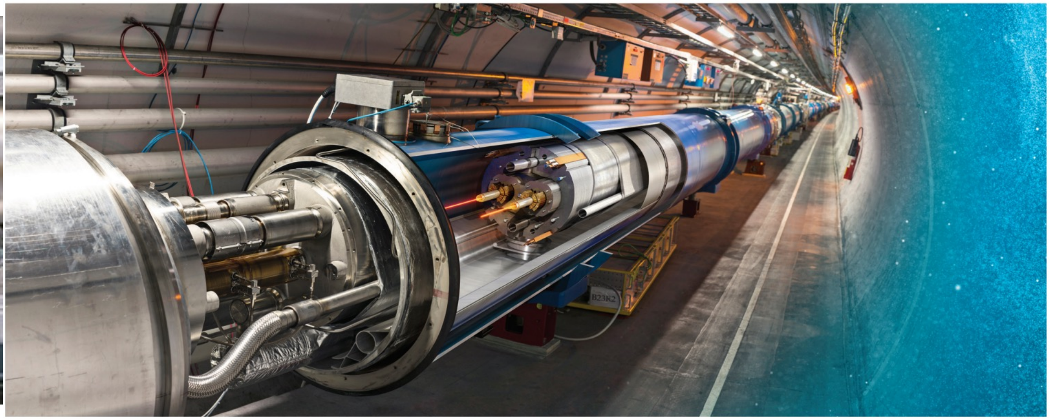
Bayesian optimization applications at the Argonne Wakefield Accelerator Facility

Juan Pablo Gonzalez-Aguilera*, Young-Kee Kim
UChicago

Ryan Roussel
SLAC

* jpga@uchicago.edu

Optimization and tuning in accelerators



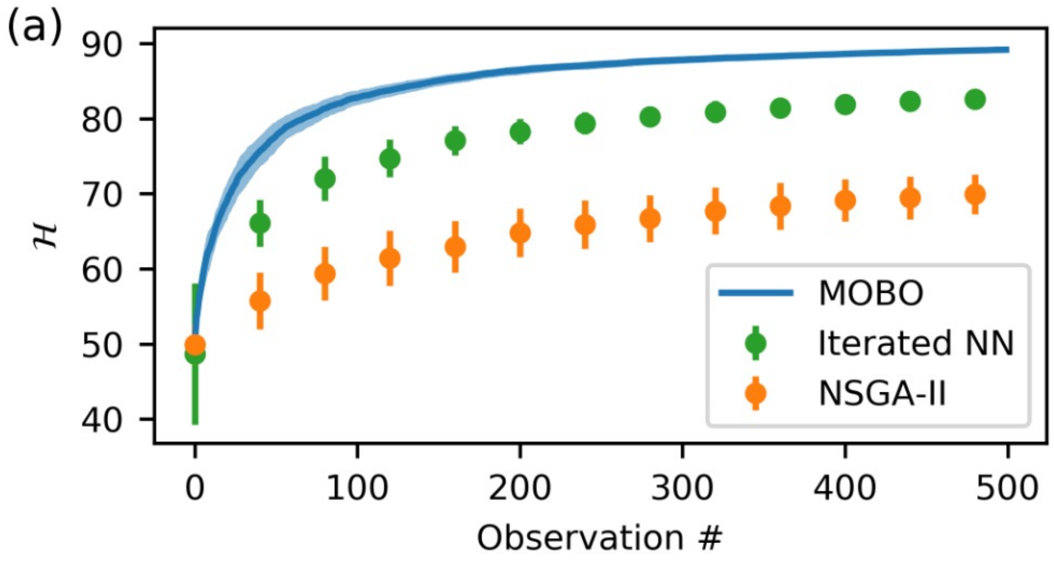
- Must meet various beam quality objectives
- Have many components and constraints
- Observations are limited

https://lcls.slac.stanford.edu/overview

https://home.cern/sites/home.web.cern.ch/files/2018-07/factsandfigures-en_0.pdf

Multi-objective optimization in accelerators

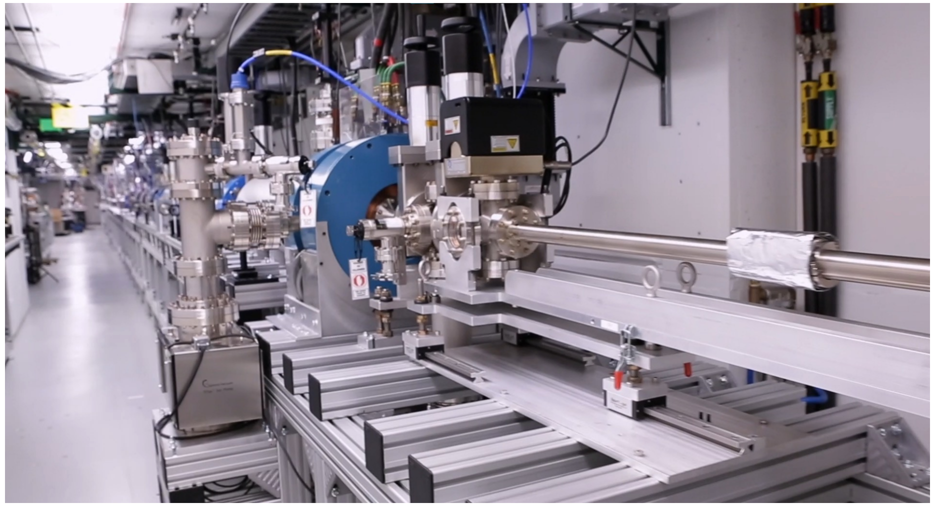
- Current methods require multiple observations / parallel computing
- Converge to Pareto front slowly when serialized (thousands of observations)
- Not ideal for online tuning

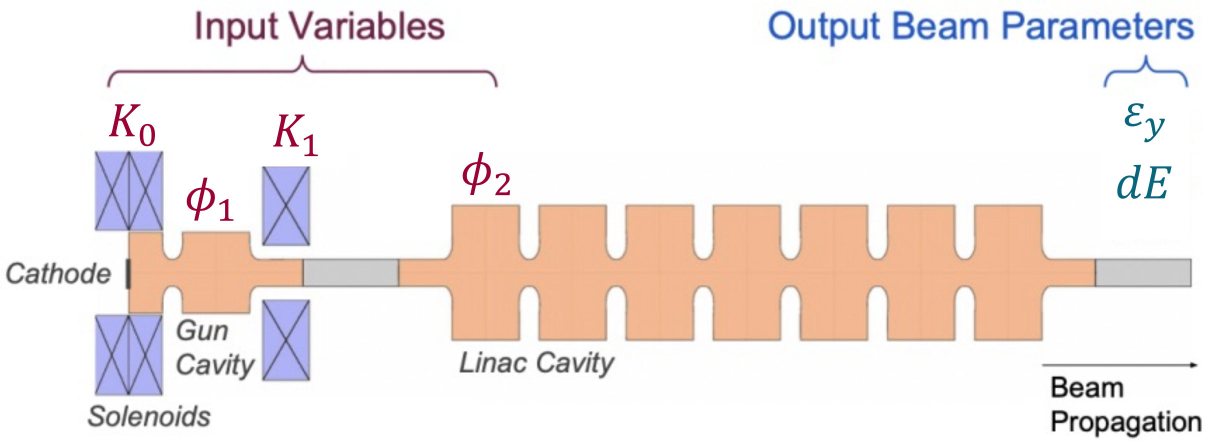


Multi-Objective Bayesian Optimization: Reduces observations needed to converge by at least one order of magnitude

https://doi.org/10.1103/PhysRevAccelBeams.24.062801

Experiment at AWA photoinjector





Adapted from: https://doi.org/10.1103/PhysRevAccelBeams.23.044601

- Input variables:

$$\{K_0, K_1, \phi_1, \phi_2\}$$

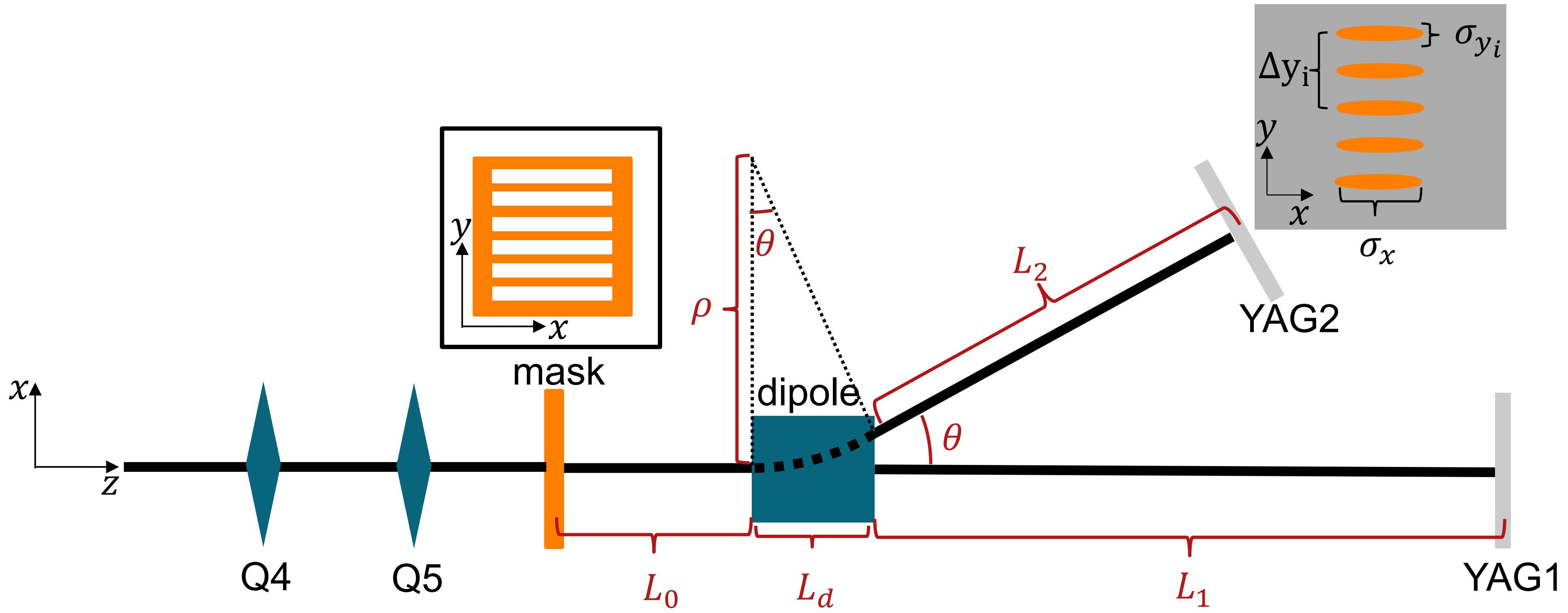
- Minimize objectives:

$$\{\varepsilon_y, dE\}$$

(Simultaneously and with few observations)

$$Q = 5\ \mathrm{nC}$$
$$\sigma_t = 6\ \mathrm{ps}$$
$$\langle E \rangle = 42\ \mathrm{MeV}$$

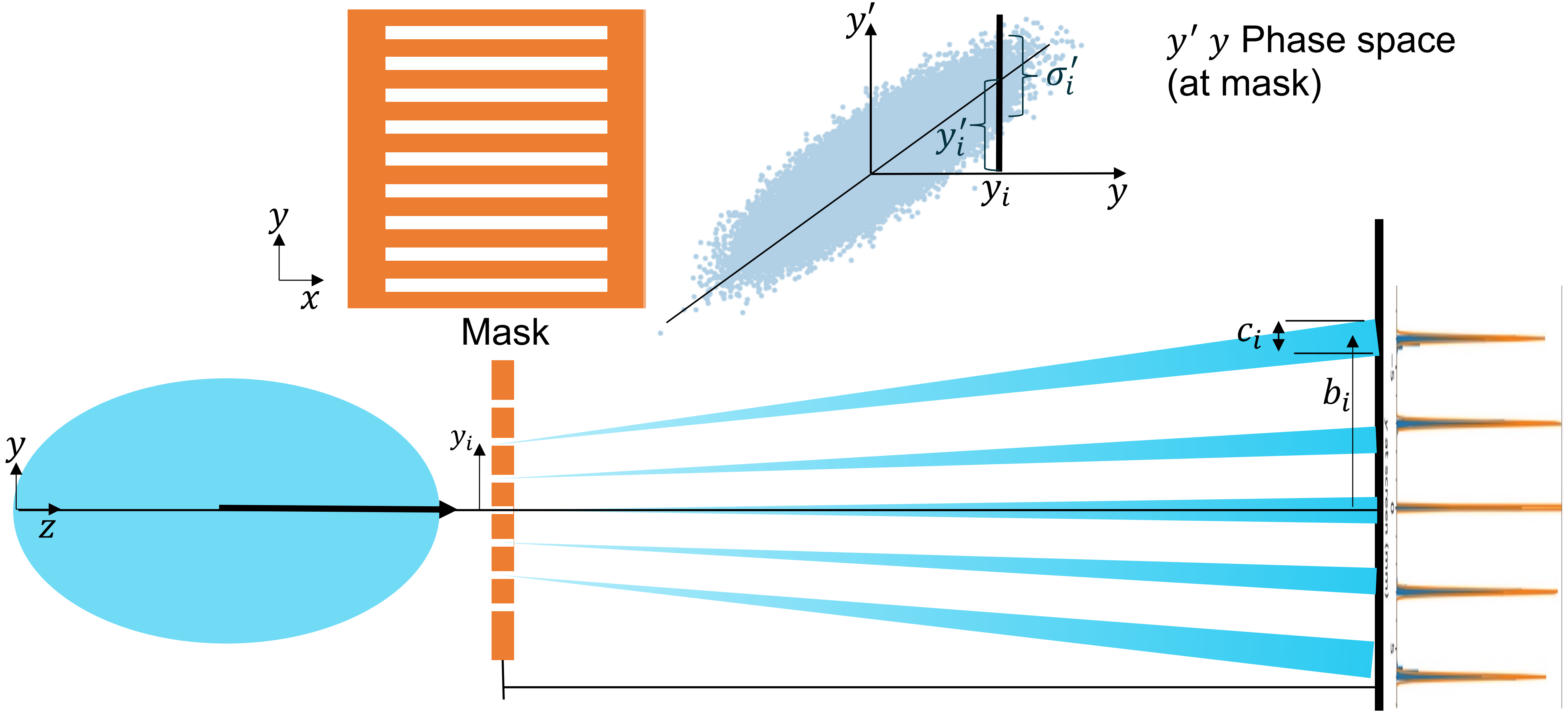
Diagnostics



$L_0 = 379.5$ mm
$L_d = 251.0$ mm
$L_1 = 2174.5$ mm
$L_2 = 974.4$ mm
$\theta = 20.0^o$
$\rho = 807.0$ mm

Emittance measurement

- Beamlets' intensities, location and width at YAG determine emittance at mask
- Explore parameters: solenoids (+ quads)



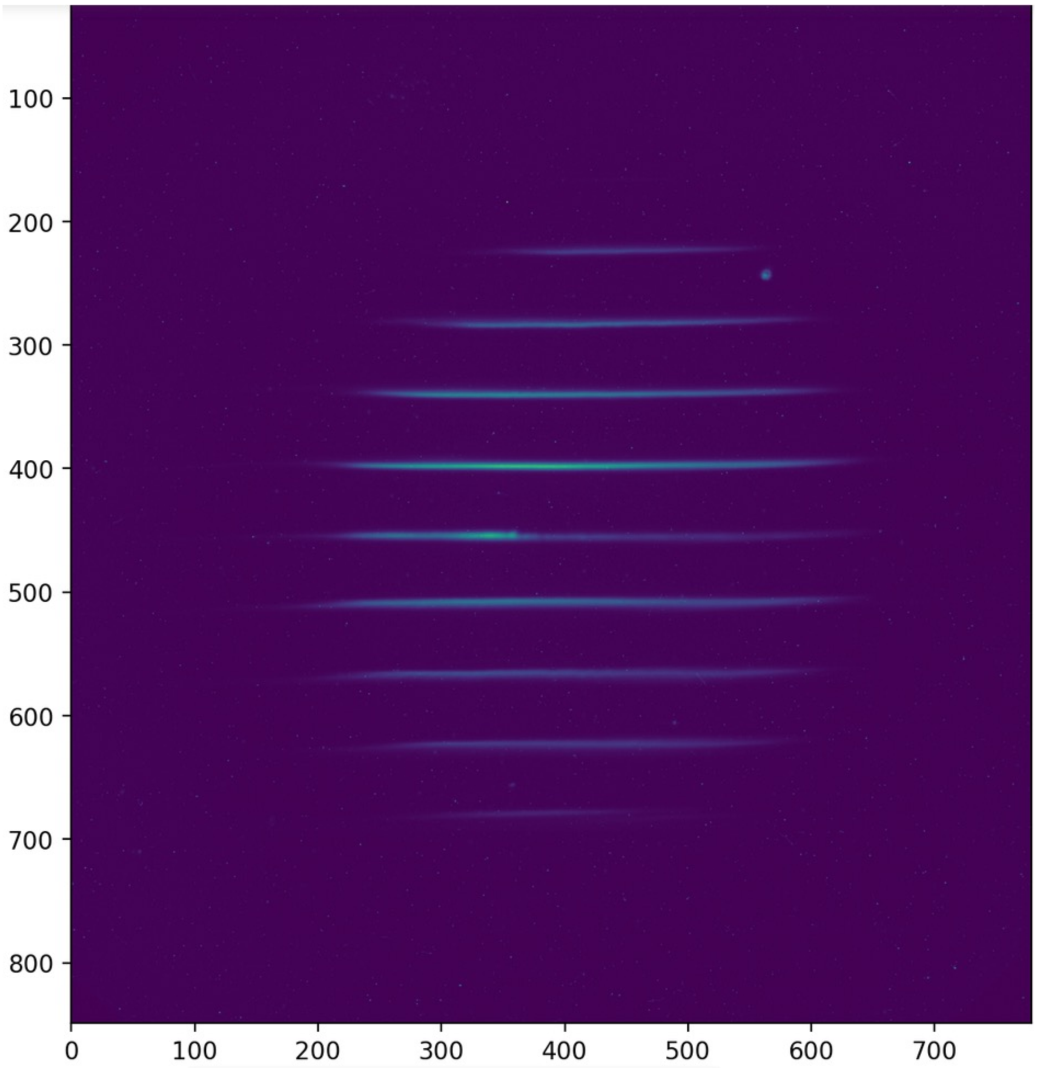
Measurement validity

ε_y measurement is valid if:

- Number of blobs > 3
- Blobs don't touch
- No blobs outside the screen

ε_y is difficult to explore and characterize:

- ε_y and validity depend on input variables
- No prior knowledge of where is the valid region in input space



Exploration challenges

- Exploring input space is difficult
  - Emittance measurement is highly constrained
  - Limited prior knowledge on the beam response/measurement validity with respect to inputs
  - Result: Grid scans are inefficient

- What if we adapt Bayesian optimization to maximize information gain?

"Bayesian exploration"


nature COMMUNICATIONS

ARTICLE

Check for updates

https://doi.org/10.1038/s41467-021-25757-3 OPEN

Turn-key constrained parameter space exploration for particle accelerators using Bayesian active learning

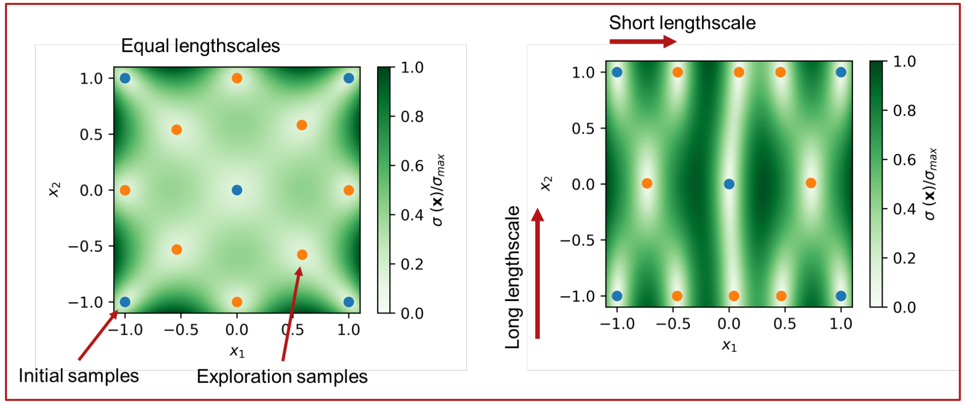
Ryan Roussel[1✉], Juan Pablo Gonzalez-Aguilera[1], Young-Kee Kim[1], Eric Wisniewski[2], Wanming Liu[2], Philippe Piot[2,3], John Power[2], Adi Hanuka[4] & Auralee Edelen[4]

https://doi.org/10.1038/s41467-021-25757-3

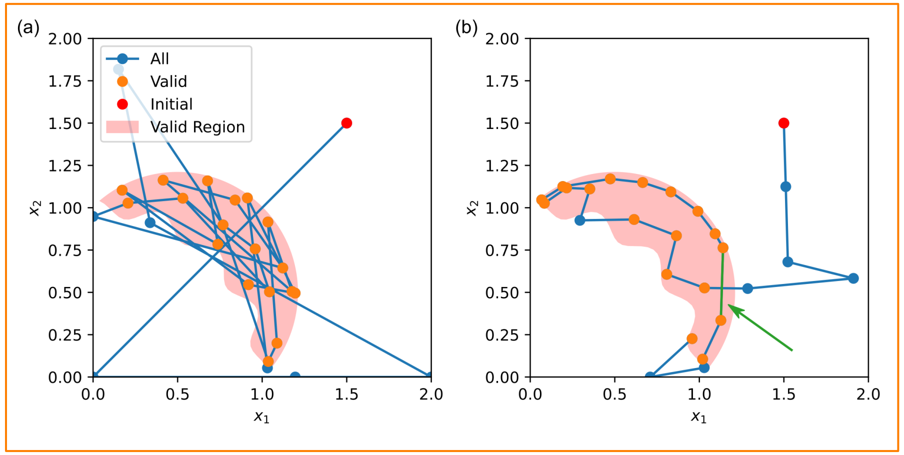
Constrained Proximal Bayesian Exploration

$$\alpha(\boldsymbol{x}) = \sigma(\boldsymbol{x}) \prod_{i=1}^{N} p_i(g_i(\boldsymbol{x}) \geq h_i)\, \Psi(\boldsymbol{x}, \boldsymbol{x_0})$$

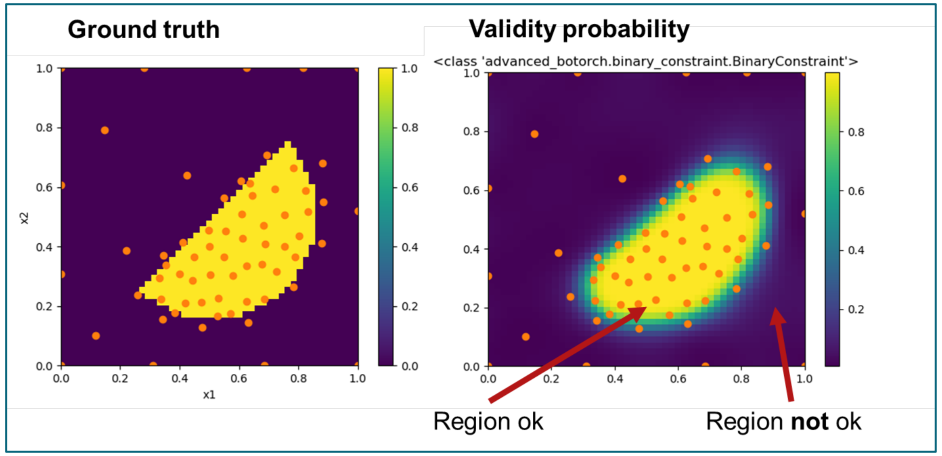
Adaptative sampling



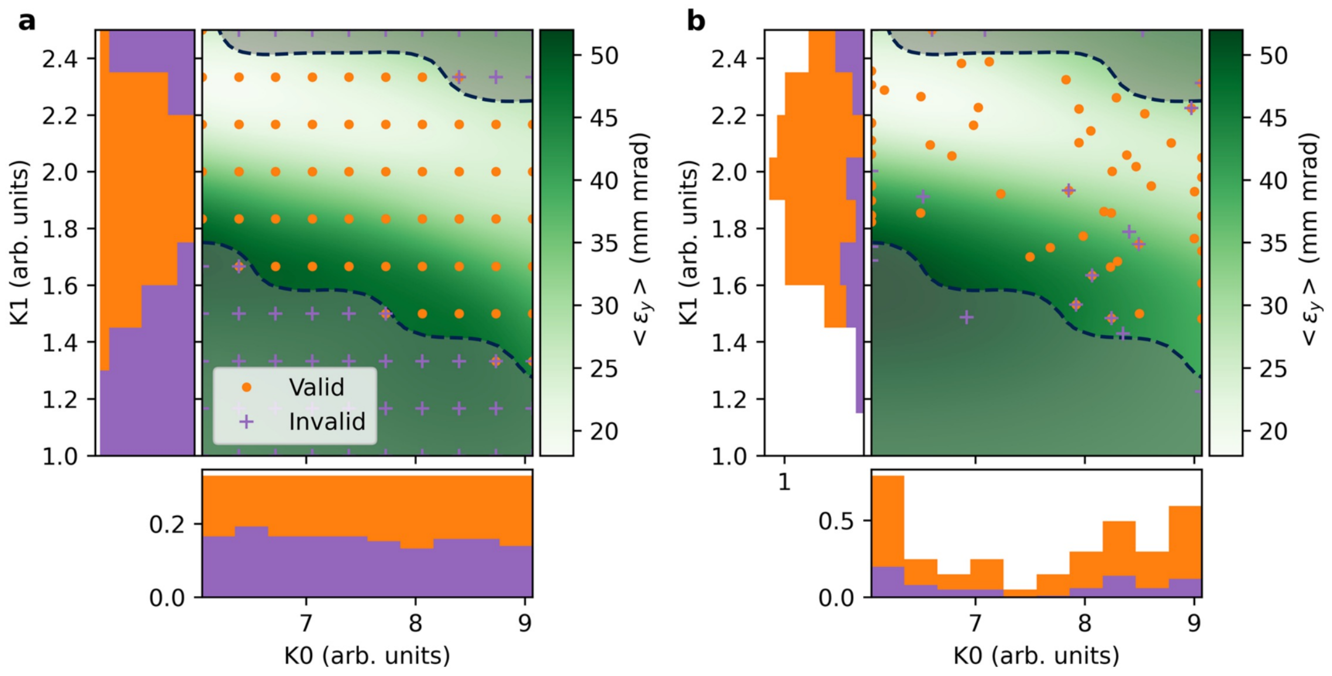
Proximal biasing



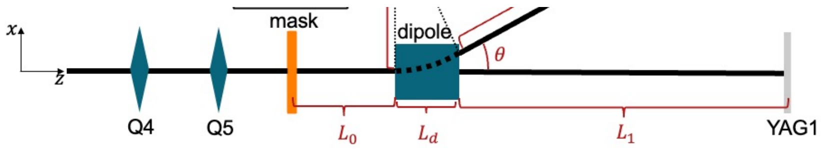
Unknown constraints



Characterizing Emittance at AWA

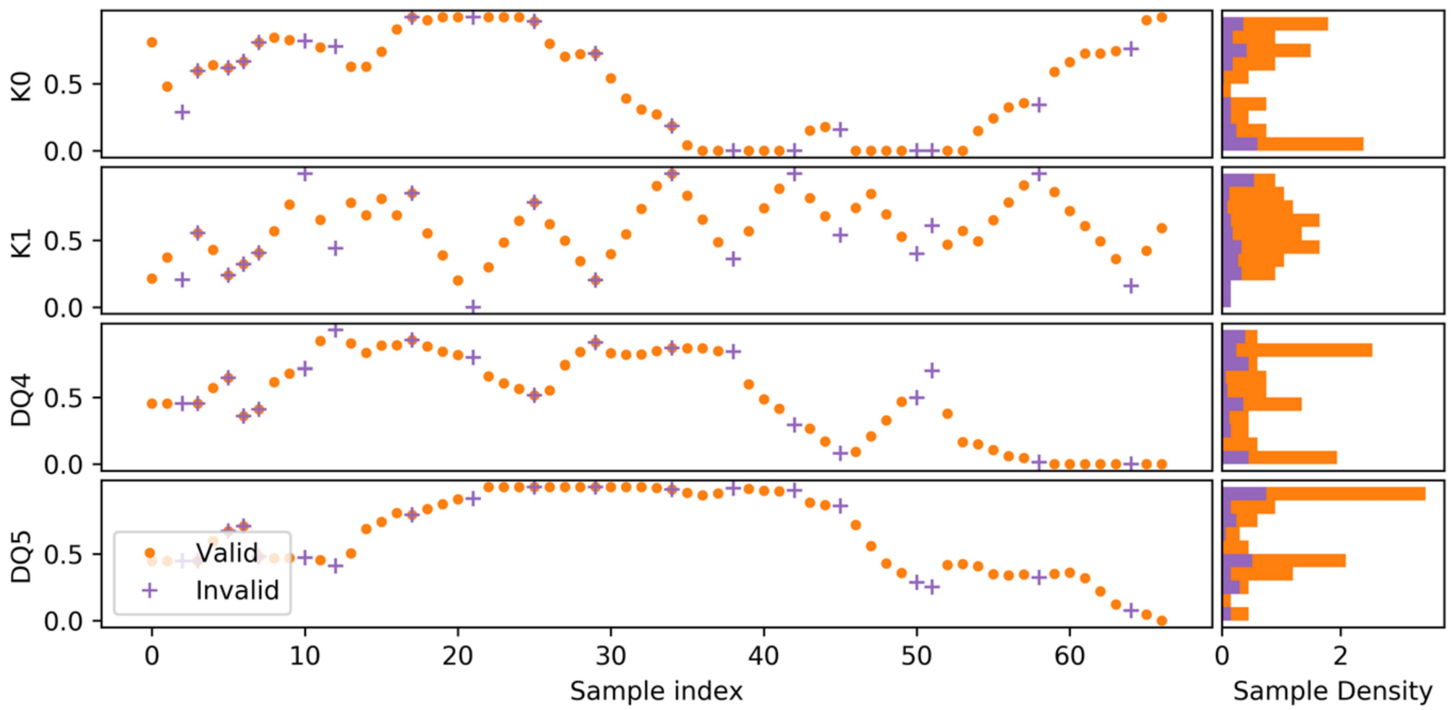


- Avoids invalid regions (77% vs 52% validity)
- Samples valid region with shorter length scales over K_1 (identifies that ε_y varies faster along K_1)
- Improves characterization of functional dependence
- Allows exploration of 4D input space (solenoids + quads)



https://doi.org/10.1038/s41467-021-25757-3

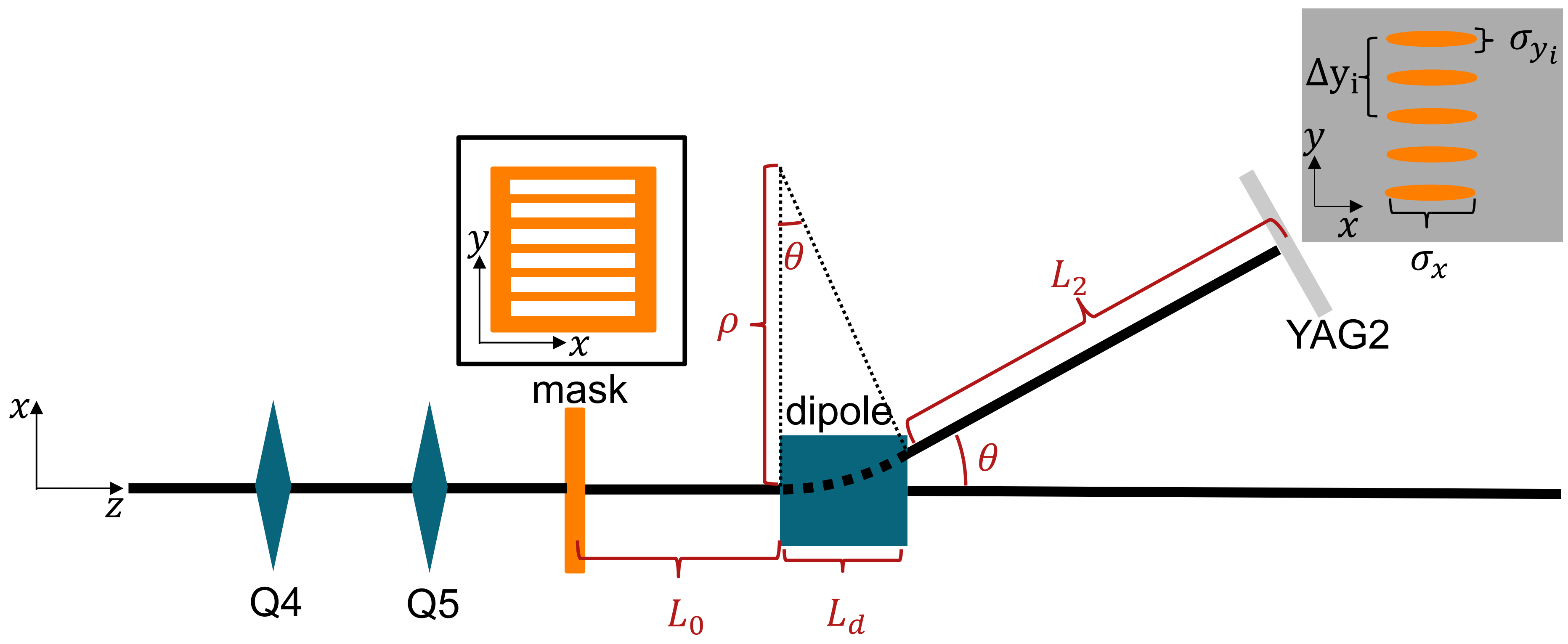
Input space scan



- Scans K_1 back and forth across valid input region
- Reduces redundant measurements on other inputs with longer length scales

https://doi.org/10.1038/s41467-021-25757-3

Energy spread measurement



$$\sigma_x^2 = \eta^2\,\sigma_\delta^2 + \epsilon(x_0, x_0')$$
$$\min_{Q4,Q5} \sigma_x^2 = \eta^2\,\sigma_\delta^2$$

- Must minimize σ_x for each measurement → Quality-aware B.O.

Questions about BO

Questions?

Autodifferentiable accelerator modeling (Project just started)

Juan Pablo Gonzalez-Aguilera*, Young-Kee Kim
UChicago

Ryan Roussel, Auralee Edelen, Christopher Mayes
SLAC

* jpga@uchicago.edu

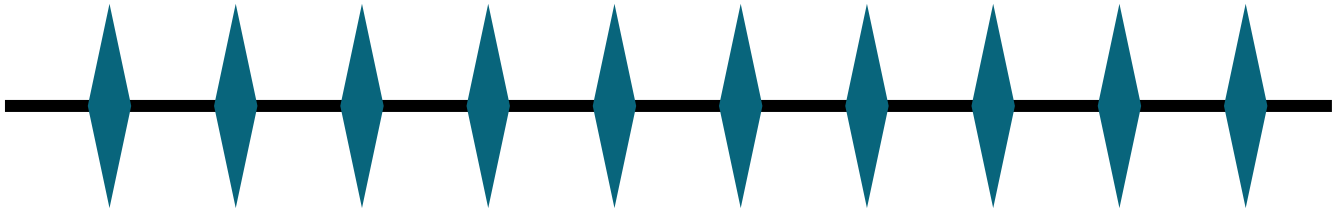
Automatic differentiation

- Computers execute sequence of elementary operations/functions

  (+, -, ×, ÷, e, log, sin, cos, …)
- Computer functions are composed of these sequences
- Autodiff uses the derivatives of these elementary operations/functions and the chain rule to evaluate the derivative of a computer function
- This results in fast and accurate derivatives

Accelerator modeling using autodiff

Toy-model: lattice composed of 10 quads separated by drifts.



- Want to optimize some beam property downstream (e.g., σ_x)
- but… high-dimensional input space
- We can test gradient-based optimization methods providing gradient information using autodiff

End

Thanks!