Probability of Matching for Batch Multi-Objective Bayesian Optimization

Mingqian Li
Sina Hossein Zadeh
Raymundo Arroyave
Xiaoning Qian

Texas A&M University

ICML 2026

Main objective of Multi-objective optimization

- Multi-objective BO aims to recover the Pareto optimal set:

$$\mathcal{X}^* = \{\boldsymbol{x} \in \mathcal{X} : \nexists \boldsymbol{x}' \in \mathcal{X} \text{ such that } f(\boldsymbol{x}') \succ f(\boldsymbol{x})\}.$$

- In batch MOBO, each iteration selects

$$\boldsymbol{X} = \{\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(q)}\}.$$

- A good batch should be both:
  - high-quality: close to the Pareto optimal set;
  - diverse: covering different Pareto optimal regions.
- Problem: Existing methods cannot adapt the diversity-quality balance to match the true Pareto-set distribution.

Main idea: Probability of Matching

- We directly model the probability that the acquired batch matches the true Pareto set:

$$P(\boldsymbol{X} = \mathcal{X}^*).$$

- Decompose matching into Pareto optimality and coverage:

$$P(\boldsymbol{X} = \mathcal{X}^*) = P(\boldsymbol{X} \subseteq \mathcal{X}^*)P(\mathcal{X}^* \subseteq \boldsymbol{X} \mid \boldsymbol{X} \subseteq \mathcal{X}^*).$$

- $P(\boldsymbol{X} \subseteq \mathcal{X}^*)$: selected candidates are Pareto optimal.
- $P(\mathcal{X}^* \subseteq \boldsymbol{X} \mid \boldsymbol{X} \subseteq \mathcal{X}^*)$: selected batch covers the Pareto set.

Pareto optimality and coverage

- Pareto optimality:

$$P(\boldsymbol{X} \subseteq \mathcal{X}^*) = \prod_{\boldsymbol{x}' \in \boldsymbol{X}} P\left(f(\boldsymbol{x}') \not\prec f(\boldsymbol{x})\right).$$

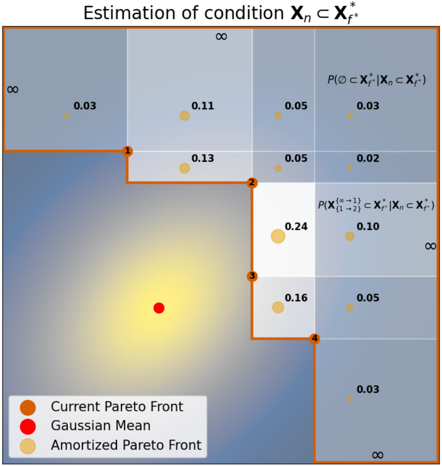
- Coverage estimator:

$$P(\mathcal{X}^* \subseteq \boldsymbol{X}) \approx \frac{1}{\mu(\mathcal{X})} \int_{\mathcal{X}} \mathbb{1}\left(\exists \boldsymbol{x}' \in \boldsymbol{X} \text{ s.t. } f(\boldsymbol{x}) \preceq f(\boldsymbol{x}')\right) d\mu(\boldsymbol{x}).$$

- Two practical variants:
  - POMH: hypervolume-based coverage;
  - POMC: CDF-based coverage.

Pareto optimal region sampling

- Standard posterior sampling may waste samples in regions with no Pareto-front potential.
- We sample only from Pareto-possible boxes.
- This keeps both backbone construction and Pareto search global.



Pareto optimal region sampling

- For the k-th box $l_k < Y < u_k$, define

$$\alpha_{k,m} = \frac{l_{k,m} - \mu_m}{\sigma_m}, \qquad \beta_{k,m} = \frac{u_{k,m} - \mu_m}{\sigma_m}.$$

- Box probability:

$$p_k = \prod_{m=1}^{M} \left[\Phi(\beta_{k,m}) - \Phi(\alpha_{k,m})\right].$$

- Box expected location:

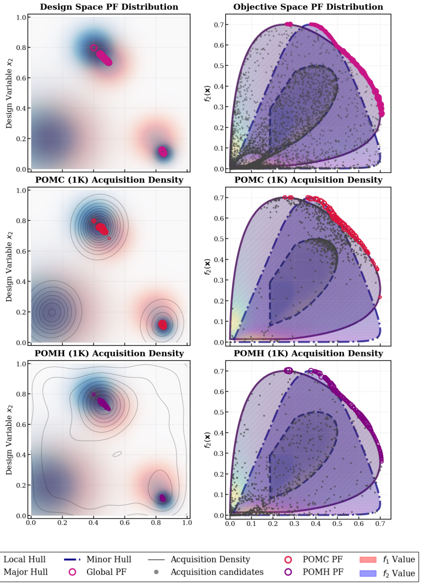
$$\bar{Y}_k = \mu + \sigma \frac{\phi(\alpha_k) - \phi(\beta_k)}{\Phi(\beta_k) - \Phi(\alpha_k)}.$$

- Weighted box locations define a multinomial distribution:

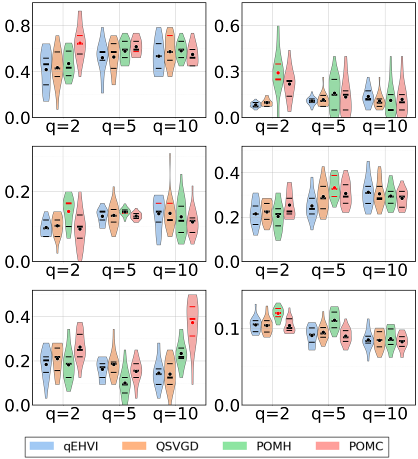
$\bar{Y}_k$ is selected with probability p_k, $k \sim \mathrm{Multinomial}(p_1, \ldots, p_K)$

POMH and POMC

- POMH gives stronger uniform Pareto-front coverage in objective space.
- POMC emphasizes denser and easier-to-sample Pareto regions.
- Both recover disconnected Pareto regions better than local oversampling.



- Six material properties: SFE, C11, HC, TC, SR, RTD.
- Tasks include bi-objective, tri-objective, and six-objective optimization.
- Metric: Pareto optimal solution rediscovery ratio.
- POM consistently rediscovers more Pareto optimal alloy compositions.



Conclusion

- POM directly targets the desired batch event:

$$P(\boldsymbol{X} = \mathcal{X}^*) = P(\boldsymbol{X} \subseteq \mathcal{X}^*)P(\mathcal{X}^* \subseteq \boldsymbol{X} \mid \boldsymbol{X} \subseteq \mathcal{X}^*).$$

- It combines:
  - Pareto optimality: build the Pareto-front backbone.
  - Pareto set coverage: distribute solutions along the backbone.
  - Pareto-possible region sampling: search globally.
- POMH improves objective-space uniformity.
- POMC reflects true Pareto front density.
- Both improve performance under:
  - Small batch sizes
  - Multiple Pareto regions