PAUL SCHERRER INSTITUT

ETH zürich

Sebastian Heinekamp, A. Adelmann

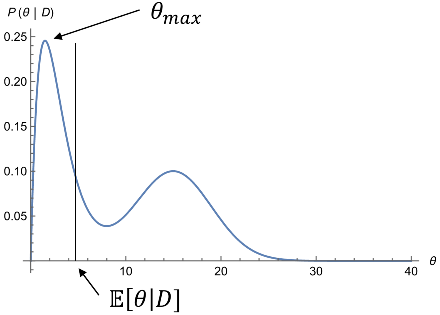
Bayesian Model Calibration

AWANOW, 11th August 2023

Contact: sebastian.heinekamp@psi.ch

Overview

Bayesian Inference vs Parameter Estimation



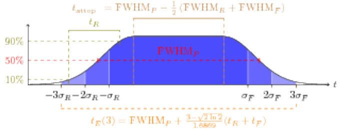
Bayes' Theorem

$$P(\theta|D) = \frac{P(D|\theta)P(\theta)}{P(D)}$$

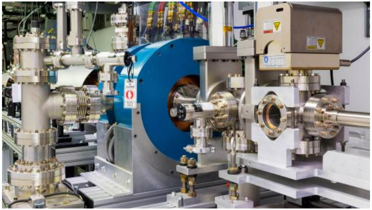
Goal:
Obtain full distribution
$P(\theta|D)$
not just $\mathbb{E}[\theta|D]$ or θ_{max}

Exact profile of laser in the AWA electron gun

Parameter θ e.g.:



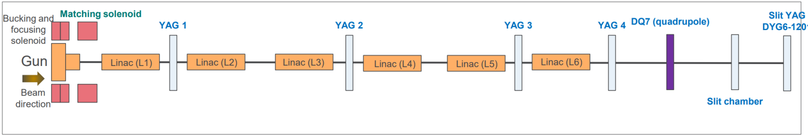
Profile of laser



Electron Gun

t_R : intensity build up from 10 - 90%

t_l : pulse length at 50% maximum



Application to AWA

Distribution e.g. of the time rise t_R and pulse length t_l can be used to

- improve simulation predictions
- calculate uncertainties
- refine the experimental setup

Bayesian Model Calibration [1]

Statistical Model

$$z_i = \zeta(\xi_i) + \epsilon_i = \rho\eta(x_i, \theta) + \delta(x_i) + \epsilon_i$$

$\zeta(\xi_i)$	Real world experiment
$\eta(x_i, \theta)$	simulation
$\delta(x_i)$	Code inadequacy
ρ	regression coefficient
ϵ	measurement error

θ	parameter
z_i	output
x_i	input

Quantities of Interest and Measurement Data

$$z_i = \zeta(\xi_i) + \epsilon_i = \rho\eta(x_i, \theta) + \delta(x_i) + \epsilon_i$$

Quantities of Interest

Experimental Inputs	Parameters	Measurement Outputs
x / ξ	θ	z
known; can be changed	distribution of interest	can be measured
ILS, IBF, Φ_{Gun}, SigXY, ...	trise t_r, pulse length t_l, ...	Energy, ΔE, σ_x, σ_y, ...

Types of Variables in Calibration Process

Measurement Data	Simulation Data (OPAL [2])
very expensive, little control	more control, less expensive
$N \approx 20 - 100$	$N \approx 1000 - 10000$

Surrogates in Bayesian Modelling

$$z_i = \zeta(\xi_i) + \epsilon_i = \rho\eta(x_i, \theta) + \delta(x_i) + \epsilon_i$$

Need functional representations or surrogates for

Simulation $\eta(x_i, \theta)$: from simulation data, independent of measurements

Code inadequacy $\delta(x_i)$: from simulation data and measurements

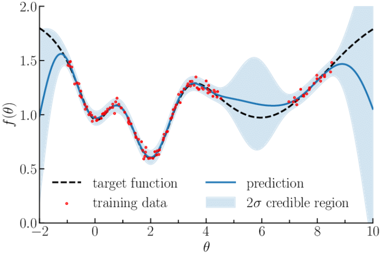
Error ϵ: $\sim \mathcal{N}(\mu, \sigma)$

Gaussian Processes [4]

Default surrogate:

$$f(\theta) \sim GP(m(\theta), cov(\theta, \theta'))$$

$m(\theta)$ - mean function
$cov(\theta, \theta')$ - covariance function



Gaussian Process with Variance [3]

GP vs other Surrogates

Gaussian Process

Advantage:

- Variance at each point
- Good for few data points
- little training (per-calculation) required

Challenge:

- Inversion of kernel matrix grows $\mathcal{O}(N^3)$

Other ML Surrogates

Advantage:

- Fast evaluation
- Already in some cases exist i.e. Bellotti et al. [5]
- All of ML tools available

Challenge:

- Variance not available

Alternatives to Plain GP

Neural Networks

- Learn variance with a second NN
- Bayesian Neural Net
- Mix using transfer learning

Advanced GP

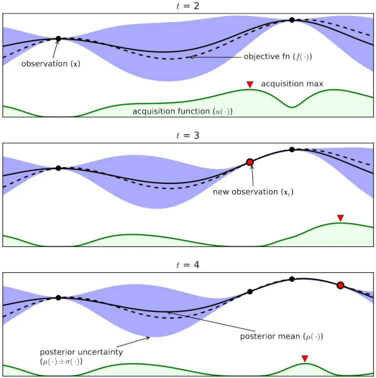
- Deep GP
- Bayes Committee
- Subset GP

These surrogates promise to allow for higher dimensional models and make use of more data

Towards Automated Experimental Design

Apart from the mentioned use in prediction improvement:

- a high dimensional model with inferred parameter distribution
- use Bayesian Optimization for experimental design



Bayesian Optimisation [6]

Experimental Training Data

Start of 2024:

Beam time at AWA

to collect experimental data

to train surrogate models

and calibrate OPAL for AWA electron gun

Goals

Increase the model dimensionality in close to real time calibration using alternative surrogates

Use a high-dimensional Bayesian Model of AWA for experimental design

Thank you for your attention.

Questions?

References

[1] Marc C. Kennedy and Anthony O'Hagan. Bayesian calibration of computer models. 63(3):425–464.

[2] Andreas Adelmann, Pedro Calvo, Matthias Frey, Achim Gsell, Uldis Locans, Christof Metzger-Kraus, Nicole Neveu, Chris Rogers, Steve Russell, Suzanne Sheehy, Jochem Snuverink, and Daniel Winklehner. OPAL a Versatile Tool for Charged Particle Accelerator Simulations.

[3] Florent Leclercq. Bayesian optimization for likelihood-free cosmological inference. 98(6):063511.

[4] Carl Edward Rasmussen and Christopher K. I. Williams. *Gaussian Processes for Machine Learning*. The MIT Press.

[5] Renato Bellotti, Romana Boiger, and Andreas Adelmann. Fast, Efficient and Flexible Particle Accelerator Optimisation Using Densely Connected and Invertible Neural Networks. 12(9):351.

[6] Bobak Shahriari, Kevin Swersky, Ziyu Wang, Ryan P. Adams, and Freitas N. Taking the Human Out of the Loop: A Review of Bayesian Optimization. 104(1):148–175.

Extra Slides

Posterior KOH Probability

Posterior probability using Gaussian Process

Bayes' Theorem

$$P(\theta|D) = \frac{P(D|\theta)P(\theta)}{P(D)}$$

Posterior $\propto$ Prior $\times$ Likelihood

$$p(\theta, \beta, \phi|d) \propto p(\theta)p(\phi)|V_d(\theta)|^{-\frac{1}{2}} exp[-\frac{1}{2}(d - m_d(\theta))^T V_d(\theta)(d - m_d(\theta))]$$