Mathematical Modeling and Global Optimization of Large-Scale Extended Pooling Problems with the (EPA) Complex Emissions Constraints

Ruth Misener, Chrysanthos E. Gounaris, and Christodoulos A. Floudas*
Department of Chemical Engineering
Princeton University
Princeton, NJ 08544-5263

February 8, 2010

Abstract

Environmental Protection Agency (EPA) *Title 40 Code of Federal Regulations Part 80.45: Complex Emissions Model* [40CFR80.45, 2007] codifies a mathematical model of gasoline emissions for reformulated gasoline (RFG) as a function of eleven fuel properties. In this paper we propose an *extended pooling problem* to maximize the profit of blending reformulated gasoline on a predetermined network structure of feed stocks, intermediate storage tanks, and gasoline products subject to applicable environmental standards. A mixed-integer nonlinear programming (MINLP) model is introduced which is nonconvex due to the presence of bilinear, polynomial, and fractional power terms. A mixed-integer linear programming (MILP) relaxation of the extended pooling problem is proposed and we introduce several test cases from small to medium to large scale and solve them to global optimality. The large-scale test case involves fourteen feed stocks, five pools, and ten products and consists of 1104 continuous variables, 150 binary variables, and 640 nonlinear terms. The nonconvexities in the large-scale case study include 410 bilinear terms, 40 polynomial terms, and 10 terms raised to a fractional power.

Keywords: large-scale optimization; global optimization; MINLP; bilinear; polynomial; pooling problem

1 Introduction

Environmental Protection Agency (EPA) *Title 40 Code of Federal Regulations Part 80.45: Complex Emissions Model* [40CFR80.45, 2007] codifies a mathematical model of gasoline emissions for reformulated gasoline (RFG). The RFG program, which impacts roughly 75 million people, was developed to reduce smog and airborne toxic pollutants (*e.g.*, benzene, a human carcinogen) in accordance with the Clean Air Act [EPA, 2008].

The Complex Emissions Model calculates volatile organic, nitrous oxide (NO_X), and airborne toxic emissions using functions of the eleven fuel qualities listed in Table 1. The three emissions models form the basis for other legislation, such as *Title 40 Code of Federal Regulations Part 80.41: Standards and requirements for compliance* [40CFR80.41, 2008], to set emissions standards. Final products exiting an oil refinery must comply with these standards, or upper bounds, on volatile organic ($\mathrm{VOC_{MAX}}$), NO_X ($\mathrm{NOX_{MAX}}$) and airborne toxic ($\mathrm{TOX_{MAX}}$) emissions.

*To whom all correspondence should be addressed (floudas@titan.princeton.edu; Tel: (609) 258-4595; Fax: (609) 258-0211).

Table 1: EPA Complex Emissions Model fuel components bounded by the limits of RFG model accuracy.

	Var	Fuel Quality	Bounds	Units
1	OXY	oxygen	0.0-4.0	wt%
2	SUL	sulfur	0.0-500.0	ppm
3	RVP	Reid Vapor Pressure	6.4-10.0	psi
4	E200	200^oF distillation fraction	30.0-70.0	vol%
5	E300	300^oF distillation fraction	70.0-100.0	vol%
6	ARO	aromatics	0.0-50.0	vol%
7	BEN	benzene	0.0-2.0	vol%
8	OLE	olefins	0.0-25.0	vol%
9	MTB	methyl tertiary butyl ether (MTBE)	0.0-4.0	wt% O_2
10	ETB	ethyl tertiary butyl ether (ETBE)	0.0-4.0	wt% O_2
11	ETH	ethanol	0.0-4.0	wt% O_2

The *pooling problem* maximizes profit (equivalently, minimizes cost) by optimally selecting flow rates on a predetermined network structure of feed stocks, pooling tanks, and final products. A common refining application of the pooling problem addresses the temporary storage of intermediate feed stocks exiting processing units [Visweswaran, 2009]. The temporary storage tanks or *pools*, which are subsequently mixed into final products, are monitored to ensure that the concentration of regulated qualities does not exceed environmental limits in the final products [Misener and Floudas, 2009].

The *extended pooling problem* appends the EPA Complex Emissions Model and associated constraints to a standard pooling problem. The goal is to comply with reformulated gasoline standards while maximizing profitability. Integrating recently-developed relaxation techniques into a global optimization algorithm [Meyer and Floudas, 2005, Wicaksono and Karimi, 2008, Gounaris et al., 2009], we globally optimize three case studies including a large-scale example with fourteen feed stocks, five pools, and ten products which consists of 1104 continuous variables, 150 binary variables, and 640 nonlinear terms. The nonconvexities in the large-scale test case include 410 bilinear, 40 polynomial, and 10 fractional power terms.

After reviewing the literature in Section 2 and presenting the nomenclature in Table 2, we formulate a mixed-integer nonlinear model (MINLP) in Section 3 to represent the extended pooling problem. Section 4 discusses a linear relaxation of the MINLP and Section 5 presents an algorithm to converge on the global solution. We demonstrate the algorithm by posing a number of test cases in Section 6 and solving them to optimality in Section 7. The values of the EPA-defined parameters are listed in Appendix A. Appendices B – D contain the case study parameters.

Although Furman and Androulakis [2008] have previously formulated an MINLP of the EPA Complex Emissions Model, this paper presents an alternative representation. Our alternative formulation of the EPA Complex Emissions Model is valuable because:

1 This formulation integrates the EPA Complex Emissions Model into a pooling problem framework by appending the emissions model to each exiting RFG product, while Furman and Androulakis [2008] analyze the emissions for a single fuel. Therefore, the variables natural to our model differ from the Furman and Androulakis [2008] formulation. Our model has 87 continuous variables per RFG product in comparison to 78 continuous variables for Furman and Androulakis [2008]. Twenty continuous variables arise in our model from equivalently representing eight bilinear terms in the Furman and Androulakis [2008] formulation as piecewise linear (see Item 3). Excluding those twenty, we have reduced the number of variables by using items (*i.e.*, $u_{j,k}$) that are natural to the pooling problem.

2 Our model has 15 binary variables per RFG product in comparison to the 21 binary variables in the Furman and Androulakis [2008] model. Reduction in the number of binary variables comes from

eliminating redundancy. For example, Furman and Androulakis [2008] define three binary variables (Y_n^{SUL}, $Y_n^{\mathrm{SUL}'}$, and $Y_n^{\mathrm{SUL}''}$) to represent the two EPA-defined disjunctions at 10 and 450 ppm sulfur. We use two binary variables ($y^N_{\mathrm{SUL},\,10,\,j}$ and $y^N_{\mathrm{SUL},\,450,\,j}$) to represent the two disjunctions rather than using three variables to represent the three segments created by the two disjunctions.

3 Eight piecewise-linear functions in the EPA Complex Emissions Model are formulated by Furman and Androulakis [2008] as piecewise-nonlinear functions. The MINLP model of Furman and Androulakis [2008] uses bilinear terms to express the functions illustrated in Figures 1(a) & 2. Although the functions delineated in Figures 1(a) & 2 can be modeled as the product of two variables, at least one of the variables is constant in every region of the domain, so the functions can be modeled without approximation as piecewise linear. Our model uses piecewise linear equations with slack variables (see Item 1) to avoid using nonconvex bilinear terms to equivalently represent the expressions.

4 Furman and Androulakis [2008] explain that:

> due to discrepancies between [40CFR80.45] and the spreadsheet model ... on the EPA website, the spreadsheet has been chosen as a source.

Although there *are* discrepancies between the spreadsheet model the EPA provides (accessible at `http://www.epa.gov/otaq/rfg.htm`) and 40CFR80.45 [2007] (*e.g.*, the maximum oxygen concentration is 4.0 wt.% in 40CFR80.45 [2007] but 3.7 wt.% in the spreadsheet and the winter time Reid Vapor Pressure is fixed to 8.7 psi in 40CFR80.45 [2007] but variable in the spreadsheet), Furman and Androulakis [2008] introduce model elements departing from both 40CFR80.45 [2007] and the spreadsheet model that we downloaded on July 23, 2008. For example, Furman and Androulakis [2008] define additional disjunctions like model variable $\Delta\mathrm{E200}_v$ defined as $(\Delta\mathrm{E200}_v = \mathrm{E200}_t - 65.52)$ when $(\mathrm{E200}_t > 65.52)$ even though 40CFR80.45 [2007] for volatile organic emissions says:

> If the E200 level of the target fuel equals or exceeds 33 volume percent, then ΔE200 shall be set equal to zero.

The spreadsheet model initializes the $(\Delta\mathrm{E200}_v = \mathrm{E200}_t - 65.52)$ computation, but a subsequent cell sets $\Delta\mathrm{E200}_v$ equal to zero if $\Delta\mathrm{E200}_v$ is positive. Therefore, they have introduced disjunctions not present in either the law or the spreadsheet. The $\Delta\mathrm{E200}_v$ case is one of the model extension variables where we avoid using bilinear terms, so we never explicitly introduce a variable $\Delta\mathrm{E200}_v$, but an equivalent piecewise linear formulation is introduced in Eqs. (50) – (53). In cases where the spreadsheet and 40CFR80.45 [2007] deviate, our model follows the rules in 40CFR80.45 [2007].

2 Literature Review

This section discusses the pooling problem generally, introduces algorithms commonly used to approach the pooling problem, and specifically covers scholarship related to the EPA Complex Emissions Model [40CFR80.45, 2007]. Refer to our recent review [Misener and Floudas, 2009] for more comprehensive coverage of the pooling problem. To understand the importance of the pooling problem within the context of the chemical processing industry, see the review of Kallrath [2000] or the process synthesis applications addressed by Floudas and co-workers [Floudas and Paules, 1988, Kokossis and Floudas, 1991, Ciric and Floudas, 1989, Kokossis and Floudas, 1994, Lin and Floudas, 2001]. For a general discussion of global optimization, the reader is directed to the book of Floudas [2000], the conference proceedings edited by Floudas and Pardalos [1995], or recent reviews of Floudas et al. [2005] and Floudas and Gounaris [2009].

Ever since Haverly [1978] introduced the pooling problem and used a local search algorithm to optimize it, the pooling problem has enjoyed nearly continuous interest from optimization researchers. Significant

contributions include that of Lasdon et al. [1979], who employed successive linear programming; Floudas et al. [1989] and Floudas and Aggarwal [1990], who designed an algorithm based on Generalized Benders' Decomposition; Floudas and Visweswaran [Floudas and Visweswaran, 1990, 1993, Visweswaran and Floudas, 1990, 1993], who developed the first global optimization approach, GOP, based on duality theory; Foulds et al. [1992], who implemented the bilinear envelopes of McCormick [1976]; Ben-Tal et al. [1994], who introduced the q-formulation; Adhya et al. [1999], who explored Lagrangian approaches; Quesada and Grossmann [1995], who used the reformulation-linearization technique (RLT) of Sherali and Alameddine [1992][1]; Tawarmalani and Sahinidis [2002], who proposed the pq-formulation; Audet et al. [2004], who applied the branch-and-cut RLT-based QCQP algorithm of Audet et al. [2000]; Pham et al. [2009], who designed a fast-solving method to reach near-optimal solutions; and Almutairi and Elhedhli [2009], who suggested a new Lagrangian relaxation for the pooling problem and demonstrated that their relaxation is often tighter than previously-developed Lagrangian relaxations. Other research has focused on a combinatorially complex generalization of the pooling problem where the topology becomes a decision variable [Audet et al., 2004, Meyer and Floudas, 2006, Karuppiah and Grossmann, 2006].

Global optimization algorithms applicable to the pooling problem include the integration of termwise convex envelopes into a branch-and-bound method [Al-Khayyal and Falk, 1983], the duality-based global optimization algorithm, GOP, for biconvex programs [Floudas and Visweswaran, 1990, 1993, Visweswaran and Floudas, 1990, 1993], an RLT-based branch-and-cut method [Audet et al., 2000], the branch-and-bound procedure of Linderoth [2005] that generates tight relaxations by partitioning two dimensional regions into triangles and rectangles, an integration of RLT and semidefinite programming [Anstreicher, 2009], and a cutting plane algorithm that generates a multiterm relaxation [Bao et al., 2009]. In this paper, we will approach the extended pooling problem using a branch-and-bound method based on (a) piecewise relaxation of bilinear terms, (b) convex envelopes of polynomial terms generated using the edge-concave paradigm, and (c) outer approximation of convex terms. The *ab initio* domain partitioning of variables participating in bilinear terms, which tightens the problem relaxation, was first introduced by Meyer and Floudas [2006] and Karuppiah and Grossmann [2006] in the context of tightly underestimating large-scale wastewater treatment and water networks problems, respectively. Other groups who have successfully integrated the piecewise-linear relaxations into global optimization algorithms include Bergamini et al. [2008] and Saif et al. [2008]. In this paper, we will introduce the findings of Wicaksono and Karimi [2008] and some of our own work [Gounaris et al., 2009] to choose a suitable formulation for the underestimators.

Furman and Androulakis [2004, 2008] presented an MINLP representation of the EPA Complex Emissions Model. Earlier approaches used local, realtime solvers to model blending [Treiber et al., 1998] or integrated heuristics, such as the observation that sulfur has a significant impact on the NO_X model [Rhodes, 1998]. We have also previously discussed a piecewise-linear approximation of the EPA Complex Emissions Model in the context of an approximation algorithm [Misener and Floudas, 2010].

Gounaris and Floudas [2007] discussed preliminary advances in relaxing a pooling problem with the EPA Complex Emissions Model appended, but this paper represents (to the best of our knowledge) the first success in globally optimizing large-scale pooling problems with the EPA Model appended. To integrate the EPA Complex Emissions Model into the problem formulation, we use several relaxation techniques including outer approximation, recursive arithmetic [Maranas and Floudas, 1995, Ryoo and Sahinidis, 2001], and the edge-concave paradigm [Tardella, 1988/89, 2003, 2008, Meyer and Floudas, 2005]. The application of these techniques will be further described in Section 4.

Reformulated gasoline emissions standards (*e.g.*, 40CFR80.41 [2008]) are regulated using a *command-and-control* policy on environmental pollutants, so we model the EPA standards as constraints and place *profitability* in the objective function. Malcolm et al. [2006] discussed the economic and environmental dif-

[1]Also see Sherali and Adams [1999] for a comprehensive study of the reformulation-linearization technique developed by Sherali and co-workers.

ferences between regulatory policies (*i.e.*, command-and-control, environmental taxes, and cap-and-trade) and Lim and Park [2008] addressed the idea of minimizing environmental impact rather than cost. The mathematical complexity of the pooling problem with the EPA Complex Emissions Model appended would not change between different regulatory policies.

Table 2: Extended Pooling Problem Notation

Indices	$i \in 1, \ldots, \mathrm{I}$	feed stocks from refinery
	$j \in 1, \ldots, \mathrm{J}$	final products
	$k \in 1, \ldots, \mathrm{K}$	fuel qualities in EPA Model
	$l \in 1, \ldots, \mathrm{L}$	pools
	$e \in 1,\ 2$	normal ($e = 1$) and higher ($e = 2$) emitters defined by the EPA
Sets	T_X	pairs with a connection between feed i and pool l
	T_Y	pairs with a connection between pool l and product j
	T_Z	pairs with a connection between feed i and product j
Network Variables	$q_{i,l}$	proportional flow from feed i to pool l
	$y_{l,j}$	flow from pool l to product j
	$z_{i,j}$	flow from feed i to product j
	of_j	outflow rate of product j
	$u_{j,k}$	quality k of product j
	$\widehat{u_{j,3}}$	auxiliary variable used as a placeholder for $u_{j,3}^{1.25}$
	$n = 1, \cdots, N$	segments for piecewise underestimation of $q_{i,l} \cdot y_{l,j}$
	$\lambda_{i,\,l}(n)$	binary variable activates the correct $q_{i,l}$ domain section
	$w_{i,\,l,\,j}^{q,\,y}$	auxiliary variable replaces nonconvex term $q_{i,l} \cdot y_{l,j}$
	$\Delta y_{i,\,l,\,j}(n)$	auxiliary variable activates the appropriate envelope
	$w_{j,\,k}^{u,\,of}$	auxiliary variable replaces nonconvex term $u_{j,k} \cdot (of_j)$
Binary Switches	$y_{\mathrm{E300},\,j}$	binary switch for product j common to all three models
	$y_{\mathrm{ARO},\,j}$	binary switch for product j common to all three models
	$y_{\mathrm{E200},\,33,\,j}^{V}$	binary switch for product j sets a breakpoint at $u_{j,4} = 33$
	$y_{\mathrm{E200},\,65.52,\,j}^{V}$	binary switch for product j sets a breakpoint at $u_{j,4} = 65.52$
	$y_{\mathrm{E300},\,72,\,j}^{V}$	binary switch for product j sets a breakpoint at $u_{j,5} = 72$
	$y_{\mathrm{ARO},\,18,\,j}^{V}$	binary switch for product j sets a breakpoint at $u_{j,6} = 18$
	$y_{\mathrm{ARO},\,46,\,j}^{V}$	binary switch for product j sets a breakpoint at $u_{j,6} = 46$
	$y_{\mathrm{STAR},\,j}^{V}$	binary switch for product j represents aromatics & E300 interaction
	$y_{\mathrm{STAR},\,\mathrm{E300},\,j}^{V}$	binary switch for product j for the variable $\mathrm{E300}_j^V$ bound
	$y_{\mathrm{SUL},\,10,\,j}^{N}$	binary switch for product j sets a breakpoint at $u_{j,2} = 10$
	$y_{\mathrm{SUL},\,450,\,j}^{N}$	binary switch for product j sets a breakpoint at $u_{j,2} = 450$
	$y_{\mathrm{ARO},\,18,\,j}^{N}$	binary switch for product j sets a breakpoint at $u_{j,6} = 18$
	$y_{\mathrm{ARO},\,36.8,\,j}^{N}$	binary switch for product j sets a breakpoint at $u_{j,6} = 36.8$
	$y_{\mathrm{OLE},\,3.77,\,j}^{N}$	binary switch for product j sets a breakpoint at $u_{j,8} = 3.77$
	$y_{\mathrm{OLE},\,19,\,j}^{N}$	binary switch for product j sets a breakpoint at $u_{j,8} = 19$
EPA Model Variables	VOC_j	volatile organic emissions in each product j ($\frac{\mathrm{mg}}{\mathrm{mile}}$)
	NOX_j	$\mathrm{NO_X}$ emissions in each product j ($\frac{\mathrm{mg}}{\mathrm{mile}}$)
	TOX_j	toxics emissions in each product j ($\frac{\mathrm{mg}}{\mathrm{mile}}$)

continued on the next page

EPA Model Variables	VOCE_j	exhaust volatile organic emissions in each product j ($\frac{\text{mg}}{\text{mile}}$)
	VOCNE_j	non-exhaust volatile organic emissions in each product j ($\frac{\text{mg}}{\text{mile}}$)
	BENZ_j	exhaust benzene component of TOX_j ($\frac{\text{mg}}{\text{mile}}$)
	FORM_j	formaldehyde component of TOX_j ($\frac{\text{mg}}{\text{mile}}$)
	ACET_j	acetaldehyde component of TOX_j ($\frac{\text{mg}}{\text{mile}}$)
	BUTA_j	1,3-butadiene component of TOX_j ($\frac{\text{mg}}{\text{mile}}$)
	NEBENZ_j	nonexhaust benzene component of TOX_j ($\frac{\text{mg}}{\text{mile}}$)
	POM_j	polycyclic organic matter component of TOX_j ($\frac{\text{mg}}{\text{mile}}$)
	OXY_j	oxygen index in product j
	RVP_j	Reid Vapor Pressure index in product j
	BEN_j	benzene index in product j
	MTB_j	MTBE index in product j
	ETB_j	ETBE index in product j
	ETH_j	ethanol index in product j
	SUL_j^M	sulfur index in product j for models $M = V,\ N,\ T$
	E200_j^M	E200 index in product j for models $M = V,\ N,\ T$
	E300_j^M	E300 index in product j for models $M = V,\ N,\ T$
	ARO_j^M	aromatics index in product j for models $M = V,\ N,\ T$
	OLE_j^M	olefins index in product j for models $M = V,\ N,\ T$
	$\text{E300}_{\text{STAR},\,j}^V$	VOC model auxiliary variable bounds the range of E300_j
	$t_{V,\,e,\,j}$	auxiliary variable used to compute VOCE_j
	$t_{N,\,e,\,j}$	auxiliary variable used to compute NOX_j
	$t_{BE,\,e,\,j}$	auxiliary variable used to compute BENZ_j
	$t_{F,\,e,\,j}$	auxiliary variable used to compute FORM_j
	$t_{A,\,e,\,j}$	auxiliary variable used to compute ACET_j
	$t_{BU,\,e,\,j}$	auxiliary variable used to compute BUTA_j
	$f_{\text{EXT},\,e,\,1}^V(u_{j,4})$	extrapolation extends the accurate range of the VOC model
	$f_{\text{EXT},\,e,\,2}^V(u_{j,5},\ u_{j,6})$	extrapolation extends the accurate range of the VOC model
	$\Delta\text{E300}_j,\ \Delta\text{ARO}_j$	auxiliary variables for the definitions of $f_{\text{EXT},\,e,\,2}^V(u_{j,5},\ u_{j,6})$
	$f_{\text{EXT},\,e,\,1}^N(u_{j,2})$	extrapolation extends the accurate range of the NOX model
	$f_{\text{EXT},\,e,\,2}^N(u_{j,6})$	extrapolation extends the accurate range of the NOX model
	$f_{\text{EXT},\,e,\,3}^N(u_{j,8})$	extrapolation extends the accurate range of the NOX model
	$s_{\text{E200, LO},\,e,\,j}^+,\ s_{\text{E200, LO},\,e,\,j}^-$	slacks activating the lower section of $f_{\text{EXT},\,e,\,1}^V(u_{j,4})$
	$s_{\text{E200, HI},\,e,\,j}^+,\ s_{\text{E200, HI},\,e,\,j}^-$	slacks activating the upper section of $f_{\text{EXT},\,e,\,1}^V(u_{j,4})$
	$s_{\text{SUL, LO},\,e,\,j}^+,\ s_{\text{SUL, LO},\,e,\,j}^-$	slacks activating the lower section of $f_{\text{EXT},\,e,\,1}^N(u_{j,2})$
	$s_{\text{SUL, MD},\,e,\,j}^+,\ s_{\text{SUL, MD},\,e,\,j}^-$	slacks activating the middle section of $f_{\text{EXT},\,e,\,1}^N(u_{j,2})$
	$s_{\text{SUL, HI},\,e,\,j}^+,\ s_{\text{SUL, HI},\,e,\,j}^-$	slacks activating the upper section of $f_{\text{EXT},\,e,\,1}^N(u_{j,2})$
	$s_{\text{ARO, LO},\,e,\,j}^+,\ s_{\text{ARO, LO},\,e,\,j}^-$	slacks activating the lower section of $f_{\text{EXT},\,e,\,2}^N(u_{j,6})$
	$s_{\text{ARO, MD},\,e,\,j}^+,\ s_{\text{ARO, MD},\,e,\,j}^-$	slacks activating the middle section of $f_{\text{EXT},\,e,\,2}^N(u_{j,6})$
	$s_{\text{ARO, HI},\,e,\,j}^+,\ s_{\text{ARO, HI},\,e,\,j}^-$	slacks activating the upper section of $f_{\text{EXT},\,e,\,2}^N(u_{j,6})$
	$s_{\text{OLE, LO},\,e,\,j}^+,\ s_{\text{OLE, LO},\,e,\,j}^-$	slacks activating the lower section of $f_{\text{EXT},\,e,\,3}^N(u_{j,8})$
	$s_{\text{OLE, HI},\,e,\,j}^+,\ s_{\text{OLE, HI},\,e,\,j}^-$	slacks activating the upper section of $f_{\text{EXT},\,e,\,3}^N(u_{j,8})$

continued on the next page

Network Parameters	c_i	cost of feed stock i
	d_j	revenue from product j
	A_i^L - A_i^U	availability bounds on feed i
	$C_{i,k}$	quality k of feed i
	D_j^L - D_j^U	demand bounds for product j
	$P_{j,k}^L$ - $P_{j,k}^U$	bounds on quality k for product j
	S_l	volumetric capacity of pool l
EPA Model Parameters	$\text{VOC}_{\text{j, MAX}}$	VOC emissions constraint for each product j ($\frac{\text{mg}}{\text{mile}}$)
	$\text{NOX}_{j,\text{MAX}}$	NO_X emissions constraint for each product j ($\frac{\text{mg}}{\text{mile}}$)
	$\text{TOX}_{j,\text{MAX}}$	toxics emissions constraint for each product j ($\frac{\text{mg}}{\text{mile}}$)
	VOC(b)	baseline emissions of volatile organics ($\frac{\text{mg}}{\text{mile}}$)
	NOX(b)	baseline emissions of nitrous oxides ($\frac{\text{mg}}{\text{mile}}$)
	BENZ(b)	baseline emissions of exhaust benzene ($\frac{\text{mg}}{\text{mile}}$)
	ACET(b)	baseline emissions of acetaldehyde ($\frac{\text{mg}}{\text{mile}}$)
	FORM(b)	baseline emissions of formaldehyde ($\frac{\text{mg}}{\text{mile}}$)
	BUTA(b)	baseline emissions of 1,3-butadiene ($\frac{\text{mg}}{\text{mile}}$)
	w_e^V, w_e^N, w_e^T	VOC, NO_X, and toxics weighting factors for emitter e
	$c_{e,1}^V, \cdots, c_{e,10}^V$	coefficients participating in $t_{V,e,j}$
	$c_{e,1}^N, \cdots, c_{e,10}^N$	coefficients participating in $t_{N,e,j}$
	$c_{e,1}^{BE}, \cdots, c_{e,5}^{BE}$	coefficients participating in $t_{BE,e,j}$
	$c_{e,1}^F, \cdots, c_{e,4}^F$	coefficients participating in $t_{F,e,j}$
	$c_{e,1}^A, \cdots, c_{e,7}^A$	coefficients participating in $t_{A,e,j}$
	$c_{e,1}^{BU}, \cdots, c_{e,6}^{BU}$	coefficients participating in $t_{BU,e,j}$
	$\alpha_1^V, \cdots, \alpha_3^V$	coefficients participating in VOCNE_j
	α^{POM}	coefficient participating in POM_j
	$\alpha_1^{NB}, \cdots, \alpha_7^{NB}$	coefficients participating in NEBENZ_j
	α_4'	parameter determines the extent of edge concavity in NEBENZ_j
	$e^{v_e(b)}$	baseline exhaust VOC normalization term for emitter e
	$e^{n_e(b)}$	baseline nitrous oxides normalization term for emitter e
	$e^{b_e(b)}$	baseline exhaust benzene normalization term for emitter e
	$e^{f_e(b)}$	baseline formaldehyde normalization term for emitter e
	$e^{a_e(b)}$	baseline acetaldehyde normalization term for emitter e
	$e^{d_e(b)}$	baseline 1,3-butadiene normalization term for emitter e
	$U_{\text{E200},e,j}$	big-M parameter to bound the $s_{\text{E200, LO/HI},e,j}^{+/-}$ slacks
	$U_{\text{SUL},e,j}$	big-M parameter to bound the $s_{\text{SUL, LO/MD/HI},e,j}^{+/-}$ slacks
	$U_{\text{ARO},e,j}$	big-M parameter to bound the $s_{\text{ARO, LO/MD/HI},e,j}^{+/-}$ slacks
	$U_{\text{OLE},e,j}$	big-M parameter to bound the $s_{\text{OLE, LO/HI},e,j}^{+/-}$ slacks

3 MINLP Formulation

Formulating the entire MINLP model of the extended pooling problem requires a standard pooling problem backbone, a formulation of the EPA Complex Emissions Model, and relevant nonlinear blending rules:

- The standard pooling problem backbone, which acts as the basis for all the extensions, is defined in Section 3.1 by Eqs. (1) – (9). Following precedent in the pooling problem literature, our formulation

assumes *perfect mixing* at each intermediate and output node and *linear blending* of all the qualities in Table 1 except for Reid Vapor Presure (RVP).

- The EPA Complex Emission Model, discussed in Section 3.2, has three components: volatile organics, NO_X, and toxics (Sections 3.2.1 – 3.2.3). Adding the EPA model extension to the pooling problem means appending Eqs. (10) – (134) to standard pooling problem backbone. The EPA Complex Emissions Model requires exact quality inputs, so this work assumes accurate knowledge of fuel concentrations. Variations or uncertainties in the fuel qualities are addressed through sensitivity analysis of the EPA Model or parametric optimization of the entire network.

- Reid Vapor Pressure (RVP), one of the components monitored by the EPA Complex Emissions Model, blends nonlinearly. Eqs. (139) & (140) from Section 3.3 formulate the blending rule using the *Chevron Method* [Visweswaran, 2009].

To give a snapshot of the model complexity before immersion into a detailed analysis of the problem formulation, the MINLP has the following characteristics:

- Number of continuous variables: $\|T_X\| + \|T_Y\| + \|T_Z\| + 89 \cdot \mathrm{J} + \mathrm{J} \cdot \mathrm{K}$ (including the $87 \cdot \mathrm{J}$ and J continuous variables in the EPA Model and nonlinear blending extensions, respectively). The small, medium, and large case studies in Section 6 have 214, 331, and 1104 continuous variables.

- Number of binary variables: $15 \cdot \mathrm{J}$. The test cases introduced in this study have 30, 45, and 150 binary variables.

- Number of bilinear terms in the standard pooling problem:

$$\left\lfloor \sum_{\substack{(i,l)\in T_X,\\ j:(l,j)\in T_Y}} 1 \right\rfloor + \mathrm{J} \cdot \mathrm{K}$$

  For the topologies we address in Section 6, this is equivalent to $\mathrm{J} \cdot (\|T_X\| + \mathrm{K})$ or equal to 30 bilinear terms for the small test case, 63 for the mid-size case study, and 250 for the large-scale instance.

- Total nonconvex terms in the standard pooling, model extension, and nonlinear blending portions:

$$\left\lfloor \sum_{\substack{(i,l)\in T_X,\\ j:(l,j)\in T_Y}} 1 \right\rfloor + \mathrm{J} \cdot \mathrm{K} + 39 \cdot \mathrm{J}$$

  This includes the bilinear terms in the standard pooling backbone, the various nonlinear terms in the EPA Model extension ($12 \cdot \mathrm{J}$ exponential, $16 \cdot \mathrm{J}$ bilinear, $6 \cdot \mathrm{J}$ quadratic, and $4 \cdot \mathrm{J}$ higher-order polynomial), and the J terms raised to the 1.25 power in the nonlinear blending extension. In total, the small, medium, and large test cases introduced in this paper have 108, 180, and 640 nonlinear terms, respectively. The nonconvexities in the large-scale case study include 410 bilinear terms, 40 polynomial terms, and 10 terms raised to a fractional power.

As described in Table 2, T_X, T_Y, and T_Z represent sets of connections between the network nodes and I, J, K, and L represent the number of feed stocks, final products, fuel qualities, and pool nodes, respectively. The MILP relaxation of the MINLP, which is explicitly described in Section 4, uses a substantially larger number of equations and auxiliary variables to tightly relax the original problem.

3.1 Standard Pooling Problem Backbone

The standard backbone of the extended pooling problem can be expressed using a number of formulations. Because there are eleven fuel qualities, the 'Q'–formulation will require fewer bilinear terms, but will likely be a looser relaxation than the 'P'–formulation [Ben-Tal et al., 1994, Gounaris et al., 2009]. Combining the advantages of a small formulation and a tight relaxation, we choose the 'PQ'–formulation [Sherali and Alameddine, 1992, Quesada and Grossmann, 1995, Tawarmalani and Sahinidis, 2002]. The objective is shown in Eq. 1:

$$\min_{\substack{q_{i,l},\, y_{l,j},\\ z_{i,j}}} \sum_{\substack{(i,l)\in T_X\\(l,j)\in T_Y}} c_i \cdot q_{i,l} \cdot y_{l,j} - \sum_{(l,j)\in T_Y} d_j \cdot y_{l,j} - \sum_{(i,j)\in T_Z} (d_j - c_i) \cdot z_{i,j} \tag{1}$$

Eqs. (2) – (6) represent the constraints. Eq. (5) is only valid for the linearly blending qualities. The corresponding quality constraint for RVP ($k = 3$) is developed in Section 3.3.

$$A_i^L \leq \sum_{\substack{l:(i,l)\in T_X\\(l,j)\in T_Y}} q_{i,l} \cdot y_{l,j} + \sum_{j:(i,j)\in T_Z} z_{i,j} \leq A_i^U,\ \forall\, i \tag{2}$$

$$\sum_{j:(l,j)\in T_Y} y_{l,j} \leq S_l,\ \forall\, l \tag{3}$$

$$of_j = \sum_{l:(l,j)\in T_Y} y_{l,j} + \sum_{i:(i,j)\in T_Z} z_{i,j} \quad \forall\, j \tag{4}$$

$$(u_{j,k}) \cdot (of_j) = \sum_{\substack{l:(l,j)\in T_Y\\i:(i,l)\in T_X}} C_{i,k} \cdot q_{i,l} \cdot y_{l,j} + \sum_{i:(i,j)\in T_Z} C_{i,k} \cdot z_{i,j} \quad \forall\, j, \forall\, k \neq 3 \tag{5}$$

$$\sum_{i:(i,l)\in T_X} q_{i,l} = 1,\ \forall\, l \tag{6}$$

The additional cut introduced by the 'PQ'–formulation is:

$$\sum_{i:(i,l)\in T_X} q_{i,l} \cdot y_{l,j} = y_{l,j} \quad \forall\, (l,\, j) \in T_Y \tag{7}$$

and the hard bounds associated with the problem are:

$$\text{Hard Bounds} \begin{cases} 0 \leq & q_{i,l} & \leq 1 & \forall\, (i,\, l) \in T_X \\ 0 \leq & y_{l,j} & \leq \min\{S_l,\, D_j^{\mathrm{U}},\, \sum\limits_{i:(i,l)\in T_X} A_i^{\mathrm{U}}\} & \forall\, (l,\, j) \in T_Y \\ 0 \leq & z_{i,j} & \leq \min\{A_i^{\mathrm{U}},\, D_j^{\mathrm{U}}\} & \forall\, (i,\, j) \in T_Z \\ D_j^{\mathrm{L}} \leq & of_j & \leq D_j^{\mathrm{U}} & \forall\, j \\ P_{j,k}^{\mathrm{L}} \leq & u_{j,k} & \leq P_{j,k}^{\mathrm{U}} & \forall\, j,\, k \\ \min\limits_i C_{i,k} \leq & u_{j,k} & \leq \max\limits_i C_{i,k} & \forall\, j, \forall\, k \neq 3 \end{cases} \tag{8}$$

Finally, we augment the hard bounds (8), with topological restrictions derived from the network structure. These hard bounds, which were determined by studying the interaction between the variables in Eqs. (2) – (6), tighten the MILP relaxation of the bilinear terms by reducing the domain size:

$$
\begin{array}{rcl}
1 - \sum_{i \neq i} q^{\mathrm{U}}_{i,l} \;\leq\; q_{i,l} \;\leq\; 1 - \sum_{i \neq i} q^{\mathrm{L}}_{i,l} & & \forall\, (i,\, l) \in T_X \\
y_{l,j} \;\leq\; of^{\mathrm{U}}_j - \sum_{i:(i,j)\in T_Z} z^{\mathrm{L}}_{i,j} & & \forall\, (l,j) \in T_Y \\
z_{i,j} \;\leq\; of^{\mathrm{U}}_j - \sum_{l:(l,j)\in T_Y} y^{\mathrm{L}}_{l,j} & & \forall (i,\, j) \in T_Z \\
\sum_{l:(l,j)\in T_Y} y^{\mathrm{L}}_{l,j} + \sum_{i:(i,j)\in T_Z} z^{\mathrm{L}}_{i,j} \;\leq\; of_j \;\leq\; \sum_{l:(l,j)\in T_Y} y^{\mathrm{U}}_{l,j} + \sum_{i:(i,j)\in T_Z} z^{\mathrm{U}}_{i,j} & & \forall\, j
\end{array}
\tag{9}
$$

3.2 EPA Model Extension

The EPA Complex Emissions model appends three constraints to each product j:

$$
\begin{aligned}
\mathrm{VOC_j} &\leq \mathrm{VOC}_{j,\,\mathrm{MAX}} && (10)\\
\mathrm{NOX_j} &\leq \mathrm{NOX}_{j,\,\mathrm{MAX}} && (11)\\
\mathrm{TOX_j} &\leq \mathrm{TOX}_{j,\,\mathrm{MAX}} && (12)
\end{aligned}
$$

where $\mathrm{VOC}_{j,\,\mathrm{MAX}}$, $\mathrm{NOX}_{j,\,\mathrm{MAX}}$, and $\mathrm{TOX}_{j,\,\mathrm{MAX}}$ are parameters satisfying applicable legislation for each product [40CFR80.41, 2008, 40CFR80.45, 2007]. This section develops the expressions for $\mathrm{VOC_j}$, $\mathrm{NOX_j}$, and $\mathrm{TOX_j}$.

Eqs. (10) – (12) are functions of the 11 fuel components presented in Table 1. These inputs to the Complex Emissions Model are defined by the EPA as functions of the outflow fuel qualities $u_{j,k}$. Eqs. (13) & (14) relate the fuel components specified by the EPA to the outflow fuel qualities $u_{j,k}$. For oxygen, benzene, MBTE, ETBE, and ethanol, the EPA specifies that the component is equal to the fuel quality:

$$
\begin{array}{llll}
\mathrm{OXY}_j = u_{j,\,1}\ \forall j & \mathrm{BEN}_j = u_{j,\,7}\ \forall j & \mathrm{MTB}_j = u_{j,\,9}\ \forall j \\
\mathrm{ETB}_j = u_{j,\,10}\ \forall j & \mathrm{ETH}_j = u_{j,\,11}\ \forall j, &
\end{array}
\tag{13}
$$

but the value of RVP_j depends on the time of year:

$$
\mathrm{RVP}_j = \begin{cases} u_{j,\,3} & \text{Summer} \\ 8.7 & \text{Winter} \end{cases} \ \forall j. \tag{14}
$$

The EPA stipulates oxygen content is considered to be the sum of only four components: methyl tertiary butyl ether (MTBE), ethyl tertiary butyl ether (ETBE), tertiary amyl methyl ether (TAME), and ethanol. Because tracking all four additives and the oxygen content across the intermediate and output nodes is redundant, we reduce the number of variables and bilinear terms in the model by never explicitly considering TAME. All oxygenated compounds other than MBTE, ETBE, TAME, and ethanol are evaluated as if they were one of the four compounds specified by the EPA (for details, see CFR80.45). To enforce this restriction, we write:

$$
\mathrm{OXY}_j \geq \mathrm{MTB}_j + \mathrm{ETB}_j + \mathrm{ETH}_j \quad \forall\, j. \tag{15}
$$

Additionally, two binary variables, representing quality breakpoints defined by the EPA, are common to all three models:

$$
\begin{aligned}
u_{j,\,5} - 95 &\geq (u^{\mathrm{L}}_{j,\,5} - 95) \cdot y_{\mathrm{E300},\,j} && (16)\\
u_{j,\,5} - 95 &\leq (u^{\mathrm{U}}_{j,\,5} - 95) \cdot (1 - y_{\mathrm{E300},\,j}) && (17)
\end{aligned}
$$

$$u_{j,6} - 10 \;\geq\; (u^{\mathrm{L}}_{j,6} - 10) \cdot y_{\mathrm{ARO},j} \tag{18}$$
$$u_{j,6} - 10 \;\leq\; (u^{\mathrm{U}}_{j,6} - 10) \cdot (1 - y_{\mathrm{ARO},j}) \tag{19}$$

However, the EPA specifies functions for 5 fuel components (SUL, E200, E300, ARO, OLE) differently according to the specific emissions model (volatile organic, NO_X, or toxics). Sections 3.2.1 – 3.2.3 present the elements of the model that are specific to each emissions type (the models are volatile organic V, NO_X N, and toxics T).

3.2.1 Volatile Organic Compounds Emissions Model

The Volatile Organic Compounds (VOC) Emissions Model is a function of 7 fuel components presented in Table 1 (OXY_j, SUL^V_j, RVP_j, $\mathrm{E200}^V_j$, $\mathrm{E300}^V_j$, ARO^V_j & OLE^V_j). In this model, exhaust (Eq. 48) and nonexhaust (Eq. 72) components are summed to find the total volatile organics emissions (Eq. 47). As Rhodes [1998] explains, RVP is the most significant variable in the volatile organic emissions model followed by the distillation fractions at 200 and 300^{o}F. Table 8 shows that these three variables have relatively large coefficients in the exponential expression for exhaust emissions (Eq. 48). Further, the nonexhaust emissions (Eq. 72) are only a function of RVP. The solutions to our mid-size and large case studies in Section 7 match the suggestion of Rhodes [1998] to meet EPA standards by removing light ends from the fuel. Figures 8 and 9, which show the flowrates of feed stocks into intermediate storage nodes, show that the optimal solution uses very little butane.

These seven inputs to the VOC emissions equation are, in turn, functions of the outflow fuel qualities $u_{j,k}$. For the qualities in Eqs. (13) & (20), the two are identical:

$$\mathrm{SUL}^V_j = u_{j,2}\ \forall j \quad \mathrm{OLE}^V_j = u_{j,8}\ \forall j \tag{20}$$

To define the three other fuel indices ($\mathrm{E200}^V_j$, $\mathrm{E300}^V_j$, & ARO^V_j), we implement the logical disjunctions defined by the EPA as binary variables. The volatile organic compounds component of the EPA Complex Emissions Model defines two logical disjunctions at 200^{o}F distillation fractions of 33 & 65.52 vol%:

$$u_{j,4} - 33 \;\geq\; (u^{\mathrm{L}}_{j,4} - 33) \cdot y^V_{\mathrm{E200},33,\,j} \tag{21}$$
$$u_{j,4} - 33 \;\leq\; (u^{\mathrm{U}}_{j,4} - 33) \cdot (1 - y^V_{\mathrm{E200},33,j}) \tag{22}$$

$$u_{j,4} - 65.52 \;\geq\; (u^{\mathrm{L}}_{j,4} - 65.52) \cdot y^V_{\mathrm{E200},65.52,\,j} \tag{23}$$
$$u_{j,4} - 65.52 \;\leq\; (u^{\mathrm{U}}_{j,4} - 65.52) \cdot (1 - y^V_{\mathrm{E200},65.52,j}). \tag{24}$$

In addition to the logical disjunction defined by Eqs. (16) & (17), there is an E300 breakpoint at 72 vol%:

$$u_{j,5} - 72 \;\geq\; (u^{\mathrm{L}}_{j,5} - 72) \cdot y^V_{\mathrm{E300},72,\,j} \tag{25}$$
$$u_{j,5} - 72 \;\leq\; (u^{\mathrm{U}}_{j,5} - 72) \cdot (1 - y^V_{\mathrm{E300},72,j}) \tag{26}$$

The aromatics concentrations also have more breakpoints than were defined in Eqs. (18) & (19):

$$u_{j,6} - 18 \;\geq\; (u^{\mathrm{L}}_{j,6} - 18) \cdot y^V_{\mathrm{ARO},18,\,j} \tag{27}$$
$$u_{j,6} - 18 \;\leq\; (u^{\mathrm{U}}_{j,6} - 18) \cdot (1 - y^V_{\mathrm{ARO},18,j}) \tag{28}$$
$$u_{j,6} - 46 \;\geq\; (u^{\mathrm{L}}_{j,6} - 46) \cdot y^V_{\mathrm{ARO},46,\,j} \tag{29}$$
$$u_{j,6} - 46 \;\leq\; (u^{\mathrm{U}}_{j,6} - 46) \cdot (1 - y^V_{\mathrm{ARO},46,j}). \tag{30}$$

Finally, the VOC model uses binary switch $y^V_{\mathrm{STAR},\,j}$ to better represent the interaction between E300 and aromatics. We define $y^V_{\mathrm{STAR},\,j} = 0$ when $(0.385 \cdot u_{j,\,6} + 79.75) < 94$ and 1 otherwise.

$$(0.385 \cdot u_{j,\,6} + 79.75) - 94 \geq ((0.385 \cdot u^{\mathrm{L}}_{j,\,6} + 79.75) - 94) \cdot (1 - y^V_{\mathrm{STAR},\,j}) \tag{31}$$

$$(0.385 \cdot u_{j,\,6} + 79.75) - 94 \leq ((0.385 \cdot u^{\mathrm{U}}_{j,\,6} + 79.75) - 94) \cdot y^V_{\mathrm{STAR},\,j} \tag{32}$$

In the VOC component of the Complex Emissions Model, the binary switches in Eqs. (21) – (32) are used to define the fuel indices ($\mathrm{E200}^V_j$, $\mathrm{E300}^V_j$, ARO^V_j) and model extensions, which will be developed in Eqs. (50) – (71). When the 200^oF distillation fraction is less than 33 vol% or greater than 65.52 vol%, $\mathrm{E200}^V_j$ is fixed at 33 and 65.52, respectively.

$$\text{Lower Range} \quad \begin{cases} \mathrm{E200}^V_j - 33 & \leq (u^{\mathrm{U}}_{j,\,4} - 33) \cdot (1 - y^V_{\mathrm{E200},\,33,\,j}) \\ \mathrm{E200}^V_j - 33 & \geq (u^{\mathrm{L}}_{j,\,4} - 33) \cdot (1 - y^V_{\mathrm{E200},\,33,\,j}) \end{cases} \tag{33}$$

$$\text{Middle Range} \quad \begin{cases} \mathrm{E200}^V_j - u_{j,\,4} & \leq (u^{\mathrm{U}}_{j,\,4} - u^{\mathrm{L}}_{j,\,4}) \cdot (y^V_{\mathrm{E200},\,33,\,j} + 1 - y^V_{\mathrm{E200},\,65.52,\,j}) \\ \mathrm{E200}^V_j - u_{j,\,4} & \geq (u^{\mathrm{L}}_{j,\,4} - u^{\mathrm{U}}_{j,\,4}) \cdot (y^V_{\mathrm{E200},\,33,\,j} + 1 - y^V_{\mathrm{E200},\,65.52,\,j}) \end{cases} \tag{34}$$

$$\text{Higher Range} \quad \begin{cases} \mathrm{E200}^V_j - 65.52 & \leq (u^{\mathrm{U}}_{j,\,4} - 65.52) \cdot y^V_{\mathrm{E200},\,65.52,\,j} \\ \mathrm{E200}^V_j - 65.52 & \geq (u^{\mathrm{L}}_{j,\,4} - 65.52) \cdot y^V_{\mathrm{E200},\,65.52,\,j}. \end{cases} \tag{35}$$

Similarly, the value ARO^N_j is usually $u_{j,\,6}$ but is fixed to 18 vol% when $u_{j,\,6} \leq 18$ and to 46 vol% when $u_{j,\,6} \geq 46$:

$$\text{Lower Range} \quad \begin{cases} \mathrm{ARO}^V_j - 18 & \leq (u^{\mathrm{U}}_{j,\,6} - 18) \cdot (1 - y^V_{\mathrm{ARO},\,18,\,j}) \\ \mathrm{ARO}^V_j - 18 & \geq (u^{\mathrm{L}}_{j,\,6} - 18) \cdot (1 - y^V_{\mathrm{ARO},\,18,\,j}) \end{cases} \tag{36}$$

$$\text{Middle Range} \quad \begin{cases} \mathrm{ARO}^V_j - u_{j,\,6} & \leq (u^{\mathrm{U}}_{j,\,6} - u^{\mathrm{L}}_{j,\,6}) \cdot (y^V_{\mathrm{ARO},\,18,\,j} + 1 - y^V_{\mathrm{ARO},\,46,\,j}) \\ \mathrm{ARO}^V_j - u_{j,\,6} & \geq (u^{\mathrm{L}}_{j,\,6} - u^{\mathrm{U}}_{j,\,6}) \cdot (y^V_{\mathrm{ARO},\,18,\,j} + 1 - y^V_{\mathrm{ARO},\,46,\,j}) \end{cases} \tag{37}$$

$$\text{Higher Range} \quad \begin{cases} \mathrm{ARO}^V_j - 46 & \leq (u^{\mathrm{U}}_{j,\,6} - 46) \cdot y^V_{\mathrm{ARO},\,46,\,j} \\ \mathrm{ARO}^V_j - 46 & \geq (u^{\mathrm{L}}_{j,\,6} - 46) \cdot y^V_{\mathrm{ARO},\,46,\,j}. \end{cases} \tag{38}$$

But the representation of $\mathrm{E300}^V_j$ is not as straightforward. Because the upper bound of $\mathrm{E300}^V_j$ depends on aromatics in the EPA model, we define auxiliary variable $\mathrm{E300}^V_{\mathrm{STAR},\,j}$ to represent the variable bound of $\mathrm{E300}^V_j$:

$$\text{Lower Range} \quad \begin{cases} \mathrm{E300}^V_{\mathrm{STAR},\,j} - 0.385 \cdot 10 - 79.75 & \leq 0.385 \cdot (u^{\mathrm{U}}_{j,\,6} - 10) \\ & \qquad \cdot (y^V_{\mathrm{STAR},\,j} + 1 - y_{\mathrm{ARO},\,j}) \\ \mathrm{E300}^V_{\mathrm{STAR},\,j} - 0.385 \cdot 10 - 79.75 & \geq 0.385 \cdot (u^{\mathrm{L}}_{j,\,6} - 10) \\ & \qquad \cdot (y^V_{\mathrm{STAR},\,j} + 1 - y_{\mathrm{ARO},\,j}) \end{cases} \tag{39}$$

$$\text{Middle Range} \quad \begin{cases} \mathrm{E300}^V_{\mathrm{STAR},\,j} - 0.385 \cdot u_{j,\,6} - 79.75 & \leq 0.385 \cdot (u^{\mathrm{U}}_{j,\,6} - u^{\mathrm{L}}_{j,\,6}) \\ & \qquad \cdot (y^V_{\mathrm{STAR},\,j} + y_{\mathrm{ARO},\,j}) \\ \mathrm{E300}^V_{\mathrm{STAR},\,j} - 0.385 \cdot u_{j,\,6} - 79.75 & \geq 0.385 \cdot (u^{\mathrm{L}}_{j,\,6} - u^{\mathrm{U}}_{j,\,6}) \\ & \qquad \cdot (y^V_{\mathrm{STAR},\,j} + y_{\mathrm{ARO},\,j}) \end{cases} \tag{40}$$

$$\text{Higher Range} \quad \begin{cases} \text{E300}^V_{\text{STAR},\,j} - 94 & \leq (0.385 \cdot u^{\text{U}}_{j,\,6} + 79.75 - 94) \cdot (1 - y^V_{\text{STAR},\,j}) \\ \text{E300}^V_{\text{STAR},\,j} - 94 & \geq (0.385 \cdot u^{\text{L}}_{j,\,6} + 79.75 - 94) \cdot (1 - y^V_{\text{STAR},\,j}). \end{cases} \tag{41}$$

We also define binary switch $y_{\text{STAR, E30},\,j} = 1$ when $\text{E300}^V_{\text{STAR},\,j} > u_{j,5}$ and 0 otherwise:

$$u_{j,5} - \text{E300}^V_{\text{STAR},\,j} \geq (u^{\text{L}}_{j,\,5} - 0.385 \cdot u^{\text{U}}_{j,\,6} - 79.75) \cdot y^V_{\text{STAR, E300},\,j} \tag{42}$$

$$u_{j,5} - \text{E300}^V_{\text{STAR},\,j} \leq (u^{\text{U}}_{j,\,5} - 0.385 \cdot 10 - 79.75) \cdot (1 - y^V_{\text{STAR, E300},\,j}) \tag{43}$$

Now, $\text{E300}^V_j = u_{j,\,6}$ when $u_{j,\,6}$ is between 72 vol% and $\text{E300}^V_{\text{STAR},\,j}$ and fixed to one of the two bounds otherwise:

$$\text{Lower Range} \begin{cases} \text{E300}^V_j - 72 & \leq (u^{\text{U}}_{j,\,5} - 72) \cdot (1 - y^V_{\text{E300, 72},\,j}) \\ \text{E300}^V_j - 72 & \geq (u^{\text{L}}_{j,\,5} - 72) \cdot (1 - y^V_{\text{E300, 72},\,j}) \end{cases} \tag{44}$$

$$\text{Middle Range} \quad \begin{cases} \text{E300}^V_j - u_{j,\,5} & \leq (u^{\text{U}}_{j,\,5} - u^{\text{L}}_{j,\,5}) \cdot (y^V_{\text{E300, 72},\,j} + 1 - y^V_{\text{STAR, E300},\,j}) \\ \text{E300}^V_j - u_{j,\,5} & \geq (u^{\text{L}}_{j,\,5} - u^{\text{U}}_{j,\,5}) \cdot (y^V_{\text{E300, 72},\,j} + 1 - y^V_{\text{STAR, E300},\,j}) \end{cases} \tag{45}$$

$$\text{Higher Range} \quad \begin{cases} \text{E300}^V_j - \text{E300}^V_{\text{STAR},\,j} & \leq (u^{\text{U}}_{j,\,5} - 0.385 \cdot 10 - 79.75) \cdot y^V_{\text{STAR, E300},\,j} \\ \text{E300}^V_j - \text{E300}^V_{\text{STAR},\,j} & \geq (u^{\text{L}}_{j,\,5} - 0.385 \cdot u^{\text{U}}_{j,\,6} - 79.75) \cdot y^V_{\text{STAR, E300},\,j}. \end{cases} \tag{46}$$

The fuel indices developed in Eqs. (33) – (46) are used in the VOC component of the VOC model, which is the sum of exhaust (VOCNE_j) and non-exhaust (VOCNE_j) components:

$$\text{VOC}_j = \text{VOCE}_j + \text{VOCNE}_j. \tag{47}$$

The exhaust component has an exponentiated linear term multiplied by some model extensions (see Eqs. 50 & 71):

$$\text{VOCE}_j = \sum_{e=1}^{2} \frac{\text{VOC}(b) \cdot w^V_e}{e^{v_e(b)}} \cdot \exp\{t_{V,\,e,\,j}\} \cdot \{1 + f^V_{\text{EXT},\,e,\,1}(\text{E200}_j) + f^V_{\text{EXT},\,e,\,2}(\text{E300}_j,\, \text{ARO}_j)\}. \tag{48}$$

where the exponentiated linear term $t_{V,\,e,\,j}$ is a function of the fuel components:

$$\begin{aligned} t_{V,\,e,\,j} = & c^V_{e,\,1} \cdot \text{OXY}_j + c^V_{e,\,2} \cdot \text{SUL}^V_j + c^V_{e,\,3} \cdot \text{RVP}_j + c^V_{e,\,4} \cdot \text{E200}^V_j + c^V_{e,\,5} \cdot \text{E300}^V_j + c^V_{e,\,6} \cdot \text{ARO}^V_j + \\ & c^V_{e,\,7} \cdot \text{OLE}^V_j + c^V_{e,\,8} \cdot (\text{E200}^V_j)^2 + c^V_{e,\,9} \cdot (\text{E300}^V_j)^2 + c^V_{e,\,10} \cdot \text{ARO}^V_j \cdot \text{E300}^V_j \end{aligned} \tag{49}$$

The final component of Eq. (48), the summation of model extensions, increase the region of model applicability. Although the text of CFR80.45 presents the $f^V_{\text{EXT},\,e,\,1}(u_{j,\,4})$ extension as a bilinear function, closer examination (Figure 1) shows that the function is, without approximation, piecewise-linear. However, the extensions $f^V_{\text{EXT},\,1,\,2}(u_{j,\,5},\, u_{j,\,6})$ and $f^V_{\text{EXT},\,2,\,2}(u_{j,\,5},\, u_{j,\,6})$ are bilinear in some domain regions.

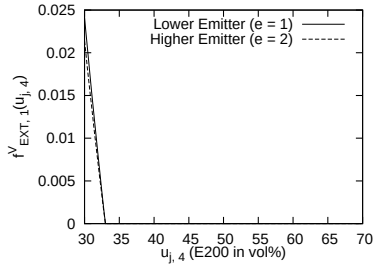
To represent the piecewise-linear nature of $f^V_{\text{EXT},\,e,\,1}(u_{j,\,4})$, slack variables activate a specific area of the domain [Floudas, 1995]:

$$s^+_{\text{E200, LO},\,e,\,j} + s^-_{\text{E200, LO},\,e,\,j} \leq U_{\text{E200},\,e,\,j} \cdot (1 - y^V_{\text{E200, 33},\,j}) \tag{50}$$

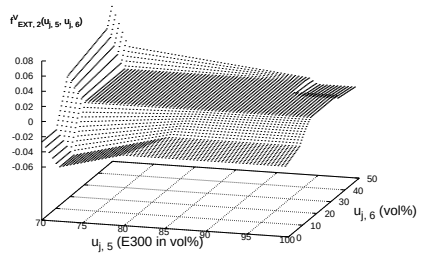
$$s^+_{\text{E200, HI},\,e,\,j} + s^-_{\text{E200, HI},\,e,\,j} \leq U_{\text{E200},\,e,\,j} \cdot y^V_{\text{E200, 33},\,j} \tag{51}$$

$$f^V_{\mathrm{EXT},\,e,\,1}(u_{j,\,4}) + s^+_{\mathrm{E200,\,LO},\,e,\,j} - s^-_{\mathrm{E200,\,LO},\,e,\,j} = (2 \cdot c^V_{e,8} \cdot 33 + c^V_{e,4}) \times (u_{j,\,4} - 33) \quad (52)$$

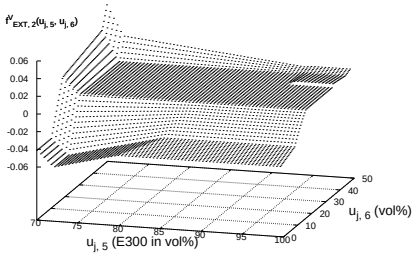
$$f^V_{\mathrm{EXT},\,e,\,1}(u_{j,\,4}) + s^+_{\mathrm{E200,\,HI},\,e,\,j} - s^-_{\mathrm{E200,\,HI},\,e,\,j} = 0.0. \quad (53)$$



(a) $f^V_{\mathrm{EXT},\,e,\,1}(u_{j,\,4})$



(b) $f^V_{\mathrm{EXT},\,1,\,2}(u_{j,\,5},\,u_{j,\,6})$



(c) $f^V_{\mathrm{EXT},\,2,\,2}(u_{j,\,5},\,u_{j,\,6})$

Figure 1: Representation of the piecewise VOC model extension functions

Because $f^V_{\mathrm{EXT},\,1,\,2}(u_{j,\,5},\,u_{j,\,6})$ is bilinear, we define two auxiliary variables, $\Delta\mathrm{ARO}_j$ & $\Delta\mathrm{E300}_j$ where:

$$\Delta\mathrm{E300}_j - u_{j,\,5} + 72 \leq (\Delta\mathrm{E300}^{\mathrm{U}}_j - u^{\mathrm{L}}_{j,\,5} + 72) \cdot (1 - y^V_{\mathrm{E300,\,72},\,j}) \quad (54)$$

$$\Delta\mathrm{E300}_j - u_{j,\,5} + 72 \geq (\Delta\mathrm{E300}^{\mathrm{L}}_j - u^{\mathrm{U}}_{j,\,5} + 72) \cdot (1 - y^V_{\mathrm{E300,\,72},\,j}) \quad (55)$$

$$\Delta\mathrm{E300}_j \leq \Delta\mathrm{E300}^{\mathrm{U}}_j \cdot (y^V_{\mathrm{E300,\,72},\,j} + y^V_{\mathrm{STAR},\,j}) \quad (56)$$

$$\Delta\mathrm{E300}_j \geq \Delta\mathrm{E300}^{\mathrm{L}}_j \cdot (y^V_{\mathrm{E300,\,72},\,j} + y^V_{\mathrm{STAR},\,j}) \quad (57)$$

$$\Delta\mathrm{E300}_j \leq \Delta\mathrm{E300}^{\mathrm{U}}_j \cdot (y^V_{\mathrm{E300,\,72},\,j} + 1 - y^V_{\mathrm{STAR,\,E300},\,j}) \quad (58)$$

$$\Delta\mathrm{E300}_j \geq \Delta\mathrm{E300}^{\mathrm{L}}_j \cdot (y^V_{\mathrm{E300,\,72},\,j} + 1 - y^V_{\mathrm{STAR,\,E300},\,j}) \quad (59)$$

$$\Delta\mathrm{E300}_j - u_{j,\,5} + 94 \leq (\Delta\mathrm{E300}^{\mathrm{U}}_j - u^{\mathrm{L}}_{j,\,5} + 94) \cdot (y^V_{\mathrm{STAR,\,E300},\,j} + 1 - y^V_{\mathrm{STAR},\,j}) \quad (60)$$

$$\Delta\mathrm{E300}_j - u_{j,\,5} + 94 \geq (\Delta\mathrm{E300}^{\mathrm{L}}_j - u^{\mathrm{U}}_{j,\,5} + 94) \cdot (y^V_{\mathrm{STAR,\,E300},\,j} + 1 - y^V_{\mathrm{STAR},\,j}) \quad (61)$$

$$\Delta\mathrm{E300}_j \leq 1 \quad (62)$$

and:

$$\Delta\mathrm{ARO}_j + 8 \leq (\Delta\mathrm{ARO}^{\mathrm{U}}_j + 8) \cdot (1 - y_{\mathrm{ARO},\,j}) \quad (63)$$

$$\Delta\mathrm{ARO}_j + 8 \geq (\Delta\mathrm{ARO}^{\mathrm{L}}_j + 8) \cdot (1 - y_{\mathrm{ARO},\,j}) \quad (64)$$

$$\Delta\mathrm{ARO}_j - u_{j,\,6} + 18 \leq (\Delta\mathrm{ARO}^{\mathrm{U}}_j - u^{\mathrm{L}}_{j,\,6} + 18) \cdot (y_{\mathrm{ARO},\,j} + 1 - y^V_{\mathrm{ARO,\,18},\,j}) \quad (65)$$

$$\Delta\mathrm{ARO}_j - u_{j,\,6} + 18 \geq (\Delta\mathrm{ARO}^{\mathrm{L}}_j - u^{\mathrm{U}}_{j,\,6} + 18) \cdot (y_{\mathrm{ARO},\,j} + 1 - y^V_{\mathrm{ARO,\,18},\,j}) \quad (66)$$

$$\begin{aligned}
\Delta\mathrm{ARO}_j &\leq \Delta\mathrm{ARO}_j^{\mathrm{U}} \cdot (y^V_{\mathrm{ARO},\,18,\,j} + 1 - y^V_{\mathrm{ARO},\,46,\,j}) && (67)\\
\Delta\mathrm{ARO}_j &\geq \Delta\mathrm{ARO}_j^{\mathrm{L}} \cdot (y^V_{\mathrm{ARO},\,18,\,j} + 1 - y^V_{\mathrm{ARO},\,46,\,j}) && (68)\\
\Delta\mathrm{ARO}_j - u_{j,\,6} + 46 &\leq (\Delta\mathrm{ARO}_j^{\mathrm{U}} - u_{j,\,6}^{\mathrm{L}} + 46) \cdot y^V_{\mathrm{ARO},\,46,\,j} && (69)\\
\Delta\mathrm{ARO}_j - u_{j,\,6} + 46 &\geq (\Delta\mathrm{ARO}_j^{\mathrm{L}} - u_{j,\,6}^{UB} + 46) \cdot y^V_{\mathrm{ARO},\,46,\,j}. && (70)
\end{aligned}$$

The upper (UB) and lower (LB) bounds on each of the variables is controlled through preprocessing analysis. Using the two new auxiliary variables $\Delta\mathrm{E300}_j$ and $\Delta\mathrm{ARO}_j$, we can define the model extension:

$$\begin{aligned}
f^V_{\mathrm{EXT},\,e,\,2}(u_{j,\,5},\, u_{j,\,6}) =& (2 \cdot c^V_{e,\,9} \cdot \mathrm{E300}^V_j + c^V_{e,\,5} + c^V_{e,\,10} \cdot \mathrm{ARO}^V_j) \cdot \Delta\mathrm{E300}_j \\
&+ (c^V_{e,\,10} \cdot \mathrm{E300}^V_j + c^V_{e,\,6}) \cdot \Delta\mathrm{ARO}_j.
\end{aligned} \tag{71}$$

Eq. 71 completes the definition of VOCE_j. The other component of VOC_j is defined as follows:

$$\mathrm{VOCNE}_j = \begin{cases} 0.0 & \text{Winter} \\ \alpha^V_1 + \alpha^V_2 \cdot \mathrm{RVP}_j + \alpha^V_3 \cdot \mathrm{RVP}^2_j & \text{Summer} \end{cases} \forall\, j. \tag{72}$$

3.2.2 $\mathrm{NO_X}$ Emissions Model

The $\mathrm{NO_X}$ (NOX) Emissions Model is a function of 7 types of fuel components presented in Table 1 (OXY_j, SUL^N_j, RVP_j, $\mathrm{E200}^N_j$, $\mathrm{E300}^N_j$, ARO^N_j & OLE^N_j). Equations (74) – (96) and (99) – (114) define various disjunctions and edge cases, but Eqs. (97) and (98) represent the exponential expression at the backbone of the $\mathrm{NO_X}$ model. Sulfur is the fuel quality with the largest contribution to this model, but RVP, aromatics, and olefins also play an important role [Rhodes, 1998].

These seven inputs to the NOX emissions equation are, in turn, functions of the outflow fuel qualities $u_{j,k}$. For the distillation fraction at 200^{o}F, the two are identical:

$$\mathrm{E200}^N_j = u_{j,\,4}\ \forall j \tag{73}$$

To define the four other fuel indices, we implement the logical disjunctions defined by the EPA as binary variables. The nitrous oxides component of the EPA Complex Emissions Model defines two logical disjunctions at sulfur concentrations of 10 & 450 ppm:

$$\begin{aligned}
u_{j,\,2} - 10 &\geq (u^{\mathrm{L}}_{j,\,2} - 10) \cdot y^N_{\mathrm{SUL},10,\ j} && (74)\\
u_{j,\,2} - 10 &\leq (u^{\mathrm{U}}_{j,\,2} - 10) \cdot (1 - y^N_{\mathrm{SUL},\,10,\,j}) && (75)\\
u_{j,\,2} - 450 &\geq (u^{\mathrm{L}}_{j,\,2} - 450) \cdot y^N_{\mathrm{SUL},450,\ j} && (76)\\
u_{j,\,2} - 450 &\leq (u^{\mathrm{U}}_{j,\,2} - 450) \cdot (1 - y^N_{\mathrm{SUL},\,450,\,j}) && (77)
\end{aligned}$$

In addition to the aromatics breakpoint defined in Eqs. (18) & (19), the $\mathrm{NO_X}$ model defines logical disjunctions at aromatics concentrations of 18 & 36.8 vol%:

$$\begin{aligned}
u_{j,\,6} - 18 &\geq (u^{\mathrm{L}}_{j,\,6} - 18) \cdot y^N_{\mathrm{ARO},18,\ j} && (78)\\
u_{j,\,6} - 18 &\leq (u^{\mathrm{U}}_{j,\,6} - 18) \cdot (1 - y^N_{\mathrm{ARO},\,18,\,j}) && (79)\\
u_{j,\,6} - 36.8 &\geq (u^{\mathrm{L}}_{j,\,6} - 36.8) \cdot y^N_{\mathrm{ARO},36.8,\ j} && (80)\\
u_{j,\,6} - 36.8 &\leq (u^{\mathrm{U}}_{j,\,6} - 36.8) \cdot (1 - y^N_{\mathrm{ARO},\,36.8,\,j}). && (81)
\end{aligned}$$

Finally, breakpoints are defined at olefins concentrations of 3.77 & 19 vol%:

$$u_{j,8} - 3.77 \geq (u^{\mathrm{L}}_{j,8} - 3.77) \cdot y^{N}_{\mathrm{OLE},3.77,\,j} \tag{82}$$

$$u_{j,8} - 3.77 \leq (u^{\mathrm{U}}_{j,8} - 3.77) \cdot (1 - y^{N}_{\mathrm{OLE},\,3.77,\,j}) \tag{83}$$

$$u_{j,8} - 19 \geq (u^{\mathrm{L}}_{j,8} - 19) \cdot y^{N}_{\mathrm{OLE},19,\,j} \tag{84}$$

$$u_{j,8} - 19 \leq (u^{\mathrm{U}}_{j,8} - 19) \cdot (1 - y^{N}_{\mathrm{OLE},\,19,\,j}). \tag{85}$$

In the $\mathrm{NO_X}$ component of the EPA model, these breakpoints are used to define the fuel indices (SUL^N_j, $\mathrm{E300}^N_j$, ARO^N_j & OLE^N_j) and model extensions, which will be developed in Eqs. (99) – (114). When the concentration of sulfur ($u_{j,2}$) is less than 10 ppm, the fuel index (SUL^N_j) is set to 10 ppm for the purpose of calculating the model. Similarly, SUL^N_j is 450 ppm when $u_{j,2}$ is greater than 450 ppm:

$$\text{Lower Range} \quad \begin{cases} \mathrm{SUL}^N_j - 10 \leq (u^{\mathrm{U}}_{j,2} - 10) \cdot (1 - y^{N}_{\mathrm{SUL},\,10,\,j}) \\ \mathrm{SUL}^N_j - 10 \geq (u^{\mathrm{L}}_{j,2} - 10) \cdot (1 - y^{N}_{\mathrm{SUL},\,10,\,j}) \end{cases} \tag{86}$$

$$\text{Middle Range} \quad \begin{cases} \mathrm{SUL}^N_j - u_{j,2} \leq (u^{\mathrm{U}}_{j,2} - u^{\mathrm{L}}_{j,2}) \cdot (y^{N}_{\mathrm{SUL},\,10,\,j} + 1 - y^{N}_{\mathrm{SUL},\,450,\,j}) \\ \mathrm{SUL}^N_j - u_{j,2} \geq (u^{\mathrm{L}}_{j,2} - u^{\mathrm{U}}_{j,2}) \cdot (y^{N}_{\mathrm{SUL},\,10,\,j} + 1 - y^{N}_{\mathrm{SUL},\,450,\,j}) \end{cases} \tag{87}$$

$$\text{Higher Range} \quad \begin{cases} \mathrm{SUL}^N_j - 450 \leq (u^{\mathrm{U}}_{j,2} - 450) \cdot y^{N}_{\mathrm{SUL},\,450,\,j} \\ \mathrm{SUL}^N_j - 450 \geq (u^{\mathrm{L}}_{j,2} - 450) \cdot y^{N}_{\mathrm{SUL},\,450,\,j} \end{cases} \tag{88}$$

$\mathrm{E300}^N_j$ is fixed to 95 vol% when $u_{j,5} > 95$:

$$\text{Lower Range} \quad \begin{cases} \mathrm{E300}^N_j - 95 \leq (u^{\mathrm{U}}_{j,5} - 95) \cdot y_{\mathrm{E300},\,j} \\ \mathrm{E300}^N_j - 95 \geq (u^{\mathrm{L}}_{j,5} - 95) \cdot y_{\mathrm{E300},\,j} \end{cases} \tag{89}$$

$$\text{Higher Range} \quad \begin{cases} \mathrm{E300}^N_j - u_{j,5} \leq (u^{\mathrm{U}}_{j,5} - u^{\mathrm{L}}_{j,5}) \cdot (1 - y_{\mathrm{E300},\,j}) \\ \mathrm{E300}^N_j - u_{j,5} \geq (u^{\mathrm{L}}_{j,5} - u^{\mathrm{U}}_{j,5}) \cdot (1 - y_{\mathrm{E300},\,j}) \end{cases} \tag{90}$$

and the value ARO^N_j is usually $u_{j,6}$ but is fixed to 18 vol% when $u_{j,6} \leq 18$ and to 36.8 vol% when $u_{j,6} \geq 36.8$:

$$\text{Lower Range} \quad \begin{cases} \mathrm{ARO}^N_j - 18 \leq (u^{\mathrm{U}}_{j,6} - 18) \cdot (1 - y^{N}_{\mathrm{ARO},\,18,\,j}) \\ \mathrm{ARO}^N_j - 18 \geq (u^{\mathrm{L}}_{j,6} - 18) \cdot (1 - y^{N}_{\mathrm{ARO},\,18,\,j}) \end{cases} \tag{91}$$

$$\text{Middle Range} \quad \begin{cases} \mathrm{ARO}^N_j - u_{j,6} \leq (u^{\mathrm{U}}_{j,6} - u^{\mathrm{L}}_{j,6}) \cdot (y^{N}_{\mathrm{ARO},\,18,\,j} + 1 - y^{N}_{\mathrm{ARO},\,36.8,\,j}) \\ \mathrm{ARO}^N_j - u_{j,6} \geq (u^{\mathrm{L}}_{j,6} - u^{\mathrm{U}}_{j,6}) \cdot (y^{N}_{\mathrm{ARO},\,18,\,j} + 1 - y^{N}_{\mathrm{ARO},\,36.8,\,j}) \end{cases} \tag{92}$$

$$\text{Higher Range} \quad \begin{cases} \mathrm{ARO}^N_j - 36.8 \leq (u^{\mathrm{U}}_{j,6} - 36.8) \cdot y^{N}_{\mathrm{ARO},\,36.8,\,j} \\ \mathrm{ARO}^N_j - 36.8 \geq (u^{\mathrm{L}}_{j,6} - 36.8) \cdot y^{N}_{\mathrm{ARO},\,36.8,\,j}. \end{cases} \tag{93}$$

Finally, $\mathrm{OLE}^N_j = u_{j,6}$ except when $u_{j,8} \leq 3.77$ and it is fixed to 3.77 vol% or $u_{j,8} \geq 19$ and it is fixed to 19 vol%:

$$\text{Lower Range} \quad \begin{cases} \mathrm{OLE}^N_j - 3.77 \leq (u^{\mathrm{U}}_{j,8} - 3.77) \cdot (1 - y^{N}_{\mathrm{OLE},\,3.77,\,j}) \\ \mathrm{OLE}^N_j - 3.77 \geq (u^{\mathrm{L}}_{j,8} - 3.77) \cdot (1 - y^{N}_{\mathrm{OLE},\,3.77,\,j}) \end{cases} \tag{94}$$

$$\text{Middle Range} \quad \begin{cases} \text{OLE}_j^N - u_{j,8} & \leq (u_{j,8}^{\text{U}} - u_{j,6}^{\text{L}}) \cdot (y_{\text{OLE},\,3.77,\,j}^N + 1 - y_{\text{OLE},\,19,\,j}^N) \\ \text{OLE}_j^N - u_{j,8} & \geq (u_{j,8}^{\text{L}} - u_{j,8}^{\text{U}}) \cdot (y_{\text{OLE},\,3.77,\,j}^N + 1 - y_{\text{OLE},\,19,\,j}^N) \end{cases} \tag{95}$$

$$\text{Higher Range} \quad \begin{cases} \text{OLE}_j^N - 19 & \leq (u_{j,8}^{\text{U}} - 19) \cdot y_{\text{OLE},\,19,\,j}^N \\ \text{OLE}_j^N - 19 & \geq (u_{j,8}^{\text{L}} - 19) \cdot y_{\text{OLE},\,19,\,j}^N. \end{cases} \tag{96}$$

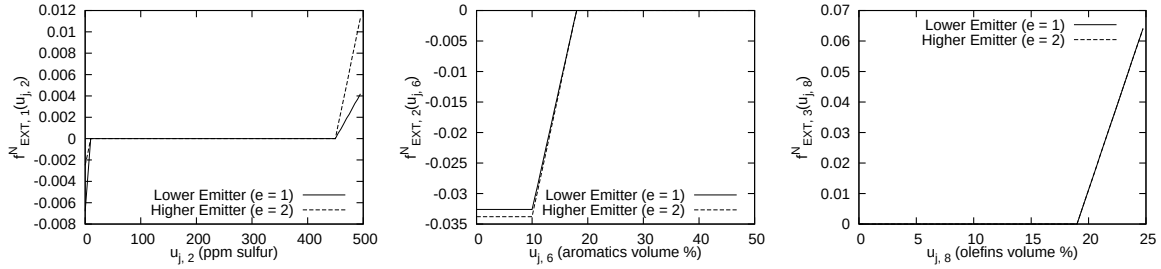
The fuel indices developed in Eqs. (86) – (96) are used in the NO_X component of the Complex Emissions Model:

$$\begin{aligned} \text{NOX}_j = \sum_{e=1}^{2} \frac{\text{NOX}(b) \cdot w_e^N}{e^{n_e(b)}} \cdot \exp\{t_{N,\,e,\,j}\} \cdot \{1 + f_{\text{EXT},\,e,\,1}^N(u_{j,\,2}) + f_{\text{EXT},\,e,\,2}^N(u_{j,\,6}) \\ + f_{\text{EXT},\,e,\,3}^N(u_{j,\,8})\} \end{aligned} \tag{97}$$

as the components of the exponentiated term:

$$\begin{aligned} t_{N,\,e,\,j} = & c_{e,\,1}^N \cdot \text{OXY}_j + c_{e,\,2}^N \cdot \text{SUL}_j^N + c_{e,\,3}^N \cdot \text{RVP}_j + c_{e,\,4}^N \cdot \text{E200}_j^N + c_{e,\,5}^N \cdot \text{E300}_j^N + \\ & c_{e,\,6}^N \cdot \text{ARO}_j^N + c_{e,\,7}^N \cdot \text{OLE}_j^N + c_{e,\,8}^N \cdot (\text{SUL}_j^N)^2 + c_{e,\,9}^N \cdot (\text{ARO}_j^N)^2 + c_{e,\,10}^N \cdot (\text{OLE}_j^N)^2 \end{aligned} \tag{98}$$

The final component of the NO_X model are the extrapolations the EPA specifies to widen the region of model applicability. Although a cursory reading of CFR80.45 makes these extrapolations ($f_{\text{EXT},\,e,\,1}^N(u_{j,\,2})$, $f_{\text{EXT},\,e,\,2}^N(u_{j,6})$, & $f_{\text{EXT},\,e,\,3}^N(u_{j,\,8})$) look like bilinear terms, closer examination (illustrated in Figure 2) demonstrates that the model extrapolations are, in fact, piecewise linear.



(a) $f_{\text{EXT},\,e,\,1}^N(u_{j,\,2})$ (b) $f_{\text{EXT},\,e,\,2}^N(u_{j,\,6})$ (c) $f_{\text{EXT},\,e,\,3}^N(u_{j,\,8})$

Figure 2: Representation of the piecewise-linear NO_X model extension functions

Because each of the functions in Figure 2 is piecewise-defined, we introduce slack variables to activate a particular segment of the domain. For sulfur, there three relevant regions:

$$s_{\text{SUL,\,LO},\,e,\,j}^+ + s_{\text{SUL,\,LO},\,e,\,j}^- \leq U_{\text{SUL},\,e,\,j} \cdot (1 - y_{\text{SUL},\,10,\,j}^N) \tag{99}$$

$$s_{\text{SUL,\,MD},\,e,\,j}^+ + s_{\text{SUL,\,MD},\,e,\,j}^- \leq U_{\text{SUL},\,e,\,j} \cdot (1 - y_{\text{SUL},\,450,\,j}^N + y_{\text{SUL},\,10,\,j}^N) \tag{100}$$

$$s_{\text{SUL,\,HI},\,e,\,j}^+ + s_{\text{SUL,\,HI},\,e,\,j}^- \leq U_{\text{SUL},\,e,\,j} \cdot y_{\text{SUL},\,450,\,j}^N \tag{101}$$

$$f_{\text{EXT},\,e,\,1}^N(u_{j,\,2}) + s_{\text{SUL,\,LO},\,e,\,j}^+ - s_{\text{SUL,\,LO},\,e,\,j}^- = (2 \cdot c_{e,8}^N \cdot 10 + c_{e,2}^N) \times (u_{j,\,2} - 10) \tag{102}$$

$$f_{\text{EXT},\,e,\,1}^N(u_{j,\,2}) + s_{\text{SUL,\,MD},\,e,\,j}^+ - s_{\text{SUL,\,MD},\,e,\,j}^- = 0.0 \tag{103}$$

$$f_{\text{EXT},\,e,\,1}^N(u_{j,\,2}) + s_{\text{SUL,\,HI},\,e,\,j}^+ - s_{\text{SUL,\,HI},\,e,\,j}^- = (2 \cdot c_{e,8}^N \cdot 450 + c_{e,2}^N) \times (u_{j,\,2} - 450). \tag{104}$$

In the case of the aromatics extrapolation, there are also three segments to consider:

$$
\begin{aligned}
s^{+}_{\mathrm{ARO},\,\mathrm{LO},\,e,\,j} + s^{-}_{\mathrm{ARO},\,\mathrm{LO},\,e,\,j} &\leq U_{\mathrm{ARO},\,e,\,j} \cdot (1 - y_{\mathrm{ARO},\,j}) && (105)\\
s^{+}_{\mathrm{ARO},\,\mathrm{MD},\,e,\,j} + s^{-}_{\mathrm{ARO},\,\mathrm{MD},\,e,\,j} &\leq U_{\mathrm{ARO},\,e,\,j} \cdot (1 - y^{N}_{\mathrm{ARO},\,18,\,j} + y_{\mathrm{ARO},\,j}) && (106)\\
s^{+}_{\mathrm{ARO},\,\mathrm{HI},\,e,\,j} + s^{-}_{\mathrm{ARO},\,\mathrm{HI},\,e,\,j} &\leq U_{\mathrm{ARO},\,e,\,j} \cdot y^{N}_{\mathrm{ARO},\,18,\,j} && (107)\\
f^{N}_{\mathrm{EXT},\,e,\,2}(u_{j,\,6}) + s^{+}_{\mathrm{ARO},\,\mathrm{LO},\,e,\,j} - s^{-}_{\mathrm{ARO},\,\mathrm{LO},\,e,\,j} &= (2 \cdot c^{N}_{e,9} \cdot 18 + c^{N}_{e,6}) \times (-8) && (108)\\
f^{N}_{\mathrm{EXT},\,e,\,2}(u_{j,\,6}) + s^{+}_{\mathrm{ARO},\,\mathrm{MD},\,e,\,j} - s^{-}_{\mathrm{ARO},\,\mathrm{MD},\,e,\,j} &= (2 \cdot c^{N}_{e,9} \cdot 18 + c^{N}_{e,6}) \times (u_{j,\,6} - 18) && (109)\\
f^{N}_{\mathrm{EXT},\,e,\,2}(u_{j,\,6}) + s^{+}_{\mathrm{ARO},\,\mathrm{HI},\,e,\,j} - s^{-}_{\mathrm{ARO},\,\mathrm{HI},\,e,\,j} &= 0.0 && (110)
\end{aligned}
$$

but the olefins model is extrapolated using only one logical disjunction:

$$
\begin{aligned}
s^{+}_{\mathrm{OLE},\,\mathrm{LO},\,e,\,j} + s^{-}_{\mathrm{OLE},\,\mathrm{LO},\,e,\,j} &\leq U_{\mathrm{OLE},\,e,\,j} \cdot (1 - y^{N}_{\mathrm{OLE},\,19,\,j}) && (111)\\
s^{+}_{\mathrm{OLE},\,\mathrm{HI},\,e,\,j} + s^{-}_{\mathrm{OLE},\,\mathrm{HI},\,e,\,j} &\leq U_{\mathrm{OLE},\,e,\,j} \cdot y^{N}_{\mathrm{OLE},\,19,\,j} && (112)\\
f^{N}_{\mathrm{EXT},\,e,\,3}(u_{j,\,8}) + s^{+}_{\mathrm{OLE},\,\mathrm{LO},\,e,\,j} - s^{-}_{\mathrm{OLE},\,\mathrm{LO},\,e,\,j} &= 0.0 && (113)\\
f^{N}_{\mathrm{EXT},\,e,\,3}(u_{j,\,8}) + s^{+}_{\mathrm{OLE},\,\mathrm{HI},\,e,\,j} - s^{-}_{\mathrm{OLE},\,\mathrm{HI},\,e,\,j} &= (2 \cdot c^{N}_{e,10} \cdot 19 + c^{N}_{e,7}) \times (u_{j,\,8} - 19). && (114)
\end{aligned}
$$

Notice that these model extensions are first-order Taylor expansions of the primary component of the model, Eq. (98), at predetermined breakpoints.

3.2.3 Toxics Emissions Model

In addition to the 6 fuel components introduced in Eqs. (13) & (14), the toxics (TOX_j) Emissions Model is a function of 5 other fuel components presented in Table 1 (SUL^{T}_{j}, $\mathrm{E200}^{T}_{j}$, $\mathrm{E300}^{T}_{j}$, ARO^{T}_{j}, & OLE^{T}_{j}). The toxic air emissions model is ultimately the sum of six individual toxic air emissions (Eq. 124), the largest of which is usually the exhaust benzene emissions (Eq. 125). Therefore, Rhodes [1998] qualitatively suggests removing benzene precursors to meet EPA standards.

The eleven variables input to the TOX_j emissions equation are, in turn, functions of the outflow fuel qualities $u_{j,k}$. For some of the qualities, the two are identical:

$$\mathrm{SUL}^{T}_{j} = u_{j,\,2}\ \forall j, \quad \mathrm{E200}^{T}_{j} = u_{j,\,4}\ \forall j, \quad \mathrm{OLE}^{T}_{j} = u_{j,\,8}\ \forall j, \tag{115}$$

but the value $\mathrm{E300}^{T}_{j}$ is fixed to 95 vol% when $u_{j,\,5} \geq 95$:

$$
\begin{aligned}
\mathrm{E300}^{T}_{j} - 95 &\leq (u^{\mathrm{U}}_{j,\,5} - 95) \cdot y_{\mathrm{E300},\,j} && (116)\\
\mathrm{E300}^{T}_{j} - 95 &\geq (u^{\mathrm{L}}_{j,\,5} - 95) \cdot y_{\mathrm{E300},\,j} && (117)\\
\mathrm{E300}^{T}_{j} - u_{j,\,5} &\leq (u^{\mathrm{U}}_{j,\,5} - u^{\mathrm{L}}_{j,\,5}) \cdot (1 - y_{\mathrm{E300},\,j}) && (118)\\
\mathrm{E300}^{T}_{j} - u_{j,\,5} &\geq (u^{\mathrm{L}}_{j,\,5} - u^{\mathrm{U}}_{j,\,5}) \cdot (1 - y_{\mathrm{E300},\,j}), && (119)
\end{aligned}
$$

and the value ARO^{T}_{j} is fixed to 10 vol% when $u_{j,\,6} \leq 10$:

$$
\begin{aligned}
\mathrm{ARO}^{T}_{j} - 10 &\leq (u^{\mathrm{U}}_{j,\,6} - 10) \cdot (1 - y_{\mathrm{ARO},\,j}) && (120)\\
\mathrm{ARO}^{T}_{j} - 10 &\geq (u^{\mathrm{L}}_{j,\,6} - 10) \cdot (1 - y_{\mathrm{ARO},\,j}) && (121)\\
\mathrm{ARO}^{T}_{j} - u_{j,\,6} &\leq (u^{\mathrm{U}}_{j,\,6} - u^{\mathrm{L}}_{j,\,6}) \cdot y_{\mathrm{ARO},\,j} && (122)\\
\mathrm{ARO}^{T}_{j} - u_{j,\,6} &\geq (u^{\mathrm{L}}_{j,\,6} - u^{\mathrm{U}}_{j,\,6}) \cdot y_{\mathrm{ARO},\,j}. && (123)
\end{aligned}
$$

Toxics emission (TOX_j) is the sum of six components: exhaust benzene (BENZ_j), formaldehyde (FORM_j), acetaldehyde (ACET_j), 1,3-butadiene (BUTA_j), nonexhaust benzene (NEBENZ_j), and polycyclic organic matter (POM_j). The toxics emissions model is:

$$\begin{aligned}\mathrm{TOX}_j =& \mathrm{BENZ}_j + \mathrm{FORM}_j + \mathrm{ACET}_j + \mathrm{BUTA}_j + \\ & 10 \cdot \mathrm{NEBENZ}_j + \mathrm{POM}_j \quad \forall j\end{aligned} \tag{124}$$

The six components of the toxics emission model are presented in Eqs. (125) – (134). The model coefficients, which vary according to the time of year, region of the country, and emitter type, are presented in Tables 6 – 7. The models for BENZ_j, FORM_j, ACET_j, and BUTA_j involve an exponentiated linear term:

$$\mathrm{BENZ}_j = \sum_{e=1}^{2} \frac{\mathrm{BENZ}(b) \cdot w_e^T}{e^{b_e(b)}} \times \exp\{t_{BE,\,e,\,j}\}, \tag{125}$$

$$\mathrm{FORM}_j = \sum_{e=1}^{2} \frac{\mathrm{FORM}(b) \cdot w_e^T}{e^{f_e(b)}} \times \exp\{t_{F,\,e,\,j}\} \tag{126}$$

$$\mathrm{ACET}_j = \sum_{e=1}^{2} \frac{\mathrm{ACET}(b) \cdot w_e^T}{e^{a_e(b)}} \times \exp\{t_{A,\,e,\,j}\} \tag{127}$$

$$\mathrm{BUTA}_j = \sum_{e=1}^{2} \frac{\mathrm{BUTA}(b) \cdot w_e^T}{e^{d_e(b)}} \times \exp\{t_{BU,\,e,\,j}\} \tag{128}$$

where the linear term is a function of the fuel indices:

$$t_{BE,\,e,\,j} = c^{BE}_{e,\,1}\mathrm{OXY}_j + c^{BE}_{e,\,2}\mathrm{SUL}^T_j + c^{BE}_{e,\,3}\mathrm{E300}^T_j + c^{BE}_{e,\,4}\mathrm{ARO}^T_j + c^{BE}_{e,\,5}\mathrm{BEN}_j \tag{129}$$

$$t_{F,\,e,\,j} = c^{F}_{e,\,1}\mathrm{E300}^T_j + c^{F}_{e,\,2}\mathrm{ARO}^T_j + c^{F}_{e,\,3}\mathrm{OLE}^T_j + c^{F}_{e,\,4}\mathrm{MTB}_j \tag{130}$$

$$\begin{aligned}t_{A,\,e,\,j} =& c^{A}_{e,\,1}\mathrm{SUL}^T_j + c^{A}_{e,\,2}\mathrm{RVP}_j + c^{A}_{e,\,3}\mathrm{E300}^T_j + c^{A}_{e,\,4}\mathrm{ARO}^T_j + c^{A}_{e,\,5}\mathrm{MTB}_j \\ & + c^{A}_{e,\,6}\mathrm{ETB}_j + c^{A}_{e,\,7}\mathrm{ETH}_j\end{aligned} \tag{131}$$

$$\begin{aligned}t_{BU,\,e,\,j} =& c^{BU}_{e,\,1}\mathrm{OXY}_j + c^{BU}_{e,\,2}\mathrm{SUL}^T_j + c^{BU}_{e,\,3}\mathrm{E200}^T_j + c^{BU}_{e,\,4}\mathrm{E300}^T_j + c^{BU}_{e,\,5}\mathrm{ARO}^T_j \\ & + c^{BU}_{e,\,6}\mathrm{OLE}^T_j .\end{aligned} \tag{132}$$

POM_j is a linear multiple of VOCE_j, which was previously defined in Eq. (48):

$$\mathrm{POM}_j = \alpha^{POM} \cdot \mathrm{VOCE}_j \tag{133}$$

and NEBENZ_j is a function of the RVP, benzene, and MBTE fuel indices:

$$\mathrm{NEBENZ}_j = \begin{cases} 0.0 & \text{Winter} \\ \begin{cases} \alpha^{NB}_1 \cdot \mathrm{BEN}_j + \alpha^{NB}_2 \cdot \mathrm{RVP}_j \cdot \mathrm{BEN}_j + \\ \alpha^{NB}_3 \cdot \mathrm{BEN}_j \cdot \mathrm{MTB}_j + \alpha^{NB}_4 \cdot \mathrm{RVP}^2_j \cdot \mathrm{BEN}_j + \\ \alpha^{NB}_5 \cdot \mathrm{RVP}_j \cdot \mathrm{BEN}_j \cdot \mathrm{MTB}_j + \alpha^{NB}_6 \cdot \mathrm{RVP}^3_j \cdot \mathrm{BEN}_j + \\ \alpha^{NB}_7 \cdot \mathrm{RVP}^2_j \cdot \mathrm{BEN}_j \cdot \mathrm{MTB}_j \end{cases} & \text{Summer} \end{cases} \forall\, j. \tag{134}$$

3.3 Equations for RVP Blending

Assuming that i denotes a set of streams, x_i is the flow into a node, RVP_i is the RVP of each entering stream, and RVP is the RVP value in the node, the *Chevron Method* blending rule for RVP is [Visweswaran, 2009]:

$$(\sum_i x_i) \cdot \mathrm{RVP}^{1.25} = \sum_i x_i \cdot \mathrm{RVP}_i^{1.25}. \tag{135}$$

Using the the notation developed in this document, the RVP value of each pool could be calculated as:

$$\mathrm{RVP}_{\mathrm{pool},\,l}^{1.25} \cdot \left(\sum_{i:(i,l)\in T_X} q_{i,l} \right) = \sum_{i:(i,l)\in T_X} C_{i,3}^{1.25} \cdot q_{i,l} \quad \forall\, l \tag{136}$$

and the RVP of each product as:

$$u_{j,3}^{1.25} \cdot (of_j) = \sum_{l:(l,j)\in T_Y} \mathrm{RVP}_{\mathrm{pool},\,l}^{1.25} \cdot y_{l,j} + \sum_{i:(i,j)\in T_Z} C_{i,3}^{1.25} \cdot z_{i,j} \quad \forall\, j \tag{137}$$

but, because $\sum_{i:(i,l)\in T_X} q_{i,l} = 1\ \forall\, l$, it would be much simpler to substitute Eq. (136) into Eq. (137):

$$u_{j,3}^{1.25} \cdot (of_j) = \sum_{i:(i,l)\in T_X} \sum_{l:(l,j)\in T_Y} C_{i,3}^{1.25} \cdot q_{i,l} \cdot y_{l,j} + \sum_{i:(i,j)\in T_Z} C_{i,3}^{1.25} \cdot z_{i,j} \quad \forall\, j. \tag{138}$$

Eq. (138) is equivalent to Eq. (5) for $k = 3$. Better yet, we define auxiliary variable:

$$\widehat{u_{j,3}} = u_{j,3}^{1.25} \tag{139}$$

and get:

$$\widehat{u_{j,3}} \cdot (of_j) = \sum_{i:(i,l)\in T_X} \sum_{l:(l,j)\in T_Y} C_{i,3}^{1.25} \cdot q_{i,l} \cdot y_{l,j} + \sum_{i:(i,j)\in T_Z} C_{i,3}^{1.25} \cdot z_{i,j} \quad \forall\, j. \tag{140}$$

In sum, the MINLP model for the extended pooling problem consists of Eqs. (1) – (134), (139) & (140). The model appends the EPA Complex Emissions Model and the nonlinear blending of RVP to a standard pooling problem framework.

4 MILP Relaxation

Section 3 introduced the full MINLP model for the extended pooling problem. To solve the extended pooling problem to global optimality, we now present a linear relaxation of the nonlinear equations in the MINLP formulation. The nonlinear model equations in the extended pooling problem and their type are:

- The bilinear terms of the type $q_{i,l} \cdot y_{l,j}$ in Eqs. (1), (2), (5), (7) & (140) are underestimated using piecewise-linear relaxation [Gounaris et al., 2009]. The underestimators are formulated according to Section 4.1.
- The bilinear terms $u_{j,k} \cdot (of_j)$ in Eq. (5) are relaxed using termwise envelopes as described in Section 4.2 [McCormick, 1976].
- The convex nonlinear terms in Eqs. (48), (97), (125) – (128) & (139) are underestimated using outer approximation. The relaxation of these terms, which have the form $\exp\{t_{M,\,e,\,j}\}$ where $M \in \{V,\, N,\, BE,\, F,\, A,\, BU\}$, is described in Section 4.3.

- Eqs. (72) & (139) are underestimated using outer approximation on the convex side of the equality and a secant line on the concave side, but the nonlinear coefficient α_3^V in Eq. (72) is relatively small (see Table 7) and $\widehat{u_{j,3}}$ never deviates far from $u_{j,3}$ in Eq. (139). Because Eqs. (72) & (139) have relatively low curvature, we relaxed them using 2-3 supporting hyperplanes and a single secant line rather than employing the scheme described in Section 4.3.

- As described in Section 4.4, Eq. (134) is underestimated using the edge-concave paradigm [Meyer and Floudas, 2005].

- The remaining nonconvexities in Eqs. (48) – (49), (71), (97) – (98) & (140) are underestimated using recursive application of bilinear envelopes as described by Maranas and Floudas [1995] and Ryoo and Sahinidis [2001].

4.1 Underestimating $q_{i,l} \cdot y_{l,j}$ Bilinear Terms

The success that Meyer and Floudas [2006] and Karuppiah and Grossmann [2006] had in tightly relaxing large-scale bilinear programs led Wicaksono and Karimi [2008] and Gounaris et al. [2009] to investigate a variety of piecewise-linear formulations which underestimate bilinear terms. In underestimating the bilinear terms of type $q_{i,l} \cdot y_{l,j}$, we construct piecewise-linear underestimators for the bilinear terms. Specifically, we use the representation denoted **nf4r** based on our recent extensive computational study [Gounaris et al., 2009]. The bilinear terms $q_{i,l} \cdot y_{l,j}$ are replaced with a placeholder variable:

$$q_{i,l}^{\mathrm{L}} \cdot y_{l,j}^{\mathrm{L}} \leq w_{i,l,j}^{q,y} \leq q_{i,l}^{\mathrm{U}} \cdot y_{l,j}^{\mathrm{U}}. \tag{141}$$

To construct piecewise-linear bilinear underestimators, each flow proportion $q_{i,l}$ is *ab initio* partitioned into N segments:

$$q_{i,l}(n) = q_{i,l}^{\mathrm{L}} + \frac{n}{N} \cdot (q_{i,l}^{\mathrm{U}} - q_{i,l}^{\mathrm{L}}) \quad \forall\, i,\, l,\, n = 0, \cdots, N \tag{142}$$

and an SOS1 variable $\lambda_{i,l}(n)$ is introduced to activate one and only one domain segment:

$$\lambda_{i,l}(n) = \begin{cases} 1 & \text{if } q_{i,l}(n-1) \leq q_{i,l} \leq q_{i,l}(n) \\ 0 & \text{else} \end{cases} \quad \forall\, i,\, l,\, n = 1, \cdots, N \tag{143}$$

$$\sum_{n=1}^{N} q_{i,l}(n-1) \cdot \lambda_{i,l}(n) \leq q_{i,l} \leq \sum_{n=1}^{N} q_{i,l}(n) \cdot \lambda_{i,l}(n) \quad \forall\, i,\, l \tag{144}$$

$$\sum_{n=1}^{N} \lambda_{i,l}(n) = 1 \quad \forall\, i,\, l. \tag{145}$$

Continuous variable $\Delta y_{i,l,j}(n),\, n = 1, \cdots, N$ is a place holder for the flow rate $y_{l,j}$:

$$y_{l,j} = y_{l,j}^{\mathrm{L}} + \sum_{n=1}^{N} \Delta y_{i,l,j}(n) \quad \forall\, i \tag{146}$$

$$0 \leq \Delta y_{i,l,j}(n) \leq (y_{l,j}^{\mathrm{U}} - y_{l,j}^{\mathrm{L}}) \cdot \lambda_{i,l}(n) \quad \forall\, i,\, n = 1, \cdots, N. \tag{147}$$

The final relaxation of the bilinear terms $\forall\, i,\, l,\, j$ is:

$$
\begin{aligned}
w^{q,y}_{i,l,j} &\geq y^{\mathrm{L}}_{l,j} \cdot q_{i,l} + \sum_{n=1}^{N} q_{i,l}(n-1) \cdot \Delta y_{i,l,j}(n) \\
w^{q,y}_{i,l,j} &\geq y^{\mathrm{U}}_{l,j} \cdot q_{i,l} + \sum_{n=1}^{N} \left\{ q_{i,l}(n) \cdot \left(\Delta y_{i,l,j}(n) - (y^{\mathrm{U}}_{l,j} - y^{\mathrm{L}}_{l,j}) \cdot \lambda_{i,l}(n) \right) \right\} \\
w^{q,y}_{i,l,j} &\leq y^{\mathrm{L}}_{l,j} \cdot q_{i,l} + \sum_{n=1}^{N} \left\{ q_{i,l}(n-1) \cdot \left(\Delta y_{i,l,j}(n) - (y^{\mathrm{U}}_{l,j} - y^{\mathrm{L}}_{l,j}) \cdot \lambda_{i,l}(n) \right) \right\} \\
w^{q,y}_{i,l,j} &\leq y^{\mathrm{U}}_{l,j} \cdot q_{i,l} + \sum_{n=1}^{N} q_{i,l}(n) \cdot \Delta y_{i,l,j}(n).
\end{aligned}
\qquad \forall\, i,\, l,\, j \tag{148}
$$

4.2 Underestimating $u_{j,k} \cdot (of_j)$ Bilinear Terms

Observing in Eq. (5) that the $u_{j,k} \cdot (of_j)$ bilinear terms are closely related to the $q_{i,l} \cdot y_{l,j}$ terms, we found that adding piecewise-linear relaxations to the $u_{j,k} \cdot (of_j)$ terms significantly increased the problem solution time without notable change in the objective function. Therefore, we chose not to construct a piecewise-linear relaxation of the $(u_{j,k}) \cdot (of_j)$ terms. Instead, we replaced each of the $(u_{j,k}) \cdot (of_j)$ bilinear terms with the continuous variable $w^{u,of}_{j,k}$ and underestimated the terms using the convex envelope [McCormick, 1976]:

$$
\begin{aligned}
w^{u,of}_{j,k} &\geq u^{\mathrm{L}}_{j,k} \cdot of_j + u_{j,k} \cdot of^{\mathrm{L}}_j - u^{\mathrm{L}}_{j,k} \cdot of^{\mathrm{L}}_j \\
w^{u,of}_{j,k} &\geq u^{\mathrm{U}}_{j,k} \cdot of_j + u_{j,k} \cdot of^{\mathrm{U}}_j - u^{\mathrm{U}}_{j,k} \cdot of^{\mathrm{U}}_j \\
w^{u,of}_{j,k} &\leq u^{\mathrm{L}}_{j,k} \cdot of_j + u_{j,k} \cdot of^{\mathrm{U}}_j - u^{\mathrm{L}}_{j,k} \cdot of^{\mathrm{U}}_j \\
w^{u,of}_{j,k} &\leq u^{\mathrm{U}}_{j,k} \cdot of_j + u_{j,k} \cdot of^{\mathrm{L}}_j - u^{\mathrm{U}}_{j,k} \cdot of^{\mathrm{L}}_j
\end{aligned}
\qquad \forall\, j,\, k \tag{149}
$$

4.3 Outer Approximation of Exponential Terms

To relax monotonic convex function $f(x)$, we can construct $(N^{\mathrm{OA}} + 1)$ supporting hyperplanes by taking $(N^{\mathrm{OA}} + 1)$ domain points:

$$
x(n^{\mathrm{OA}}) = f^{-1}\left(f^{\mathrm{L}} + \frac{n^{\mathrm{OA}}}{N^{\mathrm{OA}}} \cdot \left(f^{\mathrm{U}} - f^{\mathrm{L}} \right) \right), \quad n^{\mathrm{OA}} = 0, \ldots, N^{\mathrm{OA}} \tag{150}
$$

and generating $(N^{\mathrm{OA}} + 1)$ tangent planes:

$$
f(x) \geq \frac{\partial f(n^{\mathrm{OA}},\, x)}{\partial x} \cdot \left(x - x(n^{\mathrm{OA}}) \right) + f\left(x(n^{\mathrm{OA}}) \right), \quad n^{\mathrm{OA}} = 0, \ldots, N^{\mathrm{OA}} \tag{151}
$$

We also include a secant tie line between the function end points:

$$
f(x) \leq f^{\mathrm{L}} + (x - x^{\mathrm{L}}) \cdot \frac{f^{\mathrm{U}} - f^{\mathrm{L}}}{x^{\mathrm{U}} - x^{\mathrm{L}}} \tag{152}
$$

Although this hyperplane generation method of underestimating would not work for functions such as $f(x) = x^2,\, x \in [-1,\, 1]$ because it assumes that $f : \mathbb{R} \mapsto \mathbb{R}$ is monotonic, it is perfect for convex functional forms such as those found in Eqs. (48), (97), (125) – (128) & (139). The convex terms in these functions, $\exp\{t_{M,\,e,\,j}\}$ where $M \in \{V,\, N,\, BE,\, F,\, A,\, BU\}$ represents the different emissions types, can be underestimated using the following scheme:

$$
t_{M,\,e,\,j}(n^{\mathrm{OA}}) = \ln\left[e^{t^{\mathrm{L}}_{M,\,e,\,j}} + \frac{n^{\mathrm{OA}}}{N^{\mathrm{OA}}} \cdot \left(e^{t^{\mathrm{U}}_{M,\,e,\,j}} - e^{t^{\mathrm{L}}_{M,\,e,\,j}} \right) \right], \quad n^{\mathrm{OA}} = 0, \ldots, N^{\mathrm{OA}} \tag{153}
$$

$$e^{t_{M,\,e,\,j}} \geq e^{t_{M,\,e,\,j}(n^{\mathrm{OA}})} \cdot \left[t_{M,\,e,\,j} - t_{M,\,e,\,j}(n^{\mathrm{OA}}) + 1\right], \quad n^{\mathrm{OA}} = 0, \ldots, N^{\mathrm{OA}} \tag{154}$$

$$e^{t_{M,\,e,\,j}} \leq e^{t_{M,\,e,\,j}(n^0)} + (t_{M,\,e,\,j} - t_{M,\,e,\,j}(n^0)) \cdot \frac{e^{t_{M,\,e,\,j}(N^{\mathrm{OA}})} - e^{t_{M,\,e,\,j}(n^0)}}{t_{M,\,e,\,j}(N^{\mathrm{OA}}) - t_{M,\,e,\,j}(n^0)} \tag{155}$$

Eqs. (153) & (154) are especially well-suited for exponential equations because the grid points evenly partition the function range rather than the domain.

4.4 Underestimating Eq. (134) with Edge-Concave Properties

The properties of edge-concave functions, that is, functions admitting a vertex polyhedral envelope, were explored by Tardella [1988/89, 2003, 2008]. Meyer and Floudas [2005] developed an easy-to-implement algorithm which, given an edge-concave function of three dimensions or fewer, computes the facets of the convex envelope. Although Eq. (134), representing nonexhaust benzene emissions, is not edge-concave, we use the paradigm of edge-concavity to efficiently generate a tight lower bound on NEBENZ_j. A function on a box is edge-concave if and only if it is component-wise concave [Tardella, 2003], that is:

$$\begin{aligned} \frac{\partial^2 \mathrm{NEBENZ}_j}{\partial \mathrm{RVP}_j^2} &= 2\alpha_4^{NB} \cdot \mathrm{BEN}_j + 6\alpha_6^{NB} \cdot \mathrm{RVP}_j \cdot \mathrm{BEN}_j \\ &\quad + 2\alpha_7^{NB} \cdot \mathrm{MTB}_j \cdot \mathrm{BEN}_j \leq 0 \quad \forall\, j \end{aligned} \tag{156}$$

$$\frac{\partial^2 \mathrm{NEBENZ}_j}{\partial \mathrm{MTB}_j^2} = 0 \leq 0 \quad \forall\, j \tag{157}$$

$$\frac{\partial^2 \mathrm{NEBENZ}_j}{\partial \mathrm{BEN}_j^2} = 0 \leq 0 \quad \forall\, j \tag{158}$$

Since Eq. (157) and (158) are always true, the remaining task is to see when Eq. (156) is negative. Because $\frac{\partial^2 \mathrm{NEBENZ}_j}{\partial \mathrm{RVP}_j^2} \not\leq 0$, Eq. (134) is not edge-concave. However:

$$\mathrm{NEBENZ}_j - \alpha_4' \cdot \mathrm{RVP}_j^2 \cdot \mathrm{BEN}_j \tag{159}$$

is edge-concave when:

$$\alpha_4' = \alpha_4^{NB} + 3 \cdot \alpha_6^{NB} \cdot \mathrm{RVP}_j^{\mathrm{L}} + \alpha_7^{NB} \cdot \mathrm{MTB}_j^{\mathrm{L}}, \tag{160}$$

so we underestimate Expression (159) using the facets of the convex envelope of the edge concave function, and the remainder ($\alpha_4' \cdot \mathrm{RVP}_j^2 \cdot \mathrm{BEN}_j$) using recursive arithmetic techniques [Maranas and Floudas, 1995, Ryoo and Sahinidis, 2001]. The improvement we obtained using the properties of edge-concave functions gave us a 22% advantage over standard recursive arithmetic. The edge-concave relaxation of Eq. (134) has a lower bound of -10.61 and -9.01 (in EPA-defined Regions 1 and 2, respectively), while the looser recursive arithmetic relaxations are -13.56 and -11.49. We generated these edge-concave and recursive arithmetic relaxations using the fuel quality bounds listed in Table 1.

We cannot explicitly state the twelve linear equations composing the convex envelope of Eq. (159) because they change as bounds are tightened within a global optimization algorithm. Refer to Meyer and Floudas [2005] for the explicit method to determine the convex envelope of a three-dimension edge-concave function. However, a typical example of the lower bounding facets for Eq. (159) is:

$$\text{Eq. (159)} \geq \begin{cases} 0.00 + 0.00 \cdot \text{RVP}_j + 0.00 \cdot \text{MTB}_j - 2.08 \cdot \text{BEN}_j \\ 0.00 + 0.00 \cdot \text{RVP}_j + 0.00 \cdot \text{MTB}_j - 2.08 \cdot \text{BEN}_j \\ 0.00 + 0.00 \cdot \text{RVP}_j - 0.06 \cdot \text{MTB}_j - 1.93 \cdot \text{BEN}_j \\ 4.13 - 0.64 \cdot \text{RVP}_j - 0.06 \cdot \text{MTB}_j - 0.77 \cdot \text{BEN}_j \\ 4.48 - 0.70 \cdot \text{RVP}_j - 0.01 \cdot \text{MTB}_j - 0.77 \cdot \text{BEN}_j \\ 4.48 - 0.70 \cdot \text{RVP}_j + 0.00 \cdot \text{MTB}_j - 0.82 \cdot \text{BEN}_j \end{cases} \tag{161}$$

5 Global Optimization Algorithm

To globally optimize the three case studies presented in Section 6, we implemented a branch-and-bound algorithm in C++. At each node in the branch-and-bound tree, we minimized a linear relaxation of the node using the MILP solver CPLEX [ILOG, 2007] and initialized a local solve using the NLP solver MINOS [Murtagh et al., 2004] through a system call to the modeling language GAMS [Brooke et al., 2007]. We choose to implement the linear relaxation in C++ because solving the lower bounding nodes to optimality was the consistent bottleneck in the solution process, but we did not implement the upper bounding portion in a solver based in C++ because the local solves converged quickly.

Specifically, we used the following algorithm to optimize the test cases presented in the following section. We terminate at an optimality gap of $\epsilon = 0.005$:

1 To generate a good initial upper bound, we solve the lower bounding problem several times ($N = 16$ and $N^{\text{OA}} = \{1,\ 3,\ 5,\ 7,\ 9,\ 11\}$) and use the resulting solutions to initialize local solves in MINOS [Murtagh et al., 2004]. MINOS is an NLP solver, so we fix the binary variables inherent in the EPA Complex Emissions Model. Fixing the binary variables implicitly bounds the $u_{j,\,k}$ quality variables, so we appropriately reduce the variable bounds through logical inference before performing each local solve. For example, if $y^{V}_{\text{E200},\,33,\,j}$ is fixed to 0 and $y^{V}_{\text{E200},\,65.52,\,j}$ to 1, then we satisfy Eqs. (21) – (24) & (34) in the local NLP solve by reducing the $u_{j,k}$ bounds to $[33,\ 65.52]$.

2 After multistarting to obtain a good upper bound, we initialize the branch and bound tree. Before solving a node, we use optimality-based tightening to contract the bounds of the $q_{i,\,l}$, $y_{l,\,j}$, and $u_{j,\,k}$ variables [Floudas, 2000, Tawarmalani and Sahinidis, 2002]. To reduce the time needed to solve the bound contraction problems, we set the bilinear partitioning level to $N = 1$ (equivalent to termwise McCormick envelopes) and the outer approximation partitioning to $N^{\text{OA}} = 8$. Then the node itself is solved using a finer partitioning scheme ($N = 16$ and $N^{\text{OA}} = 8$). These values of N and N^{OA} were obtained from experience with the small and mid-size test cases presented in Table 3, but they still perform well for the largest test case.

3 The node relaxation found in Step 2 is used to initialize a local solve. The binary variables are fixed as in Step 1 and the $u_{j,\,k}$ variables are contracted through logical inference. If the local solve yields a result less than the current upper bound, it becomes the new upper bound. Then nodes greater than or sufficiently close to the current upper bound ($\text{LB}^{\text{NODE}} \cdot (1 + \epsilon) \geq \text{UB}$) are fathomed.

4 If the node relaxed in Step 2 survives Step 3, it begets two child nodes by splitting the of_j flow variable satisfying:

$$\arg\max_{j} \sum_{k} |w^{u,\,of}_{j,\,k} - u_{j,\,k} \cdot of_j|$$

in half ($\frac{of^{\text{L}}_j + of^{\text{U}}_j}{2}$) and adding the two nodes to the tree. We chose to branch specifically on the (of_j) variables because they participate in non-partitioned bilinear terms $u_{j,\,k} \cdot (of_j)$ and there are fewer

(of_j) than $u_{j,k}$ variables. The relaxation constructed in Step 2 is not preserved from parent to child, so child nodes may have weaker lower bounds than parent nodes. See the work of Gounaris et al. [2009] and Karuppiah and Grossmann [2006] for a more detailed explanation regarding the non-monotonic trends of piecewise-linear underestimation.

5 After updating the tree in Step 3, the node with the least lower bound for further branching (Step 2). The algorithm terminates when there are no nodes remaining in the tree.

6 Case Studies

Table 3 displays the complexity of the three test cases illustrated in Figures 3 – 5. The parameters relevant to the test cases are recorded in Appendices B – D. The feed stocks represent the characteristics of intermediate stocks leaving the processing units of a refinery. Each feed stock is estimated to have a market value based on its composition. Transportation considerations normally require additives to be mixed into gasoline at a distribution station (*e.g.*, ethanol is rarely transported by pipeline with other gasoline components), but the test cases described in this study simplify the problem by assuming that additives are blended into final products at the refinery.

Table 3: Overview of the Three Case Studies

	Topology							# Variables		# Nonlinear Terms	
	I	J	K	L	$\|T_X\|$	$\|T_Y\|$	$\|T_Z\|$	Contin	Binary	Bilinear Only	All
Case 1	7	2	11	1	4	2	8	214	30	62	108
Case 2	14	3	11	3	10	9	12	331	45	111	180
Case 3	14	10	11	5	14	50	40	1104	150	410	640

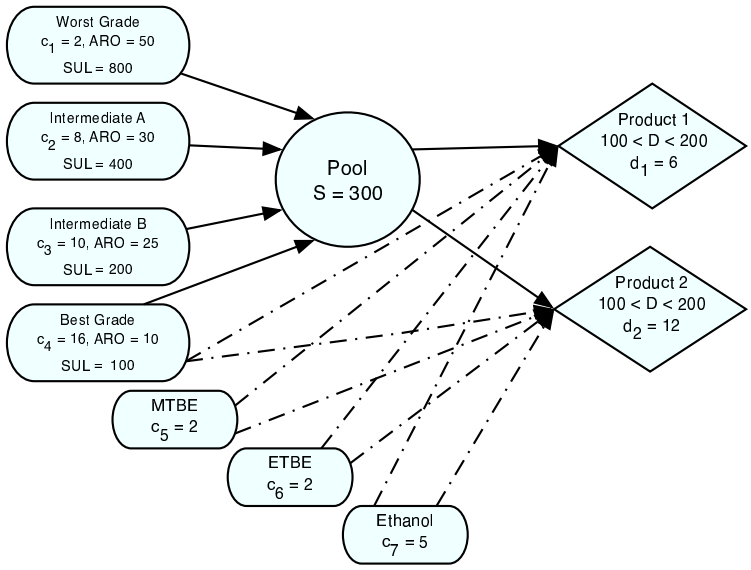


Figure 3: Test Case 1 [Gounaris and Floudas, 2007]

The emissions limits for the three test cases are listed in Table 4. In the three test cases, we assumed that the emissions limits are identical for all products, but the complexity of the problems would not change if each product was given a different target emissions level. Further, each of the three test cases assumes

that the refineries are blending products during the summer because considering the summer allows us to approach the most challenging version of the EPA Complex Emissions Model (the winter model eliminates the non-exhaust volatile organic and non-exhaust benzene equations and fixes RVP). The difference between Region 1 and 2 is only a change in parameters (see Appendix A), so our test cases consider both the Region 1 and 2 cases.

Table 4: Emissions Limits of the Three Case Studies

	$\mathrm{VOC}_{j,\mathrm{MAX}}$ $(\frac{\mathrm{mg}}{\mathrm{mile}})$	$\mathrm{NOX}_{j,\mathrm{MAX}}$ $(\frac{\mathrm{mg}}{\mathrm{mile}})$	$\mathrm{TOX}_{j,\mathrm{MAX}}$ $(\frac{\mathrm{mg}}{\mathrm{mile}})$
Case 1	1200	1300	90
Case 2	1700	1400	95
Case 3	1600	1300	95

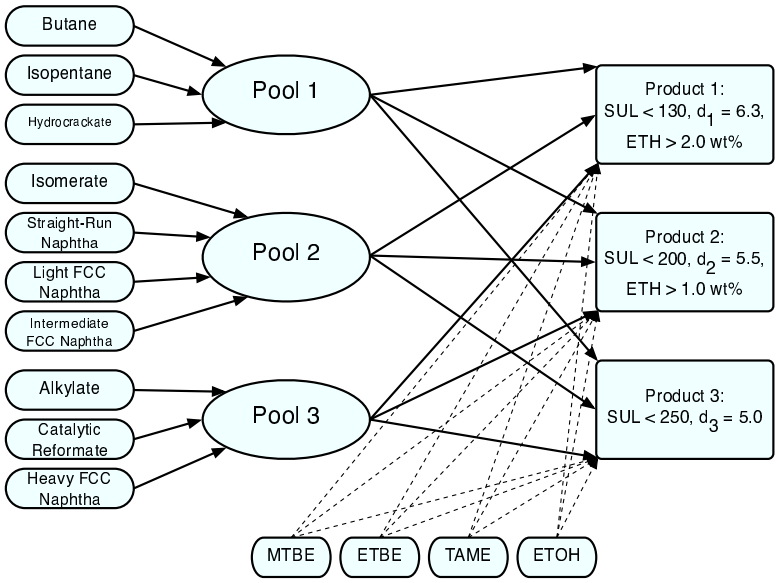


Figure 4: Test Case 2

7 Computational Results

The test cases were solved on a Linux workstation with an Intel Core 2 Quad processor containing four 2.83 GHz cores. The optimization process was run serially within using CPLEX Version 11.0 [ILOG, 2007] for the node relaxations and MINOS Version 5.51 [Murtagh et al., 2004] for the local upper bounding solves. Table 5 lists the results of the optimization runs. There are few nodes in the branch-and-bound trees because the underestimators described in Section 4 tightly relax the problem. For example, Figure 6 depicts the branch-and-bound tree generated by the mid-size test case. Each node in Figure 6 corresponds to the MILP relaxation described in Section 4 on the relevant domain. Figures 7 – 9 list some of the key flow rates in the best-found upper bound for each topology in Region 1.

In the second and third test cases, where the costs of the four additives are substantially higher than the costs of the feed stocks exiting the refinery (see Table 18), the optimization algorithm effectively serves to minimize unnecessary purchase of additives (MTBE, ETBE, TAME, and ethanol). According to the EPA Complex Emissions model, the four additives reduce all three types of emissions, so a feasible but sub-optimal solution could meet EPA standards by giving away valuable additives.

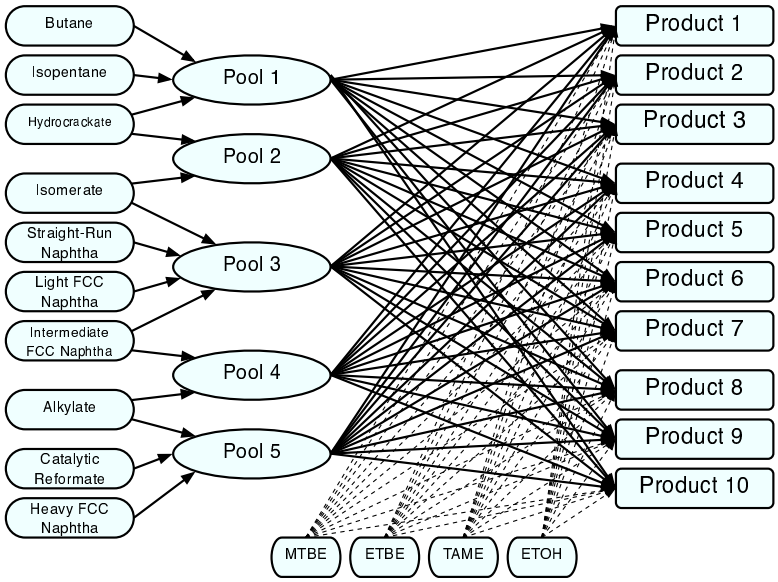


Figure 5: Test Case 3

Table 5: Branch & Bound Optimization of the Three Test Cases

	Region	Root Node Relaxation	# Nodes	Termination			CPU Time (s)
				Lower Bnd	Upper Bnd	% Gap	
Case 1	1	-4.731×10^2	7	-2.808×10^2	-2.794×10^2	0.5	4.6
	2	-4.731×10^2	7	-2.818×10^2	-2.808×10^2	0.5	4.5
Case 2	1	-4.598×10^3	9	-4.587×10^3	-4.567×10^3	0.4	292
	2	-4.598×10^3	9	-4.588×10^3	-4.567×10^3	0.5	263
Case 3	1	-1.500×10^4	11	-1.498×10^4	-1.490×10^4	0.5	5274
	2	-1.500×10^4	1	-1.500×10^4	-1.496×10^4	0.2	411

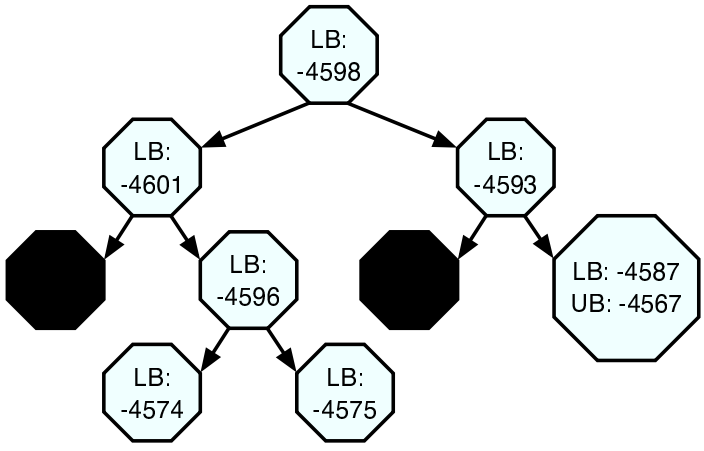


Figure 6: Branch-and-bound tree for Test Case 2. Cuts are dropped from parent to child, possibly generating looser lower bounds. The blackened nodes have been fathomed. The algorithm terminates at a 0.5% gap in 487 CPU s.

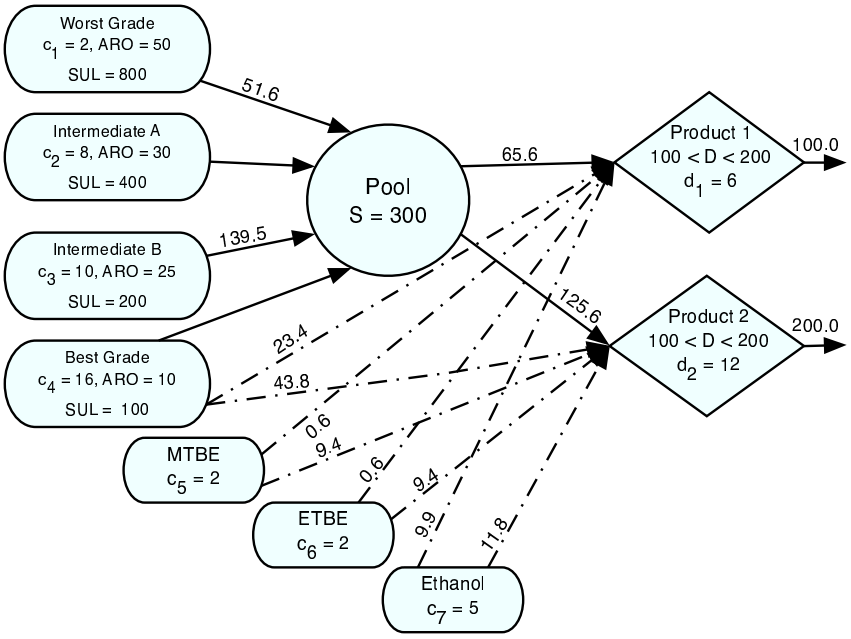


Figure 7: Best Upper Bound found for Test Case 1 [Gounaris and Floudas, 2007]

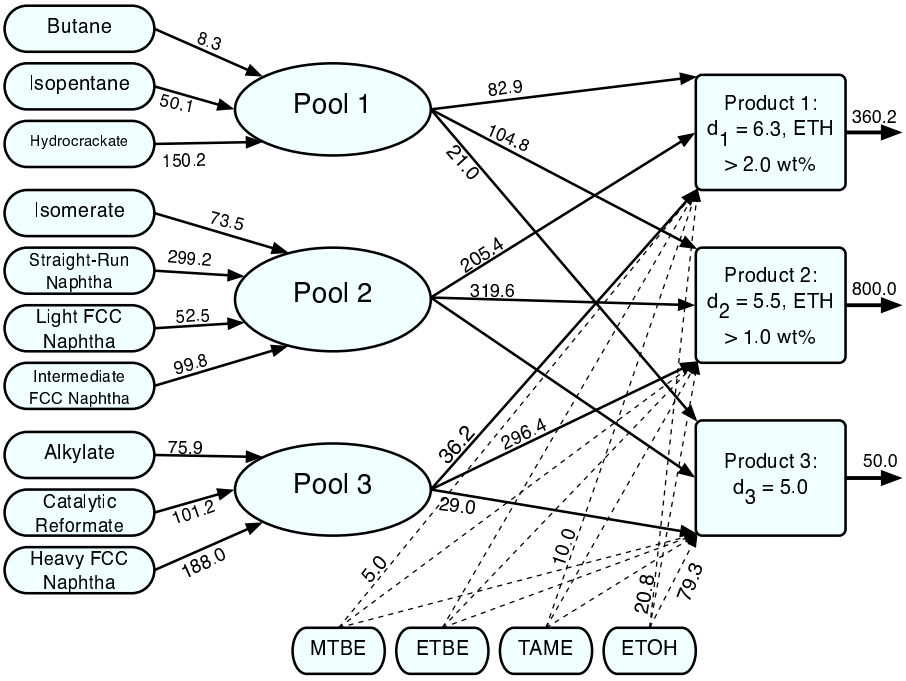


Figure 8: Best Upper Bound found for Test Case 2

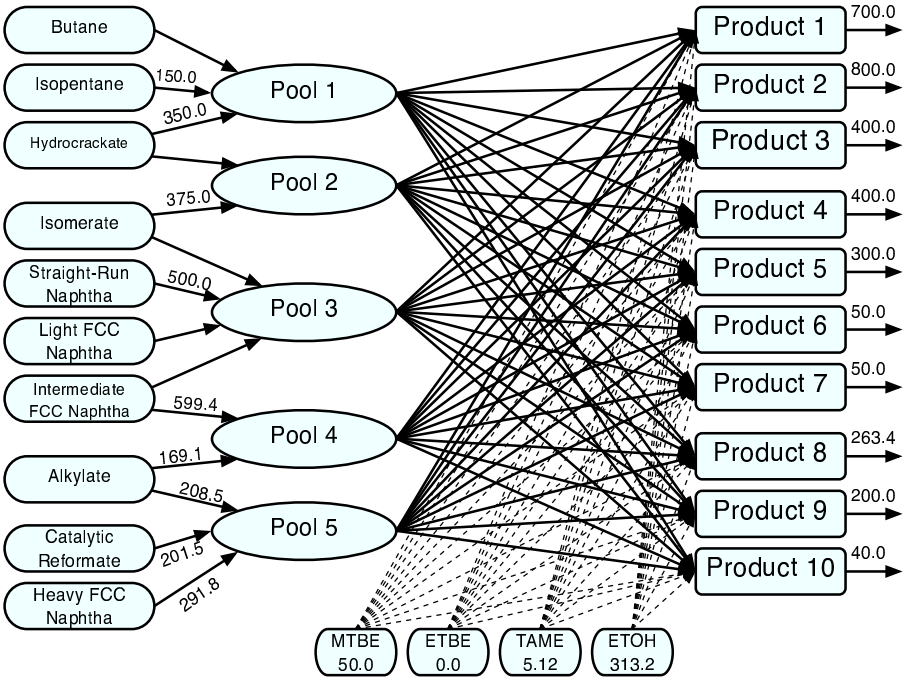


Figure 9: Best Upper Bound found for Test Case 3

8 Conclusion

We have introduced a MINLP mathematical model for the extended pooling problem, an augmentation to the pooling problem that appends the EPA Complex Emission Model [40CFR80.45, 2007] to the constraint set. This extension is valuable because it integrates boutique fuel blending into the pooling problem, which addresses the temporary storage and subsequent blending of intermediate feed stocks. We have proposed a MILP relaxation of extended pooling problem, integrated the relaxation into a branch-and-bound global optimization algorithm, and demonstrated the success of the method for three test cases including a large-scale case study which features 1104 continuous variables, 150 binary variables, and 640 nonlinear terms. The nonconvexities in the large-scale case study include 410 bilinear terms, 40 polynomial terms, and 10 terms raised to a fractional power.

Acknowledgments: The authors gratefully acknowledge support from the National Science Foundation (CBET – 0827907). R.M. is further thankful for her NSF Graduate Research Fellowship.

References

40CFR80.41. Code of Federal Regulations: Standards and requirements for compliance, 2008. http://www.gpoaccess.gov/cfr/retrieve.html.

40CFR80.45. Code of Federal Regulations: Complex emissions model, 2007. http://www.gpoaccess.gov/cfr/retrieve.html.

N. Adhya, M. Tawarmalani, and N. V. Sahinidis. A Lagrangian approach to the pooling problem. *Ind. Eng. Chem. Res.*, 38(5):1965–1972, 1999.

F. A. Al-Khayyal and J. E. Falk. Jointly constrained biconvex programming. *Math. Oper. Res.*, 8(2):273–286, 1983.

H. Almutairi and S. Elhedhli. A new Lagrangean approach to the pooling problem. *J. Global Optim.*, 2009. Forthcoming.

K. M. Anstreicher. Semidefinite programming versus the reformulation-linearization technique for nonconvex quadratically constrained quadratic programming. *J. Global Optim.*, 43(2-3):471–484, 2009.

C. Audet, P. Hansen, B. Jaumard, and G. Savard. A branch and cut algorithm for nonconvex quadratically constrained quadratic programming. *Math. Program.*, 87(1):131 – 152, 2000.

C. Audet, J. Brimberg, P. Hansen, S. Le Digabel, and N. Mladenovic. Pooling problem: Alternate formulations and solution methods. *Manage. Sci.*, 50(6):761 – 776, 2004.

X. Bao, N. V. Sahinidis, and M. Tawarmalani. Multiterm polyhedral relaxations for nonconvex, quadratically-constrained quadratic programs. *Optimization Methods and Software*, 24(4 - 5):485–504, 2009.

A. Ben-Tal, G. Eiger, and V. Gershovitz. Global minimization by reducing the duality gap. *Math. Program.*, 63:193–212, 1994.

M. L. Bergamini, I. Grossmann, N. Scenna, and P. Aguirre. An improved piecewise outer-approximation algorithm for the global optimization of MINLP models involving concave and bilinear terms. *Comput. Chem. Eng.*, 32(3):477 – 493, 2008.

A. Brooke, D. Kendrick, and A. Meeraus. General algebraic modeling language (GAMS). `http://www.gams.com/`, 2007. Version 22.6.

A. R. Ciric and C. A. Floudas. A retrofit approach for heat exchanger networks. *Comput. Chem. Eng.*, 13 (6):703 – 715, 1989.

EPA. Reformulated gas. http://www.epa.gov/otaq/rfg.htm, 2008. Accessed 08/24/09.

C. A. Floudas. *Nonlinear and Mixed-Integer Optimization: Fundamentals and Applications*. Oxford University Press, New York, NY, 1995.

C. A. Floudas. *Deterministic Global Optimization : Theory, Methods and Applications*. Nonconvex Optimization and Its Applications. Kluwer Academic Publishers, Dordrecht, Netherlands, 2000.

C. A. Floudas and A. Aggarwal. A decomposition strategy for global optimum search in the pooling problem. *ORSA J. Comput.*, 2, 1990.

C. A. Floudas and C. E. Gounaris. A review of recent advances in global optimization. *J. Global Optim.*, 45 (1):3 – 38, 2009.

C. A. Floudas and P. M. Pardalos. State-of-the-art in global optimization - Computational methods and applications - Preface. *J. Glob. Optim.*, 7(2):113, 1995.

C. A. Floudas and G. E. Paules. A mixed-integer nonlinear programming formulation for the synthesis of heat-integrated distillation sequences. *Comput. Chem. Eng.*, 12(6):531–546, 1988.

C. A. Floudas and V. Visweswaran. A global optimization algorithm (GOP) for certain classes of nonconvex NLPs: I. Theory. *Comput. Chem. Eng.*, 14(12):1397–1417, 1990.

C. A. Floudas and V. Visweswaran. Primal-relaxed dual global optimization approach. *J. of Optim. Theory and Appl.*, 78(2):187 – 225, 1993.

C. A. Floudas, A. Aggarwal, and A. R. Ciric. Global optimum search for nonconvex NLP and MINLP problems. *Comput. Chem. Eng.*, 13(10):1117 – 1132, 1989.

C. A. Floudas, I. G. Akrotirianakis, S. Caratzoulas, C. A. Meyer, and J. Kallrath. Global optimization in the 21st century: Advances and challenges. *Comput. Chem. Eng.*, 29:1185 – 1202, 2005.

L. R. Foulds, D. Haughland, and K. Jornsten. A bilinear approach to the pooling problem. *Optim.*, 24: 165–180, 1992.

K. C. Furman and I. P. Androulakis. A novel MINLP-based representation of the original complex model for predicting gasoline emissions. In *Sixth International Conference on Foundations of Computer-Aided Process Design*, pages 525–528, Princeton, NJ, 2004. FOCAPD.

K. C. Furman and I. P. Androulakis. A novel MINLP-based representation of the original complex model for predicting gasoline emissions. *Comput. Chem. Eng.*, 32:2857–2876, 2008.

C. E. Gounaris and C. A. Floudas. Formulation and relaxation of an extended pooling problem. In *2007 AIChE Annual Meeting*, Salt Lake City, Utah, 2007. AIChE.

C. E. Gounaris, R. Misener, and C. A. Floudas. Computational comparison of piecewise-linear relaxations for pooling problems. *Ind. Eng. Chem. Res.*, 48(12):5742 – 5766, 2009.

C. A. Haverly. Studies of the behavior of recursion for the pooling problem. *ACM SIGMAP Bulletin*, 25:19 – 28, 1978.

ILOG. CPLEX. http://www.ilog.com/products/cplex/, 2007. Version 11.0.0.

J. Kallrath. Mixed integer optimization in the chemical process industry: Experience, potential and future perspectives. *Chemical Engineering Research and Design*, 78(6):809 – 822, 2000.

R. Karuppiah and I. E. Grossmann. Global optimization for the synthesis of integrated water systems in chemical processes. *Comput. Chem. Eng.*, 30:650–673, 2006.

A. C. Kokossis and C. A. Floudas. Synthesis of isothermal reactor–separator–recycle systems. *Chem. Eng. Sci.*, 46(5-6):1361 – 1383, 1991.

A. C. Kokossis and C. A. Floudas. Optimization of complex reactor networks–II. Nonisothermal operation. *Chem. Eng. Sci.*, 49(7):1037 – 1051, 1994.

L. S. Lasdon, A. D. Waren, S. Sarkar, and F. Palacios. Solving the pooling problem using generalized reduced gradient and successive linear programming algorithms. *ACM SIGMAP Bulletin*, 27, 1979.

S.-R. Lim and J. M. Park. Synthesis of an environmentally friendly water network system. *Ind. Eng. Chem. Res.*, 47(6):1988 – 1994, 2008.

X. Lin and C. A. Floudas. Design, synthesis and scheduling of multipurpose batch plants via an effective continuous-time formulation. *Comput. Chem. Eng.*, 25(4-6):665 – 674, 2001.

J. Linderoth. A simplicial branch-and-bound algorithm for solving quadratically constrained quadratic programs. *Math. Program.*, 103(2):251–282, 2005.

A. Malcolm, L. Zhang, and A. A. Linninger. Design of environmental regulatory policies for sustainable emission reduction. *AIChE J.*, 52(8):2792 – 2804, 2006.

C. D. Maranas and C. A. Floudas. Finding all solutions of nonlinearly constrained systems of equations. *J. Global Optim.*, 7:143–182, 1995.

G. P. McCormick. Computability of global solutions to factorable nonconvex programs: Part 1-convex underestimating problems. *Math. Program.*, 10(1):147–175, 1976.

C. A. Meyer and C. A. Floudas. Global optimization of a combinatorially complex generalized pooling problem. *AIChE J.*, 52(3):1027 – 1037, 2006.

C. A. Meyer and C. A. Floudas. Convex envelopes for edge-concave functions. *Math. Program.*, 103(2): 207–224, 2005.

R. Misener and C. A. Floudas. Advances for the pooling problem: Modeling, global optimization, and computational studies. *Applied and Computational Mathematics*, 8(1):3 – 22, 2009.

R. Misener and C. A. Floudas. Piecewise-Linear Approximations of Multidimensional Functions. *J. Optim. Theory Appl.*, In Press (DOI 10.1007/s10957-009-9626-0), 2010.

B. A. Murtagh, M. A. Saunders, W. Murray, M. A. Saunders, P. E. Gill, R. Raman, and E. Kalvelagen. MINOS. http://www.gams.com/dd/docs/solvers/minos.pdf, 2004.

V. Pham, C. Laird, and M. El-Halwagi. Convex hull discretization approach to the global optimization of pooling problems. *Ind. Eng. Chem. Res.*, 48:1973 – 1979, 2009.

I. Quesada and I. E. Grossmann. Global optimization of bilinear process networks with multicomponent flows. *Comput. Chem. Eng.*, 19:1219–1242, 1995.

A. K. Rhodes. U.s. refiners make complex-model RFG as they prepare for the next hurdle. *Oil & Gas Journal*, 96:22 – 27, 1998.

H. S. Ryoo and N. V. Sahinidis. Analysis of bounds for multilinear functions. *J. Global Optim.*, 19(4): 403–424, 2001.

Y. Saif, A. Elkamel, and M. Pritzker. Global optimization of reverse osmosis network for wastewater treatment and minimization. *Ind. Eng. Chem. Res.*, 47(9):3060 – 3070, 2008.

H. D. Sherali and W. P. Adams. *A Reformulation-Linearization Technique for Solving Discrete and Continuous Nonconvex Problems*. Nonconvex Optimization and Its Applications. Kluwer Academic Publishers, Dordrecht, Netherlands, 1999.

H. D. Sherali and A. Alameddine. A new reformulation-linearization technique for bilinear programming problems. *J. Global Optim.*, 2:379–410, 1992.

F. Tardella. On a class of functions attaining their maximum at the vertices of a polyhedron. *Discret. Appl. Math.*, 22:191–195, 1988/89.

F. Tardella. On the existence of polyhedral convex envelopes. In C. A. Floudas and P. M. Pardalos, editors, *Frontiers in Global Optimization*, pages 563–573. Kluwer Academic Publishers, 2003.

F. Tardella. Existence and sum decomposition of vertex polyhedral convex envelopes. *Optim. Lett.*, 2: 363–375, 2008.

M. Tawarmalani and N. V. Sahinidis. *Convexification and Global Optimization in Continuous and Mixed-Integer Nonlinear Programming: Theory, Applications, Software, and Applications*. Nonconvex Optimization and Its Applications. Kluwer Academic Publishers, Norwell, MA, USA, 2002.

S. Treiber, R. S. McLeod, Y. Faitakis, and R. L. Hutchings. RFG: The challenge to conventional blending technology. *Hydrocarbon Processing*, 77(6):101 – 103, 1998.

V. Visweswaran. MINLP: Applications in blending and pooling. In C. A. Floudas and P. M. Pardalos, editors, *Encyclopedia of Optimization*, pages 2114–2121. Springer Science, 2 edition, 2009.

V. Visweswaran and C. A. Floudas. A global optimization algorithm (GOP) for certain classes of nonconvex NLPs: II. application of theory and test problems. *Comput. Chem. Eng.*, 14(12):1419–1434, 1990.

V. Visweswaran and C. A. Floudas. New properties and computational improvement of the GOP algorithm for problems with quadratic objective functions and constraints. *J. Global Optim.*, 3:439 – 462, 1993.

D. S. Wicaksono and I. A. Karimi. Piecewise MILP under-and overestimators for global optimization of bilinear programs. *AIChE J.*, 54(4):991–1008, 2008.

A Reformulated Gasoline Parameters

Table 6: Baseline emission values for Eqs. (48), (97) & (125) – (128)

	Summer ($\frac{\text{mg}}{\text{mile}}$)	Winter ($\frac{\text{mg}}{\text{mile}}$)
VOC(b)	907.00	1341.00
NOX(b)	1340.00	1540.00
BENZ(b)	53.54	77.62
ACET(b)	4.44	7.25
FORM(b)	9.70	15.34
BUTA(b)	9.38	15.84

Table 7: Coefficients for Eqs. (72), (133) & (134)

Coefficient	Region 1	Region 2	
α_1^V	1.2269	1.0633	
α_2^V	-0.3534	-0.3008	Eq. (72)
α_3^V	0.0318	0.0270	
α^{POM}	0.003355		Eq. (133)
α_1^{NB}	1.7502	1.5210	
α_2^{NB}	-0.6031	-0.5161	
α_3^{NB}	-0.0403	-0.0352	
α_4^{NB}	0.0738	0.0628	Eq. (134)
α_5^{NB}	0.0116	0.0100	
α_6^{NB}	-0.0026	-0.0022	
α_7^{NB}	-0.0010	-0.0009	

Table 8: Constant Coefficients of Eqs. (49), (98) & (129) – (132)

Constant	Emitter 1 ($e = 1$)	Emitter 2 ($e = 2$)	
$c_{e,1}^{V}$	-0.003641	-0.003626	Eq. (49)
$c_{e,2}^{V}$	0.0005219	-0.000054	
$c_{e,3}^{V}$	0.0289749	0.043295	
$c_{e,4}^{V}$	-0.01447	-0.013504	
$c_{e,5}^{V}$	-0.068624	-0.062327	
$c_{e,6}^{V}$	0.0323712	0.0282042	
$c_{e,7}^{V}$	-0.002858	-0.002858	
$c_{e,8}^{V}$	0.0001072	0.000106	
$c_{e,9}^{V}$	0.0004087	0.000408	
$c_{e,10}^{V}$	-0.0003481	-0.000287	
$c_{e,1}^{N}$	0.0018571	-0.00913	Eq. (98)
$c_{e,2}^{N}$	0.0006921	0.000252	
$c_{e,3}^{N}$	0.0090744	-0.01397	
$c_{e,4}^{N}$	0.000931	0.000931	
$c_{e,5}^{N}$	0.000846	-0.00401	
$c_{e,6}^{N}$	0.0083632	0.007097	
$c_{e,7}^{N}$	-0.002774	-0.00276	
$c_{e,8}^{N}$	-0.000000663	0.0	
$c_{e,9}^{N}$	-0.000119	-0.00007995	
$c_{e,10}^{N}$	0.0003665	0.0003665	
$c_{e,1}^{BE}$	0.0	-0.096047	Eq. (129)
$c_{e,2}^{BE}$	0.0006197	0.000337	
$c_{e,3}^{BE}$	-0.003376	0.011251	
$c_{e,4}^{BE}$	0.02655	0.011882	
$c_{e,5}^{BE}$	0.22239	0.222318	
$c_{e,1}^{F}$	-0.010226	-0.010226	Eq. (130)
$c_{e,2}^{F}$	-0.007166	-0.007166	
$c_{e,3}^{F}$	0.0	-0.031352	
$c_{e,4}^{F}$	0.0462131	0.0462131	
$c_{e,1}^{A}$	0.0002631	0.0002627	Eq. (131)
$c_{e,2}^{A}$	0.039786	0.0	
$c_{e,3}^{A}$	-0.012172	-0.012157	
$c_{e,4}^{A}$	-0.005525	-0.005548	
$c_{e,5}^{A}$	-0.009594	-0.05598	
$c_{e,6}^{A}$	0.31658	0.3164665	
$c_{e,7}^{A}$	0.24925	0.2493259	
$c_{e,1}^{BU}$	0.0	-0.060771	Eq. (132)
$c_{e,2}^{BU}$	0.0001552	0.0	
$c_{e,3}^{BU}$	-0.007253	-0.007311	
$c_{e,4}^{BU}$	-0.014866	-0.008052	
$c_{e,5}^{BU}$	-0.004005	-0.004005	
$c_{e,6}^{BU}$	0.028235	0.043696	

Table 9: Weighting Factors for Eqs. (48), (97) & (125) – (128)

Constant	Emitter 1 ($e = 1$)	Emitter 2 ($e = 2$)
w_e^V	0.444	0.556
w_e^N	0.738	0.262
w_e^T	0.444	0.556

Table 10: Normal ($e = 1$) and high ($e = 2$) emitter values for Eqs. (48), (97) & (125) – (128)

	Summer		Winter	
	$e = 1$	$e = 2$	$e = 1$	$e = 2$
$e^{v_e(b)}$	0.0621	0.1038	0.0579	0.0971
$e^{n_e(b)}$	1.6438	0.8353	1.6664	0.8420
$e^{b_e(b)}$	3.5308	5.8617	3.0231	5.6183
$e^{f_e(b)}$	0.3403	0.2550	0.3542	0.2439
$e^{a_e(b)}$	0.4715	0.3337	0.4862	0.3441
$e^{d_e(b)}$	0.2600	0.4995	0.2688	0.5382

B Case Study 1 Parameters

This case study was taken from Gounaris and Floudas [2007]. The singular pool has capacity $S_1 = 300$.

Table 11: Quality Bounds k on Product j ($P_{j,k}^L$ & $P_{j,k}^U$) and Quality k of Raw Material i ($C_{i,k}$)

	k										
	1	2	3	4	5	6	7	8	9	10	11
$P_{1,k}^L$	0.3	50	6.4	30	70	0.0	0.0	0	0.1	0.1	0.1
$P_{2,k}^L$	0.3	50	6.4	30	70	0.0	0.0	0	0.1	0.1	0.1
$P_{1,k}^U$	4.0	500	10.0	70	100	30	2.0	25	4.0	4.0	4.0
$P_{2,k}^U$	4.0	250	8.0	60	85	25	0.5	10	4.0	4.0	4.0
$C_{1,k}$	0.1	800	6.0	20	70	50	0.0	10	0	0	0
$C_{2,k}$	0.2	400	8.8	60	85	30	0.8	15	0	0	0
$C_{3,k}$	0.4	200	8.0	55	80	25	1.0	15	0	0	0
$C_{4,k}$	0.7	100	8.0	50	75	10	0.2	5	0	0	0
$C_{5,k}$	18.15	0	8.4	100	100	0	0.0	0	18.15	0	0
$C_{6,k}$	15.66	0	8.0	100	100	0	0.0	0	0	15.66	0
$C_{7,k}$	34.73	0	9.6	100	100	0	0.0	0	0	0	34.73

Table 12: Cost (c_i) and Availability (A_i^L & A_i^U) of Feed i

	i						
	1	2	3	4	5	6	7
c_i	2	8	10	16	2	2	5
A_i^L	50	0	0	0	0	0	0
A_i^U	400	200	200	100	10	10	50

Table 13: Price (d_j) and Demand (D_j^L & D_j^U) of Product j

j	d_j	D_j^L	D_j^U
1	6	100	200
2	12	100	200

C Case Study 2 Parameters

Table 14: Price (d_j) and Demand (D_j^L & D_j^U) of Product j

	j		
	1	2	3
d_j	6.3	5.5	5.0
D_j^L	100	0	50
D_j^U	700	800	400

Table 15: Capacity S_l of Pool l

	l		
	1	2	3
S_1	400	575	500

Table 16: Quality k of Raw Material i ($C_{i,k}$)

	k										
	1	2	3	4	5	6	7	8	9	10	11
$C_{1,k}$	0.00	0	60.0	100	100	0.0	0	0	0.00	0.00	0.00
$C_{2,k}$	0.00	0	21.0	100	100	0.0	0	0	0.00	0.00	0.00
$C_{3,k}$	0.00	0	7.4	50	95	0.0	0	0	0.00	0.00	0.00
$C_{4,k}$	0.00	50	10.0	100	100	0.0	0	0	0.00	0.00	0.00
$C_{5,k}$	0.00	100	9.0	70	100	7.5	2	37	0.00	0.00	0.00
$C_{6,k}$	0.00	15	3.4	60	85	3.2	0	12	0.00	0.00	0.00
$C_{7,k}$	0.00	200	10.2	85	100	10.0	1	60	0.00	0.00	0.00
$C_{8,k}$	0.00	400	8.2	45	80	35.0	3	20	0.00	0.00	0.00
$C_{9,k}$	0.00	700	2.1	15	60	65.0	4	15	0.00	0.00	0.00
$C_{10,k}$	0.00	10	7.4	30	70	60.0	5	3	0.00	0.00	0.00
$C_{11,k}$	18.15	0	8.8	100	100	0.0	0	0	18.15	0.00	0.00
$C_{12,k}$	15.66	0	5.7	95	100	0.0	0	0	0.00	15.66	0.00
$C_{13,k}$	15.66	0	2.7	70	100	0.0	0	0	0.00	0.00	0.00
$C_{14,k}$	34.73	0	23.0	100	100	0.0	0	0	0.00	0.00	34.73

Table 17: Quality Bounds k on Product j ($P^L_{j,k}$ & $P^U_{j,k}$)

	k										
	1	2	3	4	5	6	7	8	9	10	11
$P^L_{1,k}$	0.0	0.0	6.4	30	70	0	0	0	0.0	0.0	2.0
$P^L_{2,k}$	0.0	0.0	6.4	30	70	0	0	0	0.0	0.0	1.0
$P^L_{3,k}$	0.0	0.0	6.4	30	70	0	0	0	0.0	0.0	0.0
$P^U_{1,k}$	3.7	130	10	70	100	50	2	25	3.7	3.7	3.7
$P^U_{2,k}$	3.7	200	10	70	100	50	2	25	3.7	3.7	3.7
$P^U_{3,k}$	3.7	250	10	70	100	50	2	25	3.7	3.7	3.7

Table 18: Cost (c_i) and Availability (A^L_i & A^U_i) of Feed i

	i													
	1	2	3	4	5	6	7	8	9	10	11	12	13	14
c_i	3.0	2.0	3.5	2.0	1.0	3.0	0.7	0.5	0.3	2.5	7.5	10.5	8.5	5.5
A^L_i	0	0	0	0	0	0	0	10	20	10	0	0	0	0
A^U_i	75	50	75	75	300	150	50	100	200	100	5	10	10	100

D Case Study 3 Parameters

The raw material qualities ($C_{i,k}$) are identical in Case Study 2 & 3, so Table 16 records these values.

Table 19: Price (d_j) and Demand (D^L_j & D^U_j) of Product j

	j									
	1	2	3	4	5	6	7	8	9	10
d_j	8.0	7.5	6.5	6.0	6.5	5.5	5.0	5.5	5.5	5.0
D^L_j	100	0	50	50	50	50	50	50	100	40
D^U_j	700	800	400	400	300	200	400	300	200	400

Table 20: Capacity S_l of Pool l

	l				
	1	2	3	4	5
S_1	900	575	500	800	900

Table 21: Cost (c_i) and Availability (A^L_i & A^U_i) of Feed i

	i													
	1	2	3	4	5	6	7	8	9	10	11	12	13	14
c_i	3.0	2.0	3.5	2.0	1.0	3.0	0.7	0.5	0.3	2.5	7.5	10.5	8.5	5.5
A^L_i	0	0	0	0	0	0	0	100	50	10	0	0	0	0
A^U_i	175	150	375	375	900	350	250	600	500	200	50	100	100	400

Table 22: Quality Bounds k on Product j ($P^{L}_{j,k}$ & $P^{U}_{j,k}$)

	k										
	1	2	3	4	5	6	7	8	9	10	11
$P^{L}_{1,k}$	0.0	0.0	6.4	30.0	70.0	0.0	0.0	0.0	0.0	0.0	2.0
$P^{L}_{2,k}$	0.0	0.0	6.4	30.0	70.0	0.0	0.0	0.0	0.0	0.0	2.0
$P^{L}_{3,k}$	0.0	0.0	6.4	30.0	70.0	0.0	0.0	0.0	0.0	0.0	2.0
$P^{L}_{4,k}$	0.0	0.0	6.4	30.0	70.0	0.0	0.0	0.0	0.0	0.0	2.0
$P^{L}_{5,k}$	0.0	0.0	6.4	30.0	70.0	0.0	0.0	0.0	0.0	0.0	2.0
$P^{L}_{6,k}$	0.0	0.0	6.4	30.0	70.0	0.0	0.0	0.0	0.0	0.0	1.0
$P^{L}_{7,k}$	0.0	0.0	6.4	30.0	70.0	0.0	0.0	0.0	0.0	0.0	1.0
$P^{L}_{8,k}$	0.0	0.0	6.4	30.0	70.0	0.0	0.0	0.0	0.0	0.0	1.0
$P^{L}_{9,k}$	0.0	0.0	6.4	30.0	70.0	0.0	0.0	0.0	0.0	0.0	1.0
$P^{L}_{10,k}$	0.0	0.0	6.4	30.0	70.0	0.0	0.0	0.0	0.0	0.0	0.0
$P^{U}_{1,k}$	3.7	130	10	70	100	50	2	25	3.7	3.7	3.7
$P^{U}_{2,k}$	3.7	150	10	70	100	50	2	25	3.7	3.7	3.7
$P^{U}_{3,k}$	3.7	170	10	70	100	50	2	25	3.7	3.7	3.7
$P^{U}_{4,k}$	3.7	190	10	70	100	50	2	25	3.7	3.7	3.7
$P^{U}_{5,k}$	3.7	150	10	70	100	50	2	25	3.7	3.7	3.7
$P^{U}_{6,k}$	3.7	150	10	70	100	50	2	25	3.7	3.7	3.7
$P^{U}_{7,k}$	3.7	150	10	70	100	50	2	25	3.7	3.7	3.7
$P^{U}_{8,k}$	3.7	200	10	70	100	50	2	25	3.7	3.7	3.7
$P^{U}_{9,k}$	3.7	200	10	70	100	50	2	25	3.7	3.7	3.7
$P^{U}_{10,k}$	3.7	250	10	70	100	50	2	25	3.7	3.7	3.7