Applied Energy 233–234 (2019) 1065–1077

Contents lists available at ScienceDirect

Applied Energy

journal homepage: www.elsevier.com/locate/apenergy

Stochastic planning of electricity and gas networks: An asynchronous column generation approach

Carlos Saldarriaga-Cortés[a,*], Harold Salazar[a], Rodrigo Moreno[b,c], Guillermo Jiménez-Estévez[d,e]

[a] Electrical Engineering Department, Universidad Tecnológica de Pereira, Pereira 660003, Colombia
[b] Department of Electrical Engineering, Universidad de Chile, Santiago 8370451, Chile
[c] Department of Electrical and Electronic Engineering, Imperial College London, London SW7 2AZ, UK
[d] Energy Center, Universidad de Chile, Santiago 8370451, Chile
[e] Department of Electrical and Electronics Engineering, Universidad de Los Andes, Bogotá 111711, Colombia

HIGHLIGHTS

- This paper presents enhanced modeling of gas network within planning models.
- Solves the gas-electricity planning problem appling Dantzig-Wolfe decomposition.
- Presents a scalable planning problem under long-term uncertainties.
- Highlights the importance of multi-stage-stochastic gas-electricity planning problem.

ARTICLE INFO

Keywords:
Integrated planning
Natural gas and electricity systems
Stochastic programming
Dantzig-Wolfe decomposition

ABSTRACT

Planning networks within a multi-stage stochastic framework is becoming critical for improving the economic performance of investment decisions against the present levels of uncertainty. This problem, however, has been proved extremely challenging to be solved on real networks, especially when considering the interactions among various energy vectors. In this context, this paper proposes the use of Dantzig-Wolfe decomposition and parallel asynchronous column generation to solve a multi-stage stochastic planning of an integrated power and natural gas system, including non-linear effects of gas compressors reformulated in a mixed integer linear programming fashion. We compare the computational performance of the proposed approach against two alternatives: a parallel synchronous column generation approach and the counterfactual, monolithic approach, where the mixed integer linear program (without decomposition) is directly solved by a commercial solver. Our sources of long-term uncertainty are the locations and volumes of (i) new renewable generation (which may depend on policy objectives, regulatory incentives, etc. that are constantly evolving) and (ii) new demands. The model also ensures that the planned energy infrastructure can effectively be operated reliably against a large array of operating conditions originated by high variability of renewable generation outputs, multiple demand levels and hydro inflows. Through various case studies, we discuss and demonstrate the importance of stochastic and integrated planning of electricity and natural gas systems along with the benefits of asynchronous algorithms and decomposition techniques that can be parallelized.

In this paper we use the following rules to simplify the notation:

- Variables are presented in italic and bold, and constants are in italics. For instance, $\boldsymbol{V}_{i,j}^{k}$ refers to a variable, $V_{i,j}^{k}$ is a parameter, and k, i and j are indexes within particular sets. We use lower case to refer generically to an index and upper case if an index is equal to a particular value in a set. For example, if we want to refer to the element where $j = J0$, we will write $V_{i,J0}^{k}$.
- We use a hat and an English letter in upper case to refer to sets as in $\widehat{A}$. Sets can also be index dependent as in $\widehat{A}_n$.
- Upper and lower bounds are denoted with a bar above (upper bound) or below (lower bound) the name of the variable, e.g.,$\bar{V}$ and $\underline{V}$ are the upper and lower limits of variable $\boldsymbol{V}_{i,j}^{k}$.

* Corresponding author.
E-mail addresses: casaldarriaga@utp.edu.co (C. Saldarriaga-Cortés), hsi@utp.edu.co (H. Salazar), rmorenovieyra@ing.uchile.cl, r.moreno@imperial.ac.uk (R. Moreno), gjimenez@centroenergia.cl (G. Jiménez-Estévez).

https://doi.org/10.1016/j.apenergy.2018.09.148
Received 3 August 2018; Accepted 15 September 2018
Available online 10 November 2018
0306-2619/

Nomenclature

Indexes

h	index that refers to a hydrological scenario
i	index that refers to a specific asset in the class r
j	index that refers to type, i.e., new asset capacity
k	index that refers to an electrical or natural gas network bus
l	index that refers to a segment of a piecewise linear approximation
n	index that refers to a scenario-tree node
o	index that refers to an operating condition
r	index that refers to a generation technology or transmission asset. For generation technology, it can be hydropower (H), renewable (RNW), dual-fuel (DF) [which can be, in turn, coal (DFC), oil (DFO) or natural gas (DFG)], non-dual-fuel (NDF). For a transmission asset, it can be a power transmission line (L), pipeline (P), natural gas compressor (CP), natural gas well (W), or liquefied-natural gas regasification terminal (LNG)
s	index that refers to an electrical system (ES) or natural gas system (GS)
w	Index that refers to wind availability scenario

Sets

$\widehat{F}^r$	set of all new and existing assets of class r
$\widehat{G}_{h,w}$	set of all operating conditions associated with a hydrological scenario h and a wind availability scenario w
$\widehat{H}$	set of all thehydrological scenarios
$\widehat{I}^r$	set of discretized capacities –or types– associated with all assets in class r
$\widehat{M}^r$	set of the new assets of class r that can be built
$\widehat{\mathcal{N}}$	set of all scenario-tree nodes
$\widehat{O}$	set of the operating conditions
$\widehat{S}^s$	set of all buses of system s
$\widehat{W}$	set of all the wind availability scenarios

Parameters

$a^r_{i,k}$	incidence matrix entry of the electrical network of class r, asset i, bus k
$b^r_{i,k}$	incidence matrix entry of the natural gas network of class r, asset i, bus k
$BIGM^r$	scalar of large magnitude of class r
CAP^r_j	maximum capacity of a new asset of class r, type j
$IC^r_{i,j}$	annuitized investment cost (IC) of a new class r, asset i, type j
$in^r_{i,k}$, $out^r_{i,k}$	constants that indicate if the natural gas flows into or out of the bus k of the asset i, class r
$k^r_{0,i}$, $k^r_{1,i,l}$	constants that depend on the physical characteristics of the asset i, class r, segment l
$m_{i,l}$, $n_{i,l}$	weighting factor of the asset i, segment l
NLS^r	number of segments of the piecewise linear approximation of class r
UC^r_i	fuel cost per unit of class r, asset i
UUC^s	unsupplied energy unit cost of system s
$PL_{k,o,n}$	active power consumption at bus k, operating condition o, scenario-tree node n
$GL_{k,o,n}$	natural gas consumption at bus k, operating condition o, scenario-tree node n
$X0_i$	reactance of an existing transmission line, asset i
$X_{i,j}$	reactance of a potential new transmission line, asset i, capacity j
$\alpha_{0,i,l}$, $\alpha_{1,i,l}$, $\alpha_{2,i,l}$	constants that depend on the physical characteristics of compressors, asset i, segment l
$\bar{\beta}_i$	maximum compression ratio of the asset i
ρ_n	probability that scenario-tree node n occurs
λ_n	discount rate of the scenario-tree node n
τ_o	duration –in hours- of an operating condition o

Binary Variables

$\boldsymbol{S}_{i,o,n}$	natural gas flow direction (1 if positive or 0 if negative) through a pipeline or a compressor, asset i, operating condition o, scenario-tree node n
$\boldsymbol{z}_{i,l,o,n}$	auxiliary binary variable of the piecewise linear approximation of the pipeline, asset i, segment l, operating condition o, scenario-tree node n
$\boldsymbol{\Lambda'}^r_{i,j,n}$	granting decision variables of class r, asset i, type j, scenario-tree node n
$\boldsymbol{\Lambda}^r_{i,j,n}$	investment decision variables of class r, asset i, type j, scenario-tree node n
$\boldsymbol{\sigma}^r_{i,o,n}$	commitment variable of class r, asset i, operating condition o, scenario-tree node n

Positive Variables

$\boldsymbol{PGF}_{i,o,n}$	natural gas flow for a positive direction of the asset i, operating condition o, scenario-tree node n
$\boldsymbol{NGF}_{i,o,n}$	natural gas flow for a negative direction of the asset i, operating condition o, scenario-tree node n
$\boldsymbol{UE}^s_{k,o,n}$	unsupplied energy in system s, at bus k, operative condition o, scenario-tree node n
$\boldsymbol{\Delta}^r_{i,l,o,n}$	auxiliary variable of the piecewise linear approximation of class r, asset i, segment l, operating condition o, scenario-tree node n
$\boldsymbol{\Psi}^r_{i,o,n}$	amount of primary energy consumed by a power generator or transported by a transmission asset of class r, asset i, operating condition o, scenario-tree node n
$\boldsymbol{\pi}_{k,o,n}$	natural gas pressure at bus k, operating condition o, scenario-tree node n
$\boldsymbol{\pi}^{in}_{i,o,n}$	natural gas pressure at the input node of the pipeline of the asset i, operating condition o, scenario-tree node n
$\boldsymbol{\pi}^{out}_{i,o,n}$	natural gas pressure at the output node of the pipeline of the asset i, operating condition o, scenario-tree node n

Free Variables

$\boldsymbol{LCP}^{CPC}_{i,o,n,l}$	natural gas consumption of a compressor for a positive direction of the natural gas flow of the asset i, segment l, operative condition o, scenario-tree node n
$\boldsymbol{LCN}^{CPC}_{i,o,n,l}$	natural gas consumption of a compressor for a negative direction of the natural gas flow of the asset i, segment l, operative condition o, scenario-tree node n
$\boldsymbol{P}^r_{i,o,n}$	active power flow through asset i, class r, operative condition o of the scenario-tree node n
$\boldsymbol{\Delta\pi}_{i,o,n}$	difference between nodal pressures in the compressor i, operative condition o of the scenario-tree node n
$\boldsymbol{\theta}_{\boldsymbol{k,o,n}}$	voltage angle, at bus k, operative condition o of the scenario-tree node n

1. Introduction

Nowadays, energy planners need to deal with both the increasing amounts of renewable generation in the power network and the unprecedented levels of uncertainty in the long term originated, among other reasons, by evolving policy and market conditions.

Regarding the increased amounts of renewables, it is important to ensure that the planned network infrastructure can effectively be operated reliably against a large array of operating conditions due to the high variability of renewable generation outputs. This is exacerbated in hydro-thermal systems, where there is also a need to consider a variety of hydro inflows that can change every year/season. For instance, the need to adequately face a dry year/season (driven, for example, by a climate phenomenon like *el niño* in South America) might ultimately drive the need for further thermal generation capacity. In this context, gas-fired generation technologies are attractive because these can properly provide generation adequacy in hydro-thermal power systems and because of their lower CO2 emission levels that are in line with the future decarbonization energy policy. Proper consideration of gas-fired power plants, however, requires careful planning of further critical infrastructure (which is also beyond the electricity system) such as the gas network (including pipelines, compressors, LNG regasification terminals, etc.). Hence, a proper energy plan may require recognition of the interactions between the gas and the electricity systems, and this has been already acknowledged by planners and regulators in jurisdictions like Colombia [1], Chile, [2], and CAISO [3].

Regarding the treatment of long-term uncertainty, planners need to consider a variety of future scenarios that may happen and thus plan infrastructure accordingly. In this context, investment decisions need to be sufficiently flexible in order to adapt to multiple scenarios in the long-term future, avoiding to lock into inefficient network plans that are usually determined through deterministic models. In fact, deterministic planning, like that proposed in [4–9], can lead to network investments that may be very efficient only if one given scenario realizes in the future (the one that is considered in the deterministic plan), but can also lead to an extremely poor economic and reliability performance if other realizations occur [10]. In contrast, planning through stochastic optimization models (such as [11–14], for a comprehensive review in electricity networks see [15]) can endogenously capture uncertainty, finding the true optimal solutions that can properly hedge against the uncertain future by determining investment plans that are flexible and adaptable in the long term. Furthermore, stochastic planning can find solutions that remain hidden by deterministic plans, irrespective of the parameters used [15]. In other words, there are solutions that are efficient and are revealed only under uncertainty and these cannot be found by deterministic models that do not acknowledge the presence of uncertainty. Although endogenous recognition of uncertainty in planning models is paramount, in particular in the case of the integrated electricity and gas systems as indicated in the latest literature such as [16–18], stochastic models still remain difficult to be solved due to the large amounts of variables and constraints needed in order to capture the occurrence of multiple scenarios in the future. This difficulty is compounded by the presence of network elements that require non-linear equations to be properly represented. In this context, our paper provides a new multi-stage stochastic model to deal with the integrated planning of power and natural gas systems with compressors (that are usually ignored due to their non-linear equations) that is solved through an asynchronous column generation approach.

1.1. Literature review

There are several papers that have proposed integrated electricity and gas expansion planning problems. From a deterministic point of view, static [4–6] and dynamic [7–9] integrated planning models have been proposed, demonstrating the advantages of the combined optimization of the power and natural gas systems. Beyond deterministic models, only a few papers have incorporated uncertainties in the power and natural gas systems. In this context, Refs. [16–22] proposed both static (for instance [21,22]) and dynamic or multi-stage (for instance [16–20]) planning models through robust and stochastic programs, considering various uncertainties such as the load demand growth, fuel availability, fuel costs, hydro inflows, and wind power outputs. Due to the inherent computational complexity associated with uncertainty, some of these papers (such as [19–21]) proposed meta-heuristic techniques that do not guarantee optimal solutions (although these can determine sufficiently good solutions in real, large-scale case studies). Robust optimization has been also used to deal with reliability and resilience within the integrated planning of the electricity and gas systems. For instance [16,17], used robust optimization to determine secured plans against various events such as the occurrence of extreme events (superstorms, earthquakes and floods), wind power availability, and system contingencies.

As far as we know, only Refs. [18,22] have proposed stochastic optimization programs to deal with the combined planning of electricity and gas networks under uncertainty that can be optimally solved through mixed integer linear programming (MILP) techniques. In fact, Ref. [22] presented a two-stage stochastic optimization model to plan electricity and gas networks under uncertainties related to electricity and natural gas demands. This reference also highlights the need for advanced algorithms to solve problems at a larger scale. Likewise, Ref. [18] presented a multi-stage stochastic approach, demonstrating the advantages of the multi-stage approach against a two-stage approach. Our paper complements the previous literature as indicated next.

1.2. Paper's contributions

The two main contributions of our paper are as follows:

1. Enhanced modelling of natural gas network components within planning models: We model two features of the natural gas compressors: their capability to transfer gas in two directions and the effect of the compression rate on the natural gas consumption. These features are often ignored in planning models. In fact, the former is usually addressed under the assumption that natural gas flows only in a single direction. The latter is often assumed proportional to the flow magnitude through the compressor when, in reality, these relations are non-linear. So, neither of these is accurate. In this vein, we proposed a linear representation of both aforementioned features so that (i) the natural gas can flow freely in both directions over the network infrastructure and (ii) gas consumptions from compressors are more accurately represented. Additionally, we propose a new piecewise linear approximation of the natural gas flow through a pipeline that has a lower error when comparing with other academic literature. As a consequence, we believe that our piecewise linear approximations make the modelling of the entire natural gas network more realistic, accurate and therefore superior to previous works (e.g. [4–28]) since the greater levels of details in compressors and pipeline's model have an important impact on network operation, which may, in turn, affect investment decisions.
2. A scalable gas-electricity network expansion problem: We apply Dantzig-Wolfe decomposition and parallel asynchronous column generation (expanding from [12]) to solve the integrated planning of electricity and gas systems under long-term uncertainties such as demand growth, volumes and locations of new renewable generation plants (that may depend on evolving policy, regulatory incentives, etc.) and under short-term variability of operating conditions originated by multiple demand levels, renewable generation outputs, hydro conditions, etc. Particularly, we use the split-variable formulation and Dantzig-Wolfe reformulation of the capacity planning problem for electricity distribution proposed in reference [12] and adapted it to our integrated gas-electricity network expansion problem, solving the reformulated master-slave problem through an

asynchronous column generation algorithm undertaken in a parallel computing fashion. We also compare the computational performance of the asynchronous column generation algorithm against that of the synchronous version. To our knowledge, this is the first time that these techniques are used to solve this problem.

Finally, it is worth to mention that we also contribute by adding more quantitative evidence to the academic literature (like that in [17–19]) that demonstrates the importance of multi-stage stochastic planning of integrated electricity and gas systems. We believe this is a key discussion since, in practice, network planners are still running deterministic approaches to determine network investment decisions, which significantly endangers the cost-effective transition towards a more sustainable and low-carbon energy system. These deterministic planning practices though may be justified by the lack of adequate computational tools and algorithms to tackle multi-stage stochastic planning problems in real life, which we attempt to attack in this paper.

This paper is organized as follows. Section 2 provides an overview of the problem and the main features of our proposal. Sections 3 and 4 present our mathematical program and its solution methodology, respectively. Section 5 shows the main results, illustrating (i) the features of the stochastic planning of electricity and natural gas systems and (ii) the computational performance of the proposed approach. Finally, Section 6 concludes.

2. Problem overview

We seek to plan both electricity and natural gas networks in an integrated fashion and across various years (so-called stages due to the stochastic nature of the proposed mathematical program) when facing long-term uncertainty. In this paper, our sources of long-term uncertainty are the locations and volumes of (i) new renewable generation, for instance, wind based power generation (WG) [which may depend on policy objectives, regulatory incentives, etc. that are constantly evolving] and (ii) new demands (ND); although other sources of uncertainty can be included without undertaking significant changes in the proposed mathematical program. In our approach, we minimize the coupled, total cost of electricity and natural gas systems (which is the sum of the investment, operational and unsupplied demand costs in both electricity and natural gas networks), while investing in new infrastructure such as transmission lines, transformers, conventional generating units, gas pipelines, LNG regasification terminals and, importantly, compressors.

In the short-term, we model various operating conditions within every node of the scenario tree in order to determine operational and unsupplied demand costs. These operating conditions also ensure that the infrastructure planned can be operated against an array of different levels of wind power generation, demand and hydro inflows.

An important feature of the proposed model is also the coupling between electricity and gas sectors, and this is extremely important when the planner can invest in gas generating units that need significant network infrastructure (in both electricity and gas networks) to properly operate the systems. Importantly, this is undertaken within a stochastic framework with multiple long-term scenarios and operating conditions in the short-term, which clearly increases the model complexity.

In this context, we present 2 variants of a master-slave, column generation algorithm (based on a Dantzig-Wolfe decomposition –DWD– technique proposed in [12]): parallel asynchronous and synchronous. The mathematical formulation and the solution methodology are presented next.

3. Mathematical formulation

The formulation corresponds to an integrated electricity and natural gas expansion planning problem under uncertainty, which is tackled through a multi-stage stochastic MILP model, whose objective function and constraints are explained next.

3.1. Objective function

The model minimizes the expected present value of operational, investment and unsupplied energy costs of both power and natural gas networks. The objective function is shown in (1a) which is composed of three terms –IC_n, $OC_{o,n}$, $UEC_{o,n}$– that are detailed in (1b), (1c) and (1d), respectively. In this paper, an operating condition o is a combination of one hydrological/inflow scenario, one discretized demand level, and one wind availability level.

$$\upsilon = \sum_{n\in\hat{\mathcal{N}}} \rho_n \lambda_n \cdot \left(\boldsymbol{IC}_n + \sum_{o\in\hat{O}} \tau_o \cdot (\boldsymbol{OC}_{o,n} + \boldsymbol{UEC}_{o,n}) \right) \tag{1a}$$

Eq. (1b) is the investment cost of expanding the current power and natural gas infrastructures to supply future demands at the scenario-tree node n. $\boldsymbol{\Lambda}'$ and $\boldsymbol{\Lambda}$ are the granted and required investment decision variables at the scenario-tree node n that are detailed in Section 3.2.3.

$$\boldsymbol{IC}_n = \sum_{r\in\hat{A}} \sum_{i\in\hat{M}^r} \sum_{j\in\hat{I}^r} IC^r_{i,j} \cdot \boldsymbol{\Lambda}'^r_{i,j,n};\ \hat{A} = \{H, DF, NDF, L, P, CP, LNG\} \tag{1b}$$

Eq. (1c) is the operational costs –$\boldsymbol{OC}$– of an operating condition o at node n of the scenario tree. It takes into account the cost of different types of fuel to supply power and natural gas demands.

$$\boldsymbol{OC}_{o,n} = \sum_{r\in\hat{B}} \sum_{i\in\hat{F}^r} UC^r_i \cdot \boldsymbol{\Psi}^r_{i,o,n};\ \hat{B} = \{DFO, DFC, NDF, W, LNG\} \tag{1c}$$

Finally, (1d) is the unsupplied energy cost of the power system (PS) and the natural gas system (GS) of an operating condition o at node n of the scenario tree.

$$\boldsymbol{UEC}_{o,n} = \sum_{s\in\{PS,GS\}} \sum_{k\in\hat{S}^s} UUC^s \cdot \boldsymbol{UE}^s_{k,o,n} \tag{1d}$$

3.2. Main optimization constraints

For the sake of clarity, power and natural gas system constraints are explained separately next. Non-anticipativity constraints are presented at the end.

3.3. Power system constraints

Eq. (2a) is the active power balance equation at each bus of the power network.

$$\sum_{r\in\hat{C}} \sum_{i\in\hat{F}^r} a^r_{i,k} \cdot \boldsymbol{P}^r_{i,o,n} = PL_{k,o,n} - \boldsymbol{UE}^{PS}_{k,o,n};\ \forall\ k \in \hat{S}^{PS};$$
$$\hat{C} = \{DFO, DFC, DFG, NDF, H, RNW, L\} \tag{2a}$$

Eqs. (2b)–(2d) represent the linear DC power flow equations associated with the existing and potentially new network infrastructure. $\boldsymbol{PF}_{i,j,o,n}$ is an auxiliary variable that includes the multiple effects of a transmission line on the disjunctive DC model.

$$\boldsymbol{P}^L_{i,o,n} = \frac{1}{X0_i} \cdot \sum_{k\in\hat{S}^{PS}} a^L_{i,k} \cdot \theta_{k,o,n};\ \forall\ \{i \in \hat{F}^L\} \& \{i \notin \hat{M}^L\} \tag{2b}$$

$$-BIGM^L \cdot \left(1 - \sum_{j\in\hat{I}^L} \boldsymbol{\Lambda}^L_{i,j,n}\right) \le \sum_{j\in\hat{I}^L} X_{i,j} \cdot \boldsymbol{PF}_{i,j,o,n} - \sum_{k\in\hat{S}^{PS}} a^L_{i,k} \cdot \theta_{k,o,n} \le BIGM^L \cdot \left(1 - \sum_{j\in\hat{I}^L} \boldsymbol{\Lambda}^L_{i,j,n}\right);\ \forall\ i \in \hat{M}^L \tag{2c}$$

$$\boldsymbol{P}^{L}_{i,o,n} = \sum_{j \in \hat{I}^{L}} \boldsymbol{PF}_{i,j,o,n};\ \forall\ i \in \hat{M}^{L} \tag{2d}$$

A piecewise linear representation of the fuel consumption of thermal generators is shown in (2e)–(2h). Note that this representation corresponds to a linearization of the quadratic fuel consumption function, often used for power generators [4]. It also models the dual-fuel capability of thermal generators that it is the ability to utilize more than one type of fuel.

$$\boldsymbol{\Psi}^{r}_{i,o,n} = k^{r}_{0,i} \cdot \boldsymbol{\sigma}^{r}_{i,o,n} + \sum_{l=1}^{NLS^{r}} k^{r}_{1,i,l} \cdot \boldsymbol{\Delta}^{r}_{i,l,o,n};\ \forall\ r \in \hat{D},\ \forall\ i \in \hat{F}^{r} \tag{2e}$$

$$\boldsymbol{P}^{r}_{i,o,n} = \underline{P}^{r}_{i} \boldsymbol{\sigma}^{r}_{i,o,n} + \sum_{l=1}^{NLS^{r}} \boldsymbol{\Delta}^{r}_{i,l,o,n};\ \forall\ r \in \hat{D},\ \forall\ i \in \hat{F}^{r} \tag{2f}$$

$$0 \le \boldsymbol{\Delta}^{r}_{i,l,o,n} \le \bar{\Delta}^{r}_{i,l} \boldsymbol{\sigma}^{r}_{i,o,n};\ \forall\ r \in \hat{D},\ \forall\ i \in \hat{F}^{r},\ \forall\ l \in \{1, 2, ..., NLS^{r}\} \tag{2g}$$

$$\boldsymbol{\sigma}^{DFC}_{i,o,n} + \boldsymbol{\sigma}^{DFO}_{i,o,n} + \boldsymbol{\sigma}^{DFG}_{i,o,n} \le 1;\ \forall\ i \in \hat{F}^{DF};\ \hat{D} = \{NDF, DFC, DFO, DFG\} \tag{2h}$$

Eq. (2i) is the water discharge ($\boldsymbol{\Psi}^{H}_{i,o,n}$) as a function of the power output ($\boldsymbol{P}^{H}_{i,o,n}$) and a water discharge constant (κ^{H}_{i}). Eq. (2j) limits the energy storage capacity ($\boldsymbol{E}^{H}_{i,h}$) of a water reservoir for hydropower generation.

$$\boldsymbol{\Psi}^{H}_{i,o,n} = \kappa^{H}_{i} \cdot \boldsymbol{P}^{H}_{i,o,n};\ \forall\ i \in \hat{F}^{H} \tag{2i}$$

$$\sum_{o \in \hat{G}_{h,w}} \tau_{o} \cdot \boldsymbol{\Psi}^{H}_{i,o,n} \le \boldsymbol{E}^{H}_{i,h};\ \forall\ i \in \hat{F}^{H},\ \forall\ h \in \hat{H},\ \forall\ w \in \hat{W} \tag{2j}$$

Operational limits are represented as follows. Eq. (2k) limits the unsupplied energy at each bus, (2l) limits the power output of all generators, (2m)–(2n) limit the active power flow through the power network assets.

$$0 \le \boldsymbol{UE}^{PS}_{k,o,n} \le PL_{k,o,n};\ \forall\ k \in \hat{S}^{PS} \tag{2k}$$

$$\underline{P}^{r}_{i} \cdot \boldsymbol{\sigma}^{r}_{i,o,n} \le \boldsymbol{P}^{r}_{i,o,n} \le \bar{P}^{r}_{i} \cdot \boldsymbol{\sigma}^{r}_{i,o,n};\ \forall\ r \in \{NDF, DFC, DFO, DFG, H, RNW\},\ \forall\ i \in \hat{F}^{r} \tag{2l}$$

$$-\bar{P}^{L}_{i} \le \boldsymbol{P}^{L}_{i,o,n} \le \bar{P}^{L}_{i};\ \forall\ \{i \in \hat{F}^{L}\} \& \{i \notin \hat{M}^{L}\} \tag{2m}$$

$$-CAP^{L}_{j} \cdot \boldsymbol{\Lambda}^{L}_{i,j,n} \le \boldsymbol{PF}_{i,j,o,n} \le CAP^{L}_{j} \cdot \boldsymbol{\Lambda}^{L}_{i,j,n};\ \forall\ i \in \hat{M}^{L},\ \forall\ j \in \hat{I}^{L} \tag{2n}$$

Finally, the maximum number of assets that can be built is limited by (2o), while (2p)–(2q) ensure that a new asset can be utilized only after being built. All new assets have a construction time (lag between decision and implementation times) of one stage.

$$\sum_{j \in \hat{I}^{r}} \boldsymbol{\Lambda}^{r}_{i,j,n} \le 1;\ \forall\ r \in \hat{A},\ \forall\ i \in \hat{M}^{r} \tag{2o}$$

$$\boldsymbol{P}^{r}_{i,o,n} \le \sum_{j \in \hat{I}^{r}} CAP^{r}_{j} \cdot \boldsymbol{\Lambda}^{r}_{i,j,n};\ \forall\ r \in \{H, NDF, L\},\ \forall\ i \in \hat{M}^{r} \tag{2p}$$

$$\sum_{r \in \{DFO, DFC, DFG\}} \boldsymbol{P}^{r}_{i,o,n} \le \sum_{j \in \hat{I}^{DF}} CAP^{DF}_{j} \cdot \boldsymbol{\Lambda}^{DF}_{i,j,n};\ \forall\ i \in \hat{M}^{DF} \tag{2q}$$

3.4. Natural gas system constraints

For the natural gas system, (3a) represents the nodal balance equation that includes the natural gas consumption of the compressor.

$$\sum_{r \in \hat{E}} \sum_{i \in \hat{F}^{r}} b^{r}_{i,k} \cdot \boldsymbol{\Psi}^{r}_{i,o,n} = GL_{k,o,n} - \boldsymbol{UE}^{GS}_{k,o,n};$$

$$\forall\ k \in \hat{S}^{GS};\ \hat{E} = \{P, CP, CPC, NDF, DFG, W, LNG\} \tag{3a}$$

Eqs. (3b)–(3q) are the linear piecewise representations of the natural gas flows through existing or potentially new pipelines. The corresponding non-linear formulation can be found in [4,23–28] and the proof of (3b)–(3q) is shown in appendix A1.

$$\boldsymbol{\pi}^{out}_{i,o,n} = \boldsymbol{\pi}^{in}_{i,o,n} + \sum_{l=1}^{NLS} m_{i,l} \cdot \boldsymbol{\Delta}^{P}_{i,l,o,n};\ \forall\ i \in \hat{F}^{P} \tag{3b}$$

$$\boldsymbol{\Psi}^{r}_{i,o,n} = \boldsymbol{PGF}^{r}_{i,o,n} - \boldsymbol{NGF}^{r}_{i,o,n};\ \forall\ r \in \{P, CP\},\ \forall\ i \in \hat{F}^{r} \tag{3c}$$

$$\boldsymbol{PGF}^{P}_{i,o,n} + \boldsymbol{NGF}^{P}_{i,o,n} = \sum_{l=1}^{NLS} \boldsymbol{\Delta}^{P}_{i,l,o,n};\ \forall\ i \in \hat{F}^{P} \tag{3d}$$

$$0 \le \boldsymbol{PGF}^{P}_{i,o,n} \le BIGM^{P} \cdot \boldsymbol{S}_{i,o,n};\ \forall\ i \in \hat{F}^{P} \tag{3e}$$

$$0 \le \boldsymbol{NGF}^{P}_{i,o,n} \le BIGM^{P} \cdot (1 - \boldsymbol{S}_{i,o,n});\ \forall\ i \in \hat{F}^{P} \tag{3f}$$

$$\sum_{u=1}^{l} \boldsymbol{\Delta}^{P}_{i,u,o,n} \le \frac{1}{n_{i,l}} \cdot \left(\boldsymbol{\pi}^{in}_{i,o,n} + \sum_{u=1}^{l} m_{i,u} \cdot \boldsymbol{\Delta}^{P}_{i,u,o,n}\right);\ \forall\ i \in \hat{F}^{P},\ \forall\ l \in \{1, ..., NLS^{P}\} \tag{3g}$$

$$\boldsymbol{\xi}_{i,l,o,n} \le \sum_{u=1}^{l-1} \boldsymbol{\Delta}^{P}_{i,u,o,n};\ \forall\ i \in \hat{F}^{P},\ \forall\ l \in \{2, ..., NLS^{P}\} \tag{3h}$$

$$\boldsymbol{\Delta}^{P}_{i,u,o,n} \le BIGM^{P} \cdot \boldsymbol{z}_{i,l,o,n};\ \forall\ i \in \hat{F}^{P},\ \forall\ l \in \{2, ..., NLS^{P}\} \tag{3i}$$

$$-BIGM^{P} \cdot (1 - \boldsymbol{z}_{i,l,o,n}) \le n_{i,l-1} \cdot \boldsymbol{\xi}_{i,l,o,n} - \left(\boldsymbol{\pi}^{in}_{i,o,n} + \sum_{u=1}^{l-1} m_{i,u} \cdot \boldsymbol{\Delta}^{P}_{i,u,o,n}\right) \le BIGM^{P} \cdot (1 - \boldsymbol{z}_{i,l,o,n});\ \forall\ i \in \hat{F}^{P},\ \forall\ l \in \{2, ..., NLS^{P}\} \tag{3j}$$

$$-BIGM^{P} \cdot \boldsymbol{z}_{i,l,o,n} \le \boldsymbol{\xi}_{i,l,o,n} \le BIGM^{P} \cdot \boldsymbol{z}_{i,l,o,n};\ \forall\ i \in \hat{F}^{P},\ \forall\ l \in \{2, ..., NLS^{P}\} \tag{3k}$$

$$-BIGM^{P} \cdot (2 - \boldsymbol{S}_{i,o,n} - \boldsymbol{\Lambda}^{P}_{i,n}) \le \sum_{k \in \hat{S}^{GS}} in^{P}_{i,k} \cdot \boldsymbol{\pi}_{k,o,n} - \boldsymbol{\pi}^{in}_{i,o,n} \le BIGM^{P} \cdot (2 - \boldsymbol{S}_{i,o,n} - \boldsymbol{\Lambda}^{P}_{i,n});\ \forall\ i \in \hat{F}^{P} \tag{3l}$$

$$-BIGM^{P} \cdot (2 - \boldsymbol{S}_{i,o,n} - \boldsymbol{\Lambda}^{P}_{i,n}) \le \sum_{k \in \hat{S}^{GS}} out^{P}_{i,k} \cdot \boldsymbol{\pi}_{k,o,n} - \boldsymbol{\pi}^{out}_{i,o,n} \le BIGM^{P} \cdot (2 - \boldsymbol{S}_{i,o,n} - \boldsymbol{\Lambda}^{P}_{i,n});\ \forall\ i \in \hat{F}^{P} \tag{3m}$$

$$-BIGM^{P} \cdot (1 + \boldsymbol{S}_{i,o,n} - \boldsymbol{\Lambda}^{P}_{i,n}) \le \sum_{k \in \hat{S}^{GS}} out^{P}_{i,k} \cdot \boldsymbol{\pi}_{k,o,n} - \boldsymbol{\pi}^{in}_{i,o,n} \le BIGM^{P} \cdot (1 + \boldsymbol{S}_{i,o,n} - \boldsymbol{\Lambda}^{P}_{i,n});\ \forall\ i \in \hat{F}^{P} \tag{3n}$$

$$-BIGM^{P} \cdot (1 + \boldsymbol{S}_{i,o,n} - \boldsymbol{\Lambda}^{P}_{i,n}) \le \sum_{k \in \hat{S}^{GS}} in^{P}_{i,k} \cdot \boldsymbol{\pi}_{k,o,n} - \boldsymbol{\pi}^{out}_{i,o,n} \le BIGM^{P} \cdot (1 + \boldsymbol{S}_{i,o,n} - \boldsymbol{\Lambda}^{P}_{i,n});\ \forall\ i \in \hat{F}^{P} \tag{3o}$$

$$\boldsymbol{z}_{i,l,o,n} \le \boldsymbol{\Lambda}^{P}_{i,n};\ \forall\ i \in \hat{F}^{P} \tag{3p}$$

$$\boldsymbol{\Lambda}^{P}_{i,n} = 1;\ \forall\ i \in \hat{M}^{P} \tag{3q}$$

Eqs. (4a)–(4h) are the linear piecewise representations of the natural gas flows through existing or potentially new compressors. The corresponding non-linear formulation can be found in [23–26], and the proof of (4a)–(4h) is shown in appendix A2.

$$\sum_{k\in\hat{S}^{GS}} b^{CP}_{i,k}\cdot\boldsymbol{\pi}_{k,o,n} = \boldsymbol{\Delta\pi}_{i,o,n};\ \forall\ i\in\hat{F}^{CP} \tag{4a}$$

$$(1-\bar{\beta}_i)\cdot\sum_{k\in\hat{S}^{GS}} out^{CP}_{i,k}\cdot\boldsymbol{\pi}_{k,o,n} \le \boldsymbol{\Delta\pi}_{i,o,n} \le (\bar{\beta}_i-1)\cdot\sum_{k\in\hat{S}^{GS}} in^{CP}_{i,k}\cdot\boldsymbol{\pi}_{k,o,n};\ \forall\ i\in\hat{F}^{CP} \tag{4b}$$

$$-(\bar{\pi}-\underline{\pi})\cdot\boldsymbol{\sigma}^{CP}_{i,o,n} \le \boldsymbol{\Delta\pi}_{i,o,n} \le (\bar{\pi}-\underline{\pi})\cdot\boldsymbol{\sigma}^{CP}_{i,o,n};\ \forall\ i\in\hat{F}^{CP} \tag{4c}$$

$$0 \le \boldsymbol{\Psi}^{CPC}_{i,o,n} \le BIGM^{CP}\cdot\boldsymbol{\sigma}^{CP}_{i,o,n};\ \forall\ i\in\hat{F}^{CP} \tag{4d}$$

$$\boldsymbol{LCP}^{CPC}_{i,o,n,l} \le \boldsymbol{\Psi}^{CPC}_{i,o,n};\ \forall\ l\in\{1,...,NLS^{CP}\} \tag{4e}$$

$$\boldsymbol{LCN}^{CPC}_{i,o,n,l} \le \boldsymbol{\Psi}^{CPC}_{i,o,n};\ \forall\ l\in\{1,...,NLS^{CP}\} \tag{4f}$$

$$-BIGM^{CP}(1-\boldsymbol{\sigma}^{CP}_{i,o,n}) \le \boldsymbol{LCP}^{CPC}_{i,o,n,l} - \alpha_{2,i,l}\boldsymbol{\Psi}^{CP}_{i,o,n} - \alpha_{1,i,l}\boldsymbol{\Delta\pi}_{i,o,n} + \alpha_{0,i,l}\cdot\sum_{k\in\hat{S}^{GS}} in^{CP}_{i,k}\cdot\boldsymbol{\pi}_{k,o,n} \le BIGM^{CP}\cdot(1-\boldsymbol{\sigma}^{CP}_{i,o,n});$$

$$\forall\ l\in\{1,...,NLS^{CP}\},\ \forall\ i\in\hat{F}^{CP} \tag{4g}$$

$$-BIGM^{CP}\cdot(1-\boldsymbol{\sigma}^{CP}_{i,o,n}) \le \boldsymbol{LCN}^{CPC}_{i,o,n,l} + \alpha_{2,i,l}\cdot\boldsymbol{\Psi}^{CP}_{i,o,n} + \alpha_{1,i,l}\cdot\boldsymbol{\Delta\pi}_{i,o,n} + \alpha_{0,i,l}\cdot\sum_{k\in\hat{S}^{GS}} out^{CP}_{i,k}\cdot\boldsymbol{\pi}_{k,o,n} \le BIGM^{CP}\cdot(1-\boldsymbol{\sigma}^{CP}_{i,o,n});$$

$$\forall\ l\in\{1,...,NLS^{CP}\},\ \forall\ i\in\hat{F}^{CP} \tag{4h}$$

The operational limits associated with the natural gas system are presented in (5a)–(5c). Eq. (5a) limits the unsupplied natural gas volumes per bus, (5b) constrains the production of natural gas wells and LNG-RT and the natural gas flow through the compressors, and (5c) limits the natural gas pressure at each bus.

$$0 \le \boldsymbol{UE}^{GS}_{k,o,n} \le GL_{k,o,n};\ \forall\ k\in\hat{S}^{GS} \tag{5a}$$

$$\underline{\Psi}^{r}_{i} \le \boldsymbol{\Psi}^{r}_{i,o,n} \le \bar{\Psi}^{r}_{i};\ \forall\ r\in\{W,\ LNG,\ CP\},\ \forall\ \{i\in\hat{F}^{r}\}\&\{i\notin\hat{M}^{GNL}\} \tag{5b}$$

$$\underline{\pi} \le \boldsymbol{\pi}_{k,o,n} \le \overline{\pi};\ \forall\ k\in\hat{S}^{GS} \tag{5c}$$

Eqs. (5d)–(5e) ensure that a new natural gas asset (e.g. pipelines, LNG-TR, etc.) can be utilized only after being built.

$$\sigma^{CP}_{i,o,n} \le \sum_{j\in\hat{I}^{CP}} \boldsymbol{\Lambda}^{CP}_{i,j,n};\ \forall\ i\in\hat{M}^{CP} \tag{5d}$$

$$0 \le \boldsymbol{\Psi}^{LNG}_{i,o,n} \le \sum_{j\in\hat{I}^{GNL}} CAP^{LNG}_{j}\cdot\boldsymbol{\Lambda}^{LNG}_{i,j,n};\ \forall\ i\in\hat{M}^{GNL} \tag{5e}$$

3.5. Non-anticipativity constraints

Eq. (6) –adapted from [12]– shows the non-anticipativity constraints in a compact form, where $\mathbb{X}_n$ is the binary vector composed of all investment decision variables $\boldsymbol{\Lambda}^{r}_{i,j,n}$ and $\mathbb{X}'_n$ is the binary vector composed of all "granting" decision variables that are penalized in (1b), i.e., $\boldsymbol{\Lambda}''^{r}_{i,j,n}$. Note that vectors are highlighted in double script.

$$\mathbb{X}_n \le \sum_{n'\in\hat{\tau}_n} \mathbb{X}'_{n'};\ \forall\ n\in\widehat{\mathscr{N}} \tag{6}$$

Note that $\mathbb{X}_n$ establishes a request of network expansion at each node $n\in\widehat{\mathscr{N}}$ of the scenario tree and those requirements are granted by the vector $\mathbb{X}'_{n'}$ at some predecessor node $n'\in\hat{\tau}_n$.

4. Proposed methodology

This section presents the algorithm used to determine the optimal solution of the abovementioned problem, which is based on a Dantzig-Wolfe, column generation method introduced in [12]. Here, we present 2 implementations of a column generation: parallel asynchronous and synchronous. Also, we briefly introduce the serial case in order to help the reader understand our proposal. Next, we introduce some fundamental definitions that serve, afterwards, to explain the algorithm implementation.

4.1. Fundamental definitions

Next, we re-write the entire problem in its compact form in order to introduce, in a more straightforward manner, the master and slave subproblems used to determine the optimal solution.

4.2. The compact problem

For the sake of simplicity, the complete formulation of the model in its compact form is re-written and presented in (7a)–(7c), where $\mathbb{X}'_{n'}$ is a vector that corresponds to the investment decisions at node $n'\in\widehat{\mathscr{N}}$, $\mathbb{Y}_n$ is a vector that corresponds to the operational variables at node$n\in\widehat{\mathscr{N}}$, and $\mathscr{F}_n$ represents the feasible solution space formulated in detail previously in (2a)–(5e).

$$\min_{\mathbb{X}'_n,\mathbb{X}_n,\mathbb{Y}_n}\left\{\sum_{n\in\widehat{\mathscr{N}}} \mathbb{CO}^{\mathrm{T}}_n\mathbb{Y}_n + \mathbb{CI}^{\mathrm{T}}_n\mathbb{X}'_n\right\} \tag{7a}$$

s.t.

$$\mathbb{X}_n \le \sum_{n'\in\tau_n} \mathbb{X}'_{n'};\ \forall\ n\in\widehat{\mathscr{N}} \tag{7b}$$

$$\{\mathbb{X}_n,\ \mathbb{Y}_n\}\in\mathscr{F}_n;\ \forall\ n\in\widehat{\mathscr{N}} \tag{7c}$$

In (7b), $\mathbb{X}_n$ is a vector that facilitates the formulation of the master-slave subproblems below and corresponds to, as explained in [17], the required investments to deal only with the operating conditions that compose node $n\in\widehat{\mathscr{N}}$, ignoring other nodes of the scenario tree.

4.3. The master problem (MP)

The MP is presented in (8a)–(8e), where investment decisions are made based on a linear convex combination (with $\mathbb{L}^{p}_{n}$ being the weighting variables) of the extreme points $\{\widehat{\mathbb{X}}^{p}_{n},\ \widehat{\mathbb{Y}}^{p}_{n}\}$ that represents (5c). Each *p-th* extreme point of $\mathscr{F}_n$ is obtained by running the slave subproblems (presented next). The dual variables $\mathbb{W}_n$ and $\boldsymbol{\mu}_n$ are obtained from a fully linear, relaxed MP that ignores the integrality of (8d)–(8e).

$$\upsilon^{MP} = \min_{\mathbb{L}^{p}_{n},\mathbb{X}'_n}\left\{\sum_{n\in\widehat{\mathscr{N}}}\left(\mathbb{CI}^{\mathrm{T}}_n\mathbb{X}'_n + \mathbb{CO}^{\mathrm{T}}_n\sum_{p\in\widehat{\mathscr{P}}}\boldsymbol{L}^{p}_{n}\widehat{\mathbb{Y}}^{p}_{n}\right)\right\} \tag{8a}$$

[DualVariables]

s.t.

$$\sum_{p\in\widehat{\mathscr{P}}}\boldsymbol{L}^{p}_{n}\widehat{\mathbb{X}}^{p}_{n} \le \sum_{n'\in\hat{\tau}_n}\mathbb{X}'_{n'};\ \forall\ n\in\widehat{\mathscr{N}}\ \ [\mathbb{W}_n] \tag{8b}$$

$$\sum_{p\in\widehat{\mathscr{P}}}\boldsymbol{L}^{p}_{n} = 1;\ \forall\ n\in\widehat{\mathscr{N}}\ \ [\boldsymbol{\mu}_n] \tag{8c}$$

$$\mathbb{X}'_n\in\{0,1\};\ \forall\ n\in\widehat{\mathscr{N}} \tag{8d}$$

$$\boldsymbol{L}^{p}_{n}\in\{0,1\};\ \forall\ n\in\widehat{\mathscr{N}},\ \forall\ p\in\widehat{\mathscr{P}} \tag{8e}$$

4.4. The slave subproblem (SSP)

The SSP per node $n\in\widehat{\mathscr{N}}$ is presented in (9a)–(9b), where dual variables $\mathbb{W}_n$ and $\boldsymbol{\mu}_n$ (imported from the master problem), are used.

$$\min_{\mathbb{X}_n, \mathbb{Y}_n} \left\{ \sum_{n \in \widehat{\mathcal{N}}} \mathbb{CO}_n^{\mathrm{T}} \mathbb{Y}_n - \mathbb{W}_n^{\mathrm{T}} \mathbb{Y}_n - \boldsymbol{\mu}_n \right\} \tag{9a}$$

s.t.

$$\{\mathbb{X}_n, \mathbb{Y}_n\} \in \mathscr{F}_n \tag{9b}$$

4.5. Algorithm

(1) *General description*

The MP and SSP previously introduced are run by using the following 8-step algorithm.

i. By using $\boldsymbol{\mu}_n = 0$ and $\mathbb{W}_n = \mathbb{CI}_n$ for all $n \in \widehat{\mathcal{N}}$, solve (9a)–(9b) and determine optimal values of $\mathbb{X}_n^*$ and $\mathbb{Y}_n^*$ for all $n \in \widehat{\mathcal{N}}$. Also, define k: =1, $\widehat{\mathscr{P}}$: ={k}, and, for all $n \in \widehat{\mathcal{N}}$, $\widehat{\mathbb{X}}_n^{p=k}$: =$\mathbb{X}_n^*$ and $\widehat{\mathbb{Y}}_n^{p=k}$: =$\mathbb{Y}_n^*$.
ii. By using defined $\widehat{\mathbb{X}}_n^p$ and $\widehat{\mathbb{Y}}_n^p$ for all $n \in \widehat{\mathcal{N}}$ and $p \in \widehat{\mathscr{P}}$, solve (8a)–(8e), relaxing the integrality constraints associated with (8d)–(8e). Determine the value of the dual variables for all $n \in \widehat{\mathcal{N}}$ ($\mathbb{W}_n$ and $\boldsymbol{\mu}_n$), the objective function υ^{MP^*}, and define υ^{MP-LP}: =υ^{MP^*}.
iii. If optimal solution from step ii is integer, then define $\upsilon^{MP-MILP}$: =υ^{MP-LP} and go to v. Otherwise, go to iv.
iv. By using defined $\widehat{\mathbb{X}}_n^p$ and $\widehat{\mathbb{Y}}_n^p$ for all $n \in \widehat{\mathcal{N}}$ and $p \in \widehat{\mathscr{P}}$, solve (8a)–(8e) (without linear relaxation), determine the optimal value of the objective function υ^{MP^*}, and define $\upsilon^{MP-MILP}$: =υ^{MP^*}.
v. By using $\mathbb{W}_n$ and $\boldsymbol{\mu}_n$ for all $n \in \widehat{\mathcal{N}}$ from step ii, solve (9a)–(9b) and determine optimal values of $\mathbb{X}_n^*$ and $\mathbb{Y}_n^*$ for all $n \in \widehat{\mathcal{N}}$.
vi. Define d_n: =$\mathbb{CO}_n^{\mathrm{T}} \mathbb{Y}_n^* - \mathbb{W}_n^{*\mathrm{T}} \mathbb{X}_n^* - \mu_n^*$ for all $n \in \widehat{\mathcal{N}}$ and calculate (10):

$$GAP = \frac{\upsilon^{MP-MILP} - (\upsilon^{MP-LP} + \sum_{n \in \mathcal{N}} d_n)}{\upsilon^{MP-LP} + \sum_{n \in \mathcal{N}} d_n} \tag{10}$$

vii. If *GAP* is equal to or lower than a certain (very small) number, then STOP. Otherwise, go to viii.
viii. For all $n \in \widehat{\mathcal{N}}$ do: if $d_n < 0$ then: k: =k + 1, $\widehat{\mathscr{P}}$: =$\widehat{\mathscr{P}} \cup \{k\}$, $\widehat{\mathbb{X}}_n^{p=k}$: =$\mathbb{X}_n^*$ and $\widehat{\mathbb{Y}}_n^{p=k}$: =$\mathbb{Y}_n^*$. Go to ii.

Although the above procedure does not ensure *GAP* = 0, in practice, we always obtain *GAP* < 0.05% in our case studies.

(2) *Algorithm implementations*

We implement the previous algorithm in 2 different running modes: parallel synchronous and parallel asynchronous. We also present the serial case in order to help the reader understand our proposal. These modes differ in the aforementioned step v (i.e. execution of SSP) as explained next:

a. Serial (S), where (9a)–(9b) is solved for all $n \in \widehat{\mathcal{N}}$, one after the other.
b. Parallel synchronous (PS), where execution of (9a)–(9b) is started for all $n \in \widehat{\mathcal{N}}$. Execution of (9a)–(9b) for each node n is undertake in parallel. We move from step v to vi, once (9a)–(9b) has been solved for all $n \in \widehat{\mathcal{N}}$.
c. Parallel asynchronous (PAS), where execution of (9a)–(9b) is started for all $n \in \widehat{\mathcal{N}}$. We move from step v to vi after a critical waiting time is reached (reset every time a new set of subproblems is executed) and (9a)–(9b) has been solved for, at least, one node n (while solutions for other nodes remain being executed). In this case, MP can be run by adding new columns associated with $d_n < 0$, while slave subproblems continue being executed. Note that there is a lag to compute (10) since, to do so, we have to wait for solutions from (9a)–(9b) for all $n \in \widehat{\mathcal{N}}$ whose executions started at the same time. Next, we show that, despite this lag, this asynchronous mode presents the fastest times to solve the problem.

It is important to emphasize that we selected a Dantzig–Wolfe-based approach rather than a (nested) Benders-based decomposition approach (or any variants), because the latter does not allow straightforward treatment of integer variables in a multi-stage stochastic setting. In this vein, the work in [29] that developed the stochastic dual dynamic integer programming (SDDiP) concept is promising. Comparisons between our proposal and the recent work in [29] are beyond the scope of this paper.

5. Tests and results

In this section, two test systems are used to illustrate the advantages of the proposed stochastic framework and solution methodology. The first system is shown in Fig. 1(a) along with its scenario tree in Fig. 2(b) and this is used to demonstrate how the proposed integrated stochastic approach is paramount to capture uncertainty and thus determine flexible, first-stage adaptive investment solutions that can be

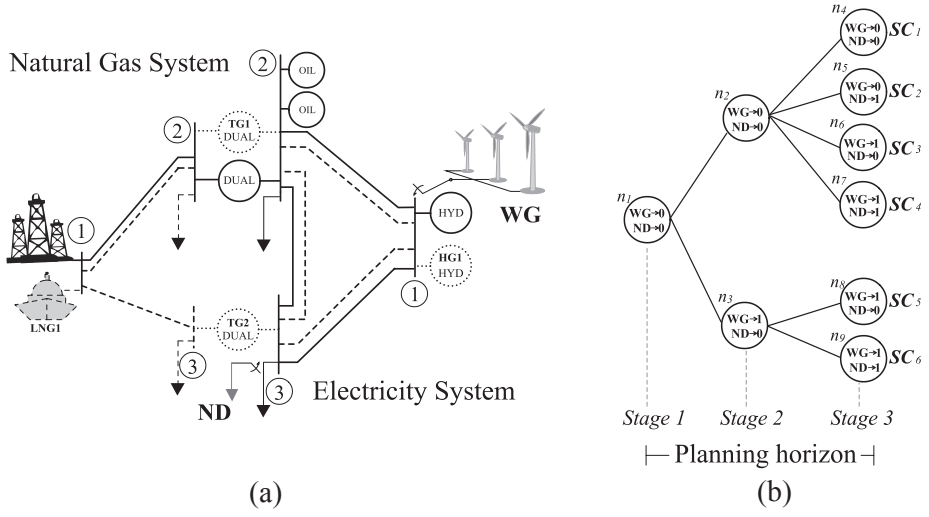

Fig. 1. (a) 3-bus test system at the stage 1 (where continuous and dashed lines indicate existing and candidate infrastructure, respectively) and (b) scenario tree.

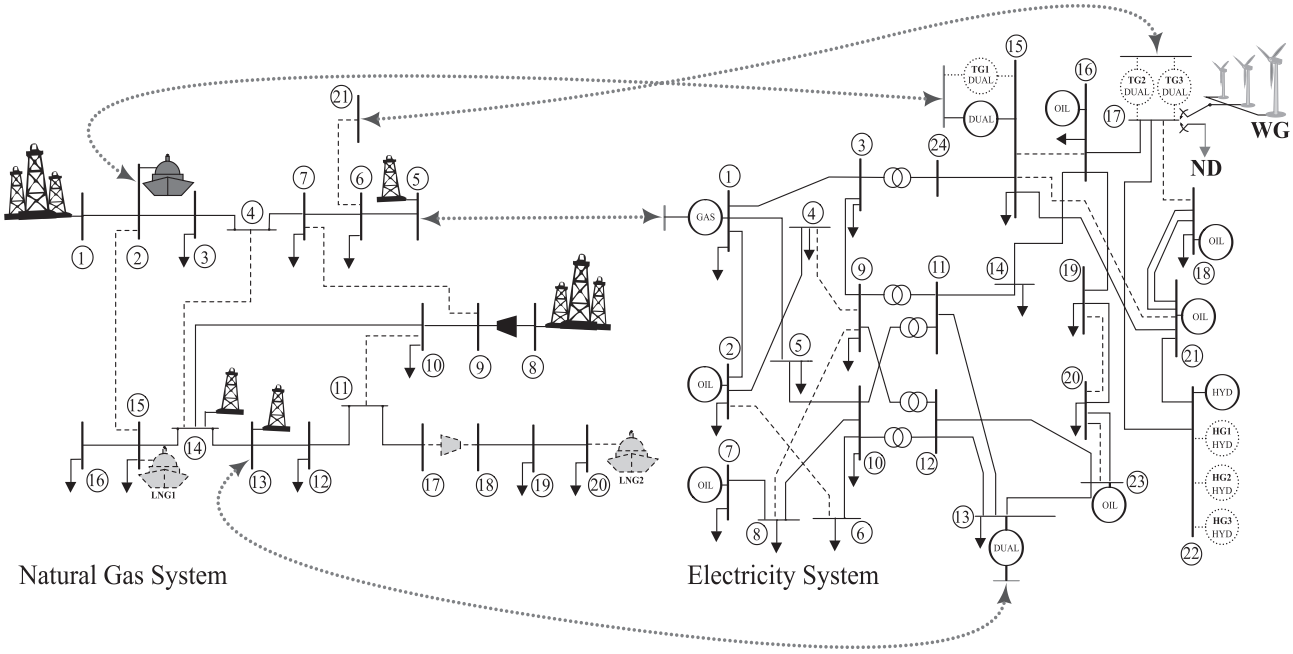

Fig. 2. Natural gas (left) and electricity (right) systems at the stage 1. Continuous and dashed lines indicate existing and candidate infrastructure, respectively.

complemented later on by further investments, when more information become available. The second test system is shown in Fig. 2, which is used to demonstrate the need for decomposition approaches to make the problem scalable. Here, we show that the proposed asynchronous approach is critical to reach solutions within a reasonable execution time. Note that the second test system presents a compressor in the gas network, which is (as explained earlier) represented through a piecewise mathematical representation. Although this representation increases the number of binary variables in the problem, we demonstrated that the proposed approach can successfully deal with that as shown next. For both test systems, all new assets have a construction time (lag between decision and implementation times) of one stage.

The scenario tree used for both test systems is described in Fig. (2b), where two types of uncertainty are considered: (1) the future connection of wind generation capacity (WG) that depends on evolution of regulatory incentives and policy, and (2) the future connection of new demand (ND).

All results are obtained on a 64-bit personal computer with a 3.2-GHz processor Intel(R) Core(TM) i5-4570 CPU and 8 GB of RAM. The proposed algorithms were implemented in GAMS version 24.5.4, the MILP sub-problems were solved by using CPLEX 12.6.2.0, and the LP master problem was solved by using KNITRO 9.1.0.

5.1. 3-bus case study: perfect information vs. Stochastic solution

(1) *Input data*

The 3-bus integrated system is composed of a 3-bus power system and a 3-bus natural gas system (see Fig. 1a). At stages 1 and 2, the power system is composed of 3 thermal generators (1 dual-fuel unit and 2 oil units), 1 hydro generator, 3 transmission lines and 2 power demands located at bus 2 (50 MW) and bus 3 (700 MW). Likewise, the natural gas system is composed of 1 pipeline, 1 well and no natural gas demand, except for that from the dual-fuel unit. At stage 3, the natural gas system will present 2 new demands ($20 \times 10^3 m^3/h$) at buses 2 and 3 (equal in volume), and, in the power system, 50 MW of additional demand will be connected to bus 2. Additionally, it is probable that a new WG of 600 MW might be connected to bus 1 and a ND of 300 MW might be connected to bus 3, and all the possible scenarios of ND and WG are shown in the scenario tree of Fig. 1b.

Table 1
Investment Cost – 3-Bus Test System.

	L		DF	H	P		LNG-RT	
Type[a]	1	2	1	1	1	2	1	2
Capacity[b]	300	1000	500	10	115	190	0.15	0.3
Annuitized cost[c]	0.013	0.044	31	2	0.09	0.15	2.66	5.33

[a] Indicates the size of the new investment.
[b] In MW for electricity infrastructure, Mm^3 for LNG terminal, and mm for pipelines (diameter).
[c] In $/km for electricity lines and pipelines, and in $ for generators and LNG terminals.

In order to supply the future demand in the electricity and natural gas systems, various candidate assets are proposed, which are plotted in dotted lines in Fig. 1a. The capacity and annuitized cost of each candidate asset is presented in Table 1.

(2) *Results*

To demonstrate the advantages of the proposed stochastic framework, the stochastic (STC) solution is compared against the optimal investment plan under perfect information, i.e., the optimal plan that results when the planner has full information on future generation and demand realizations [11]. The optimal infrastructure found for each scenario (SC) under perfect information and the optimal infrastructure proposed by the STC, are presented in Table 2.

Table 2 shows that under perfect information, there are decisions that are not optimal under the stochastic approach. Conversely and more importantly, in the true stochastic solution, there are optimal investments that are not found in any scenario under the deterministic approach. For example, deterministic analysis can justify investment in pipe 1–3 of up to 190 mm, while in the stochastic approach investment in the same pipe can be only up to 115 mm. Similarly, investments in line 1–2 and generator TG1 are not determined under any scenario in the deterministic approach; however, they are clearly part of the optimal stochastic solution.

Table 3 shows the benefits of the stochastic solution in terms of its lower expected cost against the uncertain future. In effect, Table 3 shows that although the realization of a particular scenario can be faced

Table 2
Optimal infrastructure for 3-Bus Test System.

		SC1[a]	SC2[a]	SC3[a]	SC4[a]	SC5[a]	SC6[a]	STC[b]
Line	**1–2**	0	0	0	0	0	0	1@n2
	1–3	0	0	1@st2	1@st2	1@st1	1@st1	1@n1
	2–3	0	1@st2	0	0	0	0	1@n1
Generator	**TG1**	0	0	0	0	0	0	1@n1
	TG2	1@st1	1@st1	1@st1	1@st1	0	1@st2	1@n2
	HG1	0	0	0	0	0	0	0
Pipe	**1–2**	0	1@st2	0	0	0	0	2@n1
	1–3	2@st1	2@st1	2@st1	2@st1	1@st2	1@st2	1@n2,3
LNG-RT	**LNG1**	2@st1	2@st1	1@st1	1@st1	0	1@st2	2@n1

[a] x@y indicates: element type x decided at stage y (installed one stage later).
[b] x@z indicates: element type x decided at scenario tree node z (installed one stage later).

Table 3
Operational and Investment Cost for 3-Bus Test System MMUSD.

		Realization of						
		SC1	SC2	SC3	SC4	SC5	SC6	STC
Optimal plan	**SC1**	**$ 14.2**	$ 18.8	$ 9.4	$ 12.0	$ 5.0	$ 7.6	$ 11.2
	SC2	$ 14.5	**$ 17.5**	$ 9.5	$ 11.8	$ 5.2	$ 7.5	$ 11.0
	SC3	$ 14.5	$ 20.1	**$ 8.6**	$ 10.5	$ 4.3	$ 6.1	$ 10.7
	SC4	$ 14.5	$ 20.1	$ 8.6	**$ 10.5**	$ 4.3	$ 6.1	$ 10.7
	SC5	$ 15.1	IS[a]	IS	IS	**$ 2.2**	$ 4.6	IS
	SC6	$ 17.5	IS	IS	IS	$ 2.5	**$ 4.3**	IS
	STC	$ 14.5	$ 17.6	$ 9.3	$ 11.0	$ 3.8	$ 5.7	$ 10.3

[a] IS: infeasible solution.

at lower costs by deterministic decisions, under uncertainty, the stochastic solution features the smallest average cost. In fact, optimal plan SC1, for example, that is optimal only if scenario SC1 occurs in the future, can perform very poorly if scenario SC6 happens (rather than SC1). Hence, every deterministic plan performs the best under the scenario for which it was designed, but poorly under another scenario. In this context, the stochastic solution performs reasonably well under every individual scenario and presents the lowest expected cost.

5.2. Larger scale study: scalability and computational performance

(1) *Input data*

The second test system is composed of the IEEE 24-bus power system coupled with a modified Belgian 21-bus natural gas system (see Fig. 2). For the power system, it is expected that a ND of 100 MW and a new WG of 1000 MW may be connected to power bus 17; the possible scenarios of ND and WG are shown in the scenario tree depicted in Fig. 1b.

Table 4 shows the computational performance of the monolithic, complete formulation (CF) of the model represented by (5) [without using decomposition methods], and the PS and PAS implementations presented in Section 3.2. We ignored the serial case, since this proved extremely slow. To analyze the advantages of each implementation, 9 cases were defined with increasing size (escalating the number of operating conditions, i.e. combination of different demand levels, hydro inflows and wind availabilities). In this context, case 1 considers only one operating condition in the entire year and, evidently, corresponds to the smallest size in terms of the number of variables and constraints. In contrast, case 9 features 150 operating conditions in a year which represent the combination of 5 demand levels, 3 hydro conditions and 10 wind levels.

(2) *Results*

Table 4 shows that solving the problem through the CF can be significantly problematic since we could only find solutions for the firsts 3 cases. For those of increased size, only PS and PAS are capable to find an optimal solution, where PAS features a clear advantage in terms of execution times. Interestingly, note that the advantages of CF for extremely small-size cases are quickly lost when the problem increases its size.

6. Conclusions

We presented an integrated electricity and gas multi-stage stochastic mathematical program, which considers (i) long-term uncertainty associated with locations and volumes of new renewable generation and new demands, (ii) short-term variability associated with various operating conditions from multiple levels of wind, demand and hydro inflows in operational timescales, and (iii) a more adequately representation of the physics of the gas network by including the natural gas consumption from compressors through a piecewise formulation of their non-linear equations. The complete formulation is broken down through a Dantzig-Wolfe decomposition and a parallel asynchronous column generation algorithm that proves efficient for a larger instance of the planning problem.

Through our model, we discussed the importance to properly recognize uncertainty, variability and the coupling between the gas and electricity networks when planning new infrastructure in the energy sector. Failing to do so can lead to solutions extremely exposed to higher costs in both the short term (since the network has not been properly designed to deal with variable resources) and long term (since investments cannot be easily adapted to various scenarios that may occur in the future). We argue that these types of models and the necessary algorithms to efficiently solve large instances of the presented expansion planning problem are extremely necessary given the very high levels of uncertainty faced by energy planners currently, the

Table 4
IEEE 24-Bus and Belgium Natural Gas Test System.

Case	Demand levels	Hydrology conditions	Wind levels	Number of operating conditions	Number of continuous variables	Number of binary variables	Number of constraints	Simulation time (h)			Costs		
								CF	PS	PAS	Total cost (MUSD)	Investment cost (MUSD)	Operational cost (MUSD)
1	1	1	1	1	5135	1622	5601	0.02	0.08	0.07	18596.463	302.061	18294.402
2	1	1	2	2	9807	2782	10,745	0.07	0.11	0.09	18645.986	309.534	18336.452
3	2	2	2	8	37,839	9742	41,481	20.31	5.72	4.38	19001.165	328.779	18672.386
4	3	3	3	27	126,607	31,782	138,769	[a]	13.99	9.49	19713.900	329.711	19384.189
5	5	3	2	30	140,623	35,262	154,009	[a]	11.45	7.49	19993.769	329.091	19664.678
6	5	3	3	45	210,703	52,662	230,785	[a]	22.07	15.69	19846.431	329.156	19517.275
7	5	3	5	75	350,863	87,462	384,337	[a]	42.91	27.21	19898.514	329.156	19569.358
8	5	5	5	125	584,463	145,462	640,257	[a]	71.91	41.22	19898.389	329.504	19568.885
9	5	3	10	150	701,263	174,462	768,217	[a]	75.12	55.16	19891.603	330.021	19561.582

[a] More than 80 h.

necessity to ensure that decisions need to be both robust in the short-term (so as to deal with variable conditions) and flexible and adaptive in the long-term (so as to take appropriate first-stage decisions that can be optimally complemented later on, while the future unfolds), and the evident necessity to solve these problems in practice, at least over simplified, equivalent networks that are large enough for policy and long-term planning studies.

Acknowledgements

This work was supported by National Department of Science, Technology and Research (COLCIENCIAS) of Colombia under Grant COLCIENCIAS 159-2015, 286-2018, and "Programa de Formación Becas Doctorales – Doctorados nacionales convocatoria 617 - 2013" scholarship. We also thank the Pacific Alliance in its student and academic mobility platform, the Chilean Agency for International Cooperation (AGCI), Energy Center FCFM at University of Chile. The authors gratefully acknowledge the Complex Engineering Systems Institute (CONICYT-PIA-FB0816; ICM P-05-004-F), the financial support from Conicyt-Chile (through grants Fondecyt/1181928, Newton-Picarte/MR/N026721/1, SERC Fondap/15110019) and "Sustainable gas pathways for Brazil; from microcosm to macrocosm" project supported by Natural Environment Research Council - NE/N018656/1.

Appendix A1: A linear approximation of the natural gas compressor

Eqs. (a.1)–(a.5) represent the relationship among natural gas flow $\boldsymbol{\Psi}$, the pressure at both ends of the compressor ($\boldsymbol{\pi}_{in}$ and $\boldsymbol{\pi}_{out}$) and the compression ratio of a natural gas compressor $\boldsymbol{\beta}$. $\boldsymbol{LC}$ is the load consumption that represents the natural gas that is required for the compressor to operate. The complete model is in [23–26].

$$\boldsymbol{\pi}_{in}\cdot\boldsymbol{\beta} = \boldsymbol{\pi}_{out} \tag{a.1}$$

$$\boldsymbol{LC} = \gamma\cdot\left(\left(\frac{\boldsymbol{\pi}_{out}}{\boldsymbol{\pi}_{in}}\right)^{\alpha} - 1\right)\cdot|\boldsymbol{\Psi}| \tag{a.2}$$

$$1 \le \boldsymbol{\beta} \le \bar{\beta} \tag{a.3}$$

$$\underline{\pi} \le \boldsymbol{\pi}_{in},\ \boldsymbol{\pi}_{out} \le \bar{\pi} \tag{a.4}$$

$$-\bar{\Psi} \le \boldsymbol{\Psi} \le \bar{\Psi} \tag{a.5}$$

Let $\boldsymbol{\Delta\pi}$ be the difference between nodal pressures, i.e., $\boldsymbol{\Delta\pi} = \boldsymbol{\pi}_{out} - \boldsymbol{\pi}_{in}$. Note that delta can also be expressed as a function of the compressor ratio, i.e., $\boldsymbol{\Delta\pi} = (\boldsymbol{\beta} - 1)\cdot\boldsymbol{\pi}_{in}$, and the upper and lower bounds of $\boldsymbol{\Delta\pi}$ are given by 0 and $(\bar{\beta} - 1)\cdot\boldsymbol{\pi}_{in}$ respectively. The first order approximation of (a.2) is given by (a.6) where $\partial LC\big/\partial\Psi = \gamma\cdot\left((\Delta\pi_0/\pi_0 + 1)^{\frac{\alpha}{2}} - 1\right)$, and $\partial LC/\partial\Delta\pi = \alpha\cdot\Psi_0/2\cdot(\pi_0 + \Delta\pi_0)\cdot(\partial LC/\partial\Psi + \gamma)$.

$$\boldsymbol{LC} \approx \frac{\partial LC}{\partial\Psi}\cdot\boldsymbol{\Psi} + \frac{\partial LC}{\partial\Delta\pi}\cdot\left(\boldsymbol{\Delta\pi} - \frac{\Delta\pi_0}{\pi_0}\cdot\boldsymbol{\pi}_{in}\right) \tag{a.6}$$

Eq. (a.6) is only valid for a given direction of $\boldsymbol{\Delta\pi}$ and $\boldsymbol{\Psi}$. That is, it is valid if $\boldsymbol{\Delta\pi} > 0$ and $\boldsymbol{\Psi} > 0$ that in practical terms mean that the natural gas can only flow through the compressor from node i to node j. However, the compressor in reality can redirect the flow so that natural gas might flow from node j to node i. In such a case, (a.6) must be adapted so that the bi-directional ability of the compressor can be adequately model.

Let us assume, without loss of generality, that natural gas flows through the compressor from node i to node j that implies that $\boldsymbol{\pi}_{in} = \boldsymbol{\pi}_i$ and $\boldsymbol{\pi}_{out} = \boldsymbol{\pi}_j$. Hence, Eqs. (a.7)–(a.10) is a linear approximation of the compressor that models the bi-directional ability where $\boldsymbol{LCP}$ and $\boldsymbol{LNP}$ denote the load consumption for a positive direction ($\boldsymbol{\Delta\pi} > 0$ and $\boldsymbol{\Psi} > 0$) and a negative direction ($\boldsymbol{\Delta\pi} < 0$ and $\boldsymbol{\Psi} < 0$)respectively.

$$\boldsymbol{\Delta\pi} = \boldsymbol{\pi}_j - \boldsymbol{\pi}_i \tag{a.7}$$

$$\underline{\pi} \le \boldsymbol{\pi}_i,\ \boldsymbol{\pi}_j \le \bar{\pi} \tag{a.8}$$

$$-(\bar{\beta} - 1)\cdot\boldsymbol{\pi}_j \le \boldsymbol{\Delta\pi} \le (\bar{\beta} - 1)\cdot\boldsymbol{\pi}_i \tag{a.9}$$

$$\boldsymbol{LC} = \begin{cases} \boldsymbol{LCP} = \frac{\partial LC}{\partial\Psi}\cdot\boldsymbol{\Psi} + \frac{\partial LC}{\partial\Delta\pi}\cdot\left(\boldsymbol{\Delta\pi} - \frac{\Delta\pi_0}{\pi_0}\cdot\boldsymbol{\pi}_i\right) & , if \quad (\boldsymbol{\Delta\pi} > 0)\&(\boldsymbol{\Psi} > 0) \\ \boldsymbol{LCN} = -\frac{\partial LC}{\partial\Psi}\cdot\boldsymbol{\Psi} + \frac{\partial LC}{\partial\Delta\pi}\cdot\left(-\boldsymbol{\Delta\pi} - \frac{\Delta\pi_0}{\pi_0}\cdot\boldsymbol{\pi}_j\right) & , if \quad (\boldsymbol{\Delta\pi} < 0)\&(\boldsymbol{\Psi} < 0) \\ 0 & , if \quad (\boldsymbol{\Delta\pi} = 0)|(\boldsymbol{\Psi} = 0) \end{cases} \tag{a.10}$$

Eqs. (a.11)–(a.16) replace (11) to model the compressor in optimization problems. Note that the formulation (a.11)–(a.16) incorporate multiple hyperplanes that correspond to different linear approximation around several points for improving the linear approximation. Sigma ($\boldsymbol{\sigma}$) is used to indicate the state of the compressor (on for $\sigma = 1$ and off for $\sigma = 0$), $BIGM$ is a scalar of large magnitude, and NLS is the number of points. Sigma ($\boldsymbol{\sigma}$) is a binary decision variable.

$$0 \le \boldsymbol{LC} \le BIGM\cdot\boldsymbol{\sigma} \tag{a.11}$$

$$-(\bar{\pi} - \underline{\pi})\cdot\boldsymbol{\sigma} \le \boldsymbol{\Delta\pi} \le (\bar{\pi} - \underline{\pi})\cdot\boldsymbol{\sigma} \tag{a.12}$$

$$\boldsymbol{LCP}_z \le \boldsymbol{LC};\ \forall\ z \in \{1,...,NLS\} \tag{a.13}$$

$$\boldsymbol{LCN}_z \le \boldsymbol{LC};\ \forall\ z \in \{1,...,NLS\} \tag{a.14}$$

$$-BIGM\cdot(1 - \boldsymbol{\sigma}) \le \boldsymbol{LCP}_z - \left(\frac{\partial LC}{\partial\Psi_z}\cdot\boldsymbol{\Psi} + \frac{\partial LC}{\partial\Delta\pi_z}\cdot\left(\boldsymbol{\Delta\pi} - \frac{\Delta\pi_z}{\pi_z}\cdot\boldsymbol{\pi}_i\right)\right) \le BIGM\cdot(1 - \boldsymbol{\sigma});\ \forall\ z \in \{1,...,NLS\} \tag{a.15}$$

$$-BIGM\cdot(1-\boldsymbol{\sigma}) \le \boldsymbol{LCN}_z - \left(-\frac{\partial LC}{\partial \Psi_z}\cdot\boldsymbol{\Psi} + \frac{\partial LC}{\partial \Delta\pi_z}\cdot\left(-\boldsymbol{\Delta\pi} - \frac{\Delta\pi_z}{\pi_z}\cdot\boldsymbol{\pi}_j\right)\right) \le BIGM\cdot(1-\boldsymbol{\sigma});\ \forall\ z \in \{1,...,NLS\} \tag{a.16}$$

Finally, and for illustration purposes, Fig. 3(a) is the ***LC***-Eq. (A2)- for a compressor, and Fig. 3(b) shows our proposed approximation –Eqs. (a.7)–(a.9) and (a.11)–(a.16)– if NLS = 2 and the input pressure is assumed to be constant. Fig. 3(c) shows what is often used in the academic literature as a linear approximation for the compressor that assumes a consumption between 3% and 5% of natural gas flow through it [4,17]. Observe that our approach leads to a lower error that in fact is shown in Fig. 4. Our approach, therefore, is closer -and more realistic- to the real operational cost that, in fact, provides more accurate results.

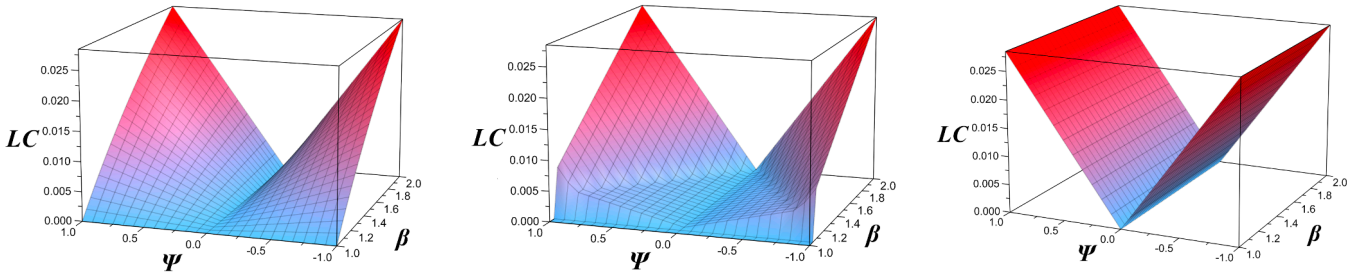

Fig. 3. ***LC*** comparison between (a) real consumption, (b) our approach, and (c) academic literature assumption.

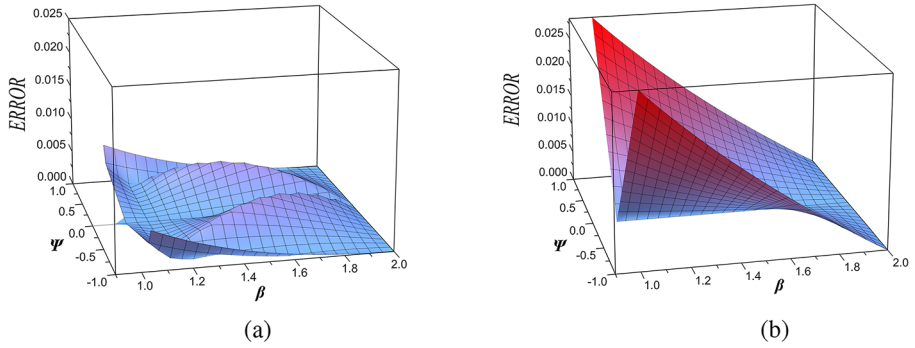

Fig. 4. Error comparison between (a) our approach and (b) academic literature assumption.

Appendix A2: A linear approximation of the pipeline

Eq. (b.1) shows the natural gas flow ($\boldsymbol{\Psi}$) through a pipeline as a function of nodal pressure ($\boldsymbol{\pi}$) and the duct resistance (C) [4,23–28]. A standard linear approximation is in [23,27,28] that requires a considerable number of linear segments -and binary variables- to have a low error. As a result, computation times of using this approximation are high. Our approach, on the contrary, utilizes line segments of the different size that implies lower approximation error, as it proves in this appendix.

$$sign(\boldsymbol{\Psi})\cdot\boldsymbol{\Psi}^2 = C^2\cdot(\pi_{in}^2 - \pi_{out}^2) \tag{b.1}$$

Let us define an isobaric curve (IC) that is obtained from (b.1) and resulting from solving π_{out} for a constant value of π_{in} i.e., $\pi_{out}=(\pi_{in}^2 - \boldsymbol{\Psi}^2/C^2)^{0.5} = (K - \boldsymbol{\Psi}^2/C^2)^{0.5}$ where K denotes a constant value. Fig. 5 shows various IC –continuous blue lines– where each of them is drawn for a particular value of K. The green IC corresponds to $K = \bar{\pi}$.

The dotted red line in Fig. 5 is a linear piecewise approximation of an IC that is obtained as indicated by (b.2). The parameter m_l is the slope of an approximation segment and $\boldsymbol{\Delta}_l$ is its amplitude. Let us call the dotted black lines as "cut," and each cut is drawn from the origin to one point of the green curve. *NLS* is the number of segments or the number of cuts plus one, and it is an input parameter for the linear approximation.

$$\pi_{out} = \pi_{in} + \sum_{l=1}^{NLS} m_l\cdot\boldsymbol{\Delta}_l \tag{b.2}$$

Note that Eq. (b.2) can approximate any IC, as long as the $\boldsymbol{\Delta}_l$ strictly meet the following conditions:

(1) The summation of $\boldsymbol{\Delta}_l$ must be equal to the absolute value of the pipeline flow (Eqs. (b.3)–(b.6)).

$$\boldsymbol{PF} - \boldsymbol{NF} = \boldsymbol{\Psi} \tag{b.3}$$

$$\boldsymbol{PF} + \boldsymbol{NF} = \sum_{l=1}^{NLS} \boldsymbol{\Delta}_l \tag{b.4}$$

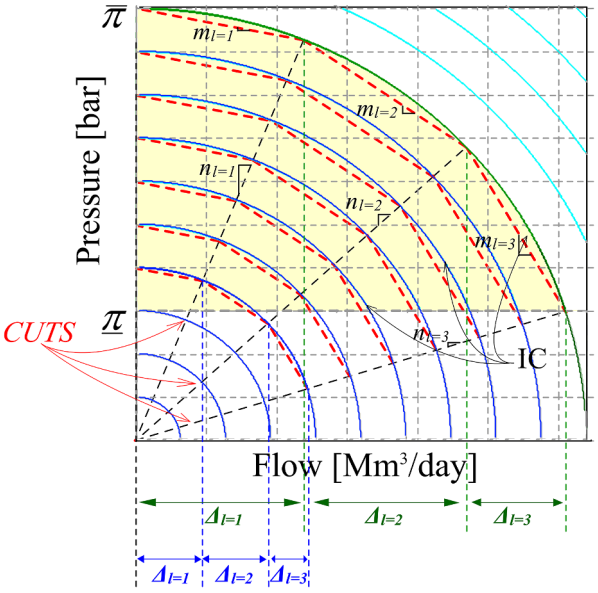

Fig. 5. Isobaric curves for a pipeline.

$$0 \leq \boldsymbol{PF} \leq BIGM \cdot \boldsymbol{S} \tag{b.5}$$

$$0 \leq \boldsymbol{NF} \leq BIGM \cdot (1 - \boldsymbol{S}) \tag{b.6}$$

(2) $\boldsymbol{\Delta}_l$ must be positive (Eq. (b.7)), and the upper limit of summation of the $\boldsymbol{\Delta}_l$ (Eq. (b.8)) is equal to the value given by the intersection between the approximation segments and the cuts as shown in Fig. 5. n_l is the slope of the cut.

$$0 \leq \boldsymbol{\Delta}_l; \forall\ l \tag{b.7}$$

$$\sum_{k=1}^{l} \boldsymbol{\Delta}_k \leq \frac{1}{n_l} \cdot \left(\pi_{in} + \sum_{k=1}^{l} m_k \cdot \boldsymbol{\Delta}_k \right); \forall\ l \tag{b.8}$$

(3) The next segment of $\boldsymbol{\Delta}_l$ is added if and only if all the previous segments are at its upper bound. This requirement implies the use of auxiliary binary variables (denoted as $\boldsymbol{z}_l$) and continuous variables (denoted as $\boldsymbol{\xi}_l$) as indicated by (b.9)–(b.12). Note that (B9)–(20) also define the lower bound of the summation of the $\boldsymbol{\Delta}_l$.

$$\boldsymbol{\xi}_l \leq \sum_{k=1}^{l-1} \boldsymbol{\Delta}_k; \forall\ l > 1 \tag{b.9}$$

$$\boldsymbol{\Delta}_l \leq BIGM \cdot \boldsymbol{z}_l; \forall\ l > 1 \tag{b.10}$$

$$-BIGM \cdot (1 - \boldsymbol{z}_l) \leq n_{l-1} \cdot \boldsymbol{\xi}_l - \left(\pi_{in} + \sum_{k=1}^{l-1} m_k \cdot \boldsymbol{\Delta}_k \right) \leq BIGM \cdot (1 - \boldsymbol{z}_l); \forall\ l > 1 \tag{b.11}$$

$$-BIGM \cdot \boldsymbol{z}_l \leq \boldsymbol{\xi}_l \leq BIGM \cdot \boldsymbol{z}_l; \forall\ l > 1 \tag{b.12}$$

On the other hand, Eqs. (b.13)–(b.16) relate the nodal pressures at both ends of the pipeline to the pressures through it. It is assumed, without loss of generality, that the natural gas flow from node i to j.

$$-BIGM \cdot (1 - \boldsymbol{S}) \leq \pi_i - \pi_{in} \leq BIGM \cdot (1 - \boldsymbol{S}) \tag{b.13}$$

$$-BIGM \cdot (1 - \boldsymbol{S}) \leq \pi_j - \pi_{out} \leq BIGM \cdot (1 - \boldsymbol{S}) \tag{b.14}$$

$$-BIGM \cdot \boldsymbol{S} \leq \pi_j - \pi_{in} \leq BIGM \cdot \boldsymbol{S} \tag{b.15}$$

$$-BIGM \cdot \boldsymbol{S} \leq \pi_i - \pi_{out} \leq BIGM \cdot \boldsymbol{S} \tag{b.16}$$

Finally, Eqs. (b.2)–(b.16) are our piecewise-linear representation of (b.1). Our model shows a lower approximation error as shown in Fig. 6. Observe how the error is substantially reduced for low pressures and low flows when compared with the traditional approximation [23,27,28]. The black lines are our approach and the red lines are what it is often used as the approximation. Additionally, we used fewer binary variables than [23,27,28] as a consequence of (b.9)–(b.12). It implies lower computational time than the traditional approach. Fig. 6 uses NLS = 3 for comparison purposes.

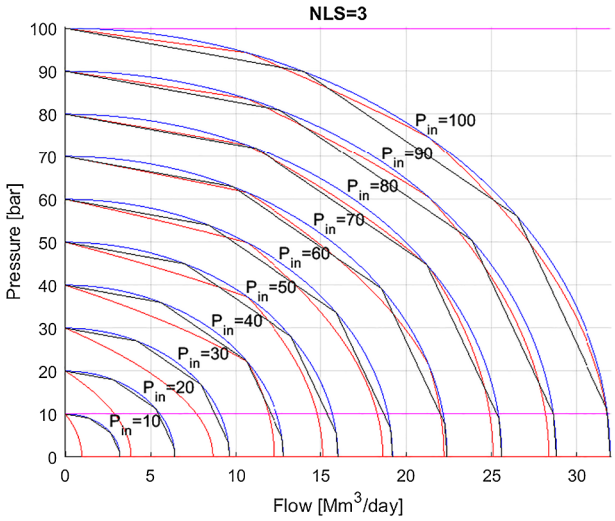

Fig. 6. A piecewise-linear representation comparison.

References

[1] UPME. Energy Planning Unit of Colombia. Plan Energético Nacional Colombia - Ideario Energético 2050. UPME. Bogotá, Colombia. [Online]; 2015, Jan. Available: < https://www.iea.org/media/pams/colombia/EnergyPlan2050PlanEnergeticoNacionalColombia2050.pdf > [in Spanish].

[2] Moreno R, Matus M, Flores A, Püschel S. Análisis Económico del Despacho Eléctrico de Generadores con Contratos de Suministro de Combustible GNL Take or Pay. Centro de Energía- Universidad de Chile. Santiago, Chile. [Online]; 2014, Dec. Available: < https://www.cne.cl/wp-content/uploads/2015/07/CNE-CE-ToP-InformeFinal_vf.pdf > .

[3] Schlag N, Olson A, Kwok G, Ming Z, Bolze M, Hemingway K. Natural gas infrastructure adequacy in the western interconnection: an electric system perspective. energy and environmental economics, Inc. San Francisco, CA. [Online]; 2014, July. Available: < http://westernenergyboard.org/wp-content/uploads/2014/07/E3_WIEB_Ph2_Report_full_7-28-2014.pdf > .

[4] Saldarriaga CA, Salazar H. Security of the Colombian Energy Supply: the need for liquefied natural gas regasification terminals for power and natural gas sectors. Energy 2016;100:349–62. https://doi.org/10.1016/j.energy.2016.01.064.

[5] Sánchez CB, Bent R, Backhaus S, Blumsack S, Hijazi H, Hentenryck Pv, Convex optimization for joint expansion planning of natural gas and power systems. In: 2016 49th Hawaii international conference on system sciences (HICSS), Koloa, HI; 2016. p. 2536–45. 10.1109/HICSS.2016.317.

[6] Zhang X, Shahidehpour M, Alabdulwahab A, Abusorrah A. Optimal Expansion planning of energy hub with multiple energy infrastructures. IEEE Trans Smart Grid 2015;6(5):2302–11. https://doi.org/10.1109/TSG.2015.2390640.

[7] Unsihuay-Vila C, Marangon-Lima JW, de Souza ACZ, Perez-Arriaga IJ, Balestrassi PP. A model to long-term, multiarea, multistage, and integrated expansion planning of electricity and natural gas systems. IEEE Trans Power Syst 2010;25(2):1154–68. https://doi.org/10.1109/TPWRS.2009.2036797.

[8] Saldarriaga CA, Hincapié RA, Salazar H. A holistic approach for planning natural gas and electricity distribution networks. IEEE Trans Power Syst 2013;28(4):4052–63. https://doi.org/10.1109/TPWRS.2013.2268859.

[9] Zhang X, Shahidehpour M, Alabdulwahab AS, Abusorrah A. Security-constrained co-optimization planning of electricity and natural gas transportation infrastructures. IEEE Trans Power Syst 2015;30(6):2984–93. https://doi.org/10.1109/TPWRS.2014.2369486.

[10] Munoz FD, Hobbs BF, Ho JL, Kasina S. An engineering-economic approach to transmission planning under market and regulatory uncertainties: WECC case study. IEEE Trans Power Syst 2014;29(1):307–17. https://doi.org/10.1109/TPWRS.2013.2279654.

[11] Konstantelos I, Strbac G. Valuation of flexible transmission investment options under uncertainty. IEEE Trans Power Syst Mar. 2015;30(2):1047–55. https://doi.org/10.1109/TPWRS.2014.2363364.

[12] Singh KJ, Philpott AB, Wood RK. Dantzig-Wolfe decomposition for solving multistage stochastic capacity-planning problems. Oper Res 2009;57(5):1271–86.

[13] Park H, Baldick R. Transmission planning under uncertainties of wind and load: sequential approximation approach. IEEE Trans Power Syst 2013;28(3):2395–402. https://doi.org/10.1109/TPWRS.2013.2251481.

[14] Moreira A, Strbac G, Moreno R, Street A, Konstantelos I. A five-level MILP model for flexible transmission network planning under uncertainty: a min-max regret approach. IEEE Trans Power Syst 2018;33(1):486–501. https://doi.org/10.1109/TPWRS.2017.2710637.

[15] Moreno R, Street A, Arroyo JM, Mancarella P. Planning low-carbon electricity systems under uncertainty considering operational flexibility and smart grid technologies. Phil Trans R Soc A 2017;375:20160305. https://doi.org/10.1098/rsta.2016.0305.

[16] Shao C, Shahidehpour M, Wang X, Wang X, Wang B. Integrated planning of electricity and natural gas transportation systems for enhancing the power grid resilience. IEEE Trans Power Syst 2017;32(6):4418–29. https://doi.org/10.1109/TPWRS.2017.2672728.

[17] He C, Wu L, Liu T, Bie Z. Robust co-optimization planning of interdependent electricity and natural gas systems with a joint N-1 and probabilistic reliability criterion. IEEE Trans Power Syst 2017;PP(99). https://doi.org/10.1109/TPWRS.2017.2727859. 1-1.

[18] Ding T, Hu Y, Bie Z. Multi-stage stochastic programming with nonanticipativity constraints for expansion of combined power and natural gas systems. IEEE Trans Power Syst 2018;33(1):317–28. https://doi.org/10.1109/TPWRS.2017.2701881.

[19] Qiu J, et al. Multi-stage flexible expansion co-planning under uncertainties in a combined electricity and gas market. IEEE Trans Power Syst 2015;30(4):2119–29. https://doi.org/10.1109/TPWRS.2014.2358269.

[20] Saldarriaga-Cortés CA, Salazar H, Moreno R, Jiménez-Estévez G. Integrated planning of electricity and natural gas systems under uncertain hydro inflows: a multi-objetive approach. In: Proc IEEE/PES general meeting; 2017. p. 1–5.

[21] Qiu J, Dong ZY, Zhao JH, Meng K, Zheng Y, Hill DJ. Low Carbon oriented expansion planning of integrated gas and power systems. IEEE Trans Power Syst 2015;30(2):1035–46. https://doi.org/10.1109/TPWRS.2014.2369011.

[22] Zhao B, Conejo AJ, Sioshansi R. Coordinated expansion planning of natural gas and electric power systems. IEEE Trans Power Syst 2017;PP(99). https://doi.org/10.1109/TPWRS.2017.2759198. 1-1.

[23] Correa-Posada CM. Optimal security-constrained model for the integrated power and natural-gas system, [PhD dissertation]. Madrid, Spain: Universidad Pontificia Comillas de Madrid Escuela Técnica Superior de Ingeniería (ICAI); 2015. Available from: < https://repositorio.comillas.edu/xmlui/handle/11531/2042?locale-attribute=en > [accessed 27.06.18].

[24] Hu Y, Bie Z, Ding T, Lin Y. An NSGA-II based multi-objective optimization for combined gas and electricity network expansion planning. Appl Energy 2015;167:280–93. https://doi.org/10.1016/j.apenergy.2015.10.148.

[25] Qiao Z, Guo Q, Sun H, Pan Z, Liu Y, Xiong W. An interval gas flow analysis in natural gas and electricity coupled networks considering the uncertainty of wind power. Appl Energy 2016;201:343–53. https://doi.org/10.1016/j.apenergy.2016.12.020.

[26] Qiu J, Zhao J, Yang H, Wang D, Dong ZY. Planning of solar photovoltaics, battery energy storage system and gas micro turbine for coupled micro energy grids. Appl Energy 2017;219:361–9. https://doi.org/10.1016/j.apenergy.2017.09.066.

[27] Correa-Posada CM, Sánchez-Martin P. Security-constrained optimal power and natural-gas flow. IEEE Trans Power Syst 2014;29(4):1780–7.

[28] Correa-Posada CM, Sánchez-Martín P. Integrated power and natural gas model for energy adequacy in short-term operation. IEEE Trans Power Syst 2015;30(6):3347–55.

[29] Zou J, Ahmed S, Sun XA. Stochastic dual dynamic integer programming. Math Program 2018:1–42.