Fast Bayesian Optimization with Dimension Scheduling

Imperial College London

Doniyor Ulmasov, Ruth Misener, Marc Deisenroth

Department of Computing, Imperial College London

Abstract

- Data-Efficient, Global Black-box Optimization Algorithm
- Optimizes along a limited set of dimensions, selected randomly, at each iteration
- Conceptually straightforward to parallelize/distribute

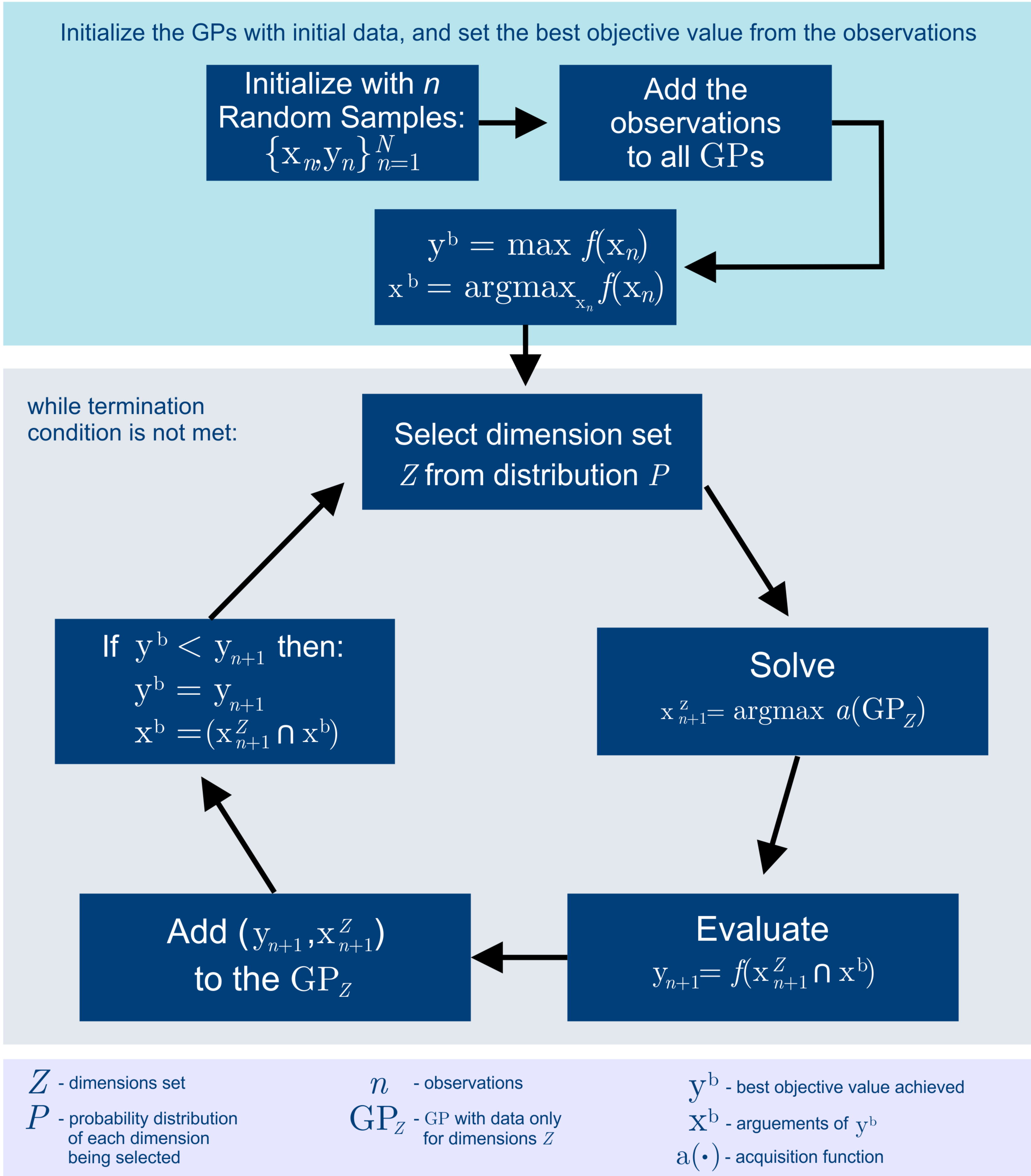
Motivation

- In Bayesian Optimization (BO) we are interested in finding

$$x^* \in \arg\min_{x \in B \in \mathbb{R}^d} f(x)$$

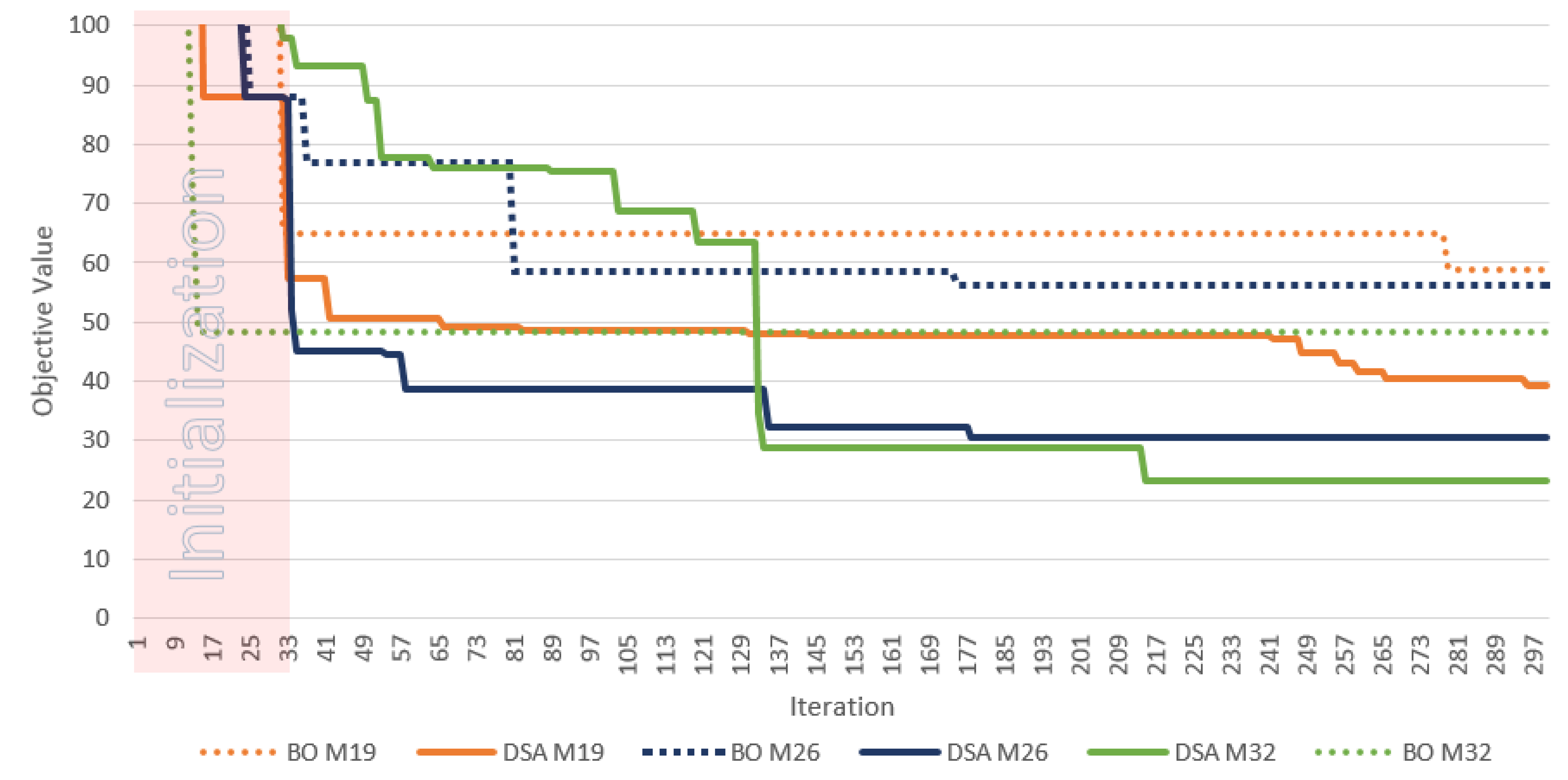
- The black-box function may be non-convex and gradient information is not available
- Random sampling methods too slow or/and computationally expensive
- BO performance computation time degrades as number of observations increases.
- Critical: GP evaluation time when optimizing the fitness function
- DSA optimizes the fitness function only along a small set of dimensions at each iteration.
- Each subset contains its own GP with reduced input dimensions
- As a result each GP contains fewer observation points
- Each iteration the dimension subset is selected randomly from a probability distribution

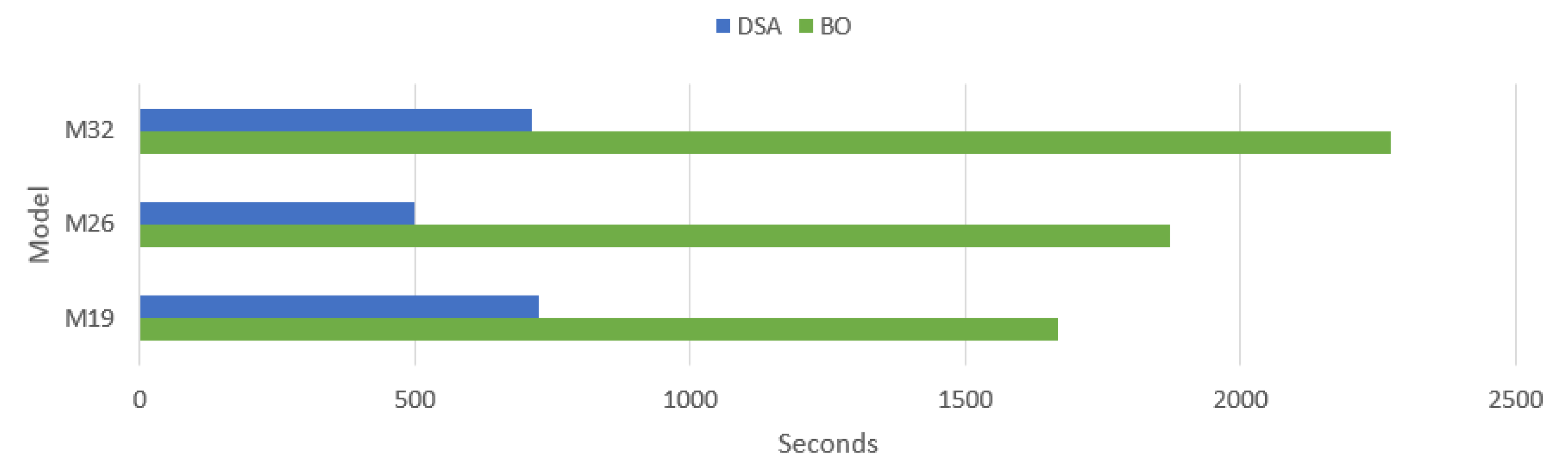
The Algorithm



- The probability distribution can be updated online from the data observed
- The gray section is parallelizable
- Kernels and acquisition functions used in the BO can be used with DSA

Results





Experiment Results based on dynamic models of biological processes:

- Top: DSA achieves lower objective values in all three models presented
- Bottom: (Sequential Implementation) DSA (blue) finishes faster than the BO (green) method

Conclusion

- **Faster** than the traditional BO methods
- **Near-Optimal Objective Value**
- Possible **distributed** or/and **parallelized implementation**