Optimizing High Dimensional Physics Simulations with Composite Bayesian Optimization

Wesley Maddox[1], Qing Feng[2], Max Balandat[2]

[1] NEW YORK UNIVERSITY [2] Meta

NEURAL INFORMATION PROCESSING SYSTEMS

Paper: https://arxiv.org/abs/2111.14911

Methods

We use Matheron's rule to draw samples from the HOGP posterior.

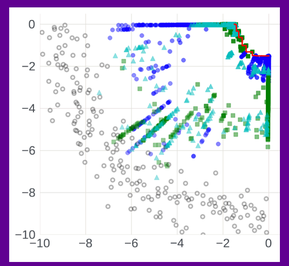
$f^* \mid (Y + \epsilon = y) = f^* + K_{x^*X}(K_{XX} + \sigma^2 I)^{-1}(y - Y - \epsilon)$,

Where $(f^*, Y) \sim \mathcal{N}(0, K_{joint})$

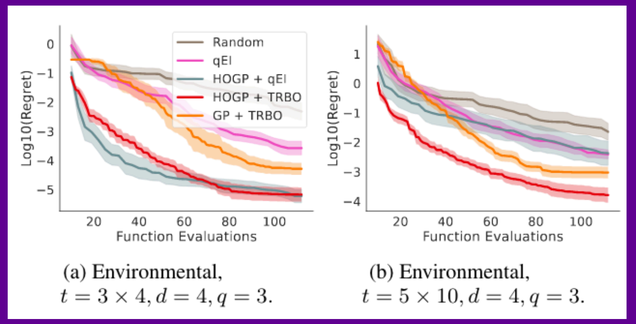
Exploiting Kronecker structure to enable n^3 + t^3 sampling time [1].

We use HOGPs [2] which extend MTGPs to tensor outputs.

To scale to high-dimensions, we use MORBO [2], which uses trust regions and discrete hyper volume computations to scale BO to large dimensions and several objectives.



Each trust region explores a different section of the Pareto front. From [2].



(a) Environmental, $t = 3 \times 4, d = 4, q = 3$.

(b) Environmental, $t = 5 \times 10, d = 4, q = 3$.

The HOGP + TRBO combination tends to outperform non-composite approaches on single-objective problems.

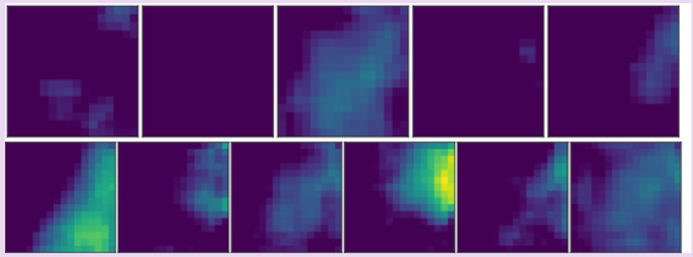
References:

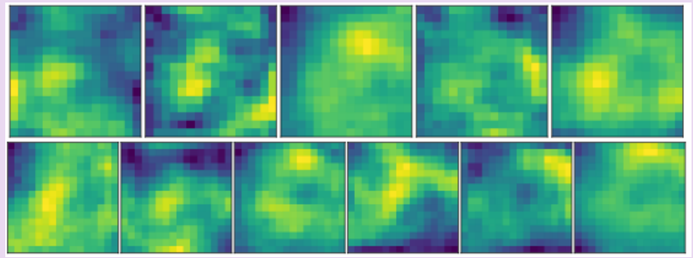
[1] Bayesian Optimization with High Dimensional Outputs, Maddox et al, NeurIPS, '21.

[2] Multi-Objective Bayesian Optimization over High-Dimensional Search Spaces, Daulton et al, https://arxiv.org/abs/2109.10964, '21+

[3] Scalable High-Order Gaussian Processes, AISTATS, '19.

Un-optimized images from an optical design problem



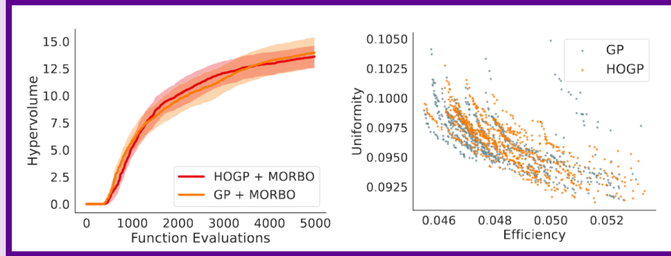


Optimized images

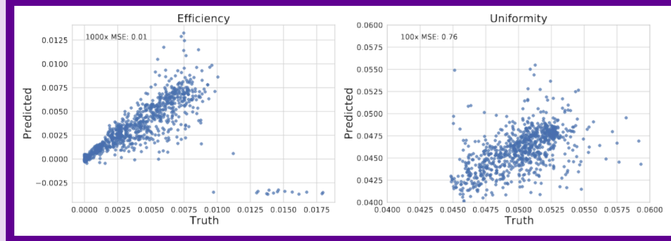
Use Matheron's rule to sample High Order Gaussian processes (HOGP) alongside multi-objective TRBO to perform scalable Bayesian optimization.

Enables **multi-objective Bayesian optimization** with HOGPs **over tens of thousands of outputs.**

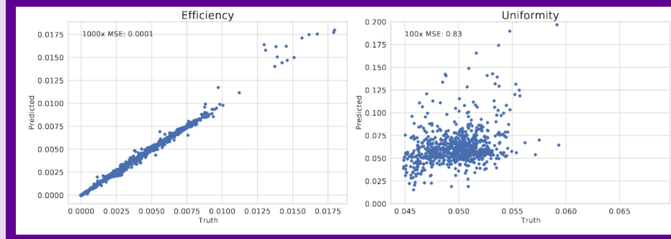
MORBO Experiment



HOGP + MORBO is slightly outperformed by the GP on the optical design problem, despite better exploring the uniformity metric.



This underperformance is because the HOGP is a worse model of the efficiency metric.



By comparison, the GP is a much worse model of the uniformity metric, but is almost perfect when predicting efficiency.

In future: See if there's a mixed composite / non-composite strategy that can outperform GP + MORBO only.