Pipeline Network Design for Gathering Unconventional Oil and Gas Production using Mathematical Optimization

Agustín F. Montagna[a,b], Diego C. Cafaro[a,b,*], Ignacio E. Grossmann[c], Damian Burch[d], Yufen Shao[e], Xiao-Hui Wu[f], Kevin Furman[e]

[a] *Facultad de Ingeniería Química, Universidad Nacional del Litoral, Santiago del Estero 2829, 3000 Santa Fe, Argentina*
[b] *INTEC-Conicet - Güemes 3450, 3000 Santa Fe, Argentina*
[c] *Dept. of Chem. Eng., Carnegie Mellon University, 5000 Forbes Ave, Pittsburgh, PA 15213, United States*
[d] *(Previously) - ExxonMobil Global Projects Co, 22777 Springwoods Village Parkway, Spring, TX 77389, United States*
[e] *ExxonMobil Upstream Research Co, 22777 Springwoods Village Parkway, Spring, TX 77389, United States*
[f] *ExxonMobil Upstream Integrated Solutions Co, 22777 Springwoods Village Parkway, Spring, TX 77389, United States*

Corresponding Author e-mail address: dcafaro@fiq.unl.edu.ar

Abstract.
The optimal design of gathering networks for the unconventional oil and gas production is a relevant problem, particularly from the shale boom. In this work, we address an optimal pipeline network design problem in which the strategic decisions are the location and sizing of separation tank batteries together with the pipeline connections. The goal is to gather the oil, gas and water production from a large number of wellpads, being also possible to install junction nodes to merge the production in strategic points of the field. One major challenge is due to the steep production decline curves of unconventional wells, requiring continued turn-in-lines of new wells. To address this problem, we develop a complex multiperiod formulation to address the varying flows over the planning horizon. We aim to develop an optimization framework to obtain efficient solutions within reasonable computational times. To circumvent the computational burden due to the large-scale nonlinear and non-convex formulation required to model fluid dynamics in multiphase flows, we propose a solution algorithm, based on a bi-level decomposition. The results yield near optimal solutions for real-world instances, suggesting that the infrastructure planning can considerably improve the economics of unconventional projects.

Keywords.
Multi-phase gathering pipeline network, supply chain design, unconventional production, mathematical optimization

1. Introduction

The paradigm for managing upstream and midstream operations has changed radically since the shale revolution, whose impact has reached and transformed the oil and gas (O&G) industry. Shale energy has become profitable after the improvement in the efficiency and costs of the hydraulic fracturing techniques, lowering O&G prices, and promoting multi-million projects including pipeline networks, production separation and petrochemical plants. The mid and long-term projections for energy consumptions expect a large growth in natural gas production, being a relatively clean source of energy towards reducing CO_2 emissions. This increase in natural gas consumption will be driven by industrial and electrical sectors, and in this context, over 90% of O&G production in the U.S. is expected to come from unconventional fields by 2050 (U.S. Energy Information Administration, 2018)

In addition to the fact that the hydraulic fracturing and stimulation techniques are significantly more costly than conventional operations, the production profiles show sharp decline, requiring the continued drilling of new wells

to maintain the global production. On the one hand, this leads to much larger materials requirements and logistics operations. This in turn requires advanced optimization models to design integrated supply chain networks for upstream operations that can accomplish effective planning of materials and services flows (Montagna and Cafaro, 2019). On the other hand, the steady development of new wellpads to maintain the production levels points to the importance of efficiently and systematically managing "producing" wells (Rexilius, 2015), for which a large planning horizon should be considered in order to develop reasonable infrastructure investment plans.

This work aims at optimally designing pipelines networks for gathering the unconventional oil and gas production of large fields with many wellpads under development. The production has to be separated into its individual components (oil, gas and water) in tank batteries to be subsequently sent to centralized delivery points. As previously mentioned, the production profiles in unconventional fields show a significant peak followed by a steep decline, thus posing the challenge of keeping the infrastructure with reasonable utilization levels, motivated by the large capital expenditures involved. If the flow rate of any of the individual components exceeds the tank battery capacity, the wells connected to it must be choked back, which is undesirable. The number of tank batteries, their size and location should be optimally determined to comply with the well development plan while minimizing investment and operating costs. Moreover, pipeline diameters and network connections should be properly selected.

Another important aspect in the problem is related with the fluid dynamics, including pressure estimations at every node and pipelines transportation capacities, which are rather complex calculations when different phases share the same pipeline section. Likewise, all the investments are assessed under the net present cost concept, mainly due to the large capital expenditures required and the significant impact of postponing certain decisions to later periods. In conclusion, it is well-known that integrating the planning and development of unconventional fields is key in order to avoid missed production and project overruns. Therefore, the design of integrated supply chain networks is one of the most important challenges for modern O&G companies (Accenture, 2014; Stevens, 2012).

The outline of this paper is the following. First, a literature review is presented, analyzing the most important contributions to both the pipeline network design problem and the multiphase flow modelling. Moreover, the main insights about these works are noted. Second, we present the background to the problem, including technical details, statement and modeling assumptions. Next, a general MINLP (mixed integer non-linear programming) formulation is developed to tackle the problem. This is followed by an overall optimization framework including a bi-level decomposition involving NLP and MILP subproblems. We describe some strategies aimed at improving the general performance of the optimization procedure. Both, an illustrative example and a real-world case study are solved to assess the performance of the proposed formulation and the proposed solution strategy. Finally, conclusions and future work perspectives are outlined.

1.1. Literature Review

The literature addressing the planning and design of the O&G production gathering through pipeline networks is rather extensive, being a topic that has been studied for a long time using different approaches and methodologies. However, most of the literature is not targeted to address the particular conditions and assumptions required for the unconventional developments. The work of Mah and Shacham (1978) is a pioneering contribution in the area, clearly conceptualizing the elements of a network, characterizing flows, stating alternative formulations and even a very interesting analogy with electrical circuits. Moreover, different NLP solution methods are proposed in that work, with alternatives for large problems and sensitivity analysis. More recently, a group of works (Duran and

Grossmann, 1986; Guerra et al., 2019; Peng et al., 2020; Cafaro and Grossmann, 2014; Drouven and Grossmann, 2016) address the design of supply chain networks for the shale gas development, each of them with some specific features. In general terms, every work determines a network of pipelines to gather the products of the field development, selecting the pipeline diameters and lengths, compressors power and locations, wellpads location, drilling and stimulation scheduling, water management, among other features. However, the separation of the flow streams into oil, gas and water is not considered in these contributions. On the other hand, the resulting models are largely complex MINLP formulations, for which the authors propose different solution strategies, such as bi-level decomposition, piecewise linear approximations, linearization techniques, among others. This clearly shows the computational difficulty of the pipeline network design and infrastructure planning problem.

Multiproduct gathering optimization has been studied by Gupta and Grossmann (2012) who address the oil, gas and water production and transportation in offshore fields taking into account the evolution of Gas-to-Oil and Water-to-Oil ratios (GOR and WOR) over the reservoir behavior to plan the optimal infrastructure. However, the declining curves are not comparable to those of unconventional wells, and the pipelines capacities are not addressed in this work. Another closely related work is the one by Hong et al. (2019) aimed at the design of pipeline networks, particularly considering hydraulic conditions. The work specifically highlights the fact that previous studies have ignored the hydraulic aspects when designing gathering networks. The work proposes an MILP formulation to minimize the total costs taking into account technical constraints, topological profiles, pipeline diameters and wellhead pressures. The non-linear hydraulic equations are approximated through a piecewise linearization approximation. A similar approach is proposed by Zhang et al. (2017), where oil and gas gathering networks are optimized in order to minimize the development cost, accounting for specific connection structures, terrain conditions and limits in the pressure drops by considering friction losses and potential energy changes by linear simplified equations. Also Dbouk et al., (2020) optimize the design of gathering pipeline networks paying special attention to topological complexities and using a shortest path algorithm. Nevertheless, all these works are not focused on unconventional production, proposing static formulations that do not account for the continued turn-in-line of new wells over the planning horizon. More recently, Hong et al. (2020) address the multi-period development of onshore gas pipeline systems by means of a MILP formulation, although it comprises a single-phase problem without separation infrastructure.

Regarding the fluid dynamics, Liu et al. (2020) study the dynamic optimization of natural gas pipeline networks, also accounting for the uncertainties in the demand and production composition. They propose a robust optimization algorithm with rigorous thermodynamic equations, accurately computing the pressures at all the nodes of the network. Focusing on hydraulics and pressures optimization, the work by Drouven and Grossmann (2017) determines the line pressure profile within a given shale gas gathering network, and the required compression power. However, the optimal design of the network is out of the scope in the two latter contributions.

1.2. New contributions of this work

After reviewing the main literature related to the design of gathering networks for unconventional resources, it is important to highlight that this work presents an optimization framework aimed at solving a problem that has not been previously addressed. The main novelties of the work are therefore the following: (1) the handling of multiphase flows through the pipeline network to account for the oil, gas and water streams from unconventional production, which are modelled through complex empirical correlations and industry guidelines, (2) the separation infrastructure planning, accounting for tank battery capacities in a context of steep declining production from every single well, and (3) the formulation of multi-period models to address the steady development of unconventional resources, dealing with the early production peaks from new wells. An efficient

optimization framework is presented to face the complexity of the general problem, which turns to be a large-scale MINLP formulation.

2. Background

The problem is concerned with the optimal planning of the gathering infrastructure of an unconventional oil reservoir with hundreds of wells to be developed over a long-term horizon. The problem can be regarded as a supply chain design problem, one of whose main complexities lies on the multi-phase flows (oil, gas and water) produced by the wellpads. Multi-phase flows have to be gathered through a series of pipelines, and then routed to tank batteries (TB) where the streams are separated into individual components before sending them off to centralized delivery points (CDP). In their way to the tank batteries, multi-phase flows merge at junction nodes, which allow for the consolidation of the production coming from many wellpads, in order to make a more efficient use of the pipeline's transportation capacity. Figure 1 illustrates a simple scheme of the supply chain to design.
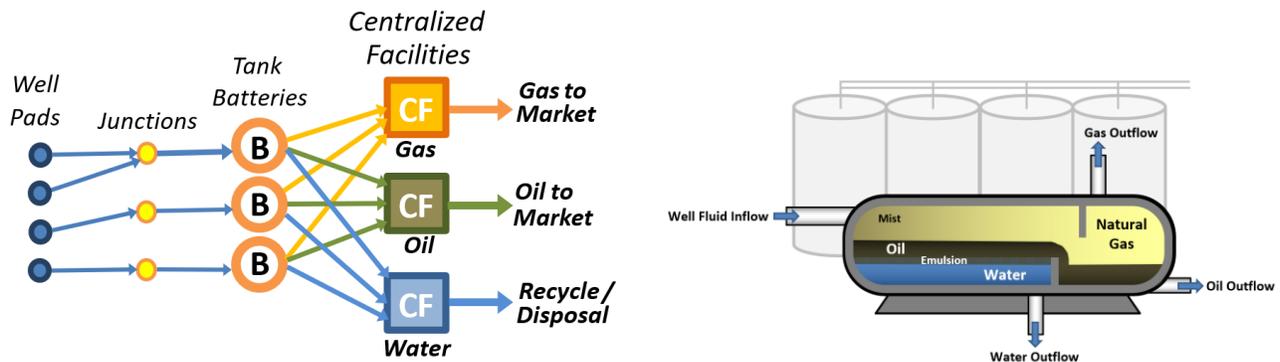
Figure 1. Left: Simple Schema of the network to be designed. Right: tank battery illustrative diagram.

Due to operational reasons, the production of a wellpad has to be routed to one and only one tank battery. This assumption follows from: (1) contract conditions with the land owners, (2) flow measuring and integrity requirements, and (3) difficulties when splitting a three-phase flow with unsteady pressure conditions (Alvarez et al., 2010; Azzopardi and Smith, 1992; Worthen and Henkes, 2015). The difficulty of the problem increases with the number of wellpads, and in practice, the straightforward solution is to increase the number of tank batteries (TB), even though they become underutilized in a few months when the wells production declines. The aim of this work is to look for a better TB utilization over the planning horizon, and to offer insights for facility planners on how to bring production coming from different wellpads into a single battery. To accomplish this goal, this work develops a rigorous mathematical formulation and a series of solution strategies.

2.1. Oil Gathering Networks

The design of gathering networks for the production of unconventional oil is rigorously described in this section through the definition of the elements that play a key role in the system. This includes transportation of three-phase flows though pipelines, separation at tank batteries, determination of pressure profiles and topological constrains.

Main Components in the Network. A complete oil gathering network comprises the elements depicted in the Figure 2, which include: (a) a series of wellpads (nodes $p \in P$) whose production start date has been previously scheduled for a given month in the current planning horizon, (b) potential locations for *junction nodes* (nodes $j \in J$) where the unconventional production coming from different wellpads converges, (c) a group of potential or existing tank battery locations ($b \in B$) whose aim, if installed, is to gather the three-phase flows from junctions to separate them into three individual components: oil, gas and water; (d) potential and/or existing pipelines of different diameters ($d \in D$) transporting oil, gas and water from wellpads p to junctions nodes j, and from these to tank batteries b; (e) potential and/or existing junction nodes ($bj \in BJ$) (usually comprising pump and/or compressor stations) where the single-phase flows converge after separation; (f) potential and/or existing single-phase pipelines of different diameters d transporting single-phase flows from tank batteries to centralized delivery points; and (g) centralized delivery points ($cdp \in CDP$) receiving the flows to be commercialized).
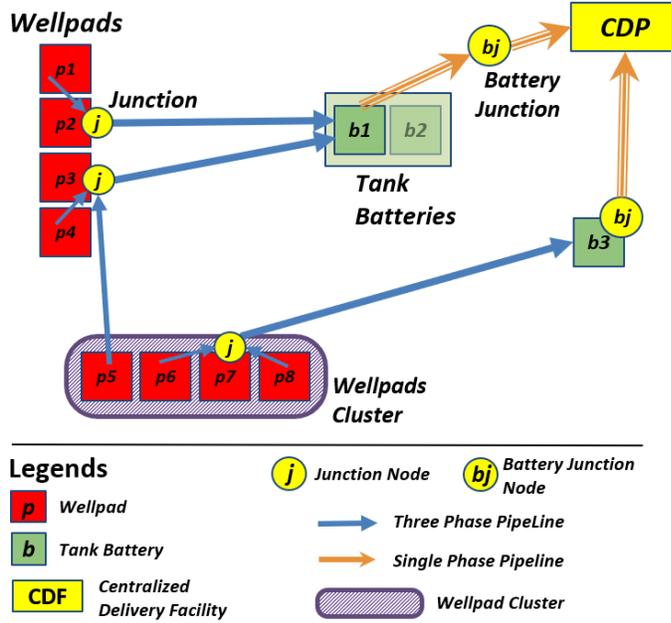
Figure 2 shows an illustrative gathering network chosen from a superstructure of alternatives, where a very large number of pipelines connecting nodes p, j, b, bj and cdp may be adopted, which gives rise to the combinatorial complexity to the problem. Another important concept introduced in Figure 2 is the *cluster of wellpads* ($k \in K$) as a group of wellpads in a common area. It is important to note that all the wellpads start their production at a given time period, and their productivity profiles are assumed to be known. There is no decision to be made on how many and which wells are to be developed. Therefore, the total income at every time period is a fixed term in the objective function.

Figure 2. Typical Oil gathering network infrastructure.

Tank Batteries. Tank batteries can be arranged in two different ways. The first refers to isolated locations where it is possible to install a single tank battery. The second arrangement refers to a *cluster of batteries* (see *b1* and *b2* in Fig. 2) which comprises a single point where several facilities can be installed. In both cases, compressors and pumps are usually installed in an adjacent location (called the tank battery junction) to gain the required energy to reach the delivery location. The cost of these facilities is included in the battery capital expenditure. Regarding the size of a tank battery, it has to be selected from a set of available alternatives ($bz \in BZ$). Each alternative size bz has its own processing capacity $batcap_{bz}^c$ for every component c and a capital expenditure equal to i_bat_{bz}. Companies normally have two or three alternative designs for tank batteries, according to the average flow compositions, and use them over the exploitation areas. The capital expenditure of these facilities is usually several times larger than a normal length pipeline installation cost. Therefore, their optimal location and size is a very relevant decision. In turn, we assume that unit processing costs at tank batteries are invariant with their size and location, from which operating costs for TB represent a constant term in the objective function.

Junction Nodes. A group of accessories and equipment, such as T-junctions or manifolds, are used to connect two or more pipes to route the gathered flow towards a single destination. The potential locations for *junction nodes* are usually the wellpads themselves. The corresponding capital expenditures and operational costs are

included into the objective function to be minimized. Nevertheless, it should be noted that the latter are relatively small in comparison with the other network elements.

Pipelines. Pipeline installation and operating costs are the other important terms to be minimized. The required investment for every mile of a pipeline of diameter d is given by the parameter i_pipe_d. Usually, a limited set of alternative diameters is available. Regarding the pipeline transportation capacity, it is necessary to distinguish between single phase and multi-phase flows and handle them separately. The aim is to determine the maximum flow $\boldsymbol{maxflow_{i,j,d}}$ admitted by every potential pipeline of diameter d linking nodes i and j in the network. This is strongly dependent on the distance between the nodes, the pressure drop being assumed, and the flow composition. Figure

Figure 3. Multi-phase flow patterns according to the gas to liquid ratios and velocities.

3 shows different multi-phase flow patterns according to the liquid to gas ratio and velocities, which trigger different head losses (Oddie et al., 2003; Spedding et al., 2005). In the particular case of multi-phase flows, a reduction in the transportation capacity is observed when compared with the same amount of gas or liquid as moving alone in a pipeline.

For single-phase pipelines the computation is relatively straightforward as reported in the literature (Green and Southard, 2018). In Appendix A we present the main assumptions and equations used for estimating the transportation capacity of single-phase pipelines. In our model, liquid phase pipelines transport oil or water from TB to CDP passing through battery junctions. Since, for this specific case a maximum velocity is imposed, typically 1.5 m/s (Society of Petroleum Engineers, 2006), the maximum flow is assumed to be directly proportional to the pipeline section. The case of gas-phase pipelines is intrinsically more complex due to gas compressibility. A detailed formulation is also given in Appendix A, according to the pressure levels and other specific gas properties following the Weymouth correlation. By means of this model, the maximum flow of a gas pipeline of diameter d is proportional to the pipeline diameter raised to the power of 2.667 (Duran and Grossmann, 1986; Weymouth, 1942).

In the case of multi-phase flows, the transportation capacity of the pipelines is more complex, requiring some important assumptions to be made and then solving a system of nonlinear equations. In fact, the calculation of pressure drops in multi-phase flows is usually based on empirical relationships to consider the effects caused by pressure and temperature changes, the relative velocities of the streams, the flow pattern and the elevation variations. Even though many different models have been proposed to represent multi-phase flows, the Lockhart-Martinelli correlation (Lockhart and Martinelli, 1949) and the SPE guidelines (Society of Petroleum Engineers, 2006), both widely known and used in practice, are taken into account in this model. We address this point with more detail in following sections.

Three-Phase Production Profiles. Production from unconventional wellpads p is characterized by specific gas to oil (gor_p) and water to oil (wor_p) ratios. For simplicity, both are assumed to be deterministic and constant during the wellpad lifespan. Regarding the productivity profile, the production rate starts with a peak followed by a steep decline. This is explained because of the sudden release of the gas, oil and water stored in the pores after the hydraulic fractures are opened. The steady decline in the flowrate is due to the pressure loss and the low permeability of these formations (Baihly et al., 2015; Tan et al., 2018; Vanorsdale, 2013). In fact, the production profile can be represented by a decreasing hyperbolic function of the wellpad age (see Figure 4). The parameter $v_{p,\tau}^c$ is the production of component c at wellpad p in its τth period after the *starttime_p*. It is usually given in

barrels per day for liquid phases, and in standard cubic feet per day for the gas phase. Finally, either from empirical or theoretical models, the long-term production estimates for every wellpad p are assumed to be given.

Single Pad Productivity
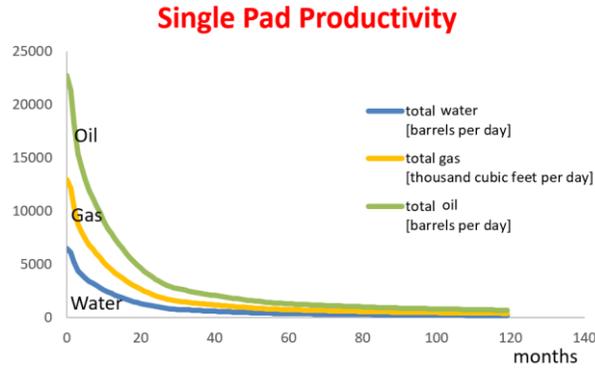
Figure 4. Single Wellpad Productivity Profile for each component: oil, gas and water.

Pressure Estimations. We assume that the pressure at the wellheads is high enough to make oil, gas and water flow towards the tank batteries. Let us consider the production pressure P_p to be known at the outlet of the wellpad p. The inlet pressure at tank batteries has to be at least P_b, while P_j indicates the minimum pressure at junction nodes. For the latter, as the distance between wellpads and junctions is usually short, the value of P_j is close to P_p. We also assume that the pipelines from junction nodes to batteries operate at medium-low pressures. After the stream flows are separated in the tank batteries, the pressure is assumed to drop to P_{bj} at the tank battery junctions, where pumps and compressors direct the fluids to the CDP at high pressure conditions. The pressure after the gas compressor reaches P_{gc}, which is assumed to drop to P_{cdp} at the centralized delivery point. The pipeline capacity computation together with the network design problem leads to a complex non-linear model, which is presented as the general formulation in the next sections. However, if we fix the maximum pressure drops between each pair of nodes in the network, we can predetermine the pipelines capacity, yielding a linear approximation of the original problem. The challenges of each of the formulations, namely the MINLP and MILP, are further fully described.

Time representation. A long term planning horizon is involved (from 5 to 10 years), which is discretized in monthly periods. Despite drilling and completion tasks requiring between 60 and 90 days, and the water flowback requiring 10 to 20 more days, it is common to have a group of drilling and fracturing crews starting to produce one or more new well every month. The exact period (namely, *starttime$_p$*) when a wellpad p starts producing is a given data for this problem. As already mentioned, capital expenditures are very significant and postponing investment decisions over the planning horizon is crucial. Hence, it is required to assess the model results in terms of the *net present cost,* and also to allow the formulation to decide on the investments at any period, thus leading to a multiperiod decision model.

2.2. An Illustrative Example

In order to show one of the tradeoffs, an illustrative example comparing two optional gathering network designs is presented in Figure 5. The unconventional O&G produced by eight different wellpads has to be gathered, separated and sent to a single delivery point. The alternative design A proposes the use of two tank batteries (TB), each one near each group of four wellpads, whereas the alternative B centralizes the separation process at a single TB. Assuming that the processing capacities are met, it can be seen at first sight that the total length of pipelines (particularly those of low pressure) installed in case A is much lower than in case B, but in the latter

case the capital investment in TB is reduced by half. Also, as the distances covered by the multi-phase pipelines are shorter in case A, the pipeline diameters can be much smaller. In a real world case study, these typical tradeoffs have to be integrated with the TB processing constraints and the production scheduling from the wellpads, among other features.
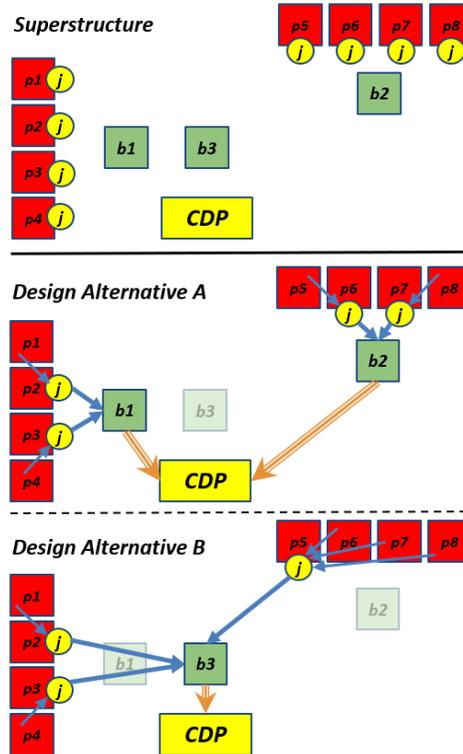
Figure 5. *Network superstructure and two alternative designs*

3. Problem Statement

Based on the description so far, the aim of this section is to formally state the problem. Specifically, given is:

(a) A number of **time periods** (usually months) comprising a long-term time horizon.

(b) A set of **wellpads**, their production **start times** and expected **productivity profiles** (including *gor* and *wor*).

(c) **Potential location** and existing sites for tank batteries, their **alternative capacities** per component and the corresponding **capital expenditure** requirements and **operational costs**.

(d) Potential location for **junction nodes**.

(e) Potential and existing locations for **Centralized Delivery Points (CDP)** and their **available capacities**.

(f) Alternative **pipeline diameters** to use over the network, their **installation costs per unit of length** and the **operational expenditures**.

(g) The **estimated pressure** at the well-heads, the maximum admitted pressure at tank batteries and the compression factor achieved by the compressor at batteries junction nodes.

The objective is to minimize the net present cost of the capital investments required to build the unconventional oil gathering network and the corresponding operational costs by determining: (a) the number of tank batteries to be installed, their location, sizes and period of installation, (b) the junctions to be installed in the network and

their connections, (c) the location and diameter of the pipelines gathering production flows from pads to junctions, from junctions to tank batteries, and finally to CDP. Furthermore, it is necessary to determine the optimal timing for installing facilities over the planning horizon.

In the most general case, the pressures in every node in the network, at every monthly period, are also variables of the problem, which are computed according to the flow properties, the selected diameters and the distances between nodes by applying fluid-dynamics equations. These pressures are determined from the network design since the transportation capacity is a function of these pressures (among other factors). This situation leads to a highly non-linear and non-convex mixed-integer nonlinear programming (MINLP) formulation, which is presented in the next sections.

3.1. Model Assumptions

The following are the main assumptions made for the proposed MINLP model:

(a) Constant and deterministic *gor* and *wor* are assumed for all the wellpads, all over the time horizon.

(b) For every wellpad *p*, the production peak is reached in the first period after completion, that is, in the corresponding start time *starttime$_p$*.

(c) Production from many wellpads may converge to a junction node, and many junction nodes may be connected to a single tank battery. However, splitting flows is not allowed: every pad should be assigned to a single junction node, and every junction node to a single battery. This is mainly due to operating constraints, the need to track the production streams, and the physical/technical issues arising when splitting a three-phase flow. The mathematical implication is the elimination of the bi-linearities required when tracking flow compositions in the network (Drouven and Grossmann, 2017). Even if splitting were permitted, constant *gor* and *wor* would allow the tracking of only one component of the flow in the three-phase pipelines.

(d) Single-phase streams (after separation at the tank batteries) are treated differently according to the component. For liquid phase pipelines (oil/water), the maximum rate is proportional to the pipeline section considering a maximum average velocity of 1.5 m/s, while the gas pipelines are modeled through the Weymouth correlation (see Appendix A).

(e) Multi-phase flows are modeled by the Lockhart-Martinelli correlation (Lockhart & Martinelli, 1949) and the SPE guidelines (Society of Petroleum Engineers, 2006), both widely known and used in practice. Indeed, pressure drop calculations are based on these empirical correlations considering the effects of pressure and temperature changes, relative velocities, flow pattern and the elevation variations (see Appendix A).

(f) Installation or construction times are disregarded. We assume that the capacities of the facilities are available in the same time period the investments are scheduled.

4. General Mathematical Formulation

In this section, an overall MINLP mathematical formulation for the optimal design of the unconventional oil gathering network is formally presented. It aims to account for numerous trade-offs, selecting the most convenient location for facilities and pipeline connections, together with the determination of nodes pressures and flow conditions that minimize the total present cost from both capital and operational expenditures over the planning horizon. A mathematical formulation is proposed by considering a rigorous multiperiod model in which the installation decisions can be made at any time period over the planning horizon.

The aim is to design a gathering network for the production of a set $P = \{1, 2, ..., p\}$ of unconventional oil wellpads, whose production is scheduled to start at a given period $starttime_p$ over the planning horizon. The time horizon is comprised by $T= \{1, 2, ..., t\}$ periods (usually months). Firstly, the production of every wellpad has to be transported to one and only one potential junction node from the set $J= \{1, 2, ..., j\}$, where the production from other wellpads may be also gathered before flowing to one and only tank battery from the potential set $B= \{1, 2, ..., b\}$. After the separation of each component $c \in C=\{oil, gas, water\}$, the individual phases are transported in single-phase pipelines to centralized delivery points $cdp \in CDP=\{1, 2, ..., cdp\}$ being previously gathered in single-component junction nodes $bj \in BJ=\{1, 2, ..., bj\}$. Moreover, the production flow through every pair of nodes implies the decision of the installation of a pipeline of a certain diameter $diam_d$ from the available set $d \in D=\{1, 2, ..., d\}$. It is important to mention that the wellpads are usually grouped into clusters k by geographical or operational reasons, which is used later to reduce the problem size. More specifically, every wellpad belongs to one and only one cluster $k \in K$. The constraints of the mathematical formulation can be divided into the flow balances, facilities installation and sizing, topological constraints and integer cuts.

<u>Flow Balances at Critical Time Points</u>. Eq. 1 imposes that the wellpad production has to be delivered toward the selected junction node at every time period after production starts. Note that $v_{p,\tau}^c$ is the expected production of component c at pad p, τ periods after its start date (in barrels or thousand-cubic feet per day, according to the component). The continuous variable $FP_{p,j,t}^c$ represents the flow of component c from pad p to junction node j during period t. Eqs. 2, 3 and 4 are the mass balances for every single component at junction nodes, tank batteries and battery junctions, respectively, while the continuous variables $FPB_{j,b,t}^c$, $FBJ_{b,bj,t}^c$ and $FCDP_{bj,cdp,t}^c$ are the flows of component c moving between nodes j, b, bj and cdp during period t according to the segment being linked in the hierarchical topology of the network.

$$v_{p,t-starttime_p+1}^c = \sum_{j \in J_p} FP_{p,j,t}^c \quad \forall p, c, t \in TP, t \geq starttime_p \tag{1}$$

$$\sum_{p \in P_j} FP_{p,j,t}^c = \sum_{b \in B_j} FPB_{j,b,t}^c \quad \forall j, c, t \in TP \tag{2}$$

$$\sum_{j \in J_b} FPB_{j,b,t}^c = \sum_{bj \in BJ_b} FBJ_{b,bj,t}^c \quad \forall b, c, t \in TP \tag{3}$$

$$\sum_{b \in B_{bj}} FBJ_{b,bj,t}^c = \sum_{\substack{cdp \in CPD_{bj} \\ \cap CDP_c}} FCDP_{bj,cdp,t}^c \quad \forall bj, c, t \in TP \tag{4}$$

Note that it is not strictly necessary to control the flow between every pair of nodes when no wellpad increases its production rate, which are the most critical periods in terms of transportation and processing requirement. Therefore, some important subsets are introduced in the equations above in order to limit the size of the model and the connectivity of the network: (a) **TP** includes the periods t with some wellpad increasing its production rate regarding the previous period, which in our case involves the initial period when the production peak is reached; (b) **TS** includes the periods in which some wellpad starts production and the infrastructure investment might be planned, (c) **J_p** and **P_j** are subsets of wellpads-junctions to be potentially connected, (d) **J_b** and **B_j** are junctions-tank batteries that might be linked, (e) **BJ_b** and **B_{bj}** are batteries-battery junctions potential connections, (f) **CDP_{bj}** and **BJ_{cdp}** are subsets of battery junctions and centralized facilities to be potentially linked, and (g) **CDP_c** includes the centralized delivery nodes available for component c. Subsets (c) to (f) are usually built from geographical proximity criteria. It is relevant to note that **J_p** and **P_j** are built with the following logic: if a wellpad

p and the junction j are part of the same cluster k, then the pair p-j is included, whereas if p and j belong to different clusters, but these clusters are close enough, then the pair p-j is also added to the corresponding subset. Otherwise, the pair p-j is not included among the possible connections. Similar logic is applied to build the ordered sets J_b, B_j, BJ_b, B_{bj}, CDP_{bj} and BJ_{cdp}. However, as later discussed, tank batteries capital expenditures are determinant to the total investment, often requiring to build some pipelines that at first sight look counterintuitive and costly.

Processing and Delivery Capacities. Tank batteries installation and capacity control constraints are shown in Eq. 5, where the set $BZ=\{1, 2, ..., bz\}$ comprises the alternative battery capacities ($batcap_{bz}^c$), being BZ_b the capacities bz allowed at location b. Likewise, the corresponding control for the centralized delivery points is given by Eq. 6, with $cdpcap^c$ being the maximum flowrate of component c allowed by the centralized facility cdp. As observed, two main binary variables are introduced to model the tank batteries and centralized facilities installation decisions: (a) $\hat{z}bat_{b,bz,t}$ is equal to 1 if a battery of capacity bz is installed in the location b at period t, and (b) $\hat{z}cdp_{cdp,t}$ that equals 1 if the delivery point cdp is installed at period t. Note that the tank batteries and centralized delivery points can only be installed in the time periods when a new wellpad starts its production ($t' \in TS$). On the other hand, $zbat_{b,bz,t}$ and $zcdp_{cdp,t}$ determine that the corresponding facilities are already available at period t. A value equal to 1 in $zbat_{b,bz,t}$ or $zcdp_{cdp,t}$ implies that in the same or in a previous period the corresponding installation has already been made, which is captured by Eq. 7 and illustrated in Figure 6.

$$\sum_{j \in JB} FPB_{j,b,t}^c \leq \sum_{bz \in BZ_b} batcap_{bz}^c \cdot zbat_{b,bz,t} \qquad \forall b, c, t \in TP \tag{5}$$

$$\sum_{bj \in DF_{bj,cdp}} FCDP_{bj,cdp,t}^c \leq cdpcap^c \cdot zcdp_{cdp,t} \qquad \forall c, cdp \in CDP_c, t \in TP \tag{6}$$

$$zbat_{b,bz,t} = \sum_{\substack{t' \leq t \\ t' \in TS}} \hat{z}bat_{b,bz,t'} \quad \forall b, bz \in BZ_b, t \in TP \qquad zcdp_{cdp,t} = \sum_{\substack{t' \leq t \\ t' \in TS}} \hat{z}cdp_{cdp,t'} \quad \forall cdp, t \in TP \tag{7}$$

To reduce the model size, since **gor** and **wor** are assumed to be constant, Eq. 5 can be imposed only for the most restricting component, by comparing **gor** and **wor** with the single component capacities of the tank batteries. These straightforward strategies can significantly reduce the model size.
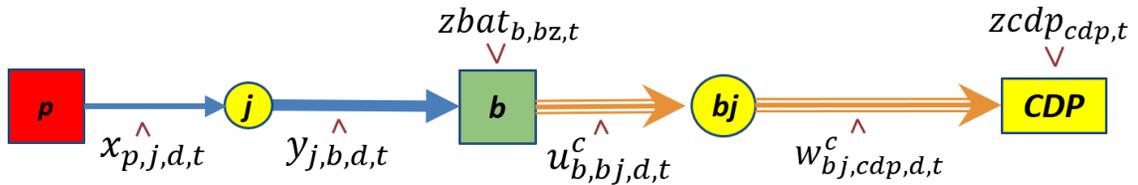
Figure 6. *Binary variables indicating the availability of a facility in the superstructure.*

It is also interesting to note the relationship of this formulation with the knapsack problem from Integer Programming (Wolsey, 1998), with every component of the flow from the wellpads representing the "packages" that have to be allocated in the TB (knapsack). However, in this case there are the following additional conditions: (1) every component has its own processing capacity in the TB, (2) the magnitude of the flows changes with time because of the declining production curves and the drilling of new unconventional wells, and (3) the number of TB to be installed, its size and location is not pre-determined.

Pipeline Selection. Pipeline sizing and installation decision variables are also divided into two groups, one determining the period in which it is installed and the other one to state the availability of the resource over time. The first one, whose elements can be active for at most one period of the planning horizon, comprises the following variables: (a) $\hat{x}_{p,j,d,t}$ being 1 if a pipeline of diameter d connecting pad p to junction node j is installed at period t, (2) $\hat{y}_{j,b,d,t}$ equals 1 if a pipeline of diameter d connecting node j to battery b is installed at period t, (c) $\hat{u}^c_{b,bj,d,t}$ is 1 if a pipeline of diameter d for component c connecting battery b and junction bj is installed at period t, and finally (d) $\hat{w}^c_{bj,cdp,d,t}$ taking a value of 1 if a pipeline of diameter d for component c connecting junction bj with the delivery point cdp is installed at period t. Similar to the tank batteries and delivery points, pipelines can only be installed in those time periods in which a new wellpad starts its production ($t' \in TS$). The second group of variables consists of: $x_{p,j,d,t}$, $y_{j,b,d,t}$, $u^c_{b,bj,d,t}$ and $w^c_{bj,cdp,d,t}$ which are equal to one if the corresponding pipeline is already installed at period t. Eqs. 8 and 9 relate both groups of variables and Figure 6 illustrates them.

$$x_{p,j,d,t} = \sum_{\substack{t' \le t \\ t' \in TS}} \hat{x}_{p,j,d,t'} \quad \forall p, j \in J_p, d, t \in TP \qquad\qquad y_{j,b,d,t} = \sum_{\substack{t' \le t \\ t' \in TS}} \hat{y}_{j,b,d,t'} \quad \forall j, b \in B_j, d, t \in TP \tag{8}$$

$$u^c_{b,bj,d,t} = \sum_{\substack{t' \le t \\ t' \in TS}} \hat{u}^c_{b,bj,d,t'} \quad \forall c, b, bj \in BJ_b, d, t \in TP$$

$$w^c_{bj,cdp,d,t} = \sum_{\substack{t' \le t \\ t' \in TS}} \hat{w}^c_{bj,cdp,d,t'} \quad \forall c, bj, cdp \in CDP_{bj}, d, t \in TP \tag{9}$$

Regarding the pipeline transportation constraints, the total flow from a pad to a junction node, and from a junction node to a tank battery should not exceed the selected pipeline capacity. This is achieved by controlling the pressures drops for each segment and diameter selected, which is further elaborated in the *Multi-phase Pipelines Capacity* sub-section. After the multi-phase flows are separated, single-product pipelines connect batteries to battery junction nodes, and the latter to centralized delivery points. Single-product pipelines are sized as given by Eq. 10 and Eq. 11. Parameters ϕ^c are unit-transformation coefficients that depend on the fluid-dynamic model used to calculate pipeline transportation capacity.

$$\phi^c FBJ^c_{b,bj,t} \le \sum_d maxflow^c_{b,bj,d} \cdot u^c_{b,bj,d,t} \qquad \forall b, bj \in BJ_b, c, t \in TP \tag{10}$$

$$\phi^c FCDP^c_{bj,cdp,t} \le \sum_d maxflow^c_{bj,cdp,d} \cdot w^c_{bj,cdp,d,t} \qquad \forall bj, cdp \in CDP_{bj}, c, t \in TP \tag{11}$$

In a given superstructure with alternative connections, the maximum transportation capacities $\boldsymbol{maxflow^c_{b,bj,d}}$ and $\boldsymbol{maxflow^c_{bj,cdp,d}}$ are the corresponding single-phase capacities for component c. As mentioned in the previous sections and further discussed in Appendix A, the capacity transportation for liquid-phase pipelines transporting oil or water is determined by the Eq. 12, while Eq. 13 represents the maximum flow for gas-phase pipelines. In Eq. 12, $vmax^{LP}$ is the fixed maximum velocity (1.5 m/s) for the liquid phase, while in both equations μ is used for units conversion. Observe that, for liquid-phase pipelines, the maximum flow does not depend on the distance between the nodes being linked, in contrast with the inverse relation existing for gas pipelines (Weymouth, 1942). In the most general case, the inlet and outlet pressures are also decision variables.

$$maxflow_d^{oil/water} = \mu^{oil/water} \cdot vmax_{LP} \cdot \frac{\pi}{4} \cdot \left(diam_d^2\right) \quad \forall d \tag{12}$$

$$maxflow_d^{gas} = \frac{\mu^{gas}}{dist^{gas}} \cdot (P_{in}^2 - P_{out}^2)^{0.5} \cdot diam_d^{2.667} \quad \forall d \tag{13}$$

Multi-phase Pipelines Capacity. A series of nonlinear equations aim to control the transportation capacities allowed for multiphase (oil, gas, water) pipelines of diameter d linking every pair of nodes i and j. The corresponding transportation capacity is strongly dependent on the **gor** and **wor**, or, in simpler terms, dependent on the ratio between the liquid and the gas phases. Even if we assume constant temperature and horizontal pipelines, the resulting flow patterns are quite complex depending on the liquid-gas ratio and physical properties of the fluids. According to the literature on multiphase flows (Al-Hadhrami et al., 2014; Petalas and Aziz, 2000; Spedding et al., 2006), there are many theoretical models and empirical correlations to compute the flow conditions and pressure drop in different lab conditions, but just a few are useful for O&G production (Edwards et al., 2018; Zaghloul et al., 2008). Tools to predict flow regimes vary from large computational methods and algorithms (usually in the form of commercial codes) to very simple graphical procedures (flow pattern maps). In this model, two blocks of equations have been selected to represent every triplet origin-destination-diameter. The first block is developed under the Lockhart-Martinelli (LM) correlation for liquid-gas flows, whereas the second one includes SPE guidelines for multi-phase pipeline sizing.

In order to state the equations more clearly, we define them for a generic variable flowrate FP^c of every component c instead of referencing it for a pair of nodes and a period of time. Also, the equations are stated for every potential diameter d. More specifically, the set of equations 14 to 28 should be imposed for every pair of nodes (wellpads-junctions ($FP_{p,j,t}$) / junctions-tank batteries ($FPB_{j,b,t}$)) and every pipeline diameter. Also, it would be required to multiply the flowrates by the corresponding binary variables ($x_{p,j,d,t}$ or $y_{j,b,d,t}$) to account for the selected pipeline diameter. Although these bilinear terms can be linearized, they are retained to gain clarity in the presentation.

Lockhart-Martinelli (LM) Correlation. This procedure is for predicting the pressure drop for fully developed horizontal gas-liquid flows, and is widely used in the Chemical Engineering field (Green and Southard, 2018; Lockhart and Martinelli, 1949). It is based on the calculation of the pressure drops that would be expected for each of the phases as if flowing alone through the pipeline. Then, the Lockhart-Martinelli parameter is computed as the square root of the ratio between the individual expected pressure drops. Finally, the overall two-phase pressure drop is estimated by using phase-specific multipliers that are obtained from nonlinear correlations with the LM parameter. We assume oil and water streams as part of a common liquid phase. Even though this is not entirely valid since oil and water phases are not miscible, generating particularly complex flow-patterns, the assumption is made to gain simplicity. In fact, as reviewed by Edwards et al. (2018), the major effect on the pressure drops could be the apparent viscosity of the liquid mixture, which depend on the water fraction. In our view, this assumption does not excessively affect the pressure drops computations since the *gor* and *wor* are constant terms in our model.

From Eq. 14 we obtain the ratios between the component flow rates. Eq. 15 computes the superficial velocities for liquid (VS^{LP}) and gas (VS^{GP}) phases, while Eq. 16 determines the estimated Reynolds numbers for liquid and gas phases (RY^{LP} - RY^{GP}). The remaining parameters account for: $diam_d$, the pipeline diameter [m], ρ and μ, the average density [kg/m³] and viscosity [Pa.s] of each phase, respectively. The friction factor HFF^{LP} for the liquid phase is computed through the Haaland correlation in Eq. 17, while the pressure drop gradient $\Delta P^{LP}/L$ is estimated by Eq. 18, as if the liquid phase were flowing alone in the pipeline. The parameter ε accounts for the

pipeline internal surface roughness [m]. For the gas phase, the Weymouth equation (Weymouth, 1942) estimates a pressure gradient along the pipeline, as in Eq. 19.

$$wor \cdot FP^{oil} = FP^{water} \qquad gor \cdot FP^{oil} = FP^{gas} \tag{14}$$

$$VS_d^{LP} = \frac{FP^{water} + FP^{oil}}{\pi diam_d^2/4} \qquad VS_d^{GP} = \frac{FP^{gas}}{\pi diam_d^2/4} \tag{15}$$

$$RY_d^{LP} = \frac{\rho^{LP} \cdot VS^{LP} \cdot diam_d}{\mu^{LP}} \qquad RY_d^{GP} = \frac{\rho^{GP} \cdot VS^{GP} \cdot diam_d}{\mu^{GP}} \tag{16}$$

$$\frac{1}{(HFF_d^{LP})^{0.5}} = -1.8 \, log \left(\frac{\varepsilon}{3.7 \, diam_d} + \frac{6.9}{RY_d^{LP}} \right) \tag{17}$$

$$\frac{\Delta P_d^{LP}}{L} = \frac{HFF_d^{LP}}{2} \cdot \frac{\rho^{LP} \, VS_d^{LP \, 2}}{diam_d} \tag{18}$$

$$FP^{gas} = dist^{-0.5} \cdot diam_d^{2.667} \left\{ \frac{(P_{in}^2 - P_{out}^2)}{(\rho^{GP} T[P_0/0.375 T_0]^2)} \right\}^{0.5} \tag{19}$$

$$\frac{\Delta P^{GP}}{L} = \frac{(P_{in} - P_{out})}{dist} \tag{20}$$

Note that in the most general case, only the pressure at the wellpads is given. As the gas flow FP^{gas} increases, the square pressure difference $(P^2_{in} - P^2_{out})$ also increases, leading to a larger single-phase pressure gradient $\Delta P^{GP}/L$ (Eq. 20). The other parameters account for the average gas temperature T [K], the standard temperature T_0 [K] and pressure P_0 [MPa] conditions. Also, the inlet pressure P_{in} has an upper limit according to the segment of network being considered. It can be easily proved that $\Delta P^{LP}/L$ and $\Delta P^{GP}/L$ are strictly increasing functions with the oil flow rate, for fixed *gor* and *wor*.

$$X_{LM} = \left(\frac{\Delta P^{LP}/L}{\Delta P^{GP}/L} \right)^{0.5} \tag{21}$$

$$Y^{GP} = \left(1 + X_{LM}^{\frac{2}{n}} \right)^n \quad Y^{LP} = \frac{Y^{GP}}{X_{LM}^2} \tag{22}$$

The Lockhart-Martinelli parameter X_{LM} is a measure of the relation between the single-phase pressure gradients, which is computed by Eq. 21. The corresponding multipliers Y^{LP} and Y^{GP} are obtained in Eq. 22 to determine how much the single-phase pressure gradients are increased when both phases are flowing together in a two-phase flow. The latter equation follows the Wilkes approach (Wilkes, 2005) to fit the curves of the LM correlation (see Figure 7). The parameter *n* depends on the flow regime of each phase, being equal to 4.12 if both liquid and gas phases have a turbulent regime, as is usually seen in practice for unconventional streams. The overall multiphase pressure gradient $\Delta P/L$ is computed by Eq. 23 and the total pressure drop ΔP^{Total} is given by Eq. 24 while imposing a maximum pressure drop Δp^{max}. This limit in the pressure drops is set according the operating conditions of the manifolds and the tank batteries. Note that the overall pressure gradient $\Delta P/L$ is also a strictly increasing function with the oil flow rate, for fixed *gor* and *wor*.

$$\frac{\Delta P}{L} = Y^{LP} \cdot \frac{\Delta P^{LP}}{L} \tag{23}$$

$$\Delta P^{Total} = \frac{\Delta P}{L} \cdot dist \ \leq \Delta p^{max} \tag{24}$$

SPE (Society of Petroleum Engineers) Guidelines. According to the SPE (Society of Petroleum Engineers, 2006), certain recommendations should be taken into account to determine the diameter of a multiphase pipeline. These are related to maximum and minimum velocity constraints to prevent erosion, corrosion, noise or water hammer effects. The maximum recommended velocity is 60 ft/sec to inhibit noise, and 50 ft/sec for CO_2 corrosion inhibition. In turn, a minimum velocity of 10 to 15 ft/sec is usually set to prevent surging and keep the line swept clear of solids.
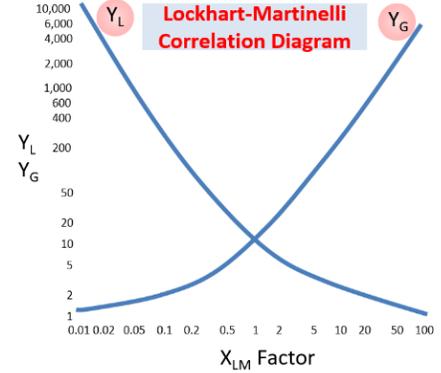
Figure 7. Liquid and gas pressure gradient multipliers vs. Lockhart-Martinelli parameter.

The SPE guidelines for multiphase pipeline design and sizing are then added to the MINLP formulation. Eq. 25 limits the flow in order not to exceed the maximum velocity v_{max} [ft/s], according to the given pipeline diameter. The parameter v_{max} is the maximum velocity according to SPE procedure [ft/s] and it is computed as the division between the constant ζ and the square root of the average density ρ_{avg}. ζ is an SPE specific constant with a value of 150 for solids-free fluids and continuous service operation. The average density is also a parameter computed as shown in Eq. 26. Moreover, z is the compressibility factor, r the gas/liquid ratio [ft³/bbl], T the gas/liquid flowing temperature [°R], P_{in} the inlet pressure in PSI, sg the specific gravity of the liquid phase relative to water, and s the specific gravity of the gas, relative to air.

$$FP^{oil} + FP^{water} \leq \frac{1000 \cdot v_{max} \cdot (0.0254\,diam_d)^2}{\left(11.9 + \dfrac{z\,r\,T}{16.7\,P_{in}}\right)} \tag{25}$$

$$v_{max} = \frac{\zeta}{\rho_{avg}^{0.5}} \quad ; \quad \rho_{avg} = \frac{12409\,sg\,P_{in} + 2.7\,r\,s\,P_{in}}{198.7\,P_{in} + z\,r\,T} \tag{26}$$

Topological Constraints. A set of constraints is introduced in order to account for the topological conditions of the network. First, it is necessary to impose that the outgoing flows cannot be split at any node in the network (Eq. 27 and Eq. 28). These constraints also enforce that only a single diameter be selected for any pipeline connection, and that the installation can be made at a single period within the subset *TS*. Note that for the first segment in the network it is possible to tighten the constraint by making it strictly equal to one (production needs to be sent to a junction node). On the other hand, the last inequality in constraint 28 also establishes that in a potential tank battery location only one battery capacity can be installed. Note that the investments in infrastructure over the network can only be scheduled for those time periods in which a new wellpad starts its production.

$$\sum_{\substack{j\,\in\,J_p \\ d,t\,\in\,TS}} \hat{x}_{p,j,d,t} = 1 \quad \forall p \qquad\qquad \sum_{\substack{b\,\in\,B_j \\ d,t\,\in\,TS}} \hat{y}_{j,b,d,t} \leq 1 \quad \forall j \tag{27}$$

$$\sum_{\substack{bj\,\in\,BJ_b \\ d,t\,\in\,TS}} \hat{u}^c_{b,bj,d,t} \leq 1 \quad \forall b,c \qquad \sum_{\substack{cdp\,\in\,CPD_{bj} \\ d,t\,\in\,TS}} \widehat{w}^c_{bj,cdp,d,t} \leq 1 \quad \forall bj,c \qquad \sum_{\substack{bz\,\in\,BZ_b \\ t\,\in\,TS}} \hat{z}bat_{b,bz,t} \leq 1 \quad \forall b \tag{28}$$

15

Integer Cuts. Further constraints are included in order to tighten the mathematical formulation. Tank battery location and delivery point selection restrain pipeline connections, as given by Eqs. 29 and 30. In other words, no pipeline can be connected to a node that is not selected. Similarly, Eq. 31 relates the binary variables accounting for upstream and downstream flows in every node: no downstream pipeline can leave a node if no upstream pipeline reaches the same node.

$$y_{j,b,d,t} \leq \sum_{bz \, \epsilon \, BZ_b} zbat_{b,bz,t} \quad \forall j, b \in B_j, d, t \, \epsilon \, TP \qquad u^c_{b,bj,d,t} \leq \sum_{bz \, \epsilon \, BZ_b} zbat_{b,bz,t} \quad \forall c, b, bj \in BJ_b, d, t \, \epsilon \, TP \tag{29}$$

$$w^c_{bj,cdp,d,t} \leq zcdp_{cdp,t} \quad \forall c, bj, cdp \in CDP_{bj} \cap CDP_c, d, t \, \epsilon \, TP \tag{30}$$

$$\sum_{b \, \epsilon \, B_{j,d}} y_{j,b,d,t} \leq \sum_{p \, \epsilon \, P_{j,d}} x_{p,j,d,t} \quad \forall j, t \, \epsilon \, TP \qquad \sum_{bj \, \epsilon \, BJ_{b,d}} u^c_{b,bj,d,t} \leq \sum_{j \, \epsilon \, J_{b,d}} y_{j,b,d,t} \quad \forall b, t \, \epsilon \, TP$$

$$\sum_{cdp \, \epsilon \, CDP_{bj,d}} w^c_{bj,cdp,d,t} \leq \sum_{b \, \epsilon \, B_{bj,d}} u^c_{b,bj,d,t} \quad \forall bj, t \, \epsilon \, TP \tag{31}$$

Symmetry Breaking Constraints. Many nodes b for TB installation may comprise the same physical location (a cluster of batteries). In fact, there may be a maximum of tank batteries that can be installed in the location, which induces several equivalent or symmetric solutions when varying the individual element being selected and the corresponding flow allocation. Therefore, we systematically activate the tank batteries within the same location following the lexical order of the individual candidates, yielding Eq. 32.

$$\sum_{bz \, \epsilon \, BZ_{b+1}} zbat_{b+1,bz,t} \leq \sum_{bz \, \epsilon \, BZ_b} zbat_{b,bz,t} \quad \forall b, b + 1 \, \epsilon \, BatCluster, t \, \epsilon \, TP \tag{32}$$

Objective Function. The goal of the mathematical formulation is to minimize the total present cost of capital and operating expenditures involved in the network design and operation. Without loss of generality, we assume that the availability of the installed facilities is instantaneous so as to be able to discount the cost appropriately. The objective function shown in Eq. 33 accounts for: (1) the tank batteries installation costs, being i_bat_{bz} the cost of a battery of capacity bz [in thousands USD], (2) the investment required for pipelines in the network, being i_pipe_d the unit cost of a pipeline of diameter d [thousand USD/mile] and $dist_{p,j}$ / $dist_{j,b}$ / $dist_{b,bj}$ / $dist_{bj,cdp}$ the euclidean distances between nodes [miles], (3) the junction capital requirements, where i_jun_j and i_junb_{bj} represent the installation costs of a junction, while $\widehat{xy}_{j,t}$ and $\widehat{bbj}_{bj,t}$ are new binary variables defining if the corresponding junctions are installed at period t ($\widehat{xy}_{j,t} \geq \hat{x}_{p,j,d,t}$, and $\widehat{bbj}_{bj,t} \geq \hat{u}^c_{b,bj,d,t}$), (4) pipelines operational expenditures, where c_pipe_d is the unit cost of using a pipeline per unit of volume being transported [thousand USD/(bbl-tcf³)], and (5) tank batteries and junctions operational costs, where c_bat_{bz} and c_jun are the corresponding costs per unit of volume being processed, and $xy_{j,t}$ and $bbj_{bj,t}$ determine the use of a junction during a period t. For all cases, i is the interest rate considered to assess the capital flows through the planning horizon. Strictly speaking, and for simplicity, all the operating cost terms are given by non-linear terms involving a binary variable and a continuous one. These conditional terms can be easily linearized by introducing a new positive variable and a set of new linear constraints.

$$Min\ z = \sum_t \frac{1}{(1+i)^{t-1}} \left\{ \sum_{b,bz} \left(i\_bat_{bz} \cdot \hat{z}bat_{b,bz,t} + zbat_{b,bz,t} \cdot c\_bat_{bz} \sum_{c,j} FPB^c_{j,b,t} \right) \right.$$

$$+ \sum_{p,j,d} \left(dist_{p,j} \cdot i\_pipe_d \cdot \hat{x}_{p,j,d,t} + x_{p,j,d,t} \cdot c\_pipe_d \sum_c FP^c_{p,j,t} \right)$$

$$+ \sum_{j,b,d} \left(dist_{j,b} \cdot i\_pipe_d \cdot \hat{y}_{j,b,d,t} + y_{j,b,d,t} \cdot c\_pipe_d \sum_c FPB^c_{j,b,t} \right) \qquad (33)$$

$$+ \sum_{c,b,bj,d} \left(dist_{b,bj} \cdot i\_pipe_d \cdot \hat{u}^c_{b,bj,d,t} + u^c_{b,bj,d,t} \cdot c\_pipe_d \cdot FBJ^c_{b,bj,t} \right)$$

$$+ \sum_{c,bj,cdp,d} \left(dist_{bj,cdp} \cdot i\_pipe_d \cdot \hat{w}^c_{bj,cdp,d,t} + w^c_{bj,cdp,d,t} \cdot c\_pipe_d \cdot FCDP^c_{bj,cdp,t} \right)$$

$$\left. + \sum_j \left(i\_jun_j \cdot \widehat{xy}_{j,t} + xy_{j,t} \cdot c\_jun \sum_{c,p} FP^c_{p,j,t} \right) + \sum_{bj} \left(i\_junb_{bj} \cdot \widehat{bb}J_{bj,t} + bbj_{bj,t} \cdot c\_jun \sum_{c,b} FBJ^c_{b,bj,t} \right) \right\}$$

5. Optimization Framework

After having presented a general MINLP formulation for the unconventional production gathering network design, a global solution strategy is developed in this section. The MINLP model proves to be computationally intractable even for rather small instances, because of the following reasons: (1) the highly nonlinear and non-convex nature of the constraints, (2) the combinatorial complexity due to the large number of alternatives when dealing with case studies of reasonable size, and (3) the loose NLP relaxation that yields poor bounds. In fact, for a case study resulting in a MINLP model with 6526 single equations, 3250 nonnegative variables and 350 binary variables, after one day of CPU processing, not even a feasible solution had been found using BARON global optimization solver.

The proposed solution strategy is composed of a series of approximations derived from the MINLP model, which when solved in a specified sequence, can provide efficient near-optimal and feasible solutions for the overall problem. The overall solution strategy is based on a bi-level decomposition consisting of: (1) an NLP formulation comprising all the fluid dynamic equations to estimate the transportation capacity of the pipelines, and (2) a MILP formulation for the gathering network design.

5.1. NLP Formulation for Estimation of the Maximum Flow in Multiphase Pipelines

To separate all the nonlinear and nonconvex elements of the MINLP formulation, an NLP model comprising the group of equations from Eq. 14 to Eq. 26 is developed. By carefully analyzing these equations, it is important to highlight that, for any potential multiphase pipeline in the network (defined by the triplet origin-destination-diameter), we can setup an NLP model whose objective is to maximize the value of the variable ***maxflow***$_{i,j,d}$ computing the total maximum flow. Also note the following key aspects: (1) if the inlet and outlet pressures are assumed to be given for each segment, the NLP model leads to the transportation capacity of the pipeline for the given pressure drop Δp^{max}, (2) the resulting NLP formulation is a system of nonlinear equations that can be solved separately for every triplet *i-j-d*, resulting in a large number of small NLP models; and (3) segments with multiphase flows are the ones linking wellpads *p* with junction nodes *j*, and junctions *j* with tank batteries *b*.

In summary, the NLP aims at maximizing the multiphase flowrate for every potential pipeline segment in the network, given by $Max \; z = FP^{oil}$, subject to Eqs. 14 to Eq. 26. In other words, we want to setup a reasonable capacity for the segment *i-j* with diameter *d* that can be designed in practice from an engineering viewpoint. Note that the oil flowrate is used for reference as the *gor* and *wor* parameters are assumed to be constant. It is important to mention that the proposed NLP formulation comprises both the Lockhart-Martinelli procedure and the SPE guidelines. In general terms, a separate comparison of LM and SPE procedures leads to the conclusion that the first one is more restrictive for longer pipelines, while the second one imposes tighter upper bounds on the flowrate as the pipeline becomes shorter. In fact, in contrast with the LM correlation, the SPE guidelines are independent of the pipeline length because they are not based on the calculation of the pressure drop. In Appendix B, an illustrative example is presented to show the formulation results and its understanding.

5.2. Multiperiod MILP Formulation: A Rigorous Approach for Investment Planning

Based on the previous NLP formulation, the resulting parameter ***maxflow***$_{i,j,d}$ is introduced into an MILP formulation for the optimal design of the gathering network. In particular, the new Eqs. 34 and 35 for pipelines connecting wellpads with junction nodes, and junctions with tank batteries, are added to the group of equations given by Eq. 1 to 13 and Eq. 27 to 33, leading to a MILP model. Its objective function to be minimized is given by Eq. 33, which is subject to Eqs. 1 to 13, 27 to 32, 34 and 35.

$$\sum_c \phi^c \cdot FP^c_{p,j,t} \leq \sum_d maxflow_{p,j,d} \cdot x_{p,j,d,t} \quad \forall \; p,j \in J_p, t \in TP \tag{34}$$

$$\sum_c \phi^c \cdot FPB^c_{j,b,t} \leq \sum_d maxflow_{j,b,d} \cdot y_{j,b,d,t} \quad \forall \; j,b \in B_j, t \in TP \tag{35}$$

Pre-processing step. According to the maximum flowrate estimation, for both multiphase and single-phase pipeline segments, straightforward logic can be applied to reduce the size of the problem without affecting the quality of the solutions being obtained:

- For every wellpad-junction connection, the maximum flowrate over time is known in advance according to the wellpad production forecast. Hence, it is possible to minimize the pipeline costs by selecting the minimum diameter *d* that is able to handle the production peak of the wellpad. This is achieved by fixing the $x_{p,j,d'}$ variables to zero for all $d' \neq d$.

- For connections between junction nodes and tank batteries, the minimum flow to be transported is that coming from the less productive wellpad associated to *j*. Then, it is possible to omit from the network superstructure every diameter *d* that is not able to handle at least the production of that single wellpad.

- Also, it is possible to assume that no more than a certain number of wellpads will be connected to the junction node, according to the maximum processing capacity of a single tank battery. From that, one can omit alternative diameters that are larger than required.

5.3. MILP Formulation with investments at initial time

A new strong assumption is introduced in this section to reduce the size of the MILP formulation, and thereby improving its computational performance. We assume that all the gathering facilities selected by the model are installed and available at time zero, yielding a compact formulation with a large reduction in the model size, both in variables and equations.

Note that every feasible solution of this compact MILP is also a feasible solution of the previous MILP. However, in some cases, it might be beneficial to delay some capital expenditures at the expense of smaller ones being brought forward, which is disregarded in this simplified model.

Model Reformulation. The main changes with regards to the previous MILP are as follows. Regarding the binary variables, all decisions related to the installation of a facility after the first period are omitted, as stated by Eqs. 36 and 37.

$$\hat{x}_{p,j,d,t} = 0 \quad \forall p,j,d,t \in TS, t > 1 \;;\; \hat{y}_{j,b,d,t} = 0 \quad \forall j,b,d,t \in TS, t > 1 \;;\; \hat{u}^c_{b,bj,d,t} = 0 \quad \forall b,c,bj,d,t \in TS, t > 1 \quad (36)$$

$$\hat{w}^c_{bj,cdp,d,t} = 0 \quad \forall bj,c,cdp,d,t \in TS, t > 1 \quad ; \quad \hat{z}bat_{b,bz,t} = 0 \quad \forall b,bz,t \in TS, t > 1 \quad (37)$$

Moreover, after solving the simplified MILP, it is important to compute the actual net present cost of every facility by postponing the capital expenditures to the time they are used for the first time, according to the network design. These computations are introduced to determine the *net present cost* of the solution, which is comparable with those yielded by the previous MILP.

It is also interesting to note that this formulation yields feasible solutions for the multiperiod MILP model presented in section 5.2. This can be easily proved by considering that: (a) every solution of the reduced MILP model is a solution in the multiperiod approach, and (b) the objective function when the capital expenditures are discounted back in the multiperiod MILP is always lower or equal to the total cost in the reduced MILP. The implications are: (1) the reduced MILP formulation with investments at time zero is a valid starting point for augmented problem, and (2) imposes a valid upper bound to the optimal solution.

5.4. Bounds on the Tank Batteries

In order to find better solutions and to improve the efficiency of the mathematical formulations, an iterative procedure is developed to establish the minimum and maximum number of tank batteries among which the optimal solution lies. This reduces the feasible region of the problem, limits the combination of potential locations to be activated, and can also improves the search for an overall lower bound, as explained in the *Solution Algorithm* section.

Heuristic Determination of the Number of Tank Batteries. The minimum number of tank batteries required *MinBat* is determined by adding Eq. 38 to the MILP, where the parameter *nTB* is iteratively increased (starting at 1) until the problem becomes feasible, as shown in Figure 8. In other words, it seeks to determine the minimum number of TB at which the problem becomes feasible, regardless of the quality of the solution. Since this strategy relies on a feasibility search, many of the constraints of the original formulation can be relaxed in order to improve the procedure. Particularly, it is possible to relax pipeline sizing conditions, solving a pad-battery assignment problem, without affecting the feasibility search.

On the other hand, by imposing a minimum number of batteries to use (through the same parameter *nTB*) we can determine the number of tank batteries *MaxBat* above which the solution is guaranteed to be suboptimal (Eq. 39). Taking into consideration the best integer solution found so far, the heuristic procedure starts with a large value for *nTB* (in an extreme case, all potential tank batteries are enforced to be used), and compares the lower bound of the branch and bound search with the best solution. If the lower bound exceeds the net present cost of best solution, it is guaranteed that the total number of batteries has to be smaller than the imposed value, interrupting the execution, reducing the value of *nTB* by 1 and repeating the procedure, as observed in Figure 8.

$$\sum_{b,bz \, \in BZ_b} zbat_{b,bz} \leq nTB \qquad (38)$$

$$\sum_{b,bz \, \in BZ_b} zbat_{b,bz} \geq nTB \qquad (39)$$
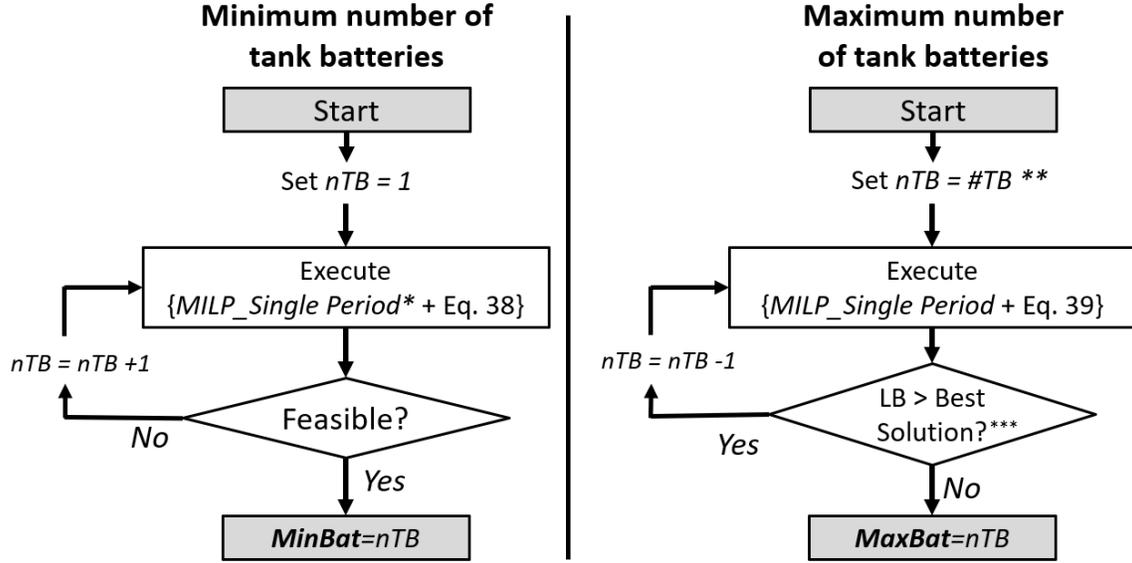
*Figure 8. Heuristic Determination of the number of tank batteries. *Adapted for feasibility search. ** #TB=total of potential TB locations. ***Checked during the first 2 hours of CPU computation for every iteration.*

5.5. Lower Bound Improvement Strategies

In this section, we introduce the concept of SuperBatteries and further topological constraints are added for cutting off suboptimal solutions. Both ideas aim to provide and guarantee better optimality gaps for the overall problem, as explained in the *Solution Algorithm* section.

SuperBatteries (*SB*) Concept. The idea is about the aggregation of the tank batteries capacities within the same geographical location, looking for a valid relaxation whose optimal solution has a much faster convergence. Through the synthesis of multiple batteries as a whole, the element *SB* replaces the tank batteries in the same site with a SuperBattery with a capacity that is an integer multiple of a single battery, which is captured by the aggregate integer variable *nBat* (see Figure 9).
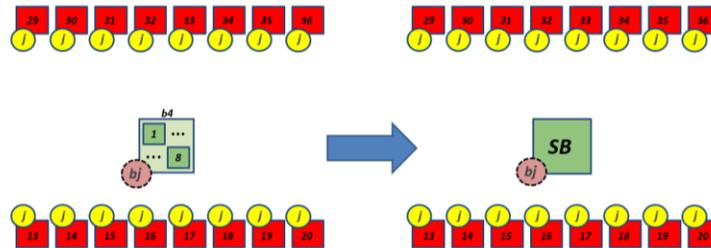
Figure 9. Original topology (left) and SuperBattery Topology (right) topology of the problem with the SuperBattery concept

We prove that this aggregation provides a valid lower bound for the original problem. Formally: (1) every solution in the original problem is also a solution of the SB relaxation (it is simply obtained by assigning single flows to the corresponding SB); and (2) the objective function of the SB relaxation is always less or equal to that in the original formulation. More specifically, from the original Eq. 5, adding up the incoming flows to the

potential tank batteries in every SuperBattery location *SB* (as in Eq. 40 and Eq. 41), it is possible to obtain the integer variable $nbat_{sb,bc}$ as the sum of the corresponding $zbat_{b,bc}$ binary variables. Clearly, the number of SuperBatteries may be significantly smaller than the tank batteries in the original formulation, which allows to: (1) significantly reduce the model size, and (2) notably decrease the problem complexity by cutting off the number of potential links between junctions and tank batteries (now *SuperBatteries)*.

$$\sum_{b\in SB}\sum_{j\in J_b} FPB_{j,b,t}^c \leq \sum_{b\in SB}\sum_{bz\in BZ_{sb}} batcap_{bz}^c \cdot zbat_{b,bz} = \sum_{bz\in BZ} batcap_{bz}^c \sum_{b\in SB} zbat_{b,bz} \quad \forall sb, c, t \in TP \tag{40}$$

$$\sum_{j\in J_b} FPB_{j,sb,t}^c \leq \sum_{bz\in BZ_{sb}} batcap_{bz}^c \cdot nbat_{sb,bz} \quad \forall sb, c, t \in TP \tag{41}$$

However, the relaxation gives rise to two important issues: first, it may lead to *almost* feasible solutions, which is illustrated in Figure 10. In that example, 11 wellpads are developed and assigned month after month to a double-capacity *SuperBattery*, but it turns to be impossible to re-assign the flows to the individual tank batteries without exceeding their single capacities. Secondly, it tends to consolidate flows from many wellpads into a single junction before sending them to the *SB*, as observed in Figure 11 (left). That solution is infeasible in terms of the original formulation since a single flow towards the SB exceeds the capacity of a single tank battery. However, Eq. 42 can be added to the relaxation to overcome this issue, which also turns out to be an excellent tightening constraint.
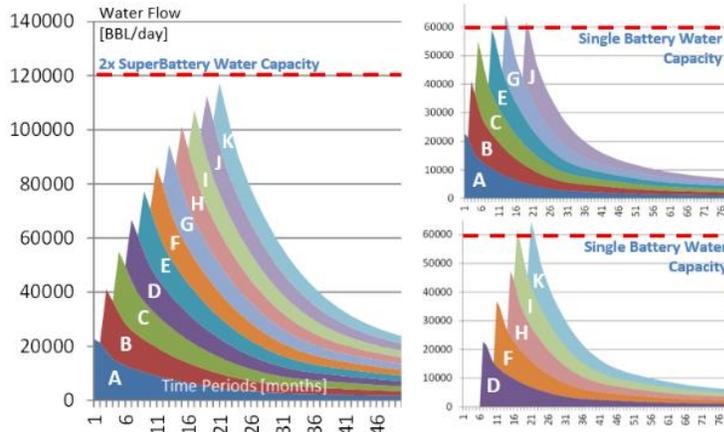
Figure 10. *Aggregate/Disaggregate assignment of wellpads when the SuperBattery relaxation is applied.*
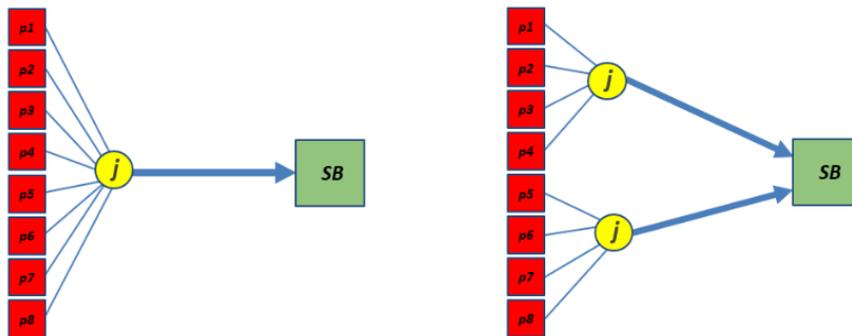
Figure 11. *Infeasible gathering of the production into a single pipeline (left), which is overcome by limiting the intermediate pipeline's capacity (right).*

$$FPB_{j,\boldsymbol{sb},t}^{c} \leq max_{bz}\{batcap_{bz}^{c}\} \sum_{d} y_{j,sb,d,t} \qquad \forall \boldsymbol{sb}, c, j, t \in TP \tag{42}$$

New Topological Constraints to Cut Off Suboptimal Solutions. It is also worth to analyze how we can gather the production in a given cluster of wellpads in order to study the viability of adding special topological constraints. First of all, it is possible to determine if the total production from a cluster of wellpads can be sent to a single battery, or, in other words, determine the minimum number of junction nodes required per cluster.
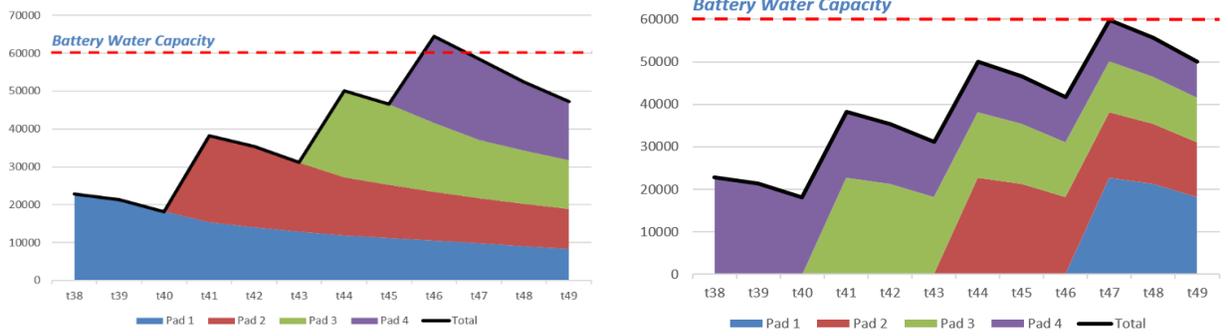
Figure 12 shows two illustrative cases gathering the production of 4 wellpads (each of them with an initial production peak of 22,750 BBL/day). At the left of Figure 12, the production start dates are in periods *38, 41, 44* and *46*, yielding an excess of 7.3% in the capacity of the largest tank battery.
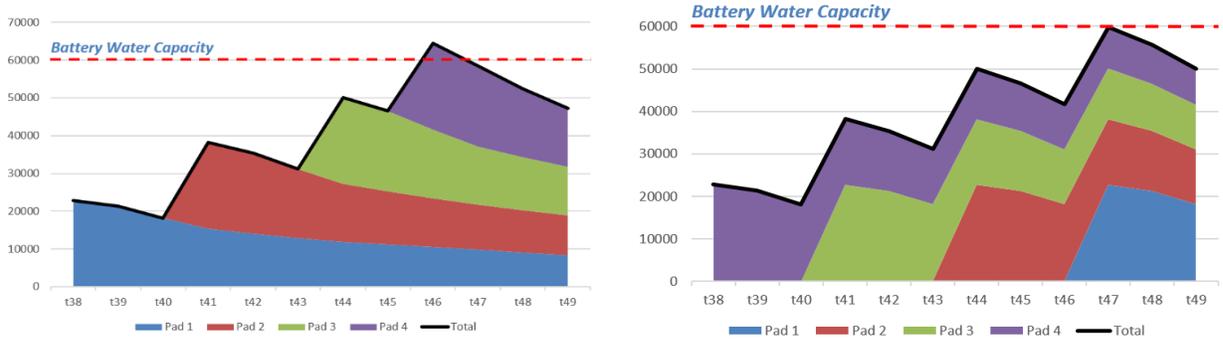
Figure 12. *Effects of gathering the complete production of a cluster of wellpad and send it to a single tank battery.*

Therefore, installing a single junction node in this cluster is an infeasible solution. Two ways to avoid this are illustrated in Figure 13. The first is by installing at least two junction nodes in the cluster, and the second is by sending the production of at least one wellpad to a junction of another cluster of wellpads. This can be generalized and applied to the SB relaxation by adding equations in two ways. The first method (Eq. 43) makes use of an auxiliary binary variable to select the strategy for every cluster k in the subset K'. The second one (Eq. 44) limits the internal pipelines connected to every junction to be fewer than the number of wellpads nw_k in k. The latter approach involves a larger number of equations when compared to Eq. 43, but it avoids introducing a further binary variable. Eq. 43 and Eq. 44 are designed for clusters with 4 wellpads (as observed in Figure 13), but the generalization is straightforward.
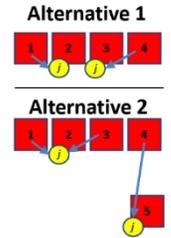
Figure 13. *Logic of topological cuts.*

$$(a) \sum_{j \in K,b,d} y_{j,b,d} \geq 2 - aux_k \ ; (b) \sum_{j \in K,b,d} x_{p,j,d} \leq nw_k - aux_k \qquad \forall k \in K' \tag{43}$$

$$\sum_{p \in K,d} x_{p,j,d} \leq nw_k - 1 \quad \forall j, k \in K'$$

(44)

5.6. Solution Algorithm

In this section we integrate all the elements previously presented in order to develop a global solution strategy capable of providing very good solutions in reasonable CPU times. The algorithm, presented in Figure 14, is based on a bi-level decomposition of the MINLP general formulation, being the NLP model used to estimate the maximum admissible flows in multiphase pipelines at the first level, and the MILP model for selecting the gathering network design in the second level. The algorithm, that includes special strategies in each step, is as follows (starting with the iteration counter at *n=0*):

First Step. The NLP formulation is solved for every potential triplet origin-destination-diameter in the superstructure, giving as a result the values of the corresponding maximum admissible flows ***maxflow$_{i,j,d}$***. Details on how to estimate pressure levels are described in Appendix C, based on conservative considerations of the pressure drop.

Second Step. A pre-processing step aims to reducing the size of the problem by cutting off infeasible combinations.

Third Step. This step aims to determine the minimum number of TB required, as an input for the following phases. It consists of an iterative feasibility search to be performed before the MILP formulation (see Figure 8).

Fourth Step. Increase the iteration counter to *n+1*. The MILP formulation is solved within a limit of H_n CPU hours of computation. The MILP formulations are, as shown in the case studies, very difficult to solve, often returning only feasible solutions. The reasons of this complexity come from the closeness with the knapsack problem in the Integer Programming field, which is proved to be NP-hard. Therefore, the strategy consists in limiting the computational time to H_n and, if the solution proves to satisfy a small tolerance criterion (optimality gap < 10^{-3}), the algorithm terminates. Otherwise, we save the best solution Z_n found so far and proceed with the next step. It is important to mention that the parameter H_n can be modified to generate many iterations in shorter CPU times (more feasible solutions to reduce the size of the model in future instances), or fewer iterations in longer times, searching for more efficient solutions.

Fifth Step. Determine the maximum number of TB required according to the best solution found so far Z_n, as shown at the right of Figure 8.

Sixth Step. By introducing the concept of SuperBatteries and new topological constraints, a relaxed problem *RP* is solved to optimality (Z_n^{*RP}). We search for tighter lower bounds on the minimum net present cost to improve the optimality gap and the MILP execution performance. After this step, a new iteration can be performed by returning to step 4.

Solution Algorithm

Initialize
Iteration $n=0$
lower bound $LB_0 = -inf$ | upper bound $UB_0 = inf$
$MinBat_0 = 1$ | $MaxBat_0 = \#TB$
best solution $Z_0 = inf$

↓

Solve NLP Formulation for every potential
triplet in the superstructure. Set $MaxFlow_{i,j,d}$

↓

Pre-processing
Cut-off infeasible combinations

↓

Determine bound $MinBat_0$ (feasibility proof)
on the number of TB to install → set LB_0

↓

Iteration $n=n+1$

↓

Solve MILP formulation with a limit of H_n CPU hours
of computation → **New global solution** Z_n

↓

Is Z_n optimal?* — Yes → **End** $Z_n = Z^*$

No ↓

Determine bound $MaxBat_n$ (against Z_n) on the
TB to install by the heuristic procedure

↓

Lower Bound improvement strategy
SuperBattery and Topological Constraints.
Solve the relaxed problem (**RP**) applying $MinBat_n$ and
$MaxBat_n$ parameters. Solution: Z_n^{*RP} → $LB_n = Z_n^{*RP}$

Figure 14. *Solution algorithm in detail. *Optimality Gap=0%*

In summary, a special strategy has been proposed to obtain very good solutions and, at the same time, close up the optimality gap into reasonable values, all in moderate computational processing time, as can be observed in the next section. Future works are aimed to improve and expand the scope of the algorithm.

6. Results and Discussion

Two case studies are addressed in this section to assess the capabilities of the proposed mathematical formulation and solution strategy to design gathering networks for unconventional production. All the case studies are coded in GAMS 30.3 and solved using IPOPT 30.3 for the NLPs, and GUROBI 9.0.1 for the MILPs, on an Intel Core i7 8th gen with 16gb RAM under deterministic, non parallel mode. Based on information from our industry partner, additional assumptions are made on both instances: (1) capital and operating costs of junction nodes are

neglected, and (2) operating costs for the pipelines and tank batteries are assumed to be constant terms in the objective function. The latter is justified by the fact that the total volume to be transported is fixed by the well development plan, and pressure drops are set at specific values for every point in the network according to the solution strategy.

Note that, for both case studies, the production of every wellpad reaches its peak as soon as it starts. In this case, the subsets of time periods *TP* and *TS* coincide. If the production profiles were different, the constraints should be expressed for an extended set of periods *TP*.

6.1. Case Study 1: Small Illustrative Example

The proposed approach is first applied to a small illustrative example to show the major trade-offs of the problem. As seen in Figure 15, the first case study comprises 12 wellpads with their starting production dates, 2 potential locations for tank batteries (one of them allowing up to 2 single batteries), two battery junction nodes, and a centralized delivery point (CDP). Tank batteries of two sizes *bz1* and *bz2* are available, whereas alternative pipeline diameters of 8, 12, 16 and 20 inches are considered. Table 1 shows the capital expenditures for both pipelines and tank batteries, along with its processing capacity per component.

All the wellpads are considered identical in terms of their production curves (Figure 16). The gas-to-oil and water-to-oil ratios *gor$_p$* and *wor$_p$* are 2 and 3.5, respectively, being constant for all the wellpads throughout the entire time horizon. Regarding the junction nodes, every wellpad is proposed as a potential location for them. The potential locations for tank batteries include a battery junction from where the production can be transported by single phase pipelines to the CDP. All the potential connections among wellpads, junction nodes and batteries are allowed in this case study. More detailed information for this problem is given in the supplementary material.
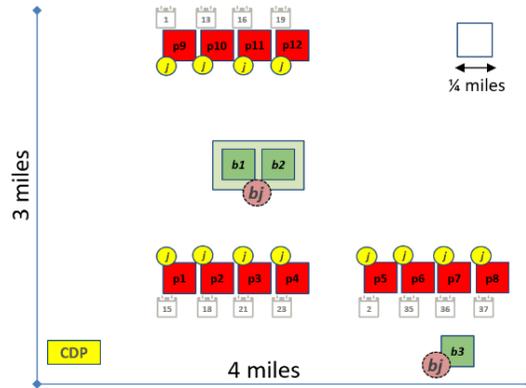
Figure 15. *Case Study 1 Superstructure: wellpads locations, scheduled production start dates, potential junctions and tank batteries.*

According to the solution algorithm, the pipeline transportation capacity (***maxflow***) is computed by solving the NLP formulation for every potential connection wellpad-junction and junction-battery, considering alterative pipeline diameters. The NLP models are solved by using GAMS-IPOPT 30.3 in less than 1 CPU second per instance, amounting to 1080 CPUs for all possible combinations. The parameters for these calculations are also given in the supplementary material. After pre-processing (Step 2), a significant reduction in the number of variables is obtained. For instance, consider the pad *p5* in Figure 15. If the model decides to link *p5* with *p1* or *p2*, a 16" diameter pipeline will be required, while a 12" pipeline will be large enough to connect *p5* with *p3, p4, p6, p7* or *p8*. Then, for any other combination, the binary variables $x_{p,j,d}$ are set to zero. In the case of pipelines

25

connecting junctions with tank batteries, 20" pipelines are also blocked because the capacity of the 16" pipelines is large enough to transport more production than the flow that can be processed by the tank batteries in this case.
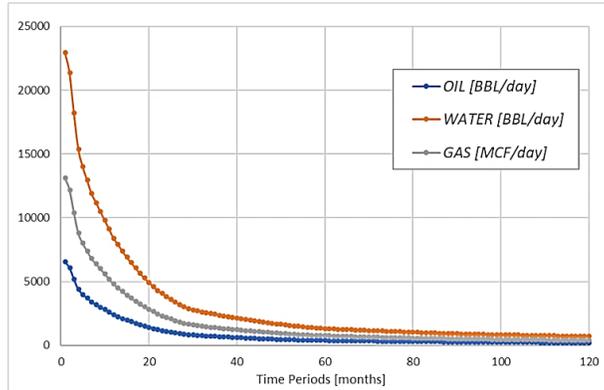
Figure 16. *Characteristic production curve for wellpads in Case Study 1*

Table 1. *Case Study 1 Costs and Parameters*

Tank Battery Capacity	Capital Expenditures [MM USD]	Water capacity [BBL/day]	Oil Capacity [BBL/day]	Gas capacity [thousand cf/day]
bz1	10	30,000	10,000	30,000
bz2	20	60,000	20,000	60,000
Pipeline Diameter [inches]	**8**	**12**	**16**	**20**
Capital Expenditures [10^3 USD/mile]	201.72	819.49	1954.16	3053.36

Investments at the Initial Time. MILP Results. The MILP model with investments at the initial time comprises 4,157 single equations, 2,348 nonnegative variables and 265 binary variables after the solver pre-processing, and is solved to optimality in 218.5 CPU seconds. The optimal solution amounts to a total investment of 53.449 MM USD, while the relaxed LP reports a solution of 43.33 MM USD (a 18.9% integrality gap). From the total investment, 40 MM USD are related to TB and the remaining amount to pipelines. The net present cost of this solution, assuming a 10% annual discount rate, turns out to be 52.502 MM USD. No further steps of the solution algorithm are required because of the optimality achieved at the first iteration.

The optimal network design is shown in Figure 17, where it is possible to note that: (a) two tank batteries of capacity *bz2* are installed, both of them in the central node, collecting the production from 6 junction nodes, (b) the optimal network of pipelines it is not composed of the "shortest-path" options (some crossings are evident) since they are driven by the usage of tank batteries; (c) all of the multiphase pipelines between wellpads and tank batteries are of 12" of diameter; and (d) the single-phase pipelines after separation are of 8" for gas, 12" for oil and 16" for water.
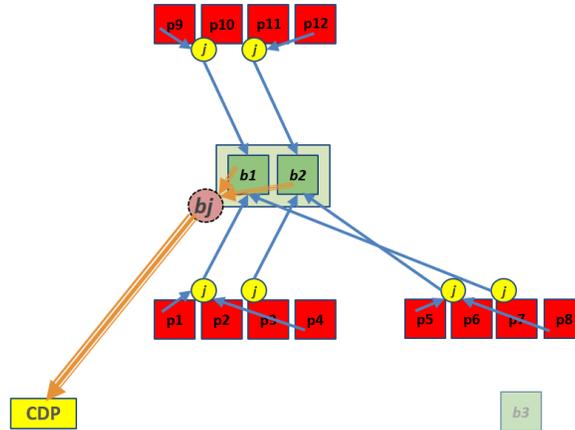
Figure 17. *Case Study 1 Network Solution.*

Figure 18 shows the tank battery utilization profiles for the water component, the most restrictive one. Both batteries are used from the early stages, with a clear underutilization until the 16th period, which could indicate an opportunity for improvement when considering the decisions over time. However, from the 16th period onwards, dealing with production peaks is not trivial, as seen in Figure 18. More detailed results including individual flows over the time horizon are shown in the supplementary material.
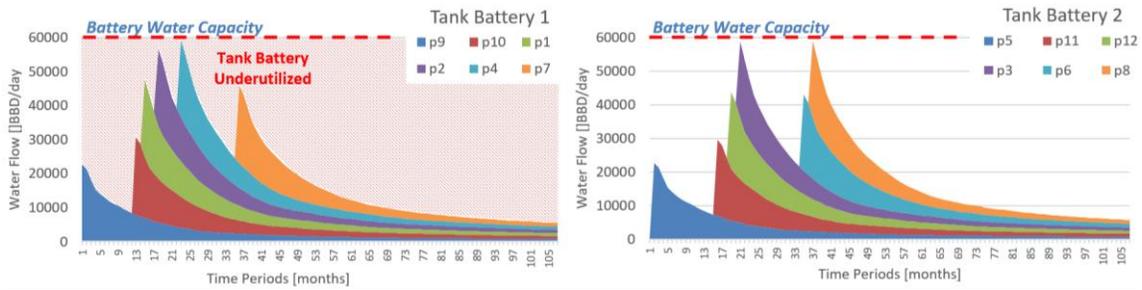
Figure 18. *Utilization of TB 1 (left) and TB 2 (right) for processing water flows (most restrictive component) distinguishing between the wellpads being linked throughout the planning horizon.*

Multiperiod MILP Formulation. When solving the multiperiod investment model, it is possible to place the investments in any period along the planning horizon. In this case, the total present cost improves by 2%, being reduced to 51.472 MM USD. The computational performance of the model clearly deteriorates because of the growth in the model size. The model comprises a similar number of continuous variables, but equations and binary variables increase by an order of magnitude. The total time spent to close the optimality gap to 0% is 12,887 CPUs (3.58 h), sixty times the CPU time for the initial time approach. As in the previous case, the optimal solution is found, and no further steps are required in the solving strategy.

Figure 19 depicts the new network obtained, while Figure 20 shows the tank battery utilization for water flows. Note that the new solution delays by 13 periods the installation of the second tank battery. As observed in Figure 19, even though the network of pipelines is crisscrossed and tangled, which increases the pipeline investment costs, the reduction in the total net present cost of the second tank battery is more significant.

27
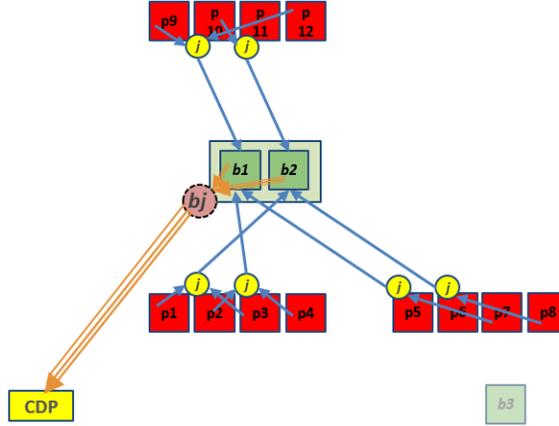
Figure 19. *Multiperiod Solution for Case Study 1*
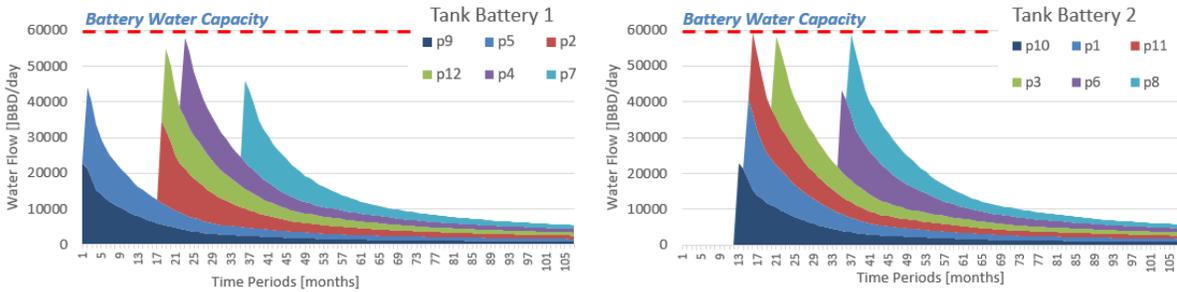
Figure 20. *Utilization of TB 1 (left) and TB 2 (right) in the multiperiod optimal solution. For simplicity, only water flows (the most restrictive component) are represented, differentiating among the wellpads connected to every TB.*

6.2. Case Study 2: Real World Problem

This case study is a larger size problem from the O&G industry with a challenging superstructure. As seen in Figure 21, it comprises 40 wellpads with their corresponding scheduled production start dates over a planning horizon of 50 months. Potential locations for tank batteries are 5, 3 of which account for a single unit, and the remaining 2 for sites where a maximum of 4 units are allowed. Also, a potential junction node is assumed in each wellpad. Finally, the set of pipeline diameters and tank battery capacities considered, as well as their related installation costs, are similar to the ones in Case Study 1 (see Table 1). Large tank batteries are the only allowed for all the potential locations, except for node *b5* where only a small tank battery can be installed. Given that the *gor* and the *wor* are assumed to be constant for all the wellpads and over the whole time horizon, we focus on water flows, the most restrictive component according to the tank batteries capacities. The production profile of every wellpad is similar to the one predicted for Case Study 1. Figure 21 also shows the clusters of wellpads in the area (*c1* to *c8*), each of them comprising from 4 to 8 wellpads. According to the solution algorithm, the first step is to solve the NLP formulations in order to determine the maximum pipeline transportation capacities $maxflow_{i,j,d}$ for every triplet *i-j-d* in the superstructure of alternatives for the pipeline network design. For every segment, the NLP formulation involves 21 variables and 15 constraints. By using GAMS-IPOPT 30.3, the maximum admissible flow for every pipeline segment is found in less than 1 CPU second, solving a total of 2,000 NLPs in about 15 minutes of CPU). The pre-processing stage is applied afterwards, and the heuristic procedure determines that a feasible solution requires at least 6 tank batteries (see Table 3).
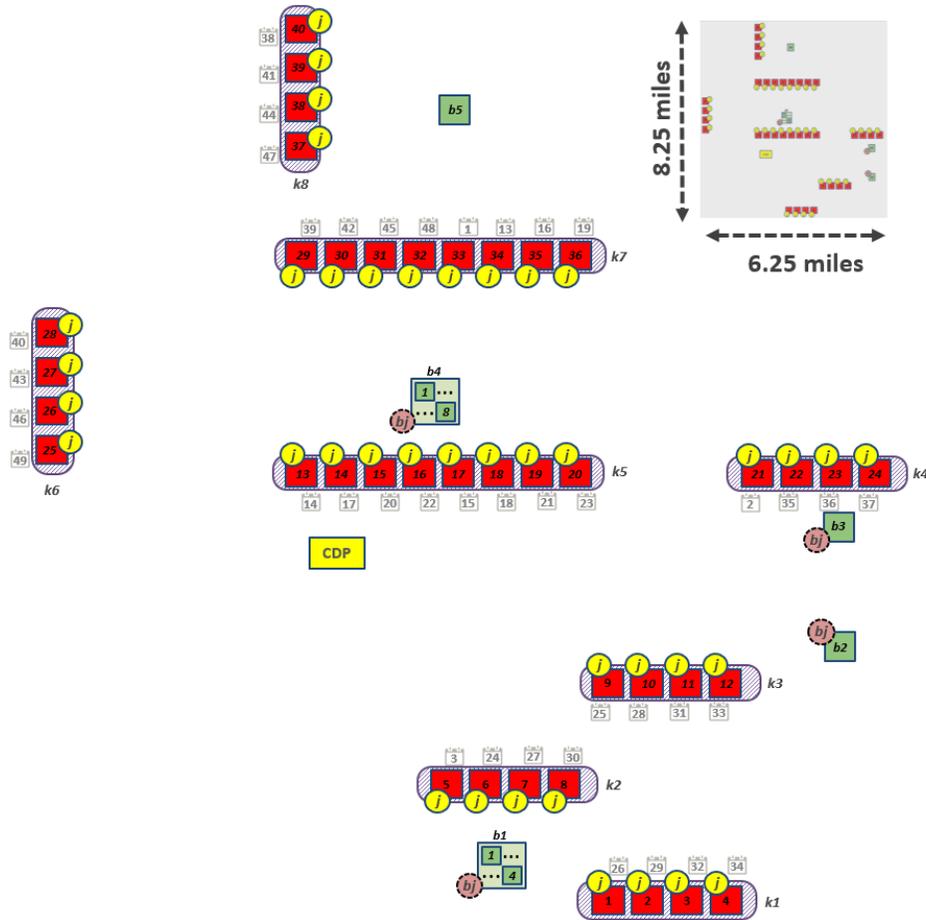
Figure 21. *Case study 2 superstructure wellpads locations, scheduled production start dates, potential junctions and TB.*

Next, the MILP formulation with investments at the initial time is solved with a limit of 10 hours of CPU. The MILP model is applied to two different instances. The first one, assuming no limitations in the potential connections among the nodes, resulting in a model with a total of 29,441 equations, 24,600 positive variables and 2,250 binary variables (after solver pre-processing). The second instance is a variant assuming a reduced connectivity among the elements of the network. Table 2 shows the possible connections among the elements of the network, which have been selected with the premise that two extremely distant elements (like connections between wellpads and junctions in opposite positions) are unlikely to be part of the optimal solution. Also, the potential number of junctions are reduced by half, allowing at most 2 nodes every cluster of 4 wellpads. The model size for the second instance comprises 8,323 equations, 6,062 positive variables and 488 binary variables.

Table 2. *Connectivity between clusters of wellpads and with tank batteries. A number 1 means available wellpad-junction connections between clusters (left) and allowed tank batteries to receive production from a junction in a cluster (right).*

Cluster	k1	k2	k3	k4	k5	k6	k7	k8
k1	1	1	1					
k2	1	1	1					
k3	1	1	1	1	1			
k4				1	1			
k5				1	1			
k6						1	1	1
k7							1	
k8							1	1

Cluster/Battery	k1	k2	k3	k4	k5	k6	k7	k8
b1	1	1	1	1		1		
b2	1	1	1	1	1			
b3	1	1	1	1	1	1	1	
b4	1	1	1	1	1	1	1	1
b5							1	1

For the first instance, after 3 hours of CPU the model reports a solution of 185.94 MM USD, with a 25.4% optimality gap, which is reduced to 21% after 10 hours of computation. It is worth to mention that the best solution amounting to 176.75 MM USD is found in the last 100 seconds, which highlights the difficulty of finding better integer solutions. In the second instance, despite not improving the solution significantly (176.56 MM USD) the network design is found after 1 hour of CPU and the MILP solver reports an optimality gap of 18.2% after the 10 CPU hours. In both cases, 120 MM USD are the capital expenditures in tank batteries and slightly more than 56 MM USD are paid for pipelines. The total present cost of the capital expenditures in the network is 157 MM USD, considering a 10% annual discount rate. Regarding the LP relaxation, an integrality gap of 29% is observed when compared with the best solution found. To improve this value through tighter formulations is the key to achieve a better computational performance.

The gathering network suggested by these solutions, which is shown in Figure 22, comprises a total of 6 tank batteries, 4 in the central location, 1 in the south and 1 in the east of the map. Two junction nodes are planned for almost every cluster of wellpads, whereas in the case of the cluster $k2$ on the south, only one junction node is required. Pipelines of 12" are the most common option, while 16" pipelines are installed to satisfy pressure drop constraints for longer distances. Regarding the single-phase pipelines, 8" high pressure (HP) pipelines are used for natural gas, 20" for water and 16" for the oil phase. It is important to note that a rather counterintuitive pipeline links a junction node in the west with the southern tank battery, yielding an investment of more than 10 MM USD. Guided by a simple analysis, despite the high cost of this pipeline, it is more convenient than adding a new tank battery. However, there is an open question about its need, since the optimality gap is still large. Moreover, it is interesting to observe how many wellpads finally send their production to every tank battery. For example, tank batteries $b1$ and $b3$ gather the production from 7 wellpads.
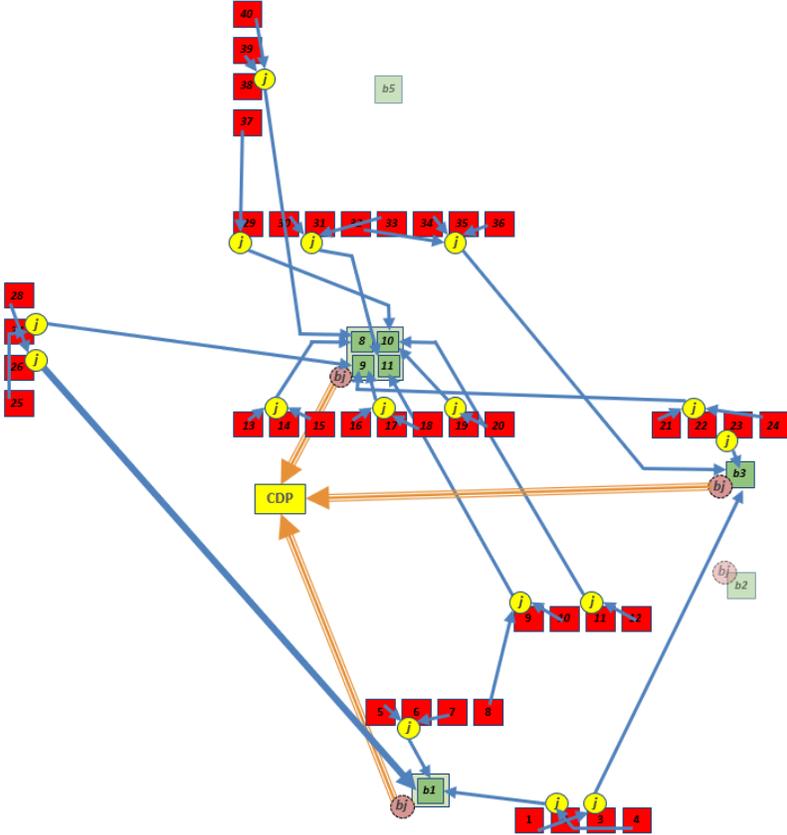
Figure 22. Network obtained by general formulation instances (a) and (b).

SuperBattery Relaxation. When solving the relaxed version of the model considering the notion of SuperBatteries, a 0% optimality gap is reached after 732 CPU seconds. The model comprises 4,285 continuous variables, 588 binary variables and a total of 6,734 constraints. The relaxed solution of 149.2 MM USD yields a valid lower bound for the overall MILP formulation previously addressed. The solution is obtained in 1.2 CPU hours and is represented in Figure 23. Even though this bound reduces the optimality gap to 15%, it is still a rather poor quality measure.
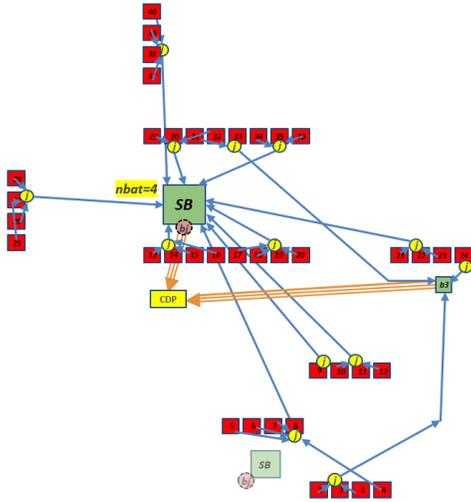
Figure 23. Solution for the SB relaxation
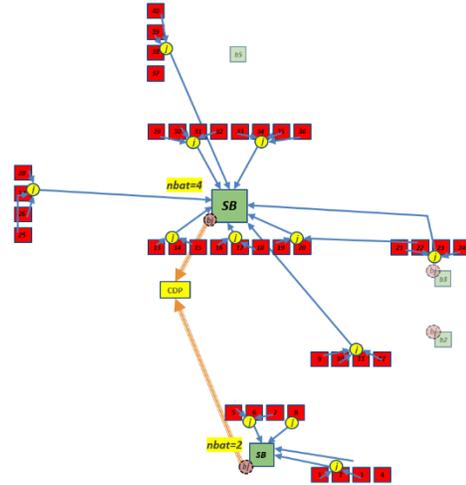
Figure 24. SuperBattery RP solution with the addition of the MinBat and MaxBat parameters

Limits on the Tank Batteries and Lower Bound Improvement Strategy. Back to the optimization algorithm, as the MILP solution has not closed the optimality gap to 0%, the next step is to determine the MaxBat parameter. As seen in Table 3, the heuristic procedure determines that the number of TB to install in the optimal solution will finally have a total of 6 or 7 batteries.

*Table 3. MinBat and MaxBat determination by the heuristic procedure. *value adopted for the corresponding parameter.*

MinBat Determination			MaxBat Determination		
MaxBat Value	**Report**	**CPU Time [seconds]**	**MinBat Value**	**LB found [MM USD]**	**CPU Time [seconds]**
3	Infeasible Instance	1.92	10	218.63	12
4	Infeasible Instance	2.28	9	198.6	20
5	Infeasible Instance	25	8	178.63	17
6*	Feasible Solutions Found	310	7*	171	7200

Based on the *MinBat* and *MaxBat* parameters, the next step of the algorithm is executed, forcing the installation of at least 6 batteries, but no more than 7. Hence, the problem converges to optimality (10^{-9} optimality tolerance) after 7.75 hours of computation, yielding a solution of 162.61 MM USD. This new valid LB for the MILP problem represents a better and tighter bound, 7.9% below the best solution found so far. As seen in Figure 24, the solution involves the installation of 2 superbatteries, one composed of 2 single TB and the other of 4.

MILP at the second iteration. The next iteration is executed (*n=2*) by solving again the MILP formulation including the updated information about *MinBat*, *MaxBat,* lower and upper bounds. After 10 hours of CPU, a substantial improvement of the solution is achieved, reaching a total cost of 171.7 MM USD. A global reduction of 4.8 MM USD (6.6 MM USD in terms of the present cost) is achieved, representing a relative improvement of

2.7%. In terms of optimality gap, note that this solution shows a 5.2% of optimality gap when compared with the lower bound from the SB relaxation.

The new gathering network obtained is shown in Figure 25, comprising the same number of tank batteries in the same locations, but notably improving the pipelines network, which is the driver of the total cost reduction. The most important aspect to highlight is the elimination of the counterintuitive pipeline linking the junction node in the west with the southern tank battery. The rest of the results can be observed in the supplementary material.
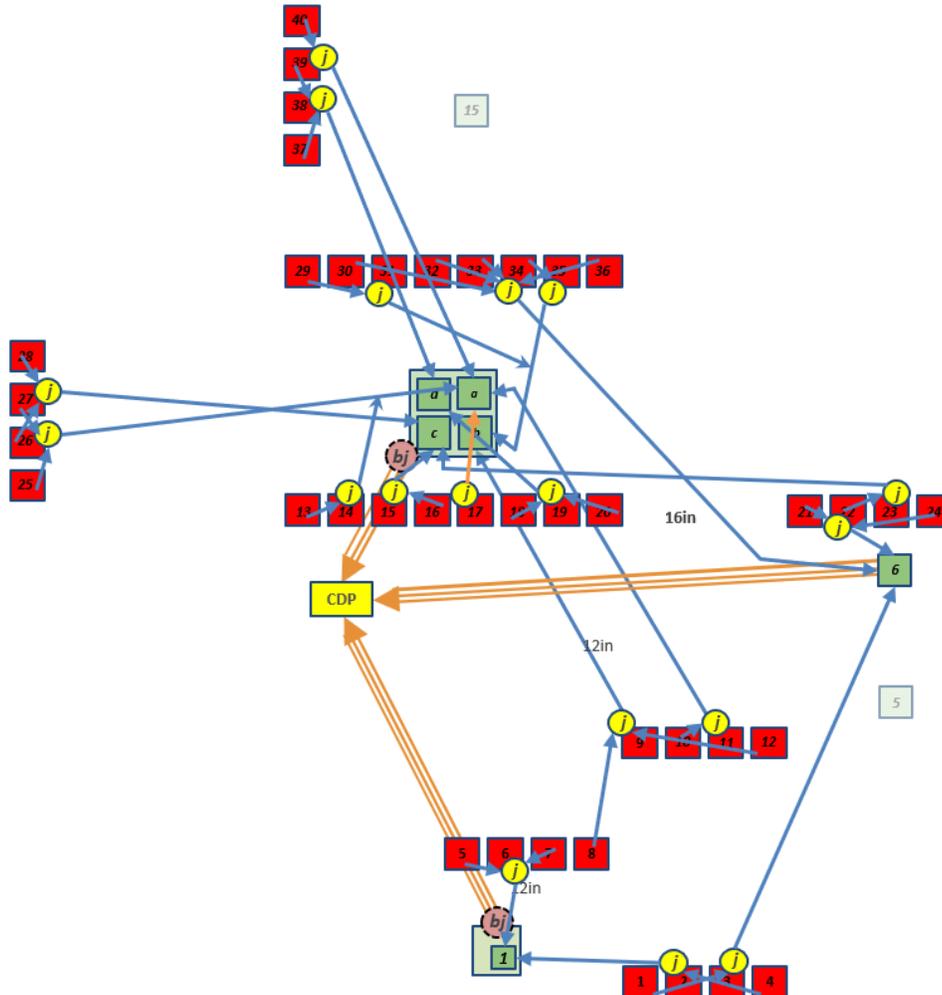
Figure 25. *Gathering network obtained after solving the MILP formulation in the second iteration.*

Final Analysis. A real world case study has been addressed by formally applying the solution algorithm proposed by this work, combining a bi-level decomposition into an NLP-MILP sequence together with tailored strategies to improve the solutions and reduce de optimality gap. The procedure has been able to reduce the gap to a 5.2%, which, despite being much lower still does not guarantee that there are no better alternatives for the unconventional gathering network design.

7. Conclusions and Future Work

We have developed a rigorous approach for the design of supply chain networks for gathering the production of unconventional oil and gas, yielding a complex MINLP formulation. The complexity is due to the integration of

empirical correlations and industry guidelines to model the oil, gas and water multiphase flows of the unconventional production, within the network infrastructure design and planning problem. The model takes into account infrastructure capacity limitations of tank batteries and the multi-period nature of the problem, required to address steep-declining production profiles and steady turn-in-line of new wellpads. Overcoming the problem complexity is one of the major contributions of this work, facing the challenge of developing an optimization framework to obtain efficient solutions in reasonable times, also providing a valid reference on how much it might be improved (optimality gap).

An iterative solution algorithm, based on a bi-level decomposition and lower bound improvement strategies, has been developed and tested both in an illustrative case study and in a real world problem, to determine the number, size location and period of installation of tank batteries and pipelines to gather the production from multiple wellpads to centralized delivery points. The real world case study, an extremely challenging instance comprising 40 wellpads over a planning horizon of 50 months whose production has to be routed to tank batteries to be located within 5 potential locations. After executing the solution algorithm, a near optimal solution is obtained, yielding an efficient network configuration. These results clearly show the potential of the proposed optimization framework.

It is important to mention that by assuming minimum pressures at the nodes, the transportation capacity of multiphase pipelines is conservatively computed by solving a large number of NLP models, accounting for the Lockhart-Martinelli correlation and SPE guidelines for pipeline sizing. This implies that the worst conditions for pressure losses are imposed for every pipeline segment, resulting in solutions that are an upper bound on the minimum total cost for the overall MINLP.

Future work is to assess the impact of deferring the start of the well production, which could lead to a reduction in the net present value of the incomes, but may significantly improve the usage of the TB capacities. The integration of the wellpads development planning together with the supply chain network design optimization configures an extremely challenging problem. On the other hand, we aim to overcome some of the limitations of the proposed optimization framework, to enhance the convergence of the solution algorithm.

Acknowledgments

Appendix A: Maximum Flow in Single Phase Pipelines

The objective of this appendix is to deepen the calculations needed to estimate the maximum flow in single phase pipelines, both for the liquid and the gas phase.

Liquid Phase Pipelines. One of the most important objectives when designing liquid phase pipelines is to avoid the causes generating corrosion, erosion, water hammer effect, among other negative phenomena. One of the specific actions in this sense is to impose a maximum mean velocity, which is set to 1.5 m/s (Society of Petroleum Engineers, 2006). Then, according to typical flow computations, Eq. A1 determines the maximum flow (in m³/s) for oil and water phases when linking tank batteries with battery junction nodes and these with central delivery points by a pipeline of diameter $diam_d$ (in m). The maximum flow is directly proportional to the pipeline section.

$$maxflow_d^{oil/water} = vmax_{LP} \cdot \frac{\pi}{4} \cdot \left(Diam_d^2\right) \tag{A1}$$

Gas Phase Pipelines. The compressibility of the gas phase makes the calculation of the transportation capacity much more complex. Similar to Cafaro & Grossmann (2014) and (Duran & Grossmann (1986), assuming first a mixture of ideal gases being transported, the head loss in a given pipeline segment with diameter $diam_d$ (in m) is assumed to follow the Weymouth correlation (Weymouth, 1942) for gas pipelines sizing, represented by Eq. A2 and A3. The value of the gas density ρ^{GP} is 0.729 kg/m³ for shale gas in standard conditions (*P0*=0.1013 MPa; *T0*=288.9 K). *T* is the average gas temperature (fixed at 288.9K) and the value of α is 3/16. Input and output pressures (*P_i* and *P_j*, in MPa) are variables in the case of the general formulation, whereas in the proposed optimization framework are assumed to be known (see Appendix C). The pipeline segment length is given by *dist* (in km). Merging A2 and A3, the final Eq. A4 represents the gas flow computation (in 10⁶ m³/day).

$$diam_d = dist^\alpha \left(P_{in}^2 - p_{out}^2\right)^{-\alpha} B^\alpha \tag{A2}$$

$$B^\alpha = \rho^{GP} T [P_0/0.375 T_0]^2 \, maxflow_d^{gas} \tag{A3}$$

$$maxflow_d^{gas} = dist^{-0.5} \cdot diam_d^{2.667} \left\{ \frac{\left(P_{in}^2 - p_{out}^2\right)}{\left(\rho^{GP} T [P_0/0.375 T_0]^2\right)} \right\}^{0.5} \tag{A4}$$

As described in the general mathematical formulation, the $maxflow_d^{gas}$ is proportional to: (1) the pipeline diameter raised to the power of 2.667, and (2) the difference of the inlet and outlet pressures individually squared. Then, Eq. A5 represents the final flow computation according to the assumptions made.

$$maxflow_d^{gas} = \frac{\mu^{gas}}{dist^{gas}} \left(P_{in}^2 - p_{out}^2\right) diam_d^{2.667} \quad \forall d \tag{A5}$$

Appendix B: Examples of Multiphase Maximum Flow Computation with the NLP Formulation

A group of illustrative examples is developed to show the NLP formulation results and its understanding when applied to a specific pipeline segment with given minimum inlet and outlet pressures.

Illustrative Example 1. To illustrate the general behavior of the NLP formulation to estimate the pipeline transportation capacity, assume two given nodes: (1) the origin representing a wellpad with a pressure of 250 PSI (1.72 MPa), connected by a pipeline of diameter *d* and length *dist* with (2) a junction node with an outlet pressure of 200 PSI (1.37 MPa) (see Figure B1). The NLP formulation is solved assuming ten different segment lengths: 250, 500, 750, 1000, 1250, 1500, 2000, 3000, 4500 and 6000 meters. *gor* and *wor* values are fixed at 2 and 3.5 respectively, whereas the pipeline diameters considered are of 8", 12" and 16".
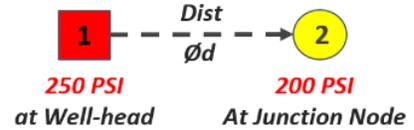
Figure B1. *Pressures for the Example 1*

Figure B2 shows the results after solving the NLP formulation (using IPOPT as solver) for the maximum flow determination. For every diameter, the maximum flowrate is a decreasing function with the length. Some conclusions can be derived from analyzing the data. First, it is possible to observe an important increase in the maximum flowrate with the diameter. Moreover, for short lengths, the maximum flowrate is large, which decreases rapidly when increasing slightly the distance, while, for larger distances, the reduction in the maximum flowrate is not so notorious. For every diameter, the SPE guideline generates a unique result independent of the distance, which is evident in the totally horizontal segment of each curve. This indicates that, for short distances, the SPE guidelines are more restrictive than the Lockhart-Martinelli (LM) correlation (dotted lines). The red circles represent the distance at which LM determines the maximum admitted flow. Then, it is possible to conclude that, for longer distances, the LM procedure is more restrictive because of being based in the head losses, which are larger with the pipeline length, while for short distances SPE guidelines impose because of limiting the maximum velocity to avoid damage to the infrastructure. Finally, it is observed that the SPE guidelines become more restrictive as the diameter increases.
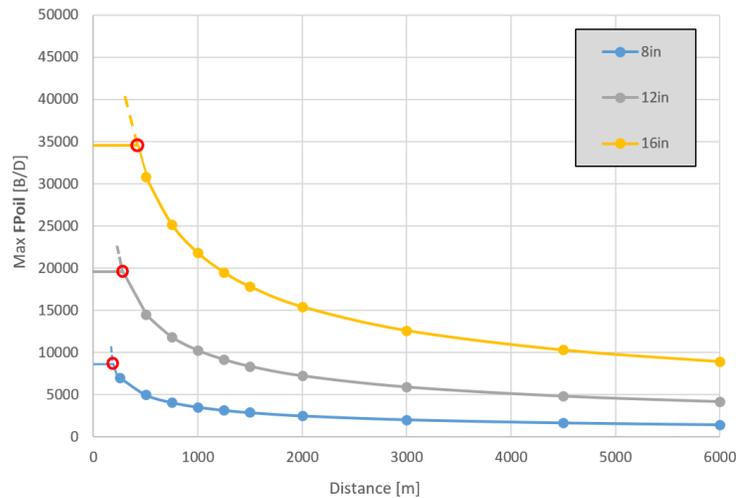
Figure B2. *Example 1 results for different pipeline diameters. As representative of the multiphase flow, oil phase maxflow is reported.*

Illustrative Example 2. Keeping the first example configuration, we only change the *gor* and *wor* values with the aim of study the effect of the phases ratio. The three scenarios studied were: (1) *gor* =2 and *wor* = 3.5 (Phases Ratio = 444 [ft^3/BBL]), (2) *gor* =3 and *wor* = 2,(Phases Ratio = 1000 [ft^3/BBL]), (3) *gor* =4 and *wor* =2 (Phases Ratio = 1333 [ft^3/BBL]). Figure B3 below shows, for an 8in diameter pipeline, the comparative **MaxFlows** obtained for each individual phase and the corresponding total flow. Being a predefined parameter, the total head losses are constant for all the cases, which makes possible to observe that the maximum total flows are rather similar, with a slight difference in favor of the 1st scenario (lowest gas/liquid ratio). It is interesting to note that, for the 1st and 2nd scenarios, the oil maximum flow are practically similar with the corresponding variations in gas and water results. However, in the 3rd scenario, in order to achieve the greater proportion of gas in the flow, the oil flow decreases.
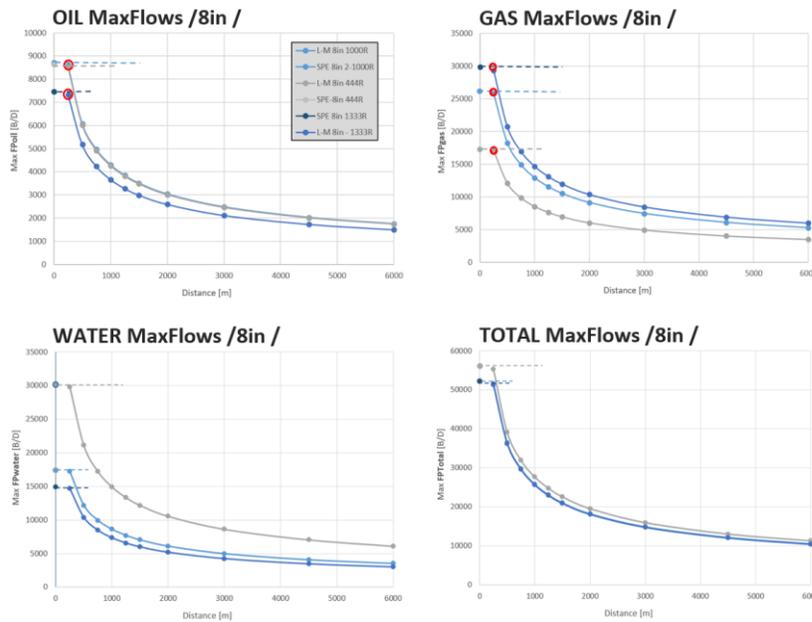
Figure B3. *Results for the Example 2.*

Next, we study the behavior of the models when varying the *gor* and *wor* toward extreme values, in particular a *gor* =10 and a *wor* = 0.5. When comparing with the original case study (*gor* =2, *wor* = 3.5), it is observed in the Figure B4 below how the maximum oil flow decreases at the expense of a larger increase in the gas flow. Additionally, the break-even point is slightly moved to the right, meaning that the SPE model imposes a maximum bound for the flows that is valid for larger segment distances before changing to the LM bounds.
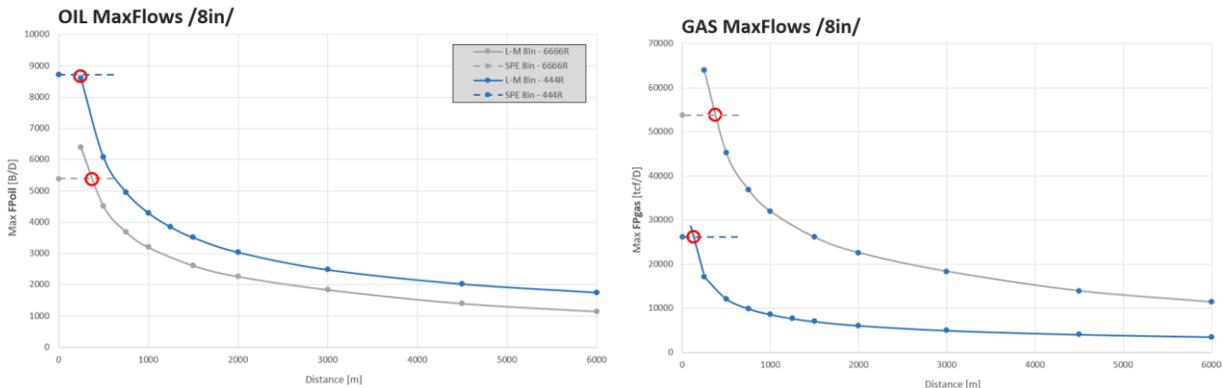
Figure B4. *Results for the Example 2 when varying **gor** and **wor** to extreme values.*

36

Appendix C: Pressure Estimations for the Case Studies

The values in Figure C1 are the minimum pressures admitted for every type of node in the system when the optimization framework is performed. Consider that, for the general mathematical formulation, the unique known pressure is the one at the wellpads, being the rest variables of the formulation. We assume that the pressure at the wellheads is high enough to make oil, gas and water flow towards the tank batteries. Even though in certain cases the pressure reaches 650 PSI (4.48 MPa) a conservative value of 250 PSI (1.72 MPa) is assumed at the outlet of the wellpads. The inlet pressure at tank batteries has to be at least of 80 PSI (0.55 MPa), resulting a maximum allowable pressure drop of 170 PSI. The distance between wellpads and junctions will be short, leading to set the maximum pressure drop to 50 PSI (0.34 MPa) in the potential link. From junction nodes to batteries the pipelines operate at medium-low pressures, being the maximum head loss equal to 120 PSI (0.82 MPa). After the stream flows are separated in the tank batteries, the pressure is assumed to drop to 40 PSI at the tank battery junctions, where pumps and compressors push the fluids to the CDP at high pressure conditions. The pressure after the gas compressor reaches 1100 PSI, which is assumed to drop to 580 PSI at the centralized delivery point. As a result, if pipelines are not used at maximum capacity in some periods, the inlet/outlet pressure at each node can be larger but never smaller.
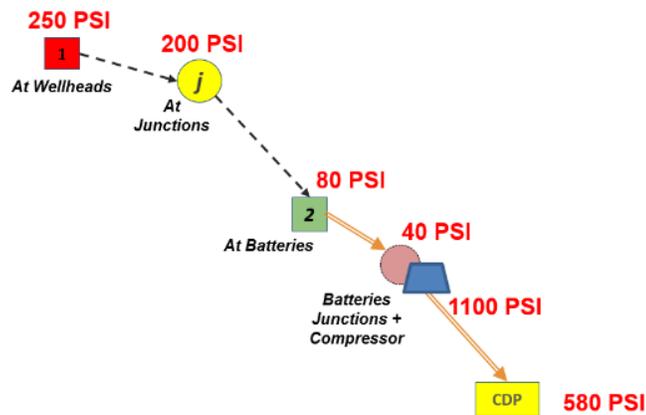
Figure C1. *Fixing Pressures at the Nodes of the Network*

References

Accenture, 2014. Achieving high performance in unconventional operations.

Al-Hadhrami, L.M., Shaahid, S.M., Tunde, L.O., Al-Sarkhi, A., 2014. Experimental study on the flow regimes and pressure gradients of air-oil-water three-phase flow in horizontal pipes. Sci. World J. 2014. https://doi.org/10.1155/2014/810527

Alvarez, L.P., Mohan, R.S., Shoham, O., Avila, C., 2010. Multiphase Flow Splitting in Parallel/Looped Pipelines, in: SPE Annual Technical Conference and Exhibition. Society of Petroleum Engineers. https://doi.org/10.2118/135723-MS

Azzopardi, B.J., Smith, P.A., 1992. Two-phase flow split at T junctions: Effect of side arm orientation and downstream geometry. Int. J. Multiph. Flow 18, 861–875. https://doi.org/10.1016/0301-9322(92)90064-N

Baihly, J., Malpani, R., Altman, R.M., Lindsay, G., Clayton, R., 2015. Shale Gas Production Decline Trend Comparison Over Time and Basins Revisited, in: Proceedings of the 3rd Unconventional Resources Technology Conference. American Association of Petroleum Geologists, Tulsa, OK, USA. https://doi.org/10.15530/urtec-2015-2172464

Cafaro, D.C., Grossmann, I.E., 2014. Strategic Planning , Design , and Development of the Shale Gas Supply Chain Network. AIChE J. 60, 2122–2142.

Dbouk, H.M., Hayek, H., Ghorayeb, K., 2020. Modular approach for optimal pipeline layout. J. Pet. Sci. Eng. 107934. https://doi.org/10.1016/j.petrol.2020.107934

Drouven, M.G., Grossmann, I.E., 2017. Mixed-integer programming models for line pressure optimization in shale gas gathering systems. J. Pet. Sci. Eng. 157, 1021–1032. https://doi.org/10.1016/j.petrol.2017.07.026

Drouven, M.G., Grossmann, I.E., 2016. Multi-Period Planning , Design , and Strategic Models for Long-Term , Quality-Sensitive Shale Gas Development 62. https://doi.org/10.1002/aic

Duran, M.A., Grossmann, I.E., 1986. A mixed-integer nonlinear programming algorithm for process systems synthesis. AIChE J. 32, 592–606. https://doi.org/10.1002/aic.690320408

Edwards, L., Dhanpat, D., Chakrabarti, D.P., 2018. Hydrodynamics of three phase flow in upstream pipes. Cogent Eng. 5. https://doi.org/10.1080/23311916.2018.1433983

Green, D.W., Southard, M.Z., 2018. Perry's Chemical Engineers' Handbook, 9th ed. McGraw-Hill Education.

Guerra, O.J., Calderón, A.J., Papageorgiou, L.G., Reklaitis, G. V, 2019. Integrated shale gas supply chain design and water management under uncertainty. AIChE J. 65, 924–936. https://doi.org/10.1002/aic.16476

Gupta, V., Grossmann, I.E., 2012. An efficient multiperiod MINLP model for optimal planning of offshore oil and gas field infrastructure. Ind. Eng. Chem. Res. 51, 6823–6840. https://doi.org/10.1021/ie202959w

Hong, B., Li, X., Di, G., Li, Y., Liu, X., Chen, S., Gong, J., 2019. An integrated MILP method for gathering pipeline networks considering hydraulic characteristics. Chem. Eng. Res. Des. 152, 320–335. https://doi.org/10.1016/j.cherd.2019.08.013

Hong, B., Li, X., Di, G., Song, S., Yu, W., Chen, S., Li, Y., Gong, J., 2020. An integrated MILP model for optimal planning of multi-period onshore gas field gathering pipeline system. Comput. Ind. Eng. 146, 106479. https://doi.org/10.1016/j.cie.2020.106479

Liu, K., Biegler, L.T., Zhang, B., Chen, Q., 2020. Dynamic optimization of natural gas pipeline networks with demand and composition uncertainty. Chem. Eng. Sci. 215, 115449. https://doi.org/10.1016/j.ces.2019.115449

Lockhart, R., Martinelli, R.C., 1949. Proposed Correlation of Data for Isothermal Two-Phase, Two-Component Flow in Pipes. Chem. Eng. Prog. 45, 38–48.

Lockhart, R.W., Martinelli, R.C., 1949. Proposed Correlation of Data for Isothermal Two-Phase, Two-

Component Flow in Pipes. Chem. Eng. Prog. 45, 38–48.

Mah, R.S.H., Shacham, M., 1978. Pipeline Network Design and Synthesis. pp. 125–209. https://doi.org/10.1016/S0065-2377(08)60133-7

Montagna, A.F., Cafaro, D.C., 2019. Supply chain networks servicing upstream operations in oil and gas fields after the shale revolution. AIChE J. 65. https://doi.org/10.1002/aic.16762

Oddie, G., Shi, H., Durlofsky, L.J., Aziz, K., Pfeffer, B., Holmes, J.A., 2003. Experimental study of two and three phase flows in large diameter inclined pipes. Int. J. Multiph. Flow 29, 527–558. https://doi.org/10.1016/S0301-9322(03)00015-6

Peng, Z., Li, C., Grossmann, I.E., Kwon, K., Ko, S., Shin, J., Feng, Y., 2020. Multi-period Design and Planning Model of Shale Gas Field Development. AIChE J.

Petalas, N., Aziz, K., 2000. A Mechanistic Model for Multiphase Flow in Pipes. J. Can. Pet. Technol. 39, 43–55. https://doi.org/10.2118/00-06-04

Rexilius, J., 2015. The Well Factory Approach to Developing Unconventionals: A Case Study from the Permian Basin Wolfcamp Play. SPE/CSUR Unconv. Resour. Conf.

Society of Petroleum Engineers, 2006. Petroleum Engineering Handbook.

Spedding, P.L., Benard, E., Donnelly, G.F., 2006. Prediction of pressure drop in multiphase horizontal pipe flow. Int. Commun. Heat Mass Transf. 33, 1053–1062. https://doi.org/10.1016/j.icheatmasstransfer.2006.05.004

Spedding, P.L., Donnelly, G.F., Cole, J.S., 2005. Three phase oil-water-gas horizontal co-current flow: I. Experimental and regime map. Chem. Eng. Res. Des. 83, 401–411. https://doi.org/10.1205/cherd.02154

Stevens, P., 2012. The ' Shale Gas Revolution ': Developments and Changes. Chatham House.

Tan, L., Zuo, L., Wang, B., 2018. Methods of Decline Curve Analysis for Shale Gas Reservoirs. Energies 11, 552. https://doi.org/10.3390/en11030552

U.S. Energy Information Administration (EIA)., 2018. Annual Energy Outlook 2018 with projections to 2050. Washington, DC.

Vanorsdale, C.R., 2013. Production Decline Analysis Lessons from Classic Shale Gas Wells, in: SPE Annual Technical Conference and Exhibition. Society of Petroleum Engineers. https://doi.org/10.2118/166205-MS

Weymouth, T., 1942. Problems in natural gas engineering. ASME Trans. 34, 185–234.

Wilkes, J.O., 2005. Fluid Mechanics for Chemical Engineers, 2nd ed. Prentice Hall.

Wolsey, L.A., 1998. Integer Programming, 1st ed. Wiley-Interscience.

Worthen, R.A., Henkes, R.A., 2015. CFD for the Multiphase Flow Splitting From a Single Flowline Into a Dual Riser. SPE Annu. Tech. Conf. Exhib.

Zaghloul, J., Adewumi, M., Ityokumbul, M.T., 2008. Hydrodynamic modeling of three-phase flow in production and gathering pipelines. J. Energy Resour. Technol. Trans. ASME 130, 0430041–0430048. https://doi.org/10.1115/1.3000135

Zhang, H., Liang, Y., Zhang, W., Wang, B., Yan, X., Liao, Q., 2017. A unified MILP model for topological structure of production well gathering pipeline network. J. Pet. Sci. Eng. 152, 284–293. https://doi.org/10.1016/j.petrol.2017.03.016

Abbreviations

O&G	Oil and Gas
MINLP	Mixed Integer Non-Linear Programming
NLP	Non-Linear Programming
MILP	Mixed Integer Linear Programming
GOR	Gas-to-Oil ratio
WOR	Water-to-Oil ratio
TB	Tank Battery
CDP	Centralized Delivery Point
LM	Lockhart-Martinelli
SB	SuperBattery

Nomenclature

Sets and Indices

B	tank batteries potential location
BJ	battery junctions potential nodes
BZ	batteries capacities
C	components in the flow {oil, gas, water}
CDP	centralized delivery points
D	pipeline diameters
J	potential wellpad junction nodes
K	clusters of wellpads
P	wellpads
T	time in months
B_k	batteries able to receive production from wellpads/junctions of cluster k
BZ_b	batteries capacities allowed at location b
P_k	wellpads that belong to cluster k
P_t	wellpads that starts its production at period t
J_k	wellpad junctions that belong to cluster k
CDP_c	includes the centralized delivery nodes available for component c
P_j - J_p	wellpad - junctions possible connections
J_b - B_j	junctions - batteries possible connections
$B_{bj} - BJ_b$	batteries - batteries junctions possible connections
CDP_{bj} - BJ_{cdp}	batteries junctions - centralized facilities possible connections
TP	subset of periods with some wellpad increasing its production or reaching production peak
TS	subset of periods with some wellpad starting its production

Parameters

General Parameters

$dist_{p,j}$ / $dist_{j,b}$ / $dist_{b,bj}$ / $dist_{bj,cdp}$	euclidean distances between nodes [miles]
i_jun_j / i_junb_{bj}	installation costs of a junction [thousands USD]
c_jun	costs per unit of volume being processed in a junction node [thousand USD/(bbl-tcf^3)]
P_j	minimum pressure at junction nodes
P_{bj}	estimated pressure after separation at battery junction nodes
P_{gc}	pressure after the gas compressor
P_{cdp}	pressure at the centralized delivery point
μ	synthesizes unit conversion and fluid parameters
i	interest rate considered to assess the capital flows

ε	roughness of the pipe [m]
ρ^{LP}	average density of the liquid phase [kg/m^3]
ρ^{GP}	density of the gas phase [kg/m^3]
Δp^{max}	maximum pressure drop admitted between the nodes
μ^{LP}	average viscosity of the liquid phase [Pa.s]
μ^{GP}	viscosity of the gas phase [Pa.s]
T_0	standard temperature conditions [ºK]
P_0	standard pressure conditions [MPa]
T	average flow temperature [ºK]
P_{in}	upper pressure limit at the inlet of the pipeline according to the pipeline segment
z	compressibility factor, dimensionless
r	gas/liquid ratio [ft3/bbl]
sg	specific gravity of liquid phase relative to water
s	specific gravity of the gas relative to air
ζ	SPE specific constant according to a flow regimen

Wellpads Data

$starttime_p$	start time (in months) for the production of wellpad p
gor_p	Gas to Oil Ratio from wells of pad p
wor_p	Water to Oil Ratio from wells of pad p
$v_{p,\tau}^c$	production of component c at wellpad p, t periods after its start date [B/D] - [mcf/d]
P_p	production pressure at the outlet of the wellpad p

Tank Batteries and CDP Data

$batcap_{bz}^c$	maximum flow rate of component c allowed by a battery with capacity bz.
i_bat_{bz}	cost of a battery of type bz [thousands USD]
c_bat_{bz}	costs per unit of volume being processed in a battery of size bz [thousand USD/(bbl-tcf^3)]
P_b	minimum inlet pressure at tank batteries

Pipelines Data

$diam_d$	diameter of the pipeline d [in]
i_pipe_d	investment unit cost of a pipeline of diameter d [thousand USD/mile]
c_pipe_d	unit cost of using a pipeline per unit of volume being transported [thousand USD/(bbl-tcf^3)]
$vmax^{LP}$	liquid-phase maximum velocity [m/s]

Binary Variables

$\hat{x}_{p,j,d,t}$	1 if a pipeline of diameter d connecting pad p to junction node j is installed at period t
$\hat{y}_{j,b,d,t}$	1 if a pipeline of diameter d connecting node j to battery b is installed at period t
$\hat{u}_{b,bj,d,t}^c$	1 if a pipeline of diameter d for component c connecting TB b and junction bj is installed at period t
$\hat{w}_{bj,cdp,d,t}^c$	1 if a pipeline of diameter d for component c connecting junction bj with cdp is installed at period t
$\hat{z}bat_{b,bz,t}$	1 if a battery of capacity bz is installedse in the location b at period t
$\hat{z}cdp_{cdp,t}$	1 if the delivery point cdp is installed at period t
$zbat_{b,bz,t}$	1 if a battery of capacity bz is available at location b during period t
$zcdp_{cdp,t}$	1 if the delivery point cdp is active during period t
$x_{p,j,d,t}$	1 if a pipeline of diameter d connecting pad p to junction node j is available at period t
$y_{j,b,d,t}$	1 if a pipeline of diameter d connecting node j to battery b is available at period t
$u_{b,bj,d,t}^c$	1 if a pipeline of diameter d for component c connecting TB b and junction bj is available at period t
$w_{bj,cdf,d,t}^c$	1 if a pipeline of diameter d for component c connecting junction bj with cdp is available at period t

Positive Continuous Variables

$FP^c_{p,j,t}$	flow of component c from wellpad p to junction node j during period t
$FPB^c_{j,b,t}$	flow of component c from junction node j to battery b during period t
$FBJ^c_{b,bj,t}$	flow of component c from battery b to battery junction bj during period t
$FCDP^c_{bj,cdp,t}$	flow of component c from battery junction bj to delivery point cdp during period t
VS^{LP}	superficial velocity of liquid phase (water + oil)
VS^{GP}	superficial velocity of gas phase [m/s]
RY^{LP}	Reynolds Number for Liquid Phase
RY^{GP}	Reynolds Number for Gas Phase
HFF^{LP}	friction factor for gas phase following Haland Equation
$maxflow^c_d$	single-phase capacities for component c in a pipeline of diameter d
$\Delta P^{LP}/L$	pressure drop gradient for liquid phase
$\Delta P^{GP}/L$	pressure drop gradient for gas phase
X_{LM}	Lockhart-Martinelli parameter
Y^{LP}	liquid phase multiplier for Lockhart-Martinelli
Y^{GP}	gas phase multiplier for Lockhart-Martinelli
$\Delta P/L$	pressure drop gradient for multiphase flow per distance unit
ΔP^{Total}	multiphase total pressure drop in the pipeline segment
P_{in}	inlet pressure
P_{out}	outlet pressure