NEURAL INFORMATION PROCESSING SYSTEMS

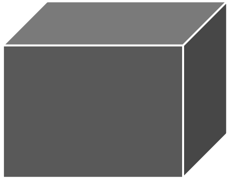
Risk-averse Heteroscedastic Bayesian Optimization

Anastasia Makarova, Ilnura Usmanova, Ilija Bogunovic, Andreas Krause

ETH zürich

Black-box optimization arises in high-stakes applications

Black Box

x →  → $y = f(x) + \xi$

$f(x)$

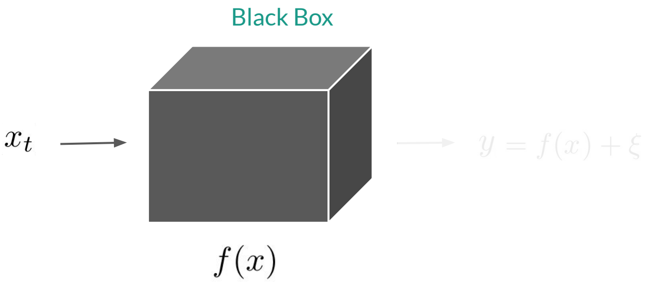
Drug discovery

→ Compound efficiency

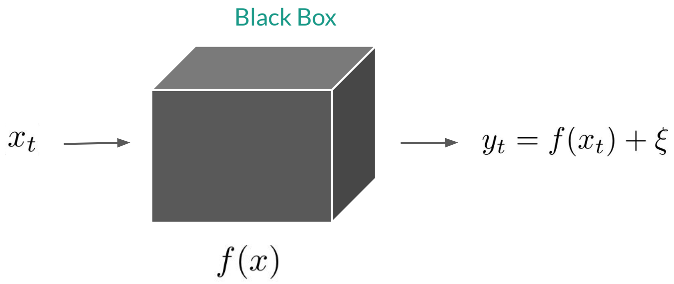
Tuning accelerators

Beam positions → → Pulse energy

Black-box optimization arises in high-stakes applications



Black-box optimization arises in high-stakes applications



Drug discovery

Compound efficiency

Tuning accelerators

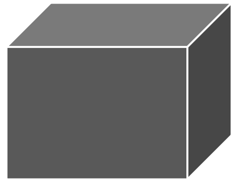
Beam positions

Pulse energy

Black-box optimization arises in high-stakes applications

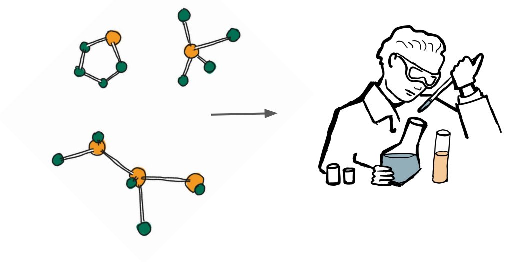
where one needs to trade off attaining high utility vs minimizing risk

Black Box

x_t →  → $y_t = f(x_t) + \xi$

$f(x)$

Drug discovery

 → Compound efficiency

Tuning accelerators

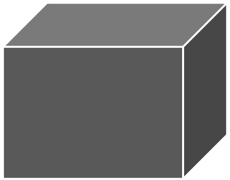
Beam positions → → Pulse energy

Black-box optimization arises in high-stakes applications

where one needs to trade off attaining high utility vs minimizing risk

Black Box

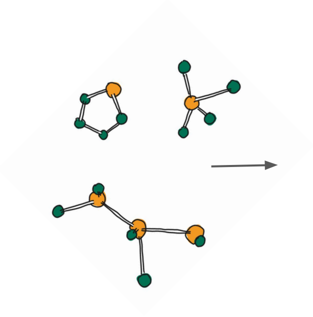
x_t →



→ $y_t = f(x_t) + \xi$

$f(x)$

Drug discovery

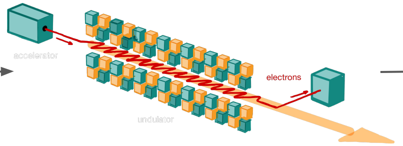


→

→ Compound efficiency

Tuning accelerators

Beam positions →

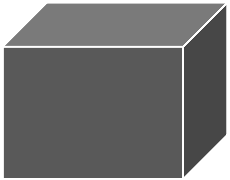


→ Pulse energy

Black-box optimization arises in high-stakes applications

Black Box

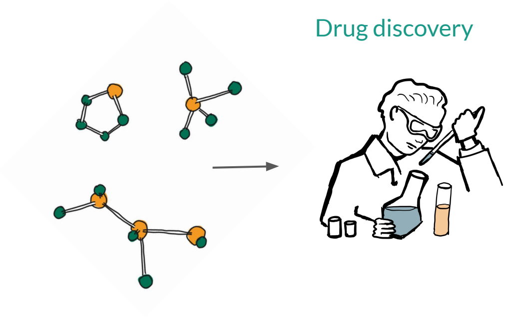
x_t



$y_t = f(x_t) + \xi$

$f(x)$

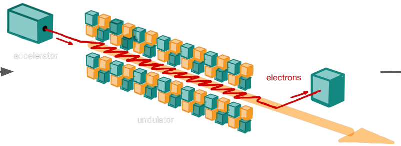
Drug discovery



Compound efficiency

Tuning accelerators

Beam positions



Pulse energy

In these applications, one needs to trade off attaining high utility vs minimizing risk

Bayesian optimization is a powerful framework for black-box functions

trading off exploration & exploitation

objective $f(x)$

$$x^* \in \arg\max_{x \in \mathcal{X}} f(x)$$

regret $R_T = \sum_{t=1}^{T} \left[f(x^*) - f(x_t) \right]$

Noise-perturbed evaluations

$$y_t = f(x_t) + \xi(x_t)$$

ρ -sub-Gaussian noise

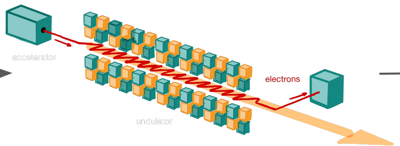
Optimism under uncertainty via GP Upper Confidence bound

$$x_t \in \arg\max_{x \in \mathcal{X}} \underbrace{\mu_{t-1}(x) + \beta_t \sigma_{t-1}(x)}_{=:\mathrm{ucb}_t^f(x)}$$

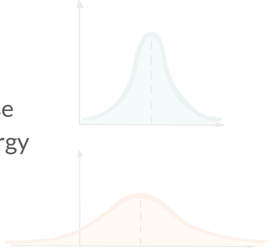
GP posterior mean and variance

Tuning accelerators

Beam positions x_1^* x_2^*



Pulse energy



Two different inputs have **similar expected values**, but one produces **much noisier realizations**

Bayesian optimization is a powerful framework for black-box functions

trading off exploration & exploitation

objective

$$f(x)$$

$$x^* \in \arg\max_{x \in \mathcal{X}} f(x)$$

regret

$$R_T = \sum_{t=1}^{T} \left[f(x^*) - f(x_t) \right]$$

Noise-perturbed evaluations

$$y_t = f(x_t) + \xi(x_t)$$

ρ -sub-Gaussian noise

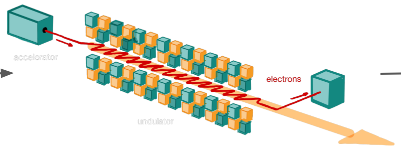
Optimism under uncertainty via GP Upper Confidence bound

$$x_t \in \arg\max_{x \in \mathcal{X}} \underbrace{\mu_{t-1}(x) + \beta_t \sigma_{t-1}(x)}_{=:\mathrm{ucb}_t^f(x)}$$

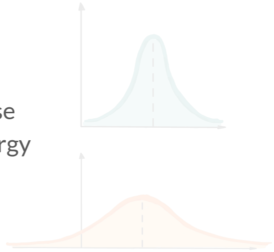
GP posterior mean and variance

Tuning accelerators

Beam positions x_1^* x_2^*



Pulse energy



Two different inputs have **similar expected values**, but one produces **much noisier realizations**

Bayesian optimization is a powerful framework for black-box functions

trading off exploration & exploitation

objective

$$f(x)$$

$$x^* \in \arg\max_{x \in \mathcal{X}} f(x)$$

regret

$$R_T = \sum_{t=1}^{T} \left[f(x^*) - f(x_t) \right]$$

Noise-perturbed evaluations

$$y_t = f(x_t) + \xi(x_t)$$

ρ -sub-Gaussian noise

Optimism under uncertainty via GP Upper Confidence bound

$$x_t \in \arg\max_{x \in \mathcal{X}} \underbrace{\mu_{t-1}(x) + \beta_t \sigma_{t-1}(x)}_{=: \mathrm{ucb}_t^f(x)}$$

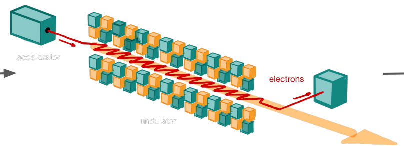
GP posterior mean and variance

Tuning accelerators

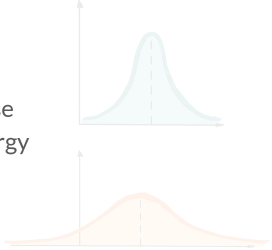
Beam positions

x_1^*

x_2^*



Pulse energy



Two different inputs have **similar expected values**, but one produces **much noisier realizations**

Bayesian optimization is a powerful framework for black-box functions

trading off exploration & exploitation

objective $f(x)$

$$x^* \in \arg\max_{x \in \mathcal{X}} f(x)$$

regret $R_T = \sum_{t=1}^{T} \left[f(x^*) - f(x_t) \right]$

Noise-perturbed evaluations

$$y_t = f(x_t) + \xi(x_t)$$

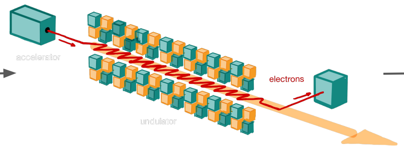
ρ -sub-Gaussian noise

Optimism under uncertainty via GP Upper Confidence bound

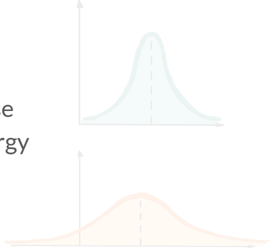
$$x_t \in \arg\max_{x \in \mathcal{X}} \underbrace{\mu_{t-1}(x) + \beta_t \sigma_{t-1}(x)}_{=:\mathrm{ucb}_t^f(x)}.$$

Tuning accelerators

Beam positions x_1^* x_2^*



Pulse energy



Two different inputs have **similar expected values**, but one produces **much noisier realizations**

Bayesian optimization is a powerful framework for black-box functions
trading off exploration & exploitation

objective $f(x)$

$$x^* \in \arg\max_{x \in \mathcal{X}} f(x)$$

regret $$R_T = \sum_{t=1}^{T} \left[f(x^*) - f(x_t) \right]$$

Noise-perturbed evaluations

$$y_t = f(x_t) + \xi(x_t)$$

ρ -sub-Gaussian noise

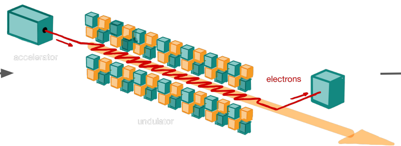
Optimism under uncertainty via GP Upper Confidence bound

$$x_t \in \arg\max_{x \in \mathcal{X}} \underbrace{\mu_{t-1}(x) + \beta_t \sigma_{t-1}(x)}_{=:\mathrm{ucb}_t^f(x)}$$

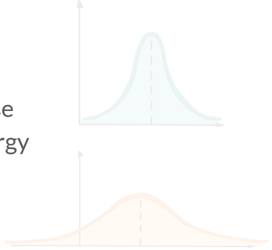
GP posterior
mean and variance

Tuning accelerators

Beam positions



Pulse energy



Two different inputs have **similar expected values**, but one produces **much noisier realizations**

Bayesian optimization is a powerful framework for black-box functions

trading off exploration & exploitation

objective $f(x)$

$$x^* \in \arg\max_{x \in \mathcal{X}} f(x)$$

regret
$$R_T = \sum_{t=1}^{T} \left[f(x^*) - f(x_t) \right]$$

Noise-perturbed evaluations

$$y_t = f(x_t) + \xi(x_t)$$

ρ -sub-Gaussian noise

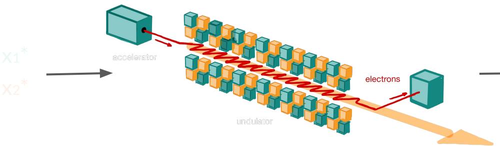
Optimism under uncertainty via GP Upper Confidence bound

$$x_t \in \arg\max_{x \in \mathcal{X}} \underbrace{\mu_{t-1}(x) + \beta_t \sigma_{t-1}(x)}_{=:\mathrm{ucb}_t^f(x)}$$

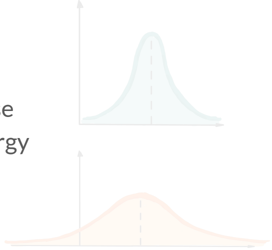
GP posterior
mean and variance

Tuning accelerators

Beam
positions



Pulse
energy



Two different inputs have **similar expected values**,
but one produces **much noisier realizations**

Bayesian optimization is a powerful framework for black-box functions

trading off exploration & exploitation

objective $f(x)$

$$x^* \in \arg\max_{x\in\mathcal{X}} f(x)$$

regret
$$R_T = \sum_{t=1}^{T}\left[f(x^*) - f(x_t)\right]$$

Noise-perturbed evaluations

$$y_t = f(x_t) + \xi(x_t)$$

ρ -sub-Gaussian noise

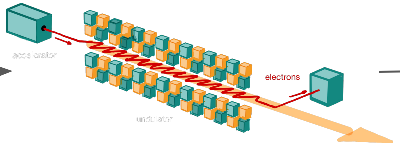
Optimism under uncertainty via GP Upper Confidence bound

$$x_t \in \arg\max_{x\in\mathcal{X}} \underbrace{\mu_{t-1}(x) + \beta_t \sigma_{t-1}(x)}_{=:\mathrm{ucb}_t^f(x)}$$

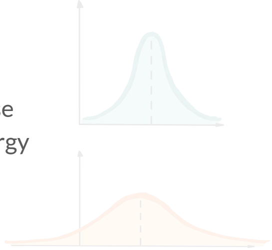
GP posterior
mean and variance

Tuning accelerators

Beam positions x_1^* x_2^*



Pulse energy



Two different inputs have **similar expected values**,
but one produces **much noisier realizations**

Bayesian optimization is a powerful framework for black-box functions
trading off exploration & exploitation

objective $f(x)$

$$x^* \in \arg\max_{x \in \mathcal{X}} f(x)$$

regret
$$R_T = \sum_{t=1}^{T} \left[f(x^*) - f(x_t) \right]$$

Noise-perturbed evaluations

$$y_t = f(x_t) + \xi(x_t)$$

ρ -sub-Gaussian noise

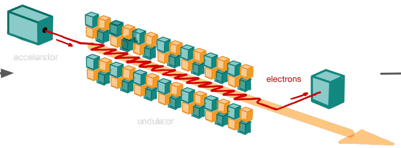
Optimism under uncertainty via GP Upper Confidence bound

$$x_t \in \arg\max_{x \in \mathcal{X}} \underbrace{\mu_{t-1}(x) + \beta_t \sigma_{t-1}(x)}_{=:\mathrm{ucb}_t^f(x)}$$

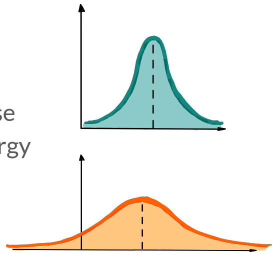
GP posterior
mean and variance

Tuning accelerators

Beam positions x_1^* x_2^*



Pulse energy



Two different inputs have **similar expected values**, but one produces **much noisier realizations**

Heteroscedastic Bayesian optimization is a powerful framework for black-box functions trading off exploration & exploitation

objective $f(x)$

$$x^* \in \arg\max_{x \in \mathcal{X}} f(x)$$

regret $$R_T = \sum_{t=1}^{T} \left[f(x^*) - f(x_t) \right]$$

Noise-perturbed evaluations

$$y_t = f(x_t) + \xi(x_t)$$

$\rho(x)$-sub-Gaussian noise

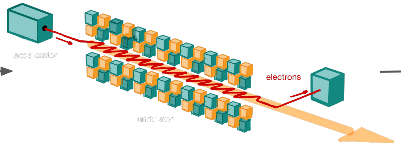
Optimism under uncertainty via GP Upper Confidence bound

$$x_t \in \arg\max_{x \in \mathcal{X}} \underbrace{\mu_{t-1}(x) + \beta_t \sigma_{t-1}(x)}_{=:\mathrm{ucb}_t^f(x)}$$

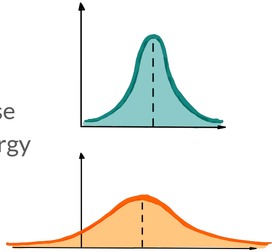
GP posterior mean and variance

Tuning accelerators

Beam positions x_1^* x_2^*



Pulse energy



Two different inputs have **similar expected values**, but one produces **much noisier realizations**

Heteroscedastic Bayesian optimization is a powerful framework for black-box functions trading off exploration & exploitation **& risk**

Mean-Variance objective

$$\mathrm{MV}(x) = f(x) - \alpha\rho^2(x)$$

$$x^* \in \arg\max_{x \in \mathcal{X}} \mathrm{MV}(x)$$

Risk-averse regret

$$R_T = \sum_{t=1}^{T}\left[\mathrm{MV}(x^*) - \mathrm{MV}(x_t)\right]$$

Noise-perturbed evaluations

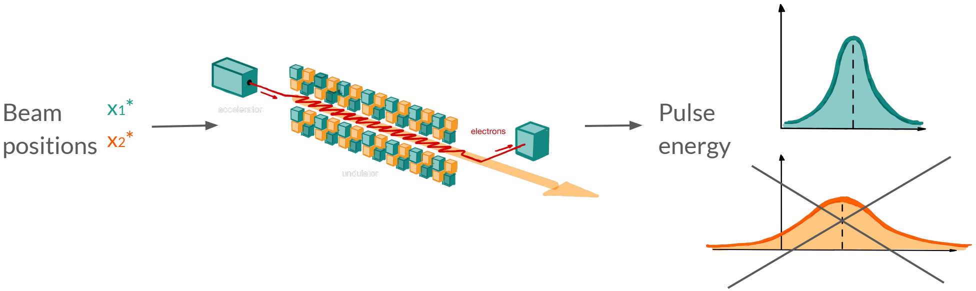
$$y_t = f(x_t) + \xi(x_t)$$

$\rho(x)$-sub-Gaussian noise

Optimism under uncertainty via RAHBO:

$$x_t \in \arg\max_{x \in \mathcal{X}} \ \mathrm{ucb}^f_{t-1}(x) - \alpha\, \mathrm{lcb}^{\rho^2}_{t-1}(x)$$

Tuning accelerators



RAHBO idea: Generalize Bayesian optimization to **trade mean and input-dependent variance** of the objective, despite both of them being unknown a priori.

Risk-averse Heteroscedastic Bayesian optimization (RAHBO) is a framework for black-box functions trading off exploration & exploitation **& risk**

Mean-Variance objective

$$\mathrm{MV}(x) = f(x) - \alpha\rho^2(x)$$

$$x^* \in \arg\max_{x\in\mathcal{X}} \mathrm{MV}(x)$$

Risk-averse regret

$$R_T = \sum_{t=1}^{T}\left[\mathrm{MV}(x^*) - \mathrm{MV}(x_t)\right]$$

Noise-perturbed evaluations

$$y_t = f(x_t) + \xi(x_t)$$

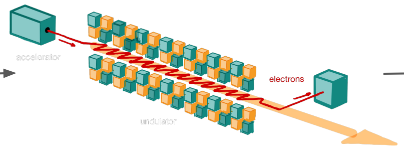
$\rho(x)$-sub-Gaussian noise

Optimism under uncertainty via RAHBO:

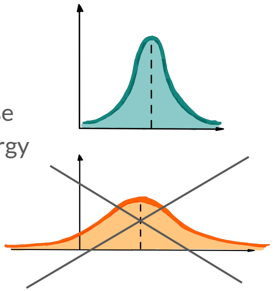
$$x_t \in \arg\max_{x\in\mathcal{X}} \ \mathrm{ucb}^f_{t-1}(x) - \alpha\ \mathrm{lcb}^{\rho^2}_{t-1}(x)$$

Tuning accelerators

Beam positions x_1^* x_2^*



Pulse energy



RAHBO idea: Generalize Bayesian optimization to **trade mean and input-dependent variance** of the objective, despite both of them being unknown a priori.

Risk-averse Heteroscedastic Bayesian optimization (RAHBO) is a framework for black-box functions trading off exploration & exploitation **& risk**

Mean-Variance objective

$$\mathbf{MV}(x) = f(x) - \alpha\rho^2(x)$$

$$x^* \in \arg\max_{x \in \mathcal{X}} \mathbf{MV}(x)$$

Risk-averse regret

$$R_T = \sum_{t=1}^{T} \left[\mathbf{MV}(x^*) - \mathbf{MV}(x_t) \right]$$

Noise-perturbed evaluations

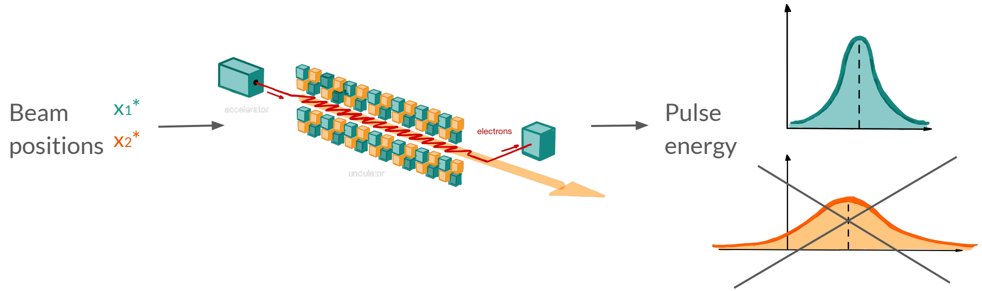
$$y_t = f(x_t) + \xi(x_t)$$

$\rho(x)$-sub-Gaussian noise

Optimism under uncertainty via RAHBO:

$$x_t \in \arg\max_{x \in \mathcal{X}} \ \mathrm{ucb}^f_{t-1}(x) - \alpha \ \mathrm{lcb}^{\rho^2}_{t-1}(x)$$

Tuning accelerators



RAHBO idea: Generalize Bayesian optimization to **trade mean and input-dependent variance** of the objective, despite both of them being unknown a priori.

Risk-averse Heteroscedastic Bayesian optimization (RAHBO) is a framework for black-box functions trading off exploration & exploitation **& risk**

coefficient of absolute risk tolerance

Mean-Variance objective

$$\mathbf{MV}(x) = f(x) - \alpha\rho^2(x)$$

$$x^* \in \arg\max_{x\in\mathcal{X}} \mathbf{MV}(x)$$

Risk-averse regret

$$R_T = \sum_{t=1}^{T}\left[\mathbf{MV}(x^*) - \mathbf{MV}(x_t)\right]$$

Noise-perturbed evaluations

$$y_t = f(x_t) + \xi(x_t)$$

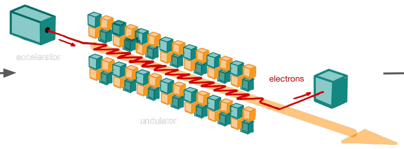
$\rho(x)$-sub-Gaussian noise

Optimism under uncertainty via RAHBO:

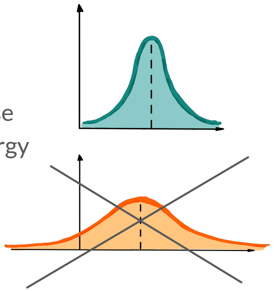
$$x_t \in \arg\max_{x\in\mathcal{X}} \ \mathrm{ucb}^f_{t-1}(x) - \alpha \ \mathrm{lcb}^{\rho^2}_{t-1}(x)$$

Tuning accelerators

Beam positions x_1^* x_2^*



Pulse energy



RAHBO idea: Generalize Bayesian optimization to **trade mean and input-dependent variance** of the objective, despite both of them being unknown a priori.

Risk-averse Heteroscedastic Bayesian optimization (RAHBO) is a framework for black-box functions trading off exploration & exploitation **& risk**

Mean-Variance objective

$$\mathbf{MV}(x) = f(x) - \alpha\rho^2(x)$$

$$x^* \in \arg\max_{x\in\mathcal{X}} \mathbf{MV}(x)$$

Risk-averse regret

$$R_T = \sum_{t=1}^{T}\Big[\mathbf{MV}(x^*) - \mathbf{MV}(x_t)\Big]$$

Noise-perturbed evaluations

$$y_t = f(x_t) + \xi(x_t)$$

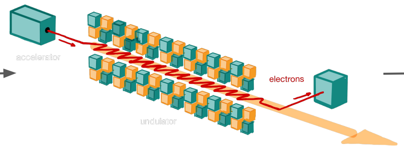
$\rho(x)$-sub-Gaussian noise

Optimism under uncertainty via RAHBO:

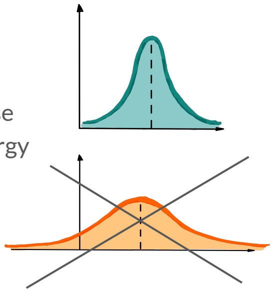
$$x_t \in \arg\max_{x\in\mathcal{X}} \ \mathrm{ucb}^f_{t-1}(x) - \alpha\ \mathrm{lcb}^{\rho^2}_{t-1}(x)$$

Tuning accelerators

Beam positions x_1^* x_2^*



Pulse energy



RAHBO idea: Generalize Bayesian optimization to **trade mean and input-dependent variance** of the objective, despite both of them being unknown a priori.

Risk-averse Heteroscedastic Bayesian optimization (RAHBO) is a framework for black-box functions trading off exploration & exploitation **& risk**

Mean-Variance objective

$$\mathbf{MV}(x) = f(x) - \alpha\rho^2(x)$$

$$x^* \in \arg\max_{x \in \mathcal{X}} \mathbf{MV}(x)$$

Risk-averse regret

$$R_T = \sum_{t=1}^{T} \Big[\mathbf{MV}(x^*) - \mathbf{MV}(x_t)\Big]$$

Noise-perturbed evaluations

$$y_t = f(x_t) + \xi(x_t)$$

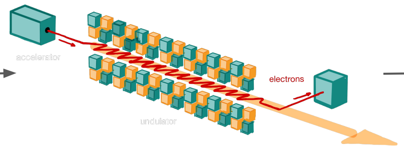
$\rho(x)$-sub-Gaussian noise

Optimism under uncertainty via RAHBO:

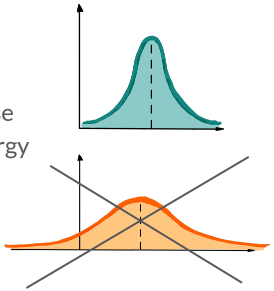
$$x_t \in \arg\max_{x \in \mathcal{X}} \ \mathrm{ucb}^f_{t-1}(x) - \alpha\, \mathrm{lcb}^{\rho^2}_{t-1}(x)$$

Tuning accelerators

Beam positions x_1^* x_2^*



Pulse energy

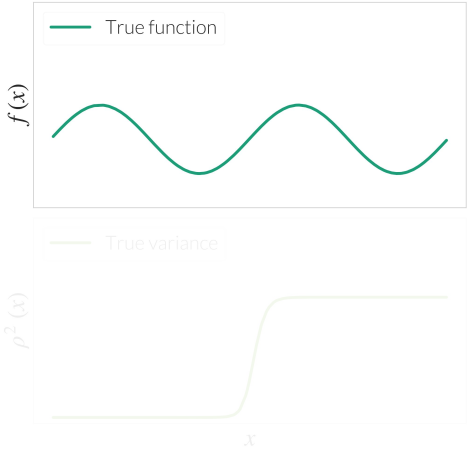


RAHBO idea: Generalize Bayesian optimization to **trade mean and input-dependent variance** of the objective, despite both of them being unknown a priori.

Failure of GP-UCB

In the risk-averse setting the maximizers for mean-variance (RAHBO) and for mean (GP-UCB) might not coincide

Unknown objective (sine) with two global optimizers
-- but one has much higher noise variance



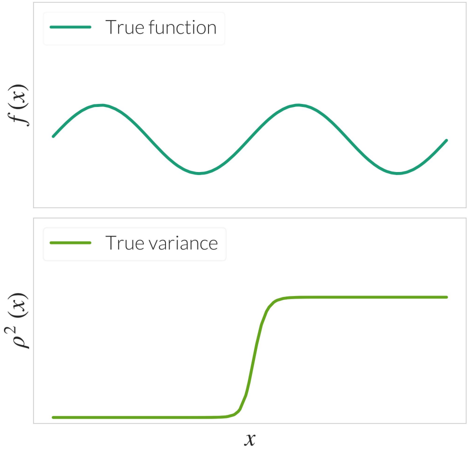
RAHBO: exploration & exploitation **& risk** tradeoff

GP-UCB: exploration & exploitation tradeoff

Failure of GP-UCB

In the risk-averse setting the maximizers for mean-variance (RAHBO) and for mean (GP-UCB) might not coincide

Unknown objective (sine) with two global optimizers
-- but one has much higher noise variance



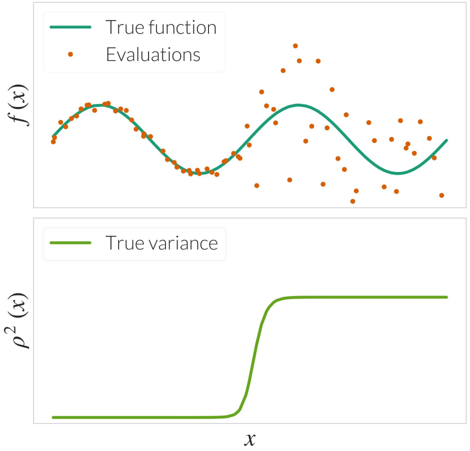
RAHBO: exploration & exploitation **& risk** tradeoff

GP-UCB: exploration & exploitation tradeoff

Failure of GP-UCB

In the risk-averse setting the maximizers for mean-variance (RAHBO) and for mean (GP-UCB) might not coincide

Unknown objective (sine) with two global optimizers
-- but one has much higher noise variance



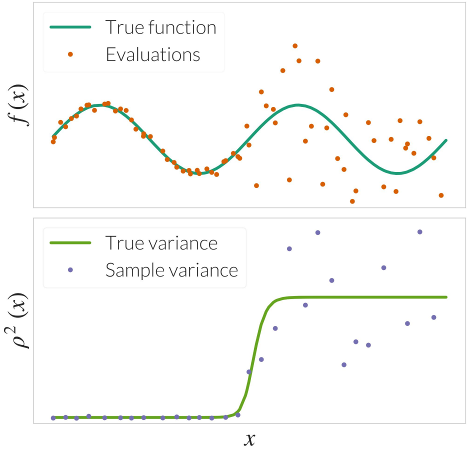
RAHBO: exploration & exploitation **& risk** tradeoff

GP-UCB: exploration & exploitation tradeoff

Failure of GP-UCB

In the risk-averse setting the maximizers for mean-variance (RAHBO) and for mean (GP-UCB) might not coincide

Unknown objective (sine) with two global optimizers
-- but one has much higher noise variance



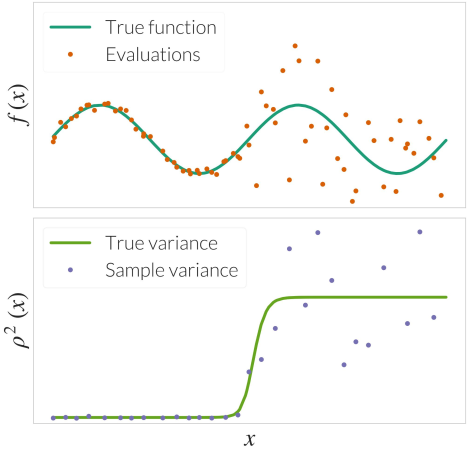
RAHBO: exploration & exploitation **& risk** tradeoff

GP-UCB: exploration & exploitation tradeoff

Failure of GP-UCB

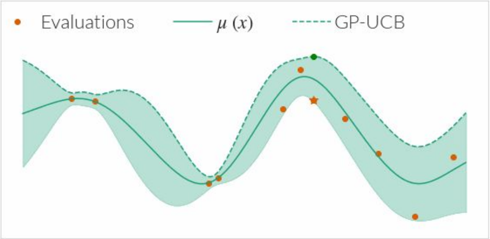
In the risk-averse setting the maximizers for mean-variance (RAHBO) and for mean (GP-UCB) might not coincide

Unknown objective (sine) with two global optimizers
-- but one has much higher noise variance



RAHBO: exploration & exploitation **& risk** tradeoff

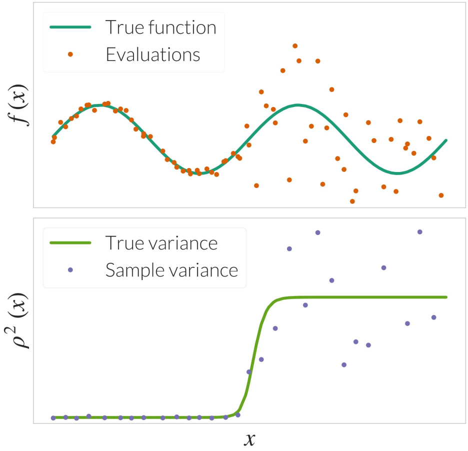
GP-UCB: exploration & exploitation tradeoff



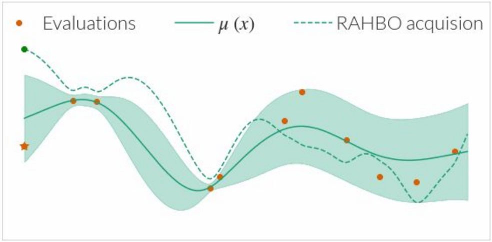
Failure of GP-UCB

In the risk-averse setting the maximizers for mean-variance (RAHBO) and for mean (GP-UCB) might not coincide

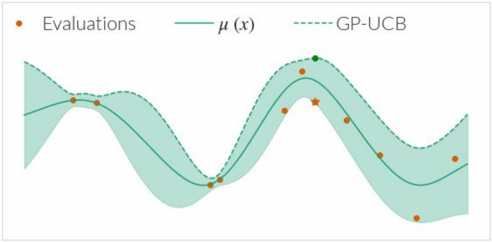
Unknown objective (sine) with two global optimizers
-- but one has much higher noise variance



RAHBO: exploration & exploitation **& risk** tradeoff



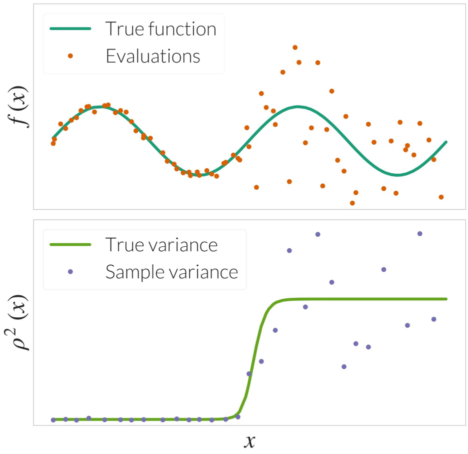
GP-UCB: exploration & exploitation tradeoff



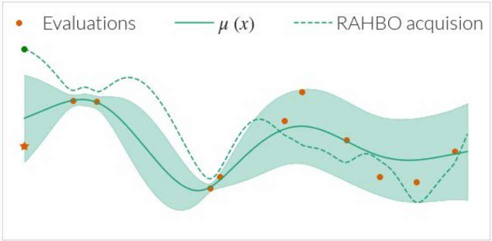
Failure of GP-UCB

The maximizers for mean-variance (RAHBO) and for mean (GP-UCB) might not coincide

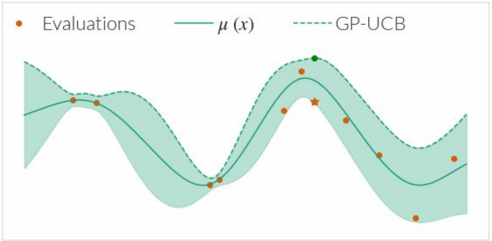
Unknown objective (sine) with two global optimizers -- but one has much higher noise variance



RAHBO: exploration & exploitation **& risk** tradeoff

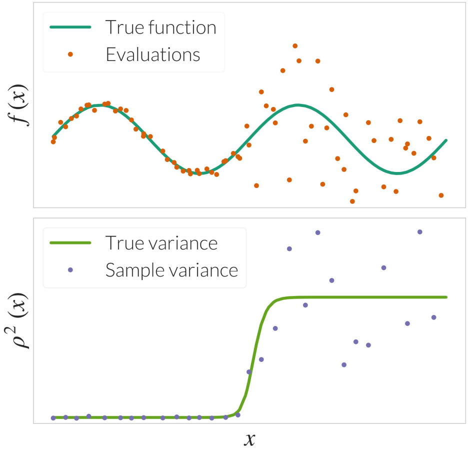


GP-UCB: exploration & exploitation tradeoff



The Risk-averse Heteroscedastic Bayesian Optimization (RAHBO) Algorithm

RAHBO attaints sublinear regret $R_T = \mathcal{O}(\sqrt{T})$

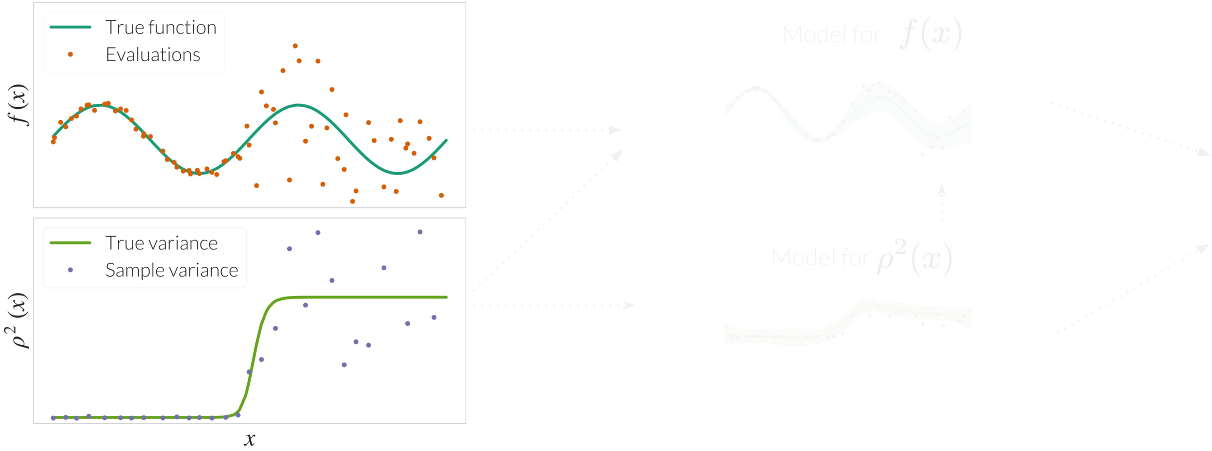


Model for $f(x)$

Model for $\rho^2(x)$

$$x_t \in \arg\max_{x \in \mathcal{X}} \ \mathrm{ucb}^{f}_{t-1}(x) - \alpha \ \mathrm{lcb}^{\rho^2}_{t-1}(x)$$

The Risk-averse Heteroscedastic Bayesian Optimization (RAHBO) Algorithm

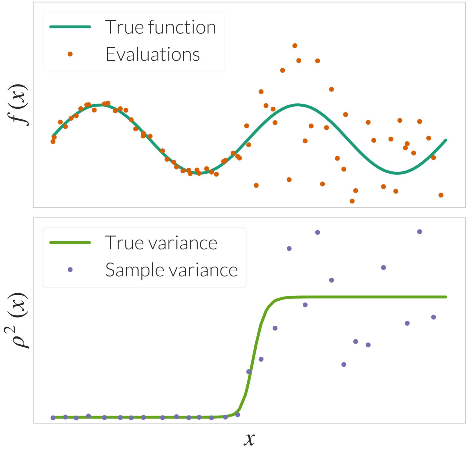


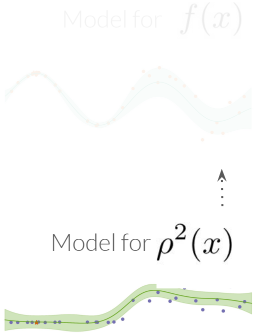
Optimism under uncertainty

$$x_t \in \arg\max_{x \in \mathcal{X}} \ \mathrm{ucb}^{f}_{t-1}(x) - \alpha \ \mathrm{lcb}^{\rho^2}_{t-1}(x)$$

The Risk-averse Heteroscedastic Bayesian Optimization (RAHBO) Algorithm

RAHBO attaints sublinear regret $R_T = \mathcal{O}(\sqrt{T})$



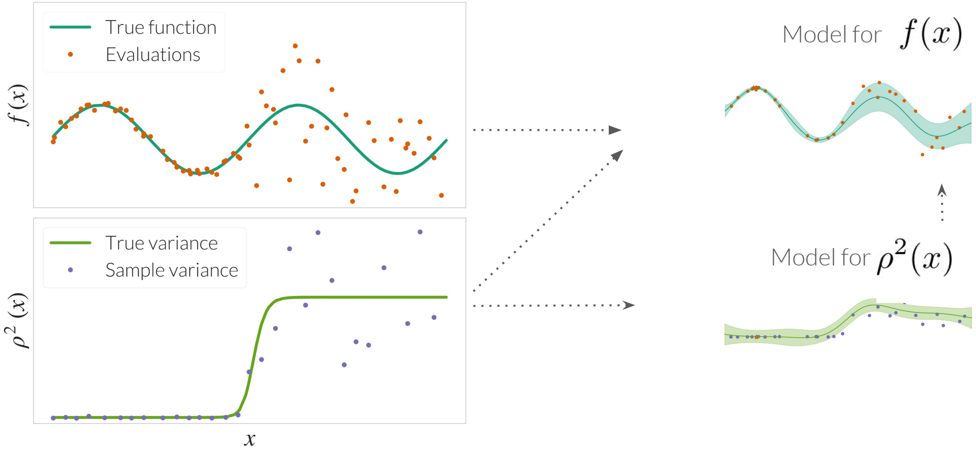


Optimism under uncertainty

$$x_t \in \arg\max_{x \in \mathcal{X}} \ \mathrm{ucb}^{f}_{t-1}(x) - \alpha \ \mathrm{lcb}^{\rho^2}_{t-1}(x)$$

The Risk-averse Heteroscedastic Bayesian Optimization (RAHBO) Algorithm

RAHBO attaints sublinear regret $R_T = \mathcal{O}(\sqrt{T})$

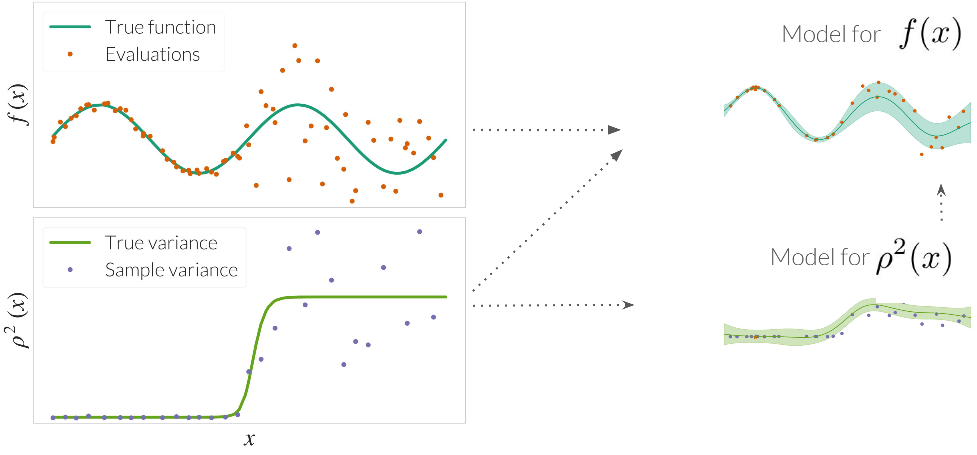


Optimism under uncertainty

$$x_t \in \arg\max_{x \in \mathcal{X}} \; \mathrm{ucb}^{f}_{t-1}(x) - \alpha \, \mathrm{lcb}^{\rho^2}_{t-1}(x)$$

The Risk-averse Heteroscedastic Bayesian Optimization (RAHBO) Algorithm

RAHBO attaints sublinear regret $R_T = \mathcal{O}(\sqrt{T})$

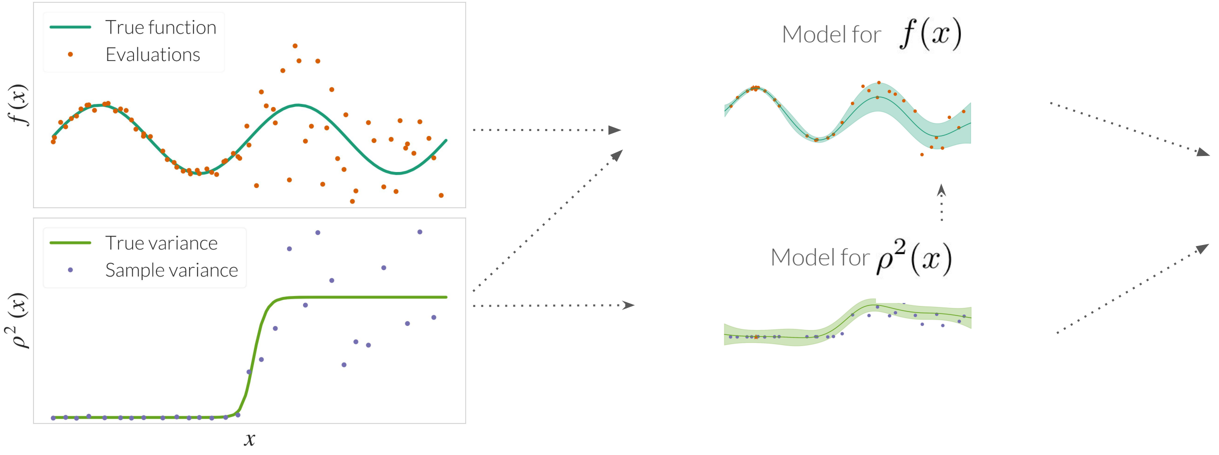


Optimism under uncertainty

$$x_t \in \arg\max_{x \in \mathcal{X}} \ \mathrm{ucb}^{f}_{t-1}(x) - \alpha \, \mathrm{lcb}^{\rho^2}_{t-1}(x)$$

The Risk-averse Heteroscedastic Bayesian Optimization (RAHBO) Algorithm

Convergence guarantees (informal):
Under certain assumptions (see the paper for details),
RAHBO attains sublinear regret.

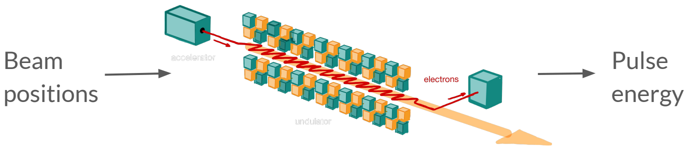


Optimism under uncertainty

$$x_t \in \arg\max_{x \in \mathcal{X}} \ \mathrm{ucb}^{f}_{t-1}(x) - \alpha \ \mathrm{lcb}^{\rho^2}_{t-1}(x)$$

RAHBO results for SwissFEL

Tuning Swiss Free Electron Laser



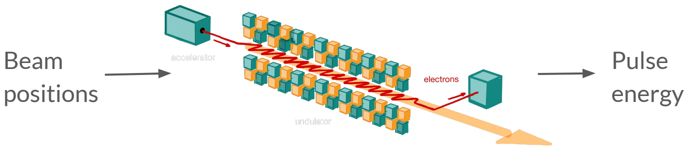
Empirical distribution of true values at acquired points

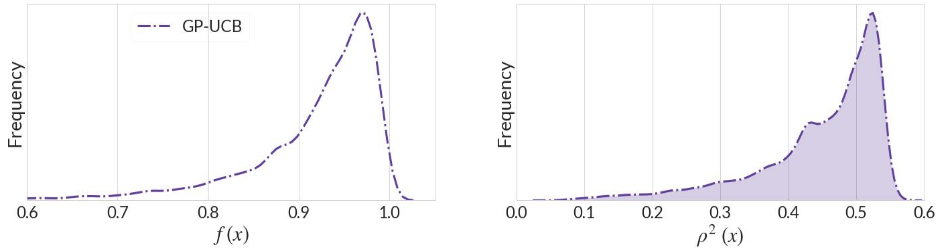
(b) Mean-variance tradeoff (FEL) (c) Cum. regret ($\alpha = 0.5$) (d) Cum. regret ($\alpha = 1$)

GP-UCB tends to query points with higher noise, while RAHBO achieves substantially reduced variance and minimal reduction in mean performance.

RAHBO results for SwissFEL

Tuning Swiss Free Electron Laser



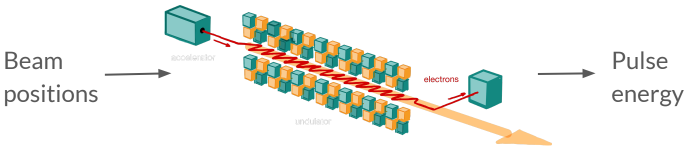


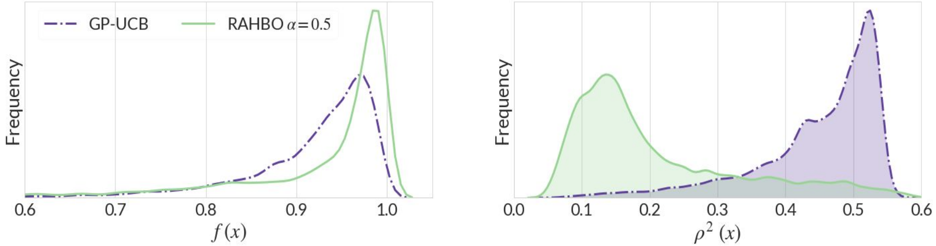
Empirical distribution of true values at acquired points

GP-UCB tends to query points with higher noise, while RAHBO achieves substantially reduced variance and minimal reduction in mean performance.

RAHBO results for SwissFEL

Tuning Swiss Free Electron Laser





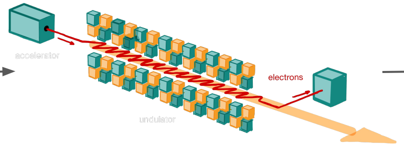
Empirical distribution of true values at acquired points

GP-UCB tends to query points with higher noise, while RAHBO achieves substantially reduced variance and minimal reduction in mean performance.

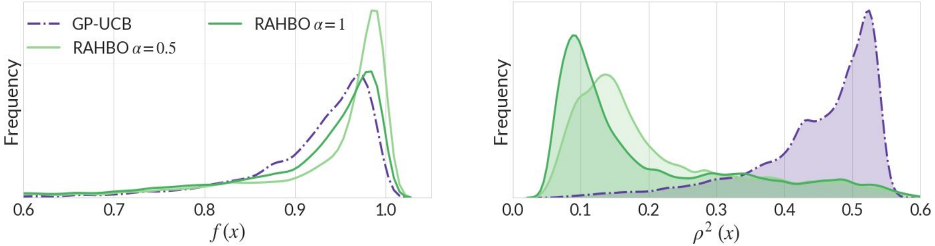
RAHBO results for SwissFEL

Tuning Swiss Free Electron Laser

Beam positions



Pulse energy

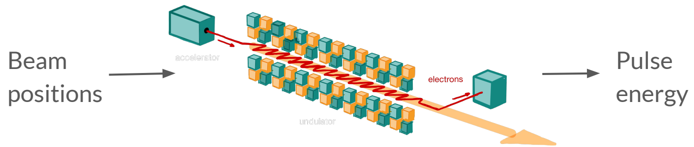


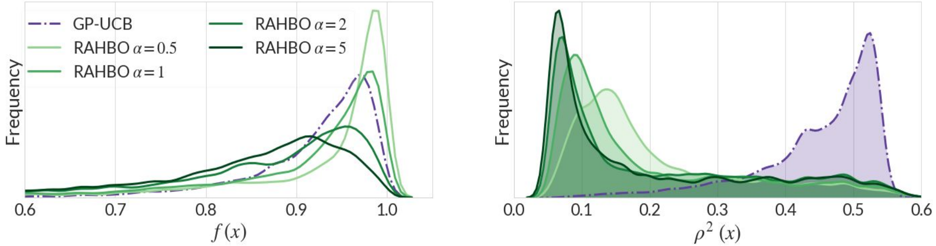
Empirical distribution of true values at acquired points

GP-UCB tends to query points with higher noise, while RAHBO achieves substantially reduced variance and minimal reduction in mean performance.

RAHBO results for SwissFEL

Tuning Swiss Free Electron Laser





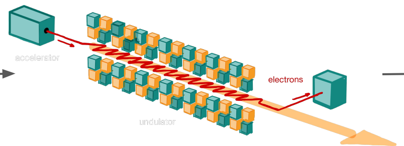
Empirical distribution of true values at acquired points

GP-UCB tends to query points with higher noise, while RAHBO achieves substantially reduced variance and minimal reduction in mean performance.

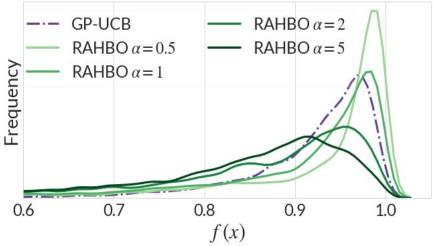
RAHBO results for SwissFEL

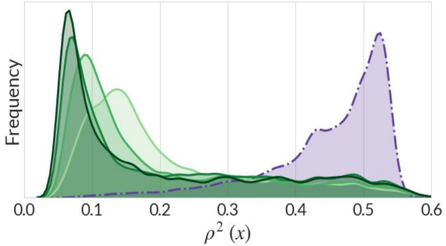
Tuning Swiss Free Electron Laser

Beam positions

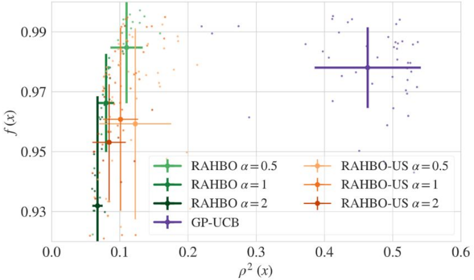


Pulse energy





Empirical distribution of true values at acquired points



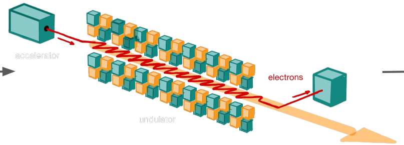
Mean-variance trade off

GP-UCB tends to query points with higher noise, while RAHBO achieves substantially reduced variance and minimal reduction in mean performance.

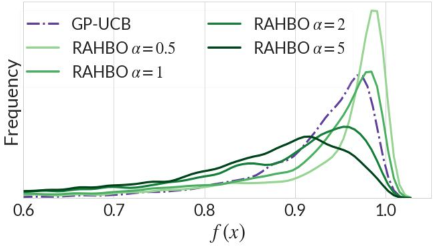
RAHBO results for SwissFEL

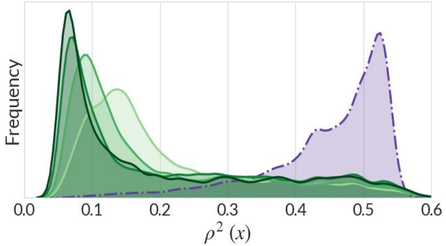
Tuning Swiss Free Electron Laser

Beam positions

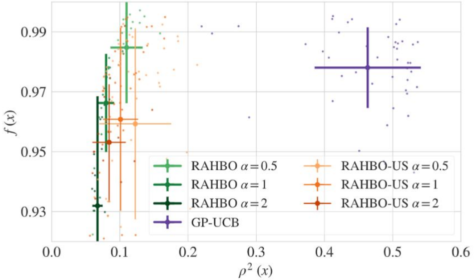


Pulse energy





Empirical distribution of true values at acquired points

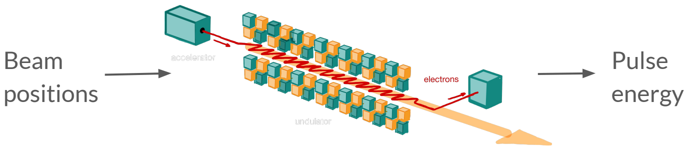


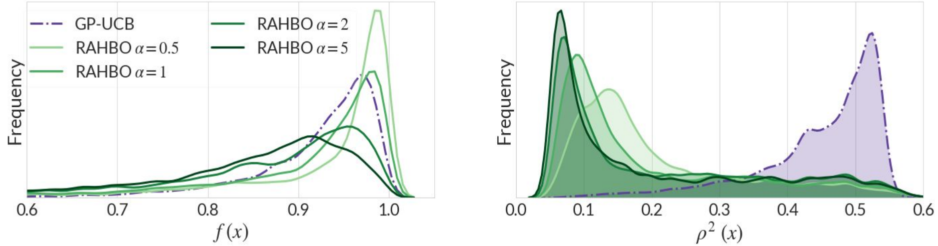
Mean-variance trade off

GP-UCB tends to query points with higher noise, while RAHBO achieves substantially reduced variance and minimal reduction in mean performance.

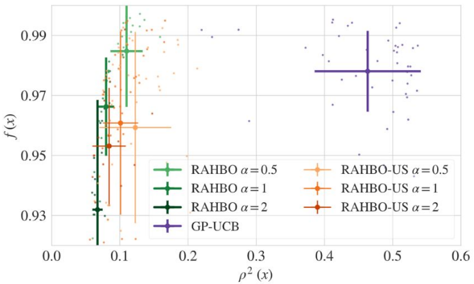
RAHBO results for Swiss FEL

Tuning Swiss Free Electron Laser

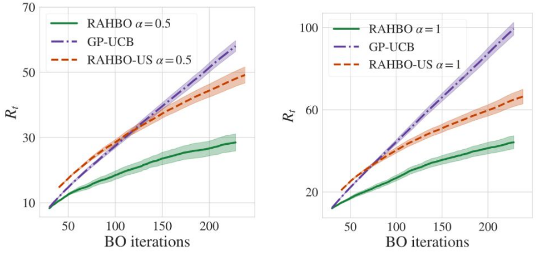




Empirical distribution of true values at acquired points



Mean-variance trade off



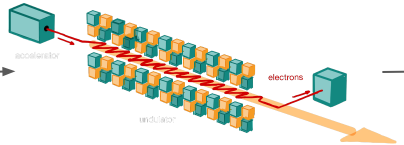
Cumulative regret

GP-UCB tends to query points with higher noise, while RAHBO achieves substantially reduced variance and minimal reduction in mean performance.

Summary

Tuning Swiss Free Electron Laser

Beam positions



Pulse energy

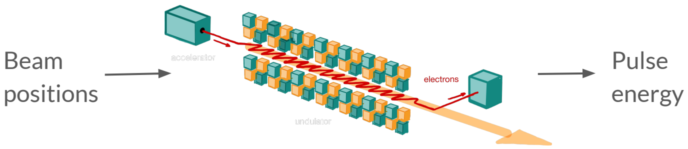
Goal:

- Incorporate risk into exploration-exploitation trade-off

Our contributions:

- Mean-variance approach for Bayesian optimization
- Practical algorithm based on optimism under the face of uncertainty
- Theoretical regret bounds
- Empirical results on SwissFEL simulator and ML model tuning

Tuning Swiss Free Electron Laser



Summary

Goal:

- Incorporate risk into exploration-exploitation trade-off

Our contributions:

- Mean-variance approach for Bayesian optimization
- Practical algorithm based on optimism under the face of uncertainty
- Theoretical regret bounds
- Empirical results on SwissFEL simulator and ML model tuning

Summary

Tuning Swiss Free Electron Laser

Beam positions

Pulse energy

Goal:

- Avoid cost of failure due to noisy realizations in high-stakes applications
- Incorporate risk into exploration-exploitation trade-off

Drop by our poster for more details :)

Our contributions:

- Mean-variance approach for Bayesian optimization
- Practical algorithm based on optimism under the face of uncertainty
- Theoretical regret bounds
- Empirical results on SwissFEL simulator and ML model tuning

Paper ID 26309