Inferring Black-Box Function Statistics under a Limited Computational Budget

SIAM CSE 2019
Spokane, WA

Piyush Pandita Nimish Awalgaonkar
Ilias Bilionis Jitesh Panchal
School of Mechanical Engineering,
Purdue University, West Lafayette, IN

Motivating example

- Expensive computer codes ($50 per simulation/experiment)
- Limited computational budget ($1000)
- Limited initial observations (10 observations for 5 dim. input space)
- **Quantity of Interest (QoI): function of the code output**

Frame of Reference

- What is a Black-Box function ?
  - Computationally expensive computer codes
  - Time-taking laboratory experiments
- What are different kinds of Uncertainty ?
  - Stochastic input(s)
  - Measurement noise (experiments)

State-of-the-art

- Classical methods
  - Factorial designs
  - Alphabetical designs (A, D etc.)
  - e.g. Latin-hypercube (LHS)
  - Usefulness
    - Fixed sample size
- Sequential methods
  - Bayesian Alphabetical designs
  - Bayesian global optimization
  - e.g. efficient global optimization, entropy search
  - Usefulness
    - Varying sample size
    - Expensive objective functions

PREDICTIVE SCIENCE LABORATORY

PURDUE UNIVERSITY

General Sequential DOE Algorithm

1. Start from some initial simulations/observations.
2. Train Bayesian surrogates of the output(s).
3. Find the design point $\mathbf{x}_{\text{next}}$ that maximizes some sampling criterion.
4. Perform simulations at $\mathbf{x}_{\text{next}}$ to get output(s) $\boldsymbol{y}_{\text{next}}$.
4. Add $(\mathbf{x}_{\text{next}}, \mathbf{y}_{\text{next}})$ to observed data.
5. If budget is exhausted STOP. Otherwise GOTO 2.

Fully Bayesian Surrogates

- State-of-the-art
  - Polynomial Chaos expansion
  - Gaussian Process (GP) regression
- GP regression
  - Stationary covariance kernels

$$f \sim \mathrm{GP}(0, k)$$

PREDICTIVE SCIENCE LABORATORY

PURDUE UNIVERSITY

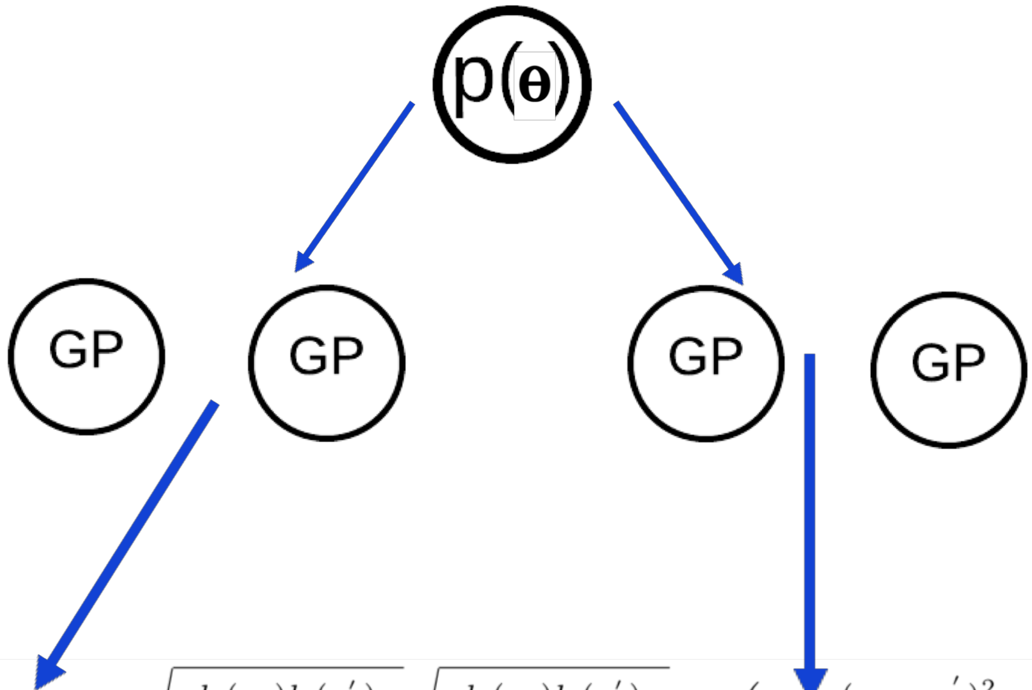
Inferring Local Properties

$$k(x,x') = s^2 \exp\left(-\frac{(x-x')^2}{2l^2}\right)$$

$$k(x,x') = s(x)s(x')\sqrt{\frac{l(x)l(x')}{l^2(x)+l^2(x')}}\exp\left(-\frac{(x-x')^2}{l^2(x)+l^2(x')}\right)$$

$$k(\mathbf{x},\mathbf{x}') = \prod_{i=1}^{d} s(x_i)s(x_i')\sqrt{\frac{l(x_i)l(x_i')}{l^2(x_i)+l^2(x_i')}}\exp\left(-\frac{(x_i-x_i')^2}{l^2(x_i)+l^2(x_i')}\right)$$

Fully Bayesian Inference

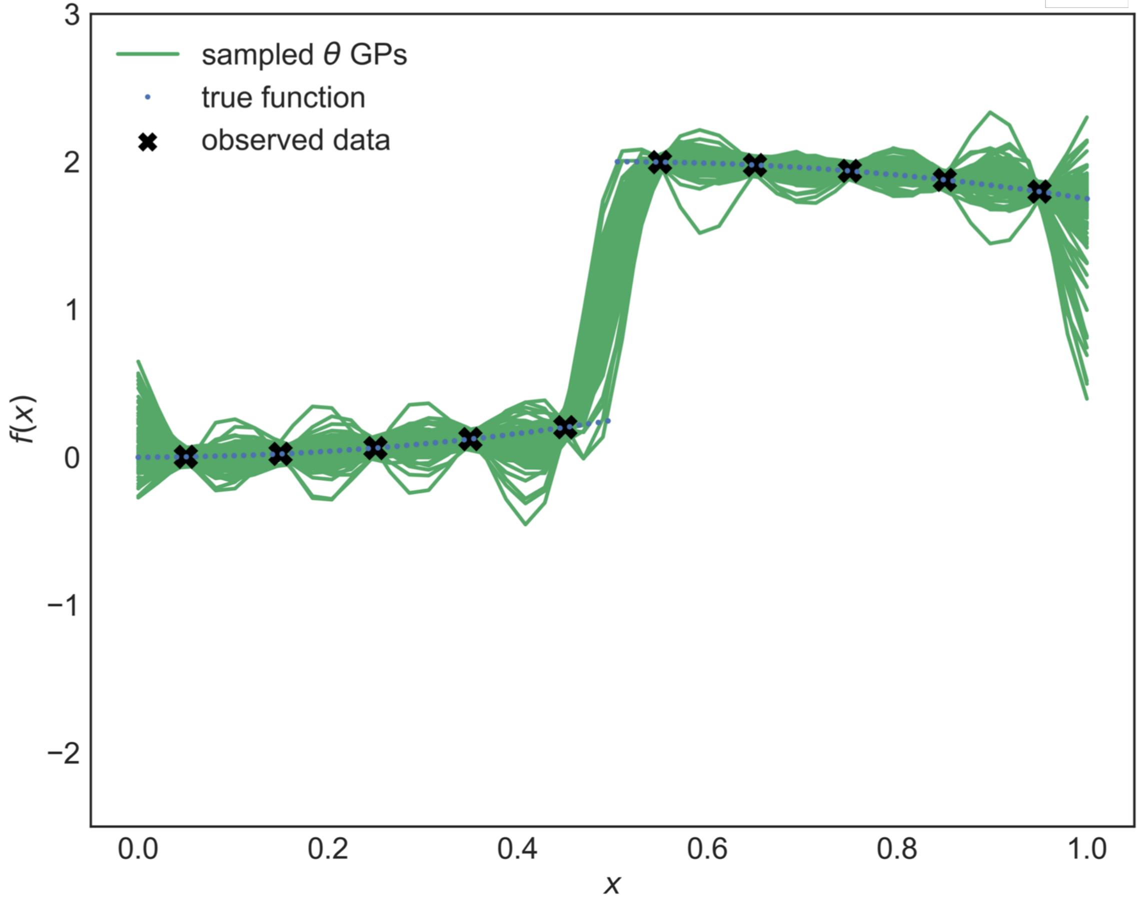


$$k(\mathbf{x}, \mathbf{x}') = s_1(x_1)s_1(x_1')s_2(x_2)s_2(x_2')\sqrt{\frac{l_1(x_1)l_1(x_1')}{l_1^2(x_1)+l_1^2(x_1')}}\sqrt{\frac{l_2(x_2)l_2(x_2')}{l_2^2(x_2)+l_2^2(x_2')}}\exp\left(-\frac{(x_1-x_1')^2}{l_1^2(x_1)+l_1^2(x_1')}\right)\exp\left(-\frac{(x_2-x_2')^2}{l_2^2(x_2)+l_2^2(x_2')}\right)$$

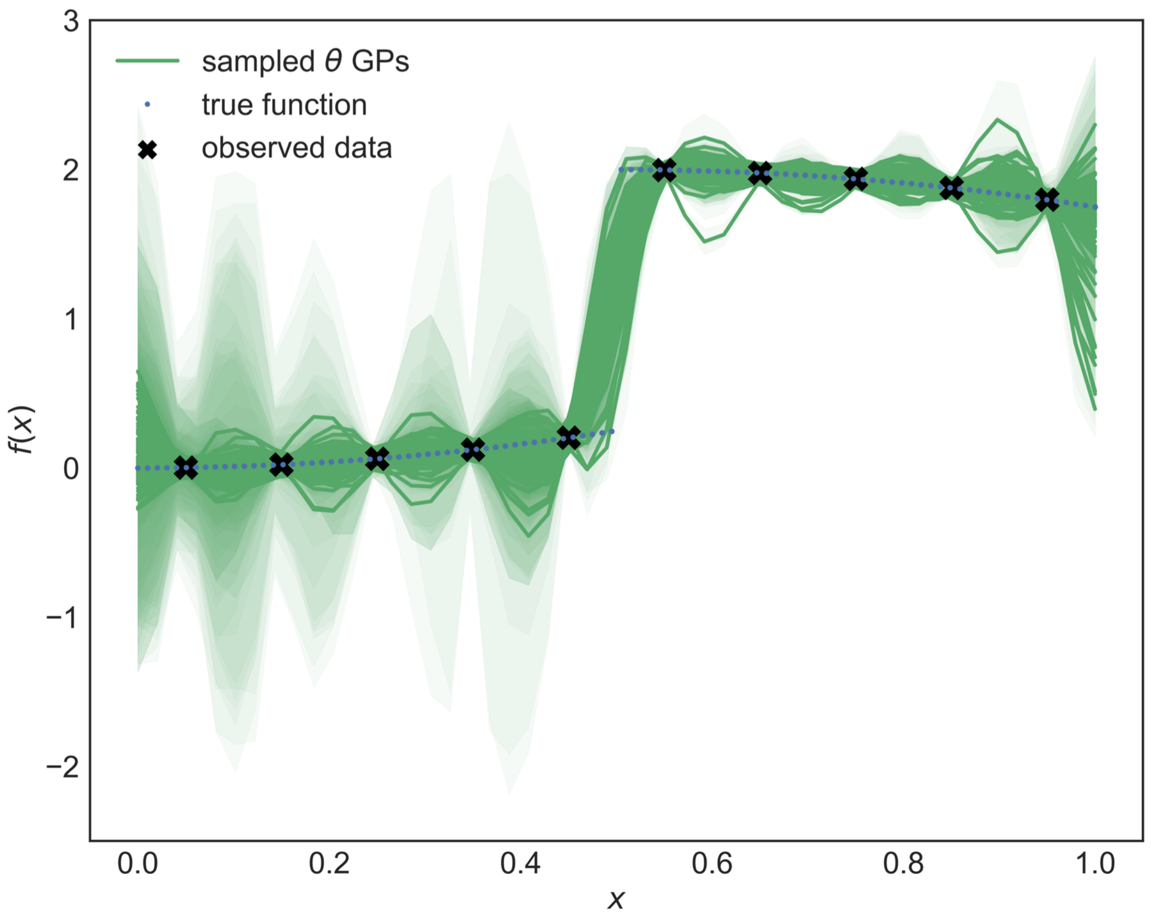
$$f \sim \mathrm{GP}(0, k)$$

Fully Bayesian Inference

posterior samples via MCMC: $\boldsymbol{\theta}$



Fully Bayesian Inference



Quantifying Information Gain

Want to infer: $Q[f]$

given **posterior samples** via MCMC: $\boldsymbol{\theta}$

$$p(Q|\boldsymbol{\theta},\mathbf{D}_n)=\mathbb{E}[\delta(Q-Q[f])|\boldsymbol{\theta},\mathbf{D}_n]$$

$$p(Q|\boldsymbol{\theta},\mathbf{D}_n)=\mathcal{N}(Q|\mu_1,\sigma_1^2)$$ (MC)

* Pandita, Piyush, Ilias Bilionis, and Jitesh Panchal. "Bayesian Optimal Design of Experiments For Inferring The Statistical Expectation Of A Black-Box Function."

Quantifying Information Gain

hypothetical experiment $\tilde{\mathbf{x}} \in \mathrm{R}^d$

$$p(Q|\boldsymbol{\theta}, \mathbf{D}_n, \tilde{\mathbf{x}}, \tilde{y}) = \mathbb{E}\left[\delta\left(Q - Q[f]\right)|\boldsymbol{\theta}, \mathbf{D}_n, \tilde{\mathbf{x}}, \tilde{y}\right]$$

since,

$$p(\tilde{y}|\tilde{\mathbf{x}}, \mathbf{D}_n, \boldsymbol{\theta}) = \mathcal{N}\left(\tilde{y}\middle|m_n(\tilde{\mathbf{x}};\boldsymbol{\theta}), \sigma_n^2(\tilde{\mathbf{x}};\boldsymbol{\theta})\right)$$

$$p(Q|\boldsymbol{\theta}, \mathbf{D}_n, \tilde{\mathbf{x}}, \tilde{y}) = \mathcal{N}(Q|\mu_2(\tilde{\mathbf{x}}, \tilde{y}), \sigma_2^2(\tilde{\mathbf{x}}))$$ (MC)

* Pandita, Piyush, Ilias Bilionis, and Jitesh Panchal. "Bayesian Optimal Design of Experiments For Inferring The Statistical Expectation Of A Black-Box Function."

PREDICTIVE SCIENCE LABORATORY

PURDUE UNIVERSITY

Quantifying Information Gain

Information Gain in a hypothetical experiment $\tilde{\mathbf{x}}$ is the KLD between

$p(Q|\boldsymbol{\theta}, \mathbf{D}_n)$ and $p(Q|\boldsymbol{\theta}, \mathbf{D}_n, \tilde{\mathbf{x}}, \tilde{y})$

$$G(\tilde{\mathbf{x}}, \tilde{y}; \boldsymbol{\theta}) = \int_{-\infty}^{\infty} p(Q|\boldsymbol{\theta}, \mathbf{D}_n, \tilde{\mathbf{x}}, \tilde{y})) \log \frac{p(Q|\boldsymbol{\theta}, \mathbf{D}_n, \tilde{\mathbf{x}}, \tilde{y}))}{p(Q|\boldsymbol{\theta}, \mathbf{D}_n)} dQ$$

$$G(\tilde{\mathbf{x}}) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} G(\tilde{\mathbf{x}}, \tilde{y}; \boldsymbol{\theta}) p(\tilde{y}|\boldsymbol{\theta}, \tilde{\mathbf{x}}, \mathbf{D}_n) p(\boldsymbol{\theta}|\mathbf{D}_n) d\tilde{y} d\boldsymbol{\theta}$$

$G(\tilde{\mathbf{x}})$ approximated via. MC sampling

$$\mathbf{x}_{n+1} = \arg\max_{\tilde{\mathbf{x}}} G(\tilde{\mathbf{x}})$$

PREDICTIVE SCIENCE LABORATORY

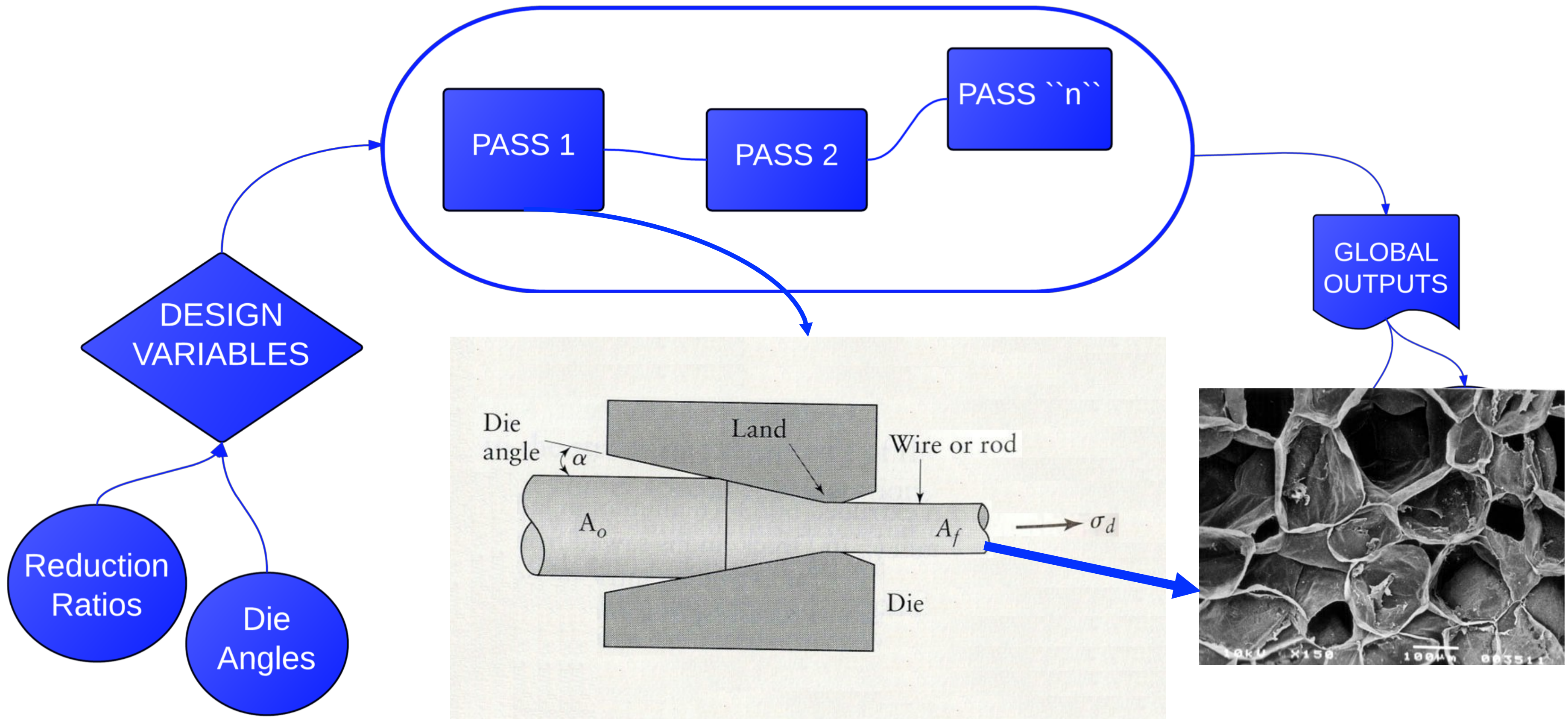
PURDUE UNIVERSITY

Algorithm

1. Start with some initial simulations/observations.
2. Train Fully Bayesian surrogate(s) of the objective(s) using MCMC (Hamiltonian Monte Carlo).
3. Find the design point $\mathbf{x}_{\text{next}}$ that maximizes:
   $\boldsymbol{G}(\tilde{\mathbf{x}})$ (using Bayesian Global Optimization)
4. Perform simulations at $\mathbf{x}_{\text{next}}$ to get objectives $\boldsymbol{y}_{\text{next}}$.
5. Add $(\mathbf{x}_{\text{next}}, \mathbf{y}_{\text{next}})$ to observed data.
6. If relative KL below threshold or budget is exhausted STOP. Otherwise GOTO 2.

1. Pandita, Piyush, Ilias Bilionis, and Jitesh Panchal. "Bayesian Optimal Design of Experiments For Inferring The Statistical Expectation Of A Black-Box Function."

Application: Wire Drawing Process



courtesy TATA Research, Pune, India

TATA
TATA CONSULTANCY SERVICES

Wire Drawing Problem

QoI to be inferred: $Q[f(\mathbf{x})]$

where,

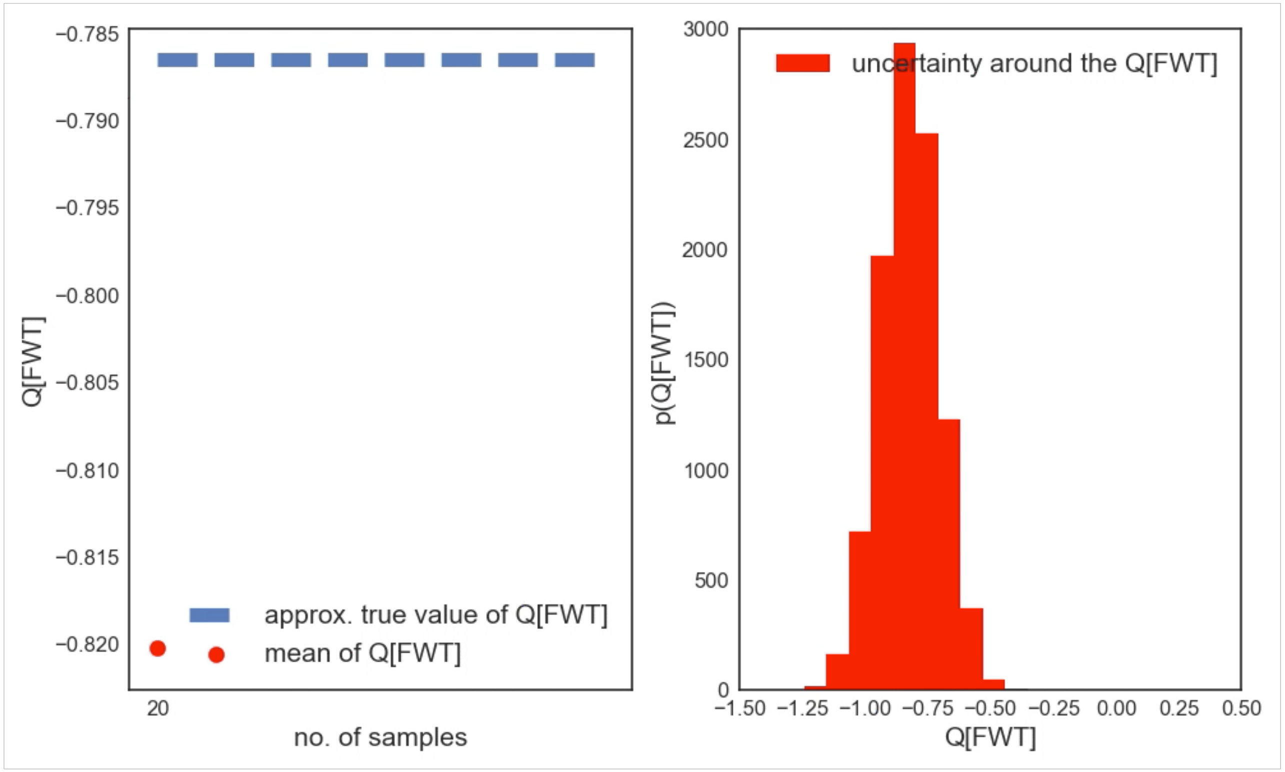
$f(\mathbf{x})$ = Friction work done (FWT) expensive-to-evaluate code

$\mathbf{x}^* \in R^8$, (8 die angles)

Wire Drawing Problem

Inferring statistical expectation

Mean $Q[f]$ Variance $Q[f]$



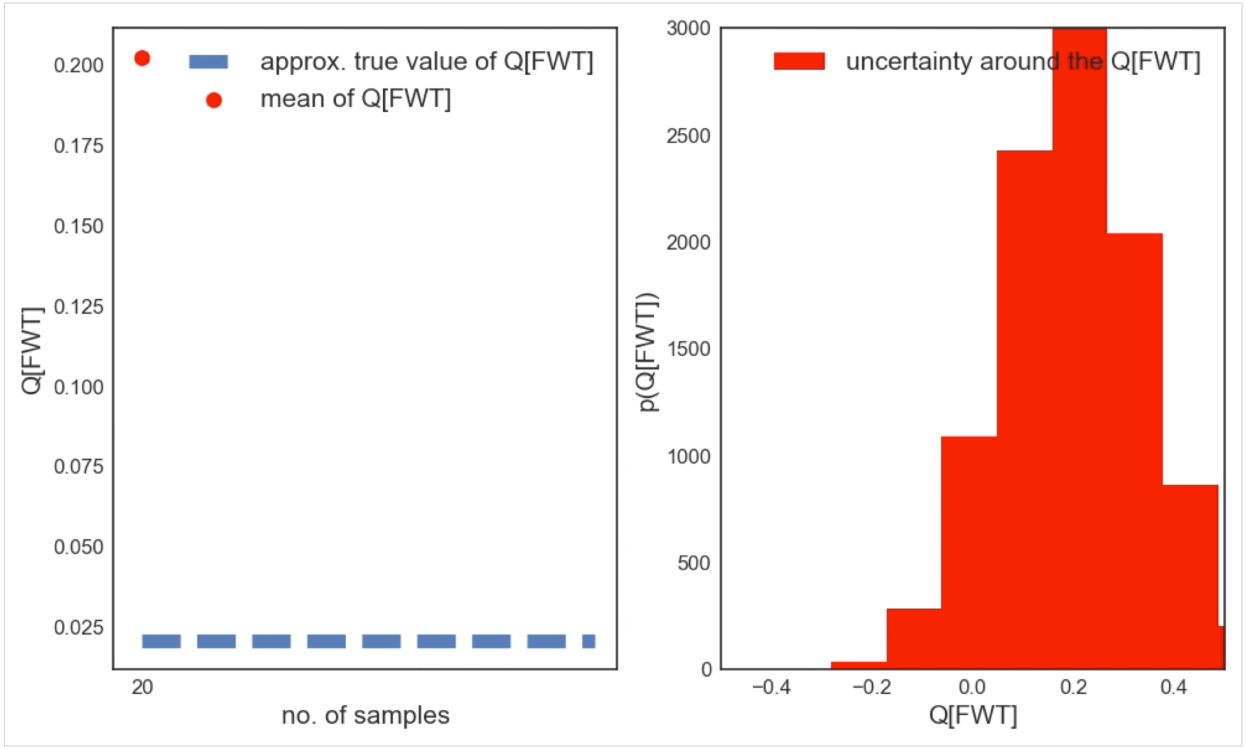
PREDICTIVE SCIENCE LABORATORY

PURDUE UNIVERSITY

Wire Drawing Problem

Inferring variance

Mean $Q[f]$ Variance $Q[f]$



TATA CONSULTANCY SERVICES

PREDICTIVE SCIENCE LABORATORY

PURDUE UNIVERSITY

Future Work

- More challenging QoIs
- Parallelizing the algorithm

PREDICTIVE SCIENCE LABORATORY

PURDUE UNIVERSITY

Thank You

References

- Kullback, Solomon, and Richard A. Leibler. "On information and sufficiency." *The annals of mathematical statistics* 22, no. 1 (1951): 79-86.
- Pandita, Piyush, Ilias Bilionis, and Jitesh Panchal. "Deriving Information Acquisition Criteria For Sequentially Inferring The Expected Value Of A Black-Box Function." ASME IDETC 2018
- Huan, Xun, and Youssef Marzouk. "Gradient-based stochastic optimization methods in Bayesian experimental design." *International Journal for Uncertainty Quantification* 4, no. 6 (2014).
- Hennig, Philipp, and Christian J. Schuler. "Entropy search for information-efficient global optimization." *Journal of Machine Learning Research* 13, no. Jun (2012): 1809-1837.