

# A New Optimization Model and a Customized Solution Method for Natural Gas Production Network Design and Operation <sup>†</sup>

Dan Li and Xiang Li\*

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

E-mail: xiang.li@queensu.ca

## Abstract

This paper proposes to tackle integrated design and operation of natural gas production networks under uncertainty, using a new two-stage stochastic programming model, a novel reformulation strategy, and a customized global optimization method. The new model addresses material balances for multiple key gas components, pressure flow relationships in gas wells and pipelines, and compressor performance. This model is a large-scale nonconvex mixed-integer nonlinear programming problem that cannot be practically solved by existing global optimization solvers or decomposition-based optimization methods. With the new reformulation strategy, the reformulated model has a better decomposable structure, and then a new decomposition-based global optimization method is developed for efficient global optimization. In the case study of an

---

<sup>†</sup>This is the peer-reviewed version of the following article:  
Li, D. and X. Li, "A new optimization model and a customized solution method for natural gas production network design and operation", *AIChE Journal*, DOI: 10.1002/aic.15428, 2016,  
which has been published in final form at <http://onlinelibrary.wiley.com/doi/10.1002/aic.15428/abstract>.  
This article may be used for noncommercial purposes in accordance with Wiley Terms and Conditions for Self-Archiving.

industrial natural gas production system, it is shown that the proposed modeling and optimization methods enable efficient solution, and the proposed optimization method is faster than a state-of-the-art decomposition method by at least an order of magnitude.

## Introduction

As one of the major components of global energy, natural gas was the fastest-growing of the fossil fuels in last year.<sup>1</sup> Compared with other fossil fuels such as coal and oil, natural gas yields relatively low emission of greenhouse gases and pollutants, i.e., sulfide, nitrogen oxides and particulates. The increased demand for natural gas has motivated the development of new natural gas production networks or expansion of existing ones. A gas production network usually requires large investment costs and has a long life time, so considering the design and the long-term operating problems together in an optimization framework can help to achieve the most economic designs. In addition, the many uncertain factors associated with the new network often need to be explicitly included in the optimization framework, in order to ensure that the network can provide enough gas products under different uncertainty scenarios while optimizing the expected total profit.

In a natural gas production network, raw natural gases from different gas wells are boosted via compressors, mixed at preprocessing facilities or simple pipeline junctions, and finally gathered at terminal processing plants that produce the final gas products. Preprocessing facilities, typically gas platforms at reservoirs, are used for oil separation and water removal. Oil separation is to separate dissolved natural gas from oil, and water removal (dehydration) is to reduce corrosion and make raw natural gas amenable to transport. Raw natural gases from different reservoirs usually contain different amounts of contaminants (such as CO, N<sub>2</sub>, CO<sub>2</sub>, H<sub>2</sub>S, and water vapor),<sup>2</sup> which can hinder natural gas processing and pollute the environment. Therefore, contents of the contaminants need to be kept below certain thresholds before the gases enter the terminal processing plants. To obtain the desired dry natural gas products (mainly methane), the inlet flow pressure to a preprocessing or a termi-

nal processing facility needs to be within a certain range.<sup>3-6</sup> Moreover, the whole production network has to be robust against the fluctuation in raw gas compositions and the large variation in demands for the gas products. Ideally, an optimization model should describe gas flow compositions, pressures, and uncertainties explicitly, so that the optimization result can represent realistically feasible and optimal decisions.

Optimization models and methods have long been adopted for decision making of natural gas production systems. Initially, the application of optimization was limited due to lack of fast computing hardware and efficient numerical optimization methods. The design and operation problems were considered separately in order for practical solution. For design problems, classical ways are to develop linear programming (LP) or mixed-integer linear programming (MILP) models. For example, Aronofsky and Williams<sup>7</sup> solved an underground oil production problem through a LP formulation with fixed operation profile. Bohannon<sup>8</sup> and Sullivan<sup>9</sup> developed MILP models for specific infrastructure design and production planning problems. On the other hand, operation problems are usually cast as nonlinear programming (NLP) problems due to the need to address pressure constraints. For example, Rothfarb et al.<sup>10</sup> developed a NLP model for natural gas transmission systems and solved the NLP problem via fixing pressures at each node and applying a branch strategy for the tree structure of offshore natural gas pipeline systems; Murtagh<sup>11</sup> then solved a similar NLP problem by utilizing duality theory of nonlinear programming; Soliman and Murtagh<sup>12</sup> included pressure constraints and compressors in a large-scale gas pipeline network design problem and solved the problem with limited computing capability; Selot et al.<sup>13</sup> cast a short-term natural gas production system operation problem as a nonconvex mixed-integer nonlinear programming (MINLP) problem, which addresses nonlinear pressure-flow relations, well performance and compressor performance submodels, and production-sharing contracts.

With rapid improvement of computing hardware and significant advances in optimization theories and algorithms over the last several decades, solving the design and operation problems together becomes more and more realistic. In the process system engineering

community, efforts have been taken to develop novel models and optimization methods for integrated design and operation of oil and gas production systems. For example, Van den Heever et al.<sup>14</sup> developed a MINLP model for integrated design and operation of offshore oil and gas production systems, and nonlinearity of the model comes from the nonlinear reservoir submodel involved; Gupta and Grossmann<sup>15</sup> proposed a multiperiod MINLP model for planning of offshore oil and gas field infrastructure, where the nonlinearity comes from high order polynomials. Uncertainty and other problem specific features of oil and gas systems are also intensively studied via stochastic programming approaches, such examples include Goel and Grossmann<sup>16</sup>, You et al.<sup>17</sup>, Gupta and Grossmann<sup>18</sup>, Gupta and Grossmann<sup>19</sup> and Tarhan et al.<sup>20</sup>. Among these studies, Li et al.<sup>21</sup> developed a two-stage stochastic programming model that tracks the compositions of the gas flows throughout the gas production network, for different uncertainty realizations. The resulting MINLP problem is called the stochastic pooling problem, and this problem includes a large number of bilinear terms that introduce strong nonconvexity. Effort has also been taken to address integrated design and operation of natural gas transmission networks, where the pressures are more important quantities to track than gas compositions. In order to describe the relationship between pressure and gas flow, nonconvex nonlinear gas flow equations need to be included in the model, which can be solved via solving a convex optimization problem (as discussed by Raghunathan<sup>22</sup>, Misra et al.<sup>23</sup>).

The optimization models for integrated design and operation of natural gas production systems are usually nonconvex MINLP models, for which obtaining a global optimal solution is usually challenging. When the uncertainties are considered using the classical scenario approach,<sup>24</sup> the nonconvex MINLP can have a very large size and therefore cannot be solved practically by any state-of-the-art optimization solver. A viable approach to solve the large-scale nonconvex MINLP is to exploit its structure through a decomposition-based optimization method. There are mainly two types of rigorous decomposition methods. One is Lagrangian decomposition (LD),<sup>25,26</sup> which needs to be performed together with a branch-

and-bound framework to guarantee global optimality. The other is Benders decomposition (BD)<sup>27</sup> and generalized Benders decomposition (GBD),<sup>28</sup> which do not need an explicit branch-and-bound search framework, but are limited to problems with certain partially convex structures. Recently, Li et al.<sup>29</sup> developed a significant extension of BD and GBD, called nonconvex generalized Benders decomposition (NGBD), which can solve a wider range of nonconvex MINLPs. NGBD does not require the problem to have the special partially convex structure, but it requires that the "complicating" variables be all integers (in order to guarantee finite termination). Both LD and NGBD can be very efficient if the nonconvex subproblems to be solved are relatively easy problems; but if the nonconvex subproblems are challenging, the overall LD or NGBD procedure can be very slow.

This paper considers the integrated design and operation problem in which the directions of gas flows in the pipelines are known. This problem is cast as a large-scale nonconvex MINLP through the scenario-based two-stage stochastic programming approach, and this optimization model describes gas compositions, pressure flow relationships, and compressor performances, at different parts of the gas production network. This large-scale nonconvex MINLP cannot be solved by existing decomposition methods efficiently, because each operation subproblem (for each uncertainty scenario) is a very challenging nonconvex MINLP. In order to overcome this difficulty, the nonconvex MINLP formulation is reformulated via several steps, and the reformulation procedure is rigorous under mild assumptions. The resulting reformulation significantly reduces nonconvexity, and its nonconvex part and convex part are separable, allowing efficient solution of each operation subproblem via GBD. To the best of our knowledge, it is the first application of GBD to deterministic operation problem in process systems engineering. With the integration of the classical NGBD and the GBD procedure for solving the operation subproblems, a multi-loop NGBD method is developed to efficiently solve the proposed optimization model. It is the first time the multi-loop NGBD method is developed and applied.

The remaining part of the paper is organized as follows. First, the two-stage stochastic

programming model is presented in the context of a real industrial natural gas production system. Next, the model reformulation procedure is explained step by step, and the structure of the final reformulated model is analyzed. Then the multi-loop NGBD method is developed to exploit the structure of the reformulated model, and its advantage over the classical NGBD and the monolith approach is demonstrated via case study results. The paper ends with conclusions and discussions for future work.

## The Optimization Model for Integrated Design and Operation Under Uncertainty

The optimization model is presented for a real and representative industrial natural gas production system, called Sarawak Gas Production System (SGPS). This system was originally studied by Selot et al.<sup>13</sup> for optimal short-term operation, and then by Li et al.<sup>21</sup> for integrated design and operation through a stochastic pooling model. Figure 1 illustrates the superstructure of SGPS. Each ellipse in the figure refers to a gas field modeled as a gas platform, which contains multiple gas wells that are not explicitly showed in the figure but are listed in the appendix. Each gray circle in the figure refers to remote gas production subnetwork connected to SGPS, and it can also be simplified as a gas platform with available flow rate and inlet pressure ranges. Each rectangle refers to a gas platform, a slug-catcher,<sup>13</sup> or a simple pipeline conjunction, at which gas flows are mixed and split. Although a gas platform often involves oil-gas separation and water removal, the platform does not remove major impurities in the gas significantly. The three liquefied natural gas (LNG) plants at the end of the network superstructure, labeled as LNG1, LNG2, and LGN3, purify and convert raw gases into final LNG products. For a LNG plant to work properly, the composition of gas entering the plant needs to satisfy certain specification.

As we only consider the cases in which the directions of gas flows are known, the gas production network can be viewed as a directed graph  $(V, E)$ , where  $V$  is the set of all vertices

and  $E$  is the set of all directed edges. The vertices represent gas wells, platforms, pipeline junctions, and terminal LNG plants, and the directed edges represent pipelines connecting two vertices. The proposed optimization model is constructed based on previously published work.<sup>13,21</sup> It contains a stochastic pooling model,<sup>21</sup> which is a two-stage, scenario-based stochastic programming model that describes the topology relationships among the vertices and edges, and the material balances for each major components of the gas flows. It also combines a detailed operation model for each uncertainty scenario based on the work by Selot et al.<sup>13</sup>, which describes the relationships between pressures, flows, and energy consumptions in the system. The objective of optimization is to determine the system network structure that maximize the expected net present value (NPV) of the system over its life span. Table 1 shows all the symbols used in the model. The set elements and parameter values are showed in the appendix. Next, each part of the optimization model is presented step by step.

## The Stochastic Pooling Model

The stochastic pooling model is a two-stage, scenario based stochastic programming model. In this model, the first-stage decisions are network structure decisions that need to be implemented before the construction of the real physical system. The second-stage decisions are operational decisions on flow rates, which are included to predict the profits of operating the system under different uncertainty scenarios. Each uncertainty scenario is indexed by subscript  $w$ . This model consists of the following Eq. (1-12).

$$f_{j,k,q,w} = s_{j,k,w} \left( \sum_{i \in \{i | (i,j) \in E\}} f_{i,j,q,w} \right), \quad (1)$$

$$\forall (j,k) \in \{(j,k) \in E \mid j \notin V^W \cup V^S\}, \quad \forall q \in \Omega, \quad \forall w \in \Phi$$

$$s_{j,k,w} \geq 0, \quad \forall (j,k) \in \{(j,k) \in E \mid j \notin V^W \cup V^S\}, \quad \forall w \in \Phi \quad (2)$$

$$\sum_{k \in \{k | (j,k) \in E\}} s_{j,k,w} = 1, \quad \forall j \in V \setminus (V^W \cup V^S \cup V^T) \quad (3)$$

$$y_i^v \geq y_{i,j}, \quad \forall (i,j) \in E \quad (4)$$

$$y_j^v \geq y_{i,j}, \quad \forall (i,j) \in E \quad (5)$$

$$y_i^v \leq \sum_{j \in \{j | (i,j) \in E\}} y_{i,j}, \quad \forall i \in V \setminus V^T \quad (6)$$

$$y_j^v \leq \sum_{i \in \{i | (i,j) \in E\}} y_{i,j}, \quad \forall j \in V \setminus (V^S \cup V^W) \quad (7)$$

$$y_{i,j} F_{i,j}^{L B} \leq \sum_{q \in \Omega} f_{i,j,q,w} \leq y_{i,j} F_{i,j}^{U B}, \quad \forall (i,j) \in E, \quad \forall w \in \Phi \quad (8)$$

$$y_i^v Z_{i,w}^{L B} \leq \sum_{q \in \Omega} \sum_{j \in \{j | (i,j) \in E\}} f_{i,j,q,w} \leq y_i^v Z_{i,w}^{U B}, \quad \forall i \in V \setminus V^T, \quad \forall w \in \Phi \quad (9)$$

$$y_k^v D_{k,w}^{L B} \leq \sum_{q \in \Omega} \sum_{j \in \{j | (j,k) \in E\}} f_{j,k,q,w} \leq y_{k,w} D_{k,w}^{U B}, \quad \forall k \in V^T, \quad \forall w \in \Phi \quad (10)$$

$$\left( \sum_{q \in \Omega} \sum_{j \in \{j | (j,k) \in E\}} f_{j,k,q,w} \right) K_{k,q}^{L B} \leq \sum_{j \in \{j | (j,k) \in E\}} f_{j,k,q,w}, \quad \forall k \in V^T, \quad \forall q \in \Omega, \quad \forall w \in \Phi \quad (11)$$

$$\left( \sum_{q \in \Omega} \sum_{j \in \{j | (j,k) \in E\}} f_{j,k,q,w} \right) K_{k,q}^{U B} \geq \sum_{j \in \{j | (j,k) \in E\}} f_{j,k,q,w}, \quad \forall k \in V^T, \quad \forall q \in \Omega, \quad \forall w \in \Phi \quad (12)$$

Eq. (1-3) stand for the material balances for different components (indexed by  $q$ ) of the gas flows. Eq. (1) reflects the fact that different components of gas are split at a vertex with the same ratios (denoted by  $s$ ). Specifically, the rate of component  $q$  in a flow leaving a vertex  $j$  equals to the rate of the same component in all flows entering node  $j$  times a split fraction for this flow. This relationship is included in the model for all gas flows except those coming out of a starting vertex (i.e., gas wells or remote subnetworks), as the ratios of their components are parameters that are independent of the optimization decisions. Eq. (2-3) mean that the split fractions are nonnegative, and the sum of the split fractions of all flows leaving vertex  $j$  equals to 1. Note that, Eq. (1-3) generally introduce nonconvexity, but they can be reduced to a set of linear constraints for a vertex  $j$ , if this vertex has only one outlet flow.

Eq.(4-5) denote topology restriction of the network structure, they mean that edge  $(i, j)$  can be developed only when both vertex  $i$  and vertex  $j$  are developed. Eq.(6-7) are introduced to prevent developing a vertex that is not connected to any other vertex. Note that an optimal solution cannot contain an isolated vertex (with certain investment cost) no matter whether Eq.(6-7) are included in the model, but including these constraints can expedite the solution of the optimization problem.

Eq. (8-12) pose bounds on the gas flow rates in the pipelines, the total gas flow rates going through a vertex, and component contents of gas in the terminal vertex. When a vertex or edge is not to be developed, the related flow rates are forced to be zero by the bounding constraints.

## The Pressure Flow Relationships

Gas flows are driven in the gas production network by pressure differentials, so they are not only restricted by the capacities of gas wells and pipelines, but also restricted by reservoir pressures and pressure rises that can be provided by the compressors. In Li et al.<sup>21</sup>, pressure flow relationships are not considered in the integrated design and operation model, as they can make the optimization model too complicated for practical solution. Selot et al.<sup>13</sup> considers the pressure flow relationships for deterministic operation problems only, and the resulting problem is already very challenging to solve. We incorporate into our proposed optimization model equations describing the pressure-flow-energy relationships in wells, pipelines and compressors (called well performance, pipeline performance, and compression performance models, respectively), based on the work in Selot et al.<sup>13</sup>. In the next section we will show how reformulation of these equations can lead to a less nonconvex and better structured problem.

**Well Performance Model** consists of the following Eq. (13-15). Raw natural gas is corrosive and possibly sticky before going into preprocessing facilities from wells, and a first-principle well performance model can be overly complicated for optimization. Therefore, the

well performance is approximated by Eq. (13).<sup>13</sup> In this equation, Darcy's constant  $\alpha_{i,w}$  and non-Darcy correction factor  $\beta_{i,w}$  reflect in-flow performance,  $\vartheta_{i,w}$  and  $\lambda_{i,w}$  reflect vertical lift performance,  $P_{i,w}^{out}$  and  $\pi_{i,w}$  represent tubing head and reservoir pressure,  $Q_{i,j,w}$  represents volumetric flow rate from well  $i$  to a platform  $j$ . If the gas well  $y_i$  is not to be developed,  $y_{i,j}$  will be zero and the volumetric gas flow rate in the well will be forced to be zero. With Eq. (13), equations (14) and (15) imply each other so only one of them is needed for optimization. However, they are both included for more efficient global optimization.

$$\alpha_{i,w}Q_{i,j,w} + (\beta_{i,w} + \vartheta_{i,w})Q_{i,j,w}^2 = y_{i,j}(\pi_{i,w}^2 - \lambda_{i,w}(P_{i,w}^{out})^2), \quad \forall (i,j) \in E^W, \quad \forall w \in \Phi \quad (13)$$

$$Q_{i,j,w} \geq 0, \quad \forall (i,j) \in E^W, \quad \forall w \in \Phi \quad (14)$$

$$P_{i,w}^{out} \leq \pi_{i,w}, \quad \forall i \in \{i | (i,j) \in E^W\}, \quad \forall w \in \Phi \quad (15)$$

**Pipeline Performance Model** consists of Eq. (16-17). In a gas pipeline, the gas flow usually includes both gas and liquid phases and there may be mass transfer between the two phases. In order to simplify the model, gas flows are assumed to contain only gas phase and behave like ideal gas. With these assumptions and under steady-state operation, the pressure flow relationship in a long transmission pipeline that is longer than 20 km can be well approximated by Eq. (16).<sup>13</sup> Here  $P_{i,w}^{out}$  denotes the pressure at the outlet of vertex  $i$ ,  $P_{i,w}^{in}$  denotes the pressure at the inlet of vertex  $i$ , set  $E^L \subset E$  includes edges representing long pipelines. This equation also enforces the volumetric flow rate to be zero if the long pipeline is not to be developed. When a pipeline is short (such as one connects a gas well to a gas platform), the pressure drop along the pipeline is usually negligible, so it can be described by Eq. (17).

$$y_{i,j}((P_{i,w}^{out})^2 - (P_{j,w}^{in})^2) = \kappa_{i,j}Q_{i,j,w}^2, \quad \forall (i,j) \in E^L, \quad \forall w \in \Phi \quad (16)$$

$$y_{i,j}(P_{i,w}^{out} - P_{j,w}^{in}) = 0, \quad \forall (i,j) \in E \setminus E^L, \quad \forall w \in \Phi \quad (17)$$

**Compression Performance Model** consists of Eq. (18-20). Eq. (18) relates the power consumption  $W$  with the outlet-inlet pressure ratio and the molar flow rates of gas compressed, where  $\sigma_j$  and  $\nu$  are the parameters obtained from process data,<sup>13</sup> and set  $V^C \subset V$  includes vertices having onsite compressors. Eq. (19) ensures that the inlet pressure of a compressor is no larger than the outlet pressure. Eq. (20) is to bound the power consumption of a compressor and enforce the power consumption to be zero when the gas platform containing the compressor is not to be developed.

$$W_{j,w} - \sigma_j \sum_{i \in \{i | (i,j) \in E\}} \sum_{q \in \Omega} f_{i,j,q,w} \left[ \left( \frac{P_{j,w}^{out}}{P_{j,w}^{in}} \right)^\nu - 1 \right] = 0, \quad \forall j \in V^C, \quad \forall w \in \Phi \quad (18)$$

$$P_{i,w}^{in} \leq P_{i,w}^{out}, \quad \forall i \in V^C, \quad \forall w \in \Phi \quad (19)$$

$$y_i^v \Psi_i^{LB} \leq W_{i,w} \leq y_i^v \Psi_i^{UB}, \quad \forall i \in V^C, \quad \forall w \in \Phi \quad (20)$$

**Other Pressure Constraints** are given in Eq. (21-23). Eq. (21) means that there is no pressure rise for a vertex having no onsite compressor. If the vertex is a well, this relationship is already described by Eq. (15). Eq. (22-23) pose bounds on pressures at the inlet of a vertex (excluding the well) and at the outlet of a vertex, respectively.

$$P_{i,w}^{in} = P_{i,w}^{out}, \quad \forall i \in V \setminus (V^C \cup V^W), \quad \forall w \in \Phi \quad (21)$$

$$y_i^v \Gamma_{i,w}^{in, LB} \leq P_{i,w}^{in} \leq y_i^v \Gamma_{i,w}^{in, UB}, \quad \forall i \in V \setminus V^W, \quad \forall w \in \Phi \quad (22)$$

$$y_i^v \Gamma_{i,w}^{out, LB} \leq P_{i,w}^{out} \leq y_i^v \Gamma_{i,w}^{out, UB}, \quad \forall i \in V, \quad \forall w \in \Phi \quad (23)$$

## Equations for Unit Conversion and the Economic Objective

Molar component flow rates and volumetric total flow rates are used in the stochastic pooling model and in the process flow relationship constraints, respectively. The following Eq. (24-25) link the two types of flow rates, under the ideal gas assumption.  $\phi$  is the unit conversion parameter<sup>13</sup> and  $U_{i,q,w}$  is the mole fraction of component  $q$  at well or remote subnetwork platform  $i$  in scenario  $w$ .

$$f_{i,j,q,w} = \phi U_{i,q,w} Q_{i,j,w}, \quad \forall (i,j) \in \{(i,j) \in E \mid i \in V^W \cup V^S\}, \forall w \in \Phi \quad (24)$$

$$\sum_{q \in \Omega} f_{i,j,q,w} = \phi Q_{i,j,w}, \quad \forall (i,j) \in E, \forall w \in \Phi \quad (25)$$

The objective of the optimization is to minimize the negative expected NPV of the system over its life span,  $C^{(NPV)}$ , as defined in Eq. (26).  $C^{(CC)}$  represents the total capital cost, which is calculated using Eq. (27),  $AP_w$  represents the annual operation profit for scenario  $w$ , which is calculated using Eq. (28),  $p_w$  denotes the probability for uncertainty scenario  $w$ ,  $L$  denotes the system life span, and  $r$  denotes the annual discount rate.

$$C^{(NPV)} = C^{(CC)} - \sum_{t=1}^L \sum_w \frac{p_w AP_w}{(1+r)^t} \quad (26)$$

$$C^{(CC)} = \sum_{i \in V} y_i^v I_i^{(v,CC)} + \sum_{(i,j) \in E} y_{i,j} I_{i,j}^{(CC)} \quad (27)$$

$$\begin{aligned} AP_w = & \tau \sum_{k \in V^T} \left( \sum_{q \in \Omega} \sum_{j \in \{j \mid (j,k) \in E\}} f_{j,k,q,w} \right) UP_{k,w} \\ & - \tau \sum_{i \in V^W} \left( \sum_{q \in \Omega} \sum_{j \in \{j \mid (i,j) \in E\}} f_{i,j,q,w} \right) UC_{i,w} \\ & - \sum_{j \in V^C} \tau \gamma_w W_{j,w} \end{aligned} \quad (28)$$

## The Optimization Model

When considering only the stochastic pooling model, the optimization problem can be stated as:

$$\begin{aligned}
 \min \quad & C^{(NPV)} \\
 \text{s.t.} \quad & \text{Eq. (1 – 12)} \\
 & \text{Eq. (26 – 27)} \\
 & \text{Eq. (28*)}
 \end{aligned} \tag{I}$$

where Eq. (28\*) is the Eq. (28) that does not include the power costs. This model is relatively easy for global optimization, but may lead to a solution in which the gas flow rates planned for the operation are not realistic due to the lack of sufficient pressure differential (as will be showed in the case study). Therefore, we propose in this paper to use the following optimization model that addresses the pressure flow relationships explicitly:

$$\begin{aligned}
 \min \quad & C^{(NPV)} \\
 \text{s.t.} \quad & \text{Eq. (1 – 12)} \\
 & \text{Eq. (13 – 25)} \\
 & \text{Eq. (26 – 28)}
 \end{aligned} \tag{II}$$

Note that in the right-hand-side of Eq. (13), the binary variable  $y_{i,j}$  and the expression  $(\pi_{i,w}^2 - \lambda_{i,w}(P_{i,w}^{out})^2)$  can be further separated by the well-known big-M method,<sup>30</sup> and this reformulation can reduce the nonconvexity of the problem. For simplicity of the subsequent discussion, we do not explicitly present this reformulation here.

Formulation (II) is much more challenging to solve than Formulation (I), because it contains more variables and nonconvex constraints. In order to practically solve the proposed optimization model, we reformulate it in the next section so that: (a) some nonconvex constraints can be rewritten into convex constraints, and (b) the convex and nonconvex

parts of the model are separable so that an operation subproblem for a scenario can be solved by a classical decomposition method (such as GBD).

## Reformulation of the Optimization Model

In this section, the proposed optimization formulation (II) is reformulated in three steps. In the first step, we propose to consider the effect of pressure regulators, which can be readily built in different parts of a gas production network but are omitted in (II). In the second step, it will be seen that, after the inclusion of pressure drop from pressure regulators, some nonlinear equality constraints in formulation (II) (which is nonconvex) can be transformed into nonlinear inequality constraints and some of the reformulated nonlinear inequality constraints are convex. In the third step, the "convex part" of the reformulated nonconvex nonlinear constraints can be separated from its "nonconvex part", under a mild assumption on the structure of the network.

### Inclusion of Pressure Drops in Pressure Regulators

In Formulation (II), pressures at vertices are uniquely determined by gas flow rates and compressor powers, i.e., there is no flexibility in adjusting the pressures with fixed gas flow rates and compressor powers. This clearly does not represent real gas production systems, as pressure regulators that can provide extra pressure drops are normally installed at different parts of the system. We first consider pressure drop  $\Delta_{i,w}$  that is enabled by a pressure regulator near the exit of a gas well  $i$ , as illustrated in Figure 2. The pressure regulator can reduce the pressure at the end of the well pipeline from  $P_{i,w}^{out}$  to  $\hat{P}_{i,w}^{out}$ , without incurring significant operating cost. Similarly, pressure drop can also be enabled near the exit of a platform (with or without an onsite compressor) by a pressure regulator, as illustrated in Figure 3. As a result, the well performance equation (13) and the compression performance

equation (18), can be rewritten as:

$$\begin{aligned} \alpha_{i,w} Q_{i,j,w} + (\beta_{i,w} + \vartheta_{i,w}) Q_{i,j,w}^2 &= y_{i,j} (\pi_{i,w}^2 - \lambda_{i,w} (\hat{P}_{i,w}^{out} + \Delta_{i,w})^2), \\ \forall (i,j) \in E^W, \quad \forall w \in \Phi \end{aligned} \quad (29)$$

$$W_{j,w} - \sigma_j \sum_{i \in \{i | (i,j) \in E\}} \sum_{q \in \Omega} f_{i,j,q,w} \left[ \left( \frac{\hat{P}_{j,w}^{out} + \Delta_{j,w}}{P_{j,w}^{in}} \right)^\nu - 1 \right] = 0, \quad \forall j \in V^C, \quad \forall w \in \Phi \quad (30)$$

In both equations, the pressure after the regulator is used instead of the pressure before the regulator. In this way, the outlet pressures of vertices in the model all become pressures after the pressure regulators, and the previous pressure constraints Eq. (15), (19), (21), (23) should now be written as:

$$\hat{P}_{i,w}^{out} \leq \pi_{i,w}, \quad \forall i \in V^W, \quad \forall w \in \Phi \quad (31)$$

$$P_{i,w}^{in} \leq \hat{P}_{i,w}^{out}, \quad \forall i \in V^C, \quad \forall w \in \Phi \quad (32)$$

$$P_{i,w}^{in} \geq \hat{P}_{i,w}^{out}, \quad \forall i \in V \setminus (V^C \cup V^W), \quad \forall w \in \Phi \quad (33)$$

$$y_i^v \Gamma_{i,w}^{out, LB} \leq \hat{P}_{i,w}^{out} \leq y_i^v \Gamma_{i,w}^{out, UB}, \quad \forall i \in V, \quad \forall w \in \Phi \quad (34)$$

And we should also require that the pressure drop be nonnegative:

$$\Delta_{i,w} \geq 0, \quad \forall i \in V, \quad \forall w \in \Phi \quad (35)$$

We can also consider pressure regulators in pipelines. Without loss of generality, assume the pressure regulator in a pipeline is located at the starting vertex of the pipeline, as illustrated in Figure 4.  $\delta_{i,j,w}$  denotes the pressure drop introduced by the pressure regulator. As a result, the pipeline performance equations (16) and (17) can be rewritten as:

$$y_{i,j}((\hat{P}_{i,w}^{out} - \delta_{i,j,w})^2 - (P_{j,w}^{in})^2) = \kappa_{i,j} Q_{i,j,w}^2, \quad \forall (i,j) \in E^L, \quad \forall w \in \Phi \quad (36)$$

$$y_{i,j}((\hat{P}_{i,w}^{out} - \delta_{i,j,w}) - P_{j,w}^{in}) = 0, \quad \forall (i,j) \in E \setminus E^L, \quad \forall w \in \Phi \quad (37)$$

and the pressure drop and the pressure before the pressure regulator have to be nonnegative:

$$\delta_{i,j,w} \geq 0, \quad \forall (i,j) \in E, \quad \forall w \in \Phi \quad (38)$$

$$\hat{P}_{i,w}^{out} - \delta_{i,j,w} \geq 0, \quad \forall (i,j) \in E, \quad \forall w \in \Phi \quad (39)$$

## Removal of Pressure Drops

Usually investments and operating costs for pressure regulators are negligible compared to the costs for other major units in a gas production network (such as long pipelines and compressors), so we do not need to include the costs incurred by the pressure drops in the objective function. As a result,  $\Delta_{i,w}$  does not need to be included in the model explicitly.

For example, given that  $\Delta_{i,w} \geq 0$ , constraint (30) can be transformed into:

$$W_{j,w} - \sigma_j \sum_{i \in \{(i,j) \in E\}} \sum_{q \in \Omega} f_{i,j,q,w} \left[ \left( \frac{\hat{P}_{j,w}^{out}}{P_{j,w}^{in}} \right)^\nu - 1 \right] \geq 0, \quad \forall j \in V^C, \quad \forall w \in \Phi \quad (40)$$

(40) is equivalent to (30) in the following sense: for any  $(W_{j,w}, f_{i,j,q,w}, \hat{P}_{j,w}^{out}, P_{j,w}^{in})$  that is feasible for (40), there exists  $\Delta_{i,w} (\geq 0)$  such that  $(W_{j,w}, f_{i,j,q,w}, \hat{P}_{j,w}^{out}, P_{j,w}^{in}, \Delta_{i,w})$  is feasible for (30); and for any  $(W_{j,w}, f_{i,j,q,w}, \hat{P}_{j,w}^{out}, P_{j,w}^{in}, \Delta_{i,w})$  that is feasible for (30) (and  $\Delta_{i,w} \geq 0$ ),  $(W_{j,w}, f_{i,j,q,w}, \hat{P}_{j,w}^{out}, P_{j,w}^{in})$  is feasible for (40). Note that, this reformulation is exact, no matter whether the lower bound on  $W_{j,w}$  is 0 or a positive value. However, this reformulation is not exact if  $\frac{\hat{P}_{j,w}^{out}}{P_{j,w}^{in}}$  has to be greater than 1; because in this case, the model needs to include an extra constraint  $\frac{\hat{P}_{j,w}^{out} + \Delta_{j,w}}{P_{j,w}^{in}} \geq \text{ratio}_{j,w}$ , which cannot be replaced by  $\frac{\hat{P}_{j,w}^{out}}{P_{j,w}^{in}} \geq \text{ratio}_{j,w}$ . The

model does not need this extra constraint if  $\text{ratio}_{j,w} = 1$ , as in this case it is already implied by constraint (32).

Following the same logic, constraint (29) can be equivalently transformed into:

$$\begin{aligned} \alpha_{i,w}Q_{i,j,w} + (\beta_{i,w} + \vartheta_{i,w})Q_{i,j,w}^2 &\leq y_{i,j}(\pi_{i,w}^2 - \lambda_{i,w}(\hat{P}_{i,w}^{out})^2), \\ \forall (i,j) \in E^W, \quad \forall w \in \Phi \end{aligned} \quad (41)$$

Note that Eq. (41) is convex when the binary variable  $y_{i,j}$  is fixed (and  $y_{i,j}$  can be separated from the quadratic expression by the big-M method). Before the reformulation, Eq. (29) is nonconvex with fixed  $y_{i,j}$ . Therefore, this reformulation reduces the nonconvexity of the optimization model.

Similarly,  $\delta_{i,j,w}$  does not need to be included in the model, and constraints (36-39) can be replaced by:

$$y_{i,j}((\hat{P}_{i,w}^{out})^2 - (P_{j,w}^{in})^2) \geq \kappa_{i,j}Q_{i,j,w}^2, \quad \forall (i,j) \in E^L, \quad \forall w \in \Phi \quad (42)$$

$$y_{i,j}(\hat{P}_{i,w}^{out} - P_{j,w}^{in}) \geq 0, \quad \forall (i,j) \in E \setminus E^L, \quad \forall w \in \Phi \quad (43)$$

Note that Eq. (42) can be written as convex constraints when the binary variable  $y_{i,j}$  is fixed. It is obvious when  $y_{i,j} = 0$ . When  $y_{i,j} = 1$ , the constraint can be written as the following second-order cone constraint:<sup>31</sup>

$$\hat{P}_{i,w}^{out} \geq \sqrt{(P_{j,w}^{in})^2 + \kappa_{i,j}Q_{i,j,w}^2}, \quad \forall (i,j) \in E^L, \quad \forall w \in \Phi \quad (44)$$

After this step, Formulation (II) can be written into the following optimization formula-

tion:

$$\begin{aligned}
& \min && C^{(NPV)} \\
& s.t. && \text{Eq. (1 – 12)} \\
& && \text{Eq. (14), (20), (22), (24 – 25)} \\
& && \text{Eq. (26 – 28)} \\
& && \text{Eq. (31 – 34), (41 – 43)}
\end{aligned} \tag{III}$$

The equations in the last row of Formulation (III) are the new equations developed in the two reformulation steps, in place of some constraints in Formulation (II). The operation subproblems of Formulation (III), i.e., the subproblems resulting from fixing the design decision variables  $y_i^v, y_{i,j}$ , are less nonconvex compared to those of Formulation (II). This is because the well performance and pipeline performance constraints are now convex. However, the compression performance constraint Eq. (40) is still highly nonconvex for each  $j \in V^C$  and  $w \in \Phi$ . In addition, these constraints hinder efficient application of decomposition methods for solving each operation subproblem, as the constraint links major variables in the operation problem.

## Separation of Flow Rates and Pressures

Here we explain how to separate the pressure variables in the compression performance constraint (40), so that the constraints involving pressure variables are all convex. This way, we can separate more convex part out of the nonconvex model when applying decomposition based optimization. We define new variables  $P_{j,w}^{in,c} \triangleq (P_{j,w}^{in})^2, P_{j,w}^C$ , and rewrite Eq. (40) and Eq. (32) into:

$$W_{j,w} - \sigma_j \sum_{i \in \{i | (i,j) \in E\}} \sum_{q \in \Omega} f_{i,j,q,w} P_{j,w}^C \geq 0, \quad \forall j \in V^C, \quad \forall w \in \Phi \tag{45}$$

$$P_{j,w}^C \geq \left( \frac{\hat{P}_{j,w}^{out}}{P_{j,w}^{in}} \right)^\nu - 1, \quad \forall j \in V^C, \quad \forall w \in \Phi \tag{46}$$

$$P_{j,w}^C \geq 0, \quad \forall j \in V^C, \quad \forall w \in \Phi \quad (47)$$

and Eq. (46) can be further written into

$$(P_{j,w}^C + 1)^{2/\nu} \geq \frac{(\hat{P}_{j,w}^{out})^2}{P_{j,w}^{in,c}}, \quad \forall j \in V^C, \quad \forall w \in \Phi \quad (48)$$

Note that Eq. (48) is convex if  $P_{j,w}^C$  is fixed.<sup>32</sup> The advantage of Eq. (45), (47), (48) over Eq. (40) and (32) is that the pressures are separated from the molar flow rates. Following this, we seem to be able to separate the pressure flow relationships (which seem to be convex now) from the stochastic pooling submodel (which is nonconvex), but we also need to examine whether this reformulation may complicate other pressure related constraints. Consider a long pipeline connecting the compressor on platform  $j$  and an upstream vertex  $i$ , as shown in Figure 5. Due to the inclusion of new variable  $\hat{P}_{j,w}^{in,c}$ , the pipeline performance constraint Eq. (44) needs to be reformulated, and Figure 5 shows three different reformulations. In reformulation (a), the nonconvex constraint  $P_{j,w}^{in,c} = (P_{j,w}^{in})^2$  is explicitly included, which destroys the convexity of the pressure related part of the model. In reformulation (b),  $P_{j,w}^{in,c}$  is used in place of  $(P_{j,w}^{in})^2$ , but the constraint becomes nonconvex. In reformulation (c), a new variable  $\hat{P}_{i,w}^{out,c} \triangleq (\hat{P}_{i,w}^{out})^2$  is introduced, which leads to a convex constraint.

From reformulation (c), we can use  $\hat{P}_{i,w}^{out,c}$  instead of  $\hat{P}_{i,w}^{out}$  for all the vertices that are predecessors of a compressor, in order to make the pressure relations related to these vertices convex, if none of these vertices contain a compressor. Note that a predecessor of a vertex  $A$  in a directed graph, is a vertex from which there exists a directed path to  $A$ . Therefore, we make the following assumption to enable the desired reformulation:

**Assumption 1.** *For any vertex that has an onsite compressor, none of its predecessor vertices have an onsite compressor.*

For cases where Assumption 1 is satisfied, we define a new vertex subset  $\bar{V} \subset V$  that includes all vertices which are a predecessor of a compressor. According to Assumption 1, the vertices in this subset do not contain a compressor, so the relevant pressure constraints

in Formulation (III) are Eq. (22), (31), (33-34), (41-43). They now can be rewritten as:

$$y_i^v(\Gamma_{i,w}^{in,LB})^2 \leq P_{i,w}^{in,c} \leq y_i^v(\Gamma_{i,w}^{in,UB})^2, \quad \forall i \in (V^C \cup \bar{V}) \setminus V^W, \quad \forall w \in \Phi \quad (49)$$

$$\hat{P}_{i,w}^{out,c} \leq (\pi_{i,w})^2, \quad \forall i \in V^W \cap \bar{V}, \quad \forall w \in \Phi \quad (50)$$

$$P_{i,w}^{in,c} \geq \hat{P}_{i,w}^{out,c}, \quad \forall i \in \bar{V} \setminus V^W, \quad \forall w \in \Phi \quad (51)$$

$$y_i^v(\Gamma_{i,w}^{out,LB})^2 \leq \hat{P}_{i,w}^{out,c} \leq y_i^v(\Gamma_{i,w}^{out,UB})^2, \quad \forall i \in \bar{V}, \quad \forall w \in \Phi \quad (52)$$

$$\begin{aligned} \alpha_{i,w}Q_{i,j,w} + (\beta_{i,w} + \vartheta_{i,w})Q_{i,j,w}^2 &\leq y_{i,j}(\pi_{i,w}^2 - \lambda_{i,w}\hat{P}_{i,w}^{out,c}), \\ \forall(i,j) \in \{(i,j)|(i,j) \in E^W, i \in \bar{V}\}, &\quad \forall w \in \Phi \end{aligned} \quad (53)$$

$$y_{i,j}(\hat{P}_{i,w}^{out,c} - P_{j,w}^{in,c}) \geq \kappa_{i,j}Q_{i,j,w}^2, \quad \forall(i,j) \in \{(i,j)|(i,j) \in E^L, i \in \bar{V}\}, \quad \forall w \in \Phi \quad (54)$$

$$y_{i,j}(\hat{P}_{i,w}^{out,c} - P_{j,w}^{in,c}) \geq 0, \quad \forall(i,j) \in \{(i,j)|(i,j) \in E \setminus E^L, i \in \bar{V}\}, \quad \forall w \in \Phi \quad (55)$$

Also according to Assumption 1, any successor vertex of a compressor does not have an onsite compressor. Note that a successor of a vertex  $A$  in a directed graph, is a vertex to which a directed path from  $A$  exists. Therefore, for the vertices that are not included in subset  $\bar{V}$ , the relevant pressure constraints are all pipeline performance constraints which do not need to be changed:

$$y_i^v\Gamma_{i,w}^{in,LB} \leq P_{i,w}^{in} \leq y_i^v\Gamma_{i,w}^{in,UB}, \quad \forall i \in V \setminus (V^W \cup V^C \cup \bar{V}), \quad \forall w \in \Phi \quad (56)$$

$$\hat{P}_{i,w}^{out} \leq \pi_{i,w}, \quad \forall i \in V^W \setminus \bar{V}, \quad \forall w \in \Phi \quad (57)$$

$$P_{i,w}^{in} \geq \hat{P}_{i,w}^{out}, \quad \forall i \in V \setminus (V^W \cup V^C \cup \bar{V}), \quad \forall w \in \Phi \quad (58)$$

$$y_i^v\Gamma_{i,w}^{out,LB} \leq \hat{P}_{i,w}^{out} \leq y_i^v\Gamma_{i,w}^{out,UB}, \quad \forall i \in V \setminus \bar{V}, \quad \forall w \in \Phi \quad (59)$$

$$\begin{aligned} \alpha_{i,w}Q_{i,j,w} + (\beta_{i,w} + \vartheta_{i,w})Q_{i,j,w}^2 &\leq y_{i,j}(\pi_{i,w}^2 - \lambda_{i,w}(\hat{P}_{i,w}^{out})^2), \\ \forall (i,j) \in \{(i,j) | (i,j) \in E^W, i \in V \setminus \bar{V}\}, \quad \forall w \in \Phi \end{aligned} \quad (60)$$

$$y_{i,j}((\hat{P}_{i,w}^{out})^2 - (P_{j,w}^{in})^2) \geq \kappa_{i,j}Q_{i,j,w}^2, \quad \forall (i,j) \in \{(i,j) | (i,j) \in E^L, i \in V \setminus \bar{V}\}, \quad \forall w \in \Phi \quad (61)$$

$$y_{i,j}(\hat{P}_{i,w}^{out} - P_{j,w}^{in}) \geq 0, \quad \forall (i,j) \in \{(i,j) | (i,j) \in E \setminus E^L, i \in V \setminus \bar{V}\}, \quad \forall w \in \Phi \quad (62)$$

After this reformulation step, the optimization model (III) can be transformed into the following model:

$$\begin{aligned} \min \quad & C^{(NPV)} \\ \text{s.t.} \quad & \text{Eq. (1 – 12)} \\ & \text{Eq. (14), (20), (24 – 25)} \\ & \text{Eq. (26 – 28)} \\ & \text{Eq. (45), (47 – 48), (49 – 62)} \end{aligned} \quad (IV)$$

Figure 6 shows two example networks. The one in Figure 6(a) satisfies Assumption 1, so the above reformulation procedure can be applied. The one in Figure 6(b) does not satisfy Assumption 1, but a subpart of it (i.e., the part excluding the compressor near the bottom) does. Therefore, the reformulation can still be applied, but only for this subpart of the network.

## A Customized Global Optimization Method

### The Decomposable Structure of Formulation (IV)

This section discusses an efficient global optimization method for solving the proposed optimization formulation (IV) to  $\epsilon$ -optimality. For convenience of subsequent discussion, we

rewrite Formulation (IV) into the following form:

$$\begin{aligned}
& \min_{x_0, x_1, \dots, x_s} \sum_{w=1}^s p_w [f_{0,w}(x_0) + f_{1,w}(x_w^{nc}) + f_{2,w}(x_w^c)] \\
& s.t. \quad g_{0,w}(x_0) + g_{1,w}(x_w^{nc}) + g_{2,w}(x_w^c) \leq 0, \quad w = 1, \dots, s \\
& \quad \quad x_w^{nc} \in X_w^{nc}, \quad w = 1, \dots, s \\
& \quad \quad x_w^c \in X_w^c, \quad w = 1, \dots, s \\
& \quad \quad x_0 \in X_0
\end{aligned} \tag{P}$$

Here  $p_w$  represents the probability for scenario  $w$ .  $x_0$  includes design decisions (i.e., the "y"s in Formulation (IV)) or so-called first-stage decisions in the context of stochastic programming, and  $X_0 \subset \{0, 1\}^{n_0}$  is nonempty.  $x_w = (x_w^{nc}, x_w^c)$  includes operation decisions, or called second-stage decisions.  $x_w^{nc}$  includes the operation decisions involved in nonconvex functions (mostly "s"s, "f"s, "P<sup>C</sup>"s and "W"s in Formulation (IV)), which are also called nonconvex variables in the paper, and the set  $X_w^{nc} \subset \mathbb{R}^{n_{nc}}$  is nonempty and compact.  $x_w^c$  includes the operation decisions only involved in convex functions (i.e., "P"s and "Q"s in Formulation (IV)), which are also called convex variables in this paper, and  $X_w^c \subset \mathbb{R}^{n_c}$  is nonempty, compact and convex. Functions  $f_{2,w}$ ,  $g_{2,w}$  are convex functions on  $X_w^c$ .

Figure 7 illustrates the decomposable structure of Problem (P). The first block in Figure 7(a) represents the first group of constraints in Formulation (P), which links the design decisions and operation decisions in different scenarios. The second block in Figure 7(a) represents the second and third groups of constraints in Formulation (P), which consists of  $s$  operation subproblems that are independent of one another. The third block in Figure 7(a) represents the last group of constraint in Formulation (P) that restricts the choice of  $x_0$ . Intuitively, problem (P) is decomposable in the sense that it can be separated into a set of operation subproblems via fixing  $x_0$  or dualizing the linking constraints. If each operation subproblem can be solved quickly, then Problem (P) can be solved by either LD (together with explicit branch-and-bound search) or NGBD. Each operation subproblem here is a

large nonconvex MINLP which state-of-the-art global optimization solvers may not be able to solve quickly.<sup>13</sup>

Fortunately, due to the reformulation introduced in the last section, each operation problem in Problem (P) also has a decomposable structure, as illustrated by Figure 7(b). The operation problem consists of a convex part (i.e., pressure flow submodel), a nonconvex part (i.e., the pooling submodel), and a set of complicating constraints (including unit conversion constraints and compressor constraints (48)) that link the convex and the nonconvex parts. Considering that existing state-of-the-art global optimization solvers (such as BARON,<sup>33</sup> ANTIGONE<sup>34</sup>) can solve pooling problems very quickly, the operation subproblem may be solved efficiently through a GBD procedure. Following this idea, we show in the next subsection how to develop a variant of NGBD for efficiently solving Problem (P) to  $\epsilon$ -optimality. The proposed variant of NGBD can be applied not only to Formulation (IV), but also to any other problems that can be written as Problem (P).

## The Multi-Loop NGBD Method

NGBD was originally developed for solving problems with the structure shown in Figure 7(a). The basic idea is to generate a sequence of upper bounding problem via fixing  $x_0 = x_0^l$  at each NGBD iteration  $l$ . The upper bounding problem is called **Primal Problem**, which can be naturally decomposed into the following subproblem for scenario  $w = 1, \dots, s$ :

$$\begin{aligned}
 \min_{x_w} \quad & p_w[f_{0,w}(x_0^l) + f_{1,w}(x_w^{nc}) + f_{2,w}(x_w^c)] \\
 \text{s.t.} \quad & g_{0,w}(x_0^l) + g_{1,w}(x_w^{nc}) + g_{2,w}(x_w^c) \leq 0 \\
 & x_w^{nc} \in X_w^{nc} \\
 & x_w^c \in X_w^c
 \end{aligned} \tag{PP}_w^l$$

On the other hand, a **Lower Bounding Problem** (LBP) is constructed via replacing the nonconvex functions in Problem (P) with their convex relaxations. Problem (LBP) is a large-scale MILP or convex MINLP which can be solved via GBD efficiently, and it is

enhanced with an integer cut after each GBD iteration. Therefore, a sequence of lower bounds are generated via the solution of Problem (LBP) over the NGBD iterations. The NGBD algorithm terminates in a finite number of iterations when the upper and lower bounds converge to an  $\epsilon$ -optimal solution or the infeasibility of the problem is indicated by an infeasible lower bounding problem. Figure 8(a) shows a simplified diagram for the standard NGBD algorithm. More details about NGBD subproblems and algorithm can be found in Li et al.<sup>29</sup>. In this paper, the standard NGBD algorithm is called single-loop NGBD, as there is only one GBD loop in one NGBD iteration.

For Formulation (IV), a primal subproblem ( $PP_w^l$ ) is an operation problem for scenario  $w$ . As discussed in the previous subsection, each operation problem is further decomposable and can be solved via GBD. Therefore, it is natural to embed another GBD loop in each NGBD iteration, in order to solve the primal subproblems. This leads to a multi-loop NGBD as shown in Figure 8(b), which includes 1 GBD loop for Problem (LBP) and  $s$  GBD loops for the primal subproblems at each NGBD iteration.

Here we provide more details on the GBD procedure for solving each primal subproblem ( $PP_w^l$ ). We view  $x_w^{nc}$  as the complicating variable, since Problem ( $PP_w^l$ ) is convex if  $x_w^{nc}$  is fixed. Therefore, at each GBD iteration  $k$  for solving Problem ( $PP_w^l$ ), we fix  $x_w^{nc} = x_w^{nc,k}$  to construct an upper bounding problem, called **Benders Primal Problem**, as follows:

$$\begin{aligned}
\min_{x_w^c} \quad & p_w[f_{0,w}(x_0^l) + f_{1,w}(x_w^{nc,k}) + f_{2,w}(x_w^c)] \\
s.t. \quad & g_{0,w}(x_0^l) + g_{1,w}(x_w^{nc,k}) + g_{2,w}(x_w^c) \leq 0 \\
& x_w^c \in X_w^c
\end{aligned} \tag{BPP_w^{l,k}}$$

Note that if Problem  $BPP_w^{l,k}$  is a convex NLP, it needs to satisfy some constraint qualification for the GBD algorithm to be valid, as GBD relies on strong duality of each upper bounding problem. The solution of Problem  $BPP_w^{l,k}$  yields an upper bound for Problem ( $PP_w^l$ ). If

Problem  $\text{BPP}_w^{l,k}$  is infeasible for  $x_w^{nc,k}$ , then the following feasibility problem is solved instead:

$$\begin{aligned}
& \min_{x_w^c, z} \quad \|z\|_1 \\
& s.t. \quad g_{0,w}(x_0^l) + g_{1,w}(x_w^{nc,k}) + g_{2,w}(x_w^c) \leq z \\
& \quad \quad x_w^c \in X_w^c \\
& \quad \quad z \geq 0
\end{aligned} \tag{BFP}_w^{l,k}$$

where  $\|\cdot\|_1$  denotes the 1-norm.

After Problem  $(\text{BPP}_w^{l,k})$  or  $(\text{BFP}_w^{l,k})$  is solved, the following lower bounding problem, called **Benders Relaxed Master Problem**, is solved to yield a lower bound for Problem  $(\text{PP}_w^l)$ :

$$\begin{aligned}
& \min_{\eta_w, x_w^{nc}} \quad \eta_w \\
& s.t. \quad \eta_w \geq \text{obj}_{\text{BPP}_w^{l,j}} + (f_{1,w}(x_w^{nc}) - f_{1,w}(x_w^{nc,j})) + (\lambda_w^{l,j})^T (g_{1,w}(x_w^{nc}) - g_{1,w}(x_w^{nc,j})), \quad \forall j \in T^k \\
& \quad \quad 0 \geq \text{obj}_{\text{BFP}_w^{l,i}} + (\mu_w^{l,i})^T (g_{1,w}(x_w^{nc}) - g_{1,w}(x_w^{nc,i})), \quad \forall i \in S^k \\
& \quad \quad x_w^{nc} \in X_w^{nc} \\
& \quad \quad \text{UBDPP}_w^{l,k} \geq \eta_w \quad \text{UBD}_w^l \geq \eta_w
\end{aligned} \tag{BRMP}_w^{l,k}$$

The solution of  $(\text{BRMP}_w^{l,k})$  yields a lower bound for Problem  $(\text{PP}_w^l)$ , called  $\text{LBDPP}_w^{l,k}$ ; If  $(\text{BRMP}_w^{l,k})$  or  $(\text{BFP}_w^{l,k})$  is infeasible, then Problem  $(\text{PP}_w^l)$  is infeasible. In  $(\text{BRMP}_w^{l,k})$ , the first group of constraints are called optimality cuts. The parameters  $\text{obj}_{\text{BPP}_w^{l,j}}$  and  $\lambda_w^{l,j}$  are the optimal objective value and Lagrange multipliers for subproblem  $(\text{BPP}_w^{l,k})$ , respectively. Set  $T^k$  includes indexes of previous iterations at which subproblem  $(\text{BPP}_w^{l,k})$  is feasible. The second group of constraints are called feasibility cuts. The parameters  $\text{obj}_{\text{BFP}_w^{l,i}}$  and  $\mu_w^{l,i}$  are the optimal objective value and Lagrange multipliers for subproblem  $(\text{BFP}_w^{l,k})$ , respectively. Set  $S^k$  includes indexes of previous iterations at which subproblem  $(\text{BPP}_w^{l,k})$  is infeasible (and therefore subproblem  $(\text{BFP}_w^{l,k})$  is solved). Since Problem  $(\text{BRMP}_w^{l,k})$  is nonconvex and its

solution time dominates the GBD solution time, we add extra bounding constraints to expedite the solution of Problem (BRMP $_{w}^{l,k}$ ). Two upper bounds are used in the two bounding constraints. UBDPP $_{w}^{l,k}$  is the current best upper bound for Problem (PP $_{w}^l$ ), obtained from the best optimal objective value of the previously solved Benders primal problem. This upper bound surely does not exclude an optimal solution of Problem (PP $_{w}^l$ ). UBD $_{w}^l$  is obtained from primal problems solved in the previous NGBD iterations, as explained in Li et al.<sup>29</sup>. This bound can exclude solutions of NGBD primal problem that are worse than the current NGBD upper bound. This upper bound can exclude an optimal solution of Problem (PP $_{w}^l$ ), but not an optimal solution of Problem (P), so it is also a valid bound for the multi-loop NGBD.

Figure 9 shows the algorithm flowchart for the customized GBD procedure of solving each NGBD primal subproblem (PP $_{w}^l$ ).

## Case Study

### Case Study Problem Statement and Implementation

In this case study, we consider an integrated design and operation problem for the SGPS introduced in the second section. In this problem, part of the SGPS is already developed, as shown by the solid lines in Figure 10. This existing part includes 5 gas fields (SC, E11, F6, F23SW and F23) that contain 40 wells, 3 remote subnetwork vertices (D35, BY, BN), 4 gas platforms (BYP, E11P, F23P, E11R-A), 1 plant slug catcher (SC-1), and 1 LNG plant. In order to meet the increased demand, the system is to be expanded, and the superstructure of the part that can be developed for the expansion is showed by the dashed lines in Figure 10. This superstructure includes 6 gas fields (B11, SE, M3, M4, M1, JN) that contain 31 wells, 1 remote subnetwork vertex (HL), 5 platforms (B11P, M3P, M1P, E11R-B, E11R-C), 1 pipeline junction (T), 2 plant slug catchers (SC-2, SC-3), and 2 LNG plants. Note that the existing part and potential part of the system include 5 platforms (E11P, F6, B11P, M1P,

M3P) that have an onsite compressor. If gas field B11 is to be developed, then platform B11P needs to be developed as well, and vice versa. The same relationship exists between M3 and M3P, M1 and M1P, SC-2 and LNG2, SC-3 and LNG3.

The following assumptions are made for developing the optimization model.

1. If a gas field is to be developed, all wells in the field are to be developed and the fractions of components for gas flows from these wells are the same. With this assumption, there is no pooling effect at each gas field, and the number of integer variables is reduced to 38.
2. The total mass loss during gas pre-processing (e.g., water removal, oil/gas separation, etc.) is negligible.
3. The pooling subproblem only tracks the  $CO_2$  and  $H_2S$  contents in the gas flows, as only bounds on  $CO_2$  and  $H_2S$  contents need to be observed.

One uncertain parameter is considered explicitly in the optimization model, which is the  $CO_2$  content of gas from gas field M1. This parameter is assumed to follow a normal distribution, with the mean  $\mu = 5.04$  mol% and the standard deviation  $\sigma = 0.1$  mol%. A simple scenario sampling rule is used to generate  $n$  uncertainty scenarios for the case study:

1. The scenarios are sampled in the  $6\sigma$  range and the value of  $x$  for  $i$ th scenario is:  

$$x_i = -3\sigma + \mu + 3\sigma/n + (i - 1)6\sigma/n.$$
2. The probability of the  $i$ th scenario can be calculated by equally dividing interval  $[\mu - 3\sigma, \mu + 3\sigma]$  into  $n$  subintervals and calculating the probability for each subinterval using cumulative probability distribution function  $\Phi$  for the standard normal distribution, i.e.,

$$p(x = x_i) = \frac{\Phi(-3\sigma + \mu + 6\sigma i/n) - \Phi(-3\sigma + \mu + 6\sigma(i - 1)/n)}{1 - 2\Phi(-3\sigma + \mu)} \quad (63)$$

The goal of the optimization is to maximize the expected NPV of the system over the next

25 years, while satisfying the demands for different uncertainty scenarios. Other parameters in this model are provided in the appendix.

We consider all the four formulations presented in the paper, Formulations (I-IV), for the case study. Formulation (I), the stochastic pooling model from the literature, does not consider pressure flow relationships. Formulation (II) includes both the pooling submodel and the pressure flow relationships, but it does not include flexible pressure drops that can be provided by pressure regulators. Formulation (III) allows pressure drops from pressure regulators and therefore is more realistic and practical than Formulation (II). Formulation (IV) is a rigorous reformulation of Formulation (III), which includes separable convex and nonconvex parts that allow further decomposition of the operation subproblems. The rigorous reformulation to Formulation (IV) is possible as the SGPS system satisfies Assumption 1. The sizes of the four formulations are shown in Table 2, where  $b$  represents the number of scenarios addressed in the problem formulation.

All simulations are run on a virtual machine with Ubuntu 12.04, one 3.40GHz CPU core, and 4GB RAM. All problems/subproblems are modeled using GAMS 24.2.3. The nonconvex MINLP/NLP problems/subproblems are solved with BARON 12.7.7, the convex NLP subproblems are solved with CONOPT 3.15P, and the MILP/LP subproblems are solved with CPLEX 12.6. The relative and absolute optimality tolerances are  $10^{-2}$  for decomposition. The convex relaxations required in single-loop and multi-loop NGBDs are constructed using the standard McCormick approach.<sup>35</sup>

## Results and Discussion

### Computational Results

We first solve the four optimization formulations for the expected value case (i.e., an one-scenario case where the value of the uncertain parameter is its mean value). This problem can be solved by single-loop or multi-loop NGBD. We compare the solution results of single-loop NGBD for the four formulations in Table 3. We set a 120-hour time limit for all

the problems; for those that cannot be solved within this time limit, the best objective value obtained before the termination and the optimality gap are reported in the table. Note that only Formulation (IV) can be solved by multi-loop NGBD, because the operation subproblems in other formulations are not separable in convex and nonconvex variables and GBD cannot be applied for solving these operation subproblems.<sup>28</sup>

It can be seen from Table 3 that, Formulation (I) can be solved by single-loop NGBD very quickly. However, this formulation ignores the pressure flow relationships and therefore may generate a solution that is unrealistic. We will give further discussion on this later. While Formulation (II-IV) are more realistic than Formulation (I), they cannot be solved to the  $10^{-2}$  tolerance by single-loop NGBD within 120 hours. Compared to Formulation (III), Formulation (IV) is more difficult to solve with single-loop NGBD. This is because Formulation (IV) includes more variables and constraints in order to achieve a better decomposable structure, and this structure is not fully exploited by single-loop NGBD. When multi-loop NGBD is applied, the decomposable structure of Formulation (IV) is better exploited and therefore a  $10^{-2}$ -optimal solution can be obtained within 10 hours.

In order to further compare the performance of single-loop NGBD and multi-loop NGBD, we solve Formulation (IV) for a case where 9 uncertainty scenarios (sampled using the rule explained before) are explicitly included in the model. The sizes of subproblems to be solved in single-loop NGBD and multi-loop NGBD, are provided in Table 4 and 5, respectively. The computational results are summarized in Table 6. It can be seen that the single-loop NGBD cannot solve the problem within 120 hours; at the end of the 120th hour, the gap is still as large as 43.2 % and the expected NPV achieved with the intermediate solution is way worse than the globally optimal expected NPV. Multi-loop NGBD, on the other hand, can return a global optimal solution (a gap smaller than 0.1 %) within 30 hours. The reason for multi-loop NGBD being faster than single-loop NGBD is that the operation subproblems (i.e., Problem (PP<sub>w</sub><sup>l</sup>)) are solved via GBD rather than directly by BARON. Due to its nice decomposable structure, each operation subproblem can be solved by GBD efficiently.

In order to see why GBD is efficient for the operation subproblems, we show in Table 7 computational results for an operation subproblem. In this subproblem, the design decisions  $x_0$  are fixed to the optimal design decisions and the scenario to be considered is the nominal scenario (i.e., the scenario in which the uncertainty parameter takes its mean value). From the table, the monolith approach (i.e., solving the problem by BARON directly) cannot guarantee the global optimal solution within 900 seconds, and the gap at the 900th second is larger than the predefined tolerance  $10^{-3}$ . On the other hand, GBD can return a solution with a gap smaller than the tolerance, using less than 20 seconds. When using GBD, the CPU time is dominated by the time for solving nonconvex relaxed master problems ( $\text{BRMP}_w^{l,k}$ ); each such nonconvex subproblem can be solved much more quickly than the operation problem itself (on average around 0.5 second per subproblem), this is because these subproblems have smaller problem sizes. In fact, the monolith approach generates the optimal solution within a second, but it cannot close the gap and verify the global optimality of the solution quickly, while in the GBD approach the lower bound approaches the optimal objective value quickly although the optimal solution is generated at the termination of the GBD procedure (about the 18th second). Figure 11 compares the lower bounds generated from the monolith and the GBD approaches over time during the solution procedures.

## Production Network Design Results

Here we compare the design results from Formulation (I) and Formulation (IV) in order to demonstrate the need for including pressure-flow relationship in the optimization model. Figure 12 and Figure 13 show the designed SGPS networks from using Formulation (IV) and Formulation (I), respectively. In the former design, gas fields M3, M1 and JN containing 23 wells in total are developed to supply raw natural gas, while in the latter design only gas field M1 is developed. Although the latter design seems to be more economic as it includes fewer gas fields and platforms, the compressor on platform M1P cannot provide sufficient pressure rise in order to drive 2292 Mmol gas per day to the LNG plants (under the nominal

scenario), in order to meet the product demands at the plants. This is the reason why in the former design, only about 1014 Mmol gas per day is planned to be generated from M1, and other raw gas is planned to be generated from M3 and JN.

## Conclusions

In this paper, a two-stage stochastic programming model is proposed for integrated design and operation of natural gas production systems under uncertainty, which describes the material balances for key gas components as well as the pressure flow relationships in gas wells, long and short pipelines, and compressors. This is the first time in the literature that this model is proposed and practically solved for gas production networks. Although the model is proposed for a typical industrial gas production system, most equations developed in this paper are also valid for other gas production systems. While the importance of considering material balances for key gas components has been demonstrated in the literature,<sup>21</sup> the case study results in this paper demonstrate the importance of including pressure flow relationships for achieving a realistic optimization result.

In order to solve the large-scale nonconvex MINLP from the proposed model, the model is reformulated in a number of steps such that the model becomes 'more convex' (i.e., the number of nonconvex constraints is reduced) and the convex and nonconvex parts are separable (which enables the application of GBD for operation subproblems). This reformulation is rigorous if a mild assumption on the compressor locations is satisfied, for the whole production network or for a subpart of it. This novel reformulation strategy should also be applicable to other types of networks where pressures play an important role (such as water networks), and it is interesting to generalize the proposed reformulation strategy to other networks in the future.

Since the reformulation allows the further decomposition of the operation subproblems, a variant of NGBD, called multi-loop NGBD method, is developed in this paper for efficiently

solving the proposed reformulated model to global optimality. In the multi-loop NGBD, each primal subproblem (which is an operation subproblem) is solved using GBD rather than using a global optimization solver directly, and it is faster than the standard single-loop NGBD by at least an order of magnitude. This is because each operation subproblem can be solved by GBD much more efficiently. This result indicates that, while GBD has been long considered as an efficient solution approach for multi-scenario problems, it can also be an efficient approach for one-scenario/deterministic optimization problems, such as deterministic optimal operation problems that are very common in chemical engineering practice. Successful application of GBD to a deterministic optimization problem may require novel reformulation of the optimization model in order to achieve a decomposable structure. If the operation problem for a system is decomposable, the integrated design and operation problem (under uncertainty) for the system can be solved by the multi-loop NGBD developed in this paper.

## Acknowledgment

The authors are grateful to the financial support from National Sciences and Engineering Research Council of Canada (RGPIN 418411-13).

## Literature Cited

- (1) International Energy Agency, *World Energy Outlook, Executive Summary*; 2015.
- (2) Mokhatab, S.; Poe, W. A. *Handbook of natural gas transmission and processing*; Gulf Professional Publishing, 2012.
- (3) Haugen, E. L. Alternative CO<sub>2</sub> Removal Solutions for the LNG Process on an FPSO. **2011**,

- (4) Rabeau, P.; Paradowski, H.; Launois, J. How to reduce CO<sub>2</sub> emissions in the LNG chain. LNG15 International Conference. 2007.
- (5) Jaramillo, P.; Griffin, W. M.; Matthews, H. S. Comparative Life Cycle Carbon Emissions of LNG Versus Coal and Gas for Electricity Generation. **2007**,
- (6) Ebenezer, S. A.; Gudmunsson, J. Removal of Carbon dioxide from natural gas for LPG production. *Semester project work* **2005**,
- (7) Aronofsky, J.; Williams, A. The Use of Linear Programming and Mathematical Models in Under-Ground Oil Production. *Management Science* **1962**, *8*, 394–407.
- (8) Bohannon, J. M. A linear programming model for optimum development of multi-reservoir pipeline systems. *Journal of Petroleum Technology* **1970**, *22*, 1–429.
- (9) Sullivan, J. A computer model for planning the development of an offshore gas field. *Journal of Petroleum Technology* **1982**, *34*, 1555–1564.
- (10) Rothfarb, B.; Frank, H.; Rosenbaum, D.; Steiglitz, K.; Kleitman, D. Optimal design of offshore natural-gas pipeline systems. *Operations research* **1970**, *18*, 992–1020.
- (11) Murtagh, B. A. An approach to the optimal design of networks. *Chem. Eng. Sci.* **1972**, *27*, 1131.
- (12) Soliman, F. I.; Murtagh, B. A. The Solution of Large-scale Gas Pipeline Design Problems. *Engineering Optimization* **1982**, *6*, 77–83.
- (13) Selot, A.; Kuok, L. K.; Robinson, M.; Mason, T. L.; Barton, P. I. A short-term operational planning model for natural gas production systems. *AIChE Journal* **2008**, *54*, 495–515.
- (14) Van den Heever, S. A.; Grossmann, I. E.; Vasantharajan, S.; Edwards, K. A Lagrangean decomposition heuristic for the design and planning of offshore hydrocarbon field in-

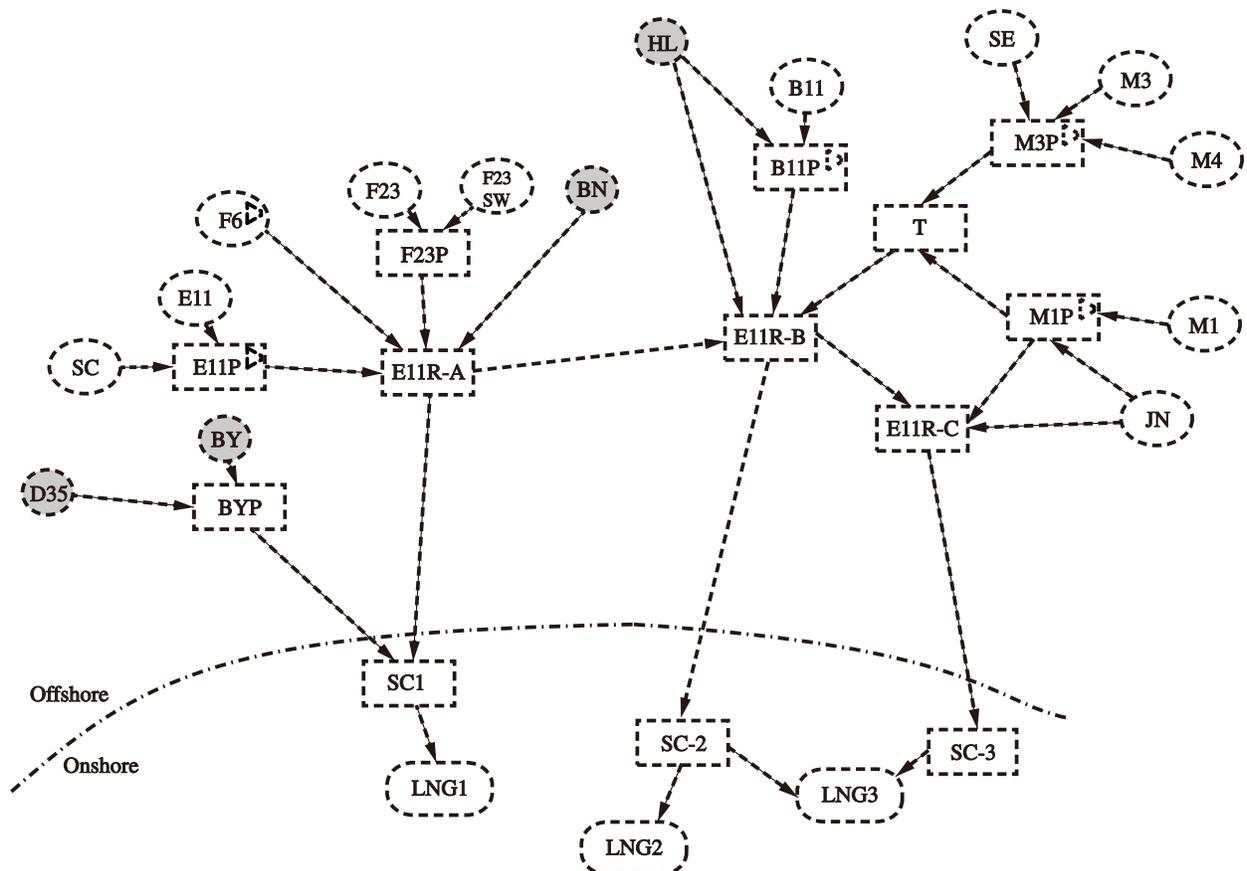
- frastructures with complex economic objectives. *Industrial & engineering chemistry research* **2001**, *40*, 2857–2875.
- (15) Gupta, V.; Grossmann, I. E. An efficient Multiperiod MINLP Model for Optimal Planning of Offshore Oil and Gas Field Infrastructure. *Industrial & Engineering Chemistry Research* **2012**, *51*, 6823–6840.
- (16) Goel, V.; Grossmann, I. E. A stochastic programming approach to planning of offshore gas field developments under uncertainty in reserves. *Computers & chemical engineering* **2004**, *28*, 1409–1429.
- (17) You, F.; Pinto, J. M.; Grossmann, I. E.; Megan, L. Optimal distribution-inventory planning of industrial gases. II. MINLP models and algorithms for stochastic cases. *Industrial & Engineering Chemistry Research* **2011**, *50*, 2928–2945.
- (18) Gupta, V.; Grossmann, I. E. Multistage Stochastic Programming Approach for Offshore Oilfield Infrastructure Planning under Production Sharing Agreements and Endogenous Uncertainties. *Journal of Petroleum Science and Engineering* **2014**, *124*, 180–197.
- (19) Gupta, V.; Grossmann, I. E. Offshore Oilfield Development Planning under Uncertainty and Fiscal Considerations. **2011**,
- (20) Tarhan, B.; Grossmann, I. E.; Goel, V. Stochastic programming approach for the planning of offshore oil or gas field infrastructure under decision-dependent uncertainty. *Industrial & Engineering Chemistry Research* **2009**, *48*, 3078–3097.
- (21) Li, X.; Armagan, E.; Tomasgard, A.; Barton, P. I. Stochastic pooling problem for natural gas production network design and operation under uncertainty. *AIChE Journal* **2011**, *57*, 2120–2135.
- (22) Raghunathan, A. U. Global Optimization of Nonlinear Network Design. *SIAM Journal on Optimization* **2013**, *23*, 268–295.

- (23) Misra, S.; Fisher, M. W.; Backhaus, S.; Bent, R.; Chertkov, M.; Pan, F. Optimal Compression in Natural Gas Networks: A Geometric Programming Approach. *IEEE Transactions on Control of Network Systems* **2015**, *2*, 47–56.
- (24) Birge, J. R.; Louveaux, F. *Introduction to stochastic programming*; Springer Science & Business Media, 2011.
- (25) Caroe, C. C.; Schultz, R. Dual decomposition in stochastic integer programming. *Operations Research Letters* **1999**, *24*, 37 – 45.
- (26) Karuppiah, R.; Grossmann, I. E. A Lagrangean based branch-and-cut algorithm for global optimization of nonconvex mixed-integer nonlinear programs with decomposable structures. *Journal of global optimization* **2008**, *41*, 163–186.
- (27) Benders, J. F. Partitioning procedures for solving mixed-variables programming problems. *Numerische mathematik* **1962**, *4*, 238–252.
- (28) Geoffrion, A. M. Generalized benders decomposition. *Journal of optimization theory and applications* **1972**, *10*, 237–260.
- (29) Li, X.; Tomasgard, A.; Barton, P. I. Nonconvex generalized Benders decomposition for stochastic separable mixed-integer nonlinear programs. *Journal of optimization theory and applications* **2011**, *151*, 425–454.
- (30) Bemporad, A.; Morari, M. Control of systems integrating logic, dynamics, and constraints. *Automatica* **1999**, *35*, 407 – 427.
- (31) Lobo, M. S.; Vandenberghe, L.; Boyd, S.; Lebret, H. Applications of second-order cone programming. *Linear Algebra and its Applications* **1998**, *284*, 193 – 228, International Linear Algebra Society (ILAS) Symposium on Fast Algorithms for Control, Signals and Image Processing.
- (32) Boyd, S.; Vandenberghe, L. *Convex optimization*; Cambridge university press, 2004.

- (33) Sahinidis, N. V.; Tawarmalani, M. BARON 7.2. 5: Global optimization of mixed-integer nonlinear programs. *User manual* **2005**,
- (34) Misener, R.; Floudas, C. A. ANTIGONE: Algorithms for Continuous/Integer Global Optimization of Nonlinear Equations. *Journal of Global Optimization* **2014**, *59*, 503–526.
- (35) McCormick, G. P. Computability of global solutions to factorable nonconvex programs: Part I convex underestimating problems. *Mathematical programming* **1976**, *10*, 147–175.

## List of Figures

1	The superstructure of Sarawak Gas Production System (SGPS) . . . . .	38
2	Well pipeline with regulator . . . . .	39
3	Platform with regulator . . . . .	40
4	Pipeline with regulator . . . . .	41
5	Pipeline performance reformulations . . . . .	42
6	Example gas networks . . . . .	43
7	The decomposable structure of Problem (P)/Formulation (IV) . . . . .	44
8	Comparison of the two NGBD diagrams . . . . .	45
9	Flowchart of GBD for solving $(PP_w^l)$ . . . . .	46
10	The superstructure of SGPS for the case study . . . . .	47
11	Comparison of lower bounds from GBD and the monolith approach . . . . .	48
12	The design result with Formulation (IV) including 9 scenarios . . . . .	49
13	The design result with Formulation (I) including 9 scenarios . . . . .	50



(V, E)	Edges(E)	Vertices(V)				
		Gas fields	Gas platforms	Terminal plants	Remote subnetwork	Compressors
Units	Pipelines					
Symbols	$\dashrightarrow$	$\circ$	$\square$	$\circ$	$\bullet$	$\triangleright$

Figure 1: The superstructure of Sarawak Gas Production System (SGPS)

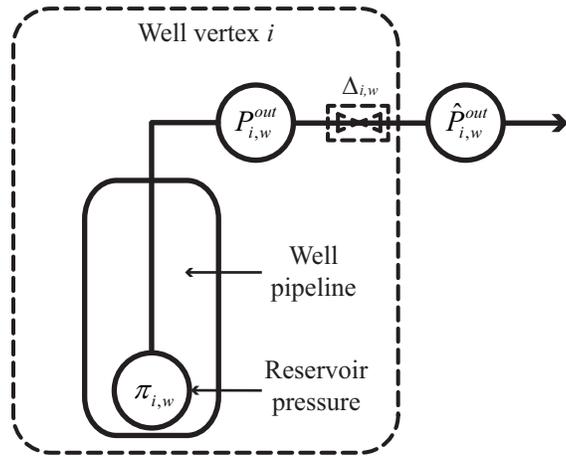


Figure 2: Well pipeline with regulator

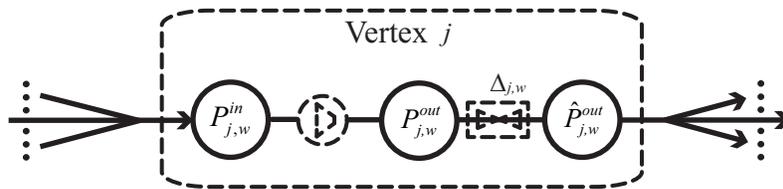


Figure 3: Platform with regulator

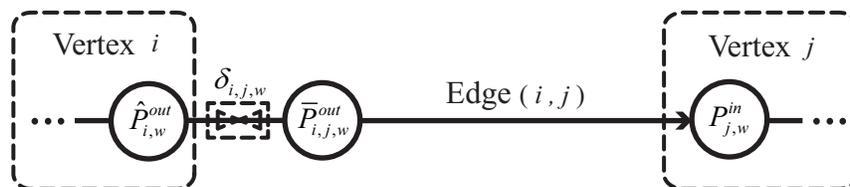
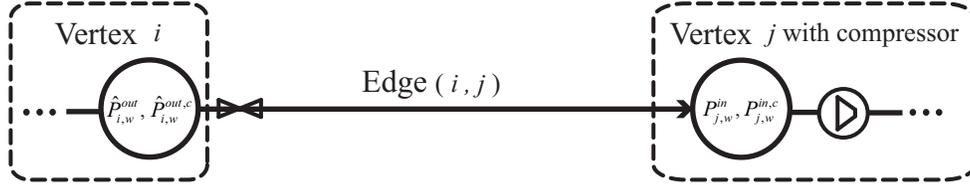


Figure 4: Pipeline with regulator



Different reformulations for pipeline performance:

- (a).  $P_{j,w}^{in,c} = (P_{j,w}^{in})^2$ ,  $\hat{P}_{i,w}^{out} \geq \sqrt{(P_{j,w}^{in})^2 + \kappa_{i,j} Q_{i,j,w}^2}$ , nonconvex
- (b).  $P_{i,w}^{out} \geq \sqrt{P_{j,w}^{in,c} + \kappa_{i,j} Q_{i,j,w}^2}$ , nonconvex
- (c).  $\hat{P}_{i,w}^{out,c} \geq P_{j,w}^{in,c} + \kappa_{i,j} Q_{i,j,w}^2$ , convex

Figure 5: Pipeline performance reformulations

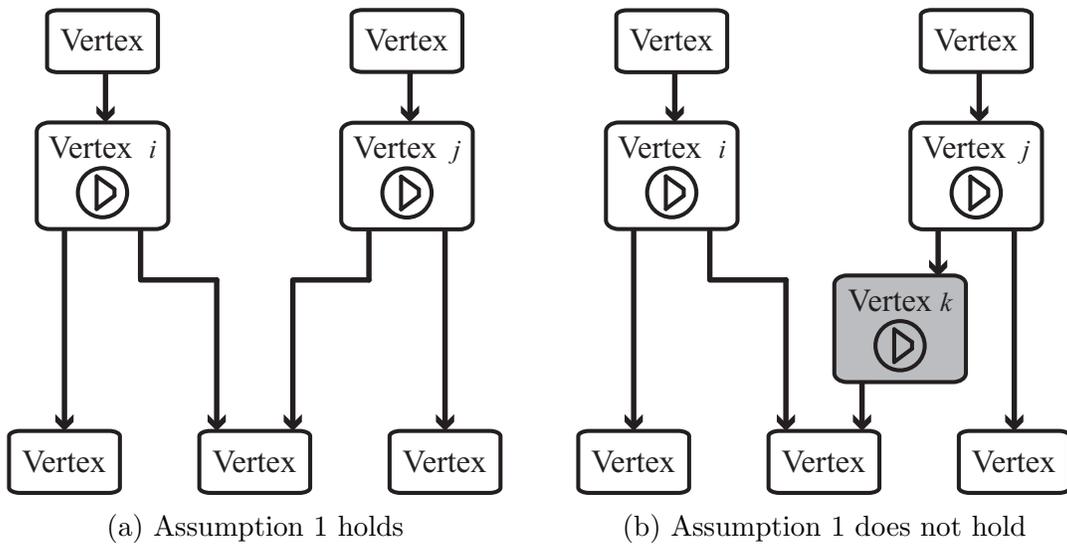


Figure 6: Example gas networks

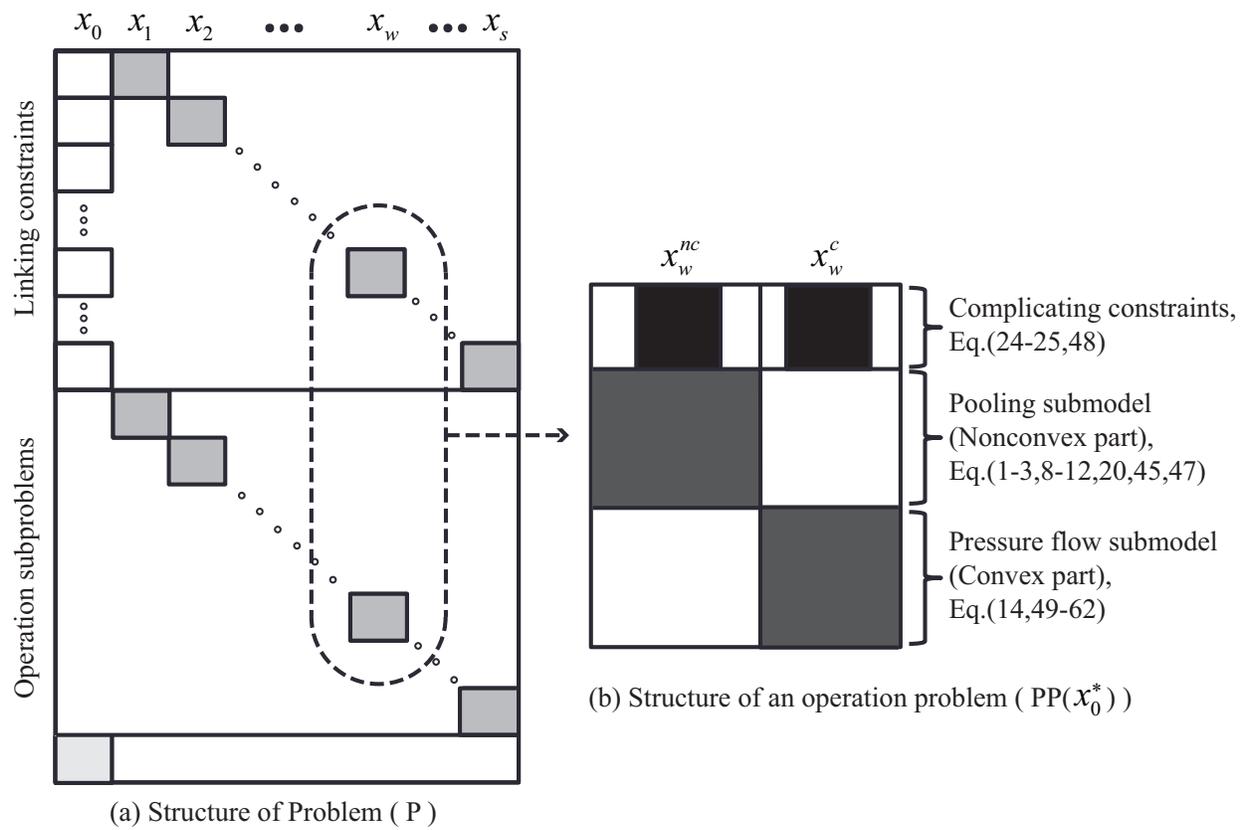
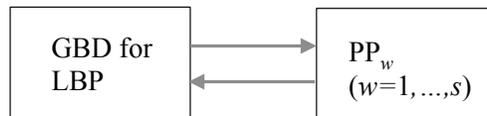
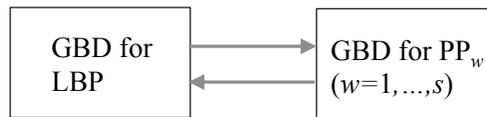


Figure 7: The decomposable structure of Problem (P)/Formulation (IV)



(a) Simplified diagram for single-loop NGBD



(b) Simplified diagram for multi-loop NGBD

GBD: Generalized Benders decomposition  
 LBP: Lower bounding problem  
 PP<sub>w</sub>: Primal subproblem for scenario  $w$

Figure 8: Comparison of the two NGBD diagrams

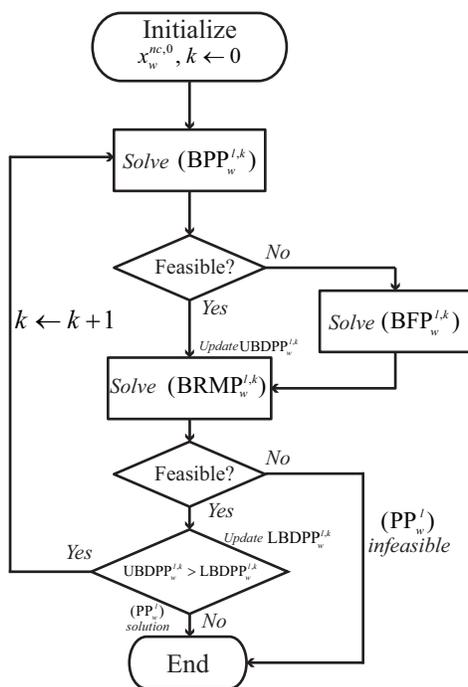
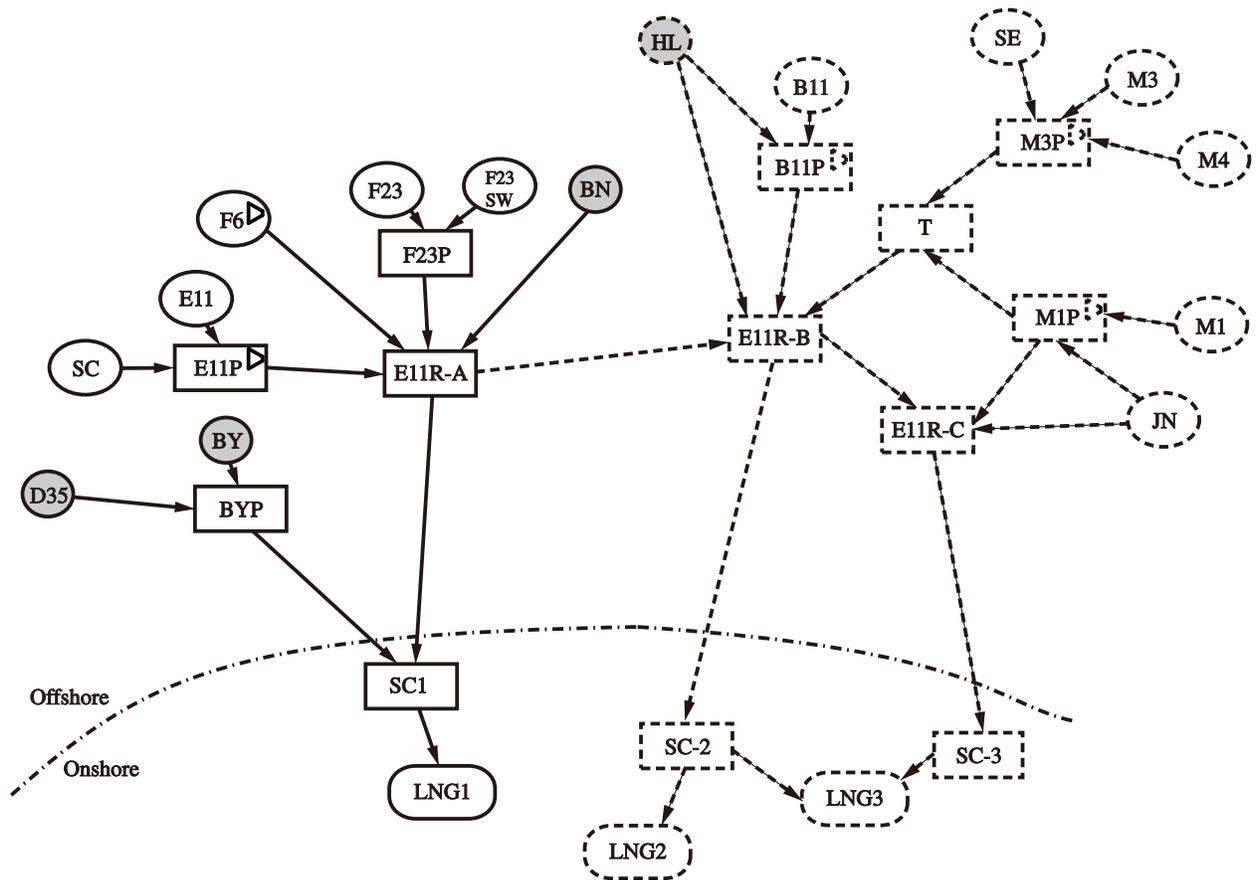


Figure 9: Flowchart of GBD for solving  $(PP_w^l)$



(V, E)	Edges(E)	Vertices(V)				
Units	Pipelines	Gas fields	Gas platforms	Terminal plants	Remote subnetwork	Compressors
Existed structures	→	○	□	◌	●	▷
Available structures	->	○	□	◌	●	▷

Figure 10: The superstructure of SGPS for the case study

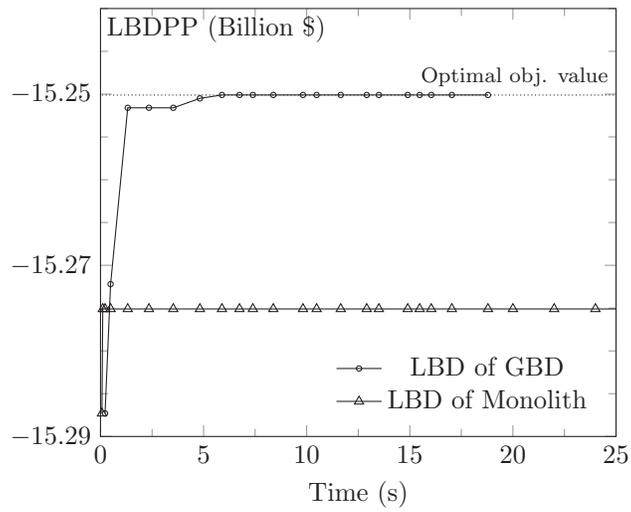
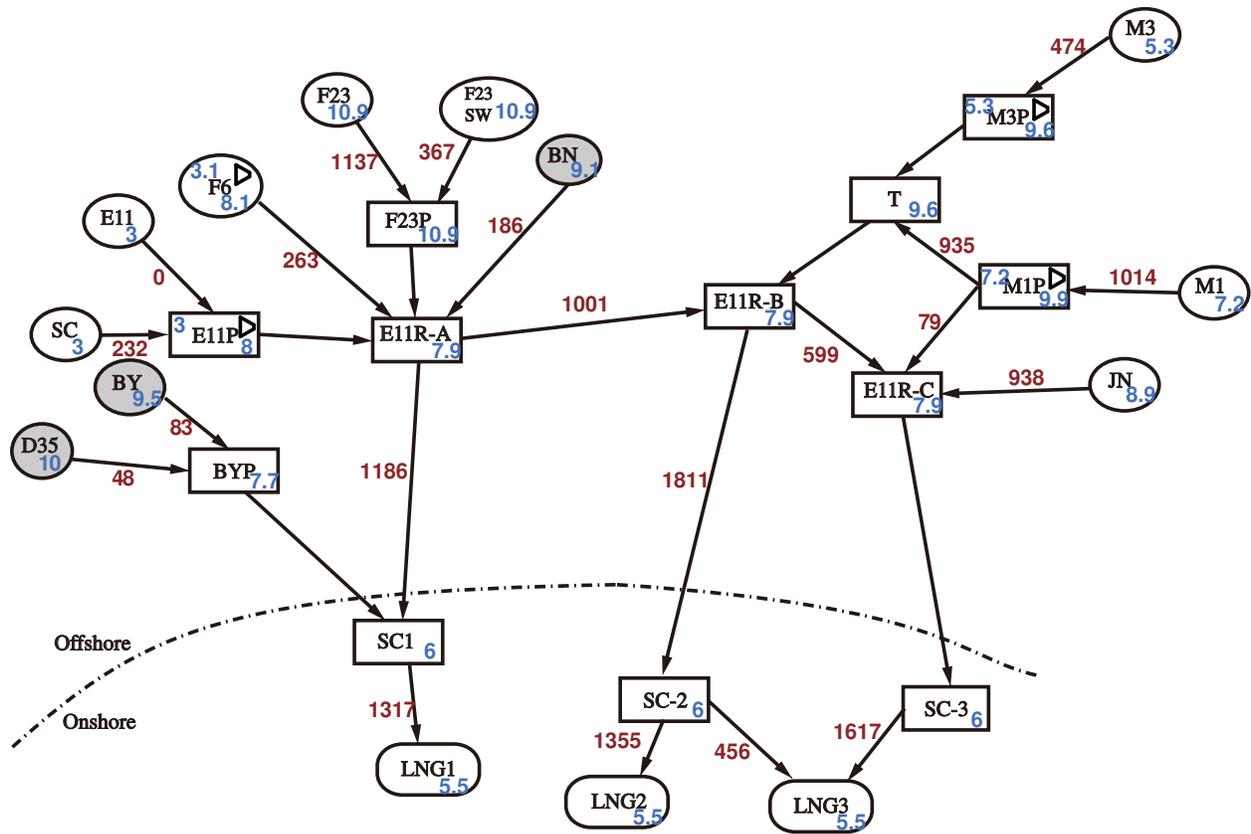


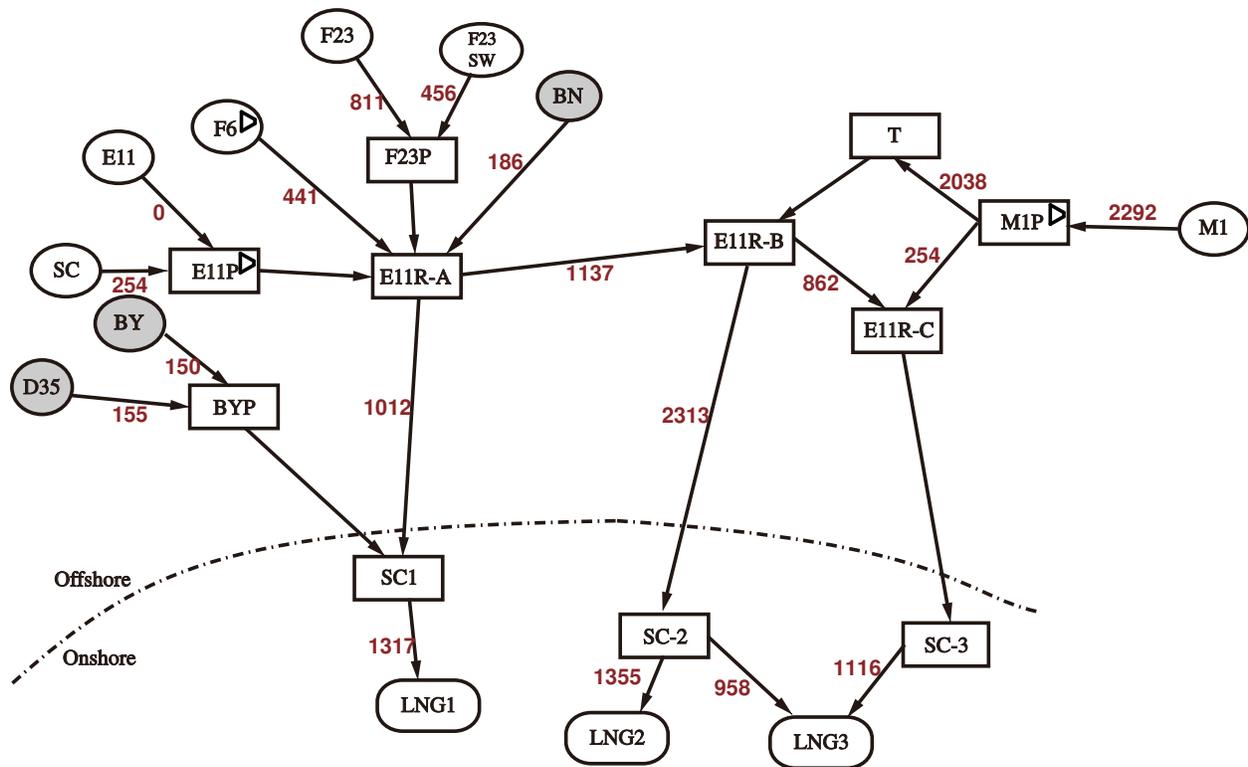
Figure 11: Comparison of lower bounds from GBD and the monolith approach



(V, E)	Edges(E)	Vertices(V)				
Units	Pipelines	Gas fields	Gas platforms	Terminal plants	Remote subnetwork	Compressors
Symbols	gas flow (Mmol/d)					

Note: The value near the each vertex is the optimal flow (Mmol/d) under the nominal scenario.

Figure 12: The design result with Formulation (IV) including 9 scenarios



(V, E)	Edges(E)	Vertices(V)				
Units	Pipelines	Gas fields	Gas platforms	Terminal plants	Remote subnetwork	Compressors
Symbols	gas flow (Mmol/d)					

Note: The value near the each vertex is the optimal flow (Mmol/d) under the nominal scenario, and no compressor models are considered.

Figure 13: The design result with Formulation (I) including 9 scenarios

## List of Tables

1	Descriptions of Symbols . . . . .	52
2	Problem sizes of Formulation (I)-(IV) <sup>†</sup> . . . . .	54
3	Computational results for the expected value problem . . . . .	55
4	Sizes of subproblems in single-loop NGBD for Formulation (IV) <sup>†</sup> . . . . .	56
5	Sizes of subproblems in multi-loop NGBD for Formulation (IV) <sup>†</sup> . . . . .	57
6	Computational results for Formulation(IV) with 9 scenarios . . . . .	58
7	Computational Results for an operation subproblem . . . . .	59

Table 1: Descriptions of Symbols

Symbol	Type	Descriptions
$E$	set	edges in the directed flow graph
$E^L$	subset of $E$	edges representing long pipelines
$E^W$	subset of $E$	edges connecting wells and platforms
$V$	set	vertices in the directed flow graph
$\bar{V}$	subset of $V$	vertices located before compressors
$V^C$	subset of $V$	vertices with onsite compressors
$V^S$	subset of $V$	vertices representing remote subnetwork
$V^T$	subset of $V$	vertices representing terminal plants
$V^W$	subset of $V$	vertices representing wells
$\Omega$	set	index set for components in flow
$\Phi$	set	index set for scenarios
$in$	superscript	indicator of the inlet pressures
$out$	superscript	indicator of the outlet pressures
$v$	superscript	indicator of vertices
$CC$	superscript	indicator of the capital cost
$LB$	superscript	indicator of the lower bounds
$NPV$	superscript	indicator of the net present value
$UB$	superscript	indicator of the upper bounds
$q$	subscript	index of components
$w$	subscript	index of scenarios
$f$	variable	molar flow rate, Mmol/d
$s$	variable	flow split ratio
$y$	binary variable	decision variable for development of vertex or edge
$AP$	variable	annual operation profit, Million \$/a
$C$	variable	cost, Million \$
$P$	variable	pressure before pressure regulator on a vertex, kPa
$\hat{P}$	variable	pressure after pressure regulator on a vertex, kPa
$\bar{P}$	variable	pressure after pressure regulator on an edge, kPa
$Q$	variable	volumetric flow rate, m <sup>3</sup> /d
$W$	variable	compressor power, MW
$\delta$	variable	pressure drop due to pressure regular at an edge, kPa
$\Delta$	variable	pressure drop due to pressure regular at a vertex, kPa
$r$	parameter	annual discount rate
$p$	parameter	probability of scenarios
$D$	parameter	demands of product at terminal plants, Mmol/d
$F$	parameter	flow bounds for edges, Mmol/d
$I$	parameter	investment cost, Million \$
$K$	parameter	composition bounds for terminal plants (%)
$M$	parameter	the big-M in constraints
$U$	parameter	component percentage at wells, (%)

Symbol	Type	Descriptions
$UC$	parameter	unit operation cost, Million \$/Mmol
$UP$	parameter	unit product profit, Million \$/Mmol
$Z$	parameter	flow bounds for vertices, Mmol/d
$\alpha$	parameter	Darcy's constant, $\text{Pa}^2 \cdot \text{d}^2 / \text{m}^6$
$\beta$	parameter	non-Darcy correction factor, $\text{Pa}^2 \cdot \text{d}^2 / \text{m}^6$
$\gamma$	parameter	power cost, Million \$/MW · d
$\tau$	parameter	time conversion constant, d/a
$\kappa$	parameter	coefficients for long pipelines, $\text{Pa}^2 \cdot \text{d}^2 / \text{m}^6$
$\lambda$	parameter	well performance coefficients
$\nu$	parameter	coefficient for compression model
$\pi$	parameter	reservoir pressure, MPa
$\sigma$	parameter	compressors capability ratio
$\vartheta$	parameter	well performance coefficients, $\text{Pa}^2 \cdot \text{d}^2 / \text{m}^6$
$\phi$	parameter	unit conversion constant, mol/m <sup>3</sup>
$\Gamma$	parameter	pressure bounds for vertices, MPa
$\Psi$	parameter	bounds for compressor power, MW

Table 2: Problem sizes of Formulation (I)-(IV)<sup>†</sup>

Formulation	(I)	(II)	(III)	(IV)
Number of binary var.	38	38	38	38
Number of continuous var.	$209+419b$	$209+662b$	$209+662b$	$209+667b$
Number of linear constr.	$363+461b$	$363+1180b$	$363+1180b$	$363+1180b$
Number of convex nonlinear constr.	0	0	$87b$	$87b$
Number of nonconvex constr.	$87b$	$179b$	$92b$	$97b$

<sup>†</sup> Sizes of the problems may vary on GAMS, depending on the GAMS preprocessing procedure.  $b$  is number of scenarios (i.e.,  $b = |\Phi|$ ).

Table 3: Computational results for the expected value problem

Formulation	Solution Method	Best obj. (Billion \$)	Gap	CPU time (h)
(I)	Single-loop NGBD	16.1	< 1 %	< 0.001
(II)	Single-loop NGBD	10.5	39.7 %	120
(III)	Single-loop NGBD	14.9	6.4%	120
(IV)	Single-loop NGBD	10.2	57.2%	120
(IV)	Multi-loop NGBD	15.1	< 1%	9.3

Table 4: Sizes of subproblems in single-loop NGBD for Formulation (IV)<sup>†</sup>

Subproblems	PP <sub>w</sub> <sup>l</sup>	LBP
Number of binary var.	0	38
Number of continuous var.	876	209+1074 <i>b</i>
Number of linear constr.	1207	363+2793 <i>b</i>
Number of convex nonlinear constr.	87	87 <i>b</i>
Number of nonconvex constr.	97	0

<sup>†</sup> Size of Problem LBP may vary on GAMS, depending on convex relaxations used and the GAMS preprocessing procedure. *b* is number of scenarios (i.e.,  $b = |\Phi|$ ).

Table 5: Sizes of subproblems in multi-loop NGBD for Formulation (IV)<sup>†</sup>

Subproblems	BPP <sub>w</sub> <sup>l,k</sup>	BRMP <sub>w</sub> <sup>l,k</sup>	LBP
Number of binary var.	0	0	38
Number of continuous var.	522	360	209+1074 <i>b</i>
Number of linear constr.	707	500+ <i>k</i>	363+2793 <i>b</i>
Number of convex constr.	92	0	87 <i>b</i>
Number of nonconvex constr.	0	97	0

<sup>†</sup> The sizes of all subproblems may vary on GAMS, depending on convex relaxations used and the GAMS preprocessing procedure. The sizes of BPP<sub>w</sub><sup>l,k</sup> and BRMP<sub>w</sub><sup>l,k</sup> are also dependent on the values of binary variables ( $x_0^l$ ), and the ones provided in the table are the largest possible sizes during the NGBD procedure.  $b$  is number of scenarios (i.e.,  $b = |\Phi|$ ).

Table 6: Computational results for Formulation(IV) with 9 scenarios

Method	Best obj. (Billion \$)	Gap	Total CPU time (h)	CPU Time for PP <sup>†</sup> (h)	Num. of PP
Single-loop NGBD	10.2	43.2%	120	120	22
Multi-loop NGBD	15.0	< 0.1 %	29.5	29.5	36

† PP refers to primal problem (or equivalently, operation subproblem), which can be decomposed for each scenario  $w$  as (PP <sub>$w$</sub>  <sup>$l$</sup> )

Table 7: Computational Results for an operation subproblem

Method	Best obj. (Billion \$)	Gap Gap	Total CPU time (s)	CPU Time for BRMP (s)	Num. of BRMP <sup>†</sup>
Monolith <sup>‡</sup>	-15.25	0.16%	900	-	-
GBD	-15.25	< 0.01 %	18.79	18.61	34

<sup>†</sup> BRMP refers to  $(\overline{\text{BRMP}}_w^{l,k})$ , the relaxed master problem in a GBD procedure for solving a primal problem.

<sup>‡</sup> Monolith refers to solving the problem directly using BARON.