



*Research article*

## **Nonlinear modeling and interior point algorithm for the material flow optimization in petroleum refinery**

**Fenglian Dong<sup>1,2</sup>, Dongdong Ge<sup>3,4,\*</sup>, Lei Yang<sup>1,2</sup>, Zhiwei Wei<sup>1,2</sup>, Sichen Guo<sup>4</sup> and Hekai Xu<sup>3</sup>**

<sup>1</sup> Petrochina Planning and Engineering Institute, Beijing 100086, China

<sup>2</sup> CNPC Laboratory of Oil & Gas Business Chain Optimization, Beijing 100086, China

<sup>3</sup> Cardinal Operations, Beijing 100102, China

<sup>4</sup> Shanghai University of Finance and Economics, Shanghai 200082, China

\* **Correspondence:** Email: [ge.dongdong@mail.shufe.edu.cn](mailto:ge.dongdong@mail.shufe.edu.cn); Tel: 15801738716.

**Abstract:** This paper established a mathematical model with nonconvex bilinear terms. It formulated the complex material flow in the petroleum refinery scenario based on the concept of the “P model”. The mathematical model described the nonlinear constraints such as linear and nonlinear mass and volume intersection flow blending of crude and middle material physical properties. Additionally, it described the complex inflow and outflow in secondary devices as nonlinear constraints such as delta-base structure and physical property transfer. It is highly difficult to determine the direction and quantity of each material in the network of refineries. An improved interior point algorithm with an initial point strategy was proposed to find a high-quality feasible solution in a short time. The real instances from the petroleum refinery were employed to compare and analyze the solutions from the improved algorithm and commercial solver. The experimental results show that the proposed algorithm framework can balance the solution quality and computational efficiency and perform well in different scenarios of refinery material flow networks.

**Keywords:** material flow; process control; intersection mixture; interior point method; nonconvex optimization

---

### **1. Introduction**

The pooling problem, also known as the mixed-flow problem, is a classic revenue maximization problem involving supply, storage, demand, and specified flow mixture constraints and usually appears in petroleum refinery and wastewater treatment scenarios [1]. Crude oils flowing into the petroleum refinery network are refined and processed to produce gasoline, kerosene, diesel, and other liquid fuels, hundreds of kinds of lubricating oil, paraffin, asphalt, and other refinery products, and tens of

thousands of synthetic resin, synthetic rubber, synthetic fiber, and other chemical products. Various physical and chemical processes, including crude oil separation, light oil, and gas processing, should be performed for the crude oil processing of various products. Each of these processes is relatively independent in constructing a refinery production unit network, such as the atmospheric-vacuum distillation unit and catalytic cracking, catalytic reforming, hydrocracking, coking, and other secondary devices. Refinery optimization refers to the overall optimization of the whole process of refinery enterprises, including raw material procurement, production, processing, and product delivery, for maximizing the economic benefits of enterprises while satisfying various production constraints. In terms of the pooling problem, the mixture of quality attributes often leads to nonconvex bilinear constraints, causing significant decision-making difficulties. Therefore, the research mainly focuses on the formulation and solution of nonlinear models.

There are three classical modeling methods for the pooling problem: the “P(proportional) model” proposed by Haverly [2], the “Q(quantity) model” improved by Ben-Tal [3], and the “PQ(proportional and quantity) model” further modified by Quesada and Grossmann [4]. The “Q model” takes actual material inflow and product outflow of the processing system as the decision variables. The “P model” replaces the amount of material inflow and outflow by the transformation proportion relationship between all processes to obtain the equivalent form of the “Q model” with fewer variables. Based on the “Q model”, the “PQ model” adds constraints of proportional variables and product output to obtain a tighter convex envelope relaxation model [5].

The global optimization algorithms for solving the pooling problem in academia and industry are listed as follows: the nonlinear programming method based on Lagrange relaxation proposed by Ben-Tal et al. [3,6], the branch-and-bound method based on convex relaxation proposed by Zamora and Grossmann [7], the branch-and-bound method for convex relaxation with enhanced lower bound based on discretization proposed by Ruiz and Grossmann [8], reformulation-linearization-technique(RLT) nonlinear method proposed by Meyer and Floudas [9], and external approximation nonlinear method proposed by Bergamini [10]. Although these global optimization algorithms can theoretically guarantee the optimality of the results, their solution’s efficiency cannot meet the practical operation requirements for complex problems.

Compared with the global optimization algorithms, several local optimization algorithms have been proposed to solve the pooling problem. These algorithms include the continuous linear programming solution (SLP) designed for refinery scenarios [11–14], the nonmonotonic continuous linear programming (NMSLP) method [15], the nonlinear term discretization method [16–18], the Benders decomposition method proposed by Floudas and Aggarwal [19], the distributed recursive method for solving pooling problems in commercial software [20–22], and an interior-point method for solving nonlinear problems in commercial software [23,24]. Among them, the Benders decomposition and interior point methods fail to converge even for a long time. The continuous linear programming and distributed recursive methods easily fall into the local optimum.

This paper proposes a formulation of a nonlinear model suitable for refinery mixed flow based on the modeling of the “P model” in the pooling problem combined with complex real business scenarios. Furthermore, an improved interior point method algorithm is presented, which can find a high-quality feasible solution quickly. This algorithm has contributed to the generalized mixed-flow processing problems considering the complex nonlinear constraints in the refinery process and the generality of the interior point method.

The main contributions of this paper are:

- 1) A nonconvex and nonlinear model is established considering complex processes, such as mass linear, mass nonlinear, volume linear, and volume nonlinear intersection flow blending of crude and

middle material physical properties with delta-base structure and physical property transfer.

2) An improved interior point method is presented to solve the nonconvex and nonlinear model in the allowable time.

3) Numerical experiment shows that the result of the improved interior point method is superior to that of the commercial solver.

## 2. Materials and methods

### 2.1. Material flow in crude oil refinery

The main difficulty of production optimization in the crude oil refinery mainly lies in the significant amount of different raw materials and final products and the complex mutual supply relationship between various devices. Also, to meet different requirements for different final products, it is necessary to perform secondary processing of raw materials and intermediate components for selecting different processing options. It is challenging to construct a detailed optimization model for the above system. Therefore, this paper simplifies the entire industrial production process into five parts: raw material procurement, atmospheric and vacuum unit processing, secondary unit processing, blending, and product sales. By refining the transformation relationships within and between these five parts, the entire system can be well described through a relatively simple and neat model. It is noticeable that no necessary procedure is omitted. The performed simplifications include a summary and refinement of the actual process. The material quantity of inflow and outflow of each device is the main decision variable in the optimization model. These decision variables should meet specific equality or inequality relations caused by production requirements. Meanwhile, there are limited materials processed in each device and system due to the device capacity, production technology, and product delivery indicators. In addition, the inflow and outflow of devices should meet specific constraints caused by processing capacity and physical indicators. Finally, raw materials and final products should meet the business operation constraints, such as purchasing, sales, and warehouse storage. Even after some simplification, the great number of raw materials and final products, the complex processing and transformation relations, and the embedded scheme selection will still lead to a nonconvex and nonlinear model, resulting in significant difficulties for the model's formulation and solution. This paper will give the model formulation in detail in Sections 2.2.1–2.2.6 and the solution algorithm of the optimization model in Section 2.3.

### 2.2. Modeling the pooling problem

Symbols	Definition
$E$	set of device types
$E^s$	set of devices in link $S$
$I^M$	$I^M = (e, m_{in})$ is the set of materials flowing into the device
$I^G$	$I^G = (e, g_{in})$ is the set of the utility system flowing into the device
$G$	set of the utility system types
$H$	set of unit processing capacity
$H^s$	set of unit processing capacity in link $S$
$M$	set of material types
$M^s$	set of material types that should be processed in link $S$

*Continued on next page*

Symbols	Definition
$\mathcal{O}^M$	$\mathcal{O}^M = (e, m_{out})$ is the set of device output material
$\mathcal{O}^G$	$\mathcal{O}^G = (e, g_{out})$ is the set of the device output utility system
$X_m^s$	The amount of material $m$ in link $S$ (104 ton), $s \in \{gm, xs\}$
$\overline{X_m^s}, \overline{X_g^s}$	The upper limit of material $m$ (utility system $g$ ) in link $S$ (104 tons), $s \in \{gm, xs\}$
$\underline{X_m^s}, \underline{X_g^s}$	The lower limit of material $m$ (utility system $g$ ) in link $S$ (104 tons), $s \in \{gm, xs\}$
$X_{e, m_{in}}^s$	The amount of material $m$ flowing into device $e$ in link $S$ is the raw material consumption of device $e$ in process $S$ (104 tons), $s \in \{cy, ec, th\}$
$X_{e, m_{out}}^s$	The amount of material $m$ outflowing into device $e$ in link $S$ is the raw material consumption of device $e$ in process $S$ (104 tons), $s \in \{cy, ec, th\}$
$X_{e, m_{in}, p}^s$	The amount of material $m$ in link $S$ flowing into device $e$ by the processing scheme $P$ , $s \in \{cy, ec, th\}$
$X_{e, g_{in}}^s$	The amount of the utility system $g$ of link $S$ flowing into device $e$ is the utility system consumption of device $e$ of process $S$ (104 tons), $s \in \{cy, ec, th\}$
$X_{e, g_{out}}^s$	The amount of the utility system $g$ of link $S$ outflowing into device $e$ is the utility system consumption of device $e$ of process $S$ (104 tons), $s \in \{cy, ec, th\}$
$X_{e, m_{in}, m_{out}}^s$	The amount of product $m_{out}$ processed from raw material $m_{in}$ in device $e$ of link $S$ (104 tons), $s \in \{cy, ec, th\}$
$Y_{m_{in}, q}^s$	$q$ value of the quality attribute flowing into material $m$ in link $S$ , $s \in \{cy, ec, th\}$
$Y_{m_{out}, q}^s$	$q$ value of the quality attribute outflowing material $m$ in link $S$ , $s \in \{cy, ec, th\}$
$Y_{e, m_{in}, q}^s$	$q$ value of the quality attribute of material $m$ flowing into device $e$
$\overline{Y_{m, q}^s}, \underline{Y_{m, q}^s}$	upper and lower limits of the $q$ value of the quality attribute of material $m$ in link $S$ . $s \in \{cy, ec, th\}$
<b>superscript</b>	
$S$	$s \in \{gm, cy, ec, th, xs\}$ represents raw material procurement, atmospheric-vacuum distillation unit processing, secondary device processing, blending, and product sales.
<b>subscript</b>	
$g$	Consumption of the utility system
$m$	amount of material used

In the model, the material set is  $M$ , the utility system set is  $G$ , and the device set is  $E$ . The superscript  $s \in \{gm, cy, ec, th, xs\}$  can identify the corresponding production parts, representing raw material procurement, atmospheric-vacuum distillation unit processing, secondary device processing, blending, and product sales, respectively. For example,  $M^s$  represents the materials that should be processed in production part  $S$  and  $E^s$  represents the devices employed in production part  $s$ . In this study, the consumption pairs of the device, material, and utility system comprise set I and the device's output pairs. Material and utility systems consist of sets  $O$ ,  $I^M(O^M)$ , and  $I^G(O^G)$  as the consumption (output) sets of devices and materials and utility systems, respectively. The quantity of some materials is denoted as  $X$ .  $X_{e, m_{in}}^s$  ( $X_{e, g_{in}}^s$ ) and  $X_{e, m_{out}}^s$  ( $X_{e, g_{out}}^s$ ) represent the inflow and outflow of one material (utility system) in device  $e$  in the production part  $S$ , respectively. In the model, subscripts *in* and *out* represent the inflow and outflow properties of the material or utility system.

2.2.1. Objective function

The business requirement of refinery optimization implies optimizing the whole process from raw material procurement, production, and processing to product delivery to maximize the economic benefits of enterprises. Therefore, the following objective function is considered for the model to maximize profits:

$$\max \sum_{m \in M^{xs}} p_m X_m^{xs} + \sum_{g \in G^{xs}} p_g X_g^{xs} - \sum_{m \in M^{gm}} c_m X_m^{gm} - \sum_{g \in G^{gm}} c_g X_g^{gm} \tag{1}$$

where  $p_m$  is the selling price of material  $m$ ,  $p_g$  is the selling price of utility system  $g$ ,  $c_m$  is the purchase price of raw material  $m$ , and  $c_g$  is the purchase price of utility system  $g$ .

2.2.2. Material balance limit

Any material must meet the material quantity balance constraint across the whole plant; that is, the sum of its purchased quantity and its quantity produced by all devices should equal the sum of the sales quantity of the material and the quantity of the material consumed by all devices:

$$X_m^{gm} + \sum_{s \in \{cy, ec, th\}} \sum_{e \in E^s} X_{e, m_{out}}^s = X_m^{xs} + \sum_{s \in \{cy, ec, th\}} \sum_{e \in E^s} X_{e, m_{in}}^s \tag{2}$$

The atmospheric-vacuum distillation unit is the leading device in refinery enterprises. With crude oil as the main raw material, the output of each side of the atmospheric vacuum device is determined by the amount of processed crude oil and the yield of each fraction of crude oil after distillation.  $\alpha_{e, m_{in}, m_{out}}^{cy}$  is the yield of the fraction  $m_{out}$  corresponding to the crude oil  $m_{in} \in M^{cy}$ . For each atmospheric-vacuum distillation unit, the amount of material in and out should meet the following requirements:

$$\begin{aligned} X_{e, m_{out}}^{cy} &= \sum_{m_{in} \in M^{cy}} \alpha_{e, m_{in}, m_{out}}^{cy} X_{e, m_{in}}^{cy} \\ &:= \sum_{m_{in} \in M^{cy}} X_{e, m_{in}, m_{out}}^{cy}, \forall e \in E^{cy}, (e, m_{in}) \in I^M, (e, m_{out}) \in O^M \end{aligned} \tag{3}$$

The secondary unit is a general term for other processing units in refinery enterprises except for the atmospheric-vacuum distillation unit, including residual oil hydrogenation, catalytic cracking, catalytic reforming, hydrocracking, delayed coking, and other types. A set of secondary devices can have many production schemes, where the production process and the nature of raw materials determine each production scheme of raw material consumption and product yield. The delta-base structure of the secondary device should be established to describe the above process accurately [25,26].

$$X_{e, p}^{ec} = \sum_{m_{in} \in M^{ec}} X_{e, m_{in}, p}^{ec}, \forall e \in E^{ec}, (e, m_{in}) \in I^M \tag{4}$$

$$X_{e, m_{out}}^{ec} = \sum_{p \in P} \eta_{e, p, m_{out}} X_{e, p}^{ec} + \sum_{(p, d, m_{in}, q)} \gamma_{e, p, d, m_{in}} (Y_{m_{in}, q} - \delta_{e, p, d, m_{in}, q}) X_{e, p}^{ec}, \forall (e, m_{out}) \in O^M \tag{5}$$

where  $P$  is the set of alternative processing options,  $q$  is the quality attribute,  $d$  is the delta-base,  $\eta_{e, p, m_{out}}$  is the benchmark yield of material  $m$  processed by scheme  $P$  of refinery unit  $e$ ,

$\delta_{e,p,d,m_{in},q}$  is the reference value of the quality attribute of feed, and  $\gamma_{e,p,d,m_{in}}$  is the change rate of the product yield with the feed's quality attribute. The above constraints indicate that the product yield is based on the base and can change with the feed's quality attribute.

The feed of the atmospheric-vacuum distillation unit and the secondary device is usually mixed with multiple strands of materials with different properties. Limited by the production process, the properties of mixed raw materials should meet upper and lower ranges. Additionally, gasoline, diesel, and other final products produced by refinery enterprises are usually mixed with various components. The mixed product properties should meet specific upper and lower limits to meet national or industrial standards. In this paper, the above mixing processes are collectively referred to as blending and are characterized by a virtual device called blending pools. For each blending pool, the amount of material in and out should meet the following constraints:

$$X_{e,m_{out}}^{th} = \sum_{m_{in}} X_{e,m_{in}}^{th}, \forall e \in E^{th} \quad (6)$$

### 2.2.3. Processing capacity of the device

The processing capacity of the device defines the maximum and minimum processing capacities of the production device. The device is not allowed to operate beyond the upper or lower capacities in the actual production process to ensure safe production.

In this paper, the processing capacity set of the device is denoted by  $H$  and  $H^s$  is the processing capacity set of all devices in link  $S$ . The upper and lower limits of processing capacity  $h$  of device  $e$  are denoted as  $\overline{J}_h^s, \underline{J}_h^s$ . Generally, the consumption of primary raw materials characterizes the processing capacity of a device, then for each device, the processing capacity required for processing all major raw materials should not exceed the upper and lower limits; that is, it must meet the following constraints:

$$\overline{J}_h^s \geq \sum_{m_{in}} X_{e,m_{in}}^s \geq \underline{J}_h^s, \forall s \in \{cy, ec\}, \forall h \in H^s, e \in E^s, (e, m_{in}) \in I^M \quad (7)$$

### 2.2.4. Physical properties limit

Since the quality attribute calculation between the lateral line and the blending discharge is similar, the blending equation is adopted to characterize the quality attribute calculation. Specifically, for the atmospheric-vacuum distillation unit, the quality attribute of one lateral line is equal to the weighted average of the quality attribute of the lateral line produced by each processed crude oil. For the blending pooling, the quality attributes of the discharge are equal to the weighted average of the quality attributes of each feed.

In this paper,  $Q$  is the set of quality attributes and  $Y_{m,q}^s$  is the  $q$  value of the quality attribute of material  $m$  in link  $S$ . The set of quality attributes satisfying the blending relation of the weight is denoted as  $Q^{weight}$  and the set of quality attributes satisfying the volume blending relation is  $Q^{volume}$ . The quality attribute calculation in the blending process should satisfy the following equality constraints:

$$Y_{m_{out},q}^s \sum_{e \in E^s} X_{e,m_{out}}^s = \sum_{(e,m_{in})} X_{e,m_{in},m_{out}}^s \Gamma_q^s (Y_{e,m_{in},q}), \forall q \in Q^{weight} \quad (8)$$

$$Y_{m_{out},q}^s \sum_{e \in E^s} Z_{e,m_{out}}^s = \sum_{(e,m_{in})} Z_{e,m_{in},m_{out}}^s \Gamma_q^s(Y_{e,m_{in},q}), \forall q \in Q^{volume} \quad (9)$$

$$Z_{e,m_{out}} Y_{e,m_{out},\rho} = X_{e,m_{out}}^s \quad (10)$$

$$Z_{e,m_{in},m_{out}} Y_{e,m_{in},\rho} = X_{e,m_{in},m_{out}}^s \quad (11)$$

where  $Y_{e,m,\rho}$  is the material density. The introduction of variable volume  $Z$  instead of a fractional expression (volume= weight/density) reduces the nonlinear constraints.  $\Gamma_q^s(Y)$  is the  $q$  value of the quality attribute after link  $s$ . If the quality attributes conform to a linear blending law,  $\Gamma_q^s(Y) = Y$ ; otherwise, this function is determined by the specific physical law.

According to the requirements of the factory standards or the device's tolerance to the raw materials, the physical properties of the blended materials should meet specific upper and lower limits.  $\overline{Y}_{m,q}^s$  and  $\underline{Y}_{m,q}^s$  are the upper and lower limits of the  $q$  value of the quality attribute of material  $m$  in link  $s$ . The material's physical value should meet the following constraints:

$$\underline{Y}_{m_{out},q}^s \leq Y_{m_{out},q}^s \leq \overline{Y}_{m_{out},q}^s, \forall s \in \{cy, th\}, e \in E^s, \forall q \in Q \quad (12)$$

The nature of the secondary plant discharge, the primary feed source for the blending pool, usually varies with the processed raw material nature. It is necessary to establish the quality attribute transfer structure of the secondary device to describe the above process accurately [25,26]:

$$Y_{m_{out},q}^{ec} = \sum_{(e,m_{in})} a_{e,m_{in},m_{out},q} Y_{m_{in},q}^{ec} + b_{e,m_{in},m_{out},q} \quad (13)$$

where  $a_{e,m_{in},m_{out},q}$  and  $b_{e,m_{in},m_{out},q}$  are the transfer coefficients of the transfer equation of the quality attribute  $q$  between the feed  $m_{in}$  and the discharge  $m_{out}$  of secondary device  $e$ . Physical properties in secondary devices are not subject to upper and lower limits.

### 2.2.5. Restrictions on the utility system

The utility system is a general term for auxiliary facilities to maintain the normal operation of refinery production and processing equipment, also known as supporting works, usually covering the water supply and drainage system, power supply system, compressed air system, steam power system, and wind power system. The production and consumption of the utility system occupy the main variable processing cost of the refinery, which should be modeled and optimized. To reflect the production and marketing process of the refinery utility system, the structures of procurement, utility sales and production, and consumption of the utility system are established in the optimization model of the refinery planning. Common utility systems include fresh water, recycled water, desalted water, deoxygenated water, power, steam, industrial air, instrument air, nitrogen, fuel gas, and fuel oil. In this scenario, the utility system is produced and consumed only in the atmospheric-vacuum distillation unit and secondary plant. The output or consumption of the utility system and the material consumption of the process are expressed with linear relations:

$$X_{e,g}^{cy} = \sum_{m_{in} \in M^{cy}} X_{e,m_{in}}^{cy} \sigma_{e,g}^{cy}, \forall e \in E^{cy}, (e, m_{in}) \in I^{cy}, g \in G \quad (14)$$

$$X_{e,g}^{ec} = \sum_{m_{in} \in M^{cy}} X_{e,m_{in}}^{ec} \sigma_{e,g}^{ec}, \quad \forall e \in E^{ec}, (e, m_{in}) \in I^{ec}, g \in G \quad (15)$$

where  $\sigma$  represents the correspondence between the produced or consumed utility system and the number of processed materials.

The utility system of the whole plant should balance the output and consumption:

$$\sum_{e|(e,g_{in}) \in I^G} X_{e,g_{in}}^s = \sum_{e|(e,g_{out}) \in O^G} X_{e,g_{out}}^s \quad (16)$$

### 2.2.6. Operational limits

In practice, raw material procurement and finished product sales should be within a given upper and lower range:

$$\bar{X}_m^s \geq X_m^s \geq \underline{X}_m^s, \quad \forall m \in M^s, s \in \{gm, xs\} \quad (17)$$

The set of the utility system to be procured is  $G^{gm}$  and the set of the utility system available for sale is  $G^{xs}$ . The procurement and sales volume of the utility system should meet the upper and lower limits:

$$\bar{X}_g^s \geq \sum_{e \in E^{gy}} X_{e,g}^{cy} + \sum_{e \in E^{gy}} X_{e,g}^{ec} \geq \underline{X}_g^s, \quad \forall g \in G^s, s \in \{gm, xs\} \quad (18)$$

Accordingly, the optimization model for this problem is established. The solution algorithm is presented in the next section.

### 2.3. Algorithm design

To facilitate the discussion of algorithm design, the optimization model is summarized in the following form:

$$\begin{aligned} (P) \quad & \max_X \quad aX + b \\ \text{s.t.} \quad & f_1(X) = 0 \\ & f_2(X) \leq 0 \\ & h_1(Y) = 0 \\ & h_2(Y) \leq 0 \\ & g_1(X, Y) = 0 \\ & X \in R^+, Y \in R \end{aligned}$$

where  $a$  and  $b$  are parameters in Eq (1); the decision variable  $X$  corresponds to  $X_m^s$ ,  $X_{e,m}^s$ ,  $X_{e,m,p}^s$ ,  $X_g^s$ ,  $X_{e,g}^s$  in the model; the decision variable  $Y$  corresponds to  $Y_{m,q}^s$  and  $Y_{e,m,q}^s$  in the model;  $f_1$  is a linear equality constraint on  $X$  corresponding to Eqs (2)–(4), (6) and (14)–(16);  $f_2$  is a linear inequality constraint on  $X$  corresponding to Eqs (7), (17) and (18);  $h_1$  is a linear equality constraint on  $Y$  corresponding to Eq (13);  $h_2$  is a linear inequality constraint on  $Y$  corresponding to Eq (12); and  $g_1$  is a nonlinear equality constraint on  $X$  and  $Y$  corresponding to Eqs (5) and (8)–(11).



The optimization model is nonconvex and nonlinear because the chemical reactions can be generally described with a nonlinear model in actual industrial production, such as the delta-base structure in secondary devices and the quality attribute calculation of blending pools. The above problem coupled with the huge number of raw materials and final products makes it challenging to solve the model. Although many commercial or open-source solvers provide nonconvex optimization model solving, their solving efficiency is insufficient to deal with large industrial optimization problems. Therefore, to reduce the solving time of the model, a high-quality initial solution-finding method is proposed based on the problem structure and actual business logic, then this method is combined with the classical interior point method to solve the problem. The interior point method has a guarantee of convergence efficiency [27,28], especially if the initial solution is good.

The similarity between the initial and optimal solutions can significantly influence the efficiency of the interior point method. This paper generates the initial solution from two linear sub-models, whose solutions are combined to form the initial solution. The proposed algorithm consists of three parts: the material (linear model) for solving the initial value corresponding to model ( $P_2$ ), the solution of initial values of quality attributes (linear model) corresponding to questions ( $P_3$ ), and the entire model corresponding to model ( $P$ ). Among them, the model ( $P_2$ ) ignores the quality attribute based on the model  $P$ . Based on constraints such as the raw material purchasing prices, selling price of the product, upper and lower limits of purchase quantity, and upper and lower limits of sales, material balance, the raw material purchasing amount to maximize revenue, product sales, and the flow distribution between devices are determined. Model ( $P_3$ ) relaxes the nonlinear constraints based on model ( $P$ ), and to avoid obtaining solutions that are grossly infeasible, we refer to the idea of penalty functions [29]. Additionally, based on the material quantity calculated in the ( $P_2$ ) model, the quality attribute value that meets the quality attribute constraints is calculated. Finally, the model  $P$  takes the solutions of ( $P_2$ ) and ( $P_3$ ) as the initial points and employs the interior point method. The ( $P_2$ ) and ( $P_3$ ) models are described as follows:

$$\begin{aligned}
 (P_2) \quad & \max_X \quad aX + b \\
 \text{s.t.} \quad & f_1(X) = 0 \\
 & f_2(X) \leq 0 \\
 & X \in R^+
 \end{aligned}$$

$$\begin{aligned}
 (P_3) \quad & \min_{Y, SL_1^+, SL_1^-} \quad SL_1^+ + SL_1^- \\
 \text{s.t.} \quad & h_1(Y) = 0 \\
 & h_2(Y) \leq 0 \\
 & g_1(\hat{X}, Y) + SL_1^+ - SL_1^- = 0 \\
 & SL_1^+, SL_1^- \in R^+
 \end{aligned}$$

where ( $P_2$ ) is the initial model  $P$  without quality attribute variables and constraints;  $P_3$  is a model in which the initial model  $P$  fixes a certain set of material values and only considers quality attribute constraints. In the model ( $P_3$ ),  $SL_1^+$  and  $SL_1^-$  are the positive and negative relaxation variables of function  $g_1$ , respectively. The relaxation cost is taken as the objective function. Minimizing the

relaxation cost ensures the satisfaction of the physical constraints as much as possible. The physical constraints can be satisfied by minimizing the slack cost under the determined raw material, product yield, and flow rate between the devices.

For the complete algorithm framework based on the interior point method, we present the solution approach in the form of the following pseudocode:

---

Improved interior point method algorithm

---

Initialization: model ( $P$ ) is established according to actual business scenarios and parameters.

- 1) Based on model  $P$ , the models ( $P_2$ ) and ( $P_3$ ) are generated
- 2) The optimal solution  $\hat{X}$  is obtained by solving model ( $P_2$ )
- 3)  $\hat{X}$  is kept fixed, and the optimal solution  $\hat{Y}$  is obtained by solving model ( $P_3$ )
- 4)  $X^0 \leftarrow \hat{X}$ ;  $Y^0 \leftarrow \hat{Y}$
- 5) Taking  $X^0$  and  $Y^0$  as the initial solutions of the interior point method, the model ( $P$ ) is solved.

Output: the optimal solution  $X^*$  and  $Y^*$

---

### 3. Case study

This section employs some actual data to establish the model and applies the above algorithm to solve the model. The simulation results indicate the superiority of the proposed algorithm to commercial solvers.

#### 3.1. Experimental instances

The examples come from two refiners owned by an oil company. These refiners should consider more than 30 kinds of physical properties when optimizing the production plans, including density, sulfur content, acid content, residual carbon content, metal content, octane number, olefin content, aromatics content, freezing point, and viscosity. This paper selects the refining and petrochemical processes of two refineries A and B for the optimization process. Refinery A contains 36 kinds of physical properties and 79 sets of units, including 3 sets of atmospheric-vacuum distillation units, 50 sets of secondary units, and 26 sets of blending pools. Refinery B contains 54 kinds of physical properties and 58 sets of units, including 2 sets of atmospheric-vacuum distillation units, 40 sets of secondary units, and 16 sets of blending pools. After applying the proposed model, Case 1 (refinery A) has 7229 continuous variables and 3926 constraints, 1827 of which are bilinear. Case 2 (refinery A) has 7230 continuous variables and 3308 constraints, 1851 of which are bilinear. Case 3 (refinery B) includes 5365 continuous variables and 2339 constraints, 379 of which are bilinear. The above information is summarized in Table 3.

**Table 1.** The model scale.

Case ID	The number of physical properties	The number of devices	The number of variables	The number of constraints	The number of bilinear constraints
1	22	58	7229	3926	1827
2	36	79	7230	3308	1851
3	22	58	5365	2339	379

### 3.2. Result

As shown in Table 1, all three examples in the numerical experiment have more than 5000 continuous variables and more than 2000 constraints, many of which are bilinear. Since a nonlinear model with such a scale is extremely challenging for a general solver, this paper selects the Gurobi 9.5.1 for comparative evaluation. It improves the original spatial branch-and-bound algorithm and optimizes the solving efficiency of the pooling problem in the refinery. Gurobi's built-in spatial branch-and-bound algorithm can guarantee global optimality theoretically. The proposed algorithm is evaluated by combining the initial solution module with the interior point method solver IPOPT 3.14.5. The test environment is an Intel 2GHz i5 processor with 16 GB memory and MacOS operating system. The results are displayed in Table 2. For Cases 1 and 2, Gurobi cannot find feasible solutions within 1000 s, while the proposed improved interior point algorithm provides feasible solutions within 1 min. Meanwhile, the results of model 3 indicate that the proposed algorithm only needs 8 s to find a feasible solution, and the objective function value of the algorithm is nearly 10% higher than that obtained by Gurobi. Moreover, for Case 3, the theoretical upper bound of the spatial branch-and-bound is also obtained to evaluate the optimality of the proposed algorithm. The difference between the objective function value and the theoretical optimal value is less than or equal to 2.3%.

**Table 2.** Comparison between the results obtained by the solver and the proposed algorithm.

Case ID	Objective function_solver (RMB)	Upper limit of the objective function_solver (RMB)	Objective function_algorithm (RMB)	Solution time_solver (s)	Solution time_algorithm (s)
1	—	—	85,920.37	1000	16.73
2	—	—	1,443,410.4	1000	63.75
3	34,326.95	38,700.55	37,801.59	1000	7.8

Furthermore, in the experimental process, this paper also verifies the improvement degree of the initial solution module to the solving efficiency of the interior point method. As shown in Table 3, when the initial solution is not set in Cases 1–3 while directly employing the interior point method to solve the problem, a feasible solution cannot be obtained within 1000 s. It can be concluded that the designed initial solution module significantly improves the computational efficiency of the overall interior point method.

**Table 3.** Comparison of the results of the interior point method algorithm with and without the initial solution.

Case ID	Objective function without solver (RMB)	Objective function solver (RMB)	Solution time without solver (s)	Solution time solver (s)
1	—	85,920.37	1000	16.73
2	—	1,443,410.4	1000	63.75
3	—	37,801.59	1000	7.8

## 4. Conclusions

This paper employed the traditional P model to formulate the pooling structure in the complex refinery material flow network. Furthermore, an algorithm framework composed of an initial solution strategy and interior point method was designed to obtain and improve the efficiency of the solution.

The main conclusions are as follows:

a) Compared with the commercial solver (Gurobi), the proposed solution approach can produce high-quality feasible solutions of the instances we adopt in 1 minute, increasing the revenue of the refinery plant by nearly 10%.

b) The experimental results demonstrate robustness, which improves the stability of the solution under complex scenarios with differences.

c) Although we can generate satisfactory solutions in a reasonable time, there is still much work to improve efficiency and optimality in the initial points generation strategy.

### Use of AI tools declaration

The authors declare that they have not used Artificial Intelligence (AI) tools in the creation of this article.

### Conflict of interest

The authors declare there is no conflict of interest.

### References

1. M. Ali, Pooling problem: Modeling, global optimization and computational studies survey, *J. Comput. Appl. Math.*, (2021), 478.
2. C. A. Haverly, Studies of the behavior of recursion for the pooling problem, *Acm Sigmap Bull.*, **25** (1978), 19–28. <https://doi.org/10.1145/1111237.1111238>
3. A. Ben-tal, G. Eiger, V. Gershovitz, Global minimization by reducing the duality gap, *Math. Program.*, **63** (1994), 193–212. <https://doi.org/10.1007/BF01582066>
4. I. Quesada, I. E. Grossmann, Global optimization of bilinear process networks with multicomponent flows, *Comput. Chem. Eng.*, **19** (1995), 1219–1242. [https://doi.org/10.1016/0098-1354\(94\)00123-5](https://doi.org/10.1016/0098-1354(94)00123-5)
5. M. Tawarmalani, N. V. Sahindis, Convexification and global optimization in continuous and mixed-integer nonlinear programming, 2022, Boston, MA: Springer US. <https://doi.org/10.1007/978-1-4757-3532-1>
6. T. Kuno, T. Utsunomiya, A Lagrangian based branch-and-bound algorithm for production-transportation problems, *J. Global Optim.*, **18** (2000), 59–73. <https://doi.org/10.1023/A:1008373329033>
7. J. M. Zamora, I. E. Grossmann, A branch and contract algorithm for problems with concave univariate, bilinear and linear fractional terms, *J. Global Optim.*, **14** (1999), 217–249. <https://doi.org/10.1023/A:1008312714792>
8. J. P. Ruiz, I. E. Grossmann, Strengthening of lower bounds in the global optimization of Bilinear and Concave Generalized Disjunctive Programs, *Comput. Chem. Eng.*, **34** (2010), 914–930. <https://doi.org/10.1016/j.compchemeng.2009.10.016>
9. C. A. Meyer, C. A. Floudas, Global optimization of a combinatorially complex generalized pooling problem, *AIChE J.*, **52** (2006), 1027–1037. <https://doi.org/10.1002/aic.10717>
10. M. L. Bergamini, P. Aguirre, I. Grossmann, Logic-based outer approximation for globally optimal synthesis of process networks, *Comput. Chem. Eng.*, **29** (2005), 1914–1933. <https://doi.org/10.1016/j.compchemeng.2005.04.003>

11. R. E. Griffith, R. A. Stewart, A nonlinear programming technique for the optimization of continuous processing systems, *Manage. Sci.*, **7** (1961), 379–392. <https://doi.org/10.1287/mnsc.7.4.379>
12. F. Palacios-Gomez, L. Lasdon, M. Engquist, Nonlinear optimization by successive linear programming, *Manage. Sci.*, **28** (1982), 1106–1120. <https://doi.org/10.1287/mnsc.28.10.1106>
13. T. E. Baker, L. S. Lasdon, Successive linear programming at Exxon, *Manage. Sci.*, **31** (1985), 264–274. <https://doi.org/10.1287/mnsc.31.3.264>
14. L. S. Lasdon, A. D. Waren, S. Sarkar, Solving the pooling problem using generalized reduced gradient and successive linear programming algorithms, *ACM Sigmap Bull.*, **27** (1979), 9–15. <https://doi.org/10.1145/1111246.1111247>
15. Y. H. Dai, R. Diao, K. Fu, Complexity analysis and algorithm design of pooling problem, *J. Oper. Res. Soc. China*, **6** (2018), 249–266. <https://doi.org/10.1007/s40305-018-0193-7>
16. V. Pham, C. Laird, M. El-Halwagi, Convex hull discretization approach to the global optimization of pooling problems, *Ind. Eng. Chem. Res.*, **48** (2009), 1973. <https://doi.org/10.1021/ie8003573>
17. A. Gupte, S. Ahmed, S. S. Dey, Relaxations and discretizations for the pooling problem, *J. Global Optim.*, **67** (2017), 631–669. <https://doi.org/10.1007/s10898-016-0434-4>
18. M. Alfaki, D. Haugland, Comparison of discrete and continuous models for the pooling problem. in *11th Workshop on Algorithmic Approaches for Transportation Modelling, Optimization, and Systems*, 2011. <https://doi.org/10.4230/OASIS.ATMOS.2011.112>
19. C. A. Floudas, A. Aggarwal, A decomposition strategy for global optimum search in the pooling problem, *ORSA J. Comput.*, **2** (1990), 225–235. <https://doi.org/10.1287/ijoc.2.3.225>
20. *How Distributed Recursion Solves the Pooling Problem*, 2022. Available from: <https://www.haverly.com/kathy-blog/578-blog-4-howdrsolvespooling>.
21. Y. R. He, H. Zhang, Solution of pooling problem in planning model with distributed recursive method, *Pet. Process. Petrochem.*, **27** (1996), 4.
22. Y. R. He, J. Z. Xie, Application cases of refinery's production and operation plan optimization, Beijing: China Petrochemical Press, 2018.
23. A. Wachter, L. T. Biegler, Failure of global convergence for a class of interior point methods for nonlinear programming, *Math. Program.*, **88** (2000), 565–574. <https://doi.org/10.1007/PL00011386>
24. A. Wachter, L. T. Biegler, On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming, *Math. Program.*, **106** (2006), 25–57. <https://doi.org/10.1007/s10107-004-0559-y>
25. J. B. Guo, M. S. Yang, Chemical industry's production planning and scheduling optimization, Beijing: Chemical Industry Press, 2006.
26. J. B. Guo, Linear programming technique in petroleum sector: a review, *Comput. Appl. Chem.*, **21** (2004), 1–5.
27. R. H. Byrd, M. E. Hribar, J. Nocedal, An interior point algorithm for large-scale nonlinear programming, *SIAM J. Optim.*, **9** (1999), 877–900. <https://doi.org/10.1137/S1052623497325107>
28. O. Güler, Barrier functions in interior point methods, *Math. Oper. Res.*, **21** (1996), 860–885. <https://doi.org/10.1287/moor.21.4.860>
29. G. D. Pillo, L. Grippo, Exact penalty functions in constrained optimization, *SIAM J. Control Optim.*, **27** (1989), 1333–1360. <https://doi.org/10.1137/0327068>

