I&EC research
Industrial & Engineering Chemistry Research

ARTICLE

pubs.acs.org/IECR

Multisite Capacity, Production, and Distribution Planning with Reactor Modifications: MILP Model, Bilevel Decomposition Algorithm versus Lagrangean Decomposition Scheme

Fengqi You,[†,§] Ignacio E. Grossmann,[†,§,*] and John M. Wassick[‡]

[†]Department of Chemical Engineering, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213

[‡]The Dow Chemical Company, Midland, Michigan 48674

ABSTRACT: We propose a multiperiod mixed-integer linear programming (MILP) model for the simultaneous capacity, production, and distribution planning for a multisite system including a number of production sites and markets. Multiple products are produced in several production trains that are located in different sites. The unique feature of the proposed model is that it considers the construction times of capacity modifications and takes into account the option of capacity transformation by modifying the reactor in a production train from producing one product family to producing another one. To solve the resulting large-scale MILP model, we present solution techniques based on Lagrangean decomposition and bilevel decomposition. Numerical examples are presented to illustrate the applicability of the model and the performance of the algorithms. It is shown that the bilevel decomposition is the superior solution approach in terms of faster computational times and smaller optimality gaps for the problem addressed in this work.

1. INTRODUCTION

Inherent in the strategic planning for a global multiproduct specialty chemicals business is the long-term capital spending plan for adjusting the capacities of the manufacturing facilities to meet the anticipated demands of the product portfolio. The plan, which projects out 5−10 years, involves modifying the global multisite production facilities to meet the demands of multiple geographically distributed markets. Potential capacity modifications include building new production sites or facilities, adding additional capacity to an existing site, shutting down selected facilities and modifying the reactor of a production facility in order to switch from producing one set of products to another set. The reactor modification may be realized for instance through changes in piping, the addition of side arm heat exchangers, overhead condensers, and rundown tanks. These capacity modifications are necessary because of anticipated changes in the demand and/or prices of each of the products in the portfolio. The plan may include new product introductions and the elimination of nonprofitable products. The relative demand for individual products can fluctuate widely over time as conditions change in the many markets that are served. The price and profitability among products varies greatly and can shift over time. Shipping costs from a production site to a customer region may also be taken into account. The resulting planning decisions will take place over a long time horizon, where specific decisions must be made at each period along the horizon. All these conditions create a challenging strategic planning environment for enterprise-wide optimization.[1−3]

The objective of this paper is to first develop a multiperiod mixed-integer linear programming (MILP) model for the planning and coordination of modifications for capacity (including adding a production facility, shutting down a facility, and modifying the reactor of a facility from producing one product family to another), and production, transportation, and sale activities of geographically distributed multisite production facilities. Our formulation takes into account multiple trade-offs, and simultaneously determines the optimal capacity expansion, capacity reduction, and capacity transformation decisions, and the detailed production profiles of each processing facilities in each production site, as well as the accurate distribution and sale profiles of each product in each market. The unique feature of the proposed model is that it considers the construction lead times of capacity modifications and takes into account the option of capacity transformation by modifying the reactor in a production facility from producing one product family to producing another one. To solve the resulting MILP problem efficiently, we propose two solution techniques based on Lagrangean decomposition and bilevel decomposition, respectively. The proposed model and solution techniques are applied to four example problems in order to illustrate the application of the model and to assess the performance of the proposed algorithms.

This work involves novel features in both the modeling part and the algorithmic part. The main challenge of modeling this problem arises from the potential capacity modification that involves modifying the reactor of a production facility from producing one product family to producing another product family. Although this type of modification usually happens in the specialty chemicals industry, it has not been addressed previously in the literature to the best of our knowledge. Besides, explicitly

Special Issue: Puigjaner Issue

Received: March 9, 2010
Accepted: May 4, 2010
Revised: April 23, 2010
Published: June 03, 2010

considering the construction lead times of capacity modifications introduces significant complexities in the planning model and it is another novelty of this work. To tackle the combinatorial complexities of the resulting large-scale optimization problem, we developed a bilevel decomposition algorithm and a Lagrangean decomposition method to address this computational challenge. A comparison between the performances of these two decomposition methods is also somewhat unique and has not been reported before.

The paper is organized as follows. In section 2, we review some literature closely related to this work. A general problem statement is provided in section 3, and it is followed by the detailed model formulation in section 4. Two solution methods, the Lagrangean decomposition scheme and the bilevel decomposition algorithm are presented in section 5. In section 6, we present four examples to illustrate the application of the model and to assess the performance of the proposed algorithms. The conclusion of this paper is then given in the last section.

2. LITERATURE REVIEW

Strategic planning typically covers a time horizon of a few years, and decisions cover issues such as capacity, production, and distribution. The area of strategic supply chain design largely falls into this category. However, most of the work on capacity planning focuses on capacity expansion, and to the best of our knowledge no previous work has so far taken into account the potential capacity modification actions of modifying an existing facility to produce other products. Long-range multiperiod capacity planning models for chemical process industry was first studied by Sahinidis et al.[4,5] and later by Liu and Sahinidis.[6] In these works, the authors developed optimal planning procedures for production and capacity expansions using mixed-integer linear programming techniques. Bok et al.[7] extended the model proposed by Liu and Sahinidis[6] to consider the robust investment for long-term capacity expansion of chemical complexes and developed a mixed-integer nonlinear programming (MINLP) model to trade-off the expected net present value (NPV), the expected square of deviation of NPV, and the expected square of excess capacity. Lee et al.[8] developed an MINLP model that integrates the capacity expansion model with production and distribution considerations. They reformulate a convex MINLP model by using exponential transformation of the variables to eliminate the bilinear terms in the original model, and used the outer-approximation algorithm for the solution. Later, Tsiakis et al.[9] proposed an MILP model for the optimal design of multiechelon supply chain networks under demand uncertainty. Papageorgiou et al.[10] presented a multiperiod MILP model for managing product portfolios in the pharmaceutical industry. Their model addresses product development and introduction along with capacity expansion planning. A simulation-based multiobjective optimization framework was proposed by Cheng et al.[11] for simultaneous capacity planning and production-inventory control under uncertainty. Oh and Karimi[12] proposed a MILP model for the capacity expansion planning problem on a global scale while accounting for the effect of multiple key regulatory factors. A two-stage multiscenario MILP model is proposed by Levis et al.[13] for multisite capacity planning under uncertainty in the pharmaceutical industry. A hierarchical algorithm is developed to reduce the computational effort needed for the solution of the resulting large-scale MILP problem. You and Grossmann[14] proposed a bicriterion MINLP model for the design of responsive chemical supply chains. The model simultaneously optimizes the capacity, production, distribution, and inventory planning decisions of a multiproduct chemical supply chain, and a hierarchical algorithm is also developed to improve the computational efficiency. Recently, Naraharisetti et al.[15] presented a novel MILP model to determine the optimal asset management and capital budgeting decisions for the capacity expansion planning and supply chain redesign.

Another body of research closely related to this work is about decomposition methods for solving the capacity planning problems. Lagrangean decomposition is recognized as an efficient tool for solving large-scale optimization problems with special structures. The Lagrangean relaxation and subgradient optimization were first discussed by Fisher.[16,17] Later, Guignard and Kim[18] proposed the well-known Lagrangean decomposition method that yields stronger bounds than the Lagrangean relaxation algorithm.

A large number of applications of Lagrangean-based algorithms for capacity planning and supply chain design problems have been reported in the past 20 years. Sridharan[19] implemented the Lagrangean relaxation method for the plant location problem with consideration of capacity issues. Van den Heever et al.[20] developed a Lagrangean heuristic method for the design and planning of offshore oil fields. Maravelias and Grossmann[21] proposed a heuristic algorithm based on Lagrangean decomposition for solving a multiperiod MILP model for simultaneous optimization of resource-constrained scheduling of testing tasks in new product development and design/planning of batch manufacturing facilities. A Lagrangean decomposition algorithm for multisite production and distribution planning was proposed by Jackson and Grossmann.[22] They also compared the performance of temporal decomposition approach with the spatial decomposition, and illustrated the effectiveness of the proposed temporal decomposition. Neiro and Pinto[23] applied the Lagrangean-based method to a petroleum supply chain planning model. The results showed that significant improvement in computational efficiency can be achieved by using Lagrangean decomposition. This work was further extended by Chen and Pinto[24] for supply chain planning and scheduling of flexible process networks. Recently, You and Grossmann[25,26] proposed two modified Lagrangean decomposition methods for solving large-scale MINLP models for supply chain network design with stochastic inventories. In addition to Lagrangean decomposition, bilevel decomposition has also been applied to efficiently solve large-scale capacity planning problems. Iyer and Grossmann[27] proposed a rigorous bilevel decomposition algorithm to reduce the computational effort of solving a multiperiod capacity expansion planning model for process networks. The decomposition algorithm solves a master problem in the reduced variable space of binary variables to determine the selection of processes and an upper bound to the NPV. A detailed planning model is then solved for the selected processes to determine the expansion plan and a lower bound to the objective function. Similar bilevel decomposition schemes have also been implemented by Bok et al.[28] for a supply chain optimization problem, and by Erdirik-Dogan and Grossmann[29] for the simultaneous planning and scheduling of single-stage continuous multiproduct plants. Despite the increasing success of implementing decomposition methods for solving large-scale capacity planning problems, to the best of our knowledge, a comparison between Lagrangean decomposition and bilevel decomposition has not been addressed in the existing literature.

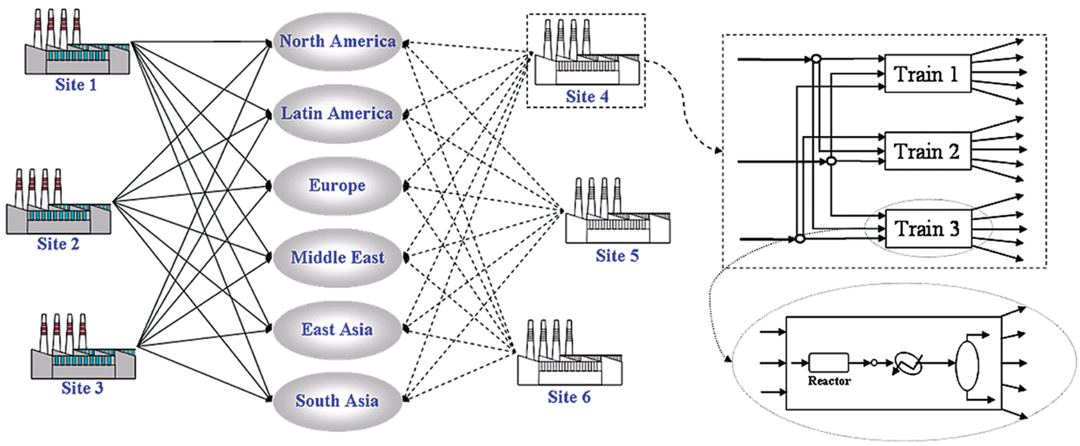

Figure 1. Multisite production−distribution network.

3. PROBLEM STATEMENT

Given is a set of products $p \in P$ to be produced and distributed in a given multisite production−distribution network as shown in Figure 1. The products are produced in several sites $s \in S$, each of which contains a set of production trains $i \in I$ whose capacities and operations are optimized over $t \in T$ time periods.

Some production sites, such as Sites 1−3 in Figure 1, have some existing production trains $i \in I_E(s)$, while others have available space to install new trains $i \in I_N(s)$. Each production train, including a reactor and some downstream processing facilities, can produce a specific product family $j \in J$, which includes a number of products $p \in P(j)$. There are finite number of potential capacity levels (or size of the production train) $k \in K$, which correspond to different capacities $(\overline{Q}_{j,k})$ of a production train if it is producing product family j. Note that the production capacity of a train is measured in terms of the main product $p \in PM(j)$ of the product family j, and a coefficient $\rho_{j,p}$ links the production rate of product $p \in P(j)$ and the main product $p \in PM(j)$ of this product family. An existing production train i at site s with capacity level k can be modified from producing product family j' at time $t-1$ to producing product family j, after modifying the reactor and downstream processing facilities of a train. The modification starts at time period t and spans over $\tau_{s,i,j,j'}$ time periods, which is the modification lead time. We are also given the construction lead time $(\lambda_{s,i})$ to install a new production train as train i in site s, and the shut down of the train action is assumed to be effective in the same time period as the decision is made. Note that no product can be produced during the installation or modification process of a production train. We are also given the cost $(\beta^{SD}_{s,i,k,t})$ of shutting down train i in site s with capacity level k at time t, the cost $(\beta^{0}_{s,i,k,t})$ of installing a new train with capacity level k at time period t, and the cost $(\beta^{T}_{s,i,j,j',k,t})$ of modifying an existing train starting at the beginning of time period t from producing product family j' at time $t-1$ to producing product family j. There is a fixed operational cost $(\beta^{P}_{s,i,k,t})$ for train i in site s operating with capacity level k to produce product family j at time period t. We assume that all the investments on capacity modifications happen at the beginning of the corresponding time periods.

The products are sold to several global markets $m \in M$. On the basis of the market forecast, upper and lower bounds for the demand at each time period t are specified for each product p in each market m. We are given the sale prices $(\text{price}_{m,p,t})$ for each product, the unit shipping costs $(\eta_{s,m,p,t})$, and unit production cost $(\sigma_{s,i,j,p,t})$, including discounted raw material cost.

We assume that production rates, raw material consumption rates, and certain operating conditions at each train are constant within a given time period. Transportation costs are assumed to be linear. All sites may supply all markets, but a production site may not be able to produce all products. We do not account for transportation delays in the network due to the long-term planning horizon. The length of the time period is such that the effects of transportation delays are minimal.

The problem is to determine at each time period which products to manufacture in each train and which sites will supply products to each market, as well as to establish an optimal capacity modification plan of capital outlays and optimal operating conditions for each site over the total time horizon, such that future demand is satisfied consistently with the strategic marketing plan. The objective is to maximize the net present value of the entire production−distribution network over the planning horizon.

4. MODEL FORMULATION

The model will be formulated as a multiperiod MILP problem, which will predict the detailed capacity modification plan, production profiles, and distribution amounts of a production−distribution network. The objective function of this model is to maximize the NPV as given in eq 33. Constraints (1)−(5) refer to the logic constraints for capacity expansion, constraints (6)−(7) are used for the capacity reduction by shutting down production trains, constraints (8)−(14) determine the existence of production trains, constraints (15)−(24) are for the capacity transformation with reactor conversion, constraints (25)−(26) determine the production capacities, constraints (28)−(31) are production−distribution planning constraints, constraints (32) refer to non-negative constraints, and constraints (34)−(36) are for linearization. The detailed mathematical formulation is presented in this section and a list of indices, sets, parameters, and variables is given in the Appendix.

Logical Constraints for the Installation of a Production Train. The first type of constraints is for the installation of production facilities (trains), that is, capacity expansion constraints. If a new train is selected to be installed as train i of site s, then only one type of train size (with capacity level k) can be selected. Thus, we have the following constraint.

$$A_{s,i,t} = \sum_k C_{s,i,k,t}, \quad \forall\, s \in S,\ i \in I_N(s),\ t \in T \tag{1}$$

where $C_{s,i,k,t}$ is a binary variable to indicate if a new train with capacity level k is selected to be installed as train i in site s at time period t, and $A_{s,i,t}$ is another binary variable (that can be treated as continuous variable), which indicates whether there is a new train installed as train i in site s at time period t. $I_N(s)$ is the set of potential new production trains considered for site s.

If a new train is selected to install as train i of site s, it should be installed in a certain time period t within the planning horizon.

$$D_{s,i,k} = \sum_t C_{s,i,k,t}, \quad \forall\, s \in S,\ i \in I_N(s),\ k \in K \tag{2}$$

where $D_{s,i,k}$ is a binary variable (that can also be treated as continuous variable), which equals to 1 if a new train with capacity level k is selected to install as train i in site s.

Constraint 3 states that we have the logic relationship that at most one type of train can be installed as train i of site s over the planning horizon.

$$\sum_k \sum_t C_{s,i,k,t} \le 1, \quad \forall\, s \in S,\ i \in I_N(s) \tag{3}$$

Note that for an existing production train i of site s, we set $A_{s,i,t} = 0$, $C_{s,i,k,t} = 0$, and $D_{s,i,k'} = 1$ and $D_{s,i,k} = 0$, $\forall k \neq k'$, where k' is the given capacity level.

If there is an investment at the beginning of time period t to install a new production facility as train i in site s, it implies that this train does not exist in time period t' prior to $t + \lambda_{s,i}$, where $\lambda_{s,i}$ is the construction lead time to install a new production train. It also implies that this production train exists in time period $t + \lambda_{s,i}$. These relationships can be represented through the following logic propositions.

$$A_{s,i,t} \Rightarrow \neg X_{s,i,t'}, \quad \forall\, s \in S,\ i \in I_N(s),\ t' \le t + \lambda_{s,i} - 1$$

$$A_{s,i,t} \Rightarrow X_{s,i,t+\lambda_{s,i}}, \quad \forall\, s \in S,\ i \in I_N(s),\ t \in T$$

These can be further transformed into inequalities as described in Raman and Grossmann.[30]

$$A_{s,i,t} \le 1 - X_{s,i,t'}, \quad \forall\, s \in S,\ i \in I_N(s),\ t' \le t + \lambda_{s,i} - 1 \tag{4}$$

$$A_{s,i,t} \le X_{s,i,t+\lambda_{s,i}}, \quad \forall\, s \in S,\ i \in I_N(s),\ t \in T \tag{5}$$

where $X_{s,i,t}$ is a binary variable, that is equal to 1 if train i in site s exists at time period t.

Logical Constraints for Shutting down Production Trains. The second type of constraints is for capacity reduction by shutting down production trains. If a train i in site s is shut down at time period t, it implies that this train does not exist in time period t and all the time periods after t. It also implies that this train exists in time period $t - 1$. Thus, we have the following logic propositions:

$$B_{s,i,t} \Rightarrow \neg X_{s,i,t'}, \quad \forall\, s \in S,\ i \in I,\ t' \ge t$$

$$B_{s,i,t} \Rightarrow X_{s,i,t-1}, \quad \forall\, s \in S,\ i \in I,\ t \ge 2$$

which are equivalent to

$$B_{s,i,t} \le 1 - X_{s,i,t'}, \quad \forall\, s \in S,\ i \in I,\ t' \ge t \tag{6}$$

$$B_{s,i,t} \le X_{s,i,t-1}, \quad \forall\, s \in S,\ i \in I,\ t \ge 2 \tag{7}$$

where $B_{s,i,t}$ is a binary variable (that can also be treated as a continuous variable), which is equal to 1 if train i in site s is shut down at time period t.

Logical Constraints for the Existence of Production Trains. If a new train i in site s exists at time period t, then either it starts to be installed at time period $t - \lambda_{s,i}$, where $\lambda_{s,i}$ is the installation time or it exists in the pervious time period.

$$X_{s,i,t} \Rightarrow A_{s,i,t-\lambda_{s,i}} \vee X_{s,i,t-1}, \quad \forall\, s \in S,\ i \in I_N(s),\ t \in T$$

The above logic proposition can be transformed to the following constraint for new production trains, that is, $\forall i \in I_N(s)$:

$$X_{s,i,t} \le A_{s,i,t-\lambda_{s,i}} + X_{s,i,t-1}, \quad \forall\, s \in S,\ i \in I_N(s),\ t \in T \tag{8}$$

For existing trains $i \in I_E(s)$, the logic proposition can be transformed to the following constraint (9):

$$X_{s,i,t+1} \le X_{s,i,t}, \quad \forall\, s \in S,\ i \in I_E(s),\ t \in T \tag{9}$$

If a train i in site s exists at time period t, then it will either be shut down in the next time period, or it will still exist in the next time period. This relationship leads to the following logic proposition.

$$X_{s,i,t} \Rightarrow B_{s,i,t+1} \vee X_{s,i,t+1}, \quad \forall\, s \in S,\ i \in I,\ t \in T$$

which is equivalent to

$$X_{s,i,t} \le B_{s,i,t+1} + X_{s,i,t+1}, \quad \forall\, s \in S,\ i \in I,\ t \in T \tag{10}$$

Similarly, we have the logic relationship that the existing trains should either exist or be shut down in the first time period.

$$1 \le B_{s,i,1} + X_{s,i,1}, \quad \forall\, s \in S,\ i \in I_E(s) \tag{11}$$

In addition, we can only shut down train i in site s at most once throughout the entire planning horizon.

$$\sum_t B_{s,i,t} \le 1, \quad \forall\, s \in S,\ i \in I \tag{12}$$

For a potential train $i \in I_N(s)$ in site s, if it has been installed at any time period before $t - \lambda_{s,i}$, it either must exist in time period t or had been shut down at a time period not later than t. On the other hand, if this new train exists in time period t or had been shut down in time period t or earlier, then it must have been installed at any time period before $t - \lambda_{s,i}$. This complex relationship leads to the following logic proposition.

$$\bigvee_{t' \le t} B_{s,i,t'} \vee X_{s,i,t} \Leftrightarrow \bigvee_{t' \le t} A_{s,i,t'-\lambda_{s,i}},$$
$$\forall\, s \in S,\ i \in I_N(s),\ t \in T$$

which can be further transformed to the following equation

$$X_{s,i,t} = \sum_{t'=1}^{t} (A_{s,i,t'-\lambda_{s,i}} - B_{s,i,t'}), \quad \forall\, s \in S,\ i \in I_N(s),\ t \in T \tag{13}$$

For an existing train $i \in I_N(s)$ in site s, if it has been shut down before time period t, it would not exist in time period t. If it exists in time period t, then it has not been shut down before. Thus, we have

$$\bigvee_{t' \leq t} B_{s,i,t'} \Leftrightarrow \neg X_{s,i,t}, \quad \forall\, s \in S,\ i \in I_E(s),\ t \in T, t \geq t'$$

which is equivalent to

$$X_{s,i,t} = 1 - \sum_{t'=1}^{t} B_{s,i,t'}, \quad \forall\, s \in S,\ i \in I_E(s),\ t \in T, t \geq t' \quad (14)$$

Constraints (13) and (14) guarantee $X_{s,i,t}$ must be either 0 or 1.

Logical Constraints for Capacity Transformation with Reactor Modification. Capacity transformation is very similar to production scheduling in terms of model formulation. To mathematically describe the logical relationships inherent in the capacity transformation, we first introduce a binary variable $Y_{s,i,j,t}$ that is equal to 1 if train i in site s is producing product family j at time period t, and another binary variable $Z_{s,i,j,j',t}$ (which can be treated as a continuous variable), such that $Z_{s,i,j,j',t} = 1$ if train i in site s produces product family j' at time period $t - 1$ and there is an investment starting at the beginning of time period t to modify this train to produce product family j. With these two types of variables, the logical constraints for reactor modification are given below.

If train i in site s exists, it can produce at most one product family j (not producing any product during the train modification process); if the train does not exist, no products are produced. This relationship yields the following constraint:

$$\sum_j Y_{s,i,j,t} \leq X_{s,i,t}, \quad \forall\, s \in S,\ i \in I,\ t \in T \quad (15)$$

If a reactor modification happens in train i site s starting from the beginning of time period t, from producing product family j' at time period $t - 1$ to produce product family j, it implies that this train produces product family j' at time period $t - 1$, and produces product family j at time period $t + \tau_{s,i,j,j'}$, where $\tau_{s,i,j,j'}$ is the construction lead time for reactor modification. Also, during the train modification, this train can not produce any products. Thus, we have the following logic propositions:

$$Z_{s,i,j,j',t} \Rightarrow Y_{s,i,j',t-1}, \quad \forall\, s \in S,\ i \in I,\ j \in J,\ j' \in J,\ t \geq 2$$

$$Z_{s,i,j,j',t} \Rightarrow Y_{s,i,j,t+\tau_{s,i,j,j'}}, \quad \forall\, s \in S,\ i \in I,\ j \in J,\ j' \in J,\ t \in T$$

$$Z_{s,i,j,j',t} \Rightarrow \neg Y_{s,i,j'',t'}, \quad \forall\, s \in S,\ i \in I,\ j \in J,\ j' \in J,\ \forall\, j'' \in J,\ t \leq t' < t + \tau_{s,i,j,j'}$$

They can be further transformed into the following constraints.

$$Z_{s,i,j,j',t} \leq Y_{s,i,j',t-1}, \quad \forall\, s \in S,\ i \in I,\ j \in J,\ j' \in J,\ t \geq 2 \quad (16)$$

$$Z_{s,i,j,j',t} \leq Y_{s,i,j,t+\tau_{s,i,j,j'}}, \quad \forall\, s \in S,\ i \in I,\ j \in J, j' \in J,\ t \in T \quad (17)$$

$$Z_{s,i,j,j',t} \leq 1 - Y_{s,i,j'',t'}, \quad \forall\, s \in S,\ i \in I,\ j \in J,\ j' \in J,\ \forall\, j'' \in J,\ t \leq t' < t + \tau_{s,i,j,j'} \quad (18)$$

On the other hand, if production train i in site s produces product family j' at time period $t - 1$ and produces product family j at time period $t + \tau_{s,i,j,j'}$ there should be an investment for reactor modification in this train at the beginning of time period t. The relationship leads to the following logic proposition.

$$Y_{s,i,j',t-1} \wedge Y_{s,i,j,t+\tau_{s,i,j,j'}} \Rightarrow Z_{s,i,j,j',t}, \quad \forall\, s \in S,\ i \in I,\ j \in J,\ j' \in J,\ t \geq 2$$

So we have

$$Z_{s,i,j,j',t} \geq Y_{s,i,j',t-1} + Y_{s,i,j,t+\tau_{s,i,j,j'}} - 1, \quad \forall\, s \in S,\ i \in I,\ j \in J,\ j' \in J,\ t \geq 2 \quad (19)$$

If production train i in site s produces product family j' at time period $t - 1$, and it also exists at time period t, then there must be an investment for reactor modification at the beginning of time period t so that this train is transformed from producing product family j' at time period $t - 1$ to product family j. Thus we can have the following logic proposition:

$$Y_{s,i,j',t-1} \wedge X_{s,i,t} \Rightarrow \bigvee_{j \in J} Z_{s,i,j,j',t}, \quad \forall\, s \in S,\ i \in I,\ j \in J,\ j' \in J,\ t \geq 2$$

which yields constraint (20):

$$Y_{s,i,j',t-1} + X_{s,i,t} - 1 \leq \sum_{j \in J} Z_{s,i,j,j',t}, \quad \forall\, s \in S,\ i \in I,\ j \in J,\ j' \in J, t \geq 2 \quad (20)$$

Note that product family j could be the same as product family j', that is, no modification. In this model we set both modification lead time and cost as 0 when modifying a train from product family j to the same product family j. This is similar to the transitions in production scheduling.

If production train i in site s produces product family j at time period t, it implies that this train is either newly installed (installation starting at the beginning of time period $t - \lambda_{s,i}$) or just finishes the reactor modification starting from the beginning of time period $t - \tau_{s,i,j,j'}$ from producing product family j' (at time period $t - \tau_{s,i,j,j'} - 1$) to product family j at time period t. Thus, we have the following logic proposition:

$$Y_{s,i,j,t} \Rightarrow \bigvee_{j' \in J} Z_{s,i,j,j',t-\tau_{s,i,j,j'}} \vee A_{s,i,t-\lambda_{s,i}}, \quad \forall\, s \in S,\ i \in I,\ j \in J,\ j' \in J,\ t \geq \max\{\tau_{s,i,j,j'}, \lambda_{s,i}\}$$

which can be transformed to the following constraint:

$$Y_{s,i,j,t} \leq \sum_{j' \in J} Z_{s,i,j,j',t-\tau_{s,i,j,j'}} + A_{s,i,t-\lambda_{s,i}}, \quad \forall\, s \in S,\ i \in I,\ j \in J,\ j' \in J,\ t \geq \max\{\tau_{s,i,j,j'}, \lambda_{s,i}\} \quad (21)$$

In addition to the aforementioned constraints, the "boundary conditions" of variables $Y_{s,i,j,t}$ and $Z_{s,i,j,j',t}$ for existing trains should be considered. For production $i \in I_E(s)$ at site s that produces product family j' at the first time period $t = 1$, we have the following constraints:

$$Z_{s,i,j,j',1} = 0, \quad \forall\, s \in S,\ i \in I_E(s),\ j \neq j' \quad (22)$$

$$X_{s,i,1} \leq \sum_{j \in J} Z_{s,i,j,j',1}, \quad \forall\, s \in S,\ i \in I_E(s) \quad (23)$$

$$Z_{s,i,j,j',1} = Y_{s,i,j,1+\tau_{s,i,j,j'}}, \quad \forall\, s \in S,\ i \in I_E(s),\ j \neq j' \quad (24)$$

Note that in constraints (22)−(24), the index j' is fixed.

Capacity Constraints. If a train i in site s with capacity level k is producing product family j at time period t ($V_{s,i,j,k,t}$), it is equivalent that this train is producing product family j at time period t ($Y_{s,i,j,t}$) and it is installed with train type k ($D_{s,i,k}$). So we have $V_{s,i,j,k,t} = D_{s,i,k}Y_{s,i,j,t}$, which can be linearized with the following constraints.

$$V_{s,i,j,k,t} \le D_{s,i,k}, \quad \forall s \in S, i \in I, j \in J, k \in K, t \in T \quad (25.1)$$

$$V_{s,i,j,k,t} \le Y_{s,i,j,t}, \quad \forall s \in S, i \in I, j \in J, k \in K, t \in T \quad (25.2)$$

$$V_{s,i,j,k,t} \ge D_{s,i,k} + Y_{s,i,j,t} - 1, \quad \forall s \in S, i \in I, j \in J, k \in K, t \in T \quad (25.3)$$

Therefore, the capacity of train i in site s in terms of product family j at time period t is given by

$$Q_{s,i,j,t} = Y_{s,i,j,t}\sum_k \overline{Q}_{j,k}D_{s,i,k} = \sum_k \overline{Q}_{j,k}V_{s,i,j,k,t}, \quad \forall s \in S, i \in I, j \in J, t \in T \quad (26)$$

Constraint (26) defines the capacity of each production train. It states that if a production train does not produce product family j, the corresponding capacity in terms of product family j is 0. If the train is producing product family j, and it was installed with capacity level k, the associated capacity in terms of product family j should be $\overline{Q}_{j,k}$ (capacity of new production train with capacity level k if producing product family j).

Production and Distribution Planning Constraints. The production amount of product p at train i in site s with product family j at time t ($W_{s,i,j,p,t}$) should not exceed the product of production rate of this product ($r_{s,i,j,p,t}$) and its production time ($\theta_{s,i,j,p,t}$),

$$W_{s,i,j,p,t} \le r_{s,i,j,p,t}\theta_{s,i,j,p,t}, \quad \forall s \in S, i \in I, j \in J, p \in P(j), t \in T \quad (27.1)$$

Similar to the work by Sahinidis and Grossmann.,[4,5] the production rate of product p at train i in site s with product family j at time t ($r_{s,i,j,p,t}$) is proportional to the associated capacity through a relative production coefficient ($\rho_{j,p}$).

$$r_{s,i,j,p,t} = \rho_{j,p}Q_{s,i,j,t}, \quad \forall s \in S, i \in I, j \in J, p \in P(j), t \in T \quad (27.2)$$

The total production time of all the products on a train in a time period should not exceed the duration of the time period (H_t),

$$\sum_{p \in P(j)} \theta_{s,i,j,p,t} \le H_t, \quad \forall s \in S, i \in I, j \in J, t \in T \quad (27.3)$$

Since the right-hand side of constraint 27.1 is the product of two variables, this constraint is nonlinear. We follow the same procedure as in Sahinidis and Grossmann[4,5] to linearize this constraint. First, substituting (27.2) into (27.1) yields the following constraint.

$$W_{s,i,j,p,t}/\rho_{j,p} \le Q_{s,i,j,t}\theta_{s,i,j,p,t}, \quad \forall s \in S, i \in I, j \in J, p \in P(j), t \in T \quad (27.4)$$

Summing both sides of the above constraint over all the products $p \in P(j)$, we have

$$\sum_{p \in P(j)} \frac{W_{s,i,j,p,t}}{\rho_{j,p}} \le \sum_{p \in P(j)} Q_{s,i,j,t}\theta_{s,i,j,p,t} = Q_{s,i,j,t}\sum_{p \in P(j)} \theta_{s,i,j,p,t}, \quad \forall s \in S, i \in I, j \in J, t \in T \quad (27.5)$$

Substituting (27.3) into (27.5) leads to the following constraint:

$$\sum_{p \in P(j)} \frac{W_{s,i,j,p,t}}{\rho_{j,p}} \le Q_{s,i,j,t}H_t, \quad \forall s \in S, i \in I, j \in J, t \in T \quad (28)$$

which is a linear constraint and equivalent to constraints (27.1)−(27.3), although the bilinear terms and the variables for production rates ($r_{s,i,j,p,t}$) and production times ($\theta_{s,i,j,p,t}$) are eliminated.

The production amount at train i in site s should be equal to the total shipping amount from this site ($F_{s,m,p,t}$) to all the markets m,

$$\sum_i \sum_j W_{s,i,j,p,t} = \sum_m F_{s,m,p,t}, \quad \forall s \in S, p \in P, t \in T \quad (29)$$

The summation of the shipping amount from all sites should be equal to the sale amount in each market ($\mathrm{Sa}_{m,p,t}$).

$$\mathrm{Sa}_{m,p,t} = \sum_s F_{s,m,p,t}, \quad \forall m \in M, p \in P, t \in T \quad (30)$$

The total sale ($\mathrm{Sa}_{m,p,t}$) of product p to market m at time periods t should lie between the lower and upper bounds of the demands,

$$d^{\mathrm{L}}_{m,p,t} \le \mathrm{Sa}_{m,p,t} \le d^{\mathrm{U}}_{m,p,t}, \quad \forall m \in M, p \in P, t \in T \quad (31)$$

Non-negative Constraints. All continuous variables must be nonnegative and the binary variables should be integer:

$$0 \le A_{s,i,t},\ B_{s,i,t},\ D_{s,i,k},\ Z_{s,i,j,j',t} \le 1 \quad (32.1)$$

$$Q_{s,i,j,t},\ V_{s,i,j,k,t},\ W_{s,i,j,p,t},\ F_{s,m,p,t},\ \mathrm{Sa}_{m,p,t} \ge 0 \quad (32.2)$$

$$C_{s,i,k,t},\ X_{s,i,t},\ Y_{s,i,j,t} \in \{0,1\} \quad (32.3)$$

Note that the binary variables $A_{s,i,t}$, $B_{s,i,t}$, $D_{s,i,k}$, and $Z_{s,i,j,j',t}$ are relaxed as continuous variables as given in eq (32.1).

Objective Function. The objective of this model is to maximize the net present value,[31] which takes into account the income from sales, the costs from transportation, operations, and capacity modifications, and the salvage values of the facilities after the planning horizon.

$$\max: \ \mathrm{NPV} = \mathrm{income} - C_{\mathrm{transport}} - C_{\mathrm{operate}} - C_{\mathrm{capital}} + \mathrm{salvage} \quad (33)$$

where

$$\mathrm{income} = \sum_s \sum_m \sum_p \sum_t \frac{\mathrm{price}_{m,p,t} \cdot \mathrm{Sa}_{s,m,p,t}}{(1+ir)^t} \quad (33.1)$$

$$C_{\mathrm{transport}} = \sum_s \sum_m \sum_p \sum_t \frac{\eta_{s,m,p,t} \cdot F_{s,m,p,t}}{(1+ir)^t} \quad (33.2)$$

$$C_{\mathrm{operate}} = \sum_s \sum_i \sum_j \sum_p \sum_t \frac{\sigma_{s,i,j,p,t} \cdot W_{s,i,j,p,t}}{(1+ir)^t} + \sum_s \sum_i \sum_k \sum_t \frac{\beta^{\mathrm{P}}_{s,i,k,t} \cdot D_{s,i,k} \cdot X_{s,i,t}}{(1+ir)^t}$$

$$C_{\mathrm{capital}} = \sum_s \sum_i \sum_k \sum_t \frac{\beta^0_{s,i,k,t} \cdot C_{s,i,k,t}}{(1+ir)^{t-1}} + \sum_s \sum_i \sum_j \sum_{j'} \sum_k \sum_i \frac{\beta^{\mathrm{T}}_{s,i,j,j',k,t} \cdot Z_{s,i,j,j',t} \cdot D_{s,i,k}}{(1+ir)^{t-1}} + \sum_s \sum_i \sum_k \sum_t \frac{\beta^{\mathrm{SD}}_{s,i,k,t} \cdot B_{s,i,t} \cdot D_{s,i,k}}{(1+ir)^{t-1}}$$

$$\mathrm{salvage} = \sum_s \sum_{i \in I_{\mathrm{N}}(s)} \sum_k \sum_t \frac{SV^{\mathrm{N}}_{s,i,k,t} \cdot C_{s,i,k,t}}{(1+ir)^{|\mathrm{T}|}} + \sum_s \sum_{i \in I_{\mathrm{E}}(s)} \sum_k \frac{SV^{\mathrm{E}}_{s,i,k} \cdot D_{s,i,k}}{(1+ir)^{|\mathrm{T}|}} \quad (33.3)$$

where ir is the interest rate, $\text{price}_{m,p,t}$ is the product price, $\sigma_{s,i,j,p,t}$ is the unit production cost, $\beta^{P}_{s,i,k,t}$ is the operating cost of the production train per unit time period, $\eta_{s,m,p,t}$ is the unit transportation cost, $\beta^{0}_{s,i,k,t}$ is the installation cost of production trains, $\beta^{T}_{s,i,j,j',k,t}$ is the reactor modification cost, $\beta^{SD}_{s,i,k,t}$ is the shutdown cost of production trains, $SV^{N}_{s,i,k,t}$ is the salvage value of new train i in site s with capacity level k that is installed at time t, and $SV^{E}_{s,i,k}$ is the salvage value of the existing train i in site s with capacity level k. Note that the income from sales, the operational cost, and the transportation cost are accounted at the end of each time period, while the capital investment cost is paid at the beginning of each time period. The discounted values of facilities at the end of the planning horizon are considered as the salvage values.

Although we do not consider the strategic inventory cost in the NPV formulation, the average inventory in a production site can be estimated as a function of the corresponding production amount per year multiplied by the average storage period, which is introduced to cover fluctuations in both supply and demand as well as plant interruptions.[32] Alternatively, a stochastic inventory model can be employed to estimate the strategic inventory cost by integrating demand and supply uncertainty with strategic capacity planning.[25,26,33]

Let $Z_{s,i,j,j',t} \cdot D_{s,i,k} = ZD_{s,i,j,j',k,t}$, $B_{s,i,t} \cdot D_{s,i,k} = BD_{s,i,k,t}$ and $D_{s,i,k} \cdot X_{s,i,t} = DX_{s,i,k,t}$, we can use the following linear constraints to replace these three nonlinear equations:

$$ZD_{s,i,j,j',k,t} \le Z_{s,i,j,j',t}, \quad \forall\, s \in S,\ i \in I,\ j \in J,\ j' \in J,\ t \in T \tag{34.1}$$

$$ZD_{s,i,j,j',k,t} \le D_{s,i,k}, \quad \forall\, s \in S,\ i \in I,\ j \in J,\ j' \in J,\ t \in T \tag{34.2}$$

$$ZD_{s,i,j,j',k,t} \ge Z_{s,i,j,j',t} + D_{s,i,k} - 1, \quad \forall\, s \in S,\ i \in I,\ j \in J,\ j' \in J,\ t \in T \tag{34.3}$$

$$ZD_{s,i,j,j',k,t} \ge 0, \quad \forall\, s \in S,\ i \in I,\ j \in J,\ j' \in J,\ t \in T \tag{34.4}$$

$$BD_{s,i,k,t} \le B_{s,i,t}, \quad \forall\, s \in S,\ i \in I,\ k \in K,\ t \in T \tag{35.1}$$

$$BD_{s,i,k,t} \le D_{s,i,k}, \quad \forall\, s \in S,\ i \in I,\ k \in K,\ t \in T \tag{35.2}$$

$$BD_{s,i,k,t} \ge B_{s,i,t} + D_{s,i,k} - 1, \quad \forall\, s \in S,\ i \in I,\ k \in K,\ t \in T \tag{35.3}$$

$$BD_{s,i,k,t} \ge 0, \quad \forall\, s \in S,\ i \in I,\ k \in K,\ t \in T \tag{35.4}$$

$$DX_{s,i,k,t} \le X_{s,i,t}, \quad \forall\, s \in S,\ i \in I,\ k \in K,\ t \in T \tag{36.1}$$

$$DX_{s,i,k,t} \le D_{s,i,k}, \quad \forall\, s \in S,\ i \in I,\ k \in K,\ t \in T \tag{36.2}$$

$$DX_{s,i,k,t} \ge X_{s,i,t} + D_{s,i,k} - 1, \quad \forall\, s \in S,\ i \in I,\ k \in K,\ t \in T \tag{36.3}$$

$$DX_{s,i,k,t} \ge 0, \quad \forall\, s \in S,\ i \in I,\ k \in K,\ t \in T \tag{36.4}$$

Then the operational cost and capital cost can be reformulated as follows:

$$C_{\text{operate}} = \sum_s \sum_i \sum_j \sum_p \sum_t \frac{\sigma_{s,i,j,p,t} \cdot W_{s,i,j,p,t}}{(1+ir)^t} + \sum_s \sum_i \sum_j \sum_k \sum_t \frac{\beta^{P}_{s,i,k,t} \cdot DX_{s,i,k,t}}{(1+ir)^t} \tag{33.4}$$

$$C_{\text{capital}} = \sum_s \sum_i \sum_k \sum_t \frac{\beta^{0}_{s,i,k,t} \cdot C_{s,i,k,t}}{(1+ir)^{t-1}} + \sum_s \sum_i \sum_j \sum_{j'} \sum_k \sum_i \frac{\beta^{T}_{s,i,j,j',k,t} \cdot ZD_{s,i,j,j',k,t}}{(1+ir)^{t-1}} + \sum_s \sum_i \sum_k \sum_t \frac{\beta^{SD}_{s,i,k,t} \cdot BD_{s,i,k,t}}{(1+ir)^{t-1}} \tag{33.5}$$

5. SOLUTION METHODS

The capacity planning model for a typical industrial single site (e.g., 3 production trains, 6 product families, and 36 products) may include more than 300 binary variables, 20 000 variables and 20 000 constraints for a 10-year time horizon (10 time periods). The complicating constraints, which lead to significant computational complexity, are given in the capacity transformation constraints (15)−(24). The multisite capacity, production, and distribution planning model described in the previous section contains additional interconnections between the different sites and markets in the network through product flows. Therefore, the problem size of the corresponding mixed-integer linear programming model can greatly increase when several sites and markets are considered. One approach of expediting the solution of the MILP problem is to specify a branching priority for the binary variables during the branch-and-bound search according to their contribution to the objective function. In addition, nonzero tolerance for the relative optimality criterion can be used to reduce the computation time for large problems. While the above schemes can help to reduce the computational time, they may not be enough to effectively tackle large problems. To address the computational challenge, we exploit the problem structure and apply two solution methods, bilevel decomposition and Lagrangean decomposition.

In the bilevel decomposition algorithm,[27] we first construct an aggregated upper level problem by neglecting the capacity transformation constraints (15)−(24). The solution of the upper level problem provides an upper bound to the optimal NPV. We then solve a lower level detailed planning problem in the reduced space of binary variables for selected production trains to determine the detailed capacity, production, and distribution decisions, and to provide a lower bound to the optimal NPV. The two decomposed problems are iteratively solved and integer cuts, superset cuts, and subset cuts are added to the upper level aggregate problem at each iteration, until an optimality tolerance is reached.

In the Lagrangean decomposition method,[18] the mass balances between sites and markets are dualized and the relaxed problem is then decomposed into a number of subproblems for production sites and markets. A lower bound of optimal NPV can be obtained by solving these Lagrangean subproblems. With the optimal solutions of the Lagrangean subproblems, we can obtain an upper bound of optimal NPV by solving the original problem in the reduced variable space. The subgradients are updated in each iteration and the algorithm keeps iterating until the stopping criterion is reached. Note that, because the capacity planning problem results in the formulation of an MILP model, the solution may exhibit a duality gap with Lagrangean decomposition.

The detailed implementations of these two decomposition algorithms are presented in the following sections.

Bilevel Decomposition. In the bilevel decomposition algorithm, we "decompose" the original problem (P) into an aggregated problem (AP) and a detailed problem (DP). The formulation of (AP) is given as follows:

(AP)

$$\max: \ \mathrm{NPV} = \sum_s\sum_m\sum_p\sum_t \frac{\mathrm{price}_{m,p,t}\cdot \mathrm{Sa}_{s,m,p,t}}{(1+ir)^t} - \sum_s\sum_m\sum_p\sum_t \frac{\eta_{s,m,p,t}\cdot F_{s,m,p,t}}{(1+ir)^t} - \sum_s\sum_i\sum_j\sum_p\sum_t \frac{\sigma_{s,i,j,p,t}\cdot W_{s,i,j,p,t}}{(1+ir)^t} - \sum_s\sum_i\sum_k\sum_t \frac{\beta^{\mathrm{P}}_{s,i,k,t}\cdot DX_{s,i,k,t}}{(1+ir)^t} - \sum_s\sum_i\sum_k\sum_t \frac{\beta^{0}_{s,i,k,t}\cdot C_{s,i,k,t}}{(1+ir)^{t-1}} - \sum_s\sum_i\sum_k\sum_t \frac{\beta^{\mathrm{SD}}_{s,i,k,t}\cdot BD_{s,i,k,t}}{(1+ir)^{t-1}} + \sum_s\sum_{i\in I_{\mathrm{N}}(s)}\sum_k\sum_t \frac{SV^{\mathrm{N}}_{s,i,k,t}\cdot C_{s,i,k,t}}{(1+ir)^{|T|}} + \sum_s\sum_{i\in I_{\mathrm{E}}(s)}\sum_k \frac{SV^{\mathrm{E}}_{s,i,k}\cdot D_{s,i,k}}{(1+ir)^{|T|}} \tag{37}$$

subject to

$$Q_{s,i,j,t} \le Q^{\mathrm{U}}_{s,i,j,t}\cdot X_{s,i,t}, \quad \forall\ s\in S, i\in I, j\in J,\ t\in T \tag{38}$$

$$\sum_k\sum_j \frac{QD_{s,i,j,k,t}}{\overline{Q}_{j,k}} \le X_{s,i,t}, \quad \forall\ s\in S, i\in I, t\in T \tag{39}$$

$$QD_{s,i,j,k,t} + QD1_{s,i,j,k,t} = Q_{s,i,j,t}, \quad \forall\ s\in S, i\in I, j\in J, t\in T \tag{40.1}$$

$$QD_{s,i,j,k,t} \le Q^{\mathrm{U}}_{s,i,j,t}\cdot D_{s,i,k}, \quad \forall\ s\in S, i\in I, j\in J, t\in T \tag{40.2}$$

$$QD1_{s,i,j,k,t} \le Q^{\mathrm{U}}_{s,i,j,t}\cdot(1-D_{s,i,k}), \quad \forall\ s\in S, i\in I, j\in J, t\in T \tag{40.3}$$

$$QD_{s,i,j,k,t} \ge 0, \quad QD1_{s,i,j,k,t} \ge 0, \quad \forall\ s\in S, i\in I, j\in J, t\in T \tag{40.4}$$

and constraints (1)−(14), (28)−(32) and (35)−(36).

In model (AP), the objective function is to maximize an "approximate" NPV as given in eq 38, which neglects the cost from capacity transformations. Some constraints of this model are also included in the original model (P): logic constraints for capacity expansion (1)−(5), capacity reduction constraints (6) and (7), production trains existence constraints (8)−(14), production−distribution planning constraints (28)−(31), non-negative constraints (32), and linearization constraints (34)−(36).

Note that (AP) does not consider capacity transformation and it aggregates constraints (15)−(26), which are for the capacity transformation and for defining the production capacities. The 0−1 variables $Y_{s,i,j,t}$ and continuous variables $Z_{s,i,j,j',t}$, which relate the capacity transformation with reactor modification, are not included in (AP). Thus, the cost of capacity transformation $(\Sigma_s\Sigma_i\Sigma_j\Sigma_{j'}\Sigma_k\Sigma_i(\beta^{\mathrm{T}}_{s,i,j,j',k,t}\cdot ZD_{s,i,j,j',k,t})/((1+ir)^{t-1}))$ is not included in the objective function of (AP), which is given in eq (37).

In addition, we relax the capacity constraint (26) to eliminate variable $Y_{s,i,j,t}$ in (AP). Since $D_{s,i,k}$ can only take the value of 0 or 1, if $\Sigma_k D_{s,i,k} \neq 0$, constraint (26) can be reformulated as follows:

$$Y_{s,i,j,t} = \frac{Q_{s,i,j,t}}{\sum_k \overline{Q}_{j,k}D_{s,i,k}} = Q_{s,i,j,t}\sum_k \frac{D_{s,i,k}}{\overline{Q}_{j,k}} = \sum_k \frac{Q_{s,i,j,t}\cdot D_{s,i,k}}{\overline{Q}_{j,k}} \tag{41.1}$$

Due to constraint (15), we can relax the above equation to the following constraint,

$$\sum_j Y_{s,i,j,t} = \sum_j\sum_k \frac{Q_{s,i,j,t}\cdot D_{s,i,k}}{\overline{Q}_{j,k}} \le X_{s,i,t} \tag{41.2}$$

Let $QD_{s,i,j,k,t} = Q_{s,i,j,t}\cdot D_{s,i,k}$. The above inequality leads to constraint (39) and the nonlinear equation is linearized through constraints (40). Note that if $\Sigma_k D_{s,i,k} = 0$, it implies that the $D_{s,i,k} = 0$, $\forall k \in K$, which also satisfies constraints (39)−(40).

Since (AP) is a relaxation of the original problem (P), the following property holds:. *Property 1:* Problem (AP) yields an upper bound to the solution of problem (P).

In the detailed problem (DP), the multisite production−distribution network is optimized for selected production sites and trains determined by (AP). The formulation of (DP) is given as follows:

(DP)

$$\max: \ \mathrm{NPV} = \sum_s\sum_m\sum_p\sum_t \frac{\mathrm{price}_{m,p,t}\cdot \mathrm{Sa}_{s,m,p,t}}{(1+ir)^t} - \sum_s\sum_m\sum_p\sum_t \frac{\eta_{s,m,p,t}\cdot F_{s,m,p,t}}{(1+ir)^t} - \sum_s\sum_i\sum_j\sum_p\sum_t \frac{\sigma_{s,i,j,p,t}\cdot W_{s,i,j,p,t}}{(1+ir)^t} - \sum_s\sum_i\sum_k\sum_t \frac{\beta^{\mathrm{P}}_{s,i,k,t}\cdot DX_{s,i,k,t}}{(1+ir)^t} - \sum_s\sum_i\sum_k\sum_t \frac{\beta^{0}_{s,i,k,t}\cdot C_{s,i,k,t}}{(1+ir)^{t-1}} - \sum_s\sum_i\sum_j\sum_{j'}\sum_k\sum_i \frac{\beta^{\mathrm{T}}_{s,i,j,j',k,t}\cdot ZD_{s,i,j,j',k,t}}{(1+ir)^{t-1}} - \sum_s\sum_i\sum_k\sum_t \frac{\beta^{\mathrm{SD}}_{s,i,k,t}\cdot BD_{s,i,k,t}}{(1+ir)^{t-1}} + \sum_s\sum_{i\in I_{\mathrm{N}}(s)}\sum_k\sum_t \frac{SV^{\mathrm{N}}_{s,i,k,t}\cdot C_{s,i,k,t}}{(1+ir)^{|T|}} + \sum_s\sum_{i\in I_{\mathrm{E}}(s)}\sum_k \frac{SV^{\mathrm{E}}_{s,i,k}\cdot D_{s,i,k}}{(1+ir)^{|T|}} \tag{42}$$

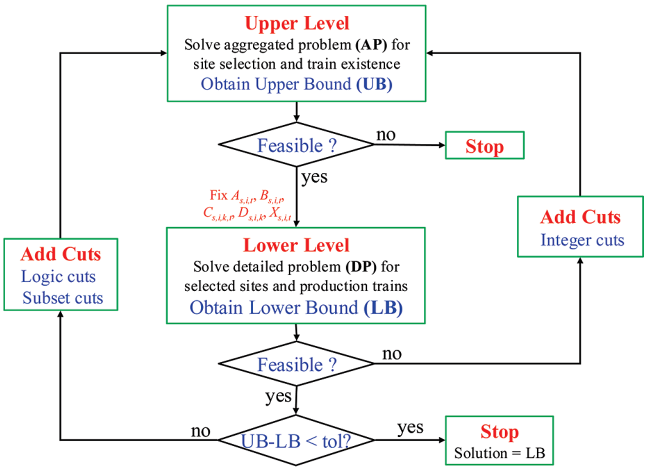

Figure 2. Bilevel decomposition algorithm flowchart.

subject to constraints (15)−(26), (28)−(32) and (34)−(36).

Note that in (DP), the variables $A_{s,i,t}$, $B_{s,i,t}$, $C_{s,i,k,t}$, $D_{s,i,k}$, and $X_{s,i,t}$ are treated as fixed values, which are given by (AP) in each iteration of the algorithm. The objective function of model (DP) given in (42) is similar to the objective function (33) of the original model (P), except that those binary variables discussed above are fixed in (42). Major constraints in this model include constraints (15)−(24) for capacity transformation with reactor modification, constraints (25)−(26) for defining the production capacities, constraints (28)−(31) for production−distribution planning, constraints (34)−(36) for linearization, and non-negative constraints (32). Note that binary variables $A_{s,i,t}$, $B_{s,i,t}$, $C_{s,i,k,t}$, $D_{s,i,k}$, and $X_{s,i,t}$ are all fixed, thus those logic constraints (1)−(14) for capacity expansion, capacity reduction, and train existence are neglected in this model.

Since any feasible solution of (DP) is also a feasible solution of the original problem (P), the following property holds:. *Property 2:* Problem (DP) yields a lower bound to the solution of problem (P).

In summary, the upper level problem (AP) neglects the capacity transformations due to reactor modification and provides an upper bound to the optimal NPV. The detailed problem (DP) is solved in the reduced space of binary variables for selected production trains to determine the detailed capacity, production, and distribution decisions, and provides a lower bound to the objective. As shown in Figure 2, the bilevel decomposition algorithm solves iteratively the upper level problem (AP) and the lower level problem (DP). The relaxed problem (AP) is updated at each iteration by adding the integer cuts, superset cuts, and subset cuts in order to accelerate the convergence. Convergence is achieved when the global lower and global upper bounds lie within a specified tolerance or the global upper bound is less than the global upper bound. Note that due to the inclusion of integer cuts, the global upper bound might be less than the global upper bound after sufficiently large number of iterations. If at some iteration the global upper bound is less than the global lower bound, it implies that no better solution can be obtained in the future iterations and the algorithm should stop at this iteration.

The steps of the bilevel decomposition algorithm are presented below.

Step 1: Initialization. Set iteration count iter = 1, global lower bound GLB = −∞, and global upper bound GUB = +∞.

Step 2. At iteration iter, solve the upper level aggregated problem (AP). Denote the optimal objective function value (upper bound obtained in this iteration) as φ^{iter} and the optimal solution of variables $A_{s,i,t}$, $B_{s,i,t}$, $C_{s,i,k,t}$, $D_{s,i,k}$, and $X_{s,i,t}$ as ($A^*_{s,i,t}$ iter, $B^*_{s,i,t}$ iter, $C^*_{s,i,k,t}$ iter, $D^*_{s,i,k}$ iter, $X^*_{s,i,t}$ iter). Define

$$C1_{\text{iter}} = \{(s,i,k,t)|C^{*\text{iter}}_{s,i,k,t} = 1\} \tag{43}$$

$$C0_{\text{iter}} = \{(s,i,k,t)|C^{*\text{iter}}_{s,i,k,t} = 0\} \tag{44}$$

$$X1_{\text{iter}} = \{(s,i,t)|X^{*\text{iter}}_{s,i,t} = 1\} \tag{45}$$

$$X0_{\text{iter}} = \{(s,i,t)|X^{*\text{iter}}_{s,i,t} = 0\} \tag{46}$$

If φ^{iter} < GUB, then update GUB = φ^{iter}.

Step 3. For fixed ($A^*_{s,i,t}$ iter, $B^*_{s,i,t}$ iter, $C^*_{s,i,k,t}$ iter, $D^*_{s,i,k}$ iter, $X^*_{s,i,t}$ iter), solve (DP). If (DP) is infeasible, add the following integer cut to (AP), and then go to Step 2.

$$\sum_{(s,i,k,t)\in C1_{\text{iter}}} C_{s,i,k,t} + \sum_{(s,i,t)\in X1_{\text{iter}}} X_{s,i,t} - \sum_{(s,i,k,t)\in C0_{\text{iter}}} C_{s,i,k,t} - \sum_{(s,i,t)\in X0_{\text{iter}}} X_{s,i,t} \le |C1_{\text{iter}}| + |X1_{\text{iter}}| - 1 \tag{47}$$

Step 4. If (DP) is feasible, denote the optimal objective function value (lower bound obtained in this iteration) as Φ^{iter} and the optimal solution as ($Y^*_{s,i,j,t}$ iter, $Z^*_{s,i,j,j',t}$ iter, $V^*_{s,i,j,k,t}$iter, $Q^*_{s,i,j,t}$ iter, $W^*_{s,i,j,p,t}$ iter, $\theta^*_{s,i,j,p,t}$ iter, $F^*_{s,m,p,t}$ iter, $Sa^*_{m,p,t}$ iter). If Φ^{iter} > GLB, then update GLB = Φ^{iter} and store the current optimal solution of the (DP). Add the following integer cuts (48.1)−(48.3) to (AP).

$$W_{s,i,j,p,t} \ge W^{*\text{iter}}_{s,i,j,p,t}\cdot\left(\sum_{(s,i,k,t)\in C1_{\text{iter}}} C_{s,i,k,t} + \sum_{(s,i,t)\in X1_{\text{iter}}} X_{s,i,t} - \sum_{(s,i,k,t)\in C0_{\text{iter}}} C_{s,i,k,t} - \sum_{(s,i,t)\in X0_{\text{iter}}} X_{s,i,t} - |C1_{\text{iter}}| - |X1_{\text{iter}}| + 1\right),$$
$$\forall\ s\in S, i\in I, j\in J, p\in P, t\in T \tag{48.1}$$

$$\theta_{s,i,j,p,t} \ge \theta^{*\text{iter}}_{s,i,j,p,t}\cdot\left(\sum_{(s,i,k,t)\in C1_{\text{iter}}} C_{s,i,k,t} + \sum_{(s,i,t)\in X1_{\text{iter}}} X_{s,i,t} - \sum_{(s,i,k,t)\in C0_{\text{iter}}} C_{s,i,k,t} - \sum_{(s,i,t)\in X0_{\text{iter}}} X_{s,i,t} - |C1_{\text{iter}}| - |X1_{\text{iter}}| + 1\right),$$
$$\forall\ s\in S, i\in I, j\in J, p\in P, t\in T \tag{48.2}$$

$$F_{s,m,p,t} \ge F^{*\text{iter}}_{s,m,p,t}\cdot\left(\sum_{(s,i,k,t)\in C1_{\text{iter}}} C_{s,i,k,t} + \sum_{(s,i,t)\in X1_{\text{iter}}} X_{s,i,t} - \sum_{(s,i,k,t)\in C0_{\text{iter}}} C_{s,i,k,t} - \sum_{(s,i,t)\in X0_{\text{iter}}} X_{s,i,t} - |C1_{\text{iter}}| - |X1_{\text{iter}}| + 1\right),$$
$$\forall\ s\in S, m\in M, p\in P, t\in T \tag{48.3}$$

Besides, the superset cuts (49.1)−(49.2) and subset cuts (50.1)−(50.2) are also added to (AP) to preclude the supersets and subsets of $X1_{\text{iter}} \cup C1_{\text{iter}}$.

$$\sum_{(s,i,k,t) \in C1_{\text{iter}}} C_{s,i,k,t} + \sum_{(s,i,t) \in X1_{\text{iter}}} X_{s,i,t} + C_{s,i,k,t} \le |C1_{\text{iter}}| + |X1_{\text{iter}}|, \quad \forall s \in S, i \in I, k \in K, t \in T \tag{49.1}$$

$$\sum_{(s,i,k,t) \in C1_{\text{iter}}} C_{s,i,k,t} + \sum_{(s,i,t) \in X1_{\text{iter}}} X_{s,i,t} + X_{s,i,t} \le |C1_{\text{iter}}| + |X1_{\text{iter}}|, \quad \forall s \in S, \quad i \in I, \quad t \in T \tag{49.2}$$

$$\sum_{(s,i,k,t) \in C0_{\text{iter}}} C_{s,i,k,t} + \sum_{(s,i,t) \in X0_{\text{iter}}} X_{s,i,t} + C_{s,i,k,t} \ge 1, \quad \forall s \in S, \quad i \in I, \quad k \in K, \quad t \in T \tag{50.1}$$

$$\sum_{(s,i,k,t) \in C0_{\text{iter}}} C_{s,i,k,t} + \sum_{(s,i,t) \in X0_{\text{iter}}} X_{s,i,t} + X_{s,i,t} \ge 1, \quad \forall s \in S, i \in I, t \in T \tag{50.2}$$

Step 5. If GUB − GLB ≤ ε (e.g., 10^{-3}), stop and output the stored optimal solution of (DP); otherwise, set iter = iter + 1 and go to Step 2.

Note that all the cuts are generated for the binary variables $C_{s,i,k,t}$ and $X_{s,i,t}$, because they are the only binary variables of (AP) and the values of $A_{s,i,t}$, $B_{s,i,t}$ and $D_{s,i,k}$ can be determined if $C_{s,i,k,t}$ and $X_{s,i,t}$ are fixed. Constraint (47) is the integer cut, which is employed to exclude the set of solutions explored by the algorithm in previous iterations. Constraints (48.1)−(48.3) are logic cuts according to the following property.

Property 3: Constraints (48.1)−(48.3) are valid integer cuts for the problem (AP).

Proof. The proof of this property is analogous to that of Iyer and Grossmann.[27] For given ($A^*_{s,i,t}$iter, $B^*_{s,i,t}$ iter, $C^*_{s,i,k,t}$iter, $D^*_{s,i,k}$iter, $X^*_{s,i,t}$ iter), the solution to (DP) gives an optimal $W^*_{s,i,j,p,t}$ iter, which is the optimal value for the original problem (P) when variables $C_{s,i,k,t}$ and $X_{s,i,t}$ are fixed to corresponding values. For the selected value $C_{s,i,k,t}$*iter and $X^*_{s,i,t}$ iter, the right-hand side of (48.1) is equal to $W^*_{s,i,j,p,t}$iter multiplied by 1. Therefore, the inequality $W_{s,i,j,p,t} \ge W^*_{s,i,j,p,t}$iter is enforced through the above equation. Clearly, these inequalities are valid cuts since the objective function has $W_{s,i,j,p,t}$ with negative coefficients. Note that for any other choice of $C^*_{s,i,k,t}$iter and $X^*_{s,i,t}$ iter, $W_{s,i,j,p,t} \ge 0$ dominates the equation, due to constraint (47). Therefore, we show that (48.1) is a valid integer cut for (AP). Using a similar approach, we can show that (48.2) and (48.3) are also valid integer cuts for (AP).

Constraints (49)−(50) are superset and subset cuts to reduce the search space and the number of iterations in the decomposition procedure. These cuts are motivated by the cuts proposed by Iyer and Grossmann.[27] Property 4 establishes the basis for the derivation of such cuts.

Property 4. Let $\text{BIN}_1^{\text{iter}} = C1_{\text{iter}} \cup X1_{\text{iter}}$ and $\text{BIN}_0^{\text{iter}} = C0_{\text{iter}} \cup X0_{\text{iter}}$ correspond to the optimal solution of (DP) in iteration iter. For all iterations IT > iter, if $\text{BIN}_1^{\text{iter}}$ is feasible in (AP), then any solution $\text{BIN}_1^{\text{IT}} \supset \text{BIN}_1^{\text{iter}}$ will result in a solution of (AP) such that $\text{NPV}_{(\text{AP})}^{\text{IT}} \le \text{NPV}_{(\text{AP})}^{\text{iter}}$ as in constraints (49.1)−(49.2). Furthermore, the subsets $\text{BIN}_1^{\text{IT}} \subset \text{BIN}_1^{\text{iter}}$ can be excluded from (DP) at iteration IT > iter as in constraints (50.1)−(50.2).

Proof. The proof of this property is also analogous to that of Iyer and Grossmann.[27] If any solution BIN_1^{IT}, such that $\text{BIN}_1^{\text{IT}} \supset \text{BIN}_1^{\text{iter}}$, is not selected before iteration iter, then it implies that the selection of any additional production time period or capacity level does not result in an increase in the NPV of the upper level problem (AP) due to constraints (1)−(14). Therefore, $\text{NPV}_{(\text{AP})}^{\text{IT}} \le \text{NPV}_{(\text{AP})}^{\text{iter}}$, and any superset of the optimal configuration $\text{BIN}_1^{\text{iter}}$ can be excluded from further consideration at the upper level for all iterations IT > iter.

Now consider any $\text{BIN}_1^{\text{IT}} \subset \text{BIN}_1^{\text{iter}}$. BIN_1^{IT} may either be feasible or infeasible in upper level problem (AP). If BIN_1^{IT} is infeasible in (AP), then we may exclude all these infeasible solutions for consideration in problem (DP). Consider only those solutions in BIN_1^{IT} that are feasible in (AP). Since the optimal solution of (AP) will consider as candidates $\text{BIN}_1^{\text{iter}}$ and all possible subsets BIN_1^{IT}, it follows that BIN_1^{IT} can be excluded from further consideration in (DP) for all iterations IT > iter.

Lagrangean Decomposition. Because constraints (1)−(26), (28), (29), and (34)−(36) are for each production site s, while constraints (30)−(31) are only for each market m, an intuitively obvious approach is to spatially decompose the network nodes by using Lagrangean decomposition. The idea is to dualize the interconnection constraints between the sites and markets in order to optimize each site and market individually. Specifically, the flows $F_{s,m,p,t}$ of product p between site s and market m at time period t are severed; we duplicate the flow variables so that one copy $F_{s,m,p,t}$ is used for the site constraints and one copy $F_{s,m,p,t}$ for the market constraints.[18] The new interconnection constraints are added as shown in eq (51).

$$\tilde{F}_{s,m,p,t} = F_{s,m,p,t} \tag{51}$$

We now rewrite the affected constraints (29) and (30) in terms of the proper interconnection variables. Constraint (29a) remains the same, while constraint (52) contains the duplicated flow variable $\tilde{F}_{s,m,p,t}$.

$$\sum_i \sum_j W_{s,i,j,p,t} = \sum_m F_{s,m,p,t}, \quad \forall s \in S, p \in P, t \in T \tag{29a}$$

$$S_{m,p,t} = \sum_s \tilde{F}_{s,m,p,t}, \quad \forall m \in M, \quad p \in P, \quad t \in T \tag{52}$$

We then apply Lagrangean decomposition by dualizing the interconnection constraints (51), which involves removing them from the constraint set and adding them to the objective function multiplied by the Lagrange multipliers $\lambda_{s,m,p,t}$. The Lagrangean relaxation problem (LRP) is given as follows.

(LRP)

$$\begin{aligned}
\max:\ \text{L_NPV} = {} & \sum_s \sum_m \sum_p \sum_t \frac{\text{price}_{m,p,t} \cdot \text{Sa}_{s,m,p,t}}{(1+ir)^t} \\
& - \sum_s \sum_m \sum_p \sum_t \frac{\eta_{s,m,p,t} \cdot F_{s,m,p,t}}{(1+ir)^t} - \sum_s \sum_i \sum_j \sum_p \sum_t \frac{\sigma_{s,i,j,p,t} \cdot W_{s,i,j,p,t}}{(1+ir)^t} \\
& - \sum_s \sum_i \sum_k \sum_t \frac{\beta^{\text{P}}_{s,i,k,t} \cdot DX_{s,i,k,t}}{(1+ir)^t} - \sum_s \sum_i \sum_k \sum_t \frac{\beta^{0}_{s,i,k,t} \cdot C_{s,i,k,t}}{(1+ir)^{t-1}} \\
& - \sum_s \sum_i \sum_j \sum_{j'} \sum_k \sum_t \frac{\beta^{\text{T}}_{s,i,j,j',k,t} \cdot ZD_{s,i,j,j',k,t}}{(1+ir)^{t-1}} - \sum_s \sum_i \sum_k \sum_t \frac{\beta^{\text{SD}}_{s,i,k,t} \cdot BD_{s,i,k,t}}{(1+ir)^{t-1}} \\
& + \sum_s \sum_{i \in I_N(s)} \sum_k \sum_t \frac{SV^{\text{N}}_{s,i,k,t} \cdot C_{s,i,k,t}}{(1+ir)^{|T|}} + \sum_s \sum_{i \in I_E(s)} \sum_k \frac{SV^{\text{E}}_{s,i,k} \cdot D_{s,i,k}}{(1+ir)^{|T|}} \\
& + \sum_s \sum_m \sum_p \sum_t \lambda_{s,m,p,t} \cdot (F_{s,m,p,t} - \tilde{F}_{s,m,p,t})
\end{aligned} \tag{53}$$

subject to site constraints (1)−(26), (28), (29), and (34)−(36), market constraints (31) and (52), variable bounds given in constraints (32.1)−(32.3).

For fixed multipliers $\lambda_{s,m,p,t}$, the model is now separable in sites s and all the markets m. The resulting subproblems for each site

$s \in S$ and all the markets are given below by (SP^s) and (MP), respectively, which can be solved independently for fixed multipliers $\lambda_{s,m,p,t}$. Note that (SP^s) involves a MILP subproblem for each site s, while (MP) is a linear programming (LP) subproblem for all the markets. The formulation of (SP^s) and (MP) are given as follows.

(SP^s)

$$\max: \text{S\_NPV}_s = \sum_m \sum_p \sum_t \lambda_{s,m,p,t} \cdot F_{s,m,p,t} - \sum_i \sum_j \sum_p \sum_t \frac{\sigma_{s,i,j,p,t} \cdot W_{s,i,j,p,t}}{(1+ir)^t} - \sum_i \sum_k \sum_t \frac{\beta^{\mathrm{P}}_{s,i,k,t} \cdot DX_{s,i,k,t}}{(1+ir)^t} - \sum_i \sum_k \sum_t \frac{\beta^{0}_{s,i,k,t} \cdot C_{s,i,k,t}}{(1+ir)^{t-1}} - \sum_i \sum_j \sum_k \sum_i \frac{\beta^{\mathrm{T}}_{s,i,j,j',k,t} \cdot ZD_{s,i,j,j',k,t}}{(1+ir)^{t-1}} - \sum_i \sum_k \sum_t \frac{\beta^{\mathrm{SD}}_{s,i,k,t} \cdot BD_{s,i,k,t}}{(1+ir)^{t-1}} \tag{54}$$

subject to constraints (1)−(26), (28), (29), (32), and (34)−(36) for a given site s.

(MP)

$$\max: \text{M\_NPV} = \sum_s \sum_m \sum_p \sum_t \frac{\text{price}_{m,p,t} \cdot \text{Sa}_{s,m,p,t}}{(1+ir)^t} - \sum_s \sum_m \sum_p \sum_t \frac{\eta_{s,m,p,t} \cdot F_{s,m,p,t}}{(1+ir)^t} - \sum_s \sum_m \sum_p \sum_t \lambda_{s,m,p,t} \cdot \tilde{F}_{s,m,p,t} \tag{55}$$

subject to

$$S_{m,p,t} = \sum_s \tilde{F}_{s,m,p,t}, \quad \forall\ m \in M, p \in P, t \in T \tag{52a}$$

$$d^{\mathrm{L}}_{m,p,t} \le \text{Sa}_{m,p,t} \le d^{\mathrm{U}}_{m,p,t}, \quad \forall\ m \in M, p \in P, t \in T \tag{31a}$$

$$\sum_m F_{s,m,p,t} \le \sum_i \sum_j \rho_{j,p} \theta_{s,i,j,p,t}, \quad \forall\ s, p, t \tag{56}$$

$$\sum_i \sum_j \tilde{W}_{s,i,j,p,t} = \sum_m F_{s,m,p,t}, \quad \forall\ s \in S, p \in P, t \in T \tag{57}$$

$$\sum_{p \in P(j)} \frac{\tilde{W}_{s,i,j,p,t}}{\rho_{j,p}} \le \max_{k \in K}\{\overline{Q}_{j,k}\} H_t, \quad \forall\ s \in S, i \in I, j \in J, t \in T \tag{58}$$

Note that constraints (56)−(58) are redundant in the Lagrangean relaxation problem (LRP) but may provide a tight relaxation for the market subproblem (MP). Specifically, constraint (56) comes from constraint (29), and constraints (57) and (58) come from constraint (28). Note that $\tilde{W}_{s,i,j,p,t}$ is an auxiliary variable and is not the same as the production amount $W_{s,i,j,p,t}$.

Let us denote the optimal objective function value of problems (LRP), (SP^s), and (MP) as L_NPV, S_NPV_s and M_NPV, respectively. Then we have

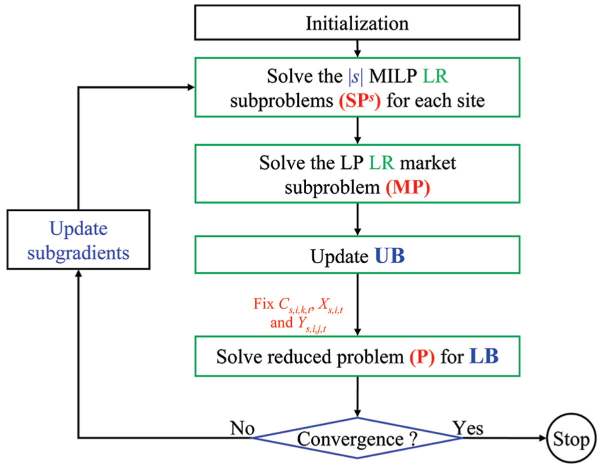

Figure 3. Lagrangean decomposition algorithm flowchart.

$$\text{L\_NPV} = \sum_s \text{S\_NPV}_s + \text{M\_NPV} \tag{58a}$$

Note that L_NPV provides an upper bound to the optimal NPV of the original problem (P), because L_NPV is the optimal objective function value of the Lagrangean relaxation problem (LRP).

A feasible solution of the original problem (P) naturally provides a valid lower bound to its optimal objective function value. To obtain a feasible solution, we fix all the binary variables to the values from the optimal solutions of the site subproblems (SP^s), and solve the original problem (P) in the reduced space to obtain a feasible solution. If the solution is feasible, we update the lower bound to the original problem; if not, we move on to the next iteration. It is worth mentioning that due to the potential infeasibility of the reduced problem, this heuristic might not be most efficient way to obtain the lower bound. However, we use it in this work because it is easy to implement and computationally efficient.

The algorithm iterates between the Lagrangean relaxation subproblems and the reduced problem. It stops when the maximum number of iterations has been reached or the algorithm converges to a solution within a predefined optimality tolerance. Due to the presence of integer variables in the model, a duality gap may exist and a solution within a predefined optimality tolerance might not be able to be obtained. Thus, an additional stopping criterion is that the change of Lagrange multiplier is sufficiently small.

To summarize, the steps of the Lagrangean decomposition algorithm are given as follows (also given in Figure 3):

Step 1: (Initialization). Initial multipliers $\lambda^0_{s,m,p,t}$ are chosen to be zero. Let the upper bound be UB = $+\infty$, lower bound be LB = $-\infty$, and iteration counter be iter = 0. Set the step length parameter $\mu = 2$.

Step 2. With fixed Lagrange multipliers $\lambda^{\text{iter}}_{s,m,p,t}$, first solve all the site subproblems (SP^s) and the market subproblem (MP) of the Lagrangean relaxation. Denote the optimal objective function value of site subproblems (SP^s) as $\text{S\_NPV}^{\text{iter}}_s$, the optimal objective function value of the market subproblem (MP) as $\text{M\_NPV}^{\text{iter}}$, the optimal solutions of the binary variables in (SP^s) as $(A^{*\text{iter}}_{s,i,t}, B^{*\text{iter}}_{s,i,t}, C^{*\text{iter}}_{s,i,k,t}, D^{*\text{iter}}_{s,i,k}, X^{*\text{iter}}_{s,i,t}, Y^{*\text{iter}}_{s,i,j,t}, Z^{*\text{iter}}_{s,i,j,j',t})$, and the optimal values of the interconnection variables $F_{s,m,p,t}$ and $\tilde{F}_{s,m,p,t}$ as $F^{*\text{iter}}_{s,m,p,t}$ and $\tilde{F}^{*\text{iter}}_{s,m,p,t}$, respectively.

dx.doi.org/10.1021/ie100559y | Ind. Eng. Chem. Res. 2011, 50, 4831–4849

Calculate

$$\text{L\_NPV}^{\text{iter}} = \sum_s S\_\text{NPV}_s + \text{M\_NPV}$$

If $\text{L\_NPV}^{\text{iter}} < \text{UB}$, update upper bound by setting $\text{UB} = \text{L\_NPV}^{\text{iter}}$. If more than two iterations of the subgradient procedure[16] are performed without any reduction of UB, then halve the step length parameter by setting $\mu = \mu/2$.

Step 3. Fixing the binary variable values as $C_{s,i,k,t} = C^{*\text{iter}}_{s,i,k,t}$, $X_{s,i,t} = X^{*\text{iter}}_{s,i,t}$ and $Y_{s,i,j,t} = Y^{*\text{iter}}_{s,i,j,t}$, and then solve the original problem (P) in the reduced space of binary variables.

If the reduced problem is feasible, denote the objective function as NPV^{iter}. If $\text{NPV}^{\text{iter}} > \text{LB}$, update the lower bound by setting $\text{LB} = \text{NPV}^{\text{iter}}$ and store the current solution. If the reduced problem is infeasible, set $\text{NPV}^{\text{iter}} = \text{NPV}^{\text{iter}-1}$.

Step 4. If gap = $(\text{UB} - \text{LB})/\text{UB} < \text{tol}$ (e.g., 10^{-3}), or $\|\lambda^t - \lambda^{t-1}\|^2 < \text{tol}$ (e.g., 10^{-2}) or the maximum number of iterations has been reached, set LB as the optimal objective function value and the current stored solution as the optimal solution.

Else, go to Step 5.

Step 5. Calculate the step size *Step*,[16,17]

$$\text{Step}^{\text{iter}} = \frac{\mu \cdot (\text{L\_NPV}^{\text{iter}} - \text{LB})}{\sum_s \sum_m \sum_p \sum_t (\tilde{F}^{*\text{iter}}_{s,m,p,t} - F^{*\text{iter}}_{s,m,p,t})^2}$$

Update the multipliers:[16−18]

$$\lambda^{\text{iter}+1}_{s,m,p,t} = \lambda^{\text{iter}}_{s,m,p,t} + \text{Step}^{\text{iter}} \cdot (\tilde{F}^{*\text{iter}}_{s,m,p,t} - F^{*\text{iter}}_{s,m,p,t}),$$
$$\forall\ s \in S,\ m \in M,\ p \in P,\ t \in T$$

Increment iter as iter + 1 and go to Step 2.

Note that this algorithm is based on the decomposition of the network nodes, that is, "spatial" decomposition. One drawback of this method is that the major network flows, specifically the product flows, are disconnected, and therefore solving the original problem (P) in the reduced space of binary variables can be infeasible because the solution of binary variables is for the Lagrangean relaxation of the original problem. The potential infeasibility might cause the completion of the algorithm to be difficult. Note that the potential infeasibility is due to the heuristic approach used to generate the lower bound. There could be other heuristic methods that are more efficient in generating lower bound and avoiding the infeasibility, but we use the one discussed above because it is easy to implement and computationally efficient. As discussed in Jackson and Grossmann,[22] one approach to overcome this issue is to use a temporal decomposition method, where the Lagrangean relaxation problem is decomposed by time periods. While this scheme can readily be applied in multiperiod planning problems with inventories, which only relate two adjacent time periods, most variables in this problem are related to capacity planning, thereby affecting multiple time periods simultaneously. Although the temporal decomposition scheme can still be applied to this problem, it would not be a straightforward task since it would require the introduction of many auxiliary linking variables. Thus, "temporal" decomposition is not a suitable approach for this problem.

6. NUMERICAL EXAMPLES

To illustrate the application of the proposed model and the performance of the proposed solution strategies, we present four numerical examples. The number of production sites, production trains at each site, markets, products, product families, and capacity levels considered in each example are given in Table 1. There are some existing production trains in each example, and the detailed information will be presented in the following sections. We can see that examples 1 and 2 are relatively small examples for illustration, while examples 3 and 4 are large-scale problems representative of real-world cases. In all the examples, we consider 10 time periods corresponding to 10 year planning horizon, although more time periods can be specified.

Table 1. Numbers of Production Sites, Trains, Markets, Products, Product Families and Capacity Levels of the Four Examples

	sites	trains per site	markets	products	product families	capacity levels
example 1	3	2	4	6	4	3
example 2	4	2	4	6	4	4
example 3	6	2	6	16	6	4
example 4	6	3	6	36	6	4

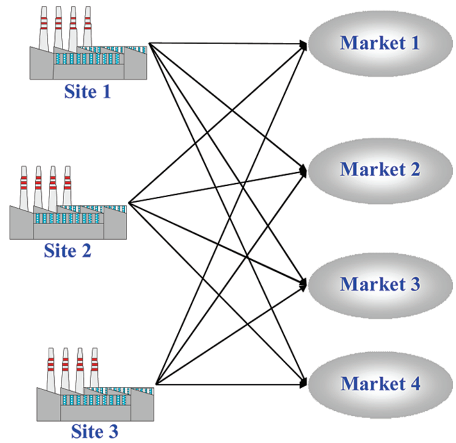

Figure 4. Network structure of example 1.

All the computational experiments are performed on an IBM T400 laptop with Intel 2.53 GHz CPU and 2 GB RAM. The proposed solution procedures are coded in GAMS 23.2.1,[34] and all the LP and MILP problems are solved using CPLEX 12.1.

The detailed problem sizes will be discussed in each example. In general, the full problem includes $|S|\cdot|T|\cdot(|I|+|I|\cdot|J|+|I_N|\cdot|K|)$ binary variables, The upper level aggregate problem (AP) in the bilevel decomposition has $|S|\cdot|T|\cdot(|I|+|I_N|\cdot|K|)$ binary variables, the lower level detailed problem (DP) includes $|S|\cdot|T|\cdot|I|\cdot|J|$ binary variables, the Lagrangean subproblem for each site includes $|T|\cdot(|I|+|I|\cdot|J|+|I_N|\cdot|K|)$ binary variables, the Lagrangean subproblem for markets and the original problem (P) in the reduced variable space with all the binary variables fixed do not include any binary variables.

Example 1. In the first example, we consider a production−distribution network with three production sites and four markets (as in Figure 4). Each production site has an existing production train and the space to install another production train. There are three possible capacity levels for the new trains to be installed. Six chemicals, A, B, C, D, E, and F are manufactured and distributed through this network. These six chemicals belong to four product families, and their relative production coefficients are given in Table 2.

The original problem (P) in the full variable space includes 390 binary variables, 9768 continuous variables, and 20074 constraints. We first solve the problem directly with GAMS/CPLEX 12.1 and it takes 92 CPU seconds to obtain an optimal solution with 0% optimality margin.

This example leads to a maximum NPV of $4.460 billion. The optimal solutions are given in Figure 5a,b. Figure 5a shows the optimal total production amounts of all the six chemicals in the 10-year planning horizon. We can see that the total production amount keeps increasing from around 340 Kton in year 1 to around 470 Kton in year 10, although the detailed production amount of each chemical varies over the years. The solution shows that chemical A should be consistently produced in the 10 years due to its profitability. It is also profitable to produce chemical B, but only up to the seventh year. We only need to produce a certain amount of chemical C in 7 out of the 10 years. This is presumably due to its relative narrow margin. Chemicals D and E are also produced over the planning horizon, except in year 9 and year 5, respectively. It is interesting to note that chemical F is only produced in years 7, 9 and 10, presumably because it is not profitable to introduce this chemical to the markets in the early stage of the planning horizon.

Table 2. Relative Production Coefficients of Chemicals in Each Product Family For Example 1

	A	B	C	D	E	F
product family 1	1	0.9		0.95		
product family 2		1	0.7		1.05	
product family 3			1			0.8
product family 4					0.95	1

The total capacity of the entire production–distribution network for each product family in each year is given in Figure 5b. The solution shows that total capacity should increase from around 350 Kton/year in year 1 to around 460 Kton/year in year 10. Specifically, the following are the changes in production capacity:

Year 1: two production trains are installed in sites 2 and 3 with 100 Kton/year capacity each for product family 1, in addition to the existing capacities for product families 1–3.

Year 2: no change in terms of production capacity.

Year 3: a new production train with 100 Kton/year capacity for product family 2 will be installed in site 1.

Year 4: a production train at site 3 that produces product family 3 with 30 Kton/year capacity will be modified to produce product family 1. This modification leads to zero capacity for product family 1.

Year 5: no change in terms of production capacity.

Year 6: a production train at site 3 that produces product family 1 with 100 Kton/year capacity will be modified to produce product family 3.

Year 7: no change in terms of production capacity.

Year 8: a production train at site 3 will be modified to produce product family 4. This change leads to zero capacity for product family 3.

Year 9: the production train that has been installed at site 2 in year 1 with 100 Kton/year capacity is modified from producing product family 1 in years 1–8 to producing product family 3 in years 9.

Year 10: no change in terms of production capacity.

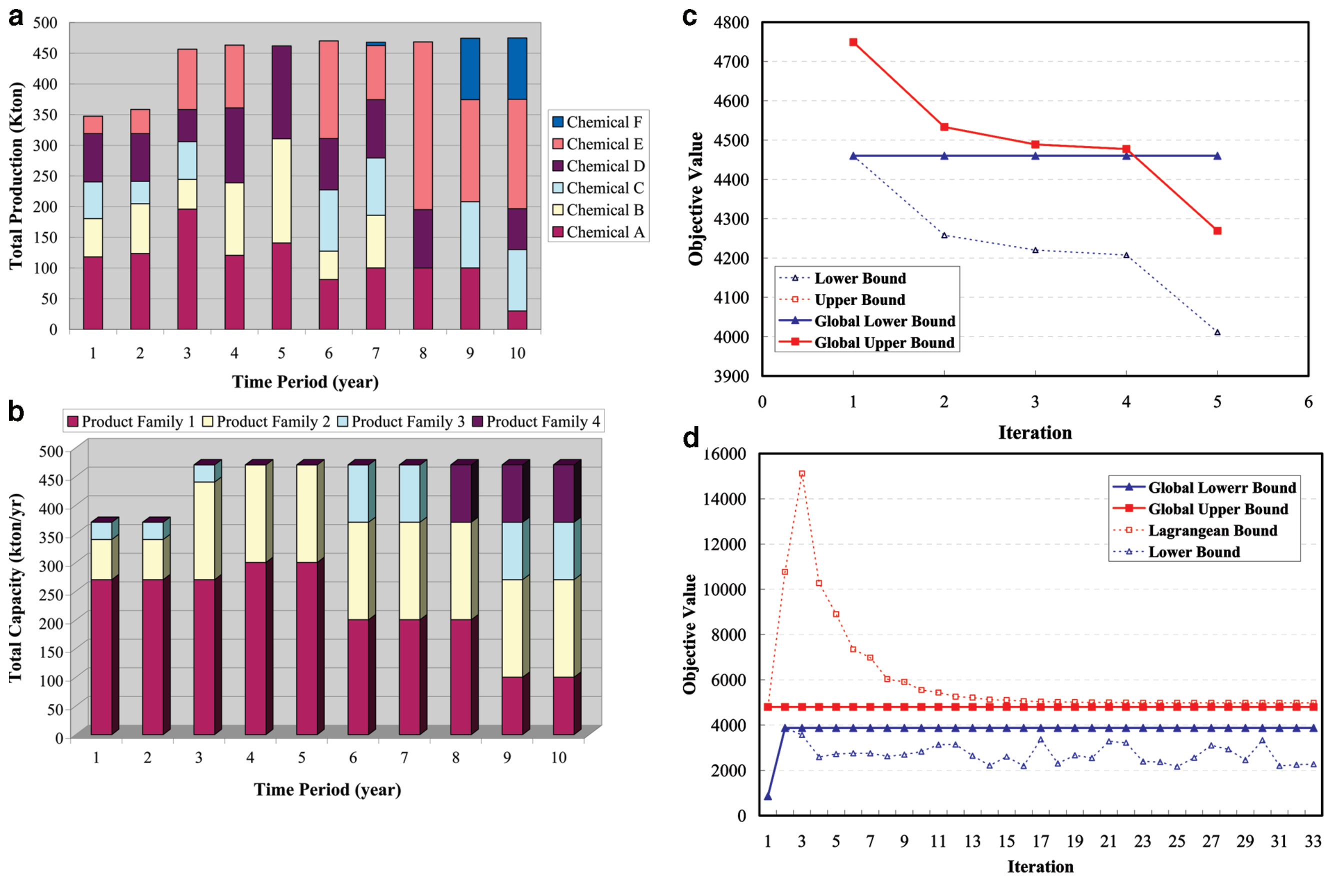

Figure 5. Optimal solution of example 1: (a) total production/sale amout of each chemical in example 1; (b) total capacity of each product family in example 1; (c) upper and lower bounds of bilevel decomposition in example 1; (d) upper and lower bounds of Lagrangean decomposition in example 1.

dx.doi.org/10.1021/ie100559y |Ind. Eng. Chem. Res. 2011, 50, 4831–4849

Table 3. Problem Sizes for the Four Examples[a]

	original problem (P)			aggregated problem (AP)			detailed problem (DP)			site subproblem (SP^s)		
	(full space)			(bilevel decomposition—iter. 1)			(bilevel decomposition—iter. 1)			(Lagrangean decomposition)		
	bin. var.	con. var.	const.	bin. var.	con. var.	const.	bin. var.	con. var.	const.	bin. var.	con. var.	const.
example 1	390	9 768	20 074	150	5 680	5720	240	8 700	18 435	130	2 936	6 406
example 2	650	16 350	36 050	330	9 116	10 046	320	15 040	33 924	150	3 800	8 768
example 3	1 120	61 684	103 782	400	33 098	20 054	720	52 320	90 490	140	9 206	16 610
example 4	1 620	158 106	184 056	540	97 146	39 326	1 080	124 920	149 230	210	21 730	28 816

[a] bin. var. = binary variable; con. var. = continuous variable; const. = constraints.

Table 4. Comparison of the Performance of the Algorithms for the Four Examples

problem	solving (P) with CPLEX 12.1 directly			bilevel decomposition				Lagrangean decomposition			
	solution	gap	time (s)	solution	gap	time (s)	iter.	solution	gap	time (s)	iter.
example 1	UB: 4.460 LB: 4.460	0.0%	92	UB: 4.461 LB: 4.460	0.0%	114	5	UB: 4.800 LB: 3.875	23.8%	186	16
example 2	UB: 6.272 LB: 6.272	0.0%	656	UB: 6.272 LB: 6.272	0.0%	241	3	UB: 6.801 LB: 5.323	27.8%	695	34
example 3	UB: 27.229 LB: 21.512	26.6%	36000	UB: 22.365 LB: 21.792	2.6%	1605	5	UB: 22.606 LB: 17.550	28.8%	1815	27
example 4	UB:123.908 LB: 86.217	43.72%	36000	UB: 90.998 LB: 90.009	1.1%	10529	2	UB: 91.820 LB: 79.138	16.0%	12627	31

The results show that it is optimal to install the capacities for product families 1 and 2 in the first few years, but in the later stage of the planning horizon some production facilities for producing product family 1 are modified to produce product family 4, since those products of product family 4 are more profitable. In summary, capacity expansions will happen at the beginning of years 1 and 4, and the capacity transformations with reactor modifications will take place in years 4, 6, 8, and 9. Capacity transformations do not change the total capacity, but increase and decrease the capacities for some product families, revealing that there are complex trade-offs between capital cost and the production cost.

We also use this example to test the computational performances of the proposed bilevel decomposition algorithm and the "spatial" Lagrangean decomposition method. The detailed problem sizes for the subproblems are given in Table 3. For the subproblems of the bilevel decomposition, we report the problem sizes for the aggregated problem (AP) and detailed problem (DP) in the first iteration, that is, without adding integer cuts. For the Lagrangean subproblems, we only report the problem sizes of the site subproblems (SP^s), since the market subproblem (MP) is a linear programming problem, which is much easier to solve. We can see that after the decomposition, we obtain smaller MILP problems that are less difficult to solve. The computational performance of the two types of decomposition methods are given in Figure 5c,d, as well as in Table 4. From Figure 5c, we can see that both upper bound and lower bound decrease as the iteration number increases. The reason is that we add integer cuts to the aggregated problem (AP) that causes the decrease of the upper bound. Although global upper bound is the same as the upper bound in each iteration, the global lower bound stays to be the same as the first lower bound, which has the maximum NPV among all the lower bounds. Table 4 shows that bilevel decomposition requires five iterations to converge, and the solution takes 114 CPUs, which is longer than solving the original problem (P) in the full variable space with CPLEX 12.1 directly (92 CPUs). However, as will be discussed in the following examples, the bilevel decomposition turns out to be more effective than the full space solution for large scale problems. The upper and lower bounds in each iteration of the Lagrangean decomposition algorithm are given in Figure 5d. We can see that the global upper and lower bounds do not change starting from iteration 2. This large duality gap (23.8%) is presumably due to both the large number of integer variables, and the fact that the overall mass balance is violated in the Lagrangean subproblems. Similar performance of Lagrangean decomposition method was repeated in the work by Chen and Pinto.[24]

Example 2. In the second example, we consider a relatively larger problem with four production sites and four markets (as in Figure 6). Each site can install at most two production trains and three production sites have an existing production train. There are four possible capacity levels for a new train to be installed. Six chemicals, belonging to four product families, are manufactured and distributed through this network. The relative production coefficients between chemicals and product families are given in Table 5.

As can be seen from Table 3, the original problem (P) in the full variable space includes 650 binary variables, 16350 continuous variables, and 36050 constraints. The aggregated problem (AP) at the first iteration of the bilevel decomposition has 330 binary variables, 9116 continuous variables, and 10046 constraints. The detailed problem (DP) has a similar number of binary variables, and a few more continuous variables and constraints. The site subproblem (SP^s) has relatively small size, with only 150 binary variables, 3800 continuous variables, and 8768 constraints. The

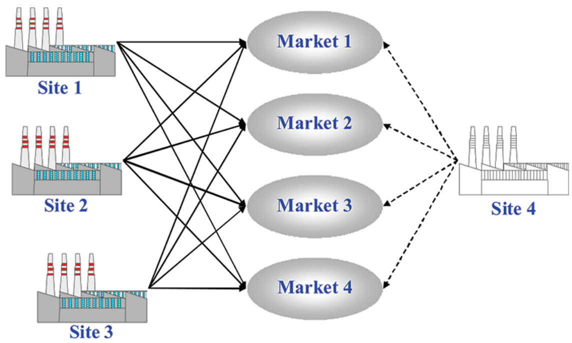

Figure 6. Network structure of example 2.

Table 5. Relative Production Coefficients of Chemicals in Each Product Family for Example 2

	A	B	C	D	E	F
product family 1	1	0.9		0.95		
product family 2		1	0.7		1.05	
product family 3	1.2		1		0.9	0.8
product family 4					0.95	1

computational performance is given in Table 4. For this example, the full space solution with CPLEX 12.1 takes 656 CPUs for optimal solution with 0% gap, but the bilevel decomposition requires only three iterations and 241 CPUs to obtain the same optimal solution. The performance of Lagrangean decomposition for this problem is not as good as bilevel decomposition—it takes 695 CPUs to obtain a suboptimal solution with 27.8% gap.

The maximum NPV of this example is $6.272 billion. Figure 7a,b shows the optimal solutions. The optimal total production amounts of all the chemicals in the 10-year planning horizon are given in Figure 7a. It is interesting to see that the total production amount increases from around 310 Kton in year 1 to around 680 Kton in year 3, and then oscillates between 620 Kton/year and 680 Kton/year from year 4 to year 10. The solution also shows that Chemical A should be consistently produced in the 10 years due to its profitability, although its production amount varies year by year. We only need to produce chemical B in years 2, 3, and 5, and to produce chemical A in years 1, 2, and 6. Chemical D is not produced or sold due to its relatively low margin. Chemicals E and F are only produced in the past few years, since their prices increase dramatically in the latter part of the planning horizon.

Figure 7b shows the total capacity of the entire production−distribution network for each product family in each year. We can see that the optimal total capacity increases from around 280 Kton/year in year 1 to around 580 Kton/year in year 3 and then keeps on the same level in the remaining years. Most of the production capacity is for product family 2, and there is zero capacity for product family 1. The capacity of product family 2 increases from zero in year 1 to around 330 Kton/year in year 8. There is only a small portion of the total capacity in years 6, 7, 9, and 10 for product family 4. We can see clearly that capacity expansion will happen in the first 3 years, while capacity transformations with reactor modifications will happen in years 4, 5, 7, 8, and 9. The result implies that it is optimal to expand the capacity at the beginning and then transform the capacity in the latter years for the adjustments of the capacity and production levels to achieve maximum profitability.

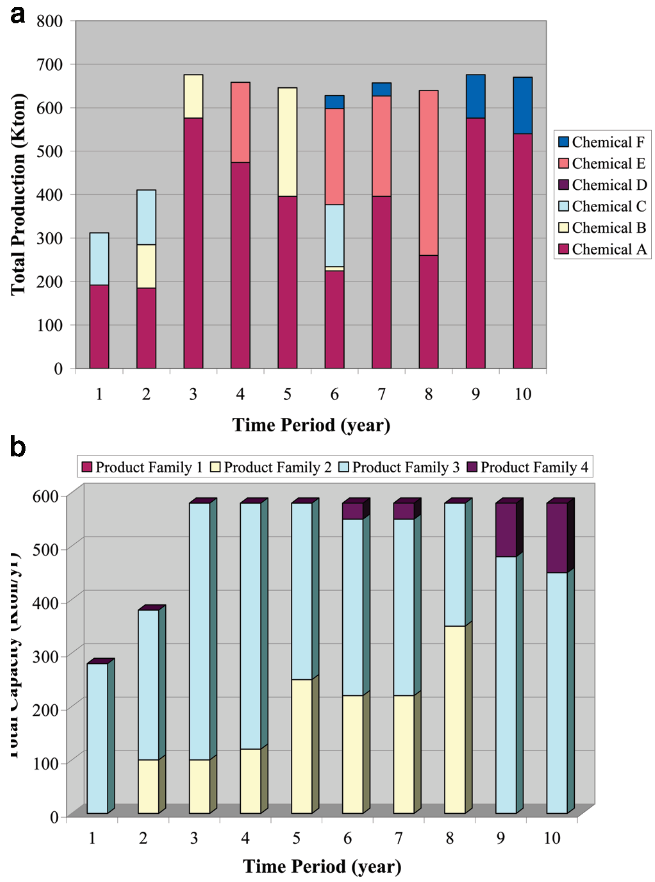

Figure 7. Optimal solution of example 2: (a) total production/sale amount of each chemical of example 2; (b) total production capacity of each product family of example 2.

Table 6. Relative Production Coefficients of Chemicals in Each Product Family for Example 3

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
product family 1	1				0.95											
product family 2		1									1			1.05		
product family 3			1													0.78
product family 4						1							1.3		1	
product family 5				0.7			0.6		1							
product family 6								1.1		1		0.9				

Example 3. The third example is for a relatively large production−distribution network with six production sites and six markets (as in Figure 1). Each site can install at most two production trains and sites 1, 2, and 3 have an existing production train each. There are four possible capacity levels for a new train to be installed at each site. Sixteen chemicals, belonging to six product families, are manufactured and distributed through this network. The relative production coefficients between chemicals and product families are given in Table 6.

The sizes of the full space problem and the decomposed subproblems are given in Table 3, and the computational performance of the algorithms can be found in Table 4. We can see that the full MILP problem includes 1120 binary variables, 61684 continuous variables, and 103782 constraints. After decomposition, each subproblem has relatively small size. Solving the full problem directly with CPLEX 12.1 leads to a suboptimal solution with 26.6% optimality gap after 10 CPU hours. However, the bilevel decomposition needs five iterations

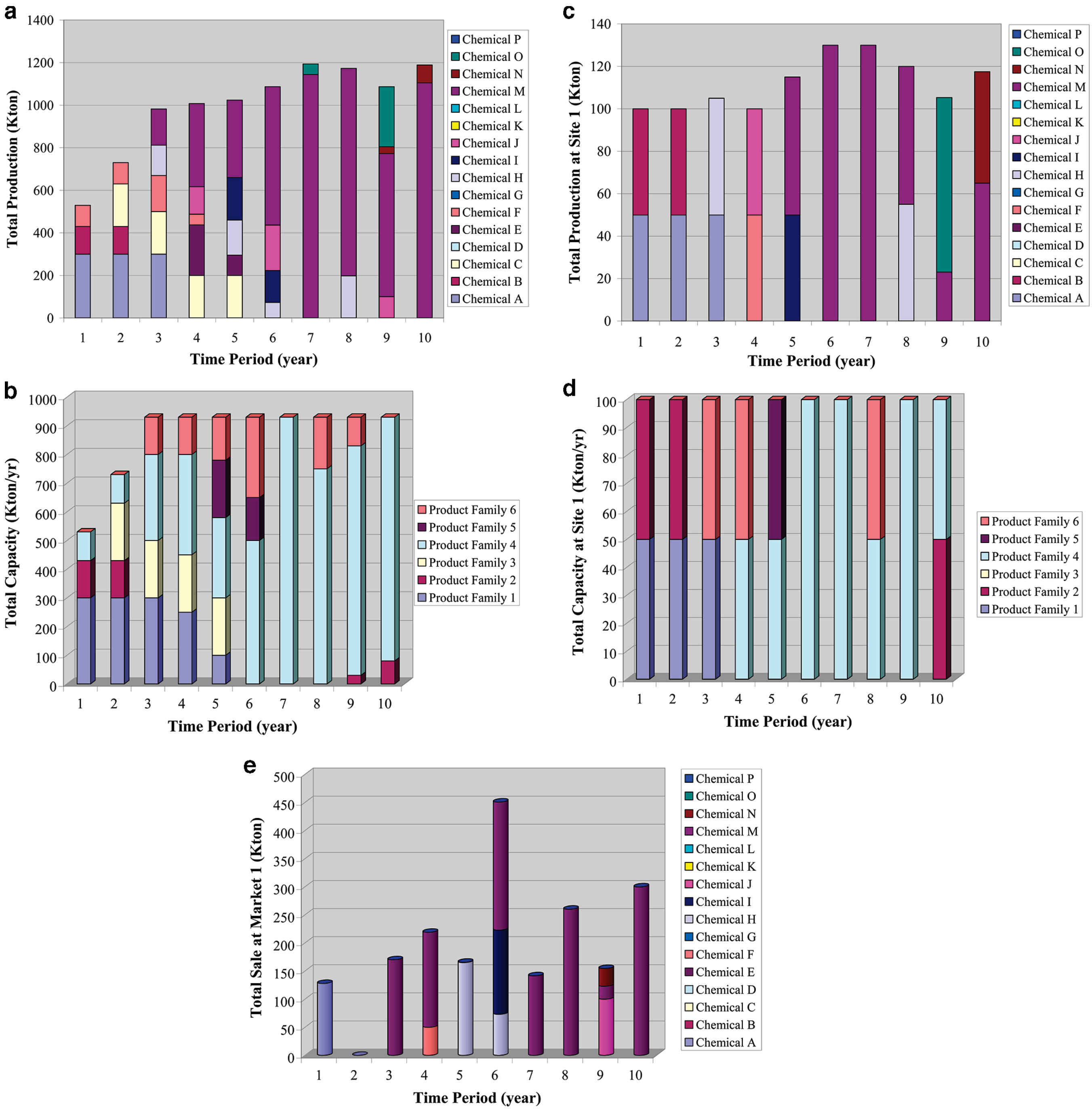

Figure 8. Optimal solution of example 3: (a) total production/sale amout of each chemical of example 3; (b) total production capacity of each product family of example 3; (c) total production amout of each chemical at site 1; (d) total production capacity of each product family at site 1; (e) total sale amout of each chemical at market 1.

and 1605 CPUs for a solution with 2.6% optimality gap. It takes Lagrangean decomposition 1815 CPUs and 27 iterations for sub-optimal solution with 28.8% gap. The comparison between these solution approaches shows that bilevel decomposition has much better performance than full space solution and Lagrangean decomposition, in terms of both computational time and optimality gap.

The best known solution for this example leads to a NPV of \$21.792 billion. The optimal total production amounts of all the chemicals in the 10 year planning horizon are given in Figure 8a. We can see that the total production/sale amount increases from around 460 Kton/year in year 1 to around 1190 Kton/year in year 10. No chemical is produced throughout the 10-year horizon, but a large portion of the total production amounts from year 3 to year 10 is dedicated to Chemical M. The total capacity of the entire production−distribution network of each product family in each year is given in Figure 8b. We can see that the total capacity increases from around 510 Kton/year in year 1 to around 920 Kton/year in year 10. It is interesting to note that the production capacity of product family 1 decreases from year 1 to year 5, and the capacity of product family 4 increases from year

Table 7. Relative Production Coefficients of Chemicals in Each Product Family for Example 4

	A	B	C	D	E	F	G	H	I
product family 1	1	0.95	0.7	1.2	0.95	0.54			
product family 2							1	1.3	0.88
	J	K	L	M	N	O	P	Q	R
product family 2	0.56	1.2	0.74						
product family 3				1	0.78	1.05	0.97	1.23	0.66
	S	T	U	V	W	X	Y	Z	AA
product family 4	1	1.3	0.95	1.02	1.1	0.77			
product family 5							1	1.2	0.74
	AB	AC	AD	AE	AF	AG	AH	AI	AJ
product family 5	2.41	1.1	1.06						
product family 6				1	0.9	1.1	0.84	0.67	1.03

1 to year 10. The detailed production amount of each chemical in site 1, and the detailed capacity of each product family in site 1 are given in Figure 8c,d. Specifically, the following are the changes in production amount and capacity at site 1 throughout the planning horizon:

Year 1: a new train with 50 Kton/year capacity for product family 2 will be installed at site 1, in addition to an existing train for product family 1 with 50 Kton/year capacity. This is the only capacity expansion to happen at site 1 throughout the planning horizon. As shown in Figure 8c, the capacity of site 1 leads to a production amount of 50 Ktons of chemical A and 50 Ktons of chemical B.

Year 2: no change in terms of production amount and capacity.

Year 3: the train producing product family 2 is modified to produce product family 6. This change yields a production amount of 55 Ktons of chemical H.

Year 4: the existing train producing product family 1 will be modified to produce product family 4. The consequence is that the total capacity at site 1 does not change, but we produce 50 Ktons of chemicals F and J each.

Year 5: we further transform the capacity of product family 6 to product family 5, and this change yields the production of chemicals I and M.

Year 6: all the capacity at site 1 is transformed to product family 4, and chemical M is the only chemical produced at this site.

Year 7: no change in terms of production amount and capacity.

Year 8: part of the capacity of product family 4 will be transformed to product family 6, and 55 Ktons of chemical H are produced.

Year 9: we produce chemical O instead of chemical H due to the change of prices.

Year 10: chemical N is introduced to the market after part of the capacity for product family 4 is transformed to product family 2.

The total sale amounts of each chemical in market 1 are given in Figure 8e. Although there is a combination of chemicals sold to this market in each year, we can see that a majority of chemical M is sold to this market from year 3 to 10, due to its relatively high prices in those years at this market. Besides, it is interesting to see that we do not sell any chemicals to this markets in year 2. These zero sale amounts show that the total demand from all the markets is higher than the total production capacity, and thus the optimal sale profiles show that best product portfolio.

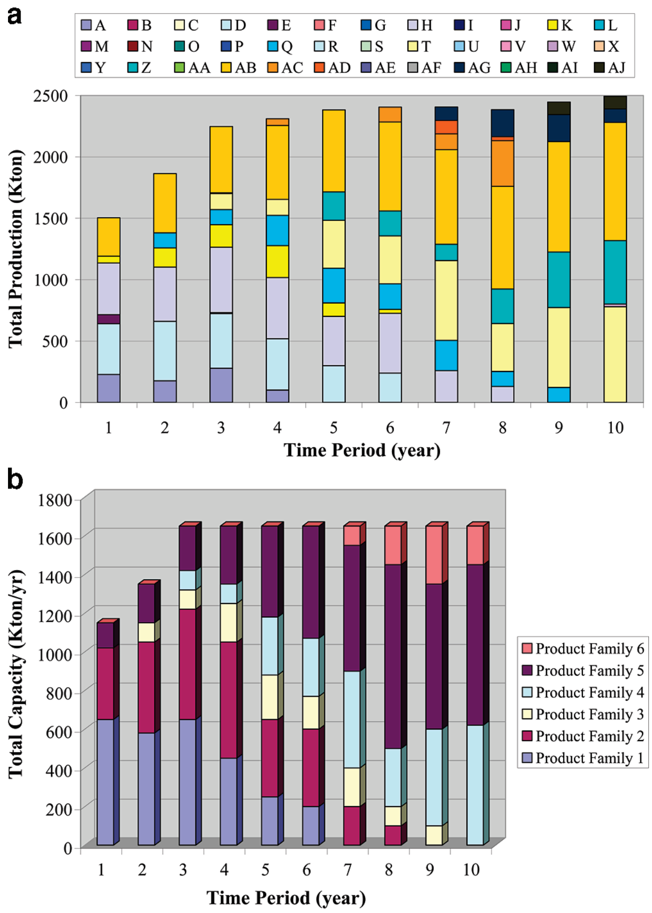

Figure 9. Optimal solution of example 4: (a) total production/sale amout of each chemical of example 4; (b) total production capacity of each product family of example 4.

Example 4. In the last example, we consider a large-scale problem with 6 production sites and 6 markets (as in Figure 1) for the production and distribution of 36 chemicals, which are categorized into 6 product families. Each site can install at most 3 production trains and 3 production sites have an existing production train each. There are 4 possible capacity levels for a new train to be installed at each site. The relative production coefficients between chemicals and product families are given in Table 7.

As shown in Table 3, the original problem (P) in the full variable space includes 1620 binary variables, 158 106 continuous variables, and 184 056 constraints. The aggregated problem (AP) at the first iteration of the bilevel decomposition has 540 binary variables, 97 146 continuous variables, and 39 326 constraints. The detailed problem (DP) has 1080 binary variables, 124 920 continuous variables, and 149 230 constraints. The site subproblem (SP^s) has relatively small size, with only 210 binary variables, 21 730 continuous variables, and 28 816 constraints. The computational performance is given in Table 4. For this example, solving the full problem directly with CPLEX 12.1 leads to a suboptimal solution with 43.72% optimality gap after 10 CPU hours. However, bilevel decomposition needs two iterations and 10 529 CPUs for a solution with 1.1% optimality gap. It takes Lagrangean decomposition 12 627 CPUs and 31 iterations for suboptimal solution with 16.0% gap. The results show that bilevel decomposition is much more effective than full space solution or Lagrangean decomposition for large scale problems.

The best known solution of this example has a maximum NPV of $90.009 billion. The optimal total production/sale amounts of all the chemicals in the 10 year planning horizon are given in Figure 9a. We can see that the total production amount increases from around 1500 Kton in year 1 to around 2500 Kton in year 10. It is interesting to note that at most eight chemicals are produced in the same year, although the production facilities can produce all the 36 chemicals. Figure 9a reveals the optimal product portfolio under the demand and price forecasting. Figure 9b shows the total capacity of the entire production−distribution network for each product family in each year. We can see that the optimal total capacity increases from around 1120 Kton/year in year 1 to around 1620 Kton/year in year 3 and then is kept at the same level in the remaining years. The result implies that it is optimal to expand the capacity in the first few years and then transform the capacity in the latter years for the adjustments of the capacity and production levels to achieve maximum profitability. Besides, we can also see that the capacities for product families 1 and 2 decrease over the years, while the capacities for product families 4 and 5 increase from year 3 to year 10. The adjusting of the production capacities is presumably due to the anticipated changes in the demand and/or prices of each of the products in the portfolio.

7. CONCLUSION

In this paper, we have proposed a multiperiod capacity, production, and distribution planning model for multisite networks consisting of several production trains for families of products. We formulated this model as an MILP and took into account three types of capacity modification actions: capacity expansion over time, plant shut down, and capacity transformation due to reactor modification, which is commonly used to adjust the capacities of production facilities in the specialty chemical industry. The MILP model takes into account multiple trade-offs and simultaneously predicts the optimal capacity adjustment plan, production levels, and sale profiles. Bilevel decomposition method and spatial decomposition schemes based on Lagrangean decomposition were developed to avoid simultaneously solving the resulting large-scale multiperiod MILP problem. Four example problems have been formulated and solved using the proposed methods. Numerical results illustrate the applicability of the proposed model and the benefits of applying the proposed decomposition approaches. Computational experiments clearly show that bilevel decomposition requires smaller computational times for all the examples, leading to solutions that are much closer to the global optimum, compared to full space solution and Lagrangean decomposition.

■ AUTHOR INFORMATION

Corresponding Author

*Tel.: 412-268-3642. Fax: 412-268-7139. E-mail: grossmann@cmu.edu.

Present Addresses

[§]Present address: Argonne National Laboratory, 9700 South Cass Avenue, Building 240, Argonne, IL 60439.

■ ACKNOWLEDGMENT

The authors gratefully acknowledge financial support from The Dow Chemical Company, the Pennsylvania Infrastructure Technology Alliance (PITA), and the National Science Foundation under Grant No. CMMI-0556090.

■ NOMENCLATURE

Sets/Indices

S = set of production sites indexed by s
I = set of production trains indexed by i
T = set of time periods indexed by t, t'
P = set of products indexed by p
J = set of product families indexed by j, j', j''
K = set of possible capacity levels for installation indexed by k
M = set of markets indexed by m

Subsets

$I_N(s)$ = subset of potential new production trains at site s, subset of I, indexed by i
$I_E(s)$ = subset of existing production trains at site s, subset of I, indexed by i
$P(j)$ = subset of products belonging to product family j, subset of P, indexed by p
$PM(j)$ = main product of product family j, subset of P

Parameters

$\rho_{j,p}$ = relative production coefficient of product family j and product p
$\overline{Q}_{j,k}$ = capacity in terms of product family j for capacity level k
$\lambda_{s,i}$ = construction lead time to install a new train as train i in site s
$\tau_{s,i,j,j'}$ = construction lead time to modify production train i in site s from producing product family j' at time $t-1$ to produce product family j
$\text{price}_{m,p,t}$ = price of product p at customer market m at time period t
$\sigma_{s,i,j,p,t}$ = unit production cost (includes raw material costs) of product p in train i site s when producing product family j at time period t
$\eta_{s,m,p,t}$ = unit transportation cost of product p from site s to market m at time t
$\beta^{P}_{s,i,k,t}$ = fixed cost of operating train type k as train i in site s at time period t
$\beta^{0}_{s,i,k,t}$ = cost of installing capacity level k as train i in site s at time period t
$\beta^{T}_{s,i,j,j',k,t}$ = cost of modifying train i in site s with capacity level k starting at the beginning of time t from producing product family j' at time $t-1$ to product family j
$\beta^{SD}_{s,i,k,t}$ = cost of shutting down train i in site s with capacity level k at time t
$SV^{N}_{s,i,k,t}$ = salvage value of train i in site s which is installed with capacity level k at time t
$SV^{E}_{s,i,k}$ = salvage value of the existing train i in site s with capacity level k
ir = interest rate

Binary Variables (0−1)

$C_{s,i,k,t}$ = 0−1 variable. Equal to 1 if capacity level k is selected to install at train i in site s at time t
$X_{s,i,t}$ = 0−1 variable. Equal to 1 if train i in site s exists at time period t
$Y_{s,i,j,t}$ = 0−1 variable. Equal to 1 if train i in site s is producing product family j at time period t
$A_{s,i,t}$ = 0−1 variable but can be relaxed as continuous variable. Equal to 1 if a new train is installed as train i in site s at time period t
$B_{s,i,t}$ = 0−1 variable but can be relaxed as continuous variable. Equal to 1 if train i in site s is shut down at time period t

$D_{s,i,k}$ = 0−1 variable but can be relaxed as continuous variable. Equal to 1 if capacity k is selected to install at train i site s

$Z_{s,i,j,j',t}$ = 0−1 variable but can be relaxed as continuous variable. Equal to 1 if train i in site s produces product family j' at time period $t-1$ and there is a modification for this train starting at the beginning of time period t for producing product family j

$V_{s,i,j,k,t}$ = 0−1 variable but can be relaxed as continuous variable. Equal to 1 if train i with capacity level k in site s is producing product family j at time period t

Continuous Variables (0 to $+\infty$)

$Q_{s,i,j,t}$ = capacity of train i in site s to produce product family j at time t

$W_{s,i,j,p,t}$ = production amount of product p at train i site s for product family j at time t

$F_{s,m,p,t}$ = shipping amount of product p from site s to market m at time t

$S_{m,p,t}$ = sale amount of product p in market m at time t

$\theta_{s,i,j,p,t}$ = production time of product p at train i site s for product family j at time t

$r_{s,i,j,p,t}$ = production rate of product p at train i site s for product family j at time t

Auxiliary Variables (0 to $+\infty$)

$ZD_{s,i,j,j',k,t}$ = continuous variables for the product of $Z_{s,i,j,j',t}$ and $D_{s,i,k}$

$BD_{s,i,k,t}$ = continuous variables for the product of $B_{s,i,t}$ and $D_{s,i,k}$

$DX_{s,i,k,t}$ = continuous variables for the product of $X_{s,i,t}$ and $D_{s,i,k}$

$QD_{s,i,j,k,t}$ = continuous variables for the product of $Q_{s,i,j,t}$ and $D_{s,i,k}$

$\tilde{W}_{s,i,j,p,t}$ = continuous variable for Lagrangean decomposition

REFERENCES

(1) Wassick, J. M. Enterprise-wide optimization in an integrated chemical complex. *Comput. Chem. Eng.* **2009**, *33*, 1950–1963.

(2) Grossmann, I. E. Enterprise-wide optimization: A new frontier in process systems engineering. *AIChE J.* **2005**, *51*, 1846–1857.

(3) Varma, V. A.; Reklaitis, G. V.; Blau, G. E.; Pekny, J. F. Enterprise-wide modeling & optimization—An overview of emerging research challenges and opportunities. *Comput. Chem. Eng.* **2007**, *31*, 692–711.

(4) Sahinidis, N. V.; Grossmann, I. E. Optimization model for the long range planning in the chemical industry. *Comput. Chem. Eng.* **1989**, *13* (9), 1049–1063.

(5) Sahinidis, N. V.; Grossmann, I. E. Multiperiod investment model for processing networks with dedicated and flexible plants. *Ind. Eng. Chem. Res.* **1991**, *30* (6), 1165–1171.

(6) Liu, M. L.; Sahinidis, N. V. Optimization in process planning under uncertainty. *Ind. Eng. Chem. Res.* **1995**, *35* (11), 4154–4165.

(7) Bok, J.-K.; Lee, H.; Park, S. Robust investment model for long-range capacity expansion of chemical processing networks under uncertain demand forecast scenarios. *Comput. Chem. Eng.* **1998**, *22* (7−8), 1037–1049.

(8) Lee, H.-K.; Lee, I.-B.; Reklaitis, G. V. Capacity expansion problem of multisite batch plants with production and distribution. *Comput. Chem. Eng.* **2000**, *24*, 1597–1602.

(9) Tsiakis, P.; Shah, N.; Pantelides, C. C. Design of multi-echelon supply chain networks under demand uncertainty. *Ind. Eng. Chem. Res.* **2001**, *40*, 3585–3604.

(10) Papageorgiou, L. G.; Rotstein, G. E.; Shah, N. Strategic supply chain optimization for the pharmaceutical industries. *Ind. Eng. Chem. Res.* **2001**, *40* (1), 275–286.

(11) Cheng, L.; Subrahmanian, E.; Westerberg, A. W. Multi-objective decisions on capacity planning and production-inventory control under uncertainty. *Ind. Eng. Chem. Res.* **2004**, *43*, 2192–2208.

(12) Oh, H.-C.; Karimi, I. A. Regulatory factors and capacity-expansion planning in global chemical supply chains. *Ind. Eng. Chem. Res.* **2004**, *43*, 3364–3380.

(13) Levis, A. A.; Papageorgiou, L. G. A hierarchical solution approach for multisite capacity planning under uncertainty in the pharmaceutical industry. *Comput. Chem. Eng.* **2004**, *28*, 707–725.

(14) You, F.; Grossmann, I. E. Design of responsive supply chains under demand uncertainty. *Comput. Chem. Eng.* **2008**, *32* (12), 2839–3274.

(15) Naraharisetti, P. K.; Karimi, I. A.; Srinivasana, R. Supply chain redesign through optimal asset management and capital budgeting. *Comput. Chem. Eng.* **2008**, *32*, 3153–3169.

(16) Fisher, M. L. The Lagrangean relaxation method for solving integer programming problems. *Manage. Sci.* **1981**, *27* (1), 18.

(17) Fisher, M. L. An application oriented guide to Lagrangian relaxation. *Interfaces* **1985**, *15* (2), 2–21.

(18) Guignard, M.; Kim, S. Lagrangean decomposition: A model yielding stronger Lagrangean bounds. *Math. Program.* **1987**, *39*, 215–228.

(19) Sridharan, R. The capacitated plant location problem. *Eur. J. Oper. Res.* **1993**, *87*, 203–213.

(20) van den Heever, S. A.; Grossmann, I. E.; Vasantharajan, S. A Lagrangean decomposition heuristic for the design and planning of offshore hydrocarbon field infrastructures with complex economic objectives. *Ind. Eng. Chem. Res.* **2001**, *40* (13), 2857–2875.

(21) Maravelias, C. T.; Grossmann, I. E. Simultaneous Planning for New Product Development and Batch Manufacturing Facilities. *Ind. Eng. Chem. Res.* **2001**, *40* (26), 6147–6164.

(22) Jackson, J. R.; Grossmann, I. E. Temporal decomposition scheme for nonlinear multisite production planning and distribution models. *Ind. Eng. Chem. Res.* **2003**, *42* (13), 3045–3055.

(23) Neiro, S. M. S.; Pinto, J. M. A general modeling framework for the operational planning of petroleum supply chains. *Comput. Chem. Eng.* **2004**, *28*, 871–896.

(24) Chen, P.; Pinto, J. M. Lagrangean-based techniques for the supply chain management of flexible process networks. *Comput. Chem. Eng.* **2008**, *32* (11), 2505–2528.

(25) You, F.; Grossmann, I. E. Mixed-integer nonlinear programming models and algorithms for large-scale supply chain design with stochastic inventory management. *Ind. Eng. Chem. Res.* **2008**, *47* (20), 7802–7817.

(26) You, F.; Grossmann, I. E. Integrated multi-echelon supply chain design with inventories under uncertainty: minlp models, computational strategies. *AIChE J.* **2010**, *56*, 419–440.

(27) Iyer, R. R.; Grossmann, I. E. A bilevel decomposition algorithm for long-range planning of process networks. *Ind. Eng. Chem. Res.* **1998**, *37*, 474–481.

(28) Bok, J.-K.; Grossmann, I. E.; Park, S. Supply chain optimization in continuous flexible process networks. *Ind. Eng. Chem. Res.* **2000**, *39* (5), 1279–1290.

(29) Dogan, M. E.; Grossmann, I. E. A decomposition method for the simultaneous planning and scheduling of single-stage continuous multiproduct plants. *Ind. Eng. Chem. Res.* **2006**, *45* (1), 299.

(30) Raman, R.; Grossmann, I. E. Symbolic integration of logic in mixed integer linear programming techniques for process synthesis. *Comput. Chem. Eng.* **1993**, *27*, 909.

(31) Bagajewicz, M. J. On the use of net present value in investment capacity planning models. *Ind. Eng. Chem. Res.* **2008**, *47*, 9413–9416.

(32) Almansoori, A.; Shah, N. Design and operation of a future hydrogen supply chain. *Chem. Eng. Res. Des.* **2006**, *84*, 423–438.

(33) You, F.; Grossmann, I. E. Stochastic inventory management for tactical process planning under uncertainties: MINLP models and algorithms. AIChE J., In press. DOI: 10.1002/aic.12244.

(34) Rosenthal, R. E. *GAMS: A User's Manual*; GAMS Development Corporation: Washington, DC, 2008.