Adaptive Ranking and Selection Based Genetic Algorithm for Data-driven Problems

EDWARD P. FITTS DEPARTMENT OF INDUSTRIAL AND SYSTEMS ENGINEERING
NC STATE UNIVERSITY

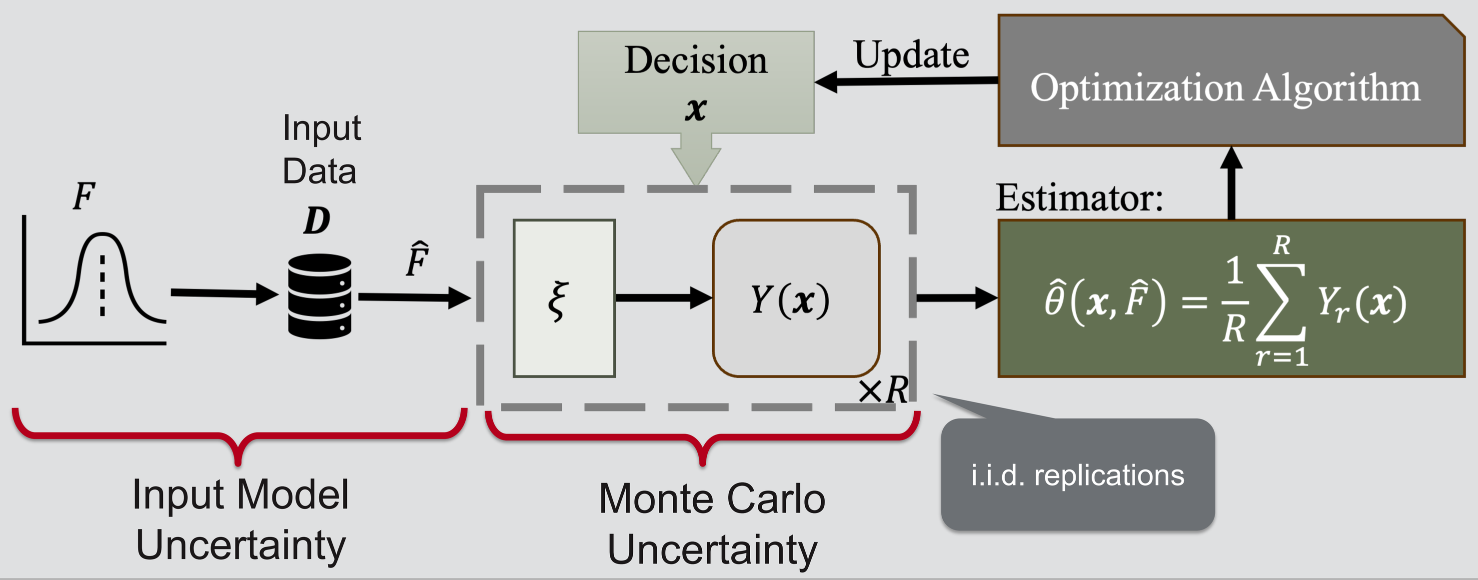
Kimia Vahdat, Ph.D.

Introduction

Simulation Optimization (SO) is an effective tool for solving data-driven problems.

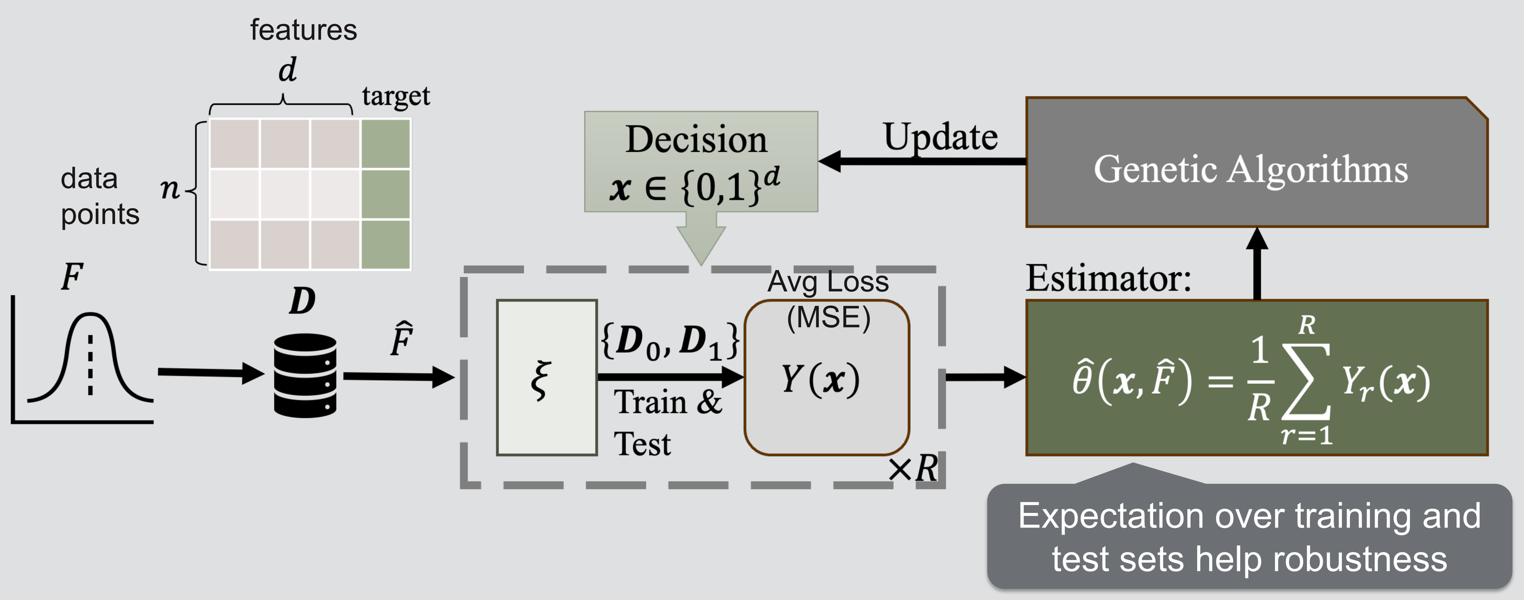
Problem Statement:

$$\min_x \theta(x, F) = \mathrm{E}[Y(x, F)]$$



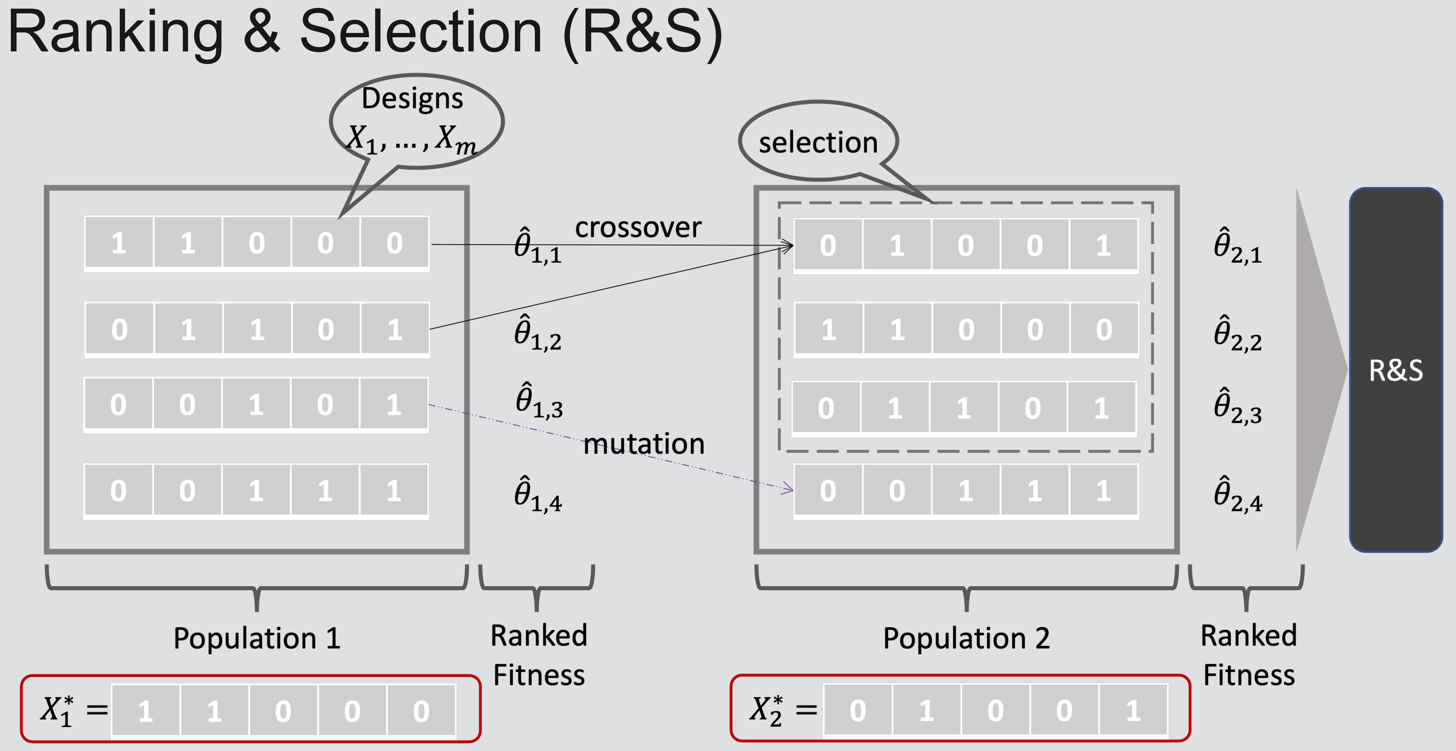
Genetic Algorithms

We focus on feature selection (FS) as an instance of machine learning (ML) optimization:



Genetic Algorithms provide a flexible framework for data-driven stochastic zero-one optimization.

- Ranking & Selection (R&S)

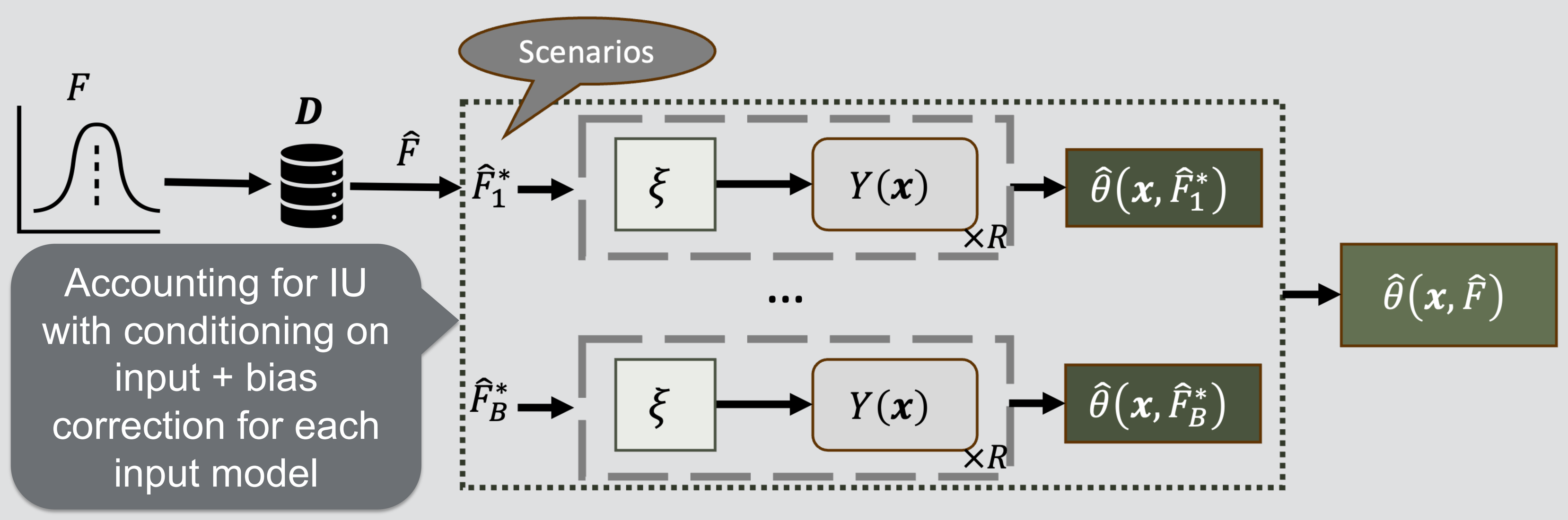


We develop an ***inexpensive bias correction*** method that reduces bias error to $\mathcal{O}(1/n^3)$ while reducing variance by a constant factor, incorporate it into SO for robustness, and couple it with ***adaptive sampling*** to enhance its efficiency.

Bias Estimation

We can write:

$$\hat{\theta}(x, \hat{F}) - \theta(x, F) \approx \mathrm{N}(\underbrace{\textbf{Bias}}_{\mathcal{O}(1/n)}, \underbrace{IU\ Var.}_{\mathcal{O}(1/n)} + \underbrace{MC\ Var.}_{\mathcal{O}(1/R)})$$



The conditional bias corrected estimator becomes:

$$\hat{\theta}^{bc}(x, \hat{F}) = \mathrm{E}\left[\hat{\theta}(x, \hat{F}^*) - bias(\hat{F}^*) \middle| \hat{F}^*\right]$$
$$\approx 2\hat{\theta}(x, \hat{F}) - 2\hat{\theta}(x, \hat{F}^{**}) + \hat{\theta}(x, \hat{F}^{***})$$

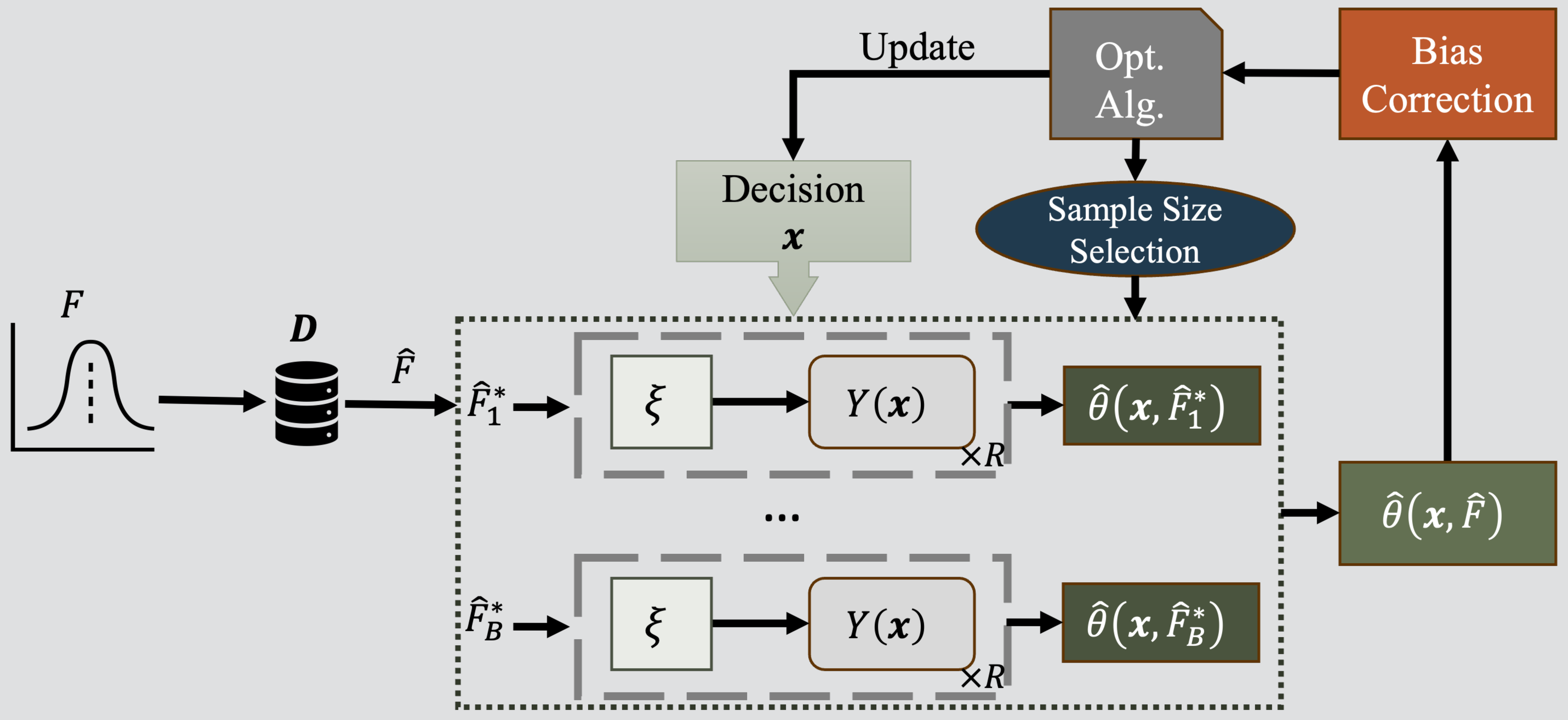
- ✓ Derived from iterative bootstrapping,
- ✓ **It's fast**: number of required inner bootstraps = 1,
- ✓ **Has reduced variance**: Var($\hat{\theta}^{bc}(x, \hat{F})$) = Var($\hat{\theta}^{IB}(x, \hat{F})$) / c.
- ✓ Theorem: asymptotic CI convergence

$$\lim_{B\to\infty} \mathrm{P}\{\theta(x, F) \in [\hat{\theta}^{bc}(x, \hat{F}) \pm \tau^*_{\alpha/2}\sqrt{n}]\} = 1 - \alpha/2$$

Bootstrap quintile

Adaptive Sampling and Budget Allocation

Bias estimation with high precision is expensive!



Increase the computation budget only if:

Optimality gap
$$\left|\frac{\hat{\theta}(x^*, \hat{F})}{\frac{1}{m}\sum_{t=1}^{m} \hat{\theta}(x_t, \hat{F})} - 1\right| \leq \alpha \times \hat{\sigma}_{pop.}$$
Population std. dev.

Which *input model* and *design* should receive extra budget?

✓ Robust Optimal Computing Budget Allocation (R-OCBA)

Numerical Results and Conclusion

We compare three methods:

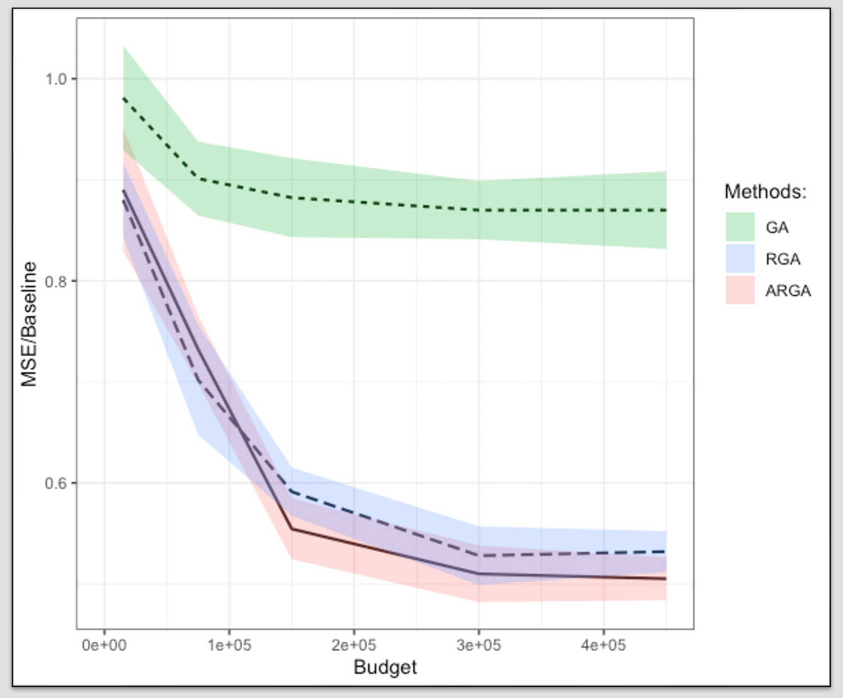
GA | Robust GA | Adaptive RGA

- Y-axis: MSE of selected subset of features divided by MSE of all features on a holdout set
- X-axis: total spent budget

✓ Fixed total computation budget

✓ Bias correction makes a significant difference!

✓ Adaptive bias correction improves the chance of finding x^*



References:

Gao, S., Xiao, H., Zhou, E., and Chen, W. (2017). Robust ranking and selection with optimal computing budget allocation. *Automatica*, 81:30 – 36.
Chang, J. and Hall, P. (2015). Double-bootstrap methods that use a single double-bootstrap simulation. *Biometrika*, 102(1):203–214.
Vahdat, K. and Shashaani, S. (2021). Non-parametric uncertainty bias and variance estimation via nested bootstrapping and influence functions. In *Proceedings of the 2021 WSC*, pages 1–12
Vahdat, K. and Shashaani, S. (2022). Robust prediction error estimation with monte-carlo methodology. *arXiv preprint arXiv:2207.13612*.