Scalable Combinatorial Bayesian Optimization with Tractable Models

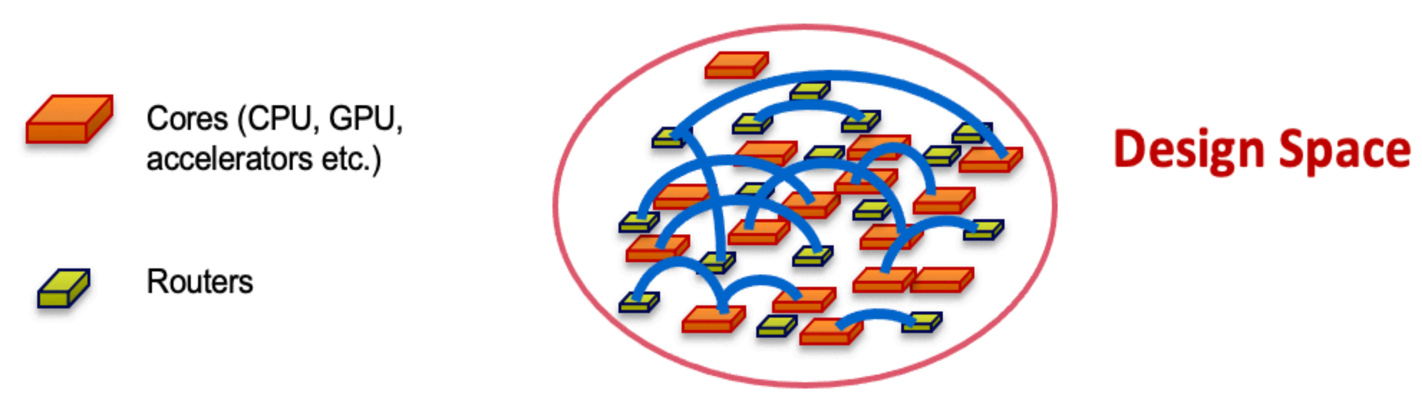
Aryan Deshwal, Syrine Belakaria, Janardhan Rao Doppa
Washington State University

NEURAL INFORMATION PROCESSING SYSTEMS

Introduction

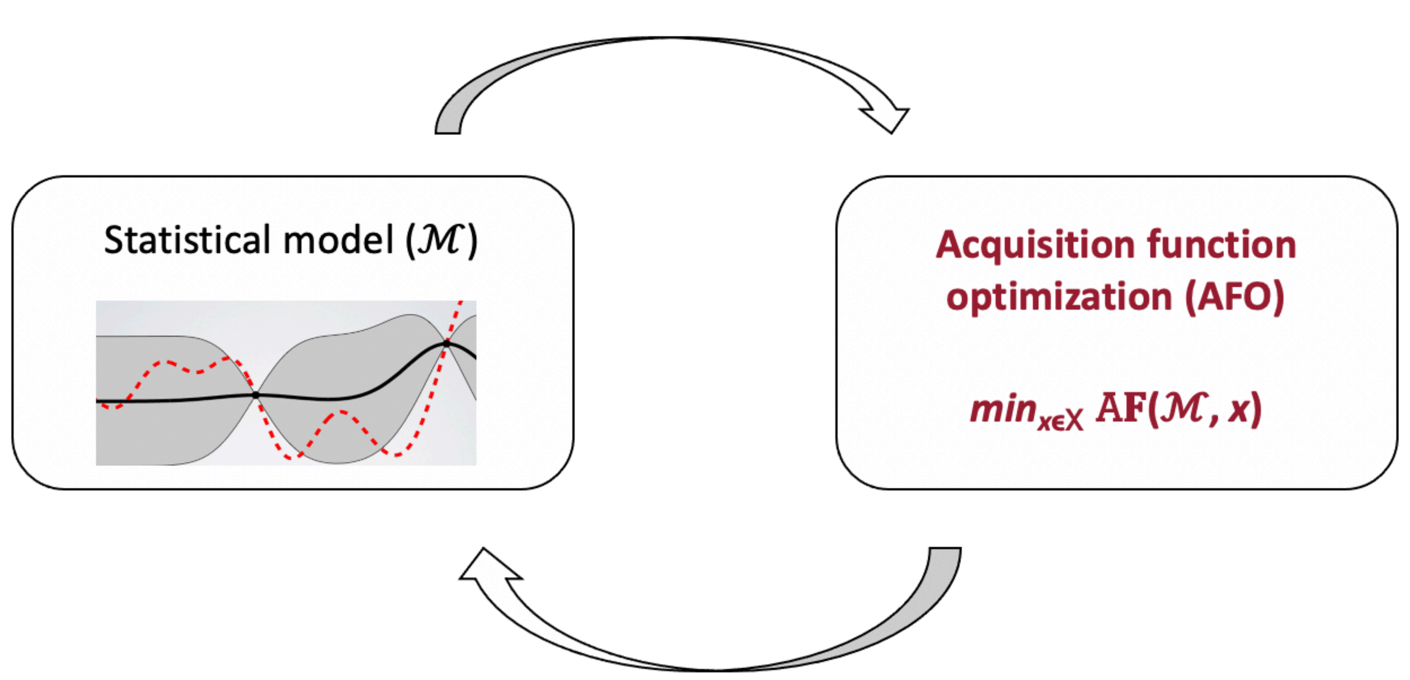
➢ **Problem**: Optimize *expensive black-box* functions over *combinatorial* spaces (e.g. sets, sequences, graphs)

➢ Many applications in science and engineering. For example, multi-core chip design, materials design etc.



➢ **Goal**: find a structure that optimizes a black-box function by *minimizing the number of function evaluations*

Bayesian Optimization Framework



➢ **Key challenge**: Balancing tradeoff between Statistical model and Acquisition function optimization

Parametrized Submodular Relaxation

➢ **Statistical model** (Baptista R. And Poloczek M.)

$$- f_\alpha(x) = \alpha_0 + \sum_j \alpha_j x_j + \sum_{i,j>i} \alpha_{ij} x_i x_j;\ x \in \{0,1\}^n$$

➢ **Acquisition function** : Thompson Sampling

➢ **Acquisition function optimization** (AFO)

$$\min_{x \in \{0,1\}^n} f_{\alpha_t}(x)$$

$$\min_{x \in \{0,1\}^n} x^T A x + b^T x$$

Binary quadratic programming (BQP) problem

➢ Construct a Λ-parametrized submodular of the BQP objective

$$h_\Lambda(x) + x^T A^- x \quad \leq \quad x^T A x + b^T x$$

➢ Relaxation solvable via scalable & efficient minimum graph cut algorithms

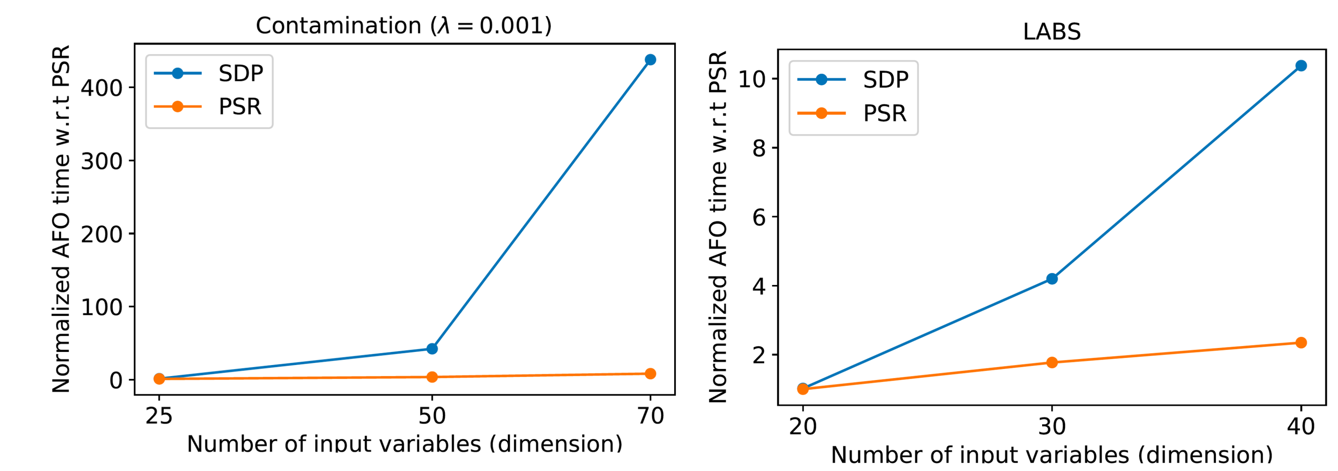
➢ Optimize the relaxation over Λ

$$h_{\Lambda_1}(x) + x^T A^- x \quad \leq \quad h_{\Lambda_2}(x) + x^T A^- x \quad \leq \ \ldots. \qquad x^T A x + b^T x$$

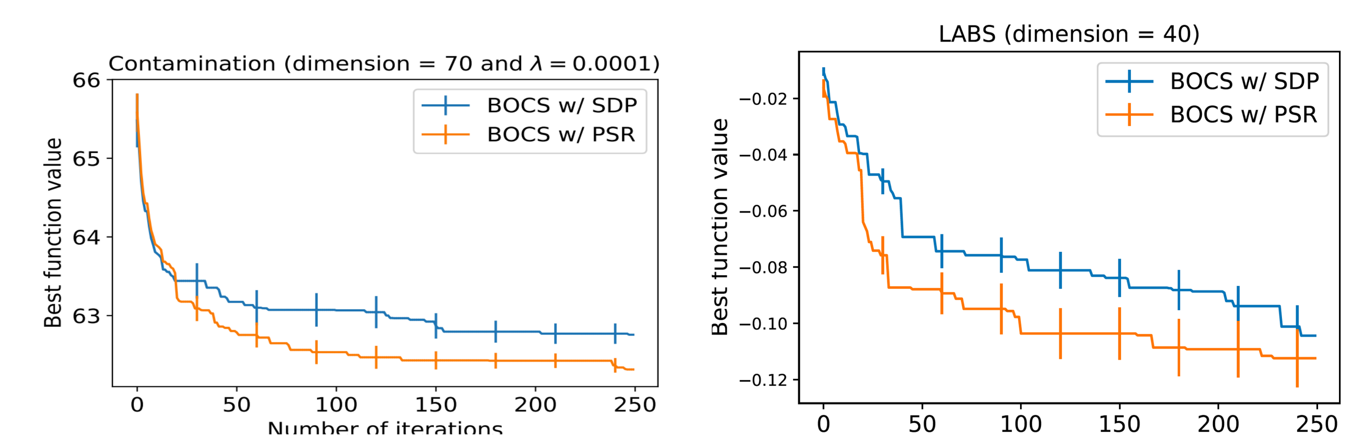
➢ Inspired by work on prescriptive price optimization (Ito S. and Fujimaki R.)

Experimental Results

➢ Scalability
- PSR scales better than SDP



➢ Overall BO performance
- PSR finds better AFO solutions



References

Baptista, R. & Poloczek, M.. (2018). Bayesian optimization of combinatorial structures. Proceedings of the *35th International Conference on Machine Learning*, in PMLR 80:462-471

Ito, S., & Fujimaki, R. (2016). Large-scale price optimization via network flow. In *Advances in Neural Information Processing Systems* (pp. 3855-3863).