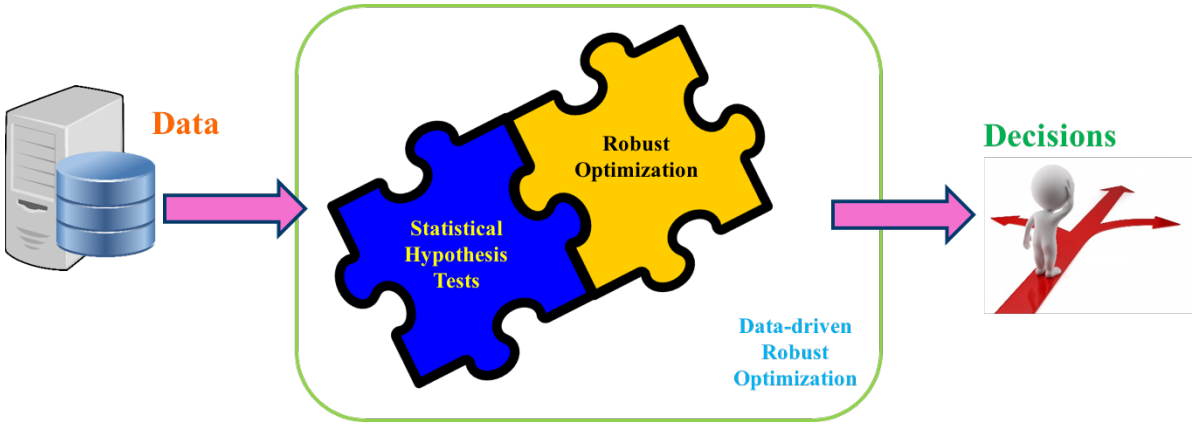
Big-Data-Driven Robust Optimization of Process Operations

Dr. Fengqi You, Roxanne E. and Michael J. Zak Professor
Process-Energy-Environmental Systems Engineering
fengq.you@cornell.edu | you.cbe.cornell.edu

Robust optimization is becoming a popular paradigm for optimization under uncertainty. In robust optimization, uncertainty is modeled using an uncertainty set. The choice of uncertainty sets is important in robust optimization. With a well-chosen set, the resulting robust optimization is computationally tractable, and yields less conservative robust solutions. In the era of big data, a large amount of data is available for the decision makers. A general data-driven framework for designing uncertainty sets from big data is needed. Data-driven uncertainty sets should be able to truthfully capture a variety of features of data, such as heavy-tails and correlations.



In this project, we will use statistical hypothesis tests, such as Kolmogorov-Smirnov test, to construct data-driven uncertainty sets. Bootstrapping and Gaussian approximation are further used to improve the performance. The robust counterpart with these data-driven uncertainty sets will be derived. The condition under which robust counterparts are computationally tractable will be investigated. Numerical examples will be used to illustrate the superiority of the data-driven method.

The statistical hypothesis test based method assumes that the data for uncertainty are drawn i.i.d. from a distribution. Therefore, it is not suitable for the cases with non-stationary data that exhibit significant autocorrelations. We will also study how these statistical hypothesis tests coupled with time series models can be used to construct uncertainty sets for time series data. Some real-world data, such as electricity price data, are used to demonstrate the effectiveness of the developed methods.