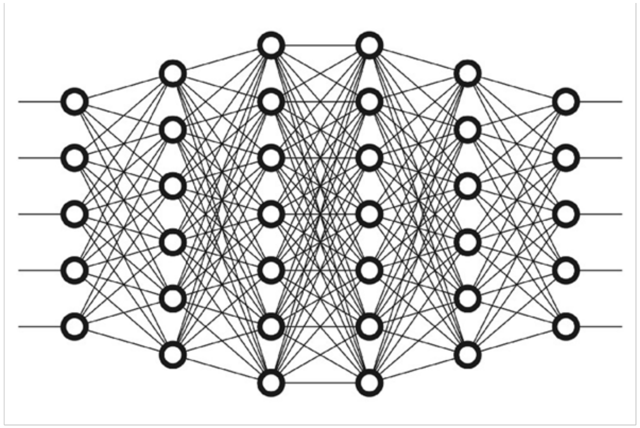
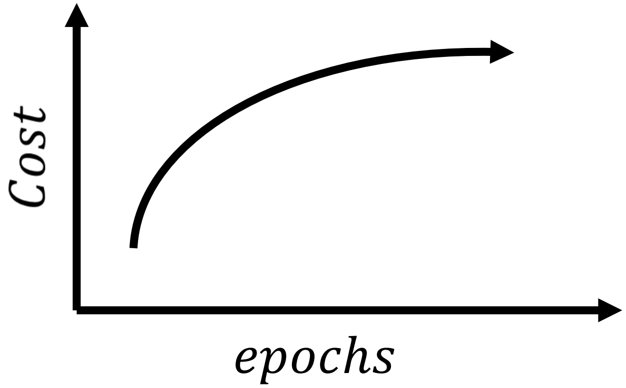
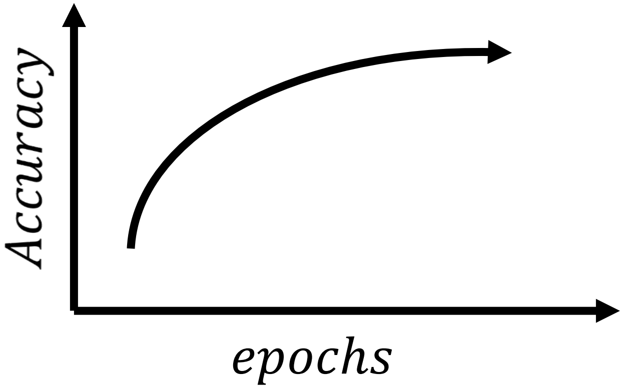
Multi-Fidelity Bayesian Optimization

Application #1: Auto ML and Hyperparameter Tuning

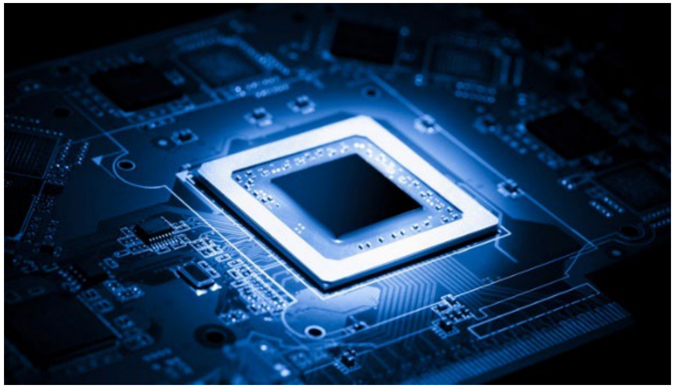


Cost vs. Accuracy trade-offs in evaluating hyperparameter configurations

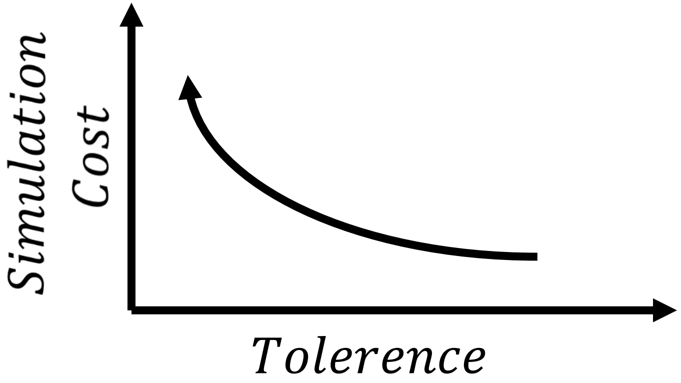


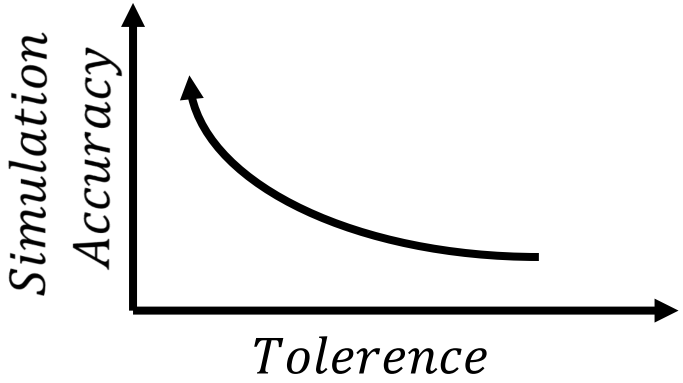


Application #2: Hardware Design via Simulations

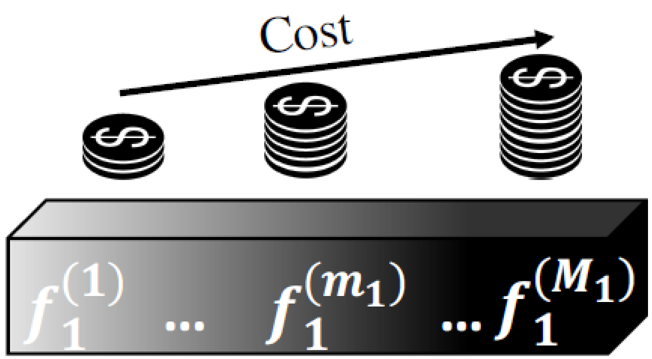


Cost vs. Accuracy trade-offs in evaluating hardware designs

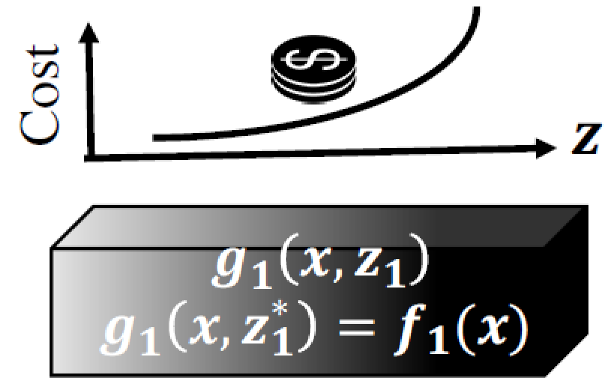




Multi-Fidelity BO: The Problem



Discrete fidelity



Continuous fidelity

- Cost vs. accuracy trade-offs for function approximations
- Continuous-fidelity is the most general case
  - Discrete-fidelity is a special case

- Goal: (approximately) optimize the highest-fidelity function by minimizing the resource cost of experiments

Multi-Fidelity BO: Key Challenges

- Intuition: use cheap (low-fidelity) experiments to gain information and prune the input space; and use costly (high-fidelity) experiments on promising candidates

- Modeling challenge: How to model multi-fidelity functions to allow information sharing?

- Reasoning challenge: How to select the input design and fidelity pair in each BO iteration?

Multi-Fidelity GPs for Modeling

- Desiderata: model relationship/information sharing between different fidelities

- Solution: multi-output GPs with vector-valued kernels

$$k(\{x, z\}, \{x', f\}) = k(x, x')k_F(z, f)$$

- Provides a prediction μ and uncertainty σ for each input and fidelity pair

EI Extension for Multi-Fidelity BO

- Multi-fidelity expected improvement (MF-EI)
  - Extension of EI for multi-fidelity setting
  - Applicable for discrete-fidelity setting

$$EI(x, z) = E\left[\max\left(\tau - y^f\right)\right] cov\left[y^z, y^f\right] C_f / C_z$$

- Acquisition function optimization
  - Enumerate each fidelity z and find the best x fixing z

Information-Theoretic Extensions for Multi-Fidelity BO

$$AF(x) = H(\alpha \mid D) - E_y[H(\alpha | D \cup \{x, y\})]$$
$$= \text{Information Gain}(\alpha; y)$$

- Design choices of α leads to different algorithms
- α as input location of optima x^*
  - Entropy Search (ES) / Predictive Entropy Search (PES)
  - Intuitive but requires expensive approximations
- α as output value of optima y^*
  - Max-value Entropy Search (MES) and it's variants
  - Computationally cheaper and more robust

Information-Theoretic Extensions for Multi-Fidelity BO

$$AF(x,z) = H(\alpha \mid D) - E_y[H(\alpha|D \cup \{x,z,y\})]$$
$$= \text{Information Gain per Unit Cost}(\alpha; y)$$

- Design choices of α leads to different algorithms

- α as input location of optima x^*
  - MF-Predictive Entropy Search (MF-PES)
  - Intuitive but requires expensive approximations

- α as output value of optima y^*
  - MF Max-value Entropy Search (MF-MES)
  - Computationally cheaper and more robust

Continuous-Fidelity BO: BOCA Algorithm

- Two step procedure to select input x and fidelity z separately

- **Selection of input $\boldsymbol{x}$**
  - Optimize UCB $(y^f(x) + \beta\, \sigma^f(x))$ of highest fidelity

- **Selection of fidelity $\boldsymbol{z}$**
  - Reducing fidelity space: $Z_t = \{f\} \cup \{z: \sigma^z(x_{opt}) \geq \gamma(z)\}$
  - If Z_t is not empty, select the cheapest fidelity from it
  - Otherwise, select the highest-fidelity

Code and Software

- Multi-fidelity modeling
  - https://mlatcl.github.io/mlphysical/lectures/05-02-multifidelity.html
- BOTorch
  - https://botorch.org/tutorials/discrete_multi_fidelity_bo

Questions ?