SAND2022-4929C

Exploring Bayesian Optimal Experimental Designs for High-dimensional Combustion Systems

Jim Oreluk, Leonid Sheps, Habib N. Najm

SIAM UQ22

joreluk@sandia.gov
Sandia National Laboratories, Livermore, CA, USA

Sandia National Laboratories

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

Outline

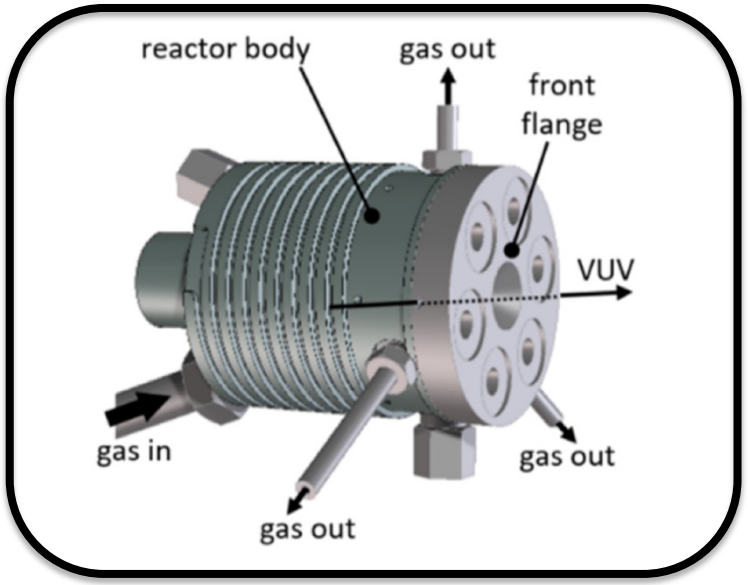
- Overview of a high-dimensional combustion system
  - time-of-flight mass spectrometry
- Bayesian optimal experimental design
- Challenges associated with high-dimensional models
  - finding low-dimensional representations
- Example
- Conclusion & Future work

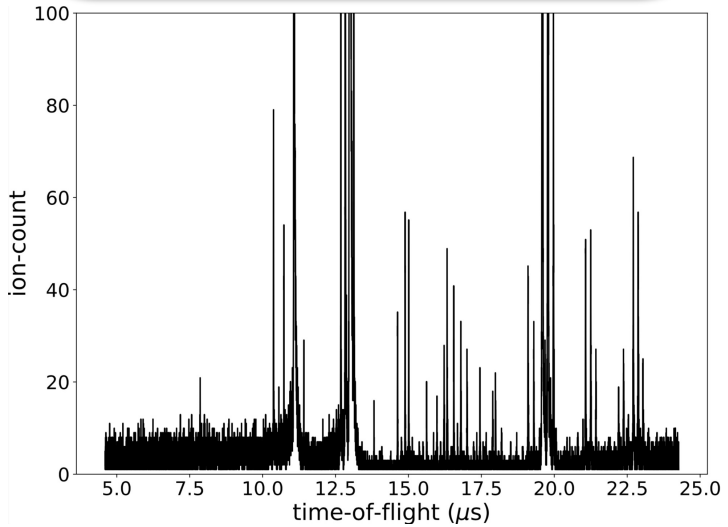
High-pressure photolysis reactor

- Mass spectrometry is an analytical method to identify compounds in an unknown sample

High-pressure photolysis reactor experiment

- Premixed gaseous mixture flows into a constant pressure reactor
- Photolysis laser fires instantaneously irradiating the gas mixture
  - Chemical precursor broken down initiating a reaction
- Synchrotron tunable vacuum-ultraviolet (VUV) photoionization mass spectrometry
  - Measurement of time-of-flight mass spectrum taken across kinetic times and VUV energies





Time-of-flight mass spectrum at a fixed VUV energy and kinetic time

L. Sheps, I. Antonov, K. Au. Sensitive mass spectrometer for time-resolved gas-phase chemistry studies at high pressures. *The Journal of Physical Chemistry A* 123.50 (2019) 10804-10814.

Modeling the high-pressure photolysis reactor

Data model:

$$z(\boldsymbol{d}, \boldsymbol{x}) = \xi(\boldsymbol{d}, \boldsymbol{x}) + \epsilon(\boldsymbol{x})$$

$$z(\boldsymbol{d}, \boldsymbol{x}) = f(\boldsymbol{\theta}, \boldsymbol{d}, \boldsymbol{x}) + \delta(\boldsymbol{x}) + \epsilon(\boldsymbol{x})$$

$$\boldsymbol{x} = [\tau, t, E]$$

$$\delta(\boldsymbol{x}) \sim GP(\mu_\delta(\boldsymbol{x}), \Sigma_\delta(\boldsymbol{x}, \boldsymbol{x}')), \; \epsilon(\boldsymbol{x}) \sim \mathcal{N}(0, s(\boldsymbol{x})^2)$$

$\boldsymbol{d}$: design conditions

$\boldsymbol{\theta}$: model parameters

$\boldsymbol{x}$: spatial/temporal coordinates

$z(\boldsymbol{d}, \boldsymbol{x})$: ion-count data

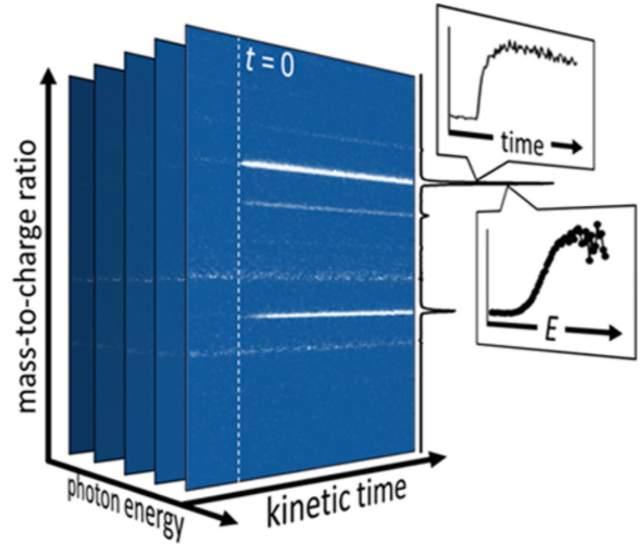
$\xi(\boldsymbol{d}, \boldsymbol{x})$: true physical process

$f(\boldsymbol{\theta}, \boldsymbol{d}, \boldsymbol{x})$: physics and instrument model

$\delta(\boldsymbol{x})$: model error

$\epsilon(\boldsymbol{x})$: observation noise

- **Physics model**
  - Zero-dimensional reactor
  - Photolysis laser model
- **Chemical model**
  - C0-C3 chemical mechanism
  - 171 species / 1143 reactions
- **Instrument model**
  - Maps concentrations to ion counts
  - Peaks idealized as Gaussian distributions



Visualization of the measurement tensor, $z(\boldsymbol{d}, \boldsymbol{x})$

Modeling the high-pressure photolysis reactor

Data model:

$$z(\boldsymbol{d}, \boldsymbol{x}) = \xi(\boldsymbol{d}, \boldsymbol{x}) + \epsilon(\boldsymbol{x})$$

$$z(\boldsymbol{d}, \boldsymbol{x}) = f(\boldsymbol{\theta}, \boldsymbol{d}, \boldsymbol{x}) + \delta(\boldsymbol{x}) + \epsilon(\boldsymbol{x})$$

$$\boldsymbol{x} = [\tau, t, E]$$

$$\delta(\boldsymbol{x}) \sim GP(\mu_\delta(\boldsymbol{x}), \Sigma_\delta(\boldsymbol{x}, \boldsymbol{x}')), \; \epsilon(\boldsymbol{x}) \sim \mathcal{N}(0, s(\boldsymbol{x})^2)$$

$\boldsymbol{d}$: design conditions

$\boldsymbol{\theta}$: model parameters

$\boldsymbol{x}$: spatial/temporal coordinates

$z(\boldsymbol{d}, \boldsymbol{x})$: ion-count data

$\xi(\boldsymbol{d}, \boldsymbol{x})$: true physical process

$f(\boldsymbol{\theta}, \boldsymbol{d}, \boldsymbol{x})$: physics and instrument model

$\delta(\boldsymbol{x})$: model error

$\epsilon(\boldsymbol{x})$: observation noise

- **Physics model**
  - Zero-dimensional reactor
  - Photolysis laser model
- **Chemical model**
  - C0-C3 chemical mechanism
  - 171 species / 1143 reactions
- **Instrument model**
  - Maps concentrations to ion counts
  - Peaks idealized as Gaussian distributions

High-dimensional output tensor

$$z(\boldsymbol{d}, \boldsymbol{x}) \in \mathbb{R}^{25000 \times 240 \times 85}$$

$$(25000 \times 240 \times 85) = 5.1 \times 10^8 \text{ elements}$$

High-dimensional parameter space

$$\boldsymbol{\theta} \in \mathbb{R}^{1151}$$

Motivation

- Identify key operating conditions to study specific chemical rate constant measurements (model parameters)
  - reactor temperature, reactor pressure, mixture composition

Why is this important?

- Operation of the real experiment is costly and laborious
  - Initial setup time for the apparatus
  - Daily calibration experiments necessary before any measurements are made
- Limited time to run experiments
  - Advanced Light Source, Lawrence Berkeley National Laboratory

Bayesian optimal experimental design

Objective

Find a set of experimental conditions that maximizes the expected utility

- Goal of the experiment is to learn chemical rate constant measurements of interest

$$d^* = \arg\max_{d \in \mathcal{D}} U(d)$$

where,

$$U(d) = \int_{y \in \mathcal{Y}} \int_{\theta \in \Theta} u(y, d, \theta) p(\theta, y|d) \mathrm{d}\theta \mathrm{d}y$$

$$= \int_{y \in \mathcal{Y}} \int_{\theta \in \Theta} u(y, d, \theta) p(\theta|y, d) p(y|d) \mathrm{d}\theta \mathrm{d}y$$

Notation

d : design conditions

θ : model parameters

y : data

Choice of utility function

Select a utility function to satisfy a particular modeling goal

- Parameter inference
  - Information gain of an experiment is closely related to minimizing the parameter uncertainty
  - Kullback-Leibler divergence can be used to measure what we can learn from the experimental data

$$u(y, d, \theta) = D_{\mathrm{KL}}(p(\theta|y, d)||p(\theta)) = \int p(\theta|y, d) \log \left[\frac{p(\theta|y, d)}{p(\theta)} \right] \mathrm{d}\theta$$

Choice of utility function

Select a utility function to satisfy a particular modeling goal

- Parameter inference
  - Information gain of an experiment is closely related to minimizing the parameter uncertainty
  - Kullback-Leibler divergence can be used to measure what we can learn from the experimental data

$$u(y, d, \theta) = D_{\mathrm{KL}}(p(\theta|y, d)||p(\theta)) = \int p(\theta|y, d) \log \left[\frac{p(\theta|y, d)}{p(\theta)} \right] \mathrm{d}\theta$$

$$U(d) = \int_{y \in \mathcal{Y}} \int_{\theta \in \Theta} \log \left[\frac{p(\theta|y, d)}{p(\theta)} \right] p(\theta|y, d) \mathrm{d}\theta p(y|d) \mathrm{d}y$$

How do we compute $U(d)$?

Approximating the expected utility

Numerical approximation:

$$U(d) \approx \frac{1}{N} \sum_{i=0}^{N} \left[\log p(y^{(i)}|\theta^{(i)}, d) - \log \underbrace{p(y^{(i)}|d)}_{???} \right] \quad \text{where, } \theta^{(i)} \sim p(\theta), \; y^{(i)} \sim p(y|\theta^{(i)}, d)$$

Several approaches to estimate the marginal likelihood

- **Monte Carlo sampling**
- Laplace approximation
- Importance sampling
- Variational methods

$$p(y^{(i)}|d) = \int p(y^{(i)}|\theta, d) p(\theta) \mathrm{d}\theta$$
$$\approx \frac{1}{M} \sum_{j=0}^{M} p(y^{(i)}|\theta^{(j)}, d), \quad \text{where, } \theta^{(j)} \sim p(\theta)$$

N. Friel, J. Wyse, Estimating the evidence–a review, *Statistica Neerlandica* 66.3 (2012) 288-308.
A. Gelman, X. Meng, Simulating normalizing constants: From importance sampling to bridge sampling to path sampling, *Statistical science* (1998) 163-185.

Approximating the expected utility

Numerical approximation:

$$U(d) \approx \frac{1}{N}\sum_{i=0}^{N}\left[\log p(y^{(i)}|\theta^{(i)},d) - \log \underbrace{p(y^{(i)}|d)}_{???}\right] \quad \text{where, } \theta^{(i)} \sim p(\theta),\ y^{(i)} \sim p(y|\theta^{(i)},d)$$

Nested Monte Carlo

$$U(d) \approx \frac{1}{N}\sum_{i=0}^{N}\left[\log p(y^{(i)}|\theta^{(i)},d) - \frac{1}{M}\sum_{j=0}^{M}\log p(y^{(i)}|\theta^{(j)},d)\right]$$

T. Rainforth et al, On nesting monte carlo estimators, *International Conference on Machine Learning*. PMLR, 2018.
K.J. Ryan, Estimating expected information gains for experimental designs with application to the random fatigue-limit model, *Journal of Computational and Graphical Statistics* 12 (2003) 585–603.

Maximizing the expected utility, $U(d)$

$$d^* = \arg\max_{d \in \mathcal{D}} \; U(d)$$

Bayesian Optimization

- Construct a Gaussian process model of the unknown objective function $U(d)$

$$U(d) \sim \mathcal{N}(\mu(d), K(d, d'))$$

- Use an *acquisition function* $\alpha(d)$ to select new samples, trading-off between exploration and exploitation

- Select next sample as: $$d_t = \arg\max_{d \in \mathcal{D}} \alpha_t(d)$$

- Evaluate utility function at $U(d_t)$

N. Srinivas, A. Krause, S. Kakade, & M. Seeger. Gaussian process optimization in the bandit setting: No regret and experimental design. *arXiv preprint arXiv:0912.3995* (2009).

Maximizing the expected utility, $U(d)$

$$d^* = \arg\max_{d \in \mathcal{D}} \; U(d)$$

Bayesian Optimization

- Construct a Gaussian process model of the unknown objective function $U(d)$

$$U(d) \sim \mathcal{N}(\mu(d), K(d, d'))$$

- Use an *acquisition function* $\alpha(d)$ to select new samples, trading-off between exploration and exploitation

---

Gaussian Process Upper Confidence Bound

$$\alpha_t(d) = \mu_{t-1}(d) + \sqrt{\beta_t}\sigma_{t-1}(d)$$

where t is the optimization iteration and $\sigma_{t-1}(d) = \sqrt{K(d, d)}$.

- Exploits regions with a high mean and explores regions of high uncertainty

N. Srinivas, A. Krause, S. Kakade, & M. Seeger. Gaussian process optimization in the bandit setting: No regret and experimental design. *arXiv preprint arXiv:0912.3995* (2009).

Challenges

Computational limitations

- High-fidelity physics-based simulations can be expensive to evaluate
- At most, NM evaluations needed to estimate $U(d)$, assuming no reuse of data
- Memory limitations storing a $(N \times J)$ sparse matrix, with $J = 5.1 \times 10^8$

- Constructing a surrogate model addresses the computational cost
  - number of model outputs remains problematic
- Can we find a low-dimensional representation of the high-dimensional model output?

Reducing output dimensionality

Goal: Map model output from high-dimensional space to a lower-dimensional space while minimizing loss of information

Truncated SVD

At a fixed design d,

- Draw n sample of $\boldsymbol{\theta}^{(i)} \sim p(\boldsymbol{\theta})$
- Evaluate model $f(\boldsymbol{\theta}, \boldsymbol{d}, \boldsymbol{x}) + \delta(\boldsymbol{x}) + \epsilon(x)$
- Construct output matrix $\boldsymbol{Z} = \boldsymbol{U}\boldsymbol{S}\boldsymbol{V}^T$, where $\boldsymbol{Z} \in \mathbb{R}^{n \times J}, J = 5.1 \times 10^8$
- Retain only top K singular values of $\boldsymbol{S}$
- Low-rank approximation: $\boldsymbol{Z}_K = \boldsymbol{U}_K \boldsymbol{S}_K \boldsymbol{V}_K^T$

Transformation:

$$q(\boldsymbol{\theta}, \boldsymbol{d}, \boldsymbol{x}) = z(\boldsymbol{d}, \boldsymbol{x})\boldsymbol{V}_K$$

$$\underbrace{q(\boldsymbol{\theta}, \boldsymbol{d}, \boldsymbol{x})}_{(1\times K)} = \underbrace{[f(\boldsymbol{\theta}, \boldsymbol{d}, \boldsymbol{x}) + \delta(\boldsymbol{x}) + \epsilon(\boldsymbol{x}))]}_{(1\times J)} \underbrace{\boldsymbol{V}_K}_{(J\times K)}$$

Reducing output dimensionality

Construct K surrogate models, one for each of the low-dimensional QOIs

$$g_k(\boldsymbol{\theta}) \approx q_k(\boldsymbol{\theta}, \boldsymbol{d}, \boldsymbol{x}), \text{ for } k = 1, \ldots, K$$

How should we represent the likelihood in the low-dimensional space?

Reducing output dimensionality

Construct K surrogate models, one for each of the low-dimensional QOIs

$$g_k(\boldsymbol{\theta}) \approx q_k(\boldsymbol{\theta}, \boldsymbol{d}, \boldsymbol{x}), \text{ for } k = 1, \ldots, K$$

Recall,

$$z(\boldsymbol{d}, \boldsymbol{x}) = \xi(\boldsymbol{d}, \boldsymbol{x}) + \epsilon(\boldsymbol{x})$$
$$z(\boldsymbol{d}, \boldsymbol{x}) = f(\boldsymbol{\theta}, \boldsymbol{d}, \boldsymbol{x}) + \delta(\boldsymbol{x}) + \epsilon(\boldsymbol{x})$$

$$\delta(\boldsymbol{x}) \sim GP(\mu_\delta(\boldsymbol{x}), \Sigma_\delta(\boldsymbol{x}, \boldsymbol{x}')), \epsilon(\boldsymbol{x}) \sim \mathcal{N}(0, s(\boldsymbol{x})^2)$$

Therefore,

$$\mu = \mathbb{E}[z] = f(\boldsymbol{\theta}, \boldsymbol{d}, \boldsymbol{x}) + \mu_\delta(\boldsymbol{x})$$
$$\Sigma = \text{Var}[z] = \Sigma_\delta(\boldsymbol{x}, \boldsymbol{x}') + \text{diag}(s(\boldsymbol{x})^2)$$

Given a linear transformation of z,

$$\mu_q = \mathbb{E}[q] = \mu \boldsymbol{V}_k$$
$$\Sigma_q = \text{Var}[q] = \boldsymbol{V}_k^T \Sigma \boldsymbol{V}_k$$

Evaluating $U(d)$

$$U(d) \approx \frac{1}{N}\sum_{i=1}^{N}\left[\log p(z^{(i)}|\theta^{(i)},d) - \frac{1}{M}\sum_{j=1}^{M}\log p(z^{(i)}|\theta^{(j)},d)\right]$$

where $z^{(i)} \sim p(z|\theta^{(i)},d)$ and $p(z^{(i)}|\theta^{(i)},d) \sim \mathcal{N}(\mu,\Sigma)$.

Rewriting the data,

$$U(d) \approx \frac{1}{N}\sum_{i=1}^{N}\left[\log p(q^{(i)}\boldsymbol{V}_K^T|\theta^{(i)},d) - \frac{1}{M}\sum_{j=1}^{M}\log p(q^{(i)}\boldsymbol{V}_K^T|\theta^{(j)},d)\right]$$

Taking a linear combination of the data,

$$U(d) \approx \frac{1}{N}\sum_{i=1}^{N}\left[\log p(q^{(i)}|\theta^{(i)},d) - \frac{1}{M}\sum_{j=1}^{M}\log p(q^{(i)}|\theta^{(j)},d)\right]$$

where $q^{(i)} \sim p(q|\theta^{(i)},d)$ and $p(q^{(i)}|\theta^{(i)},d) \sim \mathcal{N}(\mu\boldsymbol{V}_K, \boldsymbol{V}_K^T\Sigma\boldsymbol{V}_K)$.

Example: Simplified reactor model

Original data model,

$$z(\boldsymbol{d}, \boldsymbol{x}) = \xi(\boldsymbol{d}, \boldsymbol{x}) + \epsilon(\boldsymbol{x})$$
$$z(\boldsymbol{d}, \boldsymbol{x}) = f(\boldsymbol{\theta}, \boldsymbol{d}, \boldsymbol{x}) + \delta(\boldsymbol{x}) + \epsilon(\boldsymbol{x})$$

$$\boldsymbol{x} = [\tau, t, E]$$

$$\delta(\boldsymbol{x}) \sim GP(\mu_\delta(\boldsymbol{x}), \Sigma_\delta(\boldsymbol{x}, \boldsymbol{x}')), \ \epsilon(\boldsymbol{x}) \sim \mathcal{N}(0, s(\boldsymbol{x})^2)$$

$z(\boldsymbol{d}, \boldsymbol{x})$: ion-count data

$\xi(\boldsymbol{d}, \boldsymbol{x})$: true physical process

$f(\boldsymbol{\theta}, \boldsymbol{d}, \boldsymbol{x})$: physics and instrument model

$\delta(\boldsymbol{x})$: model error

$\epsilon(\boldsymbol{x})$: observation noise

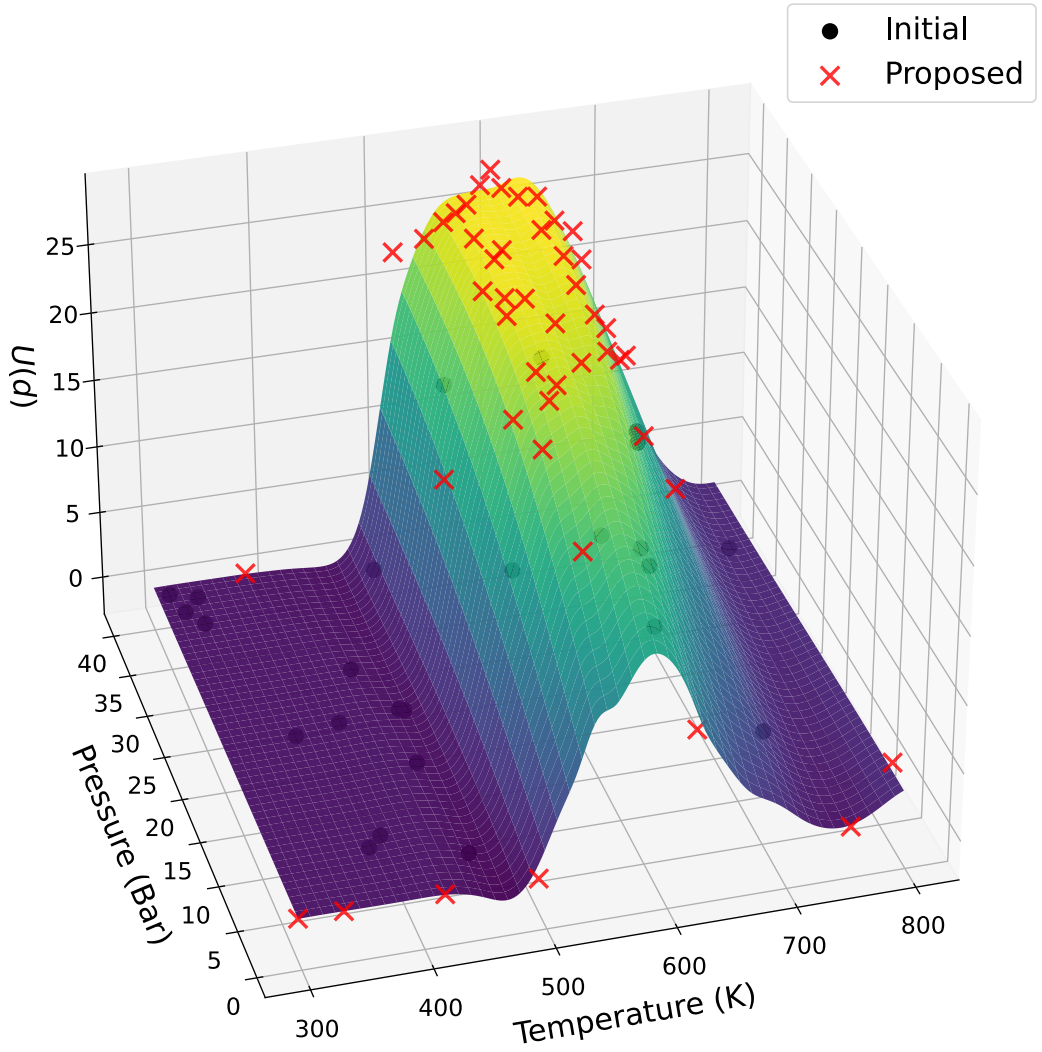
Simplifying assumption:

Only a small subset of the model parameters are considered uncertain

- 4 / 1143 reaction rates uncertain, all other reactions are at their nominal values

1) $\mathrm{O} + \mathrm{H_2} \rightarrow \mathrm{OH} + \mathrm{H}$
2) $\mathrm{C_3H_8} + \mathrm{H} \rightarrow \mathrm{CH_3CH_2CH_2} + \mathrm{H_2}$
3) $\mathrm{O_2} + \mathrm{CH_3CH_2CH_2} \rightarrow \mathrm{OH} + \mathrm{C{-}CH_2OCH(CH_3)}$
4) $\mathrm{CH_3CH(OOH)CH_2} \rightarrow \mathrm{OH} + \mathrm{C{-}CH_2OCH(CH_3)}$

Results



Solid surface is the mean function from a Gaussian process model representing $U(d)$. Evaluations of the utility function are shown as black points or red crosses.

Fixed design parameters: $\chi_{C_3H_8} = 8.3 \times 10^{-7}$
$\chi_{O_2} = 2.5 \times 10^{-2}$
$\chi_{pre} = 1.9 \times 10^{-4}$

# of utility samples: $N = 1 \times 10^4, M = 1 \times 10^4$

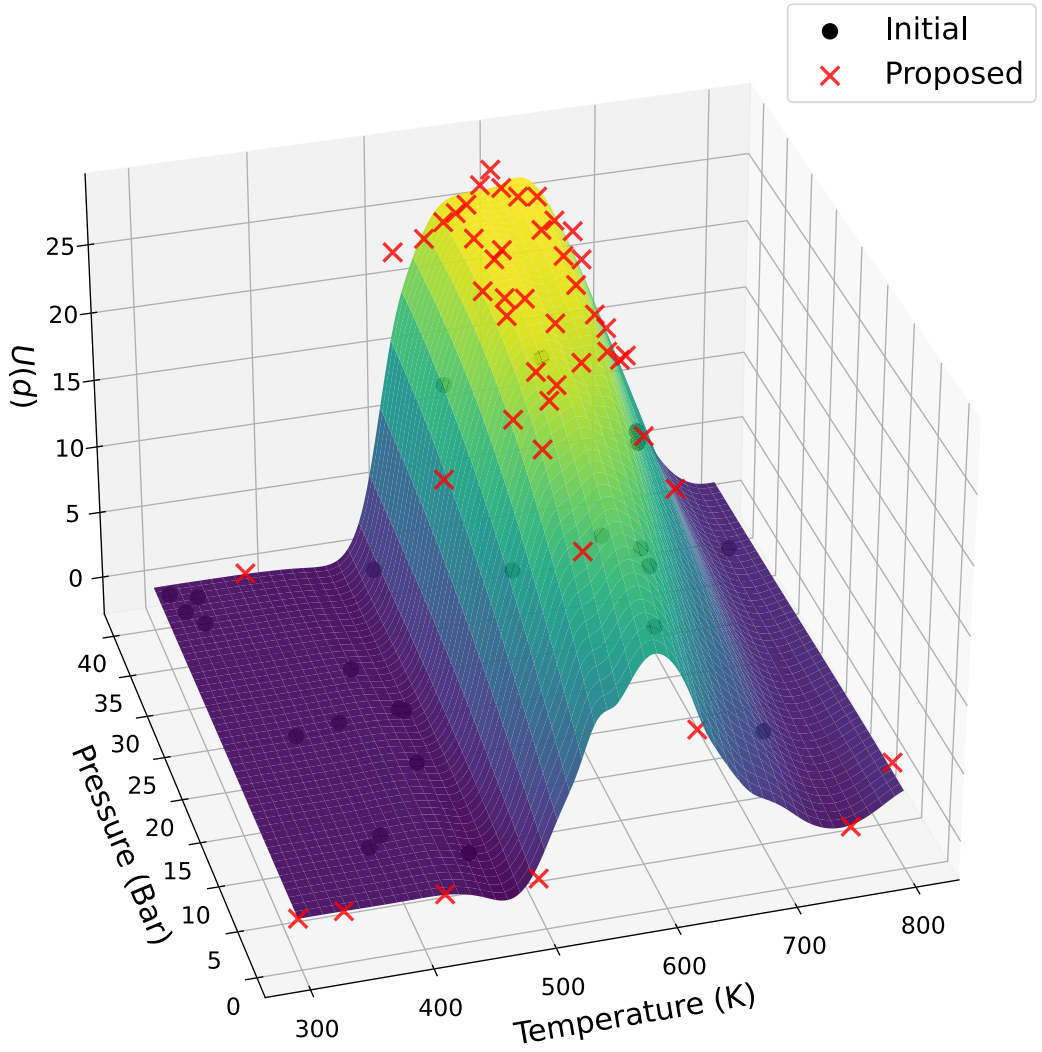
Optimization method: Bayesian Optimization

Acquisition function: UCB, with $\sqrt{\beta_t} = 2.5$

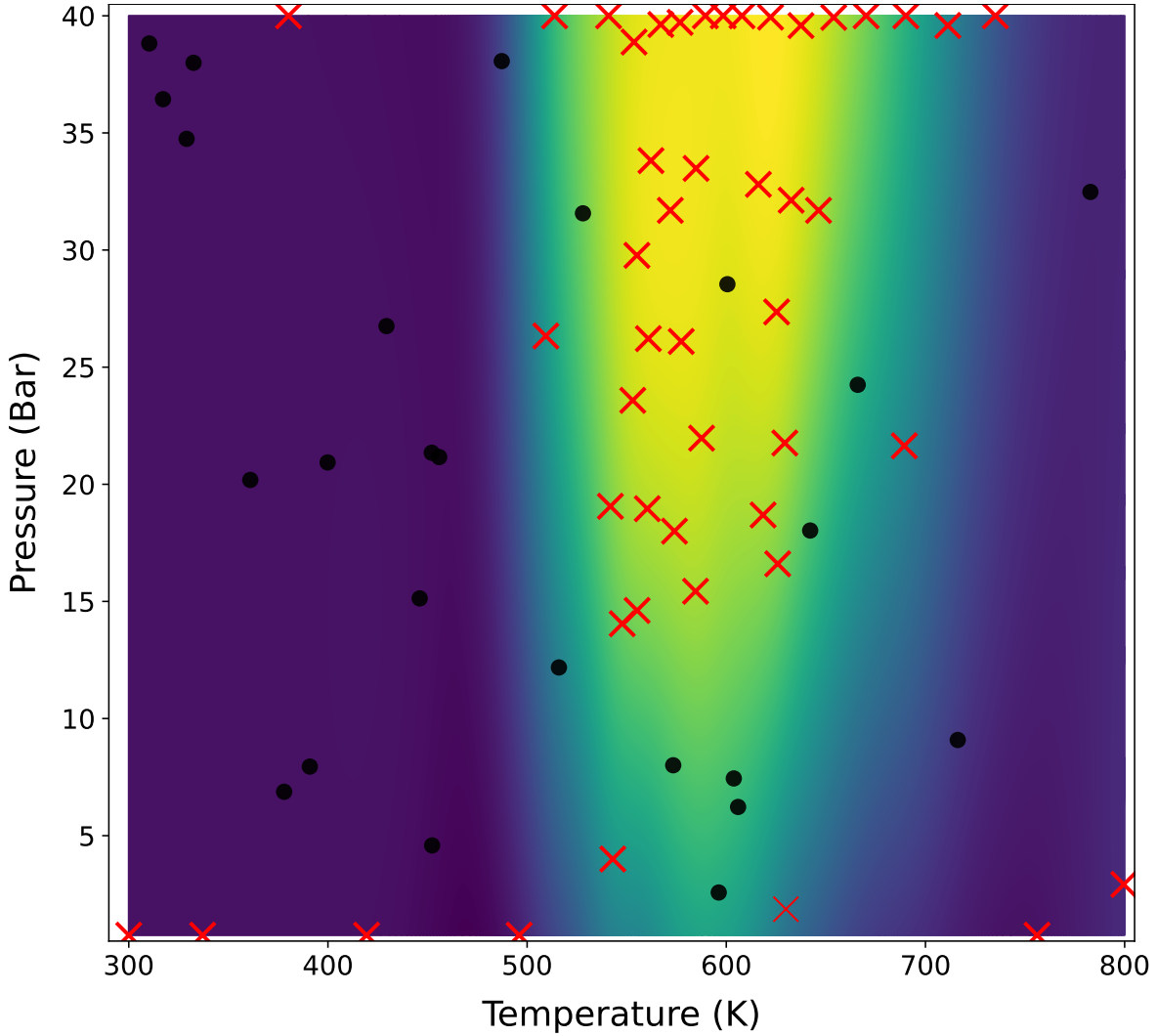
Dimension reduction: $K = 20$ components

25 Latin-Hypercube samples (black points)
50 proposal samples (red crosses)

Results

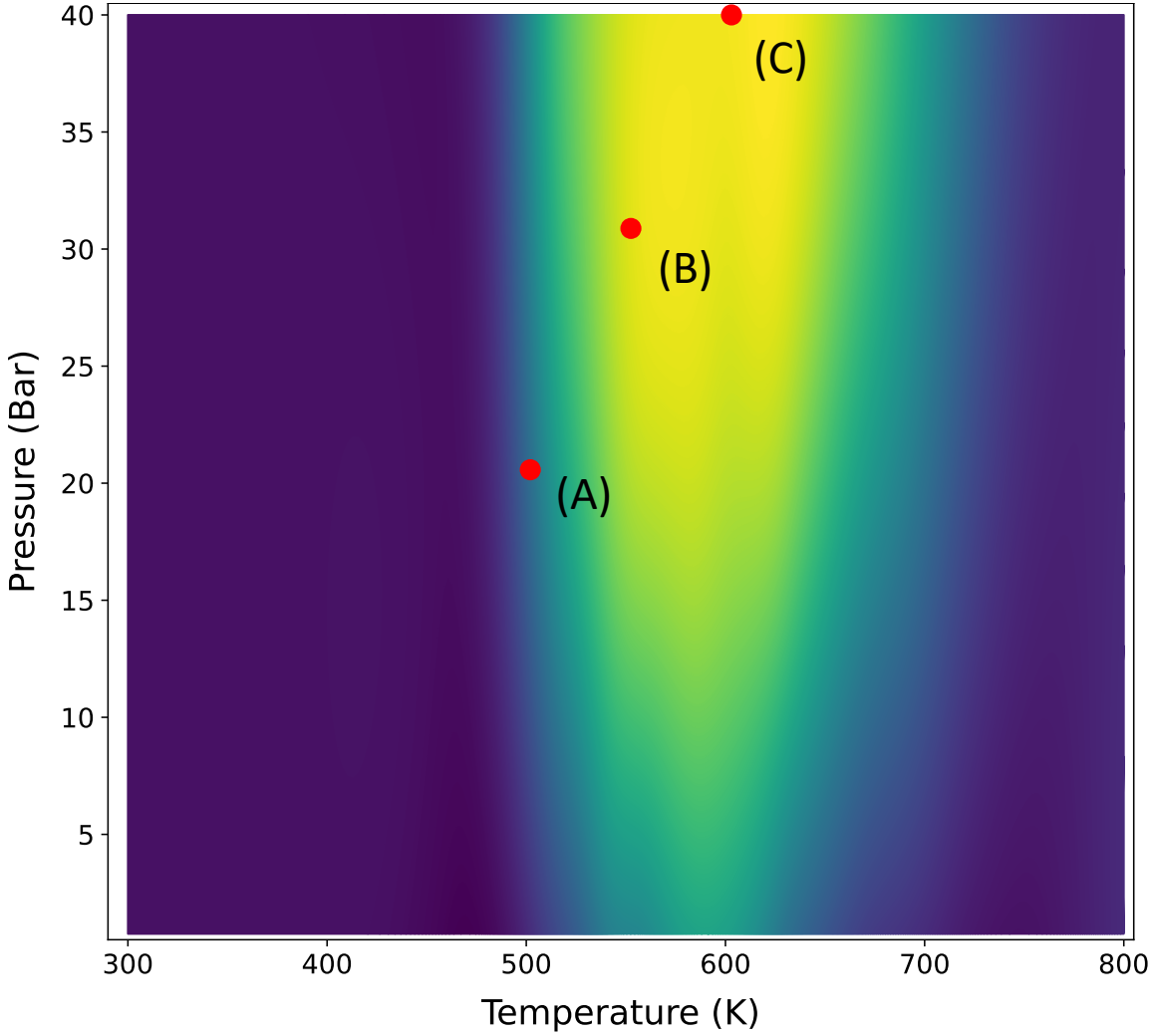


Solid surface is the mean function from a Gaussian process model representing $U(d)$. Evaluations of the utility function are shown as black points or red crosses.

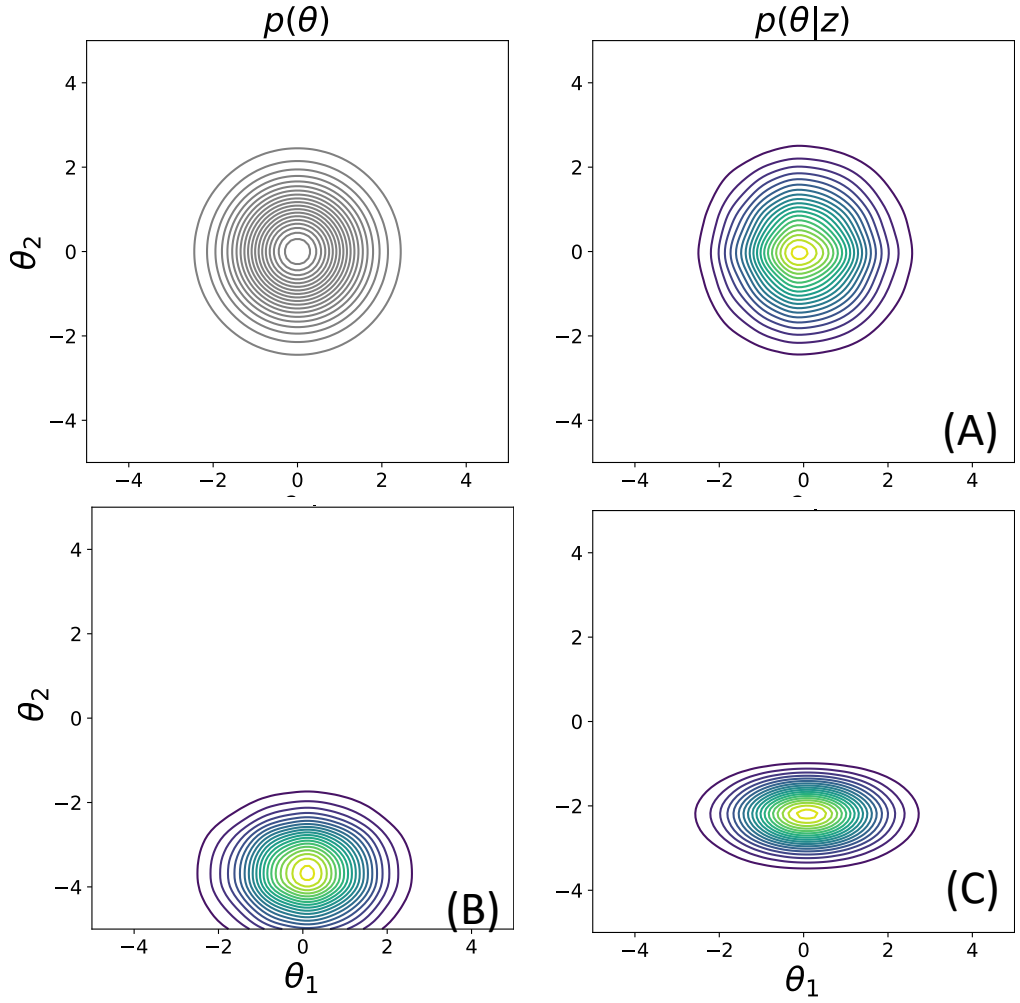


Initial proposals by the acquisition function were near the maximum, (T, p) = (598.4 K, 40 bar)

Results



Heatmap of the estimated utility function using a Gaussian process mean function. Data sets are generated at each of the design points (A, B, C) and ae used to infer the posterior density.



Bayesian inference was performed to observe change in the joint posterior density given plausible data sets at each of the design points.

Conclusion & Future Work

High-dimensional combustion systems can be challenging for OED studies

- Demonstrated feasibility of BOED on the high-pressure photolysis reactor
- Bayesian optimization shown to be efficient optimizing noisy utility functions
- Low-dimensional representation of the model output gives significant and necessary computational savings

Future Work

- Relaxing assumptions on number of uncertain model parameters
  - Preliminary work shows set of influential model parameters can greatly change across the design space (in total $\sim 40 - 100$ relevant model parameters)

- Perform several iterations of the OED loop, collecting data at the optimal design
  - Improve model error, allowing design dependence
  - Compare performance of experiments at optimal designs to random designs

Acknowledgements

We would like to thank Oscar Diaz-Ibarra, Kyungjoo Kim, Arun Hegde, Cosmin Safta, and Khachik Sargsyan for helpful conversations about this work.

This work was supported by the US Department of Energy (DOE), Office of Basic Energy Sciences (BES) Division of Chemical Sciences, Geosciences, and Biosciences. Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC., a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA-0003525. The work of the HP-PIMS experiment is supported through the Argonne-Sandia Consortium on High-Pressure Combustion Chemistry.

Sandia National Laboratories

Additional Slides

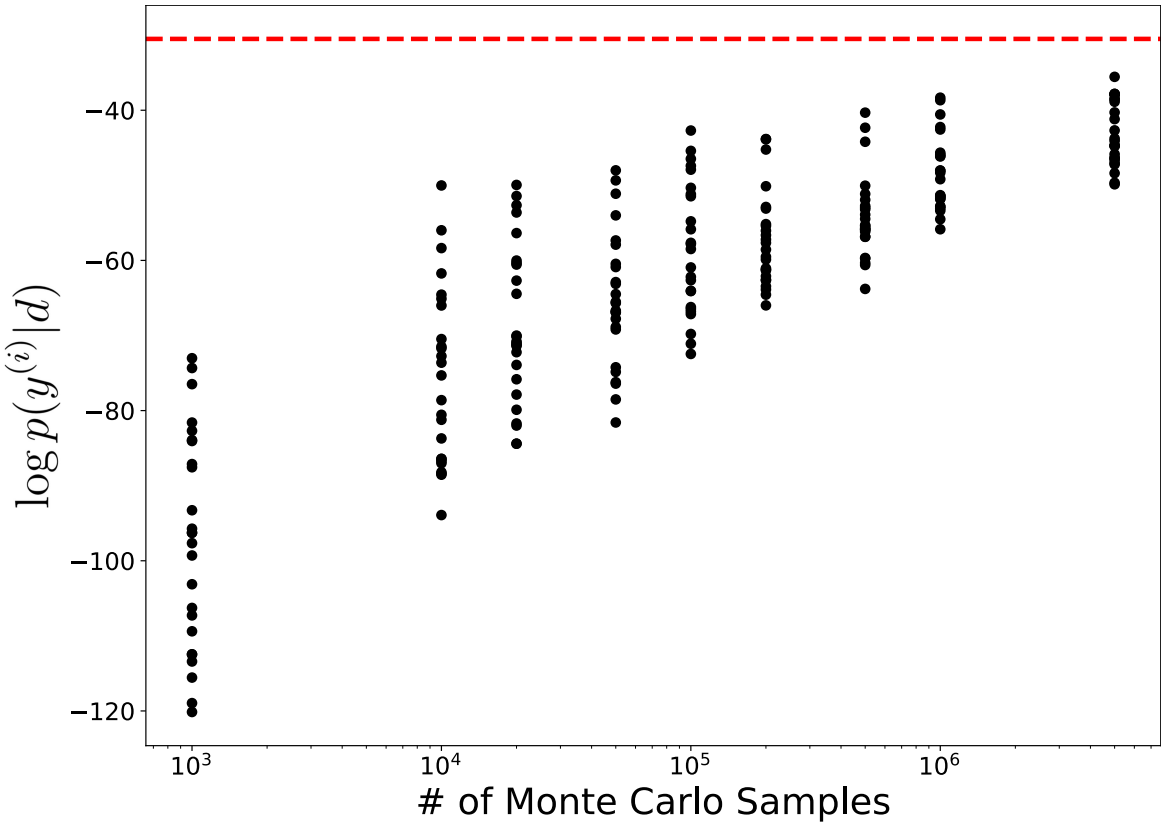
Estimating the marginal likelihood

Linear Model

$$
\begin{aligned}
y &= G(\theta) + \epsilon \\
G(\theta) &= A\theta \qquad\qquad y \in \mathbb{R}^{15} \\
\theta &\sim \mathcal{N}(\mu_0, \Sigma_0) \qquad \theta \in \mathbb{R}^{25} \\
\epsilon &\sim \mathcal{N}(0, \Sigma_\epsilon)
\end{aligned}
$$

Monte Carlo estimation

$$
\begin{aligned}
p(y^{(i)}|d) &= \int p(y^{(i)}|\theta, d)p(\theta)\mathrm{d}\theta \\
&\approx \frac{1}{M}\sum_{j=0}^{M} p(y^{(i)}|\theta^{(j)}, d), \quad \text{where, } \theta^{(j)} \sim p(\theta)
\end{aligned}
$$



Monte Carlo estimate of the log marginal likelihood converges to the true value, shown as a red dashed line, as number of samples goes to infinity.

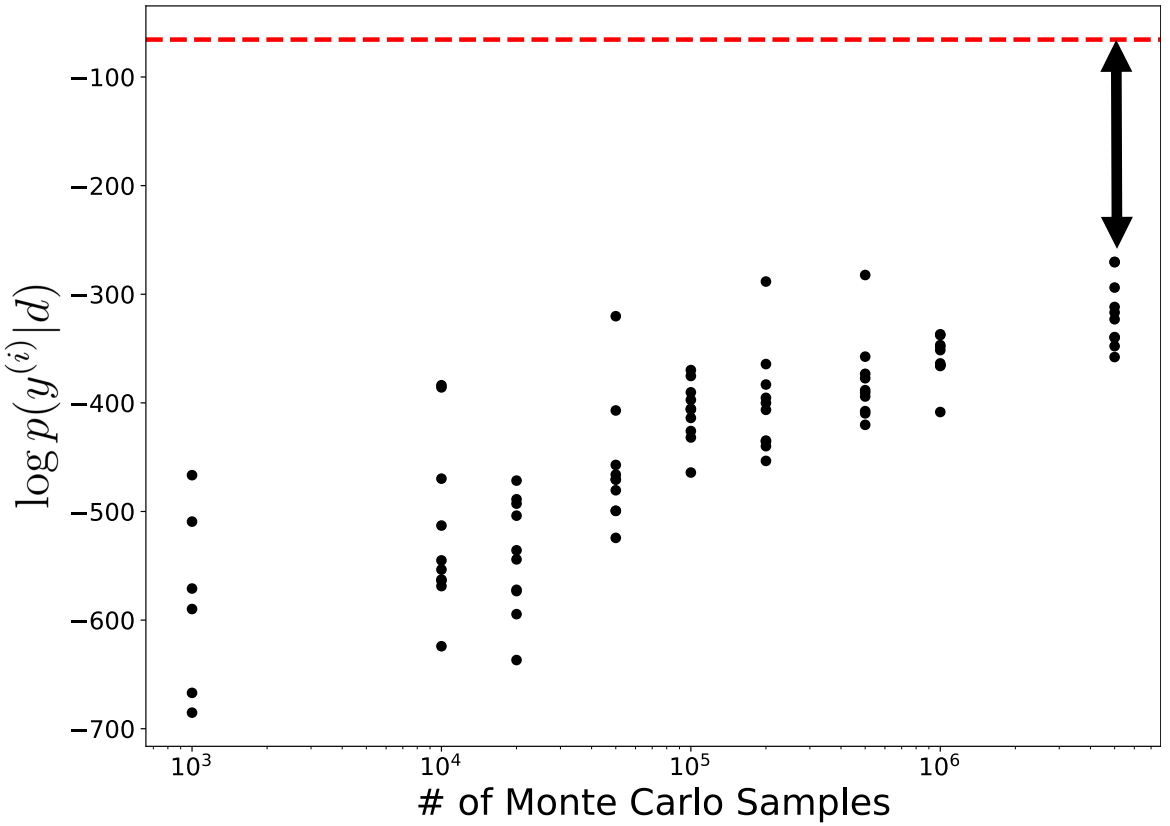
Estimating the marginal likelihood

Linear Model

$$y = G(\theta) + \epsilon$$
$$G(\theta) = A\theta$$
$$\theta \sim \mathcal{N}(\mu_0, \Sigma_0)$$
$$\epsilon \sim \mathcal{N}(0, \Sigma_\epsilon)$$

$$y \in \mathbb{R}^{30}$$
$$\theta \in \mathbb{R}^{50}$$

- As dimensionality increases, numerous samples are necessary to converge to the true marginal likelihood value



Significant error in the estimate of the log marginal likelihood as compared to the lower dimensional problem at a fixed number of samples.

Numerical approximation

$$u(y,d,\theta) = D_{KL}(p(\theta|y,d)||p(\theta)) = \int_\Theta p(\theta|y,d)\log\left[\frac{p(\theta|y,d)}{p(\theta)}\right]d\theta = u(y,d)$$

$$\begin{aligned} U(d) &= \int_\mathcal{Y}\int_\Theta u(y,d)p(\theta|y,d)d\theta p(y|d)dy \\ &= \int_\mathcal{Y} u(y,d)p(y|d)dy \\ &= \int_\mathcal{Y}\left(\int_\Theta p(\theta|y,d)\log\left[\frac{p(\theta|y,d)}{p(\theta)}\right]d\theta\right)p(y|d)dy \end{aligned}$$

$$\text{Using } p(\theta|y,d) = p(y|\theta,d)p(\theta)/p(y|d),$$

$$\begin{aligned} U(d) &= \int_\mathcal{Y}\int_\Theta \log\left[\frac{p(y|\theta,d)}{p(y|d)}\right]p(y|\theta,d)p(\theta)d\theta dy \\ &= \int_\mathcal{Y}\int_\Theta \left[\log p(y|\theta,d) - \log p(y|d)\right]p(y|\theta,d)p(\theta)d\theta dy \end{aligned}$$

Modeling the high-pressure photolysis reactor

Data model:

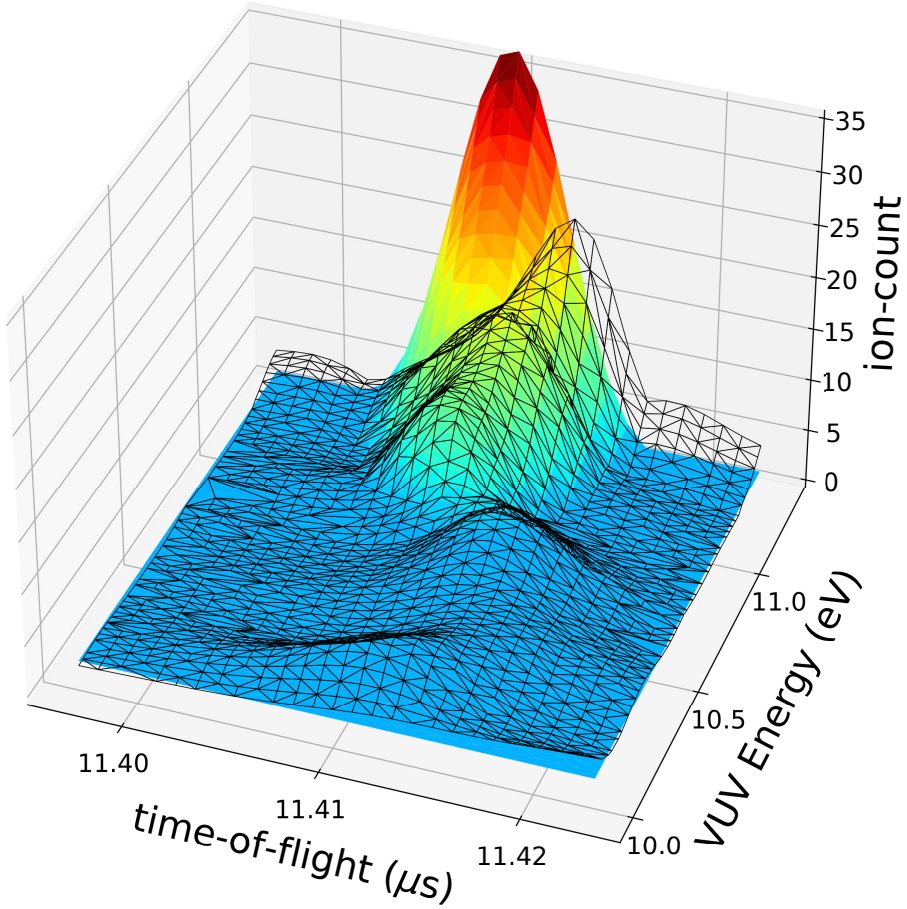
$$z(\boldsymbol{d}, \boldsymbol{x}) = \xi(\boldsymbol{d}, \boldsymbol{x}) + \epsilon(\boldsymbol{x})$$

$$z(\boldsymbol{d}, \boldsymbol{x}) = f(\boldsymbol{\theta}, \boldsymbol{d}, \boldsymbol{x}) + \delta(\boldsymbol{x}) + \epsilon(\boldsymbol{x})$$

$$\boldsymbol{x} = [\tau, t, E]$$

$$\delta(\boldsymbol{x}) \sim GP(\mu_\delta(\boldsymbol{x}), \Sigma_\delta(\boldsymbol{x}, \boldsymbol{x}')), \; \epsilon(\boldsymbol{x}) \sim \mathcal{N}(0, s(\boldsymbol{x})^2)$$

- **Physics model**
  - Zero-dimensional reactor
  - Photolysis laser model
- **Chemical model**
  - C0-C3 chemical mechanism
  - 171 species / 1143 reactions
- **Instrument model**
  - Maps concentrations to ion counts
  - ***Peaks idealized as Gaussian distributions***



Solid surface is the prediction of $f(\theta_{MAP}, d, x)$ for one of the peaks in the time-of-flight spectrum (H_2O_2).
Mesh surface shows the prediction with model error, $f(\theta_{MAP}, d, x) + \mu_\delta(x)$ which increases the fidelity of the predictive model.