Natural gas production network infrastructure development under uncertainty

Citation: Li, Xiang, Asgeir Tomasgard, and Paul I. Barton. "Natural Gas Production Network Infrastructure Development under Uncertainty." Optimization and Engineering 18.1 (2017): 35–62.

As Published: http://dx.doi.org/10.1007/s11081-016-9322-5

Publisher: Springer US

Persistent URL: http://hdl.handle.net/1721.1/107637

Version: Author's final manuscript: final author's manuscript post peer review, without publisher's formatting or copy editing

Noname manuscript No.
(will be inserted by the editor)

Natural gas production network infrastructure development under uncertainty

Xiang Li · **Asgeir Tomasgard** · **Paul I. Barton**

Received: date / Accepted: date

Abstract Mathematical programming has been widely applied for the planning of natural gas production infrastructure development. As the production infrastructure involves large investments and is expected to remain in operation over several decades, the factors that will impact the gas production but cannot be foreseen before the development of the infrastructure need to be taken into account in the planning. Therefore, two scenario-based two-stage stochastic programming models are developed to facilitate natural gas production infrastructure development under uncertainty. One is called the stochastic pooling model, which tracks the qualities of gas streams throughout the production network via a generalized pooling model. The other is an enhancement of the stochastic pooling model with the consideration of pressure. Either model results in a large-scale nonconvex mixed-integer nonlinear programming (MINLP) problem, for which a global optimal solution, although very important for a problem that involves large investments, is very difficult to obtain. A novel optimization method, called nonconvex generalized Benders decomposition (NGBD), is developed for efficient global optimization of the large-scale nonconvex MINLP. Case studies of a real industrial natural gas production system show the advantages of the proposed stochastic programming models over deterministic optimization models, as well as

This work was supported by Statoil and the research council of Norway (project nr176089/S60) as part of the paired Ph.D. research program in gas technologies between MIT and NTNU.

Xiang Li
Process Systems Engineering Laboratory, Department of Chemical Engineering, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA, 02139, USA E-mail: xiangli@mit.edu
Present address: Department of Chemical Engineering, Queen's University, 19 Division Street, Kingston, ON, K7L3N6, Canada E-mail: xiang.li@queensu.ca

Asgeir Tomasgard
Department of Industrial Economics and Technology Management, Norwegian University of Science and Technology, 7491 Trondheim, Norway E-mail: asgeir.tomasgard@iot.ntnu.no

Paul I. Barton (Corresponding Author)
Process Systems Engineering Laboratory, Department of Chemical Engineering, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA, 02139, USA E-mail: pib@mit.edu

the dramatic computational advantages of NGBD over a state-of-the-art global optimization solver.

Keywords Natural gas production network · Mixed-integer nonlinear programming · Stochastic programming · Benders decomposition · Global optimization

Mathematics Subject Classification (2000) MSC 90C11 · MSC 90C15 · MSC 90C26 · MSC 90C90

1 Introduction

Natural gas currently contributes about a fifth of global energy demand and it is projected to play an increasingly important role in the global energy economy through 2035 [30]. To meet increasing demands for natural gas, new gas production systems need to be developed or existing production systems need to be expanded. The natural gas infrastructure development problem can be cast as an integrated design and operation problem, as it is natural to consider the long-term production plan when considering the infrastructure development.

The development of oil or gas production infrastructure involves large investments, and the infrastructure often remains in operation over the entire life span of the project (which can be several decades), so even small fractional performance gains made in the design can translate into significant increases in profits. Therefore, mathematical programming has long been adopted to facilitate decision-making for oil or gas production infrastructure development. With appropriate assumptions and approximations, an infrastructure development problem can be formulated as a linear programming (LP) [4] or mixed-integer linear programming (MILP) problem [10] [58] [50] [49] [31], for which reliable and efficient commercial solvers are available (e.g., CPLEX [29]). However, the optimality or even the feasibility of the solution may be lost due to the linear approximation of the inherently nonlinear production system. In order to reduce the model mismatch, appropriate nonlinearities may be introduced into the model, such as nonlinear reservoir or well performance models [13] [45] [66] [65] [43], nonlinear compression model [13], and nonlinear relationships between gas flow rates and pressures [13] [43]. The introduction of nonlinearities usually leads to mixed-integer nonlinear programming problems (MINLP), which are more difficult to solve than MILPs. Readers can refer to a survey paper by Grossmann [23] for details on MINLP solution technologies. Notice that for a nonconvex MINLP , i.e., a MINLP involving nonconvex functions, global optimization techniques [28] are required to guarantee that the solution obtained is a global optimum, which is especially important for problems involving large capital investment. Typical global optimization methods for nonconvex MINLPs include branch-and-reduce [61], SMIN-αBB and GMIN-αBB [1], and nonconvex outer approximation [35] (which is an extension of traditional outer approximation methods [11] [14] to programs with nonconvex functions participating). All these methods require solving a sequence of subproblems, and the sizes and/or the number of the subproblems to be solved rely heavily on the size of the original problem.

Before the development of an oil or gas production system, there is always unknown or uncertain information that can significantly affect the future operation of the system. For example, the quality and capacity of a reservoir may not be known exactly, and the prices of the petroleum products and customer demands in the future are uncertain. It has been widely recognized that considering these uncertainties in the planning of the infrastructure development can help to reduce risks and improve expected profits, so the use of stochastic programming [9] models has attracted growing interest, especially two-stage or multistage scenario-based models [32] [21] [22] [60] [39]. A scenario-based stochastic program includes a set of scenarios, each of which represents a possible and representative group of values for uncertain parameters that may be realized in the future. The uncertain parameters may include capacity and quality of reservoir, costumer demands and gas prices The number of the scenarios required by the formulation depends on the number of uncertain parameters and the characteristics of the uncertainties; for real-world industrial problems, this number can be very large. Therefore, with the decisions and submodels for each of the scenarios included, the scenario-based program is likely to be a large-scale mathematical program. On the other hand, these large-scale mathematical programs retain a decomposable structure that can be exploited by specialized optimization methods for efficient solution, e.g., Benders decomposition [7] (also called *L*-shaped model in the stochastic programming literature [57]) for linear models, and generalized Benders decomposition [20] for a class of nonnlinear models. However, these methods require strong duality [8] for convergence to a global optimum, which prevents them from solving nonconvex MINLPs such as those arising in natural gas infrastructure development under uncertainty. Lagrangian decomposition [34] has also been applied to exploit the structure of scenario-based problems. However, if strong duality does not hold, this method needs branch-and-bound operations in the full variable space, the dimension of which depends on the number of scenarios, to guarantee global optimality [34] [36].

In all the mentioned work on oil or gas production infrastructure development, product qualities (e.g., composition) are not included in the optimization models. One reason is that most of the work focuses on oil production systems, where the quality of the crude oil is not a key quantity to be controlled in the production system. In fact, oil product qualities are ensured by processing in the downstream refineries. However, gas products from natural gas production systems will be sent to customers with little further processing, so they have to satisfy strict specifications. As the raw natural gas streams entering a production network may come from different reservoirs with different levels of impurities (e.g., CO_2, H_2S), modeling the composition of the gas products is not trivial. Considering that these raw gas streams are mixed and split thought the production network with little processing, the qualities of gas products can be evaluated through a pooling model [24] [47], which was originally studied for gasoline blending in oil refineries. The pooling model has been adopted for operational optimization of natural gas production [55] [64] [52], but not yet for the infrastructure development (except in the authors' recent work [39] [42] [41]). The pooling problem is a class of bilinear optimization problem, which is highly nonconvex and difficult to solve. Solution methods for the pooling problem have been studied for decades. The methods that locate a local optimum include recursive guessing [24] [25], successive LP [38], and generalized Benders decomposition [15], and the methods that guarantee a global optimum include GOP [16] [67] and branch-and-bound [17] [6] [61] [48]. In addition, reformulation-linearization techniques [56] [51] and piecewise relaxation methods [33] [46] [68] have been proposed to generate tighter relaxations to improve global optimization of the pooling problem.

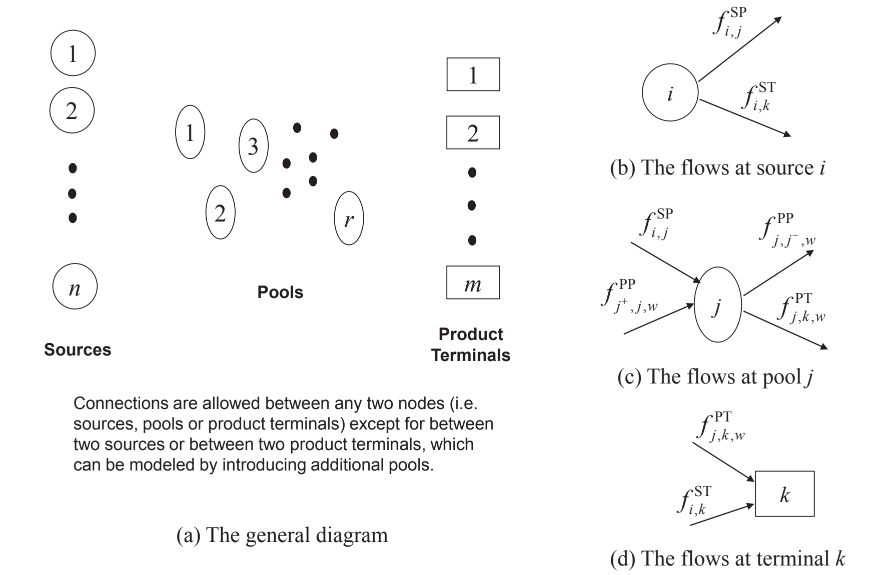
This paper presents mathematical programming models and solution methods, for the planning of natural gas infrastructure development under uncertainty, in a unified framework. The contents of this paper are based on the results of the authors' recent work [39] [42] [41]. Here the goal of the planning is, to determine the optimal system design decisions and long-term operating conditions for natural gas production that maximize the expected profitability of developing and operating the system, while satisfying the product-specific constraints for all the uncertainty scenarios addressed. A scenario-based, two-stage stochastic programming model is developed, where the first stage decisions determine whether or not gas fields/wells, gas platforms or trunklines in the production network are to be developed, and the second stage decisions determine different long-term operating conditions for different realizations of the uncertain parameters, known after the development of the production system. This stochastic programming model is primarily based on a generalized pooling model that tracks the qualities of the gas streams throughout the production network, but it also allows the consideration of pressure via integrating well performance models, compression models and trunkline pressure-flow relationships. The resulting mathematical programming problem is a potentially large-scale nonconvex MINLP. As mentioned above, traditional global optimization methods may have to solve a large number of large-scale subproblems to locate a global optimum of this problem, and no valid decomposition methods in the literature can take advantage of the special problem structure. To this end, a novel decomposition method, called nonconvex generalized Benders decomposition (NGBD), is developed for efficient global optimization of the stochastic nonconvex MINLP. This method only requires solution of a sequence of subproblems whose sizes are independent of the number of scenarios addressed, and computational experience shows that number of the subproblems to be solved does not increase significantly with the number of scenarios.

The rest of the paper is organized as follows: Section 2 describes the stochastic nonconvex MINLP model for natural gas production infrastructure development under uncertainty; Section 3 introduces the NGBD method for the global optimization of the stochastic MINLP; Section 4 presents two industrial case studies, one addresses stream composition for the production network, but not the effects of pressure, and the other addresses both; Section 5 summarizes the results and suggests future work.

2 Mathematical Programming Models

2.1 General Settings

A natural gas production system is viewed as a generalized pooling system here. Fig. 1 illustrates the system, which has n sources (labelled from 1 to n) that supply materials into the system, r pools (labelled from 1 to r) where different materials or



(a) The general diagram

(b) The flows at source *i*

(c) The flows at pool *j*

(d) The flows at terminal *k*

Fig. 1 The generalized pooling system.

intermediate products are mixed or blended, *m* product terminals (labelled from 1 to *m*) that yield final products. For a natural gas production system, the sources can be gas fields or individual gas wells in the gas fields, the pools can be production platforms, riser platforms or simple mixing and splitting units, and the product terminals can be liquefied natural gas (LNG) plants that produce LNG for long distance transportation or dry gas terminals supplying end customers through gas transmission systems. Different from the conventional pooling system, this system not only allows connections between a source and a product terminal, a source and a pool, or a pool and a product terminal, but also allows connections between two pools. Connections between two sources or between two product terminals, which are also possible in many engineering problems, are not considered explicitly in this paper for simplicity of the resulting mathematical model. But these connections can be modeled within the proposed general diagram by introducing additional (virtual) pools.

Two types of formulations can be used to model the generalized pooling system. One is to formulate the mass balance equations with total flows and component compositions, and the other is to express the mass balances with individual component flows. As has been well recognized [51], these two formulations have their own pros and cons respectively. In general, the first formulation will lead to more bilinear terms if the total number of mixing flows entering the pools are more than the total number of splitting flows leaving the pools; the second formulation will lead to more bilinear terms otherwise. Natural gas production systems usually collect gas from certain numbers of gas wells through smaller numbers of gas platforms and to smaller numbers of product terminals, and merging instead of dividing the gas flows in the transport is preferred as it helps reduce the pipeline investment cost. Therefore, natural gas production systems often involve more mixing flows than splitting flows, so the second formulation is adopted here for fewer bilinear terms. Note that the global optimization method proposed in this paper is applicable to either of the two formulations, so the choice of the second formulation does not affect the generality of the subsequent discussion on global optimization.

Figs. 1 (b)-(d) give more details on the total flows or individual component flows entering or leaving the sources, pools and product terminals. Fig. 1 (b) shows that a flow coming out of source i may either go to a pool j (denoted by $f^{\text{SP}}_{i,j}$), or go to a product terminal k (denoted by $f^{\text{ST}}_{i,k}$). Note that since the component compositions of source i are parameters, there is no need to model the individual component flows explicitly for a flow coming out of source i (which avoids introducing more bilinear terms). Fig. 1 (c) shows that a flow entering a pool j may come from a source i, or come from another pool j^+ (whose flow of component w is denoted by $f^{\text{PP}}_{j^+,j,w}$). Also, a flow leaving pool j may go to another pool j^- (whose flow of component w is denoted by $f^{\text{PP}}_{j,j^-,w}$), or go to a product terminal k (whose flow of component w is denoted by $f^{\text{PT}}_{j,k,w}$). The subscript $w \in \{1,...,l\}$ indicates the different components. Fig. 1 (d) shows that a flow entering a product terminal k may come from a pool j or a source i.

A two-stage stochastic programming problem is basically a bilevel optimization problem whose inner optimization problems represent the second-stage decision-making for different realizations of the uncertain parameters. As has been widely recognized, this bilevel optimization problem can be naturally reformulated into an equivalent single-level optimization problem, called the deterministic equivalent program (of the stochastic program) [9]. For a succinct presentation, this paper derives the stochastic programming model in the form of its deterministic equivalent program directly. Two stochastic programming models are developed in the following two subsections, respectively. The first model, called the stochastic pooling model, assumes that compressors can be installed anywhere in the system with unlimited capacity and their investment and operating costs have little impact on the design decisions of the system (e.g. these costs are negligible compared to the total infrastructure investment cost and revenue). With this assumption, pressure can be ignored in the model because any desired flow rates of the gas streams obtained from the model can be realized by placing compressors at appropriate locations in the system. The second model is developed based on the stochastic pooling model, and it considers the impact of pressure by including additional submodels that relate the pressures with the gas flow rates and energy consumption.

2.2 The Stochastic Pooling Model

The equality and inequality constraints of the stochastic pooling problem represent mass balances and physical limits at the sources, pools and product terminals, which are developed below step by step. After that, the expected net present value of the system is given as the objective function. The symbols used in the stochastic pooling model are summarized in Table 1.

Table 1 List of Symbols for the Stochastic Pooling Model

Symbol	Type	Description
a	Parameter	Discount rate for calculating net present value
f	Variable	Flow rate
h	Subscript	Index for scenarios, $h \in \{1,...,s\}$
i	Subscript	Index for sources, $i \in \{1,...,n\}$
j	Subscript	Index for pools, $j \in \{1,...,r\}$
j^+	Subscript	Index for the pools whose outlet flows enter a particular pool
j^-	Subscript	Index for the pools whose inlet flows are from a particular pool
k	Subscript	Index for product terminals, $k \in \{1,...,m\}$
l	Parameter	Number of quality components
m	Parameter	Number of product terminals
n	Parameter	Number of sources
p	Parameter	Probability of scenario
q	Variable	Ratio of a flow leaving a pool to the total flow entering the pool
r	Parameter	Number of pools
s	Parameter	Number of scenarios
t	Parameter	year
w	Subscript	Index of quality, $w \in \{1,...,l\}$
y	Binary variable	Decision on source, pool, product terminal or trunkline investment
C	Parameter	Economic coefficient with cost and price information
(Cap)	Superscript	Indicator of capital cost
D	Parameter	Demand at product terminal
F	Parameter	Bound on flow rate
L	Parameter	Life span of the system
LB	Superscript	Indicator of lower bound
N	Parameter	Number of objectives in multi-objective optimization
(OC)	Superscript	Indicator of annual cost and revenue information related to operation
P	Superscript	Indicator of pool
PP	Superscript	Indicator of flow from pool to pool
PP+	Superscript	Indicator of flow from pool entering a particular pool
PP–	Superscript	Indicator of flow from a particular pool entering a pool
PT	Superscript	Indicator of flow from pool to product terminal
S	Superscript	Indicator of source related quantity
SP	Superscript	Indicator of flow from source to pool
ST	Superscript	Indicator of flow from source to product terminal
T	Superscript	Indicator of product terminal
U	Parameter	Quality of materials at source
UB	Superscript	Indicator of upper bound
V	Parameter	Quality bound at product terminal
Z	Parameter	Source outlet flow bound
Π	Set	Index set for sources and pools connected to a product terminal
Θ	Set	Index set for pools and product terminals connected to a source
Ω	Set	Index set for sources, pools and terminals connected to a pool

Submodel for the sources - The variables related to the sources are subject to the following inequalities:

$$y_i^{\mathrm{S}} Z_{i,h}^{\mathrm{LB}} \leq \sum_{j \in \Theta_i^{\mathrm{SP}}} f_{i,j,h}^{\mathrm{SP}} + \sum_{k \in \Theta_i^{\mathrm{ST}}} f_{i,k,h}^{\mathrm{ST}} \leq y_i^{\mathrm{S}} Z_{i,h}^{\mathrm{UB}}, \tag{1}$$

$$y_{i,j}^{\mathrm{SP}} F_{i,j}^{\mathrm{SP,LB}} \leq f_{i,j,h}^{\mathrm{SP}} \leq y_{i,j}^{\mathrm{SP}} F_{i,j}^{\mathrm{SP,UB}}, \tag{2}$$

$$y_{i,k}^{\mathrm{ST}} F_{i,k}^{\mathrm{ST,LB}} \leq f_{i,k,h}^{\mathrm{ST}} \leq y_{i,k}^{\mathrm{ST}} F_{i,k}^{\mathrm{ST,UB}}, \tag{3}$$

$$y_i^{\text{S}} \geq y_{i,j}^{\text{SP}}, \quad y_i^{\text{S}} \geq y_{i,k}^{\text{ST}}, \quad y_i^{\text{S}}, y_{i,j}^{\text{SP}}, y_{i,k}^{\text{ST}} \in \{0,1\}, \tag{4}$$
$$\forall i \in \{1,...,n\},\ \forall j \in \Theta_i^{\text{SP}},\ \forall k \in \Theta_i^{\text{ST}},\ \forall h \in \{1,...,s\}.$$

Here the subscript h is attached to each scenario-dependent variable or parameter to index different scenarios and s is the total number of the scenarios. The binary variable y_i^{S} represents whether source i is to be developed or not. If $y_i^{\text{S}} = 1$, then source i will be developed, and Eq. (1) denotes that the total flow coming out of a source j is subject to a lower bound $Z_{i,h}^{\text{LB}}$ (which is due to non-negativity of flow or other system requirements) and an upper bound $Z_{i,h}^{\text{UB}}$ (which is due to the source capacity or other system requirements) for scenario h. Θ_i^{SP} is an index set containing the indices of the pools that can connect to source i and Θ_k^{ST} is an index set containing the indices of the product terminals that can connect to source i. If $y_i^{\text{S}} = 0$, Eq. (1) forces the total flow coming out of source i to be zero. Eq. (2) denotes that the flow in the trunkline connecting source i and pool j is subject to relevant bounds (that are due to the trunkline capacity, non-negativity of flow, or other system requirements) for scenario h, if this trunkline is to be developed (i.e., $y_i^{\text{SP}} = 1$); otherwise, the flow is forced to be zero. Eq. (3) denotes the similar limit for the flow in the trunkline connecting source i and product terminal k. Eq. (4) denotes the topological restrictions on the sources and the trunklines connecting to them, i.e., the trunkline between a source and a pool or a product terminal can be developed only when the source is developed.

Submodel for the pools - The submodel for the pools can be written as follows:

$$f_{j,j^-,w,h}^{\text{PP}} = q_{j,j^-,h}^{\text{PP}} \left(\sum_{i \in \Omega_j^{\text{SP}}} f_{i,j,h}^{\text{SP}} U_{i,w,h} + \sum_{j^+ \in \Omega_j^{\text{PP+}}} f_{j^+,j,w,h}^{\text{PP}} \right), \tag{5}$$

$$f_{j,k,w,h}^{\text{PT}} = q_{j,k,h}^{\text{PT}} \left(\sum_{i \in \Omega_j^{\text{SP}}} f_{i,j,h}^{\text{SP}} U_{i,w,h} + \sum_{j^+ \in \Omega_j^{\text{PP+}}} f_{j^+,j,w,h}^{\text{PP}} \right), \tag{6}$$

$$\sum_{j^- \in \Omega_j^{\text{PP-}}} q_{j,j^-,h}^{\text{PP}} + \sum_{k \in \Omega_j^{\text{PT}}} q_{j,k,h}^{\text{PT}} = 1, \quad q_{j,j^-,h}^{\text{PP}}, q_{j,k,h}^{\text{PT}} \geq 0, \tag{7}$$

$$y_{j,j^-}^{\text{PP}} F_{j,j^-}^{\text{PP,LB}} \leq \sum_{w \in \{1,...,l\}} f_{j,j^-,w,h}^{\text{PP}} \leq y_{j,j^-}^{\text{PP}} F_{j,j^-}^{\text{PP,UB}}, \tag{8}$$

$$y_{j,k}^{\text{PT}} F_{j,k}^{\text{PT,LB}} \leq \sum_{w \in \{1,...,l\}} f_{j,k,w,h}^{\text{PT}} \leq y_{j,k}^{\text{PT}} F_{j,k}^{\text{PT,UB}}, \tag{9}$$

$$f_{j,j^-,w,h}^{\text{PP}}, f_{j,k,w,h}^{\text{PT}} \geq 0, \tag{10}$$

$$y_j^{\text{P}} \geq y_{i,j}^{\text{SP}}, \quad y_j^{\text{P}} \geq y_{j,j^-}^{\text{PP}}, \quad y_{j^-}^{\text{P}} \geq y_{j,j^-}^{\text{PP}}, \quad y_j^{\text{P}} \geq y_{j,k}^{\text{PT}}, \quad y_j^{\text{P}}, y_{j,j^-}^{\text{PP}}, y_{j,k}^{\text{PT}} \in \{0,1\}, \tag{11}$$
$$\forall j \in \{1,...,r\},\ \forall j^- \in \Omega_j^{\text{PP-}},\ \forall k \in \Omega_j^{\text{PT}},\ \forall w \in \{1,...,l\},\ \forall h \in \{1,...s\}.$$

Here fractional variables $q_{j,j^-,h}^{\text{PP}}$ and $q_{j,k,h}^{\text{PT}}$ are introduced to model the mass balances at pool j. They denote the ratio of the flow from pool j to pool j^- to the total flow entering pool j for scenario h, and the ratio of the flow from pool j to product terminal k to the total flow entering pool j for scenario h, respectively. Then, each individual component flow of an outlet flow to another pool can be represented as Eq. (5) by definition, and each individual component flow of an outlet flow to a pool as Eq. (6), where parameter $U_{i,w,h}$ denotes the fraction of component w in the flow from source i for scenario h, Ω_j^{SP} is an index set containing the indices of the sources from which an inlet flow to pool j can come, $\Omega_j^{\mathrm{PP+}}$ is an index set containing the indices of the pools from which an inlet flow to pool j can come, $\Omega_j^{\mathrm{PP-}}$ is an index set containing the indices of the pools where an outlet flow from pool j can go to, and Ω_j^{PT} is an index set containing the indices of the product terminals where an outlet flow from pool j can go to. According to their definition, the split fraction variables for pool j should be nonnegative and their sum should equal unity because of mass balance at pool j, and these relationship are described by Eq. (7). Eqs. (8) and (9) enforce the relevant limit on the flow (or the sum of its individual component flows) in a trunkline connecting pool j and another pool or a product terminal, if the trunkline is to be developed; otherwise, they force the relevant flow to be zero. Eq. (10) enforces the non-negativity of each the individual component flows related to pool j . Eq. (11) represents topological restrictions, i.e., a trunkline connecting to a pool can be developed only when that pool is to be developed, where the binary decision variable y_j^{P} determines whether pool j is to be developed or not, y_{j,j^-}^{PP} determines whether the trunkline between pool j and a downstream pool j^- is to be developed or not, and $y_{j,k}^{\mathrm{PT}}$ determines whether the trunkline between pool j and product terminal k is to be developed or not.

Submodel for the product terminals - The variables related to the product terminals are subject to the following inequalities:

$$y_k^{\mathrm{T}} D_k^{\mathrm{LB}} \leq \sum_{j\in\Pi_k^{\mathrm{PT}}} \sum_{w\in\{1,\ldots,l\}} f_{j,k,w,h}^{\mathrm{PT}} + \sum_{i\in\Pi_k^{\mathrm{ST}}} f_{i,k,h}^{\mathrm{ST}} \leq y_k^{\mathrm{T}} D_{k,h}^{\mathrm{UB}}, \tag{12}$$

$$\begin{aligned} &\left(\sum_{j\in\Pi_k^{\mathrm{PT}}} \sum_{w\in\{1,\ldots,l\}} f_{j,k,w,h}^{\mathrm{PT}} + \sum_{i\in\Pi_k^{\mathrm{ST}}} f_{i,k,h}^{\mathrm{ST}} \right) V_{k,w}^{\mathrm{UB}} \geq \\ &\sum_{j\in\Pi_k^{\mathrm{PT}}} f_{j,k,w,h}^{\mathrm{PT}} + \sum_{i\in\Pi_k^{\mathrm{ST}}} f_{i,k,h}^{\mathrm{ST}} U_{i,w,h} \\ &\geq \left(\sum_{j\in\Pi_k^{\mathrm{PT}}} \sum_{w\in\{1,\ldots,l\}} f_{j,k,w,h}^{\mathrm{PT}} + \sum_{i\in\Pi_k^{\mathrm{ST}}} f_{i,k,h}^{\mathrm{ST}} \right) V_{k,w}^{\mathrm{LB}} , \end{aligned} \tag{13}$$

$$y_k^{\mathrm{T}} \geq y_{i,k}^{\mathrm{ST}}, \quad y_k^{\mathrm{T}} \geq y_{j,k}^{\mathrm{PT}}, \quad y_k^{\mathrm{T}} \in \{0,1\}, \tag{14}$$

$$\forall k \in \{1,\ldots,m\},\ \forall w \in \{1,\ldots,l\},\ \forall h \in \{1,\ldots,s\},$$

where the binary variable y_k^{T} determines whether product terminal k is to be developed. Eq. (12) means that if product terminal k is to be developed (i.e., $y_k^{\mathrm{T}} = 1$), the total flow (i.e., the the sum of all individual component flows) coming into a product terminal k for scenario h is subject to lower and upper bounds, which are related to the minimum supply required by contract and the maximum possible demand from the market or plant capacity for the scenario, respectively; otherwise, the relevant individual component flows are all zero. Π_k^{PT} is an index set containing the indices of the pools where an inlet flow to product terminal k can come from, and Π_k^{ST} is an index set containing the indices of the sources where an inlet flow to product terminal k can come from. The gas streams coming out of product terminal k are almost always subject to quality requirements imposed by contracts, technological limitations or laws, which are usually ranges for the percentages of specific components allowed in the product. Thefore, Eq. (13) enforces the quality requirements for scenario h, where $V^{\text{LB}}_{k,w,h}$ and $V^{\text{UB}}_{k,w,h}$ are the lower and upper bounds, respectively, on the fraction of component w in the final product at product terminal k for scenario h. Eq. (14) enforces topology restrictions, i.e., a trunkline connecting to a product terminal can be developed only when the product terminal is to be developed.

Objective function - There are several economic measures for the profitability of developing the natural gas production system, such as annualized profit, net present value (NPV), and internal rate of return. Here the expect NPV over the scenarios is adopted. Accordingly, the objective of the optimization is:

$$\max \quad -C^{(Cap)} + \left(\sum_{t\in\{1,\dots,L\}} \frac{1}{(1+a)^t} \right) \left(\sum_{h\in\{1,\dots,b\}} p_h C_h^{(OC)} \right), \tag{15}$$

where a is the annual discount rate, L is the lifespan of the system, t denotes the time (in years) after the development of the system. $C^{(Cap)}$ denotes the total capital cost and

$$\begin{aligned} C^{(Cap)} = & \sum_{i\in\{1,\dots,n\}} C_i^{\text{S},(Cap)} y_i^{\text{S}} + \sum_{j\in\{1,\dots,r\}} C_j^{\text{P},(Cap)} y_j^{\text{P}} + \sum_{k\in\{1,\dots,m\}} C_k^{\text{T},(Cap)} y_k^{\text{T}} \\ & + \sum_{i\in\{1,\dots,n\}} \sum_{j\in\Theta_i^{\text{SP}}} C_{i,j}^{\text{SP},(Cap)} y_{i,j}^{\text{SP}} + \sum_{i\in\{1,\dots,n\}} \sum_{k\in\Theta_i^{\text{ST}}} C_{i,k}^{\text{ST},(Cap)} y_{i,k}^{\text{ST}} \\ & + \sum_{j\in\{1,\dots,r\}} \sum_{j^-\in\Omega_j^{\text{PP-}}} C_{j,j^-}^{\text{PP},(Cap)} y_{j,j^-}^{\text{PP}} + \sum_{j\in\{1,\dots,r\}} \sum_{k\in\Omega_j^{\text{PT}}} C_{j,k}^{\text{PT},(Cap)} y_{j,k}^{\text{PT}}, \end{aligned} \tag{16}$$

where $C_i^{\text{S},(Cap)}$, $C_j^{\text{P},(Cap)}$, $C_k^{\text{T},(Cap)}$ are the investment costs of source i, pool j and product terminal k, respectively, and $C_{i,j}^{\text{SP},(Cap)}$, $C_{i,k}^{\text{ST},(Cap)}$, $C_{j,j^-}^{\text{PP},(Cap)}$, $C_{j,k}^{\text{PT},(Cap)}$ are the investment costs of the trunklines connecting different units in the network. Eq. (16) implies that all the capital investment are incurred at the same time. If it is not the case, the time value of the investment cost needs to be considered in the evaluation of the net present value. $C_h^{(OC)}$ denotes the annual net income from operating the production system, which depends on the scenario h with probability p_h, and

$$\begin{aligned} C_h^{(OC)} = & \sum_{i\in\{1,\dots,n\}} -C_i^{\text{S},(OC)} \left(\sum_{j\in\Theta_i^{\text{SP}}} f_{i,j,h}^{\text{SP}} + \sum_{k\in\Theta_i^{\text{ST}}} f_{i,k,h}^{\text{ST}} \right) \\ & + \sum_{k\in\{1,\dots,m\}} C_{k,h}^{\text{T},(OC)} \left(\sum_{j\in\Pi_k^{\text{PT}}} \sum_{w\in\{1,\dots,l\}} f_{j,k,w,h}^{\text{PT}} + \sum_{i\in\Pi_k^{\text{ST}}} f_{i,k,h}^{\text{ST}} \right), \end{aligned} \tag{17}$$

where $C_i^{\text{S},(OC)}$ denotes the annual cost related to the operation of source i per unit of gas produced, $C_{k,h}^{\text{T},(OC)}$ denotes the annual revenue related to the operation of product terminal k in scenario h per unit of gas produced. Notice that the costs incurred by

Table 2 Additional Symbols for Consideration of Pressure

Symbol	Type	Description
P	Variable	Pressure at a pool without compression or at a pool after compression
P_b	Variable	Bottom hole pressure
P_c	Variable	Pressure at a pool containing compressors (before compression)
P_t	Variable	Flowing tubing head pressure
W	Variable	Compression power
α	Parameter	Coefficient in well performance model
β	Parameter	Coefficient in well performance model
θ	Parameter	Coefficient in well performance model
κ	Parameter	Coefficient in trunkline pressure-flow relationship
λ	Parameter	Exponent in well performance model
ν	Parameter	Coefficient in compression model
π_r	Parameter	Reservoir pressure in the vicinity of a well
ω	Parameter	Coefficient in compression model
Δ	Set	Index set for pools containing compressors

operating the trunklines and pools are not considered here, but they can be easily incorporated into the objective function as necessary.

As a summary of the above discussion, when maximizing the expected net present value, the stochastic pooling problem to be solved is

obj: Eq. (15)
s.t. Eqs. (16-17) and (1-14),

where obj stands for the objective of the optimization problem. Note the this formulation can be enhanced with extra (but redundant) reformulation-linearization constraints, for efficient global optimization [61]. These constraints are provided in the online supplementary material.

2.3 Consideration of Pressure

If the effects of pressure need to be considered for planning natural gas production infrastructure development, then additional submodels need to be incorporated into the stochastic pooling model to reflect the relationship between the pressures in the system and the gas flows and energy consumption. These submodels were originally developed in [55] [54] for a short-term operational model, and they are tailored here in the context of the stochastic programming model for infrastructure development. Table 2 summarizes the additional symbols introduced in these submodels.

Well performance model - Reservoir pressure can vary at different wells even for the wells in the same field. These wells need to be addressed separately if pressure needs to be considered, so each of them can be deemed as a individual source in a generalized pooling system. The reservoir pressure in the vicinity of a well (i.e., a source) i, π_r, is assumed to be constant over the lifespan of the system for simplicity. And a reservoir pressure model that describes the change of the reservoir pressure over the lifespan of the system, if available, can be easily incorporated into the problem. There are two other pressures associated to each producing well i. One is the bottom hole pressure $P_{b,i}$, which is the pressure at the bottom of the well bore. The other is the flowing tubing head pressure $P_{t,i}$, which is the pressure at the well head.

These pressures are subject to the following equations:

$$f^{S}_{i,h} = \sum_{j\in\Theta_i^{SP}} f^{SP}_{i,j,h} + \sum_{k\in\Theta_i^{ST}} f^{ST}_{i,k,h}, \tag{18}$$

$$\alpha_{i,h} f^{S}_{i,h} + \beta_{i,h}(f^{S}_{i,h})^2 = \pi^2_{r,i,h} - P^2_{b,i,h}, \tag{19}$$

$$\theta_{i,h}(f^{S}_{i,h})^2 = P^2_{b,i,h} - \lambda_{i,h} P^2_{t,i,h}, \tag{20}$$

$$P_{b,i,h}, P_{t,i,h} \geq 0, \tag{21}$$

$$\forall i \in \{1,...,n\}, \quad \forall h \in \{1,...,s\}.$$

Again, the subscript h indexes different scenarios. Eq. (18) relates the total outlet flow to each outlet stream for well i and scenario h. Eqs. (19) and (20) relate the pressures and the total outlet flow for well i and scenario h. Eq. (21) enforces non-negativity for the pressures.

Implicit choke assumption - This assumption on the well head is that the pressure at the common header of a pool must be less than the flowing tubing head pressure of all wells connecting to the pool. In reality, this is achieved by a choke valve at each wellhead, but an explicit model of the choke valve is not considered here. This assumption can be modelled using the following inequalities:

$$P_{c,j,h} - P_{t,i,h} \leq 0, \quad \forall j \in \Delta,\ i \in \Omega_j^{SP}, \tag{22}$$

$$P_{j,h} - P_{t,i,h} \leq 0, \quad \forall j \in \{1,...,r\} \setminus \Delta,\ i \in \Omega_j^{SP}, \tag{23}$$

$$P_{c,i,h}, P_{i,h} \geq 0, \tag{24}$$

$$\forall h \in \{1,...,s\}.$$

Here Eq. (22) is for pools with compression, where the common header is the pressure before the compression and it is denoted by $P_{c,j,h}$ for pool j and scenario h. Index set $\Delta \subset \{1,...,r\}$ contains indices of the pools containing compressors. Eq. (23) is for pools without compression, where the common header is the pressure at the pool and it is presented by $P_{j,h}$ for pool j and scenario h. Eq. (24) enforces non-negativity for the pressures.

Compression model - It is assumed that the compression equation is given by the polytropic work of compression. The relationship between the compression power and the inlet and the outlet pressures of the pool is as follows:

$$W_{j,h} = \omega_j f^{P}_{j,h}\left[\left(\frac{P_{j,h}}{P_{c,j,h}}\right)^{\nu} - 1\right], \tag{25}$$

$$f^{P}_{j,h} = \sum_{i\in\Omega_j^{SP}} f^{SP}_{i,j,h} + \sum_{w\in\{1,...,l\}} \sum_{j^+\in\Omega_j^{PP+}} f^{PP}_{j^+,j,w,h}, \tag{26}$$

$$\forall j \in \Delta, \quad \forall h \in \{1,...,s\},$$

where $W_{j,h}$ denotes the compression power required for pool j and scenario h, $f^{P}_{j,h}$ denotes the total flow coming through pool j in scenario h, ω_j and ν are compression parameters that can be calculated as explained in [55]. The required compression power may be purchased from some energy suppliers or generated by burning gas produced from the gas production system. When the energy cost is explicitly considered as part of the operating cost, Eq. (17) becomes,

$$
\begin{aligned}
C_h^{(OC)} = & \sum_{i\in\{1,\dots,n\}} -C_i^{\mathrm{S},(OC)} \left(\sum_{j\in\Theta_i^{\mathrm{SP}}} f_{i,j,h}^{\mathrm{SP}} + \sum_{k\in\Theta_i^{\mathrm{ST}}} f_{i,k,h}^{\mathrm{ST}} \right) \\
& + \sum_{k\in\{1,\dots,m\}} C_{k,h}^{\mathrm{T},(OC)} \left(\sum_{j\in\Pi_k^{\mathrm{PT}}} \sum_{w\in\{1,\dots,l\}} f_{j,k,w,h}^{\mathrm{PT}} + \sum_{i\in\Pi_k^{\mathrm{ST}}} f_{i,k,h}^{\mathrm{ST}} \right) - \sum_{j\in\Delta} C_{j,h}^{\mathrm{W}} W_{j,h} \,,
\end{aligned} \tag{27}
$$

where $C_{j,h}^{\mathrm{W}}$ denotes the energy cost .

Trunkline pressure-flow relationship - The relationship between pressures at the units connected by a trunkline and the flow rate in the trunkline is described based on a standard gas flow equation [55] here. This equation is tailored for the stochastic programming model as follows:

$$
P_{i,h}^2 - P_{c,j,h}^2 = \kappa_{i,j,h} \left(f_{i,j,h}^{\mathrm{SP}} \right)^2, \quad \forall j \in \Delta,\ i \in \Omega_j^{\mathrm{SP}}, \tag{28}
$$

$$
P_{i,h}^2 - P_{j,h}^2 = \kappa_{i,j,h} \left(f_{i,j,h}^{\mathrm{SP}} \right)^2, \quad \forall j \in \{1,\dots,r\} \setminus \Delta,\ i \in \Omega_j^{\mathrm{SP}}, \tag{29}
$$

$$
P_{i,h}^2 - P_{k,h}^2 = \kappa_{i,k,h} \left(f_{i,k,h}^{\mathrm{ST}} \right)^2, \quad \forall i \in \{1,\dots,n\},\ k \in \Theta_i^{\mathrm{ST}}, \tag{30}
$$

$$
P_{j^+,h}^2 - P_{c,j,h}^2 = \kappa_{j^+,j,h} \left(\sum_{w\in\{1,\dots,l\}} f_{j^+,j,w,h}^{\mathrm{PP}} \right)^2, \quad \forall j \in \Delta,\ j^+ \in \Omega_j^{\mathrm{PP+}}, \tag{31}
$$

$$
P_{j^+,h}^2 - P_{j,h}^2 = \kappa_{j^+,j,h} \left(\sum_{w\in\{1,\dots,l\}} f_{j^+,j,w,h}^{\mathrm{PP}} \right)^2, \quad \forall j \in \{1,\dots,r\} \setminus \Delta,\ j^+ \in \Omega_j^{\mathrm{PP+}}, \tag{32}
$$

$$
P_{j,h}^2 - P_{k,h}^2 = \kappa_{j,k,h} \left(\sum_{w\in\{1,\dots,l\}} f_{j,k,w,h}^{\mathrm{PT}} \right)^2, \quad \forall j \in \{1,\dots,r\},\ k \in \Omega_j^{\mathrm{PT}}. \tag{33}
$$

Eq. (28) is for any trunkline connecting a source and a pool with compression, and the pool pressure used is the one before the compression, while Eq. 28 is for any trunkline connecting a source and a pool without compression. Eq. (30) is for any trunkline connecting a source and a product terminal. Eq. (31) or (32) is for any trunkline connecting a pool to a pool with or without compression. Eq.(33) is for any trunkline connecting a pool and a product terminal.

As a summary, with the consideration of pressure, the stochastic programming problem for natural gas production infrastructure development is:

obj: Eq. (15)
s.t. Eqs. (1-14), (16) and (18-33).

3 Decomposition-Based Global Optimization

This section describes the NGBD method, which can solve the proposed stochastic programming models to global optimality via solving a sequence of subproblems whose sizes are independent of the number of scenarios. The discussion is based on a more general MINLP formulation as follows:

$$\begin{aligned}
\min_{x_1,\dots,x_s,y} \quad & \sum_{h=1}^{s} w_h \left(c_h^{\mathrm{T}} y + f_h(x_h)\right) \\
\text{s.t.} \quad & g_h(x_h) + B_h y \le 0, \quad \forall h \in \{1,\dots,s\}, \\
& x_h \in X_h, \quad \forall h \in \{1,\dots,s\}, \\
& y \in Y,
\end{aligned} \tag{P}$$

where $X_h = \{x_h \in \Pi_h \subset \mathbb{R}^{n_x} : p_h(x_h) \le 0\}$, $Y = \{y \in \{0,1\}^{n_y} : Ay \le d\}$, Π_h is convex, and functions $f_h : \Pi_h \to \mathbb{R}$, $g_h : \Pi_h \to \mathbb{R}^m$ and $p_h : \Pi_h \to \mathbb{R}^{m_p}$ are continuous. Assume at least one function in Problem (P) is nonconvex. In the stochastic pooling problem, the nonconvex functions are the bilinear functions. The consideration of pressure in the model results in additional quadratic equations and power functions that also render nonconvexity. The binary variables y represent design decisions, such as whether or not to build a platform or a trunkline. The continuous variables x_h represent the operational decisions, such as gas flow rates and pressures, for each scenario h.

3.1 Decomposition strategy in NGBD

The reformulation and decomposition of Problem (P) in NGBD is described here. First, Problem (P) is relaxed via convex relaxation of the nonconvex functions into the following lower bounding problem:

$$\begin{aligned}
\min_{\substack{x_1,\dots,x_s,\\ e_1,\dots,e_s,y}} \quad & \sum_{h=1}^{s} w_h \left(c_h^{\mathrm{T}} y + u_{f,h}(x_h, e_h)\right) \\
\text{s.t.} \quad & u_{g,h}(x_h, e_h) + B_h y \le 0, \quad \forall h \in \{1,\dots,s\}, \\
& (x_h, e_h) \in D_h, \quad \forall h \in \{1,\dots,s\}, \\
& y \in Y,
\end{aligned} \tag{LBP}$$

where $D_h = \{(x_h, e_h) \in \Pi_h \times \Theta_h : u_{p,h}(x_h, e_h) \le 0, u_{e,h}(x_h, e_h) \le 0\}$, Θ_h is convex, and functions $u_{f,h} : \Pi_h \times \Theta_h \to \mathbb{R}$, $u_{g,h} : \Pi_h \times \Theta_h \to \mathbb{R}^m$, $u_{p,h} : \Pi_h \times \Theta_h \to \mathbb{R}^{m_p}$, and $u_{e,h} : \Pi_h \times \Theta_h \to \mathbb{R}^{m_e}$ are convex on $\Pi_h \times \Theta_h$. So Problem (LBP) is a convex MINLP or a MILP. This problem involves auxiliary variables e_h and constraints $u_{e,h}(x_h, e_h) \le 0$ that may be required to construct smooth relaxations. Several convex relaxation techniques are available to generate Problem (LBP), e.g., McCormick relaxation [44] and outer linearization [61] for factorable nonconvex functions, which usually introduce additional variables and constraints for differentiable relaxations, and αBB for twice-differentiable nonconvex functions [2], which does not require additional variables and constraints. Readers can refer to [19] for more discussions on the convex relaxation techniques.

If Problem (LBP) is assumed to satisfy a constraint qualification (which implies strong duality) for any fixed $y \in Y$ for which Problem (LBP) is feasible, then it can be equivalently transformed into a master problem via the principles of projection and dualization [20] [41]. The number of variables in the master problem is independent of the number of scenarios, but it contains an infinite number of constraints and is usually difficult to solve directly. Instead, it is solved via solving a sequence of **Primal Bounding Problems (PBP)**, **Feasibility Problems (FP)** and **Relaxed Master Problems (RMP)**, which are much easier to solve. The primal bounding problem is constructed at each iteration k by restricting the binary variables to specific values, say $y = y^{(k)}$, in (LBP), whose solution yields a valid upper bound for the lower bounding problem (and hence the master problem). Furthermore, it can be naturally decomposed into s subproblems of the following form:

$$\begin{aligned} obj_{\mathrm{PBP}_h}(y^{(k)}) = \min_{x_h, e_h} & \; w_h \left(c_h^{\mathrm{T}} y^{(k)} + u_{f,h}(x_h, e_h) \right) \\ \text{s.t.} & \; u_{g,h}(x_h, e_h) + B_h y^{(k)} \le 0, \\ & (x_h, e_h) \in D_h, \end{aligned} \tag{PBP$_h^k$}$$

where $obj_{\mathrm{PBP}_h}(y^{(k)})$ denotes the optimal objective value of Problem (PBP$_h^k$), $h = 1, ..., s$. Obviously, $\sum_{h=1}^{s} obj_{\mathrm{PBP}_h}(y^{(k)}) = obj_{\mathrm{PBP}}(y^{(k)})$ where $obj_{\mathrm{PBP}}(y^{(k)})$ is the optimal objective value of the primal bounding problem for $y = y^{(k)}$. When the primal bounding problem is infeasible for $y = y^{(k)}$, a corresponding feasibility problem is solved, which can be also decomposed into s subproblems of the following form:

$$\begin{aligned} obj_{\mathrm{FP}_h}(y^{(k)}) = \min_{x_h, e_h, z_h} & \; w_h ||z_h|| \\ \text{s.t.} & \; u_{g,h}(x_h, e_h) + B_h y^{(k)} \le z_h, \\ & (x_h, e_h) \in D_h, \quad z_h \in Z_h, \end{aligned} \tag{FP$_h^k$}$$

where $obj_{\mathrm{FP}_h}(y^{(k)})$ denotes the optimal objective value of Problem (FP$_h^k$), $h = 1, ..., s$. Obviously, $\sum_{h=1}^{s} obj_{\mathrm{FP}_h}(y^{(k)}) = obj_{\mathrm{FP}}(y^{(k)})$ where $obj_{\mathrm{FP}}(y^{(k)})$ is the optimal objective value of the feasibility problem. The relaxed master problem is constructed at each iteration k by relaxing the master problem with a finite number of constraints that are derived according to the solution information of all the previously solved primal bounding and feasibility problems, as follows:

$$\begin{aligned} \min_{\eta, y} \quad & \eta \\ \text{s.t.} \quad & \eta \ge obj_{\mathrm{PBP}}(y^{(j)}) + \left(\sum_{h=1}^{s} \left(w_h c_h^{\mathrm{T}} + \left(\lambda_h^{(j)} \right)^{\mathrm{T}} B_h \right) \right) \left(y - y^{(j)} \right), \quad \forall j \in T^k, \\ & 0 \ge obj_{\mathrm{FP}}(y^{(i)}) + \left(\sum_{h=1}^{s} \left(\mu_h^{(i)} \right)^{\mathrm{T}} B_h \right) \left(y - y^{(i)} \right), \quad \forall i \in S^k, \\ & \sum_{r \in R_1^t} y_r - \sum_{r \in R_0^t} y_r \le |R_1^t| - 1, \quad \forall t \in T^k \cup S^k, \\ & y \in Y, \; \eta \in \mathbb{R}, \end{aligned} \tag{RMP^k}$$

where the index sets

$$T^k = \{j \in \{1,...,k\} : \text{Problem (LBP) is feasible for } y = y^{(j)}\},$$
$$S^k = \{i \in \{1,...,k\} : \text{Problem (LBP) is infeasible for } y = y^{(i)}\},$$
$$R_1^t = \{r \in \{1,\cdots,n_y\} : y_r^{(t)} = 1\},\ R_0^t = \{r \in \{1,\cdots,n_y\} : y_r^{(t)} = 0\}.$$

$\lambda_h^{(j)}$ are the Lagrange multipliers for Problem (PBP$_h^j$), which form an optimality cut for iteration j ($\forall j \in T^k$). $\mu_h^{(i)}$ are the Lagrange multipliers for Problem (FP$_h^i$), which form a feasibility cut for iteration i ($\forall i \in S^k$). The last group of constraints represent a set of canonical integer cuts that prevent the previously examined integer realizations from becoming a solution [5]. The solution of the relaxed master problem yields a valid lower bound for the master problem (and therefore Problem (P)) augmented with the integer cuts. Notice that the relaxed master problem is a MILP whose size is determined by the current iteration number but is independent of the number of scenarios. In case $T^k = \emptyset$, the relaxed master problem is unbounded, so the following feasibility version of it is solved to make the algorithm proceed:

$$\begin{aligned} \min_{y} \quad & \sum_{i=1}^{n_y} y_i \\ \text{s.t.} \quad & 0 \geq obj_{\text{FP}}(y^{(i)}) + \left(\sum_{h=1}^{s} \left(\mu_h^{(i)}\right)^{\text{T}} B_h\right)\left(y - y^{(i)}\right), \quad \forall i \in S^k, \\ & \sum_{r \in R_1^t} y_r - \sum_{r \in R_0^t} y_r \leq |R_1^t| - 1, \quad \forall t \in T^k \cup S^k, \\ & y \in Y,\ \eta \in \mathbb{R}. \end{aligned} \tag{FRMP^k}$$

On the other hand, a restriction of Problem (P), which is called the **Primal Problem (PP)**, is constructed at iteration l by restricting $y = y^{(l)}$ in Problem (P), whose optimal objective value is a valid upper bound for Problem Problem (P). The primal problem can be further decomposed into s subproblems of the following form:

$$\begin{aligned} obj_{\text{PP}_h}(y^{(l)}) = \min_{x_h} & \ w_h \left(c_h^{\text{T}} y^{(l)} + f_h(x_h)\right) \\ \text{s.t.} & \ g_h(x_h) + B_h y^{(l)} \leq 0, \\ & \ x_h \in X_h, \end{aligned} \tag{PP$_h^l$}$$

where $obj_{\text{PP}_h}(y^{(l)})$ denotes the optimal objective value of Problem (PP$_h^l$), $h = 1,...,s$. Obviously, $\sum_{h=1}^{s} obj_{\text{PP}_h}(y^{(l)}) = obj_{\text{PP}}(y^{(l)})$ where $obj_{\text{PP}}(y^{(l)})$ is the optimal objective value of the primal problem. The nonconvex nonlinear programming (NLP) problem (PP$_h^l$) can be solved to ε-optimality in finite time by state-of-the-art branch-and-bound global optimization solvers, such as BARON [62], and the solution can be significantly accelerated by adding an additional cut derived from the solution of the previously solved subproblems [41].

Table 3 The NGBD Algorithm

Initialize:

1. Iteration counters $k = 0$, $l = 1$, and the index sets $T^0 = \emptyset$, $S^0 = \emptyset$, $U^0 = \emptyset$.
2. Upper bound on the problem $UBD = +\infty$, upper bound on the lower bounding problem $UBDPB = +\infty$, lower bound on the lower bounding problem $LBD = -\infty$.
3. Set tolerances ε_h and ε such that $\sum_{h=1}^{s} \varepsilon_h \leq \varepsilon$.
4. Integer realization $y^{(1)}$ is given.

repeat
  if $k = 0$ or (Problem (RMP^k) is feasible and $LBD < UBDPB$ and $LBD < UBD - \varepsilon$) **then**
    repeat
      Set $k = k+1$
      1. Solve the decomposed primal bounding subproblem (PBP$_h^k$) for each scenario $h = 1,...,s$ sequentially. If Problem (PBP$_h^k$) is feasible and has duality multipliers $\lambda_h^{(k)}$ for all the scenarios, add an optimality cut to the relaxed master problem (RMP^k) according to the multipliers $\lambda_1^{(k)},...,\lambda_s^{(k)}$, set $T^k = T^{k-1} \cup \{k\}$. If $obj_{\mathrm{PBP}}(y^{(k)}) = \sum_{h=1}^{s} obj_{\mathrm{PBP}_h}(y^{(k)}) < UBDPB$, update $UBDPB = obj_{\mathrm{PBP}}(y^{(k)})$, $y^* = y^{(k)}$, $k^* = k$.
      2. If Problem (PBP$_h^k$) is infeasible for scenario $\hat{h}$, stop solving it for scenarios $\hat{h}+1,...,s$ and set $S^k = S^{k-1} \cup \{k\}$. Then, set $\mu_h^{(k)} = 0$ for $h = 1,...,\hat{h}-1$, and solve the decomposed feasibility subproblem (FP$_h^k$) for $h = \hat{h},...,s$ and obtain the corresponding Lagrange multipliers $\mu_h^{(k)}$. Add a feasibility cut to Problem (RMP^k) according to $\mu_1^{(k)},...,\mu_s^{(k)}$.
      3. If $T^k = \emptyset$, solve the feasibility relaxed master problem (FRMP^k); otherwise, solve Problem (RMP^k). In the latter case, set LBD to the optimal objective value of Problem (RMP^k) if Problem (RMP^k) is feasible. In both cases, set $y^{(k+1)}$ to the y value at the solution of either problem.
    until $LBD \geq UBDPB$ or (Problem (RMP^k) or Problem (FRMP^k) is infeasible).
  end if
  if $UBDPB < UBD - \varepsilon$ **then**
    1. Solve the decomposed primal subproblem (PP$_h^*$) (i.e., for $y = y^*$) to ε_h-optimality for each scenario $h = 1,...,s$ sequentially. Set $U^l = U^{l-1} \cup \{k^*\}$. If Problem (PP$_h^*$) has optimum x_h^* for all the scenarios and $obj_{\mathrm{PP}}(y^*) = \sum_{h=1}^{s} obj_{\mathrm{PP}_h}(y^*) < UBD$, update $UBD = obj_{\mathrm{PP}}(y^*)$ and set $y_p^* = y^*$, $x_{p,h}^* = x_h^*$ for $h = 1,...,s$.
    2. If $T^k \setminus U^l = \emptyset$, set $UBDPB = +\infty$.
    3. If $T^k \setminus U^l \neq \emptyset$, pick $i \in T^k \setminus U^l$ such that $obj_{\mathrm{PBP}}(y^{(i)}) = \min_{j \in T^k \setminus U^l} \{obj_{\mathrm{PBP}}(y^{(j)})\}$. Update $UBDPB = obj_{\mathrm{PBP}}(y^{(i)})$, $y^* = y^{(i)}$, $k^* = i$. Set $l = l+1$.
  end if
until $UBDPB \geq UBD - \varepsilon$ and (Problem (RMP^k) or Problem (FRMP^k) is infeasible or $LBD \geq UBD - \varepsilon$).
An ε-optimal solution of Problem (P) is given by $(y_p^*, x_{p,1}^*, ..., x_{p,s}^*)$ or Problem (P) is infeasible.

3.2 NGBD Algorithm

The pseudocode of the NGBD algorithm is given in Table 3 and its finite convergence is stated in the following Theorem 1.

Theorem 1 *If all the subproblems can be solved to ε-optimality in a finite number of steps, then the NGBD algorithm terminates in a finite number of steps with an ε-optimal solution of Problem (P) or an indication that Problem (P) is infeasible.*

Proof Detailed proof can be found in [41]

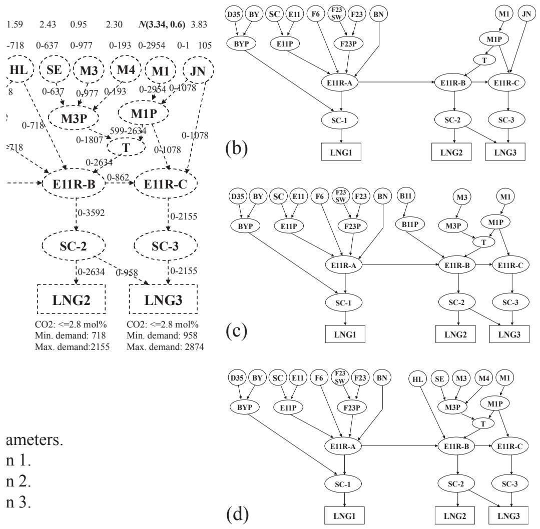


Fig. 2 SGPS Problem 1 and the different design results.

4 Case Studies

4.1 SGPS Problem 1

4.1.1 Problem Description

This case study is inspired by the Sarawak Gas Production System (SGPS) [55]. The supper structure of the system is given in Figure 2(a). The units and the connecting trunklines with solid lines in the figure represent the existing part of the system, which has eight gas fields (D35, BY, SC, E11, F6, F23SW, F23 and BN) as sources, four platforms (BYP, E11P, F23P and E11R-A) and one plant slug-catcher (SC-1) as pools, and one LNG plant (LNG1) as product terminal. Due to expansion of the market, more gas fields, platforms, trunklines need to be developed to feed gas to two potential LNG plants. The potential units and the connecting trunklines of the new part of the system are shown in dashed lines in the figure, including seven gas fields (B11, HL, SE, M3, M4, M1, JN) as sources, five platforms (B11P, M3P, M1P, E11R-B, E11R-C), and one trunkline connection (T) and two plant slug-catchers (SC-2, SC-3) as pools, and two LNG plants (LNG2, LNG3) as product terminals. The gas platform B11P is designated to locate at the gas field B11, which should at least serve gas from B11. This means B11 must be developed if B11P is developed and vice versa. The same relationship exists between M3 and M3P, M1 and M1P, SC-2 and LNG2 and SC-3 and LNG3 (where SC-2 or SC-3 is part of plant LNG2 or LNG3) as well. Such relationships are enforced by additional topology constraints to the model. The goal of optimization is to maximize the expected NPV of the system over the next 25 years. Other information on the optimization model, including how the model is implemented and solved, is provided in the online supplementary material.

It is assumed that any desired flow rate in a particular trunkline (within the trunkline capacity) can be achieved by adding a compressor at an upstream platform, so it suffices to use the stochastic pooling model to plan the infrastructure development. The following three formulations are compared in the case study:

Formulation 1 - This is a deterministic formulation that considers only the expected values of uncertain parameters, and it does not have any quality constraints.

Formulation 2 - This is a deterministic formulation that considers only the expected values of uncertain parameters, but it has quality constraints on the final products at the product terminals.

Formulation 3 - The stochastic pooling problem formulation.

4.1.2 Results and Discussion

First, the design results from Formulations 1, 2 and 3 are compared for the situation where the uncertain parameter is the CO_2 mole percentage of gas from M1. The uncertain parameter obeys a normal distribution with a mean of 3.34 mol% and a standard deviation of 0.6 mol%. Five scenarios of the uncertain M1 quality are selected according to the sampling rule described the online supplementary material. Values of non-economic parameters for this problem are shown in Figure 2 (a).

Figs. 2 (b)-(d) show the system designs using the three formulations. Since Formulation 1 does not consider the quality constraints at the LNG plants, the new part of the system designed with this formulation contains gas fields M1 and JN for the lowest investment cost. However, this design is infeasible for some scenarios considering the quality constraints at the LNG plants, because the quality of gas field JN severely violates the quality upper bounds, and the quality of gas field M1 violates the bounds as well in some scenarios. Formulation 2 observes the quality constraints, so the new part of the system designed by this formulation has gas fields B11, M3 and M1 instead. The blending of the gases from these fields can satisfy the quality constraints at the LNG plants in the deterministic case. The drawback of this design is that the quality of M1 may be so high that M1 cannot supply as much gas for blending to final products as anticipated by the deterministic formulation; in this situation, gas field M1 will be of little use and the investment in it is not profitable. When considering the quality uncertainty explicitly in Formulation 3, the designed system is different from the one designed with Formulation 2, where gas fields HL, SE, M4 are developed instead of B11. Although these gas fields are more expensive to develop than B11, they can serve gas flows with much better qualities, so M1 can still supply a substantial amount of gas for blending final products when its quality is much worse than the mean quality value.

The advantage of the stochastic pooling problem formulations over the two deterministic formulations can be further recognized with Table 4, which summarizes the design and operation results of SGPS Problem 1 with different formulations. For each formulation, the total capital cost is calculated according to each designed system and shown in the table; the net present value of each scenario is calculated and the average over the five scenarios is shown in the table. Since the operating cost is not included in all the formulations, all the net present values shown in Table 4

Table 4 Design and Operation Results With Different Formulations for SGPS Problem 1

	Average Net Present Value (Billion $)	Satisfaction of Product Quality[a] Bound at Each LNG Plant for the Five Scenarios			Capital Cost (Billion $)
		LNG 1	LNG 2	LNG 3	
Formulation 1	33.1[b]	Y/Y/Y/Y/N[c]	Y/Y/N/N/N	Y/N/N/N/N	20.8
Formulation 2	29.0	Y/Y/Y/Y/Y	Y/Y/Y/Y/Y	Y/Y/Y/Y/Y	21.1
Formulation 3	32.2	Y/Y/Y/Y/Y	Y/Y/Y/Y/Y	Y/Y/Y/Y/Y	21.6

[a]Product quality means the percentage of CO_2 in the product.

[b]This net present value cannot be achieved in reality because of the violation of the quality bound.

[c]'Y' or 'N' indicates whether the quality upper bound is satisfied or not for each of the five scenarios.

Table 5 Computational Results for SGPS Problem 1 With the Stochastic Pooling Model (Time Unit: Second)

Number of Scenarios	1	16	81	256	625
Number of Variables	38/59[a]	38/929	38/4699	38/14849	38/36251
Time With BARON 11.3	1.3	766.5	–[b]	–	–
Time With NGBD [c]	3.4	14.8	60.9	320.5	617.4

[a]Number of binary variables/number of continuous variables.

[b]No solution is returned after 100,000 seconds.

[c]This is the total time for solving all the subproblems in the local solvers.

will be higher than the real ones. Table 4 also shows whether the product quality upper bound is satisfied at the LNG plants for each scenario. It can be found that with Formulation 1, the product quality at either of the two new LNG plants violates the bound for four of the five scenarios, so the net present value calculated by this formulation is meaningless for the real problem. Formulations 2 and 3 observe the product quality constraints, but Formulation 3 achieves better average net present values than Formulation 2 (with a large improvement of $3.2 billion).

Second, the computational efficiencies of BARON and the NGBD algorithm are compared via solving the stochastic pooling problem (Formulation 3) with different numbers of scenarios, for SGPS Problem 1 in the situation where four independent uncertain parameters are present. These uncertain parameters are the qualities of gas fields M1, JN and the maximum demands at LNG plants 2 and 3, which obey normal distributions with means 5.04 mol%, 2.63 mol%, 1736 Mmol/day and 2275 Mmol/day, and standard deviations 1 mol%, 0.4 mol%, 144 Mmol/day and 239 Mmol/day, respectively. 1, 2, 3, 4 and 5 realizations are generated for each uncertain parameter in the way described before, which lead to problems with 1, 16, 81, 256 and 625 scenarios. These five problems are solved with both BARON and NGBD, and the total solver times are displayed in Table 5. Although NGBD is slower than BARON when the number of scenarios is 1 (i.e., a deterministic formulation is solved), it is much faster when more scenarios are addressed, and the solver time with NGBD increases moderately with the number of scenarios. On the other hand, BARON cannot obtain a solution for the problem within 100,000 seconds when 3 or more realizations are addressed for each uncertain parameters (i.e., 81 or more scenarios for the problem).

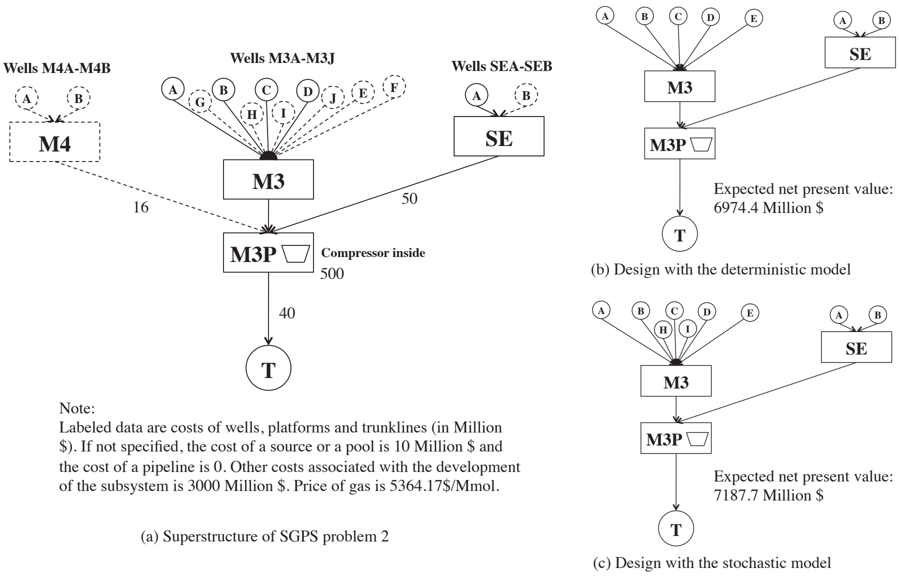


Fig. 3 SGPS Problem 2 and the different design results.

4.2 SGPS Problem 2

4.2.1 Problem Description

This case study considers the situation where the effects of pressure need to be considered, and the stochastic programming model developed in Section 2.3 applies. A subsystem of the SGPS is considered for this study,which includes gas fields SE, M3, M4, gas platform M3P, (intermediate) gas terminal T, and the relevant trunklines. Due to the consideration of pressure, the gas wells in the gas fields are considered explicitly. Fig. 3 (a) shows the superstructure of the subsystem, where the solid lines indicate the part of the subsystem that has to be developed and dashed lines indicate the part of the subsystem that may be developed for the system. Gas fields M4, M3, and SE have 2, 10 and 2 gas wells, respectively, and gas platform M3P has a compressor installed. Fig. 3 (a) also shows some economic information for the subsystem, where other costs associated with the subsystem may be understood as the investments for the facilities indispensable for the raw gas to be delivered to end customers (e.g., transmission and processing facilities). The goal of the optimization is again to maximize the expected NPV over the next 25 years. Other information on the optimization model, including how the model is implemented and solved, is provided in the online supplementary material.

Two formulations are compared for SGPS Problem 2. One is the deterministic formulation that only considers the expected value realizations of the uncertain parameters, and the other is the stochastic programming formulation developed in Section 2.3, with the scenarios generated via the sampling rule described in the online supplementary material.

Table 6 Computational Results for SGPS Problem 2 With the Stochastic Programming Model (Time Unit: Second)

Number of Scenarios	1	27	125	343	729	1331
Number of variables [a]	20/110	20/2970	20/13750	20/37730	20/80190	20/146410
Time with BARON 11.3	1.1	447.0	31841.1	- [b]	-	-
Time with NGBD [c]	0.3	30.6	167.6	431.4	951.9	1606.5

[a] Number of binary variables/Number of continuous variables.

[b] No solution returned within 100,000 seconds.

[c] This is the total time for solving all the subproblems in the local solvers.

4.2.2 Results and Discussion

Figs. 3 (b) and (c) show the design results using the deterministic model and the stochastic model, respectively. It is clear that more gas wells are to be developed with the stochastic model, which are used to hedge the risks that may come from the uncertain factors. As a result, the subsystem designed with the stochastic model achieves better expected net present value (with a improvement of more than 200 Million dollars).

Table 6 compares the computational efficiencies of BARON and NGBD, for SGPS Problem 2 with different numbers of scenarios. 1, 3, 5, 7, 9, 11 realizations are generated for each of the three uncertain parameters, which leads to 1, 27, 125, 343, 729, 1331 scenarios in the stochastic programming model. It can be found that even for the case with 1 scenario (i.e., essentially a deterministic formulation is used), NGBD is much faster than BARON. With more scenarios addressed, BARON cannot return a global solution within 10,000 seconds, while NGBD can solve the problems in reasonable time. In addition, the solution time with NGBD increases moderately with the number of scenarios. Notice that when 1331 scenarios are addressed, the problem contains nearly 150,000 variables, but NGBD can obtain a global optimum of this nonconvex MINLP within only 80 minutes.

5 Concluding Remarks

Two stochastic programming models are developed for natural gas production infrastructure development under uncertainty. One is the stochastic pooling model that treats the production system as a generalized pooling system in order to track the qualities of the gas streams and observe their bounds. The other enhances the stochastic pooling model by including additional submodels to describe the effects of pressure on the system. Due to their explicit consideration of uncertainties, both models show advantages over deterministic optimization models in the case studies.

On the other hand, each stochastic programming model results in a nonconvex MINLP whose sizes depend on the number of scenarios addressed, which is very difficult to solve especially when the number of scenarios is large. The proposed NGBD method, however, can solve this problem to global optimality via solving a finite number of subproblems whose sizes are much smaller and independent of the number of scenarios. As a result, NGBD is advantageous over general-purpose branch-and-bound type of global optimization methods, such as BARON as shown by the case study results. In addition, the efficiency of NGBD can be further improved by parallel computation, as most of the subproblems solved in a NGBD iteration do not rely on the solution results of other subproblems.

As the scenarios in the stochastic models are generated via a naive approach in this research, advanced scenario generation techniques may be introduced to favor a better model for stochastic programming, such as scenario decomposition [27], scenario reduction [26], and sample average approximation [37].

While the second-stage variables in Problem (P) are continuous, the proposed decomposition strategy can be readily extended to solve problems involving integer second-stage variables. A variant of NGBD has been developed and successfully applied to solve some capacity planning problems that include integer second-stage variables [59] [40].

The proposed two-stage stochastic programming model assumes that the operation of the production system starts after the infrastructure has been completely developed. If the infrastructure is developed in several stages and parts of the system are to be operated before the last stage, the problem needs to be modeled as a multi-stage stochastic programming problem (such as in [21]). Generally speaking, the current NGBD method cannot be used to solve the resulting multi-stage problem directly, and it remains an interesting topic for future work.

References

1. Adjiman, C.S., Androulakis, I.P., Floudas, C.A.: Global optimization of mixed-integer nonlinear problems. AIChE Journal **46**(9), 1769–1797 (2000)
2. Adjiman, C.S., Dallwig, S., Floudas, C.A., Neumaier, A.: A global optimization method, α-BB, for general twice-differentiable constrained NLPs – I. Theoretical advances. Computers and Chemical Engineering **22**(9), 1137–1158 (1998)
3. ARKI Consulting and Development: http://www.gams.com/docs/conopt3.pdf
4. Aronofsky, J.S., Williams, A.C.: The use of linear programming and mathematical models in underground oil production. Management Science **8**, 394–407 (1962)
5. Balas, E., Jeroslow, R.: Canonical cuts on the unit hypercube. SIAM Journal on Applied Mathematics **23**(1), 61–69 (1972)
6. Ben-Tal, A., Eiger, G., Gershovitz, V.: Global minimization by reducing the duality gap. Mathematical Programming **63**, 193–212 (1994)
7. Benders, J.F.: Partitioning procedures for solving mixed-variables programming problems. Numerische Mathematik **4**, 238–252 (1962)
8. Bertsekas, D.P.: Nonlinear Programming, 2nd edn. Athena Scientific, Cambridge, MA (1999)
9. Birge, J.R., Louveaux, F.: Introduction to Stochastic Programming. Springer, New York (1997)
10. Bohannon, J.: A linear programming model for optimum development of multi-reservoir pipeline systems. Journal of Petroleum Technology **22**, 1429–1436 (1970)
11. Duran, M., Grossmann, I.E.: An outer-approximation algorithm for a class of mixed nonlinear programs. Mathematical Programming **66**, 327–349 (1986)
12. Energy Information Administration: Energy information administration webpage. http://www.eia.doe.gov
13. Flanigan, O.: Constrained derivatives in natural gas pipeline system optimization. Journal of Petroleum Technology **24**, 549–556 (1972)
14. Fletcher, R., Leyffer, S.: Solving mixed integer nonlinear programs by outer approximation. Mathematical Programming **66**, 327–349 (1994)

15. Floudas, C.A., Aggarwal, A.: A decomposition strategy for optimum search in the pooling problem. ORSA Journal on Computing **2**, 225–235 (1990)
16. Floudas, C.A., Visweswaran, V.: A global optimization algorithm (GOP) for certain classes of nonconvex NLPs - I. Theory. Computers and Chemical Engineering **14**(12), 1397–1417 (1990)
17. Foulds, L.R., Haugland, D., Jornsten, K.: A bilinear approach to the pooling problem. Optimization **24**, 165–180 (1992)
18. GAMS: General Algebraic and Modeling System. Available at http://www.gams.com/
19. Gatzke, E.P., Tolsma, J.E., Barton, P.I.: Construction of convex relaxations using automated code generation technique. Optimization and Engineering **3**, 305–326 (2002)
20. Geoffrion, A.M.: Generalized Benders decomposition. Journal of Optimization Theory and Applications **10**(4), 237–260 (1972)
21. Goel, V., Grossmann, I.E.: A stochastic programming approach to planning of offshore gas field developments under uncertainty in reserves. Computers and Chemical Engineering **28**, 1409–1429 (2004)
22. Goel, V., Grossmann, I.E., El-Bakry, A.S., Mulkay, E.L.: A novel branch and bound algorithm for optimal development of gas fields under uncertainty in reserves. Computers and Chemical Engineering **30**, 1076–1092 (2006)
23. Grossmann, I.E.: Review of nonlinear mixed-integer and disjunctive programming techniques. Optimization and Engineering **3**, 227–252 (2002)
24. Haverly, C.A.: Studies of the behaviour of recursion for the pooling problem. ACM SIGMAP Bulletin **25**, 29–32 (1978)
25. Haverly, C.A.: Behaviour of recursion model - more studies. ACM SIGMAP Bulletin **26**, 22–28 (1979)
26. Heitsch, H., Römisch, W.: Scenario reduction algorithms in stochastic programming. Computational Optimization and Applications **24**, 187–206 (2003)
27. Higle, J.L., Sen, S.: Stochastic decomposition: An algorithm for two-stage linear programs with recourse. Mathematics of Operations Research **16**, 650–669 (1991)
28. Horst, R., Tuy, H.: Global Optimization: Deterministic Approaches. Springer-Verlag, Berlin, Germany (1996)
29. IBM: IBM ILOG CPLEX: High-performance mathematical programming engine. http://www-01.ibm.com/software/integration/optimization/cplex/
30. International Energy Agency: World Energy Outlook 2011 (2011)
31. Iyer, R.R., Grossmann, I.E.: Optimal planning and scheduling of offshore oil field infrastructure investment and operations. Industrial and Engineering Chemistry Research **37**, 1380–1397 (1998)
32. Jonsbraten, T.W.: Optimization models for petroleum field exploitation. Ph.D. thesis, Stavanger College (1998)
33. Karuppiah, R., Grossmann, I.E.: Global optimization for the synthesis of integrated water systems in chemical processes. Computers and Chemical Engineering **30**, 650–673 (2006)
34. 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 **41**, 163–186 (2008)
35. Kesavan, P., Allgor, R.J., Gatzke, E.P., Barton, P.I.: Outer approximation algorithms for separable nonconvex mixed-integer nonlinear programs. Mathematical Programming, Series A **100**, 517–535 (2004)
36. Khajavirad, A., Michalek, J.J.: A deterministic Lagrangian-based global optimization approach for quasiseparable nonconvex mixed-integer nonlinear programs. Journal of Mechanical Design **131**(5), 051,009–1–051,009–8 (2009)
37. Kleywegt, A.J., Shapiro, A., Homem-De-Mello, T.: The sample average approximation method for stochastic discrete optimization. SIAM Journal on Optimization **12**, 479–502 (2001)
38. Lasdon, L.S., Waren, A.D., Sarkar, S., Palacios, F.: Solving the pooling problem using generalized reduced gradient and successive linear programming algorithms. ACM SIGMAP Bulletin **27**, 9–15 (1979)
39. Li, X., Armagan, E., Tomasgard, A., Barton, P.I.: Stochastic pooling problem for natural gas production network design and operation under uncertainty. AIChE Journal **57**, 2120–2135 (2011)
40. Li, X., Sundaramoorthy, A., Barton, P.I.: Nonconvex generallized Benders decomposition. In: T.M. Rassias, C.A. Floudas, S. Butenko (eds.) Optimization in Science and Engineering, pp. 307–331 (2014)
41. Li, X., Tomasgard, A., Barton, P.I.: Nonconvex generalized Benders decomposition for stochastic separable mixed-integer nonlinear programs. Journal of Optimization Theory and Applications **151**, 425–454 (2011)

42. Li, X., Tomasgard, A., Barton, P.I.: Decomposition strategy for the stochastic pooling problem. Journal of Global Optimization **54**, 765–790 (2012)
43. Lin, X., Floudas, C.A.: A novel continuous-time modeling and optimization framework for well platform planning problems. Optimization and Engineering **4**, 65–95 (2003)
44. McCormick, G.P.: Computability of global solutions to factorable nonconvex programs: Part I - Convex underestimating problems. Mathematical Programming **10**, 147–175 (1976)
45. McFarland, J.W., Lasdon, L., Loose, V.: Development planning and management of petroleum reservoirs using tank models and nonlinear programming. Operations Research **32**, 270–289 (1984)
46. Meyer, C.A., Floudas, C.A.: Global optimization of a combinatorially complex generalized pooling problem. AIChE Journal **52**(3), 1027–1037 (2006)
47. Misener, R., Floudas, C.A.: Advances for the pooling problem: Modeling, global optimization, and computational studies. Applied and Computational Mathematics **8**(1), 3–22 (2009)
48. Misener, R., Thompson, J.P., Floudas, C.A.: APOGEE: Global optimization of standard, generalized, and extended pooling problems via linear and logarithmic partitioning schemes. Computers and Chemical Engineering **35**, 876–892 (2011)
49. Murray III, J.E., Edgar, T.F.: Optimal scheduling of production and compression in gas fields. Journal of Petroleum Technology **30**, 109–116 (1978)
50. Nygreen, B., Christiansen, M., Haugen, K., Bjørkvoll, T., Kristiansen, Ø.: Modeling Norwegian petroleum production and transportation. Annals of Operations Research **82**, 251–267 (1998)
51. Quesada, I., Grossmann, I.E.: Global optimization of bilinear process networks with multicomponent flows. Computers and Chemical Engineering **19**, 1219–1242 (1995)
52. Rømo, F., Tomasgard, A., Hellemo, L., Fodstad, M., Eidesen, B.H., Pedersen, B.: Optimizing the Norwegian natural gas production and transport. Interfaces **39**(1), 46–56 (2009)
53. Seddon, D.: Gas Usage & Value. PennWell, Tulsa, Oklahoma (2006)
54. Selot, A.: Short-term supply chain management in upstream natural gas systems. Ph.D. thesis, Massachusetts Institute of Technology (2009)
55. 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 **54**(2), 495–515 (2008)
56. Sherali, H.D., Alameddine, A.: A new reformulation-linearization technique for bilinear programming problems. Journal of Global Optimization **2**, 379–410 (1992)
57. Slyke, R.M.V., Wets, R.: *L*-shaped linear programs with applications to optimal control and stochastic programming. SIAM Journal on Applied Mathematics **17**(4), 638–663 (1969)
58. Sullivan, J.: A computer model for planning the development of an offshore gas field. Journal of Petroleum Technology **34**, 1555–1564 (1982)
59. Sundaramoorthy, A., Li, X., Evans, J.M.B., Barton, P.I.: Capacity planning under clinical trials uncertainty in continuous pharmaceutical manufacturing, 2: Solution method. Industrial and Engineering Chemistry Research **51**, 13,703–13,711 (2012)
60. 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 and Engineering Chemistry Research **48**, 3078–3097 (2009)
61. Tawarmalani, M., Sahinidis, N.: Convexification and global optimization in continuous and mixed-integer nonlinear programming. Kluwer Academic Publishers, Dordrecht, the Netherlands (2002)
62. Tawarmalani, M., Sahinidis, N.V.: Global optimization of mixed-integer nonlinear programs: A theoretical and computational study. Mathematical Programming **99**, 563–591 (2004)
63. Tawarmalani, M., Sahinidis, N.V.: A polyhedral branch-and-cut approach to global optimization. Mathematical Programming **103**, 225–249 (2005)
64. Tomasgard, A., Rømo, F., Fodstad, M., Midthun, K.: Optimization models for the natural gas value chain. In: G. Hasle, K. Lie, E. Quak (eds.) Geometric Modelling, Numerical Simulation, and Optimization: Applied Mathematics at SINTEF. Springer, Berlin (2007)
65. van den Heever, S.A., Grossmann, I.E.: A Lagrangean decomposition heuristic for the design and planning of offshore hydrocarbon field infrastructure with complex economic objectives. Industrial and Engineering Chemistry Research **40**, 2857–2875 (2001)
66. van den Heever, S.A., Grossmann, I.E., Vasantharajan, S., Edwards, K.: Integrating complex economic objectives with the design and planning of offshore oilfield infrasturctures. Computers and Chemical Engineering **24**, 1049–1055 (2000)
67. Visweswaran, V., Floudas, C.A.: A global optimization algorithm (GOP) for certain classes of nonconvex NLPs - II. Aplications of theory and test problems. Computers and Chemical Engineering **14**(12), 1419–1434 (1990)
68. Wicaksono, D.S., Karimi, I.A.: Piecewise MILP under- and overestimators for global optimization of bilinear programs. AIChE Journal **54**(4), 991–1008 (2008)