Bayesian optimization (BO) is a framework to maximize expensive black-box functions using the following elements:

- **Candidate input** \( x \)
- **Expensive Blackbox Functions**
- **Objectives** \( f_1(x), f_2(x) \)
- **Constraints** \( c_1(x), c_2(x) \)

- **Statistical models** as a prior for the functions: Gaussian processes (GPs) can provide prediction \( \mu(x) \) and uncertainty via variance \( \sigma(x) \)
- **Acquisition function** to score the utility of evaluating input \( x \)
- **Optimization procedure** to select the best input \( x \) for evaluation

**Prior Work and Our Contributions**

- **Drawbacks of existing methods**
  - Does not handle constraints
  - Genetic algorithms assume that the functions are cheap to evaluate and require a very large number of evaluations
  - Scalarization: relies on random scalars that can be sub-optimal
  - Hypervolume improvement: not scalable for high-dimensional input spaces and large number of objective functions
- **Uncertainty and information theory**
  - They either maximize information gain about the optimal Pareto set \( X^* \) and rely on approximating a very expensive and high-dimensional distribution or minimize the uncertainty over a finite set of points.

**Our Approach:**

- **MESMOC framework** selects the candidate input \( x \) for evaluation that maximizes the information gain about the optimal Pareto front \( Y^* \)
- Equivalent to expected reduction in entropy over the Pareto front \( Y^* \)
- Relies on a computationally cheap and low-dimensional \( m, k \ll m \) distribution, where \( k \) is the number of objectives

**Key advantages of MESMOC**

- Robust to the number of samples for AF computation
- Scalable for high-dimensions via output space entropy search
- Tight approximation with closed-form expression
- Two real-world applications to show the effectiveness of our algorithm

**MESMOC Algorithm**

1. **Posterior estimation**
   - Sample a set of optimal pareto fronts \( Y^* \) using functions and constraints sampled from models
   - Define the acquisition function \( \alpha_t(x) = I(\{x, y\}; Y^* | D_t) \)
   - Evaluate the functions \( f_1, f_2, c_1, c_2 \) at \( x_t = \arg \max \alpha_t(x) \)
   - **St.** \( \mu_{c_1}, \mu_{c_2} \)

2. **Input space entropy-based acquisition function**
   - \( \alpha(x) = I(\{x, y\}, \mathcal{X}^* | D) \)
   - Requires approximation

3. **Output space entropy-based acquisition function**
   - \( \alpha(x) = H(\{x, y\}, \mathcal{X}^* | D) \)
   - Sum of truncated Gaussians

**How to sample \( Y^* \)?**

- Sample functions from posterior GPs based on random Fourier features sampling procedure. Approximate each GP prior as \( f_i = \phi(x)^T \theta_i \) and \( \tilde{c}_i = \phi(x)^T \theta_c \) where \( \theta \sim N(0, I) \)
- Solve a cheap constrained multi-objective optimization problem over the sampled functions and constraints \( f_1, f_2, \tilde{c}_1, \tilde{c}_2 \) to compute sample Pareto fronts
- For each function and constraint, select the maximum-value in the cheap Pareto front as an upper bound for the truncated Gaussian

**MESMOC’s Acquisition Function**

\[
\alpha(x) = \frac{1}{s} \sum_{t=1}^{s} \frac{1}{20} \left[ \gamma_{t}^{(s)}(x) \Phi^{(s)}(x) - \ln \Phi^{(s)}(x) \right] + \frac{1}{s} \sum_{t=1}^{s} \frac{1}{2} \left[ \gamma_{t}^{(s)}(x) \Phi^{(s)}(x) - \ln \Phi^{(s)}(x) \right]
\]