LEVERAGING THE POWER OF BIG DATA FOR ROBUST PROCESS OPERATIONS UNDER UNCERTAINTY

Chao Ning, Fengqi You*
Cornell University
Ithaca, New York 14853

Abstract

We propose a data-driven outlier-insensitive adaptive robust optimization framework that leverages big data in industries. A Bayesian nonparametric model – the Dirichlet process mixture model – is adopted to extract the information embedded within uncertainty data via a variational inference algorithm. We then devise data-driven uncertainty sets for adaptive robust optimization. This Bayesian nonparametric model is seamlessly integrated with adaptive optimization approach through a novel four-level robust optimization framework. This framework explicitly considers the correlation, asymmetry and multimode of uncertainty data, and as a result generates less conservative solutions. Additionally, this framework is robust not only to parameter variations, but also to data outliers. An efficient tailored column-and-constraint generation algorithm is proposed for the resulting problem that cannot be solved directly by any off-the-shelf optimization solvers. The effectiveness and advantages of the modeling framework and solution algorithm are demonstrated through industrial applications.

Keywords

Big data, data-driven adaptive robust optimization, Dirichlet process mixture model, column-and-constraint generation algorithm, process design and operations.

Introduction

In the past few decades, optimization of process systems under uncertainty has attracted wide attention from both academia and industry (Grossmann et al., 2016). Robust optimization (RO) emerges as a popular approach due to its strong ability to hedge against uncertainties and its computational tractability (Bel-Tal et al., 2009). It has achieved success in a broad array of applications (Bertsimas et al., 2011). Traditional RO approaches, also known as static robust optimization, make all the decisions at once. This modeling framework does not fit well with sequential decision-making problems. To this end, adaptive or adjustable robust optimization (ARO) was proposed to offer a new paradigm for optimization under uncertainty by incorporating recourse decisions (Ben-Tal et al., 2004). Due to the flexibility of adjusting some decisions to counteract uncertainties, ARO typically generates less conservative solutions than static robust optimization (Gong et al., 2016; Lappas and Gounaris, 2016). Big data is reshaping both operations research and process systems engineering (Bertsimas et al., 2013; Ning et al., 2014; Qin, 2014). More recently, dramatic progress of mathematical programming methods, coupled with recent advances in machine learning algorithms, sparks a flurry of interest in the research field of data-driven optimization (Bertsimas et al., 2013; Calfa et al., 2015).

Traditional ARO approach typically fails to take full advantage of data, and makes *a priori* and simple assumptions about uncertainty, such as independence and symmetry. These assumptions may not be reasonable for real world applications. In addition, process data are often contaminated with outliers. These outliers could undesirably affect the estimated region of uncertain parameters, and therefore would have an impact on uncertainty set.

In this paper, we propose a data-driven adaptive nested robust optimization (DDANRO) modeling framework and its solution strategy. The Dirichlet process mixture model is employed to model the uncertainty data using a variational inference algorithm (Campbell and How, 2015). We propose two novel data-driven uncertainty sets for ARO in this paper. The first one is defined as a polytope using l_1 norm. The second one is constructed as a data-driven budget based uncertainty set using l_1 and l_∞ norms. Multiple basic uncertainty sets, instead of one, are used to give a high-fidelity description of uncertainty data. The machine learning model is integrated with ARO approach seamlessly in a four-level optimization framework (a min-max-max-min problem). To solve the problem, we further propose a tailored column-and-constraint generation (C&CG) algorithm (Zeng and Zhao, 2013).

The major novelties of this article are summarized as follows:

- A novel data-driven adaptive nested robust optimization (DDANRO) framework is proposed for optimal process design and operations under uncertainty.
- A data-driven approach for defining uncertainty set is proposed for ARO.
- Since the resulting min-max-max-min problem cannot be solved directly by any off-the-shelf optimization solvers, a tailored column-and-constraint generation algorithm is developed to efficiently solve the four-level problem.

The remainder of this article is organized as follows. In the next section, we present the DDANRO modeling framework, followed by the solution strategy. We then present industrial applications, followed by the conclusion.

General Modeling Framework

A novel modeling framework DDANRO is proposed in this section.

Two-Stage Adaptive Robust Optimization

The general two-stage ARO in its compact form is given as follows (Zeng and Zhao, 2013).

$$\begin{aligned}
&\min_{\mathbf{x}} \ \mathbf{c}^T\mathbf{x} + \max_{\mathbf{u}\in\mathscr{U}} \ \min_{\mathbf{y}\in\Omega(\mathbf{x},\mathbf{u})} \ \mathbf{b}^T\mathbf{y} \\
&\text{s.t.} \quad \mathbf{A}\mathbf{x} \geq \mathbf{d}, \ \ \mathbf{x}\in\mathbf{S}_\mathbf{x} \\
&\qquad \Omega(\mathbf{x},\mathbf{u}) = \left\{\mathbf{y}\in\mathbf{S}_\mathbf{y} : \mathbf{W}\mathbf{y} \geq \mathbf{h} - \mathbf{T}\mathbf{x} - \mathbf{M}\mathbf{u}\right\}
\end{aligned} \tag{1}$$

where $\mathbf{x}$ is the vector of first-stage decision variables, which need to be made before the uncertainty $\mathbf{u}$ is known, and $\mathbf{y}$ represents the vector of second-stage decisions. Traditional ARO in this paper is referred to as the ARO with the widely adopted the budget based uncertainty set $\mathscr{U} = \left\{\mathbf{u} \,\middle|\, \mathbf{u}_i = \overline{\mathbf{u}_i} + \Delta\mathbf{u}_i \cdot z_i, -1 \leq z_i \leq 1, \ \sum_i z_i \leq \Gamma, \ \forall i\right\}$ that assumes symmetry and independence of uncertainty parameters.

Data-Driven Uncertainty Set Construction for ARO

A random draw from a Dirichlet process, i.e. DP (α, F_0), is a distribution F. The Dirichlet process mixture model adds one more level to this hierarchy. It utilizes θ_k as the parameters of the distribution of data. The Dirichlet process mixture model is summarized as follows (Campbell and How, 2015).

$$\begin{aligned}
&\beta_k \,|\, \alpha \sim Beta(1,\alpha) \\
&\theta_k \,|\, F_0 \sim F_0 \\
&l_i \,|\, \{\beta_1, \beta_2, \ldots\} \sim Mult\left(\pi(\beta)\right) \\
&o_i \,|\, l_i \sim p(o_i \,|\, \theta_{l_i})
\end{aligned} \tag{2}$$

where *Mult* denotes a multinomial distribution, and l_i is the index of mixture components to which the observation o_i is assigned. We propose a data-driven polyhedral uncertainty set for ARO below.

$$\mathscr{U} = \bigcup_{i:\,\gamma_i \geq \gamma^*} \left\{\mathbf{u} \,\middle|\, \mathbf{u} = \boldsymbol{\mu}_i + s_i \boldsymbol{\Psi}_i^{1/2}\boldsymbol{\xi}, \ \|\boldsymbol{\xi}\|_1 \leq \Gamma_i\right\} \tag{3}$$

where γ_i is the weight of the i th component, which can be calculated using $\gamma_i = \frac{\tau_i}{\tau_i + \nu_i}\prod_{j=1}^{i-1}\frac{\nu_j}{\tau_j + \nu_j}, \ \ i = 1, \ldots, M-1$ and $\gamma_M = 1 - \sum_{i=1}^{M-1}\gamma_i$ (Campbell and How, 2015). The weight γ_i indicates the probability of the corresponding mixture component. γ^* is a threshold value. $\tau_i, \nu_i, \boldsymbol{\mu}_i, \lambda_i, \omega_i, \boldsymbol{\Psi}_i$ are the inference results of the i th component using the variational inference algorithm. Γ_i satisfies the probability guarantee $\Pr\left(\|\boldsymbol{\xi}\|_2 \leq \Gamma_i\right) = \varepsilon$, $\boldsymbol{\xi} \sim St_{\omega_i + 1 - \dim(\mathbf{u})}\left(\mathbf{0}, \mathbf{I}\right)$, where *St* represents Student's t-distribution (Campbell and How, 2015). We also propose the following data-driven uncertainty set.

$$\mathscr{U} = \bigcup_{i:\,\gamma_i \geq \gamma^*} \left\{\mathbf{u} \,\middle|\, \mathbf{u} = \boldsymbol{\mu}_i + s_i \boldsymbol{\Psi}_i^{1/2}\Lambda_i\,\mathbf{z}, \ \ \|\mathbf{z}\|_\infty \leq 1, \ \ \|\mathbf{z}\|_1 \leq \Phi_i\right\} \tag{4}$$

where Λ_i is a scaling factor. This type of uncertainty set can be thought of as a data-driven version of the budget based uncertainty set.

Data-Driven Adaptive Nested Robust Optimization Modeling Framework

Data-driven adaptive nested robust optimization (DDANRO) model is proposed by integrating the optimization model with the machine learning model seamlessly. The DDANRO model using the l_1 norm based data-driven uncertainty set, denoted as (DDANRO-1), is shown as follows.

$$\text{(DDANRO-1)} \quad \begin{aligned} &\min_{\mathbf{x}} \ \mathbf{c}^T\mathbf{x} + \max_{i\in\{1,\dots,m\}} \max_{\mathbf{u}\in\mathscr{U}_i} \ \min_{\mathbf{y}\in\Omega(\mathbf{x},\mathbf{u})} \mathbf{b}^T\mathbf{y} \\ &\text{s.t.} \ \ \mathbf{Ax} \ge \mathbf{d}, \ \ \mathbf{x}\in \mathbf{S_x} \\ &\quad \mathscr{U}_i = \left\{\mathbf{u} \middle| \, \mathbf{u} = \boldsymbol{\mu}_i + s_i \boldsymbol{\Psi}_i^{1/2}\boldsymbol{\xi}, \left\|\boldsymbol{\xi}\right\|_1 \le \Gamma_i \right\} \\ &\quad \Omega(\mathbf{x},\mathbf{u}) = \left\{\mathbf{y}\in \mathbf{S_y} : \mathbf{Wy} \ge \mathbf{h} - \mathbf{Tx} - \mathbf{Mu}\right\} \end{aligned}$$

The DDANRO model using the l_1 norm and l_∞ norms, denoted as (DDANRO-inf) is as follows.

$$\text{(DDANRO-inf)} \quad \begin{aligned} &\min_{\mathbf{x}} \ \mathbf{c}^T\mathbf{x} + \max_{i\in\{1,\dots,m\}} \max_{\mathbf{u}\in\mathscr{U}_i} \ \min_{\mathbf{y}\in\Omega(\mathbf{x},\mathbf{u})} \mathbf{b}^T\mathbf{y} \\ &\text{s.t.} \ \ \mathbf{Ax} \ge \mathbf{d}, \ \ \mathbf{x}\in \mathbf{S_x} \\ &\quad \mathscr{U}_i = \left\{ \begin{aligned} &\mathbf{u} \middle| \, \mathbf{u} = \boldsymbol{\mu}_i + s_i \boldsymbol{\Psi}_i^{1/2}\Lambda_i \, \mathbf{z}, \\ &\left\|\mathbf{z}\right\|_\infty \le 1, \ \left\|\mathbf{z}\right\|_1 \le \Phi_i \end{aligned} \right\} \\ &\quad \Omega(\mathbf{x},\mathbf{u}) = \left\{\mathbf{y}\in \mathbf{S_y} : \mathbf{Wy} \ge \mathbf{h} - \mathbf{Tx} - \mathbf{Mu}\right\} \end{aligned}$$

Compared with (DDANRO-inf), (DDANRO-1) is preferable in terms of computational expenses. (DDANRO-inf) trades the computational time for the solution quality when handling high-dimensional uncertainties.

Solution Strategy

In this section, we address the computational challenge by proposing a decomposition-based solution algorithm for the DDANRO problem. This algorithm is designed in the line of the column-and-constraint generation as a master-subproblem framework (Zeng and Zhao, 2013).

First, we address the subproblems of (DDANRO-1). By using $\boldsymbol{\xi} = \Gamma_i \cdot \boldsymbol{\xi}_0$, we reformulate the set (3) into an equivalent form $\mathscr{U}_i = \left\{\mathbf{u} \middle| \, \mathbf{u} = \boldsymbol{\mu}_i + s_i \boldsymbol{\Psi}_i^{1/2} \, \Gamma_i \boldsymbol{\xi}_0, \left\|\boldsymbol{\xi}_0\right\|_1 \le 1\right\}$. Suppose a polytope $\mathscr{D} = \left\{\boldsymbol{\xi}_0 \in R^K \middle| \ \left\|\boldsymbol{\xi}_0\right\|_1 \le 1\right\}$ and its extreme points are $\mathbf{d}^{(\pm j)} = [0,\dots,0, \underbrace{\pm 1}_{j\text{ th position}}, 0,\dots,0]^T$.

The random variable $\boldsymbol{\xi}_0$ can be expressed as a convex combination of all extreme points of $\mathscr{D}$ as follows.

$$\boldsymbol{\xi}_0 = \sum_{j=1}^{K} \nu^{(-j)}\mathbf{d}^{(-j)} + \sum_{j=1}^{K} \nu^{(j)}\mathbf{d}^{(j)} \tag{5}$$

where $\nu^{(\pm j)} \in [0,1]$ are variables that satisfy $\sum_{j=1}^{K} \nu^{(\pm j)} = 1$, $\mathbf{d}^{(\pm j)} \in R^K$ (j=1, …, K) are the extreme points of polytope $\mathscr{D}$. Then the subproblem can be simplified by using $\mathbf{f}^{(i,\pm j)} = s_i \Gamma_i \mathbf{M}\boldsymbol{\Psi}_i^{1/2}\mathbf{d}^{(\pm j)}$. Note that a vector with a subscript t represents its corresponding element. For the simplified problem, the optimal value $\nu^{(\pm j)}$ has been proven to be a binary variable (Billionnet et al., 2014).

We can use Glover's linearization to linearize the bilinear term $\nu^{(\pm j)}\boldsymbol{\varphi}_t$ by the substitution $g_t^{(\pm j)} = \nu^{(\pm j)}\boldsymbol{\varphi}_t$. Given $\sum_j \sum_t \mathbf{f}_t^{(i,\pm j)} g_t^{(\pm j)} = \left(s_i \Gamma_i \mathbf{M}\boldsymbol{\Psi}_i^{1/2}\mathbf{d}^{(\pm j)}\right)^T \mathbf{g}^{(\pm j)}$, the i th subproblem of (DDANRO-1) can be reformulated as the following MILP problem (SP_i-1).

$$(\text{SP}_\text{i}\text{-1}) \quad \begin{aligned} Q_i(\mathbf{x}) = &\max_{\nu^{(\pm j)},\boldsymbol{\varphi},\mathbf{g}^{(\pm j)}} \left(\mathbf{h} - \mathbf{Tx} - \mathbf{M}\boldsymbol{\mu}_i\right)^T \boldsymbol{\varphi} \\ &- \sum_j \left(s_i \Gamma_i \mathbf{M}\boldsymbol{\Psi}_i^{1/2}\mathbf{d}^{(\pm j)}\right)^T \mathbf{g}^{(\pm j)} \\ &\text{s.t.} \ \ \mathbf{W}^T\boldsymbol{\varphi} \le \mathbf{b} \\ &\quad \boldsymbol{\varphi} \ge \mathbf{0} \\ &\quad \mathbf{0} \le \mathbf{g}^{(\pm j)} \le \boldsymbol{\varphi} \\ &\quad \mathbf{g}^{(\pm j)} \le M_0 \cdot \nu^{(\pm j)} \cdot \mathbf{e} \\ &\quad \mathbf{g}^{(\pm j)} \ge \boldsymbol{\varphi} - M_0 \cdot \left(1 - \nu^{(\pm j)}\right) \cdot \mathbf{e} \\ &\quad \textstyle\sum_{j=1}^{K} \nu^{(\pm j)} = 1, \ \ \nu^{(\pm j)} \in \{0,1\} \end{aligned}$$

where M_0 is a sufficiently large constant, and $\mathbf{e}$ represents a column vector whose elements are all ones. Note that the vector inequalities should be interpreted elementwise.

Now, we address the solution of the subproblem of (DDANRO-inf). For ease of derivation, $\mathbf{z}_j$ can be divided into two parts $\mathbf{z}_j = \mathbf{z}_j^+ - \mathbf{z}_j^-$ (Thiele et al., 2009). By using Glover's linearization for $\mathbf{G}_{tj}^+ = \mathbf{z}_j^+\boldsymbol{\varphi}_t$ and $\mathbf{G}_{tj}^- = \mathbf{z}_j^-\boldsymbol{\varphi}_t$, we can reformulate the subproblem of (DDANRO-inf). The reformulation (SP_i-inf) is shown as follows.

$$(\text{SP}_\text{i}\text{-inf}) \quad \begin{aligned} Q_i(\mathbf{x}) = &\max_{\boldsymbol{\varphi},\mathbf{z}^+,\mathbf{z}^-,\mathbf{G}^+,\mathbf{G}^-} \left(\mathbf{h} - \mathbf{Tx} - \mathbf{M}\boldsymbol{\mu}_i\right)^T \boldsymbol{\varphi} \\ &- Tr\left(\left(s_i \Lambda_i \mathbf{M}\boldsymbol{\Psi}_i^{1/2}\right)^T \left(\mathbf{G}^+ - \mathbf{G}^-\right)\right) \\ &\text{s.t.} \ \mathbf{W}^T\boldsymbol{\varphi} \le \mathbf{b} \\ &\quad \boldsymbol{\varphi} \ge \mathbf{0} \\ &\quad \mathbf{0} \le \mathbf{G}^+ \le \boldsymbol{\varphi} \cdot \mathbf{e}^T \\ &\quad \mathbf{0} \le \mathbf{G}^- \le \boldsymbol{\varphi} \cdot \mathbf{e}^T \\ &\quad \mathbf{G}^+ \le \tilde{\mathbf{e}} \cdot \left(M_0 \cdot \mathbf{z}^+\right)^T \\ &\quad \mathbf{G}^- \le \tilde{\mathbf{e}} \cdot \left(M_0 \cdot \mathbf{z}^-\right)^T \\ &\quad \mathbf{G}^+ \ge \boldsymbol{\varphi} \cdot \mathbf{e}^T - \tilde{\mathbf{e}} \cdot \left(M_0 \cdot \left(\mathbf{e} - \mathbf{z}^+\right)\right)^T \\ &\quad \mathbf{G}^- \ge \boldsymbol{\varphi} \cdot \mathbf{e}^T - \tilde{\mathbf{e}} \cdot \left(M_0 \cdot \left(\mathbf{e} - \mathbf{z}^-\right)\right)^T \\ &\quad \mathbf{e}^T\left(\mathbf{z}^+ + \mathbf{z}^-\right) \le \Phi_i \\ &\quad \mathbf{z}^+ + \mathbf{z}^- \le \mathbf{e}, \quad \mathbf{z}^+, \mathbf{z}^- \in \{0,1\}^K \end{aligned}$$

where $\mathbf{e}$ and $\tilde{\mathbf{e}}$ both represent column vectors whose elements are all ones, but they are of different dimension, i.e., $\dim(\mathbf{e}) = \dim(\mathbf{u}) = K$ and $\dim(\tilde{\mathbf{e}}) = \dim(\boldsymbol{\varphi})$. The notation Tr denotes the trace of a matrix, for example, the matrix $\mathbf{Q} = \left[q_{ij} \right] \in R^{L \times L}$, $Tr\left(\mathbf{Q}\right) = \sum_{i=1}^{L} q_{ii}$. Note that the vector or matrix inequalities should be interpreted elementwise. The master problem (MP) is shown below.

$$\textbf{(MP)} \quad \begin{aligned} \min_{\mathbf{x}} \ & \mathbf{c}^T \mathbf{x} + \eta \\ s.t. \ & \mathbf{A}\mathbf{x} \geq \mathbf{d} \\ & \eta \geq \mathbf{b}^T \mathbf{y}^l, \ l = 1, \ldots, k \\ & \mathbf{T}\mathbf{x} + \mathbf{W}\mathbf{y}^l \geq \mathbf{h} - \mathbf{M}\mathbf{u}^l, \ l = 1, \ldots, k \\ & \mathbf{x} \in \mathbf{S}_{\mathbf{x}}, \ \mathbf{y}^l \in \mathbf{S}_{\mathbf{y}}, \ l = 1, \ldots, k \end{aligned}$$

The proposed algorithm iteratively solves a sequence of master problems and subproblems until the optimality gap reduces to a predefined tolerance. The procedure of the tailored C&CG algorithm for DDANRO is given as follows. Note that (SP_i) represents (SP_i-1) or (SP_i-inf) depends on the choice of uncertainty sets.

Algorithm. Tailored C&CG algorithm for DDANRO

Set $LB \leftarrow -\infty$, $UB \leftarrow +\infty$, $k \leftarrow 0$, and $\zeta \leftarrow 10^{-3}$;
while $\left| \frac{UB - LB}{UB} \right| \geq \zeta$ **do**
  Solve (MP) to obtain $\mathbf{x}_{k+1}^*, \eta_{k+1}^*, \mathbf{y}^{1*}, \ldots, \mathbf{y}^{k*}$;
  Update $LB \leftarrow \mathbf{c}^T \mathbf{x}_{k+1}^* + \eta_{k+1}^*$;
  for $i = 1$ to m **do**
    Solve (SP_i) to obtain $\mathbf{u}_i^{k+1}$ and $Q_i\left(\mathbf{x}_{k+1}^*\right)$;
  end
  $i^* \leftarrow \arg\max_i Q_i\left(\mathbf{x}_{k+1}^*\right)$;
  $\mathbf{u}^{k+1} \leftarrow \mathbf{u}_{i^*}^{k+1}$ and $Q\left(\mathbf{x}_{k+1}^*\right) \leftarrow Q_{i^*}\left(\mathbf{x}_{k+1}^*\right)$;
  Update $UB \leftarrow \min\left\{UB, \mathbf{c}^T \mathbf{x}_{k+1}^* + Q\left(\mathbf{x}_{k+1}^*\right)\right\}$;
  Create second-stage variables $\mathbf{y}^{k+1}$ and add cuts $\eta \geq \mathbf{b}^T \mathbf{y}^{k+1}, \mathbf{T}\mathbf{x} + \mathbf{W}\mathbf{y}^{k+1} \geq \mathbf{h} - \mathbf{M}\mathbf{u}^{k+1}$ to (MP);
  $k \leftarrow k + 1$;
end
return UB;

Figure 1. The tailored column-and-constraint generation algorithm.

Industrial Applications

Application 1. Data-Driven Robust Scheduling of Batch Processes under Processing Time and Demand Uncertainties

Production scheduling plays an increasingly important role in process systems engineering for ensuring an efficient and competitive process manufacturing. In this case study, we employ the proposed DDANRO framework to an industrial multipurpose batch process in The Dow Chemical Company. Figure 2 displays the state-task network (STN) of this multipurpose batch process. The mathematical programming model for scheduling is to maximize profits while satisfying a set of constraints. Following the two-stage ARO model for batch scheduling, the processing times of all reaction tasks and demands for all products are subject to uncertainty (Shi and You, 2016). Unlike previous studies, the real batch processing time data are used to construct the uncertainty set. It is worth noting that the original data are corrupted with outliers. The main focus of this case study is to demonstrate how outliers could have a negative effect on traditional ARO solution quality.

In this case study, we set the optimality tolerance for CPLEX 12 to be 10^{-3}, and set the relative optimality tolerance for the proposed tailored C&CG algorithm to be 0.1%. The optimal number of time points turns out to be 9 for the static robust optimization. We set the time horizon to be 168 hours and the number of time points to be 9 for all methods. To make a fair comparison, we adjust the related parameters of model (DDANRO-1) to cover the same percentage of uncertainty data as the traditional ARO.

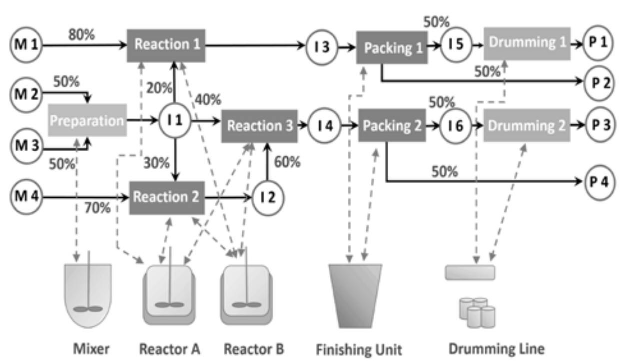

Figure 2. State task network of the multipurpose batch process.

From Figure 3, we can see that the solution (DDANRO-1) generates the highest profit, which is 36.8% higher than that of the other two methods.

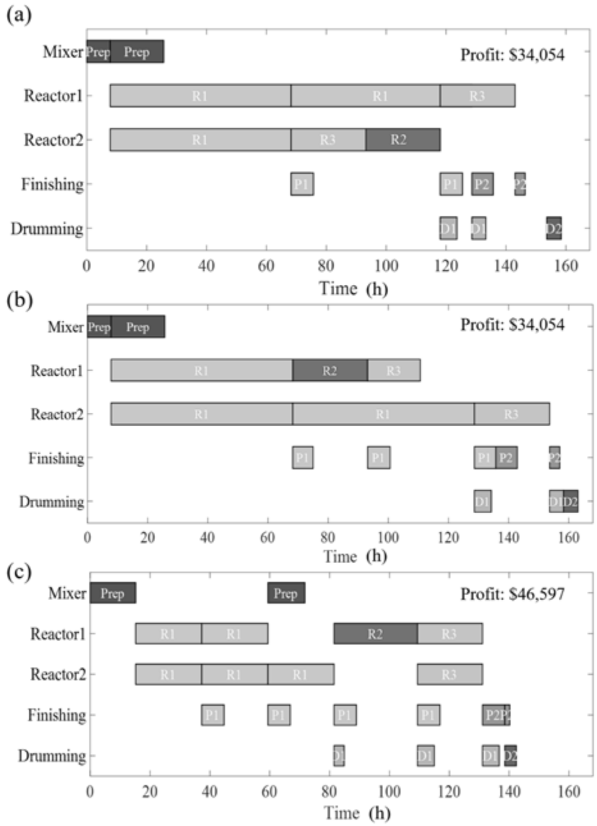

Figure 3. Gantt charts of (a) Static robust optimization, (b) Traditional ARO and (c) (DDANRO-1) when time points N=9.

The corresponding total profit is listed at the upper right corner of the Gantt charts. The DDANRO method inherits the merits of robust optimization to hedge against uncertainty and is at the same time robust to data outliers. It is interesting to note that the time horizon in Figure 3 (c) is not packed with tasks. This indicates that the time points N=9 may not be optimal for the DDANRO. As mentioned before, when we increase the number of time points to 10, the optimal objective function value of static robust optimization does not change. However, when the number of time points is 10, the traditional ARO scheduling problem cannot be solved within 20 hours.

To study the impact of time points on (DDANRO-1) solution, we increase N by one at each time for (DDANRO-1), and the corresponding Gantt charts are given in Figure 4. From Figure 4, we can see that more and more tasks can be performed within the time horizon as we increase the number of time points. Note that the problem (DDANRO-1) cannot be solved within 20 hours when the number of the time points is 13. Therefore, we only list the Gantt charts when the number of time points is 10, 11 and 12 in Figure 4.

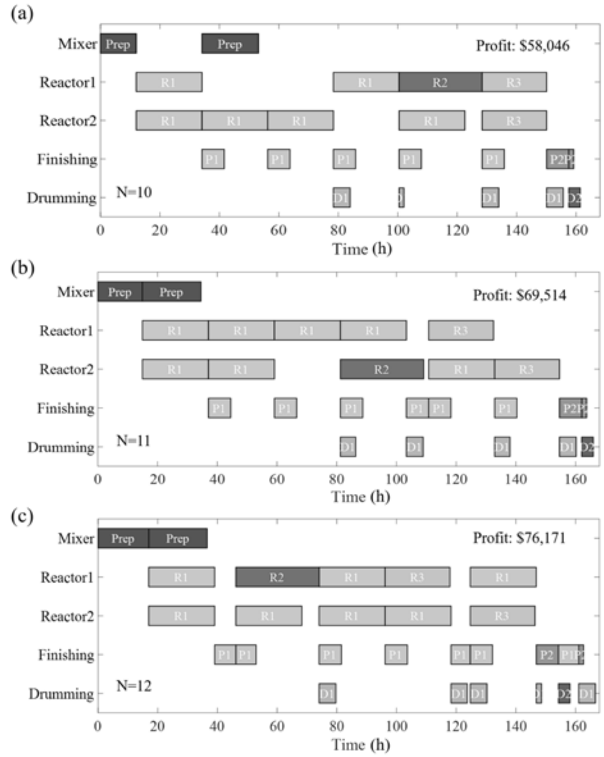

Figure 4. Gantt charts of (DDANRO-1) at different number of time points (a) N=10, (b) N=11 and (c) N=12.

Through the above discussions, the proposed DDANRO outperforms the static robust optimization and the traditional ARO when uncertainty data are corrupted with outliers.

Application 2. Data-Driven Robust Planning of Chemical Process Networks under Supply and Demand Uncertainties

In this case study, we apply the proposed modeling framework and solution algorithm to address the design and planning of a chemical process network under supply and demand uncertainty (Yue and You, 2013). This process network includes 38 processes and 28 chemicals (You and Grossmann, 2011). In this section, we apply the DDANRO approach to the problem of process network planning under supply and demand uncertainties. The decisions of the model include the selection of processes for expansion, the capacity of processes, operating levels of the installed processes, sales and purchases at each time period. Unlike the traditional ARO which assumes that uncertainties are independent, DDANRO considers correlation explicitly, thus boosting the NPV from \$761.79K to \$819.28K. Uncertainty budgets are often chosen by decision makers. In Figure 5, we display a heat map of the robust solution profile under different budgets.

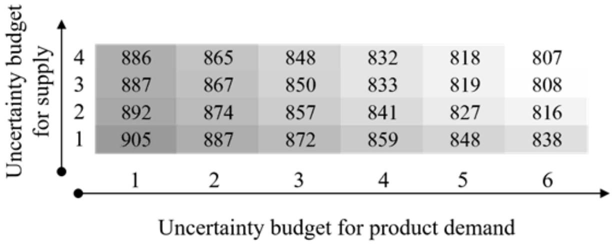

Figure 5. Heat map of solution profile for (DDANRO-inf) under different budgets.

The computational experiments are carried out on a Dell Optiplex 790 desktop with an Intel (R) Core (TM) i5-2400 3.10GHz CPU, 8GB RAM, and a Windows 7 64-bit operating system. We set the relative optimality tolerance for the proposed tailored C&CG algorithm to be 0.1%.

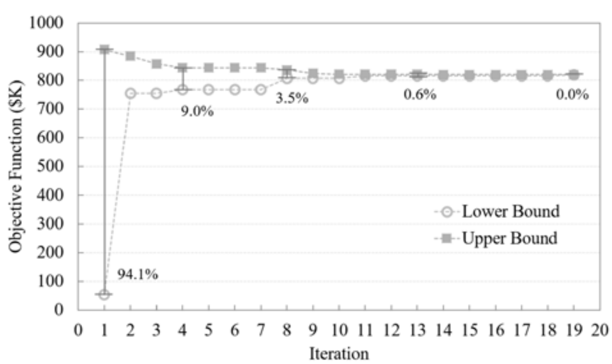

Figure 6. Lower and upper bounds of the tailored C&CG algorithm in each iteration.

To showcase the computational performance of the tailored C&CG algorithm, we display the convergence process of our proposed algorithm in Figure 6. We can see that the proposed algorithm is able to converge rapidly to a small optimality gap after the first few iterations.

Conclusion

This paper proposes a novel data-driven adaptive nested robust optimization (DDANRO) framework. It is a general framework for optimization under uncertainty. In this paper, we focus on its application in optimal design and operations. The proposed DDANRO framework inherits advantages of the machine learning model and adaptive robust optimization model. Therefore, it captures the useful information embedded within uncertainty data, and generates less conservative robust solutions.

References

Ben-Tal, A., El Ghaoui, L., Nemirovski, A. (2009). Robust Optimization. *Princeton University Press*.

Ben-Tal, A., Goryashko, A., Guslitzer, E., Nemirovski, A. (2004). Adjustable Robust Solutions of Uncertain Linear Programs. *Mathematical Programming*, 99, 351-376.

Bertsimas, D., Brown, D. B., Caramanis, C. (2011). Theory and Applications of Robust Optimization. *SIAM Review*, 53, 464-501.

Bertsimas, D., Gupta, V., Kallus, N. (2013). Data-Driven Robust Optimization. *arXiv preprint arXiv:1401.0212*.

Billionnet, A., Costa, M. C., Poirion, P. L. (2014). 2-Stage Robust MILP with Continuous Recourse Variables. *Discrete Applied Mathematics*, 170, 21-32.

Calfa, B. A., Grossmann, I. E., Agarwal, A., Bury, S. J., Wassick, J. M. (2015). Data-Driven Individual and Joint Chance-Constrained Optimization via Kernel Smoothing. *Computers & Chemical Engineering*, 78, 51-69.

Campbell, T., How, J. P. (2015). Bayesian Nonparametric Set Construction for Robust Optimization. *In Proceedings of the 2015 American Control Conference (ACC)*. Chicago, USA, 4216-4221.

Gong, J., Garcia, D. J., You, F. (2016). Unraveling Optimal Biomass Processing Routes from Bioconversion Product and Process Networks under Uncertainty: An Adaptive Robust Optimization Approach. *ACS Sustainable Chemistry & Engineering*, 4, 3160-3173.

Grossmann, I. E., Apap, R. M., Calfa, B. A., García-Herreros, P., Zhang, Q. (2016). Recent Advances in Mathematical Programming Techniques for The Optimization of Process Systems under Uncertainty. *Computers & Chemical Engineering*, 91, 3-14.

Lappas, N. H., Gounaris, C. E. (2016). Multi-Stage Adjustable Robust Optimization for Process Scheduling under Uncertainty. *AIChE Journal*, 62, 1646-1667.

Ning, C., Chen, M., Zhou, D. (2014). Hidden Markov Model-Based Statistics Pattern Analysis for Multimode Process Monitoring: An Index-Switching Scheme. *Industrial & Engineering Chemistry Research*, 53, 11084-11095.

Qin, S. J. (2014). Process Data Analytics in The Era of Big Data. *AIChE Journal*, 60, 3092-3100.

Shi, H., You, F. (2016). A Computational Framework And Solution Algorithms for Two-Stage Adaptive Robust Scheduling of Batch Manufacturing Processes under Uncertainty. *AIChE Journal*, 62, 687-703.

Thiele, A., Terry, T., Epelman, M. (2009). Robust Linear Optimization with Recourse. *Rapport Technique*, 4-37.

You, F., Grossmann, I. E. (2011). Stochastic Inventory Management for Tactical Process Planning under Uncertainties: MINLP Models and Algorithms. *AIChE Journal*, 57, 1250-1277.

Yue, D., You, F. (2013). Planning and Scheduling of Flexible Process Networks under Uncertainty with Stochastic Inventory: MINLP Models and Algorithm. *AIChE Journal*, 59, 1511-1532.

Zeng, B., Zhao, L. (2013). Solving Two-Stage Robust Optimization Problems Using A Column-and-Constraint Generation Method. *Operations Research Letters*, 41, 457-461.