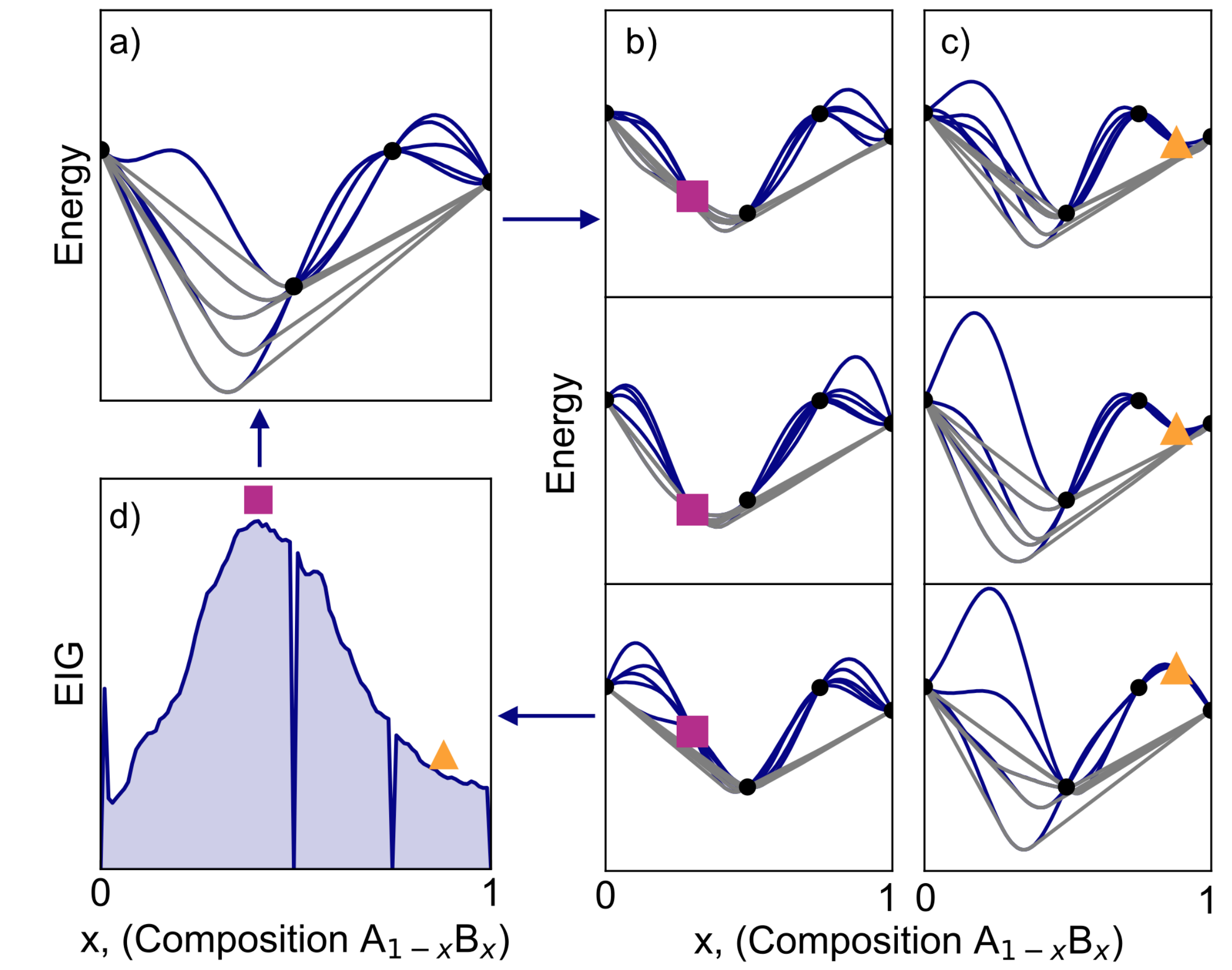
Light grey — $Er_{0.36}Ni_{0.3}Sb_{0.34}$
Dark grey — $Er_{0.2}Ni_{0.4}Sb_{0.4}$
(Maria Wrobleska, Toberer group)

## Probabilistic Convex Hull



(a) Energy surface modeled using Gaussian Process regression
(b) Three samples from GP posterior (blue). Convex hulls for each surface (grey).
(c) Binary classification of compositions being on or off hull (blue and grey).
(d) Probability of being on the hull—derived from summing across hull samples.

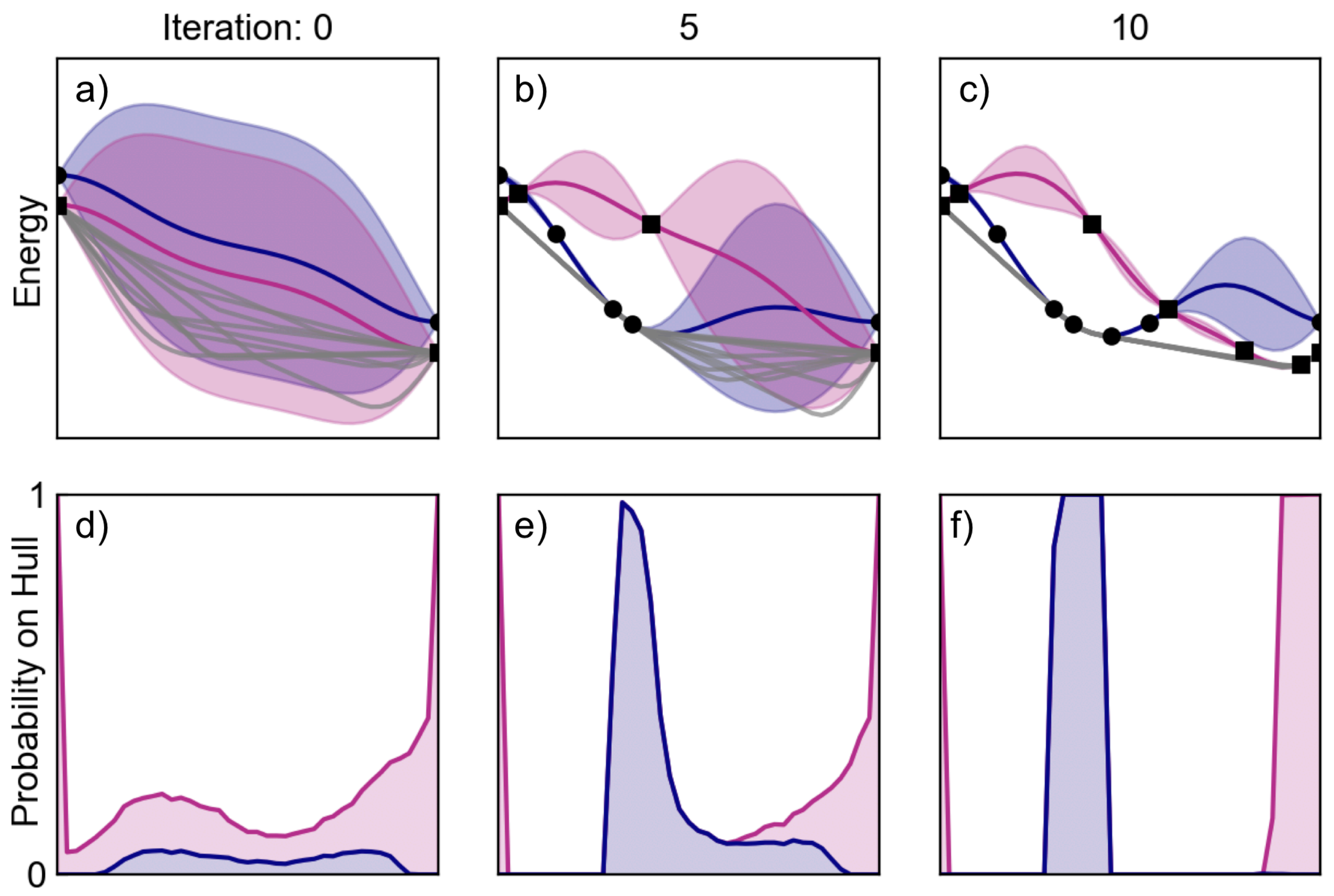
## Active Learning of Convex Hull



(a) Samples from posterior distribution of GP
(b) Hypothetical energies and convex hulls from observation at red square
(c) Hypothetical energies and convex hulls from observation at red square
(d) Expected information gain calculated across compositions.
Smallest entropy in convex hull distribution resulting from observation at red square

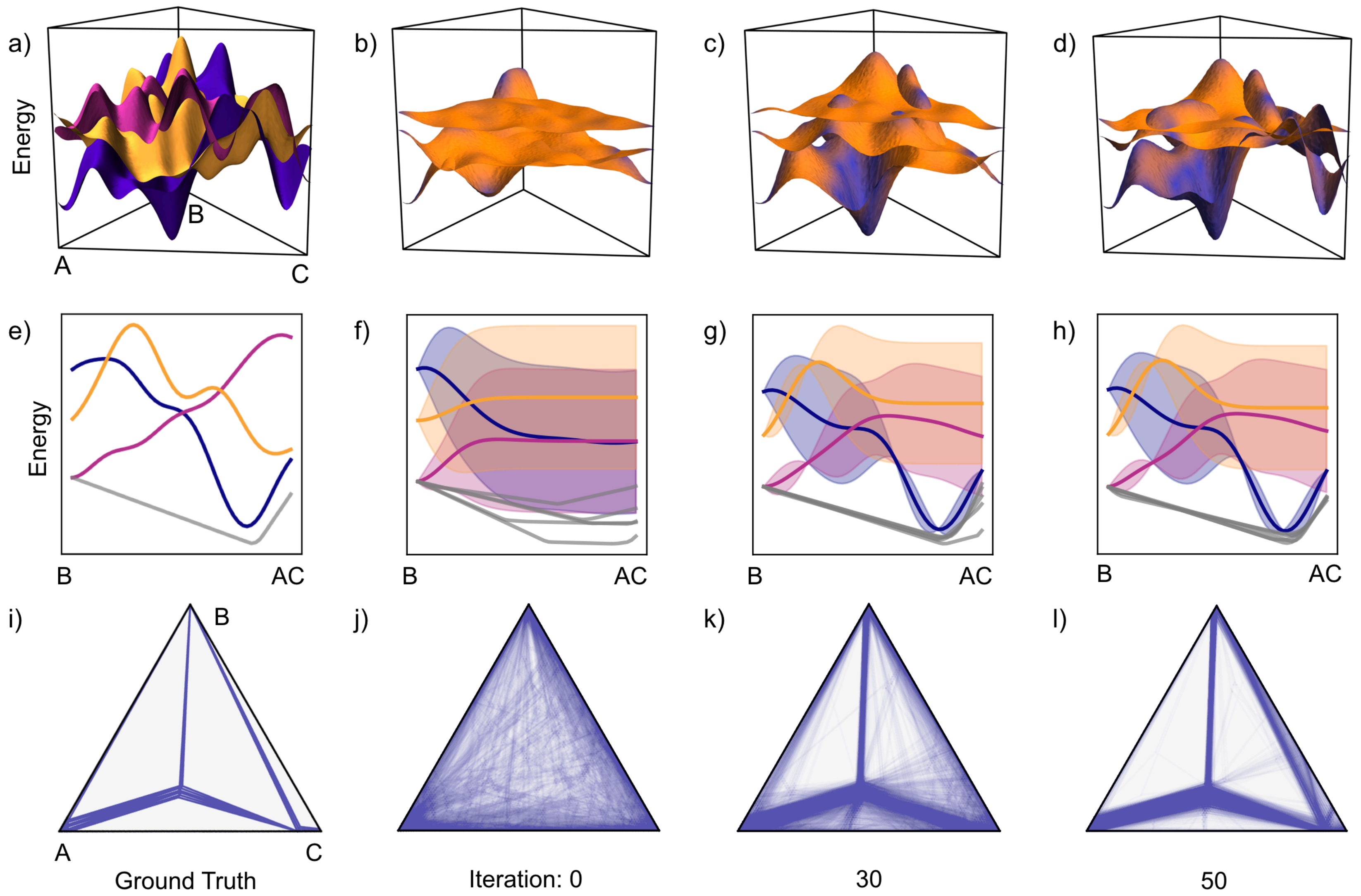
$$\mathrm{EIG}(x;\mathscr{D}) := \mathbb{S}[H_{\mathscr{D}}] - \mathbb{E}_y[\mathbb{S}[H_{\mathscr{D}\cup(x,y)}]]$$

## 1D Evolution for Two Phases

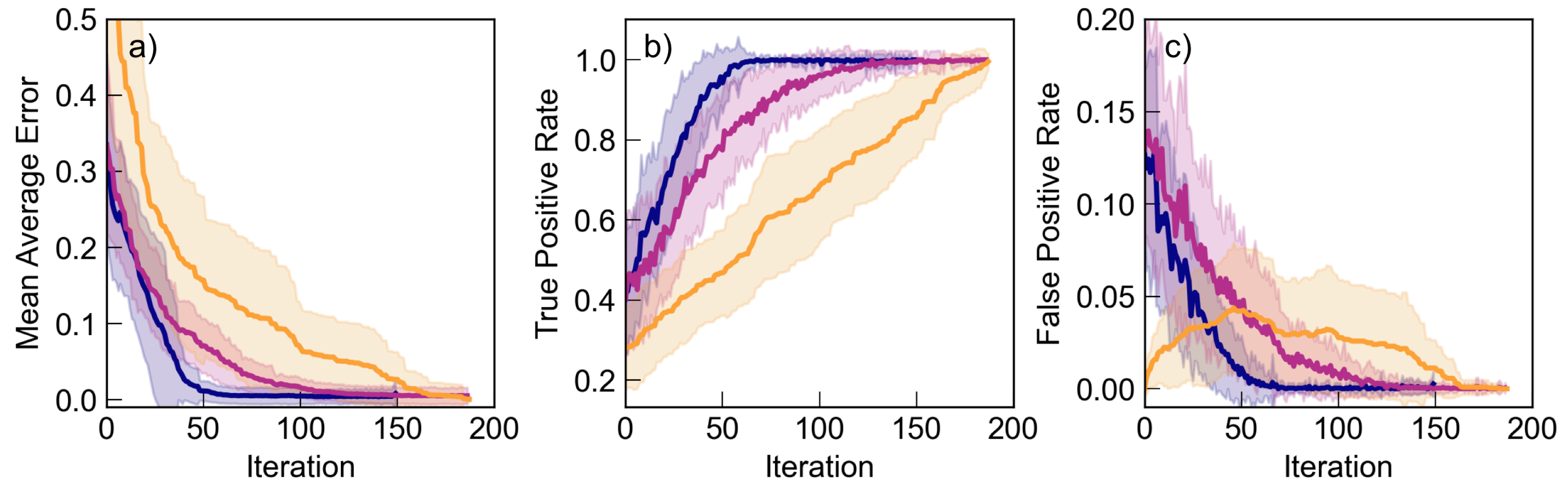


- Two different phases are represented with independent GPs (blue and pink)
- Significant uncertainty remains for composition-phase pairs irrelevant to hull

## 2D Evolution for Three Phases



- (a) Energy surface for three different phases in 2D (ternary) space
- (b-d) Mean estimate of various energy surfaces, colored by uncertainty {high:orange, low:purple}
- (f-h) GP posterior for B-AC cross section
- (j-l) 100 different convex hull samples superimposed on ternary phase diagram. Distribution of convex hulls tightens with increasing observation



- Active learning for convex hulls (blue), active learning of energy surfaces (pink), farthest point sampling (orange)
- 40 different sets of synthetic energy surfaces were used
- Convex-hull aware active learning outperforms other two policies

Funding: NSF OAC # 2118201