IEEE Cluster 2016
The Institute of Electrical and Electronics Engineers Cluster 2016
2016.09.13 – 15
Palais de Chine Hotel / Taipei, Taiwan

Active Learning in Performance Analysis

Dmitry Duplyakin

Jed Brown

Robert Ricci

09/14/2016

dmitry.duplyakin@colorado.edu

Outline

- Motivation
- Approach
- Implementation
- Datasets and Visualizations
- Evaluation
- Summary and Future Work

Motivation

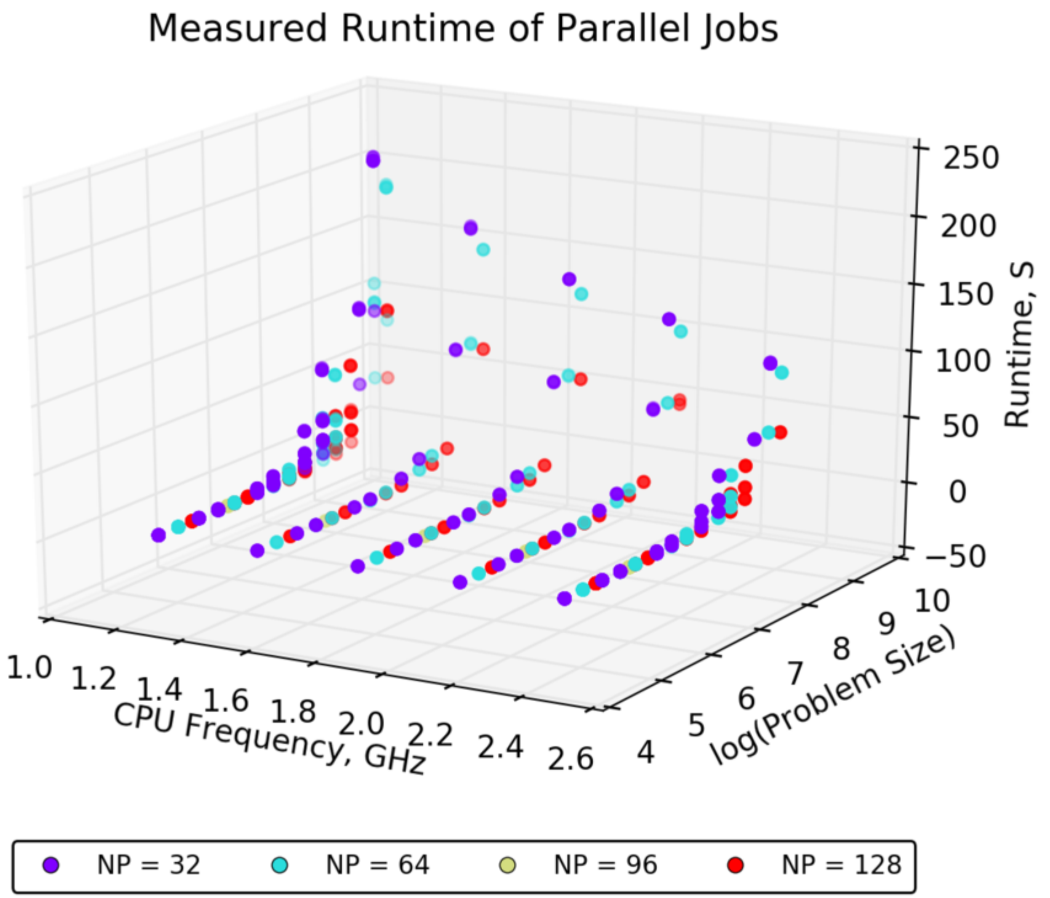
Performance Analysis:

- Take a set of measurements
- Build a model
- Understand behavior of a complex system
- Predict outcomes of future experiments

Main Challenges:

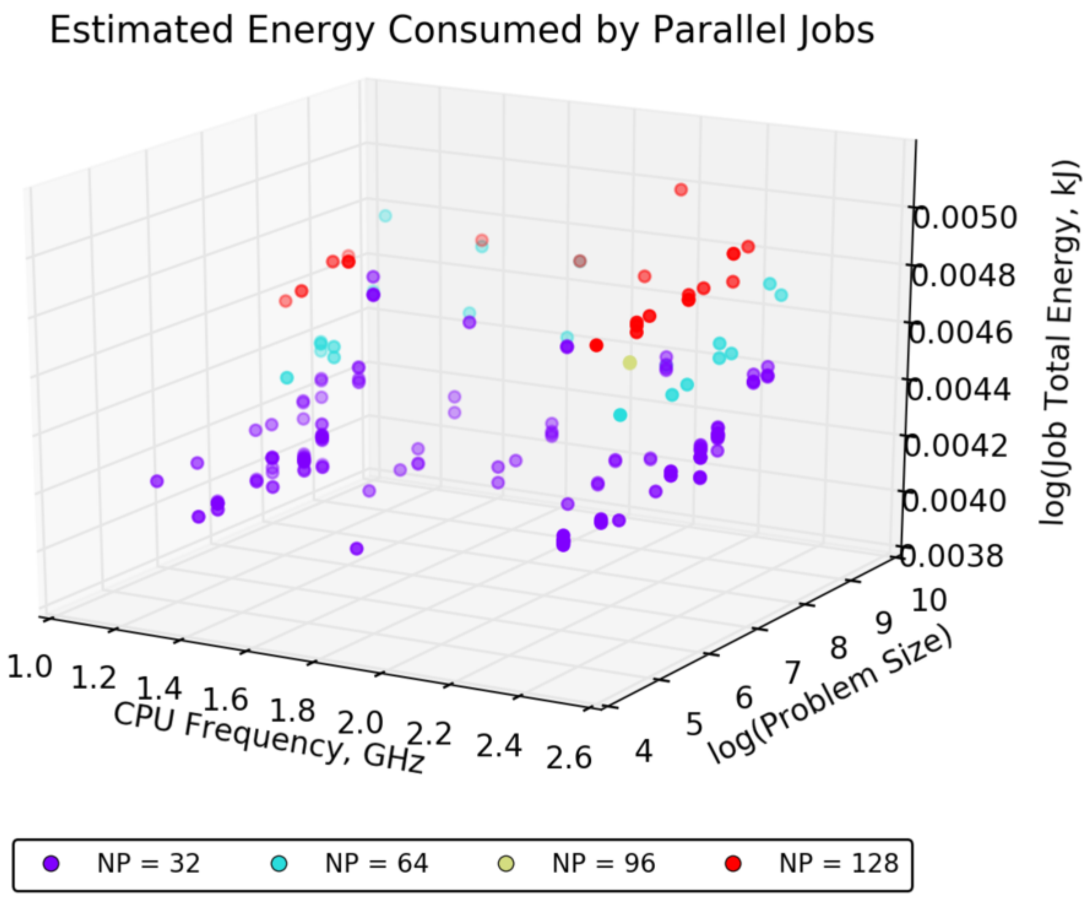
- Often too many factors
- Inability to take equal number of measurements at every configuration
- Inefficient exploration of input space

Motivation: Example 1



Each point represents a run of HPGMG-FE benchmark on a 4-node cluster provisioned on CloudLab testbed

Motivation: Example 2



Each point represents a run of HPGMG-FE benchmark on a 4-node cluster provisioned on CloudLab testbed

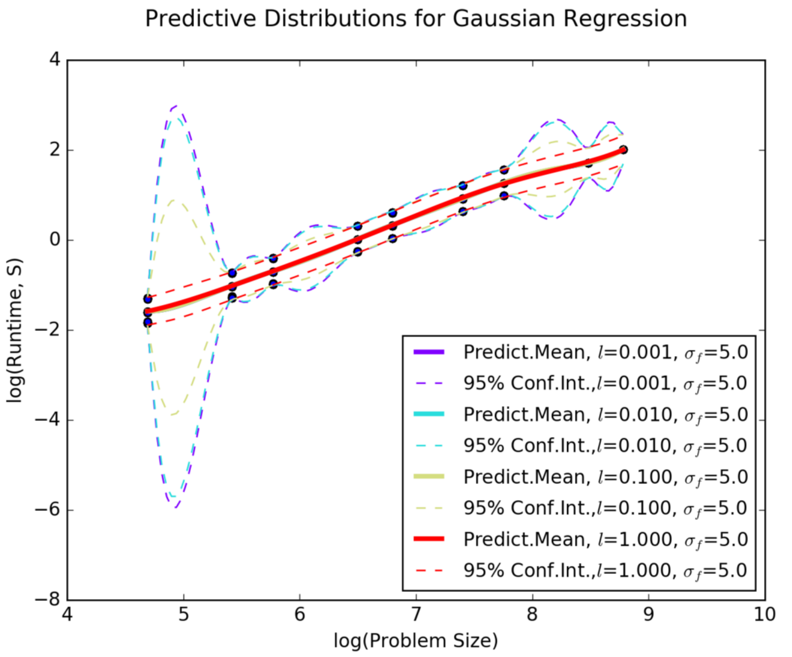
Approach: Active Learning

- Use *Active Learning* (AL) -- techniques from Machine Learning where "learner" interacts with "data source"
  1. Train a model on a small set of measurements
  2. Let the model suggest a point for the next experiment
  3. Run the suggested experiment
  4. Retrain the model with the new measurement
  5. Go back to 2 or exit

- Sometimes called: *adaptive experiment design* and *optimal experiment design*

Approach: Gaussian Process Regression

- Use *Gaussian Process Regression* (GPR) -- non-parametric non-linear interpolation technique that provides best linear unbiased prediction (under suitable assumptions)
  - Build a model for $f(x)$
  - For every new x^*, calculate estimates of $E[f(x^*)]$ and $\sigma[f(x^*)]$



- Sometimes called: *kriging* (in geostatistics) and *Wiener–Kolmogorov prediction*
- GPR works in many dimensions

Approach: Putting it Together

- Combine AL and GPR into a 2-layer system:

Upper: **AL**
Lower: **GPR**

- Optimization problem at each layer:

Upper: **Choose "best" experiment**
Lower: **Choose "best" hyperparameters**

Approach: Details

Upper: **Choose "best" experiment**

Lower: **Choose "best" hyperparameters**

Consider strategies:

Variance Reduction (VR): $x^* = \arg\max_x \left(\sigma[f(x)]\right)$

Cost Efficiency (CE): $x^* = \arg\max_x \left(\frac{\sigma[f(x)]}{f(x)}\right)$

Use: **Bayesian Model Selection**

(Marginal Likelihood Maximization)

with 3 hyperparameters:

noise level, *length scale*, and *amplitude*

Implementation

- Developed a prototype in Python which supports:
  - single realizations of AL in "offline" mode*
  - batches of realizations for comparison of *Variance Reduction* and *Cost Efficiency* strategies
- GPR: used code for Gaussian Processes in *scikit-learn* (0.18.dev0)

* Note: **offline** refers to the fact that the prototype queries a database with collected data. Future work: in **online** mode, run AL alongside the computation

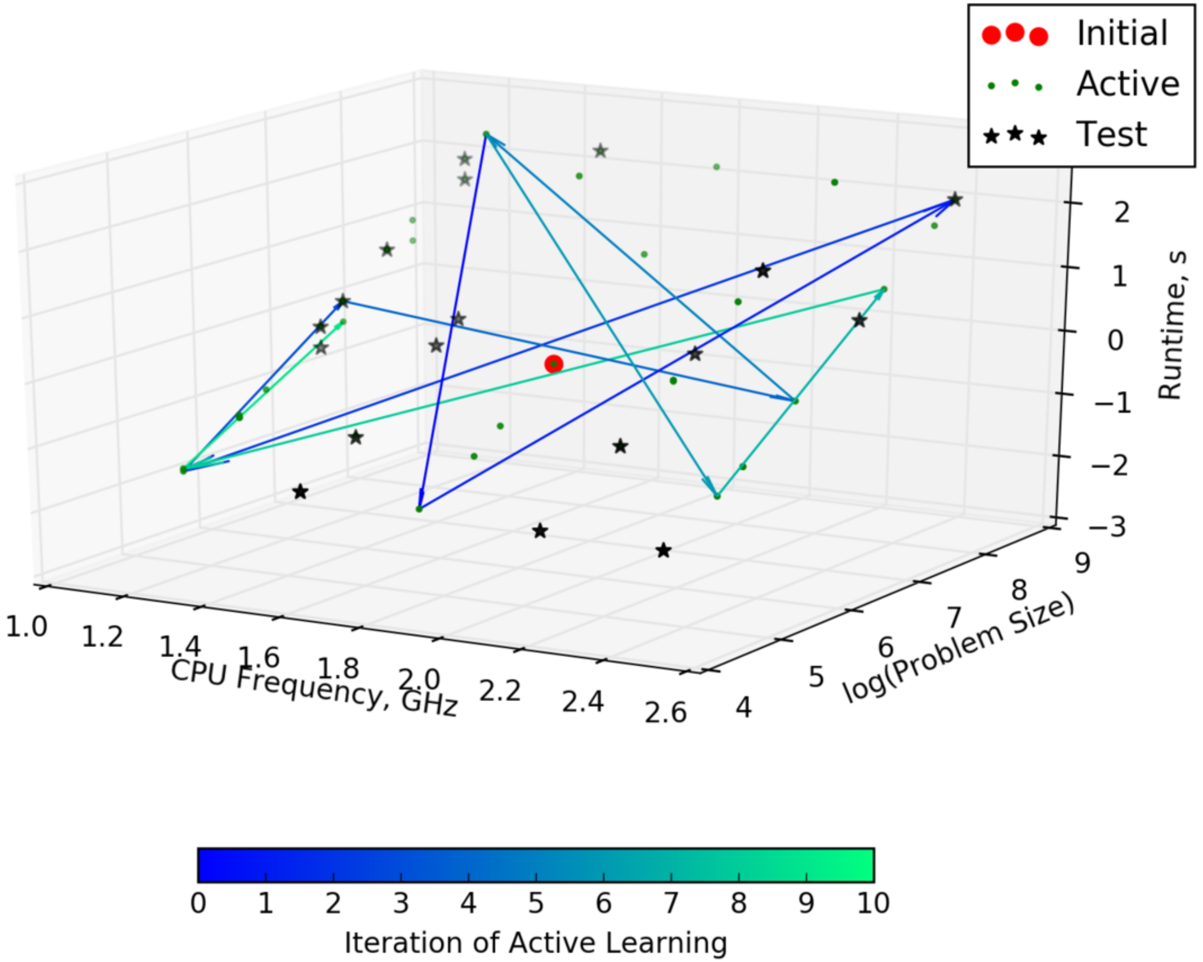
Analyzed Datasets

- Measured runtimes and estimated energy consumption for a large set of HPGMG-FE benchmark runs on a cluster provisioned on the CloudLab testbed
- Organized this data into two datasets:

	Dataset: **Performance**	Dataset: **Power**
# Jobs	3246	640
Responses	Runtime (S)	Runtime (S), Energy (J)
Runtime, S	0.005 – 458.436	0.005 – 458.436
Energy, J	–	6.4e3 – 1.1e5
Variables	Operator: poisson1,poisson2,poisson2affine Global Problem Size: 1.7e3 – 1.1e9 NP: 1,2,4,8,16,24,32,48,64,96,128 CPU Frequency (GHz): 1.2,1.5,1.8,2.1,2.4	

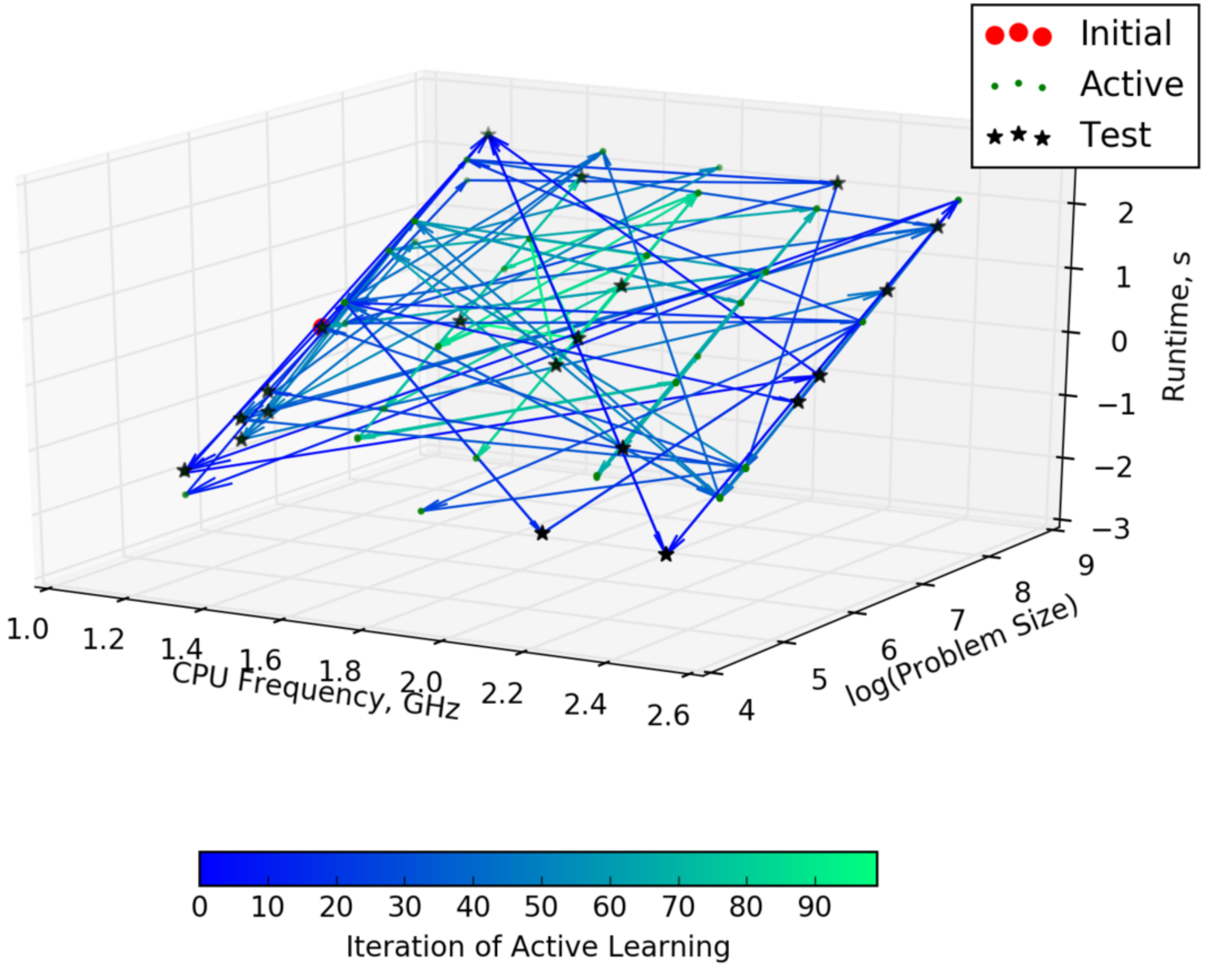
- 3d visualizations are available here

Active Learning: 10 Iterations



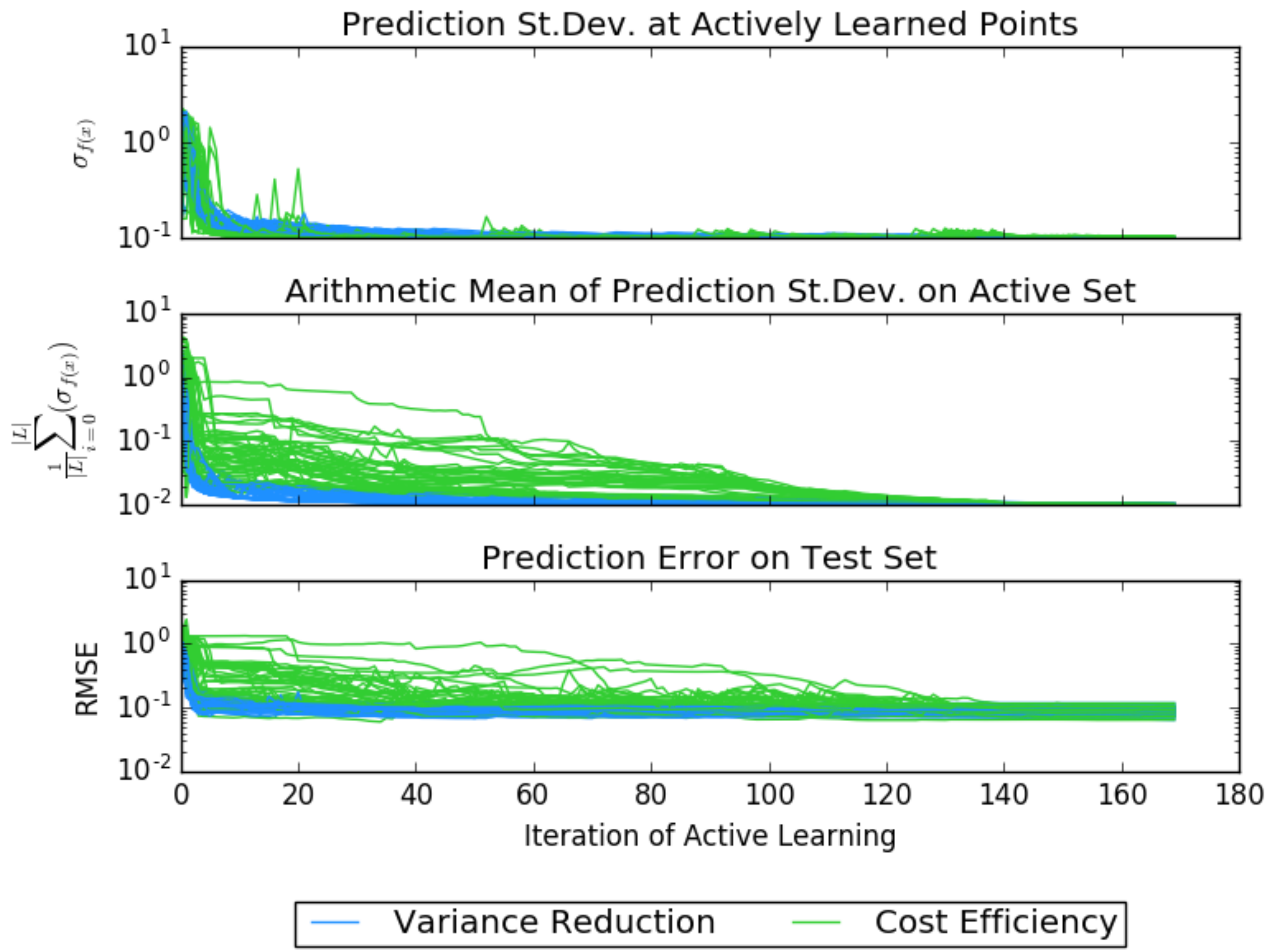
Shown points represent a subset of measurements in the **Performance** dataset; runtimes are log-transformed

Active Learning: 100 Iterations

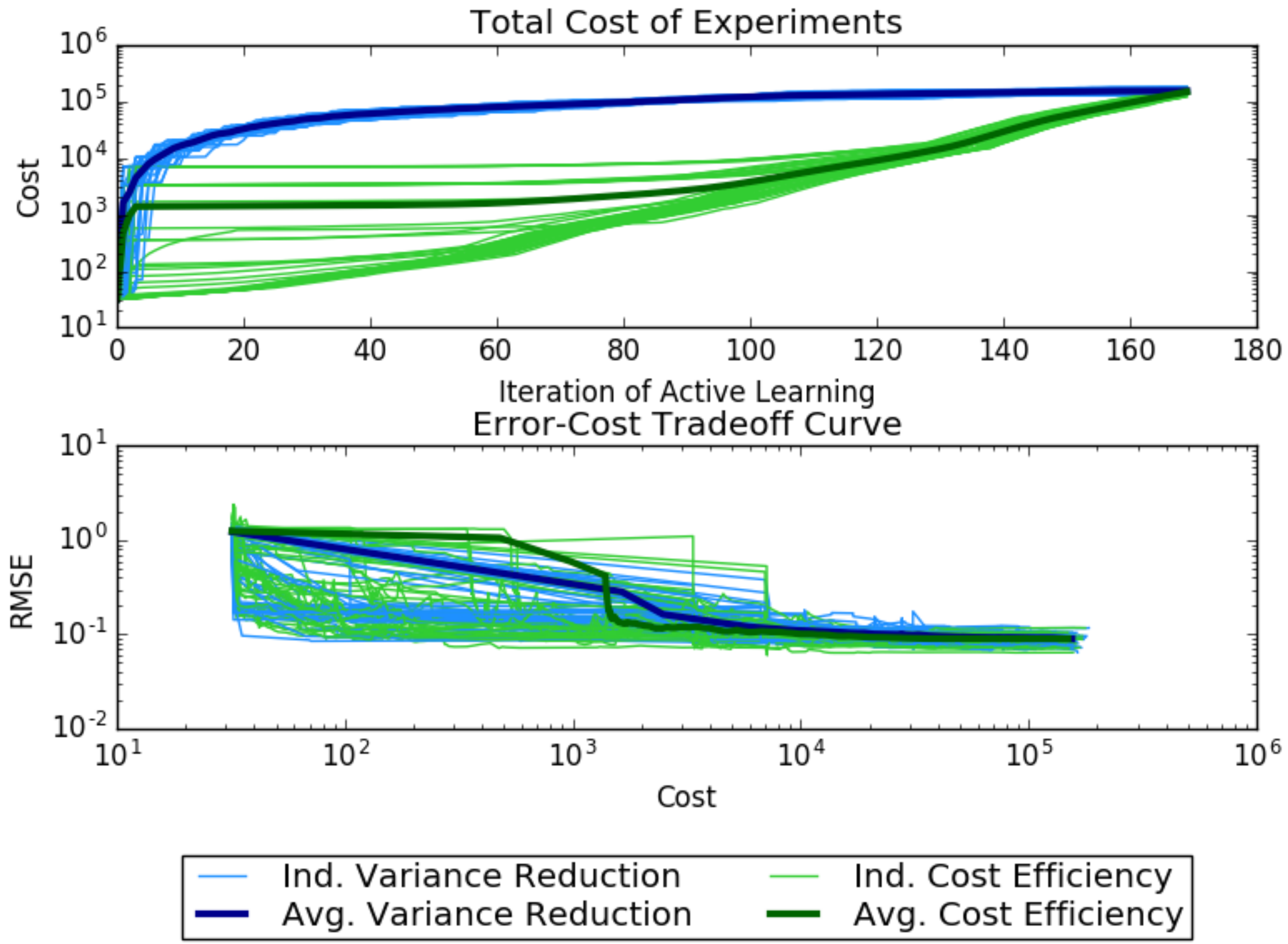


Shown points represent a subset of measurements in the **Performance** dataset; runtimes are log-transformed

Evaluation: Convergence Analysis



Evaluation: Cost Analysis



Summary and Future Work

Summary:

- Proposed using Active Learning + Gaussian Process Regression for efficient regression learning in performance analysis
- Demonstrated tradeoffs between two Active Learning algorithms, with and without adjustment for experiment cost

Future Work:

- Investigate computational requirements
- Leverage continuous optimization techniques
- Run Active Learning in the online mode

Thank you!

Questions?

dmitry.duplyakin@colorado.edu