Are All Training Data Useful?
A Empirical Revisit of Subset Selection in Bayesian Optimization

Peili Mao [1] and Ke Li [2]

[1]University of Electronic Science and Technology of China [2]University of Exeter

Background

Due to the sequential decision-making characteristics of Bayesian Optimization, the computational complexity grows exponentially.

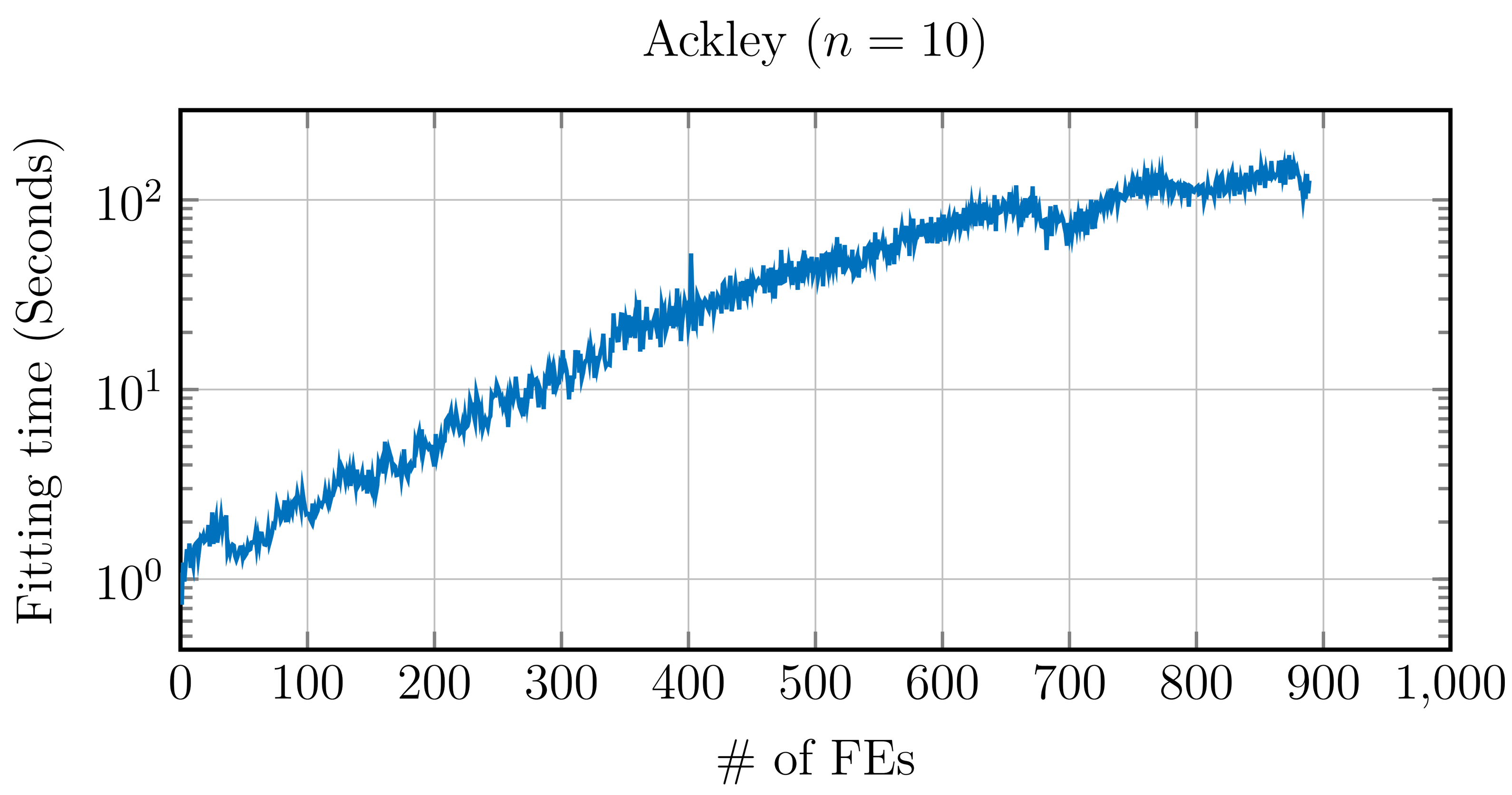


Figure 1. Fit-time of 10-dimensional Ackley optimization problem in the context of BO

Motivation

The impact of computational time reduction techniques in the context of BO have rarely been studied. In our work, we revisit the subset selection methods and approximate GP methods on synthetic problems and hyperparameter optimization problems.

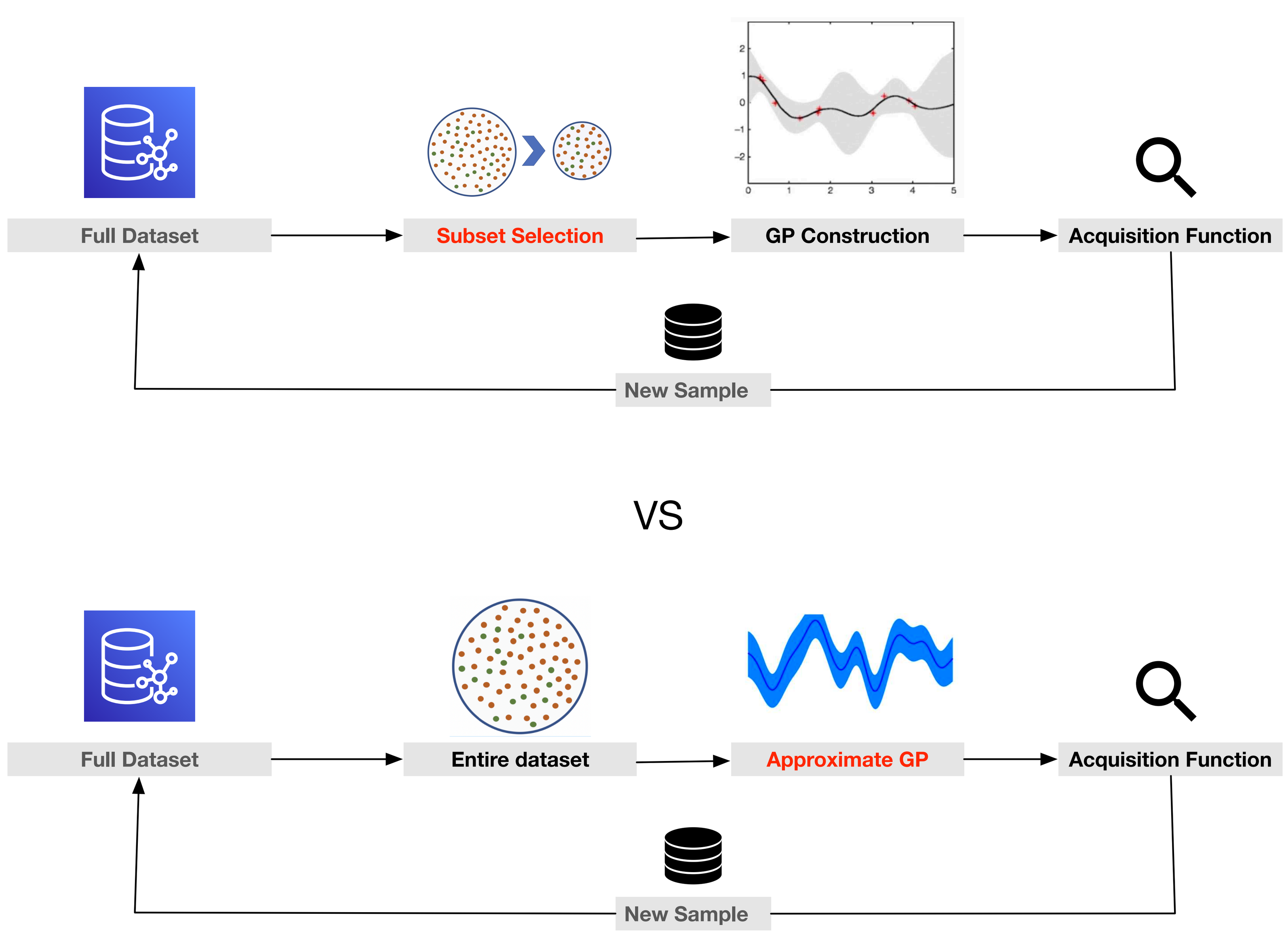
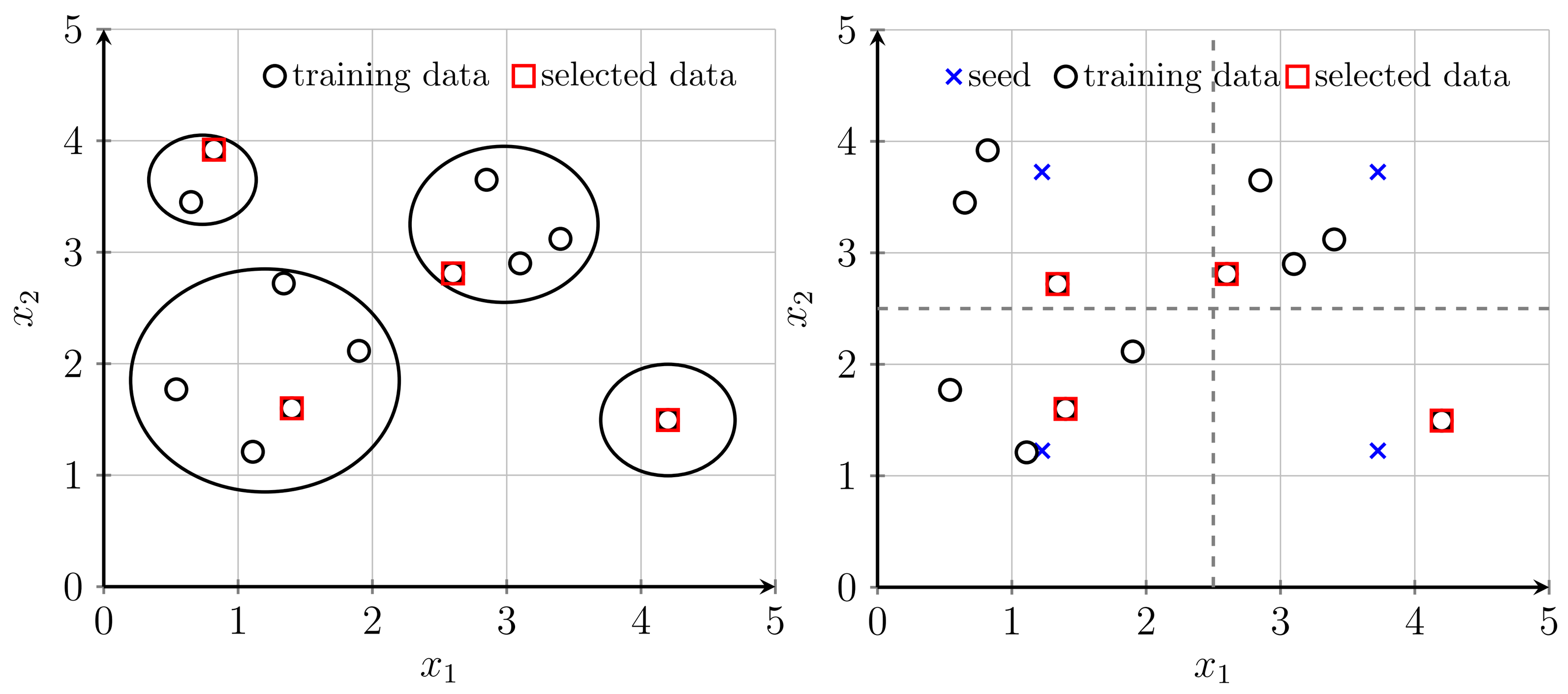


Figure 2. Bayesian optimization framework with subset selection strategy

Subset Selection Strategies

We empirically investigate three simple selection strategies as following:

- Random selection strategy (RS)
- K-means clustering selection strategy (KCS)
- Seed clustering selection strategy (SCS)



(a) K–means clustering selection strategy (b) Seed clustering selection strategy

Figure 3.

Empirical Study

We compare the subset selection strategies against the vanilla BO and GP approximation methods on synthetic problems and hyperparameter optimization problems.

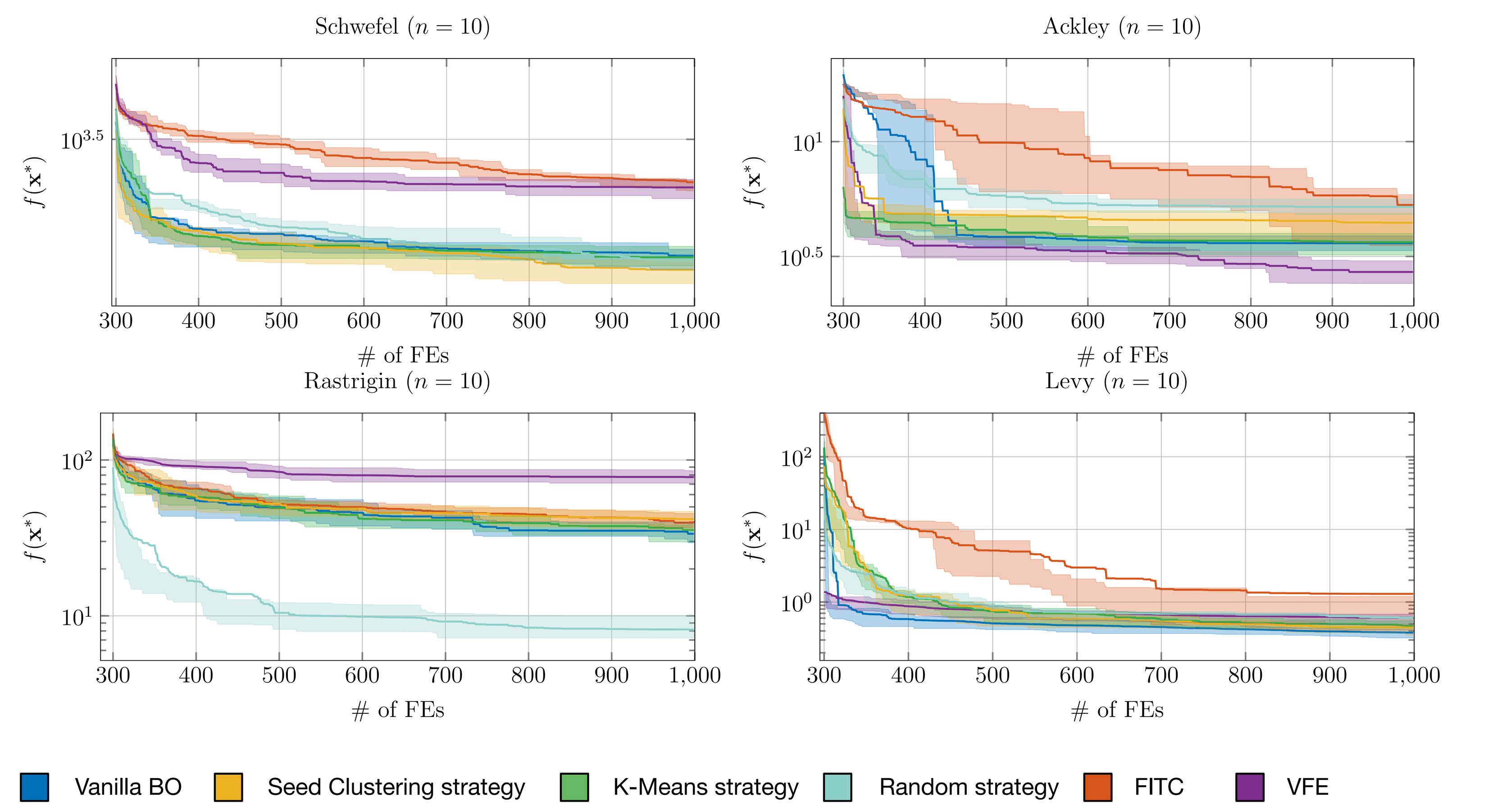


Figure 4. Plots of the convergence trajectories with confidence bounds across the optimization process obtained by different algorithms on synthetic problems

Time Reduction by Given 1000 Function Evaluations

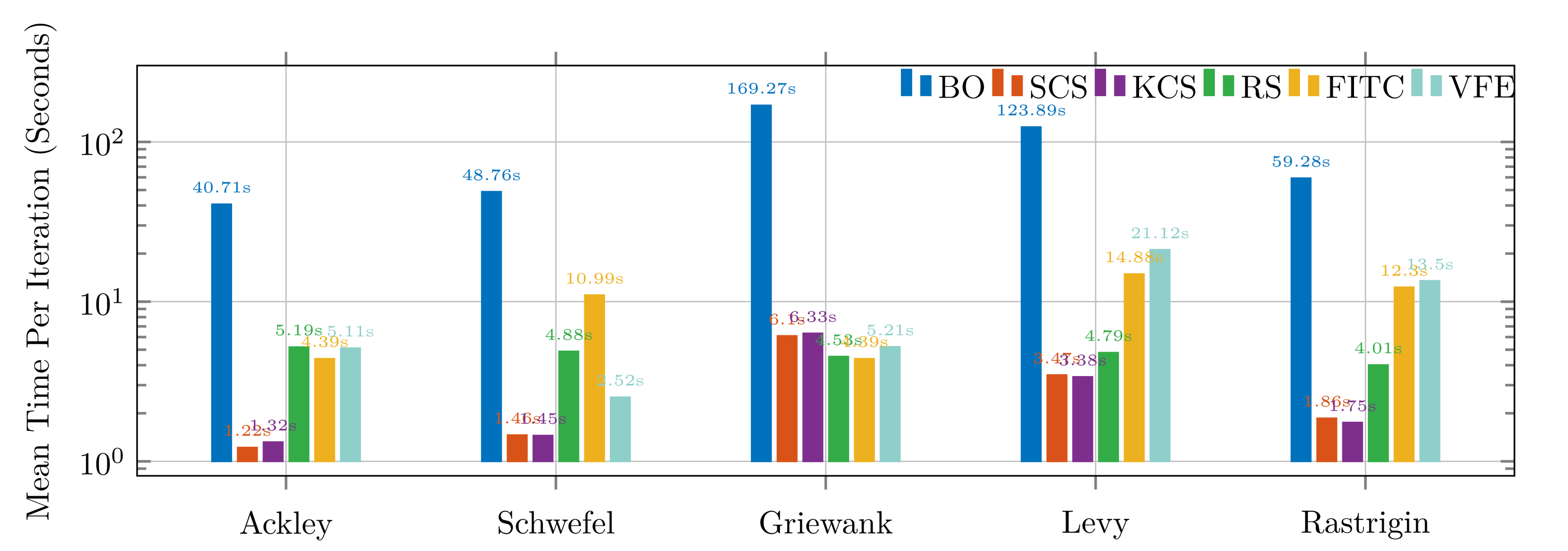


Figure 5. The mean value of computational time per iteration on synthetic problems by give 1000 function evaluations.

Performance on Tuning hyper-parameters of Neural Network

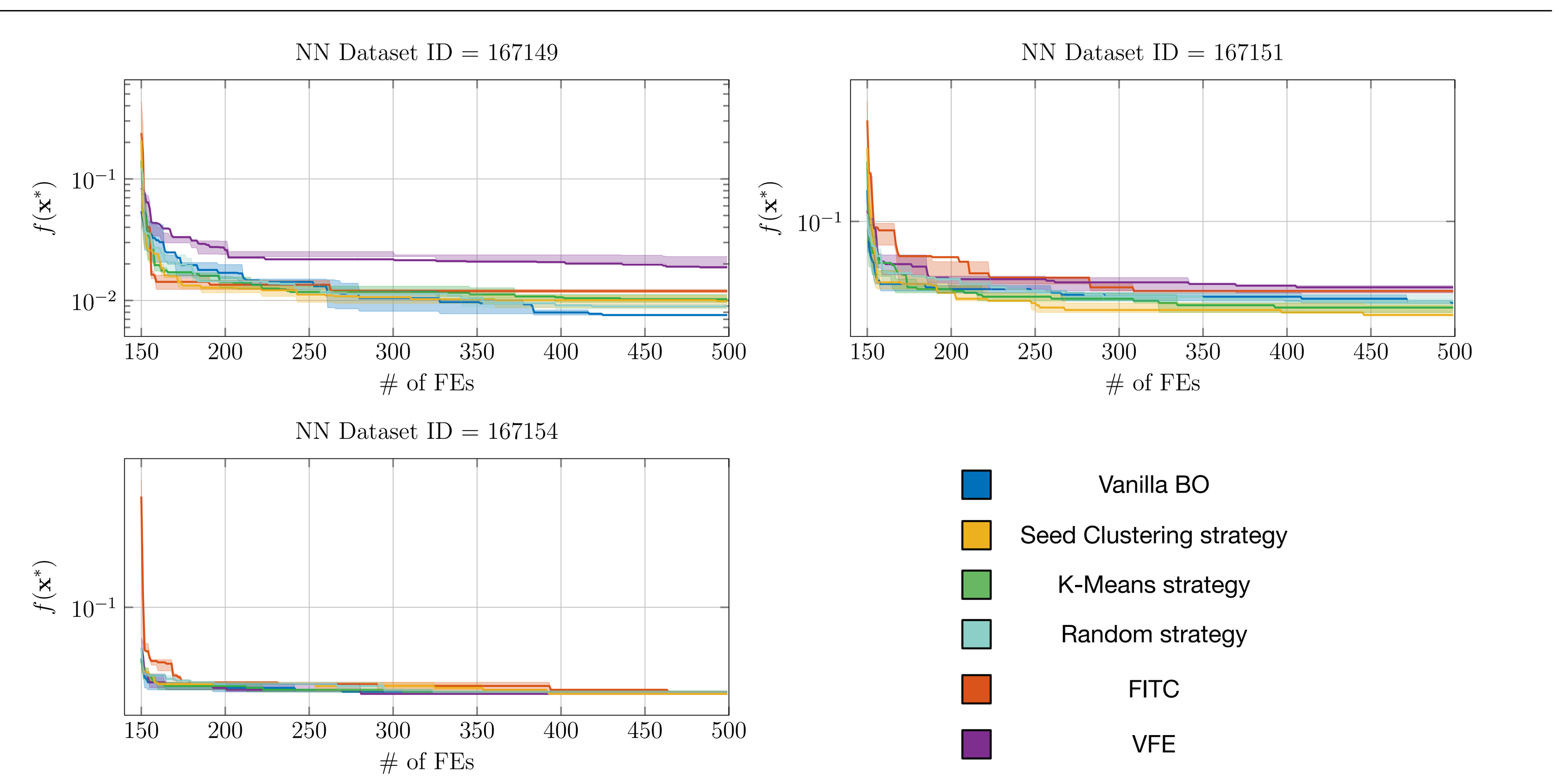


Figure 6. Plots of the convergence trajectories on hyperparameter optimization problems of neural networks

References

[1] Joaquin Quiñonero Candela and Carl Edward Rasmussen. A unifying view of sparse approximate gaussian process regression. *J. Mach. Learn. Res.*, 6:1939–1959, 2005.

[2] Michalis K. Titsias. Variational learning of inducing variables in sparse gaussian processes. In *AISTATS'03: Proc. of the 2009 12th International Workshop on Artificial Intelligence and Statistics*, pages 567–574. JMLR.org, 2009.