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

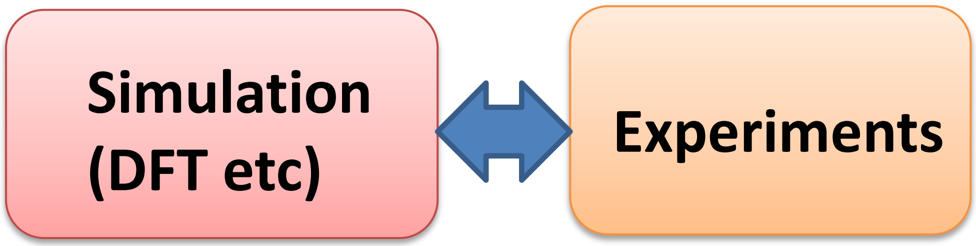
First Principles Calculations

Accurate, Slow

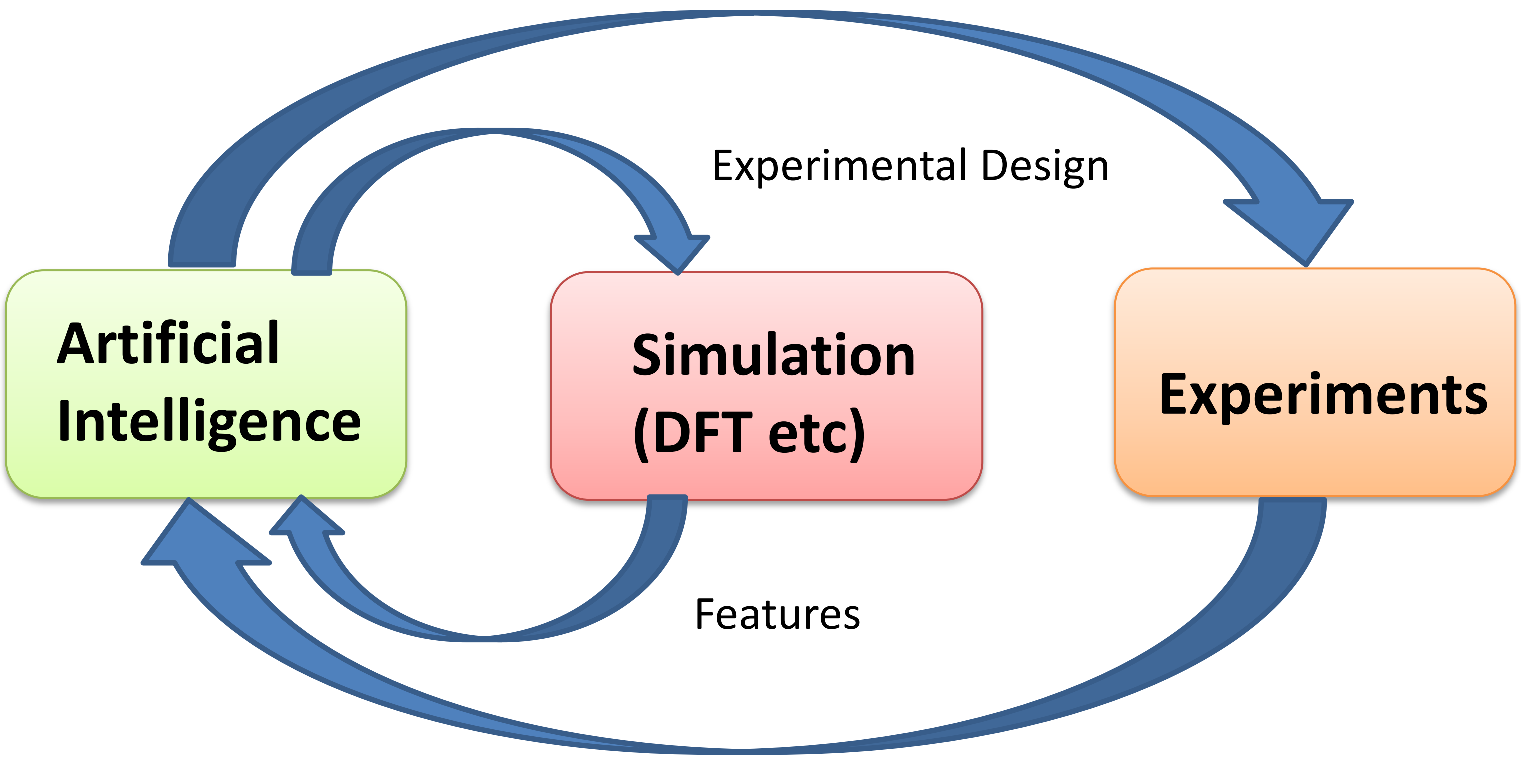
- Full configuration interaction
- Wave function based
- Density functional theory
- Semi-empirical
- Empirical potentials

Inaccurate, Fast

Old Picture



New Picture


Experimental Design
Artificial Intelligence
Simulation (DFT etc)
Experiments
Features

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.

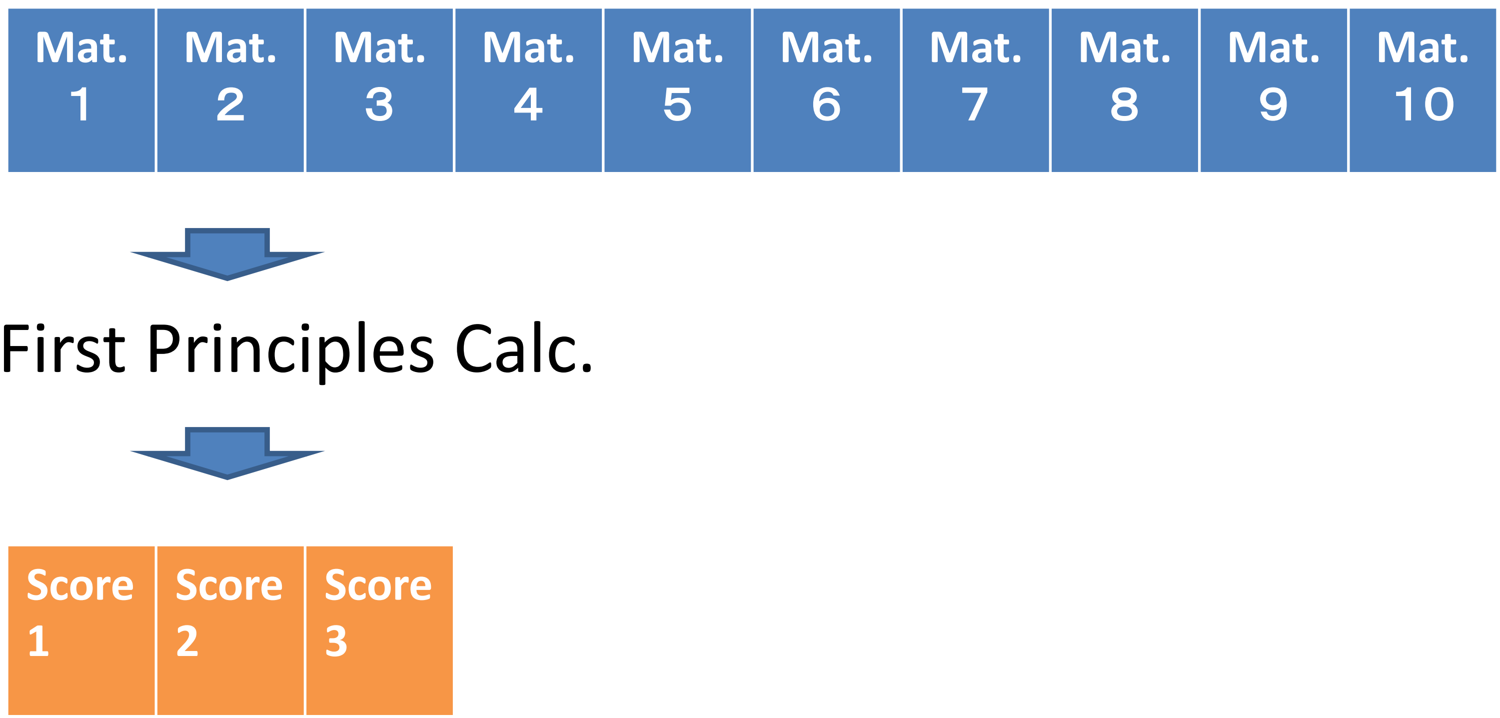
Score 1	Score 2	Score 3	Score 4	Score 5	Score 6	Score 7	Score 8	Score 9	Score 10

Bayesian Optimization
(Jones et al., 1998)

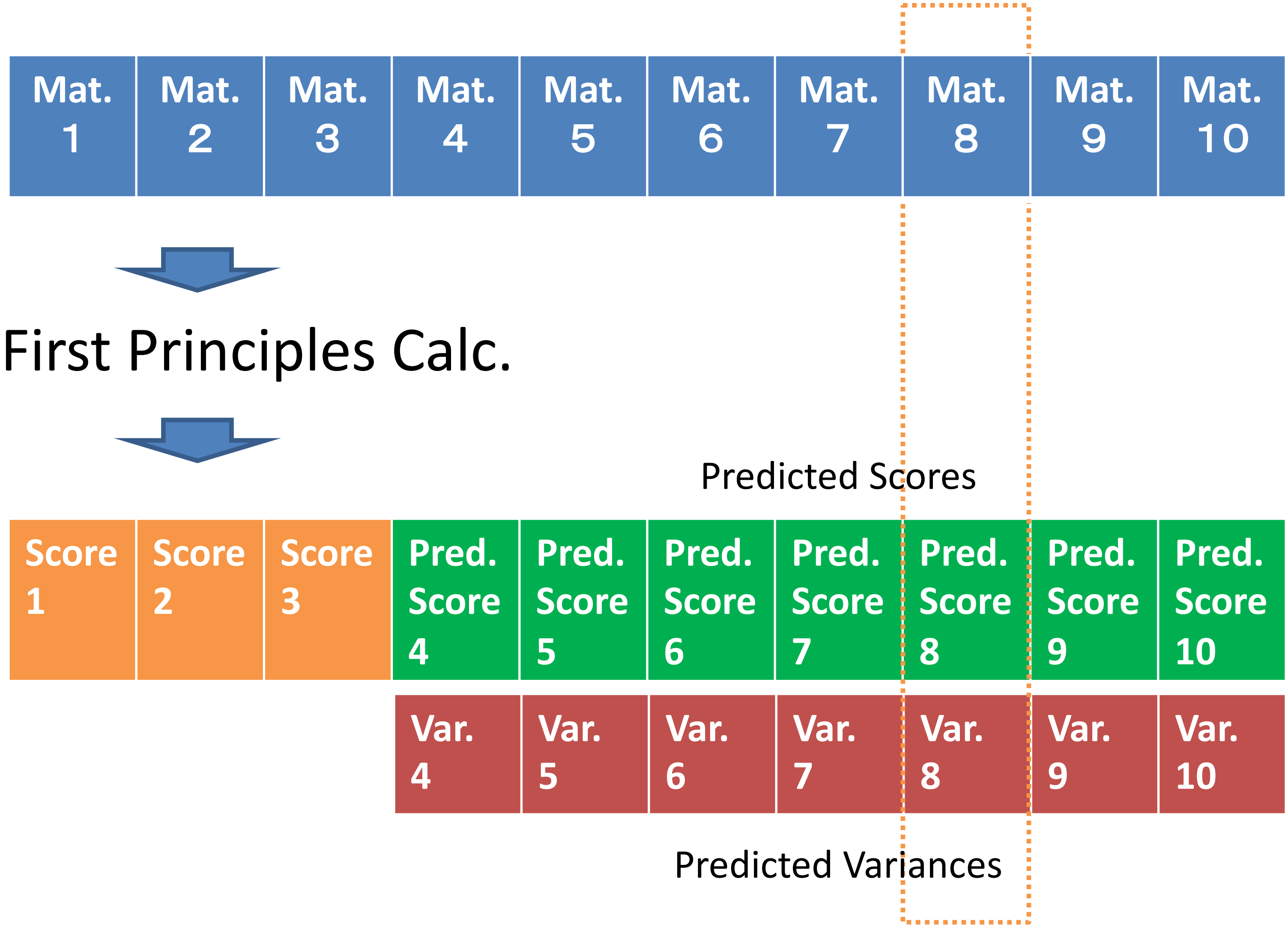
- Find best data points with minimum number of observations

- Choose next point to observe to discover the best ones as early as possible

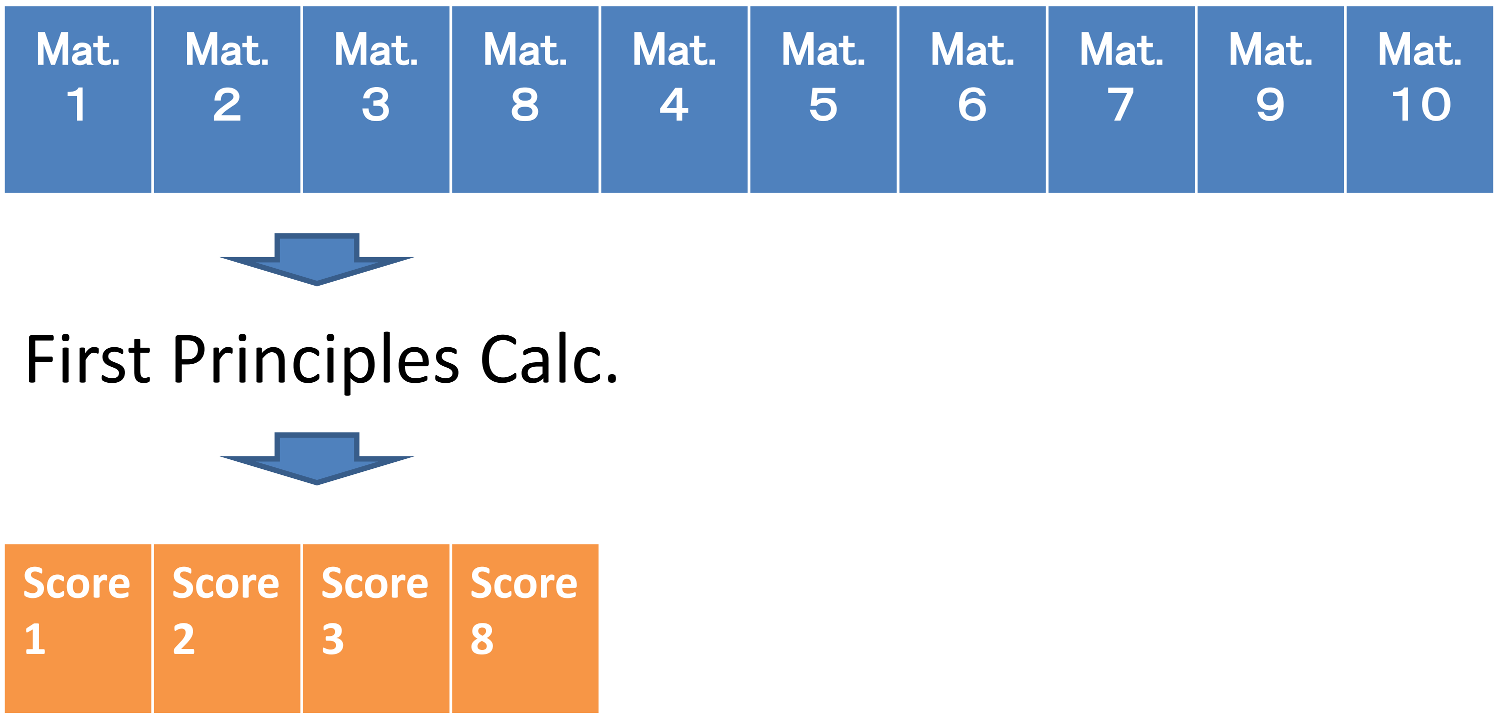
Bayesian Optimization (1)



Bayesian Optimization (2)



Bayesian Optimization (3)



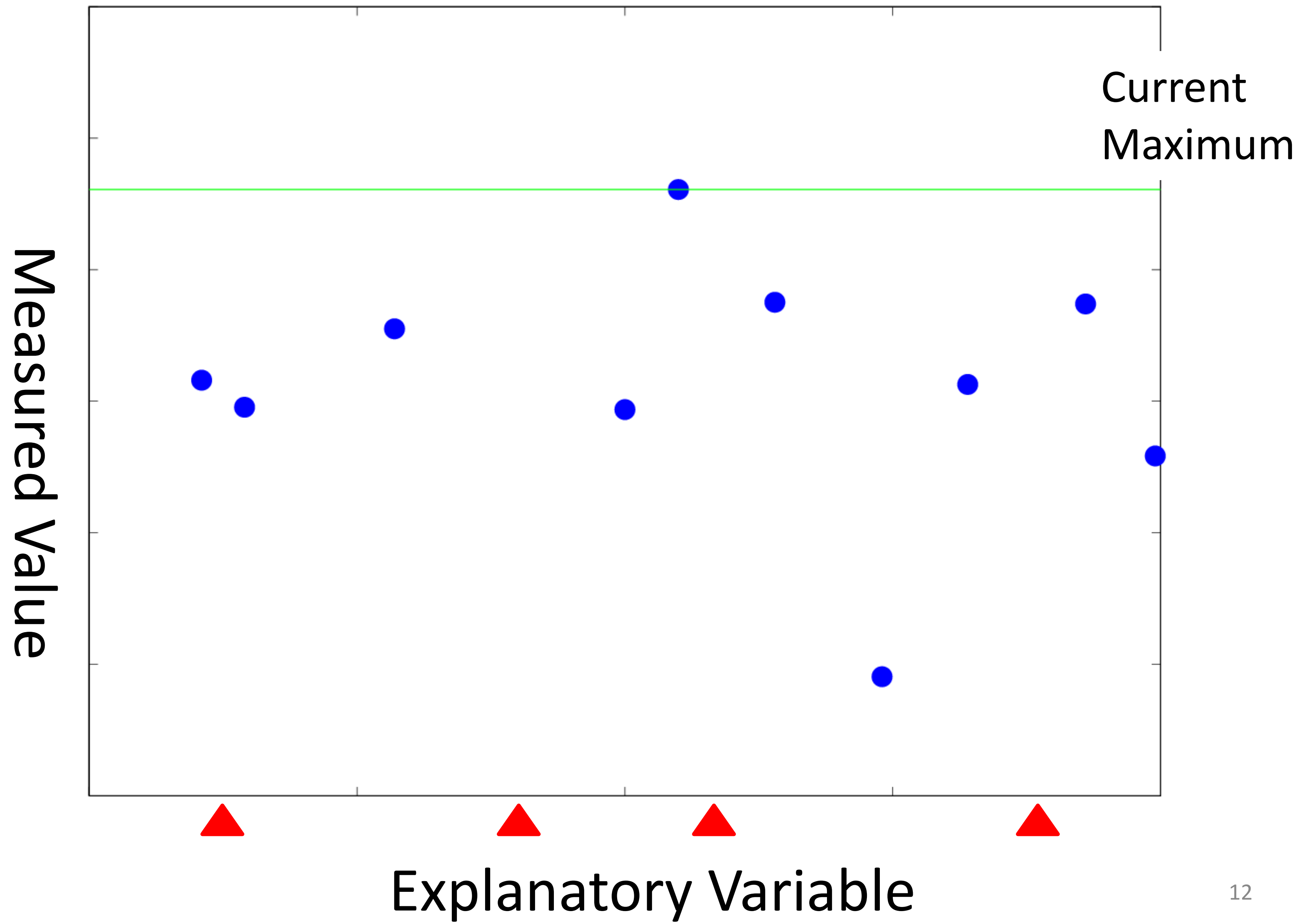
Bayesian Optimization (4)

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

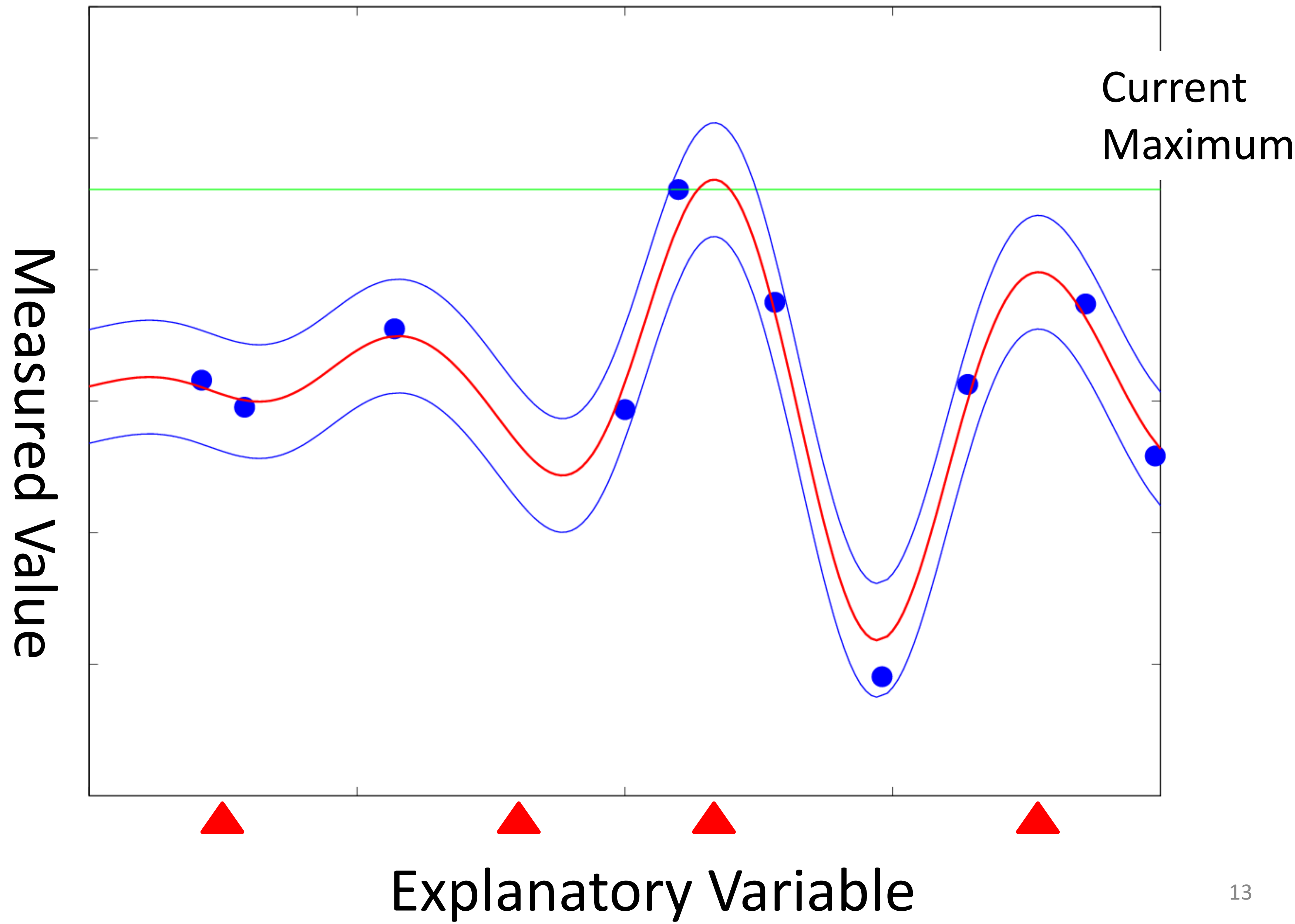
First Principles Calc.

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

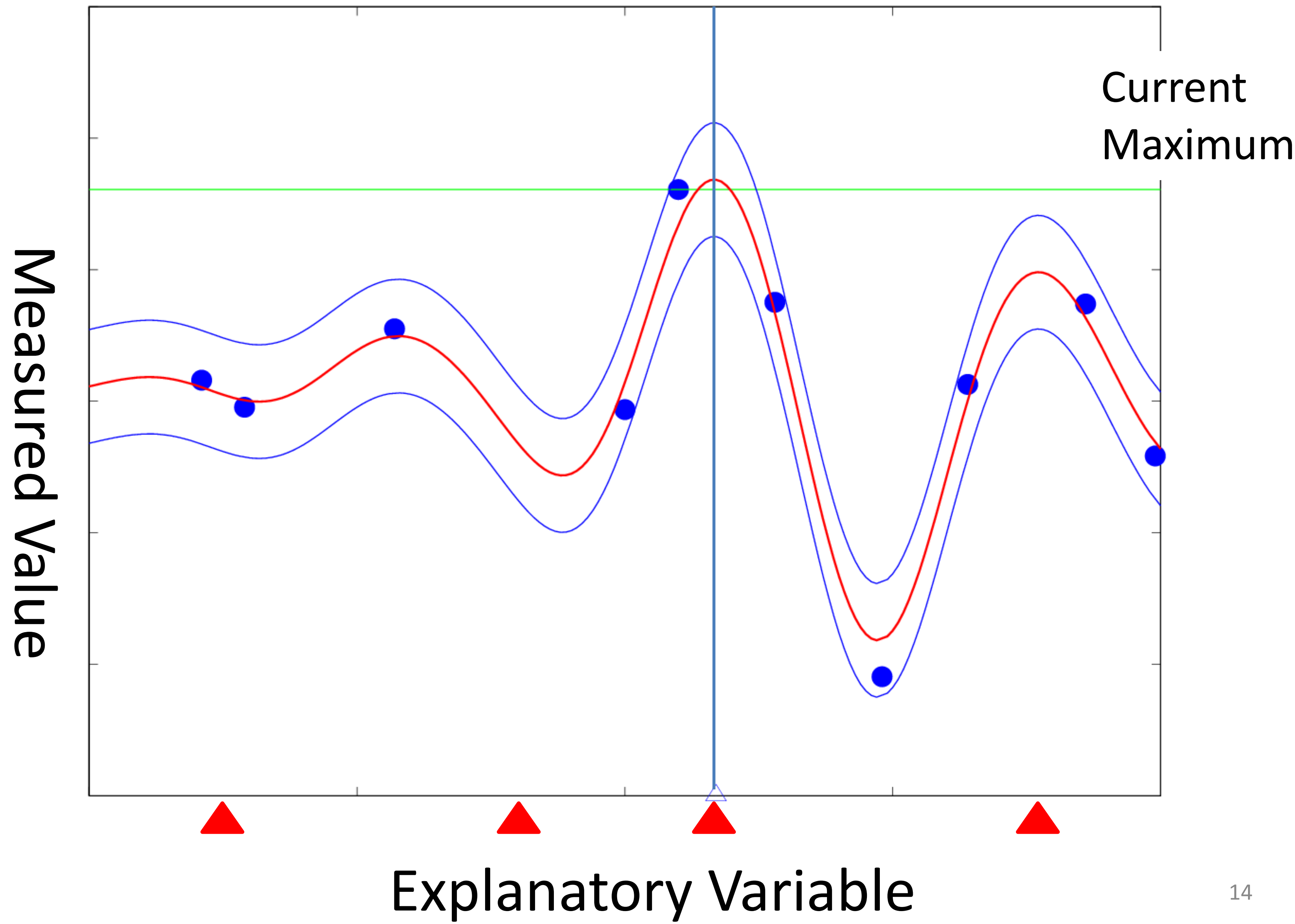
Where to observe next?



Gaussian Process



Maximum probability of improvement



Gaussian Process

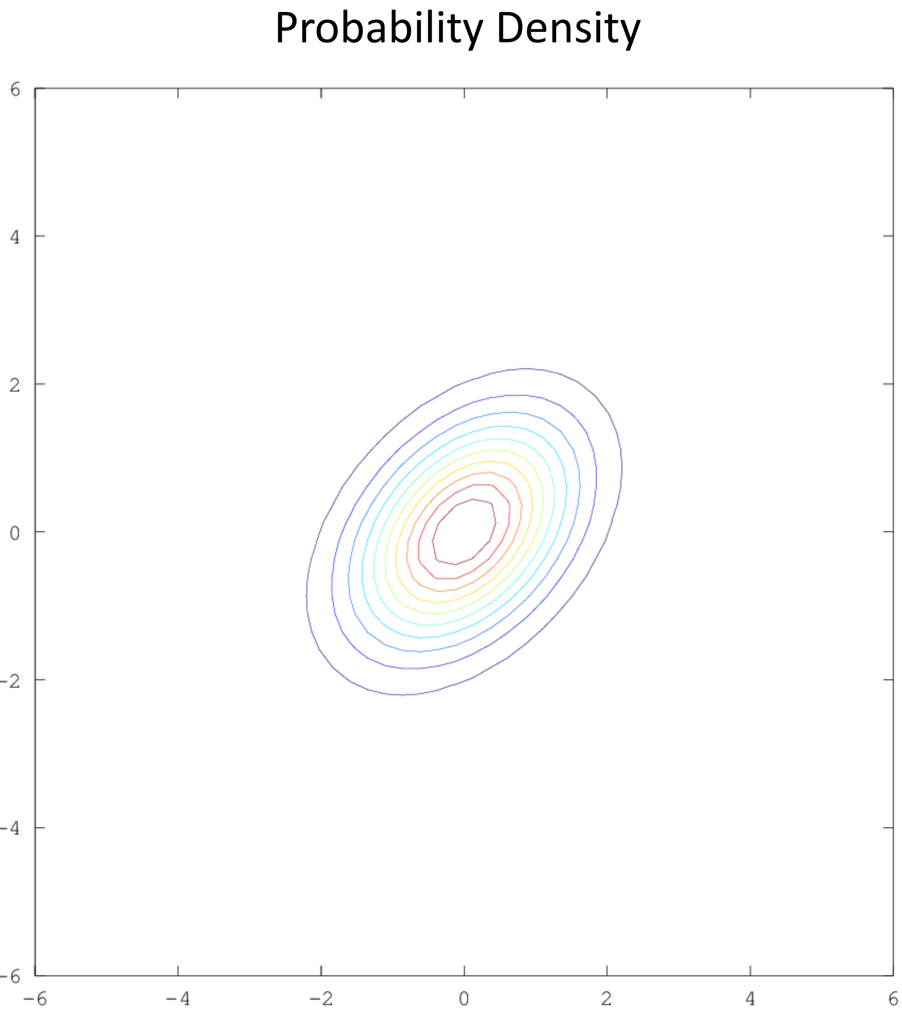
Multivariate Gaussian Distribution

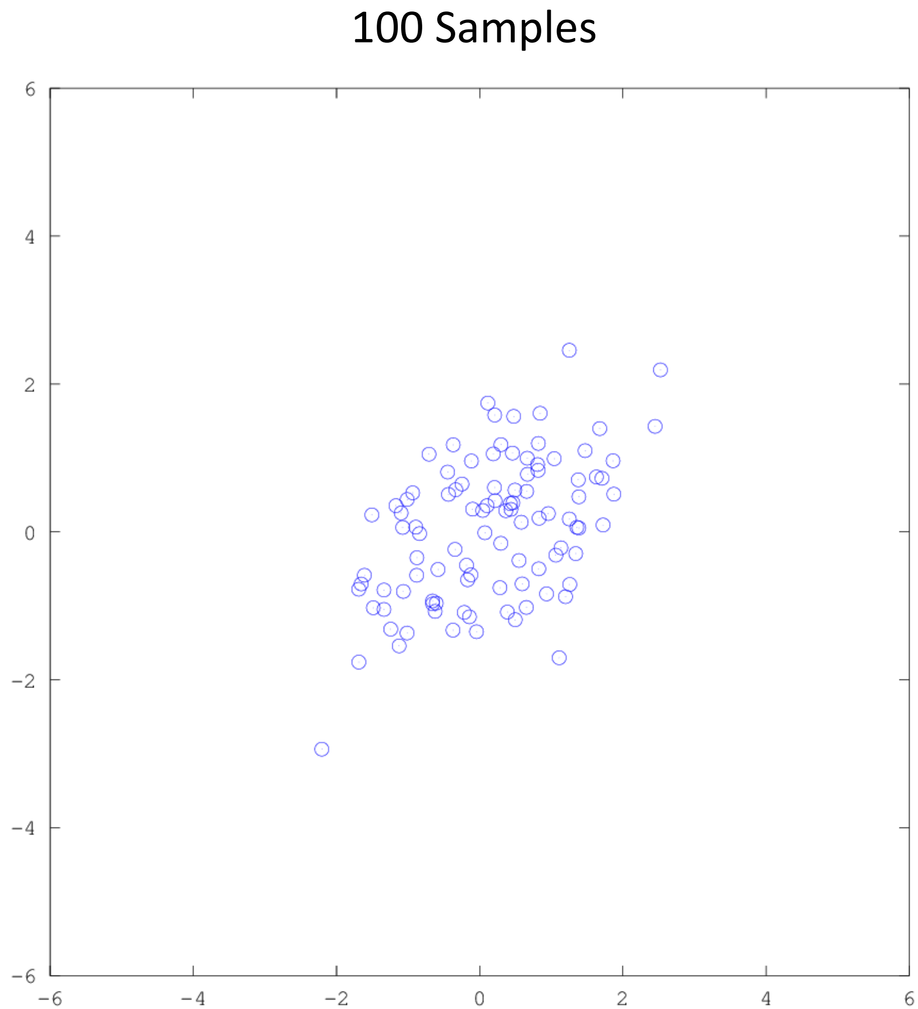
- Probability density function

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

$\boldsymbol{\mu}$ Mean

Σ Covariance Matrix





$$\boldsymbol{\mu} = (0, 0)^\top \qquad \Sigma = \begin{pmatrix} 0.4 & 1 \\ 1 & 0.4 \end{pmatrix}$$

Conditional Distribution

Mean Covariance

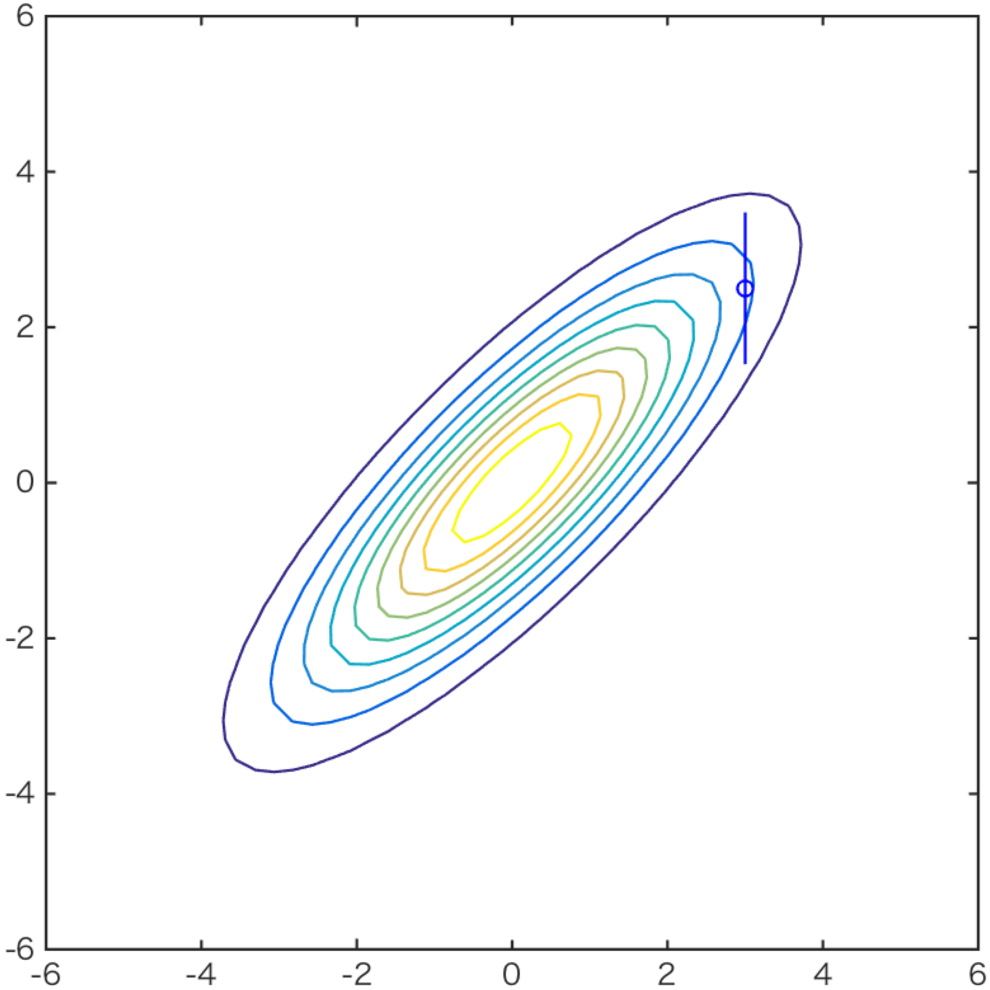
$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \quad \mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} \quad \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$$

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

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

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

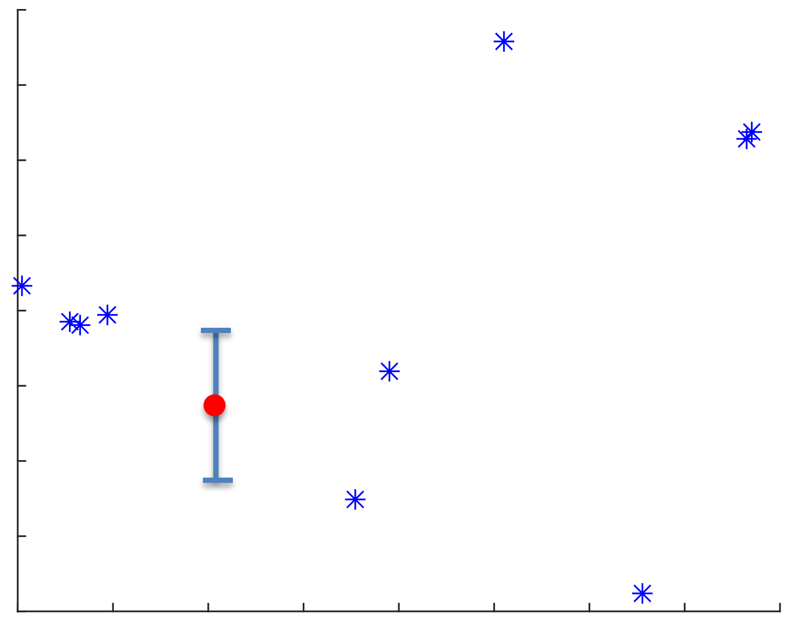
Conditional distribution at X＝3



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

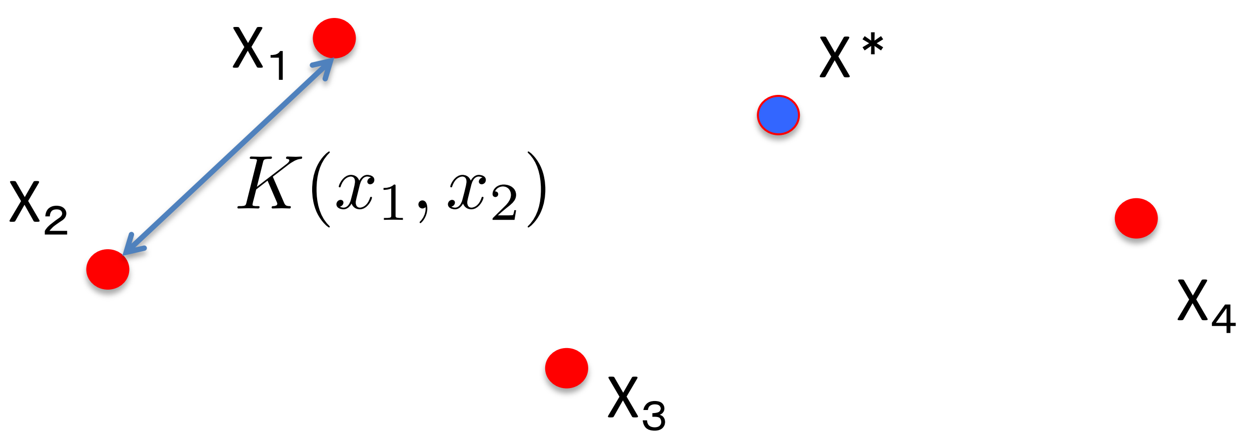
Gaussian Process

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



Gaussian Process（No noise）

- Training points $\{x_i\}_{i=1,\ldots,n}$、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
- **y**: Observed outcomes for training points
- Predicted outcome at ***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
- **y**: Observed outcomes for training points
- Predicted outcome at ***x****

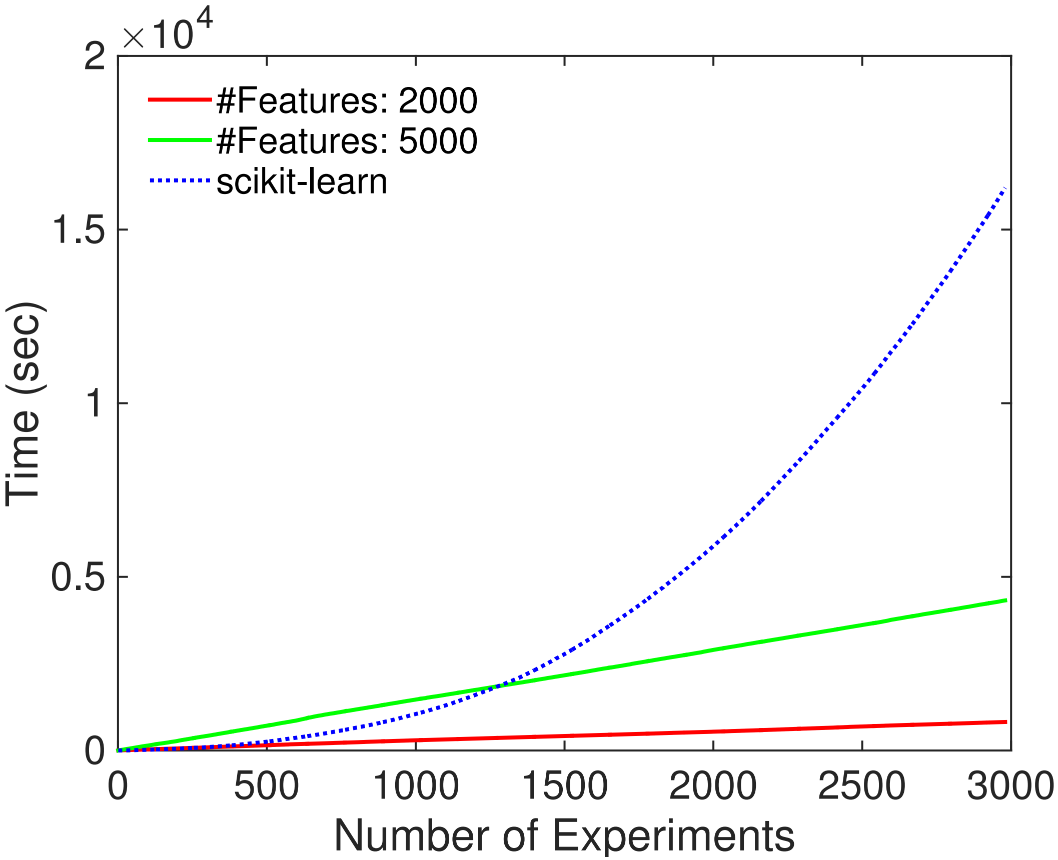
$$E[y^*] = \mathbf{k}^{*\top}(K + \sigma^2 I)^{-1}\mathbf{y}$$

- Predicted variance

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

PHYSBO (COMBO)

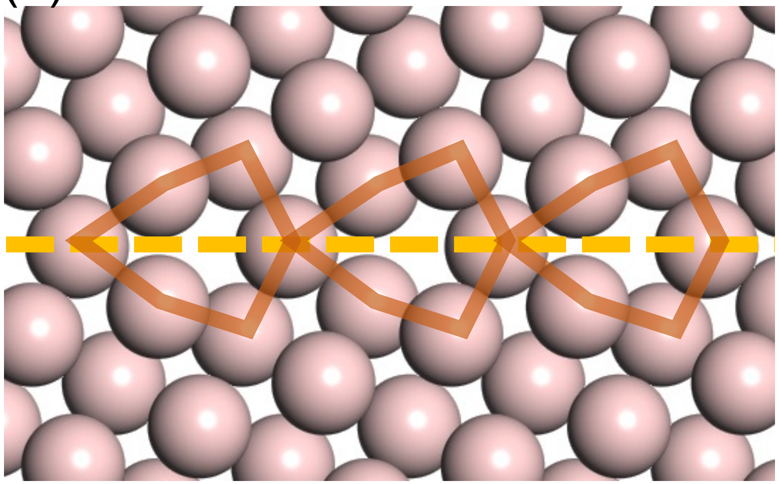
- Fast learning by random feature maps
- Automatic hyperparameter initialization & update



ベイズ最適化による粒界構造決定の高速化

Cu [001] (210) Σ5 粒界

(a)

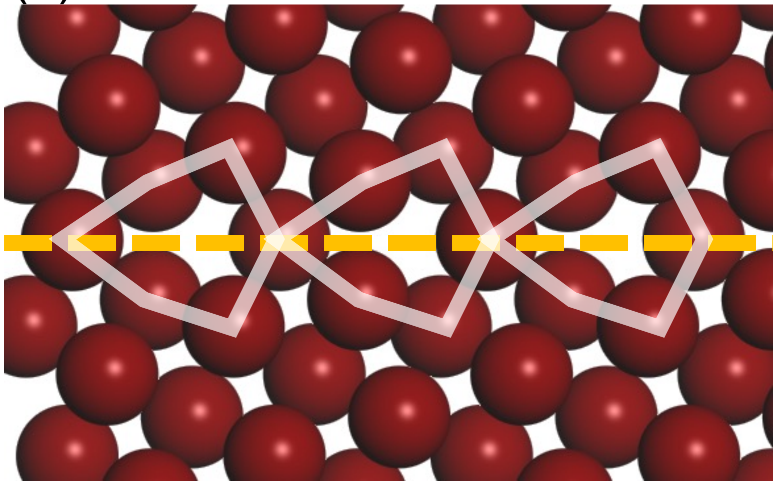


網羅的計算により決定

GB energy=0.96J/m^2

計算回数＝16,983回

(b)

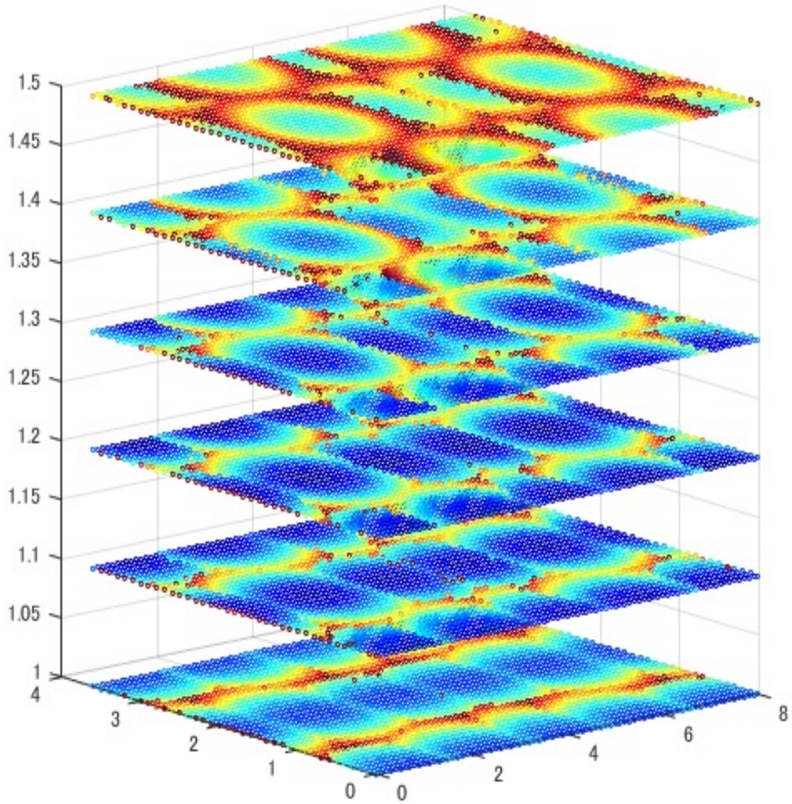


ベイズ最適化により決定

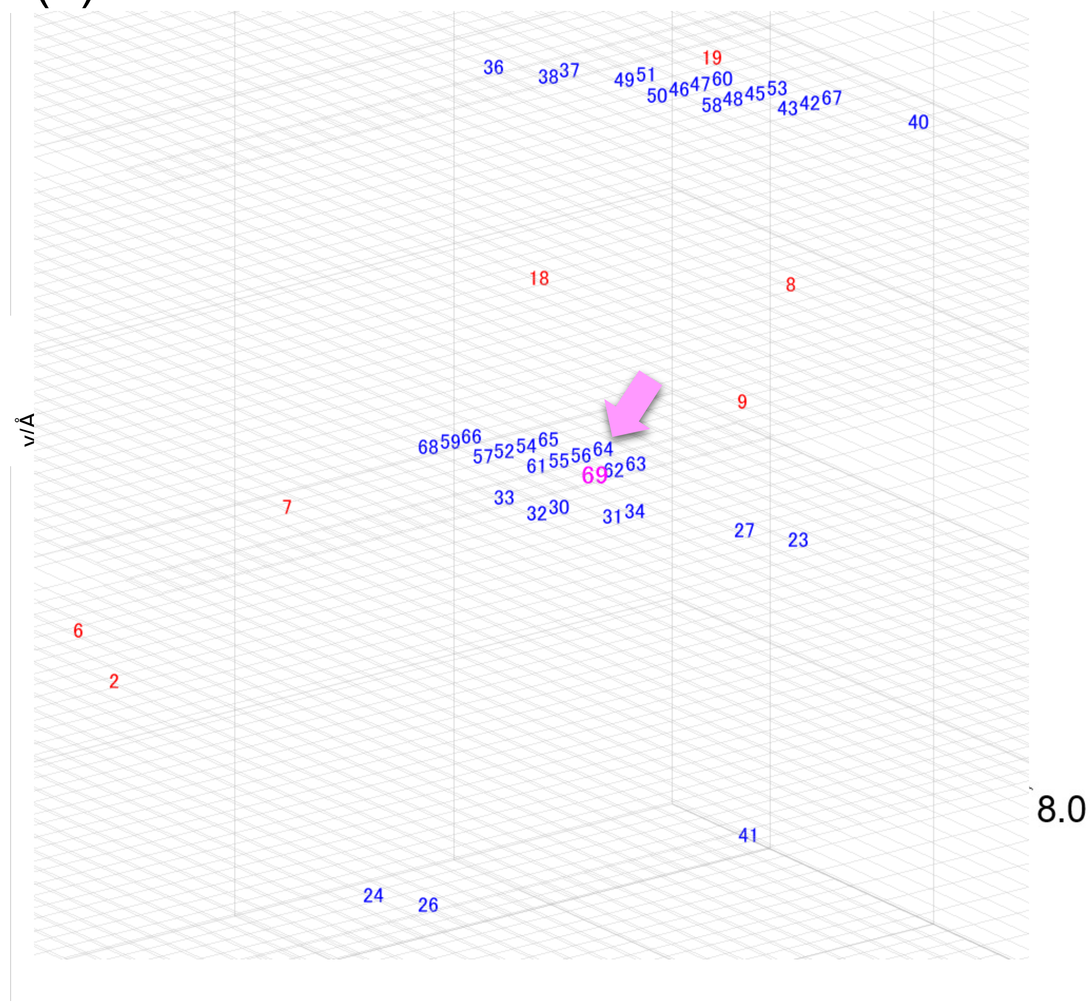
GB energy=0.96J/m^2

計算回数＝69回
(20回のinitial trial含む）

東大生研　溝口准教授との共同研究　S. Kiyohara et al., Jpn. J. Appl. Phys., 2016


1.5
1.45
1.4
1.35
1.3
1.25
1.2
1.15
1.1
1.05
1
4
3
2
1
0
0
2
4
6
8

(b)

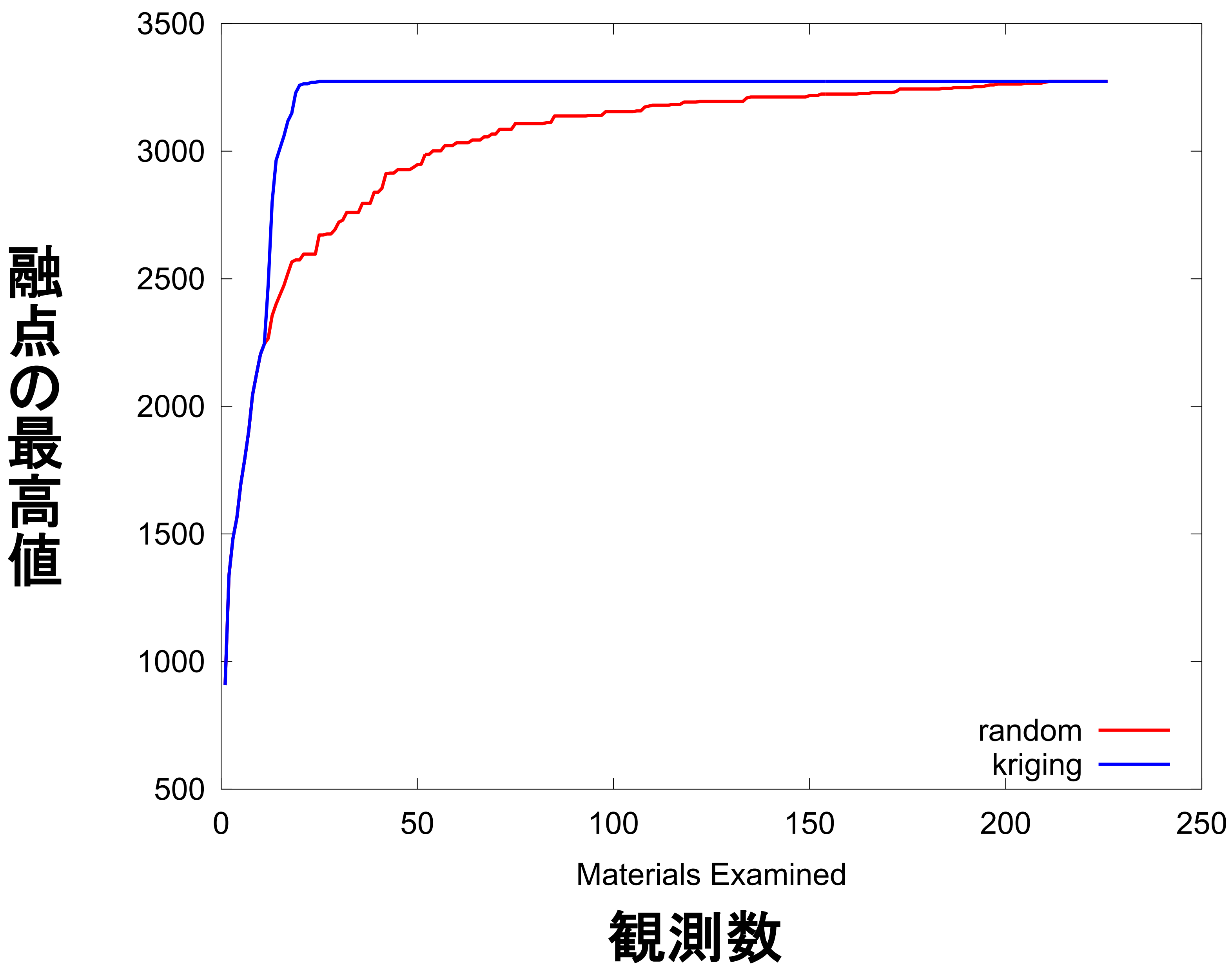

y/Å
8.0
19
36 38 37 49 51 50 46 47 60 58 48 45 53 43 42 67
40
18
8
9
68 59 66 57 52 54 65 61 55 56 64
69 62 63
7
33 32 30 31 34
27 23
6
2
41
24 26

二元化合物の融点データを用いた計算実験

- 226個ある材料のなかから、融点が最高のものを発見する
- 5%をランダムに選んで融点を観測する
- その後、ベイズ最適化を用いて、観測順を自動的に決定していく

１７個の説明変数

- #Ecoh:一原子あたりの凝集エネルギー(計算値)
- #bm:体積弾性率(計算値)
- #V:一原子あたりの格子体積(計算値)
- #NN:最近接原子間距離(計算値)
- #c:組成
- #Z1:構成元素の原子番号の二乗和
- #Z2:構成元素の原子番号の積
- #Z3:構成元素の原子番号の和
- #M1:構成元素の原子量の二乗和
- #M2:構成元素の原子量の積
- #M3:構成元素の原子量の和
- #n1:構成元素の価電子数の二乗和
- #n2:構成元素の価電子数の積
- #n3:構成元素の価電子数の和
- #p1:構成元素の周期の二乗和
- #p2:構成元素の周期の積
- #p3:構成元素の周期の和



融点の最高値
3500
3000
2500
2000
1500
1000
500
0
50
100
150
200
250
random
kriging
Materials Examined
観測数

最高融点の材料を見つけ出すまでの平均観測数

ベイズ最適化
16.1回

ランダム
133.4回

ベイズ最適化による実験順

- AlBr3 - As4S4 - GeSe - Se - BaSe - SnO2 - Sb2S3 - Sb2Te3 - Pb - SnF2 - GeBr2 - SnSe2 - BaO - BaS - SrSe - SiC - BeO - **[AlN]** - Be3N2 - Al2O3 - Si3N4 - Al4C3 - MgO - CaO - CaC2 - LiH - Cs - Be - BaH2 - Bi2O4 - K - BeF2 - Tl - RbN3 - LiF - PbTe - CsI - Li - P2O5 - Tl2O3 - BaF2 - Bi - Ba - CaS - SrO - CaSi - PbO - CaF2 - Rb - MgH2 - Si - BaSi2 - IBr - Bi2O3 - SrS - NaF - Ga2O3 - Al - TlI - CsO2 - KCl - In - I2 - BiF3 - SrF2 - LiCl - InN - CsBr - ICl - SrH2 - Pb3O4 - Na - Na2O2 - In2O3 - RbI - S - PbF2 - Bi2Te3 - Sn - CaH2 - KF - InSb - Ca - BiI3 - CsCl - K2O2 - MgF2 - Ge - PbS - SrSi2 - TeO2 - TlSe - Sr - BaI2 - AlP - Li2O - RbO2 - CsF - P4S3 - BiF5 - Mg - GeO2 - NaCl - CaSi2 - BaCl2 - Te - PbSe - TeF4 - PbI2 - TlF - KI - P - MgS - SnTe - NaO2 - GaAs - RbCl - Tl2O - SiS2 - KO2 - InAs - BaBr2 - P2S3 - Sb - KBr - TeI4 - Li3N - TeO3 - RbBr - SiI4 - LiBr - GaSb - TlCl - SeO3 - GaP - RbF - SnI4 - Cs2O - As2O3 - SrCl2 - Mg2Si - TlBr - AlAs - LiI - P4S7 - Bi2S3 - Mg2Sn - CaCl2 - AlI3 - As2O5 - SnSe - Ca3N2 - Li2S - NaBr - InI3 - BeCl2 - Sb2O3 - NaI - Mg2Ge - InI - BiBr3 - GeS - BeI2 - SeBr4 - Tl2S - InP - GaTe - P2S5 - SbF3 - K2S - BiCl3 - SrBr2 - InF3 - GeTe - SbI3 - AlSb - In2Te3 - GeF2 - Mg3Sb2 - SrI2 - PbCl2 - GaS - PI3 - Na2S - SnS - Al2S3 - GaI3 - Rb2S - GaSe - MgCl2 - TeCl4 - Rb2Se - PbBr2 - GeI4 - K2Se - CaI2 - BeBr2 - P2I4 - Sb2Se3 - CaBr2 - As2Te3 - In2Se3 - AlCl3 - InS - GeBr4 - As2S3 - Ga2Se3 - SnBr4 - InCl - As2Se3 - AsBr3 - AsI3 - GaBr3 - Al2Te3 - In2S3 - SbBr3 - MgI2 - InBr3 - GeS2 - MgBr2 - Ga2S3 - GaCl3 - SbCl3 - SnBr2 - GaCl2 - SnCl2 -