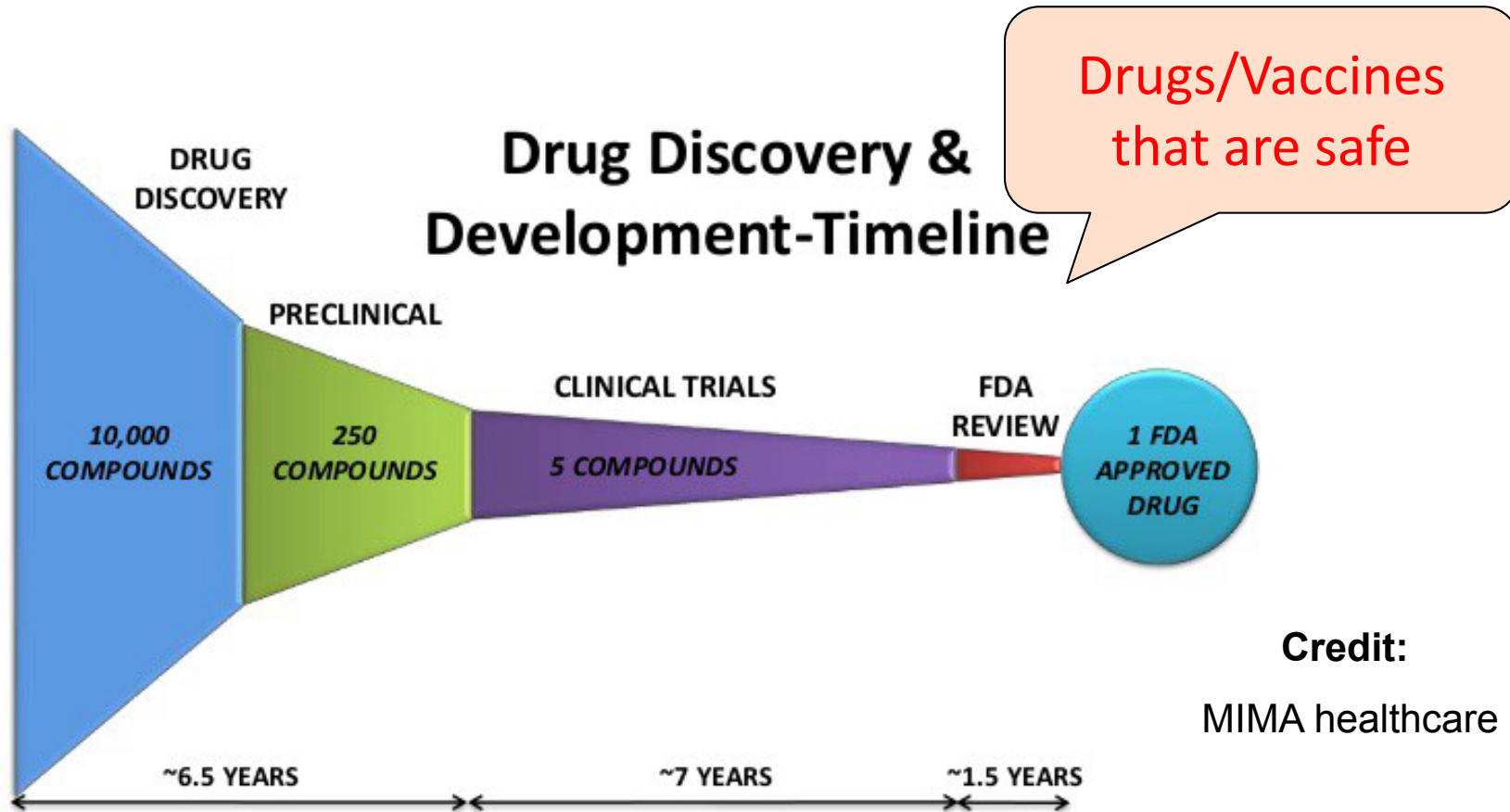


Bayesian Optimization with Black-Box Constraints

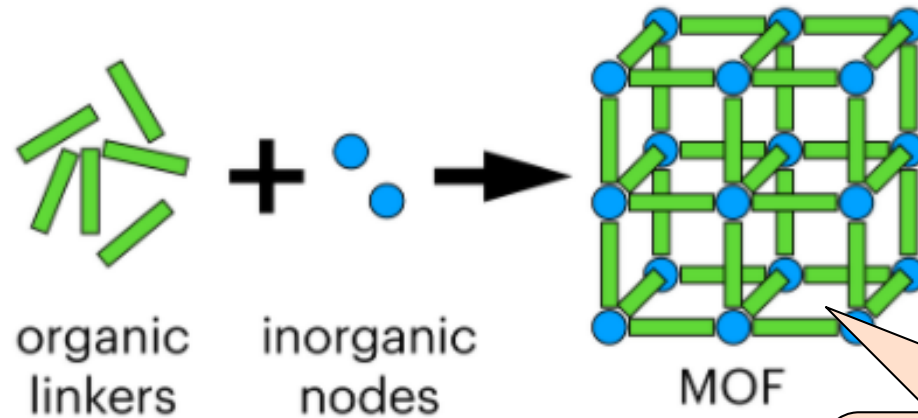


Application #1: Drug/Vaccine Design



- Accelerate the discovery of promising designs

Application #2: Nanoporous Materials Design

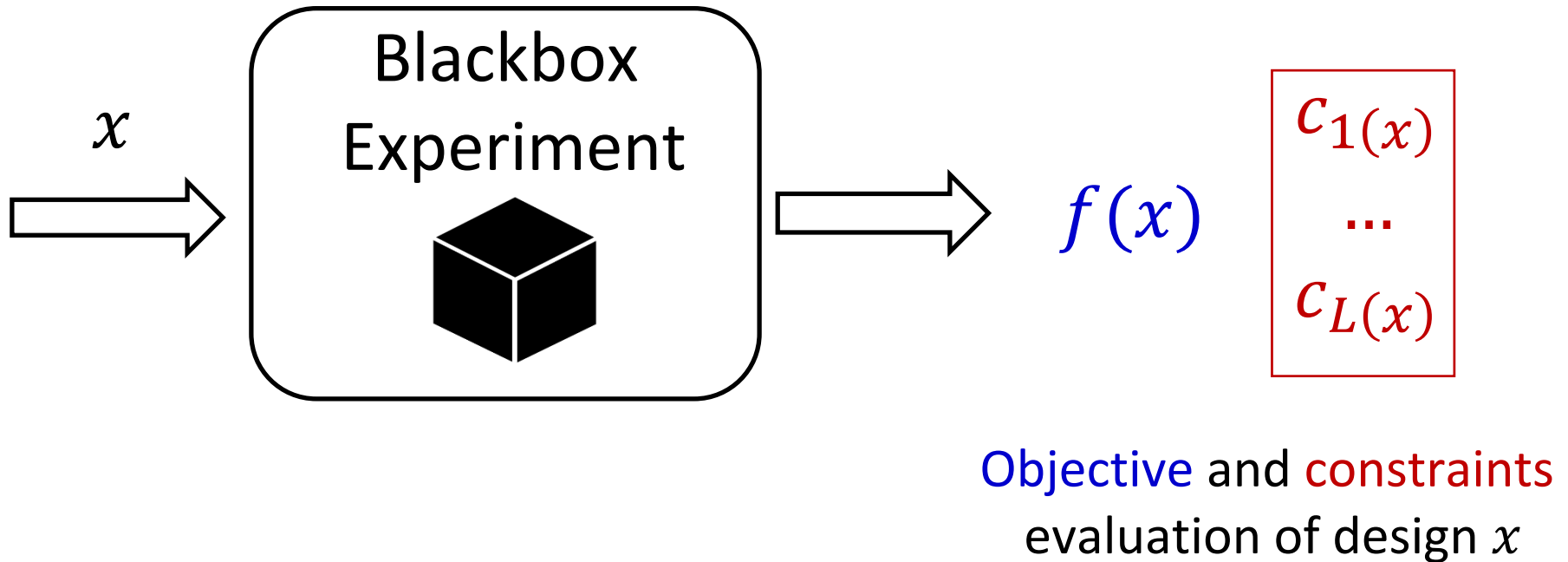


Materials that are
synthesizable

- **Sustainability applications**

- ▶ Storing gases (e.g., hydrogen powered cars)
- ▶ Separating gases (e.g., carbon dioxide from flue gas of coalfired power plants)
- ▶ Detecting gases (e.g., detecting pollutants in outdoor air)

BO with Black-Box Constraints: The Problem



- **Goal:** find the approximate optima from the constrained input space by minimizing the total cost of experiments

BO with Black-Box Constraints: Key Challenges

- **Modeling challenge:** how to model black-box constraints?
 - ▲ GP models will work
- **Reasoning challenge:** How to select the input design guided by the learned models in each BO iteration?
 - ▲ Especially, when no valid inputs (i.e., satisfies constraints) were found from past experiments

Constrained Expected Improvement (c-EI)

- Model each constraint with an independent GP
- Suppose y^{*f} is the best function value from the valid inputs (i.e., satisfies constraints) from past experiments
 - ▲ Assign zero improvement to all invalid inputs

$$EI_c(x) = EI(x) \prod_{i=1}^k P(\tilde{c}_i(x) \geq 0)$$

- When past experimental data does not contain valid inputs: y^{*f} is not defined

$$EI_c(x) = \prod_{i=1}^k P(\tilde{c}_i(x) \geq 0)$$

Constrained Predictive Entropy Search (PESC)

$$\alpha(x) = H(x^*|D) - \mathbb{E}_y[H(x^*|D \cup (x, y))]$$

- Approximating conditioned predictive distribution
 - ▲ First part has a closed-form solution
 - ▲ Second part approximated using expectation propagation

$$\alpha(x) = \log(\sigma_f^2(x)) + \sum_{k=1}^K \log(\sigma_{c_k}^2(x)) - \frac{1}{M} \left\{ \sum_{m=1}^M \log(\sigma_{f_{CPD}}^2(x|x_m^*)) + \sum_{k=1}^K \log(\sigma_{c_{k_{CPD}}}^2(x|x_m^*)) \right\}$$

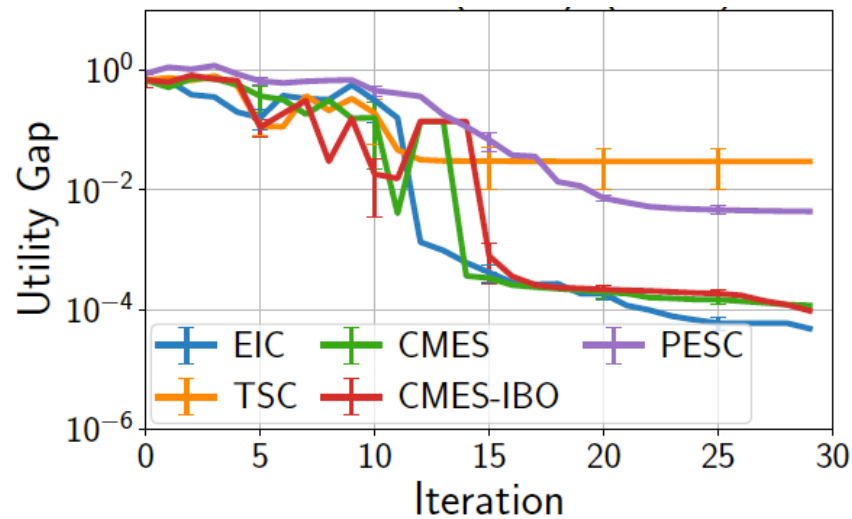
Constrained Max-value Entropy Search (CMES)

$$\alpha(x) = H(y^* | D) - \mathbb{E}_y[H(y^* | D \cup (x, y))]$$

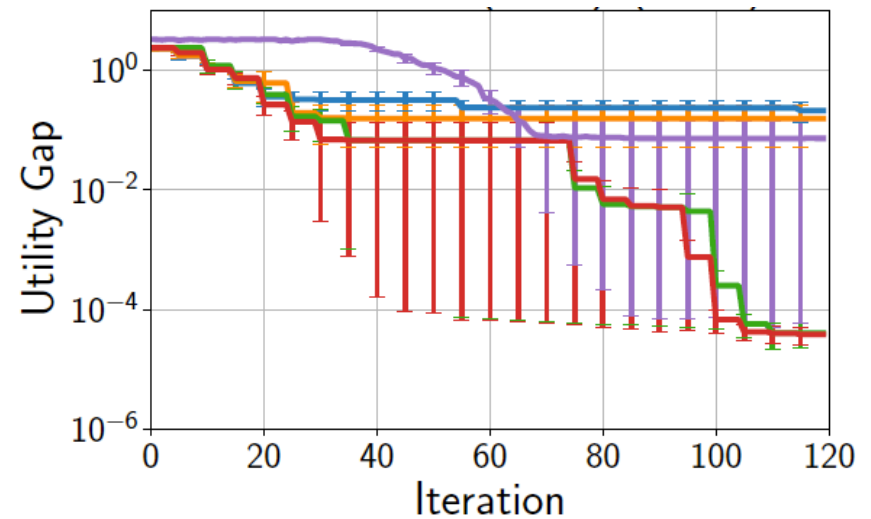
- **Truncated multivariate distribution approximation**
 - ▲ Closed-form expression
 - ▲ **Issue:** can result in negative values
- **Lower bound approximation**
 - ▲ Closed-form expression and overcomes negative values issue
 - ▲ Maximizes the probability of selecting a valid input point when no feasible path is sampled

Constrained Max-value Entropy Search: Results

Gramacy



Hartmann6



Software and Code

- PESC: github.com/HIPS/Spearmint/tree/PESC

Questions ?