

はじめに ベイズ最適化の紹介

津田 宏治

University of Tokyo / NIMS / RIKEN AIP

Discovery of new functional molecules and materials is of national importance

the WHITE HOUSE PRESIDENT BARACK OBAMA ★★★★
THE WHITE HOUSE WASHINGTON ★★★★

[Get Email Updates](#) | [Contact Us](#)

[BLOG](#) [PHOTOS & VIDEO](#) [BRIEFING ROOM](#) [ISSUES](#) the ADMINISTRATION the WHITE HOUSE our GOVERNMENT

[Home](#) • [About the Materials Genome Initiative](#)

Search WhiteHouse.gov [Search](#)

 Materials Genome Initiative

About | Goals | Examples | News & Announcements | Federal Programs | External Stakeholder Activities | Contact Us

To help businesses discover, develop, and deploy new materials twice as fast, we're launching what we call the Materials Genome Initiative. The invention of silicon circuits and lithium-ion batteries made computers and iPods and iPads possible -- but it took years to get those technologies from the drawing board to the marketplace. We can do it faster.

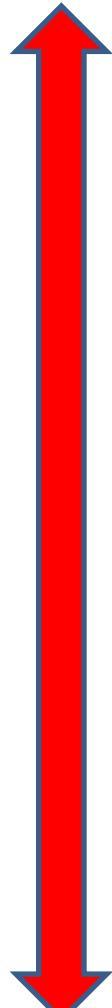
— President Obama, June 2011 at Carnegie Mellon University



A photograph showing President Barack Obama in a dark suit and tie, wearing safety goggles, standing next to a woman in a red sweater who is working with a long, glowing white tube or panel. Another man in a suit stands behind them. They appear to be in a laboratory or industrial setting with shelving in the background.

First Principles Calculations

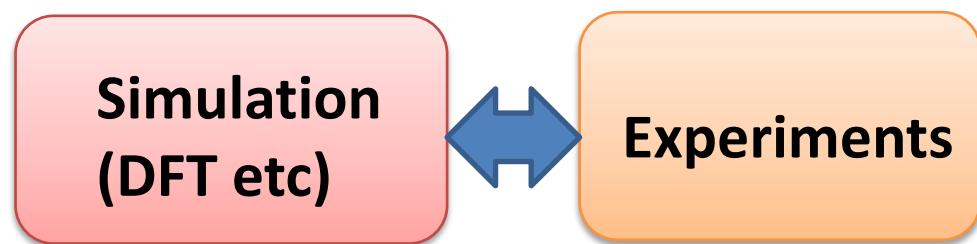
Accurate, Slow



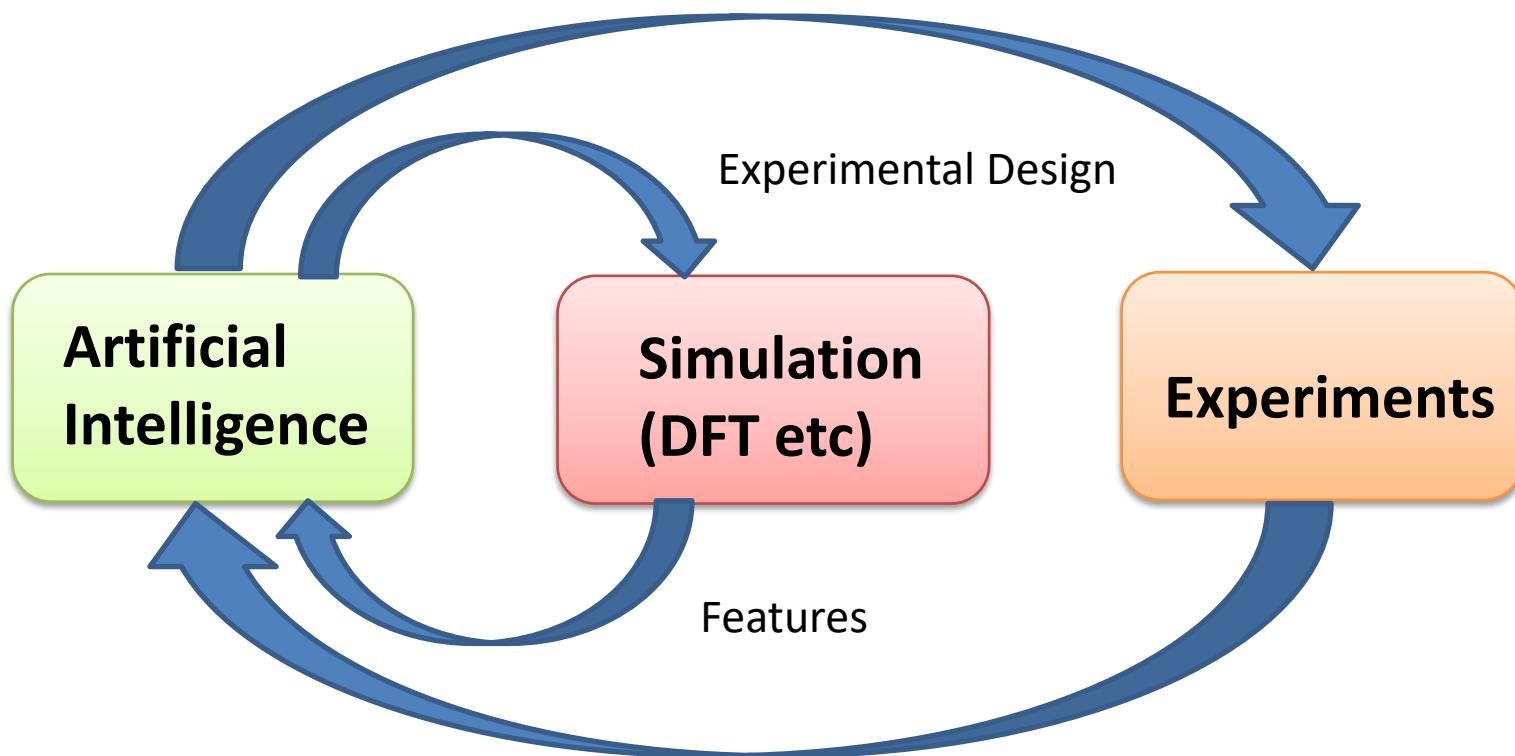
- Full configuration interaction
- Wave function based
- Density functional theory
- Semi-empirical
- Empirical potentials

Inaccurate, Fast

Old Picture



New Picture



Materials Design Examples

- Design of Si-Ge nanostructures (Phys Rev X 2017)
- De novo design of organic compounds with desired absorption wavelength

Screening by first principles calculations alone

Mat. 1	Mat. 2	Mat. 3	Mat. 4	Mat. 5	Mat. 6	Mat. 7	Mat. 8	Mat. 9	Mat. 10
-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	------------



First Principles Calc.



Score 1	Score 2	Score 3	Score 4	Score 5	Score 6	Score 7	Score 8	Score 9	Score 10
------------	------------	------------	------------	------------	------------	------------	------------	------------	-------------

Bayesian Optimization

(Jones et al., 1998)

- Find best data points with minimum number of observations
- Choose next point to observe to discover the best ones as early as possible

Bayesian Optimization (1)

Mat. 1	Mat. 2	Mat. 3	Mat. 4	Mat. 5	Mat. 6	Mat. 7	Mat. 8	Mat. 9	Mat. 10
-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	------------



First Principles Calc.



Score 1	Score 2	Score 3
------------	------------	------------

Bayesian Optimization (2)

Mat. 1	Mat. 2	Mat. 3	Mat. 4	Mat. 5	Mat. 6	Mat. 7	Mat. 8	Mat. 9	Mat. 10
-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	------------



First Principles Calc.



Predicted Scores

Score 1	Score 2	Score 3	Pred. Score 4	Pred. Score 5	Pred. Score 6	Pred. Score 7	Pred. Score 8	Pred. Score 9	Pred. Score 10
------------	------------	------------	---------------------	---------------------	---------------------	---------------------	---------------------	---------------------	----------------------

Var. 4	Var. 5	Var. 6	Var. 7	Var. 8	Var. 9	Var. 10
-----------	-----------	-----------	-----------	-----------	-----------	------------

Predicted Variances

Bayesian Optimization (3)

Mat. 1	Mat. 2	Mat. 3	Mat. 8	Mat. 4	Mat. 5	Mat. 6	Mat. 7	Mat. 9	Mat. 10
-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	------------



First Principles Calc.



Score	Score	Score	Score
1	2	3	8

Bayesian Optimization (4)

Mat. 1	Mat. 2	Mat. 3	Mat. 8	Mat. 4	Mat. 5	Mat. 6	Mat. 7	Mat. 9	Mat. 10
-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	------------

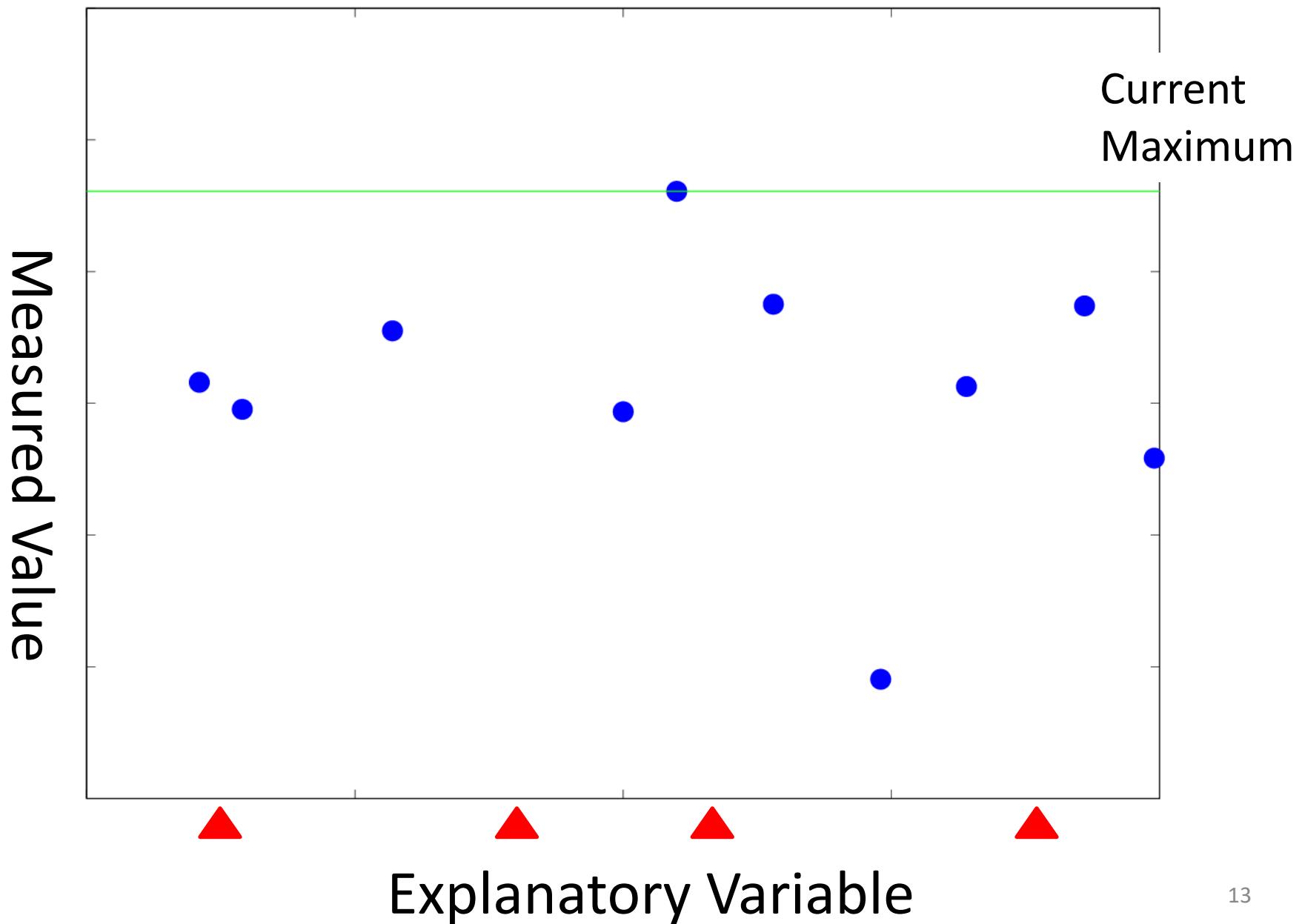


First Principles Calc.

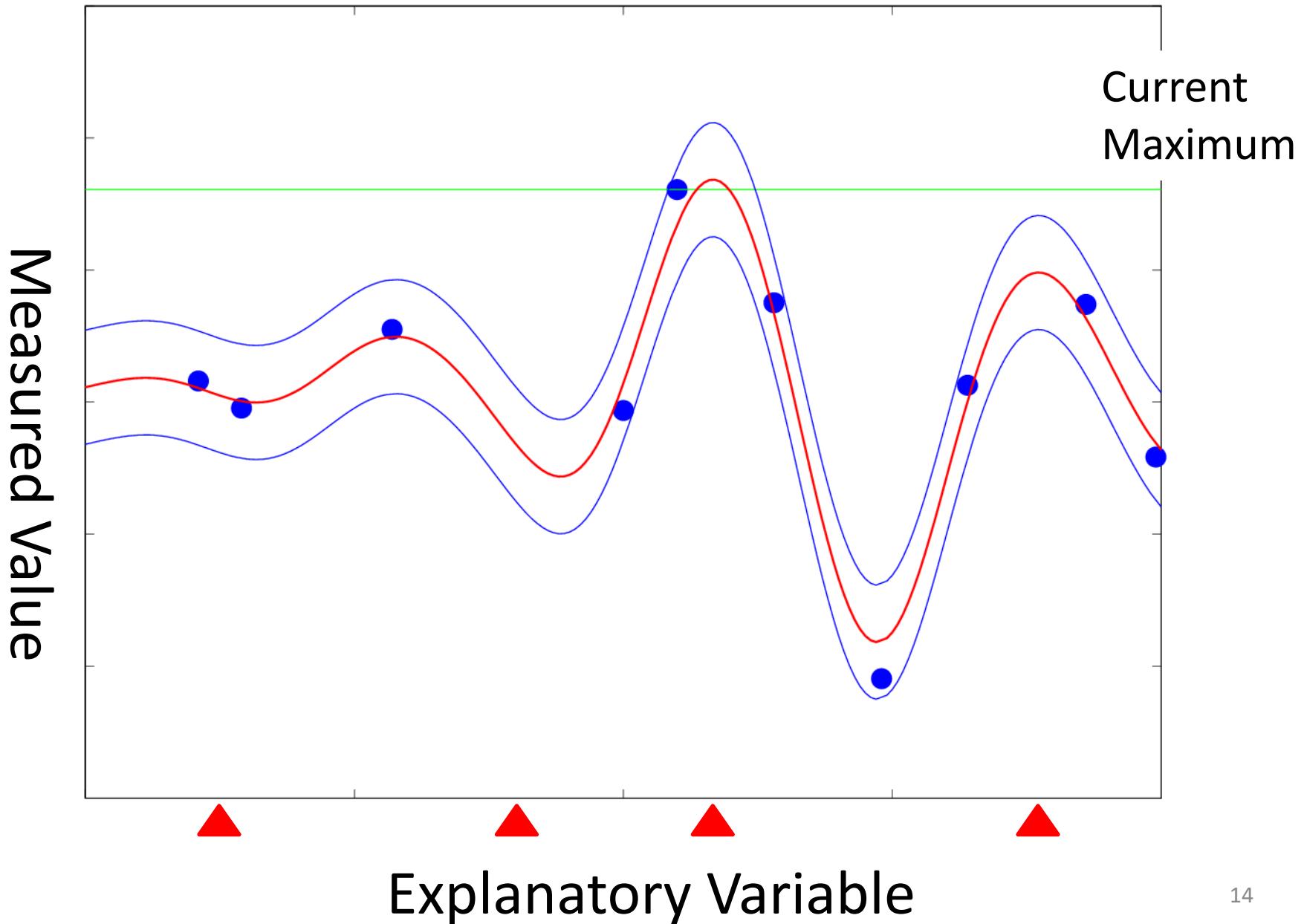


Score 1	Score 2	Score 3	Score 8	Pred. Score 4	Pred. Score 5	Pred. Score 6	Pred. Score 7	Pred. Score 9	Pred. Score 10
Var. 4	Var. 5	Var. 6	Var. 7	Var. 9	Var. 10				

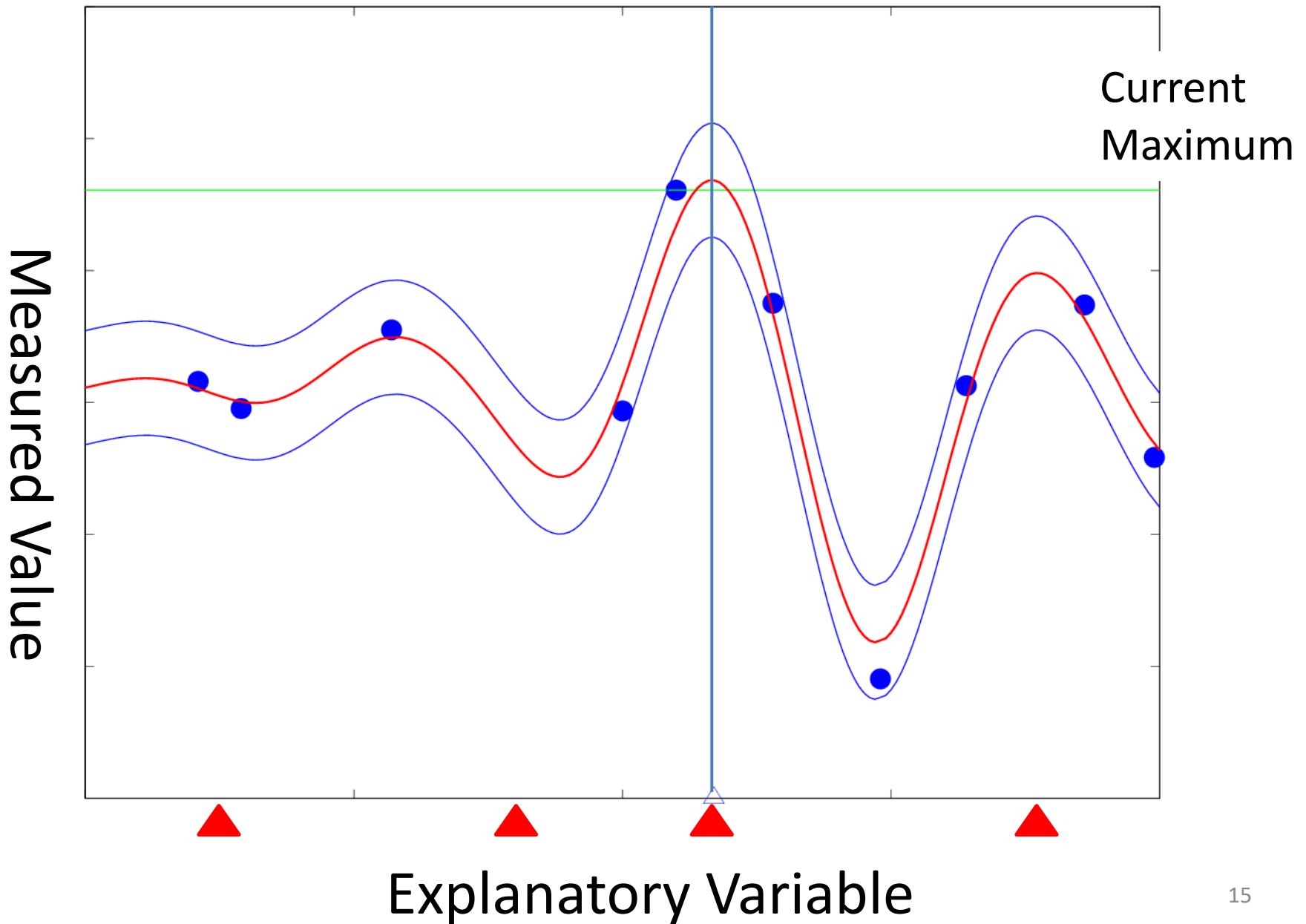
Where to observe next?



Gaussian Process



Maximum probability of improvement



Gaussian Process

Multivariate Gaussian Distribution

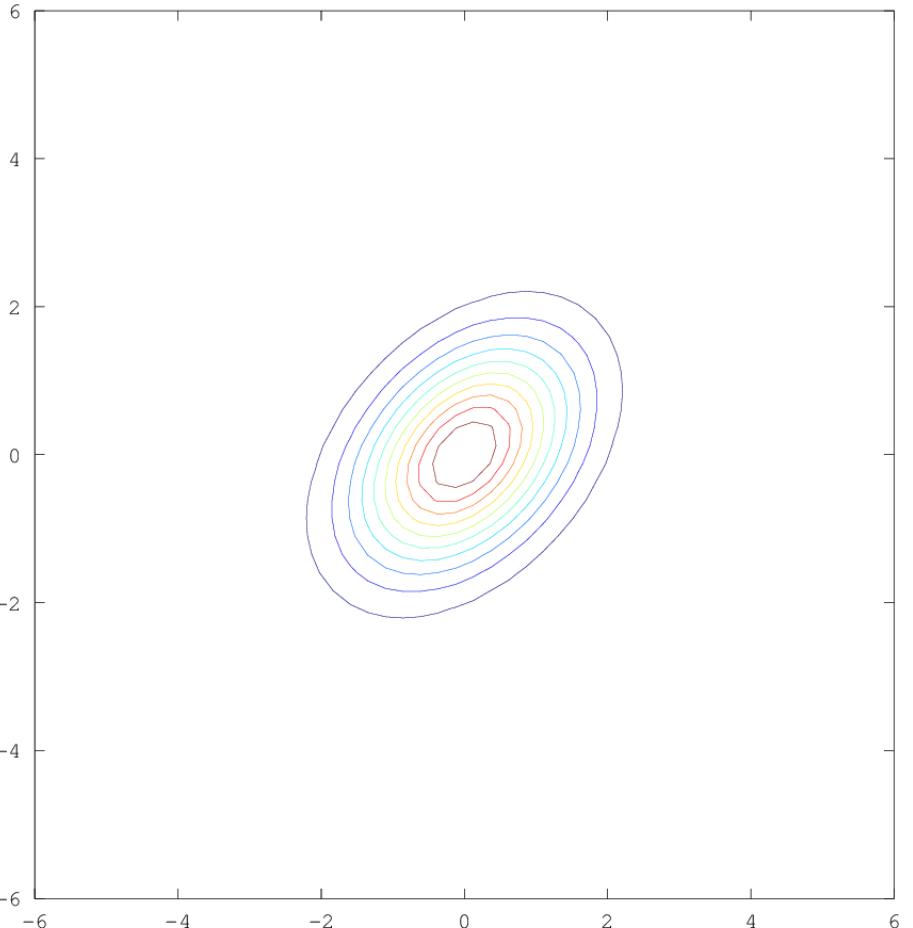
- Probability density function

$$p(\mathbf{x} \mid \mu, \Sigma) = \frac{1}{(2\pi)^{d/2} |\Sigma|^{1/2}} \exp\left(-\frac{1}{2} (\mathbf{x} - \mu)^\top \Sigma^{-1} (\mathbf{x} - \mu)\right)$$

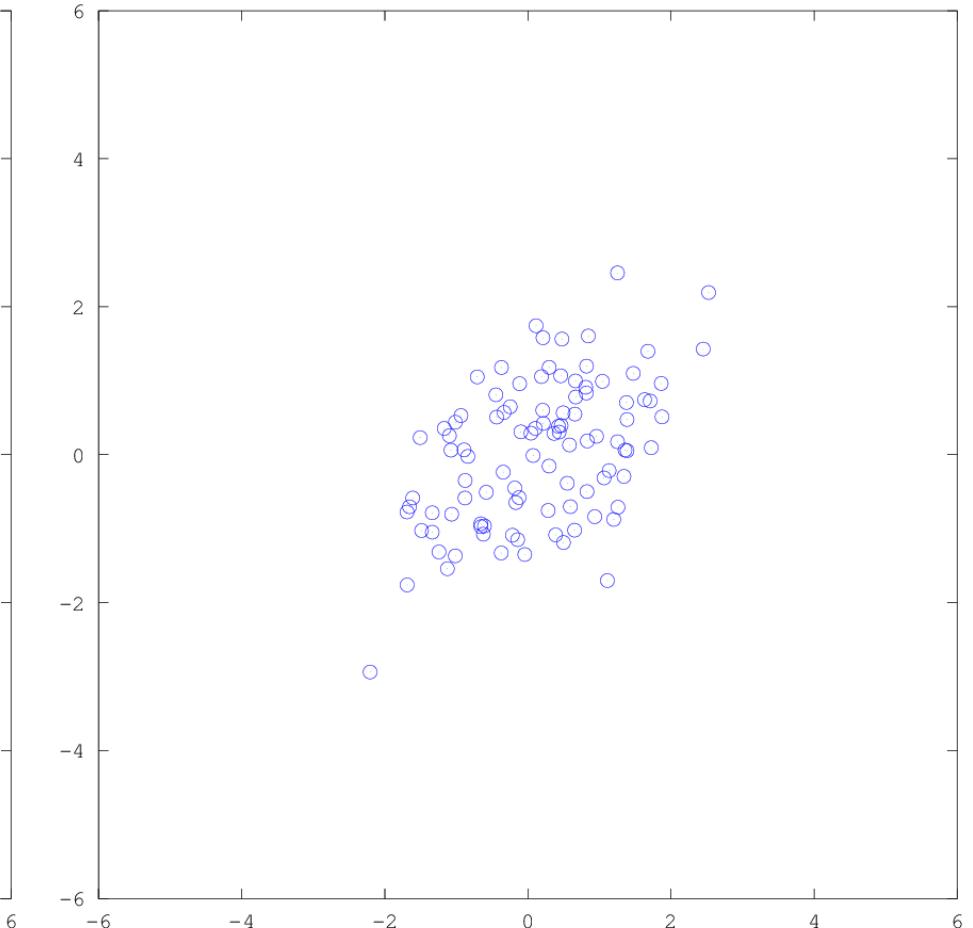
μ Mean

Σ Covariance Matrix

Probability Density



100 Samples



$$\mu = (0, 0)^T$$

$$\Sigma = \begin{pmatrix} 0.4 & 1 \\ 1 & 0.4 \end{pmatrix}$$

Conditional Distribution

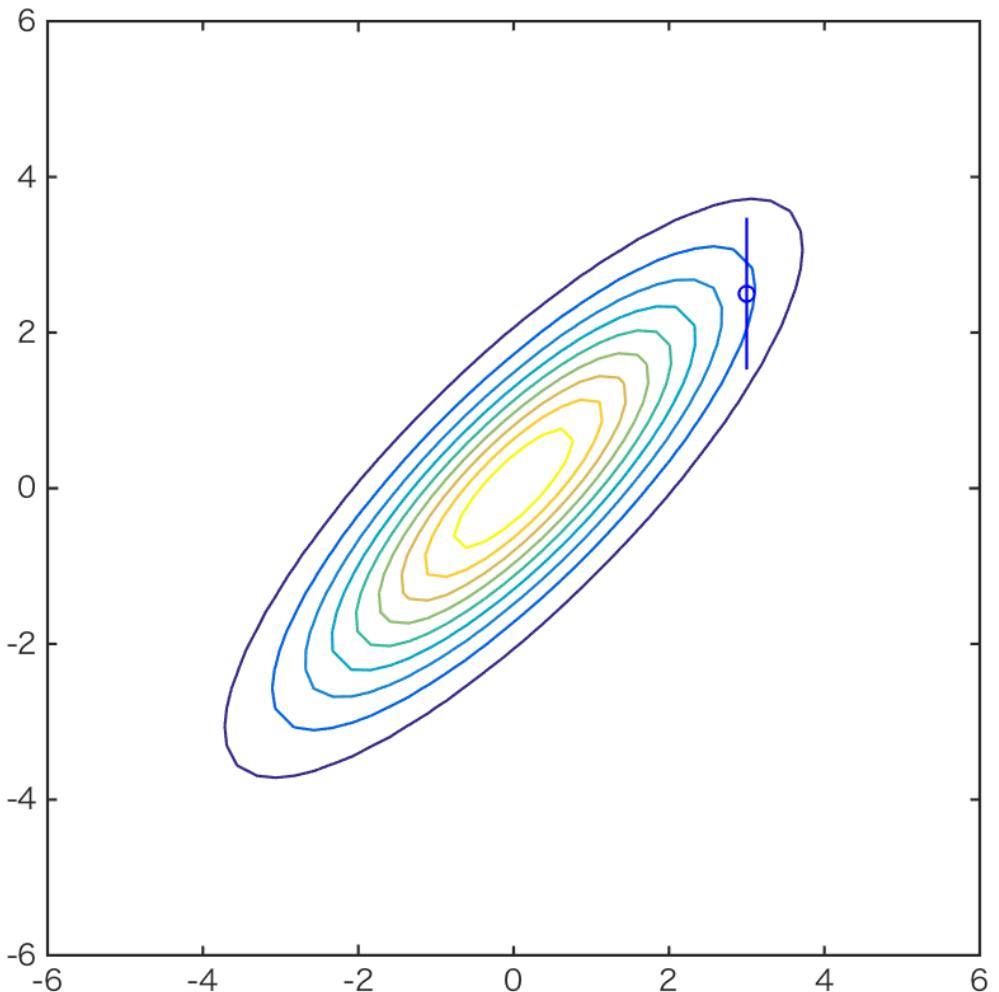
	Mean	Covariance
$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$	$\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}$	$\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$

$$P(x_1 \mid x_2 = a) = \mathcal{N}(\mu_c, \Sigma_c)$$

$$\mu_c = \mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(a - \mu_2)$$

$$\Sigma_c = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$$

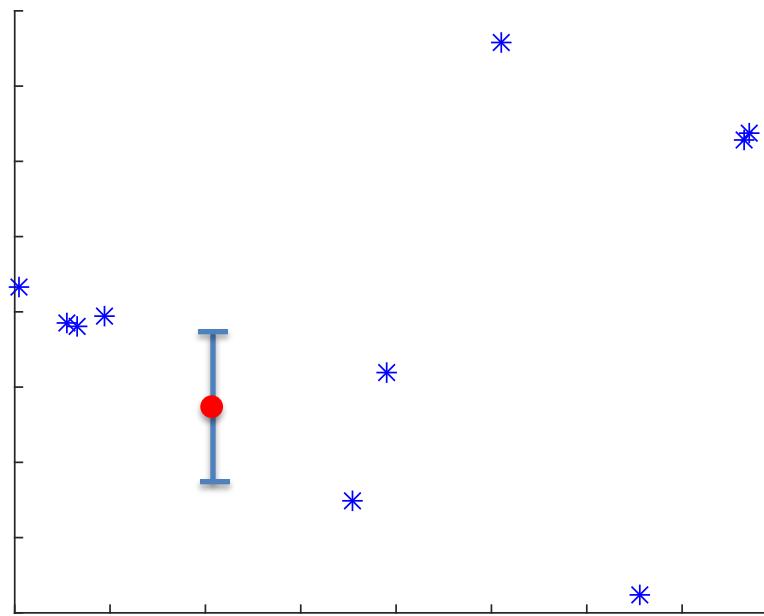
Conditional distribution at X=3



$$\Sigma = \begin{pmatrix} 3 & 2.5 \\ 2.5 & 3 \end{pmatrix}$$

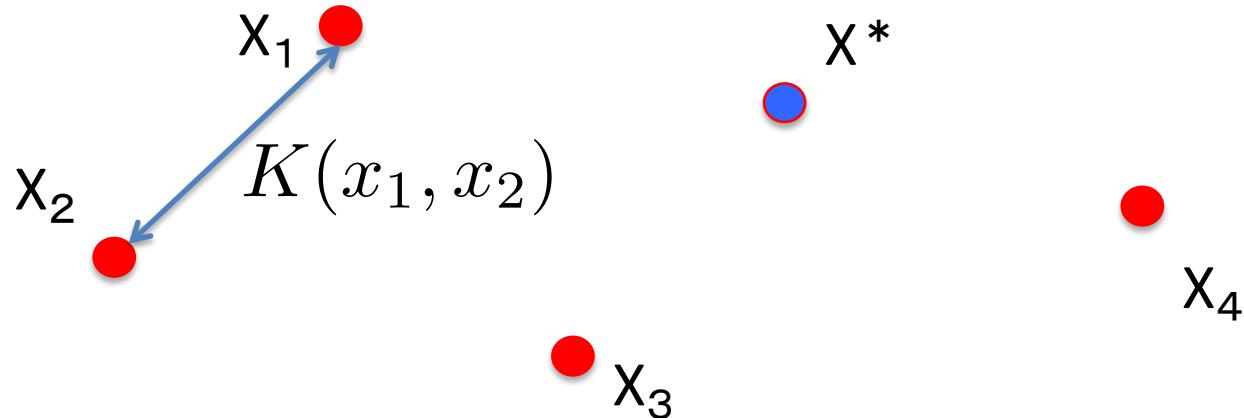
Gaussian Process

- Kernel method for regression
- Provides predictive variance in addition to regression function



Gaussian Process (No noise)

- Training points $\{x_i\}_{i=1,\dots,n}$ 、Test point x^*
- Observed outcomes y_i , y^* are subject to $n+1$ dim Gaussian
- Mean of y_i is 0.
- Covariances are given as $K(x_i, x_j)$



Covariance matrix by Gaussian kernel

$$\begin{pmatrix} k(\mathbf{x}^*, \mathbf{x}^*) & \mathbf{k}^{*\top} \\ \mathbf{k} & K \end{pmatrix}$$

$$K(x, x') = \exp(-\|x - x'\|^2/\eta)$$

Gaussian Process (No noise)

- K : Kernel matrix for training points
- \mathbf{y} : Observed outcomes for training points
- Predicted outcome at \mathbf{x}^*

$$E[y^*] = \mathbf{k}^{*\top} K^{-1} \mathbf{y}$$

- Predicted variance

$$V[y^*] = k(\mathbf{x}^*, \mathbf{x}^*) - \mathbf{k}^{*\top} K^{-1} \mathbf{k}^*$$

Gaussian process with noise

- Observed outcome include mean 0, variance σ^2 noise
- Covariance matrix

$$\begin{pmatrix} k(\mathbf{x}^*, \mathbf{x}^*) + \sigma^2 & \mathbf{k}^{*\top} \\ \mathbf{k} & K + \sigma^2 I \end{pmatrix}$$

Gaussian Process (with noise)

- K : Kernel matrix for training points
- \mathbf{y} : Observed outcomes for training points
- Predicted outcome at \mathbf{x}^*

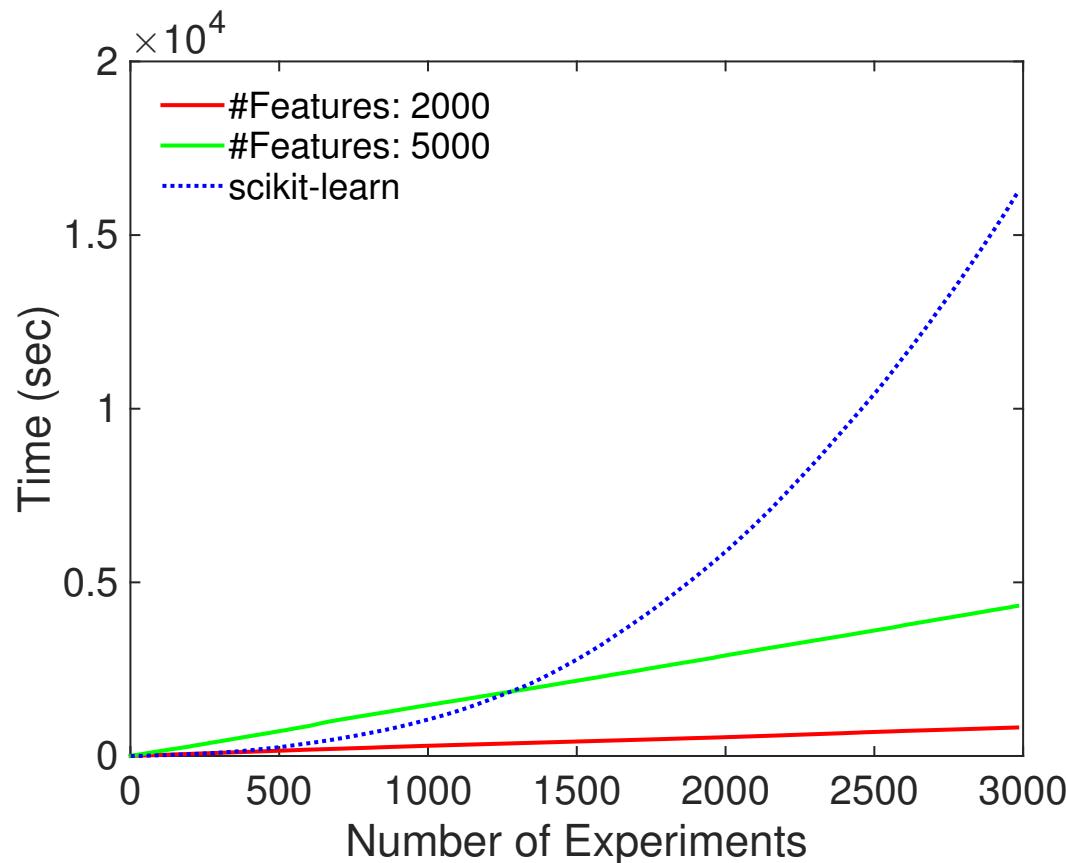
$$E[y^*] = \mathbf{k}^{*\top} (K + \sigma^2 I)^{-1} \mathbf{y}$$

- Predicted variance

$$V[y^*] = k(\mathbf{x}^*, \mathbf{x}^*) + \sigma^2 - \mathbf{k}^{*\top} (K + \sigma^2 I)^{-1} \mathbf{k}^*$$

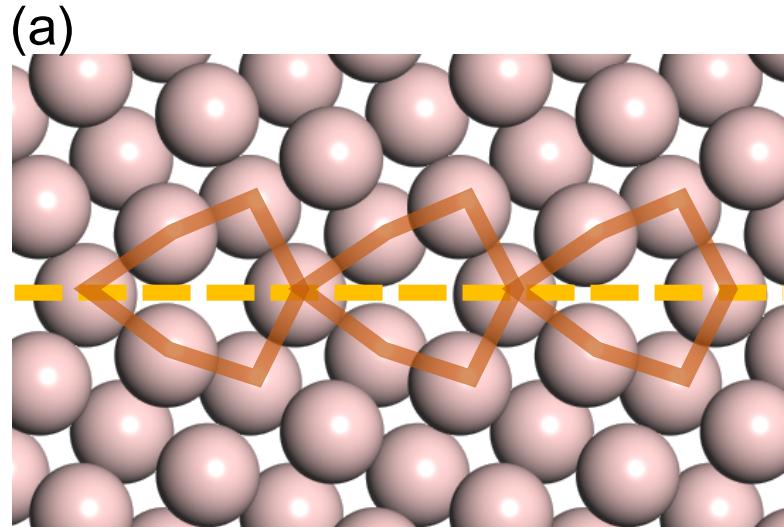
PHYSBO (COMBO)

- Fast learning by random feature maps
- Automatic hyperparameter initialization & update



ベイズ最適化による粒界構造決定の高速化

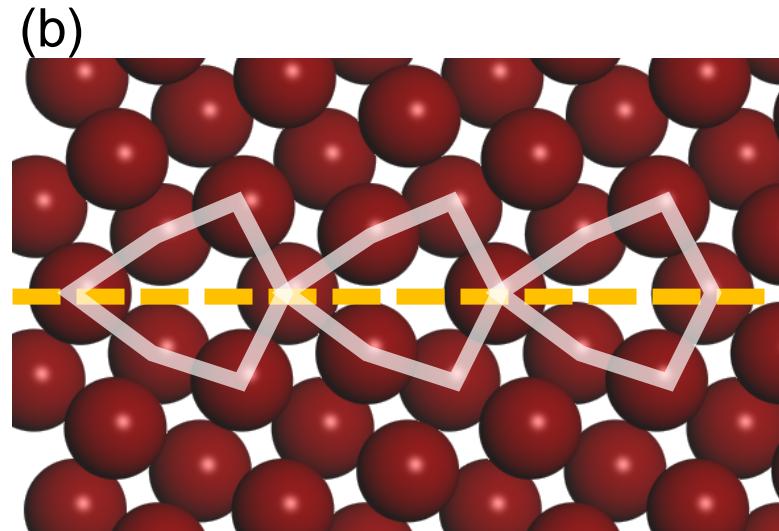
Cu [001] (210) $\Sigma 5$ 粒界



網羅的計算により決定

GB energy = 0.96J/m^2

計算回数 = 16,983回



ベイズ最適化により決定

GB energy = 0.96J/m^2

計算回数 = 69回
(20回のinitial trial含む)

(b)

