On the P-formulation and the Split-Fraction-Formulation for the Generalized Pooling Problem

Xin Cheng, Xiang Li*

Department of Chemical Engineering, Queen's University, 19 Division Street, Kingston, ON, K7L 3N6, Canada

---

Abstract

The generalized pooling problem (GPP) is a NP-hard problem for which the solution time for securing a global optimal solution heavily depends on the strength of the problem formulation. The existing GPP formulations use either quality variables (P-formulation and the variants) or split-fraction variables (SF-formulation and the variants) to model the material balance at the pools. This paper is the first attempt to develop theoretical results for comparing the strength of P-formulation and SF-formulation. It is found that, an enhanced version of P-formulation, called P$^+$-formulation, is at least as strong as SF-formulation under mild conditions. Furthermore, P$^+$-formulation becomes identical to P-formulation when the pooling network comprises only mixers and splitters. With additional conditions that are often satisfied at the root node, P-formulation is proved to be as least as strong as SF-formulation. The theoretical results are verified by the computational study of 23 problem instances.

Keywords: Pooling Problem, Bilinear programming, Global optimization, P formulation, Split fraction formulation

---

Introduction

The generalized pooling problem (GPP) is an extension of the pooling problem in that it allows the interconnection between pools. It was first considered by Audet et al. [1] and

*Corresponding author

Email addresses: x.cheng@queensu.ca (Xin Cheng), xiang.li@queensu.ca (Xiang Li)

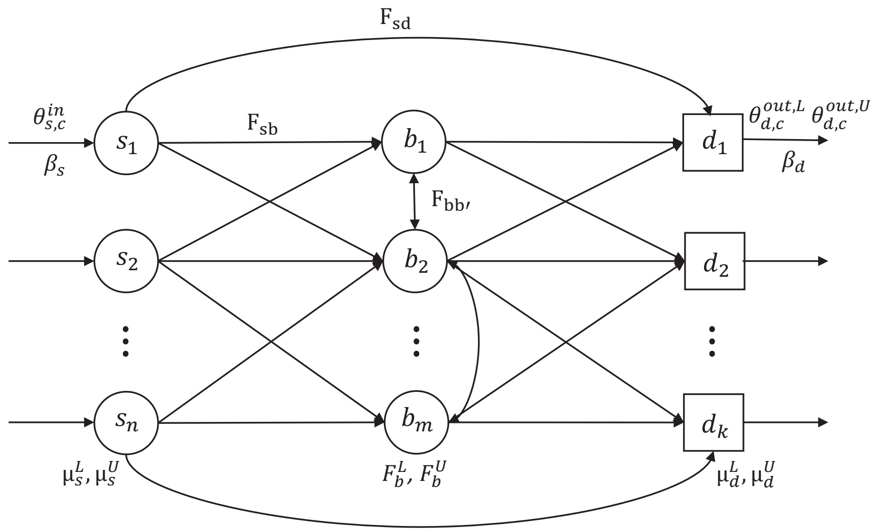


Figure 1: Network Structure of the Generalized Pooling Problem

has been widely applied to model many industrial problems, such as the crude oil blending problem [2, 3] and the natural gas transport problem [4]. The GPP can also be extended to model problems where materials can be removed or added at certain nodes of the network, such as water treatment networks [5, 6], water using networks [7, 8], and hydrogen networks [9, 10].

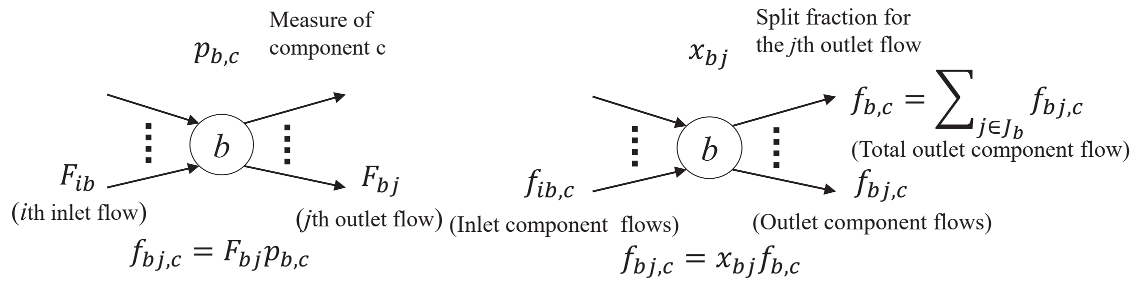
The GPP considers a network of 3 hierarchies of nodes, i.e., source nodes s, pools b and demand nodes d, as shown in Fig. 1. Each source node s provides material with several quality measures $\theta_{s,c}^{in}, \forall c \in C$ at cost β_s. A flow from a source node or a pool can go to a pool or directly to a demand tank. The demand tanks include final products that are subject to quality limitation $\theta_{d,c}^{out,L}$, $\theta_{d,c}^{out,U}$ and are sold at price β_d. There are bounds on the total flow through each arc. The objective of optimization is to find the flows in the network that maximize the profit while satisfying the quality and flow rate constraints. Table 1 summarizes the list of symbols for the GPP. The units in the table are nominal and are merely to ensure the consistency of variable and parameter units. In this paper, we assume no mass loss in the pools. We also assume linear blending, i.e., qualities of a flow are linearly dependent on the composition of the flow. For convenience, we also limit the quality measure to the component mass fraction. Our results can be readily extended to other quality measures

Table 1: List of Symbols for the Generalized Pooling Problem

Index and Sets		
$s \in S$	Subset for supply tanks	
$b \in B$	Subset for pools	
$d \in D$	Subset for demand tanks	
$i, j \in N = S \cup B \cup D$	Nodes represent supply tanks, pools and demand tanks	
$(i, j) \in A$	Allowable arcs in network	
$j \in J_i$	Set of j that $(i, j) \in A$	
$i \in I_j$	Set of i that $(i, j) \in A$	
$c \in C$	Components	
Parameters		**Units**
β_s/β_d	Cost/Profit for flow from s/to d.	\$/t
$\theta^{in}_{s,c}$	Concentration of component c in source tank s	kg/t
$\theta^{out,L}_{d,c}/\theta^{out,U}_{d,c}$	Demand tank lower/upper limit on component c	kg/t
μ^L_i/μ^U_i	Lower/Upper bound of node capacity	t/h
F^L_{ij}/F^U_{ij}	Lower/upper bound on total flow goes through each arc	t/h
Variables		**Units**
F_i	Total flow goes through node i	t/h
F_{ij}	Total flow along arc (i, j)	t/h
$f_{ij,c}$	Individual component flow along with arc (i, j)	kg/h
$f_{b,c}$	Total component coming into pool b	kg/h
$p_{b,c}$	Concentration of component c in pool b	kg/t
x_{bj}	Split fraction variable.	

because of the linear blending assumption.

The GPP problem can be formulated in different forms. There are two major types of GPP formulations in the literature. One is the P-formulation [11] and the variants. The other is the split fraction formulation [12], called SF-formulation in this paper, and the variants. The major difference between the two is how the blending at the pools is modeled. Fig. 2 compares the mass balances at a pool in the P- and the SF-formulations. In P-formulation the mass balance of a component is modeled through the total flow and the quality variable (e.g., component fraction), while in SF-formulation it is modeled through



(a) Mass balance in P-formulation

(b) Mass balance in SF-formulation

Figure 2: Pool mass balances in P-formulation and SF-formulation

individual component flows and split fraction variables.

Seeking a global optimal solution for the GPP is difficult as it is known to be NP-hard [13]. Within the classical branch-and-bound framework for global optimization, the solution efficiency largely depends on the lower bound obtained from the relaxation problem. A tighter relaxation problem provides a better lower bound and helps prune more nodes in the branch-and-bound tree. A formulation with a tighter relaxation is said to be a stronger formulation. Efforts have been made to develop stronger variants of P- and SF-formulations in order for more efficient global optimization. Alfaki et al. [14] proposed a variant of P-formulation, called multi-commodity flow formulation, where the quality variable is the commodity fraction rather than the component fraction. It is proved that, under mild bound consistency conditions, the multi-commodity flow formulation is at least as strong as the P-formulation. Boland et al. [15] further proposed other commodity based variants of P-formulation and demonstrated their computational performance through extensive simulation study. On the other hand, Lotero et al. [16] proposed a variant of SF-formulation for multi-period blending problems, and in this formulation commodity flows rather than component flows are considered in mass balance equations. They proved that the proposed formulation is at least as strong as SF-formulation. Recently, Cheng and Li [17] proposed another commodity based variant of SF-formulation for wastewater treatment network optimization and proved that it is as least as strong as SF-formulation under mild conditions.

However, there have been few results in the literature in comparing P-formulation and SF-formulation. Karuppiah and Grossmann [18] proposed to use P-formulation rather than SF-formulation for optimization of integrated water systems. They thought the SF-formulation tended to have wider ranges variable bounds and lead to numerical difficulties. Castro et al. [6] solved a wastewater treatment network problem using both P-formulation and SF-formulation and showed that the computational time required by BARON[19] is less when using SF-formulation. Castro et al. [20] solved multi-period blending problems using a decomposition algorithm based on multi-parametric disaggregation, and they showed that SF-formulation results in better computational results than P-formulation.

The original motivation of this paper was to compare the strength of P-formulation and SF-formulation systematically and explain why one formulation is better than the other for certain types of problems. We then found that theoretical comparison of P-formulation and SF-formulation was difficult. However, we have found that, a slightly stronger version of P-formulation, called P^+-formulation in this paper, can be proved to be at least as strong as SF-formulation under mild bound consistency conditions. In certain conditions, the P^+-formulation is equivalent to P-formulation. We have also compared the performance of P-formulation, SF-formulation, P^+-formulation through computational study, and the results are consistent with our theoretical results. To best of our knowledge, this is the first systematic computational comparison of P-formulation and SF-formulation for the GPP.

The rest of the paper is organized as follows. We first states the three formulations under consideration, i.e., P-formulation, SF-formulation, and the P^+-formulation first proposed in this paper. The next section proves the strength of P^+-formulation over P-formulation and SF-formulation, and it also shows that P^+-formulation is essentially P-formulation when viewing the pools as mixers connected with splitters. Then we discuss the strength of the formulations at the root node of branch-and-bound search. The next section provides and discusses the computational study results. The last section concludes the paper.

Formulations

We consider the P-formulation in Alfaki et al.[14] and the SF-formulation in Quesada and Grossmann [12]. But for convenience of comparison and discussion, we use a different set of symbols for variables and parameters (as listed in Table 1). The P-formulation is:

$$
\begin{aligned}
\text{(P):} \quad \min \quad & \sum_{s\in S}\sum_{j\in J_s} \beta_s F_{sj} - \sum_{d\in D}\sum_{i\in I_d} \beta_d F_{id} && \text{(P-1)}\\
s.t. \quad & f_{bj,c} = F_{bj} p_{b,c}, \quad \forall b \in B, j \in J_b, c \in C, && \text{(P-2)}\\
& f_{sj,c} = F_{sj} \theta^{in}_{s,c}, \quad \forall b \in B, j \in J_b, c \in C, && \text{(P-3)}\\
& \sum_{i\in I_b} f_{ib,c} = \sum_{j\in J_b} f_{bj,c}, \quad \forall b \in B, c \in C, && \text{(P-4)}\\
& \sum_{i\in I_b} F_{ib} = \sum_{j\in J_b} F_{bj}, \quad \forall b \in B, && \text{(P-5)}\\
& \sum_{i\in I_d} f_{id,c} \le \sum_{i\in I_d} F_{id} \theta^{out,U}_{d,c}, \quad \forall d \in D, c \in C, && \text{(P-6A)}\\
& \sum_{i\in I_d} f_{id,c} \ge \sum_{i\in I_d} F_{id} \theta^{out,L}_{d,c}, \quad \forall d \in D, c \in C, && \text{(P-6B)}\\
& \sum_{i\in I_b} F_{ib} = F_b, \quad \forall b \in B && \text{(P-7)}\\
& \mu^L_s \le \sum_{j\in J_s} F_{sj} \le \mu^U_s, \quad \forall s \in S && \text{(P-8)}\\
& F^L_b \le F_b \le F^U_b, \quad \forall b \in B && \text{(P-9)}\\
& \mu^L_d \le \sum_{i\in I_d} F_{id} \le \mu^U_d, \quad \forall d \in D, && \text{(P-10)}\\
& F^L_{ij} \le F_{ij} \le F^U_{ij}, \quad \forall (i,j) \in A. && \text{(P-11)}\\
& p^L_{b,c} \le p_{b,c} \le p^U_{b,c}, \quad \forall b \in B, c \in C. && \text{(P-12)}
\end{aligned}
$$

The objective function given in Eq. (P-1) is to minimize the overall cost. Eq. (P-2) contains the bilinear term and it enforces the same component concentration in all outflows. Eq. (P-5) and Eq. (P-5) give total mass balance and component mass balance around pools, respectively. Eq. (P-3) is the component mass balance around source nodes. Eq. (P-6A) and Eq. (P-6B) put quality bounds on the final products at demand nodes. Eq. (P-7) links the total flow goes through the pool F_b with the each inlet flow F_{ib}. Eq. (P-8)-Eq. (P-10) give respectively the bounds on the total flow going through source nodes, pools and demand nodes. Eq. (P-11) puts bounds on the total flow going through the pipeline at each arc. Eq. (P-12) gives the bounds of the quality variable $p_{b,c}$, which are determined by the stream quality at source nodes: $p^L_{b,c} = \min_{s \in S} \theta^{in}_{s,c}$, $p^U_{b,c} = \max_{s \in S} \theta^{in}_{s,c}$.

The SF-formulation is:

$$\text{(SF):}\quad \min \quad \sum_{s\in S}\sum_{j\in J_s}\beta_s F_{sj} - \sum_{d\in D}\sum_{i\in I_d}\beta_d F_{id} \tag{SF-1}$$

$$s.t.\quad f_{bj,c} = f_{b,c}x_{bj}, \quad \forall b\in B, j\in J_b, c\in C, \tag{SF-2}$$

$$F_{bj} = F_b x_{bj}, \quad \forall b\in B, j\in J_b \tag{SF-3}$$

$$f_{sj,c} = F_{sj}\theta^{in}_{s,c}, \quad \forall s\in S, j\in J_s, \tag{SF-4}$$

$$\sum_{j\in J_b} x_{bj} = 1, \quad \forall b\in B, \tag{SF-5}$$

$$\sum_{j\in J_b} f_{bj,c} = f_{b,c}, \quad \forall b\in B, c\in C, \tag{SF-6}$$

$$\sum_{j\in J_b} F_{bj} = F_b, \quad \forall b\in B, \tag{SF-7}$$

$$f_{b,c} = \sum_{i\in I_b} f_{ib,c}, \quad \forall b\in B, c\in C, \tag{SF-8}$$

$$F_b = \sum_{i\in I_b} F_{ib}, \quad \forall b\in B, \tag{SF-9}$$

$$\sum_{i\in I_d} f_{id,c} \le \sum_{i\in I_d} F_{id}\theta^{out,U}_{d,c}, \quad \forall d\in D, c\in C, \tag{SF-10A}$$

$$\sum_{i\in I_d} f_{id,c} \ge \sum_{i\in I_d} F_{id}\theta^{out,L}_{d,c}, \quad \forall d\in D, c\in C, \tag{SF-10B}$$

$$\mu^L_s \le \sum_{j\in J_s} F_{sj} \le \mu^U_s, \quad \forall s\in S \tag{SF-11}$$

$$F^L_b \le F_b \le F^U_b, \quad \forall b\in B \tag{SF-12}$$

$$\mu^L_d \le \sum_{i\in I_d} F_{id} \le \mu^U_d, \quad \forall d\in D, \tag{SF-13}$$

$$F^L_{ij} \le F_{ij} \le F^U_{ij}, \quad \forall (i,j)\in A. \tag{SF-14}$$

$$0 \le x_{bj} \le 1, \quad \forall b\in B, j\in J_b. \tag{SF-15}$$

Eq. (SF-1) gives the objective same as in P-formulation. Eq. (SF-2) and Eq. (SF-3) contain the bilinear term and they ensure that the same fraction of the total flow and the component flows through pool b enter arc (b,j). Eq. (SF-5) states the sum of split-fraction variable x_{bj} over all outflows must be 1. Eq. (SF-4) is the component mass balance around the source nodes. Eq. (SF-7) and Eq. (SF-9) are the total mass balance around pools. Eq. (SF-6) and Eq. (SF-8) are the component mass balance around pools. Eq. (SF-10A) and Eq. (SF-10B) are quality bounds on the final products. Eq. (SF-11) - (SF-13) put bounds on the total flow going through source nodes, pools and demand nodes respectively. Eq. (SF-14) put bounds on the total flow going through the pipeline at each arc. Eq. (SF-15) give bounds on x_{bj}.

The P$^+$-formulation that we propose in this paper includes the same objective function and all constraints in P-formulation, plus additional variables $f_{b,c}$, F_b and the relevant equations:

$$
\begin{aligned}
(\mathrm{P}^+):\quad \min\quad & (\text{P-1}) \\
s.t.\quad & (\text{P-2}) - (\text{P-12}), \\
& f_{b,c} = \sum_{j \in J_b} f_{bj,c}, \quad \forall b \in B, && (\mathrm{P}^+\text{-1}) \\
& F_b = \sum_{j \in J_b} F_{bj}, \quad \forall b \in B, && (\mathrm{P}^+\text{-2}) \\
& f_{b,c} = F_b p_{b,c}, \quad \forall b \in B, c \in C. && (\mathrm{P}^+\text{-3})
\end{aligned}
$$

Note that Eq. (P$^+$-1) and Eq. (P$^+$-2) are also included in SF-formulation. Eq. (P$^+$-3) contains bilinear term so it incurs more nonconvexity. The additional constraints are redundant for P+-formulation in the sense that they do not alter global optimal solutions, but they do improve the strength of the formulation.

Strength of the Formulations

We compare the strength of the formulations by comparing the tightness of their convex relaxations. The convex relaxation of P-formulation, the following formulation (P-R), is obtained by replacing the bilinear terms in (P-2) with their convex envelopes:

$$
\begin{aligned}
\text{(P-R):} \quad \min \quad & \text{(P-1)} \\
s.t. \quad & \text{(P-3) - (P-12)}, \\
& f_{bj,c} \leq F^U_{bj} p_{b,c} + F_{bj} p^L_{b,c} - F^U_{bj} p^L_{b,c}, && \forall b \in B, j \in J_b, c \in C, && \text{(P-2-R1)} \\
& f_{bj,c} \leq F^L_{bj} p_{b,c} + F_{bj} p^U_{b,c} - F^L_{bj} p^U_{b,c}, && \forall b \in B, j \in J_b, c \in C, && \text{(P-2-R2)} \\
& f_{bj,c} \geq F^L_{bj} p_{b,c} + F_{bj} p^L_{b,c} - F^L_{bj} p^L_{b,c}, && \forall b \in B, j \in J_b, c \in C, && \text{(P-2-R3)} \\
& f_{bj,c} \geq F^U_{bj} p_{b,c} + F_{bj} p^U_{b,c} - F^U_{bj} p^U_{b,c}, && \forall b \in B, j \in J_b, c \in C. && \text{(P-2-R4)}
\end{aligned}
$$

Similarly, the convex relaxation of P$^+$-formulation is:

$$
\begin{aligned}
\text{(P}^+\text{-R):} \quad \min \quad & \text{(P-1)} \\
s.t. \quad & \text{(P-3) - (P-12)}, \\
& \text{(P}^+\text{-1)} - \text{(P}^+\text{-2)}, \\
& \text{(P-2-R1) - (P-2-R4)}, \\
& f_{b,c} \leq F^U_{b} p_{b,c} + F_{b} p^L_{b,c} - F^U_{b} p^L_{b,c}, && \forall b \in B, c \in C, && \text{(P}^+\text{-3-R1)} \\
& f_{b,c} \leq F^L_{b} p_{b,c} + F_{b} p^U_{b,c} - F^L_{b} p^U_{b,c}, && \forall b \in B, c \in C, && \text{(P}^+\text{-3-R2)} \\
& f_{b,c} \geq F^L_{b} p_{b,c} + F_{b} p^L_{b,c} - F^L_{b} p^L_{b,c}, && \forall b \in B, c \in C, && \text{(P}^+\text{-3-R3)} \\
& f_{b,c} \geq F^U_{b} p_{b,c} + F_{b} p^U_{b,c} - F^U_{b} p^U_{b,c}, && \forall b \in B, c \in C, && \text{(P}^+\text{-3-R4)}
\end{aligned}
$$

and the convex relaxation of SF-formulation is:

$$
\begin{aligned}
\text{(SF-R):} \quad & \min \quad \text{(SF-1)} \\
& s.t. \quad \text{(SF-4) - (SF-15)}, \\
& f_{bj,c} \leq f^U_{b,c} x_{bj} + f_{b,c} x^L_{bj} - f^U_{b,c} x^L_{bj}, && \forall b \in B, j \in J_b, c \in C, && \text{(SF-2-R1)} \\
& f_{bj,c} \leq f^L_{b,c} x_{bj} + f_{b,c} x^U_{bj} - f^L_{b,c} x^U_{bj}, && \forall b \in B, j \in J_b, c \in C, && \text{(SF-2-R2)} \\
& f_{bj,c} \geq f^L_{b,c} x_{bj} + f_{b,c} x^L_{bj} - f^L_{b,c} x^L_{bj}, && \forall b \in B, j \in J_b, c \in C, && \text{(SF-2-R3)} \\
& f_{bj,c} \geq f^U_{b,c} x_{bj} + f_{b,c} x^U_{bj} - f^U_{b,c} x^U_{bj}, && \forall b \in B, j \in J_b, c \in C, && \text{(SF-2-R4)} \\
& F_{bj} \leq F^U_b x_{bj} + F_b x^L_{bj} - F^U_b x^L_{bj}, && \forall b \in B, j \in J_b, && \text{(SF-3-R1)} \\
& F_{bj} \leq F^L_b x_{bj} + F_b x^U_{bj} - F^L_b x^U_{bj}, && \forall b \in B, j \in J_b, && \text{(SF-3-R2)} \\
& F_{bj} \geq F^L_b x_{bj} + F_b x^L_{bj} - F^L_b x^L_{bj}, && \forall b \in B, j \in J_b, && \text{(SF-3-R3)} \\
& F_{bj} \geq F^U_b x_{bj} + F_b x^U_{bj} - F^U_b x^U_{bj}, && \forall b \in B, j \in J_b. && \text{(SF-3-R4)}
\end{aligned}
$$

Since problem P^+-R includes problem P-R plus additional constraints, the following proposition is obvious:

Proposition 1. *The optimal objective value of problem (P^+-R) is no less than that of problem (P-R).*

Next, we are to show P^+-formulation is at least as strong as SF-formulation. Since the two formulations include different variables, we make the following bound consistency assumptions in order for a fair comparison.

$$
\begin{aligned}
& x^L_{bj} = \frac{F^L_{bj}}{F^U_b}, && \forall b \in B, j \in J_b, && \text{(b-1)} \\
& x^U_{bj} = \min\Big(\frac{F^U_{bj}}{F^L_b}, 1\Big), && \forall b \in B, j \in J_b, && \text{(b-2)} \\
& f^L_{b,c} = F^L_b p^L_{b,c}, && \forall b \in B, c \in C, && \text{(b-3)} \\
& f^U_{b,c} = F^U_b p^U_{b,c}, && \forall b \in B, c \in C. && \text{(b-4)}
\end{aligned}
$$

The lower bound of x_{bj} given in Eq. (b-1) is a reasonable value one would infer from given variable bounds in P-formulation. Similarly, Eq. (b-2) provides a reasonable upper bound of x_{bj}. Eq. (b-3) and Eq. (b-4) provide bounds of $f_{b,c}$ that could be inferred from given variables bounds in P-formulation.

Proposition 2. *If the bounds consistency conditions Eq. (b-1)-(b-4) hold, the optimal objective value of problem (P^+-R) is no less than that of problem (SF-R).*

Proof. Let $(f_{b,c}, f_{bj,c}, F_b, F_{bj}, p_{b,c})$ be a feasible point for problem (P^+-R). Using this point we construct a point $(x_{bj}, f_{b,c}, f_{bj,c}, F_b, F_{bj})$ such that $x_{bj} = \frac{F_{bj}}{F_b}$ $(\forall b \in B, j \in J_b)$, and we show that this new point is feasible for problem (SF-R).

First, we show that the point satisfies all constraints in (SF-R) except the convex envelopes. Eq. (SF-4),(SF-6),(SF-9)-(SF-14) are same as Eq. (P-3), (P-13), (P-6)-(P-11) respectively. Eq. (SF-5) is satisfied since $\sum_{j\in J_b} x_{bj} = \frac{\sum_{j\in J_b} F_{bj}}{F_b} = 1, \forall b$. Eq. (SF-7) can be derived from Eq. (P-5) and Eq. (P-7). Eq. (SF-8) can be derived from Eq. (P-4) and Eq. (P^+-1). Eq. (SF-15) is satisfied by the construction of x_{bj}.

Next, we show that the point satisfies Eq. (SF-2-R1) - Eq. (SF-2-R4). To prove Eq. (SF-2-R1), we need to prove

$$L = f^U_{b,c} x_{bj} + f_{b,c} x^L_{bj} - f^U_{b,c} x^L_{bj} - f_{bj,c} \geq 0$$

From Eq. (P^+-2-R2) and Eq. (P^+-3-R4),

$$\begin{aligned} L &\geq f^U_{b,c}\frac{F_{bj}}{F_b} + (F^U_b p_{b,c} + F_b p^U_{b,c} - F^U_b p^U_{b,c})\frac{F^L_{bj}}{F^U_b} - f^U_{b,c}\frac{F^L_{bj}}{F^U_b} - (F^L_{bj} p_{b,c} + F_{bj} p^U_{b,c} - F^L_{bj} p^U_{b,c}) \\ &= f^U_{b,c}\frac{F_{bj}}{F_b} + F_b p^U_{b,c}\frac{F^L_{bj}}{F^U_b} - f^U_{b,c}\frac{F^L_{bj}}{F^U_b} - F_{bj} p^U_{bc}. \end{aligned}$$

According to condition (b-4), the above inequality can be further written as

$$
\begin{aligned}
L &\geq p_{b,c}^U F_b^U \frac{F_{bj}}{F_b} + F_b p_{b,c}^U \frac{F_{bj}^L}{F_b^U} - p_{b,c}^U F_b^U \frac{F_{bj}^L}{F_b^U} - F_{bj} p_{bc}^U \\
&= \frac{F_{bj} p_{b,c}^U}{F_b}(F_b^U - F_b) + \frac{F_{bj}^L p_{b,c}^U}{F_b^U}(F_b - F_b^U) \\
&= p_{b,c}^U (F_b^U - F_b)(\frac{F_{bj}}{F_b} - \frac{F_{bj}^L}{F_b^U}),
\end{aligned}
$$

which is obviously non-negative.

We can also show Eq. (SF-2-R2) is satisfied. Note that x_{bj}^U is either F_{bj}^U/F_b^L or 1 depending on which value is smaller. If $x_{bj}^U = F_{bj}^U/F_b^L$, then from Eq. (P$^+$-2-R1), Eq. (P$^+$-3-R3) and condition (b-3),

$$
\begin{aligned}
L &= f_{b,c}^L x_{bj} + f_{b,c} x_{bj}^U - f_{b,c}^L x_{bj}^U - f_{bj,c} \\
&\geq f_{b,c}^L \frac{F_{bj}}{F_b} + (F_b^L p_{b,c} + F_b p_{b,c}^L - F_b^L p_{b,c}^L)\frac{F_{bj}^U}{F_b^L} - f_{b,c}^L \frac{F_{bj}^U}{F_b^L} - (F_{bj}^U p_{b,c} + F_{bj} p_{b,c}^L - F_{bj}^U p_{b,c}^L) \\
&= \frac{F_{bj} p_{b,c}^L}{F_b}(F_b^L - F_b) + \frac{F_{bj}^U p_{b,c}^L}{F_b^L}(F_b - F_b^L) \\
&= p_{b,c}^L (F_b - F_b^L)(\frac{F_{bj}^U}{F_b^L} - \frac{F_{bj}}{F_b}) \geq 0.
\end{aligned}
$$

If $x_{bj}^U = 1$, then from Eq. (P-2-R3), Eq. (P$^+$-1), Eq. (P$^+$-2), and condition (b-3),

$$
\begin{aligned}
L &= f_{b,c} + f_{b,c}^L \frac{F_{bj}}{F_b} - f_{b,c}^L - f_{bj,c} \\
&= \sum_{j' \neq j} f_{bj',c} + f_{b,c}^L \frac{F_{bj}}{F_b} - f_{b,c}^L \\
&\geq \sum_{j' \neq j} F_{bj'} p_{b,c}^L + f_{b,c}^L \frac{F_{bj}}{F_b} - f_{b,c}^L \\
&= p_{b,c}^L (F_b - F_{bj} + F_b^L \frac{F_{bj}}{F_b} - F_b^L) \\
&= p_{b,c}^L (F_b - F_b^L)(1 - \frac{F_{bj}}{F_b}) \geq 0.
\end{aligned}
$$

Similarly, we can prove Eq. (SF-2-R3) and Eq. (SF-2-R4) are satisfied. To prove Eq. (SF-2-R3), from Eq. Eq. (P-2-R3), Eq. (P^+-3-R1), and condition (b-3), we have

$$
\begin{aligned}
L &= f_{bj,c} - f_{b,c}^L x_{bj} - f_{b,c} x_{bj}^L + f_{b,c}^L x_{bj}^L \\
&\geq (F_{bj}^L p_{b,c} + F_{bj} p_{b,c}^L - F_{bj}^L p_{b,c}^L) - f_{b,c}^L \frac{F_{bj}}{F_b} - (F_b^U p_{b,c} + F_b p_{b,c}^L - F_b^U p_{b,c}^L) \frac{F_{bj}^L}{F_b^U} + f_{b,c}^L \frac{F_{bj}^L}{F_b^U} \\
&= \frac{F_{bj} p_{b,c}^L}{F_b} (F_b - F_b^L) - \frac{F_{bj}^L p_{b,c}^L}{F_b^U} (F_b - F_b^L) \\
&= p_{b,c}^L (F_b - F_b^L) (\frac{F_{bj}}{F_b} - \frac{F_{bj}^L}{F_b^U}) \geq 0.
\end{aligned}
$$

To prove Eq. (SF-2-R4), consider the two possible values of x_{bj}^U. If $x_{bj}^U = F_{bj}^U / F_b^L$, then from Eq. (P-2-R4), Eq. (P^+-3-R2), and condition (b-4),

$$
\begin{aligned}
L &= f_{bj,c} - f_{b,c}^U x_{bj} - f_{b,c} x_{bj}^U + f_{b,c}^U x_{bj}^U \\
&\geq (F_{bj}^U p_{b,c} + F_{bj} p_{b,c}^U - F_{bj}^U p_{b,c}^U) - f_{b,c}^U \frac{F_{bj}}{F_b} - (F_b^L p_{b,c} + F_b p_{b,c}^U - F_b^L p_{b,c}^U) \frac{F_{bj}^U}{F_b^L} + f_{b,c}^U \frac{F_{bj}^U}{F_b^L} \\
&= \frac{F_{bj} p_{b,c}^U}{F_b} (F_b - F_b^U) + \frac{F_{bj}^U p_{b,c}^U}{F_b^L} (F_b^U - F_b) \\
&= p_{b,c}^U (F_b^U - F_b) (\frac{F_{bj}^U}{F_b^L} - \frac{F_{bj}}{F_b}) \geq 0.
\end{aligned}
$$

If $x_{bj}^U = 1$, then from Eq. (P-2-R2), Eq. (P^+-1), Eq. (P^+-2), and condition (b-4),

$$
\begin{aligned}
L &= -f_{b,c} - f_{b,c}^U \frac{F_{bj}}{F_b} + f_{b,c}^U + f_{bj,c} \\
&= -\sum_{j* \neq j} f_{bj*,c} - f_{b,c}^U \frac{F_{bj}}{F_b} + f_{b,c}^U \\
&\geq -\sum_{j* \neq j} F_{bj*} p_{b,c}^U - p_{b,c}^U F_b^U \frac{F_{bj}}{F_b} + p_{b,c}^U F_b^U \\
&= p_{b,c}^U (-F_b + F_{bj} - F_b^U \frac{F_{bj}}{F_b} + F_b^U) \\
&= p_{b,c}^U (F_b^U - F_b)(1 - \frac{F_{bj}}{F_b}) \geq 0
\end{aligned}
$$

Finally, we show that the point satisfies Eq. (SF-3-R1) - (SF-3-R4). To prove Eq. (SF-3-R1), we show that

$$
\begin{aligned}
L &= F_b^U x_{bj} + F_b x_{bj}^L - F_b^U x_{bj}^L - F_{bj} \\
&= F_b^U \frac{F_{bj}}{F_b} + F_b \frac{F_{bj}^L}{F_b^U} - F_b^U \frac{F_{bj}^L}{F_b^U} - F_{bj} \\
&= (F_b^U - F_b)(\frac{F_{bj}}{F_b} - \frac{F_{bj}^L}{F_b^U}) \geq 0
\end{aligned}
$$

For ((SF-3-R2),

$$
\begin{aligned}
L &= F_b^L x_{bj} + F_b x_{bj}^U - F_b^L x_{bj}^U - F_{bj} \\
&= F_b^L \frac{F_{bj}}{F_b} + F_b x_{bj}^U - F_b^L x_{bj}^U - F_{bj} \\
&= (F_b - F_b^L)(x_{bj}^U - \frac{F_{bj}}{F_b}) \\
&= (F_b - F_b^L)(\min\{\frac{F_{bj}^U}{F_b^L}, 1\} - \frac{F_{bj}}{F_b}) \geq 0
\end{aligned}
$$

For Eq. (SF-3-R3),

$$
\begin{aligned}
L &= F_{bj} - F_b^L x_{bj} - F_b x_{bj}^L + F_b^L x_{bj}^L \\
&= F_{bj} - F_b^L \frac{F_{bj}}{F_b} - F_b \frac{F_{bj}^L}{F_b^U} + F_b^L \frac{F_{bj}^L}{F_b^U} \\
&= (F_b - F_b^L)(\frac{F_{bj}}{F_b} - \frac{F_{bj}^L}{F_b^U}) \geq 0
\end{aligned}
$$

For Eq. (SF-3-R4),

$$
\begin{aligned}
L &= F_{bj} - F_b^U x_{bj} - F_b x_{bj}^U + F_b^U x_{bj}^U \\
&= F_{bj} - F_b^U \frac{F_{bj}}{F_b} - F_b x_{bj}^U + F_b^U x_{bj}^U \\
&= (F_b^U - F_b)(x_{bj}^U - \frac{F_{bj}}{F_b}) \\
&= (F_b^U - F_b)(\min\{\frac{F_{bj}^U}{F_b^L}, 1\} - \frac{F_{bj}}{F_b}) \geq 0
\end{aligned}
$$

Therefore, point $(x_{bj}, f_{b,c}, f_{bj,c}, F_b, F_{bj})$ is feasible for problem (SF-R). Since problem (SF-R) and problem (P$^+$-R) have the same objective function, point $(x_{bj}, f_{b,c}, f_{bj,c}, F_b, F_{bj})$ and $(f_{b,c}, f_{bj,c}, F_b, F_{bj}, p_{b,c})$ attain the same objective value. In other words, for any feasible point for problem (P$^+$-R), a feasible point for (SF-R) that attains the same objective function exists. Hence the optimal objective value of problem (P$^+$-R) cannot be less than that of (SF-R). □

Proposition 2 indicates that the P$^+$-formulation is as least as strong as the SF-formulation. The proof does not rely on the specific values of the variable bounds. At any node of the branch-and-bound tree, P$^+$-formulation would be better than a SF reformulation that has consistent bounds. Therefore, P$^+$-formulation is likely to have significant computational advantage over SF-formulation in a branch-and-bound framework.

P$^+$-formulation is an enhanced P-formulation because of the additional variables $f_{b,c}$, F_b and the relevant constraints. $f_{b,c}$, F_b are the sum of (component or total) flows going through pool b, which are often dispensable for establishing material balances in P-formulation. However, if we view the pool as a mixer connected with a splitter, $f_{b,c}$, F_b will naturally be included in P-formulation. As shown in Fig. 3, a mixer is a pool with multiple inflows and one outflow, while a splitter is a pool with one inflow and multiple outflows. When viewing a general pool as the two special pools, the component flows and the total flow between the two pools are $f_{b,c}$, F_b. In this case, the P-formulation is also the P$^+$-formulation, and according

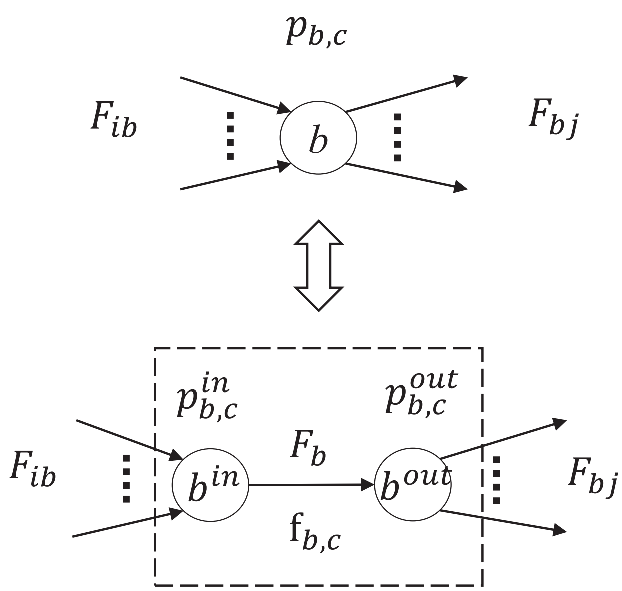


Figure 3: A Pool as a mixer connected with a splitter

to Proposition 2, it is at least as strong as SF-formulation. Many real-world GPP networks contain only mixers/splitters or only a few pools have multiple inflows and outflows, such as gasoline blending networks and natural gas transmission networks. So for these problems P-formulation is identical, or almost identical, to P$^+$-formulation, and therefore likely to be better than SF-formulation. On the other hand, when a mixer is directly connected with a splitter in a GPP network, viewing the two as a multi-input multi-output pool might not be a good idea although it would lead to a smaller problem size. In the computational study, we will see that the slight complexity introduced in the P$^+$-formulation usually leads to significant improvement in computational performance.

Strength of the Formulations at the Root Node

P$^+$-formulation usually exhibits better computational performance than P-formulation and SF-formulation, as will be seen in the computational study. However, the three formulations often achieve the same lower bound at the root gap. In this section we explain why they often have the same root gap.

Proposition 3. *Assume that at the root node* $x_{bj}^L = 0$ *and* $x_{bj}^U = 1$ *for problem (SF-R). If* *bound consistency conditions (b-1)-(b-4) hold, then the optimal objective value of Problem (P-R) is no less than that of Problem (SF-R).*

Proof. Let $(F_{bj}, f_{bj,c}, p_{b,c})$ be a feasible solution to problem (P-R). We construct a new point $(F_{bj}, f_{bj,c}, x_{bj}, F_b, f_{b,c})$ such that $x_{bj} = \frac{F_{bj}}{F_b}$ $(\forall b \in B, j \in J_b)$, $F_b = \sum_{j \in J_b} F_{bj}$ and $f_{b,c} = \sum_{j \in J_b} f_{bj,c}$ $(\forall b \in B)$. We prove the proposition by showing that the new point is feasible for problem (SF-R). Like in the proof for Proposition 2, it is easy to show that the point satisfies all constraints in (SF-R) except the convex envelopes. So it is left to show that the point satisfies the convex envelopes Eq. (SF-2-R1) - Eq. (SF-3-R4).

From Eq. (P-2-R2), $f_{bj,c} \leq F_{bj} p^U_{b,c}$, and from condition (b-4) it becomes

$$f_{bj,c} \leq F_{bj} \frac{f^U_{b,c}}{F^U_b} \leq f^U_{b,c} \frac{F_{bj}}{F_b} = f^U_{b,c} x_{bj}.$$

Note that $x^L_{bj} = 0$ at the root nodes, so we have proved Eq. (SF-2-R1). Similarly, Eq. (SF-2-R3) is satisfied because from Eq. (P-2-R3) and condition (b-3),

$$f_{bj,c} \geq F_{bj} \frac{f^L_{b,c}}{F^L_b} \geq f^L_{b,c} \frac{F_{bj}}{F_b} = f^L_{b,c} x_{bj}.$$

To prove (SF-2-R2), we need to show that

$$\begin{aligned} L &= f^L_{b,c} x_{bj} + f_{b,c} x^U_{bj} - f^L_{b,c} x^U_{bj} - f_{bj,c} \\ &= f^L_{b,c} \frac{F_{bj}}{F_b} + f_{b,c} - f^L_{b,c} - f_{bj,c} \\ &= \sum_{j' \neq j} f_{bj',c} + f^L_{b,c} \frac{F_{bj}}{F_b} - f^L_{b,c} \end{aligned}$$

is non-negative. From Eq. (P-2-R3) and condition (b-3), we get

$$\begin{aligned}
L &\geq \sum_{j'\neq j} F_{bj'} p^L_{b,c} + f^L_{b,c} \frac{F_{bj}}{F_b} - f^L_{b,c} \\
&= p^L_{b,c}(F_b - F_{bj} + F^L_b \frac{F_{bj}}{F_b} - F^L_b) \\
&= p^L_{b,c}(F_b - F^L_b)(1 - \frac{F_{bj}}{F_b}) \geq 0,
\end{aligned}$$

so Eq. (SF-2-R2) is satisfied. Similarly, to prove Eq. (SF-2-R4) we can show that, from Eq. (P-2-R2) and condition (b-4),

$$\begin{aligned}
L &= -f_{b,c} x^U_{bj} - f^U_{b,c} x_{bj} + f^U_{b,c} x^U_{b,c} + f_{bj,c} \\
&= -f_{b,c} - f^U_{b,c} \frac{F_{bj}}{F_b} + f^L_{b,c} + f_{bj,c} \\
&= -\sum_{j'\neq j} f_{bj',c} - f^U_{b,c} \frac{F_{bj}}{F_b} + f^U_{b,c} \\
&\geq -\sum_{j'\neq j} F_{bj'} p^U_{b,c} - f^U_{b,c} \frac{F_{bj}}{F_b} + f^U_{b,c} \\
&= p^U_{b,c}(-F_b + F_{bj} - F^U_b \frac{F_{bj}}{F_b} + F^U_b) \\
&= p^U_{b,c}(F^U_b - F_b)(1 - \frac{F_{bj}}{F_b}) \geq 0.
\end{aligned}$$

The proof for Eq. (SF-3-R1) and Eq. (SF-3-R4) follows a similar yet simpler procedure. So we only provide major steps here. Eq. (SF-3-R2) is satisfied because

$$F_{bj} \leq F_{bj} \frac{F^U_b}{F_b} = F^U_b x_{bj}.$$

Eq. (SF-3-R2) is satisfied because

$$F^L_b x_{bj} + F_b x^U_{bj} - F^L_b x^U_{bj} - F_{bj} = (F_b - F^L_b)(1 - \frac{F_{bj}}{F_b}) \geq 0.$$

Eq. (SF-3-R3) is satisfied because

$$F_{bj} \geq F_{bj}\frac{F_b^L}{F_b} = F_b^L x_{bj}.$$

Eq. (SF-3-R4) is satisfied because

$$-F_b^U x_{bj} - F_b x_{bj}^U + F_b^U x_{bj}^U + F_{bj} = (F_b^U - F_b)(1 - \frac{F_{bj}}{F_b}) \geq 0.$$

$\square$

Proposition 4. *The optimal objective value of Problem (P$^+$-R) is the same to that of Problem (P-R) if the following conditions hold:*

$$F_b^L \leq \sum_{j \in J_b} F_{bj}^L, \quad \forall b \in B, \tag{b-5}$$

$$F_b^U \geq \sum_{j \in J_b} F_{bj}^U, \quad \forall b \in B. \tag{b-6}$$

Proof. From Proposition 1, the optimal objective value of (P$^+$-R) is no less than that of (P-R). So we only need to prove the optimal objective value of (P-R) is no less than that of (P$^+$-R). To prove this, consider any feasible point of (P-R), say $(f_{bj,c}, F_{bj}, p_{b,c})$, and we show that we can always construct a feasible point of (P$^+$-R) that attains the same objective value. Let $f_{b,c} = \sum_{j \in J_b} f_{bj,c}$ and $F_b = \sum_{j \in J_B} F_{bj}$, and we can show that point $(f_{bj,c}, F_{bj}, p_{b,c}, f_{b,c}, F_b)$ is feasible for (P$^+$-R) as follows.

Obviously, the point satisfies the constraints that are also in the P-formulation. The point also satisfies Eq. (P$^+$-1) and Eq. (P$^+$-2) by construction. From Eq. (P-2-R1) and condition (b-6),

$$\begin{aligned} f_{b,c} &= \sum_{j\in J_b} f_{bj,c} \\ &\le \sum_{j\in J_b} F^U_{bj} p_{b,c} + \sum_{j\in J_b} F_{bj} p^L_{b,c} - \sum_{j\in J_b} F^U_{bj} p^L_{b,c} \\ &= \sum_{j\in J_b} F^U_{bj} (p_{b,c} - p^L_{b,c}) + F_b p^L_{b,c} \\ &\ge F^U_b (p_{b,c} - p^L_{b,c}) + F_b p^L_{b,c}, \end{aligned}$$

so the point satisfies Eq. (P^+-3-R1). Similarly, we can show that the point satisfies Eq. (P^+-3-R2) (because of (P-2-R2) and (b-5)), Eq. (P^+-3-R3) (because of (P-2-R3) and (b-5)), and Eq. (P^+-3-R4) (because of (P-2-R4) and (b-6)). This completes the proof. □

For many GPP problems, at the root node $F^L_b = F^L_{bj} = 0$, so condition (b-5) holds. In addition, if the capacities of pools are sufficiently large, then (b-6) also holds. Therefore, Proposition 4 explains why in many cases P^+-formulation and P-formulation have the same root gap.

Computational Study

The goal of the computational study is to compare the strength and the solution efficiency of P+, P and SF formulation. For convenience, we compare the strength of the formulations only through their root gaps. The root gap δ is defined as

$$\delta = \frac{|Obj_{gop} - Obj_{root}|}{|Obj_{gop}|} \times 100\%,$$

where Obj_{gop} is the optimal objective value of the optimization problem, and Obj_{root} is the optimal objective of the convex relaxation problem obtained at the root node of the branch-and-bound search. The solution efficiency is shown by the solution time or the optimality gap if the solver cannot converge to its global optimum within time limit.

The simulation was performed on a virtual machine with 3-core 3.40 GHz CPU, 4GB memory, and Ubuntu 16.04 operating system. All problems are formulated on GAMS 32.3.0 [21]. The NLP problems are solved by BARON 20.4.14. BARON is set to use CONOPT 4.22 as local NLP solver and CPLEX 12.10.0 as LP solver. The run time limit is set to be 1 hour, and the relative and absolute termination tolerance are set to be 10^{-2}.

We consider 23 instances as shown in Table 2, in which L1-L15 are from Alfaki and Haugland [14], X1 is Example 3 from Cheng et al. [22], X2 is modified from Case study B from Li et al. [23], and X3A-X5B are additional instances we have created. Network structures and parameters of all problem instances (except X1) can be found in the supplementary material. The computational results in Table 2 are organized as follows. The second column gives the objective value at the solution returned by BARON, either with guaranteed optimality or not (depending on whether the gap is closed at termination). The third to fifth column provide the root gaps returned by P^+-, P- and SF-formulations. The last three columns report the solution times for the three formulations. If an instance cannot be solved within the time limit, the optimality gap at termination is reported instead.

We have the following observations regarding the root gaps of the three formulations. First, P^+-formulation provides no worse root gap than P-formulation and SF-formulation, which is implied by the Proposition 1 and Proposition 2. Second, P-formulation provides no worse root gap than SF-formulation, which is implied by the Proposition 3 given that for all the problem instances the initial upper and lower bounds of split-fraction variables are 1 and 0. Third, for instances L2-L4, L6-L14 and X1, P^+ and P yield the same root gap. This is expected from Proposition 4, as these problems have sufficiently large pool capacities. On the other hand, for instances X3-A, X4-B and X5-B where conditions (b-5)-(b-6) are not satisfied, P^+ provides tighter relaxation than P-formulation at the root node.

We also have the following observations regarding the solution times. First, P^+-formulation is not always better than P-formulation, and vice versa. For instances where P^+-formulation provides better root gap (i.e., X3-A, X4-B, X5-B), P^+-formulation requires much less solu-

Table 2: Comparison of computational results of the three formulations

Instance	Obj	Root Gap δ			Sol. Time (s)/ Opt. Gap		
		P^+	P	SF	P^+	P	SF
L1	-42.58	0.98%	0.98%	1.96%	0.1	0.1	0.1
L2	-549.80	81.76%	81.76%	81.76%	7.63%	1955.8	38.60%
L3	-549.80	55.35%	55.35%	55.35%	1.21%	1776.1	32.08%
L4	-561.04	57.36%	57.36%	57.36%	36.45%	8.94%	36.45%
L5	-877.65	17.64%	17.64%	20.21%	856.7	699.7	14.52%
L6	-450.00	44.44%	44.44%	44.44%	0.4	1.0	4.7
L7	-3500.00	0.00%	0.00%	0.00%	0.1	0.1	0.1
L8	-1100.00	9.09%	9.09%	9.09%	8.33%	8.33%	8.33%
L9	-8.00	0.00%	0.00%	0.00%	0.9	0.9	0.5
L10	-8.00	0.00%	0.00%	0.00%	1.4	0.2	4.5
L11	-8.00	0.00%	0.00%	0.00%	0.1	0.1	1.4
L12	-400.00	50.00%	50.00%	50.00%	0.5	1.1	7.8
L13	-600.00	100.00%	100.00%	100.00%	0.6	0.8	24.8
L14	-750.00	16.67%	16.67%	16.67%	0.3	0.5	4.00
L15	-439182.59	44.17%	44.17%	44.17%	4.2	5.8	9.6
X1	-22.99	5.56%	5.56%	5.56%	0.9	0.7	1.4
X2	-8540.43	32.80%	32.80%	36.12%	14.3	3.6	23.14%
X3-A	-1518.67	25.73%	26.95%	30.91%	297.8	666.6	17.40%
X3-B	-1694.07	17.36%	17.36%	17.36%	26.3	873.1	275.6
X4-A	-704.98	18.18%	18.18%	18.18%	647.9	864.4	2212.0
X4-B	-472.35	40.04%	40.33%	40.33%	0.5	7.3	7.1
X5-A	-10151.30	25.62%	25.62%	25.69%	658.4	1023.9	851.7
X5-B	-7709.53	26.22%	29.87%	31.12%	106.9	2101.6	1636.1

tion time than P-formulation. If P^+ doesn't provide better root gap (i.e., same root gap as by P-formulation), then the solution time required by P^+-formulation may be much worse (L2-L4, X2), slightly worse (L5), similar, or still much better (X3-B, X4-A, X5-A) than the solution time required by P-formulation. The under-performance of P^+-formulation is not unexpected, because P^+-formulation contains more bilinear terms than P-formulation and the computational performance reflects whether the advantage of tighter relaxation outweighs the disadvantage of more nonconvex terms. Second, comparing P^+-formulation and SF-formulation, we see that P^+ always provides a better solution time. Finally, comparing P-formulation and SF-formulation, we see that for P-formulation provides better solution time than SF-formulation in all instances except X3-B, X4-B and X5-B. SF-formulation performs much better than P-formulation for instances X3-B, X4-B and X5-B in spite of worse root gaps, and the reason may be that SF-formulation provides tight relaxations at some child nodes of the branch-and-bound tree. This indicates that when comparing P-formulation and SF-formulation, a tighter root gap does not ensure a better solution time. For instances that have sufficiently large pool capacities and $F_{ij}^L = 0$ (i.e., L2-L14, X1, X2), P-formulation provides both a tighter root gap and a better solution time than SF-formulation. An explanation for this is that, for these instances the conditions (b-5)-(b-6) hold at not only the root node but also many child nodes in the branch-and-bound tree, so according to Proposition 4 P-formulation is as strong as P^+-formulation and therefore often stronger than the SF-formulation at these nodes.

Concluding Remarks

In this paper, we have developed a new GPP formulation, P^+-formulation, by adding to the classical P-formulation additional variables and bilinear terms. We have proved that P^+-formulation is at least as strong as P-formulation (Proposition 1) or SF-formulation (Proposition 2), under mild bound consistency conditions. While the P^+-formulation achieves tighter relaxation at the cost of more nonconvex constraints, it can provide better computational performance (especially for difficult problems) as shown in the computational study. Moreover, when the pooling network comprises only mixers and splitters, P^+-formulation is equivalent to P-formulation, so in this case P-formulation is at least as strong as SF-formulation. With additional conditions on the variable bounds, we can prove that P-formulation is at least as strong as SF-formulation and/or P^+-formulation (Propositions 3 and 4). These conditions usually hold at the root node of a branch-and-bound search, which explains why the root gap of P-formulation is often no worse than that of SF-formulation. If these conditions hold also at many child nodes of the branch-and-bound tree, then P-formulation can result in less solution time than SF-formulation. In the computational study, P-formulation generally results in better computational time than SF-formulation, and one possible reason is that the conditions (such as (b-5) and (b-6)) allowing P-formulation to be stronger than SF-formulation hold at many branch-and-bound nodes.

Acknowledgement

The authors are grateful to the financial support from Natural Sciences and Engineering Research Council of Canada through the Discovery Grant RGPIN-2019-05217 and the Collaborative Research and Development Grant CRDPJ 485798-15.

Literature Cited

References

[1] C. Audet, J. Brimberg, P. Hansen, S. L. Digabel, N. Mladenović, Pooling problem: Alternate formulations and solution methods, Management Science 50 (6) (2004) 761–776.

[2] P. M. Castro, I. E. Grossmann, Global optimal scheduling of crude oil blending operations with RTN continuous-time and multiparametric disaggregation, Industrial & Engineering Chemistry Research 53 (39) (2014) 15127–15145.

[3] R. Karuppiah, K. C. Furman, I. E. Grossmann, Global optimization for scheduling refinery crude oil operations, Computers & Chemical Engineering 32 (11) (2008) 2745–2766.

[4] X. Li, E. Armagan, A. Tomasgard, P. I. Barton, Stochastic pooling problem for natural gas production network design and operation under uncertainty, AIChE Journal 57 (8) (2011) 2120–2135.

[5] S. Doyle, R. Smith, Targeting water reuse with multiple contaminants, Process Safety and Environmental Protection 75 (3) (1997) 181–189.

[6] P. M. Castro, H. A. Matos, A. Q. Novais, An efficient heuristic procedure for the optimal design of wastewater treatment systems, Resources, Conservation and Recycling 50 (2) (2007) 158–185.

[7] R. Misener, J. P. Thompson, C. A. Floudas, Apogee: Global optimization of standard, generalized, and extended pooling problems via linear and logarithmic partitioning schemes, Computers & Chemical Engineering 35 (5) (2011) 876–892.

[8] N. Takama, T. Kuriyama, K. Shiroko, T. Umeda, Optimal water allocation in a petroleum refinery, Computers & Chemical Engineering 4 (4) (1980) 251–258.

[9] J. Lou, Z. Liao, B. Jiang, J. Wang, Y. Yang, Robust optimization of hydrogen network, International Journal of Hydrogen Energy 39 (3) (2014) 1210–1219.

[10] A. Fonseca, V. Sá, H. Bento, M. L. Tavares, G. Pinto, L. A. Gomes, Hydrogen distribution network optimization: a refinery case study, Journal of Cleaner Production 16 (16) (2008) 1755–1763.

[11] C. A. Haverly, Studies of the behavior of recursion for the pooling problem, ACM Sigmap Bulletin (25) (1978) 19–28.

[12] I. Quesada, I. E. Grossmann, Global optimization of bilinear process networks with multicomponent flows, Computers & Chemical Engineering 19 (12) (1995) 1219–1242.

[13] M. Alfaki, D. Haugland, Strong formulations for the pooling problem, Journal of Global Optimization 56 (3) (2013) 897–916.

[14] M. Alfaki, D. Haugland, A multi-commodity flow formulation for the generalized pooling problem, Journal of Global Optimization 56 (3) (2013) 917.

[15] N. Boland, T. Kalinowski, F. Rigterink, New multi-commodity flow formulations for the pooling problem, Journal of Global Optimization 66 (4) (2016) 669–710.

[16] I. Lotero, F. Trespalacios, I. E. Grossmann, D. J. Papageorgiou, M.-S. Cheon, An MILP-MINLP decomposition method for the global optimization of a source based model of the multiperiod blending problem, Computers & Chemical Engineering 87 (2016) 13–35.

[17] X. Cheng, X. Li, A multi-commodity flow formulation for the optimal design of wastewater treatment networks, Computers & Chemical Engineering 134 (2020) 106681.

[18] R. Karuppiah, I. E. Grossmann, Global optimization for the synthesis of integrated water systems in chemical processes, Computers & Chemical Engineering 30 (4) (2006) 650–673.

[19] N. V. Sahinidis, Baron: A general purpose global optimization software package, Journal of Global Optimization 8 (2) (1996) 201–205.

[20] P. M. Castro, New MINLP formulation for the multiperiod pooling problem, AIChE Journal 61 (11) (2015) 3728–3738.

[21] E. Bringas, R. Karuppiah, M. F. San Román, I. Ortiz, I. E. Grossmann, Optimal groundwater remediation network design using selective membranes, Industrial & Engineering Chemistry Research 46 (17) (2007) 5555–5569.

[22] X. Cheng, K. Tang, X. Li, New multi-commodity flow formulations for the pooling problem, IFAC-PapersOnLine 51 (18) (2018) 162–167.

[23] X. Li, E. Armagan, A. Tomasgard, P. I. Barton, Long-term planning of natural gas production systems via a stochastic pooling problem, in: American Control Conference (ACC), 2010, IEEE, 2010, pp. 429–435.