Título

Cristian Alejandro Blanco Martínez

Trabajo de grado presentado como requisito
parcial para optar al título de
Magíster en Ingeniería Eléctrica

Director
Ph.D David Augusto Cárdenas Peña

UNIVERSIDAD TECNOLÓGICA DE PEREIRA
Facultad de Ingenierías
Programa de Ingeniería Eléctrica
Pereira, August 20, 2024

Nota de Aceptación

_____

_____

_____

_____

_____

Ph.D David Augusto Cárdenas Peña. Director

_____

Jurado

Pereira, August 20, 2024

Título
©Cristian Alejandro Blanco Martínez

Acknowledgments

Abstract

Contents

List of Figures

Chapter 1

Introduction

1.1 Problem statement

Natural gas transportation is an integral part of the natural gas industry, relying on a pipeline network to transfer natural gas from various sources to consumers, fulfilling their demand. This network is divided into two main types: transmission and distribution. Transmission networks transport large volumes of gas at high pressure over long distances from gas sources to distribution centers. On the other hand, distribution networks deliver gas to individual consumers [2]. In general, natural gas transmission systems are composed of four fundamental elements: injection fields, responsible for injecting the hydrocarbon from extraction fields or regasification plants into the system; pipelines, which transport the gas from a sending node to a receiving node; compressors, which are responsible for raising the pressure at the outlet node relative to the inlet node; and end user [3]. Ensuring gas flow to meet end-user demand, minimizing network operating costs, and maintaining system elements within appropriate technical operating limits are critical factors in natural gas transportation. Coordinating these factors requires formulating optimization problems, which must be solved efficiently, taking into account the numerous variables and the nature of these variables [4].

The optimization problem is determining the best operational configurations to meet consumer demand while ensuring the technical and physical constraints of the natural gas transportation system. It must also be considered that these transportation systems are usually interconnected with the electricity systems since the latter usually require natural gas as fuel for the thermal power plants. These power plants are significant natural gas consumers, relying on a steady supply to generate electricity [5].

The Colombian case is no exception; although the Colombian energy matrix comprises 70% hydroelectric plants, the remainder consists primarily of thermoelectric plants [6]. These thermoelectric plants are crucial for complementing the hydro plants to meet energy demand, especially during periods of drought. They become essential during events like the El Niño phenomenon when reduced water availability limits hydroelectric generation [7]. As other studies have shown, variations in rainfall, droughts, or floods in countries with high hydroelectric power plants can significantly affect water availability for hydropower production [8].

The above situation necessitates solving the optimization problem multiple times to ensure the system's correct operation across various scenarios. Consequently, this process takes considerable time and is both resource-intensive and time-consuming. Despite the high computational cost of each model execution, the resulting solutions are not utilized in subsequent optimization processes, even in similar operational scenarios. Therefore, there is a pressing need to develop a machine learning strategy that leverages historical solutions to provide faster responses to different operational scenarios by learning from past optimization outcomes.

Although production fields, compressors, and end users of natural gas are well-represented, modeling transmission pipelines remains complex due to the nonlinear relationship between flow and pressures at its ending nodes. This complexity arises from the Weymouth equation, which includes a nonconvex and discontinuous sign function that determines flow direction based on differential pressure. These nonconvexities introduce discontinuities lead to numerical issues and optimization instability [9, 10]. Various authors have approached the challenge posed by the Weymouth equation differently. One of the most widely accepted methods involves approximating this equation due to its inherent complexity and nonconvex nature. However, since it is an approximation, this solution introduces errors that impact the accuracy of optimization outcomes. Mitigating these errors remains critical for further research and development in natural gas transportation systems [3].

1.2 Justification

Natural gas is an energy source that has gained great relevance worldwide, and this can be attributed to two fundamental causes. Firstly, it has been observed that a country's economic growth is closely related to its energy consumption [11]. Therefore, as nations develop and grow economically, it is expected that they will seek energy security to meet their own demand and continue their progress without interruptions. The second major

motivation for the use of natural gas is its lower greenhouse gas emissions compared to other fuels, making it a favorable option, especially in a context where there is a growing interest in environmental care. Natural gas emits fewer greenhouse gases compared to other fossil fuels, making it a favorable option for climate change mitigation [12]. In this context, the natural gas system plays a crucial role in providing clean and versatile energy [13]. It is an efficient and less polluting energy source compared to conventional fossil fuels, such as coal and oil [14]. Its use is essential for electricity generation, residential and industrial heating, and also for supplying energy-intensive industrial sectors.

According to figures from the U.S. Energy Information Administration (EIA), global natural gas consumption in 2015 reached 124.24 trillion cubic feet, with a projected increase of 43% by 2040, where 75% is associated with the industrial sector and electricity generation based on thermal power plants. This pronounced increase in consumption contrasts with the total volume of proven reserves worldwide, which reached 154.53 trillion cubic feet in 2015 and projects a growth of about 52% by 2014 [15]. Colombia does not have interests different from those mentioned above, especially when considering the potential environmental impacts [16]. Therefore, it is necessary for the country not only to have a national gas transportation system but also to ensure that it is operated in the best possible way, so that natural resources are maximized. In the Colombian context, natural gas is a very important energy source as it is used in various sectors such as residential, commercial, industrial, and thermal [17]. It is especially in the latter sector that this fuel becomes more relevant during dry seasons, as it is when reservoir levels drop and thus hydroelectric power generation decreases. This problem is exacerbated in years when the El Niño phenomenon occurs [18], making it of great interest to have tools that allow for the optimal injection and transportation of natural gas to fully meet demand.

Although most of the country's electricity demand is commonly met by hydroelectric plants [19], this type of generation presents a significant source of uncertainty in the energy system since its effectiveness and generation capacity are directly linked to the country's climatic and meteorological conditions, especially in extreme cases such as the El Niño phenomenon [20]. Variations in precipitation, droughts, or floods can have a significant impact on the availability of water for hydroelectric power production, affecting the balance between supply and demand in the electrical system [21]. Additionally, the increase in energy demand and the transition to renewable energy sources pose significant challenges in the efficient and reliable transportation of gas. Optimizing the natural gas transportation system, considering the uncertainty associ-

ated with renewable energy generation and demand variability, is essential to ensure a reliable, sustainable, and environmentally friendly energy supply [22].

Chapter 2

Power System - Censnet

2.1 Preliminaries

2.1.1 Graph definition

A **graph** G is a mathematical structure that represents a set of interconnected objects. These objects are known as **vertices** (or **nodes**), denoted by the set $V(G)$, and the connections between them are called **edges** (or **arcs**), denoted by the set $E(G)$. Formally, a graph is defined as an ordered pair $G = (V, E)$, where $V(G)$ is a non-empty set of vertices, and $E(G) \subseteq \{(u, v) \mid u, v \in V(G), u \neq v\}$ is a set of edges, where each edge connects two distinct vertices.

Graphs can be categorized based on the properties of their edges. An **undirected graph** has edges that do not have a direction, so the pair $(u, v) = (v, u)$ represents an edge that simply connects vertices u and v. In contrast, in a **directed graph** (or **digraph**), each edge $(u, v) \in E(G)$ has a direction, meaning it goes from vertex u to vertex v. This implies that $(u, v) \neq (v, u)$ unless $u = v$

In Figure 2.1, two graphs are represented, each composed of four nodes labeled 1, 2, 3, and 4,, and six edges labeled A, B, C, D, E, and F. The difference between them lies in the type of graph they represent. For example, in Figure 2.1a, the edge c shows a connection between nodes 2 and 4. However, in Figure 2.1b, this connection provides additional information: a direction, which, in the context of this study, could represent the direction of a specific element, such as electric power or flow.

A graph can be represented in various ways using matrices, each capturing different aspects of the graph's structure. The two most common matrix representations are the **adjacency matrix** and the **incidence matrix**.

(a) Undirected graph (b) Directed graph
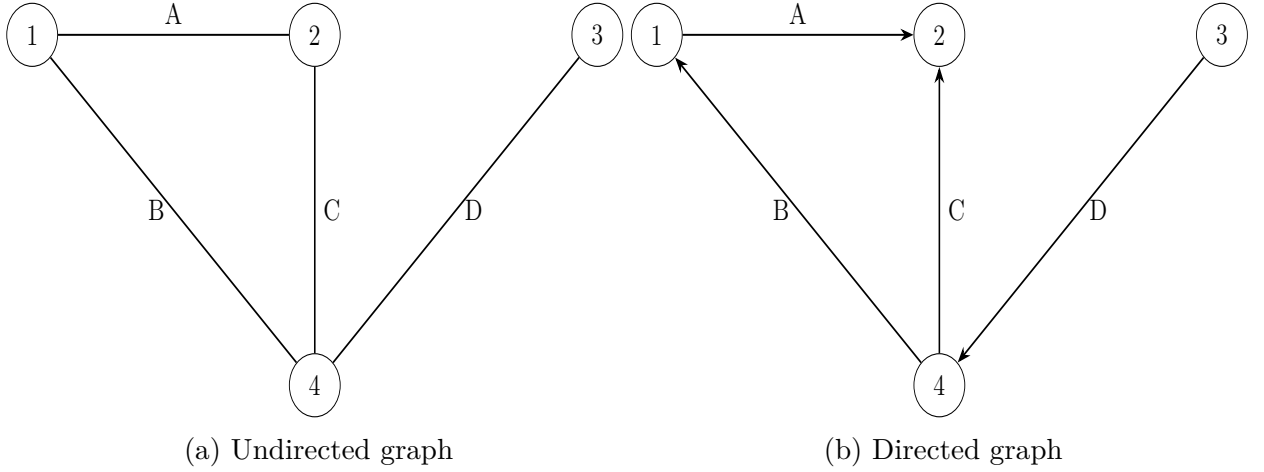
Figure 2.1: Types of graphs

The **adjacency matrix** of a graph is a square matrix used to represent the connections between vertices. For a graph G with n vertices, the adjacency matrix A is an $n \times n$ matrix where the entry a_{ij} is defined as follows:

$$a_{ij} = \begin{cases} 1 & \text{if there is an edge from vertex } i \text{ to vertex } j, \\ 0 & \text{otherwise.} \end{cases}$$

For a directed graph, the adjacency matrix captures the direction of the edges. Below is the adjacency matrix for the directed graph shown earlier:

$$A = \begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}$$

The **incidence matrix** of a graph represents the relationship between vertices and edges. For a graph G with n vertices and m edges, the incidence matrix I is an $n \times m$ matrix where the entry i_{ij} is defined as follows:

$$i_{ij} = \begin{cases} 1 & \text{if vertex } i \text{ is the starting point of edge } j \text{ in a directed graph,} \\ -1 & \text{if vertex } i \text{ is the endpoint of edge } j \text{ in a directed graph,} \\ 0 & \text{if vertex } i \text{ is not connected to edge } j. \end{cases}$$

For the directed graph previously described, the incidence matrix is given by:

$$I = \begin{pmatrix} 1 & -1 & 0 & 0 \\ -1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & -1 \end{pmatrix}$$

2.1.2 Neural networks

Multi-Layered Perceptrons

A Multilayer Perceptron (MLP) is a fundamental type of artificial neural network, often regarded as one of the building blocks of deep learning. At its core, an MLP consists of multiple layers of nodes, or neurons, where each layer is fully connected to the next one. The architecture typically includes an input layer, one or more hidden layers, and an output layer. The neurons in each layer are connected to the neurons in the subsequent layer through weighted connections, which are the key parameters learned during the training process. One of the most significant properties of an MLP is its ability to function as a universal approximator. This means that, given sufficient neurons in the hidden layers, an MLP can approximate any continuous function to an arbitrary degree of accuracy, provided the network is trained properly.

Mathematically, an MLP can be defined as follows. Let $\mathbf{x} \in \mathbb{R}^n$ represent the input vector, where n is the number of features. The output of each neuron in the first hidden layer is calculated as:

$$\mathbf{z}^{(1)} = \sigma \left(\mathbf{W}^{(1)} \mathbf{x} + \mathbf{b}^{(1)} \right)$$

where $\mathbf{W}^{(1)} \in \mathbb{R}^{m_1 \times n}$ is the weight matrix for the first hidden layer, with m_1 being the number of neurons in this layer, $\mathbf{b}^{(1)} \in \mathbb{R}^{m_1}$ is the bias vector, and $\sigma(\cdot)$ is the activation function, typically a non-linear function such as the ReLU (Rectified Linear Unit) or sigmoid function.

This process is repeated for each subsequent hidden layer k, where the output of the k-th layer is given by:

$$\mathbf{z}^{(k)} = \sigma \left(\mathbf{W}^{(k)} \mathbf{z}^{(k-1)} + \mathbf{b}^{(k)} \right)$$

Here, $\mathbf{W}^{(k)} \in \mathbb{R}^{m_k \times m_{k-1}}$ represents the weight matrix connecting layer $k-1$ to layer k, $\mathbf{b}^{(k)} \in \mathbb{R}^{m_k}$ is the bias vector for layer k, and $\mathbf{z}^{(k-1)}$ is the output of the previous layer.
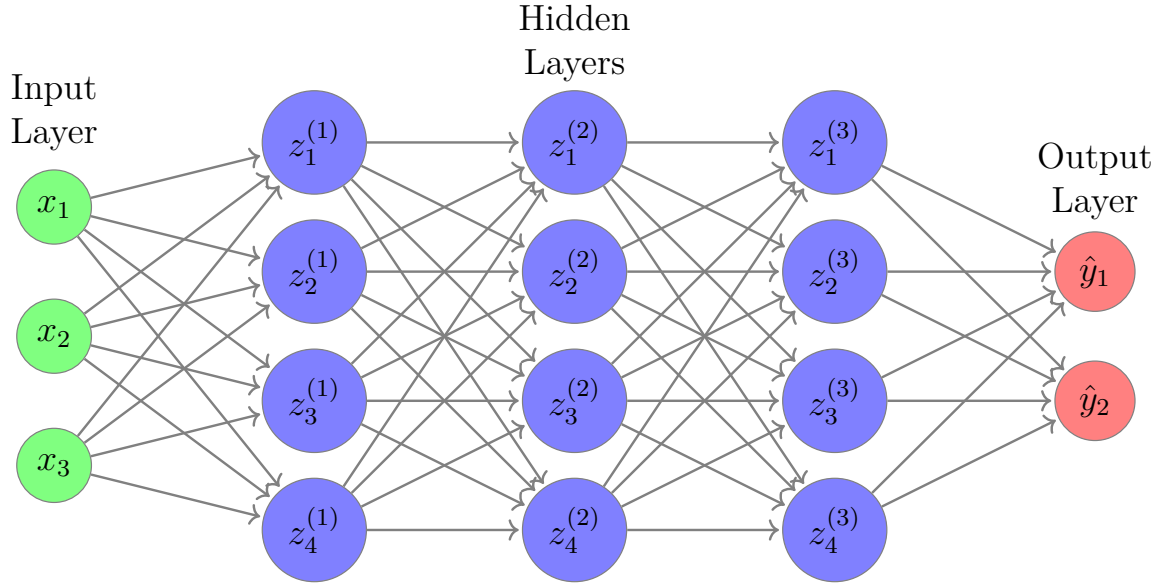
Figure 2.2

Finally, the output layer produces the final prediction $\hat{\mathbf{y}}$:

$$\hat{\mathbf{y}} = \sigma \left(\mathbf{W}^{(L)} \mathbf{z}^{(L-1)} + \mathbf{b}^{(L)} \right)$$

where L denotes the number of layers in the network, including the input and output layers. Depending on the nature of the problem (e.g., classification or regression), the activation function $\sigma(\cdot)$ used in the output layer can vary, with softmax being common in multi-class classification problems, and a linear activation for regression tasks.

The entire MLP is trained using a process called backpropagation, combined with an optimization algorithm like gradient descent, to minimize a loss function $J(\mathbf{y}, \hat{\mathbf{y}})$, which measures the difference between the true outputs $\mathbf{y}$ and the predicted outputs $\hat{\mathbf{y}}$.

Graph Neural Networks

In recent years, **Graph Neural Networks (GNNs)** have emerged as a powerful tool in machine learning, particularly for tasks involving data that can be naturally represented as graphs. Graphs are a universal data structure that can model various systems in numerous fields, including social networks, biological networks, knowledge graphs, and physical systems. Because of their ability to represent relationships and

interactions between entities, graphs are used extensively to model complex structures where the data points are not independent but interconnected.

The importance of GNNs lies in their ability to directly operate on graph-structured data, extending the success of neural networks from grid-like data structures, such as images and sequences, to more general and irregular structures. Traditional neural networks, like Convolutional Neural Networks (CNNs) or Recurrent Neural Networks (RNNs), are designed to work with data that has a fixed structure. However, many real-world problems involve data that can be better described by graphs, where nodes represent entities and edges represent relationships between those entities.

Graph Neural Networks can be broadly defined as a class of neural networks designed to perform inference on data described by graphs. Formally, let $G = (V, E)$ represent a graph, where V is the set of nodes (or vertices) and E is the set of edges. Each node $v \in V$ can be associated with a feature vector $\mathbf{x}_v$, and each edge $(u, v) \in E$ may have an associated weight or feature vector $\mathbf{e}_{uv}$. The goal of a GNN is to learn a representation for each node (or sometimes for the entire graph) by aggregating and transforming the feature information from the node's local neighborhood in the graph.

In the GNN framework, the process of message passing is understood as a series of iterations in which each node updates its representation by exchanging information with its neighbors. To move from the abstract concept to a practical implementation, it is necessary to define the specific functions used for updating and aggregating node features.

The basic message passing operation, which simplifies the original GNN model proposed by [], is expressed by the following equation:

$$
\mathbf{h}_u^{(k)} = \sigma \left(\mathbf{W}_{\text{self}}^{(k)} \mathbf{h}_u^{(k-1)} + \mathbf{W}_{\text{neigh}}^{(k)} \sum_{v \in \mathcal{N}(u)} \mathbf{h}_v^{(k-1)} + \mathbf{b}^{(k)} \right)
$$

In this equation:

- $\mathbf{h}_u^{(k)}$ represents the updated feature vector of node u at layer k.

- The term $\mathbf{W}_{\text{self}}^{(k)} \mathbf{h}_u^{(k-1)}$ applies a transformation to the node's own feature vector from the previous layer, enabling the node to retain and modify its self-information.

- The term $\mathbf{W}_{\text{neigh}}^{(k)} \sum_{v \in \mathcal{N}(u)} \mathbf{h}_v^{(k-1)}$ aggregates the feature vectors of the neighboring nodes v in the set $\mathcal{N}(u)$, and then applies a transformation via the weight matrix $\mathbf{W}_{\text{neigh}}^{(k)}$

- $\mathbf{b}^{(k)}$ is a bias term that can be added to the weighted sum, though it is sometimes omitted for simplicity.

- The non-linear function $\sigma(\cdot)$, such as ReLU or tanh, is applied elementwise to introduce non-linearity into the model, which is essential for capturing complex patterns in the data.

In the context of Graph Neural Networks (GNNs), a Graph Convolutional Network (GCN) is a specialized model that applies the concept of convolution, widely used in image processing, to graphs. First introduced by Thomas Kipf and Max Welling in 2017, GCNs offer a method to perform deep learning on graph-structured data by extending traditional convolution operations to the irregular domain of graphs.

The fundamental idea behind GCNs is to create a spectral filter that operates on graph data. The filter's purpose is to combine features from a node's local neighborhood, taking into account the graph's structure. This process is mathematically formalized in the following way:

$$\mathbf{H} = \sigma \left(\tilde{\mathbf{D}}^{-\frac{1}{2}} \tilde{\mathbf{A}} \tilde{\mathbf{D}}^{-\frac{1}{2}} \mathbf{X} \boldsymbol{\Theta} \right)$$

Where:

- $\mathbf{H}$ represents the matrix of node representations after applying the GCN layer. Each row $\mathbf{h}_u$ in $\mathbf{H}$ corresponds to the updated feature vector for node u.

- $\mathbf{X}$ is the matrix of input node features, where each row $\mathbf{x}_u$ corresponds to the feature vector for node u before applying the GCN layer.

- $\sigma(\cdot)$ denotes a non-linear activation function, such as ReLU, applied elementwise to introduce non-linearity into the model.

- $\tilde{\mathbf{A}}$ is the adjacency matrix of the graph, with added self-loops to account for the node itself in the aggregation.

- $\tilde{\mathbf{D}}$ is the degree matrix of the graph, modified to include the self-loops. The degree matrix is diagonal, with each diagonal entry $\tilde{D}_{ii}$ representing the degree of node i in the graph.

- $\boldsymbol{\Theta}$ is a matrix of trainable parameters, which is learned during the training process to optimize the model's performance.

The expression $\tilde{\mathbf{D}}^{-\frac{1}{2}}\tilde{\mathbf{A}}\tilde{\mathbf{D}}^{-\frac{1}{2}}$ is a normalized version of the adjacency matrix, ensuring that the eigenvalues of the operation are bounded between 0 and 1. This normalization step is crucial as it prevents issues such as exploding or vanishing gradients during the training of deep networks.

Specifically, the adjacency matrix $\tilde{\mathbf{A}}$ is defined as:

$$\tilde{\mathbf{A}} = \mathbf{A} + \mathbf{I}$$

where $\mathbf{A}$ is the original adjacency matrix, and $\mathbf{I}$ is the identity matrix. The identity matrix $\mathbf{I}$ ensures that each node considers its own features when aggregating information from its neighbors.

The degree matrix $\tilde{\mathbf{D}}$ is defined as:

$$\tilde{D}_{ii} = \sum_{j \in V} \tilde{A}_{ij}$$

where V represents the set of all nodes in the graph. The diagonal entries of $\tilde{\mathbf{D}}$ correspond to the degree of each node, adjusted to account for the added self-loops.

Convolution with Edge-Node Switching (CensNet)

Graph Convolutional Networks (GCNs) have demonstrated considerable success in various graph-based machine learning tasks, particularly in their ability to generalize convolution operations to non-Euclidean data structures like graphs. GCNs operate by aggregating features from a node's neighbors, thereby capturing local neighborhood information and propagating it through the network layers. Despite their effectiveness, GCNs possess certain limitations that hinder their performance in more complex scenarios.

One notable limitation of GCNs is their reliance solely on node features during the convolution process. This approach disregards the information contained within edge features. By neglecting edge features, GCNs fail to fully exploit the underlying structure of the graph, potentially missing out on critical insights that could enhance model performance.

Furthermore, GCNs typically aggregate information from immediate neighbors only, which can limit their ability to capture long-range dependencies in large or densely connected graphs. This restriction can lead to an oversimplified representation of the graph structure, particularly in cases where the graph contains intricate patterns that require deeper and more nuanced analysis.

To overcome the limitations of traditional GCNs, which focus primarily on node features, CensNet introduces a novel approach that integrates both node and edge features into the graph convolution process. The CensNet framework consists of two primary types of layers: the *node layer* and the *edge layer*. These layers work in tandem to update node and edge embeddings alternately, leveraging the information from both nodes and edges in the graph.

The propagation rules in CensNet are designed to incorporate edge features into the convolution process, enabling a more comprehensive feature propagation across the graph. We define the normalized node adjacency matrix with self-loops as follows:

$$\tilde{\mathbf{A}}_v = \mathbf{D}_v^{-\frac{1}{2}}(\mathbf{A}_v + \mathbf{I}_{N_v})\mathbf{D}_v^{-\frac{1}{2}}, \tag{2.1}$$

where $\mathbf{D}_v$ is the diagonal degree matrix of $\mathbf{A}_v + \mathbf{I}_{N_v}$.

Node Layer Propagation: In the $(l+1)$-th layer, the node features are updated using the following propagation rule:

$$H_v^{(l+1)} = \sigma\left(T\Phi\left(H_e^{(l)}P_e\right)T^\top \odot \tilde{A}_v H_v^{(l)}W_v\right)$$

Where,

- $T \in \mathbb{R}^{N_v \times N_e}$ is a binary transformation matrix that represents the connections between nodes and edges. Each element $T_{i,m}$ indicates whether edge m connects to node i. Specifically, if edge m is connected to node i, then $T_{i,m} = 1$; otherwise, $T_{i,m} = 0$. Given that each edge is formed by two nodes, every column of the matrix T will have exactly two elements equal to 1, corresponding to the two nodes that the edge connects.

- $H_e^{(l)}$ is the edge feature matrix from the l-th layer. P_e is a learnable vector of dimension d_e, which acts as a weight for the edge features. The operation $\Phi(H_e^{(l)}P_e)$ denotes the diagonalization of the vector $H_e^{(l)}P_e$, converting it into a diagonal matrix where the elements of the vector are placed on the diagonal.

- The Hadamard product, denoted by $\odot$, represents element-wise multiplication between matrices. In this context, it combines the transformed edge features with the node adjacency matrix, integrating information from both the original graph and its line graph.

- $\tilde{A}_v = D_v^{-\frac{1}{2}}(A_v + I_{N_v})D_v^{-\frac{1}{2}}$ is the normalized adjacency matrix for nodes, where A_v is the original node adjacency matrix and I_{N_v} is the identity matrix that introduces self-loops. This normalization ensures that the contributions from each node's neighbors are appropriately scaled.

- $H_v^{(l)}$ represents the node feature matrix from the l-th layer. W_v is a learnable weight matrix that is applied to the node features during the propagation process.

- The activation function σ (typically a non-linear function such as ReLU) is applied element-wise to the resulting matrix to introduce non-linearity into the model.

This expression can be interpreted as a way of fusing node and edge information. The matrix T maps edge features into the node domain, and this information is combined with the normalized node adjacency matrix $\tilde{A}_v$. This fusion creates a new node adjacency matrix that incorporates both node and edge features, which is then used to update the node embeddings.

Edge Layer Propagation: Similarly, the normalized (Laplacianized) edge adjacency matrix is defined as:

$$\tilde{A}_e = D_e^{-\frac{1}{2}}\left(A_e + I_{N_e}\right)D_e^{-\frac{1}{2}}, \tag{2.2}$$

where D_e is the degree matrix corresponding to the edge adjacency matrix $A_e + I_{N_e}$. The matrix $\tilde{A}_e$ serves as the normalized version of the edge adjacency matrix, similar to how the node adjacency matrix is normalized. This normalization ensures that the influence of each edge is scaled appropriately, which is crucial for the stability of the propagation process.

The propagation rule for edge features is defined as follows:

$$H_e^{(l+1)} = \sigma\left(T^\top \Phi\left(H_v^{(l)}P_v\right)T \odot \tilde{A}_e H_e^{(l)}W_e\right). \tag{2.3}$$

In this expression, the following components are involved:

- $T^\top$ is the transpose of the binary transformation matrix T used in the node layer propagation. The matrix $T^\top$ maps the node features back into the edge domain, allowing the edge features to be updated based on the node information.

- $H_v^{(l)}$ is the node feature matrix from the l-th layer, and P_v is a learnable weight matrix for the nodes. The operation $\Phi(H_v^{(l)}P_v)$ diagonalizes the product of node features and the learnable weights, similar to the transformation applied to edge features in the node layer propagation.

- The matrix $\tilde{A}_e$ is the normalized edge adjacency matrix, as defined above. This matrix integrates information about the connections between edges, analogous to how $\tilde{A}_v$ handles connections between nodes.

- $H_e^{(l)}$ represents the edge feature matrix from the l-th layer, while W_e is a learnable weight matrix that is applied to the edge features during the propagation.

- The Hadamard product $\odot$ element-wise multiplies the transformed node features with the edge adjacency matrix, merging the information from both domains.

- As in the node layer propagation, the activation function σ is applied element-wise to introduce non-linearity.

This propagation rule updates the edge embeddings by integrating information from the node features and the edge structure, thereby enhancing the expressiveness of the edge representations. The alternating updates between node and edge embeddings allow the model to effectively bridge signals across nodes and edges, leading to more robust and informative graph embeddings.

2.1.3 Task-Dependent Loss Functions

The output layer and corresponding loss functions in CensNet are designed to be task-dependent. For regression tasks, the loss function can be formalized as a regularized mean square error (MSE) loss. The MSE loss measures the difference between the predicted outcomes and the actual continuous values, providing a natural fit for regression problems.

We define the loss function for graph regression as follows:

$$\mathcal{L}(\Theta) = \sum_{l \in \mathcal{Y}_L} \sum_{f=1}^{F} \|Y_{lf} - \hat{Y}_{lf}\|_2^2 + \lambda \|\Theta\|_p, \tag{2.4}$$

where:

- Y_{lf} represents the true continuous value for the f-th feature of the l-th graph in the training set.

- $\hat{Y}_{lf}$ is the predicted outcome generated from the final node hidden layer of the CensNet model.

- $\|Y_{lf} - \hat{Y}_{lf}\|_2^2$ is the squared difference between the true and predicted values, summed across all features and all graphs in the training set.

- $\lambda\|\Theta\|_p$ is the regularization term, which helps control the model's complexity and prevents overfitting by penalizing large weights. The parameter λ controls the strength of the regularization, while p determines the type of regularization norm (e.g., $p = 2$ for L_2 regularization).

2.2 Linear formulation of the natural gas system

Natural gas is a widely used energy resource, particularly for electricity generation. The natural gas system consists of a network of production centers, pipelines, compressor stations, storage facilities, and distribution points that ensure reliable gas delivery from producers to consumers. Mathematically, this system can be represented as a directed graph defined as $\mathcal{G}_f = \{\mathcal{N}_f, \mathcal{E}_f\}$ where $\mathcal{N}_f$ is the set of units within the gas system, and $\mathcal{E}_f$ is the set of different elements linking them. This set of units includes gas supply nodes or wells $\mathcal{W} \subset \mathcal{N}_f$, gas demand nodes or users $\mathcal{U} \subset \mathcal{N}_f$, and gas storage facilities $\mathcal{S} \subset \mathcal{N}_f$. Similarly, the set of directed gas adjacency edges $\mathcal{A} = \{(n, m) \mid n, m \in \mathcal{N}_f\} \subset \mathcal{E}$ delineates the network structure through two kinds of transmission elements: transport pipelines $\mathcal{P} = \{p = (n, m) \mid n, m \in \mathcal{N}_f\}$ and compressing stations $\mathcal{C} = \{c = (n, m) \mid n, m \in \mathcal{N}_f\}$, so that $\mathcal{P} \cup \mathcal{C} = \mathcal{A}$ and $\mathcal{P} \cap \mathcal{C} = \emptyset$.

Natural gas transportation requires coordination to manage the flow through the different elements to maintain safe operating ranges. In optimizing this network, mathematical models minimize overall operating costs associated with the various stages of natural gas transportation, compression, storage, and handling unsupplied demand, ensuring compliance with technical and physical constraints. The function is expressed as:

$$\min_{\mathcal{P},\mathcal{F}} \quad \sum_{w\in\mathcal{W}} C_w^t f_w^t + \sum_{p\in\mathcal{P}} C_p^t f_p^t + \sum_{c\in\mathcal{C}} C_c^t f_c^t +$$

$$\sum_{u\in\mathcal{U}} C_u^t f_u^t + \quad \sum_{s\in\mathcal{S}} C_{s+}^t f_{s+}^t + \sum_{s\in\mathcal{S}} C_{s-}^t f_{s-}^t + \sum_{s\in\mathcal{S}} C_s^t V_s^t \tag{2.5}$$

The term $\sum_{w\in\mathcal{W}} C_w^t f_w^t$ represents the total cost of gas production at the wells, where C_w^t denotes the cost per unit flow of gas at a specific well w during time period t, and f_w^t corresponds to the flow of gas from well w. Similarly, the transportation of gas through pipelines is captured by the term $\sum_{p\in\mathcal{P}} C_p^t f_p^t$, where C_p^t is the cost per unit flow through pipeline p during time period t, and f_p^t represents the flow of gas through pipeline p. In addition, the total cost associated with gas compression at compressor stations is accounted for by $\sum_{c\in\mathcal{C}} C_c^t f_c^t$, where C_c^t is the cost per unit flow at compressor station c during time period t, and f_c^t is the flow of gas through compressor station c.

Beyond production, transportation, and compression, the model also considers the costs related to unmet gas demand. The term $\sum_{u\in\mathcal{U}} C_u^t f_u^t$ reflects the penalty cost associated with unsupplied gas demand, where C_u^t is the penalty cost per unit of unsupplied gas at location u during time period t, and f_u^t represents the volume of unmet demand. Additionally, the model includes costs related to storage operations. The term $\sum_{s\in\mathcal{S}} C_{s+}^t f_{s+}^t$ represents the cost of injecting gas into storage facilities, where C_{s+}^t is the cost per unit flow into storage during time period t, and f_{s+}^t denotes the flow into storage at facility s. Conversely, the cost of withdrawing gas from storage is captured by $\sum_{s\in\mathcal{S}} C_{s-}^t f_{s-}^t$, where C_{s-}^t is the cost per unit flow out of storage during time period t, and f_{s-}^t represents the flow out of storage at facility s. Finally, the model accounts for the storage cost itself through the term $\sum_{s\in\mathcal{S}} C_s^t V_s^t$, where C_s^t is the cost per unit volume of gas stored at facility s during time period t, and V_s^t denotes the volume of gas stored.

$$\underline{f_w^t} \leq f_w^t \leq \overline{f_w^t} \quad \forall \, w \in \mathcal{W} \tag{2.6}$$

$$-\overline{f_p^t} \leq f_p^t \leq \overline{f_p^t} \quad \forall \, p \in \mathcal{P} \tag{2.7}$$

$$0 \leq f_u^t \leq \overline{f_u^t} \quad \forall \, u \in \mathcal{U} \tag{2.8}$$

$$\sum_{m:(m,n)\in\mathcal{A}} f_m^t = \sum_{m':(n,m')\in\mathcal{A}} f_{m'}^t \quad \forall \, n \in \mathcal{N}_f \tag{2.9}$$

$$0 \leq f_{s+}^t \leq V_{0s} - \underline{V_s} \quad \forall \, s \in \mathcal{S} \tag{2.10}$$

$$0 \leq f_{s-}^t \leq \overline{V_s} - V_{0s} \quad \forall \, s \in \mathcal{S} \tag{2.11}$$

$$V_s^t = V_s^{t-1} + f_{s-}^{t-1} - f_{s+}^{t-1} \quad \forall \, s \in \mathcal{S} \tag{2.12}$$

$$\tag{2.13}$$

The constraint set models the gas transportation system: Equation (2.6) forces each production well to inject the flow f_w^t over the technical minimum $\underline{f_w^t}$ and under the maximum capacity $\overline{f_w^t}$. Equation (2.7) upper-bounds the gas flow through pipelines f_p^t to the structural capacity $\overline{f_p^t}$. Equation (2.8) ensures that the unsupplied demand f_u^t is lower than the corresponding user demand $\overline{f_u^t}$. The nodal gas balance in Equation (2.9) guarantees that the gas entering the node n equals the gas leaving it. Equations (2.10) and (2.11) limit the gas injection f_{s+} and extraction f_{s-} rates at storage facilities according to the feasible operating range determined by the currently stored volume V_s^t, respectively. In turn, Equation (2.12) balances the gas storage unit such that gas volume at operation period t V_s^t equals the volume from period V_s^{t-1} plus the difference between injected f_{s+}^{t-1} and extracted f_{s+}^{t-1} gas flow, a fundamental constraint for modeling the dynamics of gas storage over time.

Chapter 3

Optimization Using Mathematical Programs with Complementarity Constraints

3.1 Formulation of Interconnected Power and Gas Systems

An interconnected system can be effectively represented by a directed graph denoted as $\{\mathcal{N}, \mathcal{E}\}$, where the sets of units $\mathcal{N}$ and edges $\mathcal{E}$ consider all power and gas components along with their interconnections. On the electrical power side, the system holds power units $\mathcal{N}_P \subset \mathcal{N}$, termed buses, and power edges $\mathcal{B} \subset \mathcal{E}$ or branches. The power buses comprise generators $\mathcal{G} \subset \mathcal{N}_P$ injecting power and users $\mathcal{D} \subset \mathcal{N}_P$ demanding power [23]. The branches $\mathcal{B} = \{b = (n, m) \mid n, m \in \mathcal{N}_P\}$ connect the buses to make the electrical power flow from the generators to the users. Although the physical power flow is alternating current, the system is accurately modeled using a linear direct current (DC) approximation. The DC model ignores reactive power flows and voltage magnitude fluctuations and approximates active power flows using linear transfer distribution factors [24]. Further, the linear characteristics allow stating linear programming problems. Thus, the DC model serves as an appropriate approximation for many power system operations and planning studies, providing a balance of accuracy and computational tractability [25].

Then, the optimization problem of the interconnected system seeks to minimize the operation costs for satisfying the demands of the interconnected system while encom-

passing the power and gas constraints. Specifically, the following cost function linearly combines the flows of power and gas through the operation costs of the interconnected system elements:

$$
\min_{\mathcal{P},\mathcal{F}} \quad \sum_{g\in\mathcal{G}} C_g^t P_g^t + \sum_{d\in\mathcal{D}} C_d^t P_d^t + \sum_{w\in\mathcal{W}} C_w^t f_w^t +
$$
$$
\sum_{p\in\mathcal{P}} C_p^t f_p^t + \sum_{c\in\mathcal{C}} C_c^t f_c^t + \sum_{u\in\mathcal{U}} C_u^t f_u^t +
$$
$$
\sum_{s\in\mathcal{S}} C_{s+}^t f_{s+}^t + \sum_{s\in\mathcal{S}} C_{s-}^t f_{s-}^t + \sum_{s\in\mathcal{S}} C_s^t V_s^t
$$

(3.1)

where C_g^t denotes the generation cost by the g-th bus and C_d^t the unsupplied power demand for the d-th user. For the natural gas system, Equation (3.1) integrates the costs for injecting gas into the system using the well w, for transporting gas through the pipeline p, and for the pressure boosting of compressor c, at the time instant t, namely, C_w^t, C_p^t, and C_c^t, respectively. C_u^t denotes the penalty cost for not supplying the demanded gas to the user u. Lastly, C_{s+}^t, C_{s-}^t, and C_s^t represent the costs of injecting, extracting, and storing gas at the s-th storage station. Therefore, the decision variables for the optimization problem are P_g^t for the generated power, P_d^t for the unsupplied power, f_w^t for the inject gas flow, f_p^t and f_c^t for the transported gas through pipeline p and compressor c, f_u^t for the unsupplied gas demand, f_{s+}^t, f_{s-}^t, and f_s^t for injecting, extracting, and storing gas. Traditionally, a transported gas with a positive value of $f_p^t > 0$ moves in the predefined direction, while a negative value flows in the opposite one, with no impact on the optimization process. On the other hand, compressor stations solely allow unidirectional gas flow, expressed as $f_c^t \geq 0$. By optimizing this integrated cost function while adhering to the system's operational constraints, the proposed methodology effectively balances the demands of both energy systems, leading to a comprehensive solution that minimizes costs while ensuring reliable and efficient operation.

Optimization of the integrated cost function in Equation (3.1) while adhering to the system's operational constraints must lead to a comprehensive solution balancing the demands of both energy systems while ensuring reliable and efficient operation. Three sets of operational constraints describe the within and between power and gas interplay.

The first constraint set guarantees a stable power system operation: Equation (3.2) ensures that the generated power P_g^t lies between the technical minimum $\underline{P_g^t}$ and maximum $\overline{P_g^t}$. Equation (3.3) bounds the power flow through the transmission line P_l^t,

preventing damages, such as overheating. Equation (3.4) models the power flow over the electrical network through the reactance-based relationship of the power flow P_l^t, the line susceptance B_{nm}, and the voltage angles θ_n, θ_m at buses n, m. Equation (3.5) limits the unsupplied power P_d^t to the user demand $\overline{P_d^t}$. Equation (3.6) ensures stable operating conditions within the interconnected power grid by restricting the bus voltage angles. Equation (3.7) defines the power balance at each bus, i.e., the total input and generated power must equal the total output and unsupplied power, being $\mathcal{L}_{n+} = \{(m, n') \in \mathcal{L} : n' = n\}$ and $\mathcal{L}_{n-} = \{(n', m) \in \mathcal{L} : n' = n\}$ the set of inflow and outflow transmission lines at the n-th bus, respectively.

$$\underline{P_g^t} \leq P_g^t \leq \overline{P_g^t} \quad \forall\, g \in \mathcal{G}, \tag{3.2}$$

$$-\overline{P_l^t} \leq P_l^t \leq \overline{P_l^t} \quad \forall\, l \in \mathcal{L}, \tag{3.3}$$

$$P_l^t = B_{nm}(\theta_n - \theta_m) \quad \forall\, l = (n, m) \in \mathcal{L}, \tag{3.4}$$

$$0 \leq P_d^t \leq \overline{P_d^t} \quad \forall\, d \in \mathcal{D}, \tag{3.5}$$

$$-\overline{\theta_n^t} \leq \theta_n^t \leq \overline{\theta_n^t} \quad \forall\, n \in \mathcal{N}_P, \tag{3.6}$$

$$\sum_{\substack{l \in \mathcal{L}_{n+} \\ g=n}} P_l^t + P_g^t = \sum_{\substack{l \in \mathcal{L}_{n-} \\ d=n}} P_l^t + P_d^t \quad \forall\, n \in \mathcal{N}_P \tag{3.7}$$

The second constraint set interconnects natural gas and electrical power systems through gas-fired power plants generating electricity, as expressed by Equation (3.8) , where f_n^t stands for the natural gas fuel consumption to generate a power P_n^t at generator bus $n \in \mathcal{N}_I$, the heat-rate HR_n defines the generator efficiency, and the set $\mathcal{N}_I = \mathcal{G} \cap \mathcal{U}$ holds all the units in the interconnected system belonging to both the power generator and gas demand sets.

$$f_n^t = P_n^t \cdot \mathrm{HR}_n, \quad \forall\, n \in \mathcal{N}_I, \tag{3.8}$$

3.2 Mathematical Programming with Complementarity Constraints for Weymouth Approximation

The Weymouth equation is the fundamental model for gas flow through pipelines. However, it presents a challenge for optimal interconnected operation due to its nonlinearity,

which arises from the signum function determining the gas flow direction. This nonlinearity results from the complex physics of gas flow, making it challenging to find optimal solutions for gas transportation systems [26]. Traditional optimization approaches struggle to handle the non-convex terms within the Weymouth equation. However, recent advances in optimization techniques, particularly mathematical programs with complementary constraints (MPCC), offer a promising solution to address this issue. MPCC specializes in handling complementarity constraints and non-convexities, making it well-suited to tackle the intricacies of the Weymouth equation [27]. This type of formulation involves optimization problems of the general form:

$$\mathcal{O}: \ \min \ f(x, y) \tag{3.9a}$$
$$\text{s.t. } h_i(x, y) = 0 \tag{3.9b}$$
$$g_j(x, y) \geq 0 \tag{3.9c}$$
$$0 \leq G_k(x) \perp H_k(y) \geq 0 \tag{3.9d}$$

where $f(x, y)$ is the cost function, $h(x, y)$ and $g(x, y)$ capture equality and inequality constraints in the optimization problem $\mathcal{O}$. Equation (3.9d) represents the complementarity conditions, with the operator $\perp$ indicating that at a solution, either x or y must be zero while the other must remain non-negative. These conditions turn MPCC into a modeling tool for scenarios with variables exhibiting complementarity relationships, such as economic equilibrium [28], variational inequalities [29], and the intricate dynamics of natural gas transportation systems [30]. To deal with the non-convexity, this work rewrites the Weymouth equation as the following mathematical program with two complementarity constraints:

$$\mathcal{O}_W: \ \min_{y_p^t} -y_p^t f_p^t \tag{3.10}$$
$$\text{s.t. } y_p^t (f_p^t)^2 = K_{nm}((\pi_n^t)^2 - (\pi_m^t)^2)$$
$$-1 \leq y_p^t \leq 1$$
$$f_p^t = f_{p+}^t - f_{p-}^t$$
$$0 \leq f_{p+}^t \perp (y_p^t + 1) \geq 0$$
$$0 \leq f_{p-}^t \perp (1 - y_p^t) \geq 0$$

where $f_{p+}^t \geq 0$ and $f_{p-}^t \geq 0$ hold the positive and negative components of the gas flow in the p-th pipeline at operation period t, for assessing directional flow.

Solving MPCC presents a unique set of challenges distinguishing it from traditional optimization problems. One notable challenge is the need for regularity properties, making MPCC more complex [31]. Compared to smooth optimization problems, where gradients and Hessians provide valuable information for optimization algorithms, MPCC often lacks these properties, leading to difficulties in devising efficient numerical methods.

3.2.1 Linear Independence Constraint Qualification (LICQ)

The Linear Independence Constraint Qualification (LICQ) is a critical condition in optimization, particularly in nonlinear programming problems (NLPs). LICQ ensures the existence and uniqueness of Lagrange multipliers, simplifying their interpretation and enhancing the clarity of their role in constrained optimization. Additionally, LICQ provides a robust framework for local analysis, guaranteeing that the KKT conditions are sufficient for optimality when satisfied at a specific point. [32].

LICQ is a constraint qualification used in optimization problems to ensure that the gradients of the active inequality constraints and the gradients of the equality constraints are linearly independent at the minimizing point x^* of the original constrained optimization problem $\mathcal{P}$, understanding the set of active constraints as,

$$I(x^*) := \left\{ 1 \le l \le p \mid g_l(x^*) = 0 \right\}, \tag{3.11}$$

i.e., the inequality constraints at the point x^* that lie on its boundary. The above indicates that this constraint qualification is fulfilled when the elements of the set $\mathcal{F}$ are linearly independent at the point x^*.

$$\mathcal{F} = \left\{ (\nabla h_1(x^*)), \ldots, (\nabla h_m(x^*)), (\nabla g_n(x^*), \forall n \in I(x^*)) \right\} \tag{3.12}$$

3.2.2 Mangasarian-Fromovitz Constraint Qualification (MFCQ)

When an optimization problem does not meet the LICQrequirements, it is possible to resort to a second,less stringent criterion to check whether the KKT conditions are satisfied. This second criterion is known as Mangasarian-Fromovitz Constraint Qualification (MFCQ). The LICQ focuses on ensuring linear independence of the gradients of the active inequality and equality constraints [33]. On the other hand, the main objective of MFCQ is to guarantee that the gradients of the equality constraints are

linearly independent at the optimal point $\mathbf{x}^*$, and furthermore that there exists a vector $\mathbf{d} \in \mathbb{R}^n$ such that

$$\nabla h_i(\mathbf{x}^*)^\top \mathbf{d} < 0 \tag{3.13}$$

for all equality constraints.

$$\nabla g_j(\mathbf{x}^*)^\top \mathbf{d} < 0 \tag{3.14}$$

and for all active inequality constraints.

It is widely recognized that conventional constraint qualifications in nonlinear programming, such as LICQ and MFCQ, are typically not satisfied in the case of MPCC. As a result, KKT conditions commonly associated with MPCC may not be applicable or valid at a local minimization point [34]. Therefore, posing relaxed nonlinear programs (RNLP) deals with the numerical resolution of MPCC by introducing a positive regularization parameter $\epsilon \in \mathbb{R}^+$ that simplifies the solution and properly handles the inequalities [35]. These programs typically satisfy constraint qualifications, making them more amenable to efficient optimization techniques. Relaxing MPCC ensures that inequalities are appropriately treated as inactive, particularly when $G_k(x)H_k(y) \leq \epsilon$, enhancing their structural integrity. Besides, relaxed programs reliably approximate the original problem as $\epsilon \to 0$ [36]. Hence, instead of working with the original problem $\mathcal{O}_W$, the relaxed problem $\mathcal{O}_\epsilon$ is considered:

$$\mathcal{O}_\epsilon : \min_{y_p^t} -y_p^t f_p^t \tag{3.15a}$$

$$\text{s.t. } y_p^t (f_p^t)^2 = K_{nm}((\pi_n^t)^2 - (\pi_m^t)^2) \tag{3.15b}$$

$$f_p^t = f_{p+}^t - f_{p-}^t \tag{3.15c}$$

$$-1 \leq y_p^t \leq 1 \tag{3.15d}$$

$$f_{p+}^t(y_p^t + 1) \leq \epsilon \tag{3.15e}$$

$$f_{p-}^t(1 - y_p^t) \leq \epsilon \tag{3.15f}$$

Theoretically, the relaxed problem offers fundamental properties that tackle challenging MPCC problems [37]. Firstly, the relaxed approach guarantees the convergence to the true MPCC solution as $\epsilon \to 0$. Additionally, the boundedness of Lagrange multipliers ensures numerical stability and avoids issues with infinitely large values during optimization. Lastly, the local uniqueness of the $\mathcal{O}_\epsilon$ solution under specific conditions guarantees a single and well-defined solution. Therefore, the proposed relaxed

optimization problem deals with the non-convexity in the Weymouth equation while guaranteeing the KKT conditions around ϵ, posing a standard optimization problem, and avoiding ambiguity in interpreting results.

3.3 Case studies

The current section validates the proposed MPCC approach by comparing its performance against two well-established methods for approximating the Weymouth equation: i) The Taylor series approach that piecewise approximates Weymouth with line segments [38] and ii) The SOC programming that introduces a two-stage optimization, namely, flow direction estimation and cost minimization [39]. The validation contrasts Taylor, SOC, and MPCC approaches in three case studies of interconnected systems with different complexities.

The considered validation aims to quantify the inherent errors and the cost–error trade-off of the contrasted approaches to support its real-world pertinence. Therefore, this work reports two performance metrics: the cost function in Equation (3.1) that assesses the capacity for optimally operating an integrated system and the Weymouth error metric ($WE_p^t \in \mathbb{R}^+$) for quantifying the required flow to guarantee equality for pipeline p at time instant t in **??**, as follows:

$$WE_p^t = \left| f_p^t - \left(K_{nm} |(\pi_n^t)^2 - (\pi_m^t)^2| \right)^{1/2} \right|, \quad \forall\, p = (n, m) \in \mathcal{P}. \qquad (3.16)$$

Hence, the WE_p^t metric, measured in million standard cubic feet per day (MM-SCFD), explains the approximations' inherent sensitivity and validates the significance of their differences.

3.3.1 Case Study I: 9/8 System

The network depicted in Figure 3.1 interconnects a nine-bus power system and an eight-node natural gas network. The small size of case 9/8 enables fast execution, efficient analysis, and rigorous validation of the contrasted approaches. The 9/8 network also features a closed trajectory and bidirectional pipelines, allowing looped infrastructure with potential flow reversals.
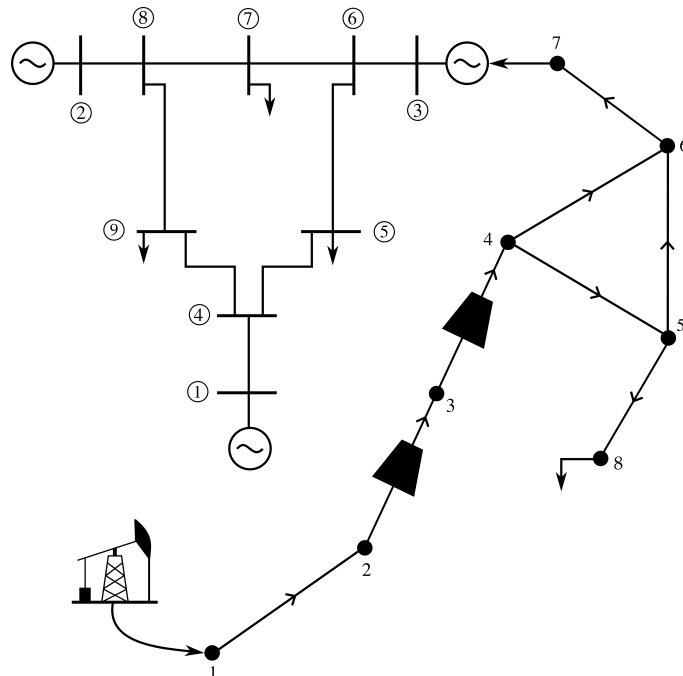
Figure 3.1: Integrated system 9/8 used in Case Study I, modified from the MPNG software. [1]

To assess the performance of Weymouth approximation approaches on the 9/8 system, a Monte Carlo experiment estimates the cost function and Weymouth error distributions by solving the optimization problem for one day ($\mathcal{T} = \{1\}$) one hundred times with uniformly sampled natural gas demands. Further network parameter details can be found in the publicly available repository OptiGasFlow (`https://github.com/cblancom/optigasflow`, accessed on 05 April 2024). Figure 3.2 depicts the cost function histogram for Taylor, SOC, and MPCC approaches. Remarkably, the three histograms evidence identical distribution patterns, leading to regular solutions across approaches.

The boxplots in Figure 3.3 show the Weymouth approximation error distribution for each pipeline using three approaches. The error distributions, including median and interquartile range, indicate that MPCC consistently maintains accuracy throughout the network. In contrast, the widely varying errors of the Taylor and SOC approaches suggest a lack of consistency in the achieved solution. Therefore, in a small network, the proposed MPCC approach converges to identical operational costs as Taylor and SOC, even in rationing, while meeting all linear constraints and improving the Weymouth
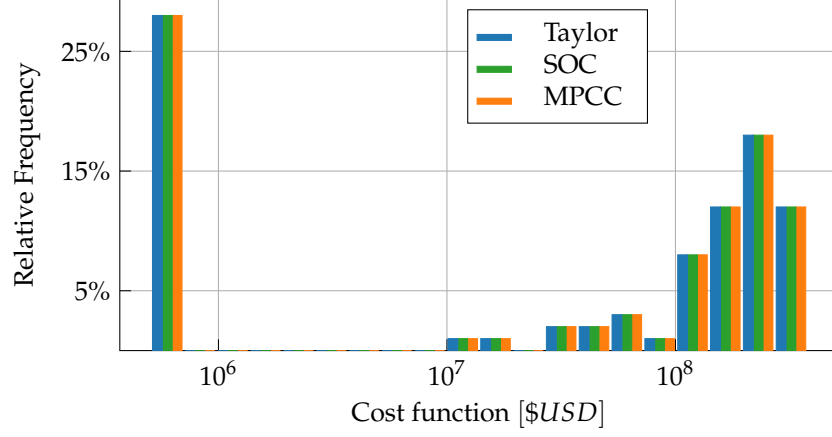
approximation.

Figure 3.2: Cost function histogram for the Taylor, SOC, and MPCC Weymouth approximation approaches in the 9-bus 8-node system.
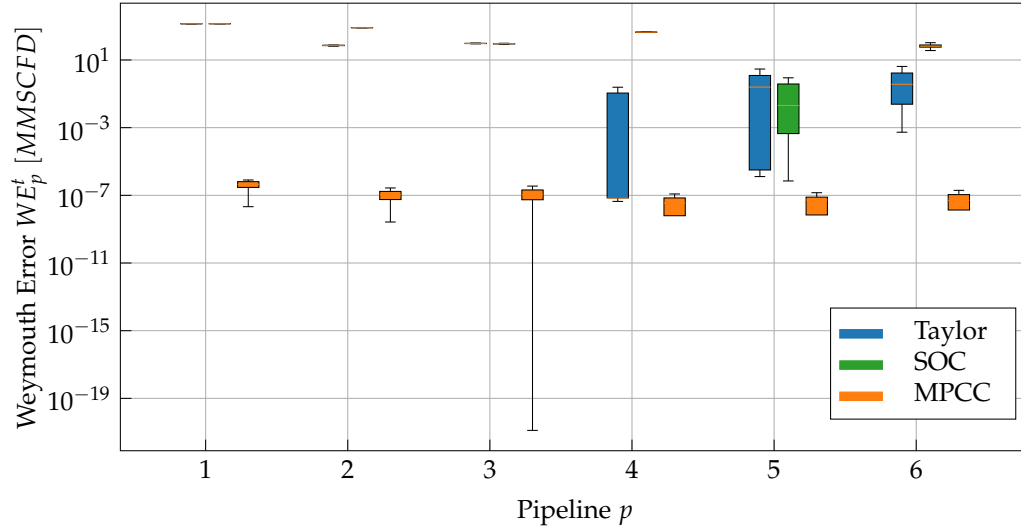

Figure 3.3: Boxplot of Weymouth error distribution for each pipeline in the 9/8 system attained by contrasted approximation approaches.

3.3.2 Case II: 118/48 System

The following case simulates a complex, large-scale electric grid system, the widely studied IEEE 118 bus system [40], consisting of 54 generator buses, 9 fed by the gas system, 186 transmission lines, and 99 users, that is, $|\mathcal{G}| = 54, |\mathcal{F}| = 186, |\mathcal{D}| = 99$. This electric grid interconnects with a 48-node natural gas system featuring 9 supply wells, 46 pipelines, eight compressor stations, and 22 user nodes through 9 connection points, i.e., $|\mathcal{W}| = 9, |\mathcal{P}| = 46, |\mathcal{C}| = 8, |\mathcal{U}| = 22, |\mathcal{I}| = 9$ [4]. The network topology deliberately introduces closed flow loops to stress the solver and the constraint approximations, as do real-world systems.

Figure 3.4 depicts the histogram of relative cost differences for the MPCC proposal to Taylor and SOC baselines from a hundred trials of the Monte Carlo experiment and a considered operation of one day ($T = 1$). It is worth noting that both baselines yielded the same cost function values. The relative difference between MPCC and the baselines is always positive, indicating that the complementarity constraint formulation consistently produces larger cost values in this system. However, the maximum difference of 6% falls within the range of real-world variations due to the dispatcher's practical decisions in line with the actual pressure–flow relationship [41].
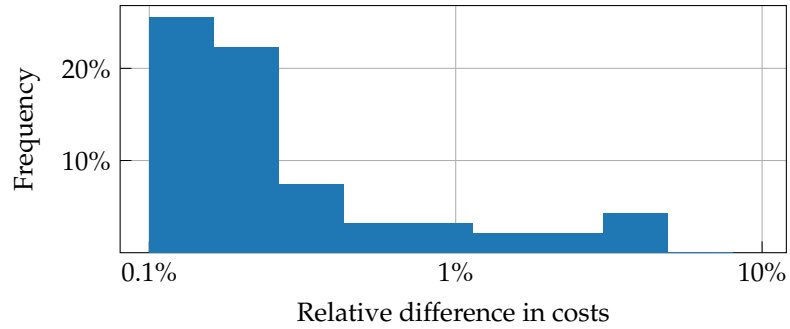

Figure 3.4: Histogram depicting the relative frequencies of cost differences obtained between MPCC and the other approaches in the 48-node 118-bus system.

Contrarily to cost function analysis, results in Figure 3.5 reveal a significant error reduction of about seven orders of magnitude (from 10^1 to 10^{-6}) under the proposed complementarity constraints. As an additional benefit, MPCC exhibits a shorter error dispersion than Taylor and SOC at most of the 46 pipelines in the network. Such behavior in the 118/48 system, also evidenced in the small 9/8 case study, proves the

reliability of MPCC in effectively addressing more complex network configurations and interconnected dynamics.

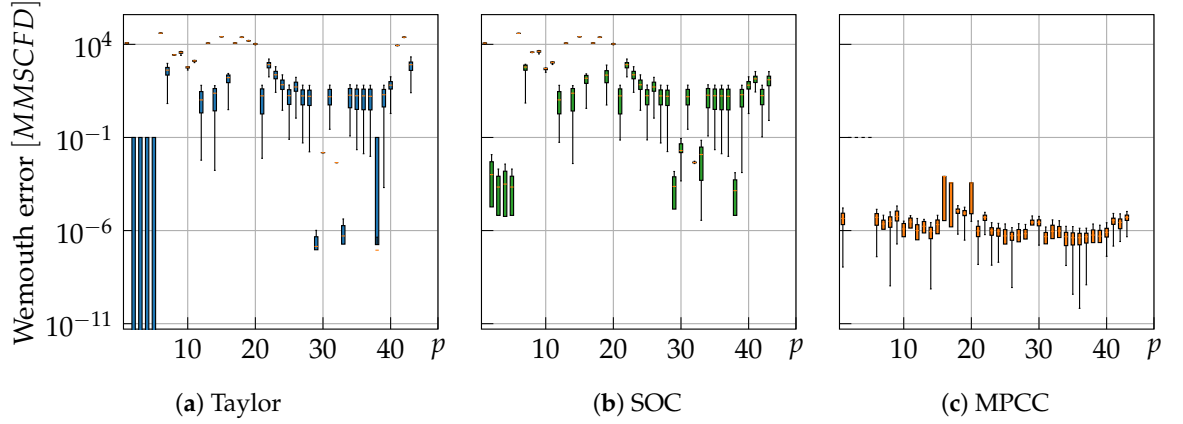**(a)** Taylor   **(b)** SOC   **(c)** MPCC

Figure 3.5: Weymouth approximation errors for each pipeline p reached by the contrasted approaches in the 118/48 study case.

3.3.3   Case Study III: 96/63 System

The last case study focuses on the Colombian power system, a complex network comprising 96 nodes ($|\mathcal{N}_P| = 96$), 49 generators ($|\mathcal{G}| = 49$), 207 transmission lines ($|\mathcal{F}| = 207$), and 80 power users ($|\mathcal{D}| = 80$). From the 49 generators, 10 are thermal power plants ($|\mathcal{I}| = 10$) fed by the natural gas transportation system, including 13 wells ($|\mathcal{W}| = 13$), 48 pipelines ($|\mathcal{P}| = 48$), 14 compressor stations ($|\mathcal{C}| = 14$), and 26 consuming users ($|\mathcal{U}| = 26$), yielding 63 nodes ($|\mathcal{N}_F| = 63$). Despite its radial structure, the gas system supports bidirectional flows in its pipelines due to the highly varying demand by thermal power plants influenced by meteorological conditions: On rainy seasons, thermal power plants dramatically reduce their demand; while on dry seasons, a large amount of gas must flow to them.

Instead of estimating the distributions of the cost function and Weymouth error as in cases 9/8 and 118/48, the 96/63 case validates the Weymouth approximations in an operation case of ten consecutive days ($|\mathcal{T}| = 10$) with randomly changing gas extraction costs. Such a complementary validation strategy allows the interconnected system to reduce gas transportation costs by exploiting its single storage station, extending the performance analysis to scheduling scenarios. Figure 3.6 illustrates the daily optimized operating cost of the integrated system over the ten-day scheduling horizon for each

tested approach. The daily cost values reveal notable similarities between the Taylor series and SOC relaxations. Nonetheless, the MPCC approach yields a 2.7% more expensive solution, from 8% cheaper to 12% more expensive, with a difference standard deviation of 6%. The above results indicate that the difference between the proposed MPCC and baseline approximations is statistically negligible and will disappear after the empirical corrections.
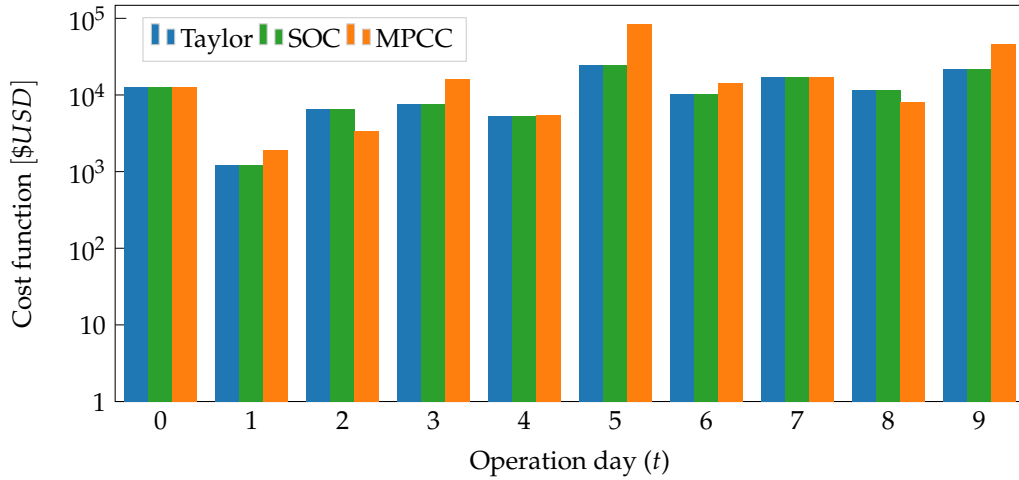

Figure 3.6: Daily operating cost obtained with each of the approaches in the 63-node 96-bus system.

Regarding the Weymouth approximation analysis, Figure 3.7 presents the error distribution and its relationship with the gas flow and the scheduled day for Taylor, SOC, and MPCC. Firstly, the error histogram in Figure 3.7a proves that the proposed MPCC formulation (in green) exhibits superior approximation accuracy to Taylor and SOC for most pipelines and days. Secondly, the scatter plot in Figure 3.7b illustrates the relationship between Weymouth error and gas pipeline flow for each approach. Note that the benchmark techniques of Taylor (blue) and SOC (orange) hold a stationary error regardless of the flow rate. In the case of MPCC (green), the larger the flow rate, the shorter the error dispersion. In addition, despite its large error dispersion at low flow rates, MPCC still delivers much lower errors than benchmark methodologies. Hence, the complementarity constraints improve the error rates of Taylor and SOC and become more reliable for higher gas flow rates.

(a) Relative Frequency   **(b)** Pipeline flow (f_m^t)   **(c)** Operation day (t)
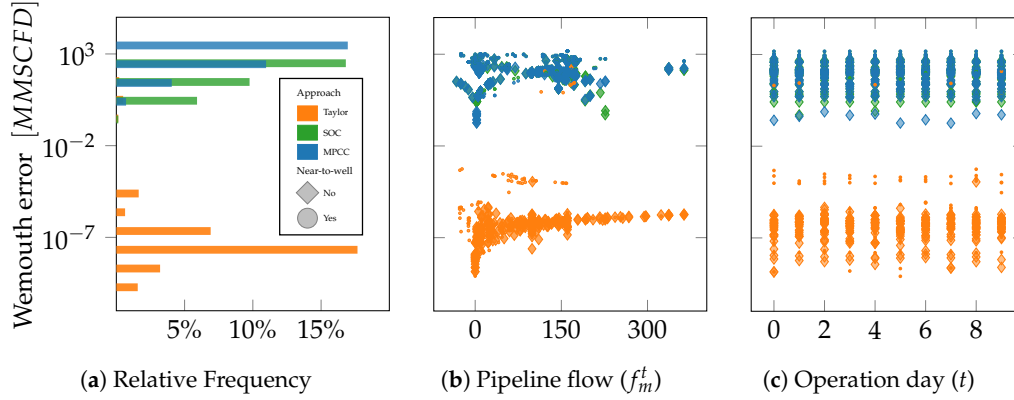
Figure 3.7: Weymouth error density on the Colombian case versus the gas flow and operation day.

Lastly, Figure 3.7c suggests independence between the Weymouth error and each scheduled day, with a stationary error distribution for all approximations. Nonetheless, MPCC holds two groups of outlying errors. The higher ones align with typical magnitudes of the benchmark techniques. The second group of errors, lying around 10^{-2}, corresponds to pipelines connected to injection wells (denoted as dots in Figure 3.8c). Since the wells are technically regulated, their fixed injection pressure hampers the flexibility of MPCC for approximating the Weymouth equation.

The heatmaps in Figure 3.8c illustrate the output-to-input pressure ratio for each of the
14 compressors over the ten days of the scheduled operation. The baseline approaches of Taylor and SOC (Figure 3.8a,b) yield constant pressure ratios stemming from an over-relaxation of the Weymouth equation that extends the feasible region to unpractical solutions. In contrast, the MPCC approach in Figure 3.8c exhibits day-to-day pressure ratio changes within each compressor. The above is because the complementarity constraints closely align with the gas transport system's real physics, restraining the range of the feasible pressure values to trade off the daily varying injection cost.
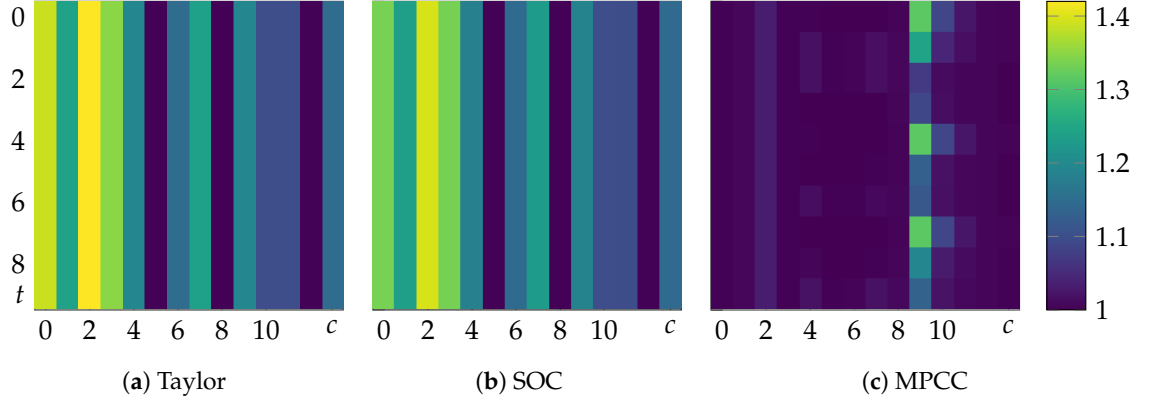
Figure 3.8: Output-to-input pressure ratio at the compressor in the 96/63 system.

As a remark, compressor nine in Figure 3.8c reaches large pressure ratios on Days 0, 1, 2, 4, and 7, overlapping with the time instants with the highest approximation errors for MPCC in Figure 3.7c. A detailed examination of these outcomes detects that compressor nine and the outlying pipeline are the two outputs of a bifurcation, the latter being followed by an injection well. Figure 3.9 exemplifies that such an interconnection is the sole over the gas network. As a hypothesis, fixing the pressure at the injection well and the flow direction at compressor nine pushes the complementarity constraints to the limits and forces the compressor to augment the pressure ratio to satisfy the forthcoming branch demand.
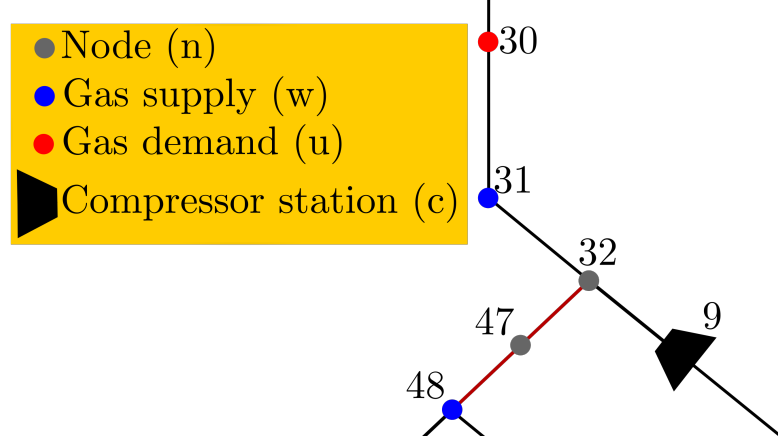
Figure 3.9: Outlying connection of well-compressor-pipeline on the system 96/63 used in Case Study III.

3.4 Conclusions

This paper presented a novel approximation for the Weymouth constraint by representing the nonconvex pressure–flow relationship as an MPCC. The MPCC-based formulation significantly benefits the optimization problems in interconnected power and gas systems using binary-behaving continuous variables related to the flow direction, which avoids costly mixed-integer approximations. Additionally, the MPCC inherently captures the complexity in the signum function, resulting in a rigorous approximation of the Weymouth equation.

The validation compared the proposed MPCC approach against the Taylor series and SOC programming approximations on optimizing the operation of interconnected power and gas transport systems. Monte Carlo experiments validated the solution reliability in two well-known case studies, while a ten-day operation planning assessed the scheduling task in a real-world case study.

Regarding cost function, the MPCC approach demonstrated a remarkable ability to balance operational costs effectively. Results on the 9/8 system proved that MPCC converges to the exact cost of Taylor and SOC in small-scale cases. For more complex networks (cases 118/48 and 96/63), MPCC yields higher operational costs than baselines due to the more rigorous Weymouth equation modeling. Nonetheless, the cost differences among approaches lie within reasonable limits and align with the dispatcher's empirical decisions.

In the case of Weymouth approximation, MPCC significantly outperforms Taylor and SOC in the tested cases. In the 118/48 and 96/63 systems, MPCC substantially reduces Weymouth approximation errors, often by several orders of magnitude, compared to traditional linearization and convex relaxation strategies. Such an accuracy improvement becomes crucial in large-scale, complex systems where precise approximation directly influences operational efficiency and system reliability. Hence, the introduced pressure–flow model mathematically benefits the optimization task, asserting its cost-effectiveness at various system scales.

The analysis of the scheduling task in the 96/63 Colombian interconnected system underscores the robustness and reliability of the MPCC approach. Despite the complexities of bidirectional flows and time-varying demand scenarios, MPCC maintains high accuracy levels in Weymouth approximation. Furthermore, the nearly negligible cost differences among approximation approaches establish MPCC as the most robust and reliable approach for short-term operational scheduling.

In conclusion, modeling the Weymouth equation as an MPCC improves the optimization of interconnected gas and power systems by balancing operational costs, minimizing approximation errors, and handling scheduling tasks. These findings establish strong evidence for the practical implementation of MPCC in gas transport optimization, particularly in scenarios demanding high accuracy and reliability in short-term operation scheduling.

Considering the current open issues on energy management, three future research directions may complement this study. Firstly, we propose to adapt MPCC to dynamic system constraints for validation in transient analysis scenarios. The second research direction accounts for the uncertainty in interconnected systems, mainly due to the growing share of low-inertia power sources, such as wind and solar, and potential gas transport failures. Hence, we plan to extend the proposed methodology to stochastic optimization, considering the varying parameters and power sources of interconnected systems. Lastly, we will integrate MPCC with distributed cooperative operation schemes considering multi-agent issues such as the lack of information due to privacy policies [42].

Chapter 4

Gas System - Censnet

4.1 Formulation of Gas System

Chapter 5

Conclusions

Bibliography

[1] S. García-Marín, W. González-Vanegas, and C. Murillo-Sánchez, "Mpng: A matpower-based tool for optimal power and natural gas flow analyses," *IEEE Transactions on Power Systems*, pp. 1–9, 2022.

[2] M. Abbasi, H. Afshari, S. Asadi, M. Bakhshayeshi Baygi, F. H. Benam, F. Daneshzand, W. Dullaert, M. H. Ebrahimi, G. Esmaeilian, B. Fahimnia, S. Fallah, R. Z. Farahani, M. Hamedi, S. Hosseini, M. Jamshidi, L. Kardar, Z. Khooban, R. Molaei, S. Mostafa Mousavi, E. Nikbakhsh, M. Parvini, M. Rajabi, F. Ranaiefar, A. Regan, S. Rezapour, Z. Rouhollahi, H. Sadjady, S.-A. Seyed-Alagheband, M. SteadieSeifi, and A. Zakery, "List of contributors," in *Logistics Operations and Management* (R. Z. Farahani, S. Rezapour, and L. Kardar, eds.), pp. xv–xvi, London: Elsevier, 2011.

[3] E. Raheli, Q. Wu, M. Zhang, and C. Wen, "Optimal coordinated operation of integrated natural gas and electric power systems: A review of modeling and solution methods," *Renewable and Sustainable Energy Reviews*, vol. 145, p. 111134, 2021.

[4] S. Chen, A. J. Conejo, R. Sioshansi, and Z. Wei, "Unit commitment with an enhanced natural gas-flow model," *IEEE Transactions on Power Systems*, vol. 34, no. 5, pp. 3729–3738, 2019.

[5] G. Byeon and P. Van Hentenryck, "Unit commitment with gas network awareness," *IEEE Transactions on Power Systems*, vol. 35, no. 2, p. 1327–1339, 2020.

[6] J. D. Morcillo, F. Angulo, and C. J. Franco, "Analyzing the hydroelectricity variability on power markets from a system dynamics and dynamic systems perspective: Seasonality and enso phenomenon," May 2020.

[7] C. Villa-Loaiza, I. Taype-Huaman, J. Benavides-Franco, G. Buenaventura-Vera, and J. Carabalí-Mosquera, "Does climate impact the relationship between the

energy price and the stock market? the colombian case," *Applied Energy*, vol. 336, p. 120800, Apr 2023.

[8] L. A. Cuartas, A. P. Cunha, J. A. Alves, L. M. Parra, K. Deusdará-Leal, L. C. Costa, R. D. Molina, D. Amore, E. Broedel, M. E. Seluchi, and et al., "Recent hydrological droughts in brazil and their impact on hydropower generation," *Water*, vol. 14, p. 601, Feb 2022.

[9] L. Yang, X. Zhao, and Y. Xu, "A convex optimization and iterative solution based method for optimal power-gas flow considering power and gas losses," *International Journal of Electrical Power & Energy Systems*, vol. 121, p. 106023, 2020.

[10] T. Jiang, C. Yuan, R. Zhang, L. Bai, X. Li, H. Chen, and G. Li, "Exploiting flexibility of combined-cycle gas turbines in power system unit commitment with natural gas transmission constraints and reserve scheduling," *International Journal of Electrical Power & Energy Systems*, vol. 125, p. 106460, 2021.

[11] M. S. Alam, S. R. Paramati, M. Shahbaz, and M. Bhattacharya, "Natural gas, trade and sustainable growth: empirical evidence from the top gas consumers of the developing world," *Applied Economics*, vol. 49, pp. 635–649, 2 2017.

[12] J.-H. Yuan, S. Zhou, T.-D. Peng, G.-H. Wang, and X.-M. Ou, "Petroleum substitution, greenhouse gas emissions reduction and environmental benefits from the development of natural gas vehicles in china," *Petroleum Science*, vol. 15, no. 3, pp. 644–656, 2018.

[13] X. Yin, K. Wen, Y. Wu, X. Han, Y. Mukhtar, and J. Gong, "A machine learning-based surrogate model for the rapid control of piping flow: Application to a natural gas flowmeter calibration system," *Journal of Natural Gas Science and Engineering*, vol. 98, p. 104384, 2022.

[14] M. Aydin, "Natural gas consumption and economic growth nexus for top 10 natural gas–consuming countries: A granger causality analysis in the frequency domain," *Energy*, vol. 165, p. 179–186, 2018.

[15] "U.s. energy information administration - eia - independent statistics and analysis."

[16] "Política nacional de cambio climático - ideam."

[17] "Estructura del sector de gas natural en colombia," Dec 2021.

[18] J. Restrepo-Trujillo, R. Moreno-Chuquen, and F. N. Jiménez-García, "Strategies of expansion for electric power systems based on hydroelectric plants in the context of climate change: Case of analysis of colombia," *International Journal of Energy Economics and Policy*, vol. 10, no. 6, p. 66–74, 2020.

[19] "Oferta y generación."

[20] "Fenomeno niño y niña."

[21] M. P. Ignacio Fariza, "La sequía desploma la generación hidroeléctrica hasta su nivel más bajo en tres décadas," Aug 2022.

[22] X. Shan, W. Yu, J. Gong, W. Huang, K. Wen, H. Wang, S. Ren, D. Wang, Y. Shi, and C. Liu, "A methodology to evaluate gas supply reliability of natural gas pipeline network considering the effects of natural gas resources," *Reliability Engineering &amp; System Safety*, vol. 238, p. 109431, 2023.

[23] Y. Wang, Y. Wang, Y. Huang, J. Yang, Y. Ma, H. Yu, M. Zeng, F. Zhang, and Y. Zhang, "Operation optimization of regional integrated energy system based on the modeling of electricity-thermal-natural gas network," *Applied Energy*, vol. 251, p. 113410, 2019.

[24] Z. Yang, K. Xie, J. Yu, H. Zhong, N. Zhang, and Q. Xia, "A general formulation of linear power flow models: Basic theory and error analysis," *IEEE Transactions on Power Systems*, vol. 34, no. 2, pp. 1315–1324, 2019.

[25] L. Yang, Y. Xu, H. Sun, and X. Zhao, "Two-stage convexification-based optimal electricity-gas flow," *IEEE Transactions on Smart Grid*, vol. 11, no. 2, pp. 1465–1475, 2020.

[26] Y. Zhang, S. Xie, and S. Shu, "Multi-stage robust optimization of a multi-energy coupled system considering multiple uncertainties," *Energy*, vol. 238, p. 122041, Jan 2022.

[27] B. Baumrucker, J. Renfro, and L. Biegler, "Mpec problem formulations and solution strategies with chemical engineering applications," *Computers & Chemical Engineering*, vol. 32, no. 12, p. 2903–2913, 2008.

[28] C. Ruiz, A. J. Conejo, J. D. Fuller, S. A. Gabriel, and B. F. Hobbs, "A tutorial review of complementarity models for decision-making in energy markets," *EURO Journal on Decision Processes*, vol. 2, p. 91–120, Oct 2014.

[29] M. Hintermüller and I. Kopacka, "Mathematical programs with complementarity constraints in function space: C- and strong stationarity and a path-following algorithm," *SIAM Journal on Optimization*, vol. 20, p. 868–902, Jan 2009.

[30] F. Hante and M. Schmidt, "Complementarity-based nonlinear programming techniques for optimal mixing in gas networks," *EURO Journal on Computational Optimization*, vol. 7, p. 299–323, Sep 2019.

[31] Y. Zhang, J. Wu, and L. Zhang, "First order necessary optimality conditions for mathematical programs with second-order cone complementarity constraints," *Journal of Global Optimization*, vol. 63, no. 2, p. 253–279, 2015.

[32] R. Bergmann and R. Herzog, "Intrinsic formulation of kkt conditions and constraint qualifications on smooth manifolds," *SIAM Journal on Optimization*, vol. 29, no. 4, p. 2423–2444, 2019.

[33] D. P. Bertsekas, *Nonlinear programming / Dimitri P. Bertsekas*. Belmont, Mass.: Athena Scientific, 2nd ed ed., 1999.

[34] G. Bouza and G. Still, "Mathematical programs with complementarity constraints: Convergence properties of a smoothing method," *Mathematics of Operations Research*, vol. 32, p. 467–483, May 2007.

[35] S. Scholtes, "Convergence properties of a regularization scheme for mathematical programs with complementarity constraints," *SIAM Journal on Optimization*, vol. 11, p. 918–936, Mar 2001.

[36] H. Scheel and S. Scholtes, "Mathematical programs with complementarity constraints: Stationarity, optimality, and sensitivity," *Mathematics of Operations Research*, vol. 25, p. 1–22, Feb 2000.

[37] D. Ralph and S. J. Wright, "Some properties of regularization and penalization schemes for mpecs," *Optimization Methods and Software*, vol. 19, no. 5, pp. 527–556, 2004.

[38] C. Ordoudis, P. Pinson, and J. M. Morales, "An integrated market for electricity and natural gas systems with stochastic power producers," *European Journal of Operational Research*, vol. 272, no. 2, pp. 642–654, 2019.

[39] A. Schwele, C. Ordoudis, J. Kazempour, and P. Pinson, "Coordination of power and natural gas systems: Convexification approaches for linepack modeling," in *2019 IEEE Milan PowerTech*, pp. 1–6, 2019.

[40] C. Wang, W. Wei, J. Wang, and T. Bi, "Convex optimization based adjustable robust dispatch for integrated electric-gas systems considering gas delivery priority," *Applied Energy*, vol. 239, pp. 70–82, 2019.

[41] Y. Zhao, M. Gong, J. Sun, C. Han, L. Jing, B. Li, and Z. Zhao, "A new hybrid optimization prediction strategy based on sh-informer for district heating system," *Energy*, vol. 282, p. 129010, 2023.

[42] B. Ding, Z. Li, Z. Li, Y. Xue, X. Chang, J. Su, X. Jin, and H. Sun, "A ccp-based distributed cooperative operation strategy for multi-agent energy systems integrated with wind, solar, and buildings," *Applied Energy*, vol. 365, p. 123275, 2024.