Resilience and coupling of interdependent critical infrastructures : models, optimization, and operations

Andrea Bellè

université PARIS-SACLAY

Resilience and coupling of interdependent critical infrastructures: models, optimization, and operations

Résilience et couplage des infrastructures critiques interdépendantes: modèles, optimisation et gestion opérationnelle

Thèse de doctorat de l'Université Paris-Saclay

École doctorale n° 573 Interfaces: matériaux, systèmes, usages
Spécialité de doctorat: Ingéniérie des systèmes complexes
Graduate School: Sciences de l'Ingénierie et des Systèmes
Référent : CentraleSupélec

Thèse préparée dans le Laboratoire Génie Industriel (Université Paris-Saclay, CentraleSupélec), sous la direction de Anne BARROS, professeure, et le co-encadrement de Zhiguo ZENG, maître de conférences

Thèse soutenue à Paris-Saclay, le 28 septembre 2022, par

Andrea BELLÈ

Composition du jury

Christophe Bérenguer Professeur, Grenoble-INP	Président & Examinateur
David W. Coit Professeur, Rutgers University	Rapporteur & Examinateur
John Andrews Professeur, University of Nottingham	Rapporteur & Examinateur
Olga Fink Professeure assistante, École Polytechnique Fédérale de Lausanne	Examinatrice
Oualid Jouini Professeur, CentraleSupélec, Université Paris-Saclay	Examinateur
Anne Barros Professeure, CentraleSupélec, Université Paris-Saclay	Directrice de thèse
Zhiguo Zeng Maître de conférences, CentraleSupélec, Université Paris-Saclay	Co-encadrant de thèse
Laurent Billet Délégue scientifique, EDF R&D	Invité
Bertrand Decocq Expert de la communauté des Réseaux du Futur, Orange	Invité
Cyril Cappi Responsable de groupe de recherche, SNCF	Invité
Bernard Yannou Professeur, CentraleSupélec, Université Paris-Saclay	Invité

THESE DE DOCTORAT
NNT : 2022UPAST105

université PARIS-SACLAY | ÉCOLE DOCTORALE Interfaces: matériaux, systèmes, usages

Titre: Résilience et couplage des infrastructures critiques interdépendantes: modèles, optimisation et gestion opérationnelle

Mots clés: Infrastructures critiques interdépendantes, résilience, interface de couplage, analyse de vulnérabilité, optimisation

Résumé: Les infrastructures critiques (IC) sont essentielles au maintien de la stabilité socio-économique et la prospérité d'une population. Assurer leur résilience face aux défaillances et aux perturbations est un enjeu majeur. De par leurs diverses relations d'interdépendance, leurs interfaces de couplage complexes, ainsi qu'une exploitation souvent réalisée par des entités indépendantes, l'analyse et l'optimisation de leur résilience est une tâche difficile. Cette thèse expose trois sujets liés à la résilience des IC interdépendantes, avec un accent particulier sur leur couplage. Premièrement, elle approfondit la vulnérabilité des réseaux ferroviaires et électriques interdépendants, en prenant en compte les interfaces de couplage réalistes et de défaillances en cascade, afin de mieux comprendre leurs risques mutuels. Deuxièmement, elle propose une approche mathématique pour améliorer la résilience des IC interdépendantes à travers l'optimisation de la topologie de leur interface de couplage. Enfin, elle suggère une approche préliminaire pour la prise de décision dans les IC interdépendantes qui tient compte des divers comportements des opérateurs indépendants dans des conditions normales et des situations de perturbation.

Title: Resilience and coupling of interdependent critical infrastructures: models, optimization, and operations

Keywords: Interdependent critical infrastructures, resilience, coupling interface, vulnerability analysis, optimization

Abstract: Critical infrastructures (CIs) are essential for maintaining the socio-economic stability and wealth of a population, and ensuring their resilience against failures and disruption is of the utmost importance. As CIs are connected with each other through various relationships of interdependency and complex coupling interfaces, and they are often operated by independent entities, their analysis and optimization is a challenging task. In this dissertation, three topics related to the resilience of interdependent CIs, with a particular focus on their coupling, are investigated. Firstly, the vulnerability of interdependent railway and power networks, accounting for realistic coupling interfaces and cascading failures, is investigated, in order to better understand the mutual risks of these interdependent CIs. Secondly, a resilience-based mathematical programming framework for the optimization of the topology of coupling interfaces between interdependent CIs is presented. Lastly, a preliminary approach for decision-making in interdependent CIs, accounting for the different behaviours of independent operators under normal conditions and in situations of disruption, is proposed.

Acknowledgments

Without them, this Ph.D. thesis would not be the same.

Loris, Susanna, Stefano, Fanny
Anne, Zhiguo, Yi-Ping, Adam

Contents

Acknowledgments i

List of Tables v

List of Figures viii

1 Introduction 1
- 1.1 Research group . . . 1
- 1.2 General context . . . 1
- 1.3 Interdependent critical infrastructures . . . 3
  - 1.3.1 Interdependencies . . . 3
  - 1.3.2 Modeling of interdependent critical infrastructures . . . 4
  - 1.3.3 Network-based models . . . 5
- 1.4 Resilience: concept and metrics . . . 8
- 1.5 Timeline and contributions of the thesis . . . 11

2 Vulnerability of railway and power networks 13
- 2.1 Introduction . . . 13
- 2.2 Modeling framework . . . 15
  - 2.2.1 Operational model . . . 15
  - 2.2.2 Cascading failure model . . . 17
- 2.3 Vulnerability analysis . . . 19
- 2.4 Case study . . . 20
- 2.5 Results . . . 22
- 2.6 On the coupling interface topology . . . 23
- 2.7 Conclusion . . . 25

3 Design and optimization of coupling interface topology 27
- 3.1 Introduction . . . 27
- 3.2 Operational modeling framework . . . 28
- 3.3 Case study . . . 29
- 3.4 Coupling interface topology - toy model . . . 31
- 3.5 Robust coupling interface . . . 33
  - 3.5.1 Defender-attacker-defender formulation . . . 33
  - 3.5.2 Solution strategy . . . 36
  - 3.5.3 Results . . . 36
  - 3.5.4 Remarks . . . 40
- 3.6 Distributionally robust coupling interface . . . 41
  - 3.6.1 Distributionally robust approach . . . 41

3.6.2 Distributionally robust optimization - toy model . . . 43
3.6.3 Distributionally robust formulation . . . 45
3.6.4 Solution strategy . . . 47
3.6.5 Results . . . 48
3.6.6 Remarks . . . 49
3.7 Allocation of redundant interdependency links . . . 50
3.7.1 Redundancy allocation problem . . . 50
3.7.2 Defender-attacker-defender formulation . . . 51
3.7.3 Results . . . 52
3.8 Computational performance . . . 54
3.9 Conclusion . . . 55

4 Joint decision-making with independent operators 57
4.1 Introduction . . . 57
4.2 Case study . . . 59
4.3 Operational models . . . 60
4.3.1 Centralized operational model . . . 60
4.3.2 Decentralized operational model . . . 61
4.4 Preliminary results . . . 64
4.5 Proposed modeling framework . . . 64
4.6 Proposed solution strategy . . . 68
4.7 Conclusion . . . 71

5 Conclusion 73
5.1 Vulnerability of railway and power networks . . . 73
5.1.1 Contributions . . . 73
5.1.2 Prospective work . . . 74
5.2 Design and optimization of coupling interface topology . . . 74
5.2.1 Contributions . . . 74
5.2.2 Prospective work . . . 75
5.3 Joint decision-making with independent operators . . . 75
5.3.1 Contributions . . . 75
5.3.2 Prospective work . . . 76

A Nested Column&Constraint Generation algorithm 77
A.1 Overview . . . 77
A.2 Inner layer . . . 77
A.3 Outer layer . . . 79

B Karush-Kuhn-Tucker conditions of the decentralized heat network model 81

C Linearization of nonlinear terms with strong duality equivalence 83

Appended papers 101

List of Tables

3.1 Cost of network metrics-based coupling interfaces. 40
3.2 Scenarios within the set $\mathcal{A}$ for the power network in Figure 3.7 and corresponding performance. 44
3.3 Upper bounds π_k^{max} for each power line k. 47
3.4 Results of the optimal allocation of redundant interdependency links in terms of worst-case combined performance and relative increase in percentage from the original case. Table from Paper V [80]. 53
3.5 Computational time in seconds of the NC&CG algorithm. Table from Paper III [10]. 54

List of Figures

1.1 Example of interdependent networks. The two networks, in red and blue, consist of nodes connected by edges. Nodes belonging to different networks are connected by dashed black edges, also referred to as interdependency links, that represent relationships of interdependency. Figure from Paper III [10]. 6
1.2 Example of CIs represented as networks. In red, the Great Britain reduced power transmission network [41] is shown. In blue, the proposition for a new high-speed British railway network made in [42] is shown. Figure based on Paper I [14]. 7
1.3 Traditional resilience curve with resilience phases. 10

2.1 Proposed cascading failure algorithm for IRPNs. Figure from Paper I [14]. 18
2.2 Geographical representation of the British IRPNs, based on Figure 3 of Paper I [14]. 21
2.3 Network-of-networks representation of the British IRPNs, based on Figure 4 of Paper I [14]. 21
2.4 Average loss of accessibility in the railway network due to different fractions of edges removed in the power network. Figure from Paper I [14]. 22
2.5 Average FDNS in the power network due to different fractions of edges removed in the power network. Figure from Paper I [14]. 23
2.6 Network-of-networks configuration of the French IRPNs, based on Figure 1 of Paper II [78]. 24
2.7 Fraction of demand not supplied $FDNS$ in the power network and traction network. Figure based on the results of Paper II [78]. 25

3.1 Geographical representation of the IPGNs. Figure from Paper IV [79]. 31
3.2 Illustrative IPGNs. The dashed lines represent the interdependency links, i.e. the coupling interface of the IPGNs. 32
3.3 Illustrative IPGNs with modified coupling interface. 32
3.4 General flowchart of the Nested Column&Constraint Generation algorithm. Picture based on Paper III, Paper IV, and Paper V [10], [79], [80]. 37
3.5 Combined performance in the worst-case feasible failure scenario with network metrics-based coupling strategies. Figure from Paper III [10]. It should be highlighted that, since multiple Euclidean coupling interfaces exist, for each value K_{att}, the plot represents the best performing Euclidean coupling interface. 39
3.6 Combined performance in the worst-case feasible failure scenario with optimal coupling strategies. Figure from Paper III [10]. 39
3.7 Toy model of a power network. 44
3.8 Results of the distributionally robust optimal coupling of IPGNs. Revised version of Figure 3 of Paper IV [79]. 48
3.9 Representative IPGNs with redundant interdependency links. 50
3.10 Results of the optimal allocation of redundant interdependency links in terms of worst-case combined performance, with different values of K_{att} and B_{ci}. Figure from Paper V [80]. . 52

3.11 Average results of the random strategy allocation of redundant interdependency links in terms of worst-case combined performance, with different values of K_{att} and B_{ci}. Figure from Paper V [80]. 53

4.1 Power network of the IPHNs. Figure based on [138]. 59
4.2 Heat network of the IPHNs. Figure based on [138]. 59
4.3 Average combined shedding in the IPHNs with different operator models and number of failed power lines. Figure adapted from Paper VI [132]. 65
4.4 Average combined cost in the IPHNs with different operator models and number of failed power lines. Figure adapted from Paper VI [132]. 65

1 - Introduction

1.1 . Research group

The work contained in this thesis was performed within the *Safety & Risks Research Group* of the Industrial Engineering Research Department of CentraleSupélec, Université Paris-Saclay, and founded by the *Chair on Risk and Resilience of Complex Systems* and its industrial partners: EDF, the major French energy utility company, Orange, the major French telecommunications company, and SNCF, the French state-owned railway company.

The activities of the Chair focus on safety and risk analysis of complex engineered systems, and they are organized into three main axes [1]:

1. *Complex systems and infrastructures*: analysis and optimization of complex engineered systems and infrastructures, which require to be treated from different perspectives.
2. *Industry 4.0 and predictive maintenance*: advanced models and optimization methods for dynamic risk management and predictive maintenance.
3. *Resilience*: assessment and optimization of the resilience of complex systems and critical infrastructures, including optimization of design, barriers, and resource allocation.

The work of this thesis gravitates around the concepts of resilience analysis and optimization of complex systems and critical infrastructures, and it belongs to Axes 1 and 3.

1.2 . General context

The socioeconomic wealth of countries and nations strongly depends on their capacity to supply essential goods, services and commodities, which is ensured by their network of critical infrastructures (CIs), such as energy systems, transportation networks, or telecommunications networks [2], [3]. In the context of the European Union, a critical infrastructure is defined as *"an asset or system which is essential for the maintenance of vital societal functions"* [4]. Moreover, *"the damage to a critical infrastructure, its destruction or disruption by natural disasters, terrorism, criminal activity or malicious behaviour, may have a significant negative impact for the security of the EU and the well-being of its citizens"* [4].

From this definition, it is clear how CIs play a key role in maintaining high standards of life and prosperity in a society, and how failures and disruption within

these systems can lead to vast and widespread negative consequences. In fact, disturbances and malfunctioning in one or several CIs can interrupt the supply of essential goods, services, and commodities, and considerably affect the daily life of citizens. A few illustrative examples are reported here:

- on September 28, 2003, failures of power transmission lines near the Swiss-Italian border caused a general blackout of the Italian peninsula which affected tens of millions of people for several hours [3], [5], [6];
- on November 4, 2006, a disconnection of high-voltage transmission lines on the Ems River, Germany, caused electrical disturbances in more than 10 million European households [7];
- on January 12, 2010, a magnitude 7.0 earthquake occurred in Haiti, causing vast disruption of the national CIs systems [8], and hampering crisis management operations.

CIs are extremely complex systems, as they are composed of a large variety of technologies and managed by various entities and stakeholders. Moreover, CIs are often interdependent on each other in terms of functionality and performance [9], and failures and disturbances within one infrastructure can propagate to other systems and cause multi-sectoral disruption [10], [11]. For example, in August 2019, British railway networks were heavily disrupted by disturbances within power networks [12], and the aforementioned Italian blackout was partially caused by a failure propagation process between interdependent power and telecommunications networks [13].

Within this context, it is clear that the analysis, optimization, and design of CIs are issues of the utmost importance. Despite interdependent CIs being an active and prolific research field, many questions remain unanswered, as their analysis and optimization are particularly challenging tasks. In fact, CIs are often dependent on each other in terms of functionality and performance, and they are often operated by independent entities and stakeholders. These factors highly increase the complexity of analyzing and designing interdependent CIs.

In this dissertation, three topics within the framework of interdependent CIs and their resilience, with a particular focus on their coupling, are addressed and investigated: vulnerability of interdependent railway and power networks, optimization of coupling interface topology, and joint decision-making with independent operators. The rest of this chapter is dedicated to the introduction of the main concepts of this dissertation. In Section 1.3, the topic of interdependent CIs, from a research perspective, is outlined; in Section 1.4, the concept of resilience and its connection to this work is addressed; in Section 1.5, the main contributions of this dissertation are highlighted.

1.3 . Interdependent critical infrastructures

1.3.1 . Interdependencies

An infrastructure A is dependent on infrastructure B if the state, functionality, and/or performance of infrastructure A depends on the state, functionality, and/or performance of infrastructure B. The interdependencies are *unidirectional* if infrastructure A is dependent on infrastructure B, but not vice versa; if infrastructures A and B are mutually dependent on each other, the interdependencies are *bidirectional* [11], [14].

Relationships of interdependencies between CIs can be classified in different ways. The most used classification of interdependencies is the one proposed in [11], where the authors identify four classes of interdependencies:

- *physical*: infrastructure A is dependent on infrastructure B through the flow of physical quantities, such as energy commodities, equipment, or goods. Typical examples are all those systems and infrastructures which are dependent on power networks in terms of electricity supply;

- *cyber*: infrastructure A is dependent on infrastructure B through the flow of data and information. Typical examples are all those systems and infrastructures which are dependent on telecommunications networks in terms of exchange of data and information;

- *geographic*: infrastructure A and infrastructure B (or some of their components) share the same geographical location, and a change in the local environment can impact both the infrastructures;

- *logic*: infrastructure A is dependent on infrastructure B through a relationship that does not belong to the previous categories. Examples can be infrastructures that are dependent on each other through human factors, regulations, policies, or financial markets.

Alternative classifications are available in the existing literature. For example, in [15], the authors analyze the existing classifications of infrastructure interdependencies, sorting them into six categories (types of interdependencies, environment, coupling and response behaviour, type of failure, infrastructure characteristics, state of operation), and propose their own classification based on the dimensions of *ontology* and *epistemology*. The dimension of ontology refers to when interdependencies between CIs exist, and it is classified into two categories [15]:

- *chronic*, when the interdependencies are permanent and exist during the whole infrastructures lifetime;

- *episodic*, when the interdependencies are temporary and exist only during particular circumstances.

The dimension of *epistemology* refers to the classification of interdependencies parallelly to the mathematical models used to describe them, and it is divided into four categories [15]:

- *hazard and exposure* models, which describe the interdependencies in terms of environment and hazard occurrence;
- *policy and control* models, which capture the interdependencies due to economic and regulatory factors;
- *operation and performance* models, which describe the operational nature of interdependencies in terms of performance, supply/demand, and cascading effect;
- *deterioration and recovery* models, which describe the temporal evolution of interdependencies and their effect on the functional state of CIs [15].

In general, relationships of interdependencies between infrastructures and systems are complex and variegate, and multiple types of interdependencies can exist simultaneously.

1.3.2 . Modeling of interdependent critical infrastructures

A suitable modeling framework is essential for an accurate analysis and/or optimization of interdependent CIs. Different modeling approaches are available in the existing literature, and a comprehensive review of the existing frameworks is presented in [9], where the author identifies six modeling categories: *empirical*, *agent-based*, *system dynamics-based*, *economic theory-based*, *network-based*, and *other approaches*. Some illustrative existing works and considerations on the modeling of interdependent CIs highlighted in [9] are recalled in this section.

Empirical approaches focus on available historical data and expert judgement to infer the nature and the strength of interdependencies between different infrastructures. These approaches are useful to identify failure patterns and failure propagation between existing infrastructures. For example, these models have been used to identify failure patterns [16], [17], empirically quantify interdependency indicators and metrics [18], or perform risk analyses [19].

Agent-based approaches found their foundations in the fact that interdependent infrastructures can be interpreted as *complex adaptive systems* [20], and their macroscopic behaviour can be modeled as the result of microscopic interactions of multiple agents [21]. For example, these models have been used to study power

systems and their dependent infrastructures [22].

System dynamics-based approaches are another class of models that describe interdependent infrastructures as complex adaptive systems in terms of concepts of feedbacks, stocks, and flows. For example, a system dynamics-based approach has been used by a consortium of US National Laboratories (Los Alamos, Sandia, and Argonne) [23] to develop a supportive tool for CIs protection [24].

Economic theory-based approaches apply economic models to the study of interdependent CIs. For example, the traditional Leontief's *input-output* model [25] has been adapted into the CIs inoperability model proposed in [26]. Input-output based models have been extended to the *Computable General Equilibrium* method to study economic resilience of CIs in case of extreme weather events [27] and terroristic attacks [28].

Network-based approaches utilize networks to represent interdependent infrastructures and study the complex correlations and behaviours that can arise. These models have been widely used in the existing literature, for analyzing and optimizing interdependent CIs with topological approaches [13], [29]–[31] and flow-based approaches [32]–[34].

Other approaches include hierarchical holographic modeling method [35], high level architecture-based method [36], petri-net-based method [37], dynamic control system theory-based method [38], and Bayesian networks [39].

The modeling of interdependent CIs include a large variety of frameworks and approaches. A comprehensive review is out of the scope of this work, and the reader is referred to the extensive review available in [9].

In the context of this thesis, network-based models have been extensively used. In the next section, the main concepts of network-based models, along with some important results of network science related to CIs, are presented.

1.3.3 . Network-based models

Network science is a branch of mathematics that focuses on the study of complex networks and their properties. A complex network is a mathematical object, represented by a graph $G = (V, E)$, usually with nontrivial topology [40], composed of a set V, containing N nodes (or vertices), and a set E, containing M edges (or arcs). In this work, each edge k is assumed to be directed and defined by an origin node $O(k)$ and a destination node $D(k)$. Interdependent networks are represented as individual networks connected by directed edges which represent interdependencies (also referred to as *interdependency links*). A representative graphical example of interdependent networks is shown in Figure 1.1.

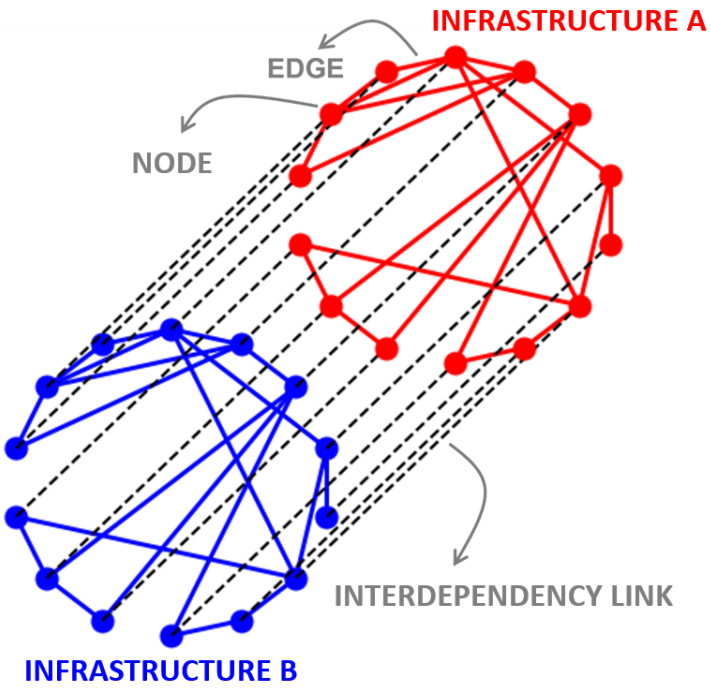

Figure 1.1: Example of interdependent networks. The two networks, in red and blue, consist of nodes connected by edges. Nodes belonging to different networks are connected by dashed black edges, also referred to as interdependency links, that represent relationships of interdependency. Figure from Paper III [10].

CIs can be easily represented by networks: nodes can be used to represent components, while edges can be used to represent connections (physical and non-physical). For example, power networks are composed of buses, which can be represented by nodes, and power lines, which can be represented by edges. Similar representations can be used for railway networks, where stations (nodes) are connected by tracks (edges), gas networks, where hubs (nodes) are connected by pipelines (edges), and many other CIs (telecommunications networks, water networks, etc.). In Figure 1.2, the Great Britain reduced power transmission network [41], composed of 29 buses connected by 50 power lines, is represented as a network of 29 nodes and 50 edges. Similarly, always in Figure 1.2, the proposition for a new high-speed British railway network contained in [42] is represented as a network composed of 16 stations connected by 21 tracks.

Networks are not only an intuitive tool for the graphical representation of CIs, but also a modeling framework for their analysis and optimization. Interdependent CIs can be modeled with two different classes of network models: *topology-based* and *flow-based* [9].

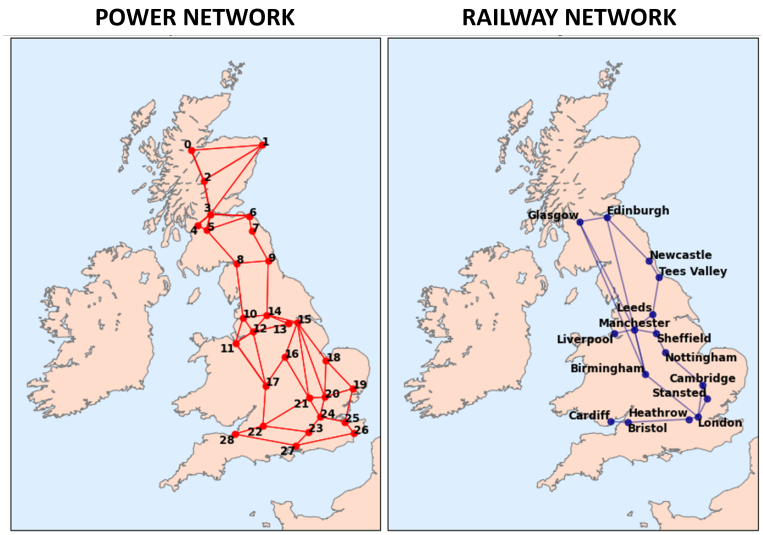

Figure 1.2: Example of CIs represented as networks. In red, the Great Britain reduced power transmission network [41] is shown. In blue, the proposition for a new high-speed British railway network made in [42] is shown. Figure based on Paper I [14].

In topology-based models, nodes and edges are considered to be homogeneous, and their physical properties, such as node production capacity or edge flow capacity, are not taken into account. CIs are modeled and analyzed only accounting for their topological properties and network metrics, such as node degree and betweenness. Node degree and betweenness belong to the so-called *centrality measures*, a group of metrics that defines the importance of each node (or edge) within a network. The degree centrality denotes the number of edges connected to a node. If the edges are directed, the degree centrality can be split into in-degree and out-degree, by distinguishing edges entering and exiting the node [43]. The betweenness centrality denotes the fraction of shortest paths within the network passing through a specific node i, and it is computed as in Equation (1.1):

$$c_B(i) = \sum_{j \neq i \neq k \in V} \frac{\sigma_{jk}(i)}{\sigma_{jk}} \tag{1.1}$$

where σ_{jk} is the number of shortest paths between nodes j and k, and $\sigma_{jk}(i)$ is the number of shortest paths between nodes j and k passing through node i [44]. Other centrality measures, such as closeness centrality [45] and graph centrality [46] exist and can be used for the analysis of networks and CIs. In general, these metrics contain valuable topological information, and they can also be used as surrogate models for physical properties of CIs. For example, betweenness centrality can be used to simulate cascading failures processes in systems with heterogeneous loads, such as power grids and Internet networks [47], [48].

Within the existing literature on CIs, topology-based models have been used to assess the resilience of telecommunications networks [49], [50], identify critical locations [51], design coupling interfaces of interdependent CIs [31], and analyze cascading failures in interdependent CIs [13].

The pioneering work of *Buldyrev et al.* [13] is of particular relevance in the context of interdependent CIs. In this paper, the authors highlight that the behaviour of interdependent networks in terms of cascading failures and failures propagation is inherently different to the behaviour of single non-interacting networks. In particular, using percolation theory as a framework to model cascading failures in networks [52], they demonstrate that interdependent networks are characterized by a first-order phase transition when subject to cascading failures, contrary to single networks which are characterized by a second-order phase transition. This feature results in an increased vulnerability of interdependent networks, and into *"the need to consider interdependent network properties in designing robust networks"* [13]. This consideration provides a solid theoretical background that justifies the study of interdependent CIs, rather than individual CIs.

Topology-based models represent a versatile modeling approach; however, topological information and metrics are often not enough for a comprehensive assessment and analysis of interdependent CIs. Flow-based models integrate the network-based representation of CIs with physical properties, such as production capacity, requested demand, and flow capacity of goods, services, or commodities supplied by CIs. Nodes and edges are heterogeneous, as they are characterized by different physical properties. The functionality and the performance of CIs are described by a flow-based approach, such as maximal flow models [34], [53]–[56], power flow models (for power networks) [57]–[59], or pressure-driven models (for water networks) [60]. Among the various existing applications, flow-based models have been applied to optimize the recovery of disrupted CIs [61]–[63], model cascading failures within power networks [64]–[66], assess the vulnerability of interdependent CIs [14], [67], [68], and enhance the resilience of interdependent CIs [34], [55], [69]–[71].

Flow-based models, as they integrate network representation and physical modeling of CIs, represent a more realistic approach, and they are applied within the work of this thesis.

1.4 . Resilience: concept and metrics

Resilience is a concept that has gained increasing attention in recent years. However, an exact definition of resilience seems to be missing, and several interpretations are present within the existing literature [72], [73]. In this dissertation, it

is interpreted as the ability of a system to *"withstand stressors, adapt, and rapidly recover from disruptions"* [74]. In general, resilience defines the behaviour of a system or a network under conditions of disruption, in terms of temporal evolution of performance. It can be described by resilience curves, such as the one shown in Figure 1.3.

As it can be clearly seen in Figure 1.3, resilience is divided into three phases: disturbance phase, degraded phase, and recovery phase [75]. A system in a stable state with nominal performance p_0, after a disruptive event at time t_e, enters the disturbance phase, which describes the disruption propagation and temporal decrease of performance $p(t)$. The disturbance phase is strictly connected to the concepts of *vulnerability* and *survivability*. Vulnerability is defined as the *"degree of loss or damage to a system when exposed to a strain of a given type and magnitude"* [68], and it can be interpreted as the loss of performance due to a specific disruptive event. Survivability is defined as *"the capability of a system to fulfill its mission in a timely manner in the presence of attacks, failures, or accidents"* [76], and it can be interpreted as the residual performance after a specific disruptive event. At time t_d, when the performance $p(t)$ reaches the minimum value, the degraded phase starts. This phase describes the time necessary for information collection, organization, and decision-making. At t_r, the recovery phase starts and lasts until time t_f, when the performance are back to the nominal value p_0.

In reality, the separation between phases is not always straightforward, and the three phases are often overlapped. However, these concepts are still useful to understand the nature of resilience in systems and networks.

The resilience of a system can be measured using different approaches, and various metrics are available in the existing literature [77]. In this section, two of the most common approaches are mentioned.

The first method consists in taking an integral approach, by measuring the area below the resilience curve (or part of it). Using the nomenclature in Figure 1.3 as a reference, the integral resilience metric $\mathcal{R}$ can be computed as in (1.2):

$$\mathcal{R} = \int_{t_e}^{t_f} p(t)\ dt. \tag{1.2}$$

This approach is equivalent to computing the area under the black curve in Figure 1.3 between t_e and t_f, and it can be interpreted as the cumulative performance of the system between t_e and t_f. A similar approach consists in computing the cumulative losses of the system, corresponding to the area between the dashed horizontal line that defines the nominal performance p_0 and the solid black curve between t_e and t_f.

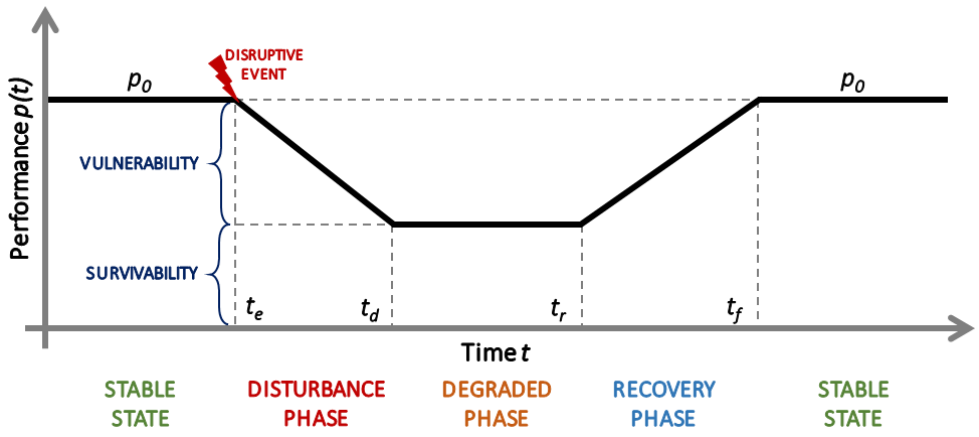

Figure 1.3: Traditional resilience curve with resilience phases.

A second method consists in computing separately different metrics related to various aspects of resilience. A renowned approach is called $\Phi\Lambda E\Pi$ (pronounced "FLEP") [75], and it consists in the computation of four different metrics [10]:

- Φ: it corresponds to the rate of performance drop in the disturbance phase. Using Figure 1.3 as a reference, it can be computed as in (1.3):

$$\Phi = \frac{p(t_e) - p(t_d)}{t_d - t_e}; \tag{1.3}$$

- Λ: it corresponds to the magnitude of the drop in performance. This metric is related to the concept of vulnerability and survivability. It can be computed as in (1.4):

$$\Lambda = p(t_e) - p(t_d); \tag{1.4}$$

- E: it corresponds to the temporal extension of the degraded phase, and it can be computed as in (1.5):

$$E = t_r - t_d; \tag{1.5}$$

- Π: it corresponds to the rate of recovery, and it can be computed as in (1.6):

$$\Phi = \frac{p(t_f) - p(t_r)}{t_f - t_r}. \tag{1.6}$$

In this dissertation, the focus is on vulnerability and survivability, strictly connected to the Λ metric of the $\Phi\Lambda E\Pi$ approach.

1.5 . Timeline and contributions of the thesis

The first year of doctorate was dedicated to the investigation of interdependent railway and power networks (IRPNs), in order to assess their mutual risks. During this phase, three literature gaps and research questions were identified:

- How to model the interconnections between railway and power networks and how do they affect the operational model?
- What is the impact of cascading failures in power networks on the dependent railway networks?
- How to assess the impact of failures in railway networks on power networks?

These questions are addressed by introducing traction networks to act as an interface between railway and power networks, by proposing a flow-based approach for cascading failures in IRPNs, and by evaluating the feedback effect of failures in railway networks on the cascading failure dynamics of power networks. The main results are highlighted in Chapter 2, which is based on the work contained in Paper I [14] and Paper II [78].

One of the main results in Paper II suggested that the topology of interdependency links between interdependent CIs, i.e. how interdependent CIs are interconnected and coupled, might play an important role in determining their resilience, and a new research question emerged:

- How to optimize the topology of interdependency links (also referred to as the *coupling interface*) in order to enhance the resilience of interdependent CIs?

The existing literature presents a limited number of works on this topic, mostly based on network metrics-based heuristic strategies. In this work, the optimization of coupling interface topology is addressed with a mathematical programming approach. Two different models (robust and distributionally robust) for the optimal design of coupling interfaces are proposed, as well as an approach for the optimal allocation of redundant interdependency links. The main results are highlighted in Chapter 3, which is based on the work contained in Paper III [10], Paper IV [79], and Paper V [80].

When optimizing the coupling interface, the interdependent CIs are assumed to be operated in a centralized way. However, interdependent CIs can also be operated in a decentralized way, in which the different operators act independently. In particular, CIs are usually operated in a decentralized way under normal conditions, when the focus of each operator is to minimize its own individual costs, and in a centralized way under conditions of disruption, when the different operators collaborate in order to minimize the negative impact on the general population. When

designing a coupling interface - or taking any other joint decision - planners and decision-makers should take into account that the behaviour of independent operators in interdependent CIs might change according to the state of their systems. Given this consideration, a new research question emerged:

- How to take joint decisions in interdependent CIs when independent operators might display different behaviours according to the systems' conditions?

This question is addressed in Chapter 4, where some preliminary results based on Paper VI are presented. Moreover, a proposal for a decision-making framework that takes into account the behaviour of independent operators (decentralized under normal conditions, centralized under conditions of disruption) of interdependent CIs is presented.

In Chapter 5, the main results and contributions of this thesis, as well as possible developments, are highlighted.

At the end of the manuscript, the aforementioned papers are appended. Paper VII, which contains the results of a distributionally robust approach with endogenous uncertainty for the optimal protection of power networks, is appended but not discussed in this manuscript.

2 - Vulnerability of railway and power networks

2.1 . Introduction

The disturbance phase of resilience is strictly connected to the concept of vulnerability, which is defined as the *"degree of loss or damage to a system when exposed to a strain of a given type and magnitude"* [68]. In other words, the vulnerability $\mathcal{V}$ of a system defines the drop in performance after a specific disruptive event, and it can be generally computed as in (2.1):

$$\mathcal{V} = \begin{cases} p(t_e) - p(t_d), & \text{if not normalized} \\ \\ \frac{p(t_e)-p(t_d)}{p(t_e)}, & \text{if normalized} \end{cases} \tag{2.1}$$

where $p(t_e)$ and $p(t_d)$, consistently with Figure 1.3, define the performance of the system before and after the disruptive event, respectively. The vulnerability $\mathcal{V}$ is a time-independent metric, and when not normalized, is equivalent to the Λ metric of the $\Phi\Lambda E\Pi$ resilience framework.

When dealing with CIs, analyzing and understanding the possible negative consequences that might arise from various disruptive events is of the utmost importance. Vulnerability analysis, defined in [68] as the process of *"systematically and comprehensively identifying the possible states a system can be put into, due to specific strains, and estimating the negative consequences associated with them"*, is a suitable framework for estimating negative impacts of disruptive events in CIs.

In the case of interdependent CIs, a comprehensive vulnerability analysis must take into account the effect of interdependencies and the mutual impact that interdependent CIs can have on each other. As it is highlighted in [67] and [68], vulnerability analysis is a suitable framework for the evaluation of interdependent CIs. However, the results of the analysis are heavily impacted by the starting assumptions and the modeling of cascading effects[1] between different CIs.

The focus of this work is on interdependent railway and power networks (IRPNs), in order to understand the mutual risks associated with these infrastructures. Consistently with the framework in [11], the interdependencies between these two infrastructures are defined as unidirectional and physical, as railway networks (when

[1]Cascading effect defines the propagation of disruption and disturbances from one infrastructure to another [14].

electrified) are dependent on power networks in terms of electricity supply.

Railway and power networks are amongst the most important infrastructures in any advanced society. As railway networks are often electrified, especially in Europe, IRPNs are a very common configuration. In addition, it is well known that failures in power networks can rapidly propagate and cause disruption in railway networks [12]. Consequently, railway operators should be aware of risks related to disruption of power supply from power networks [81]. Surprisingly, in the existing literature, vulnerability analysis of IRPNs is not treated comprehensively and sufficiently in detail. The reader is referred to Section 1.2 of Paper I for more details.

The existing literature presents some common drawbacks:

- the structure of the coupling interface, i.e. how IRPNs are coupled and interconnected, is oversimplified;
- cascading failures in power networks and their effect on dependent railway networks are overlooked or treated approximately;
- the feedback effect of failures in railway networks on the cascading failure dynamics of power networks is not evaluated.

The first drawback is related to the modeling of coupling interfaces between IRPNs, and in particular the modeling of traction networks. In fact, electrified railway networks are supplied by power networks through traction networks, composed of electrical substations fed by an external power network that regulate the electricity supply to the railway catenary[2]. In this context, the term *coupling interface* refers to traction networks, which indeed act as an interface, and the interdependency links connecting traction networks to power networks and railway networks. In the existing literature, with the exception of [67] and [68], this topological configuration is overlooked, and railway and power networks are often connected directly by interdependency links without modeling explicitly traction networks (see Paper I for more details). This might result in an unrealistic topological configuration.

The second drawback is related to the modeling of cascading failures in power networks and their consequences on railway networks in terms of vulnerability. Cascading failure is defined as a *"kind of failure in a system comprising interconnected parts, in which the failure of a part can trigger the failure of successive parts"* [82], and it can affect power networks. Cascading failures in power networks can be modeled with various approaches (the reader is referred to [83] for a comprehensive review). In the context of IRPNs, cascading failures in power networks

[2]The catenary is a dedicated power line, parallel to railway tracks, which supplies rolling stocks with electricity.

and their impact on railway networks are evaluated only in [84]–[86]. However, in these works, cascading failures are modeled with a network-based approach [87], [88], which is not suitable to capture physical interdependencies based on power supply. Network-based approaches for cascading failure modeling usually apply network metrics, such as betweenness centrality, as a surrogate model for power flows, and the results obtained by these models are locally inconsistent with the results obtained with more realistic flow-based models [89].

The third drawback is related to the modeling of the impact of failures in railway networks on power networks. Failures within railway networks modify the power demand within power networks, impacting their cascading failure dynamics. This aspect has not been addressed within the existing literature.

From this exploratory literature review on vulnerability of IRPNs, three main research questions emerged, related to the identified research gaps:

- How to model the interconnections (the coupling interface) between railway and power networks and how do they affect the operational model?
- What is the impact of cascading failures in power networks on the dependent railway networks?
- How to assess the impact of failures in railway networks on the cascading failure dynamics of power networks?

In Paper I, these drawbacks are addressed by introducing traction networks within the analysis, by proposing a cascading failure model for IRPNs based on the traditional ORNL-PSerc-Alaska (OPA) model [64]–[66], and by evaluating the feedback effect of railway network failures on the cascading failure dynamics of power networks.

In Paper II, a preliminary model which is able to account for different configurations of traction networks is proposed. This chapter focuses on the work contained in Paper I. Some insights on Paper II are available in Section 2.6.

2.2 . Modeling framework

2.2.1 . Operational model

Each network is modeled as a graph $G = (V, E)$, and characterized by the subscript PN (power network), TN (traction network), and RN (railway network).

Power networks operations are modeled with a DC Optimal Power Flow (DC-OPF) model. Each node i represents an electrical bus, containing generators j, each with power production capacity $\overline{p}_{i,j}$, and loads j, each with requested

power demand $\overline{d}_{i,j}$. Buses are connected by transmission lines, characterized by a power flow capacity $\overline{f}_k$. The DC-OPF model in (2.2)-(2.8) is implemented using Pandapower [90]:

$$\min_{\mathbf{p},\mathbf{d},\mathbf{f}} \sum_{i \in V_{PN}} \sum_{j \in G_i} p_{i,j} - W \sum_{i \in V_{PN}} \sum_{j \in L_i} d_{i,j} \tag{2.2}$$

$$0 \leq p_{i,j} \leq \overline{p}_{i,j}, \quad \forall i \in V_{PN}, \forall j \in G_i \tag{2.3}$$

$$0 \leq d_{i,j} \leq \overline{d}_{i,j}, \quad \forall i \in V_{PN}, \forall j \in L_i \tag{2.4}$$

$$-\overline{f}_k \leq f_k \leq \overline{f}_k, \quad \forall k \in E_{PN} \tag{2.5}$$

$$\sum_{j \in G_i} p_{i,j} - \sum_{j \in L_i} d_{i,j} + \sum_{k \,|\, D(k)=i} f_k - \sum_{k \,|\, O(k)=i} f_k = 0, \quad \forall i \in V_{PN} \tag{2.6}$$

$$x_k f_k - (\theta_{O(k)} - \theta_{D(k)}) = 0, \quad \forall k \in E_{PN} \tag{2.7}$$

$$-\overline{\theta} \leq \theta_i \leq \overline{\theta}, \quad \forall i \in V_{PN} \tag{2.8}$$

where $p_{i,j}$ is the power produced in generator j within the bus i, $d_{i,j}$ is the power supplied to load j within the bus i, and f_k is the power flow in power line k. The goal is to minimize the objective function in (2.2), where W is a penalty constant (in this work, W=100) that guarantees the minimization of load shedding. This objective function is based on the work in [66]. Each bus i can contain multiple generators, contained within the set G_i, and multiple loads, contained within the set L_i. The power produced in each generator $p_{i,j}$ and the power supplied to each load $d_{i,j}$ are bounded by production capacity and requested power demand, as shown in Constraints (2.3) and (2.4), respectively. As shown in Constraint (2.5), the power flow in each line f_k is bounded, in absolute value, by the flow capacity. The net power balance in each node must be 0, as enforced in Constraint (2.6). The DC power assumption is enforced in Constraint (2.7), where θ_i is the phase angle in bus i and x_k is the reactance of line k. The phase angle in each bus i is bounded, in absolute value, by a maximum value $\overline{\theta}$, as shown in Constraint (2.8). For more details, the reader is referred to Paper I, particularly Section 2.4[3].

Traction networks act as an interface between IRPNs. Traction networks consist of multiple isolated substations, such as in AC-electrified railway networks, which depend on power networks in terms of power supply. The binary functional state of each substation i is dependent on the corresponding power network. In particular, each substation i in a traction network is dependent on a load j within the bus h of the corresponding power network. The binary functional state $S_{t,i}$ of

[3]In order to maintain notations, nomenclature, and formulations consistent throughout this dissertation, they might differ from the ones proposed in the appended papers. However, notations, nomenclature, and formulations proposed in this thesis are equivalent (or at least similar) to the ones in the appended papers, and they do not alter the message conveyed by the original works.

substation i (1 if functional, 0 if failed) depends on the power supplied to load j in bus h according to the relation in Equation (2.9):

$$S_{t,i} = \begin{cases} 1, & \text{if } R_{h,j}^{t,i} \geq T_{t\leftarrow p} \text{ and } 0 < T_{t\leftarrow p} \leq 1 \\ 0, & \text{if } R_{h,j}^{t,i} < T_{t\leftarrow p} \text{ and } 0 < T_{t\leftarrow p} \leq 1 \\ 0, & \text{if } R_{h,j}^{t,i} = T_{t\leftarrow p} \text{ and } T_{t\leftarrow p} = 0 \end{cases} \tag{2.9}$$

where $R_{h,j}^{t,i}$ defines the ratio between power supplied and requested power demand of load j in bus h, as defined in Equation (2.10):

$$R_{h,j}^{t,i} = \frac{d_{h,j}}{\bar{d}_{h,j}}. \tag{2.10}$$

The parameter $T_{t\leftarrow p} \in [0, 1]$ defines the level of tolerance of traction substations to lack of electricity supply from the power network. For example, if $T_{t\leftarrow p}$=1, each substation is considered as failed if its corresponding load is not supplied with its entire requested power demand; if $T_{t\leftarrow p}$=0, each substation is considered as functional as long as its corresponding load is supplied with some electricity. For values $0 < T_{t\leftarrow p} < 1$, each substation is functional if at least a fraction $T_{t\leftarrow p}$ of electricity is supplied to its corresponding load. The values of $T_{t\leftarrow p}$ selected in Paper I are 0.0, 0.5, and 1.0.

Railway networks are composed of nodes, representing stations, connected by edges, representing tracks. They are modeled with a topological approach, and they are dependent on traction networks in terms of electricity supply. Specifically, it is assumed that each railway track k is dependent on a subset of traction substations $V_{TN}^{RN,k} \subseteq V_{TN}$. Each railway track k is functional ($S_{r,k}$=1) if all the substations in $V_{TN}^{RN,k}$ are functional, as defined in Equation (2.11):

$$S_{r,k} = \prod_{i \in V_{TN}^{RN,k}} S_{t,i}. \tag{2.11}$$

2.2.2 . Cascading failure model

Disruption within power networks can trigger cascading failure processes, which can propagate to railway networks through traction networks and the corresponding interdependency links. The proposed cascading failure modeling approach is based on the ORNL-PSerc-Alaska (OPA) model [64]–[66] (see Paper I for more details). The OPA model is a realistic flow-based modeling approach that has been validated with historical data [65], and it involves the following steps:

1. initialize network and initiating disruptive event. Go to Step 2;
2. remove failed components and go to Step 3;
3. run DC-OPF in Equations (2.2)-(2.8) and go to Step 4;

4. check the loading fraction $f_k/\overline{f}_k$ for each line k. If $f_k/\overline{f}_k \geq$0.99, line k fails with probability $p_{ol} \in [0, 1]$. Go to Step 5;

5. if at least one line failed at Step 4, return to Step 2; otherwise, the algorithm is terminated.

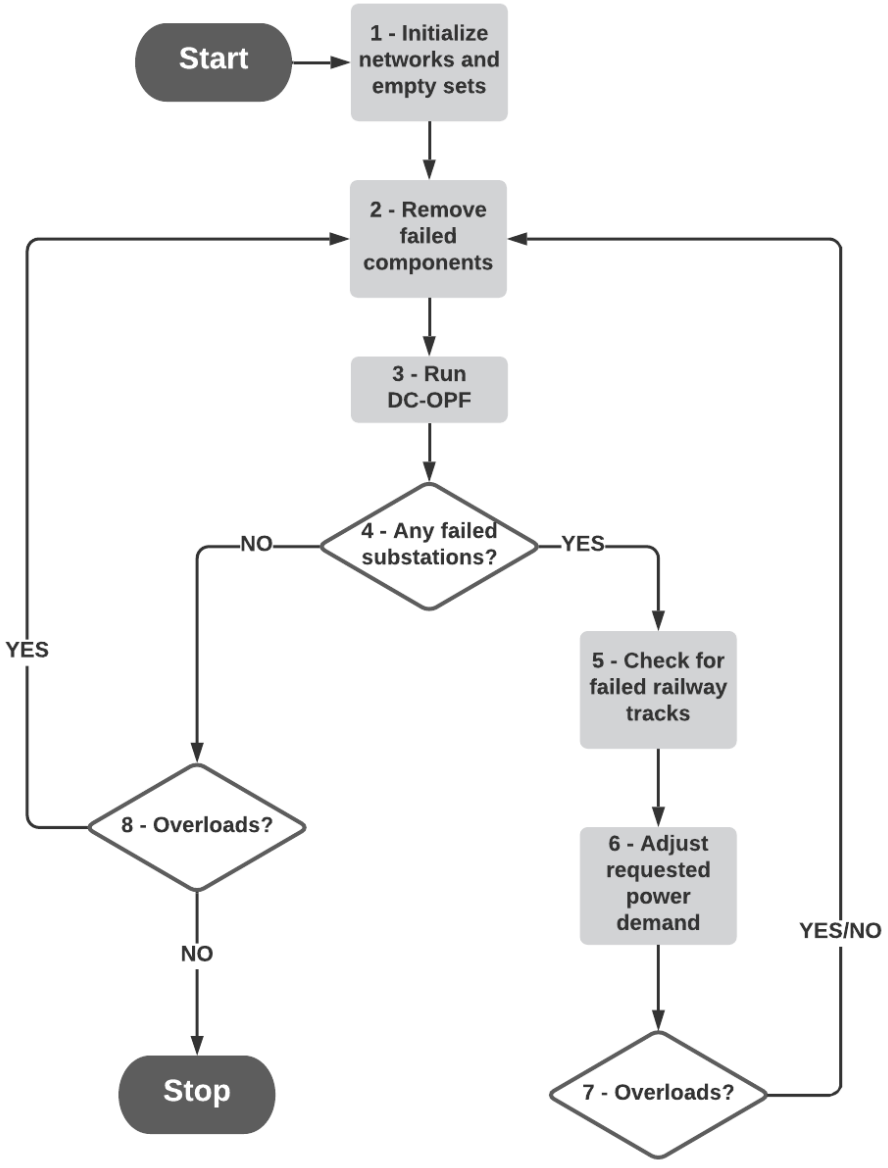

Figure 2.1: Proposed cascading failure algorithm for IRPNs. Figure from Paper I [14].

The proposed cascading failure model for IRPNs is based on the OPA model and accounts for the impact of failures in railway networks on the cascading failure dynamics of power networks. The flowchart is shown in Figure 2.1. The main

concept to highlight is that, in case of failures within railway networks, the loads in the corresponding power networks do not require any power supply. This can modify the cascading failure dynamics of power networks. The proposed approach consists of the following steps:

1. initialize networks and initiating disruptive event. Go to Step 2;
2. remove failed components and go to Step 3;
3. run DC-OPF in Equations (2.2)-(2.8) and go to Step 4;
4. assess functional states of traction substations using Equation (2.9). If at least one substation failed, go to Step 5; otherwise go to Step 8;
5. assess functional states of railway tracks using Equation (2.11) and go to Step 6;
6. adjust the requested power demand in the power network according to the failed railway tracks and go to Step 7;
7. check the loading fraction $f_k/\overline{f}_k$ for each line k. If $f_k/\overline{f}_k \geq$0.99, line k fails with probability $p_{ol} \in [0, 1]$. Return to Step 2;
8. check the loading fraction $f_k/\overline{f}_k$ for each line k. If $f_k/\overline{f}_k \geq$0.99, line k fails with probability $p_{ol} \in [0, 1]$. If at least one line failed, return to Step 2; otherwise, the algorithm is terminated.

2.3 . Vulnerability analysis

Vulnerability analysis consists in identifying the negative impact that various disruptive events can have on one or more systems or infrastructures. The three main steps of a vulnerability analysis are:

- define the initiating disruptive event;
- define the new states of the systems/infrastructures after the disruptive event;
- compute the associated negative consequences.

For a comprehensive vulnerability analysis, these three steps must be iteratively repeated for multiple types of initiating disruptive event.

In Paper I, the initiating disruptive events are simulated with the random removal of an increasing fraction of edges within the power network [67], [68]. In order to make the analysis agnostic about the cause of failures, no spatial nor temporal correlation between failed components is taken into account. This approach

is typical of network science, and it can be used to simulate initiating disruptive events with different magnitudes. The new states of the IRPNs are then computed taking into account cascading failures and cascading effects within and between networks, using the algorithm shown in Figure 2.1.

Without indulging in details, the proposed approach allows to evaluate:

- cascading failures in power networks, with a power flow-based approach, and their propagation on railway networks through traction networks;
- the impact of failures in railway networks on the cascading failure dynamics of power networks.

The negative impact of disruptive events is computed in terms of loss of performance. For example, in power networks, the negative impact can be computed in terms of fraction of demand not supplied $FDNS$, while in railway networks, it can be computed in terms of loss of topological accessibility A_r (the average fraction of stations accessible from each other [91]). The average results are evaluated for fractions of removal from 0% to 100%, with steps of 10%, and 1000 simulations per fraction of removal. In Paper I, vulnerability analyses with cascading failures included (p_{ol}=1) and not included (p_{ol}=0) are compared.

2.4 . Case study

This approach is applied to investigate the vulnerability of a British case study. The system is composed of a railway network, a traction network, and a power network. The railway network consists of 16 nodes (stations) and 21 edges (tracks), and it is based on a proposition made in [42]. The power network is based on the Great Britain reduced power network, and it is composed of 29 nodes (buses) and 50 edges (lines), 49 in double-circuit configuration and one in single-circuit configuration [41]. The railway and power networks are connected by the traction network, composed of 85 isolated nodes (substations). Each substation is connected with, and supplied by, the geographically-closest bus in the power network. Each railway track is supplied by a subset of substations. The geographical and network-based representations of this illustrative case study are shown in Figures 2.2 and 2.3, respectively.

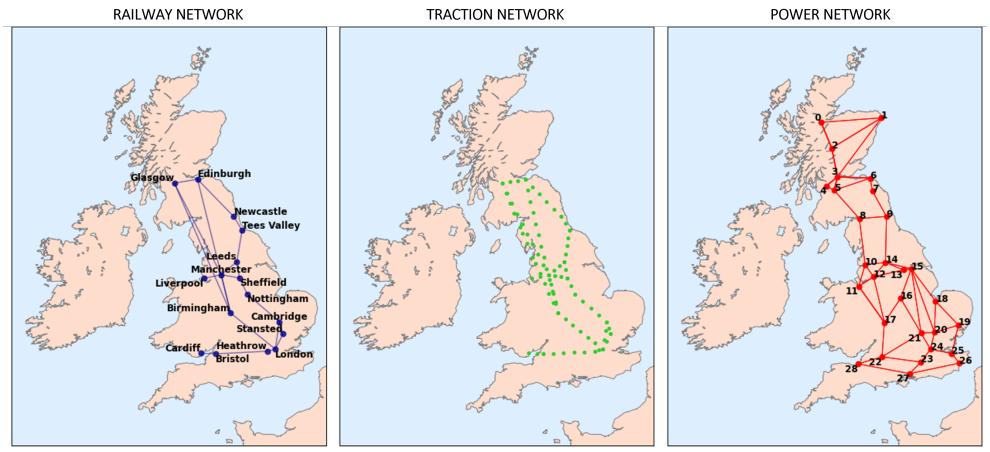

Figure 2.2: Geographical representation of the British IRPNs, based on Figure 3 of Paper I [14].

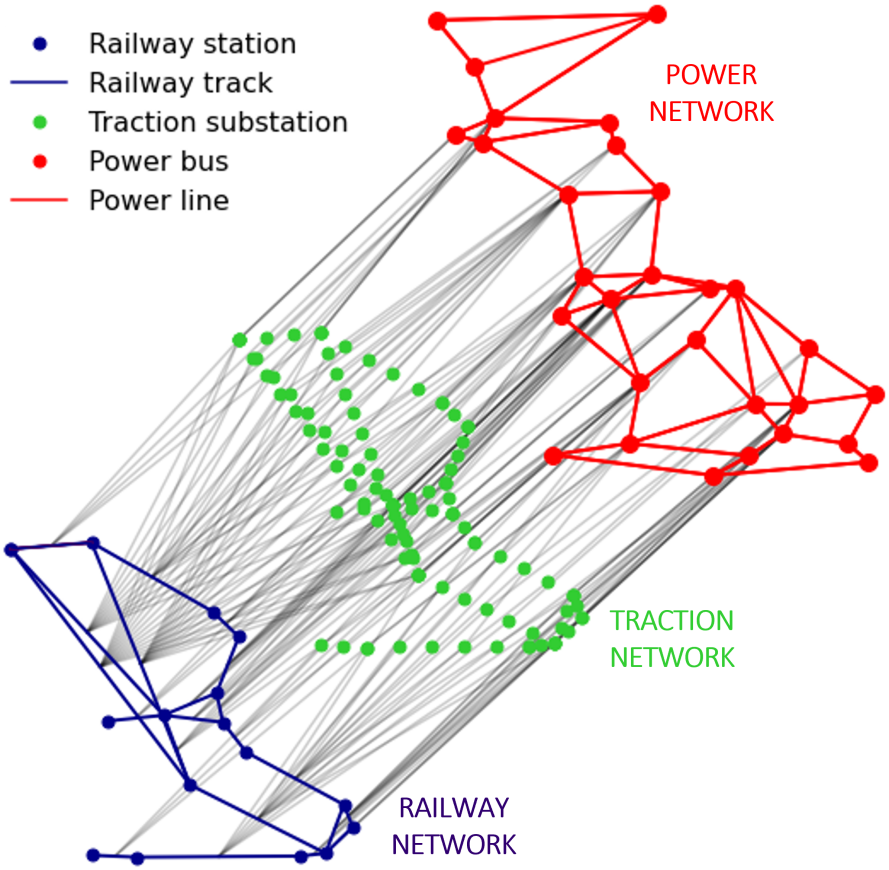

Figure 2.3: Network-of-networks representation of the British IRPNs, based on Figure 4 of Paper I [14].

2.5 . Results

The results in terms of loss of accessibility in the railway network are shown in Figure 2.4. The x-axis denotes the fraction of removed edges in the power network, which is used to simulate initiating disruptive events. The y-axis denotes the loss of accessibility. Solid curves denote results when cascading failures within power networks are included in the analysis (p_{ol}=1), while dashed curves denote results without accounting for cascading failures (p_{ol}=0). The three colors denote different levels of tolerance of traction substations to lack of electricity supply, as defined by the parameter $T_{t \leftarrow p}$.

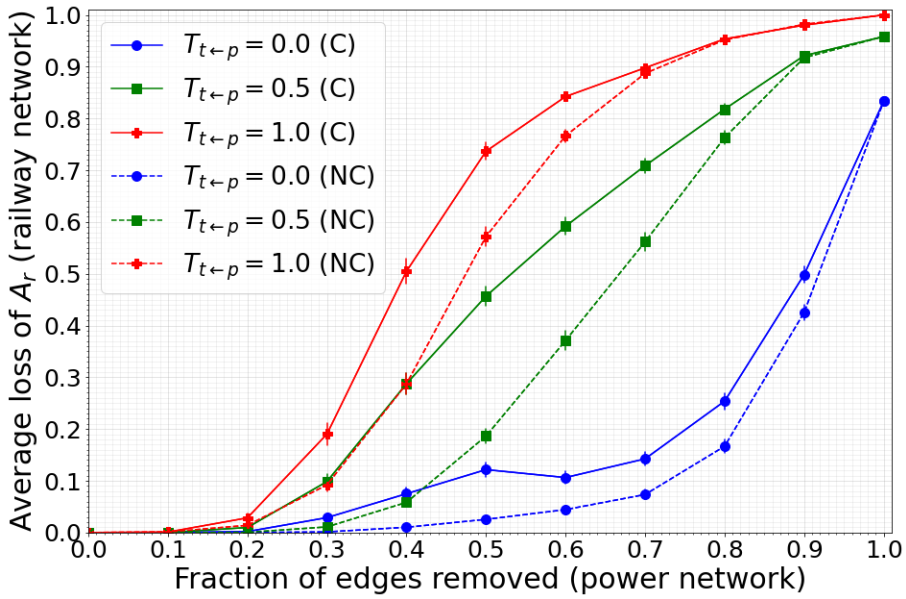

Figure 2.4: Average loss of accessibility in the railway network due to different fractions of edges removed in the power network. Figure from Paper I [14].

As it can be clearly seen, the tolerance $T_{t \leftarrow p}$ of traction substations plays a key role in terms of vulnerability of electrified railway networks, as the the results for different values of $T_{t \leftarrow p}$ vary considerably. In addition, as it can be clearly noticed by comparing solid and dashed lines, it is essential to include cascading failures within the vulnerability analysis of IRPNs, in order to avoid vulnerability underestimation.

In Figure 2.5, the results in terms of $FDNS$ in the power network are shown. The x-axis denotes the fraction of removed edges in the power network. The y-axis denotes the fraction of demand not supplied. The black curve denotes the results of a traditional OPA model, which does not account for the effect of failures in the railway network on the cascading failure dynamics of the power

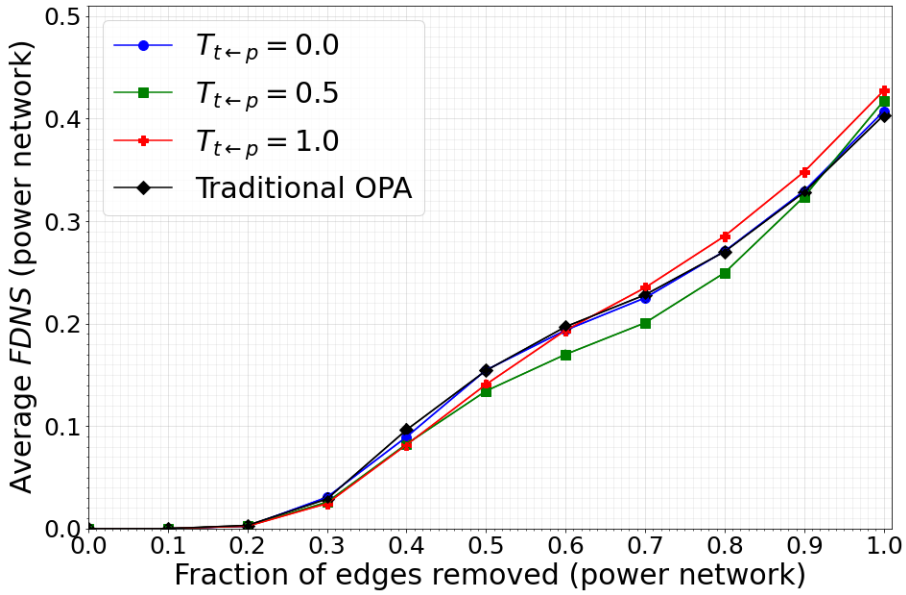

Figure 2.5: Average FDNS in the power network due to different fractions of edges removed in the power network. Figure from Paper I [14].

network. The other curves include the impact of railway network failures within the analysis. As it can be clearly seen, when including the effect of failures in the railway network on the cascading failure dynamics of the power network, the results differ from a traditional OPA model, indicating that including the feedback effect of interdependencies on cascading failure dynamics is essential in order to estimate accurately the vulnerability of IRPNs.

2.6 . On the coupling interface topology

In Paper I, traction substations are assumed to be in electrical isolation and disconnected from each other. This corresponds to the configuration of an AC-electrified railway network. For DC-electrified railway networks, traction substations are often more numerous and connected with each other. A preliminary model for evaluating the vulnerability of IRPNs with different traction network configurations is proposed in Paper II. Within this framework, when traction substations are connected to each other, such as in a DC-electrified railway network, traction networks are modeled as a power network. Specifically, they are modeled using a DC-OPF model. A preliminary vulnerability analysis is performed on the French case study shown in Figure 2.6.

The initiating disruptive events are simulated by removing, one by one, buses in the power network. The negative impact of removals is measured on the three networks. For example, in the power and traction networks, the negative impact is

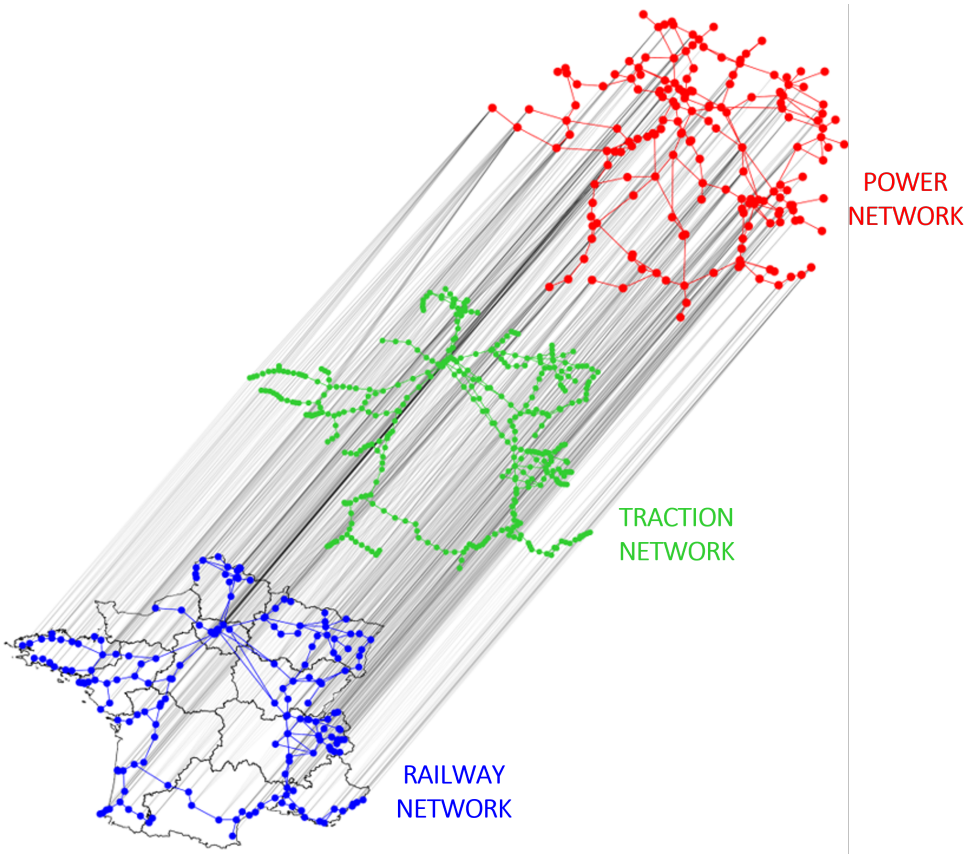

Figure 2.6: Network-of-networks configuration of the French IRPNs, based on Figure 1 of Paper II [78].

computed in terms of $FDNS$, and the results are shown in Figure 2.7. For more details on modeling and results, the reader is referred to Paper II.

As it can be seen in Figure 2.7, the worst-case scenarios for the power network and the traction network, in terms of $FDNS$, are different. In particular, the scenario which leads to the highest $FDNS$ in the power network is the failure (removal) of bus 83, which leads to $FDNS$=0.210; for the traction network, the worst scenario is the failure (removal) of bus 78, which leads to $FDNS$=0.146[4]. Failures and disruption in the power network spread to the traction network through their interdependency links, collectively referred to as the coupling interface. Clearly, as the worst-case scenario of the power network does not correspond to the worst-case scenario of the traction network, the topology of their coupling interface, i.e. where the interdependency links are present, plays a key role in terms of failures

[4]An accurate analysis of the features and the characteristics of these buses is out of the scope of this work; however, it is important to highlight that their topological positions within the interdependent networks play a key role in determining their criticality in case of failure.

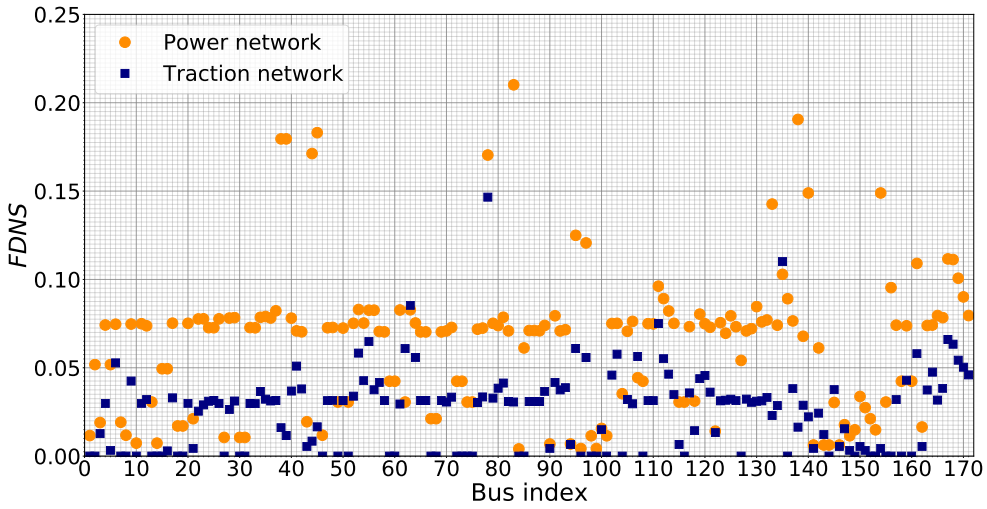

Figure 2.7: Fraction of demand not supplied $FDNS$ in the power network and traction network. Figure based on the results of Paper II [78].

and disruption propagation. This observation leads to the next research question of this dissertation:

- how to design and optimize the topology of coupling interfaces between interdependent CIs in order to enhance their resilience?

2.7 . Conclusion

In this chapter, the models for the vulnerability analysis of IRPNs proposed in Paper I and Paper II are highlighted. The main contribution of this chapter and the related papers is to propose an approach for assessing the vulnerability of IRPNs while accounting for a realistic coupling interface and cascading failures within and between networks. The obtained results demonstrate the importance of considering cascading failures within the vulnerability analysis. In Paper II, preliminary results show that the worst-case scenarios for individual interdependent networks might differ. Intuitively, this result might be strongly influenced by the coupling interface topology. Consequently, in the next chapter, the optimization of coupling interface topology and design is investigated.

3 - Design and optimization of coupling interface topology

3.1 . Introduction

As CIs are the backbone of advanced societies, it is essential to guarantee their resilience under conditions of disruption. In the previous chapter, it was highlighted that the topology of coupling interfaces, i.e. how interdependent CIs are coupled together, might play a key role in terms of vulnerability and survivability, as it defines how failures and disruption propagate between different infrastructures. Despite playing such an important role, coupling interface optimization has not been investigated thoroughly in the existing literature.

NB: *In this chapter, the term "coupling interface" strictly refers to the ensemble of interdependency links connecting two interdependent CIs.*

In the community of network science, it is already acknowledged how coupling interfaces can strongly impact failure propagation between interdependent networks. In [92], it is shown how the coupling strength between interdependent networks, i.e. the fraction of nodes in each network which is dependent on nodes of the other network, strongly impacts the failure propagation between networks. The results are evaluated in terms of transition phase using percolation theory. Similarly, in [93], the authors explore the effect of interdependencies and different coupling strategies on the vulnerability of interdependent networks. Moreover, other authors have already investigated various allocation strategies of interdependency links in interdependent networks using percolation theory as a modeling approach [94]–[96]. These works acknowledge the importance of optimizing the coupling interface topology between interdependent networks. However, their modeling framework, based on percolation theory, is not suitable for real-world decision-making, as it provides only an approximation of the realistic functionality of CIs.

In the existing literature on interdependent CIs, design and optimization of coupling interfaces have been rarely investigated. In fact, in most of the existing literature on interdependent CIs, the coupling interface topology is a constant parameter, and no sensitivity analysis nor optimization is performed. In some works, coupling strategies based on network metrics have been investigated for different types of interdependent infrastructures, such as power and water networks [97] or power and telecommunication networks [98]–[100]. While these papers investigate the impact of different coupling interface topologies on the cascading effect

in interdependent CIs, they still rely on heuristics approaches based on network metrics. Thus, these works can not provide a solid framework for decision-making, as heuristic approaches do not guarantee the optimality of solutions and are hardly generalizable.

The need to optimize coupling interfaces between interdependent infrastructures is acknowledged in [31] and [32]. In [31], the authors rely on network metrics, such as Euclidean distance, node degree, and betweenness, to identify hybrid coupling strategies between power, water and gas networks and reduce cascading effect under conditions of disruption. A similar approach is taken in [32], where the authors propose a similar framework for coupling interdependent CIs, relying on network metrics and physical properties of the CIs, in order to increase their resilience. While these works investigate different hybrid strategies for coupling interdependent infrastructures, their approaches are based on heuristics and network metrics, and they do not guarantee optimal solutions.

In this chapter, a resilience-based mathematical programming framework for the optimal coupling of interdependent CIs is proposed. Firstly, an approach for designing robust coupling interfaces, based on the content of Paper III, is presented. In this paper, the topology of coupling interfaces is optimized in order to maximize the combined survivability of interdependent CIs in the worst-case feasible failure scenario. Secondly, in Paper IV, this framework is extended including uncertainty within the analysis, leveraging on a distributionally robust optimization (DRO) approach. Lastly, in Paper V, the model is extended with the inclusion of redundant interdependency links, which exist in reality but are often overlooked in the existing literature.

3.2 . Operational modeling framework

Interdependent CIs are modeled as networks, and represented by a graph $G{=}(V, E)$, where V is the set of N nodes and E is the set of M edges. Each edge k is directed, and is defined by an origin node $O(k)$ and a destination node $D(k)$. Components belonging to different infrastructures can be connected by interdependency links, which denote relationships of interdependency. The ensemble of these links is referred to as the coupling interface.

Given a system of interdependent CIs, coupled through a coupling interface $\mathbf{y}$, their operations are defined by an operational model $\mathcal{Q}(\mathbf{y}, \boldsymbol{\xi})$, where $\boldsymbol{\xi}$ denotes the uncertainty realization. In this case, as it will be clear in the next sections, the uncertainty $\boldsymbol{\xi}$ defines the functional state of the CIs components, which depends on failures and disruption.

The operational model simulates the actions of CIs operators, and is modeled as an optimization problem. It is assumed that the operators aim at optimizing the combined performance P_{comb} of the interdependent CIs. For a set I of interdependent infrastructures, the combined performance P_{comb} can be generally defined as in Equation (3.1):

$$P_{comb} = \sum_{i \in I} w_i P_i \tag{3.1}$$

where P_i is the individual performance of infrastructure i, and w_i is its weight (i.e. its importance) when computing the combined performance.

For a fixed coupling interface $\mathbf{y}^*$ and fixed failure scenario $\mathbf{u}^*$, which defines the uncertainty realization $\boldsymbol{\xi}$, the operational model $\mathcal{Q}(\mathbf{y}, \boldsymbol{\xi})$ is defined as a maximization problem, represented in its compact matrix formulation as in (3.2)-(3.4):

$$\max_{\mathbf{h}, \boldsymbol{\delta}} \mathbf{b}^T \mathbf{h} \tag{3.2}$$

subject to:

$$\mathbf{R}\mathbf{h} \leq \mathbf{q} - \mathbf{T}\mathbf{u}^* - \mathbf{H}\mathbf{y}^* - \mathbf{W}\boldsymbol{\delta} - \mathbf{y}^{*T}\mathbf{D}\boldsymbol{\delta} \tag{3.3}$$

$$\mathbf{h} \in \mathbb{R}^{N_h},\ \boldsymbol{\delta} \in \{0,1\}^{N_d}. \tag{3.4}$$

The vectors $\mathbf{h}$ and $\boldsymbol{\delta}$, with dimensions N_h and N_d, represent the continuous and binary operational variables of the problem, respectively. The vector $\mathbf{b}$ and the matrices $\mathbf{R}$, $\mathbf{T}$, $\mathbf{H}$, $\mathbf{W}$, and $\mathbf{D}$ contain the coefficients of objective function and constraints, while the vector $\mathbf{q}$ contains the parameters of the constraints. The objective function in (3.2) represents the combined performance of interdependent infrastructures, previously shown in Equation (3.1). Equation (3.3) contains the operational constraints. The details of the optimization problem depend on the infrastructures involved and the operational assumptions. For the full formulation of the operational models, with a detailed explanation of the operational variables $\mathbf{h}$ and $\boldsymbol{\delta}$ and the related operational constraints, the reader is referred to the appended papers (e.g. Section 2 of Paper III)[5].

3.3 . Case study

In Paper III, Paper IV, and Paper V, a case study based on interdependent power and gas networks (IPGNs) is used. The geographical representation of the two infrastructures is shown in Figure 3.1. The power network, defined by the subscript PN, is based on the IEEE 14-bus system [101], composed of 14 nodes and 20 edges. The gas network, defined by the subscript GN, is based on the

[5]For simplicity, the cubic terms present in the operational model of Papers III and IV are assumed to be contained within the last term $\mathbf{y}^T\mathbf{D}\boldsymbol{\delta}$ of Equation (3.3).

IEEE 9-bus system [102], composed of 9 nodes and 9 edges. Each node i is characterized by production capacity of power or gas $\overline{p}_i$ and requested demand of power or gas $\overline{d}_i$. These energy commodities are distributed through power lines and gas pipelines, represented by edges, with a maximum flow capacity $\overline{f}_k$. The power network operations are modeled with a DC power flow model, while the gas network operations are modeled with a maximal flow approach, which represents a linear approximation of gas flow models [34]. In this case, the operational model of the IPGNs, represented by the recourse function $\mathcal{Q}(\mathbf{y}, \boldsymbol{\xi})$, denotes the maximization of the fraction of requested demand of power and gas which is possible to supply. The objective function of $\mathcal{Q}(\mathbf{y}, \boldsymbol{\xi})$ is shown in Equation (3.5):

$$P_{comb} = \frac{w_{PN}}{\overline{d}_{PN}} \sum_{i \in V_{PN}} d_i + \frac{w_{GN}}{\overline{d}_{GN}} \sum_{i \in V_{GN}} d_i. \tag{3.5}$$

The coefficients w_{PN} and w_{GN}, here assumed to be both equal to 0.5, are the weights of power network and gas network when computing the combined performance. The coefficients $\overline{d}_{PN}$ and $\overline{d}_{GN}$ are the total requested power demand in the power network and the total requested gas demand in the gas network. The variable d_i defines the supplied power or gas in each node i of the networks. The combined performance P_{comb} ranges from 0, when no power and gas demand is supplied, to 1, when 100% of the requested demand of power and gas is supplied [10].

The interdependencies between IPGNs are based on the following assumptions:

- each node in the power network with power production capacity $\overline{p}_i > 0$ contains a gas-fired power plant, and is dependent on the gas network in terms of gas supply;
- each node in the gas network needs electricity in order to guarantee the functionality of various electrical equipment, such as pumps or valves, and is dependent on the power network in terms of power supply.

These relationships of interdependency are built within the constraints of the operational model $\mathcal{Q}(\mathbf{y}, \boldsymbol{\xi})$, using mathematical assumptions which are consistent with the existing literature on interdependent CIs [34]. For example, in each node i of the power network with production capacity $\overline{p}_i > 0$, it is possible to produce electricity only if there is a functional interdependency link with a node j in the gas network, i.e. the gas-fired power plant receives the necessary gas supply. It is assumed that the interdependency link from $j \in V_{GN}$ to $i \in V_{PN}$ is functional only if the requested demand of gas in $j \in V_{GN}$ is fully satisfied. The rationale behind this assumption is that, if node j is not supplied with its entire gas demand, the dependent gas-fired power plant in $i \in V_{PN}$ might not receive the necessary gas supply. Similar assumptions are considered for the interdependencies from the power network to the gas network. For more details on the operational model

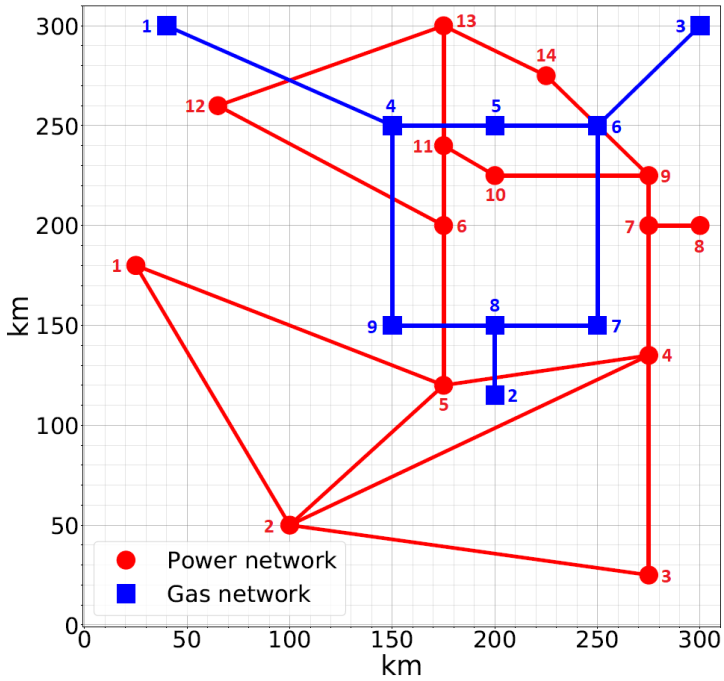

Figure 3.1: Geographical representation of the IPGNs. Figure from Paper IV [79].

and the related equations and constraints, the reader is referred to the appended papers (e.g. Section 2 of Paper III).

3.4 . Coupling interface topology - toy model

Before presenting the mathematical programming framework, it is useful to introduce an illustrative example of the coupling interface design problem and its impact on the failure propagation between interdependent CIs.

Let us have two interdependent power and gas networks, both consisting of two nodes connected by one edge, as shown in Figure 3.2. In the power network, the electricity is produced by a gas-fired power plant in node 1, and general customers (industries, households, etc.) are supplied by node 2. Similarly, in the gas network, the gas is produced (extracted) in node 1, and general customers are supplied by node 2. The gas-fired power plant in node $1 \in V_{PN}$ receives the necessary gas supply from node $2 \in V_{GN}$, while nodes $1 \in V_{GN}$ and $2 \in V_{GN}$ receive the necessary electricity from nodes $2 \in V_{PN}$ and $1 \in V_{PN}$, respectively. These relationships of interdependency define the coupling interface of the IPGNs, and they are represented by the interdependency links (dashed lines) in Figure 3.2.

Let us assume that a failure occurred in the power line connecting nodes $1 \in V_{PN}$ and $2 \in V_{PN}$. Consequently, the electricity produced in node $1 \in V_{PN}$

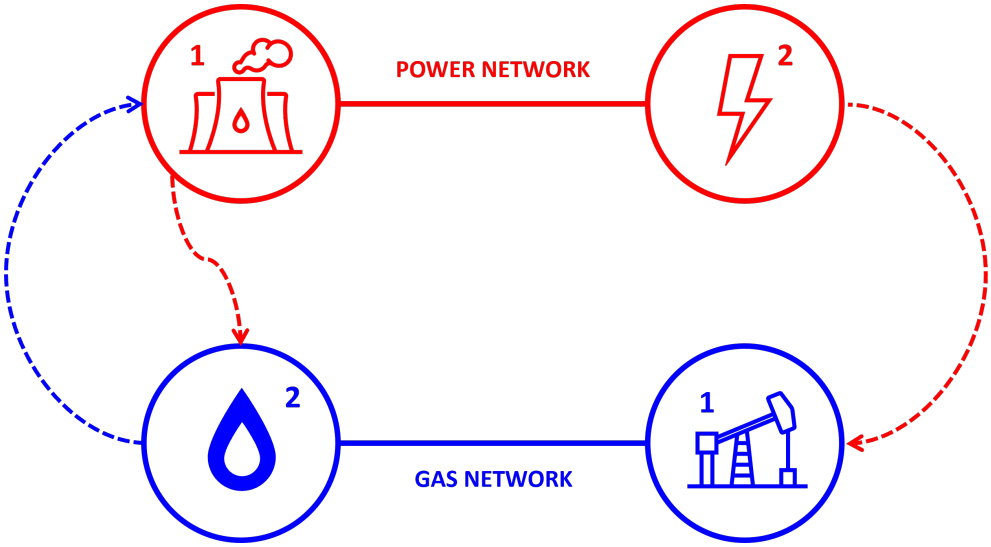

Figure 3.2: Illustrative IPGNs. The dashed lines represent the interdependency links, i.e. the coupling interface of the IPGNs.

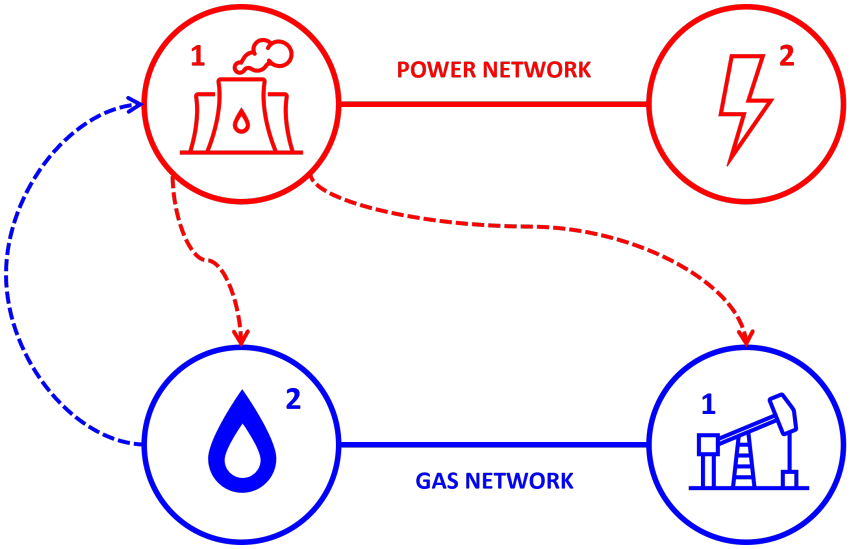

Figure 3.3: Illustrative IPGNs with modified coupling interface.

can not be delivered to node $2 \in V_{PN}$, and the interdependency link from node $2 \in V_{PN}$ to node $1 \in V_{GN}$ can be considered as nonfunctional, as no electricity is supplied to node $2 \in V_{PN}$. With no functional interdependency links from the power network to node $1 \in V_{GN}$, i.e. no power supply to node $1 \in V_{GN}$, the gas network can not produce gas, as the electrical equipment, such as pumps, valves, and compressors, is not provided with electricity. With no supply of gas, the interdependency link from node $2 \in V_{GN}$ to node $1 \in V_{PN}$ is considered nonfunctional. Consequently, it is not possible to produce electricity in the gas-fired power plant in node $1 \in V_{PN}$. The failure of the power line results, thus, in the total disruption of the IPGNs, as no power and gas is produced nor supplied.

Let us modify the coupling interface topology, by making node $1 \in V_{GN}$ coupled with, and dependent on, node $1 \in V_{PN}$, as shown in Figure 3.3. In this case, the failure of the power line would result only in partial disruption of the IPGNs. The general consumers of electricity can not be supplied, as no electricity is delivered to node $2 \in V_{PN}$ due to the power line failure. However, as in this case node $1 \in V_{GN}$ is dependent on node $1 \in V_{PN}$, the production of gas is not disrupted, and the general gas consumers, as well as the gas-fired power plant in the power network, are normally supplied by node $2 \in V_{GN}$.

As it can be clearly understood in this qualitative example, the topology of the coupling interface, which defines the specific location of interdependency links, is a driving factor of failures and disruption propagation in interdependent CIs, and its optimization is of the utmost importance.

3.5 . Robust coupling interface

3.5.1 . Defender-attacker-defender formulation

The approach taken in Paper III consists in identifying the coupling interface topology that maximizes the combined performance of the IPGNs in the worst-case feasible failure scenario. Using the resilience framework defined in Chapter 1, the problem is equivalent to identifying the coupling interface topology that maximizes the combined survivability of the IPGNs in the worst-case feasible failure scenario. The proposed approach is based on the *defender-attacker-defender* (DAD) model, which is often applied in the context of protection and resilience enhancement of CIs [55], [69], [103]–[108]. The DAD approach is a three-players sequential game: the outer defender represents planners and decision-makers, and aims at identifying the optimal planning decisions or resource allocation in order to maximize the objective function; the middle attacker represents various sources of failures and disruption, and aims at identifying the most disruptive attack plan in order to minimize the objective function; the inner defender represents the various CIs operators, and aims at maximizing the objective function by controlling the operational variables [104].

In this case, the outer defender aims at identifying the optimal coupling interface topology, in order to maximize the combined performance of the IPGNs against the optimal attack plan of the middle attacker, which aims at minimizing the combined performance of the IPGNs. The inner defender represents the operators, and it is defined by the operational model $\mathcal{Q}(\mathbf{y}, \boldsymbol{\xi})$. The problem takes the form in (3.6)-(3.12):

$$\max_{\substack{\mathbf{h}', \boldsymbol{\delta}' \in \{0,1\}^{N_d} \\ \mathbf{y} \in \{0,1\}^{N_c}}} \min_{\mathbf{u} \in \{0,1\}^{M_{PN}}} \mathcal{Q}(\mathbf{y}, \boldsymbol{\xi}) \tag{3.6}$$

subject to:

$$\sum_{j \in V_{PN}} y_{ij}^{g \leftarrow p} \leq 1, \quad \forall i \in V_{GN} \tag{3.7}$$

$$\sum_{j \in V_{GN}} y_{ij}^{p \leftarrow g} \leq 1, \quad \forall i \in V_{PN} \tag{3.8}$$

$$\sum_{\substack{i \in V_{GN} \\ j \in V_{PN}}} y_{ij}^{g \leftarrow p} \, d_{ij}^{km} \, c_{km}^{g \leftarrow p} + \sum_{\substack{i \in V_{PN} \\ j \in V_{GN}}} y_{ij}^{p \leftarrow g} \, d_{ji}^{km} \, c_{km}^{p \leftarrow g} \leq B_{ci} \tag{3.9}$$

$$\mathbf{R}\mathbf{h}' \leq \mathbf{q} - \mathbf{H}\mathbf{y} - \mathbf{W}\boldsymbol{\delta}' - \mathbf{y}^T \mathbf{D} \boldsymbol{\delta}' \tag{3.10}$$

$$\mathbf{b}^T \mathbf{h}' \geq 1 \tag{3.11}$$

$$\sum_{k \in E_{PN}} (1 - u_k) \leq K_{att}. \tag{3.12}$$

The problem presents a trilevel max-min-max formulation, where the innermost maximization is contained within the recourse function $\mathcal{Q}(\mathbf{y}, \boldsymbol{\xi})$. The outer defender aims at maximizing the combined performance of the IPGNs, represented by the objective function of the recourse function $\mathcal{Q}(\mathbf{y}, \boldsymbol{\xi})$, by allocating the coupling interface through the binary variables $y_{ij}^{g \leftarrow p}$ and $y_{ij}^{p \leftarrow g}$, contained within the vector $\mathbf{y}$ with dimension $N_c = N_{PN}^2 \times N_{GN}^2$. The binary variable $y_{ij}^{g \leftarrow p}$=1 if node $i \in V_{GN}$ is dependent on node $j \in V_{PN}$, and $y_{ij}^{g \leftarrow p}$=0 otherwise. Similarly, the binary variable $y_{ij}^{p \leftarrow g}$=1 if node $i \in V_{PN}$ is dependent on node $j \in V_{GN}$, and $y_{ij}^{p \leftarrow g}$=0 otherwise. In other words, when a y variable is equal to 1, it denotes the presence of an interdependency link between two nodes of the power and gas network. The middle attacker aims at minimizing the combined performance of the IPGNs by targeting and failing power lines of the IPGNs through the binary variables u_k, contained within the vector $\mathbf{u}$ with dimension M_{PN}. The binary variable u_k=1 when line k is functional, and u_k=0 when line k is targeted and failed by the attacker. The inner defender, represented by the recourse function $\mathcal{Q}(\mathbf{y}, \boldsymbol{\xi})$, aims at maximizing the combined performance by controlling the operational variables of the IPGNs (variables of the operational model in Section 3.2). The recourse function, as it was previously explained, depends on the coupling interface placed by the outer defender through the binary variables $\mathbf{y}$ and the uncertain functional

state of the systems $\boldsymbol{\xi}$, defined by the components targeted and failed by the middle attacker through the binary variables $\mathbf{u}$.

The optimal objective value corresponds to the maximized combined performance of the IPGNs, expressed as in Equation (3.5), in the worst-case feasible failure scenario. The coupling interface topology which maximizes the combined performance of the IPGNs in the worst-case failure scenario is obtained as a byproduct of the optimization model.

This model can be interpreted as a robust optimization problem. In fact, the attacker can target and fail a finite combinations of components. All the feasible combinations of failed components are contained within the set of feasible failure scenarios, which can be interpreted as a discrete uncertainty set. Accordingly, the outer defender seeks to optimize decisions against the worst-case scenario within the uncertainty set, i.e. the worst-case failure scenario.

The actions of both the outer defender and the attacker are regulated by some assumptions, defined by Constraints (3.7)-(3.11) for the outer defender and Constraint (3.12) for the attacker. In these equations, the coefficient d_{ij}^{km} denotes the distance in kilometer between two nodes, while the coefficients $c_{km}^{g \leftarrow p}$ and $c_{km}^{p \leftarrow g}$ denote the price per kilometer of allocating interdependency links, here both assumed to be equal to 1 \$/km. The vectors $\mathbf{h}'$ defines the continuous operational variables of the outer defender, while the vector $\boldsymbol{\delta}'$ defines the binary operational variables of the outer defender. The other vectors and matrices in Constraints (3.10) and (3.11) represent coefficients and parameters of the operational constraints of the outer defender.

The actions of the outer defender are regulated by the following assumptions:

- each node in the gas network can be dependent on a maximum of one node of the power network, and vice versa. This is defined as "single-dependency" assumption and is enforced by Constraints (3.7) and (3.8). In other words, each node in the gas network can be the receiving end of only one interdependency link from the power network, and vice versa;
- the allocation of interdependency links has a cost per kilometer, and it is limited by the available monetary budget B_{ci}, as enforced by Constraint (3.9);
- the coupling interface must be allocated in order to ensure that, under normal conditions, i.e. all the components are functional, the total requested demand of power and gas can be fully satisfied. This condition is enforced by Constraints (3.10) and (3.11) (see Section 2 of Paper III).

The actions of the attacker, similarly to other works [69], are limited by the maximum number of lines K_{att} that can be targeted and failed, as enforced by Constraint (3.12). The attacker is assumed to target only power transmission lines. With these assumptions, the attacker can choose from a limited number of attack plans, contained within the set of feasible failure scenarios $\mathcal{A}$, defined as in (3.13):

$$\mathcal{A} = \left\{ \mathbf{u} \,\middle|\, \{0,1\}^{M_{PN}}, ||\mathbb{1}^{M_{PN}} - \mathbf{u}||_1 \leq K_{att} \right\}. \tag{3.13}$$

As it is highlighted in Paper III, *"by considering the simultaneous failures of transmission lines, the model is agnostic about the source of disruption, providing a rapid and objective way of calculating the consequence of damage to any set of components"* [10].

These assumptions are used to establish the general framework, and they can be easily adapted to describe different situations. For example, the possibility for the attacker to target other components, such as nodes or interdependency links, can be easily included with additional constraints similar to Constraint (3.12) and by modifying the relative operational constraints within the inner defender recourse function $\mathcal{Q}(\mathbf{y}, \boldsymbol{\xi})$.

3.5.2 . Solution strategy

When the recourse function is extended in its explicit form, previously shown in Equations (3.2)-(3.4), the problem presents a trilevel max-min-max formulation. As the inner maximization includes binary variables, a strategy based on the dualization of the inner problem, in order to merge it with the middle minimization, is not suitable. In this work, a cutting plane strategy, called *Nested Column&Constraint Generation* (NC&CG) algorithm, is applied [109], [110]. With this approach, the original problem is extended into a four-level max-min-max-max problem by separating the binary and continuous variables of the inner level. The problem is then separated into an outer layer and an inner layer. Each layer is composed of a master problem and a subproblem. These problems and layers are solved separately and iteratively, by exchanging primal binary variables between them. The general flowchart of the NC&CG algorithm is shown in Figure 3.4. This procedure is proven to converge to the global optimal solution, and it has already been applied in similar problems [34], [104]. For more details, the reader is referred to Appendix A of this manuscript and Section 3 of Paper III.

3.5.3 . Results

The results are evaluated in terms of combined performance of the IPGNs in the worst-case feasible failure scenario. For the sake of comparison, the results obtained by the proposed DAD approach are compared with the results obtained by designing the coupling interface with network metrics-based heuristic strategies. Five network metrics-based strategies, based on Euclidean distance, node

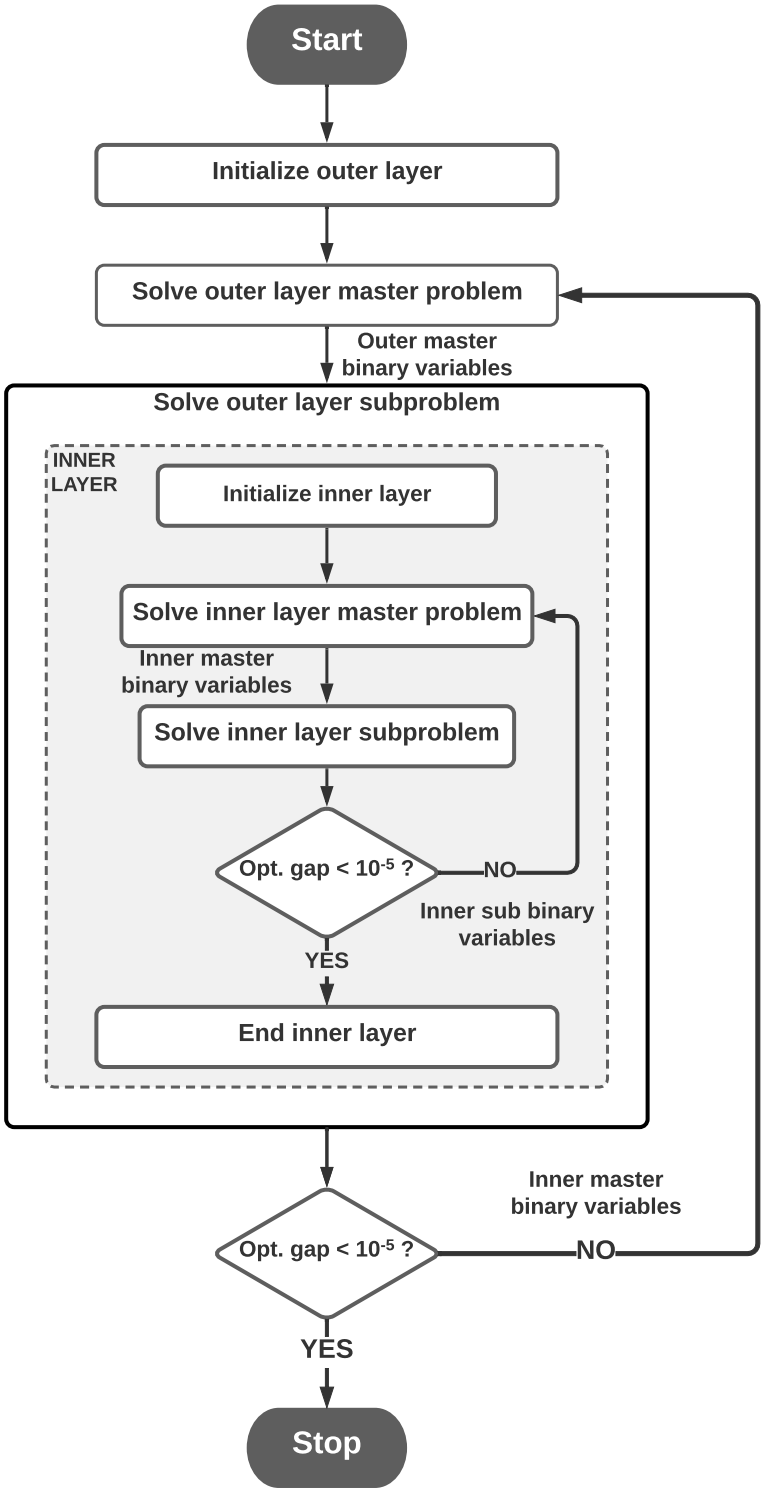

Figure 3.4: General flowchart of the Nested Column&Constraint Generation algorithm. Picture based on Paper III, Paper IV, and Paper V [10], [79], [80].

degree, and betweenness centrality, are used. The description of the five strategies contained in Paper III is here recalled [10]:

- *"$Euclidean$: each node in the power network (or gas network) is dependent on the geographically closest node in the gas network (or power network);*
- *DD_{ast}: the node with the k^{th} highest degree in the power network (or gas network) is dependent on the node with the k^{th} highest degree in the gas network (or power network);*
- *DD_{dst}: the node with the k^{th} highest degree in the power network (or gas network) is dependent on the node with the k^{th} lowest degree in the gas network (or power network);*
- *BB_{ast}:the node with the k^{th} highest betweenness in the power network (or gas network) is dependent on the node with the k^{th} highest betweenness in the gas network (or power network);*
- *BB_{dst}: the node with the k^{th} highest betweenness in the power network (or gas network) is dependent on the node with the k^{th} lowest betweenness in the gas network (or power network)."*

The results of the network metrics-based coupling interfaces and optimal coupling interfaces identified with the proposed DAD approach are shown in Figures 3.5 and 3.6, respectively. The x-axis indicates different values of K_{att}, from 1 to 5, which denotes the maximum number of power lines that can be targeted and failed by the attacker. The y-axis denotes the worst-case combined performance of the IPGNs, i.e. the combined fraction of requested demand of power and gas which is possible to supply in the worst-case scenario within the set of feasible failure scenarios $\mathcal{A}$.

In Figure 3.5, it can be clearly noticed how different coupling interfaces lead to considerably different worst-case combined performances. Particularly, the Euclidean coupling strategy, based on node geographical proximity, seems to outperform the other strategies. However, these heuristic strategies are strongly case-dependent, and for a different case study the best-performing coupling strategy might differ. In Table 3.1, the cost associated with each network metrics-based coupling strategy is computed assuming costs per kilometer $c_{km}^{g \leftarrow p}$ and $c_{km}^{p \leftarrow g}$ equal to 1 \$/km.

The results of the optimal coupling strategies, identified with the proposed approach for different available monetary budgets B_{ci}, are shown in Figure 3.6. The monetary budgets range from \$823, equivalent to the cost of Euclidean coupling

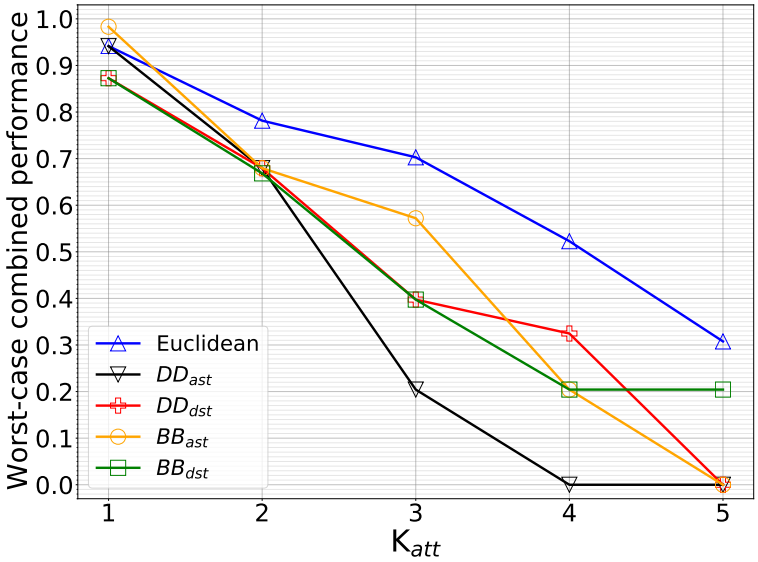

Figure 3.5: Combined performance in the worst-case feasible failure scenario with network metrics-based coupling strategies. Figure from Paper III [10]. It should be highlighted that, since multiple Euclidean coupling interfaces exist, for each value K_{att}, the plot represents the best performing Euclidean coupling interface.

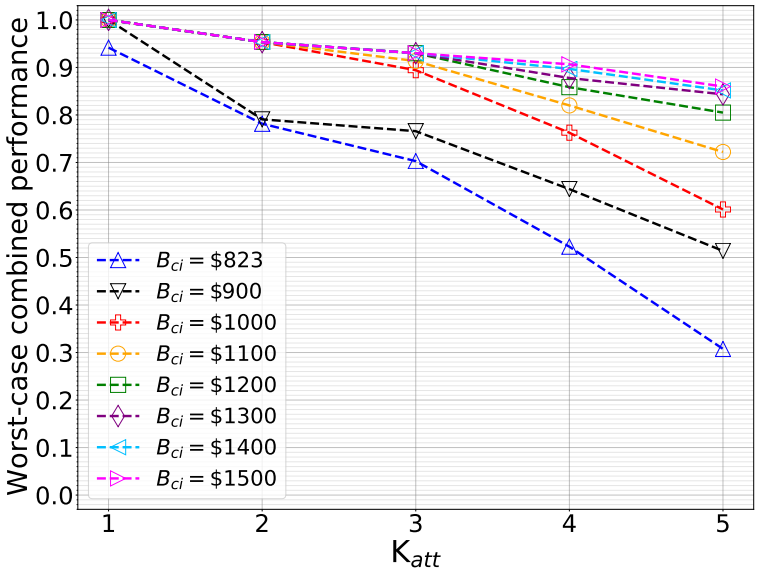

Figure 3.6: Combined performance in the worst-case feasible failure scenario with optimal coupling strategies. Figure from Paper III [10].

Table 3.1: Cost of network metrics-based coupling interfaces.

Strategy	Cost
$Euclidean$	$823
DD_{ast}	$1518
DD_{dst}	$2098
BB_{ast}	$1943
BB_{dst}	$2126

interfaces (see Table 3.1), to $1500.

Firstly, it can be noticed that the results for the Euclidean coupling interface in Figure 3.5 (blue solid line) and the results with a budget B_{ci}=$823 in Figure 3.6 (blue dashed line), are equivalent. This is because, with a budget equal to $823, the Euclidean coupling interfaces are the only ones which is possible to allocate[6]. With lower budgets, the problem is unfeasible, as it is not possible to respect Constraint (3.11).

Secondly, it can be clearly seen that the network metrics-based coupling interfaces are outperformed in terms of worst-case combined performance. For example, for the case K_{att}=5, the Euclidean coupling interface, the best-performing in Figure 3.5, leads to worst-case combined performance equal to 0.308, while an optimal coupling interface with a budget B_{ci}=$900 leads to worst-case combined performance equal to 0.514. This corresponds to a 66.9% increase in combined performance for a 9.3% increase in budget.

3.5.4 . Remarks

This work represents the first mathematical programming approach for optimizing the topology and the design of coupling interfaces between interdependent CIs.

The main contributions of this work are:

- demonstrating the importance of coupling interface optimization for increasing the survivability of interdependent CIs under conditions of disruption;
- proposing an optimization-based approach that can be used by decision-makers for analyzing, designing, or retrofitting existing and new coupling interfaces.

[6]The budget B_{ci}=$823 corresponds to the cost of Euclidean coupling interfaces in Table 3.1.

It should be mentioned that in Paper III, the proposed optimization framework also accounts for the cost of the coupling interface. In fact, for each combination of K_{att} and B_{ci}, multiple optimal coupling interfaces might exist, and the proposed approach identifies the cheapest optimal coupling interface thanks to a cost penalty term. In other words, among the set of optimal coupling interfaces that maximize the combined performance, the proposed DAD approach selects the cheapest solution. For the sake of simplicity and clarity, this part is omitted from this dissertation. The reader is referred to Paper III for more details.

The approach based on the DAD model suffers from some limitations. Particularly, similarly to other works developed with a DAD approach (or, in general, a robust optimization approach), the solutions are optimized against the worst-case scenario, and this might lead to over-conservative solutions. Despite the decision-makers having control over the size of the set of feasible failure scenarios through the parameter K_{att}, it might not suffice to avoid over-conservative solutions. In Paper IV, an alternative approach, based on distributionally robust optimization, is proposed, in order to provide decision-makers with more control over the conservativeness of the model. In this work, the occurrence probability of each failure scenario in the set $\mathcal{A}$ is taken into account.

3.6 . Distributionally robust coupling interface

3.6.1 . Distributionally robust approach

When dealing with optimization under uncertainty, the two most traditional approaches are *robust optimization* (RO) and *stochastic optimization* (SO).

In RO, the decision-maker seeks to optimize decisions against the worst-case realization of the uncertainty. For example, in the previously explained DAD approach, assuming that the uncertainty set $\mathcal{A}$ denotes the discrete set of feasible failure scenarios, the decision-makers seek to optimize the coupling interface against the worst-case failure scenario in $\mathcal{A}$. The RO framework, as the decision-makers seek to optimize against the worst uncertainty realization, is considered a risk-averse approach

In SO, decision-makers are assumed to possess full probabilistic knowledge of the uncertainty distribution, and to have a neutral attitude towards risk[7]. Decision-makers seek, thus, to identify decisions that optimize the expectation of the objective function considering the whole spectrum of uncertainty realization. In the case of coupling interface design, this situation would translate into identifying a coupling interface topology that optimizes the expected combined performance of the

[7]Risk-averse stochastic approaches also exist. In this case, a probabilistic risk measure, like Conditional Value-at-Risk (CVaR), is often optimized [111].

IPGNs under the conditions of disruption defined by the set $\mathcal{A}$. As the uncertainty set $\mathcal{A}$ is discrete, it can be described by a multinomial probability distribution. The expected performance of the IPGNs are simply computed as the weighted sum of the combined performance in each individual scenario within the set $\mathcal{A}$, where the weight of each scenario corresponds to its probability mass within the multinomial distribution. This approach relies on the assumption that the multinomial probability distribution of the set $\mathcal{A}$ of feasible failure scenarios is perfectly known.

In practice, this assumption is considerably difficult, if not impossible, to guarantee, due to various reasons:

- sparsity of data: detailed historical data on contingencies and failures are rarely available. Even when available, the quantity of events does not usually allow to estimate a probability distribution with high confidence;
- environment: the environment variability and external factors (weather, climate changes, geopolitical factors, etc.) introduce large uncertainties which make it difficult to estimate failure scenario probabilities accurately;
- unpredictability: it is often impossible to estimate the likelihood of intentional attacks, such as terroristic attacks or deliberate sabotages, against one or more CIs.

Thus, in this context, this approach is impractical and it does not represent a feasible modeling choice for the optimization of coupling interface topologies.

While it might be not feasible to identify the true multinomial probability distribution of the set $\mathcal{A}$, it is often possible to identify partial probabilistic information, such as moment information, from historical data and/or expert judgement. This available information can be exploited in a robust stochastic approach, called *distributionally robust optimization* (DRO), that protects decision-makers against the ambiguity of the probability distribution [112]. With this approach, decision-makers seek to optimize decisions against the worst probability distribution that can describe the uncertainty set. In practice, the available probabilistic information is encoded within the optimization model. The decision-makers aim at optimizing the expectation of the objective function under the worst probability distribution (which becomes a variable of the problem) that can describe the uncertainty set while respecting the encoded probabilistic information. In terms of risk-aversion, DRO lies in-between RO and SO, and it can be considered as their generalization.

The DRO approach has gained the interest of several research fields in the recent years. For example, the reader is referred to [113]–[117] for applications in various fields.

Probabilistic moments are particularly convenient information to encode within DRO approaches, as they are usually easy to estimate. In Paper IV, a DRO approach based on the upper bounds of the conditional marginal probability of each line to be in a failure state is proposed [118]. In other words, given the condition that one of the scenarios within $\mathcal{A}$ has occurred (except the "normal" scenario, where all the lines are functional), i.e. at least one line has failed, this conditional marginal probability defines the likelihood of each line k to be in a failure state (u_k=0). Accordingly, decision-makers seek to identify the coupling interface topology that maximize the expected combined performance of the IPGNs under the worst probability distribution of the set $\mathcal{A}$ that respects the encoded upper bounds on the conditional marginal probability of each line to be in a failure state.

3.6.2 . Distributionally robust optimization - toy model

The purpose of DRO is to protect decision-makers against ambiguous probability distributions. This is achieved by optimizing the expectation of an objective function accounting for the *worst probability distribution* that respects the enforced probabilistic conditions. Before presenting the DRO formulation for the optimal coupling of interdependent CIs, it is useful to introduce an illustrative example to clarify the meaning of *worst probability distribution* with upper bounds on the conditional marginal probability of each line to be in a failure state.

Let us have a power network, composed of three buses (nodes) connected by two power lines (edges), as shown in Figure 3.7. Node 1 contains a power plant with a very large production capacity $\overline{p}_1$ (enough to supply the whole network). Nodes 2 and 3 both contain a load with a requested power demand $\overline{d}_2$=$\overline{d}_3$=1 MW, for a total requested power demand $\overline{d}_{PN}$=2 MW. The two power lines are characterized by a very large flow capacity, greater than 2 MW. Under normal conditions, 2 MW are produced within node 1 and supplied to nodes 2 and 3 (1 MW each) through the power lines.

In this work, the purpose of the DRO approach is to identify a solution (the design of the coupling interface topology) that maximizes the expected performance in the worst probability distribution of the set $\mathcal{A}$ of feasible failure scenarios. The toy model in Figure 3.7, solely based on a power network, is useful to understand the concept of worst probability distribution with enforced moment information.

Assuming only failures of power lines and a value K_{att}=2, the set $\mathcal{A}$ of this toy model consists of four different scenarios $\mathbf{u}_k$, as shown in Table 3.2. The corresponding performance P_k, in terms of fraction of total requested power demand $\overline{d}_{PN}$ which is possible to supply for each scenario $\mathbf{u}_k$ are also shown in Table 3.2. In the scenario with no failures $\mathbf{u}_\emptyset$, the performance $P_\emptyset$ is equal to 1.0, as it is possible to supply 100% of the total requested power demand $\overline{d}_{PN}$. In the scenarios

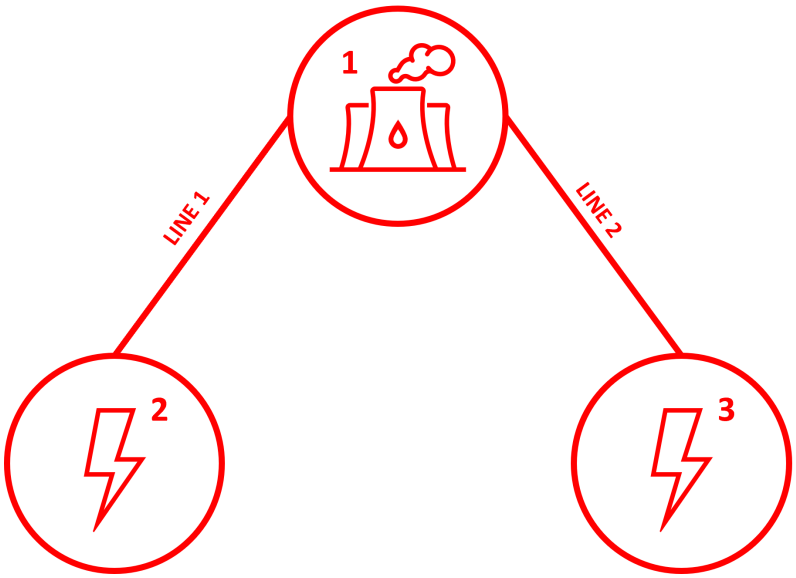

Figure 3.7: Toy model of a power network.

with one line failed, $\mathbf{u}_1$ and $\mathbf{u}_2$, the performances P_1 and P_2 are equal to 0.5, as it possible to supply either node 2 or node 3. In the scenario with two lines failed $\mathbf{u}_{1,2}$, the performance $P_{1,2}$ is equal to 0.0, as it is not possible to supply any power to nodes 2 and 3.

Table 3.2: Scenarios within the set $\mathcal{A}$ for the power network in Figure 3.7 and corresponding performance.

Scenario	**Vector $\mathbf{u_k}$**	**Performance P_k**
No failures	$\mathbf{u}_\emptyset$=[1,1]	1.0
Line 1 failed	$\mathbf{u}_1$=[0,1]	0.5
Line 2 failed	$\mathbf{u}_2$=[1,0]	0.5
Line 1 and 2 failed	$\mathbf{u}_{1,2}$=[0,0]	0.0

As the set $\mathcal{A}$ is a finite set with discrete scenarios, it can be described by a multinomial distribution $\mathbb{P}$, in which each scenario $\mathbf{u}_k$ is assigned a probability mass ϕ_k. The assigned probability masses must respect the condition in (3.14):

$$\sum_{k \in \mathcal{A}} \phi_k = 1. \tag{3.14}$$

Moreover, some probabilistic moment information are enforced to the probability distribution of $\mathcal{A}$. In this work, an upper bound π_k^{max} on the conditional marginal probability of each line k to be in a failure state[8], defined as π_k, is en-

[8]The word *conditional* refers to the condition that at least one line is in a failure state.

forced. The reason of this choice are detailed in the next section. The marginal probability π_k for line 1 and 2 are defined as in (3.15) and (3.16), respectively:

$$\pi_1 = \phi_1 + \phi_{1,2} \tag{3.15}$$

$$\pi_2 = \phi_2 + \phi_{1,2} \tag{3.16}$$

In this specific example, the distribution $\mathbb{P}$ should respect the conditions in (3.17)-(3.19):

$$\phi_\emptyset = 0 \tag{3.17}$$

$$\phi_1 + \phi_{1,2} \leq \pi_1^{max} \tag{3.18}$$

$$\phi_2 + \phi_{1,2} \leq \pi_2^{max}. \tag{3.19}$$

Equation (3.17) states that the probability mass of the normal scenario $\mathbf{u}_\emptyset$ should be 0. In other words, $\mathbb{P}$ defines a conditional distribution, where the condition is that at least one line is in a failure state. The reason of this choice is explained in the next section. Equations (3.18) and (3.19) define the upper bounds on the conditional marginal probability of each line k to be in a failure state.

Under these assumptions, the set $\mathcal{A}$ can be described by any multinomial distribution that respects the conditions in (3.17)-(3.19). The *worst probability distribution* $\widehat{\mathbb{P}}$ defines the probability distribution that respects the aforementioned conditions and leads to the lowest expected performance in the power network. Mathematically, it can be identified by solving Equation (3.20) subject to (3.14) and (3.17)-(3.19):

$$\min_{\boldsymbol{\phi} \geq 0} \sum_{k \in \mathcal{A}} \phi_k P_k. \tag{3.20}$$

For example, assuming upper bounds π_1^{max}=π_2^{max}=0.8, it can be easily verified that the worst distribution $\widehat{\mathbb{P}}$ is the one defined in (3.21), which leads to expected performance equal to 0.20.

$$\widehat{\mathbb{P}} = \begin{cases} \phi_\emptyset = 0.0 \\ \phi_1 = 0.2 \\ \phi_2 = 0.2 \\ \phi_{1,2} = 0.6 \end{cases} \tag{3.21}$$

Any other distribution would either lead to higher expected performance or fail to meet the conditions enforced in (3.14) and (3.17)-(3.19). The purpose of DRO is, thus, to identify solutions that are robust against the worst probability distribution that can characterize the set of feasible failure scenarios.

3.6.3 . Distributionally robust formulation

In Paper IV, the DRO framework for the optimal coupling of interdependent CIs is developed using IPGNs as case study. The operational model is represented by the recourse function $\mathcal{Q}(\mathbf{y}, \boldsymbol{\xi})$, which defines the maximization of the combined

performance of the IPGNs under the operational constraints. The decision-makers allocate the coupling interface in order to maximize the expected combined performance of the IPGNs under conditions of disruption, and its actions are limited by the same assumptions of Section 3.4.1, defined by Constraints (3.7)-(3.11). The middle minimization still represents the disruptive agent. However, instead of identifying the worst-case feasible failure scenario in $\mathcal{A}$, this agent aims at identifying the worst multinomial distribution which can describe the uncertainty set $\mathcal{A}$, while respecting the encoded probabilistic information.

The DRO approach for optimizing the coupling interface topology can be defined as Equation (3.22):

$$\max_{\substack{\mathbf{h}',\boldsymbol{\delta}'\in\{0,1\}^{N_d}\\ \mathbf{y}\in\{0,1\}^{N_c}}} \min_{\mathbb{P}\in\mathcal{M}} \mathbb{E}_{\mathbb{P}}\left[\mathcal{Q}(\mathbf{y},\boldsymbol{\xi})\right] \tag{3.22}$$

subject to (3.7)-(3.11) and (3.23):

$$\mathcal{M} = \left\{\mathbb{P}\in\mathcal{P}(\mathcal{A}) : 0 \leq \mathbb{E}_{\mathbb{P}}[\mathbb{1}^{M_{PN}} - \mathbf{u}] \leq \boldsymbol{\pi}^{max}\right\}. \tag{3.23}$$

The set $\mathcal{M}$, called the *ambiguity set*, contains all probability distributions on a σ-field of $\mathcal{A}$ that ensure that the conditional marginal probability of each line k to be in a failure state is lower or equal to the upper bound π_k^{max}. The upper bounds are contained within the vector $\boldsymbol{\pi}^{max}$. The term $\mathbb{1}^{M_{PN}}$ defines an M_{PN}-dimensional vector of 1s. As the middle agent in (3.22), rather than selecting the worst-case scenario in $\mathcal{A}$, aims at identifying the worst distribution in $\mathcal{M}$, Equation (3.23) replaces Constraint (3.12) of the DAD model.

The motivation for choosing an ambiguity set $\mathcal{M}$ based on the conditional marginal probability of each line k to be in a failure state is threefold:

- the *conditional* marginal probability, where the term *conditional* refers to the condition that at least one is failed, allows to optimize the expected combined performance in situations of disruption. In fact, the combined performance under normal conditions is already ensured by Constraints (3.10) and (3.11);
- this marginal probability provides a meaningful physical interpretation of the problem, as it represents the "tendency" of each line to be in a failure state;
- moment information is usually easy to estimate and encode within an optimization problem.

In summary, the ambiguity set $\mathcal{M}$ denotes the set of conditional multinomial distributions which define the probability of the system to be in a specific failure scenario, under the condition that at least one line is failed. In order to guarantee

that $\mathcal{M}$ defines the set of *conditional* multinomial distributions, i.e. the probability mass assigned to the normal scenario is equal to 0, the upper bounds $\boldsymbol{\pi^{max}}$ must respect some specific conditions (see Section 3.2 of Paper IV for more details).

In Paper IV, the upper bounds of the conditional marginal probability $\boldsymbol{\pi}^{max}$ are estimated empirically from an artificial dataset of contingency scenarios. The estimated upper bounds $\boldsymbol{\pi}^{max}$ are shown in Table 3.3. For more details, the reader is referred to Section 4.1 of Paper IV.

Table 3.3: Upper bounds π_k^{max} for each power line k.

Line	$\boldsymbol{\pi_k^{max}}$	**Line**	$\boldsymbol{\pi_k^{max}}$
1	0.064	11	0.136
2	0.012	12	0.037
3	0.089	13	0.037
4	0.037	14	0.131
5	0.164	15	0.089
6	0.012	16	0.159
7	0.084	17	0.053
8	0.043	18	0.117
9	0.031	19	0.019
10	0.089	20	0.108

The optimal objective value corresponds to the maximized expected combined performance of the IPGNs under the worst distribution in $\mathcal{M}$. The advantage of a DRO model is to identify coupling interface topologies with a robust approach that avoids the "worst-case" over-conservativeness of the DAD model.

3.6.4 . Solution strategy

The problem can be solved with an NC&CG approach. However, it needs to be recast into an equivalent tractable formulation. Particularly, the problem in (3.22), subject to (3.7)-(3.11) and (3.23), is equivalent to the problem in (3.24) subject to (3.7)-(3.12):

$$\max_{\substack{\mathbf{h}',\boldsymbol{\delta}'\in\{0,1\}^{N_d}\\ \mathbf{y}\in\{0,1\}^{N_c}\\ \boldsymbol{\beta}\geq 0}} \min_{\mathbf{u}\in\{0,1\}^{M_{PN}}} \mathcal{Q}(\mathbf{y},\boldsymbol{\xi}) + \sum_{k\in E_{PN}} \beta_k(1-u_k-\pi_k^{max}). \tag{3.24}$$

As the reformulation involves duality theory, the problem in (3.24) contains the dual variables β_k, contained within the vector $\boldsymbol{\beta}$. In this form, the problem can be solved by an NC&CG algorithm, shown in Figure 3.4. For details on the solution procedure and the derivation of the reformulation, the reader is referred to Section 3 of Paper IV.

3.6.5 . Results

The results are evaluated for a maximum number of failed edges K_{att}=3 and a monetary budget B_{ci}=\$1100, and they are shown in Figure 3.8. The results are evaluated in terms of expected combined performance of the IPGNs under conditions of disruption, as indicated by the y-axis. As indicated by the x-axis, the results are evaluated in terms of ambiguity set size by multiplying the upper bounds $\boldsymbol{\pi}^{max}$ in Table 3.3 by an increasing factor (from 1 to 6). For example, $2\boldsymbol{\pi}^{max}$ indicates that the upper bounds in Equation (3.23) are equal to the bounds in Table 3.3 multiplied by a factor of 2. The results are also evaluated when the upper bounds $\boldsymbol{\pi}^{max}$ are all set equal to 1. Results with a budget B_{ci}=\$823, corresponding to the cost of Euclidean coupling interfaces, are shown for the sake of comparison. Four curves are shown in Figure 3.8: the blue curves represent the worst expected performance associated with the optimal and Euclidean coupling interfaces, computed by the distributionally robust optimization model; the red curves represent the worst-case scenario in $\mathcal{A}$ associated to the optimal and Euclidean coupling interfaces. As it

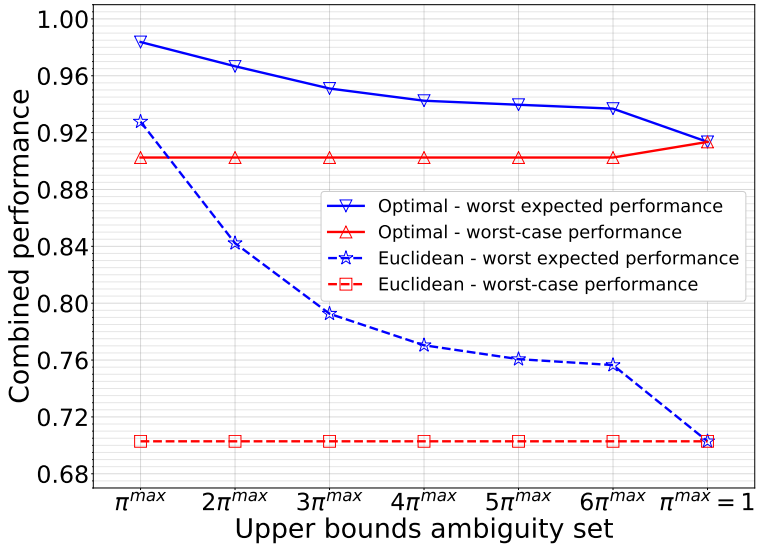

Figure 3.8: Results of the distributionally robust optimal coupling of IPGNs. Revised version of Figure 3 of Paper IV [79].

can be clearly seen in Figure 3.8, the Euclidean coupling interface is outperformed by the optimal coupling interface. Firstly, the worst expected performances are significantly higher in the optimal case. For example, using the upper bounds in Table 3.3, the optimal case leads to worst expected performance equal to 0.984, while the Euclidean case, for the same upper bounds, to worst expected performance equal to 0.927. As it is graphically evident, this difference increases as the upper bounds $\boldsymbol{\pi}^{max}$ increases. For example, using the upper bounds $\boldsymbol{\pi}^{max}$ multiplied by a factor 6, the optimal case leads to worst expected performance equal to 0.937, while the Euclidean case, for the same upper bounds, to worst expected performance equal

to 0.756. Secondly, the associated worst-case performances are also significantly better in the optimal case. For example, using the upper bounds in Table 3.3, the optimal case leads to an associated worst-case performance equal to 0.902, while the Euclidean case, for the same upper bounds, leads to an associated worst-case performance equal to 0.703.

As it was expected, increasing the upper bounds $\boldsymbol{\pi}^{max}$, by multiplying them by factors from 2 to 6, increases the conservativeness of the solution. In fact, as it can be clearly seen, the worst expected performances decrease, while the worst-case performances remain constant or increase.

If the upper bounds $\boldsymbol{\pi}^{max}$ are set to 1, worst expected performance and worst-case performance are equivalent. This corresponds to a situation where the probability mass is entirely allocated to the worst-case scenario. In other words, the worst probability distribution is the one where the worst-case scenario occurs with probability 1. Accordingly, the results of the DRO approach with the upper bounds $\boldsymbol{\pi}^{max}$ set to 1 are equivalent to the results of the DAD model explained in the previous section. This fact is easily verifiable by comparing worst expected combined performance in Figure 3.8 with $\boldsymbol{\pi}^{max}$=1 and the results of the DAD model for K_{att}=3 and B_{ci}=\$1100 in Figure 3.6. In both cases, the worst (expected) combined performances are equivalent to 0.913.

Additional results with a sensitivity analysis of the parameters K_{att} and B_{ci} are available in Paper IV [79].

3.6.6 . Remarks

The DRO approach displays good properties of robustness and control over the conservativeness of the problem. In fact, decision-makers can exploit historical contingency data and expert judgment to inform their choice on the coupling interface topology, and tune their risk attitude using the upper bounds $\boldsymbol{\pi}^{max}$.

In this specific illustrative example, the DRO approach also displays good properties in terms of worst-case scenarios. In fact, as it can be clearly seen in Figure 3.8, the worst-case scenario associated with the identified optimal coupling interface presents considerably high combined performance. However, this is not an intrinsic property of the DRO approach, but it is strictly related to the case study to which it is applied.

In both the DAD and DRO models, the coupling interface is modeled with a "single-dependency" assumption. In fact, consistently with the existing literature (e.g. [34] or [70]), each node of the gas network is dependent on a maximum of one node of the power network, and vice versa. In other words, each node in the IPGNs can be the receiving end of a maximum of one interdependency link. How-

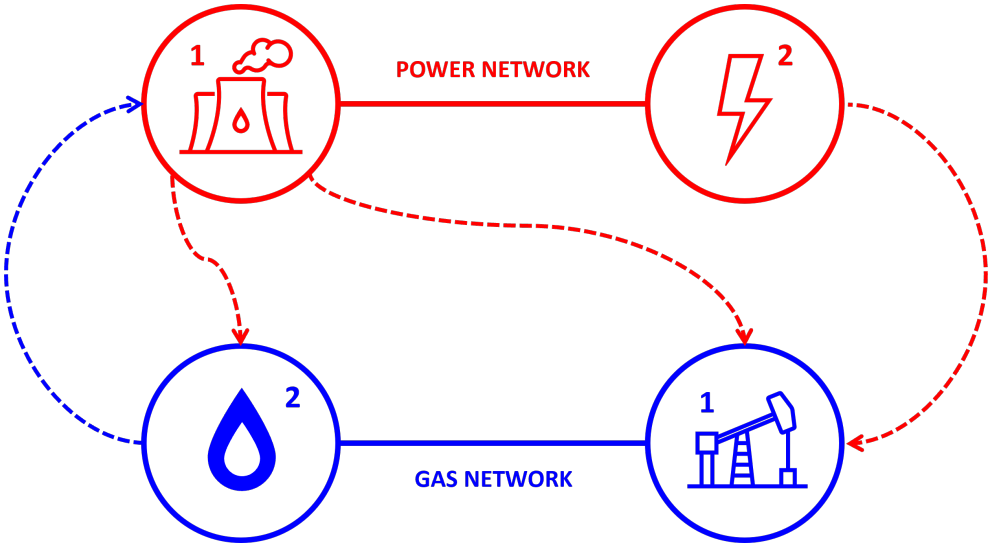

Figure 3.9: Representative IPGNs with redundant interdependency links.

ever, redundant interdependency links exist in real CIs, and decision-makers should take this possibility into account when designing and optimizing coupling interfaces.

In Paper V, the framework based on the DAD model is extended in order to include the possibility of allocating redundant interdependency links between CIs.

3.7 . Allocation of redundant interdependency links

3.7.1 . Redundancy allocation problem

In Paper III and Paper IV, each node in the gas network is assumed to be dependent on a maximum of one node of the power network, and vice versa. This assumption, previously defined as the "single-dependency" assumption, is often applied in the existing literature on interdependent CIs [34], [70], [119]–[123], even though, in reality, redundant interdependencies can be present [31], [32]. In addition, redundant relationships of interdependency are also acknowledged within the network science field [124]. In fact, each node in one CIs can be dependent on multiple nodes of another CIs. For example, in IPGNs, each node in the gas network can be supplied by multiple nodes of the power network, and be the receiving end of multiple interdependency links. These multiple interdependencies can be represented by redundant interdependency links, as shown in Figure 3.9, where node $1 \in V_{GN}$ can be supplied by both nodes $1 \in V_{PN}$ and $2 \in V_{PN}$.

The optimal allocation of redundancies is a traditional reliability problem [125]–[131]. In Paper V, the concept of optimal redundancy allocation is extended to the problem of optimizing coupling interface topologies in interdependent CIs. In

particular, the impact of redundancy allocation on the combined survivability of interdependent CIs under conditions of disruption is investigated. The model is developed with a DAD approach for illustrative purposes; however, a DRO approach can also be easily adopted.

3.7.2 . Defender-attacker-defender formulation

In Paper V, the IPGNs in Figure 3.1 are used as case study. The DAD model for the optimal allocation of redundant interdependency links is developed considering the following assumptions [80]:

- a coupling interface is already present, and it already ensures the necessary performance under normal conditions;
- the cost of allocating a redundant interdependency link depends on the distance between the two nodes;
- the middle attacker can target and fail a maximum number of power lines.

The operational model of the IPGNs is defined by the recourse function $\mathcal{Q}^r(\mathbf{y}, \boldsymbol{\xi})$, where the superscript r denotes that, in this case, the operational model accounts for the presence of redundant interdependency links (see Paper V for more details). In particular, it is assumed that each node in the gas network can receive multiple interdependency links from the power network, but only one is assumed to actively supply the electricity. The same assumption is considered for the power network and the interdependency links from the gas network.

The DAD approach for the optimal allocation of redundant interdependency links takes the form of Equation (3.25) subject to Equations (3.26)-(3.29):

$$\max_{\mathbf{y}\in\{0,1\}^{N_c}} \min_{\mathbf{u}\in\{0,1\}^{M_{PN}}} \mathcal{Q}^r(\mathbf{y}, \boldsymbol{\xi}) \tag{3.25}$$

$$y_{ij}^{g\leftarrow p} \geq \overline{y}_{ij}^{g\leftarrow p}, \quad \forall i \in V_{GN}, \forall j \in V_{PN} \tag{3.26}$$

$$y_{ij}^{p\leftarrow g} \geq \overline{y}_{ij}^{p\leftarrow g}, \quad \forall i \in V_{PN}, \forall j \in V_{GN} \tag{3.27}$$

$$\sum_{\substack{i\in V_{GN}\\ j\in V_{PN}}} y_{ij}^{g\leftarrow p}\, d_{ij}^{km}\, c_{km}^{g\leftarrow p} + \sum_{\substack{i\in V_{PN}\\ j\in V_{GN}}} y_{ij}^{p\leftarrow g}\, d_{ji}^{km}\, c_{km}^{p\leftarrow g} \leq \overline{c}_{ci} + B_{ci} \tag{3.28}$$

$$\sum_{k\in E_{PN}} (1 - u_k) \leq K_{att}. \tag{3.29}$$

Constraints (3.26) and (3.27) ensure that, if an interdependency link is already present, the corresponding binary variable is equal to 1. The presence of a pre-existing interdependency link is defined by the binary parameters $\overline{y}_{ij}^{g\leftarrow p}$ and $\overline{y}_{ij}^{p\leftarrow g}$, that are equal to 1 if the corresponding link is already existing, and equal to 0 otherwise. As enforced by Constraint (3.28), the cost of allocating the redundant

interdependency links should not exceed the budget B_{ci}. The cost of the existing coupling interface is given by $\overline{c}_{ci}$, computed as in (3.30):

$$\overline{c}_{ci} = \sum_{\substack{i \in V_{GN} \\ j \in V_{PN}}} \overline{y}_{ij}^{g \leftarrow p} \, d_{ij}^{km} \, c_{km}^{g \leftarrow p} + \sum_{\substack{i \in V_{PN} \\ j \in V_{GN}}} \overline{y}_{ij}^{p \leftarrow g} \, d_{ji}^{km} \, c_{km}^{p \leftarrow g}. \tag{3.30}$$

Including the term $\overline{c}_{ci}$ in (3.28) allows considering only the newly allocated interdependency links within the budget limitation. Constraint (3.29) controls the maximum number of power lines that can be targeted and failed by the attacker.

3.7.3 . Results

The model is solved with a NC&CG algorithm and it is applied to the IPGNs in Figure 3.1. It is assumed that a Euclidean coupling interface is already present within the IPGNs. The model is tested for values K_{att} from 1 to 5, and monetary budgets B_{ci} from \$100 to \$400. The results are shown in Figure 3.10 and Table 3.4.

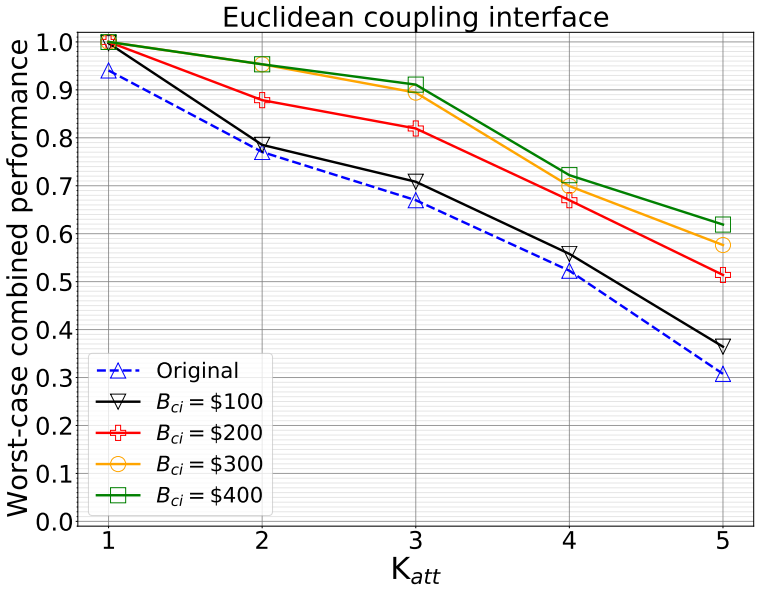

Figure 3.10: Results of the optimal allocation of redundant interdependency links in terms of worst-case combined performance, with different values of K_{att} and B_{ci}. Figure from Paper V [80].

As it can be clearly seen, the allocation of redundant interdependency links considerably improves the worst-case combined performance of the IPGNs. For example, for the case K_{att}=5, the worst-case combined performances with the original Euclidean coupling interface are equal to 0.308, while the optimal allocation of redundancies with a budget B_{ci}=\$400 leads to worst-case combined

Table 3.4: Results of the optimal allocation of redundant interdependency links in terms of worst-case combined performance and relative increase in percentage from the original case. Table from Paper V [80].

B_{ci}	$K_{att} = 1$	$K_{att} = 2$	$K_{att} = 3$	$K_{att} = 4$	$K_{att} = 5$
Original	0.940	0.770	0.670	0.523	0.308
$100	0.997 (+6.04%)	0.785 (+1.93%)	0.708 (+5.74%)	0.558 (+6.71%)	0.364 (+18.43%)
$200	1.0 (+6.33%)	0.879 (+14.03%)	0.819 (+22.33%)	0.670 (+28.13%)	0.514 (+67.04%)
$300	1.0 (+6.33%)	0.953 (+23.74%)	0.894 (+33.49%)	0.699 (+33.77%)	0.577 (+87.33%)
$400	1.0 (+6.33%)	0.953 (+23.74%)	0.911 (+35.98%)	0.722 (+38.12%)	0.619 (+101.21%)

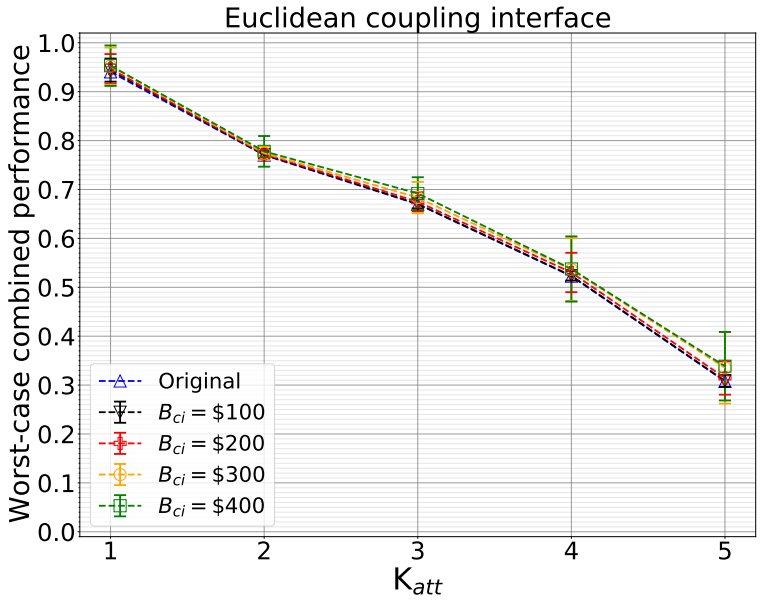

Figure 3.11: Average results of the random strategy allocation of redundant interdependency links in terms of worst-case combined performance, with different values of K_{att} and B_{ci}. Figure from Paper V [80].

performance equal to 0.619, corresponding to an increase of 101.21%.

In order to better understand the potential in terms of resilience enhancement of an optimal allocation of redundant interdependency links, the results in Figure 3.10 are compared to the results of a random allocation strategy, shown in Figure 3.11. For each combination of K_{att} and B_{ci}, 50 experiments with a random allocation strategy (more details in Appendix A of Paper V) are performed. The results

Table 3.5: Computational time in seconds of the NC&CG algorithm. Table from Paper III [10].

B_{ci}	$K_{att} = 1$	$K_{att} = 2$	$K_{att} = 3$	$K_{att} = 4$	$K_{att} = 5$
\$823	1.13	4.12	30.17	50.37	140.30
\$900	2.91	6.26	35.58	156.44	93.36
\$1000	1.73	7.80	21.64	52.67	176.26
\$1100	2.89	11.25	22.40	21.29	47.63
\$1200	2.75	9.45	53.61	97.97	53.41
\$1300	2.87	6.31	27.94	73.95	49.66
\$1400	2.89	8.03	16.48	22.89	40.37
\$1500	2.90	5.86	12.48	39.94	42.58

in Figure 3.11 show the average results and the related 95% confidence intervals. By comparing Figures 3.10 and 3.11, it can be noticed that the random allocation strategy, even when considering the upper bounds of the confidence intervals, is outperformed by the optimal allocation strategy.

Additional results with degree-based and betweenness-based pre-existing coupling interfaces are available in Paper V [80].

3.8 . Computational performance

Multi-level robust optimization problems are generally hard to solve, and often present an NP-hard complexity [109]. The NC&CG algorithm is a powerful framework to solve efficiently the class of problems addressed in Paper III, Paper IV, and Paper V.

The case studies used in this work present small-medium sizes; consequently, the computational cost is affordable. For example, Table 3.5 contains the computational cost in seconds for the optimization in Paper III, previously shown in Figure 3.6. As it can be clearly seen, the computational time is not prohibitive, as the most computationally-expensive optimization (B_{ci}=\$1000, K_{att}=5) requires less than 3 minutes. However, for larger case studies and infrastructures, the computational time required for solving the optimization might increase considerably due to the higher number of binary variables involved. Nevertheless, as it was already highlighted in Paper III, Paper IV, and Paper V, the optimization of coupling interface topologies is a problem that is meant to be solved during design phases, and high computational costs are usually not a problematic issue. Moreover, the computational complexity of the model can be reduced by limiting the number of

binary variables within the optimization problem, e.g. by limiting the number of possible coupling interfaces for the outer defender.

3.9 . Conclusion

In this chapter, some details of the work contained within Paper III, Paper IV, and Paper V are presented. The core of the chapter is to present the proposed approach, based on multilevel mathematical programming, for the optimization of coupling interface topology and design, where the term *coupling interface* denotes the ensemble of interdependency links connecting two or more interdependent CIs.

The approach proposed in Paper III is based on the traditional defender-attacker-defender approach, and it demonstrates the potential of optimized coupling interfaces in terms of improvement of combined performance under the worst feasible failure scenarios.

This approach is extended in Paper IV with a distributionally robust approach that includes ambiguous probability distributions of the set of feasible failure scenarios within the optimization framework. In addition, in Paper V, the efficacy of redundant interdependency links is investigated.

In this works, the objective functions represent the combined performance of the interdependent CIs. From a modeling perspective, it is like to assume that the interdependent CIs are operated with a centralized approach. In reality, interdependent CIs can also be operated with a decentralized approach by independent operators. In the next chapter, a preliminary approach for joint decision-making in interdependent CIs with independent operators that can display different behaviours is proposed.

4 - Joint decision-making with independent operators

4.1 . Introduction

Interdependent CIs are often operated by individual and independent entities, rather than a centralized unique operator. In the previous chapter, the focus of the proposed optimization approaches is to maximize the combined performance of the interdependent CIs. The objective functions of single infrastructures are aggregated within one single objective function through a scalarization approach, assigning a weight to each CI according to its importance (see Equation (3.1)). As the objective is to maximize their combined performance, the independent operators display a behaviour which can be categorized as *centralized* and *collaborative*, as the maximization of one CI performance is not detrimental (actually, it is often beneficial) to the performance of the other CI. For example, considering the IPGNs used in Paper III, Paper IV, and Paper V, the maximization of the power supply within the power network is beneficial to the performance of the gas network, as more electricity is available for the gas network itself. Similarly, the optimization of the gas supply is beneficial to the power network, as more gas is available for the gas-fired power plants.

Centralized approaches based on combined performance optimization have been used extensively in the existing literature, especially in the framework of resilience assessment and optimization. In fact, these models are useful to simulate the behaviour of operators in situations of disruption, when their priority is to maximize the combined performance of their CIs in order to minimize the negative impact of disruption on the general population. For example, centralized operator models have been used to enhance the resilience of interdependent CIs by optimizing protection plans and resource allocation [34], [70], [122] and by optimizing the joint restoration of disrupted interdependent CIs [61], [120], [123]. In the aforementioned works, the operators are centralized, as they can control simultaneously the ensemble of interdependent infrastructures in order to optimize an objective function, which is usually a weighted sum of the performance of each individual infrastructure.

Another class of models for independent operators consists of decentralized approaches, where operators aim at minimizing their own individual performance (usually a cost function). Decentralized models are generally useful for modeling the behaviour of operators under normal conditions. In the context of interdependent CIs, independent operators often interact by selling and purchasing their

respective goods, services, and commodities within a market-based environment. For example, in the case of IPGNs, the independent operators interact by selling and purchasing gas and electricity. In this case, the behaviour of independent operators is *decentralized* and *competitive*, as each of them aims at "selling high" and "purchasing low". The purpose of each operator is usually to optimize its own performance (usually its own cost function), and the interaction between different operators is often modeled through game-theoretic approaches. As highlighted in [132], decentralized models have been used, for example, in the optimization of the expansion planning of interdependent power and gas networks [133]–[135], the optimization of urban energy networks [136], the modeling of the security-constrained operations of integrated wind and hydrogen systems [137] and the risk assessment of interdependent power and heat networks [138].

In summary, the actions of independent operators of interdependent CIs can be modeled with two approaches:

- centralized models: the actions of operators are taken in a centralized and collaborative way, in order to optimize the combined performance of the interdependent CIs. These models are often used while assessing and enhancing the resilience of interdependent CIs under conditions of disruption;
- decentralized models: each operator acts independently in order to optimize its own performance, and the interaction between different operators is usually modeled through game-theoretic approaches. These models are often used to assess and optimize the operators behaviours under normal conditions.

These two classes of models describe two different behaviours of independent operators. When performing joint decision-making, such as the design of a coupling interface between interdependent CIs, planners should take into account the possibility for the independent operators to display different behaviours, namely decentralized under normal conditions and centralized under conditions of disruption. In the previous chapter, coupling interface topologies are optimized with a centralized approach in order to maximize the resilience, in terms of survivability, of interdependent CIs. However, the optimal coupling interface identified with a centralized operator model might be suboptimal under normal conditions, where operators might display a decentralized behaviour (and vice versa). The research question of this chapter is, thus:

- How to design a coupling interface $\mathbf{y}$ which guarantees high-quality performance in interdependent CIs both in the case of centralized and decentralized behaviour of the independent operators?

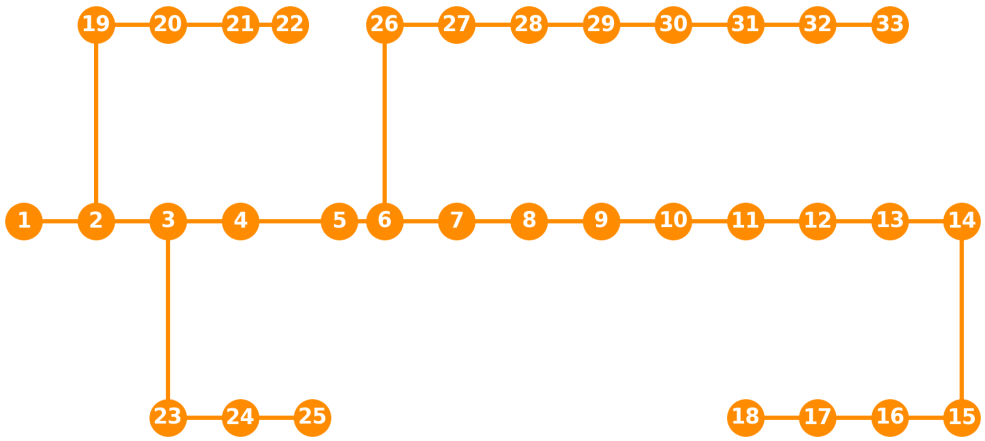

Figure 4.1: Power network of the IPHNs. Figure based on [138].

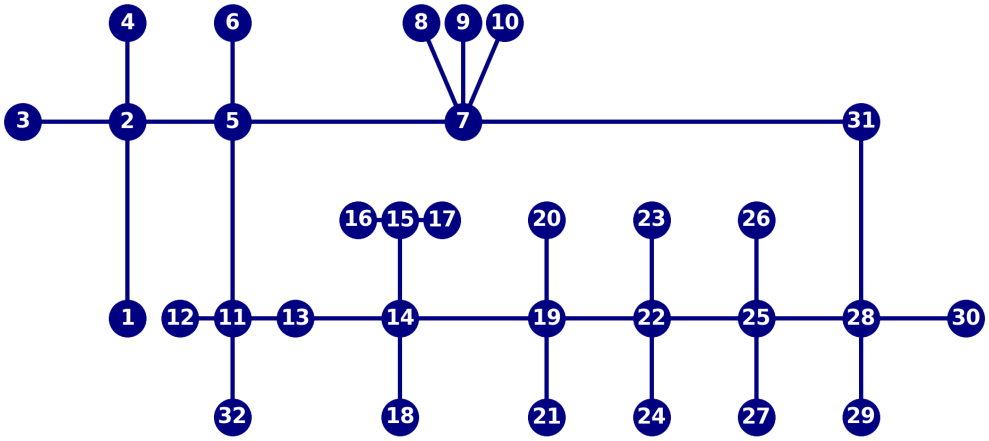

Figure 4.2: Heat network of the IPHNs. Figure based on [138].

4.2 . Case study

In this chapter, the case study of interdependent power and heat networks (IPHNs) proposed in [138] is utilized. The power network and the heat network are shown in Figures 4.1 and 4.2. The power network is based on the topology of the IEEE 33-bus system [139], and it is composed of 33 power buses connected by 32 power lines. The requested power demand in each node ranges from 0 to 0.121 MW, and the total requested power demand is 3.655 MW. The power network network is provided with a generator with a production capacity equal to 3.5 MW in node 1, and four generators with a production capacity equal to 0.5 MW in nodes 7, 15, 23, and 29. The power flow capacity of each line is 3.5 MW. The heat network is based on the topology of the district heat network of Barry Island [140], and it is composed of 32 heat nodes connected by 32 pipelines. The requested heat demand in each node ranges from 0 to 0.145 MW, with a total requested heat demand equal to 2.164 MW. The heat production is ensured by

four boilers, in nodes 1, 20, 31, 32, each with a capacity equal to 0.8 MW. Nodes 1, 31, and 32 are provided with electric boilers, supplied by nodes 5, 14, and 28 of the power network, respectively. Node 20 is provided with a gas boiler, which is assumed to have a reliable supply of gas.

4.3 . Operational models

4.3.1 . Centralized operational model

In a centralized operational model, similarly to Paper III, Paper IV, and Paper V, the interdependent CIs are operated in a centralized way, aiming at optimizing their combined performance. In this chapter, instead of maximizing the residual performance, the focus is on the minimization of loss of performance[9]. The two approaches are equivalent in terms of solution optimality, but the minimization is better suited for the proposed approach in Section 4.5, as it is merged with a cost minimization problem (more details in the next sections). Assuming that both the power network and the heat network are modeled with a linear maximal flow approach, for a fixed coupling interface $\mathbf{y}^*$ and fixed failure scenario $\mathbf{u}^*$, the centralized operational model takes the form of Equation (4.1) subject to (4.2)-(4.7):

$$\min_{\mathbf{p},\mathbf{d},\mathbf{f}} \ \frac{w_{PN}}{\overline{d}_{PN}} \sum_{i \in V_{PN}} d_i + \frac{w_{HN}}{\overline{d}_{HN}} \sum_{i \in V_{HN}} d_i \tag{4.1}$$

$$0 \leq p_i \leq \overline{p}_i, \quad \forall i \in V_{PN} \cup V_{HN} \tag{4.2}$$

$$0 \leq d_i \leq \overline{d}_i, \quad \forall i \in V_{PN} \cup V_{HN} \tag{4.3}$$

$$p_i = \eta \sum_{j \in V_{PN}} y_{ij}^{h \leftarrow p*} d_{ij}^{h \leftarrow p}, \quad \forall i \in V_{HN}^{\mathsf{eb}} \tag{4.4}$$

$$-u_k^* \overline{f}_k \leq f_k \leq u_k^* \overline{f}_k, \quad \forall k \in E_{PN} \cup E_{HN} \tag{4.5}$$

$$p_i - (\overline{d}_i - d_i) - \sum_{j \in V_{HN}^{\mathsf{eb}}} y_{ji}^{h \leftarrow p*} d_{ji}^{h \leftarrow p} + \sum_{k \,|\, D(k)=i} f_k - \sum_{k \,|\, O(k)=i} f_k = 0, \ \forall i \in V_{PN} \tag{4.6}$$

$$p_i - (\overline{d}_i - d_i) + \sum_{k \,|\, D(k)=i} f_k - \sum_{k \,|\, O(k)=i} f_k = 0, \quad \forall i \in V_{HN} \tag{4.7}$$

where d_i represents the load shedding (power or heat) in each node, p_i represents the power or heat production in each node, f_k represents the power flow in each line or the heat flow in each pipeline, and $d_{ij}^{h \leftarrow p}$ represents the power demand of node $i \in V_{HN}^{\mathsf{eb}}$ supplied by node $j \in V_{PN}$, where $V_{HN}^{\mathsf{eb}} \subseteq V_{HN}$ is the subset of nodes in the heat network containing an electric boiler. Similarly to the previous chapter, $y_{ij}^{h \leftarrow p*}$ and u_k^* define the presence of interdependency links and the binary functional state of edges, respectively. The objective function in (4.1) represents

[9]The problems of maximizing residual performance and minimizing performance loss are equivalent.

the weighted sum of the total load shedding of power and heat, where w_{PN} and w_{HN} are the weights of the power network and the heat network. The production level and the load shedding in each node are subject to Constraints (4.2) and (4.3). Moreover, the heat production in each node with an electric boiler depends on the amount of power received from the power network, as shown in Constraint (4.4), where η is the power-to-heat conversion ratio, here assumed to be equal to 0.8. The flow in each line or pipeline is subject to Constraint (4.5). The net balance in each node is enforced by Constraint (4.6) in the power network, and by Constraint (4.7) in the heat network.

The solution of this model leads to the minimization of combined loss of performance, in terms of shedding of power and heat demand.

4.3.2 . Decentralized operational model

The decentralized model is based on the game-theoretic approach proposed in [138]. Power and heat operators act independently, aiming at minimizing their own cost function, and they interact through the purchase of electricity by heat operators in order to produce heat in the electric boilers. As the operators are competing in a market-based environment (power operators aim at selling electricity at a high price, heat operators aim at purchasing it at a low price), the price of electricity and the power demand of the heat network are defined through a game-theoretic approach.

For a fixed coupling interface $\mathbf{y}^*$ and fixed failure scenario $\mathbf{u}^*$, the power network is modeled as Equation (4.8) subject to (4.9)-(4.13):

$$\min_{\mathbf{p},\mathbf{d},\mathbf{f}} \sum_{i \in V_{PN}} \gamma_i^{power} p_i + \sum_{i \in V_{PN}} \alpha_i^{power} d_i - \sum_{i \in V_{PN}} \sum_{j \in V_{HN}^{\mathsf{eb}}} \varepsilon_i d_{ji}^{h \leftarrow p} \tag{4.8}$$

$$0 \leq p_i \leq \overline{p}_i, \quad \forall i \in V_{PN} \tag{4.9}$$

$$0 \leq d_i \leq \overline{d}_i, \quad \forall i \in V_{PN} \tag{4.10}$$

$$-u_k^* \overline{f}_k \leq f_k \leq u_k^* \overline{f}_k, \quad \forall k \in E_{PN} \tag{4.11}$$

$$p_i - (\overline{d}_i - d_i) - \sum_{j \in V_{HN}^{\mathsf{eb}}} y_{ji}^{h \leftarrow p*} d_{ji}^{h \leftarrow p} + \sum_{k \,|\, D(k)=i} f_k - \sum_{k \,|\, O(k)=i} f_k = 0, \forall i \in V_{PN} \tag{4.12}$$

$$\varepsilon^{min} \leq \varepsilon_i \leq \varepsilon^{max}, \quad \forall i \in V_{PN} \tag{4.13}$$

The objective function is composed of three terms:

- the production cost of electricity, where γ_i^{power} defines the cost of producing 1 unit of power in each node i;
- the penalty cost of demand not supplied, where α_i^{power} defines the penalty cost of shedding 1 unit of power in each node i;

- the profit of selling electricity to the heat network, where ε_i defines the selling price of 1 unit of power in each node i.

Constraints (4.9)-(4.12) are equivalent to Constraints (4.2)-(4.3) and (4.5)-(4.6) of the centralized approach. The nodal price of electricity ε_i is bounded between a minimum and maximum value, as enforced by Constraint (4.13).

Similarly, the heat network is modeled as Equation (4.14) subject to (4.15)-(4.19).

$$\min_{\mathbf{p},\mathbf{d},\mathbf{f}} \sum_{i \in V_{HN}^{\mathsf{gb}}} \gamma_i^{heat} p_i + \sum_{i \in V_{HN}} \alpha_i^{heat} d_i + \sum_{i \in V_{HN}^{\mathsf{eb}}} \sum_{j \in V_{PN}} \varepsilon_j d_{ij}^{h \leftarrow p} \tag{4.14}$$

$$0 \leq p_i \leq \overline{p}_i, \quad \forall i \in V_{HN} \tag{4.15}$$

$$p_i = \eta \sum_{j \in V_{PN}} y_{ij}^{h \leftarrow p*} d_{ij}^{h \leftarrow p}, \quad \forall i \in V_{HN} \tag{4.16}$$

$$0 \leq d_i \leq \overline{d}_i, \quad \forall i \in V_{HN} \tag{4.17}$$

$$-u_k^* \overline{f}_k \leq f_k \leq u_k^* \overline{f}_k, \quad \forall k \in E_{HN} \tag{4.18}$$

$$p_i - (\overline{d}_i - d_i) + \sum_{k \,|\, D(k)=i} f_k - \sum_{k \,|\, O(k)=i} f_k = 0, \quad \forall i \in V_{HN} \tag{4.19}$$

The objective function is composed of three terms, respectively:

- the production cost of heat in the gas boilers, where γ_i^{heat} defines the cost of producing 1 unit of heat in each node with a gas boiler i, and $V_{HN}^{\mathsf{gb}} \subseteq V_{HN}$ is the subset of nodes containing a gas boiler;
- the penalty cost of demand not supplied, where α_i^{heat} defines the penalty cost of shedding 1 unit of heat in each node i;
- the cost of purchasing electricity from the power network for the electric boilers, where ε_i defines the selling price of 1 unit of power in each node $i \in V_{PN}$.

Constraints (4.15)-(4.19) are equivalent to Constraints (4.2)-(4.5) and (4.7) of the centralized approach.

As it can be clearly seen, the two problems are interconnected by the prices of electricity ε_i, which are variables of the power network model and constant coefficients of the heat network model, and the power demands of the heat network $d_{ij}^{h \leftarrow p}$, which are variables of the heat network model and constant coefficients of the power network model. In particular, power and heat operators are in competition for the price of the electricity, as the power operators aim at selling the electricity at a high price to maximize their profit, and the heat operators aim at purchasing it at a low price in order to minimize their costs. This interaction can be

modeled through a game-theoretic approach, which allows to identify the equilibrium solutions for electricity prices ε_i and power demands of the heat network $d_{ij}^{h \leftarrow p}$.

In Paper VI, similarly to [138], the interaction between power and heat operators is modeled with the assumptions of a Stackelberg game. In order to solve this problem, the two optimization models, in Equations (4.8)-(4.13) and Equations (4.14)-(4.19), need to be merged within one single problem. Under the assumption of a Stackelberg game, in which the power operators take their actions before the heat operators, i.e. the power operators are the leader and the heat operators the follower, the two problems can be merged using the Karush-Kuhn-Tucker (KKT) conditions of the heat network.

The single-level Stackelberg game between power and heat operators can be formulated as the objective function in (4.8), subject to (4.9)-(4.13), (4.15)-(4.19), which correspond to the primal feasibility of the KKT conditions of the heat network problem, and the other KKT conditions of the heat network problem, shown in Equations (4.23)-(4.25). For the sake of simplicity, the details and the full formulation of the KKT conditions are reported in Appendix B of this dissertation, and here only the compact matrix form is reported.

The compact matrix formulation of the heat network problem corresponds to Equations (4.20) subject to (4.21) and (4.22):

$$\min_{\mathbf{h}} \mathbf{b}^T \mathbf{h} \tag{4.20}$$

$$\mathbf{R}_{in}\mathbf{h} - \mathbf{q}_{in} \leq 0 \tag{4.21}$$

$$\mathbf{R}_{eq}\mathbf{h} - \mathbf{q}_{eq} = 0 \tag{4.22}$$

where (4.21) and (4.22) represent the inequality and equality constraints, previously shown in Constraints (4.15)-(4.19), and they correspond to the primal feasibility of the KKT conditions. With $\boldsymbol{\lambda}$ and $\boldsymbol{\mu}$ being the vectors of dual variables of Constraints (4.21) and (4.22), the other KKT conditions (stationarity, complementary slackness, and dual feasibility) are shown in Equations (4.23)-(4.25), respectively:

$$\nabla \mathbf{c}^T \mathbf{h} + \boldsymbol{\lambda} \nabla \mathbf{R}_{in}\mathbf{h} + \boldsymbol{\mu} \nabla \mathbf{R}_{eq}\mathbf{h} = 0 \tag{4.23}$$

$$\boldsymbol{\lambda}(\mathbf{R}_{in}\mathbf{h} - \mathbf{q}_{in}) = 0 \tag{4.24}$$

$$\boldsymbol{\lambda} \geq 0 \tag{4.25}$$

The solution of this optimization problem leads to the minimization of the individual equilibrium costs in IPHNs under the assumption of a Stackelberg game.

4.4 . Preliminary results

In Paper VI, some preliminary results for assessing the difference between centralized and decentralized operational models are presented. The IPHNs presented in Section 4.2 are used, and a vulnerability analysis, accounting for combinations of 1, 2, and 3 failed power lines, is performed. Each combination of 1, 2, and 3 failed power lines is tested, and the results are evaluated in terms of fraction of power and heat shedding and operational costs using the centralized and decentralized operational models. The cases with 1, 2, and 3 failed lines are referred to as M_{PN}-1, M_{PN}-2, and M_{PN}-3. For the centralized model, costs are computed using the equilibrium prices ε_i identified by the decentralized model under the same failure scenario.

In Figure 4.3, the results in terms of average fraction of combined shedding are shown, while in Figure 4.4, the results in terms of average combined cost are shown. In both the figures, the results are assessed in terms of average values for the three cases (M_{PN}-1, M_{PN}-2, and M_{PN}-3) and using both the centralized and decentralized models. Without indulging into details (the reader is referred to Paper VI), it can be clearly seen how the results, in terms of shedding and cost, considerably differ when using a centralized or a decentralized model. This is a clear indication that: *i)* the energy commodities, in this case power and heat, are dispatched differently in centralized and decentralized operational models, and *ii)* optimal joint decisions, such as the design of a coupling interface, might considerably differ if taken with centralized or decentralized models.

In Figure 4.3, it can be clearly seen how the centralized model leads to lower levels of average combined shedding. This result is somehow expected, as the objective function of the centralized model is the combined power and heat shedding. In Figure 4.4, it can be clearly noticed how the decentralized model leads to lower average combined costs for the case M_{PN}-2 and M_{PN}-3. However, for the case M_{PN}-1, the centralized model leads to a slightly lower cost. This is because the decentralized approach aims at minimizing the equilibrium individual cost of each infrastructure operator, rather than the combined cost. Consequently, it is possible for the centralized approach to lead to lower average combined cost. However, the individual solutions are far from the equilibrium optimality of the decentralized model. For more details, the reader is referred to Figures 4-6 of Paper VI, where it can be clearly seen that the centralized model leads to solutions far from the optimal equilibrium of the decentralized model.

4.5 . Proposed modeling framework

In order to propose a decision-making framework which accounts for the behaviours of independent operators in interdependent CIs, it is necessary to build

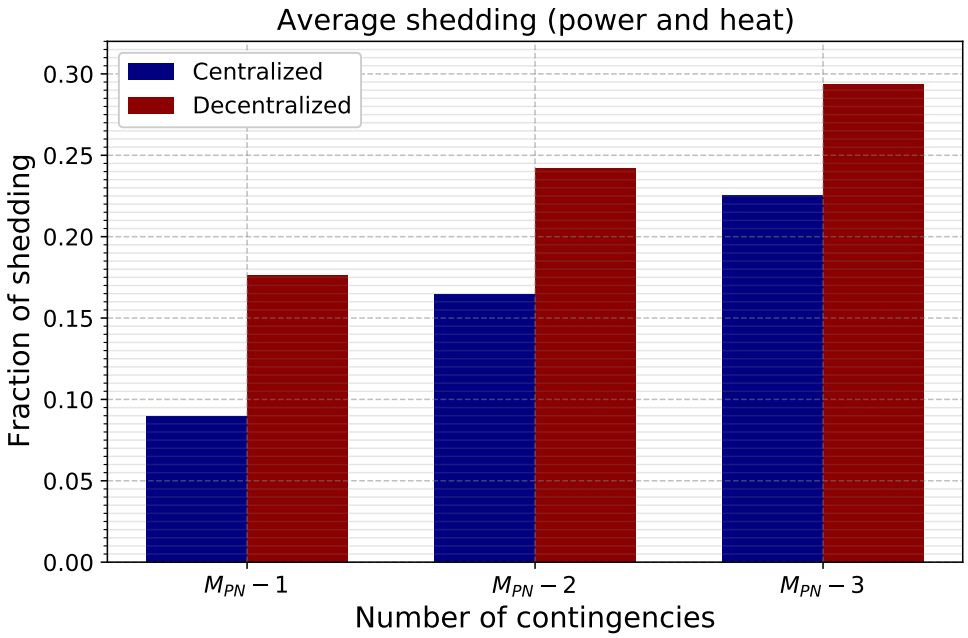

Figure 4.3: Average combined shedding in the IPHNs with different operator models and number of failed power lines. Figure adapted from Paper VI [132].

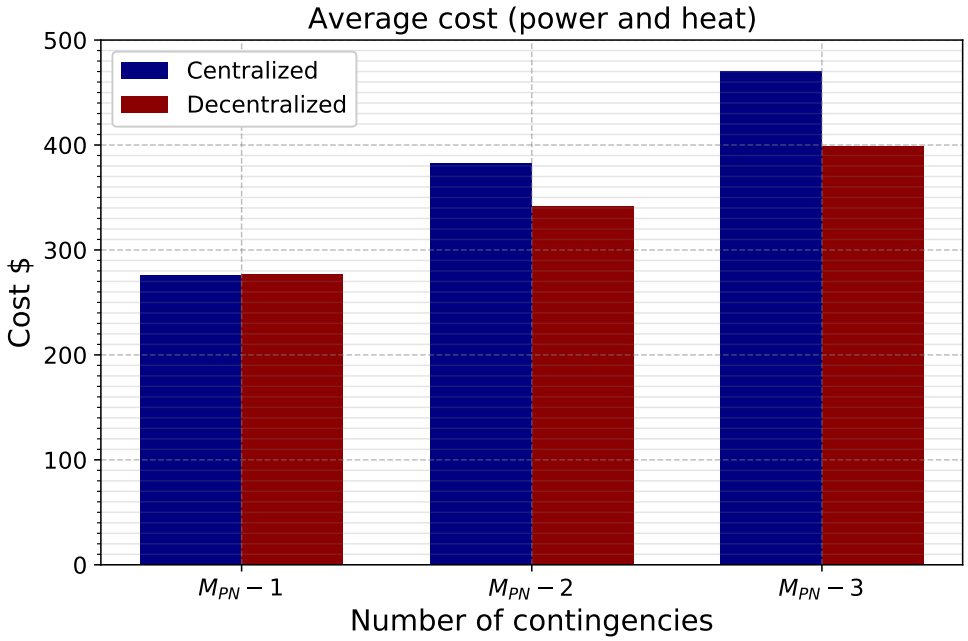

Figure 4.4: Average combined cost in the IPHNs with different operator models and number of failed power lines. Figure adapted from Paper VI [132].

an operational model that prioritizes individual costs under normal conditions, and combined performance in situations of disruptions.

For the sake of clarity, a qualitative case study is built upon the IPHNs proposed in Paper VI and [138]. Let us assume that a new heat network needs to be constructed. As it is provided with electric boilers, it must be coupled with a pre-existing power network. Consistently with the previous chapter, the decision to optimize is the topology of the coupling interface $\mathbf{y}$ between IPHNs.

A distributionally robust approach is used, and decision-makers seek to allocate the coupling interface between IPHNs accounting for an ambiguity set $\mathcal{S}$ containing all the probability distribution on a σ-field of the set of feasible failure scenarios $\mathcal{A}$. Similarly to Paper III-V, only failures of a limited number K_{att} of power lines are considered, and the sets $\mathcal{A}$ and $\mathcal{S}$ are defined as in Equations (4.26) and (4.27).

$$\mathcal{A} = \left\{ \mathbf{u} \,\middle|\, \{0,1\}^{M_{PN}}, ||\mathbb{1}^{M_{PN}} - \mathbf{u}||_1 \leq K_{att} \right\} \tag{4.26}$$

$$\mathcal{S} = \left\{ \mathbb{P} \in \mathcal{P}(\mathcal{A}) : 0 \leq \mathbb{E}_{\mathbb{P}}[\mathbb{1}^{M_{PN}} - \mathbf{u}] \leq \boldsymbol{\pi}^{max} \right\}. \tag{4.27}$$

Contrary to Paper IV and Chapter 3, the probability mass allocated to the normal scenario (every line is functional) is not enforced to be 0 [10]. In other words, the ambiguity set $\mathcal{S}$ denotes the set of multinomial distributions which define the probability of the system to be in a specific state, including the normal scenario. In Paper IV, the ambiguity set $\mathcal{M}$ denotes the set of conditional multinomial distributions which define the probability of the system to be in a specific failure state, excluding the normal scenario. In this case, the upper bounds $\boldsymbol{\pi}^{max}$ of the set $\mathcal{S}$ are not subject to the conditions of the upper bounds of the set $\mathcal{M}$ (see Paper IV for more details).

The proposed DRO approach consists of Equation (4.28) subject to (4.29)-(4.30):

$$\min_{\mathbf{y} \in \{0,1\}^{N_c}} \max_{\mathbb{P} \in \mathcal{S}} \mathbb{E}_{\mathbb{P}} \left[\mathcal{Q}(\mathbf{y}, \boldsymbol{\xi}) \right] \tag{4.28}$$

$$\sum_{j \in V_{PN}} y_{ij}^{h \leftarrow p} \leq 1, \quad \forall i \in V_{HN} \tag{4.29}$$

$$\sum_{\substack{i \in V_{HN} \\ j \in V_{PN}}} y_{ij}^{h \leftarrow p} \, d_{ij}^{km} \, c_{km}^{h \leftarrow p} \leq B_{ci} \tag{4.30}$$

[10] In Paper IV, the coupling interface is optimized in order to maximize the performance under conditions of disruption. Here, normal conditions are also taken into account, using the decentralized approach, and the normal scenario should be assigned some probability mass.

Constraint (4.29) denotes the "single-dependency" assumption, enforcing each node in the heat network to be dependent on a maximum of one node of the power network, while Constraint (4.30) states that the allocation cost of the coupling interface should not exceed the available monetary budget B_{ci}.

Differently from Paper III-V, the purpose of the decision-makers is to identify a coupling interface topology that *minimizes* the expectation of a recourse function $\mathcal{Q}(\mathbf{y}, \boldsymbol{\xi})$ under the worst probability distribution within the ambiguity set $\mathcal{M}$.

The next step is to define a recourse function $\mathcal{Q}(\mathbf{y}, \boldsymbol{\xi})$ which accounts for: *i)* decentralized cost minimization under normal conditions, and *ii)* centralized minimization of loss of combined performance under conditions of disruption. In other words, it is necessary to build a recourse function that integrates the centralized and decentralized approach previously proposed.

NB: *As at the moment of the redaction of this dissertation no final nor preliminary result is available, the final formulation of the following operational model might considerably differ from the one here proposed. In fact, the remaining of this chapter is presented as a proposal of extension of the work already contained within this thesis.*

The proposed centralized and decentralized models can be easily integrated using a binary variable w_r that defines normal conditions and conditions of disruption (w_r=1 under conditions of disruption, w_r=0 under normal conditions). For the sake of simplicity, the objective functions previously used are expressed as $\mathcal{F}^{\mathsf{cnt}}$, $\mathcal{F}^{\mathsf{dcn}}_{PN}$, and $\mathcal{F}^{\mathsf{dcn}}_{HN}$, as defined in Equations (4.31), (4.32), and (4.33):

$$\mathcal{F}^{\mathsf{cnt}} = \frac{w_{PN}}{\overline{d}_{PN}} \sum_{i \in V_{PN}} d_i + \frac{w_{HN}}{\overline{d}_{HN}} \sum_{i \in V_{HN}} d_i \tag{4.31}$$

$$\mathcal{F}^{\mathsf{dcn}}_{PN} = \sum_{i \in V_{PN}} \gamma_i^{power} p_i + \sum_{i \in V_{PN}} \alpha_i^{power} d_i - \sum_{i \in V_{PN}} \sum_{j \in V_{HN}} \varepsilon_i d_{ji}^{h \leftarrow p} \tag{4.32}$$

$$\mathcal{F}^{\mathsf{dcn}}_{HN} = \sum_{i \in V^{gb}_{HN}} \gamma_i^{heat} p_i + \sum_{i \in V_{HN}} \alpha_i^{heat} d_i + \sum_{i \in V_{HN}} \sum_{j \in V_{PN}} \varepsilon_j d_{ij}^{h \leftarrow p} \tag{4.33}$$

For a given coupling interface $\mathbf{y}^*$ and failure scenario $\mathbf{u}^*$, the proposed operational model $\mathcal{Q}(\mathbf{y}, \boldsymbol{\xi})$ consists of Equation (4.34) subject to (4.2)-(4.7), (4.13), and (4.35)-(4.39):

$$\min_{\substack{\mathbf{p},\mathbf{d},\mathbf{f} \\ w_n \in \{0,1\} \\ w_r \in \{0,1\} \\ \boldsymbol{\mu}, \boldsymbol{\lambda} \geq 0}} \quad (1 - w_r)\, \mathcal{F}^{\mathsf{dcn}}_{PN} + w_r\, \mathcal{F}^{\mathsf{cnt}} \tag{4.34}$$

$$(1 - w_r) \left(\nabla \mathcal{F}^{\mathsf{dcn}}_{HN} + \boldsymbol{\lambda} \nabla \mathbf{R}_{in} \mathbf{h} + \boldsymbol{\mu} \nabla \mathbf{R}_{eq} \mathbf{h} \right) = 0 \tag{4.35}$$

$$(1 - w_r)\boldsymbol{\lambda}(\mathbf{R}_{in}\mathbf{h} - \mathbf{q}_{in}) = 0 \tag{4.36}$$

$$w_n \leq \frac{\sum_{k \in E_{PN}} u_k}{M_{PN}} \tag{4.37}$$

$$w_r \leq M_{PN} - \sum_{k \in E_{PN}} u_k \tag{4.38}$$

$$w_r + w_n = 1 \tag{4.39}$$

The objective function in (4.34) is composed of two terms:

- the first term consists of the objective function of the power network in the decentralized approach, and it is accounted for under normal conditions (w_r=0);
- the second term consists of the objective function of the IPHNs in the centralized approach, and it is accounted for under conditions of disruption (w_r=1).

Constraints (4.2)-(4.7) and Constraint (4.13) represent the primal constraints of power and heat network. Constraints (4.35) and (4.36) define the stationarity and complementary slackness of the KKT conditions of the decentralized heat network model, and they are accounted for only under normal conditions thanks to the factor (1-w_r). Constraints (4.37)-(4.39) ensures that w_r=0 under normal conditions, and w_r=1 under conditions of disruption (when there is at least one failure).

The advantage of this model is to identify a coupling interface topology that performs well under normal conditions, where the operators display a decentralized behaviour, and in situations of disruption, where the operators display a centralized behaviour. The normal scenario is optimized using a decentralized model, while all the feasible failure scenarios are optimized with a centralized model. The weight assigned to each scenario corresponds to its probability mass in the corresponding multinomial distribution. The DRO formulation allows to tune the problem towards more risk-neutral or risk-averse solutions.

4.6 . Proposed solution strategy

The DRO model in Equations (4.28)-(4.30) must be recast into a tractable equivalent formulation. The first step is to reformulate the recourse function in order to eliminate the nonlinear terms. Three types of nonlinearities are present: multiplication between binary variables, multiplication between binary and continuous variables, and multiplication between continuous variables.

The multiplications between binary variables can be easily linearized by introducing new binary variables and additional constraints. For example, the multiplication between two binary variables x and y can be replaced by the binary variable

z and Constraints (4.40)-(4.42) [10]:

$$z \leq x \tag{4.40}$$

$$z \leq y \tag{4.41}$$

$$z \geq x + y - 1. \tag{4.42}$$

The multiplications between binary and continuous variables can be easily linearized by introducing new continuous variables and additional constraints, following the "Big-M" method. For example, the multiplication between a binary variables x and a continuous variable a can be replaced by the continuous variable c and Constraints (4.43)-(4.45):

$$c \leq a \tag{4.43}$$

$$c \leq Mx \tag{4.44}$$

$$c \geq a - (1 - x)M \tag{4.45}$$

where M represents a very large number.

The multiplications between continuous variables, in general, can not be linearized without introducing an approximation and, thus, an error. However, in this case, it is possible to obtain an exact linear reformulation. The first terms involving multiplication between binary variables are the products between primal continuous variables $\mathbf{h}$ and dual variables $\boldsymbol{\lambda}$ in the complementary slackness constraints in (4.36). Given the nature of the complementary slackness, these constraints can be linearized by introducing a binary variable and applying a "Big-M" approach [138], [141]. Complementary slackness constraints present the general form in (4.46):

$$\lambda(a - \overline{a}) = 0 \tag{4.46}$$

where a is a continuous primal variable, $\overline{a}$ is a primal parameter, and λ a continuous dual variable, and they can be replaced by Constraints (4.47) and (4.48):

$$\lambda - xM \leq 0 \tag{4.47}$$

$$\overline{a} - a - (1 - x)M \leq 0 \tag{4.48}$$

where x is a binary variable and M a very large number.

The second term containing multiplications between continuous variables is the last term of $\mathcal{F}_{PN}^{\mathsf{dcn}}$ in the objective function (4.34), which contains the product of variables ε_i and $d_{ji}^{h \leftarrow p}$. As this product is also present in the term $\mathcal{F}_{HN}^{\mathsf{dcn}}$ in (4.33), it is possible to exploit the strong duality property of the problem in (4.14)-(4.19) to find the equivalent linear reformulation in (4.49), which can be used to replace the nonlinear term in (4.34), as suggested in [138]. The full formulation of (4.49)

is shown in Appendix C of this dissertation.

$$\sum_{i\in V_{PN}}\sum_{j\in V_{HN}}\varepsilon_i d_{ji}^{h\leftarrow p}=\mathbf{q}_{in}^T\boldsymbol{\lambda}+\mathbf{q}_{eq}^T\boldsymbol{\mu}-\sum_{i\in V_{HN}^{gb}}\gamma_i^{heat}p_i-\sum_{i\in V_{HN}}\alpha_i^{heat}d_i \quad (4.49)$$

By applying these linearizations, the operational model present a linear formulation. The new linearized recourse function is referred to as $\mathcal{L}(\mathbf{y},\boldsymbol{\xi})$. The DRO problem can be formulated as Equation (4.50) subject to (4.2)-(4.7), (4.13), (4.35)-(4.39) and all the additional linearization constraints in the forms in Equations (4.40)-(4.42), (4.43)-(4.45), and (4.47)-(4.48).

$$\min_{\mathbf{y}\in\{0,1\}^{N_c}}\ \max_{\mathbb{P}\in\mathcal{S}}\ \mathbb{E}_{\mathbb{P}}\left[\mathcal{L}(\mathbf{y},\boldsymbol{\xi})\right] \quad (4.50)$$

The second step is to reformulate the new DRO model in (4.50) into an equivalent tractable reformulation, similarly to Paper IV and [118].

For a fixed coupling interface $\mathbf{y}^*$, the inner maximization in (4.50) is equivalent to the problem in (4.51) subject to (4.52)-(4.53):

$$\max_{\mathbb{P}}\int_{\mathcal{A}}\mathcal{L}(\mathbf{y},\boldsymbol{\xi})\,\mathrm{d}\mathbb{P} \quad (4.51)$$

$$\int_{\mathcal{A}}\mathrm{d}\mathbb{P}=1 \quad :\tau \quad (4.52)$$

$$\int_{\mathcal{A}}(1-u_k)\mathrm{d}\mathbb{P}\le\pi_k^{max},\quad \forall k\in E_{PN} \quad :\zeta_k. \quad (4.53)$$

The problem in (4.51)-(4.53) is convex in $\mathbb{P}$, and since Slater's conditions are satisfied, strong duality holds. The dual formulation of (4.51)-(4.53) is shown in Equation (4.54)-(4.55):

$$\min_{\tau,\boldsymbol{\zeta}\ge 0}\ \tau+\sum_{k\in E_{PN}}\zeta_k\pi_k^{max} \quad (4.54)$$

$$\tau+\sum_{k\in E_{PN}}\zeta_k\left(1-u_k^{(i)}\right)-\mathcal{L}(\mathbf{y},\boldsymbol{\xi})\ge 0,\quad \forall\mathbf{u}^{(i)}\in\mathcal{A} \quad (4.55)$$

where τ is the dual variable of Constraint (4.52) and ζ_k are the dual variables of Constraints (4.53). As (4.54) is a minimization, the optimal value $\widehat{\tau}$ of the dual variable τ is equivalent to Equation (4.56):

$$\widehat{\tau}=\max_{\mathbf{u}^{(i)}\in\mathcal{A}}\left\{\mathcal{L}(\mathbf{y},\boldsymbol{\xi})-\sum_{k\in E_{PN}}\zeta_k\left(1-u_k^{(i)}\right)\right\} \quad (4.56)$$

By substituting (4.56) in (4.54), and merging the resulting formulation with the outer level in (4.50), one can obtain Equation (4.57):

$$\min_{\substack{\mathbf{y}\in\{0,1\}^{N_c}\\ \boldsymbol{\zeta}\geq 0}} \max_{\mathbf{u}^{(i)}\in\mathcal{A}} \mathcal{L}(\mathbf{y},\boldsymbol{\xi}) - \sum_{k\in E_{PN}} \zeta_k \left(1 - u_k^{(i)} + \pi_k^{max}\right). \tag{4.57}$$

By expanding the recourse function $\mathcal{L}(\mathbf{y},\boldsymbol{\xi})$ with its explicit form, the formulation presents a min-max-min structure which can be solved efficiently by a NC&CG approach, with an algorithm similar to Figure 3.4 in the previous chapter. The full formulation of the linearized tractable equivalent formulation of the DRO model consists of Equation (4.57) subject to (4.2)-(4.7), (4.13), (4.35)-(4.39) and all the additional linearization constraints in the forms in Equations (4.40)-(4.42), (4.43)-(4.45), and (4.47)-(4.48).

4.7 . Conclusion

In this chapter, a proposal for a decision-making framework in interdependent CIs with independent operators is presented. The proposed approach accounts for decentralized cost optimization under normal conditions and centralized performance optimization under conditions of disruption. The rationale behind this modeling choice is that, under normal conditions, operators interact in a market-based environment by selling and purchasing goods, services and commodities, and they display a decentralized and competitive behaviour. However, under conditions of disruption, the priority might shift towards centralized performance optimization, in order to limit the negative consequences of disruption to the general population; in this case, operators display a centralized and collaborative behaviour.

The proposed approach accounts for the probability of the interdependent CIs to be in a specific state (normal or failure) with a distributionally robust approach. The application of a DRO approach allows to tune the conservativeness of the model towards more risk-neutral (stochastic) or risk-averse (robust) solutions.

5 - Conclusion

5.1 . Vulnerability of railway and power networks

5.1.1 . Contributions

In Chapter 2, based on the work contained in Paper I and Paper II, the vulnerability of interdependent railway and power networks (IRPNs) is investigated. The choice of these CIs is due to the industrial partners involved within this thesis. From the exploratory literature review, three main research gaps are identified:

- the coupling interface between railway and power networks is often modeled not sufficiently in detail, as traction networks are overlooked in most of the existing papers;
- the effect of cascading failure in power networks on the vulnerability of the dependent railway networks is treated rarely and approximately;
- feedback effects of failures in railway networks on the cascading failure dynamics of power networks are not evaluated in the existing literature.

In Paper I, these drawbacks are addressed by introducing the modeling of traction networks to act as an interface between railway and power networks, and by proposing a flow-based cascading failure model for IRPNs. The results in Paper I highlight that:

- it is essential to include traction networks within the modeling framework, as they define how failures and disruption propagate between networks;
- cascading failures within power networks can considerbaly increase the magnitude of cascading effect from power networks to railway networks;
- failures within railway networks can impact the cascading failures dynamic of power networks.

The results in Paper I are obtained by assuming a traction network with substations in electrical isolation. As traction networks are identified as an important factor in the modeling of IRPNs, a more general approach, with the possibility of modeling different traction network configurations, is a natural development of Paper I.

In Paper II, a preliminary proposal of a model that accounts for different configurations of traction networks is presented. The preliminary results suggest that the topology of the coupling interface between different networks might play a key role in terms of resilience and disruption propagation in interdependent CIs, and this consideration leads to the topic treated in Chapter 3.

5.1.2 . Prospective work

A possible extension of the work of Chapter 2 is the development of a more precise and comprehensive model for the vulnerability analysis of IRPNs which accounts for different configurations of traction networks, using the preliminary model in Paper II as a starting point. The topic of vulnerability of IRPNs represents a major subject for the years to come. Due to climate changes, our systems of CIs, including railway and power networks, are undergoing profound changes. In fact, rail transport, when electrified, represents an environmentally friendly means of transportation, and many countries are channeling massive investments in this direction. In addition, power networks are undergoing a profound restyling in order to decarbonize the energy sector. Given these changes, developing tools and methodologies for the vulnerability analysis of IRPNs is essential to guarantee their resilience.

Another promising direction for the vulnerability analysis of IRPNs is to exploit the mathematical approach developed in Chapter 3 to optimize the coupling interfaces between power, traction, and railway networks. This approach is of particular relevance for urban areas, where large networks of CIs are located in a relatively small area, giving the possibility of designing and allocating various coupling interface topologies.

5.2 . Design and optimization of coupling interface topology

5.2.1 . Contributions

In Chapter 3, based on the work contained in Paper III-V, the problem of designing and optimizing the topology of coupling interfaces in interdependent CIs is addressed. In the exploratory literature review, it is highlighted how this problem is acknowledged by scholars and researchers but not treated in detail. In this dissertation, contrary to the heuristics approaches available in the existing literature, a mathematical programming approach is proposed, and interdependent power and gas networks (IPGNs) are selected as representative case study.

In Paper III, a defender-attacker-defender (DAD) approach for the optimization of coupling interface topology is proposed. The results highlight the considerable potential of coupling interface optimization in terms of resilience enhancement against worst-case scenarios. To the best of our knowledge, this paper also represents the first mathematical programming approach for the optimization of coupling interfaces in interdependent CIs.

In Paper IV, a more general model, based on distributionally robust optimization (DRO), is proposed. Contrary to the DAD model, the DRO approach allows to include probability of failure scenarios within the optimization model, and the

conservativeness of the solution can be tuned by decision-makers towards more robust or stochastic solutions. The results show that the DRO approach displays good properties in terms of robustness and modeling of the attitude towards risk of decision-makers.

The proposed approach is further extended with Paper V, in which a DAD model for the optimal allocation of redundant interdependency links in existing coupling interfaces is proposed. The results confirm the potential of redundant interdependency links in terms of resilience enhancement of interdependent CIs.

The proposed models rely on a centralized operational model for the IPGNs, which is valid under the assumption of optimizing under conditions of disruption. However, under normal conditions, interdependent CIs are often operated with a decentralized approach. These classes of operational model reflect the different behaviours of independent operators, which can change according to the state of the CIs. This consideration leads to the work presented in Chapter 4.

5.2.2 . Prospective work

The models proposed in Paper III, Paper IV, and Paper V are complementary, and they could potentially be merged within one single optimization framework that allows to design and retrofit coupling interfaces while accounting for *i)* ambiguous probability of failure scenarios, *ii)* attitude towards risk of decision makers, *iii)* existing coupling interfaces, *iv)* possibility of allocating redundant interdependency links.

An interesting extension of this work is to test the potential, in terms of resilience enhancement, of the optimization of coupling interface topologies using a real case study. This would allow to validate the proposed models as an effective tool for decision-making in a real-world context.

Another possible extension of this work is the inclusion within the analysis and the optimization of nonlinear emerging behaviours of interdependent networks, such as cascading failures within and between networks. These phenomena happen in reality, and the optimization of coupling interface topologies might help increasing the robustness of interdependent CIs against them.

5.3 . Joint decision-making with independent operators

5.3.1 . Contributions

In Chapter 4, based on the work contained in Paper VI, a proposal for a decision-making framework which accounts for the behaviour of independent operators in different situations is presented. Two classes of behaviour are identified:

- *centralized*: the independent operators collaborate in order to optimize their combined performance. This behaviour is usually displayed under conditions of disruption;
- *decentralized*: each operator aims at optimizing its own performance (usually a cost), and their interaction can be modeled through game-theoretic approaches. This behaviour is usually displayed under normal conditions.

Preliminary results in Chapter 4, using interdependent power and heat networks as illustrative case study, demonstrate that centralized and decentralized operational models lead to different results. Consequently, optimal joint decisions, such as the optimal topology of coupling interfaces, might differ if taken with a centralized or decentralized model.

A proposal for a distributionally robust approach for joint decision-making in interdependent CIs with independent operators is presented. The proposed model is able to account for decentralized behaviours under normal conditions, and centralized behaviours under conditions of disruption. A tractable reformulation and a solution strategy are also presented.

5.3.2 . Prospective work

The work contained in Chapter 4 can be extended by performing some results in order to demonstrate the validity of the proposed approach.

The proposed approach is developed using a decentralized model based on the Stackelberg game assumptions. However, independent operators might display different types of decentralized behaviours (e.g. Cournot game). An interesting direction of this work would be to include within the optimization framework the possibility of having multiple decentralized behaviours, according to the specific CIs to optimize. In addition, the effect of different governmental policies, incentives, and penalties on the behaviours of independent operators could also be taken into account within this approach.

A - Nested Column&Constraint Generation algorithm

A.1 . Overview

The extended compact matrix form of the DAD model for the optimal coupling of IPGNs can be expressed as Equation (A.1) subject to Constraints (A.2)-(A.6):

$$\max_{\substack{\mathbf{h}',\boldsymbol{\delta}'\in\{0,1\}^{N_d}\\ \mathbf{y}\in\{0,1\}^{N_c}}} \min_{\mathbf{u}\in\{0,1\}^{M_{PN}}} \max_{\mathbf{h},\boldsymbol{\delta}\in\{0,1\}^{N_d}} \mathbf{b}^T\mathbf{h} \tag{A.1}$$

$$\mathbf{P}\mathbf{y} \leq \mathbf{g} \tag{A.2}$$

$$\mathbf{R}\mathbf{h}' \leq \mathbf{q} - \mathbf{H}\mathbf{y} - \mathbf{W}\boldsymbol{\delta}' - \mathbf{y}^T\mathbf{D}\boldsymbol{\delta}' \tag{A.3}$$

$$\mathbf{b}^T\mathbf{h}' \geq 1 \tag{A.4}$$

$$\sum_{k\in E_{PN}} (1-u_k) \leq K_{att} \tag{A.5}$$

$$\mathbf{R}\mathbf{h} \leq \mathbf{q} - \mathbf{T}\mathbf{u} - \mathbf{H}\mathbf{y} - \mathbf{W}\boldsymbol{\delta} - \mathbf{y}^T\mathbf{D}\boldsymbol{\delta}. \tag{A.6}$$

This problem can be solved efficiently by a *Nested Column&Constraint Generation* (NC&CG) algorithm [104], [109]. The algorithm consists in splitting the problem into an inner and outer layer. Each layer is divided into a master problem and a subproblem, which exchange primal binary variables and provide lower and upper bounds which converge to the optimal solution in finite steps [104].

A.2 . Inner layer

For a fixed coupling interface $\mathbf{y}^*$, the bilevel middle-inner problem in (A.1) identifies the optimal attack plan $\widehat{\mathbf{u}}$, i.e. the worst failure scenario within the set of feasible failure scenarios $\mathcal{A}$, defined as in (A.7):

$$\mathcal{A} = \left\{\mathbf{u} \,\middle|\, \{0,1\}^{M_{PN}}, ||\mathbb{1}^{M_{PN}} - \mathbf{u}||_1 \leq K_{att}\right\}. \tag{A.7}$$

The middle-inner problem takes the form of Equation (A.8) subject to Constraints (A.5)-(A.6).

$$\min_{\mathbf{u}\in\{0,1\}^{M_{PN}}} \max_{\mathbf{h},\boldsymbol{\delta}\in\{0,1\}^{N_d}} \mathbf{b}^T\mathbf{h} \tag{A.8}$$

with fixed outer binary variables $\mathbf{y}^*$. As the inner problem contains the binary variables $\boldsymbol{\delta}$, it is not possible to merge the inner and middle problems using duality theory. Let us fix the middle binary variables $\mathbf{u}$ and the inner binary variables $\boldsymbol{\delta}$, reducing the problem to a pure LP problem. By taking its dual form, with dual

variables $\boldsymbol{\lambda}$, it possible to write the middle-inner problem as its equivalent form in Equation (A.9) subject to (A.5), and (A.10):

$$\min_{\mathbf{u}\in\{0,1\}^{M_{PN}}} \max_{\boldsymbol{\delta}\in\{0,1\}^{N_d}} \min_{\boldsymbol{\lambda}\geq\mathbf{0}} (\mathbf{q}-\mathbf{T}\mathbf{u}-\mathbf{H}\mathbf{y}^{*}-\mathbf{W}\boldsymbol{\delta}-\mathbf{y}^{*T}\mathbf{D}\boldsymbol{\delta})^{T}\boldsymbol{\lambda} \tag{A.9}$$

$$\mathbf{R}^{T}\boldsymbol{\lambda}\leq\mathbf{b}. \tag{A.10}$$

The problem in (A.9) subject to (A.5) and (A.10) is equivalent to the problem in (A.11) subject to (A.5), (A.10), and (A.12):

$$\min_{\rho,\mathbf{u}\in\{0,1\}^{M_{PN}}} \rho \tag{A.11}$$

$$\rho\geq \max_{\boldsymbol{\delta}\in\{0,1\}^{N_d}} \min_{\boldsymbol{\lambda}\geq\mathbf{0}} (\mathbf{q}-\mathbf{T}\mathbf{u}-\mathbf{H}\mathbf{y}^{*}-\mathbf{W}\boldsymbol{\delta}-\mathbf{y}^{*T}\mathbf{D}\boldsymbol{\delta})^{T}\boldsymbol{\lambda}. \tag{A.12}$$

As the vector $\boldsymbol{\delta}$ contains N_d=N_{PN}+N_{GN} binary variables, the possible combinations of $\boldsymbol{\delta}$ variables are finite (precisely 2^{N_d}), and they are contained within the set $\mathcal{D}$. Constraint (A.12) can, thus, be rewritten as in Equation (A.13):

$$\rho\geq(\mathbf{q}-\mathbf{T}\mathbf{u}-\mathbf{H}\mathbf{y}^{*}-\mathbf{W}\boldsymbol{\delta}^{*(i)}-\mathbf{y}^{*T}\mathbf{D}\boldsymbol{\delta}^{*(i)})^{T}\boldsymbol{\lambda}^{(i)},\quad \forall\boldsymbol{\delta}^{*(i)}\in\mathcal{D}. \tag{A.13}$$

The maximization term in (A.12) can be replaced by an enumeration over the set $\mathcal{D}$ of $\boldsymbol{\delta}$ variables, while the minimization term can simply be removed as it does not affect the optimal value of Equation (A.11). Solving (A.11) subject to (A.5), (A.10), and (A.13) is possible in theory, and it would lead to identify the optimal attack plan $\widehat{\mathbf{u}}$. However, this problem presents often a very large scale due to the enumeration over the set $\mathcal{D}$, and it is, thus, often unfeasible in practice. The C&CG approach solves this problem by relying on a limited number of combinations of $\boldsymbol{\delta}$ variables. By decomposing the problem in (A.11), subject to (A.5), (A.10), and (A.13), into a master-subproblem form, it is possible to iteratively reconstruct the set $\mathcal{D}$ and converge asymptotically to the optimal solution. The following steps are employed:

1. Set $i=0$, lower bound $LB_{in}=0$, upper bound $UB_{in}=\infty$, and $\mathcal{D}_{part}=\emptyset$
2. Solve the inner master problem in Equations (A.14)-(A.17). Obtain an optimal solution $\widehat{\rho}^{(i)}$ and optimal attack plan $\widehat{\mathbf{u}}^{(i)}$. Update LB_{in}=$\widehat{\rho}^{(i)}$.

$$\min_{\substack{\rho,\boldsymbol{\lambda}^{(i)}\geq 0\\ \mathbf{u}\in\{0,1\}^{M_{PN}}}} \rho \tag{A.14}$$

subject to:

$$\rho\geq(\mathbf{q}-\mathbf{T}\mathbf{u}-\mathbf{H}\mathbf{y}^{*}-\mathbf{W}\boldsymbol{\delta}^{*(i)}-\mathbf{y}^{*T}\mathbf{D}\boldsymbol{\delta}^{*(i)})^{T}\boldsymbol{\lambda}^{(i)},\ \forall\boldsymbol{\delta}^{*(i)}\in\mathcal{D}_{part} \tag{A.15}$$

$$\mathbf{R}^{T}\boldsymbol{\lambda}^{(i)}\leq\mathbf{b},\quad\forall\boldsymbol{\delta}^{*(i)}\in\mathcal{D}_{part} \tag{A.16}$$

$$\sum_{k\in E_{PN}}(1-u_k)\leq K_{att} \tag{A.17}$$

3. Solve the inner subproblem in Equations (A.18)-(A.19) with $\widehat{\mathbf{u}}^{(i)}$=$\mathbf{u}^*$. Obtain an optimal solution $\mathbf{b}^T\widehat{\mathbf{h}}^{(i)}$ and $\widehat{\boldsymbol{\delta}}^{(i)}$. Set $UB_{in} = \min(UB_{in}, \mathbf{b}^T\widehat{\mathbf{h}}^{(i)})$.

$$\max_{\mathbf{h},\boldsymbol{\delta}\in\{0,1\}^{N_d}} \mathbf{b}^T\mathbf{h} \tag{A.18}$$

subject to :

$$\mathbf{Rh} \leq \mathbf{q} - \mathbf{Tu}^* - \mathbf{Hy}^* - \mathbf{W}\boldsymbol{\delta} - \mathbf{y}^{*T}\mathbf{D}\boldsymbol{\delta} \tag{A.19}$$

4. If $(UB_{in} - LB_{in})/LB_{in} < 10^{-5}$, $\widehat{\mathbf{u}}^{(i)}$ represents the optimal attack plan and the algorithm is terminated. Otherwise, set $\mathcal{D}_{part} = \mathcal{D}_{part} \cup \widehat{\boldsymbol{\delta}}^{(i)}$, update $i \leftarrow i + 1$, and return to step 2.

A.3 . Outer layer

The bilevel outer-middle problem identifies the optimal coupling interface topology, and it can be solved, similarly to the inner layer, by decomposing the problem into a master-subproblem form. The problem in Equations (A.1)-(A.6) can be rewritten as Equation (A.20) subject to (A.2)-(A.6) and (A.21):

$$\max_{\substack{\eta,\mathbf{h}' \\ \boldsymbol{\delta}'\in\{0,1\}^{N_d} \\ \mathbf{y}\in\{0,1\}^{N_c}}} \eta \tag{A.20}$$

$$\eta \leq \min_{\mathbf{u}\in\{0,1\}^{M_{PN}}} \max_{\mathbf{h},\boldsymbol{\delta}\in\{0,1\}^{N_d}} \mathbf{b}^T\mathbf{h} \tag{A.21}$$

By enumeration of the set of feasible failure scenarios $\mathcal{A}$, Constraint (A.21) can be replaced by Constraint (A.22):

$$\eta \leq \mathbf{b}^T\mathbf{h}^{(i)}, \quad \forall \mathbf{u}^{*(i)} \in \mathcal{A} \tag{A.22}$$

The problem can be solved by employing the following steps:

1. Set $i = 0$, lower bound $LB_{out} = 0$, upper bound $UB_{out} = \infty$, and $\mathcal{A}_{part}$=$\emptyset$

2. Solve the outer master problem in Equations (A.23)-(A.28). Obtain an optimal solution $\widehat{\eta}^{(i)}$ and optimal coupling interface $\widehat{\mathbf{y}}^{(i)}$.
   Update UB_{out}=$\min(UB_{out}, \widehat{\eta}^{(i)})$

$$\max_{\substack{\eta,\mathbf{h}',\mathbf{h}^{(i)} \\ \boldsymbol{\delta}'\in\{0,1\}^{N_d} \\ \boldsymbol{\delta}^{(i)}\in\{0,1\}^{N_d} \\ \mathbf{y}\in\{0,1\}^{N_c}}} \eta \tag{A.23}$$

$$\eta \leq \mathbf{b}^T\mathbf{h}^{(i)}, \quad \forall\, \mathbf{u}^{*(i)} \in \mathcal{A}_{part} \tag{A.24}$$

$$\mathbf{Py} \leq \mathbf{g} \tag{A.25}$$

$$\mathbf{b}^T\mathbf{h}' \geq 1 \tag{A.26}$$

$$\mathbf{Rh}' \leq \mathbf{q} - \mathbf{Hy} - \mathbf{W}\boldsymbol{\delta}' - \mathbf{y}^T\mathbf{D}\boldsymbol{\delta}' \tag{A.27}$$

$$\mathbf{Rh}^{(i)} \leq \mathbf{q} - \mathbf{Tu}^{*(i)} - \mathbf{Hy} - \mathbf{W}\boldsymbol{\delta}^{(i)} - \mathbf{y}^T\mathbf{D}\boldsymbol{\delta}^{(i)}, \quad \forall \mathbf{u}^{*(i)} \in \mathcal{A}. \tag{A.28}$$

3. Solve the outer subproblem using the inner layer in the previous subsection with $\widehat{\mathbf{y}}^{(i)}$=$\mathbf{y}^*$. Obtain an optimal solution $\widehat{\rho}^{(i)}$ and an optimal attack plan $\widehat{\mathbf{u}}^{(i)}$. Set $LB_{out} = \widehat{\rho}^{(i)}$.

4. If $(UB_{out} - LB_{out})/LB_{out} < 10^{-5}$, $\widehat{\mathbf{y}}^{(i)}$ is the optimal coupling interface and the algorithm is terminated. Otherwise, set $\mathcal{A}_{part} = \mathcal{A}_{part} \cup \widehat{\mathbf{u}}^{(i)}$, update $i \leftarrow i + 1$, and return to step 2.

B - Karush-Kuhn-Tucker conditions of the decentralized heat network model

The decentralized operational model of the heat network corresponds to Equation (B.1) subject to Constraints (B.2)-(B.6):

$$\min_{\mathbf{p},\mathbf{d},\mathbf{f}} \sum_{i \in V_{HN}^{\mathbf{gb}}} \gamma_i^{heat} p_i + \sum_{i \in V_{HN}} \alpha_i^{heat} d_i + \sum_{i \in V_{HN}} \sum_{j \in V_{PN}} \varepsilon_j d_{ij}^{h \leftarrow p} \tag{B.1}$$

$$0 \leq p_i \leq \overline{p}_i, \quad \forall i \in V_{HN} \quad : \lambda_i^1, \lambda_i^2 \tag{B.2}$$

$$p_i = \eta \sum_{j \in V_{HN}} y_{ij}^{h \leftarrow p*} d_{ij}^{h \leftarrow p}, \quad \forall i \in V_{HN} \quad : \mu_i^1 \tag{B.3}$$

$$0 \leq d_i \leq \overline{d}_i, \quad \forall i \in V_{HN} \quad : \lambda_i^3, \lambda_i^4 \tag{B.4}$$

$$-u_k^* \overline{f}_k \leq f_k \leq u_k^* \overline{f}_k, \quad \forall k \in E_{HN} \quad : \lambda_k^5, \lambda_k^6 \tag{B.5}$$

$$p_i - (\overline{d}_i - d_i) + \sum_{k \,|\, D(k)=i} f_k - \sum_{k \,|\, O(k)=i} f_k = 0, \quad \forall i \in V_{HN} \quad : \mu_i^2 \tag{B.6}$$

where $\boldsymbol{\lambda}$ and $\boldsymbol{\mu}$ are the vectors containing the corresponding dual variables.

The corresponding Karush-Kuhn-Tucker (KKT) conditions (stationarity, primal feasibility, dual feasibility, complementary slackness) are shown in Equations (B.7)-(B.17). Equations (B.18)-(B.29) contain the linearized form of the complementary slackness, where M represents a "Big-M" constant [141]. The terms u_k^* are omitted as, in this case, the failure of heat pipelines is not considered, and the terms u_k^* are always equal to 1.

Stationarity

$$\gamma_i - \lambda_i^1 + \lambda_i^2 + \mu_i^1 + \mu_i^2 = 0, \quad \forall i \in V_{HN} \tag{B.7}$$

$$\alpha_i - \lambda_i^3 + \lambda_i^4 + \mu_i^2 = 0, \quad \forall i \in V_{HN} \tag{B.8}$$

$$-\lambda_k^5 + \lambda_k^6 + \mu_{D(k)}^2 - \mu_{O(k)}^2 = 0, \quad \forall k \in E_{HN} \tag{B.9}$$

$$\varepsilon_j - y_{ij}^{h \leftarrow p} \mu_i^1 = 0, \quad \forall i \in V_{HN}, \forall j \in V_{PN} \tag{B.10}$$

Primal feasibility

Equations (B.2)-(B.6)

Dual feasibility

$$\lambda_i^1, \lambda_i^2, \lambda_i^3, \lambda_i^4, \lambda_i^5, \lambda_i^6 \geq 0, \quad \forall i \in V_{HN} \tag{B.11}$$

Complementary slackness

$$-\lambda_i^1 p_i = 0, \quad \forall i \in V_{HN} \tag{B.12}$$

$$\lambda_i^2 \left(p_i - \overline{p}_i\right) = 0, \quad \forall i \in V_{HN} \tag{B.13}$$

$$-\lambda_i^3 d_i = 0, \quad \forall i \in V_{HN} \tag{B.14}$$

$$\lambda_i^4 \left(d_i - \overline{d}_i\right) = 0, \quad \forall i \in V_{HN} \tag{B.15}$$

$$\lambda_k^5 \left(-f_k - \overline{f}_k\right) = 0, \quad \forall k \in E_{HN} \tag{B.16}$$

$$\lambda_k^6 \left(f_k - \overline{f}_k\right) = 0, \quad \forall k \in E_{HN} \tag{B.17}$$

Linearized complementary slackness

$$\lambda_i^1 - x_i^1 M \leq 0, \quad \forall i \in V_{HN} \tag{B.18}$$

$$p_i - (1 - x_i^1) M \leq 0, \quad \forall i \in V_{HN} \tag{B.19}$$

$$\lambda_i^2 - x_i^2 M \leq 0, \quad \forall i \in V_{HN} \tag{B.20}$$

$$\overline{p}_i - p_i - (1 - x_i^2) M \leq 0, \quad \forall i \in V_{HN} \tag{B.21}$$

$$\lambda_i^3 - x_i^3 M \leq 0, \quad \forall i \in V_{HN} \tag{B.22}$$

$$d_i - (1 - x_i^3) M \leq 0, \quad \forall i \in V_{HN} \tag{B.23}$$

$$\lambda_i^4 - x_i^4 M \leq 0, \quad \forall i \in V_{HN} \tag{B.24}$$

$$\overline{d}_i - d_i - (1 - x_i^4) M \leq 0, \quad \forall i \in V_{HN} \tag{B.25}$$

$$\lambda_k^5 - x_k^5 M \leq 0, \quad \forall k \in E_{HN} \tag{B.26}$$

$$\overline{f}_k + f_k - (1 - x_k^5) M \leq 0, \quad \forall k \in E_{HN} \tag{B.27}$$

$$\lambda_k^6 - x_k^6 M \leq 0, \quad \forall k \in E_{HN} \tag{B.28}$$

$$\overline{f}_k - f_k - (1 - x_k^6) M \leq 0, \quad \forall k \in E_{HN} \tag{B.29}$$

C - Linearization of nonlinear terms with strong duality equivalence

As suggested in [138], thanks to the strong duality property of the problem in (4.14)-(4.19), the equivalence in Equation (C.1), between the decentralized objective function of the heat network and its dual objective function, is satisfied:

$$\sum_{i \in V_{HN}^{\mathsf{gb}}} \gamma_i^{heat} p_i + \sum_{i \in V_{HN}} \alpha_i^{heat} d_i + \sum_{i \in V_{PN}} \sum_{j \in V_{HN}} \varepsilon_i d_{ji}^{h \leftarrow p} =$$

$$\mathbf{q}_{in}^T \boldsymbol{\lambda} + \mathbf{q}_{eq}^T \boldsymbol{\mu} =$$

$$- \sum_{i \in V_{HN}} \lambda_i^2 \overline{p_i} - \sum_{i \in V_{HN}} \lambda_i^4 \overline{d_i} - \sum_{k \in E_{HN}} \lambda_k^6 \overline{f_k} - \sum_{k \in E_{HN}} \lambda_k^7 \overline{f_k} - \sum_{i \in V_{HN}} \mu_i^2 \overline{d_i}. \quad \text{(C.1)}$$

This equivalence can be used to linearize the terms $\varepsilon_i d_{ji}^{h \leftarrow p}$ in $\mathcal{F}_{PN}^{\mathsf{dcn}}$ by substitution.

Bibliography

[1] Accessed: 06/04/2022. [Online]. Available: http://www.lgi.centralesupelec.fr/en/node/167.

[2] B. Arvidsson, J. Johansson, and N. Guldåker, "Critical infrastructure, geographical information science and risk governance: a systematic cross-field review," *Reliability Engineering & System Safety*, vol. 213, p. 107 741, 2021, issn: 0951-8320. doi: https://doi.org/10.1016/j.ress.2021.107741.

[3] Y.-P. Fang, "Critical Infrastructure Protection by Advanced Modelling, Simulation and Optimization for Cascading Failure Mitigation and Resilience," Theses, Ecole Centrale Paris, Feb. 2015. [Online]. Available: https://hal.archives-ouvertes.fr/tel-01144504.

[4] https://ec.europa.eu/home-affairs/pages/page/critical-infrastructure_en, Accessed: 07/03/2022.

[5] A. Berizzi, "The italian 2003 blackout," in *IEEE Power Engineering Society General Meeting, 2004.*, 2004, 1673–1679 Vol.2. doi: 10.1109/PES.2004.1373159.

[6] M. Sforna and M. Delfanti, "Overview of the events and causes of the 2003 italian blackout," in *2006 IEEE PES Power Systems Conference and Exposition*, 2006, pp. 301–308. doi: 10.1109/PSCE.2006.296323.

[7] E. van der Vleuten and V. Lagendijk, "Transnational infrastructure vulnerability: the historical shaping of the 2006 european "blackout"," *Energy Policy*, vol. 38, no. 4, pp. 2042–2052, 2010, Energy Security - Concepts and Indicators with regular papers, issn: 0301-4215. doi: https://doi.org/10.1016/j.enpol.2009.11.047.

[8] A. Urlainis, I. M. Shohet, R. Levy, D. Ornai, and O. Vilnay, "Damage in critical infrastructures due to natural and man-made extreme events – a critical review," *Procedia Engineering*, vol. 85, pp. 529–535, 2014, Selected papers from Creative Construction Conference 2014, issn: 1877-7058. doi: https://doi.org/10.1016/j.proeng.2014.10.580.

[9] M. Ouyang, "Review on modeling and simulation of interdependent critical infrastructure systems," *Reliability Engineering*

& System Safety, vol. 121, pp. 43–60, 2014, issn: 0951-8320. doi: https://doi.org/10.1016/j.ress.2013.06.040.

[10] A. Bellè, A. F. Abdin, Z. Zeng, Y.-P. Fang, and A. Barros, "A resilience-based framework for the optimal coupling of interdependent critical infrastructures," *Reliability Engineering & System Safety*, under review, 2022.

[11] S. Rinaldi, J. Peerenboom, and T. Kelly, "Identifying, understanding, and analyzing critical infrastructure interdependencies," *IEEE Control Systems Magazine*, vol. 21, no. 6, pp. 11–25, 2001. doi: 10.1109/37.969131.

[12] O. of Rail and Road, *Report following railway power disruption on 9th august 2019*, Accessed: 08/03/2022, 2020. [Online]. Available: https://www.orr.gov.uk/media/10752.

[13] S. V. Buldyrev, R. Parshani, G. Paul, H. E. Stanley, and S. Havlin, "Catastrophic cascade of failures in interdependent networks," *Nature*, vol. 464, no. 7291, pp. 1025–1028, 2010. doi: 10.1038/nature08932.

[14] A. Bellè, Z. Zeng, C. Duval, M. Sango, and A. Barros, "Modeling and vulnerability analysis of interdependent railway and power networks: application to british test systems," *Reliability Engineering & System Safety*, vol. 217, p. 108 091, 2022, issn: 0951-8320. doi: https://doi.org/10.1016/j.ress.2021.108091.

[15] N. Sharma, F. Nocera, and P. Gardoni, "Classification and mathematical modeling of infrastructure interdependencies," *Sustainable and Resilient Infrastructure*, vol. 6, no. 1-2, pp. 4–25, 2021. doi: 10.1080/23789689.2020.1753401.

[16] J. E. Bigger, M. G. Willingham, F. Krimgold, and L. Mili, "Consequences of critical infrastructure interdependencies: lessons from the 2004 hurricane season in florida," *International Journal of Critical Infrastructures*, vol. 5, no. 3, pp. 199–219, 2009. doi: 10.1504/IJCIS.2009.024871.

[17] S. E. Chang, T. L. McDaniels, J. Mikawoz, and K. Peterson, "Infrastructure failure interdependencies in extreme events: power outage consequences in the 1998 ice storm," *Natural Hazards*, vol. 41, no. 2, pp. 337–358, 2007. doi: 10.1007/s11069-006-9039-4.

[18] D. Mendonça and W. A. Wallace, "Impacts of the 2001 world trade center attack on new york city critical infrastructures," *Journal of Infrastructure Systems*, vol. 12, no. 4, pp. 260–270, 2006. doi: 10.1061/(ASCE)1076-0342(2006)12:4(260).

[19] G. Kjølle, I. Utne, and O. Gjerde, "Risk analysis of critical infrastructures emphasizing electricity supply and interdependencies," *Reliability Engineering & System Safety*, vol. 105, pp. 80–89, 2012, ESREL 2010, issn: 0951-8320. doi: https://doi.org/10.1016/j.ress.2012.02.006.

[20] T. Brown, W. Beyeler, and D. Barton, "Assessing infrastructure interdependencies: the challenge of risk analysis for complex adaptive systems," *International Journal of Critical Infrastructures*, vol. 1, no. 1, pp. 108–117, 2004. doi: 10.1504/IJCIS.2004.003800.

[21] K. L. S. Bernhardt and S. McNeil, "Agent-based modeling: approach for improving infrastructure management," *Journal of Infrastructure Systems*, vol. 14, no. 3, pp. 253–261, 2008. doi: 10.1061/(ASCE)1076-0342(2008)14:3(253).

[22] D. C. Barton, E. D. Eidson, D. A. Schoenwald, K. L. Stamber, and R. Reinert, "Aspen-ee: an agent-based model of infrastructure interdependency," Dec. 2000. doi: 10.2172/774027. [Online]. Available: https://www.osti.gov/biblio/774027.

[23] B. Bush, L. Dauelsberg, R. LeClaire, D. Powell, S. DeLand, and M. Samsa, *3 critical infrastructure protection decision support system (cip/dss) project overview*, 2005.

[24] N. Santella, L. J. Steinberg, and K. Parks, "Decision making for extreme events: modeling critical infrastructure interdependencies to aid mitigation and response planning," *Review of Policy Research*, vol. 26, no. 4, pp. 409–422, doi: https://doi.org/10.1111/j.1541-1338.2009.00392.x.

[25] W. Leontief, *Input-output economics*. Oxford University Press, 1986.

[26] Y. Y. Haimes and P. Jiang, "Leontief-based model of risk in complex interconnected infrastructures," *Journal of Infrastructure Systems*, vol. 7, no. 1, pp. 1–12, 2001. doi: 10.1061/(ASCE)1076-0342(2001)7:1(1).

[27] A. Rose and S.-Y. Liao, "Modeling regional economic resilience to disasters: a computable general equilibrium analysis of water service disruptions*," *Journal of Regional Science*, vol. 45, no. 1, pp. 75–112, doi: https://doi.org/10.1111/j.0022-4146.2005.00365.x.

[28] A. Rose, G. Oladosu, and S.-Y. Liao, "Regional economic impacts of terrorist attacks on the electric power system of los angeles: a computable general disequilibrium analysis," in *Second Annual*

Symposium of the DHS Center for Risk and Economic Analysis of Terrorism Events, University of Southern California, Los Angeles, CA, Citeseer, 2005.

[29] J. Gao, S. V. Buldyrev, S. Havlin, and H. E. Stanley, "Robustness of a network of networks," *Phys. Rev. Lett.*, vol. 107, p. 195 701, 19 Nov. 2011. doi: 10.1103/PhysRevLett.107.195701.

[30] X. Huang, J. Gao, S. V. Buldyrev, S. Havlin, and H. E. Stanley, "Robustness of interdependent networks under targeted attack," *Phys. Rev. E*, vol. 83, p. 065 101, 6 Jun. 2011. doi: 10.1103/PhysRevE.83.065101.

[31] J. Winkler, L. Dueñas-Osorio, R. Stein, and D. Subramanian, "Interface network models for complex urban infrastructure systems," *Journal of Infrastructure Systems*, vol. 17, no. 4, pp. 138–150, 2011. doi: 10.1061/(ASCE)IS.1943-555X.0000068.

[32] M. Ouyang and L. Dueñas-Osorio, "An approach to design interface topologies across interdependent urban infrastructure systems," *Reliability Engineering & System Safety*, vol. 96, no. 11, pp. 1462–1473, 2011, issn: 0951-8320. doi: https://doi.org/10.1016/j.ress.2011.06.002.

[33] E. E. Lee II, J. E. Mitchell, and W. A. Wallace, "Restoration of services in interdependent infrastructure systems: a network flows approach," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 37, no. 6, pp. 1303–1317, 2007. doi: 10.1109/TSMCC.2007.905859.

[34] Y.-P. Fang and E. Zio, "An adaptive robust framework for the optimization of the resilience of interdependent infrastructures under natural hazards," *European Journal of Operational Research*, vol. 276, no. 3, pp. 1119–1136, 2019, issn: 0377-2217. doi: https://doi.org/10.1016/j.ejor.2019.01.052.

[35] Y. Y. Haimes, "Models for risk management of systems of systems," *International Journal of System of Systems Engineering*, vol. 1, no. 1-2, pp. 222–236, 2008. doi: 10.1504/IJSSE.2008.018138.

[36] I. Eusgeld, C. Nan, and S. Dietz, ""system-of-systems" approach for interdependent critical infrastructures," *Reliability Engineering & System Safety*, vol. 96, no. 6, pp. 679–686, 2011, ESREL 2009 Special Issue, issn: 0951-8320. doi: https://doi.org/10.1016/j.ress.2010.12.010.

[37] M. Beccuti, S. Chiaradonna, F. Di Giandomenico, S. Donatelli, G. Dondossola, and G. Franceschinis, "Quantification of dependen-

cies between electrical and information infrastructures," *International Journal of Critical Infrastructure Protection*, vol. 5, no. 1, pp. 14–27, 2012, issn: 1874-5482. doi: https://doi.org/10.1016/j.ijcip.2012.01.003.

[38] G. D'Agostino, S. Bologna, V. Fioriti, *et al.*, "Methodologies for inter-dependency assessment," in *2010 5th International Conference on Critical Infrastructure (CRIS)*, 2010, pp. 1–7. doi: 10.1109/CRIS.2010.5617578.

[39] A. Di Giorgio and F. Liberati, "A bayesian network-based approach to the critical infrastructure interdependencies analysis," *IEEE Systems Journal*, vol. 6, no. 3, pp. 510–519, 2012. doi: 10.1109/JSYST.2012.2190695.

[40] R. Albert and A.-L. Barabási, "Statistical mechanics of complex networks," *Rev. Mod. Phys.*, vol. 74, pp. 47–97, 1 Jan. 2002. doi: 10.1103/RevModPhys.74.47.

[41] W. Bukhsh and K. McKinnon, "Power systems test case archive," 2013, Accessed: 07/03/2022. [Online]. Available: https://www.maths.ed.ac.uk/optenergy/NetworkData/fullGB/.

[42] Greengauge21, *Fast forward—a high-speed rail strategy for britain*, Accessed: 07/03/2022, 2009. [Online]. Available: http://www.greengauge21.net/fast-forward-a-high-speed-rail-strategy-for-britain/.

[43] D. Sharma and A. Surolia, "Degree centrality," in *Encyclopedia of Systems Biology*, W. Dubitzky, O. Wolkenhauer, K.-H. Cho, and H. Yokota, Eds. New York, NY: Springer New York, 2013, pp. 558–558, isbn: 978-1-4419-9863-7. doi: 10.1007/978-1-4419-9863-7_935.

[44] U. Brandes, "A faster algorithm for betweenness centrality," *Journal of mathematical sociology*, vol. 25, no. 2, pp. 163–177, 2001. doi: 10.1080/0022250X.2001.9990249.

[45] P. Hage and F. Harary, "Eccentricity and centrality in networks," *Social networks*, vol. 17, no. 1, pp. 57–63, 1995. doi: https://doi.org/10.1016/0378-8733(94)00248-9.

[46] G. Sabidussi, "The centrality index of a graph," *Psychometrika*, vol. 31, no. 4, pp. 581–603, 1966. doi: https://doi.org/10.1007/BF02289527.

[47] R. Kinney, P. Crucitti, R. Albert, and V. Latora, "Modeling cascading failures in the north american power grid," *The European Physical Journal B-Condensed Matter and Complex Systems*, vol. 46,

no. 1, pp. 101–107, 2005. doi: https://doi.org/10.1140/epjb/e2005-00237-9.

[48] Y. Fang, N. Pedroni, and E. Zio, "Optimization of cascade-resilient electrical infrastructures and its validation by power flow modeling," *Risk Analysis*, vol. 35, no. 4, pp. 594–607, 2015. doi: https://doi.org/10.1111/risa.12396.

[49] R. Cohen, K. Erez, D. Ben-Avraham, and S. Havlin, "Resilience of the internet to random breakdowns," *Physical review letters*, vol. 85, no. 21, p. 4626, 2000. doi: 10.1103/PhysRevLett.85.4626.

[50] R. Cohen, K. Erez, D. Ben-Avraham, and S. Havlin, "Breakdown of the internet under intentional attack," *Physical review letters*, vol. 86, no. 16, p. 3682, 2001. doi: 10.1103/PhysRevLett.86.3682.

[51] S. Patterson and G. Apostolakis, "Identification of critical locations across multiple infrastructures for terrorist actions," *Reliability Engineering & System Safety*, vol. 92, no. 9, pp. 1183–1203, 2007, Critical Infrastructures, issn: 0951-8320. doi: https://doi.org/10.1016/j.ress.2006.08.004.

[52] J. W. Essam, "Percolation theory," *Reports on progress in physics*, vol. 43, no. 7, p. 833, 1980. doi: https://doi.org/10.1088/0034-4885/43/7/001.

[53] A. Schrijver, "On the history of the transportation and maximum flow problems," *Mathematical programming*, vol. 91, no. 3, pp. 437–445, 2002. doi: https://doi.org/10.1007/s101070100259.

[54] S. G. Nurre, B. Cavdaroglu, J. E. Mitchell, T. C. Sharkey, and W. A. Wallace, "Restoring infrastructure systems: an integrated network design and scheduling (inds) problem," *European journal of operational research*, vol. 223, no. 3, pp. 794–806, 2012. doi: https://doi.org/10.1016/j.ejor.2012.07.010.

[55] M. Ouyang and Y. Fang, "A mathematical framework to optimize critical infrastructure resilience against intentional attacks," *Computer-Aided Civil and Infrastructure Engineering*, vol. 32, no. 11, pp. 909–929, 2017. doi: https://doi.org/10.1111/mice.12252.

[56] A. D. González, L. Dueñas-Osorio, M. Sánchez-Silva, and A. L. Medaglia, "The interdependent network design problem for optimal infrastructure system restoration," *Computer-Aided Civil and Infrastructure Engineering*, vol. 31, no. 5, pp. 334–350, 2016. doi: https://doi.org/10.1111/mice.12171.

[57] A. Conejo and J. Aguado, "Multi-area coordinated decentralized dc optimal power flow," *IEEE Transactions on Power Systems*, vol. 13, no. 4, pp. 1272–1278, 1998. doi: 10.1109/59.736264.

[58] F. Capitanescu, "Critical review of recent advances and further developments needed in ac optimal power flow," *Electric Power Systems Research*, vol. 136, pp. 57–68, 2016, issn: 0378-7796. doi: https://doi.org/10.1016/j.epsr.2016.02.008.

[59] H. W. Dommel and W. F. Tinney, "Optimal power flow solutions," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-87, no. 10, pp. 1866–1876, 1968. doi: 10.1109/TPAS.1968.292150.

[60] Y. Wu, Z. Chen, H. Gong, Q. Feng, Y. Chen, and H. Tang, "Defender–attacker–operator: tri-level game-theoretic interdiction analysis of urban water distribution networks," *Reliability Engineering & System Safety*, vol. 214, p. 107 703, 2021, issn: 0951-8320. doi: https://doi.org/10.1016/j.ress.2021.107703.

[61] E. E. Lee II, J. E. Mitchell, and W. A. Wallace, "Restoration of services in interdependent infrastructure systems: a network flows approach," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 37, no. 6, pp. 1303–1317, 2007. doi: 10.1109/TSMCC.2007.905859.

[62] B. Cavdaroglu, E. Hammel, J. E. Mitchell, T. C. Sharkey, and W. A. Wallace, "Integrating restoration and scheduling decisions for disrupted interdependent infrastructure systems," *Annals of Operations Research*, vol. 203, no. 1, pp. 279–294, 2013. doi: https://doi.org/10.1007/s10479-011-0959-3.

[63] Y.-P. Fang and G. Sansavini, "Optimum post-disruption restoration under uncertainty for enhancing critical infrastructure resilience," *Reliability Engineering & System Safety*, vol. 185, pp. 1–11, 2019, issn: 0951-8320. doi: https://doi.org/10.1016/j.ress.2018.12.002.

[64] B. A. Carreras, V. E. Lynch, I. Dobson, and D. E. Newman, "Critical points and transitions in an electric power transmission model for cascading failure blackouts," *Chaos: An interdisciplinary journal of nonlinear science*, vol. 12, no. 4, pp. 985–994, 2002. doi: 10.1063/1.1505810.

[65] B. A. Carreras, D. E. Newman, I. Dobson, and N. S. Degala, "Validating opa with wecc data," in *2013 46th Hawaii International Conference on System Sciences*, 2013, pp. 2197–2204. doi: 10.1109/HICSS.2013.594.

[66] I. Dobson, B. Carreras, V. Lynch, and D. Newman, "An initial model for complex dynamics in electric power system blackouts," in *Proceedings of the 34th annual Hawaii international conference on system sciences*, vol. 2, 2001, pp. 2017–2017. doi: 10.1109/HICSS.2001.926274.

[67] J. Johansson and H. Hassel, "An approach for modelling interdependent infrastructures in the context of vulnerability analysis," *Reliability Engineering & System Safety*, vol. 95, no. 12, pp. 1335–1344, 2010. doi: https://doi.org/10.1016/j.ress.2010.06.010.

[68] J. Johansson, H. Hassel, and A. Cedergren, "Vulnerability analysis of interdependent critical infrastructures: case study of the swedish railway system," *International journal of critical infrastructures*, vol. 7, no. 4, pp. 289–316, 2011. doi: 10.1504/IJCIS.2011.045065.

[69] Y. Fang and G. Sansavini, "Optimizing power system investments and resilience against attacks," *Reliability Engineering & System Safety*, vol. 159, pp. 161–173, 2017, issn: 0951-8320. doi: https://doi.org/10.1016/j.ress.2016.10.028.

[70] M. Ouyang, "A mathematical framework to optimize resilience of interdependent critical infrastructure systems under spatially localized attacks," *European Journal of Operational Research*, vol. 262, no. 3, pp. 1072–1084, 2017. doi: https://doi.org/10.1016/j.ejor.2017.04.022.

[71] M. Ouyang, L. Hong, Z.-J. Mao, M.-H. Yu, and F. Qi, "A methodological approach to analyze vulnerability of interdependent infrastructures," *Simulation Modelling Practice and Theory*, vol. 17, no. 5, pp. 817–828, 2009, issn: 1569-190X. doi: https://doi.org/10.1016/j.simpat.2009.02.001.

[72] W. Liu and Z. Song, "Review of studies on the resilience of urban critical infrastructure networks," *Reliability Engineering & System Safety*, vol. 193, p. 106 617, 2020, issn: 0951-8320. doi: https://doi.org/10.1016/j.ress.2019.106617.

[73] L. Petersen, D. Lange, and M. Theocharidou, "Who cares what it means? practical reasons for using the word resilience with critical infrastructure operators," *Reliability Engineering & System Safety*, vol. 199, p. 106 872, 2020, issn: 0951-8320. doi: https://doi.org/10.1016/j.ress.2020.106872.

[74] N. Sharma, A. Tabandeh, and P. Gardoni, "Resilience analysis: a mathematical formulation to model resilience of engineering

systems," *Sustainable and Resilient Infrastructure*, vol. 3, no. 2, pp. 49–67, 2018. doi: 10.1080/23789689.2017.1345257.

[75] M. Panteli, P. Mancarella, D. N. Trakas, E. Kyriakides, and N. D. Hatziargyriou, "Metrics and quantification of operational and infrastructure resilience in power systems," *IEEE Transactions on Power Systems*, vol. 32, no. 6, pp. 4732–4742, 2017. doi: 10.1109/TPWRS.2017.2664141.

[76] K. S. Trivedi, V. Jindal, and S. Dharmaraja, "Chapter 7 - stochastic modeling techniques for secure and survivable systems," The Morgan Kaufmann Series in Networking, Y. Qian, J. Joshi, D. Tipper, and P. Krishnamurthy, Eds., pp. 171–207, 2008, issn: 18759351. doi: https://doi.org/10.1016/B978-012373566-9.50009-4.

[77] C. Poulin and M. B. Kane, "Infrastructure resilience curves: performance measures and summary metrics," *Reliability Engineering & System Safety*, vol. 216, p. 107 926, 2021. doi: https://doi.org/10.1016/j.ress.2021.107926.

[78] A. Bellè, Z. Zeng, M. Sango, and A. Barros, "Towards a realistic topological and functional modeling for vulnerability analysis of interdependent railway and power networks," *Proceedings of the 31st European Safety and Reliability Conference*, pp. 2063–2070, Jan. 2021. doi: 10.3850/978-981-18-2016-8_356-cd.

[79] A. Bellè, A. F. Abdin, Y.-P. Fang, Z. Zeng, and A. Barros, "A data-driven distributionally robust approach for the optimal coupling of interdependent critical infrastructures," *European Journal of Operational Research*, under revision after first peer-review, 2022.

[80] A. Bellè, Z. Zeng, M. Sango, and A. Barros, "Resilience enhancement by optimal allocation of redundant interdependency links in interdependent critical infrastructures," *Reliability Engineering & System Safety*, to be submitted, 2022.

[81] "A european-wide power and infrastructure break-down ("black-out") and railways operators," *UIC eNews*, vol. 616, 2018, Accessed: 17/06/2022.

[82] M. S. Mahmoud and Y. Xia, *Networked control systems: cloud control and secure control*. Butterworth-Heinemann, 2019.

[83] H. Guo, C. Zheng, H. H.-C. Iu, and T. Fernando, "A critical review of cascading failure analysis and modeling of power system," *Renewable and Sustainable Energy Reviews*, vol. 80, pp. 9–22, 2017. doi: https://doi.org/10.1016/j.rser.2017.05.206.

[84] J. Zhang, B. Song, Z. Zhang, and H. Liu, "An approach for modeling vulnerability of the network of networks," *Physica A: Statistical Mechanics and its Applications*, vol. 412, pp. 127–136, 2014. doi: https://doi.org/10.1016/j.physa.2014.06.035.

[85] E. Zio and G. Sansavini, "Modeling failure cascades in critical infrastructures with physically-characterized components and interdependencies," Sep. 2010, pp. 652–661. [Online]. Available: https://hal.archives-ouvertes.fr/hal-00721060.

[86] E. Zio and G. Sansavini, "Modeling cascading failures in systems of systems with uncertain behavior," in *ICASP11*, Zurich, Switzerland, Aug. 2011, pp. 1858–1866. [Online]. Available: https://hal.archives-ouvertes.fr/hal-00658101/.

[87] A. E. Motter and Y.-C. Lai, "Cascade-based attacks on complex networks," *Physical Review E*, vol. 66, no. 6, p. 065 102, 2002. doi: 10.1103/PhysRevE.66.065102.

[88] P. Crucitti, V. Latora, and M. Marchiori, "Model for cascading failures in complex networks," *Physical Review E*, vol. 69, no. 4, p. 045 104, 2004. doi: 10.1103/PhysRevE.69.045104.

[89] V. Cupac, J. T. Lizier, and M. Prokopenko, "Comparing dynamics of cascading failures between network-centric and power flow models," *International Journal of Electrical Power & Energy Systems*, vol. 49, pp. 369–379, 2013, issn: 0142-0615. doi: https://doi.org/10.1016/j.ijepes.2013.01.017.

[90] L. Thurner, A. Scheidler, F. Schäfer, *et al.*, "Pandapower — an open-source python tool for convenient modeling, analysis, and optimization of electric power systems," *IEEE Transactions on Power Systems*, vol. 33, no. 6, pp. 6510–6521, Nov. 2018, issn: 0885-8950. doi: 10.1109/TPWRS.2018.2829021.

[91] M. Ouyang, L. Zhao, L. Hong, and Z. Pan, "Comparisons of complex network based models and real train flow model to analyze chinese railway vulnerability," *Reliability Engineering & System Safety*, vol. 123, pp. 38–46, 2014. doi: https://doi.org/10.1016/j.ress.2013.10.003.

[92] R. Parshani, S. V. Buldyrev, and S. Havlin, "Interdependent networks: reducing the coupling strength leads to a change from a first to second order percolation transition," *Physical review letters*, vol. 105, no. 4, p. 048 701, 2010. doi: 10.1103/PhysRevLett.105.048701.

[93] G. Fu, R. Dawson, M. Khoury, and S. Bullock, "Interdependent networks: vulnerability analysis and strategies to limit cascading

failure," *The European Physical Journal B*, vol. 87, no. 7, pp. 1–10, 2014. doi: https://doi.org/10.1140/epjb/e2014-40876-y.

[94] O. Yagan, D. Qian, J. Zhang, and D. Cochran, "Optimal allocation of interconnecting links in cyber-physical systems: interdependence, cascading failures, and robustness," *IEEE Transactions on Parallel and Distributed Systems*, vol. 23, no. 9, pp. 1708–1720, 2012. doi: 10.1109/TPDS.2012.62.

[95] X. Wang, W. Zhou, R. Li, J. Cao, and X. Lin, "Improving robustness of interdependent networks by a new coupling strategy," *Physica A: Statistical Mechanics and its Applications*, vol. 492, pp. 1075–1080, 2018. doi: https://doi.org/10.1016/j.physa.2017.11.037.

[96] S. Chattopadhyay, H. Dai, S. Hosseinalipour, *et al.*, "Designing optimal interlink patterns to maximize robustness of interdependent networks against cascading failures," *IEEE Transactions on Communications*, vol. 65, no. 9, pp. 3847–3862, 2017. doi: 10.1109/TCOMM.2017.2709302.

[97] S. Wang, L. Hong, and X. Chen, "Vulnerability analysis of interdependent infrastructure systems: a methodological framework," *Physica A: Statistical Mechanics and its applications*, vol. 391, no. 11, pp. 3323–3335, 2012. doi: https://doi.org/10.1016/j.physa.2011.12.043.

[98] H. Guo, S. S. Yu, H. H. Iu, T. Fernando, and C. Zheng, "A complex network theory analytical approach to power system cascading failure—from a cyber-physical perspective," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 29, no. 5, p. 053 111, 2019. doi: 10.1063/1.5092629.

[99] Z. Chen, J. Wu, Y. Xia, and X. Zhang, "Robustness of interdependent power grids and communication networks: a complex network perspective," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 65, no. 1, pp. 115–119, 2017. doi: 10.1109/TCSII.2017.2705758.

[100] D. F. Rueda and E. Calle, "Using interdependency matrices to mitigate targeted attacks on interdependent networks: a case study involving a power grid and backbone telecommunications networks," *International Journal of Critical Infrastructure Protection*, vol. 16, pp. 3–12, 2017. doi: https://doi.org/10.1016/j.ijcip.2016.11.004.

[101] Accessed: 30/03/2022. [Online]. Available: https://icseg.iti.illinois.edu/ieee-14-bus-system/.

[102] Accessed: 30/03/2022. [Online]. Available: https://icseg.iti.illinois.edu/wscc-9-bus-system/.

[103] G. G. Brown, W. M. Carlyle, J. Salmerón, and K. Wood, "Analyzing the vulnerability of critical infrastructure to attack and planning defenses," in *Emerging Theory, Methods, and Applications*, Informs, 2005, pp. 102–123. doi: 10.1287/educ.1053.0018.

[104] W. Yuan, L. Zhao, and B. Zeng, "Optimal power grid protection through a defender–attacker–defender model," *Reliability Engineering & System Safety*, vol. 121, pp. 83–89, 2014, issn: 0951-8320. doi: https://doi.org/10.1016/j.ress.2013.08.003.

[105] N. Ghorbani-Renani, A. D. González, and K. Barker, "A decomposition approach for solving tri-level defender-attacker-defender problems," *Computers & Industrial Engineering*, vol. 153, p. 107 085, 2021, issn: 0360-8352. doi: https://doi.org/10.1016/j.cie.2020.107085.

[106] Y. Xiang and L. Wang, "An improved defender–attacker–defender model for transmission line defense considering offensive resource uncertainties," *IEEE Transactions on Smart Grid*, vol. 10, no. 3, pp. 2534–2546, 2019. doi: 10.1109/TSG.2018.2803783.

[107] H. Zhang, S. Ma, T. Ding, Y. Lin, and M. Shahidehpour, "Multistage multi-zone defender-attacker-defender model for optimal resilience strategy with distribution line hardening and energy storage system deployment," *IEEE Transactions on Smart Grid*, vol. 12, no. 2, pp. 1194–1205, 2021. doi: 10.1109/TSG.2020.3027767.

[108] M. Oster, S. Chatterjee, F. Pan, C. Bakker, A. Bhattacharya, and C. Perkins, "Power system resilience through defender-attacker-defender models with uncertainty: an overview," in *2020 Resilience Week (RWS)*, 2020, pp. 11–17. doi: 10.1109/RWS50334.2020.9241279.

[109] B. Zeng and L. Zhao, "Solving two-stage robust optimization problems using a column-and-constraint generation method," *Operations Research Letters*, vol. 41, no. 5, pp. 457–461, 2013. doi: https://doi.org/10.1016/j.orl.2013.05.003.

[110] L. Zhao and B. Zeng, "Vulnerability analysis of power grids with line switching," *IEEE Transactions on Power Systems*, vol. 28, no. 3, pp. 2727–2736, 2013. doi: 10.1109/TPWRS.2013.2256374.

[111] R. T. Rockafellar, S. Uryasev, *et al.*, "Optimization of conditional

value-at-risk," *Journal of risk*, vol. 2, pp. 21–42, 2000. doi: 10.21314/JOR.2000.038.

[112] H. Rahimian and S. Mehrotra, "Distributionally robust optimization: a review," 2019. doi: 10.48550/ARXIV.1908.05659. [Online]. Available: https://arxiv.org/abs/1908.05659.

[113] Z. Wang, K. You, S. Song, and Y. Zhang, "Wasserstein distributionally robust shortest path problem," *European Journal of Operational Research*, vol. 284, no. 1, pp. 31–43, 2020. doi: https://doi.org/10.1016/j.ejor.2020.01.009.

[114] A. Saif and E. Delage, "Data-driven distributionally robust capacitated facility location problem," *European Journal of Operational Research*, vol. 291, no. 3, pp. 995–1007, 2021. doi: https://doi.org/10.1016/j.ejor.2020.09.026.

[115] Y. Liu, H. Xu, S.-J. S. Yang, and J. Zhang, "Distributionally robust equilibrium for continuous games: nash and stackelberg models," *European Journal of Operational Research*, vol. 265, no. 2, pp. 631–643, 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S037722171730677X.

[116] X. V. Doan, "Distributionally robust optimization under endogenous uncertainty with an application in retrofitting planning," *European Journal of Operational Research*, 2021. doi: https://doi.org/10.1016/j.ejor.2021.07.013.

[117] A. Arrigo, C. Ordoudis, J. Kazempour, Z. De Grève, J.-F. Toubeau, and F. Vallée, "Wasserstein distributionally robust chance-constrained optimization for energy and reserve dispatch: an exact and physically-bounded formulation," *European Journal of Operational Research*, vol. 296, no. 1, pp. 304–322, 2022. doi: https://doi.org/10.1016/j.ejor.2021.04.015.

[118] S. Babaei, R. Jiang, and C. Zhao, "Distributionally robust distribution network configuration under random contingency," *IEEE Transactions on Power Systems*, vol. 35, no. 5, pp. 3332–3341, 2020. doi: 10.1109/TPWRS.2020.2973596.

[119] X. Liu, Y.-P. Fang, and E. Zio, "A hierarchical resilience enhancement framework for interdependent critical infrastructures," *Reliability Engineering & System Safety*, vol. 215, p. 107 868, 2021. doi: https://doi.org/10.1016/j.ress.2021.107868.

[120] M. Ouyang and Z. Wang, "Resilience assessment of interdependent infrastructure systems: with a focus on joint restoration

modeling and analysis," *Reliability Engineering & System Safety*, vol. 141, pp. 74–82, 2015. doi: https://doi.org/10.1016/j.ress.2015.03.011.

[121] M. Ouyang, "Critical location identification and vulnerability analysis of interdependent infrastructure systems under spatially localized attacks," *Reliability Engineering & System Safety*, vol. 154, pp. 106–116, 2016. doi: https://doi.org/10.1016/j.ress.2016.05.007.

[122] J. Kong, C. Zhang, and S. P. Simonovic, "Optimizing the resilience of interdependent infrastructures to regional natural hazards with combined improvement measures," *Reliability Engineering & System Safety*, vol. 210, p. 107 538, 2021. doi: https://doi.org/10.1016/j.ress.2021.107538.

[123] Y. Almoghathawi, K. Barker, and L. A. Albert, "Resilience-driven restoration model for interdependent infrastructure networks," *Reliability Engineering & System Safety*, vol. 185, pp. 12–23, 2019. doi: https://doi.org/10.1016/j.ress.2018.12.006.

[124] J. Shao, S. V. Buldyrev, S. Havlin, and H. E. Stanley, "Cascade of failures in coupled network systems with multiple support-dependence relations," *Physical Review E*, vol. 83, no. 3, Mar. 2011. doi: 10.1103/physreve.83.036116.

[125] A. Peiravi, M. Karbasian, M. A. Ardakan, and D. W. Coit, "Reliability optimization of series-parallel systems with k-mixed redundancy strategy," *Reliability Engineering & System Safety*, vol. 183, pp. 17–28, 2019. doi: https://doi.org/10.1016/j.ress.2018.11.008.

[126] D. W. Coit, "Maximization of system reliability with a choice of redundancy strategies," *IIE transactions*, vol. 35, no. 6, pp. 535–543, 2003. doi: 10.1080/07408170304420.

[127] S. Kulturel-Konak, A. E. Smith, and D. W. Coit, "Efficiently solving the redundancy allocation problem using tabu search," *IIE transactions*, vol. 35, no. 6, pp. 515–526, 2003. doi: 10.1080/07408170304422.

[128] D. W. Coit, "Cold-standby redundancy optimization for nonrepairable systems," *IIE Transactions*, vol. 33, no. 6, pp. 471–478, 2001. doi: https://doi.org/10.1023/A:1007689912305.

[129] H. Garg and S. Sharma, "Multi-objective reliability-redundancy allocation problem using particle swarm optimization," *Computers & Industrial Engineering*, vol. 64, no. 1, pp. 247–255, 2013. doi: https://doi.org/10.1016/j.cie.2012.09.015.

[130] Z. Ouyang, Y. Liu, S.-J. Ruan, and T. Jiang, "An improved particle swarm optimization algorithm for reliability-redundancy allocation problem with mixed redundancy strategy and heterogeneous components," *Reliability Engineering & System Safety*, vol. 181, pp. 62–74, 2019. doi: https://doi.org/10.1016/j.ress.2018.09.005.

[131] W.-C. Yeh, "A novel boundary swarm optimization method for reliability redundancy allocation problems," *Reliability Engineering & System Safety*, vol. 192, p. 106 060, 2019. doi: https://doi.org/10.1016/j.ress.2018.02.002.

[132] A. Bellè, Z. Zeng, and A. Barros, "Vulnerability analysis of interdependent energy infrastructures with centralized and decentralized operator models," *Proceedings of the 32nd European Safety and Reliability Conference*, accepted, 2022.

[133] V. Z. Rad, S. A. Torabi, and H. Shakouri, "Joint electricity generation and transmission expansion planning under integrated gas and power system," *Energy*, vol. 167, pp. 523–537, 2019. doi: https://doi.org/10.1016/j.energy.2018.10.178.

[134] A. J. Conejo, S. Chen, and G. E. Constante, "Operations and long-term expansion planning of natural-gas and power systems: a market perspective," *Proceedings of the IEEE*, vol. 108, no. 9, pp. 1541–1557, 2020. doi: 10.1109/JPROC.2020.3005284.

[135] J. Qiu, Z. Y. Dong, J. H. Zhao, *et al.*, "Multi-stage flexible expansion co-planning under uncertainties in a combined electricity and gas market," *IEEE Transactions on Power Systems*, vol. 30, no. 4, pp. 2119–2129, 2014. doi: 10.1109/TPWRS.2014.2358269.

[136] R. Jing, M. Wang, H. Liang, *et al.*, "Multi-objective optimization of a neighborhood-level urban energy network: considering game-theory inspired multi-benefit allocation constraints," *Applied energy*, vol. 231, pp. 534–548, 2018. doi: https://doi.org/10.1016/j.apenergy.2018.09.151.

[137] M. A. Mirzaei, A. S. Yazdankhah, and B. Mohammadi-Ivatloo, "Stochastic security-constrained operation of wind and hydrogen energy storage systems integrated with price-based demand response," *International Journal of Hydrogen Energy*, vol. 44, no. 27, pp. 14 217–14 227, 2019. doi: https://doi.org/10.1016/j.ijhydene.2018.12.054.

[138] C. Wang, C. Yan, G. Li, S. Liu, and Z. Bie, "Risk assessment of integrated electricity and heat system with independent energy operators based on stackelberg game," *Energy*, vol. 198,

p. 117 349, 2020, issn: 0360-5442. doi: https://doi.org/10.1016/j.energy.2020.117349.

[139] M. Baran and F. Wu, "Network reconfiguration in distribution systems for loss reduction and load balancing," *IEEE Transactions on Power Delivery*, vol. 4, no. 2, pp. 1401–1407, 1989. doi: 10.1109/61.25627.

[140] X. Liu, J. Wu, N. Jenkins, and A. Bagdanavicius, "Combined analysis of electricity and heat networks," *Applied Energy*, vol. 162, pp. 1238–1250, 2016, issn: 0306-2619. doi: https://doi.org/10.1016/j.apenergy.2015.01.102.

[141] J. Fortuny-Amat and B. McCarl, "A representation and economic interpretation of a two-level programming problem," *Journal of the operational Research Society*, vol. 32, no. 9, pp. 783–792, 1981. doi: https://doi.org/10.1057/jors.1981.156.

Appended papers

Paper I
A. Bellè, Z. Zeng, C. Duval, M. Sango, and A. Barros, "Modeling and vulnerability analysis of interdependent railway and power networks : Application to British test systems," *Reliability Engineering & System Safety*, t. 217, p. 108091, 2022.

Paper II A. Bellè, Z. Zeng, M. Sango, and A. Barros, "Towards a Realistic Topological and Functional Modeling for Vulnerability Analysis of Interdependent Railway and Power Networks," *Proceedings of the 31st European Safety and Reliability Conference*, 2021.

Paper III A. Bellè, A. F. Abdin, Z. Zeng, Y.-P. Fang, and A. Barros, "A resilience-based framework for the optimal coupling of interdependent critical infrastructures," *Reliability Engineering & System Safety*, under review, 2022.

Paper IV A. Bellè, A. F. Abdin, Y.-P. Fang, Z. Zeng, and A. Barros, "A data-driven distributionally robust approach for the optimal coupling of interdependent critical infrastructures," *European Journal of Operational Research*, under revision after first peer-review, 2022.

Paper V A. Bellè, Z. Zeng, M. Sango, and A. Barros, "Resilience enhancement by optimal allocation of redundant interdependency links in interdependent critical infrastructures," *Reliability Engineering & System Safety*, to be submitted, 2022.

Paper VI A. Bellè, Z. Zeng, and A. Barros, "Vulnerability analysis of interdependent energy infrastructures with centralized and decentralized operator models," *Proceedings of the 32nd European Safety and Reliability Conference*, accepted, 2022.

Paper VII A. Bellè, Y.-P. Fang, Z. Zeng, and A. Barros, "A distributionally robust approach for the optimal protection of power networks with endogenous uncertainty," *Proceedings of the 18th IFAC Workshop on Control Application of Optimization*, accepted, 2022.

Paper I

A. Bellè, Z. Zeng, C. Duval, M. Sango, and A. Barros, "Modeling and vulnerability analysis of interdependent railway and power networks : Application to British test systems," *Reliability Engineering & System Safety*, t. 217, p. 108091, 2022.

Reliability Engineering and System Safety 217 (2022) 108091

Contents lists available at ScienceDirect

Reliability Engineering and System Safety

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

Modeling and vulnerability analysis of interdependent railway and power networks: Application to British test systems

Andrea Bellè [a,*], Zhiguo Zeng [a], Carole Duval [b], Marc Sango [c], Anne Barros [a]

[a] Chair on Risk and Resilience of Complex Systems, Laboratoire Génie Industriel, CentraleSupélec, Université Paris-Saclay, France
[b] Électricité de France (EDF) R&D Palaiseau, France
[c] SNCF, Direction Technologies Innovation et Projet Groupe, Innovation & Recherche, Département Physique du Système Ferroviaire, Equipe Sécurité Système, France

ARTICLE INFO

Keywords:
Critical infrastructures
Vulnerability
Interdependent networks
Cascading failures
Power network
Railway network

ABSTRACT

Railway and power networks are among the most important critical infrastructures and their vulnerability under different types of disrupting events has been analyzed extensively. However, the increasing degree of interconnection between these two critical infrastructures makes it necessary to consider the multiple interdependencies when conducting vulnerability assessments. We propose an approach for the modeling and vulnerability analysis of interdependent railway and power networks which accounts for a realistic coupling through the traction power network, which acts as an interface. Moreover, we consider cascading failures within and between networks. With the proposed model, we show that failures in the power network have a considerable negative impact on the railway network. In addition, our analysis shows that the dependency of the railway network on the power network can considerably impact vulnerability and cascading failures dynamics of the power network.

1. Introduction

1.1. Motivation

Critical infrastructures, such as transportation, healthcare, energy systems, water supply systems and telecommunication networks, are large-scale systems that provide essential goods and services for society. Their malfunctioning and failure can heavily affect the safety and the socio-economic stability of a population. Developing highly reliable infrastructures is, thus, an important issue. Lately, due to the increasing degree of interconnection and interdependency, infrastructures have transformed from isolated individual systems to highly interconnected systems-of-systems. This evolution, while leading to improvements of performance, has introduced new failure modes and threats [1]. Critical infrastructures are often analyzed in terms of their vulnerability to different types and magnitudes of disrupting events. Vulnerability analysis is defined in [2] as the process of "systematically and comprehensively identifying the possible states a system can be put into, due to specific strains, and estimating the negative consequences associated with them". This analysis has proved to be useful in design and prevention stages, as it gives an overview of how a system responds to different disruptive situations.

Transportation systems, railway networks in particular, are acknowledged as one of the most important infrastructures. It is recognized that railway systems are dependent on several internal and external subsystems and infrastructures [3]. The dependency of railway networks on external power systems is of particular relevance, as these systems supply the electricity necessary for a proper functioning. In case of a disruptive event in a power grid, this strong dependency can lead to considerable negative consequences on the dependent railway systems. In fact, power outages and blackouts have been observed to be an important cause of railway networks disruption. For example, in the UK in August 2019, issues within the power grid led to major disruption of the railway system [4,5]. It is clear that railway operators should be aware of power network-induced risks [6].

Vulnerability analysis of railway networks dependent on external power networks has been the focus of various studies. These existing works, although recognizing the importance of the power systems for the railway networks functionality, do not model in details the interconnection between these two systems. In addition, the effect of cascading failures in power networks and their repercussions on railway networks are often treated approximately or overlooked. We refer to cascading failures as the process of failure propagation within and between different systems. Cascading failures have often been observed in power networks. These failures have the potential to cause major disruption in power networks, as well as within the dependent systems, such as railway networks. Neglecting the cascading failures analysis within a vulnerability assessment of interdependent railway and power

* Corresponding author.
E-mail address: andrea.belle@centralesupelec.fr (A. Bellè).

https://doi.org/10.1016/j.ress.2021.108091
Received 27 April 2021; Received in revised form 19 August 2021; Accepted 21 September 2021
Available online 30 September 2021

List of Symbols

Railway network

$\mathbf{E_r}$	Set of edges in railway network
$\mathbf{G_r}$	Graph representing the railway network
$\mathbf{V_r}$	Set of nodes in railway network
$e_{r,i}$	Edge i in railway network, representing railway segment i
M_r	Number of edges in railway network
N_r	Number of nodes in railway network
$v_{r,i}$	Node i in the railway network, representing station i

External power network

$\mathbf{E_p}$	Set of edges in external power network
$\mathbf{G_p}$	Graph representing the external power network
$\mathbf{g_p}$	Set of generators in external power network
$\mathbf{l_p^{base}}$	Subset of loads in external power network supplying general consumers
$\mathbf{l_p^{sub}}$	Subset of loads in external power network supplying traction substations
$\mathbf{l_p}$	Set of loads in external power network
$\mathbf{V_p}$	Set of nodes in external power network
$e_{p,i}$	Edge i in external power network, representing transmission line i
g_i	Generator i in external power network
l_i	Load i in external power network
M_p	Number of edges in external power network
N_g	Number of generators in external power network
N_l	Number of loads in external power network
N_p	Number of nodes in external power network
$v_{p,i}$	Node i in the external power network, representing electrical bus i

Traction power network

$\mathbf{G_t}$	Graph representing the traction power network
$\mathbf{V_t}$	Set of nodes in traction power network
N_t	Number of nodes in traction power network
$v_{t,i}$	Node i in the traction power network, representing substation i

Interdependency

$\mathbf{E}_{r\leftarrow t}^{i\leftarrow j}$	Set of interdependency edges between railway track i and subset of substations $\mathbf{V_t^{r,i}}$
$\mathbf{V_t^{r,i}}$	Subset of substations supplying the railway track i
$e_{t\leftarrow p}^{i\leftarrow j}$	Interdependency edge between substation i and electrical bus j with load k

Initiating event

$\mathbf{N_f}$	Set of components to remove in the initiating event
f	Fraction of components to remove in the initiating event
N_{fail}	Number of elements to remove in the initiating event
p_{ie}	Probability of being selected for a removal for each component

Cascading failures

$\mathbf{E'}$	Set of edges after a disruptive event
$\mathbf{F}$	Set of line power flows
$\mathbf{G'}$	Graph after a disruptive event
$\mathbf{N_F^{dir}}$	Set containing the substations failed directly
$\mathbf{N_F^{ind}}$	Set containing the substations failed indirectly
$\mathbf{P_g}$	Set of nominal generator power
$\mathbf{P_g'}$	Set of nominal generator power after a disruptive event
$\mathbf{P_l}$	Set of nominal load power
$\mathbf{P_l'}$	Set of nominal load power after a disruptive event
$\mathbf{V'}$	Set of nodes after a disruptive event
F_i	Power flow at line i
F_i^{max}	Power flow capacity at line i
$P_{g,i}$	Nominal power of generator i
$P_{g,i}^{max}$	Maximum power of generator i
$P_{l,i}$	Supplied power demand of load i
$P_{l,i}^{max}$	Requested power demand of load i
$R_{t,i}$	Load shedding ratio of substation i
$S_{p,i}$	Binary functional state of transmission line i
$S_{r,i}$	Binary functional state of railway track i
$S_{t,i}$	Binary functional state of substation i
$T_{t\leftarrow p}$	Tolerance threshold for substation load shedding
W	Load penalty constant

Simulation

$\bar{V}$	Average vulnerability index
σ	Standard deviation estimation
A_r	Accessibility of railway network
CI_{95}	95% confidence intervals
DNS	Demand Not Supplied in megawatts
$FDNS$	Fraction of Demand Not Supplied
n_a^i	Number of stations accessible from station i
N_{exp}	Number of Monte Carlo simulations
PI	performance indicator
RPP	Railway Power Performance
TPP_i	Traction Power Performance of railway track i
V	Vulnerability index
Z	Confidence intervals constant

networks could potentially lead to a significant underestimation of the risks. With this work, we aim at complementing the available literature with a modeling approach for interdependent railway and power networks which accounts for a realistic modeling of the interface between railway and power systems and flow-based characterization of cascading failures in power networks, including their consequences in railway networks. In addition, we aim at estimating the effect of failures within railway networks on the vulnerability and cascading failures dynamics of power networks.

1.2. Related work

Power and railway networks are a traditional example of interdependent critical infrastructures. In fact, railway networks are often electrified and dependent on external power networks in terms of electricity supply. In general, two networks A and B are interdependent if the functionality of one depends on the state/output of the other one. Interdependencies are unidirectional (often simply referred to as dependencies) if network A depends on network B but not vice versa; otherwise, if networks A and B are mutually dependent on each other, the interdependencies are bidirectional (or simply referred to as interdependencies) [7,8]. The nature and the features of interdependencies

have been an important subject of research since the early 2000s, and various classifications are available in the existing literature. One the most used classification is the one proposed in [7], where the authors divide interdependencies into four categories: physical, when the state of one system is dependent on the material output of another system; cyber, when the state of one system is dependent on the information transmitted through another system; geographic, when different systems share the same location and their state can be modified by an environmental event; and logical, if the interdependency is not physical, cyber, or geographic. Within this framework, the unidirectional interdependency from power to railway networks can be defined as physical. Alternative classifications of interdependencies, which focus on different aspects, are available in the existing literature, such as in [8–10].

As critical infrastructures represent the backbone of essential societal functions, ensuring their resilience is an important issue. The resilience of a system is defined as *"its ability to withstand stressors, adapt, and rapidly recover from disruptions"* [11]. Resilience is generally defined as the combination of three phases: the disturbance progress, which describes how fast and severe are the damages to the system, the post-disturbance degraded state, which describes how extensive in time the damages are, and the restoration, which describes the system restoration [12]. Lately, the resilience of interdependent critical infrastructures has gained particular interest from researchers, and many works have addressed this topic in recent years (e.g. [13–16]). The first phase of resilience is strictly correlated with the concept of vulnerability, which can be defined as *"degree of loss or damage to a system when exposed to a strain of a given type and magnitude"* [2]. Analyzing systems in terms of their vulnerability is essential to assess their exposure to external strains and stresses, and it is an important dimension of systems resilience.

Both railway and power networks have been extensively analyzed in terms of vulnerability. For example, the vulnerability of railway networks has been assessed in terms of different performance indicators, such as topological metrics [17–19] and flow-based metrics [20,21], and different disruption scenarios, such as random and targeted failures [22,23] or natural hazards [24,25]. For more references, the reader is referred to [26,27], where the authors discuss the concepts of vulnerability and resilience in transportation networks from a research perspective, including a comprehensive literature review.

Similarly, power networks have been analyzed in several works, in terms of different disruption scenarios, such as random failures [28–30], natural hazards [31] and intentional attacks [32–34], and different functional models and metrics, such as network-based models [28,30,35] and flow-based models [28,29]. For more details, the reader is referred to the comprehensive literature review in [36].

The risk and the resilience of interdependent systems have been the subject of various recent studies [37–40], also with a focus on vulnerability analysis (e.g. [41–44]).

However, only a few existing works discussed the vulnerability of interdependent railway and power networks. In [45], the authors propose a network-based approach for modeling the vulnerability analysis of a network-of-networks using as an example interdependent transportation, power and telecommunications networks, and including a network-based cascading failures model for the power network. This work relies on these interdependent critical infrastructures to propose an approach for vulnerability analysis of networks-of-networks. A similar approach, also including a network-based cascading failures model for the power network, is proposed in [46] and in [47], where the authors consider a fictitious railway network, based on the topology of the Italian high-voltage grid, connected to the power and telecommunications networks, and they analyze its vulnerability taking into account safety margins and uncertainties. In [48], a modeling framework for the vulnerability analysis of a railway system dependent on electrical and telecommunications networks is proposed, including critical components and locations analysis. A similar approach is also proposed in [2], where the dependency of the railway network on the external power network is addressed but without considering the structure of the external power network. These two last works provide a general framework for the vulnerability analysis of interdependent critical infrastructures. In [49], the authors propose a similar approach, also including the topology of the power grid within the analysis. In [3], a mathematical framework for modeling the vulnerability analysis of a railway network considering its subsystems is presented; this study accounts for the dependency between electricity and transportation, but without considering the external power grid. In [50,51], the resilience of railway networks is studied, also accounting for the dependency on the power subsystems. In [52], the combined effects of an external event (e.g. flooding) on both the infrastructures were considered and analyzed.

These works present some common drawbacks:

- With the exception of [2,48], the structure of the interface (or interconnection) between the interdependent railway and power networks is oversimplified, and the traction power substations that act as an interface between external power networks and railway tracks are overlooked.
- The cascading failures within power networks and their consequences within railway networks are considered only in [45–47]. However, in these works, the cascading failures are modeled with a network-based approach [53,54], and this includes some disadvantages: (i) network-based approaches do not include the modeling of power flows, and they represent a computationally-cheap surrogate model unable to capture the physical features of power flows (ii) network-based and power flow-based models, under specific conditions, exhibit comparable behaviors at global scale but they are inconsistent at local scale [55]. This makes network-based model unsuitable to describe realistically interdependent systems, as the failures propagation between systems is driven by the local topology of the interdependencies.
- The effect of failure in the railway network on the power network is not addressed. In other words, only the dependent behavior is analyzed, and how failures in railway networks affect the power redistribution and the cascading failures dynamics within power networks is not considered.

It is clear that interdependent railway and power networks deserve a more detailed analysis, with a focus on a more realistic consideration of the interconnection interface and cascading failures dynamics, in order to shed some light on the mutual risks of these interconnected systems.

1.3. Contribution

In this work, we propose an approach for the modeling and vulnerability analysis of interdependent railway and power networks, which includes the modeling of the interface between the external power network and the railway network, here referred to as the traction power network, and the evaluation of cascading failures within and between networks. We model the cascading failures using a flow-based approach based on the traditional ORNL-PSerc-Alaska (OPA) model [56,57]. However, we adapt it in order to account for the effect of failures in the railway network on the cascading failures dynamics of the external power network.

In summary, the contributions of this work are the following:

- We propose a modeling approach for the vulnerability analysis of interdependent railway and power networks which accounts for a realistic coupling interface between railway and external power systems and the flow-based characterization of cascading failures within and between the infrastructures.
- We propose a performance indicator for the railway network, called *track power performance TPP*, which accounts for binary functional states and continuous degraded performance due to lack of electricity supply.

- We propose an approach which accounts for the effect of failures within the railway network on the cascading failures dynamics of the external power network.
- We evaluate the effect of cascading failures within and between networks analyzing the negative consequences on the railway and the external power network.

The rest of this work is structured as follows: in Section 2, the proposed modeling framework is presented; in Section 3, the illustrative test systems utilized in this work are described; in Section 4, the results are presented and discussed; in Section 5, final remarks and conclusions are given.

2. Modeling framework

In this section, the modeling and simulation framework is described. Each subsection describes a particular feature of the model:

- Section 2.1 describes the modeling of the topological features of railway tracks and stations (called in the paper railway network), railway traction power substations (called in the paper traction power network) and external power grid (called in the paper external power network) using network theory.
- Section 2.2 shows how to model the topological and functional features of the interdependencies between the three networks.
- Section 2.3 describes how to model and simulate initiating events as removals of elements of networks.
- Section 2.4 describes the modeling of cascading failures, i.e., failure propagation within and between networks.
- Section 2.5 describes how the analysis, in terms of performance indicators and simulation, is performed.

2.1. Network-based model for interdependent railway and power systems

Network science is often used to describe the topology of both railway and power networks, thanks to its capacity to describe complex topologies and interactions with simple mathematical artifacts such as nodes and edges. A network is defined by a graph $\mathbf{G} = (\mathbf{V}, \mathbf{E})$, with $\mathbf{V} = \{v_1, v_2, \dots, v_N\}$ representing the set of N nodes (or vertices) and $\mathbf{E} = \{e_1, e_2, \dots, e_M\}$ the set of M edges. Each edge k is also defined by a tuple $e_k = (v_i, v_j)$, which indicates the two nodes v_i and v_j connected by edge k. In this work, we identify three separate networks: the railway network, defined by the subscript r, the traction power network, defined by the subscript t, and the external power network, defined by the subscript p.

In the railway network $\mathbf{G_r} = (\mathbf{V_r}, \mathbf{E_r})$, nodes represent stations and edges represent railway tracks. Each railway edge represents a direct bi-directional physical connection between two stations.

In the traction power network $\mathbf{G_t} = (\mathbf{V_t}, \mathbf{E_t} = \emptyset)$, nodes represent railway power traction substations. In this work, we assume that each substation is in electrical isolation, and the set of edges is just an empty set. However, for some configurations of electrified railway systems, such as DC-electrified systems, substations are connected with each other. This situation is not considered in this work, but it can be modeled with an appropriate set of edges $\mathbf{E_t}$.

In the external power network $\mathbf{G_p} = (\mathbf{V_p}, \mathbf{E_p})$, nodes represent electrical buses and edges represent transmission lines. Each electrical bus can contain power production or consumption units, here referred to as generators and loads. These elements are defined by the set of generators $\mathbf{g_p} = \{g_1, g_2, \dots, g_{N_g}\}$ and the set of loads $\mathbf{l_p} = \{l_1, l_2, \dots, l_{N_l}\}$.

2.2. Interdependencies

2.2.1. Topological

Electrified railway systems are dependent on external power networks, such as transmission and distribution networks, for the electricity supply. Rolling stocks are usually supplied by the catenary, a dedicated power line which runs parallel to the railway track. Along each railway track, the catenary is divided into sections, each of them supplied by a dedicated traction power substation, which is supplied by the external power network (distribution or transmission network). Each substation also regulates the voltage level before feeding the catenary (e.g. 25 kV for AC railway systems or 1.5 kV for DC railway systems, typically). However, in this study, we do not model this feature.

The traction power network thus depends on the external power network in terms of electricity supply. Specifically, we assume that each substation in the traction power network is dependent on the geographically-closest node in the external power network. We denote this relationship with the interdependency edge $e_{t\leftarrow p}^{i\leftarrow j} = (v_{t,i}, v_{p,j})$, indicating that the node i in the traction power network is dependent on the node j in the external power network.

The traction power network is usually directly connected to the railway catenary to supply electricity to the rolling stocks. The railway network thus depends on the traction power network in terms of electricity supply. Specifically, we assume that each railway track i depends on a set of traction substations $\mathbf{V_t^{r,i}} \subseteq \mathbf{V_t}$, that are responsible for the electricity supply of that specific railway track. We denote this relationship with the set of interdependency edges $\mathbf{E}_{r\leftarrow t}^{i\leftarrow j} = (e_{r,i}, \mathbf{V_t^{r,i}})$, indicating that edge i in the railway network is dependent on the nodes within the subset $\mathbf{V_t^{r,i}} \subseteq \mathbf{V_t}$ of the traction power network. We assume that the substations in $\mathbf{V_t^{r,i}}$ are located equidistantly along the railway track. The first and the last substation, denoted as $v_{t,v_{r,j}}^{r,i}$ and $v_{t,v_{r,k}}^{r,i}$, are located within the railway stations j and k delimiting the railway track i, and they are also responsible for the electricity supply of the stations.

2.2.2. Functional

The external power network is modeled using a traditional DC power flow model (details in Section 2.4 and Appendix A). Each node represents a bus, and it can contain multiple generators, characterized by power generation $P_{g,i}$ and generation capacity $P_{g,i}^{max}$, and multiple loads, characterized by supplied power demand $P_{l,i}$ and requested power demand $P_{l,i}^{max}$. Loads can represent the power demand of the traction substations (and thus of the railway network) or general consumers (households, industries, etc.). The loads which represent traction substations are defined by the subset $\mathbf{l_p^{sub}} \in \mathbf{l_p}$, while the loads which represent general consumers are defined by the subset $\mathbf{l_p^{base}} \in \mathbf{l_p}$. Each edge represents a line, characterized by power flow F_i and flow capacity F_i^{max}.

The traction power network is dependent on the external power network in terms of electricity supply. We assume that each substation i is dependent on the geographically-closest bus of the external power network, which contains a load j, denoted as $l_j^{t,i} \in \mathbf{l_p^{sub}}$, with requested power demand $P_{l_j^{t,i}}^{max}$, which represents the power demand of the traction substation i. Each substation i is characterized by two indicators: the state $S_{t,i}$, which indicates if the station is functional or not, and the load shedding ratio $R_{t,i}$, which defines the fraction of the requested power demand which is supplied to the substation. Both these values are dependent on the power supplied to the corresponding load $l_j^{t,i}$ within the external power network. The value $R_{t,i}$ defines, as shown in Eq. (1), the ratio between the supplied and requested power demand in the corresponding load in the external power network.

$$R_{t,i} = \frac{P_{l_j^{t,i}}}{P_{l_j^{t,i}}^{max}}. \tag{1}$$

The ratio $R_{t,i}$ defines the fraction of requested demand which is supplied, and it also defines the functional state $S_{t,i}$ (1 if functional, 0 otherwise) of the substation i, according to a predefined threshold $T_{t\leftarrow p} = [0, 1]$ and Eq. (2).

$$S_{t,i} = \begin{cases} 1, & \text{if } R_{t,i} \geq T_{t\leftarrow p} \text{ and } 0 < T_{t\leftarrow p} \leq 1 \\ 0, & \text{if } R_{t,i} < T_{t\leftarrow p} \text{ and } 0 < T_{t\leftarrow p} \leq 1 \\ 0, & \text{if } R_{t,i} = T_{t\leftarrow p} \text{ and } T_{t\leftarrow p} = 0 \end{cases} \tag{2}$$

When $T_{t\leftarrow p} = 1$, it represents a "zero-tolerance" situation, where a substation is functional only as long as the entire requested power demand is satisfied. When $T_{t\leftarrow p} = 0$, it represents the opposite situation, where a substation is functional as long as some electricity is provided. The case $0 < T_{t\leftarrow p} < 1$ represents situations in-between, where a substation is considered functional as long as at least a specific fraction of the requested power demand is satisfied.

The railway network functionality depends on the traction power network. As explained in the previous section, each railway track i depends on a subset $\mathbf{V_t^{r,i}} \subseteq \mathbf{V_t}$ of traction substations. Each railway track i is defined by its functional state $S_{r,i}$ and its *track power performance* TPP_i. We assume that each railway i is functional only if all the substations in $\mathbf{V_t^{r,i}}$ are functional, as defined in Eq. (3):

$$S_{r,i} = \prod_{j \in \mathbf{V_t^{r,i}}} S_{t,j} \tag{3}$$

In fact, we assume that if one substation is not functional, the corresponding section of the railway track is not supplied with the necessary electricity, thus interrupting the continuity of the railway track. The track power performance TPP_i of each railway track i is dependent on the state $S_{r,i}$ and the ratio $R_{t,j}$ of each substation j within the subset $\mathbf{V_t^{r,i}}$, as defined in Eq. (4):

$$TPP_i = \frac{S_{r,i}}{N_{\mathbf{V_t^{r,i}}}} \sum_{j \in \mathbf{V_t^{r,i}}} R_{t,j} \tag{4}$$

where $N_{\mathbf{V_t^{r,i}}}$ is the number of substations within the subset $\mathbf{V_t^{r,i}}$. As it can be clearly seen in Eq. (4), the TPP_i of each railway track i is equal to the ratio between total supplied and requested power demand within the substations in $\mathbf{V_t^{r,i}}$ if the track is functional, or 0 otherwise. This performance indicator allows to describe the binary state of the track (functional/not functional), as well as the degraded performance due to lack of electricity supply.

In this study, we account also for the impact of failures in the railway network on the external power network. We assume that for any failed railway track i, the corresponding traction substations in $\mathbf{V_t^{r,i}}$ do not absorb electricity from the external power network. Therefore, for each substation j in $\mathbf{V_t^{r,i}}$, the corresponding load $l_j^{t,i}$ in the external power network has a reduced requested power demand $P_{l_j^{t,i}}^{max'}$. Specifically, we assume that, if the substation is located within a railway station, the requested power demand is reduced; otherwise, the requested power demand is 0. This behavior can impact the power redistribution and the cascading failures dynamics of the external power network. More details are available in Section 2.4.

2.3. Modeling initiating events

Initiating events represent single and multiple failures which might affect a system during normal operation or external strains. Vulnerability analysis investigates the system's response to different initiating events, often modeled as removals of an increasing fraction of components from the network [2,3]. The set of components to be removed is defined as $\mathbf{N_f} = \{n_{f,1}, n_{f,2}, \ldots, n_{f,N_{fail}}\}$. In this work, we assume that only edges within the external power network can be removed from the network, and the number of components N_{fail} which constitute the set $\mathbf{N_f}$ depends on the fraction f according to Eq. (5):

$$N_{fail} = \lfloor f \cdot M_p \rfloor \tag{5}$$

where M_p is the number of edges in the external power network and $\lfloor \cdot \rfloor$ is the floor function, which returns the greatest integer less than or equal to the argument of the function. When an edge fails, it is simply removed from the network. The fraction f can be interpreted as the magnitude of the initiating event. The elements selection of the set $\mathbf{N_f}$ depends on the type of initiating event. In this work, we consider only random removals as removal strategies for initiating events. Other strategies, such as spatially-localized and targeted removals, are not considered in this work.

Random removals represent a wide range of initiating events (human errors, structural defects, random sabotages, etc.), and they are useful for understanding the robustness of a network under different magnitude of strains which might impact multiple locations of the network. Given the set of network edges $\mathbf{E_p}$, we assume that each edge, has the same probability p_{ie} of being selected as part of the initiating event and removed, computed according to Eq. (6):

$$p_{ie} = \frac{1}{M_p}. \tag{6}$$

The component selection is made according to Eq. (7):

$$p_{ie} \cdot (i-1) \leq r < p_{ie} \cdot i \tag{7}$$

where $0 < r < 1$ is a random number and i is the index of the edge e_i which satisfies the above relationship, and the procedure shown in the following steps:

1. Set $\mathbf{N_f} = \emptyset$ and compute N_{fail} according to Eq. (5).
2. Generate a random number $0 < r < 1$.
3. Identify edge e_i which satisfies Eq. (7).
4. If e_i is not in $\mathbf{N_f}$, add e_i to $\mathbf{N_f}$ and go to step 5; otherwise, return to step 2.
5. If $\mathbf{N_f}$ contains N_{fail} elements, stop the selection; otherwise, return to step 2.

2.4. Modeling cascading failures

Failures and subsequent removals of failed elements can trigger a cascading failures process, which is defined as a "kind of failure in a system comprising interconnected parts, in which the failure of a part can trigger the failure of successive parts" [58]. When failures propagate within different infrastructures, it is also referred to as cascading effect, and it is often characterized by an increasing severity of damages [59]. In this work, we thus refer as cascading failures to the process of failure propagation within and between networks, triggered by an initiating event. Simulating cascading failures is an iterative simulation process. At each step, the network topology is updated, as failed components are removed. The general procedure for cascading failures simulation in single or interdependent networks includes the following steps:

1. Initialize network $\mathbf{G} = (\mathbf{V}, \mathbf{E})$.
2. Initialize initiating event $\mathbf{N_f}$.
3. Remove failed elements.
4. Update network topology.
5. Check conditions for cascading failures.
6. If there is any new failure, return to step 3; otherwise, stop the simulation.

Cascading failures have been extensively analyzed for power networks, as several approaches are available in the existing literature [60]. However, these models often focus solely on power networks, without considering interdependencies with other infrastructures. These models are not suited for evaluating the effect of the interdependencies on the cascading failures dynamics, and they allow to analyze only the effect of failures in the power network on other systems. For example, this is the approach taken in [45–47], where the cascading failures simulation in the power network is decoupled from the vulnerability analysis of the railway network and follows this general steps:

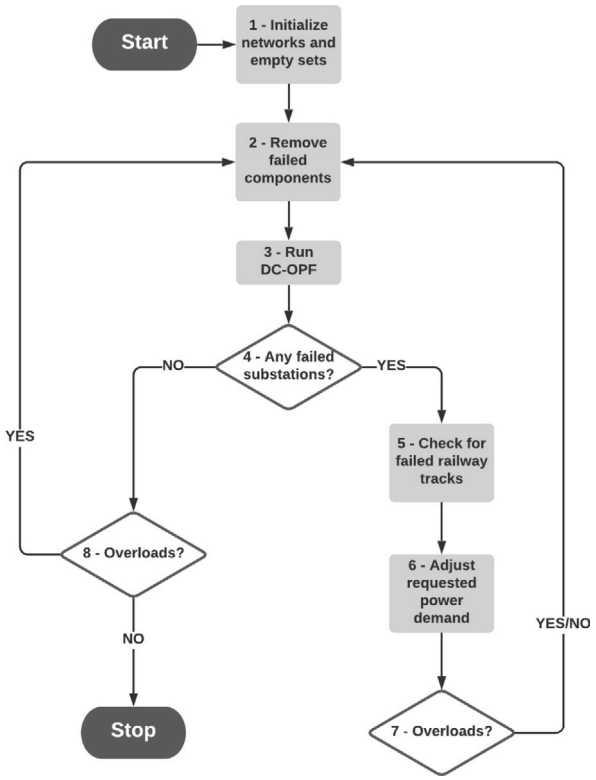

Fig. 1. Flowchart of the cascading failures algorithm for interdependent railway and power networks.

- Select initiating event.
- Run a complete cascading failures simulation with the external power network.
- From the output of the cascading failures simulation, compute the negative consequences on the railway network.

In this work, we propose a flow-based cascading failures model, based on the traditional ORNL-PSerc-Alaska (OPA) model [56,57] (see Appendix B for details), which accounts for the impact of failures in the railway network on the cascading failures dynamics of the external power network. The flowchart of the algorithm is shown in Fig. 1.

The algorithm comprises eight iterative steps, and it is based on the assumption that, when a railway track fails, the corresponding traction substations stop requesting electricity from the external power network. This impacts the requested power demand of the corresponding loads in the external power network, affecting thus the cascading failures dynamics of the external power network. After the selection of the edges to remove as initiating event (as explained in Section 2.3), the cascading failures algorithm initializes the networks and the empty sets $\mathbf{N_F^{dir}}$ and $\mathbf{N_F^{ind}}$ in Step 1, and removes the failed edges from the networks in Step 2.

In Step 3, the power supplied to each load of the external power network is computed with a DC Optimal Power Flow (DC-OPF), an optimization procedure shown in Eqs. (8)–(12):

$$\min_{\mathbf{P_g},\mathbf{P_l}} \sum_{i=1}^{N_g} P_{g,i} - W \sum_{j=1}^{N_l} |P_{l,j}| \tag{8}$$

subject to:

$$\sum_{i=1}^{N_g} P_{g,i} - \sum_{j=1}^{N_l} |P_{l,j}| = 0 \tag{9}$$

$$0 \le P_{g,i} \le P_{g,i}^{max} \tag{10}$$

$$-P_{l,j}^{max} \le P_{l,j} \le 0 \tag{11}$$

$$-F_k^{max} \le F_k \le F_k^{max}. \tag{12}$$

The objective function in (8) represents the cost to minimize, while Eqs. (9)–(12) represent the constraints. The first term of Eq. (8) represents the power production cost, which is assigned a unitary value per unit of power $P_{g,i}$ produced in each generator i. The second term represents the negative cost associated to the power $P_{l,i}$ supplied at each load i. The penalty constant W, here assumed to be equal to 100, ensures the minimization of load shedding when possible. Constraint (9) describes the power balance between power produced and consumed. Constraints (10) and (11) represent the ranges of power generation and supplied power demand, respectively. Constraint (12) represents the limit for lines power flow. The computation of the power flow is subjected to the DC power flow assumption (see Appendix A).

After the computation of the power supplied to each load in the external power network, in Step 4 we check for failed substations in the traction power network, using the relations defined in Eqs. (1) and (2).

If new failed substations are present, the algorithm proceeds with Step 5, where we check for failed railway tracks, according to Eq. (3); otherwise, it proceeds with Step 8, where we check for overloaded lines. In Step 5, if a railway track i is failed, each substation j in $\mathbf{V_t^{r,i}}$ is added to the set $\mathbf{N_f^{dir}}$ if $S_{t,j} = 0$, or to $\mathbf{N_f^{ind}}$ if $S_{t,j} = 1$. If $S_{t,j} = 0$, it means the substation j has failed directly (superscript dir) due to lack of electricity from the external power network; if $S_{t,j} = 1$, it means the substation j has failed indirectly (superscript ind) due to the direct failure of another substation k within the same subset $\mathbf{V_t^{r,i}}$. These two sets will be used to evaluate the vulnerability of the external power network (more details in Section 2.5).

The algorithm proceeds with Step 6, where the requested power demand in each load in the external power network corresponding to a substation in the traction power network is adjusted, taking into account the failures of railway tracks. For every failed railway track i, the new requested power demand $P_{l_j^{t,k}}^{max\prime}$ for each load j corresponding to a substation k within the subset $\mathbf{V_t^{r,i}}$ is set to 0, with the exception of the substations corresponding to the railway stations w and q delimiting the railway track, $v_{t,v_{r,w}}^{r,i}$ and $v_{t,v_{r,q}}^{r,i}$. In these substations, as they might supply multiple railway tracks, the requested power demand is simply reduced according to Eq. (13):

$$P_{l_j^{t,k}}^{max\prime} = P_{l_j^{t,k}}^{max} \cdot \frac{\sum_{i \in \mathbf{E_r^{v_{r,w}}}} S_{r,i}}{d(v_{r,w})} \tag{13}$$

where $\mathbf{E_r^{v_{r,w}}} \subseteq \mathbf{E_r}$ is the subset of railway tracks connected to station $v_{r,w}$, $S_{r,i}$ is the state of the railway track i and $d(v_{r,w})$ is the degree of station $v_{r,w}$ (number of edges connected to node $v_{r,w}$).

Following this procedure, we proceed with Step 7, where we check for overloaded lines, similarly to Step 8. As in the traditional OPA model, a line k is considered overloaded when its power flow F_k is within 1% of the maximum capacity of the line F_k^{max}. An overloaded line is assumed to trip, and thus fail, with probability p_{ol}. In this work, aiming for a conservative worst-case analysis, we assume $p_{ol} = 1$ [55]. The state of transmission line k is thus defined by Eq. (14):

$$S_{p,k} = \begin{cases} 1, & \text{if } \frac{F_k}{F_k^{max}} < 0.99 \\ 0, & \text{otherwise} \end{cases} \tag{14}$$

where $S_{p,k}$ is the state of line k. From Step 7, the algorithm returns to Step 2, where the new failed components are removed from the networks. New failed components comprises railway tracks i with at least one failed substation within the set $\mathbf{V_t^{r,i}}$ and overloaded power lines (if any). From Step 8, the algorithm returns to Step 2 if overloaded lines are present; otherwise, it is stopped.

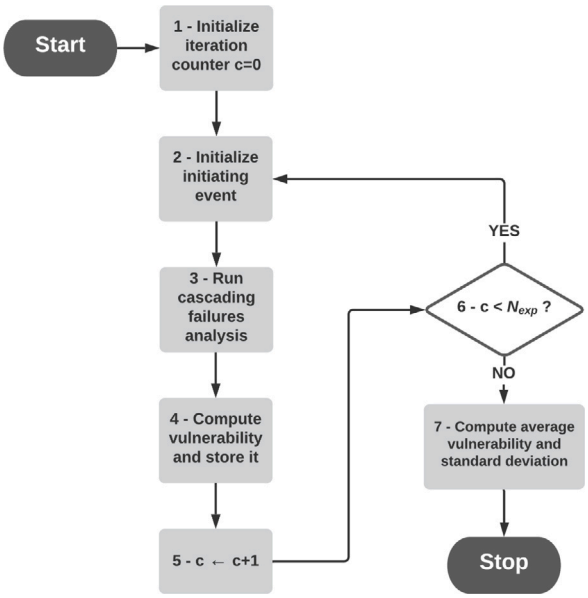

Fig. 2. Flowchart of the vulnerability analysis algorithm for each fraction of removals.

The outputs of the cascading failures algorithm are the new values of requested and supplied power demand for each load i in the external power network, respectively $P_{l,i}^{max'}$ and $P'_{l,i}$, the new railway network topology $\mathbf{G'_r} = (\mathbf{V'_r}, \mathbf{E'_r})$, and the sets of directly/indirectly failed traction substations $\mathbf{N_F^{dir}}$ and $\mathbf{N_f^{ind}}$.

2.5. Vulnerability metrics and analysis

Vulnerability analysis aims at estimating the negative consequences which arise in a system given an imposed strain [29]. Mathematically, the negative consequences on a system can be defined as the relative change of a specific performance indicator after a disruptive event, and the vulnerability V can be generally expressed as in Eq. (15):

$$V = \frac{PI - PI'}{PI} \tag{15}$$

where PI and PI' represent respectively the performance indicator before and after the disruptive event.

In order to compute an average vulnerability index, multiple iterations are needed, because different fractions of removals and different combinations of components in $\mathbf{N_f}$ leads to different vulnerability values. Average values $\bar{V}$ and 95% confidence intervals CI_{95} for each fraction of removals are computed respectively with Eqs. (16) and (17):

$$\bar{V} = \frac{\sum_{i=1}^{N_{exp}} V_i}{N_{exp}} \tag{16}$$

$$CI_{95} = \bar{V} \pm Z \cdot \frac{\hat{\sigma}}{\sqrt{N_{exp}}} \tag{17}$$

where N_{exp} is the number of experiments per fraction of removals, Z is the 95% confidence interval constant, equal to 1.96, and $\hat{\sigma}$ is the estimated standard deviation. The algorithm for computing the vulnerability indexes for each fraction of removals is shown in Fig. 2.

After the initialization of the iteration counter $c = 0$ in Step 1 and the initiating event in Step 2 (as explained in Section 2.3), the algorithm runs the cascading failures simulation explained in Section 2.4 (corresponding to the algorithm in Fig. 1). From the outputs of the cascading failures simulation, in Step 4 we compute the vulnerability indexes V_i and store them. In Step 5, we increase the iteration counter by 1, and in Step 6 we check if it is equal to the maximum number of iterations N_{exp}. If $c = N_{exp}$, in Step 7, we compute the average vulnerability indexes and the standard deviations with Eqs. (16) and (17) and we stop the algorithm; otherwise, the algorithm returns to Step 2. This procedure must be performed for every fraction of removals.

The performance indicator is selected according to the type of system under analysis. For the railway network, we rely on two performance indicators: the accessibility A_r and the *railway power performance* RPP. The accessibility is defined in Eq. (18):

$$A_r = \frac{1}{N_r} \sum_{i=1}^{N_r} \frac{n_a^i}{N_r - 1} \tag{18}$$

where N_r is the total number of stations and n_a^i is number of stations accessible from the station i. It can be interpreted as the average fraction of stations accessible from (or connected to) each other [20].

The *railway power performance* RPP is defined in Eq. (19):

$$RPP = \frac{1}{M_r} \sum_{i=1}^{M_r} TPP_i \tag{19}$$

where M_r is the number of railway tracks and TPP_i is the track power performance of railway track i, defined in Section 2.2.2. This indicator represents the average performance of the railway network tracks, where the performance of each track is equal to the average ratio $R_{l,i}$ of the supporting substation in $\mathbf{V_t^{r,i}}$ if the track is functional, or 0 if the track is not functional. This formulation allows to take into account binary states (functional/not functional), as well as degraded performance due to partial power supply.

For the power network, we utilize as performance indicator the *Demand Not Supplied* (DNS) and the *Fraction of Demand Not Supplied* (FDNS), shown in Eqs. (20) and (21).

$$DNS = \sum_{i \in \mathbf{l_p^{base}}} P_{l,i}^{max} + \sum_{i \in \mathbf{l_p^{N_F^{dir}}}} P_{l,i}^{max} - \sum_{i \in \mathbf{l_p^{base}}} P'_{l,i} - \sum_{i \in \mathbf{l_p^{N_F^{dir}}}} P'_{l,i} \tag{20}$$

$$FDNS = 1 - \frac{\sum_{i \in \mathbf{l_p^{base}}} P'_{l,i} + \sum_{i \in \mathbf{l_p^{N_F^{dir}}}} P'_{l,i}}{\sum_{i \in \mathbf{l_p^{base}}} P_{l,i}^{max} + \sum_{i \in \mathbf{l_p^{N_F^{dir}}}} P_{l,i}^{max}} \tag{21}$$

where $P_{l,i}^{max}$ represents the baseline requested power demand in each load, $P'_{l,i}$ represents the supplied power demand in each load after cascading failures simulation, $\mathbf{l_p^{base}}$ is the subset of base loads (loads which does not represent any traction substations), and $\mathbf{l_p^{N_F^{dir}}} \in \mathbf{l_p^{sub}}$ is the subset of loads representing traction substations directly failed during the cascading failures simulation, as explained in the previous section. We do not account for the loads corresponding to traction substations indirectly failed. In fact, these substations are failed due to failures of other substations, and the power not supplied in these substations is not accounted for in the computation of DNS and $FDNS$.

3. Illustrative test systems

The developed model is applied to investigate the vulnerability of an electrified railway network. For this, a simplified version of a British high-speed railway system, based on a proposition made in [61], is considered. The system comprises a railway network powered by an external power network through a traction power network. The railway network consists of 16 stations connected by 21 railways. The external power network is based on the Great Britain reduced power network [62], which represents a high-voltage transmission system (it originally represents a 400/275 kV system). It consists of 29 electrical buses, containing 29 base loads (one in each bus), representing the power demand of general consumers (households, industries, etc.), 85 loads corresponding to the traction substations and 66 generators. Buses are connected by 99 lines, most of them in redundant double circuit configuration. We assume the maximum power generation capacity of the external power network to be slightly higher than the requested power demand, respectively 20.59 GW and 20 GW. The generator

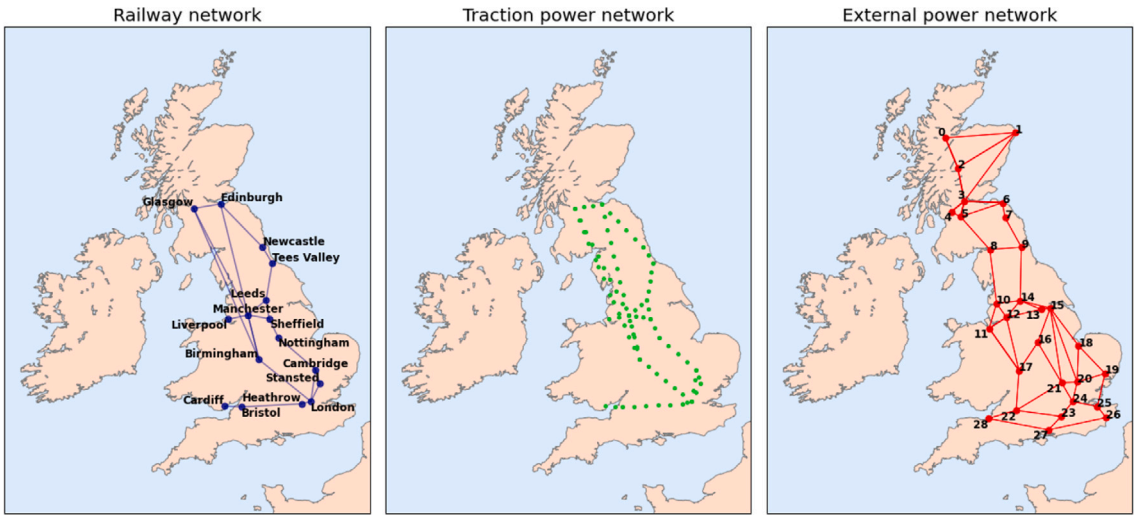

Fig. 3. Geographical representation of the railway network (blue), the traction power network (green) and external power network (red).

power capacities range from 0.003 GW to 1.958 GW. The base load power demands range from 0.033 GW to 2.724 GW. The capacities of transmission lines range from 0.046 GW to 2.471 GW. In nominal conditions, with all the elements functional, the loading percentage of lines in the external power network ranges from 1.22% to 67.82%, with an average value of 21.66%. For other electrical parameters of the external power network, please refer to the data available in [62].

The railway and external power networks are coupled by the traction power network, which consists in 85 substations, in electrical isolation, distributed equidistantly (on average one every 35 km) along the corresponding railway tracks. The requested power demand for each substation (and thus the requested power demand of each corresponding load in the external power network) is assumed to be 200 MW if the substation is located within a railway station, or 12 MW otherwise. The total requested power demand of the traction power network (and thus of the railway network) is 4028 MW, which corresponds to 20% of the total requested power demand of the external power network. We define this fraction as the *coupling strength* between the railway and the external power network. In this work, we do not perform a sensitivity analysis on the coupling strength, but it might be an interesting input for future works. As mentioned before, it is assumed that the electricity is supplied to each substation from the closest electrical bus within the external power network, which represents the power demand of the corresponding station, as denoted by the set of interdependency edges $\mathbf{E_{t\leftarrow p}}$. The geographical and network-based representations of the networks are shown in Figs. 3 and 4, respectively.

As initiating events, we remove fractions of edges from the external power network. The fraction of removals ranges from 0% to 100%, with steps of 10%. These removals can represent a wide range of disruptive events which might cause multiple failures of power transmission lines: intentional random sabotages, hidden failures due to defective relays, natural events (e.g. extreme storms and/or winds), random failures (e.g. falling trees), human mistakes (e.g. incorrect maintenance), or other events. We analyze the effect in terms of vulnerability of the railway network and external power network for three different tolerance threshold $T_{t\leftarrow p}$ values (0.0, 0.5, 1.0).

The vulnerability of the railway network is assessed in terms of accessibility A_r and railway power performance RPP. In order to highlight the importance of cascading failures analysis, the vulnerability of railway networks is assessed with and without cascading failures simulation within the external power network. The analysis without cascading failures in external power networks is performed with the algorithm is Fig. 1 and by setting the tripping probability of overloaded lines p_{ol} to 0. The vulnerability of the external power network is performed only considering cascading failures within the analysis.

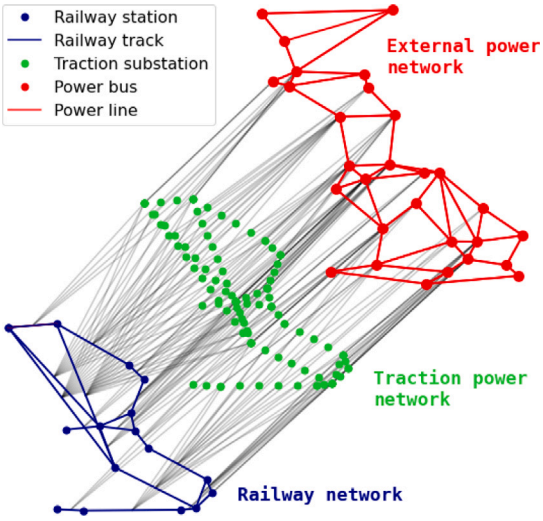

Fig. 4. Network-based representation of the railway network (blue), the traction power network (green) and external power network (red).

The vulnerability of the external power network is assessed in terms of demand not supplied DNS and fraction of demand not supplied $FDNS$.

For every fraction of removals, we compute average vulnerability indexes and standard deviations using a number of experiments N_{exp} equal to 1000, as it allows to obtain 95% confidence intervals small enough to perform a detailed and consistent analysis.

Finally, the external power network is implemented within Pandapower [63] and the power flow computations are performed using the PowerModels.jl API of Pandapower [64].

4. Results and discussion

4.1. The impact on the railway network

The impact of random removals of edges from the external power network on the railway network is assessed by the average loss of accessibility A_r and average loss of railway power performance RPP for three $T_{t\leftarrow p}$ values (0.0, 0.5 and 1.0). The results with and without cascading failures simulation within the external power network, denoted respectively as (C) and (NC), and corresponding to failure

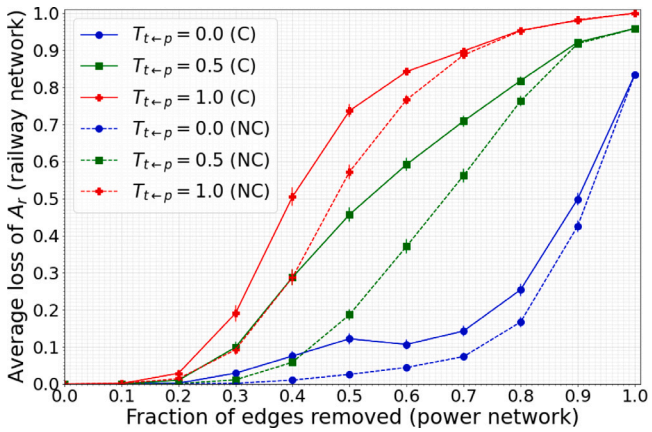

Fig. 5. Impact of removals in the power network on the railway network in terms of average loss of A_r for $T_{t\leftarrow p}$ equal to 0.0, 0.5 and 1.0. The results are evaluated including cascading failures within the external power network in the analysis ($p_{ol} = 1$), denoted as (C) in the legend, and not including them ($p_{ol} = 0$), denoted as (NC) in the legend. 95% confidence intervals are shown.

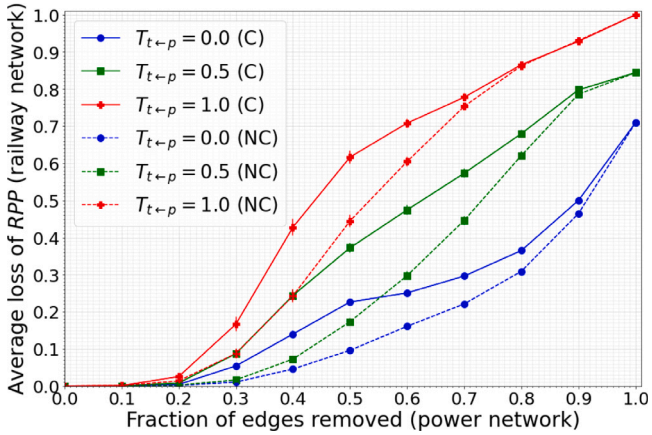

Fig. 6. Impact of removals in the power network on the railway network in terms of average loss of RPP for $T_{t\leftarrow p}$ equal to 0.0, 0.5 and 1.0. The results are evaluated including cascading failures within the external power network in the analysis ($p_{ol} = 1$), denoted as (C) in the legend, and not including them, ($p_{ol} = 0$), denoted as (NC) in the legend. 95% confidence intervals are shown.

probabilities for overloaded lines of $p_{ol} = 1$ and $p_{ol} = 0$, respectively, are presented in Figs. 5 and 6.

As it can be clearly seen, the impact on the railway network follows different patterns for different $T_{t\leftarrow p}$ values. Intuitively, the lower is the threshold $T_{t\leftarrow p}$, the less vulnerable is the railway network, as it is more tolerant to lack of electricity supply. Moreover, as expected, cascading failures considerably increase the impact of removals in the power network on the railway network. These considerations are valid for both the topological accessibility A_r and the railway power performance RPP. A direct measure of the disruption is the area below the vulnerability curves in Figs. 5 and 6. For comparison, the values of the areas below the curves for $T_{t\leftarrow p} = 0.5$ and $T_{t\leftarrow p} = 0.0$ are computed and normalized with the area below the curve for $T_{t\leftarrow p} = 1.0$ with cascading failures, for both A_r in Fig. 5 and RPP in Fig. 6. The results for A_r and RPP are shown in Tables 1 and 2, respectively.

From the results in Figs. 5 and 6 and Tables 1 and 2, we can draw two general considerations:

- The severity of disruption increases with the increase of the threshold $T_{t\leftarrow p}$.
- The inclusion of cascading failures within the analysis increases the severity of disruption within the railway network.

Table 1
Vulnerability curve areas for the accessibility A_r.

$T_{t\leftarrow p}$	Curve (NC) area	Curve (C) area
0.0	0.21	0.29
0.5	0.59	0.78
1.0	0.90	1.0

Table 2
Vulnerability curve areas for the railway power performance RPP.

$T_{t\leftarrow p}$	Curve (NC) area	Curve (C) area
0.0	0.33	0.44
0.5	0.57	0.73
1.0	0.89	1.0

These considerations are intuitive and in line with the expected results. However, it is important to quantify and discuss the difference in results within the two performance indicators when using different $T_{t\leftarrow p}$ values and considering/excluding cascading failures within the analysis.

As it can be clearly seen in Figs. 5 and 6 and Tables 1 and 2, higher tolerance thresholds $T_{t\leftarrow p}$ increase the vulnerability of the railway network. For example, in the case of accessibility A_r, the area below the continuous blue curve in Fig. 5, which corresponds to the case $T_{t\leftarrow p} = 0.0$ with cascading failures, is a fraction 0.29 of the continuous red curve within the same figure, which corresponds to the case $T_{t\leftarrow p} = 1.0$ with cascading failures. This is due to the fact that, with low $T_{t\leftarrow p}$ values within the substations of the traction power network, the railway network is more tolerant to the lack of electricity and the negative impact of disruption within the external power network on the railway network is less pronounced.

However, we can notice some differences when comparing the results for the accessibility A_r in Fig. 5 and the railway power performance RPP in Fig. 6. Firstly, the red curves ($T_{t\leftarrow p} = 1.0$) and the green curves ($T_{t\leftarrow p} = 0.5$) for A_r in Fig. 5 have higher values if compared to the same curves for RPP in Fig. 6. Secondly, the blue curves ($T_{t\leftarrow p} = 0.0$) are higher for RPP, except for a total removal of edges in the external power network (fraction of removal 1.0). As a result, the railway network is more vulnerable in terms of accessibility A_r than railway power performance RPP for high threshold values ($T_{t\leftarrow p} \geq 0.5$); on the contrary, for low threshold values (e.g. $T_{t\leftarrow p} = 0.0$), the railway network is more vulnerable in terms of RPP.

These behaviors are caused by the intrinsic differences within the two performance indicators. Firstly, the accessibility is binary-based (railway tracks are either functional or not), while the railway power performance is hybrid binary-continuous (it accounts for functional state and degraded performance). As a result, for $T_{t\leftarrow p} = 0.0$, the railway network is more vulnerable in terms of RPP. In fact, with $T_{t\leftarrow p} = 0.0$, the railway network is more tolerant in terms of lack of electricity supply and the number of not functional railway tracks is smaller if compared to higher $T_{t\leftarrow p}$ values. However, functional railways can have degraded performance: this is captured by RPP but not by A_r. As a result, for low $T_{t\leftarrow p}$, the losses in terms RPP are higher than the losses in terms of A_r.

Secondly, the failures of railways can be more impactful in terms of A_r than RPP. In fact, failed railway tracks might impact the accessibility from/to several railway stations, leading to considerable losses in terms of A_r. As a result, for high threshold values (e.g. $T_{t\leftarrow p} \geq 0.5$), the railway network is more vulnerable in terms of A_r, as the number of failed railway tracks is higher than $T_{t\leftarrow p} = 0.0$.

Both the performance indicators are strongly affected by the cascading failures. In fact, it is clearly visible in Figs. 5 and 6 that cascading failures lead to more negative consequences in terms of losses. The difference between considering or excluding cascading failures from the analysis can be quantified with Tables 1 and 2. In fact, by comparing the normalized areas for the curves with and without cascading failures

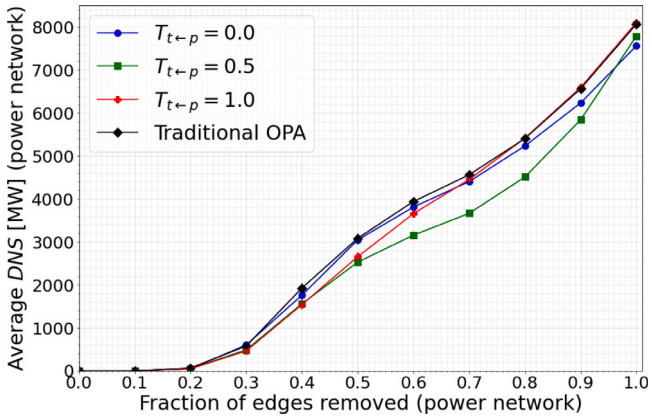

Fig. 7. Megawatts of demand not supplied within the external power network for different values of $T_{t\leftarrow p}$ and a traditional OPA model. 95% confidence intervals are shown.

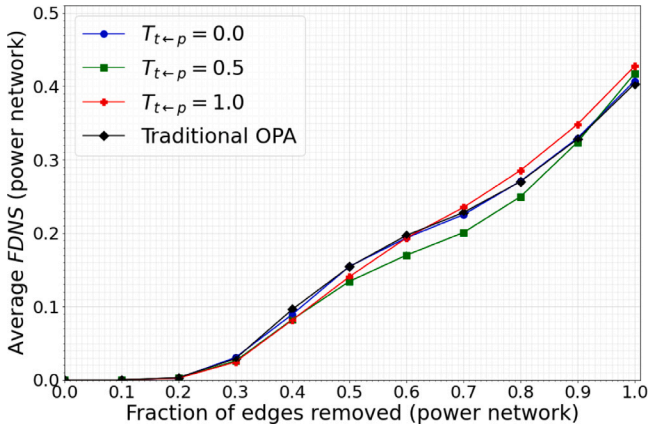

Fig. 8. Fraction of demand not supplied within the external power network for different values of $T_{t\leftarrow p}$ and a traditional OPA model. 95% confidence intervals are shown.

for the same $T_{t\leftarrow p}$, it can be seen how the cascading failures lead to greater disruption, in the order of magnitude of 10%–20% of additional losses, depending on the case.

In conclusion, failures in the power network can lead to considerable disruption within the railway network due to lack of electricity. We modeled the tolerance to lack of electricity of the railway network using the parameter $T_{t\leftarrow p}$. Intuitively, the more tolerant is the railway network, and the smaller is the disruption. In addition, we show that cascading failures can considerably increase the negative consequences on the railway network. We quantified the cascading failures effect with the normalized areas in Tables 1 and 2, and we conclude that cascading failures should be included within the vulnerability analysis for a realistic estimation of possible negative consequences on the railway network.

4.2. The impact on the power network

The dependency of the railway network on the external power network can affect the behavior of the external power network itself. In fact, as railway tracks fail, the corresponding traction substations stop absorbing power, leading to a reduction of the requested power demand within the external power network. This can impact the power redistribution and the cascading failures dynamics.

As in the previous case, a specific fraction of edges (from 0% to 100%, with steps of 10%), is randomly removed from the external power network. Through the model described in the previous section, the impact on the external power network is assessed in terms of MW of Demand Not Supplied (DNS), shown in Fig. 7, and Fraction of Demand Not Supplied ($FDNS$), shown in Fig. 8. The results are compared with the ones of a traditional OPA model (see Appendix B), which does not account for the effect of railway track failures on the external power network and the cascading failures dynamics.

The key indicator to analyze is the DNS, which provides the information on the quantity of requested power demand, in absolute value, which is not satisfied. The results are shown in Fig. 7. Two main considerations can be drawn:

- The application of a traditional OPA model leads to higher results in terms of DNS, as it can be clearly seen by comparing the black curve in Fig. 7 to the other ones.
- The effect of different $T_{t\leftarrow p}$ values is not straightforward like in the previous case.

The results in terms of demand not supplied are higher when we use the traditional OPA model. We quantify the difference using the area under the curves in Fig. 7, normalized using the area under the black curve, corresponding to the traditional OPA case. As it can be clearly seen in Table 3, the areas under the $T_{t\leftarrow p}$ curve corresponding to our model are smaller than the area under the traditional OPA curve. This trend can be characterized as clear evidence of antifragility in interdependent railway and power networks, where "*fragility is related to how a system suffers from the variability of its environment beyond a certain preset threshold [...], while antifragility refers to when it benefits from this variability*" [65]. In our case, we can extend this definition to a system-of-systems framework, and we can generalize this behavior with the following statement: *given a system-of-systems of interdependent infrastructures, sharing a unidirectional interdependency of the type supplier–consumer, stressors, strains or disruptions in the consumer system can decrease the vulnerability of the supplier system, as the margin between supply generation capacity and total consumer demand increases.*

Table 3
Vulnerability curve areas for the external power network DNS, normalized with area under the curve corresponding to the traditional OPA model.

$T_{t\leftarrow p}$	Area
0.0	0.96
0.5	0.85
1.0	0.96
OPA	1.0

However, regarding the different $T_{t\leftarrow p}$ values, the behaviors are different and less intuitive than the previous case (impact on railway network), where higher $T_{t\leftarrow p}$ corresponds to higher vulnerability.

The three $T_{t\leftarrow p}$ curves follows different patterns. The blue curve, corresponding to $T_{t\leftarrow p} = 0.0$, presents always values lower or equal to the black curve, corresponding to the traditional OPA case. For fraction of removals lower than 0.7, it presents DNS values higher than the other $T_{t\leftarrow p}$ curves.

The green curve, corresponding to $T_{t\leftarrow p} = 0.5$, presents the lowest DNS values when compared to the other curves. This is also clearly visible in Table 3, where the curve $T_{t\leftarrow p} = 0.5$ has the smallest normalized area.

The red curve, corresponding to $T_{t\leftarrow p} = 1.0$, presents DNS values lower than the traditional OPA and $T_{t\leftarrow p} = 0.0$ for fraction of removals lower than 0.7; for higher fractions, it is coincident with the traditional OPA curve.

These patterns are highly dependent on the failure modes of the traction substations: direct and indirect (defined in Section 2.4). Substations which fail directly due to lack of electricity tend to increase the DNS to values closer to a traditional OPA model; substations which fail indirectly tend to decrease the DNS to values lower than a traditional OPA model. The fractions of functional and directly/indirectly failed

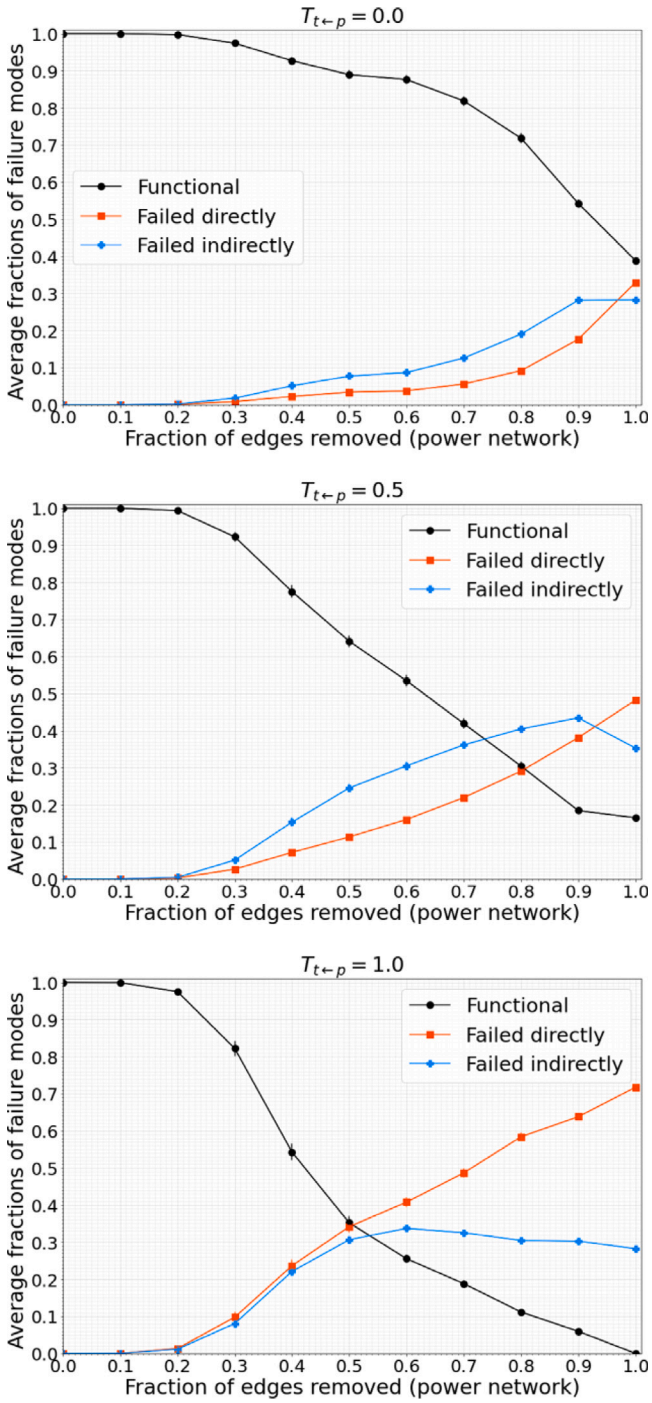

Fig. 9. Fraction of substations functional and directly/indirectly failed. 95% confidence intervals are shown.

substations play a key role in determining the patterns in Fig. 7. The average fractions of functional and directly/indirectly failed substations for different $T_{t\leftarrow p}$ values and fractions of removals are shown in Fig. 9.

For $T_{t\leftarrow p} = 0.0$, the total fraction of failed substations increases slowly and, except for a fraction of removals equal to 1.0, it is always lower than the fraction of functional substations. This explains why for low fractions of removals the DNS values are similar to a traditional OPA model. In addition, the fraction of indirectly failed substations increases faster than the fraction of directly failed substations. This explains why for high fractions of removals the traditional OPA tends to lead to larger disruption.

Table 4
Vulnerability curve areas for the external power network $FDNS$, normalized with area under the curve corresponding to the traditional OPA model.

$T_{t\leftarrow p}$	Area
0.0	0.99
0.5	0.93
1.0	1.01
OPA	1.0

For $T_{t\leftarrow p} = 0.5$, the patterns are similar to $T_{t\leftarrow p} = 0.0$ but with a faster rate. The increase of fractions of failed substations, as well as the decrease of the fraction of functional substations, occurs faster with the increase of the removals if compared to $T_{t\leftarrow p} = 0.0$. This explains the pattern in Fig. 7. As the number of indirectly failed substations increases sharply, the requested power demand in the external power network decreases. In addition, the fraction of indirectly failed substations is always higher than the directly failed ones (except for fraction of removals equal to 1), and this contributes to decrease the DNS, since the power not supplied to indirectly failed substations is not taken into account.

For $T_{t\leftarrow p} = 1.0$, the fraction of directly failed substations is always greater or equal to the fraction of indirectly failed substations, and they increase at a faster rate than the previous cases. As a consequence, for fractions of removals greater than 0.5, the DNS for $T_{t\leftarrow p} = 1.0$ and for the traditional OPA model are coincident.

While in the case of DNS the traditional OPA curve is always greater or equal to the $T_{t\leftarrow p}$ curves, the situation is different for the case of $FDNS$. As it is visible in Fig. 8, the $T_{t\leftarrow p}$ curves can be higher than the traditional OPA curve. We can appreciate the correlation with the fractions of functional/failed substations in Fig. 9. For $T_{t\leftarrow p} = 0.0$ and $T_{t\leftarrow p} = 0.5$, we can notice a perfect correlation between the $FDNS$ and the fractions of directly/indirectly failed substations. In fact, the $T_{t\leftarrow p}$ curves in Fig. 8 are greater than the traditional OPA curve for a fraction of removals equal to 1; similarly, in Fig. 9, the fraction of directly failed substations is higher than the fraction of indirectly failed substations only for a fraction of removals equal to 1.0, for both $T_{t\leftarrow p} = 0.0$ and $T_{t\leftarrow p} = 0.5$. For $T_{t\leftarrow p} = 1.0$, the correlation is not perfect, as the $T_{t\leftarrow p}$ curve in Fig. 8 is higher than the traditional OPA curve for fraction of removals greater or equal to 0.7, while in Fig. 9 the fraction of directly failed substations is strictly higher (without overlapping of confidence intervals) for fraction of removals greater or equal to 0.5. This discrepancy is due to the fact that the fraction of directly/indirectly failed substations is not the only factor to take into consideration. In fact, also the fraction of functional substations, which specific substations are failed/functional and the load shedding within the base loads in the external power network can affect the final outcome of the analysis.

It is also useful to compare the normalized area under the curves in Fig. 8. The results, normalized with the traditional OPA curve, are shown in Table 4. As it can be clearly seen, the areas for the curves corresponding to $T_{t\leftarrow p} = 0.0$ and $T_{t\leftarrow p} = 0.5$ are smaller than the area below the traditional OPA model, while the area under the curve for $T_{t\leftarrow p} = 1.0$ is slightly greater. This confirms that, depending on the $T_{t\leftarrow p}$ value, including the impact of the railway network on the external power network might impact positively or negatively the vulnerability in terms of $FDNS$.

We can conclude that failures in the railway network can affect the vulnerability of the external power network. In terms of DNS, including the impact of failures in the railway network on the external power network within the analysis reduces the vulnerability of the external power network, as the $T_{t\leftarrow p}$ curves in Fig. 7 are always smaller or equal than the traditional OPA curve. In terms of $FDNS$, including the impact of failures in the railway network on the external power network might increase or reduce the vulnerability of the external

power network, depending on the value of $T_{t\leftarrow p}$ and the fraction of removals.

In conclusion, we show that, in order to estimate precisely the vulnerability of the external power network, it is important to include the impact of failures in the railway network on the external power network, as the outcome of the analysis might differ.

5. Conclusion

In this work, we have proposed an approach for modeling and vulnerability analysis of interdependent railway-power networks which includes: (i) modeling of the interface between the external power network and the railway network through the traction power network, (ii) modeling of cascading failures dynamics within and between networks, (iii) evaluation of interdependent behaviors within the vulnerability analysis of railway and external power network. Our analysis showed that:

- The effect of cascading failures within the external power network on the railway network should be taken into account while performing a vulnerability analysis. In fact, estimations of the consequences on the railway network based on load shedding analysis which does not account for cascading failures ($p_{ol} = 0$) can lead to underestimated negative outcomes.
- It is important to include the effect of failures within the railway network on the external power network, as it might considerably change the vulnerability values of the external power network. To the best of our knowledge, this was never evaluated in previous works.

In this work, we used an illustrative case-study based on British systems. However, the initial assumptions are flexible, and the approach can be used to analyze a wide range of situations. External power networks can describe transmission or distribution systems at various voltages. Traction power networks can describe substations in electrical isolation (e.g. for AC electrified railway systems) or connected between each other (e.g. for DC electrified railway systems). Railway networks can describe different rail-based transportation means (metro, regional trains or high-speed systems).

We analyzed the vulnerability of our case-study using random removals of edges from the external power network as initiating events. However, other disruption scenarios, such as targeted removals or spatially-localized removals, can be easily analyzed. In addition, the vulnerability of each system can be analyzed with different performance indicators.

From a perspective of protecting the system, an integrated analysis accounting for interdependent behaviors between railway and external power networks is important in order to estimate correctly the vulnerabilities of the systems and plan adequate preventive measures and resource allocation. This is particularly relevant for the railway network, as the inclusion of cascading failures within the analysis can lead to considerably greater negative consequences. Railway operators should consider this while planning preventive measures, such as emergency generators allocation or maintenance scheduling, or resilience-driven design solutions, such as allocation of redundancies with the traction and external power network.

We can then conclude that vulnerability analysis of interdependent railway and power networks should include an evaluation on the possible effects of cascading failures. While the dependency is unidirectional (the railway network depends on the external power network, but not vice versa), the modeling and the analysis should be bidirectional, accounting for the effect of failures within the railway network on the cascading failures dynamics within the external power network.

CRediT authorship contribution statement

Andrea Bellè: Conceptualization, Methodology, Software, Formal analysis, Writing – original draft, Writing – review & editing, Visualization. **Zhiguo Zeng:** Writing – review & editing, Supervision. **Carole Duval:** Writing – review & editing. **Marc Sango:** Writing – review & editing. **Anne Barros:** Writing – review & editing, Supervision, Project administration.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. DC power flow model

Active and reactive power injections at bus i are defined, in the AC power flow model, respectively by Eqs. (A.1) and (A.2):

$$P_i = V_i \sum_{j=1}^{N} V_j (G_{ij} \cos\delta_{ij} + B_{ij} \sin\delta_{ij}) \tag{A.1}$$

$$Q_i = V_i \sum_{j=1}^{N} V_j (G_{ij} \sin\delta_{ij} - B_{ij} \cos\delta_{ij}) \tag{A.2}$$

where P_i and Q_i are the active and reactive power injections at bus i, V_i the voltage magnitude, $\delta_{ij} = \delta_i - \delta_j$ the voltage angle difference between buses i and j, G_{ij} and B_{ij} respectively the real and imaginary part of admittance matrix elements and N the number of buses. This formulation is non-linear, and it is usually solved by applying Gauss–Seidel or Newton–Raphson method, resulting in computationally expensive simulations. The DC power flow model represents an approximation of the aforementioned AC power flow model. It consists in a linearization of the power flow equations, and it is based on three main assumptions:

1. The electrical resistance of each line i is negligible.

$$r_i \approx 0 \tag{A.3}$$

2. The voltage magnitude at each bus i are equal to 1.

$$|V_i| = 1 \tag{A.4}$$

3. The voltage angle difference between two buses i and j, connected by the same line, is small. The trigonometric terms can thus be linearized:

$$\sin\delta_{ij} \approx \delta_i - \delta_j \tag{A.5}$$

$$\cos\delta_{ij} \approx 1. \tag{A.6}$$

Given these assumptions, the active power injection at bus i and the power flow in line k between bus i and j are expressed in Eqs. (A.7) and (A.8):

$$P_i = \sum_{j=1}^{N} B_{ij}(\delta_i - \delta_j) \tag{A.7}$$

$$F_k = \frac{\delta_i - \delta_j}{x_k} \tag{A.8}$$

where x_k is the reactance of line k. The DC power flow model can be expressed also in matrix form as following:

$$\bar{\delta} = \mathbf{B_N}^{-1} \cdot \mathbf{P_N} \tag{A.9}$$

$$\mathbf{F_l} = \mathbf{B_d} \cdot \mathbf{A} \cdot \bar{\delta} \tag{A.10}$$

where $\mathbf{B_N}$ is the admittance matrix with resistance equal to 0, $\bar{\delta}$ is the bus voltage angle vector, B_d is the diagonal line susceptance matrix and A is the line incidence matrix. For more details on derivation and application of the DC power flow model, the reader is referred to specialized literature [66–68,68,69].

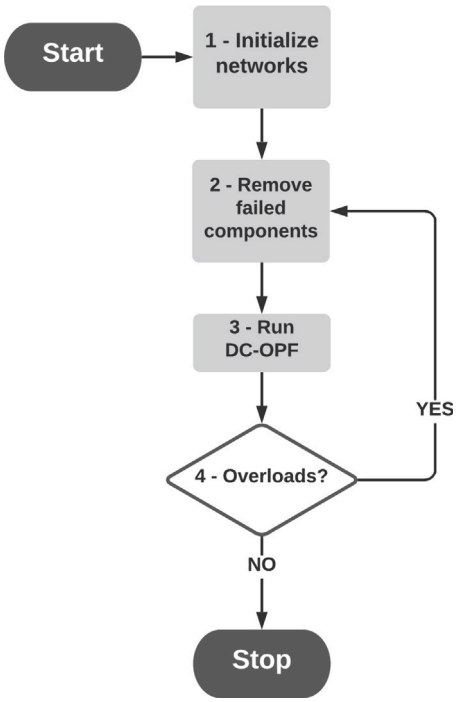

Fig. B.10. Flowchart of the traditional OPA model.

Appendix B. OPA model

The OPA model [56,57] is a flow-based model for cascading failures simulation in power networks. It comprises a slow dynamics, which accounts for the increasing in power demand over time, and a fast dynamics, which accounts for initiating events and proper cascading failures. In our work, we are interested in the fast dynamics. The flowchart of the algorithm is shown in Fig. B.10, and it comprises four iterative steps:

1. The power network is initialized.
2. The failed components, such as edges or nodes, are removed from the power network.
3. The optimal power flow is computed with a DC Optimal Power Flow, previously shown in Eqs. (8)–(12).
4. Check for overloaded lines (lines with a power flow within 1% of the maximum flow capacity). An overloaded line fails with probability p_{ol}. If there is at least one failed line, return to Step 2; otherwise, stop the algorithm.

References

[1] Buldyrev SV, Parshani R, Paul G, Stanley HE, Havlin S. Catastrophic cascade of failures in interdependent networks. Nature 2010;464(7291):1025–8.
[2] Johansson J, Hassel H, Cedergren A. Vulnerability analysis of interdependent critical infrastructures: case study of the Swedish railway system. Int J Crit Infrastruct 2011;7(4):289–316.
[3] Pant R, Hall JW, Blainey SP. Vulnerability assessment framework for interdependent critical infrastructures: case-study for Great Britain's rail network. Eur J Transp Infrastruct Res 2016;16(1).
[4] Major power failure affects homes and transport. 2020, https://www.bbc.com/news/uk-49300025?ns_mchannel=social&ns_source=twitter&ns_campaign=bbc_breaking&ns_linkname=news_central. [Accessed 02 December 2020].
[5] of Rail and Road O. Report following railway power disruption on 9th August 2019. 2020.
[6] A European-wide power and infrastructure break-down ("blackout") and railways operators. UIC ENews 2018;616.
[7] Rinaldi SM, Peerenboom JP, Kelly TK. Identifying, understanding, and analyzing critical infrastructure interdependencies. IEEE Control Syst Mag 2001;21(6):11–25.
[8] Sharma N, Nocera F, Gardoni P. Classification and mathematical modeling of infrastructure interdependencies. Sustain Resil Infrastruct 2021;6(1–2):4–25.
[9] Dudenhoeffer DD, Permann MR, Manic M. CIMS: A framework for infrastructure interdependency modeling and analysis. In: Proceedings of the 2006 winter simulation conference. IEEE; 2006, p. 478–85.
[10] Zimmerman R. Social implications of infrastructure network interactions. J Urban Technol 2001;8(3):97–119.
[11] Sharma N, Tabandeh A, Gardoni P. Resilience analysis: A mathematical formulation to model resilience of engineering systems. Sustain Resil Infrastruct 2018;3(2):49–67.
[12] Panteli M, Mancarella P, Trakas DN, Kyriakides E, Hatziargyriou ND. Metrics and quantification of operational and infrastructure resilience in power systems. IEEE Trans Power Syst 2017;32(6):4732–42.
[13] Suo W, Wang L, Li J. Probabilistic risk assessment for interdependent critical infrastructures: A scenario-driven dynamic stochastic model. Reliab Eng Syst Saf 2021;214:107730.
[14] Guidotti R, Chmielewski H, Unnikrishnan V, Gardoni P, McAllister T, van de Lindt J. Modeling the resilience of critical infrastructure: The role of network dependencies. Sustain Resil Infrastruct 2016;1(3–4):153–68.
[15] Sharma N, Tabandeh A, Gardoni P. Regional resilience analysis: A multiscale approach to optimize the resilience of interdependent infrastructure. Comput-Aided Civ Infrastruct Eng 2020;35(12):1315–30.
[16] Nan C, Sansavini G. A quantitative method for assessing resilience of interdependent infrastructures. Reliab Eng Syst Saf 2017;157:35–53.
[17] Ye Q, Kim H. Assessing network vulnerability of heavy rail systems with the impact of partial node failures. Transportation 2019;46(5):1591–614.
[18] Chen A, Yang C, Kongsomsaksakul S, Lee M. Network-based accessibility measures for vulnerability analysis of degradable transportation networks. Netw Spat Econ 2007;7(3):241–56.
[19] Zhang J, Hu F, Wang S, Dai Y, Wang Y. Structural vulnerability and intervention of high speed railway networks. Physica A 2016;462:743–51.
[20] Ouyang M, Zhao L, Hong L, Pan Z. Comparisons of complex network based models and real train flow model to analyze Chinese railway vulnerability. Reliab Eng Syst Saf 2014;123:38–46.
[21] Hong L, Ye B, Yan H, Zhang H, Ouyang M, He XS. Spatiotemporal vulnerability analysis of railway systems with heterogeneous train flows. Transp Res A 2019;130:725–44.
[22] Fang C, Dong P, Fang Y-P, Zio E. Vulnerability analysis of critical infrastructure under disruptions: An application to China Railway High-speed. Proc Inst Mech Eng O 2020;234(2):235–45.
[23] Berche B, Ferber CV, Holovatch T, Holovatch Y. Transportation network stability: a case study of city transit. Adv Complex Syst 2012;15(supp 01):1250063.
[24] Hong L, Ouyang M, Peeta S, He X, Yan Y. Vulnerability assessment and mitigation for the Chinese railway system under floods. Reliab Eng Syst Saf 2015;137:58–68.
[25] Yan Y, Hong L, He X, Ouyang M, Peeta S, Chen X. Pre-disaster investment decisions for strengthening the Chinese railway system under earthquakes. Transp Res E 2017;105:39–59.
[26] Mattsson L-G, Jenelius E. Vulnerability and resilience of transport systems–A discussion of recent research. Transp Res A 2015;81:16–34.
[27] Reggiani A, Nijkamp P, Lanzi D. Transport resilience and vulnerability: The role of connectivity. Transp Res A 2015;81:4–15.
[28] Ouyang M. Comparisons of purely topological model, betweenness based model and direct current power flow model to analyze power grid vulnerability. Chaos 2013;23(2):023114.
[29] Johansson J, Hassel H, Zio E. Reliability and vulnerability analyses of critical infrastructures: Comparing two approaches in the context of power systems. Reliab Eng Syst Saf 2013;120:27–38.
[30] Chen G, Dong ZY, Hill DJ, Zhang GH, Hua KQ. Attack structural vulnerability of power grids: A hybrid approach based on complex networks. Physica A 2010;389(3):595–603.
[31] Panteli M, Mancarella P. Influence of extreme weather and climate change on the resilience of power systems: Impacts and possible mitigation strategies. Electr Power Syst Res 2015;127:259–70.
[32] Zhu Y, Yan J, Tang Y, Sun YL, He H. Resilience analysis of power grids under the sequential attack. IEEE Trans Inf Forensics Secur 2014;9(12):2340–54.
[33] Zhao L, Zeng B. Vulnerability analysis of power grids with line switching. IEEE Trans Power Syst 2013;28(3):2727–36.
[34] Fang Y, Sansavini G. Optimizing power system investments and resilience against attacks. Reliab Eng Syst Saf 2017;159:161–73.
[35] Chang L, Wu Z. Performance and reliability of electrical power grids under cascading failures. Int J Electr Power Energy Syst 2011;33(8):1410–9.
[36] Abedi A, Gaudard L, Romerio F. Review of major approaches to analyze vulnerability in power system. Reliab Eng Syst Saf 2019;183:153–72.
[37] Wu Y, Chen Z, Zhao X, Gong H, Su X, Chen Y. Propagation model of cascading failure based on discrete dynamical system. Reliab Eng Syst Saf 2021;209:107424.

[38] Kong J, Zhang C, Simonovic SP. Optimizing the resilience of interdependent infrastructures to regional natural hazards with combined improvement measures. Reliab Eng Syst Saf 2021;210:107538.
[39] Liu X, Fang Y-P, Zio E. A hierarchical resilience enhancement framework for interdependent critical infrastructures. Reliab Eng Syst Saf 2021;215:107868.
[40] Goldbeck N, Angeloudis P, Ochieng WY. Resilience assessment for interdependent urban infrastructure systems using dynamic network flow models. Reliab Eng Syst Saf 2019;188:62–79.
[41] Ouyang M. Critical location identification and vulnerability analysis of interdependent infrastructure systems under spatially localized attacks. Reliab Eng Syst Saf 2016;154:106–16.
[42] Applegate CJ, Tien I. Framework for probabilistic vulnerability analysis of interdependent infrastructure systems. J Comput Civ Eng 2019;33(1):04018058.
[43] Lu L, Wang X, Ouyang Y, Roningen J, Myers N, Calfas G. Vulnerability of interdependent urban infrastructure networks: Equilibrium after failure propagation and cascading impacts. Comput-Aided Civ Infrastruct Eng 2018;33(4):300–15.
[44] Ouyang M, Pan Z, Hong L, He Y. Vulnerability analysis of complementary transportation systems with applications to railway and airline systems in China. Reliab Eng Syst Saf 2015;142:248–57.
[45] Zhang J, Song B, Zhang Z, Liu H. An approach for modeling vulnerability of the network of networks. Physica A 2014;412:127–36.
[46] Zio E, Sansavini G. Modeling failure cascades in critical infrastructures with physically-characterized components and interdependencies. In: ESREL 2010 annual conference. 2010. p. 652–61.
[47] Zio E, Sansavini G. Modeling cascading failures in systems of systems with uncertain behavior. In: ICASP11. 2011. p. 1858–66.
[48] Johansson J, Hassel H. An approach for modelling interdependent infrastructures in the context of vulnerability analysis. Reliab Eng Syst Saf 2010;95(12):1335–44.
[49] Svegrup L, Johansson J. Vulnerability analyses of interdependent critical infrastructures: Case study of the Swedish national power transmission and railway system. In: European safety and reliability conference. ESREL2015; 2015, p. 4499–507.
[50] Adjetey-Bahun K, Birregah B, Châtelet E, Planchet J-L. A model to quantify the resilience of mass railway transportation systems. Reliab Eng Syst Saf 2016;153:1–14.
[51] Dorbritz R. Assessing the resilience of transportation systems in case of large-scale disastrous events. In: Environmental engineering. Proceedings of the international conference on environmental engineering, vol. 8. Vilnius Gediminas Technical University, Department of Construction Economics; 2011, p. 1070.
[52] Pelzer R, Duval C, Migliorini M, Wilson F. FORTRESS - Foresight Tools for Responding to cascading effects in a crisis - Case study reports on system interdependencies. 2015, Confidential.
[53] Motter AE, Lai Y-C. Cascade-based attacks on complex networks. Phys Rev E 2002;66(6):065102.
[54] Crucitti P, Latora V, Marchiori M. Model for cascading failures in complex networks. Phys Rev E 2004;69(4):045104.
[55] Cupac V, Lizier JT, Prokopenko M. Comparing dynamics of cascading failures between network-centric and power flow models. Int J Electr Power Energy Syst 2013;49:369–79.
[56] Dobson I, Carreras B, Lynch V, Newman D. An initial model for complex dynamics in electric power system blackouts. In: Proceedings of the 34th annual hawaii international conference on system sciences, vol. 2. Citeseer; 2001, p. 2017.
[57] Carreras BA, Newman DE, Dobson I, Degala NS. Validating OPA with WECC data. In: 2013 46th hawaii international conference on system sciences. IEEE; 2013, p. 2197–204.
[58] Mahmoud MS, Xia Y. Networked control systems: cloud control and secure control. Butterworth-Heinemann; 2019.
[59] Duval C, Brinzei N, Chraibi H, Hassanaly M. Adéquation des Automates Stochastiques Hybrides pour la modélisation des conséquences d'un événement redouté. In: Lambda-Mu 23 Virtual Congress. 2020.
[60] Guo H, Zheng C, Iu HH-C, Fernando T. A critical review of cascading failure analysis and modeling of power system. Renew Sustain Energy Rev 2017;80:9–22.
[61] 21 G. Fast forward—A high-speed rail strategy for Britain. Author Kingston-upon-Thames; 2009.
[62] Bukhsh W, McKinnon K. Network data of real transmission networks. 2013, Published online at http://www.maths.ed.ac.uk/optenergy/NetworkData.
[63] Thurner L, Scheidler A, Schäfer F, Menke J, Dollichon J, Meier F, et al. pandapower — An open-source python tool for convenient modeling, analysis, and optimization of electric power systems. IEEE Trans Power Syst 2018;33(6):6510–21. http://dx.doi.org/10.1109/TPWRS.2018.2829021.
[64] Coffrin C, Bent R, Sundar K, Ng Y, Lubin M. Powermodels. jl: An open-source framework for exploring power flow formulations. In: 2018 power systems computation conference. IEEE; 2018, p. 1–8.
[65] Taleb NN, Douady R. Mathematical definition, mapping, and detection of (anti) fragility. Quant Finance 2013;13(11):1677–89.
[66] Grainger JJ, Stevenson WD, Stevenson WD, et al. Power system analysis. 2003.
[67] Van den Bergh K, Delarue E, D'haeseleer W. DC power flow in unit commitment models. 2014, TME Work. Pap. Environ. Tech. Rep.
[68] Glover JD, Sarma MS, Overbye T. Power system analysis & design, SI version. Cengage Learning; 2012.
[69] Li W. Risk assessment of power systems: models, methods, and applications. John Wiley & Sons; 2014.

Paper II

A. Bellè, Z. Zeng, M. Sango, and A. Barros, "Towards a Realistic Topological and Functional Modeling for Vulnerability Analysis of Interdependent Railway and Power Networks," *Proceedings of the 31st European Safety and Reliability Conference*, 2021.

Towards a realistic topological and functional modeling for vulnerability analysis of interdependent railway and power networks

Andrea Bellè

Chair on Risk and Resilience of Complex Systems, Laboratoire Génie Industriel, CentraleSupélec, Université Paris-Saclay, France. E-mail: andrea.belle@centralesupelec.fr

Zhiguo Zeng

Chair on Risk and Resilience of Complex Systems, Laboratoire Génie Industriel, CentraleSupélec, Université Paris-Saclay, France. E-mail: zhiguo.zeng@centralesupelec.fr

Marc Sango

SNCF, Direction Technologies Innovation et Projet Groupe, Innovation & Recherche, Département Physique du Système Ferroviaire, Equipe Sécurité Système, France. E-mail: marc.sango@sncf.fr

Anne Barros

Chair on Risk and Resilience of Complex Systems, Laboratoire Génie Industriel, CentraleSupélec, Université Paris-Saclay, France. E-mail: anne.barros@centralesupelec.fr

Railway systems and power grids are recognized as two of the most important critical infrastructures. The majority of European railway networks are electrified, and power transmission networks represent usually the main power supplier. Railway and power networks share thus a unidirectional interdependency, as the railway network functionality depends on the power network. Due to this interdependency, failures in power networks have the potential of causing vast disruption in the dependent railway networks. Despite this, the issue of modeling interdependent railway and power networks has not been addressed sufficiently carefully in the existing literature. Furthermore, the treatment of cascading failures in power networks and their consequences in railway networks is limited and approximative. In this work, we propose a modeling framework which accounts for more realistic assumptions on the interconnections topology and the cascading failures dynamics. Firstly, we model the interconnections between the railway and external power network by introducing the traction power network, which acts as a bridge between the external power grid and the railway network. Secondly, we model cascading failures in the external and traction power networks with an approach based on the DC power flow model. Thirdly, we suggest a simple approach to estimate the negative consequences on the railway network due to load shedding in the traction power network. Vulnerability analysis is performed to estimate the negative consequences in the railway network due to different failure scenarios in the external power network. Sensitivity analysis on the initial assumptions is also performed.

Keywords: Critical infrastructures, interdependent networks, power network, railway network, cascading failures, vulnerability.

1. Introduction

Critical infrastructures are large systems which provide essential services to society. Despite being technologically and functionally different, critical infrastructures are often interdependent from each other. Power and railway networks are among the most important critical infrastructures, and they share a unidirectional interdependency, as the power network supplies the necessary electricity to the railway network. This means that disturbances in power grids can propagate and cause disruption in railway systems. Despite this, the available modeling approaches for vulnerability analysis of interdependent railway and power networks are limited and approximative.

One of the major limitations of the existing modeling approaches is that the fundamental assumptions are often over-simplified, failing to fully capture the reality of the interconnection between railway and power networks. More specifically, available studies often neglect the modeling of the traction power network, which in reality plays a key role in linking the external power grid and the railway system. In fact, this subsystem is accounted for only in Johansson and Hassel (2010); Johansson et al. (2011).

Proceedings of the 31st European Safety and Reliability Conference
Edited by Bruno Castanier, Marko Cepin, David Bigaud, and Christophe Berenguer
Published by Research Publishing, Singapore.
ISBN: 978-981-18-2016-8; doi:10.3850/978-981-18-2016-8_356-cd

A second drawback of the existing literature is that the evaluation of cascading failures in power networks and their impact on the railway network is often neglected (Johansson and Hassel (2010); Johansson et al. (2011); Pant et al. (2016)) or simplified with a network-based approach (Zhang et al. (2014); Zio and Sansavini (2010, 2011)).

With this work, we propose a new approach for interdependent railway and power networks which includes the modeling of traction power networks and cascading failures in power networks. The major contributions are:

- A realistic model of interdependent railway electrical networks, which takes into account the traction power network, is proposed.
- The impact of load shedding in the external and traction power network on the performance of interdependent railway and power networks is investigated.

The focus of the study is the vulnerability analysis of the railway network given different initiating disruptive event in the external power network. The remaining part of this work is structured as follows: in section 2, the modeling framework is presented; in section 3, the case-study is described; in section 4, some preliminary results are presented; in section 5, final insights and developments are given.

2. Modeling and analysis framework

The purpose of this work is to model the impact of disturbances and disruption in external power networks, including cascading failures scenarios, to railway networks. For this, traction power networks are introduced in order to act as a bridge between external power and railway networks. The modeling and analysis framework comprises three main steps:

- Compute the negative consequences in the external power network after an initiating disruptive event and cascading failures (if any).
- Use the output from the previous step as an input for the traction power network and compute the negative consequences, including cascading failures (if any).
- Use the output from the previous step to compute performance-based vulnerability of traction power network and railway network.

In the next subsections, we address the main modeling features necessary to perform the aforementioned steps, including network-based topological modeling, modeling of electrical quantities of power networks, cascading failures within and between networks and performance-based vulnerability index.

2.1. *Network-based topological modeling*

Network science is often used to describe the topology of critical infrastructures. A network is defined by a graph $\mathbf{G} = (\mathbf{V}, \mathbf{E})$, with $\mathbf{V} = \{v_1, v_2, ..., v_N\}$ representing the set of N nodes (or vertices) and $\mathbf{E} = \{e_1, e_2, ..., e_M\}$ the set of M edges. Each edge k is also defined by a tuple $e_k = (v_i, v_j)$, which indicates the two nodes v_i and v_j connected by edge k. In this work, we identify three separate networks: the railway network, defined by the subscript R, the external power network, defined by the subscript E, and the traction power network, defined by the subscript T.

In the railway network $\mathbf{G_R} = (\mathbf{V_R}, \mathbf{E_R})$, nodes represent stations and edges represent railway tracks. Each railway edge represents a direct bi-directional physical connection between two stations.

In the external power network $\mathbf{G_E} = (\mathbf{V_E}, \mathbf{E_E})$, nodes represent electrical buses and edges represent transmission lines. Each electrical bus can contain power production or consumption units, here referred as generators and loads. These elements are defined by the set of generators $\mathbf{N_{E,G}}$ and the set of loads $\mathbf{N_{E,L}}$.

In the traction power network $\mathbf{G_T} = (\mathbf{V_T}, \mathbf{E_T})$, nodes represent electrical substations and edges represent distribution lines. Each substation contains a power production unit, which represent the available power extracted from the external power network, and a consumption unit, representing the power demand of the substation. These elements are defined by the set of generators $\mathbf{N_{T,G}}$ and the set of loads $\mathbf{N_{T,L}}$. The power production units, rather than representing real power production, they correspond to the available power extracted by each substation from the external power network.

2.2. *Interdependencies*

In most situations, the traction power network is directly connected to the external power network. However, to overcome lack of data and decrease the structural complexity of the model, we decouple and treat the two networks separately. The traction power network thus depends on the external power network in terms of electricity supply. Specifically, we assume that each substation in the traction power network is dependent on the geographically-closest load node in the external power network. We denote this relationship with interdependency edges $e_{T \leftarrow E}^{i \leftarrow j} = (v_{T,i}, v_{E,j})$, indicating that node i in the traction power network is dependent on node j in the external power network.

The traction power network is usually directly connected to the railway catenary to supply electricity to the rolling stocks. The railway network thus depends on the traction power network in

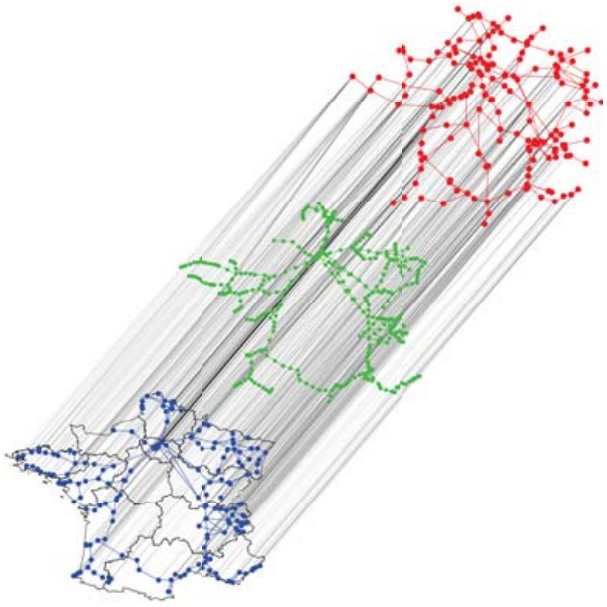

Fig. 1. Interdependent railway network (blue), traction power network (green) and external power network (red).

terms of electricity supply. Specifically, we assume that each railway track i depends on a set of substation $\mathbf{V_T^{r,i}} \subseteq \mathbf{V_T}$, that are responsible of the electricity supply of that specific railway track. In Figure 1, this relationship is graphically represented with edges between the two nodes delimiting railway track i in the railway network and the nodes of the traction power network belonging to $\mathbf{V_T^{r,i}}$.

2.3. *Modeling of electrical flows*

External and traction power networks are characterized by electrical quantities and parameters. In this work, we rely on the DC power flow model, and the main quantities to consider are power generation capacity and power demand, for generators and loads, and power flow capacity for lines. We assume that each generator and load i are characterized by a maximum generation capacity and demand $P_{G,i}^{max}$ (positive power) and $P_{L,i}^{max}$ (negative power), and each line i is characterized by a maximum flow capacity $F_{l,i}^{max}$. The flow in each line i is computed using the DC power flow model assumption, according to Equation (1):

$$F_{l,i} = B_i \, \Delta\theta_i \tag{1}$$

where B_i is the susceptance of line i and $\Delta\theta_i$ is the phase angle difference between the two nodes defining line i. For more details, the reader is referred to specialize literature (Van den Bergh et al. (2014); Li (2014)).

2.4. *Modeling of cascading failures*

2.4.1. *Power network*

External and traction power network are subjected to cascading failures, and different approaches are available (see Guo et al. (2017)). In this work, we rely on the traditional ORNL-PSerc-Alaska (OPA) model, based on DC power flow equations and linear optimization (Dobson et al. (2001); Carreras et al. (2013)). The model aims at simulating the behaviour of the power grid after an initiating disruptive event, considering electrical quantities and operator actions. After the initialization of the power network and the initiating disruptive event, the model comprises the following steps:

(1) Remove failed elements from the network. If a node is failed, remove also the connected components (generators, loads, lines).
(2) Perform a DC optimal power flow (DC-OPF).
(3) Check if there is any overloaded failed lines: if yes, go back to step 1; otherwise, stop the simulation and compute load shedding.

Firstly, at step 1 failed lines and nodes, along with the connected components, are removed from the network.

Secondly, a DC-OPF is performed in order to simulate operator actions. The cost to minimize is shown in Equation (2), and it is subjected to constraints in Equations (3)-(6).

$$\min_{\mathbf{P_G},\mathbf{P_L}} \quad \sum_{i=1}^{N_G} P_{G,i} - W \sum_{j=1}^{N_L} |P_{L,j}| \tag{2}$$

$$\sum_{i=1}^{N_G} P_{G,i} - \sum_{j=1}^{N_L} |P_{L,j}| = 0 \tag{3}$$

$$0 \leq P_{G,i} \leq P_{G,i}^{max} \tag{4}$$

$$P_{L,j}^{max} \leq P_{L,j} \leq 0 \tag{5}$$

$$-F_{l,k}^{max} \leq F_{l,k} \leq F_{l,k}^{max} \tag{6}$$

The first term of Equation (2) represents the power production cost, which is assigned a unitary value per unit of power $P_{G,i}$ produced in each generator i. The second term represents the negative cost associated to the power $P_{L,i}$ supplied at each load i. The penalty constant W, here assumed to be equal to 100, ensure the minimization of load shedding when possible. Equations (3)-(6) represent the optimization constraints. The constraint in Equation (3) represents the power balance of generation and demand in the power network, which must be always equal to 0. Equations (4) and (5) represents possible ranges of power of each generator and

load. The constraint in Equation (6) represents the maximum power flow in each transmission line. The power flow in each line F_l is computed using the DC power flow model.

Thirdly, the power network is checked for additional failures due to overloaded lines. A line is considered overloaded if its flow is within 1% of its maximum capacity. When a line is overloaded, it fails (or trips) with a probability p (in this study, we assume p=1). If there is any additional failure, the simulation goes back to step 1 for a new iteration; otherwise the simulation is stopped and the new level of production $P'_{G,i}$ for each generator and demand $P'_{L,i}$ for each load is given as output.

All the power flow computations are performed using Pandapower (Thurner et al. (2018)) and PowerModels.jl (Coffrin et al. (2018)).

2.4.2. *From external to traction power network*

The traction power network receives electricity from the external power network. This means that disruption in the external power network can decrease the power available to the traction power network. As explained in section 2.2, each substation i is dependent on the load node j in the external power network as expressed by the interdependency edge $e^{i \leftarrow j}_{T \leftarrow E}$. In general, the load j represents the power consumption of multiple users (residential, industrial, etc.) in a given area, including the substation i. If due to a disruptive event (with or without cascading failures), the power supplied at load j is not sufficient to satisfy entirely its power demand, substation i can be negatively impacted. To model this, we distinguish two separate scenarios, called the no-priority scenario and the priority scenario.

In the first one, substation i has no priority over the other users supplied by load j, and the maximum power available at substation j, defined as $P^{max'}_{T,G,i}$, is computed according to Equation (7):

$$P^{max'}_{T,G,i} = P^{max}_{T,G,i} \cdot \frac{P'_{E,L,i}}{P_{E,L,i}} \tag{7}$$

where $P^{max}_{T,G,i}$ is the nominal maximum power available at substation i. It should be noted that the power available at the substation is modeled as a generator capacity, since the two power networks are decoupled. The physical meaning of Equation (7) is that the fraction of nominal power available at substation i is equal to the fraction of nominal power demand supplied to load j.

In the second one, we assume the substation i has the priority over the other users supplied by load j. This means that, as long as there is enough power supplied at load j, avoiding load shedding in substation i is prioritized. The power available $P^{max'}_{T,G,i}$ is thus computed according to equation (8):

$$P^{max'}_{T,G,i} = \begin{cases} P^{max}_{T,G,i}, & \text{if } P'_{E,L,i} \geq A_j, \\ P^{max}_{T,G,i} \cdot B_j, & \text{otherwise,} \end{cases} \tag{8}$$

where A_j represents the sum of the nominal power available at each substation dependent on load j and B_j is the ratio between the power supplied at load j and A_j. These two parameters are expressed in Equations (9) and (10):

$$A_j = \sum_{\substack{k=1, \\ k \in \mathbf{V^j_{T \rightarrow E}}}}^{N^j_{T \rightarrow E}} P^{max}_{T,G,k} \tag{9}$$

$$B_j = \frac{P'_{E,L,i}}{A_j} \tag{10}$$

where $\mathbf{V^j_{T \rightarrow E}}$ represents the set of $N^j_{T \rightarrow E}$ substations dependent on load j.

2.5. *Modeling of vulnerability index*

The vulnerability index V can be generalized as the relative change of a system's performance indicator after a disruptive event. In this work, we consider, for external and traction power network, the fractional load shedding as vulnerability index, expressed as in Equation (11):

$$LS = 1 - \frac{\sum_{i=1}^{N_L} P'_{L,i}}{\sum_{i=1}^{N_L} P_{L,i}} \tag{11}$$

where $P'_{L,i}$ represents the power supplied to load i after the cascading failures and $P_{L,i}$ represents the actual power demand of load i. The load shedding LS represents the fraction of power demand not satisfied.

For the railway network, we consider a vulnerability index, defined as $\Delta\Phi$, which accounts for the impact of load shedding in the traction power network, as expressed in Equation (12):

$$\Delta\Phi = 1 - \frac{1}{M_R} \sum_{i=1}^{M_R} \left(\frac{\sum_{j=1}^{N_T^{R,i}} \frac{P'_{T,L,j}}{P_{T,L,j}} \prod_{k=1}^{N_T^{R,i}} S_{R,k}}{N_T^{R,i}} \right) \tag{12}$$

where M_R is the number of edges in the railway network, $N_T^{R,i}$ is the number of substation in the subset $\mathbf{V^i_{R \leftarrow T}}$ (number of substations supplying railway i) and $S_{R,k}$ is the binary state of substation k (1 if functional, 0 if failed), computed according to Equation (13):

$$S_{R,k} = \begin{cases} 1, & \text{if } \frac{P'_{T,L,k}}{P_{T,L,k}} > 0, \\ 0, & \text{otherwise.} \end{cases} \tag{13}$$

The physical meaning is that the drop of performance in each railway is equal to the average load shedding in the substations which supply the electricity to the railway, when the substations are functional. If at least one of the substations is failed, the drop of performance is total, because the railway is interrupted. In this work, we consider a substation functional as long as it receives some electricity.

3. Case-study

The external power network is based on the French 400 kV transmission power network (Fang et al. (2015)). It contains 171 electrical buses, connected by 220 transmission lines. Each bus contains either a generator, with positive electrical power (power production), or a load, with negative electrical power (power demand). There are 26 generator nodes, with powers ranging from 1.4 GW to 8.1 GW, and 145 load nodes, with powers ranging from 0.151 GW to 1.331 GW in absolute value. The maximum power production capacity and the total power demand are 85 GW and 84.988 GW, respectively. Each transmission line is assumed to have a maximum flow capacity equal to 7 GW.

The railway network, shown in blue in Figure 1, is based on the French high-speed railway systems, called *Train à Grande Vitesse* or TGV (SNCF (2014)). It is composed by 185 nodes, representing stations, connected by 214 edges, representing railway tracks.

The railway and the external power network are connected through the traction power network. Due to lack of data, we define a set of assumptions in order to build a realistic traction power network. Firstly, we design the traction power network topology starting from the railway network. Specifically, for each railway i connecting stations j and k, we assume that:

- For each station j in the railway network, a substation j is present in the traction power network at the same geographical location.
- Between each substation j and k (corresponding to stations j and k connected by railway i in the railway network), there is a number of additional substations $N^{add}_{sub,i}$ proportional to the length of railway i according to Equation (14):

$$N^{add}_{sub,i} = \begin{cases} \left\lfloor \dfrac{lt_i}{\bar{d}_{sub}} \right\rceil, & \text{if } N^{add}_{sub,i} \geq 1, \\ 1, & \text{otherwise,} \end{cases} \quad (14)$$

  where lt_i is the length of railway i in km, $\bar{d}_{sub}$ is the average distance between two consecutive substations, here assumed to be 60 km, and $\lfloor \cdot \rceil$ defines the closest integer number. These substations are located equidistantly along the trajectory of railway i and connected consecutively.

The resulting traction power network, shown in green in Figure 1, is composed by 424 nodes, representing the substations, connected by 452 edges, representing distribution lines. These elements are connected to replicate the shape of the railway network, as clearly visible in Figure 1.

Secondly, we define the electrical parameters of the traction power network. As it was explained in section 2.1, each substation contains a load, which represents the power demand of the substation itself, and a generator, which represents the maximum power available from the external power network. In this work, we assume the power demand of each load $P_{T,L,i}$ in the traction power network to be dependent on a base demand and the degree of the substation, according to Equation (15):

$$P_{T,L,i} = P^{max}_{T,L,i} = \bar{P}_{T,L,i} \cdot k(i) \quad (15)$$

where $\bar{P}_{T,L,i}$ is the base demand, here assumed to be 5 MW, and $k(i)$ is the degree (number of connected edges) of substation i. Moreover, we assume that each substation has a maximum power extraction capacity from the external power network equal to its own power demand multiplied by a safety coefficient. This feature is modeled as the maximum capacity of the substation generator $P^{max}_{T,G,i}$, according to Equation (16):

$$P^{max}_{T,G,i} = P_{T,L,i} \cdot (1 + \alpha_G) \quad (16)$$

where α_G is the safety coefficient, here assumed to be 0.5. With these assumptions, we obtain a total power demand of 4520 MW, with load powers ranging from 5 to 30 MW, and maximum extraction capacity (or generation capacity) of 6780 MW, with generator powers ranging from 7.5 to 45 MW. The power flow capacity of each distribution line $F^{max}_{T,L,i}$ is assumed to be equal to the base power demand $\bar{P}_{T,L,i}$.

We study the impact of single bus failures in the external power network on the traction power network and railway network, in terms of load shedding LS_T and loss of performance $\Delta\Phi$. The main steps of the simulation-based approach are shown in Figure 2.

4. Results

The results in terms of load shedding LS_T in the traction power network and loss of performance $\Delta\Phi$ in the railway network, in both priority and no-priority scenario, are shown in Figure 3 and 4, respectively. In both the figures, the x-axis represents the index of bus removed in the external power network as initiating event (from bus 1 to

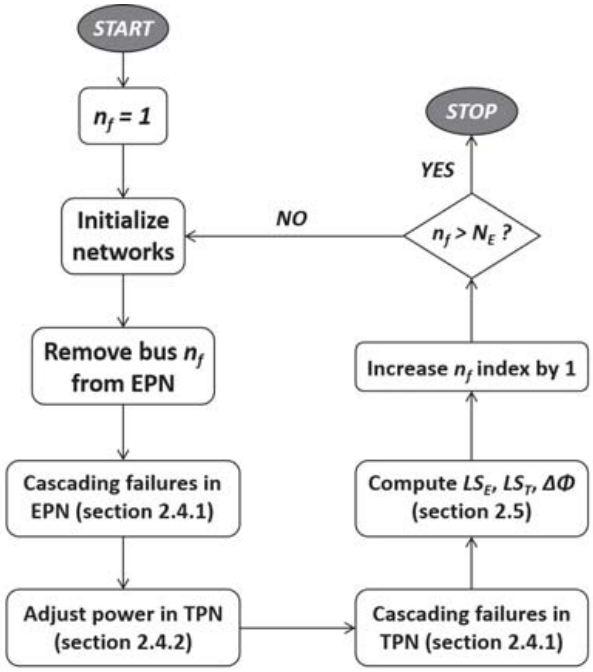

Fig. 2. Flowchart of simulation algorithm. EPN indicates the external power network, TPN indicates the traction power network.

bus 171). Blue squares represent the no-priority scenario, while red stars represent the priority scenario. Green circles represent the total load shedding LS_E in the external power network.

The first consideration to highlight is that, as it is clearly visible in both the figures, the priority scenario leads to a considerably lower negative impact in terms of load shedding in the traction power network and loss of performance in the railway network. This situation is summarized in Table 1. The average load shedding $\overline{LS_T}$ in the

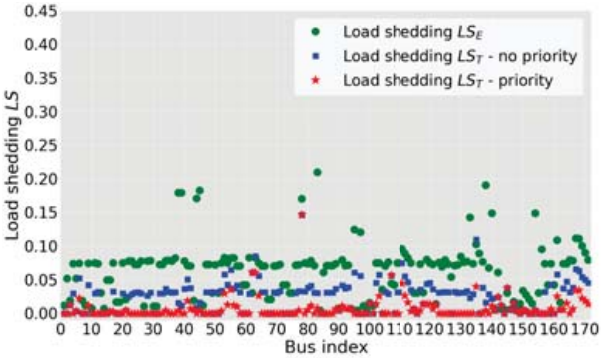

Fig. 3. Load shedding in external power network (green circles), in traction power network with no-priority (blue squares), in traction power network with priority (red stars) for single bus failures in the external power network as initiating event.

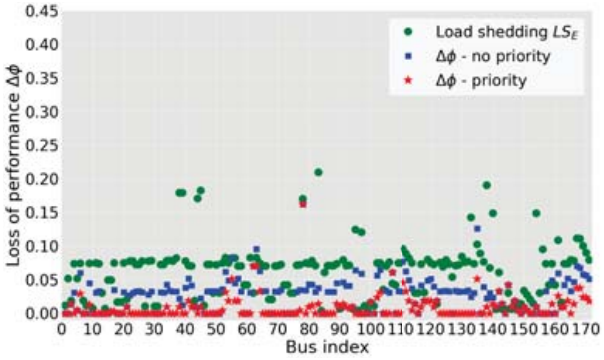

Fig. 4. Load shedding in external power network (green circles), loss of performance in railway network with no-priority (blue squares), loss of performance in railway network with priority (red stars) for single bus failures in the external power network as initiating event.

Table 1. Average load shedding and loss of performance in traction power network and railway network with 95% confidence interval.

$\overline{LS_T}$ No-priority	$\overline{LS_T}$ Priority	$\overline{\Delta\Phi}$ No-priority	$\overline{\Delta\Phi}$ Priority
2.5% $\pm$ 0.3%	0.7% $\pm$ 0.2%	2.8% $\pm$ 0.4%	0.9% $\pm$ 0.3%

traction power network with the no-priority scenario is 2.5% with 95% confidence interval equal to 2.2%-2.8%, while in the priority scenario $\overline{LS_T}$ is equal to 0.7%, with 95% confidence interval equal to 0.5%-0.9%. The same consideration can be drawn for the average loss of performance in the railway network $\overline{\Delta\Phi}$, which presents an average value of 2.8% with 95% confidence interval equal to 2.4%-3.2% in the no-priority scenario, and an average value of 0.9% with 95% confidence interval equal to 0.6%-1.2% in the priority scenario. This behaviour can be highlighted also

Table 2. Five most critical components for traction power network (no-priority and priority scenario).

Bus index No-priority	LS_T No-priority	Bus index Priority	LS_T Priority
78	14.6%	78	14.6%
135	11.0%	62	6.1%
63	8.5%	63	6.1%
111	7.5%	107	5.6%
167	6.6%	111	4.4%

Table 3. Five most critical components for railway network (no-priority and priority scenario).

Bus index No-priority	$\Delta\Phi$ No-priority	Bus index Priority	$\Delta\Phi$ Priority
78	16.2%	78	16.2%
135	12.6%	62	7.0%
63	9.5%	63	7.0%
55	8.2%	107	6.1%
111	7.7%	135	5.1%

Table 4. Five most critical components for external power network.

Bus index	LS_E
83	21.0%
138	19.0%
45	18.3%
39	17.9%
38	17.9%

by comparing results from individual initiating events. The five most critical bus failures in the external power network in terms of load shedding in the traction power network and loss of performance in the railway network are shown in Table 2 and 3, respectively. With the exception of the most critical bus failure (bus 77 for LS_T and $\Delta\Phi$, for both no-priority and priority scenario), we can observe that the results for the no-priority scenario are considerably higher. For example, for LS_T, the second most critical failure in the no-priority scenario (bus 134) leads to a load shedding of 11%, while the second most critical failure in the priority scenario (bus 62) leads to a load shedding of just 6.1%. The same considerations can be drawn for the loss of performance in the railway network. These results highlight the importance to prioritize the power supply to critical infrastructures, such as railway traction network, in case of emergency situations.

Another important aspect to highlight is the strong correlation between the three networks (external power, traction power and and railway), since the load shedding in the external power network is the primary cause of load shedding in the traction power network which can lead to loss of performance in the railway network. Despite this, it is clear that the location of the initiating event and the spatial dynamics of cascading failures play an important role. This aspect is noticeable in the case where load shedding in the external power network and in the traction power network are considerably different. For example, it is interesting to notice how the failure of bus 82 as initiating event constitutes the worst-case scenario for the external power network, leading to a load shedding LS_E of 21% (see Table 4), while in the traction power network it leads to a load shedding of just 3.8% in the no-priority scenario and 1.4% in the priority scenario. This behaviour indicates that the topological features are a key aspect in terms of negative consequences of cascading failures within and between networks.

A similar consideration can be deduced from the difference in results of LS_T and $\Delta\Phi$. Despite the fact that the computations of the two values are strongly correlated, as it is clear from Equations (11) and (12), the losses of performance in the railway network are tendentially higher than the corresponding load sheddings in the traction power network. This is clearly due to the assumption on the substation states expressed in Equation (13). However, we can notice that in some cases the difference is negligible or not present, while in other cases can be considerable, up to a few percentage points. This is again an indication of the importance of the topology, because location of failures and/or load shedding can have a strong impact on the possible negative consequences in the dependent networks.

5. Conclusion

In this work we have proposed a modeling framework for interdependent railway and power networks which accounts for an integrated evaluation of realistic interconnections, introducing traction power networks to act as a bridge, and cascading failures scenarios. The model has been used to study the impact of two different assumptions on the electricity supply priority from the external power network to the traction power network. Preliminary results have shown, as expected, that the priority scenario leads to considerably lower disruption.

Analysis on individual initiating events have shown that vulnerability index in single systems are not always correlated in terms of magnitude. For example, the worst-case scenario for the external power network leads to limited negative consequences in the dependent networks (traction power network and railway network). This behaviour indicates that the topology of the interconnections between interdependent networks plays a crucial role when performing an integrated vulnerability analysis.

Further developments of this work include a more comprehensive sensitivity analysis on the starting assumptions and the comparison, in terms of vulnerability, of different traction power network's configurations.

Acknowledgement

The authors thank Dr. Yiping Fang for providing the data of the external power network. This work is funded by the Chair on Risk and Resilience of Complex Systems (CentraleSupélec, EDF, SNCF, Orange).

References

Carreras, B. A., D. E. Newman, I. Dobson, and N. S. Degala (2013). Validating opa with wecc data. In *2013 46th Hawaii International Conference on System Sciences*, pp. 2197–2204. IEEE.

Coffrin, C., R. Bent, K. Sundar, Y. Ng, and M. Lubin (2018). Powermodels. jl: An open-source framework for exploring power flow formulations. In *2018 Power Systems Computation Conference (PSCC)*, pp. 1–8. IEEE.

Dobson, I., B. Carreras, V. Lynch, and D. Newman (2001). An initial model for complex dynamics in electric power system blackouts. In *Proceedings of the 34th annual Hawaii international conference on system sciences*, Volume 2, pp. 2017–2017. Citeseer.

Fang, Y., N. Pedroni, and E. Zio (2015). Optimization of cascade-resilient electrical infrastructures and its validation by power flow modeling. *Risk Analysis 35*(4), 594–607.

Guo, H., C. Zheng, H. H.-C. Iu, and T. Fernando (2017). A critical review of cascading failure analysis and modeling of power system. *Renewable and Sustainable Energy Reviews 80*, 9–22.

Johansson, J. and H. Hassel (2010). An approach for modelling interdependent infrastructures in the context of vulnerability analysis. *Reliability Engineering & System Safety 95*(12), 1335–1344.

Johansson, J., H. Hassel, and A. Cedergren (2011). Vulnerability analysis of interdependent critical infrastructures: case study of the swedish railway system. *International journal of critical infrastructures 7*(4), 289–316.

Li, W. (2014). *Risk assessment of power systems: models, methods, and applications*. John Wiley & Sons.

Pant, R., J. W. Hall, and S. P. Blainey (2016). Vulnerability assessment framework for interdependent critical infrastructures: case-study for great britain's rail network. *European Journal of Transport and Infrastructure Research 16*(1).

SNCF (2014). Sncf axes tgv. `https://www.oui.sncf/assets/media/2019-07/2014_Axes-TGV-c-2%20copie.pdf`. Accessed: 22/02/2021.

Thurner, L., A. Scheidler, F. Schäfer, J. Menke, J. Dollichon, F. Meier, S. Meinecke, and M. Braun (2018, Nov). pandapower — an open-source python tool for convenient modeling, analysis, and optimization of electric power systems. *IEEE Transactions on Power Systems 33*(6), 6510–6521.

Van den Bergh, K., E. Delarue, and W. D'haeseleer (2014). Dc power flow in unit commitment models. *TME Work. Pap. Environ. Tech. Rep*.

Zhang, J., B. Song, Z. Zhang, and H. Liu (2014). An approach for modeling vulnerability of the network of networks. *Physica A: Statistical Mechanics and its Applications 412*, 127–136.

Zio, E. and G. Sansavini (2010, 09). Modeling failure cascades in critical infrastructures with physically-characterized components and interdependencies. *ESREL 2010 Annual Conference*, 652 – 661.

Zio, E. and G. Sansavini (2011, 08). Modeling cascading failures in systems of systems with uncertain behavior. *ICASP11*.

Paper III

A. Bellè, A. F. Abdin, Z. Zeng, Y.-P. Fang, and A. Barros, "A mathematical framework for the optimal coupling of interdependent critical infrastructures," *Reliability Engineering & System Safety*, under review, 2022.

A resilience-based framework for the optimal coupling of interdependent critical infrastructures

Andrea Bellè[1]*, Adam F. Abdin[2], Zhiguo Zeng[1], Yi-Ping Fang[1], Anne Barros[1]

[1] *Chair on Risk and Resilience of Complex Systems, Laboratoire Génie Industriel, CentraleSupélec, Université Paris-Saclay, 3 Rue Joliot Curie, 91190 Gif-sur-Yvette, France.*
Emails: andrea.belle@centralesupelec.fr, zhiguo.zeng@centralesupelec.fr, yiping.fang@centralesupelec.fr, anne.barros@centralesupelec.fr

[2] *Laboratoire Génie Industriel, CentraleSupélec, Université Paris-Saclay, 3 Rue Joliot Curie, 91190 Gif-sur-Yvette, France.*
Email: adam.abdin@centralesupelec.fr

Abstract

As critical infrastructures (CIs) are essential for the safety and socio-economic stability of a society, ensuring their resilience is a task of the utmost importance. Critical infrastructures are often interdependent on each other, and the topology of the interdependencies between different systems, also referred to as coupling interface, plays a key role in terms of their resilience against failures. In case of failures due natural events, random disturbances, or deliberate attacks, the design of the coupling interface is a key factor for maintaining high performance within the interdependent CIs. However, in the existing literature, the issue of the coupling interface design is often addressed through heuristics. In this work, we propose an optimization-based mathematical approach for designing coupling interfaces between interdependent critical infrastructures under random failures. The proposed approach allows designing a coupling interface that is robust against the worst realization of a set of feasible failure scenarios. Using as case-study interdependent power and gas networks, we show that the proposed method outperforms existing solutions based on network metrics-based heuristics.

Keywords: Coupling interface, resilience, defender-attacker-defender, interdependent critical infrastructures, optimization

List of Symbols

Abbreviations

CI Critical infrastructure

*Corresponding author
Email address: andrea.belle@centralesupelec.fr (Andrea Bellè[1])

GN Gas network

IPGNs Interdependent power and gas networks

NC&CG Nested Column&Constraint Generation

PN Power network

Sets

E_{GN} Set of edges in the gas network

E_{PN} Set of edges in the power network

V_{GN} Set of nodes in the gas network

V_{PN} Set of nodes in the power network

Parameters and coefficients

$\overline{\theta}$ Maximum value of phase angle

$\overline{d}_i^b$ Base requested power demand of node i in the power or gas network

$\overline{d}_j^{m^3}$ Requested gas demand of node j in the gas network

$\overline{d}_j^{MW}$ Requested power demand of node j in the gas network

$\overline{d}_{GN}$ Total requested gas demand of the gas network

$\overline{d}_{PN}$ Total requested power demand of the power network

$\overline{f}_k$ Flow capacity of edge k in the power or gas network

$\overline{p}_i$ Production capacity of node i in the power or gas network

K_{att} Maximum number of attacked edges

L_{GN} Number of edges in the gas network

L_{PN} Number of edges in the power network

M_k Big-M method constant

N_{GN} Number of nodes in the gas network

N_{PN} Number of nodes in the power network

w_{GN} Weight of the gas network

w_{PN} Weight of the power network

x_k Reactance of edge k in the power network

Variables

δ_j^g Binary variable that indicates the functional state of all the interdependency links starting from node j in the gas network

δ_j^p Binary variable that indicates the functional state of all the interdependency links starting from node j in the power network

η Variable of outer layer of NC&CG algorithm

ρ Variable of inner layer of NC&CG algorithm

θ_i Phase angle of node i in the power network

d_i Supplied demand in node i in the power or gas network

f_k Flow in edge k in the power or gas network

p_i Production in node i in the power or gas network

u_k Binary variable that indicates the functional state of edge k in the power network

$y_{ij}^{g \leftarrow p}$ Binary variable that indicates if a physical link from node $j \in V_{PN}$ to node $j \in V_{GN}$ exists

$y_{ij}^{p \leftarrow g}$ Binary variable that indicates if a physical link from node $j \in V_{GN}$ to node $i \in V_{PN}$ exists

1. Introduction

1.1. Motivation

Critical infrastructures (CIs), such as power networks or transportation systems, are complex systems which supply goods, services, and commodities to people [1], [2]. Failures and disruption within CIs can lead to severe socioeconomic stress in a society [3], and ensuring their resilience against a large variety of disruptive events is an important issue [4], [5]. Moreover, CIs are increasingly interdependent on each other. This increasing degree of interdependency brings advantages in terms of functionality and efficiency, but often leads to new vulnerabilities and risks of cascading effect between interdependent infrastructures [6].

Coupling interfaces play a key role in characterizing the resilience of interdependent CIs [7], [8]. The coupling interface characterizes how the interdependent CIs are coupled together; in other words, it characterizes how the interdependent CIs are connected and what are the components in

each CI that are dependent on the other CI. When CIs are modeled as networks [9], the coupling interface simply denotes the allocation of interdependency links, as shown in Figure 1.

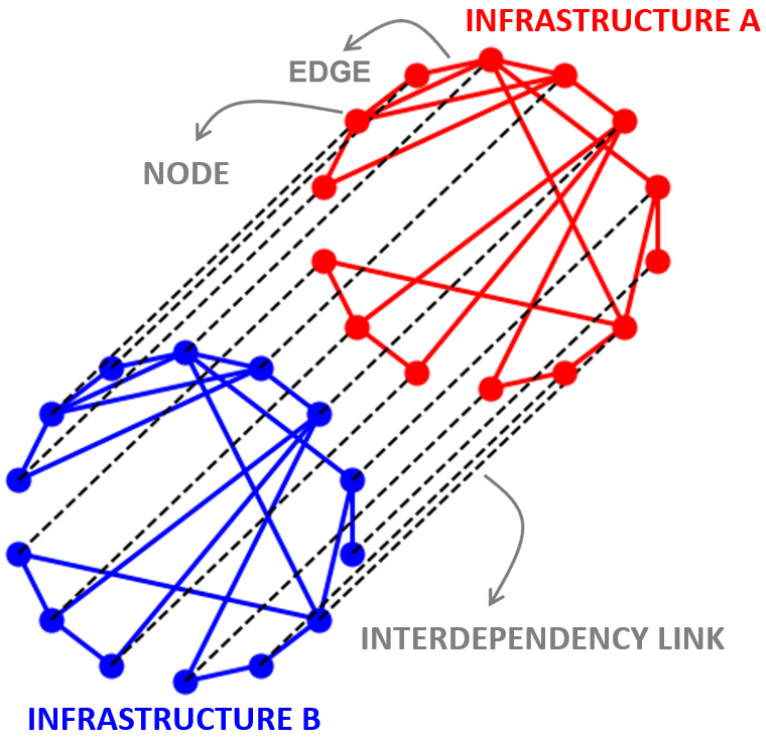

Figure 1: Network representation of two interdependent infrastructures. In evidence, we can notice the two structural components of networks (nodes and edges) and their connection (interdependency links). The ensemble of interdependency links is referred to as *coupling interface*.

In most of the existing literature on interdependent CIs, coupling interfaces are treated as a known parameter, and no optimization nor analysis is performed. Limited works try to optimize the design of the coupling interface (e.g. [7] or [8]); however, they rely on heuristic methods based on network science metrics, which do not guarantee optimal solutions nor high quality designs of coupling interface.

In this work, we propose a resilience-based mathematical framework, based on the *defender-attacker-defender* (DAD) model [10]–[13], for the optimal design of coupling interfaces in interdependent CIs. The DAD approach allows to identify solutions, in this case a coupling interface design, which are robust against the worst realization of uncertain scenarios, in this case failure scenarios.

In general, the motivations of this work are the following:

- *research*: the design of coupling interfaces between interdependent CIs has not been addressed comprehensively in the existing literature, and to the best of our knowledge, no mathemat-

ical programming approach has been proposed. As the coupling interface is a key factor of interdependent CIs and their resilience, optimizing its design is an important issue;

- *application*: due to the importance of coupling interface design, decision-makers and planners should be provided with the means and tools to evaluate and optimize the allocation of interdependency links between interdependent CIs.

As illustrative case-study, we rely on interdependent power and gas networks (IPGNs), similarly to [14], where gas networks need electricity for the functionality of their equipment (valves, pumps, compressors, etc.), and power networks need a gas supply to produce electricity in gas-fired power plants.

1.2. Related work

In the next sections we review the main works related to resilience enhancement in CIs, and design and optimization of coupling interfaces between interdependent CIs.

1.2.1. Resilience enhancement in critical infrastructures

The purpose of this section is to explain the main concepts in the context of resilience enhancement and give a general overview in order to better contextualize and position this work.

As critical infrastructures represent the backbone of essential societal functions, ensuring their resilience is a fundamental task [2]. The resilience of a system is defined as *"its ability to withstand stressors, adapt, and rapidly recover from disruptions"* [15]. Resilience refers to the behaviour of a system in disruptive conditions, it is generally represented with a resilience curve, as in Figure 2, and it is defined as the combination of three phases [2], [16]:

- the disturbance phase, which describes the speed and the severity of the disruption; this phase is strictly connected to the concepts of survivability and vulnerability[1];
- the degraded phase, which describes the temporal extension of the disruption after the disturbance phase, and it is linked to the emergency preparedness;
- the restoration phase, which describes the operations of restoration and repair.

The resilience of a system can be measured using different approaches, and various metrics are available in the existing literature [20]. A renowned approach is called $\Phi\Lambda E\Pi$ (pronounced "FLEP") [16], and it consists of the computation of four different metrics:

[1]Survivability is defined in [17] as *"the capability of a system to fulfill its mission in a timely manner in the presence of attacks, failures, or accidents"*, and it can be interpreted as the residual performance after the disturbance phase. Vulnerability is defined in [18] as *"degree of loss or damage to a system when exposed to a strain of a given type and magnitude"*, and it can be interpreted as the drop of performance due to the disturbance phase.

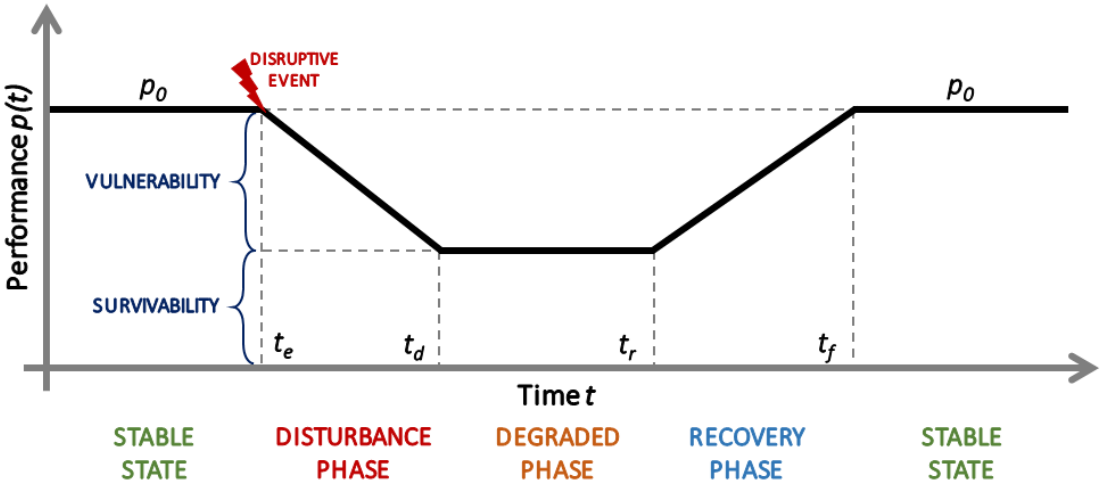

Figure 2: Qualitative representation of a resilience curve and the related phases [16], [19]

- Φ: it defines the rate of performance drop during the disturbance phase. Using Figure 2 as a reference, where $p(t)$ defines a performance indicator at time t, it can be computed as the difference in performance before and after the disruptive event divided by the duration of the event, as in Equation (1):

$$\Phi = \frac{p(t_e) - p(t_d)}{t_d - t_e}; \tag{1}$$

- Λ: it defines the magnitude of the drop in performance. This metric corresponds to the concept of vulnerability, and it strictly correlated with the survivability. It can be computed as the difference in performance before and after the disruptive event, as in Equation in (2)

$$\Lambda = p(t_e) - p(t_d); \tag{2}$$

- E: it defines the temporal extension of the degraded phase, and it can be computed as in (3):

$$E = t_r - t_d; \tag{3}$$

- Π: it defines the rate of recovery, and it can be computed as the difference in performance at the beginning and at the end of the recovery phase, divided by the duration of the recovery, as in Equation in (4):

$$\Phi = \frac{p(t_f) - p(t_r)}{t_f - t_r}. \tag{4}$$

Enhancing the resilience of systems and infrastructures by optimizing design, preventive measures and resource allocation (e.g. transmission and/or generation expansion, protection of components, allocation of recovery resources, reliable network design, etc.), is one of the most important

tasks and a major topic in the field of critical infrastructures. Within this context, several works are available, and they can be distinguished according to different characteristics: *i)* which resilience phase is optimized; *ii)* which type of infrastructures is optimized; *iii)* which type of optimization model is used.

The optimization of critical infrastructures resilience can focus on one or multiple phases: for example, in [11] and [21], the resilience of power networks is enhanced by focusing separately on the optimization of protection against the disturbance phase and recovery phase, respectively; on the contrary, in [13] and [22], the resilience of interdependent CIs is enhanced by simultaneously optimizing both the disturbance phase and recovery phase.

An important feature that distinguishes the different works is which type of infrastructure is optimized, in terms of resilience. Several authors focus on resilience of single infrastructures, such as power networks [23] or water networks [24]. However, many other authors focus on the integrated optimization of resilience of multiple interdependent CIs, such as power and gas networks [25] or power and water networks [26], accounting for their mutual interdependencies when optimizing their resilience by preventive measures and resource allocation. The type of infrastructure under consideration is a key factor, as each infrastructure is characterized by specific operational models and interdependencies on other systems.

Another important difference within the existing works is the type of optimization model used for the resilience enhancement, which strongly impacts the quality and the nature of the solution. Many authors apply multi-level approaches, such as the DAD model [27], to enhance CIs resilience. These approaches offer robust solutions, and usually affordable computational cost. Some authors also include uncertainty using a stochastic optimization approach [28], in order to enhance the resilience expectation against a known probability distribution of uncertain parameters. Moreover, heuristics approach are also used [29], in order to derive high-quality solutions with operational models which can not be solved by traditional mathematical programming approaches.

It should be highlighted than in the aforementioned works the resilience of CIs is enhanced by optimizing different preventive measures and resource allocations, such as construction of new components (generation/expansion planning), protection of components or repair scheduling. However, the coupling interface, despite being a key parameter, is not optimized. As it is explained in the next section, only a limited number of works accounts for different coupling interface designs between interdependent CIs.

1.2.2. Design and optimization of coupling interface

When the state/functionality of one infrastructure depends on the state/functionality of another one, a relationship of interdependency exists. Interdependencies are unidirectional when one infras-

tructure depends on another one, but not vice versa; otherwise, they are bidirectional [30], [31]. As interdependencies have been a subject of research since the early 2000s [2], different classifications exist in the literature [30]–[33]. One of the most used classifications is the one proposed in [30], where four categories are identified:

- *physical*, when one CI depends on another one through a physical flow (energy, goods, etc.);
- *cyber*, when one CI depends on another one through a flow of data and information;
- *geographic*, when elements of different infrastructures share the same location and they can be modified by a change in the environment conditions;
- *logical*, when a relationship which is not physical, cyber, or geographic exist.

CIs are often modeled with a network science approach [9], and the interdependencies are represented as links between components (nodes and/or edges) belonging to different infrastructures [34]. We refer to the ensemble of interdependency links as coupling interface. Its topology, i.e. where the interdependency links are present, plays a key role in terms of failure propagation between different infrastructures. Interdependency topology and design have been addressed in the field of interdependent networks, where various works focus on evaluating coupling interfaces and their impact on failure propagation [35], [36], and how coupling interfaces, if properly allocated, can increase the robustness of interdependent networks [37]–[39]. These works, despite representing a solid theoretical framework, mainly rely on percolation theory, and they fail to capture the details and the complex dynamics of real-world infrastructures.

Despite the critical role of coupling interfaces, in the existing literature they are often considered as a given parameter, and they are not analyzed nor optimized.

In some works, different network metric-based coupling strategies are tested on different interdependent CIs, such as power and water networks [40] or power and telecommunication networks [41]–[43]. In these works, the impact of different topologies is evaluated, and they demonstrate the importance of considering the coupling interface design problem within realistic CIs. However, these network-based heuristic approaches do not guarantee optimal solutions.

Similar network metrics-based approaches are also proposed in [7] and [8]. In [7], the authors propose an approach for designing coupling interfaces between urban CIs in order to increase their robustness against external attacks. The proposed strategy for designing the coupling interface is based on multiple network metrics (node degree, betweenness, clustering coefficient and Euclidean distance). In [8], the authors propose a similar approach, also accounting for physical features of the CIs, such as levels of supply and demand. However, these works still rely on network metrics as an heuristics. Consequently, they do not guarantee optimal solutions and the quality of the identified

coupling interface designs depends on the case-study considered. Moreover, these approaches are tailor-made and are not readily generalizable to other case-studies, as one specific heuristic strategy might perform well in some networks and poorly for other systems.

1.3. Contribution

In this work, a novel optimization-based approach for designing coupling interfaces between interdependent CIs is proposed. Our model ensures that coupling interface topologies are optimized in order to maximize the worst-case realization of combined performed of the interdependent infrastructures under random failures. The proposed approach is based on the DAD model, a three-stage sequential game which allows to identify robust defense strategies and/or resource allocation against a defined set of feasible attack scenarios. To demonstrate the validity of our approach, interdependent power and gas networks (IPGNs) are used as illustrative case-study.

The contributions of this papers can be summarized as follows:

- We developed a novel resilience-based optimization approach, which can be directly applied to design or retrofit new or existent coupling interfaces between interdependent CIs.
- We developed an approach for the optimization of coupling interface design that is generalizable for any case-study by selecting the appropriate operational model for the interdependent CIs.
- We demonstrated that our approach outperforms network metrics-based coupling interface strategies available in the existing literature.

The rest of this paper is organized as follows: in Section 2, the problem formulation is detailed; in Section 3, the solution strategy is explained; in Section 4, the illustrative case-study is detailed; in Section 5, results and discussion are presented; in Section 6, conclusive remarks and possible future developments are detailed.

2. Optimization problem formulation

2.1. Modeling framework

In this work, each infrastructure is modeled using a network flow-based approach [9], [44], where a network is a mathematical construct described by a graph $G = (V, E)$. The set V contains N nodes, connected by L edges, contained within the set E. Each edge k is directed and has an origin node $O(k)$ and a destination node $D(k)$. In line with a flow-based approach, we assume that commodities goods, and services are produced and consumed within nodes and distributed through

edges. Each node i has a production capacity $\overline{p}_i$ and a requested demand $\overline{d}_i$, while each edge k has a flow capacity $\overline{f}_k$.

In this work, we focus on the combined performance P_C of the interdependent CIs [14], defined as in (5):

$$P_C = \sum_{h \in H} \frac{w_h}{\overline{d}_h} \sum_{i \in V_h} d_i \tag{5}$$

where the subscripts H denotes the set of interdependent CIs, w_h represents the weight of infrastructure h when computing the combined performance, $\overline{d}_h$ is the total requested demand of goods, services, or commodities in infrastructure h, and d_i is the supplied demand of goods, services, or commodities in each node i of infrastructure h.

Considering the resilience framework described in Section 1.2.1, the combined performance in conditions of disruption represents the concept of survivability of the interdependent CIs, complementary to the concept of vulnerability and to the Λ metric of the $\Phi\Lambda E\Pi$ approach. In this work, we do not consider the restoration phase, as it is characterized by deep uncertainties and it should be optimized case-by-case according to the specific disruption and failure scenarios [21].

As illustrative case-study, we consider interdependent power and gas networks (IPGNs), which are mutually interdependent on each other with physical interdependencies. In fact, equipment in the gas network, such as valves, compressors, or pumps, needs a constant power supply; power networks, if gas-fired power plants are present, need a constant supply of gas. The combined performance of the IPGNs can be defined as in Equation (6):

$$P_{C,IPGNs} = \frac{w_{PN}}{\overline{d}_{PN}} \sum_{i \in V_{PN}} d_i + \frac{w_{GN}}{\overline{d}_{GN}} \sum_{i \in V_{GN}} d_i \tag{6}$$

where the subscripts PN and GN denote the power network and gas network, respectively, w_{PN} and w_{GN} represent the weight of power network and gas network when computing the combined performance[2], $\overline{d}_{PN}$ and $\overline{d}_{GN}$ are the total requested demand of power and gas, and d_i is the supplied power or gas in each node of the networks. The combined performance P_C ranges from 0, when no power and gas demand is supplied, to 1, when 100% of the requested demand of power and gas is supplied.

In the power network, nodes represent buses, while edges represent power lines; in the gas network, nodes represent hubs, while edges represent gas pipelines. The power network operations are simulated with a DC power flow model, while the gas network operations are simulated with a linear maximal flow model, which is a suitable approximation of flow-based infrastructures [14], [45]–[47].

[2]It should be noted that $w_{PN} + w_{GN} = 1$.

Several works analyze critical infrastructures in the context of specific types of hazards, like intentional attacks [46], spatially-localized attacks [48] and extreme natural events [14], [49]. In this work, we adopt an approach based on the maximum number of contingencies [27], [50]. For simplicity, but without loss of generality, we assume that only transmission lines (edges) in the power network can be attacked and failed. By considering the simultaneous failures of transmission lines, the present model is agnostic about the source of disruption, providing a rapid and objective way of calculating the consequence of damage to any set of components.

In this work, the following assumptions are considered:

- a single demand scenario is considered, i.e. the expected forecast of requested power and gas demand [27];
- each node in the gas network needs to receive a power supply from the power network in order to run equipment;
- each node in the power network with some production capacity is assumed to contain a gas-fired power plant and needs to receive a gas supply from the gas network;
- each node in the power network can be dependent on one, and only one node in the gas network, and vice versa;
- allocating the coupling interface has a cost that depends on the geographical distance between the two nodes connected by the interdependency link;
- the operators are perfectly aware of the status of the components within the power network and gas network [27].

The purpose of the proposed model is to design a coupling interface between IPGNs that ensures satisfactory combined performance in normal conditions (no failures) and conditions of disruption.

2.2. Defender-attacker-defender approach

The problem takes the form of a trilevel DAD optimization model, a formulation often used in the framework of optimization of defense strategies and resources in CIs (e.g. [10], [14], [27]). It is useful to imagine the problem as a three-players game: the inner defender aims at maximizing the combined performance of the IPGNs through the operational variables of the two systems; the middle attacker aims at minimizing the combined performance choosing the most disruptive attack plan; the outer defender aims at maximizing the combined performance of the IPGNs by designing a robust coupling interface that also ensures satisfactory performance in normal conditions (no failures). The full formulation is shown in (7)-(43):

$$\max_{\substack{\mathbf{p}',\mathbf{d}',\mathbf{f}',\boldsymbol{\theta}',\boldsymbol{\delta}' \\ \mathbf{y}^{g\leftarrow p}\in\{0,1\}^{N_C} \\ \mathbf{y}^{p\leftarrow g}\in\{0,1\}^{N_C}}} \min_{\mathbf{u}\in\{0,1\}^{L_{PN}}} \max_{\mathbf{p},\mathbf{d},\mathbf{f},\boldsymbol{\theta},\boldsymbol{\delta}} \frac{w_{PN}}{\overline{d}_{PN}} \sum_{i\in V_{PN}} d_i + \frac{w_{GN}}{\overline{d}_{GN}} \sum_{i\in V_{GN}} d_i$$

$$-\gamma \left(\sum_{\substack{i\in V_{GN} \\ j\in V_{PN}}} y_{ij}^{g\leftarrow p}\, d_{ij}^{km}\, c_{km}^{g\leftarrow p} + \sum_{\substack{i\in V_{PN} \\ j\in V_{GN}}} y_{ij}^{p\leftarrow g}\, d_{ji}^{km}\, c_{km}^{p\leftarrow g} \right) \tag{7}$$

subject to:

First level

$$\sum_{j\in V_{PN}} y_{ij}^{g\leftarrow p} \leq 1, \quad \forall i \in V_{GN} \tag{8}$$

$$\sum_{j\in V_{GN}} y_{ij}^{p\leftarrow g} \leq 1, \quad \forall i \in V_{PN} \tag{9}$$

$$\sum_{\substack{i\in V_{GN} \\ j\in V_{PN}}} y_{ij}^{g\leftarrow p}\, d_{ij}^{km}\, c_{km}^{g\leftarrow p} + \sum_{\substack{i\in V_{PN} \\ j\in V_{GN}}} y_{ij}^{p\leftarrow g}\, d_{ji}^{km}\, c_{km}^{p\leftarrow g} \leq B_{ci} \tag{10}$$

$$\frac{w_{PN}}{\overline{d}_{PN}} \sum_{i\in V_{PN}} d_i + \frac{w_{GN}}{\overline{d}_{GN}} \sum_{i\in V_{GN}} d_i \geq 1 \tag{11}$$

$$0 \leq p'_i \leq \overline{p}_i, \quad \forall i \in V_{TOT} \tag{12}$$

$$0 \leq d'_i \leq \overline{d}_i^{\,b} + \sum_{j\in V_{GN}} y_{ji}^{g\leftarrow p}\, \overline{d}_j^{\,MW}, \quad \forall i \in V_{PN} \tag{13}$$

$$0 \leq d'_i \leq \overline{d}_i^{\,b} + \sum_{j\in V_{PN}} y_{ji}^{p\leftarrow g}\, \overline{d}_j^{\,m^3}, \quad \forall i \in V_{GN} \tag{14}$$

$$-\overline{f}_k \leq f'_k \leq \overline{f}_k, \quad \forall k \in E_{TOT} \tag{15}$$

$$x_k f'_k - (\theta'_{O(k)} - \theta'_{D(k)}) = 0, \quad \forall k \in E_{PN} \tag{16}$$

$$p'_i - d'_i + \sum_{k|D(k)=i} f'_k - \sum_{k|O(k)=i} f'_k = 0, \quad \forall i \in V_{TOT} \tag{17}$$

$$d_i - \delta_i^{p'} \left(\overline{d_i^b} + \sum_{j \in V_{GN}} y_{ji}^{g \leftarrow p} \, \overline{d}_j^{MW} \right) \geq 0, \quad \forall i \in V_{PN} \tag{18}$$

$$d'_i - \delta_i^{g'} \left(\overline{d_i^b} + \sum_{j \in V_{PN}} y_{ji}^{p \leftarrow g} \, \overline{d}_j^{m^3} \right) \geq 0, \quad \forall i \in V_{GN} \tag{19}$$

$$p'_i - \overline{p}_i \sum_{j \in V_{GN}} y_{ij}^{p \leftarrow g} \, \delta_j^{g'} \leq 0, \quad \forall i \in V_{PN} \tag{20}$$

$$p'_i - \overline{p}_i \sum_{j \in V_{PN}} y_{ij}^{g \leftarrow p} \, \delta_j^{p'} \leq 0, \quad \forall i \in V_{GN} \tag{21}$$

$$d'_i - \left(\overline{d_i^b} + \sum_{j \in V_{PN}} y_{ji}^{p \leftarrow g} \, \overline{d}_j^{m^3} \right) \sum_{j \in V_{PN}} y_{ij}^{g \leftarrow p} \, \delta_j^{p} \leq 0, \quad \forall i \in V_{GN} \tag{22}$$

$$- \sum_{\substack{k | O(k=i) \\ j \in V_{PN}}} y_{ij}^{g \leftarrow p} \, \delta_i^{p'} \, \overline{f}_k \leq f'_k \leq \sum_{\substack{k | O(k=i) \\ j \in V_{PN}}} y_{ij}^{g \leftarrow p} \, \delta_i^{p'} \, \overline{f}_k, \quad \forall k \in E_{GN} \tag{23}$$

$$- \sum_{\substack{k | D(k=i) \\ j \in V_{PN}}} y_{ij}^{g \leftarrow p} \, \delta_i^{p'} \, \overline{f}_k \leq f'_k \leq \sum_{\substack{k | D(k=i) \\ j \in V_{PN}}} y_{ij}^{g \leftarrow p} \, \delta_i^{p'} \, \overline{f}_k, \quad \forall k \in E_{GN} \tag{24}$$

$$y_{ji}^{g \leftarrow p} \in \{0,1\}, \; y_{ij}^{p \leftarrow g} \in \{0,1\}, \quad \forall i \in V_{PN}, \forall j \in V_{GN} \tag{25}$$

$$\delta_i^{p'} \in \{0,1\}, \; \delta_j^{g'} \in \{0,1\}, \quad \forall i \in V_{PN}, \forall j \in V_{GN} \tag{26}$$

Second level

$$\sum_{k \in E_{PN}} (1 - u_k) \leq K_{att} \tag{27}$$

$$u_k \in \{0,1\}, \quad \forall k \in E_{PN} \tag{28}$$

Third level

$$0 \leq p_i \leq \overline{p}_i, \quad \forall i \in V_{TOT} \tag{29}$$

$$0 \leq d_i \leq \overline{d}_i^b + \sum_{j \in V_{GN}} y_{ji}^{g \leftarrow p} \, \overline{d}_j^{MW}, \quad \forall i \in V_{PN} \tag{30}$$

$$0 \leq d_i \leq \overline{d}_i^b + \sum_{j \in V_{PN}} y_{ji}^{p \leftarrow g} \, \overline{d}_j^{m^3}, \quad \forall i \in V_{GN} \tag{31}$$

$$-u_k \overline{f}_k \leq f_k \leq u_k \overline{f}_k, \quad \forall k \in E_{PN} \tag{32}$$

$$-\overline{f}_k \leq f_k \leq \overline{f}_k, \quad \forall k \in E_{GN} \tag{33}$$

$$\left(x_k f_k - \left(\theta_{O(k)} - \theta_{D(k)}\right)\right) u_k = 0, \forall k \in E_{PN} \tag{34}$$

$$p_i - d_i + \sum_{k|D(k)=i} f_k - \sum_{k|O(k)=i} f_k = 0, \;\; \forall i \in V_{TOT} \tag{35}$$

$$d_i - \delta_i^p \left(\overline{d}_i^b + \sum_{j \in V_{GN}} y_{ji}^{g \leftarrow p} \, \overline{d}_j^{MW}\right) \geq 0, \quad \forall i \in V_{PN} \tag{36}$$

$$d_i - \delta_i^g \left(\overline{d}_i^b + \sum_{j \in V_{PN}} y_{ji}^{p \leftarrow g} \, \overline{d}_j^{m^3}\right) \geq 0, \quad \forall i \in V_{GN} \tag{37}$$

$$p_i - \overline{p}_i \sum_{j \in V_{GN}} y_{ij}^{p \leftarrow g} \, \delta_j^g \leq 0, \quad \forall i \in V_{PN} \tag{38}$$

$$p_i - \overline{p}_i \sum_{j \in V_{PN}} y_{ij}^{g \leftarrow p} \, \delta_j^p \leq 0, \quad \forall i \in V_{GN} \tag{39}$$

$$d_i - \left(\overline{d}_i^b + \sum_{j \in V_{PN}} y_{ji}^{p \leftarrow g} \, \overline{d}_j^{m^3}\right) \sum_{j \in V_{PN}} y_{ij}^{g \leftarrow p} \, \delta_j^p \leq 0, \quad \forall i \in V_{GN} \tag{40}$$

$$-\sum_{\substack{k|O(k=i) \\ j \in V_{PN}}} y_{ij}^{g \leftarrow p} \, \delta_i^p \, \overline{f}_k \leq f_k \leq \sum_{\substack{k|O(k=i) \\ j \in V_{PN}}} y_{ij}^{g \leftarrow p} \, \delta_i^p \, \overline{f}_k, \quad \forall k \in E_{GN} \tag{41}$$

$$-\sum_{\substack{k|D(k=i) \\ j \in V_{PN}}} y_{ij}^{g \leftarrow p} \, \delta_i^p \, \overline{f}_k \leq f_k \leq \sum_{\substack{k|D(k=i) \\ j \in V_{PN}}} y_{ij}^{g \leftarrow p} \, \delta_i^p \, \overline{f}_k, \quad \forall k \in E_{GN} \tag{42}$$

$$\delta_i^p \in \{0,1\}, \; \delta_j^g \in \{0,1\}, \quad \forall i \in V_{PN}, \forall j \in V_{GN}. \tag{43}$$

Equation (7) is the objective function of the trilevel optimization problem, and it contains three terms. The first two terms correspond to the combined performance P_C, previously shown in

Equation (6). By including P_C in the objective function, we can identify a coupling interface that maximizes the combined performance of the IPGNs in the worst failure scenario; in other words, we can identify the coupling interface that maximizes the survivability of the IPGNs (or minimizes the Λ resilience metric) of the IPGNs in the worst failure scenario. The power and gas supplied to each node i are defined by the variables d_i, while the total requested demand of power and gas, denoted as $\overline{d}_{PN}$ and $\overline{d}_{GN}$, are constant parameters computed as in (44) and (45), respectively.

$$\overline{d}_{PN} = \sum_{i \in V_{PN}} \overline{d}_i^b + \sum_{j \in V_{GN}} \overline{d}_j^{MW} \tag{44}$$

$$\overline{d}_{GN} = \sum_{i \in V_{GN}} \overline{d}_i^b + \sum_{j \in V_{PN}} \overline{d}_j^{m^3} \tag{45}$$

In these equations, the constant $\overline{d}_i^b$ denotes the baseline requested demand of power or gas in each node, and it represents the consumption of various private and public consumers. The constant $\overline{d}_j^{MW}$ denotes the requested power demand of node $j \in V_{GN}$, while the constant $\overline{d}_j^{m^3}$ denotes the requested gas demand of node $j \in V_{PN}$.

The third term of the objective function ensures that, if more than one optimal coupling interface exists, the one with the lowest allocation cost is chosen. The terms within the parentheses define the cost of allocating a specific coupling interface. The binary variable $y_{ij}^{g \leftarrow p}$=1 if an interdependency link from node $j \in V_{PN}$ to node $i \in V_{GN}$ is allocated, and $y_{ij}^{g \leftarrow p}$=0 otherwise. Similarly, the binary variable $y_{ij}^{p \leftarrow g}$=1 if an interdependency link from node $j \in V_{GN}$ to node $i \in V_{PN}$ is allocated, and $y_{ij}^{p \leftarrow g}$=0 otherwise. The constant d_{ij}^{km} denotes the distance in kilometer between node $i \in V_{GN}$ and node $j \in V_{PN}$, while the constants $c_{km}^{g \leftarrow p}$ and $c_{km}^{p \leftarrow g}$ denote the cost per kilometer of allocating an interdependency link from the power network to the gas network, and from the gas network to the power network, respectively. The terms within the parentheses are multiplied by a factor γ, which represents a very small number. This factor ensures that the priority within the optimization is given to the combined performance P_C.

Equations (8)-(26) denote the constraints of the first optimization level, corresponding to the outer defender. This agent allocates the coupling interface in a way such that: *i)* the available monetary budget B_{ci} is respected, as shown in Constraint (10), and *ii)* in normal conditions (no failures), it is possible to supply the whole requested demand of power and gas (P_C=1). Consistently with the existing literature, we assume that each node in the gas network can be dependent on, and connected through an interdependency link to, only one node in the power network, and vice versa. We refer to this as the *single-dependency* assumption, and it is enforced by Constraints (8) and (9). The coupling interface, as previously explained, is allocated through the binary variables $y_{ij}^{g \leftarrow p}$ and $y_{ij}^{p \leftarrow g}$, contained within the vectors $\mathbf{y}^{g \leftarrow p}$ and $\mathbf{y}^{p \leftarrow g}$ with dimension $N_C = N_{PN} \times N_{GN}$.

The coupling interface must be allocated in order to guarantee that, in normal conditions, the requested demand of power and gas is fully satisfied, as enforced by Constraint (11). This condition depends on the first-level operational variables, contained within the vectors $\mathbf{p'}, \mathbf{d'}, \mathbf{f'}, \boldsymbol{\theta'}, \boldsymbol{\delta'}$, which represent production levels, supply demands, flows, phase angles, and interdependency links status, respectively[3].

Equations (12)-(26) contain the operational constraints of the first level. For both networks, the production level of power or gas p'_i in each node i is limited by the production capacity $\overline{p}_i$, as enforced in Constraint (12). Similarly, as shown in Constraints (13) and (14), the supplied demand of power or gas d'_i in each node i is limited by the requested demand. As it is shown on the right side of (13), the requested power demand of node $i \in V_{PN}$ is given by the sum of the baseline requested power demand $\overline{d}_i^b$ and all the requested power demands $\overline{d}_j^{MW}$ of the nodes $j \in V_{GN}$ which depend on the node $i \in V_{PN}$ for the electricity supply ($y_{ji}^{g \leftarrow p}$=1). Similarly, as it is shown on the right side of (14), the requested gas demand of node $i \in V_{GN}$ is given by the sum of the baseline requested gas demand $\overline{d}_i^b$ and all the requested gas demands $\overline{d}_j^{m^3}$ of the nodes $j \in V_{PN}$ which depend on the node $i \in V_{GN}$ for the gas supply ($y_{ji}^{p \leftarrow g}$=1).

The flow of power and gas f'_k in each edge k is limited, in absolute value, by the flow capacity $\overline{f}_k$, as shown in Constraint (15). Moreover, in each line of the power network, the power flow is subject to the DC power flow assumption, enforced by Constraint (16), where x_k represents the reactance of line k, and $\theta'_{O(k)}$ and $\theta'_{D(k)}$ are the phase angles in the origin and destination node of line k, respectively.

The net nodal balance of power and gas in each node is ensured by Constraint (17).

The operations of the IPGNs depends on the status of the interdependency links. Similarly to other existing works (e.g. [14]), we assume a binary functional status for the interdependency links (1 if functional, 0 if not functional). We assume that the binary functional status of each interdependency link starting from node $i \in V_{PN}$ is expressed by the binary variable $\delta_i^{p'}$; similarly, the binary functional status of each interdependency link starting from node $i \in V_{GN}$ is expressed by the binary variable $\delta_i^{g'}$. Each interdependency link starting from node $i \in V_{PN}$ is functional ($\delta_i^{p'}$=1) only if the requested power demand in i is fully satisfied, as enforced in Constraint (18). The rationale behind this assumption is that, if some electricity is not supplied to i, the dependent nodes within the gas network might not receive the necessary electricity. As shown in Constraint (19), the same assumption is taken for the interdependency link starting from the gas network, with a similar rationale: each interdependency link starting from node $i \in V_{GN}$ is functional ($\delta_i^{g'}$=1) only if the requested gas demand in i is fully satisfied. These assumption are consistent with the

[3]The superscript $'$ denotes the operational variables of the first level.

existing literature (e.g. [14]). However, different assumptions which are not included in this work, such as multi-discrete or continuous status for the interdependency links, can be implemented by appropriate changes of the variables $\boldsymbol{\delta}'$.

We assume that the electricity in the power network is produced by gas-fired power plants, and in each node $i \in V_{PN}$ it is possible to produce power only if a functional interdependency link with a node $j \in V_{GN}$ is present ($y_{ij}^{p \leftarrow g}$=1 and $\delta_j^{g'}$=1). This condition is enforced by Constraint (20).

We assume that gas in the gas network can be extracted (produced) and supplied only if there is enough electricity. Therefore, in each node $i \in V_{GN}$ it is possible to produce and supply gas only if a functional interdependency link with a node $j \in V_{PN}$ is present ($y_{ij}^{g \leftarrow p}$=1 and $\delta_j^{p'}$=1). These conditions are enforced by Constraints (21) and (22). Moreover, we assume that gas can flow in a pipe k only if both the origin and destination nodes present a functional interdependency link with a node $j \in V_{PN}$, as enforced by Constraints (23) and (24).

Equations (27) and (28) denote the constraints of the second level of the optimization problem, corresponding to the attacker. This agent decides which lines of the power network to target and fail through the binary variables u_k, contained within the vector $\mathbf{u}$. Each variable u_k takes the value 0 if line k is targeted and failed, and value 1 otherwise. The attacker can target and fail a maximum number K_{att} of lines in the power network, as shown in Constraint (27).

Equations (29)-(43) contain the operational constraints of the third level, corresponding to the inner defender. This agent aims at maximizing the combined performance of the IPGNs through the operational variables of the third level, contained within the vectors $\mathbf{p}, \mathbf{d}, \mathbf{f}, \boldsymbol{\theta}, \boldsymbol{\delta}$.

Constraints (29)-(43) are equivalent to the previously-explained Constraints (8)-(26). However, in the third level, we also account for the failures of power lines through the inclusion of binary variables u_k in Constraints (32) and (34). Constraint (32) ensures that the power flow in a failed power line is 0. Constraints (34) ensures that the DC power flow assumption is maintained in functional power lines and disregarded in failed power lines. Constraint (34) contains quadratic terms due to the multiplication of the binary variable u_k with the continuous variables f_k and θ_i. These quadratic terms can be linearized with a "Big-M" approach, as shown in Appendix A. The other constraints in (29)-(43) are equivalent to the ones in (8)-(26).

For simplicity, we can express the optimization problem in (7)-(43) with the compact matrix formulation in (46)-(51).

$$\max_{\substack{\mathbf{h}', \boldsymbol{\delta}' \\ \mathbf{y}^{g \leftarrow p} \in \{0,1\}^{N_C} \\ \mathbf{y}^{p \leftarrow g} \in \{0,1\}^{N_C}}} \quad \min_{\mathbf{u} \in \{0,1\}^{L_{PN}}} \quad \max_{\mathbf{h}, \boldsymbol{\delta}} \mathbf{b}^T \mathbf{h} + \mathbf{c}^T \mathbf{y} \tag{46}$$

subject to:

$$\mathbf{Py} \leq \mathbf{g} \tag{47}$$

$$\mathbf{b}^T\mathbf{h}' \geq 1 \tag{48}$$

$$\mathbf{R}'\mathbf{h}' \leq \mathbf{q}' - \mathbf{H}'\mathbf{y} - \mathbf{W}'\boldsymbol{\delta}' - \mathbf{y}^T\mathbf{D}'\boldsymbol{\delta}' \tag{49}$$

$$\mathbf{Ku} \leq \mathbf{a} \tag{50}$$

$$\mathbf{Rh} \leq \mathbf{q} - \mathbf{Tu} - \mathbf{Hy} - \mathbf{W}\boldsymbol{\delta} - \mathbf{y}^T\mathbf{D}\boldsymbol{\delta}. \tag{51}$$

The vectors $\mathbf{h}'$ and $\mathbf{h}$ contain the continuous variables of the first and third level, respectively. The other variable vectors, $\mathbf{y}$, $\boldsymbol{\delta}'$, and $\boldsymbol{\delta}$, contain binary variables (vector $\mathbf{y}$ contains vectors $\mathbf{y}^{g\leftarrow p}$ and $\mathbf{y}^{p\leftarrow g}$). The vectors $\mathbf{b}$ and $\mathbf{c}$ contains the objective function coefficients, while the vectors $\mathbf{g}$, $\mathbf{a}$ and $\mathbf{q}$ contain constraint parameters. The matrices $\mathbf{P}$, $\mathbf{R}'$, $\mathbf{H}'$, $\mathbf{W}'$, $\mathbf{D}'$, $\mathbf{K}$, $\mathbf{R}$, $\mathbf{H}$, $\mathbf{W}$, and $\mathbf{D}$ contain constraint coefficients with suitable dimensions.

Equation (46) corresponds to Equation (7); Equation (47) corresponds to Equations (8)-(10); Equation (48) corresponds to Equation (11); Equation (49) corresponds to Equations (12)-(26); Equation (50) corresponds to Equations (27)-(28); Equation (51) corresponds to Equations (29)-(43).

The optimal objective value of the trilevel optimization is the maximized combined performance of the IPGNs in the worst scenario within the set of feasible failure scenarios. An important output of the optimization problem is the optimal coupling interface design $\widehat{\mathbf{y}}$.

3. Solution strategy

3.1. Linearization

As the problem in (46)-(51) (or equivalently (7)-(43)) contains several nonlinear terms, the first step of our solution strategy involves a reformulation into an equivalent linear form. In particular, the nonlinear terms arise from the multiplications of binary variables $\mathbf{y}'$ and $\boldsymbol{\delta}'$ in (49), and $\mathbf{y}$ and $\boldsymbol{\delta}$ in (51). Products of binary variables can be easily linearized by introducing new binary variables and additional constraints. Generally, the product of two binary variables a and b is also a binary variable, here called c, subject to Constraints (52)-(54):

$$c \leq a \tag{52}$$

$$c \leq b \tag{53}$$

$$c \geq a + b - 1. \tag{54}$$

The multiplications of $\mathbf{y}'$ and $\boldsymbol{\delta}'$ in (49), and $\mathbf{y}$ and $\boldsymbol{\delta}$ in (51), can then be linearized by introducing binary variables $\mathbf{z}$ and $\mathbf{r}$ and additional constraints of the type in (52)-(54). The variables $\mathbf{z}$ are introduced to linearize the multiplication between two binary variables, while the variables $\mathbf{r}$ are introduced to linearize the multiplication between three binary variables. Constraint (49) can then be replaced by Constraints (55) and (56), while Constraint (51) can then be replaced by Constraints (57) and (58):

$$\mathbf{R}'\mathbf{h}' \leq \mathbf{q}' - \mathbf{H}'\mathbf{y} - \mathbf{W}'\boldsymbol{\delta}' - \mathbf{S}'\mathbf{z}' - \mathbf{V}'\mathbf{r}' \tag{55}$$

$$\mathbf{Q}'\mathbf{z}' + \mathbf{F}'\mathbf{r}' \leq \mathbf{t}' - \mathbf{L}'\mathbf{y} - \mathbf{J}'\boldsymbol{\delta}' \tag{56}$$

$$\mathbf{R}\mathbf{h} \leq \mathbf{q} - \mathbf{T}\mathbf{u} - \mathbf{H}\mathbf{y} - \mathbf{W}\boldsymbol{\delta} - \mathbf{S}\mathbf{z} - \mathbf{V}\mathbf{r} \tag{57}$$

$$\mathbf{Q}\mathbf{z} + \mathbf{F}\mathbf{r} \leq \mathbf{t} - \mathbf{L}\mathbf{y} - \mathbf{J}\boldsymbol{\delta} \tag{58}$$

where Equations (56) and (58) corresponds to the additional constraints of the type in (52)-(54).

The linear compact matrix formulation corresponds to Equation (46) subject to (47)-(48), (50), and (55)-(58).

3.2. Nested Column&Constraint Generation algorithm

The presence of the binary variables $\boldsymbol{\delta}$ in the third stage makes it impossible to merge the second and third stage into a single minimization problem relying on the dual formulation. Therefore, we adopt a cutting plane strategy, called Nested Column&Constraint Generation (NC&CG) algorithm. It represents an exact method, with proven convergence to the global optimum, for solving multi-level mixed-integer linear programming with recourse problems [51], [52].

Figure 3 details the flowchart with the main steps of the NC&CG algorithm. In order to adopt this strategy, the original trilevel max-min-max problem is transformed into a max-min-max-max problem, by separating binary and continuous variables in the original third stage [14]. The new fourth stage contains only continuous variables, and it is then a pure LP problem. The formulation is then transformed into a max-min-max-min through a dual reformulation of the last stage. In this form, the problem can be solved using a NC&CG algorithm, identifying an outer and inner layer which exchange primal variables in form of parameters until the convergence to the global optimum is reached.

For a more detailed explanation of the C&CG algorithm, the reader is referred to [51], [52] for a theoretical framework and [14], [27], [50] for applications.

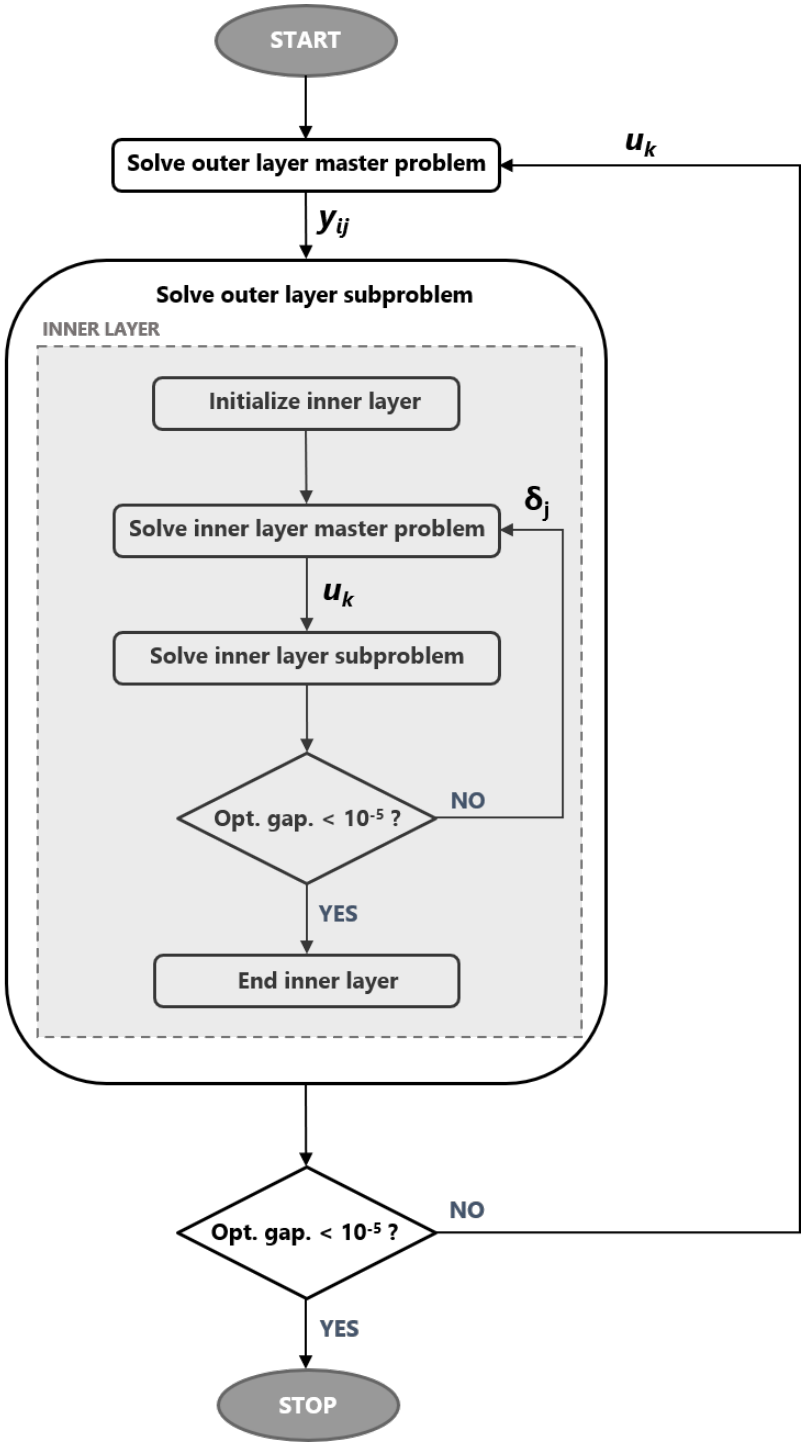

Figure 3: Flowchart of the NC&CG algorithm [14], [27].

3.3. Inner layer

The inner layer consists in solving the second and third level (min-max) in (46) with a fixed coupling interface $\mathbf{y}^*$. The output of the model is the worst-case realization of the combined performance and the associated optimal attack plan $\widehat{\mathbf{u}}$. With fixed binary variables (coupling interface $\mathbf{y}^*$, interdependency variables $\boldsymbol{\delta}^*$ and attack plan $\mathbf{u}^*$), the inner-most maximization in (46) and the relative constraints take the form in (59)-(60):

$$\max_{\mathbf{h}} \mathbf{b}^T\mathbf{h} + \mathbf{c}^T\mathbf{y}^* \tag{59}$$

subject to :

$$\mathbf{Rh} \leq \mathbf{q} - \mathbf{Tu}^* - \mathbf{Hy}^* - \mathbf{W}\boldsymbol{\delta}^* - \mathbf{y}^{*T}\mathbf{D}\boldsymbol{\delta}^* \tag{60}$$

The problem in (59)-(60) is a pure LP, and thus the introduction of variables $\mathbf{z}$ and $\mathbf{r}$ is not necessary. Thanks to its linear nature, strong duality holds and it can be transformed into its dual form in (61)-(62):

$$\min_{\boldsymbol{\lambda}} (\mathbf{q} - \mathbf{Tu}^* - \mathbf{Hy}^* - \mathbf{W}\boldsymbol{\delta}^* - \mathbf{y}^{*T}\mathbf{D}\boldsymbol{\delta}^*)^T\boldsymbol{\lambda} \tag{61}$$

subject to:

$$\mathbf{R}^T\boldsymbol{\lambda} = \mathbf{b} \tag{62}$$

As the variables $\boldsymbol{\delta}$ are binary, the number of possible combinations that they can take is equal to 2^{N_δ}, where N_δ=$N_{PN}+N_{GN}$ is the number of binary variables $\boldsymbol{\delta}$. We denote as $\mathcal{D}$ the set containing all the possible combinations of binary variables $\boldsymbol{\delta}$. The C&CG approach exploits the observation that only a partial subset $\mathcal{D}_{part} \subseteq \mathcal{D}$ is essential to compute the optimal solution. The bilevel min-max formulation can be solved by iteratively reconstructing the partial set $\mathcal{D}_{part}$ by following these steps:

1. Set $j = 0$, lower bound $LB_{in} = 0$, upper bound $UB_{in} = \infty$, and $\mathcal{D}_{part} = \emptyset$
2. Solve the inner master problem in Equations (63)-(66). Obtain an optimal solution $\widehat{\rho}^{(j)}$ and optimal attack plan $\widehat{\mathbf{u}}^{(j)}$. Update LB_{in}=$\widehat{\rho}^{(j)} + \mathbf{c}^T\mathbf{y}^*$.

$$\min_{\rho,\mathbf{u},\boldsymbol{\lambda}} \rho \tag{63}$$

subject to:

$$\rho \geq (\mathbf{q} - \mathbf{Tu} - \mathbf{Hy}^* - \mathbf{W}\boldsymbol{\delta}^{*(j)} - \mathbf{y}^{*T}\mathbf{D}\boldsymbol{\delta}^{*(j)})^T\boldsymbol{\lambda}^{(j)}, \quad \forall \boldsymbol{\delta}^{*(j)} \in \mathcal{D}_{part} \tag{64}$$

$$\mathbf{R}^T\boldsymbol{\lambda}^{(j)} = \mathbf{b}, \quad \forall \boldsymbol{\delta}^{*(j)} \in \mathcal{D}_{part} \tag{65}$$

$$\sum_{k \in E_{PN}} (1 - u_k) \leq K_{att} \tag{66}$$

3. Solve the inner subproblem in Equations (67)-(68) with $\widehat{\mathbf{u}}^{(j)}$=$\mathbf{u}^*$. Obtain an optimal solution $\mathbf{b}^T\widehat{\mathbf{h}}^{(j)}$ and $\widehat{\boldsymbol{\delta}}^{(j)}$. Set $UB_{in} = \min(UB_{in}, \mathbf{b}^T\widehat{\mathbf{h}}^{(j)} + \mathbf{c}^T\mathbf{y}^*)$.

$$\max_{\mathbf{h},\boldsymbol{\delta}} \mathbf{b}^T\mathbf{h} \tag{67}$$

subject to :

$$\mathbf{Rh} \leq \mathbf{q} - \mathbf{Tu}^* - \mathbf{Hy}^* - \mathbf{W}\boldsymbol{\delta} - \mathbf{y}^{*T}\mathbf{D}\boldsymbol{\delta} \tag{68}$$

4. If $(UB_{in} - LB_{in})/UB_{in} < 10^{-5}$, $\widehat{\mathbf{u}}^{(j)}$ represents the optimal attack and the algorithm can be terminated. Otherwise, $\mathcal{D}_{part} = \mathcal{D}_{part} \cup \widehat{\boldsymbol{\delta}}^{(j)}$. Set $j \leftarrow j + 1$ and return to step 2.

The optimal attack plan, or, in other words, the feasible combination of variables $\mathbf{u}$ which minimizes the combined performance for a fixed coupling interface $\mathbf{y}^*$, and the optimal value of the objective function represent the main outputs of the algorithm.

3.4. Outer layer

Similarly, the outer layer is solved by employing a partial set of attack scenarios $\mathcal{A}_{part} \subseteq \mathcal{A}$. The outer layer solves a bilevel max-min problem, and the minimization is solved by the inner layer algorithm.

The outer layer is solved by employing the following steps:

1. Set $j = 0$, lower bound $LB_{out} = 0$, upper bound $UB_{out} = \infty$, and $\mathcal{A}_{part} = \emptyset$
2. Solve the outer master problem in Equations (69)-(76). Obtain an optimal solution $\widehat{\eta}^{(j)} + \mathbf{c}^T\widehat{\mathbf{y}}^{(j)}$ and optimal coupling interface $\widehat{\mathbf{y}}^{(j)}$. Update UB_{out}=$\min(UB_{out}, \widehat{\eta}^{(j)} + \mathbf{c}^T\widehat{\mathbf{y}}^{(j)})$

$$\max_{\substack{\eta,\mathbf{h}^{(j)} \\ \mathbf{h}',\boldsymbol{\delta}' \\ \mathbf{y}\in\{0,1\}}} \eta + \mathbf{c}^T\mathbf{y} \tag{69}$$

$$\eta \leq \mathbf{b}^T\mathbf{h}^{(j)}, \quad \forall\, \mathbf{u}^{*(j)} \in \mathcal{A}_{part} \tag{70}$$

$$\mathbf{Py} \leq \mathbf{g} \tag{71}$$

$$\mathbf{b}^T\mathbf{h}' \geq 1 \tag{72}$$

$$\mathbf{R}'\mathbf{h}' \leq \mathbf{q}' - \mathbf{H}'\mathbf{y} - \mathbf{W}'\boldsymbol{\delta}' - \mathbf{S}'\mathbf{z}' - \mathbf{V}'\mathbf{r}' \tag{73}$$

$$\mathbf{Q}'\mathbf{z}' + \mathbf{F}'\mathbf{r}' \leq \mathbf{t}' - \mathbf{L}'\mathbf{y} - \mathbf{J}'\boldsymbol{\delta}' \tag{74}$$

$$\mathbf{Rh}^{(j)} \leq \mathbf{q} - \mathbf{Tu}^{*(j)} - \mathbf{Hy} - \mathbf{W}\boldsymbol{\delta}^{(j)} - \mathbf{Sz}^{(j)} - \mathbf{Vr}^{(j)}, \ \forall\, \mathbf{u}^{*(j)} \in \mathcal{A}_{part} \tag{75}$$

$$\mathbf{Qz}^{(j)} + \mathbf{Fr}^{(j)} \leq \mathbf{t} - \mathbf{Ly} - \mathbf{J}\boldsymbol{\delta}^{(j)}. \tag{76}$$

3. Solve the outer subproblem using the inner layer in the previous subsection with $\widehat{\mathbf{y}}^{(j)}$=$\mathbf{y}^*$. Obtain an optimal solution $\mathbf{b}^T\widehat{\mathbf{h}}^{(j)} + \mathbf{c}^T\mathbf{y}^*$ and an optimal attack plan $\widehat{\mathbf{u}}^{(j)}$. Set $LB_{out} = \mathbf{b}^T\widehat{\mathbf{h}}^{(j)} + \mathbf{c}^T\mathbf{y}^*$.
4. If $(UB_{out} - LB_{out})/UB_{out} < 10^{-5}$, $\widehat{\mathbf{y}}^{(j)}$ is the optimal coupling interface and the algorithm is terminated. Otherwise, $\mathcal{A}_{part} = \mathcal{A}_{part} \cup \widehat{\mathbf{u}}^{(j)}$, set $j \leftarrow j + 1$ and return to step 2.

The outputsof the algorithm are the optimal combined performance in the worst-case failure scenario and the related optimal coupling interface $\widehat{\mathbf{y}}$s.

4. Illustrative case-study

As illustrative case-study, a power network based on the IEEE 14-bus system [53] and a gas network based on the IEEE 9-bus system [54] are considered. As shown in Figure 4, the IPGNs are allocated within a 300×300 km area. The importance of each infrastructure is given by their weights, w_{PN} and w_{GN}, both equal to 0.5. Node 1 in the power network is chosen as the reference bus. Other parameter values are summarized in Appendix B. We test our model for values of K_{att} ranging from 1 to 5. We choose a representative interdependency cost-per-kilometer of 1 \$/km, for both $c_{km}^{g\leftarrow p}$ and $c_{km}^{p\leftarrow g}$. We assume budget values B_{ci} ranging from \$900 to \$1500 for the installment of coupling interfaces. We also consider a budget of \$823, which corresponds to the cost of the minimum-distance coupling interface, where each node in one infrastructure is dependent, if necessary, on the geographically-closest node of the other infrastructure[4]. We compare the results obtained by our model with the results obtained with network metrics-based coupling interfaces, which are identified based on different combinations of node degree (D) and betweenness (B). We distinguish four coupling interfaces using the different network metrics and the terms *assortative* (subscript *ast*) and *disassortative* (subscript *dst*). In network science, the assortativity (disassortativity) is a property that describes the tendency of the nodes of a network to be connected to nodes which are similar (different) regarding some specific properties [55]. For example, it can refer to the tendency of high degree nodes to be attached to other high degree nodes. Additionally, we identify a geographical location-based coupling interface, referred to as *Euclidean*. The five different network metrics-based interfaces used in this work are characterized by the following features:

- Euclidean: each node in the power network (or gas network) is dependent on the geographically closest node in the gas network (or power network).

[4]The cost of this coupling interface, referred to as *Euclidean* coupling interface, is, precisely, \$822.763752. For the sake of simplicity, in this work, it is approximated to \$823.

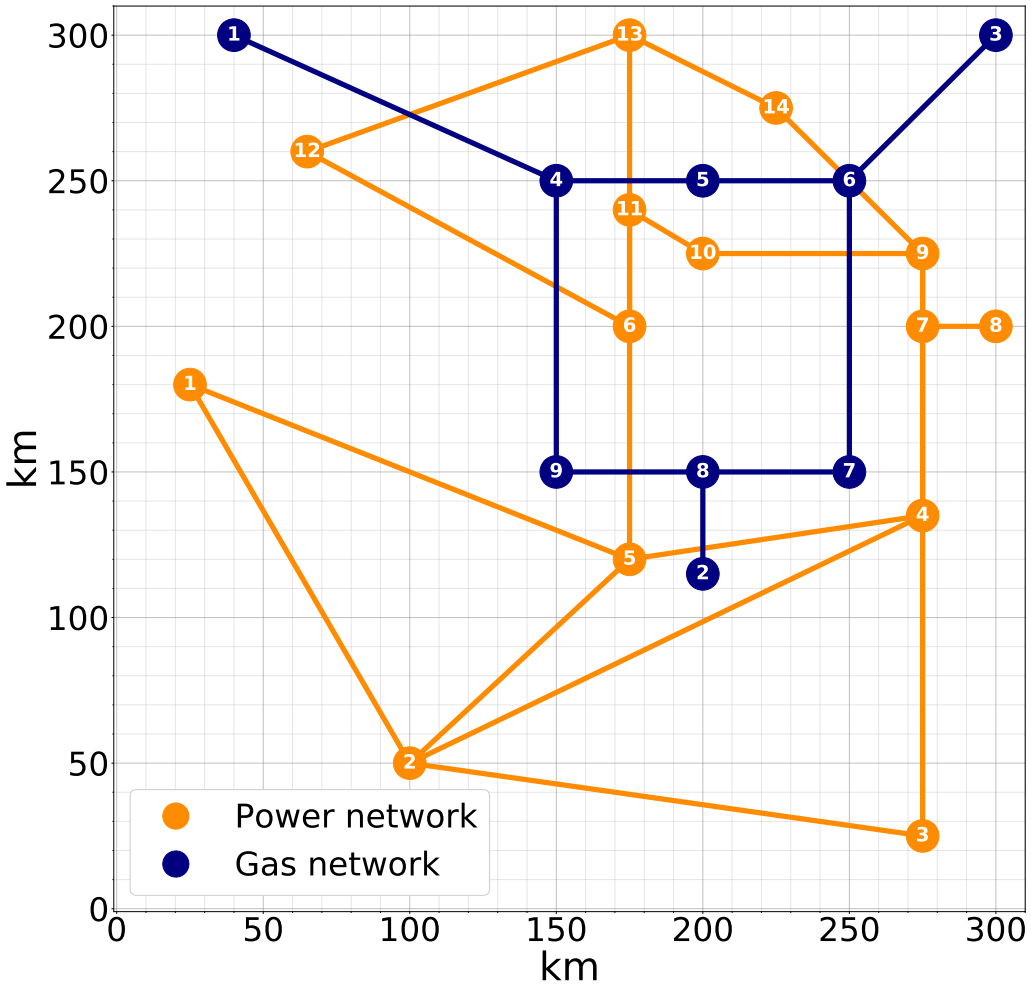

Figure 4: Interdependent power and gas networks.

- DD_{ast}: the node with the k^{th} highest degree in the power network (or gas network) is dependent on the node with the k^{th} highest degree in the gas network (or power network).
- DD_{dst}: the node with the k^{th} highest degree in the power network (or gas network) is dependent on the node with the k^{th} lowest degree in the gas network (or power network).
- BB_{ast}:the node with the k^{th} highest betweenness in the power network (or gas network) is dependent on the node with the k^{th} highest betweenness in the gas network (or power network).
- BB_{dst}: the node with the k^{th} highest betweenness in the power network (or gas network) is dependent on the node with the k^{th} lowest betweenness in the gas network (or power network).

The cost associated with each network metrics-based coupling interface is reported in Table 1.

Table 1: Cost of network metrics-based coupling interfaces. For simplicity, the costs are rounded by excess.

Interface	**Cost**
Euclidean	\$823
DD_{ast}	\$1518
DD_{dst}	\$2098
BB_{ast}	\$1943
BB_{dst}	\$2126

The optimization problem is implemented with Gurobi 9.1 [56] on a desktop PC with a 3.20 GHz CPU and 32 GB RAM.

5. Results and discussion

5.1. Combined performance

The results for the network metrics-based coupling interfaces are shown in Figure 5, while the results for the optimal coupling interfaces obtained by our approach with different budget B_{ci} are shown in Figure 6. The x-axis shows the maximum number of lines in the power network which can be attacked and failed; the y-axis shows the correspondent worst-case realization of the combined performance.

As it can be clearly seen in Figure 5, the DD_{ast} coupling interface performs quite poorly, reaching a worst-case combined performance value of 0 for K_{att}=4. The BB_{ast} coupling interface performs well for values K_{att} $\leq$4. The DD_{dst} and BB_{dst} coupling interfaces perform similarly for values K_{att} $\leq$3. For K_{att}=4, the DD_{dst} interface performs better, while for K_{att}=5, the BB_{dst} interface performs better.

The Euclidean coupling interface leads to the better performance overall: for K_{att}=3, K_{att}=4 and K_{att}=5, the Euclidean coupling interface leads worst-case combined performance of 0.703, 0.523 and 0.307. It is outperformed only by the BB_{ast} coupling interface for K_{att}=1.

These results clearly show how different coupling interfaces lead to different worst-case combined performance. In this case, the Euclidean coupling interface performs better than the other network metrics-based coupling interfaces. However, these results should not be generalized, as the performance of each network metrics-based coupling interface is strongly case-dependent. For example, if we change the geographical disposition of the nodes of the IPGNs, the Euclidean coupling interface would be different and, thus, the results would differ. Similar considerations are valid for the other network metrics-based coupling interfaces.

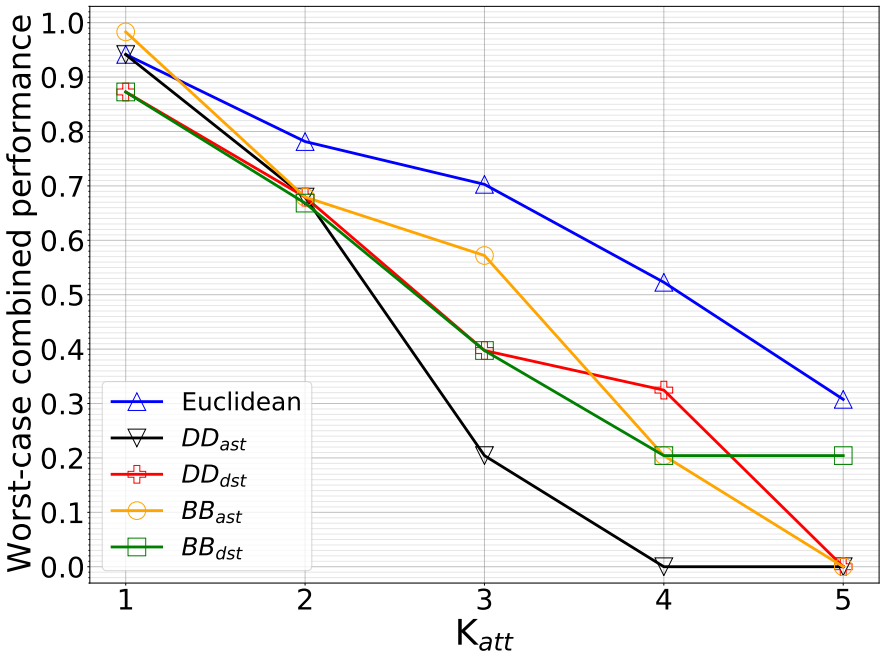

Figure 5: Worst-case combined performance for different network metrics-based coupling interface and values K_{att}.

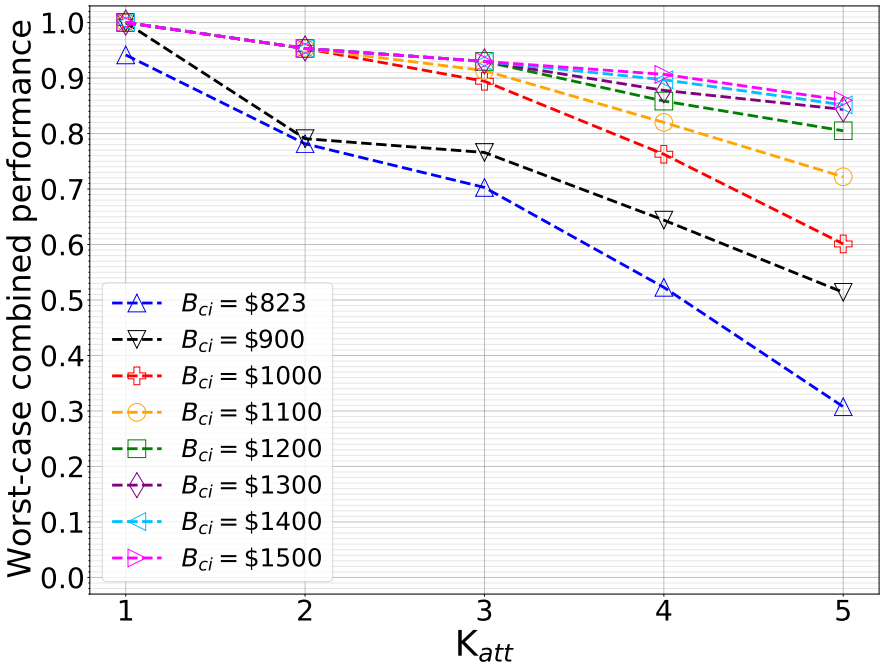

Figure 6: Worst-case combined performance for optimized coupling interface with different budgets B_{ci} and values K_{att}.

The optimal coupling interfaces, identified with the proposed optimization model, outperform the network metrics-based coupling patterns in terms of worst-case combined performance, as it can be clearly seen in Figure 6. The minimum budget which ensures the feasibility of the model is \$823, which corresponds to the cost of the Euclidean coupling interface (see Table 1). For a budget lower than \$823 it is not possible to allocate all the necessary interdependencies and to ensure satisfactory performance in normal conditions, and the optimization problem is, thus, unfeasible. The results for B_{ci}=\$823 (blue triangles in Figure 6) are equivalent to the results of the Euclidean coupling interface (blue triangles in Figure 5).

As it can be clearly seen, for values of B_{ci} greater than \$823, the traditional interfaces are outperformed by the optimal coupling interfaces identified by the proposed approach. For example, with B_{ci}=\$900 and K_{att}=3, K_{att}=4 and K_{att}=5, the worst-case combined performance are, respectively, 0.766, 0.644 and 0.514, while with B_{ci}=\$1000 and K_{att}=3, K_{att}=4 and K_{att}=5, the worst-case combined performance are 0.894, 0.763 and 0.601, respectively. These results are considerably higher than the previously explained Euclidean interface (0.703, 0.523 and 0.307, respectively).

The worst-case combined performance improves with the increasing of the budget B_{ci}. For example, with B_{ci}=\$1500 and K_{att}=3, K_{att}=4 and K_{att}=5, the worst-case combined performance are 0.930, 0.906 and 0.860, respectively. For values of B_{ci} greater than \$1500, the results do not improve. The case B_{ci}=\$1500 (pink triangles in Figure 6) leads to the best possible results for this case-study.

It is also of interest to compare optimal coupling interface designs for different B_{ci} and K_{att}. In Figure 7, the optimal coupling interfaces for B_{ci}=\$900 and B_{ci}=\$1000 with K_{att}=2 are shown. With B_{ci}=\$900 and K_{att}=2, the optimal value of the combined performance is 0.791, while with B_{ci}=\$1000 and K_{att}=2, the optimal value of the combined performance is 0.953. These values corresponds to an increase of combined performance of 20.5% for an increase of budget of 11.1%. As we can notice in Figure 7, two interdependency links from the gas network to the power network (red squares) change when passing from B_{ci}=\$900 to B_{ci}=\$1000, as it is also highlighted in Table 2. Moreover, three interdependency links from the power network to the gas network (blue squares) change when passing from B_{ci}=\$900 to B_{ci}=\$1000, as it is also highlighted in Table 3.

Table 2: Reallocation of interdependency links from the gas network to the power network (gas supply) when passing from B_{ci}=\$900 to B_{ci}=\$1000, with K_{att}=2.

Budget	**Node 6** $\in \mathbf{V_{PN}}$	**Node 8** $\in \mathbf{V_{PN}}$
\$900	Node 8 $\in V_{GN}$	Node 6 $\in V_{GN}$
\$1000	Node 9 $\in V_{GN}$	Node 7 $\in V_{GN}$

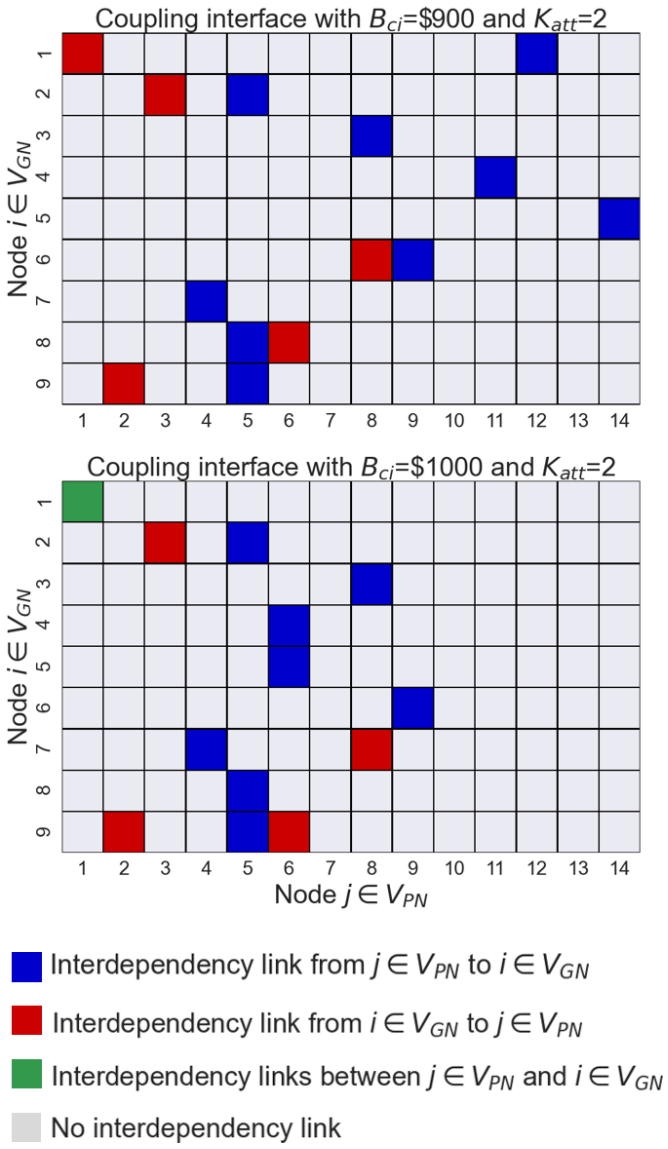

Figure 7: Example of two optimal solutions for B_{ci}=\$900 and B_{ci}=\$1000 with K_{att}=2. Blue squares represent links from the power network to the gas network (electricity supply); red squares represent links from the gas network to the power network (gas supply); green squares represent links in both the directions; grey squares represent the absence of links.

As it can be clearly seen, the reallocation of some of the interdependency links leads a considerable increase of worst-case combined performance. Moreover, it is interesting to notice that, with B_{ci}=\$1000, nodes 1, 4, and 5 of the gas network are dependent on nodes 1, 6, and 6 of the power network, respectively, and both these nodes of the power network contain a gas-fired power plant, i.e. they have some power production capacity (see Table B.5 in Appendix B). Intuitively, as in this work only failures of lines are considered, it is more convenient for nodes of the gas network to

Table 3: Reallocation of interdependency links from the power network to the gas network (electricity supply) when passing from B_{ci}=\$900 to B_{ci}=\$1000, with K_{att}=2.

Budget	**Node 1** $\in \mathbf{V_{GN}}$	**Node 4** $\in \mathbf{V_{GN}}$	**Node 5** $\in \mathbf{V_{GN}}$
\$900	Node 12 $\in V_{PN}$	Node 11 $\in V_{PN}$	Node 14 $\in V_{PN}$
\$1000	Node 1 $\in V_{PN}$	Node 6 $\in V_{PN}$	Node 6 $\in V_{PN}$

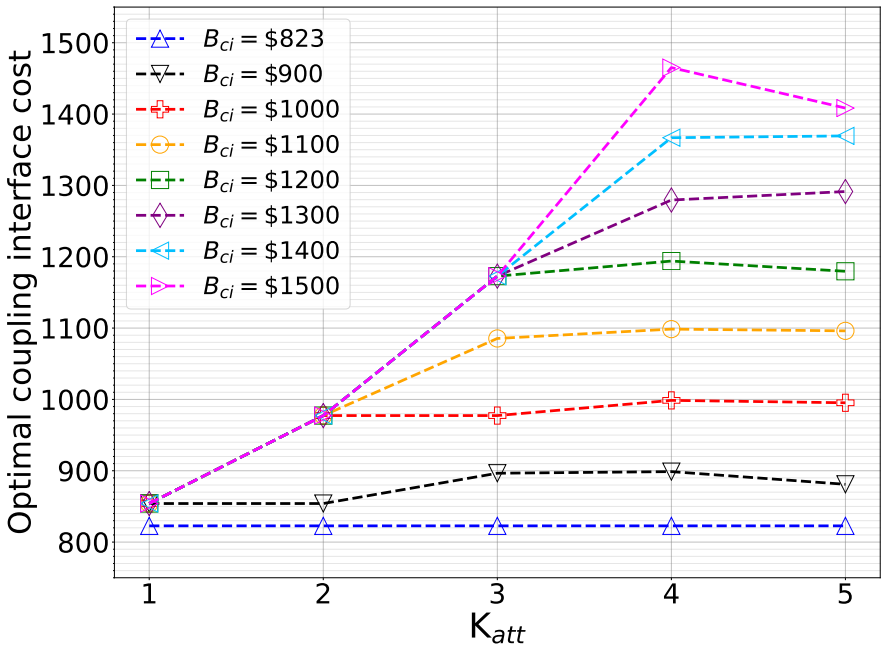

Figure 8: Cost of optimal coupling interfaces for different budgets B_{ci} and values K_{att}.

be dependent on nodes in the power network with some production capacity, and vice versa.

5.2. Coupling interface cost

In Figure 8, the results in terms of allocation cost of optimal coupling interfaces for different monetary budgets and maximum failed lines are shown. As it can be clearly seen, the network metrics-based coupling interfaces are outperformed also in terms of cost (with the exception of the Euclidean coupling interface). The cost of the network metrics-based coupling interfaces are shown in Table 1.

It is also useful to compare the increase in combined performance with the increase of cost. For example, for the case K_{att}=3, passing from B_{ci}=\$823 to B_{ci}=\$1000 (21.5% of budget increase) leads to an increase of 27.2% in worst-case combined performance (from 0.703 to 0.894). The cost of the optimal coupling interface with a budget B_{ci}=\$1000 is 977.3\$, corresponding to an increase of

cost of 18.8% from the Euclidean coupling interface. In this case, the relative increase of combined performance is greater than the relative increase of cost. However, as it can be graphically seen in Figures 6 and 8, for higher values of K_{att} and B_{ci}, the relative increase of cost is higher than the increase in performance. For example, for the case K_{att}=5, when passing from B_{ci}=\$1400 to B_{ci}=\$1500, the increase of budget is 7.1%, and the increase of actual cost is 2.8% (from \$1369.40 to \$1408.40); however, the increase of combined performance is only 0.93% (from 0.852 to 0.860).

For the case previously analyzed in Figure 7 and Table 3, an increase of 11.1% in the budget (from \$900 to \$1000) leads to an increase of 20.5% in combined performance (from 0.791 to 0.953). The actual costs of the two optimal solutions, for B_{ci}=\$900 and B_{ci}=\$1000 with K_{att}=2, are \$854.1 and \$977.3, respectively, corresponding to an increase of 14.4% in cost when passing from B_{ci}=\$900 to B_{ci}=\$1000 with K_{att}=2.

5.3. Validation

The last term in the objective function in (7) numerically pushes the optimization problem to identify the cheapest solution among the coupling interfaces that maximizes the combined performance of the IPGNs. In order to identify correctly this solution, the order of magnitude of the factor γ should be set properly, accounting for the order of magnitude of the combined performance, the monetary budget, and the optimality gap within the NC&CG algorithm. Within this paper, a value of γ=10^{-5} is used. The results are then validated by solving the optimization problem only accounting for the combined performance (γ=0), and by setting the monetary budget B_{ci} slightly below the actual cost of the optimal coupling interface, and verify that the optimal combined performance are lower.

For example, for the case B_{ci}=\$1000 and K_{att}=2, the optimal coupling interface has a cost of \$977.3 and leads to combined performance of 0.953. We can verify that the cheapest optimal solution is identified correctly by setting γ=0 and solving for a budget B_{ci}=\$977. Solving the problem with a budget B_{ci}=\$977 leads to combined performance of 0.950, lower than the optimal combined performance of 0.953. This is an indication the correct cheapest optimal solution is identified correctly.

5.4. Computational performance

The computational time in seconds of the NC&CG algorithm is shown in Table 4. In this study, the computational cost is acceptable, as the longest instance of the NC&CG algorithm occurs for B_{ci}=\$1000 and K_{att}=5, and it takes 176.26 seconds.

The illustrative case-study in this work presents a small-medium size, and the computational cost might increase considerably if larger networks are considered. However, this do not represent an issue:

Table 4: Computational time in seconds of the NC&CG algorithm.

B_{ci}	K_{att}=1	K_{att}=2	K_{att}=3	K_{att}=4	K_{att}=5
\$823	1.13	4.12	30.17	50.37	140.30
\$900	2.91	6.26	35.58	156.44	93.36
\$1000	1.73	7.80	21.64	52.67	176.26
\$1100	2.89	11.25	22.40	21.29	47.63
\$1200	2.75	9.45	53.61	97.97	53.41
\$1300	2.87	6.31	27.94	73.95	49.66
\$1400	2.89	8.03	16.48	22.89	40.37
\$1500	2.90	5.86	12.48	39.94	42.58

- the proposed model should be used for designing or retrofitting coupling interfaces, and in preliminary design phases, the computational time do not represent a critical factor;
- the computational complexity of the optimization problem can be reduced by limiting the number of binary variables of the problem. For example, the feasible allocation of interdependency links can be limited to nodes which are geographically close to each other.

6. Conclusion

CIs are essential for any advanced society, and ensuring their resilience against failures and disruption is of the utmost importance. As coupling interfaces between interdependent CIs are a key factor for maintaining high levels of resilience, optimizing their design is an important issue.

In this work, we proposed a mathematical programming approach for the resilience-based optimization of coupling interfaces between interdependent CIs that, compared to traditional network metrics-based solutions, is more generalizable and leads to better performances.

In fact, using interdependent power and gas networks as case-study, we showed how optimal coupling strategies, obtained by the proposed approach, clearly outperform traditional coupling strategies based on network metrics. In addition, the proposed approach can be easily adapted to other combinations of interdependent CIs by updating the operational model used in the optimization procedure.

In the proposed case-study, only failures of power lines are considered. However, alternative disruption scenarios, such as failure of nodes or gas pipelines, can be easily included with a similar

approach using additional binary variables.

The computational cost is affordable in this work. In general, in this kind of optimization problems, aimed at being used during design phases, the computational time does not represent a key factor.

Further improvements of this work includes the possibility of allocating redundant interdependency links within the coupling interface and the evaluation of occurrence probability of each failure scenario.

Appendix A. Linearization of DC power flow constraint

Constraint (34) can be linearized by replacing it with the equivalent Constraints (A.1) and (A.2):

$$x_k f_k - \left(\theta_{O(k)} - \theta_{D(k)}\right) \geq -M_k(1-u_k), \forall k \in E_{PN} \tag{A.1}$$

$$x_k f_k - \left(\theta_{O(k)} - \theta_{D(k)}\right) \leq M_k(1-u_k), \quad \forall k \in E_{PN} \tag{A.2}$$

where M_k is the "Big-M" constants, computed as in (A.3) as suggested in [50]:

$$M_k \geq \overline{\theta} + x_k \overline{f}_k \tag{A.3}$$

where $\overline{\theta}$ is the maximum difference of two phase angles at two connected buses, here assumed $\pi/2$.

Appendix B. IPGNs parameters

Table B.5: Production capacity and base requested demand for each node in the power network.

Node index	$\overline{\mathbf{p}}_\mathbf{i}$ [MW]	$\overline{\mathbf{d}}_\mathbf{i}^\mathbf{b}$ [MW]	$\overline{\mathbf{d}}_\mathbf{i}^{\mathbf{m^3}}$ [m^3]
1	42	8.5	3
2	42	8.5	3
3	42	8.5	3
4	0	8.5	0
5	0	8.5	0
6	42	8.5	3
7	0	8.5	0
8	42	8.5	3
9	0	8.5	0
10	0	8.5	0
11	0	8.5	0
12	0	8.5	0
13	0	8.5	0
14	0	8.5	0

Table B.6: Boundaries, maximum flow capacity and reactance for each line in the power network.

Line index	**Boundaries** $(\mathbf{i}, \mathbf{j})$	$\bar{\mathbf{f}}_{\mathbf{k}}$ [MW]	$\mathbf{x_k}$ [pu]
1	(1, 2)	30	0.05917
2	(1, 5)	30	0.22304
3	(2, 3)	30	0.19797
4	(2, 4)	30	0.17632
5	(2, 5)	30	0.17388
6	(3, 4)	30	0.17103
7	(4, 5)	30	0.04211
8	(4, 7)	30	0.20912
9	(4, 9)	30	0.55618
10	(5, 6)	30	0.24202
11	(6, 11)	30	0.1989
12	(6, 12)	30	0.25581
13	(6, 13)	30	0.13027
14	(7, 8)	30	0.17615
15	(7, 9)	30	0.11001
16	(9, 10)	30	0.0845
17	(9, 14)	30	0.27038
18	(10, 11)	30	0.19207
19	(12, 13)	30	0.19988
20	(13, 14)	30	0.34802

Table B.7: Production capacity and base requested demand for each node in the gas network.

Node index	$\overline{\mathbf{p}}_\mathbf{i}$ [m^3]	$\overline{\mathbf{d}}_\mathbf{i}$ [m^3]	$\overline{\mathbf{d}}_\mathbf{i}^{\mathbf{MW}}$ [MW]
1	15	0	7
2	15	0	7
3	15	0	7
4	0	5	7
5	0	5	7
6	0	5	7
7	0	5	7
8	0	5	7
9	0	5	7

Table B.8: Boundaries and maximum flow capacity for each line in the gas network.

Line index	**Boundaries** $(\mathbf{i}, \mathbf{j})$	$\overline{\mathbf{f}}_\mathbf{k}$ [m^3]
1	(1, 2)	15
2	(1, 5)	10
3	(2, 3)	10
4	(2, 4)	15
5	(2, 5)	10
6	(3, 4)	10
7	(4, 5)	15
8	(4, 7)	10
9	(4, 9)	10

References

[1] F. Nocera and P. Gardoni, "Selection of the modeling resolution of infrastructure," *Computer-Aided Civil and Infrastructure Engineering*, 2022.

[2] A. Bellè, Z. Zeng, C. Duval, M. Sango, and A. Barros, "Modeling and vulnerability analysis of interdependent railway and power networks: Application to british test systems," *Reliability Engineering System Safety*, vol. 217, p. 108 091, 2022, ISSN: 0951-8320.

[3] Y. Li, C. Zhang, C. Jia, X. Li, and Y. Zhu, "Joint optimization of workforce scheduling and routing for restoring a disrupted critical infrastructure," *Reliability Engineering & System Safety*, vol. 191, p. 106 551, 2019.

[4] A. Mottahedi, F. Sereshki, M. Ataei, A. N. Qarahasanlou, and A. Barabadi, "Resilience estimation of critical infrastructure systems: Application of expert judgment," *Reliability Engineering & System Safety*, vol. 215, p. 107 849, 2021.

[5] N. Sharma, A. Tabandeh, and P. Gardoni, "Regional resilience analysis: A multiscale approach to optimize the resilience of interdependent infrastructure," *Computer-Aided Civil and Infrastructure Engineering*, vol. 35, no. 12, pp. 1315–1330, 2020.

[6] S. V. Buldyrev, R. Parshani, G. Paul, H. E. Stanley, and S. Havlin, "Catastrophic cascade of failures in interdependent networks," *Nature*, vol. 464, no. 7291, pp. 1025–1028, 2010.

[7] J. Winkler, L. Dueñas-Osorio, R. Stein, and D. Subramanian, "Interface network models for complex urban infrastructure systems," *Journal of Infrastructure Systems*, vol. 17, no. 4, pp. 138–150, 2011.

[8] M. Ouyang and L. Dueñas-Osorio, "An approach to design interface topologies across interdependent urban infrastructure systems," *Reliability Engineering & System Safety*, vol. 96, no. 11, pp. 1462–1473, 2011.

[9] M. Ouyang, "Review on modeling and simulation of interdependent critical infrastructure systems," *Reliability engineering & System safety*, vol. 121, pp. 43–60, 2014.

[10] G. G. Brown, W. M. Carlyle, J. Salmerón, and K. Wood, "Analyzing the vulnerability of critical infrastructure to attack and planning defenses," in *Emerging Theory, Methods, and Applications*, Informs, 2005, pp. 102–123.

[11] T. Ding, L. Yao, and F. Li, "A multi-uncertainty-set based two-stage robust optimization to defender–attacker–defender model for power system protection," *Reliability Engineering & System Safety*, vol. 169, pp. 179–186, 2018.

[12] W. Yuan, L. Zhao, and B. Zeng, "Optimal power grid protection through a defender–attacker–defender model," *Reliability Engineering & System Safety*, vol. 121, pp. 83–89, 2014.

[13] N. Ghorbani-Renani, A. D. González, K. Barker, and N. Morshedlou, "Protection-interdiction-restoration: Tri-level optimization for enhancing interdependent network resilience," *Reliability Engineering & System Safety*, vol. 199, p. 106 907, 2020.

[14] Y.-P. Fang and E. Zio, "An adaptive robust framework for the optimization of the resilience of interdependent infrastructures under natural hazards," *European Journal of Operational Research*, vol. 276, no. 3, pp. 1119–1136, 2019.

[15] N. Sharma, A. Tabandeh, and P. Gardoni, "Resilience analysis: A mathematical formulation to model resilience of engineering systems," *Sustainable and Resilient Infrastructure*, vol. 3, no. 2, pp. 49–67, 2018.

[16] M. Panteli, P. Mancarella, D. N. Trakas, E. Kyriakides, and N. D. Hatziargyriou, "Metrics and quantification of operational and infrastructure resilience in power systems," *IEEE Transactions on Power Systems*, vol. 32, no. 6, pp. 4732–4742, 2017.

[17] K. S. Trivedi, V. Jindal, and S. Dharmaraja, "Stochastic modeling techniques for secure and survivable systems," *Information assurance: Dependability and security in networked systems. Morgan Kaufmann*, pp. 171–207, 2008.

[18] J. Johansson, H. Hassel, and A. Cedergren, "Vulnerability analysis of interdependent critical infrastructures: Case study of the swedish railway system," *International journal of critical infrastructures*, vol. 7, no. 4, pp. 289–316, 2011.

[19] D. Henry and J. E. Ramirez-Marquez, "Generic metrics and quantitative approaches for system resilience as a function of time," *Reliability Engineering & System Safety*, vol. 99, pp. 114–122, 2012.

[20] C. Poulin and M. B. Kane, "Infrastructure resilience curves: Performance measures and summary metrics," *Reliability Engineering & System Safety*, vol. 216, p. 107 926, 2021.

[21] Y.-P. Fang and G. Sansavini, "Optimum post-disruption restoration under uncertainty for enhancing critical infrastructure resilience," *Reliability Engineering & System Safety*, vol. 185, pp. 1–11, 2019.

[22] X. Liu, Y.-P. Fang, and E. Zio, "A hierarchical resilience enhancement framework for interdependent critical infrastructures," *Reliability Engineering & System Safety*, vol. 215, p. 107 868, 2021.

[23] M. Najarian and G. J. Lim, "Optimizing infrastructure resilience under budgetary constraint," *Reliability Engineering & System Safety*, vol. 198, p. 106 801, 2020.

[24] Y. Wu, Z. Chen, H. Gong, Q. Feng, Y. Chen, and H. Tang, "Defender–attacker–operator: Tri-level game-theoretic interdiction analysis of urban water distribution networks," *Reliability Engineering & System Safety*, vol. 214, p. 107 703, 2021.

[25] J. Kong, C. Zhang, and S. P. Simonovic, "Optimizing the resilience of interdependent infrastructures to regional natural hazards with combined improvement measures," *Reliability Engineering & System Safety*, vol. 210, p. 107 538, 2021.

[26] Y. Almoghathawi, K. Barker, and L. A. Albert, "Resilience-driven restoration model for interdependent infrastructure networks," *Reliability Engineering & System Safety*, vol. 185, pp. 12–23, 2019.

[27] Y. Fang and G. Sansavini, "Optimizing power system investments and resilience against attacks," *Reliability Engineering & System Safety*, vol. 159, pp. 161–173, 2017.

[28] J. Wu and P. Wang, "Risk-averse optimization for resilience enhancement of complex engineering systems under uncertainties," *Reliability Engineering & System Safety*, p. 107 836, 2021.

[29] X. Zhang, H. Tu, J. Guo, *et al.*, "Braess paradox and double-loop optimization method to enhance power grid resilience," *Reliability Engineering & System Safety*, vol. 215, p. 107 913, 2021.

[30] S. M. Rinaldi, J. P. Peerenboom, and T. K. Kelly, "Identifying, understanding, and analyzing critical infrastructure interdependencies," *IEEE control systems magazine*, vol. 21, no. 6, pp. 11–25, 2001.

[31] N. Sharma, F. Nocera, and P. Gardoni, "Classification and mathematical modeling of infrastructure interdependencies," *Sustainable and Resilient Infrastructure*, vol. 6, no. 1-2, pp. 4–25, 2021.

[32] D. D. Dudenhoeffer, M. R. Permann, and M. Manic, "Cims: A framework for infrastructure interdependency modeling and analysis," in *Proceedings of the 2006 winter simulation conference*, IEEE, 2006, pp. 478–485.

[33] R. Zimmerman, "Social implications of infrastructure network interactions," *Journal of urban technology*, vol. 8, no. 3, pp. 97–119, 2001.

[34] A. Bellè, Z. Zeng, M. Sango, and A. Barros, "Towards a realistic topological and functional modeling for vulnerability analysis of interdependent railway and power networks," *Proceedings of the 31st European Safety and Reliability Conference*, pp. 2063–2070, Jan. 2021. DOI: 10.3850/978-981-18-2016-8_356-cd.

[35] R. Parshani, S. V. Buldyrev, and S. Havlin, "Interdependent networks: Reducing the coupling strength leads to a change from a first to second order percolation transition," *Physical review letters*, vol. 105, no. 4, p. 048 701, 2010.

[36] G. Fu, R. Dawson, M. Khoury, and S. Bullock, "Interdependent networks: Vulnerability analysis and strategies to limit cascading failure," *The European Physical Journal B*, vol. 87, no. 7, pp. 1–10, 2014.

[37] O. Yagan, D. Qian, J. Zhang, and D. Cochran, "Optimal allocation of interconnecting links in cyber-physical systems: Interdependence, cascading failures, and robustness," *IEEE Transactions on Parallel and Distributed Systems*, vol. 23, no. 9, pp. 1708–1720, 2012.

[38] X. Wang, W. Zhou, R. Li, J. Cao, and X. Lin, "Improving robustness of interdependent networks by a new coupling strategy," *Physica A: Statistical Mechanics and its Applications*, vol. 492, pp. 1075–1080, 2018.

[39] S. Chattopadhyay, H. Dai, S. Hosseinalipour, *et al.*, "Designing optimal interlink patterns to maximize robustness of interdependent networks against cascading failures," *IEEE Transactions on Communications*, vol. 65, no. 9, pp. 3847–3862, 2017.

[40] S. Wang, L. Hong, and X. Chen, "Vulnerability analysis of interdependent infrastructure systems: A methodological framework," *Physica A: Statistical Mechanics and its applications*, vol. 391, no. 11, pp. 3323–3335, 2012.

[41] H. Guo, S. S. Yu, H. H. Iu, T. Fernando, and C. Zheng, "A complex network theory analytical approach to power system cascading failure—from a cyber-physical perspective," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 29, no. 5, p. 053 111, 2019.

[42] Z. Chen, J. Wu, Y. Xia, and X. Zhang, "Robustness of interdependent power grids and communication networks: A complex network perspective," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 65, no. 1, pp. 115–119, 2017.

[43] D. F. Rueda and E. Calle, "Using interdependency matrices to mitigate targeted attacks on interdependent networks: A case study involving a power grid and backbone telecommunications networks," *International Journal of Critical Infrastructure Protection*, vol. 16, pp. 3–12, 2017.

[44] N. Sharma and P. Gardoni, "Mathematical modeling of interdependent infrastructure: An object-oriented approach for generalized network-system analysis," *Reliability Engineering & System Safety*, vol. 217, p. 108 042, 2022.

[45] S. G. Nurre, B. Cavdaroglu, J. E. Mitchell, T. C. Sharkey, and W. A. Wallace, "Restoring infrastructure systems: An integrated network design and scheduling (inds) problem," *European journal of operational research*, vol. 223, no. 3, pp. 794–806, 2012.

[46] M. Ouyang and Y. Fang, "A mathematical framework to optimize critical infrastructure resilience against intentional attacks," *Computer-Aided Civil and Infrastructure Engineering*, vol. 32, no. 11, pp. 909–929, 2017.

[47] A. D. González, L. Dueñas-Osorio, M. Sánchez-Silva, and A. L. Medaglia, "The interdependent network design problem for optimal infrastructure system restoration," *Computer-Aided Civil and Infrastructure Engineering*, vol. 31, no. 5, pp. 334–350, 2016.

[48] M. Ouyang, "A mathematical framework to optimize resilience of interdependent critical infrastructure systems under spatially localized attacks," *European Journal of Operational Research*, vol. 262, no. 3, pp. 1072–1084, 2017.

[49] Y.-P. Fang, G. Sansavini, and E. Zio, "An optimization-based framework for the identification of vulnerabilities in electric power grids exposed to natural hazards," *Risk Analysis*, vol. 39, no. 9, pp. 1949–1969, 2019.

[50] L. Zhao and B. Zeng, "Vulnerability analysis of power grids with line switching," *IEEE Transactions on Power Systems*, vol. 28, no. 3, pp. 2727–2736, 2013.

[51] B. Zeng and L. Zhao, "Solving two-stage robust optimization problems using a column-and-constraint generation method," *Operations Research Letters*, vol. 41, no. 5, pp. 457–461, 2013.

[52] L. Zhao and B. Zeng, "An exact algorithm for two-stage robust optimization with mixed integer recourse problems," *submitted, available on Optimization-Online. org*, 2012.

[53] Accessed: 30/03/2022. [Online]. Available: `https://icseg.iti.illinois.edu/ieee-14-bus-system/`.

[54] Accessed: 30/03/2022. [Online]. Available: `https://icseg.iti.illinois.edu/wscc-9-bus-system/`.

[55] M. E. Newman, "Mixing patterns in networks," *Physical review E*, vol. 67, no. 2, p. 026 126, 2003.

[56] L. Gurobi Optimization, *Gurobi optimizer reference manual*, 2021. [Online]. Available: `http://www.gurobi.com`.

Paper IV

A. Bellè, A. F. Abdin, Y.-P. Fang, Z. Zeng, and A. Barros, "A data-driven distributionally robust approach for the optimal coupling of interdependent critical infrastructures," *European Journal of Operational Research*, under revision after first peer-review, 2022.

A data-driven distributionally robust approach for the optimal coupling of interdependent critical infrastructures under random failures

Andrea Bellè[1]*, Adam F. Abdin[2], Yi-Ping Fang[1], Zhiguo Zeng[1], Anne Barros[1]

[1] *Chair on Risk and Resilience of Complex Systems, Laboratoire Génie Industriel, CentraleSupélec, Université Paris-Saclay, 3 Rue Joliot Curie, 91190 Gif-sur-Yvette, France.*
Emails: andrea.belle@centralesupelec.fr, yiping.fang@centralesupelec.fr, zhiguo.zeng@centralesupelec.fr, anne.barros@centralesupelec.fr

[2] *Laboratoire Génie Industriel, CentraleSupélec, Université Paris-Saclay, 3 Rue Joliot Curie, 91190 Gif-sur-Yvette, France.*
Email: adam.abdin@centralesupelec.fr

Abstract

Critical infrastructures (CIs), such as energy systems, transportation networks and telecommunications networks, are the backbone of any advanced society. They are often interconnected to, and interdependent on, each other through complex coupling interfaces. Failures can propagate among different CIs through these coupling interfaces, causing multi-sectoral disruption. The design of the coupling interface can strongly impact the cascading effect between different CIs. In this paper, we propose a data-driven distributionally robust approach for the optimal coupling of interdependent CIs. Our model obtains an optimal coupling interface that maximizes the expected combined performance of interdependent CIs under random failure scenarios with ambiguous probability distribution. We demonstrate the validity of the proposed approach using an ambiguity set built upon a synthetic data set of historical contingency scenarios. Interdependent power and gas networks (IPGNs) are used as an illustrative case study. We show that our proposed approach leads to better coupling interfaces with higher expected performance under disruptive scenarios.

Keywords: Risk analysis, interdependent critical infrastructures, coupling interface, random failures, distributionally robust optimization

*Corresponding author
Email address: andrea.belle@centralesupelec.fr (Andrea Bellè[1])

1. Introduction

1.1. Motivation

Critical infrastructures (CIs), such as energy systems, transportation and telecommunications networks, are large and complex man-made systems which support vital societal functions and represent a driving force in the socioeconomic development (Z. Guo & Haimes, 2016). In fact, many essential services and commodities, such as electricity, public transportation, water and gas supply or telecommunications, are provided by CIs (Ouyang, 2014). Consequently, failures and disruptive events within CIs, such as intentional sabotages, extreme natural hazards or random failures, can cause disruption and considerable negative consequences within a society. For example, the blackout which affected Italy on 28 September 2003 caused over 10 hours of power outages for more than 50 millions people (Corsi & Sabelli, 2004). Thus, a risk-oriented analysis and optimization within these critical systems is an essential task (Bier & Gutfraind, 2019).

CIs are not stand-alone isolated systems, but they are interconnected with, and interdependent on, each other in terms of functionality, reliability and performance (Rinaldi et al., 2001). While these interdependencies increase the operational performance and efficiency of CIs, they often lead to an increased vulnerability (Buldyrev et al., 2010). Interdependent networks and systems are intrinsically more fragile than isolated systems, as a failure within one infrastructure can spread within other infrastructures and cause multi-sectoral disruption (Buldyrev et al., 2010; Lee II et al., 2007; Vespignani, 2010). Ensuring and optimizing the resilience of interdependent CIs are important issues, and they are the main focus of various existing works (e.g. Alkhaleel et al., 2022; Fang and Zio, 2019; Ouyang, 2017; Ouyang and Fang, 2017).

Among the various factors and parameters that can affect the resilience of interdependent systems, network science has demonstrated that the topology of the ensemble of interdependencies, here referred to as coupling interface, can strongly affect the failure propagation between different systems (Parshani et al., 2010). However, designing (or retrofitting) the topology of coupling interfaces of interdependent CIs has been investigated by very few studies.

Coupling interfaces should be designed to ensure the robustness of interdependent systems against different failure and disruption scenarios. As failure scenarios can considerably vary in terms of cause, magnitude and frequency of occurrence, optimizing coupling interfaces lies within the framework of optimization under uncertainty. In our previous work (Bellè et al., 2021), we proposed a framework to optimize the coupling interface design against the worst-case scenario within a set of feasible failure scenarios, exploiting the traditional defender-attacker-defender model. While this approach provides robust solutions, it might suffer from overconservativeness. In this paper, we propose a data-driven distributionally robust approach for the optimal coupling of interdependent critical infrastructures under random failure scenarios. Our proposed model has the advantage of

allowing to tune the conservativeness of the solution with specific parameters, in order to better reflect the attitude towards risk of the decision maker. Moreover, our model gives the possibility to exploit historical data on failures and contingencies for an informed decision making.

1.2. Related work

1.2.1. Design of coupling interface

Despite coupling interfaces playing a key role in the failure propagation between different infrastructures, their design and optimization have been seldomly treated within the existing literature. In fact, in most of the existing works on interdependent CIs, the coupling interface is given and fixed, and no sensitivity analysis of different designs is performed.

The concept of interdependencies topology is typical of network science, and in particular of interdependent networks. In this field, several works analyze the impact of different coupling interfaces on the failure propagation between different types of networks. In fact, it has been shown that the design of coupling interfaces modifies the failure propagation between interdependent networks (Fu et al., 2014; Parshani et al., 2010). Moreover, a proper coupling interface design can reduce the vulnerability and increase the robustness of interdependent networks (Chattopadhyay et al., 2017; X. Wang et al., 2018; Yagan et al., 2012). These works represent the theoretical background that justifies the study of coupling interfaces between CIs. However, the models in these works are mainly based on percolation theory, which fails to detail the complexity of real-world infrastructures. Thus, they can not be directly applied for real-world decision-making.

Some scholars leveraged on network metrics-based coupling strategies to assess the impact of different coupling interfaces in interdependent CIs, such as power and water networks (S. Wang et al., 2012) or power and telecommunication networks (Chen et al., 2017; H. Guo et al., 2019; Rueda & Calle, 2017). These strategies represent only a heuristic approach and, although they provide meaningful insights, they do not ensure solution optimality.

In (Winkler et al., 2011), the authors propose a more sophisticated heuristic model for designing optimal coupling interfaces between complex urban infrastructure systems against external attacks. The coupling interface is based on traditional network metrics, such as node degree, betweenness, clustering coefficient and Euclidean distance. Interface design between power, gas and water networks is given as an illustrative example. A similar approach is presented in (Ouyang & Dueñas-Osorio, 2011), where the authors propose a network metrics-based heuristic method to optimize coupling interfaces against cascading failures and different external attacks scenarios, also accounting for physical features of the different nodes (e.g. supply or demand nodes). However, these approaches still rely on network metrics-based heuristics, and they do not guarantee global optimality of solutions.

In our previous work (Bellè et al., 2021), we proposed a mathematical framework for the optimal coupling of interdependent CIs based on the defender-attacker-defender model (Brown et al., 2005). This model ensures a coupling interface design which is robust against the worst-case realization of the uncertainty, represented by a set of feasible failure scenarios. While this model represents the first mathematical programming approach for the optimal coupling of interdependent CIs, it does not include probabilistic information on the different feasible failure scenarios considered within the optimization. Consequently, the solutions might result overconservative, and the control of the decision-maker over the conservativeness of the model is limited to the size of the uncertainty set, i.e. the set of feasible failure scenarios. To the best of our knowledge, an optimization under uncertainty approach for optimizing the coupling interface of interdependent CIs under random failures, which avoids the the risk of overconservativeness, is missing. An approach based on distributionally robust optimization might be able to offer more control over the level of conservativeness and to leverage historical data for encoding probabilistic information within the decision-making.

1.2.2. Distributionally robust optimization in CIs

When dealing with optimization under uncertainty, different approaches are available. In *robust optimization* (RO), it is assumed that the decision-maker has no knowledge on the uncertainty distribution of the parameters, except for the support sets. The decision-maker thus seeks to optimize the solution against the worst uncertainty realization. RO represents the most risk-averse class of optimization under uncertainty, and it leads to solutions which are often highly conservative.

In *stochastic optimization* (SO), it is assumed that the decision maker has perfect distributional knowledge of the uncertainty. The goal is, thus, to identify a solution that optimizes the expectation of the objective function. SO represents a risk-neutral class of optimization under uncertainty.

In reality, it is often not possible to obtain full information on the uncertainty probability distribution. However, partial probabilistic knowledge of the uncertainty distribution is often available. While it is often insufficient for deriving a probability distribution with high confidence, this partial information can be exploited with a robust stochastic approach that protects the decision-maker from the ambiguity of the available distributional knowledge (Rahimian & Mehrotra, 2019). This approach is called *distributionally robust optimization* (DRO), and it is based on the assumption that the real distribution of uncertainty is unknown, but it is contained within an *ambiguity set*, where the available probabilistic information is encoded. In this framework, the decision-maker seeks to identify a solution that optimizes the expectation of the objective function under the worst distribution within the ambiguity set. The risk-aversion of DRO lies between RO and SO, and it can be tuned by modifying the size of the ambiguity set to correspond to one or the other.

The applications of DRO have seen an increased interest in several research domains. For

example, it has been applied in traditional operational research problems, such as the shortest path problem (Z. Wang et al., 2020), the capacitated facility location problem (Saif & Delage, 2021) or continuous games (Liu et al., 2018). Moreover, DRO has been used for real-world applications, such as retrofitting planning of transportation networks (Doan, 2021), planning of energy and reserve dispatch (Arrigo et al., 2022) or elective surgery scheduling (Shehadeh & Padman, 2021). However, the applications within the context of protection of CIs remain sparse and limited, and mostly confined within power systems applications.

Ambiguity sets can be defined using different methods, and they can be grouped within two main families: moment-based and discrepancy-based ambiguity sets. In this work, we rely on a moment-based ambiguity set, often applied within the context of power networks. For example, moment-based distributionally robust models under random contingency are applied to: configure distribution networks (Babaei et al., 2020), assess the reliability of a transmission network hardening plan (Bagheri & Zhao, 2019), define a contingency-constrained unit commitment model (Zhao & Jiang, 2017), drive the formation of microgrids for service restoration (Cai et al., 2021), plan transmission expansion with distributed energy resources (Alvarado et al., 2018), perform a multi-disaster resilience enhancement of distribution networks (Zhang et al., 2020) and plan spinning reserve in power networks (Li et al., 2022).

These existing works are a clear indication of the validity of DRO approaches with moment-based ambiguity sets within the context of protection of CIs. It is also clear that DRO approaches, in the context of critical infrastructures, have remained confined within the power systems applications, and no application in the general field of protection of interdependent CIs seems to exist.

1.3. Contribution

We propose a novel data-driven DRO model for the optimal design of coupling interfaces between interdependent CIs under random failures. The proposed approach ensures the robustness of the coupling interface in terms of worst expectation of combined performed of the interdependent infrastructures under disruption conditions. The proposed approach is based on a moment-based ambiguity set, built upon a synthetic contingency data set. Using as an illustrative case-study interdependent power and gas networks (IPGNs), our model ensures a coupling interface design which maximizes the expected combined performance of the IPGNs under the worst distribution within the ambiguity set.

The contributions of this papers can be summarized as follows:

- We propose a novel data-driven DRO-based model for the optimal coupling of interdependent CIs, which can be directly applied in real-world situations such as: *(i)* design of the coupling

interface between new CIs, *(ii)* design of the coupling interface between new and existing CIs, *(iii)* retrofitting, analysis and evaluation of existing coupling interfaces.

- We show how to build a data-driven moment-based ambiguity set using a synthetic contingency data set, generated *ad hoc* to recreate realistic historical data of failure scenarios.
- We develop a reliable and efficient solution approach based on the Nested Column&Constraint Generation (NC&CG) algorithm.
- Using IPGNs as a case-study, we show the validity of our approach comparing the optimal coupling interface designs with a traditional euclidean coupling strategy based on geographical vicinity.
- We perform a sensitivity analysis on various parameters, demonstrating the validity of our approach over a spectrum of parametric assumptions.

The rest of this work is organized as follows: in Section 2, the problem is formulated; in Section 3, the solution procedure is detailed; in Section 4, the case-study is presented in details; in Section 5, the main results are reported and discussed; in Section 6, some final remarks are given.

2. Problem formulation

2.1. Modeling framework

Critical infrastructures are interdependent if the state of one system depends on the state/output of another one. Relationships of various nature can exist between elements of different infrastructures, and the interdependencies can be divided into four categories: physical, when the state of one system is dependent on the material output of another system; cyber, when the state of one system is dependent on the information transmitted through another system; geographic, when different systems share the same location and their state can be modified by an environmental event; and logical, if the interdependency is not physical, cyber, or geographic (Rinaldi et al., 2001).

We model each infrastructure using a network flow-based approach (Ouyang, 2014). A network is a mathematical construct described by a graph $G = (V, E)$, where V is the set of N nodes and E is the set of M edges. Each edge k is directed, and it is defined by a tuple (i, j), where i and j represent, respectively, the origin node $O(k)$ and destination node $D(k)$ of the corresponding edge. We assume that each node i is characterized by a production capacity $\overline{p}_i$ (e.g. power or gas flow production) and a requested demand $\overline{d}_i$ (e.g. power or gas flow demand). Similarly, each edge k is defined by a maximum flow capacity $\overline{f}_k$.

In the context of J interdependent infrastructures, the focus of our analysis is the combined performance P_C (Fang & Zio, 2019), generally defined as in (1):

$$P_C = \sum_{j \in J} w_j \frac{\sum_{i \in V_j} d_i}{\sum_{i \in V_j} \overline{d}_i} \tag{1}$$

where w_j represents the weight of infrastructure j, V_j represents the set of nodes in infrastructure j and d_i represents the demand supplied at node i. This equation represents the fraction of requested demand that is satisfied within the interdependent CIs.

In this work, we consider interdependent power and gas networks (IPGNs), which are mutually interdependent on each other with physical interdependencies: equipment in the gas network needs to be supplied with electricity, while gas-fired power plants needs a gas flow supply.

One could imagine the problem of designing the coupling interface as a traditional reliable network design problem (e.g. Bhuiyan et al., 2020). In the reliable network design problem, a decision maker seeks to allocate the links between nodes in order to design a network which satisfies some given conditions in terms of performance and robustness against failures. In the coupling interface design problem, we seek to allocate interdependency links between nodes of different networks in order to guarantee satisfying performance in nominal conditions and robustness in conditions of disruption.

Given a power network and a gas network, we want to identify the coupling interface design which maximizes their combined performance in feasible disruption conditions. The expectation of the combined performance is computed with respect to a set of feasible failure scenarios. Each failure scenario is represented by an ambiguous probability, as defined by the ambiguity set.

In this work, we consider within the set of feasible failure scenarios only contingencies of edges of the power network, which represent power transmission lines.

2.2. Networks operational model

The power network is operated with a traditional DC power flow model, while the gas network is operated with a maximal flow model, which is a suitable linear approximation for any flow-based system or infrastructure (Fang & Zio, 2019; Nurre et al., 2012; Ouyang & Fang, 2017). For a fixed coupling interface and contingency scenario, the operational model of IPGNs is described by the optimization shown in Equation (2), subject to Constraints (3)-(18):

$$\max_{\mathbf{h},\boldsymbol{\delta}} \; w_{PN} \frac{\sum_{i \in V_{PN}} d_i}{\overline{d}_{PN}^{max}} + w_{GN} \frac{\sum_{i \in V_{GN}} d_i}{\overline{d}_{GN}^{max}} \tag{2}$$

subject to:

$$0 \leq p_i \leq \overline{p}_i, \quad \forall i \in V_{TOT} \tag{3}$$

$$0 \leq d_i \leq \overline{d}_i^b + \sum_{j \in V_{GN}} y_{ji}^1 \, \overline{d}_j^{MW}, \quad \forall i \in V_{PN} \tag{4}$$

$$0 \leq d_i \leq \overline{d}_i^b + \sum_{j \in V_{PN}} y_{ji}^2 \, \overline{d}_j^{m^3}, \quad \forall i \in V_{GN} \tag{5}$$

$$-\overline{f}_k \leq f_k \leq \overline{f}_k, \quad \forall k \in E_{GN} \tag{6}$$

$$-u_k \overline{f}_k \leq f_k \leq u_k \overline{f}_k, \quad \forall k \in E_{PN} \tag{7}$$

$$p_i - d_i + \sum_{D(k)=i} f_k - \sum_{O(k)=i} f_k = 0, \ \ \forall i \in V_{TOT} \tag{8}$$

$$x_k f_k - \left(\theta_{O(k)} - \theta_{D(k)}\right) \geq -M_k(1-u_k), \forall k \in E_{PN} \tag{9}$$

$$x_k f_k - \left(\theta_{O(k)} - \theta_{D(k)}\right) \leq M_k(1-u_k), \ \ \forall k \in E_{PN} \tag{10}$$

$$d_i - \delta_i^{PN} \left(\overline{d}_i^b + \sum_{j \in V_{GN}} y_{ji}^1 \, \overline{d}_j^{MW} \right) \geq 0, \quad \forall i \in V_{PN} \tag{11}$$

$$d_i - \delta_i^{GN} \left(\overline{d}_i^b + \sum_{j \in V_{PN}} y_{ji}^2 \, \overline{d}_j^{m^3} \right) \geq 0, \quad \forall i \in V_{GN} \tag{12}$$

$$p_i - \overline{p}_i \sum_{j \in V_{GN}} y_{ij}^2 \, \delta_j^{GN} \leq 0, \quad \forall i \in V_{PN} \tag{13}$$

$$p_i - \overline{p}_i \sum_{j \in V_{PN}} y_{ij}^1 \, \delta_j^{PN} \leq 0, \quad \forall i \in V_{GN} \tag{14}$$

$$d_i - \overline{d}_i^b \sum_{j \in V_{PN}} y_{ij}^1 \, \delta_j^{PN} - \sum_{j \in V_{PN}} y_{ji}^2 \, \overline{d}_j^{m^3} \sum_{j \in V_{PN}} y_{ij}^1 \, \delta_j^{PN} \leq 0, \quad \forall i \in V_{GN} \tag{15}$$

$$-\sum_{\substack{i=O(k) \\ j \in V_{PN}}} y_{ij}^1 \, \delta_i^{PN} \, \overline{f}_k \leq f_k \leq \sum_{\substack{i=O(k) \\ j \in V_{PN}}} y_{ij}^1 \, \delta_i^{PN} \, \overline{f}_k, \quad \forall k \in E_{GN} \tag{16}$$

$$-\sum_{\substack{i=D(k) \\ j \in V_{PN}}} y_{ij}^1 \, \delta_i^{PN} \, \overline{f}_k \leq f_k \leq \sum_{\substack{i=D(k) \\ j \in V_{PN}}} y_{ij}^1 \, \delta_i^{PN} \, \overline{f}_k, \quad \forall k \in E_{GN} \tag{17}$$

$$\delta_i^{PN}, \delta_i^{GN} \in \{0,1\}, \ \theta_j \text{ free}, \ \forall i \in V_{TOT}, \forall j \in V_{PN} \tag{18}$$

The term $\mathbf{h}=\{\mathbf{p}, \mathbf{d}, \mathbf{f}, \boldsymbol{\theta}\}$ represents the continuous operational variables of the optimization, where $\mathbf{p}$, $\mathbf{d}$, $\mathbf{f}$ and $\boldsymbol{\theta}$ are, respectively, production levels and supplied demands of power/gas in each node, power/gas flows in each edge and phase angles of the nodes in the power network. The term $\boldsymbol{\delta}$ represents the binary operational variables of the model, describing the functional states (1 if functional, 0 otherwise) of the interdependency links between the IPGNs. The term $\mathbf{y}=\{y^1_{ij}, y^2_{ij}\}$ represents binary variables which define the coupling interface design. They do not appear below the maximization in (2), as they are not variables of the operational model. The term $y^1_{ji}=1$ if the node $j \in V_{GN}$ is dependent on the node $i \in V_{PN}$ and $y^1_{ji}=0$ otherwise. Similarly, the term $y^2_{ji}=1$ if the node $j \in V_{PN}$ is dependent on the node $i \in V_{GN}$ and $y^2_{ji}=0$ otherwise. As it will be clear in the next section, they are decision variables of the outer stage of the distributionally robust model. The terms u_k represents the failure variables of the power transmission lines. The variable $u_k=0$ if line k is failed, and $u_k=1$ otherwise. Similarly to the variables $\mathbf{y}$, the variables $\mathbf{u}$ do not appear below (2), but they are present in the ambiguity set and they will be explained in the next section. The objective function in (2) represents the combined performance of the IPGNs in terms of fraction of requested demand of electricity and gas which is supplied. The subscripts PN and GN denote the power network and the gas network, respectively. The subscript TOT denotes the union of power and gas networks elements (e.g. $V_{TOT} = V_{PN} \cup V_{GN}$). The variable d_i represents the supplied demand of power/gas in the node i, while the terms $\overline{d}^{max}_{PN}$ and $\overline{d}^{max}_{GN}$ represent the total requested demand of power and gas in the power and gas network, respectively. Contrary to Equation (1), we define the requested demand in terms of networks rather than single nodes; this is because the requested demand in each node is not a constant parameter, but it depends on the coupling interface, defined by the variables $\mathbf{y}=\{y^1_{ij}, y^2_{ij}\}$. The importance of each infrastructure when computing the combined performance is given by the weights w_{PN} and w_{GN}. The constraints are shown in (3)-(18). Constraint (3) states that the production level of power/gas p_i in each node is bounded between 0 and the node production capacity $\overline{p}_i$. The supplied demand of power/gas d_i in each node is bounded between 0 and the requested demand, as enforced by Constraints (4)-(5). Contrary to traditional approaches, the requested demand in each node is not a fixed parameter but it depends on the coupling interface, as shown in (4) and (5). The requested demand in each node i, either in the power or gas network, is composed of two terms:

- The term $\overline{d}^b_i$ is the base requested demand, which represents various consumers of electricity/gas (households, industries, etc.).
- The second terms, $\sum_{j \in V_{GN}} y^1_{ji}\, \overline{d}^{MW}_j$ in (4) and $\sum_{j \in V_{PN}} y^2_{ji}\, \overline{d}^{m^3}_j$ in (5), represent the electricity/gas demand of all the nodes of the other infrastructures dependent on the node i. The terms $\overline{d}^{MW}_j$ and $\overline{d}^{m^3}_j$ represent, respectively, the electricity demand of node $j \in V_{GN}$ and the

gas flow demand of node $j \in V_{PN}$. We recall that the term y^1_{ji} has value 1 if the node $j \in V_{GN}$ is dependent on the node $i \in V_{PN}$ and 0 otherwise. Similarly, the term y^2_{ji} has value 1 if the node $j \in V_{PN}$ is dependent on the node $i \in V_{GN}$ and 0 otherwise.

Constraints (6) and Constraint (7) describes, respectively, the constraint on the flow of gas and power in each edge, which is bounded, in absolute value, by the flow capacity $\overline{f}_k$. Since in this study we consider random contingencies of power transmission lines, the variable u_k=0 when the edge k is failed and u_k=1 otherwise. The net balance between generation, demand and flow in each node is guaranteed by Constraint (8). The DC power flow assumption within the power network is enforced by Constraints (9)-(10).

The functional state of each interdependency link is described in (11) and (12). Each interdependency link from the node $i \in V_{PN}$ to the nodes $j \in V_{GN}$ is functional if the variable δ^{PN}_i is equal to 1; otherwise, δ^{PN}_i=0 and the all the nodes in V_{GN} that are depending on node i are not functional. We assume, as shown in (11), that each variable δ^{PN}_i can take the value 1 only when the requested demand in the node i is fully supplied (Fang & Zio, 2019). The same assumption is applied for the gas network and the corresponding variables δ^{GN}_i, as shown in (12).

The production level p_i in each node depends on the coupling interface. We assume that the production p_i in each node ranges between 0 and the production capacity $\overline{p}_i$ if there is one functional interdependency, and 0 otherwise. For example, as shown in (13), the production p_i in the node $i \in V_{PN}$ ranges between 0 and $\overline{p}_i$ if there is one interdependency link from a node $j \in V_{GN}$ in the gas network (y^2_{ij}=1) properly functional (δ^{GN}_j=1); otherwise, p_i takes the value 0. The rationale behind this assumption is that, in this work, each node in V_{PN} with a production capacity higher than 0 is assumed to contain a gas-fired power plant. The same assumption is applied for the gas network and the corresponding interdependency link from the power network, as shown in (14). In addition, we assume that, in the gas network, supplied demands and flows in the pipes are also dependent on the interdependency from the power network. As shown in (15), the supplied demand d_i in the node $i \in V_{GN}$ ranges between 0 and the requested demand if there is one interdependency from a node $j \in V_{PN}$ in the power network (y^1_{ij}=1) properly functional (δ^{PN}_j=1); otherwise, d_i takes the value 0. The rationale behind these assumptions is that each node in V_{GN} needs electricity to run various equipment, such as pumps, valves or compressors. Lastly, we assume that the absolute value of the flow f_k in each pipe $k \in E_{GN}$ ranges between 0 and the maximum flow capacity $\overline{f}_k$ only if both the origin and destination node of k, respectively $O(k)$ and $D(k)$, have a functional interdependency from the power network, as expressed in (16) and (17); otherwise, f_k takes the value 0.

For clarity, we can express the model in (2)-(18) with its compact matrix formulation, shown in (19)-(20):

$$\max_{\mathbf{h},\boldsymbol{\delta}} \mathbf{b}^T\mathbf{h} \tag{19}$$

subject to:

$$\mathbf{Rh} \leq \mathbf{q} - \mathbf{Tu} - \mathbf{Hy} - \mathbf{W}\boldsymbol{\delta} - \mathbf{y}^T\mathbf{D}\boldsymbol{\delta} \tag{20}$$

where $\mathbf{h}$ is the vector containing the continuous variables, $\mathbf{u}$, $\mathbf{y}$ and $\boldsymbol{\delta}$ are vector containing the binary variables, $\mathbf{b}$ is the vector containing the parameters of the objective function, $\mathbf{R}$, $\mathbf{T}$, $\mathbf{H}$, $\mathbf{W}$ and $\mathbf{D}$ are matrices containing the constraints parameters and $\mathbf{q}$ is the vector containing the constraints constants.

Due to the multiplication between $\mathbf{y}$ and $\boldsymbol{\delta}$ variables in Constraints (11)-(17), the formulation presents nonlinear terms. However, they can be linearized by introducing new binary variables, contained within the vectors $\mathbf{z}$ and $\mathbf{r}$, and new parameters matrices $\mathbf{V}$ and $\mathbf{S}$, and Constraints (20) can be replaced by (21).

$$\mathbf{Rh} \leq \mathbf{q} - \mathbf{Tu} - \mathbf{Hy} - \mathbf{W}\boldsymbol{\delta} - \mathbf{Sz} - \mathbf{Vr} \tag{21}$$

2.3. Distributionally robust formulation

We formulate a distributionally robust problem for the optimal coupling interface of IPGNs under random failures, as it is shown in (22)-(27):

$$\max_{\substack{\mathbf{h}^0,\boldsymbol{\delta}^0 \\ \mathbf{y}\in\{0,1\}}} \min_{\mathbb{P}\in\mathcal{M}} \mathbb{E}_{\mathbb{P}}\left[\mathbf{Q}(\mathbf{y},\boldsymbol{\xi})\right] \tag{22}$$

subject to:

$$\sum_{j\in V_{PN}} y_{ij}^1 \leq 1, \quad \forall i \in V_{GN} \tag{23}$$

$$\sum_{j\in V_{GN}} y_{ij}^2 \leq 1, \quad \forall i \in V_{PN} \tag{24}$$

$$\sum_{\substack{i\in V_{GN} \\ j\in V_{PN}}} y_{ij}^1\, d_{ij}^{km}\, c_{km}^1 + \sum_{\substack{i\in V_{PN} \\ j\in V_{GN}}} y_{ij}^2\, d_{ij}^{km}\, c_{km}^2 \leq B_c \tag{25}$$

$$\mathbf{R^0h^0} \leq \mathbf{q^0} - \mathbf{H^0y} - \mathbf{W^0}\boldsymbol{\delta^0} - \mathbf{y}^T\mathbf{D^0}\boldsymbol{\delta^0} \tag{26}$$

$$\mathbf{b}^T\mathbf{h^0} \geq 1 \tag{27}$$

and with the recourse function $\mathbf{Q}(\mathbf{y}, \boldsymbol{\xi})$ representing the operational model of the IPGNs, previously shown in (19) and (21).

In the first stage, a coupling interface between the IPGNs is allocated through the binary variables $\mathbf{y}=\{y_{ij}^1, y_{ij}^2\}$, with the aim of maximizing the expected combined performance of the IPGNs in disrupted conditions. The variable y_{ij}^1 equals to 1 if the node $i \in V_{GN}$ is coupled to and dependent on the node $j \in V_{PN}$, and 0 otherwise; similarly, the binary variable y_{ij}^2 equals to 1 if the node $i \in V_{PN}$ is coupled to and dependent on the node $j \in V_{GN}$, and 0 otherwise. We assume that each node in the gas network that needs electricity is dependent on one and only one node in the power network, as shown in constraint (23); similarly, each node in the power network that requires gas flow supply is dependent on one and only one node in the gas network, as shown in constraint (24). Coupling two nodes has a cost per kilometer, and the total cost of the allocated interdependencies is bounded by the available monetary budget B_c. This is expressed in Constraint (25), where d_{ij}^{km} is the distance in km between nodes i and j and the terms c_{km}^1 and c_{km}^2 are, respectively, the cost per km of placing a coupling link from the power to the gas network and from the gas to the power network.

The coupling interface must be allocated in a way such that, in nominal conditions (no components failed), the requested demands of electricity and gas are fully satisfied. This condition is enforced by Constraint (27), which depends on the operational constraints of the IPGNs, expressed in Constraint (26). The variables $\mathbf{h^0}=\{\mathbf{p^0}, \mathbf{d^0}, \mathbf{f^0}, \boldsymbol{\theta^0}\}$ and $\boldsymbol{\delta^0}$ are the operational variables of the first stage. Constraint (26) does not contain $\mathbf{u}$ variables because every line is considered as functional.

In the second stage, the probability distribution of feasible failure scenarios is chosen in order to minimize the expected combined performance of the IPGNs. The set of feasible failure scenarios $\mathcal{A}$ is defined in terms of maximum number of failed components K. In this work, we assume that only edges in the power network can fail. Each failure scenario k is defined by an M_{PN}-dimensional vector $\mathbf{u}_k$, where the i^{th} component defines the functional state (0 if failed, 1 if functional) of the power transmission line i. The set of feasible failure scenarios can be defined as in Equation (28):

$$\mathcal{A} = \left\{\mathbf{u} | \mathbf{u} \in 2^{M_{PN}}, ||\mathbf{u}||_1 \geq M_{PN} - K\right\} \tag{28}$$

where M_{PN} is the number of edges in the power network. Given the condition that one of the scenarios in $\mathcal{A}$ occurred (or, in other words, at least one edge failed), the conditional probability of each scenario $k \in \mathcal{A}$ to have occurred is defined as ϕ_k. The scenarios are mutually exclusive and their probability distribution is defined by a multinomial distribution $\mathbb{P}$ that respects the condition in (29):

$$\sum_{k \in \mathcal{A}} \phi_k = 1. \tag{29}$$

As it can be clearly seen from the formulation in (22), the distribution $\mathbb{P}$ of the scenario occurrence probabilities ϕ_k is a variable of the optimization problem. In particular, the distribution $\mathbb{P}$ is chosen among a family of distributions defined by the ambiguity set $\mathcal{M}$. This set contains all the the multinomial distributions that can describe the conditional occurrence probability of the feasible failure scenarios and respect some given conditions. Depending on the conditions which are enforced, different ambiguity set can be defined. In this work, we rely on a traditional moment-based ambiguity set, shown in Equation (30), that enforces the conditional marginal failure probability of each line $i \in E_{PN}$ to be between 0 and an upper bound π_i^{max}:

$$\mathcal{M} = \{\mathbb{P} \in \mathcal{P}(\mathcal{A}) : 0 \leq \mathbb{E}_{\mathbb{P}}[1 - \mathbf{u}] \leq \boldsymbol{\pi}^{max}\} \tag{30}$$

where $\mathcal{P}(\mathcal{A})$ defines the set of all probability distributions on a σ-algebra of $\mathcal{A}$ and $\boldsymbol{\pi}^{max}$ is the vector containing the upper bounds of the marginal failure probabilities of each power transmission line. We refer to *conditional* marginal failure probability of each line because we assume that at least one line is failed. In other words, given the condition that at least one line is failed, the conditional marginal failure probability denotes the probability of each line to be failed.

The goal of the model is to identify the coupling interface design that maximizes the expected combined performance of the IPGNs in disrupted conditions (at least one line failed) under the worst multinomial distribution within the ambiguity set.

3. Solution procedure

Our strategy involves a problem reformulation into a form which can be recast and solved through a Nested Column&Constraint Generation (NC&CG) algorithm. This tractable reformulation is shown in Proposition 1, and its proof is available in Appendix A.

Proposition 1. *The problem in* (22)-(27) *is equivalent to the problem in* (31) *subject to* (21), (23)-(27) *and* (32).

$$\max_{\substack{\mathbf{h}^0, \boldsymbol{\delta}^0 \\ \mathbf{y} \in \{0,1\} \\ \boldsymbol{\beta} \geq 0}} \min_{\mathbf{u} \in \{0,1\}} \max_{\mathbf{h}, \boldsymbol{\delta}} \mathbf{b}^T \mathbf{h} + \sum_{i \in E_{PN}} \beta_i (1 - u_i - \pi_i^{max}) \tag{31}$$

$$\sum_{k \in E_{PN}} (1 - u_k) \leq K \tag{32}$$

where (32) defines the set of feasible failure scenarios $\mathcal{A}$.

The formulation in Proposition 1 is a trilevel mixed-integer linear programming (MILP) problem, presenting a structure that can be solved by a NC&CG approach. The formulation in (31) must

be expanded into a max-min-max-max problem, by separating the binary and continuous variables in the last stage. Subsequently, the problem must be split into an outer and inner layer, that must be solved separately and iteratively. For simplicity, in the following explanation, we rely on the compact matrix formulation. The flowchart of the NC&CG algorithm is shown Figure 1.

3.1. Inner layer

The inner layer consists in the solution of the middle lower-level problem (min-max) in (31) with fixed $\mathbf{y}^*$ and $\boldsymbol{\beta}^*$ variables. The output of the model is an optimal failure scenario $\widehat{\mathbf{u}}$, which will be passed to the outer layer. For a fixed coupling interface $\mathbf{y}^*$, fixed variables $\boldsymbol{\beta}^*$, fixed interdependency variables $\boldsymbol{\delta}^*$ and fixed failure scenario $\mathbf{u}^*$, the compact form of the lower-level maximization in (31) and the relative constraints are shown in (33)-(34):

$$\max_{\mathbf{h}} \mathbf{b}^T\mathbf{h} \tag{33}$$

subject to :

$$\mathbf{Rh} \leq \mathbf{q} - \mathbf{Tu}^* - \mathbf{Hy}^* - \mathbf{W}\boldsymbol{\delta}^* - \mathbf{y}^{*T}\mathbf{D}\boldsymbol{\delta}^* \tag{34}$$

and plus the constant $\sum_{i\in E_{PN}} \beta_i^*(1-u_i^*-\pi_i^{max})$. Its dual form is expressed in (35)-(36):

$$\min_{\boldsymbol{\lambda}}(\mathbf{q} - \mathbf{Tu}^* - \mathbf{Hy}^* - \mathbf{W}\boldsymbol{\delta}^* - \mathbf{y}^{*T}\mathbf{D}\boldsymbol{\delta}^*)^T\boldsymbol{\lambda} \tag{35}$$

subject to:

$$\mathbf{R}^T\boldsymbol{\lambda} = \mathbf{b} \tag{36}$$

and plus the constant $\sum_{i\in E_{PN}} \beta_i^*(1-u_i^*-\pi_i^{max})$.

As the interdependency variables $\boldsymbol{\delta}$ are binary, the set of possible interdependency functional states $\mathcal{D}$ is finite. Following the classic C&CG approach, we can leverage on a partial set $\mathcal{D}_{part} \subseteq \mathcal{D}$ of interdependency variables combinations to identify efficiently the exact optimal solution. This is achieved employing the following steps:

1. Set $j = 0$, upper bound $UB_{inr} = \infty$, lower bound $LB_{inr} = 0$ and $\mathcal{D}_{part} = \emptyset$
2. Solve the inner master problem in Equations (37)-(40). Obtain an optimal solution $\widehat{\rho}^{(j)}$ and $\widehat{\mathbf{u}}^{(j)}$. Update LB_{inr}=$\widehat{\rho}^{(j)} + \sum_{i\in E_{PN}} \beta_i^*(1-\pi_i^{max})$.

$$\min_{\rho,\mathbf{u},\boldsymbol{\lambda}} \rho \tag{37}$$

subject to:

$$\rho \geq (\mathbf{q} - \mathbf{Tu} - \mathbf{Hy}^* - \mathbf{W}\boldsymbol{\delta}^{*(j)} - \mathbf{y}^{*T}\mathbf{D}\boldsymbol{\delta}^{*(j)})^T\boldsymbol{\lambda}^{(j)} - \sum_{i\in E_{PN}} \beta_i^* u_i, \quad \forall \boldsymbol{\delta}^{*(j)} \in \mathcal{D}_{part} \tag{38}$$

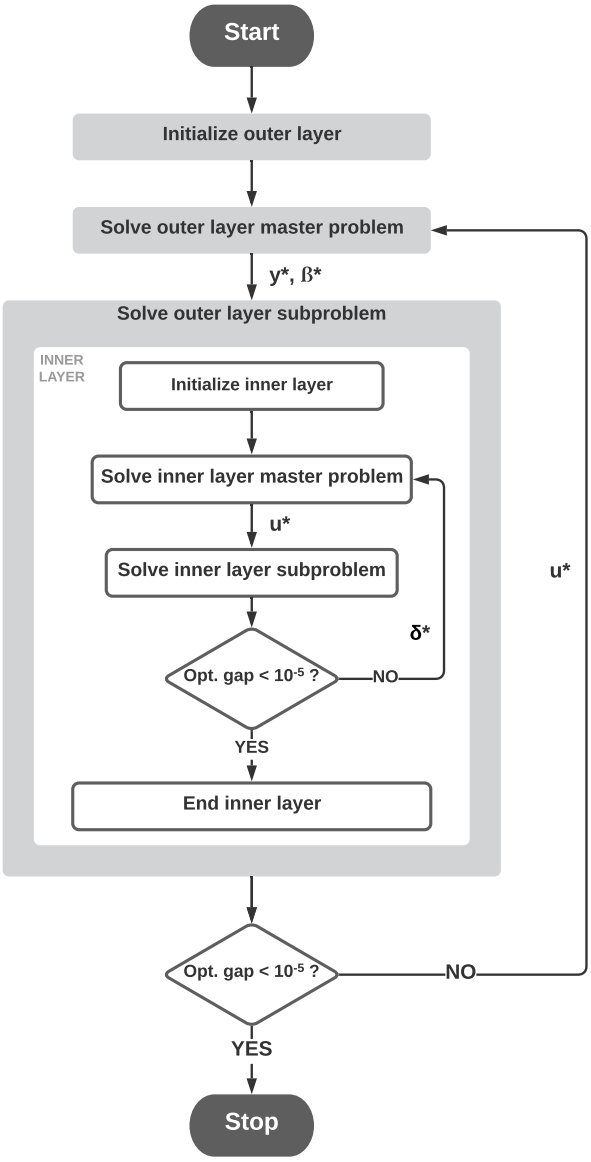

Figure 1: Flowchart of the Nested Column&Constraint Generation algorithm.

$$\mathbf{R}^T \boldsymbol{\lambda}^{(j)} = \mathbf{b}, \quad \forall \boldsymbol{\delta}^{*(j)} \in \mathcal{D}_{part} \tag{39}$$

$$\sum_{k \in E_{PN}} (1 - u_k) \leq K \tag{40}$$

3. Solve the inner subproblem in Equations (41)-(42) with $\widehat{\mathbf{u}}^{(j)}$=$\mathbf{u}^*$. Obtain an optimal solution $\mathbf{b}^T \widehat{\mathbf{h}}^{(j)}$ and $\widehat{\boldsymbol{\delta}}^{(j)}$. Set $UB_{inr} = \min\left(UB_{inr}, \mathbf{b}^T \widehat{\mathbf{h}}^{(j)} + \sum_{i \in E_{PN}} \beta_i^* (1 - u_i^* - \pi_i^{max})\right)$.

$$\max_{\mathbf{h}, \boldsymbol{\delta}} \mathbf{b}^T \mathbf{h} \tag{41}$$

subject to :

$$\mathbf{Rh} \leq \mathbf{q} - \mathbf{Tu}^* - \mathbf{Hy}^* - \mathbf{W}\boldsymbol{\delta} - \mathbf{y}^{*T}\mathbf{D}\boldsymbol{\delta} \tag{42}$$

4. If $(UB_{inr} - LB_{inr})/UB_{inr} < 10^{-5}$, the current solution $\widehat{\mathbf{u}}^{(j)}$ corresponds to the optimal attack plan, the optimal objective value $\widehat{\mathcal{H}}_{inr}(\mathbf{y}^*, \boldsymbol{\beta}^*) = UB_{inr} \approx LB_{inr}$ and the algorithm can be terminated. Otherwise, $\mathcal{D}_{part} = \mathcal{D}_{part} \cup \widehat{\boldsymbol{\delta}}^{(j)}$. Set $j \leftarrow j + 1$ and return to step 2.

This algorithm corresponds to the inner layer in Figure 1. Its output is the optimal failure scenario $\widehat{\mathbf{u}}$ that is passed to the outer layer, and the optimal objective value $\widehat{\mathcal{H}}_{inr}(\mathbf{y}^*, \boldsymbol{\beta}^*)$.

3.2. Outer layer

The upper-level problem can be solved in a similar way, by employing a partial set of feasible failure scenarios $\mathcal{A}_{part} \subseteq \mathcal{A}$. While the inner layer solves a bi-level min-max problem, the outer layer solves a bi-level max-min problem, where the minimization represents the outer subproblem, and it is solved by the inner layer in the previous section.

To be consistent with the probability framework introduced, the outer master problem should be solved with a partial set $\mathcal{A}_{part}$ that does not contain the scenario with no failed lines, defined as $\mathbf{u}_\emptyset$. Otherwise, it would be numerically possible to assign some probability mass to the scenario $\mathbf{u}_\emptyset$, violating the condition that at least one line has failed. This is in contrast with our definition of conditional occurence probability, where the condition is that at least one edge of the power network has failed. However, solving the model without the scenario $\mathbf{u}_\emptyset$ introduces some infeasibility problem within the optimization, particularly in the outer master problem in (45)-(50). An example of this situation, including an explanation on why the the scenario $\mathbf{u}_\emptyset$ has not been removed from the set $\mathcal{A}$ since the beginning, is available in the Supplementary Material.

Proposition 2. *For the problem in* (45)-(50) *to be feasible with a partial set* $\mathcal{A}_{part}$ *that does not contain the scenario* $\mathbf{u}_\emptyset$ *with no failed lines, the two following necessary conditions should be respected:*

1. *the sum of* π_i^{max} *in* (30) *must be at least 1, as shown in* (43)*:*

$$\sum_{i \in E_{PN}} \pi_i^{max} \geq 1; \tag{43}$$

2. *the partial set* $\mathcal{A}_{part}$ *must contain enough scenarios to ensure that the linear system in* (44)*, with* ϕ_k *as variables, has at least one solution:*

$$\begin{cases} \sum_{k \in \mathcal{A}_{part}} \phi_k = 1 \\ \sum_{k \in \mathcal{A}_{part}(1)} \phi_k \leq \pi_1^{max} \\ \vdots \\ \sum_{k \in \mathcal{A}_{part}(M_{PN})} \phi_k \leq \pi_{M_{PN}}^{max} \end{cases} \tag{44}$$

where $\mathcal{A}_{part}(i) \subseteq \mathcal{A}_{part}$ *is the subset containing all the scenarios where line* i *is failed and* M_{PN} *is the number of lines within the power network.*

Proof *The proofs of the two conditions mentioned above are straightforward:*

1. *the condition in* (43) *implies that the probability that at least one line is failed is 1. If the sum in* (43) *was lower than 1, it would automatically imply that, in order to respect the condition in* (29)*, some probability mass is assigned to the scenario* $\mathbf{u}_\emptyset$ *with no failed lines;*
2. *the system in* (44) *represents the conditions enforced by the ambiguity set, and it corresponds to Constraints* (A.3) *and* (A.4) *in Appendix A. The variables* ϕ_k *represent the probability of each scenario* $k \in \mathcal{A}_{part}$*. The system can have 0, 1 or infinite solutions. In case it has no solutions, it means that the probability mass can not be allocated such to respect the moment-based probabilistic conditions enforced by the ambiguity set.*

Similar to the inner layer, the outer layer is solved with a C&CG algorithm with the following steps:

1. Set $j = 0$, upper bound $UB_{otr} = \infty$, lower bound $LB_{otr} = 0$ and $\mathcal{A}_{part} = \emptyset$ and $\mathcal{A}_{part}^{\emptyset} = \{\mathbf{u}_\emptyset\}$, where $\mathbf{u}_\emptyset$ is the scenario where every line is functional.
2. Solve the outer master problem in Equations (45)-(50). If it is not feasible, solve it using $\mathcal{A}_{part}^{\emptyset}$. Obtain an optimal solution $\widehat{\eta}^{(j)}$, $\widehat{\mathbf{y}}^{(j)}$ and $\widehat{\boldsymbol{\beta}}^{(j)}$. Update UB_{otr}=min$(UB_{otr}, \widehat{\eta}^{(j)})$.

$$\max_{\substack{\eta, \mathbf{h}^{(j)}, \boldsymbol{\beta} \geq 0 \\ \mathbf{h^0}, \boldsymbol{\delta^0} \\ \mathbf{y} \in \{0,1\}}} \eta \tag{45}$$

$$\eta \leq \mathbf{b}^T \mathbf{h}^{(j)} + \sum_{i \in E_{PN}} \beta_i (1 - u_i^{*(j)} - \pi_i^{max}), \quad \forall\, \mathbf{u}^{*(j)} \in \mathcal{A}_{part} \tag{46}$$

$$\mathbf{Py} \leq \mathbf{g} \tag{47}$$

$$\mathbf{R^0 h^0} \leq \mathbf{q^0} - \mathbf{H^0 y} - \mathbf{W^0 \delta^0} - \mathbf{S^0 z^0} - \mathbf{V^0 r^0} \tag{48}$$

$$\mathbf{b}^T \mathbf{h^0} \geq 1 \tag{49}$$

$$\mathbf{Rh}^{(j)} \leq \mathbf{q} - \mathbf{Tu}^{*(j)} - \mathbf{Hy} - \mathbf{W}\boldsymbol{\delta}^{(j)} - \mathbf{Sz}^{(j)} - \mathbf{Vr}^{(j)},\ \forall\, \mathbf{u}^{*(j)} \in \mathcal{A}_{part} \tag{50}$$

where $\mathbf{P}$ and $\mathbf{g}$ are the coefficient matrix and the parameter vector of constraints in (23)-(25).

3. Solve the outer subproblem using the inner C&CG algorithm explained in the previous subsection with $\widehat{\mathbf{y}}^{(j)}$=$\mathbf{y}^*$ and $\widehat{\boldsymbol{\beta}}^{(j)}$=$\boldsymbol{\beta}^*$. Obtain an optimal attack plan $\widehat{\mathbf{u}}^{(j)}$. Update $LB_{otr} = \widehat{\mathcal{H}}_{inr}(\mathbf{y}^*, \boldsymbol{\beta^*})$.
4. If $(UB_{otr} - LB_{otr})/UB_{otr} < 10^{-5}$ and the master problem was solved with the partial set $\mathcal{A}_{part}$, the current solution $\widehat{\mathbf{y}}^{(j)}$ corresponds to the optimal coupling interface and the algorithm can be terminated. Otherwise, $\mathcal{A}_{part} = \mathcal{A}_{part} \cup \widehat{\mathbf{u}}^{(j)}$ and $\mathcal{A}^{\emptyset}_{part} = \mathcal{A}^{\emptyset}_{part} \cup \widehat{\mathbf{u}}^{(j)}$. Set $j \leftarrow j + 1$ and return to step 2.

The output represents the optimal coupling interface which maximizes the expected combined performance of the IPGNs under the worst multinomial distribution of random failures within the ambiguity set $\mathcal{M}$. The optimal objective value represents the expected combined performance in disrupted conditions.

4. Illustrative case study

As an illustrative example, we consider a power network with a topology based on the IEEE 14-bus system, and a gas network with a topology based on the IEEE 9-bus system. We assume that the two infrastructures are placed within a 300×300 km^2 geographical area, as shown in Figure 2. Each infrastructure is assumed to have an equal weight, i.e. w_{PN}=0.5 and w_{GN}=0.5. This value represents the importance of each infrastructure when computing the combined performance. Other parameter values are available in the Supplementary Material.

For the baseline case, we consider a maximum number of failed edges K=3 and a monetary budget B_c=1100\$. Sensitivity analysis on these parameters are also performed. We choose a representative interdependency cost-per-kilometer of 1 \$/km, for both c^1_{km} and c^2_{km}. Finally, all the computations are implemented in the Python API of Gurobi 9.1 (Gurobi Optimization, 2021) and performed on a desktop PC with a 3.20 GHz CPU and 32 GB RAM.

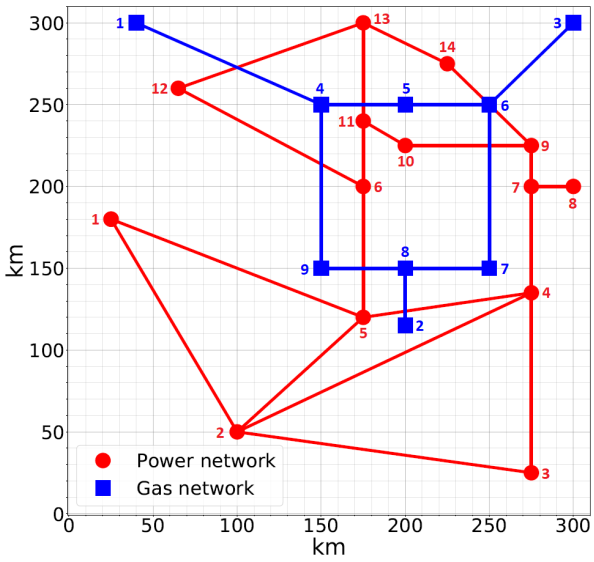

Figure 2: Geographical allocation of the interdependent power and gas network.

4.1. Synthetic data set and ambiguity set

The ambiguity set can be defined with a data-driven approach, starting from historical data on contingencies. In this illustrative case study, we generate synthetic contingencies historical data from assumed line failure rates. The following assumptions are taken for generating the data set:

- each line is characterized by an independent failure rate λ_i. For simplicity, we do not consider any correlations, and each line fails independently. The failure rates utilized in this work are shown in the Supplementary Material;
- we consider daily resolution, simulating 30 years of contingency scenarios;
- each line, when failed, is assumed to be repaired immediately;
- we compute the arrival times T_i (failures) of each line using Equation (51):

$$T_i = -\frac{\ln(s)}{\lambda_i} \tag{51}$$

  where s is a random number between 0 and 1.

We obtain a synthetic data set of 252 contingencies, including 245 scenarios $N-1$, 2 scenarios $N-2$ and 1 scenario $N-3$. From these data, we generate the upper bounds $\boldsymbol{\pi}^{max}$ for the ambiguity set with the following steps. Firstly, we compute the estimated conditional marginal failure probability $\hat{\pi}_i$ of each line i with Equation (52):

$$\hat{\pi}_i = \sum_{k \in \mathcal{A}(i)} \frac{x_k}{N_s} \tag{52}$$

where $\mathcal{A}(i) \subseteq \mathcal{A}$ is the subset which contains all the failure scenarios in which line i is failed, x_k is the number of times that scenario k has occurred, and N_s is the total number of contingency scenarios occurred. Secondly, we compute the associated 95% confidence intervals for the conditional marginal failure probability of each line i using Equation (53):

$$CI_i^{95} = Z\sqrt{\frac{\hat{\pi}_i(1-\hat{\pi}_i)}{N_s}} \tag{53}$$

where Z is the 95% confidence interval constant, equal to 1.96.

From the values $\hat{\pi}_i$ and CI_i^{95}, and aiming for a conservative solution, we rely on the upper bound of the conditional marginal failure probability π_i^{max} of each line i, computed using Equation (54):

$$\pi_i^{max} = \hat{\pi}_i + CI_i^{95}. \tag{54}$$

The obtained $\boldsymbol{\pi}^{max}$ values are shown in Table 1 and they are directly used within the definition of ambiguity set in (30).

Table 1: Upper bounds $\boldsymbol{\pi}^{max}$ for each power transmission line i.

Line	π_i^{max}	**Line**	π_i^{max}
1	0.064	11	0.136
2	0.012	12	0.037
3	0.089	13	0.037
4	0.037	14	0.131
5	0.164	15	0.089
6	0.012	16	0.159
7	0.084	17	0.053
8	0.043	18	0.117
9	0.031	19	0.019
10	0.089	20	0.108

5. Results

5.1. Baseline case

We solve the distributionally robust optimal coupling with a maximum number of failed edges K=3 and a monetary budget B_{ci}=1100\$. We evaluate the dependence of the results on the size of the ambiguity set by changing the upper bounds of the conditional marginal failure probability $\boldsymbol{\pi}^{max}$. The results of the optimal coupling interface, in terms of worst expected performance

in disrupted conditions, are compared with the ones of a coupling interface based on euclidean distance, where each node in one infrastructure is coupled with the geographically-closest node in the other infrastructure. This result is obtained simply setting the monetary budget B_c=822.76\$, corresponding to the minimum budget for the optimization to be feasible (or, in other words, the budget corresponding to the euclidean coupling interface). We also evaluate the worst-case performance associated with the optimal coupling interface and the set of feasible failure scenarios. The worst-case scenario can be easily identified by enumeration or solving the model in (55) subject to (21) and (32), also solvable by a C&CG approach.

$$\min_{\mathbf{u}\in\{0,1\}} \max_{\mathbf{h},\boldsymbol{\delta}} \mathbf{b}^T\mathbf{h}\,(\mathbf{y}^*). \tag{55}$$

where $\mathbf{y}^*$ is a fixed coupling interface.

The results are shown in Figure 3. In the x-axis, we plot different upper bounds $\boldsymbol{\pi}^{max}$, starting from the values in Table 1 and multiplying them by factors from 2 to 6. The last point in the x-axis corresponds to the case where the upper bound of each conditional marginal failure probability π_i^{max} is equal to 1. In the y-axis, we can see the performance levels in disrupted conditions, in terms of fraction of requested electricity and gas demand which is supplied. Four curves are shown: the blue curves represent the worst expected performance associated to the optimal and euclidean coupling interfaces, computed by the distributionally robust optimization model; the red curves represent the worst-case scenario associated to the optimal and euclidean coupling interfaces. As it can be clearly seen in Figure 3, the euclidean coupling interface is outperformed by the optimal coupling interface. Firstly, the worst expected performance are significantly higher in the optimal case. For example, using the upper bounds in Table 1, the optimal case leads to worst expected performance of 0.983, while the euclidean case, for the same upper bounds, to worst expected performance of 0.901. As it is graphically evident, this difference increases as the upper bounds $\boldsymbol{\pi}^{max}$ increases. Secondly, the associated worst-case performance is also significantly better in the optimal case. For example, using the upper bounds in Table 1, the optimal case leads to an associated worst-case performance of 0.902, while the euclidean case, for the same upper bounds, leads to an associated worst-case performance of 0.703. This is a clear indication of: *i)* the importance of the coupling interface in ensuring the robustness of interdependent critical infrastructures under random failures, and *ii)* a distributionally robust approach seems to lead to satisfying results both in terms of worst expected performance and associated worst-case performance.

Increasing the upper bounds $\boldsymbol{\pi}^{max}$, by multiplying them by factors from 2 to 6, increases the conservativeness of the solution. In fact, as it can be clearly seen, the worst expected performance decreases, while the worst-case performance increases. This is because increasing the upper bounds makes the worst probability distribution more disruptive. If we set the upper bounds $\boldsymbol{\pi}^{max}$ equal to

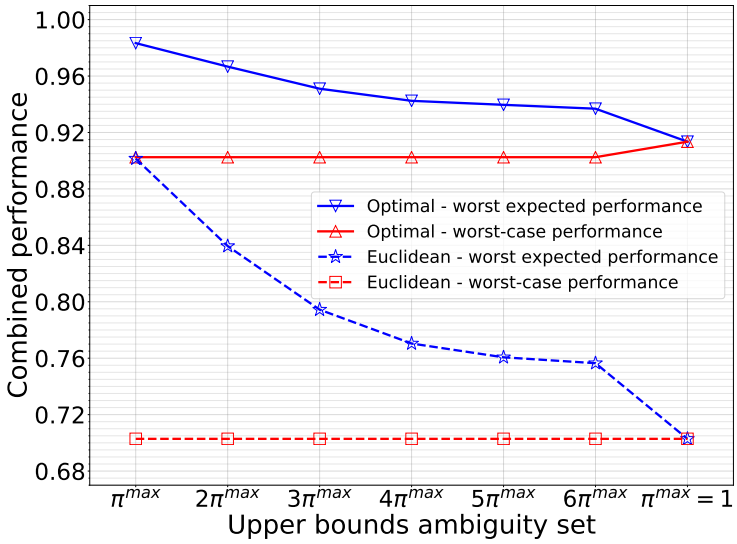

Figure 3: Results of the distributionally robust optimal coupling of IPGNs. In the x-axis we can see the upper bounds of the ambiguity set, defined in Table 1, multiplied by factors from 2 to 6. In the last point, the upper bounds $\boldsymbol{\pi}^{max}$ are simply a vector of ones. In the y-axis, we can see the combined performance of the IPGNs. The red curve denotes the worst expected combined performance, while the blue curve denotes the associated worst-case combined performance.

1, we can notice that worst expected performance and worst-case performance are equivalent. This is because this case corresponds to the situation where the probability mass is entirely allocated to the worst-case scenario. In other words, the worst-case scenario has a conditional occurrence probability equal to 1, while the other feasible scenarios have a conditional occurrence probability equal to 0. This case also corresponds to the solution that can be identified with a classic *defender-attacker-defender* framework, as it shown for example in the model in (56):

$$\max_{\substack{\mathbf{h}^0,\boldsymbol{\delta}^0 \\ \mathbf{y}\in\{0,1\}}} \min_{\mathbf{u}\in\{0,1\}} \max_{\mathbf{h},\boldsymbol{\delta}} \mathbf{b}^T\mathbf{h} \tag{56}$$

subject to (21), (23)-(27) and (32) (Bellè et al., 2021).

5.2. The effect of the available monetary budget B_c

The effect of different monetary budget B_c is evaluated for upper bounds $\boldsymbol{\pi}^{max}$ equal to the ones in Table 1 and maximum number of failed edges K=3. The results are shown in Figure 4.

Firstly, we notice the presence of an unfeasibility region, highlighted with a grey shadowed area, for budgets lower than 822.76$. As it was also mentioned in the previous section, this value is the minimum budget B_c that guarantees to respect Constraints (23)-(27), and it corresponds to the cost of the euclidean coupling interface. In other words, it is the "minimum cost" coupling interface, where each node is coupled with the geographically-closest node of the other infrastructure.

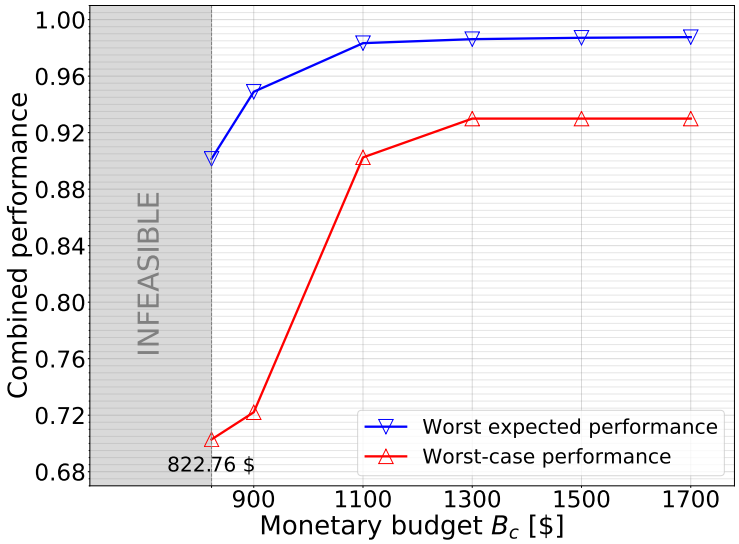

Figure 4: Results of the distributionally robust optimal coupling of IPGNs for different monetary budget values.

Secondly, we notice that, as it was expected, the results in terms of worst expected performance and associated worst-case performance improve as the monetary budget increases, because of the larger choice of coupling interface designs that is possible to implement. The improvement of results is particularly pronounced for lower budgets. For example, passing from 822.76$ to 900$, which represents a budget increase of 9.39%, leads to improvements of worst expected performance from 0.901 to 0.949, and worst-case performance from 0.703 to 0.722. These values corresponds to improvements of 5.26% and 2.74%, respectively. Increasing the budget from 900$ to 1100$, which represents an increase of 22.22%, leads to improvements of worst expected performance (from 0.949 to 0.983) and worst-case performance (from 0.722 to 0.902) of 3.63% and 24.98%, respectively. Further increasing the budget leads to only marginal improvements of performance.

5.3. The effect of the set of feasible failure scenarios

The size of the set of feasible failure scenarios is given by the parameter K, as it shown in Equation (28). This parameter, together with the upper bounds of the conditional marginal failure probability $\boldsymbol{\pi}^{max}$, reflects the decision-maker's attitude towards risk, as high K values denote a more conservative approach (and vice versa). We evaluate the effect of different K values, using the upper bounds $\boldsymbol{\pi}^{max}$ in Table 1 and a monetary budget B_c=1100$. The results are shown in Figure 5.

As it can be clearly seen, the expected performance do not decline considerably considerably as K increases. This is an indication of the validity of the DRO approach. However, it must be highlight that the results depends on the size of ambiguity set. If the ambiguity set was larger,

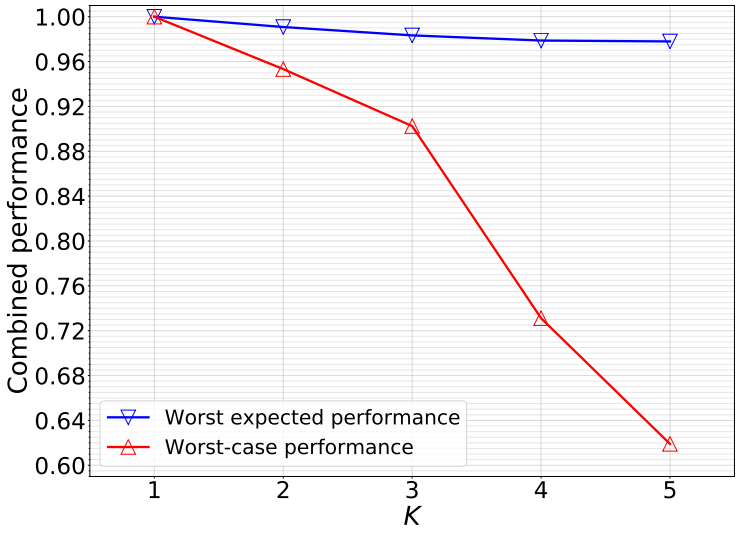

Figure 5: Results of the distributionally robust optimal coupling of IPGNs for different monetary budget values.

more probability mass could be allocated on the more disruptive scenarios which lead to lower performance, strongly impacting the results.

We can also notice that the associated worst-case performance, corresponding to the red curve in Figure 5, declines much faster than the expected performance. This is an indication that, with the current ambiguity set, it is not possible to allocate considerable probability mass on the most disruptive scenarios.

5.4. Computational performance

The computational performance for the cases presented in Sections 5.1, 5.2 and 5.3 are shown in Tables 2, 3 and 4, respectively.

Table 2: Computational cost in seconds and number of outer iterations of the NC&CG algorithm for different upper bounds $\boldsymbol{\pi^{max}}$.

Bounds	$\boldsymbol{\pi^{max}}$	$\boldsymbol{2\pi^{max}}$	$\boldsymbol{3\pi^{max}}$	$\boldsymbol{4\pi^{max}}$	$\boldsymbol{5\pi^{max}}$	$\boldsymbol{6\pi^{max}}$	$\boldsymbol{\pi^{max}}$**=1**
Time [s]	405.3	747.3	85.8	31.3	19.1	36.3	38.6
Outer iter.	29	36	18	13	10	12	8

For the baseline case presented in Section 5.1, the computational cost is highly affordable in every case.

For the sensitivity analysis of the monetary budget B_c presented in Section 5.2, the computational cost is generally higher, reaching a maximum of 2385.2 seconds in the case $B_{ci} = 900\$$.

Table 3: Computational cost in seconds and number of outer iterations of the NC&CG algorithm for different monetary budgets B_c.

Budget [$]	**822.76**	**900**	**1100**	**1300**	**1500**	**1700**
Time [s]	260.5	2385.2	405.3	435.9	342.9	203.5
Outer iter.	27	30	29	26	24	20

Table 4: Computational cost in seconds and number of outer iterations of the NC&CG algorithm for different maximum number of failed edges K.

K	**1**	**2**	**3**	**4**	**5**
Time [s]	6.1	260.4	405.3	3556.9	5276.6
Outer iter.	6	17	29	39	42

For the sensitivity analysis of the parameter K presented in Section 5.3, the computational cost increases as the parameter K increases. This is because more scenarios are included within the analysis, making the optimization more computationally-demanding.

For larger case studies, the computational time might increase considerably, due to the large number of binary variables involved. However, an increased computational cost does not represent an insurmountable problem, for the two following reasons: *i)* this model is tailored to be utilized during design phases, and long computational times do not pose particular problems, and *ii)* the complexity of the model can be reduced limiting the feasible number of coupling interfaces according to geographical and physical constraints. For example, one can assume that only nodes within a specific distance range can be coupled together; this would limit the number of $\mathbf{y}$ variables involved, thus reducing the complexity and computational cost of the model.

6. Conclusion

In this paper, we have proposed a novel distributionally robust approach for the optimal coupling of interdependent critical infrastructures, using an illustrative case study representing interdependent power and gas networks. The proposed approach clearly outperforms traditional euclidean coupling strategies based on nodes geographical vicinity, leading to better results in terms of expected performance and worst-case performance under random failures.

In summary, the contributions of this paper are:

- The development of a novel approach, based on distributionally robust optimization, for the design of coupling interfaces between interdependent CIs.

- The development of a reliable and efficient solution procedure based on the Nested Column&Constraint Generation algorithm.
- The application of the proposed framework on a case study to demonstrate the validity of the proposed approach.

The proposed case study is based on interdependent power and gas networks, modeled using a DC power flow model and a maximum flow model, respectively. However, the proposed approach can be applied to any combination of interdependent critical infrastructures, simply adapting the operational model in Section 2.2.

The conservativeness of the model can be tuned adjusting the size of the feasible attacks set, with variations of maximum number of lines attacked and failed, as well as tuning the parameters of the ambiguity set. The model can be easily adapted to different disruption scenarios, for example including failure of nodes.

Our solution strategy leads to acceptable computational times in this work. However, the computational cost might increase considerably for larger case-studies. Nevertheless, it does not represent a particular issue, since this approach aims at being used during design or retrofitting phases.

Appendix A. Proof of proposition 1

For a fixed coupling interface, the inner problem in (A.1) can be rewritten as (A.2)-(A.4) by substituting the expectation with an integral over the set of feasible failure scenarios $\mathcal{A}$ and introducing constraints (A.3) and (A.4).

$$\min_{\mathbb{P}\in\mathcal{M}} \mathbb{E}_{\mathbb{P}}\left[\mathbf{Q}(\mathbf{y},\boldsymbol{\xi})\right] \tag{A.1}$$

$$\min_{\mathbb{P}} \int_{\mathcal{A}} \mathbf{Q}(\mathbf{y},\boldsymbol{\xi})\mathrm{d}\mathbb{P} \tag{A.2}$$

subject to:

$$\int_{\mathcal{A}} \mathrm{d}\mathbb{P} = 1 \tag{A.3}$$

$$\int_{\mathcal{A}} (1-u_i)\mathrm{d}\mathbb{P} \leq \pi_i^{max}, \quad \forall i \in E_{PN}. \tag{A.4}$$

As it is easy to verify that for the problem in (A.2)-(A.4) Slater's conditions are satisfied, strong duality holds and the problem can be recast into its dual form, shown in (A.5)-(A.6), where α and β_i are the dual variables of constraints (A.3) and (A.4), respectively:

$$\max_{\alpha,\boldsymbol{\beta}\geq 0} \; -\alpha - \sum_{i\in E_{PN}} \beta_i \pi_i^{max} \tag{A.5}$$

subject to:

$$\alpha + \sum_{i\in E_{PN}} \beta_i(1-u_i) \geq -\mathbf{Q}(\mathbf{y},\boldsymbol{\xi}), \quad \forall \mathbf{u}\in\mathcal{A}. \tag{A.6}$$

Constraint in (A.6) can be rewritten as in (A.7):

$$-\alpha \leq \mathbf{Q}(\mathbf{y},\boldsymbol{\xi}) + \sum_{i\in E_{PN}} \beta_i(1-u_i), \quad \forall \mathbf{u}\in\mathcal{A}. \tag{A.7}$$

Noting that (A.5) is a maximization problem, it can be observed from Constraint (A.7) that the optimal value $\widehat{-\alpha}$ is given by (A.8):

$$\widehat{-\alpha} = \min_{\mathbf{u}\in\mathcal{A}} \; \mathbf{Q}(\mathbf{y},\boldsymbol{\xi}) + \sum_{i\in E_{PN}} \beta_i(1-u_i). \tag{A.8}$$

The reformulation of $\widehat{-\alpha}$ must be inserted in (A.5). Formulating explicitly $\mathbf{Q}(\mathbf{y},\boldsymbol{\xi})$ and merging the problem with the first stage maximization in (22) complete the proof.

References

Alkhaleel, B. A., Liao, H., & Sullivan, K. M. (2022). Risk and resilience-based optimal post-disruption restoration for critical infrastructures under uncertainty. *European Journal of Operational Research*, *296*(1), 174–202.

Alvarado, D., Moreira, A., Moreno, R., & Strbac, G. (2018). Transmission network investment with distributed energy resources and distributionally robust security. *IEEE Transactions on Power Systems*, *34*(6), 5157–5168.

Arrigo, A., Ordoudis, C., Kazempour, J., De Grève, Z., Toubeau, J.-F., & Vallée, F. (2022). Wasserstein distributionally robust chance-constrained optimization for energy and reserve dispatch: An exact and physically-bounded formulation. *European Journal of Operational Research*, *296*(1), 304–322.

Babaei, S., Jiang, R., & Zhao, C. (2020). Distributionally robust distribution network configuration under random contingency. *IEEE Transactions on Power Systems*, *35*(5), 3332–3341.

Bagheri, A., & Zhao, C. (2019). Distributionally robust reliability assessment for transmission system hardening plan under *nk* security criterion. *IEEE Transactions on Reliability*, *68*(2), 653–662.

Bellè, A., Abdin, A. F., Zeng, Z., Fang, Y.-P., & Barros, A. (2021). *A mathematical framework for the optimal coupling of interdependent critical infrastructures* [working paper or preprint]. https://hal.archives-ouvertes.fr/hal-03446712

Bhuiyan, T. H., Medal, H. R., & Harun, S. (2020). A stochastic programming model with endogenous and exogenous uncertainty for reliable network design under random disruption. *European Journal of Operational Research*, *285*(2), 670–694.

Bier, V., & Gutfraind, A. (2019). Risk analysis beyond vulnerability and resilience–characterizing the defensibility of critical systems. *European Journal of Operational Research*, *276*(2), 626–636.

Brown, G. G., Carlyle, W. M., Salmerón, J., & Wood, K. (2005). Analyzing the vulnerability of critical infrastructure to attack and planning defenses. *Emerging theory, methods, and applications* (pp. 102–123). Informs.

Buldyrev, S. V., Parshani, R., Paul, G., Stanley, H. E., & Havlin, S. (2010). Catastrophic cascade of failures in interdependent networks. *Nature*, *464*(7291), 1025–1028.

Cai, S., Xie, Y., Wu, Q., Zhang, M., Jin, X., & Xiang, Z. (2021). Distributionally robust microgrid formation approach for service restoration under random contingency. *IEEE Transactions on Smart Grid*.

Chattopadhyay, S., Dai, H., Hosseinalipour, S., et al. (2017). Designing optimal interlink patterns to maximize robustness of interdependent networks against cascading failures. *IEEE Transactions on Communications*, *65*(9), 3847–3862.

Chen, Z., Wu, J., Xia, Y., & Zhang, X. (2017). Robustness of interdependent power grids and communication networks: A complex network perspective. *IEEE Transactions on Circuits and Systems II: Express Briefs*, *65*(1), 115–119.

Corsi, S., & Sabelli, C. (2004). General blackout in italy sunday september 28, 2003, h. 03: 28: 00. *IEEE Power Engineering Society General Meeting, 2004.*, 1691–1702.

Doan, X. V. (2021). Distributionally robust optimization under endogenous uncertainty with an application in retrofitting planning. *European Journal of Operational Research*.

Fang, Y.-P., & Zio, E. (2019). An adaptive robust framework for the optimization of the resilience of interdependent infrastructures under natural hazards. *European Journal of Operational Research*, *276*(3), 1119–1136.

Fu, G., Dawson, R., Khoury, M., & Bullock, S. (2014). Interdependent networks: Vulnerability analysis and strategies to limit cascading failure. *The European Physical Journal B*, *87*(7), 1–10.

Guo, H., Yu, S. S., Iu, H. H., Fernando, T., & Zheng, C. (2019). A complex network theory analytical approach to power system cascading failure—from a cyber-physical perspective. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, *29*(5), 053111.

Guo, Z., & Haimes, Y. Y. (2016). Exploring systemic risks in systems-of-systems within a multiobjective decision framework. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, *47*(6), 906–915.

Gurobi Optimization, L. (2021). Gurobi optimizer reference manual. http://www.gurobi.com

Lee II, E. E., Mitchell, J. E., & Wallace, W. A. (2007). Restoration of services in interdependent infrastructure systems: A network flows approach. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, *37*(6), 1303–1317.

Li, R., Wang, M., Yang, M., Han, X., Wu, Q., & Wang, W. (2022). A distributionally robust model for reserve optimization considering contingency probability uncertainty. *International Journal of Electrical Power & Energy Systems*, *134*, 107174.

Liu, Y., Xu, H., Yang, S.-J. S., & Zhang, J. (2018). Distributionally robust equilibrium for continuous games: Nash and stackelberg models. *European Journal of Operational Research*, *265*(2), 631–643.

Nurre, S. G., Cavdaroglu, B., Mitchell, J. E., Sharkey, T. C., & Wallace, W. A. (2012). Restoring infrastructure systems: An integrated network design and scheduling (inds) problem. *European journal of operational research*, *223*(3), 794–806.

Ouyang, M. (2014). Review on modeling and simulation of interdependent critical infrastructure systems. *Reliability engineering & System safety*, *121*, 43–60.

Ouyang, M. (2017). A mathematical framework to optimize resilience of interdependent critical infrastructure systems under spatially localized attacks. *European Journal of Operational Research*, *262*(3), 1072–1084.

Ouyang, M., & Dueñas-Osorio, L. (2011). An approach to design interface topologies across interdependent urban infrastructure systems. *Reliability Engineering & System Safety*, *96*(11), 1462–1473.

Ouyang, M., & Fang, Y. (2017). A mathematical framework to optimize critical infrastructure resilience against intentional attacks. *Computer-Aided Civil and Infrastructure Engineering*, *32*(11), 909–929.

Parshani, R., Buldyrev, S. V., & Havlin, S. (2010). Interdependent networks: Reducing the coupling strength leads to a change from a first to second order percolation transition. *Physical review letters*, *105*(4), 048701.

Rahimian, H., & Mehrotra, S. (2019). Distributionally robust optimization: A review. *arXiv preprint arXiv:1908.05659*.

Rinaldi, S. M., Peerenboom, J. P., & Kelly, T. K. (2001). Identifying, understanding, and analyzing critical infrastructure interdependencies. *IEEE control systems magazine*, *21*(6), 11–25.

Rueda, D. F., & Calle, E. (2017). Using interdependency matrices to mitigate targeted attacks on interdependent networks: A case study involving a power grid and backbone telecommunications networks. *International Journal of Critical Infrastructure Protection*, *16*, 3–12.

Saif, A., & Delage, E. (2021). Data-driven distributionally robust capacitated facility location problem. *European Journal of Operational Research*, *291*(3), 995–1007.

Shehadeh, K. S., & Padman, R. (2021). A distributionally robust optimization approach for stochastic elective surgery scheduling with limited intensive care unit capacity. *European Journal of Operational Research*, *290*(3), 901–913.

Vespignani, A. (2010). The fragility of interdependency. *Nature*, *464*(7291), 984–985.

Wang, S., Hong, L., & Chen, X. (2012). Vulnerability analysis of interdependent infrastructure systems: A methodological framework. *Physica A: Statistical Mechanics and its applications*, *391*(11), 3323–3335.

Wang, X., Zhou, W., Li, R., Cao, J., & Lin, X. (2018). Improving robustness of interdependent networks by a new coupling strategy. *Physica A: Statistical Mechanics and its Applications*, *492*, 1075–1080.

Wang, Z., You, K., Song, S., & Zhang, Y. (2020). Wasserstein distributionally robust shortest path problem. *European Journal of Operational Research*, *284*(1), 31–43.

Winkler, J., Dueñas-Osorio, L., Stein, R., & Subramanian, D. (2011). Interface network models for complex urban infrastructure systems. *Journal of Infrastructure Systems*, *17*(4), 138–150.

Yagan, O., Qian, D., Zhang, J., & Cochran, D. (2012). Optimal allocation of interconnecting links in cyber-physical systems: Interdependence, cascading failures, and robustness. *IEEE Transactions on Parallel and Distributed Systems*, *23*(9), 1708–1720.

Zhang, G., Zhang, F., Zhang, X., Wu, Q., & Meng, K. (2020). A multi-disaster-scenario distributionally robust planning model for enhancing the resilience of distribution systems. *International Journal of Electrical Power & Energy Systems*, *122*, 106161.

Zhao, C., & Jiang, R. (2017). Distributionally robust contingency-constrained unit commitment. *IEEE Transactions on Power Systems*, *33*(1), 94–102.

Paper V

A. Bellè, Z. Zeng, M. Sango, and A. Barros, "Resilience enhancement by optimal allocation of redundant interdependency links in interdependent critical infrastructures," *Reliability Engineering & System Safety*, to be submitted, 2022.

Resilience enhancement by optimal allocation of redundant interdependency links in interdependent critical infrastructures

Andrea Bellè[1]*, Zhiguo Zeng[1], Marc Sango[2], Anne Barros[1]

[1] *Chair on Risk and Resilience of Complex Systems, Laboratoire Génie Industriel, CentraleSupélec, Université Paris-Saclay, 3 Rue Joliot Curie, 91190 Gif-sur-Yvette, France.*
Emails: andrea.belle@centralesupelec.fr, zhiguo.zeng@centralesupelec.fr, anne.barros@centralesupelec.fr

[2] *SNCF, Direction Technologies Innovation et Projet Groupe, Innovation & Recherche, Département Physique du Système Ferroviaire, Equipe Sécurité Système, France*
Email: marc.sango@sncf.fr

Abstract

Critical infrastructures (CIs) are interconnected, and they are often mutually interdependent on each other in terms of functionality and performance. CIs are often modeled using network-based approaches, and the way they are coupled is defined by a set of interdependency links, globally referred to as coupling interface. In the existing literature, each node in one CI is usually assumed to be dependent on maximum one node of another CI. However, in reality, each node in one CI can be dependent on multiple nodes of another CI through redundant interdependency links. In this paper, we explore the potential of optimal allocation of redundant interdependency links in terms of resilience enhancement of interdependent CIs. We use interdependent power and gas networks (IPGNs) as illustrative example, and we show how the optimal allocation of redundant interdependency links can considerably increase their resilience in terms of worst-case combined performance.

Keywords: Coupling interface, resilience, defender-attacker-defender, interdependent critical infrastructures, optimization, redundancy allocation

1. Introduction

Critical infrastructures (CIs), such as energy systems, transportation networks, and telecommunications systems, are often tightly coupled, and they are mutually interdependent with complex relationships of interdependencies [1]. CIs are often modeled with a network-based approach

*Corresponding author
Email address: andrea.belle@centralesupelec.fr (Andrea Bellè[1])

[2], where infrastructures are modeled as networks, connected by links that represent the different interdependencies. A qualitative representation of two interdependent CIs, modeled with a network-based approach, is shown in Figure 1. Components of CIs are represented as nodes, while their connections, physical and nonphysical, are represented as edges. Components belonging to different can be dependent on each other, and be coupled by interdependency links. The ensemble of interdependency links, which defines how two CIs are coupled, is often referred to as *coupling interface* [3]–[5].

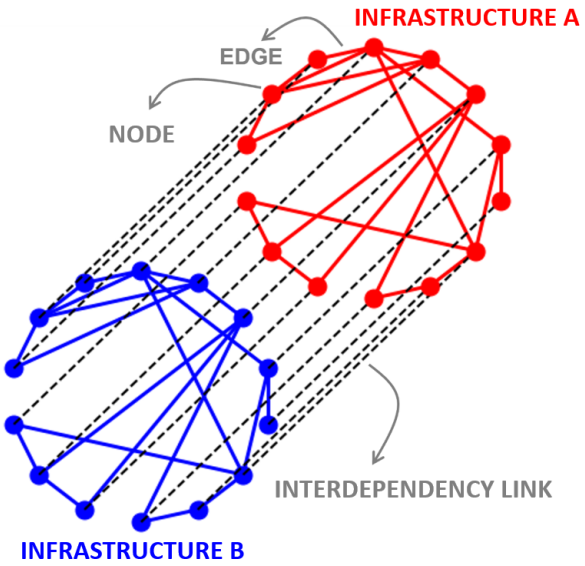

Figure 1: Network representation of two interdependent infrastructures. In evidence, we can notice the two structural components of networks (nodes and edges) and their connection (interdependency links). The ensemble of interdependency links is referred to as *coupling interface*. Figure and caption from [3].

The design and the topology of coupling interfaces heavily impacts the resilience of interdependent CIs [3]–[5], and optimizing the topology of coupling interfaces is a fundamental task. In fact, failures and disruption within one CI can propagate to other coupled CIs through the coupling interfaces. In the context of interdependent CIs, the concept of coupling interface optimization was introduced in [4] and [5]. In [4], the authors propose an heuristic approach, based on network metrics such as node distance, degree, and betweenness, to couple interdependent CIs in order to reduce failure propagation. A similar approach, which also accounts for the physical properties of the infrastructures, is proposed in [5]. In both these works, the coupling interface between CIs is acknowledged as one of the key factors for the resilience of CIs. However, as these strategies are based on network metrics, they represent an heuristic approach, which is hardly generalizable and does not ensure the optimality of solutions [3].

In our previous work, we proposed a mathematical programming approach, based on the

defender-attacker-defender (DAD) model, to enhance the resilience of interdependent CIs by optimizing the design of their coupling interface [3]. In this work, we maintained the "single-dependency" assumption[1], which states that each node can be dependent on one, and only node of the other infrastructure. This assumption is often used in the existing literature [6]–[11]. However, in reality, a node can be dependent on multiple nodes of another infrastructure, as it is taken into account in [4] and [5].

In this paper, we propose a mathematical programming approach, based on the DAD model, to allocate redundant interdependency links in interdependent CIs in order to enhance their resilience. Similarly to [3], we focus on enhancing the resilience of interdependent CIs in terms of combined survivability against random failures. The concept of survivability can be defined as *"the capability of a system to fulfill its mission in a timely manner in the presence of attacks, failures, or accidents"* [12], and it can be interpreted as the residual performance of interdependent CIs after a failure scenario [3]. Survivability is complementary to the concept of vulnerability, defined as the *"degree of loss or damage to a system when exposed to a strain of a given type and magnitude"* [13]. In other words, we seek to allocate the redundancies in order to maximize the residual performance of interdependent CIs after failures and disruptive events.

The motivations of this work are the following:

- redundancy allocation is a traditional reliability optimization problem [14]. From a research perspective, it is interesting to extend this concept to interdependent CIs and their coupling interfaces. Moreover, redundancies have already been identified to be an effective measure for enhancing the resilience of CIs [15], and investigating their potential, in terms of resilience enhancement, within interdependent CIs is of the utmost importance;

- allocating redundancies in order to retrofit existing coupling interfaces can be an effective measure to enhance the resilience of interdependent CIs [5]: planners and decision-makers should, thus, be provided with the necessary modeling tools to optimize the redundancy allocation.

This paper is presented as an extension of the work published in [3]. For more details on the relevant existing work on coupling interfaces and and their connection to resilience, the reader is

[1]The "single-dependency" assumption limits each node of one CI to be dependent on maximum another node belonging to a different CI. In other words, each node can be the receiving end of maximum one interdependency link. However, it is still possible for each node to supply multiple components of a different system. For example, assuming the case of power and gas networks, the "single-dependency" assumption defines that each gas-fired power plant is dependent on one, and only one gas node of the gas network; however, each gas node can supply electricity to multiple gas-fired power plants.

referred to Section 1 of [3].

2. Problem formulation

2.1. Modeling framework

We consider the case of interdependent power and networks (IPGNs). Particularly, we consider a power network PN, containing gas-fired power plants supplied by a gas network GN, which requires electricity from the power network in order to maintain the functionality of valves, compressors, and pumps in each node.

Each network is described by a graph $G = (V, E)$, with V the set of N nodes and E the set of M edges. Each edge k is directed, and it is defined by an origin node $O(k)$ and a destination node $D(k)$.

The power network is defined by the graph $G_{PN} = (V_{PN}, E_{PN})$, where V_{PN} is the set of N_{PN} nodes and E_{PN} is the set of M_{PN} edges. Each node i represents an electrical bus, with power production capacity $\overline{p}_i$ and requested power demand $\overline{d}_i$. Each edge k represents a power transmission line, characterized by flow capacity $\overline{f}_k$. The operation of the power network are modeled with a traditional DC power flow model.

The gas network is defined by the graph $G_{GN} = (V_{GN}, E_{GN})$, where V_{GN} is the set of N_{GN} nodes and E_{GN} is the set of M_{GN} edges. Each node i represents a gas hub, with gas production capacity $\overline{p}_i$ and requested gas demand $\overline{d}_i$. Each edge k represents a gas pipeline, characterized by flow capacity $\overline{f}_k$. The operations of the gas network are modeled with a linear maximum flow model.

The focus of our analysis is the combined performance P_C of the IPGNs, defined as in (1):

$$P_C = \frac{w_{PN}}{\overline{d}_{PN}} \sum_{i \in V_{PN}} d_i + \frac{w_{GN}}{\overline{d}_{GN}} \sum_{i \in V_{GN}} d_i \tag{1}$$

where w_{PN} and w_{GN} represent the weight of power network and gas network when computing the combined performance. The combined performance represent the fraction of requested demand of power and gas which is possible to supply. The terms $\overline{d}_{PN}$ and $\overline{d}_{GN}$ represent the total requested demand of power and gas, respectively, and they are computed as in (2) and (3):

$$\overline{d}_{PN} = \sum_{i \in V_{PN}} \overline{d}_i^b + \sum_{j \in V_{GN}} \overline{d}_j^{MW} \tag{2}$$

$$\overline{d}_{GN} = \sum_{i \in V_{GN}} \overline{d}_i^b + \sum_{j \in V_{PN}} \overline{d}_j^{m^3}. \tag{3}$$

The term $\overline{d}_i^b$ denotes the baseline requested demand (households, industries, general consumers, etc.) of power or gas in each node i. The term $\overline{d}_j^{MW}$ denotes the requested power demand of

node $j \in V_{GN}$. The term $\overline{d}_j^{m^3}$ denotes the requested gas demand of node $j \in V_{PN}$. The term d_i represents the supplied demand of power or gas in each node i.

2.2. Operational model of interdependent railway and power networks

The power network operations are modeled using a DC power flow model, while the gas network is modeled with a linear maximum flow model. For the operational model of the IPGNs, the following assumptions are considered:

- nodes in the power network with power production capacity $\overline{p}_i > 0$ contain a gas-fired power plant, and they can be supplied (and, thus, be dependent on) multiple nodes in the gas network. However, it is assumed that each node in the power network is supplied by only one node in the gas network at a time, i.e. only one interdependency link is active and the other ones are in stand-by;
- nodes in the gas network needs to receive electricity for the proper functionality of various equipment, and they can be supplied (and, thus, be dependent on) multiple nodes in the power network. However, it is assumed that each node in the gas network is supplied by only one node in the power network at a time, i.e. only one interdependency link is active and the other ones are in stand-by.

The operational model of IPGNs is described by the optimization in Equation (4), subject to Constraints (5)-(21):

$$\max_{\substack{\mathbf{p},\mathbf{d},\mathbf{f},\boldsymbol{\theta} \\ \boldsymbol{\delta}^p \in \{0,1\}^{N_{PN}} \\ \boldsymbol{\delta}^g \in \{0,1\}^{N_{GN}} \\ \boldsymbol{\mu}^{g \leftarrow p} \in \{0,1\}^{N_c} \\ \boldsymbol{\mu}^{p \leftarrow g} \in \{0,1\}^{N_c}}} \frac{w_{PN}}{\overline{d}_{PN}} \sum_{i \in V_{PN}} d_i + \frac{w_{GN}}{\overline{d}_{GN}} \sum_{i \in V_{GN}} d_i \tag{4}$$

subject to:

$$0 \leq p_i \leq \overline{p}_i, \quad \forall i \in V_{PN} \cup V_{GN} \tag{5}$$

$$0 \leq d_i \leq \overline{d}_i^b + \sum_{j \in V_{GN}} \mu_{ji}^{g \leftarrow p} \overline{d}_j^{MW}, \quad \forall i \in V_{PN} \tag{6}$$

$$0 \leq d_i \leq \overline{d}_i^b + \sum_{j \in V_{PN}} \mu_{ji}^{p \leftarrow g} \overline{d}_j^{m^3}, \quad \forall i \in V_{GN} \tag{7}$$

$$-u_k \overline{f}_k \leq f_k \leq u_k \overline{f}_k, \quad \forall k \in E_{PN} \cup E_{GN} \tag{8}$$

$$u_k \left(x_k f_k - (\theta_{O(k)} - \theta_{D(k)}) \right) = 0, \quad \forall k \in E_{PN} \tag{9}$$

$$p_i - d_i + \sum_{k|D(k)=i} f_k - \sum_{k|O(k)=i} f_k = 0, \quad \forall i \in V_{PN} \cup V_{GN} \tag{10}$$

$$d_i - \delta_i^p \left(\overline{d_i^b} + \sum_{j \in V_{GN}} \mu_{ji}^{g \leftarrow p} \overline{d}_j^{MW} \right) \geq 0, \quad \forall i \in V_{PN} \tag{11}$$

$$d_i - \delta_i^g \left(\overline{d_i^b} + \sum_{j \in V_{PN}} \mu_{ji}^{p \leftarrow g} \overline{d}_j^{m^3} \right) \geq 0, \quad \forall i \in V_{GN} \tag{12}$$

$$p_i - \overline{p}_i \sum_{j \in V_{GN}} \mu_{ij}^{p \leftarrow g} \delta_j^g \leq 0, \quad \forall i \in V_{PN} \tag{13}$$

$$p_i - \overline{p}_i \sum_{j \in V_{PN}} \mu_{ij}^{g \leftarrow p} \delta_j^p \leq 0, \quad \forall i \in V_{GN} \tag{14}$$

$$d_i - \left(\overline{d}_i^b + \sum_{j \in V_{PN}} \mu_{ji}^{p \leftarrow g} \overline{d}_j^{m^3} \right) \sum_{j \in V_{PN}} \mu_{ij}^{g \leftarrow p} \delta_j^p \leq 0, \quad \forall i \in V_{GN} \tag{15}$$

$$- \sum_{\substack{k|O(k)=i \\ j \in V_{PN}}} \mu_{ij}^{g \leftarrow p} \delta_i^p \overline{f}_k \leq f_k \leq \sum_{\substack{k|O(k)=i \\ j \in V_{PN}}} \mu_{ij}^{g \leftarrow p} \delta_i^p \overline{f}_k, \quad \forall k \in E_{GN} \tag{16}$$

$$- \sum_{\substack{k|D(k)=i \\ j \in V_{PN}}} \mu_{ij}^{g \leftarrow p} \delta_i^p \overline{f}_k \leq f_k \leq \sum_{\substack{k|D(k)=i \\ j \in V_{PN}}} \mu_{ij}^{g \leftarrow p} \delta_i^p \overline{f}_k, \quad \forall k \in E_{GN} \tag{17}$$

$$\mu_{ij}^{g \leftarrow p} \leq y_{ij}^{g \leftarrow p}, \quad \forall i \in V_{GN}, \forall j \in V_{PN} \tag{18}$$

$$\mu_{ij}^{p \leftarrow g} \leq y_{ij}^{p \leftarrow g}, \quad \forall i \in V_{PN}, \forall j \in V_{GN} \tag{19}$$

$$\sum_{j \in V_{PN}} \mu_{ij}^{g \leftarrow p} = 1, \quad \forall i \in V_{GN} \tag{20}$$

$$\sum_{j \in V_{GN}} \mu_{ij}^{p \leftarrow g} = 1, \quad \forall i \in V_{PN} \tag{21}$$

The vectors $\mathbf{p}$, $\mathbf{d}$, $\mathbf{f}$, and $\boldsymbol{\theta}$ contain the continuous variables of the problem. The variables p_i define the production of power or gas in each node, d_i defines the supplied demand of power or gas in each node, the variables f_k define the flow of power or gas in each edge, and the variablesθ_i define the phase angle in each node of the power network. The vectors $\boldsymbol{\delta}^p$, $\boldsymbol{\delta}^g$, $\boldsymbol{\mu}^{g \leftarrow p}$, and $\boldsymbol{\mu}^{p \leftarrow g}$ contain

the binary variable of the optimization problem. The dimension N_c is equal to $N_{PN} \times N_{GN}$. The binary variables δ_i^p define the functional state (1 if functional, 0 otherwise) of the interdependency links starting from the node $i \in V_{PN}$ and going to the gas network. Similarly, the binary variables δ_i^g define the functional state of the interdependency links starting from the node $i \in V_{GN}$ and going to the power network. The variable $\mu_{ij}^{g \leftarrow p}$ is equal to 1 when node $i \in V_{GN}$ is supplied by node $j \in V_{PN}$, and 0 otherwise. Similarly, the variable $\mu_{ij}^{p \leftarrow g}$ is equal to 1 when node $i \in V_{PN}$ is supplied by node $j \in V_{GN}$, and 0 otherwise.

The objective function in (4) represents the combined performance of the IPGNs in terms of fraction of requested demand of power and gas which is supplied. In each node of the IPGNs, the production of power or gas p_i is limited by a production capacity $\overline{p}_i$, as shown in Constraint (5). Similarly the supplied demand of power or gas d_i is limited by the requested demand, as shown in Constraints (6) and (7). The total requested demand of power in each node $i \in V_{PN}$ is composed of a baseline requested demand $\overline{d}_i^b$, and the sum of the power demands $\overline{d}_j^{MW}$ of all nodes $j \in V_{GN}$ which are supplied by node $i \in V_{PN}$ ($\mu_{ji}^{g \leftarrow p}$=1). Similarly, the total requested demand of gas in each node $i \in V_{GN}$ is composed of a baseline requested demand $\overline{d}_i^b$, and the sum of the gas demands $\overline{d}_j^{m^3}$ of all nodes $j \in V_{PN}$ which are supplied by node $i \in V_{GN}$ ($\mu_{ji}^{p \leftarrow g}$=1). The flow of power in each line $k \in V_{PN}$ and the flow of gas in each pipeline $k \in V_{GN}$ are limited by the flow capacity $\overline{f}_k$ if the line/pipeline k is functional (u_k=1), and forced to be 0 if the line/pipeline is failed (u_k=0), as shown in Constraint (8). The DC power flow assumption for the power network is enforced in Constraint (9). The net nodal balance of power and gas in each node is enforced by Constraint (10).

Each interdependency links from the node $i \in V_{PN}$ to the gas network is functional (δ_i^p=1) only if the requested power demand in node i is fully supplied (Constraint (11)). Similarly, each interdependency links from the node $i \in V_{GN}$ to the power network is functional (δ_i^g=1) only if the requested gas demand in node i is fully supplied (Constraint (12)). The production of power in each node $i \in V_{PN}$ is possible only if it supplied by a node $j \in V_{GN}$ ($\mu_{ij}^{p \leftarrow g}$=1) with a functional interdependency link (δ_j^g=1), as shown in Constraint (13). Similarly, as shown in Constraints (14) and (15), the production and the supply of gas in each node $i \in V_{GN}$ is possible only if it supplied by a node $j \in V_{PN}$ ($\mu_{ij}^{g \leftarrow p}$=1) with a functional interdependency link (δ_j^p=1). Moreover, as shown in Constraints (16) and (17), it is possible to flow gas in each pipeline $k \in E_{GN}$ only if both the origin and destination nodes of k are supplied by a node $j \in V_{PN}$ ($\mu_{ij}^{g \leftarrow p}$=1) with a functional interdependency link (δ_j^p=1).

Each node $i \in V_{GN}$ can be supplied by node $j \in V_{PN}$ ($\mu_{ij}^{g \leftarrow p}$=1) only if an interdependency link from j to i is present ($y_{ij}^{g \leftarrow p}$=1), as shown in Constraint (18). Each node $i \in V_{PN}$ can be supplied by node $j \in V_{GN}$ ($\mu_{ij}^{p \leftarrow g}$=1) only if an interdependency link from j to i is present ($y_{ij}^{p \leftarrow g}$=1), as

shown in Constraint (19). We assume that each node is supplied by only one node of the other infrastructure at a time, as shown in Constraints (20) and (21). In other words, if a node i presents multiple interdependency links from the other infrastructure, only one of these links is considered to be active, while the other ones are assumed to be in a stand-by state.

The problem in (4)-(21) is dependent on the coupling interface topology, defined by the binary variables $y_{ij}^{g \leftarrow p}$ and $y_{ij}^{p \leftarrow g}$, contained within the vectors $\mathbf{y}^{g \leftarrow p}$ and $\mathbf{y}^{p \leftarrow g}$, and the functional states of power lines and gas pipelines, defined by the binary variables u_k, contained within the the vector $\mathbf{u}$. The problem in (4)-(21) can be represented as a recourse function $\mathcal{Q}(\mathbf{y}^{g \leftarrow p}, \mathbf{y}^{p \leftarrow g}, \mathbf{u})$.

For the sake of simplicity, the problem in (4)-(21) can be represented with the compact matrix formulation in (22)-(23):

$$\max_{\mathbf{h}, \boldsymbol{\delta}, \boldsymbol{\mu}} \mathbf{b}^T \mathbf{h} \tag{22}$$

subject to:

$$\mathbf{Rh} \leq \mathbf{q} - \mathbf{Tu} - \mathbf{Hy} - \mathbf{W}\boldsymbol{\delta} - \mathbf{X}\boldsymbol{\mu} - \boldsymbol{\mu}^T \mathbf{D}\boldsymbol{\delta}. \tag{23}$$

where the vector $\mathbf{h}$ contains the continuous variables, the vector $\boldsymbol{\delta}$ contains the binary variables $\boldsymbol{\delta}^p$ and $\boldsymbol{\delta}^g$, the vector $\boldsymbol{\mu}$ contains the binary variables $\boldsymbol{\mu}^{g \leftarrow p}$ and $\boldsymbol{\mu}^{p \leftarrow g}$, the vectors $\mathbf{b}$ contains the objective function coefficients, the matrices $\mathbf{R}$, $\mathbf{T}$, $\mathbf{H}$, $\mathbf{W}$, $\mathbf{X}$ and $\mathbf{D}$ contain the constraints coefficients, and the vector $\mathbf{q}$ contains the constraints parameters.

2.3. Defender-attacker-defender model

The problem of allocating the redundant interdependency links is modeled as a DAD optimization problem. The inner defender corresponds to the maximization presented in the previous section, which represents the operational model of the IPGNs. The middle attacker aims at minimizing the combined performance of the IPGNs by targeting and failing some components, and it is a representation of various causes of disruption (terrorists, hackers, extreme weather events, etc.). The outer defender is interpreted as a centralized planner who allocates the redundant interdependency links, aiming at maximizing the combined performance of the IPGNs under the worst-case failure scenario. Within this work, the following assumptions are considered:

- a coupling interface is already existing, and the outer defender seeks to allocate redundant interdependency links between the power network and the gas network;
- allocating redundant interdependency links has a cost which depends on the distance between the two nodes connected by the interdependency link;
- the middle attacker can target and destroy a maximum number of components, and only power lines in the power network can be targeted.

The trilevel DAD formulation takes the form shown in (24):

$$\max_{\substack{\mathbf{y}^{g\leftarrow p}\in\{0,1\}^{N_c}\\ \mathbf{y}^{p\leftarrow g}\in\{0,1\}^{N_c}}} \min_{\mathbf{u}\in\{0,1\}^{M_{PN}}} \mathcal{Q}(\mathbf{y}^{g\leftarrow p}, \mathbf{y}^{p\leftarrow g}, \mathbf{u}) \tag{24}$$

subject to (5)-(21) and (25)-(28):

$$y_{ij}^{g\leftarrow p} \geq \overline{y}_{ij}^{g\leftarrow p}, \quad \forall i \in V_{GN}, \forall j \in V_{PN} \tag{25}$$

$$y_{ij}^{p\leftarrow g} \geq \overline{y}_{ij}^{p\leftarrow g}, \quad \forall i \in V_{PN}, \forall j \in V_{GN} \tag{26}$$

$$\sum_{\substack{i\in V_{GN}\\ j\in V_{PN}}} y_{ij}^{g\leftarrow p}\, d_{ij}^{km}\, c_{km}^{g\leftarrow p} + \sum_{\substack{i\in V_{PN}\\ j\in V_{GN}}} y_{ij}^{p\leftarrow g}\, d_{ji}^{km}\, c_{km}^{p\leftarrow g} \leq \overline{c}_{ci} + B_{ci} \tag{27}$$

$$\sum_{k\in E_{PN}} 1 - u_k \leq K_{att} \tag{28}$$

Constraints (25) and (26) ensure that any previously-existing coupling interface is considered within the optimization. The binary parameter $\overline{y}_{ij}^{g\leftarrow p}$ defines if an interdependency link from node $j \in V_{PN}$ to node $i \in V_{GN}$ is already existing ($\overline{y}_{ij}^{g\leftarrow p}$=1). Similarly, the binary parameter $\overline{y}_{ij}^{p\leftarrow g}$ defines if an interdependency link from node $j \in V_{GN}$ to node $i \in V_{PN}$ is already existing ($\overline{y}_{ij}^{p\leftarrow g}$=1). The corresponding binary variables, $y_{ij}^{g\leftarrow p}$ and $y_{ij}^{p\leftarrow g}$, are forced to be equal to 1 if interdependency links are present.

Constraint (27) enforces the cost of the allocated redundant interdependency link to be less or equal the available monetary budget B_{ci}. The parameter d_{ij}^{km} defines the distance in kilometer between nodes $i \in V_{GN}$ and $j \in V_{PN}$. The parameters $c_{km}^{g\leftarrow p}$ and $c_{km}^{p\leftarrow g}$ are the cost-per-kilometer of allocating an interdependency link. The parameter $\overline{c}_{ci}$ represents the cost of the existing coupling interface, computed as in (29):

$$\overline{c}_{ci} = \sum_{\substack{i\in V_{GN}\\ j\in V_{PN}}} \overline{y}_{ij}^{g\leftarrow p}\, d_{ij}^{km}\, c_{km}^{g\leftarrow p} + \sum_{\substack{i\in V_{PN}\\ j\in V_{GN}}} \overline{y}_{ij}^{p\leftarrow g}\, d_{ji}^{km}\, c_{km}^{p\leftarrow g} \tag{29}$$

With the term $\overline{c}_{ci}$ included, only the newly allocated interdependency links are considered within the budget limitation.

Constraint (28) enforces the number of power lines targeted and failed by the attacker to be lower or equal the parameter K_{att}. Each binary variable u_k=0 if line k is targeted and failed, and u_k=1 otherwise.

The solution of the DAD model leads to the identification of the most robust allocation of redundant interdependency links. In other words, the outer defender allocates the redundancies in order to maximize the combined performance of the IPGNs in the worst-case attack scenario that can be carried out by the attacker.

3. Solution strategy

Similarly to our previous work [3], the DAD optimization problem can be solved efficiently by implementing a *Nested Column&Constraint Generation* (NC&CG) algorithm [16], [17]. For simplicity, in the following explanation, we rely on the compact matrix formulation of the operational model in (22)-(23). If we fix the binary variables in (22)-(23), the optimization problem is a pure LP problem, and its dual formulation can be expressed as in (30)-(31)

$$\max_{\boldsymbol{\lambda}\geq 0} \left(\mathbf{q} - \mathbf{T}\mathbf{u}^* - \mathbf{H}\mathbf{y}^* - \mathbf{W}\boldsymbol{\delta}^* - \mathbf{X}\boldsymbol{\mu}^* - \boldsymbol{\mu}^{*T}\mathbf{D}\boldsymbol{\delta}^*\right)^T \boldsymbol{\lambda} \tag{30}$$

subject to:

$$\mathbf{R}^T\boldsymbol{\lambda} = \mathbf{b} \tag{31}$$

where the vector $\boldsymbol{\lambda}$ contains the dual variables.

3.1. Inner layer

With a fixed coupling interface $\mathbf{y}^*$, the middle inner-level problem (min-max) allows to identify the worst-case combined performance and the related optimal attack plan. This problem can be solved by following these steps:

1. Set $j = 0$, lower bound $LB_{in} = 0$, upper bound $UB_{in} = \infty$, and $\mathcal{D}_{part} = \emptyset$
2. Solve the inner master problem in Equations (32)-(35). Obtain an optimal solution $\widehat{\rho}^{(j)}$ and optimal attack plan $\widehat{\mathbf{u}}^{(j)}$. Update LB_{in}=$\widehat{\rho}^{(j)}$.

$$\min_{\rho,\mathbf{u},\boldsymbol{\lambda}} \rho \tag{32}$$

subject to:

$$\rho \geq \left(\mathbf{q} - \mathbf{T}\mathbf{u} - \mathbf{H}\mathbf{y}^* - \mathbf{W}\boldsymbol{\delta}^{*(j)} - \mathbf{X}\boldsymbol{\mu}^{*(j)} - \boldsymbol{\mu}^{*(j)T}\mathbf{D}\boldsymbol{\delta}^{*(j)}\right)^T \boldsymbol{\lambda}^{(j)}, \quad \forall \boldsymbol{\delta}^{*(j)}, \boldsymbol{\mu}^{*(j)} \in \mathcal{D}_{part} \tag{33}$$

$$\mathbf{R}^T\boldsymbol{\lambda}^{(j)} = \mathbf{b}, \quad \forall \boldsymbol{\delta}^{*(j)}, \boldsymbol{\mu}^{*(j)} \in \mathcal{D}_{part} \tag{34}$$

$$\sum_{k\in E_{PN}} (1 - u_k) \leq K_{att} \tag{35}$$

3. Solve the inner subproblem in Equations (36)-(37) with $\widehat{\mathbf{u}}^{(j)}$=$\mathbf{u}^*$. Obtain an optimal solution $\mathbf{b}^T\widehat{\mathbf{h}}^{(j)}$, $\widehat{\boldsymbol{\delta}}^{(j)}$, and $\widehat{\boldsymbol{\mu}}^{(j)}$. Set $UB_{in} = \min(UB_{in}, \mathbf{b}^T\widehat{\mathbf{h}}^{(j)})$.

$$\max_{\mathbf{h},\boldsymbol{\delta},\boldsymbol{\mu}} \mathbf{b}^T\mathbf{h} \tag{36}$$

subject to :

$$\mathbf{R}\mathbf{h} \leq \mathbf{q} - \mathbf{T}\mathbf{u}^* - \mathbf{H}\mathbf{y}^* - \mathbf{W}\boldsymbol{\delta} - \mathbf{X}\boldsymbol{\mu} - \boldsymbol{\mu}^T\mathbf{D}\boldsymbol{\delta} \tag{37}$$

4. If $(UB_{in} - LB_{in})/UB_{in} < 10^{-5}$, $\widehat{\mathbf{u}}^{(j)}$ represents the optimal attack and the algorithm can be terminated. Otherwise, $\mathcal{D}_{part} = \mathcal{D}_{part} \cup \widehat{\boldsymbol{\delta}}^{(j)} \cup \widehat{\boldsymbol{\mu}}^{(j)}$. Set $j \leftarrow j + 1$ and return to step 2.

3.2. Outer layer

The outer layer is solved by following these steps:

1. Set $j = 0$, lower bound $LB_{out} = 0$, upper bound $UB_{out} = \infty$, and $\mathcal{A}_{part} = \emptyset$
2. Solve the outer master problem in Equations (38)-(41). Obtain an optimal solution $\widehat{\eta}^{(j)}$, and optimal coupling interface $\widehat{\mathbf{y}}^{(j)}$. Update UB_{out}=min$(UB_{out}, \widehat{\eta}^{(j)})$

$$\max_{\eta, \mathbf{y}} \eta \tag{38}$$

$$\eta \leq \mathbf{b}^T \mathbf{h}^{(j)}, \quad \forall\, \mathbf{u}^{*(j)} \in \mathcal{A}_{part} \tag{39}$$

$$\mathbf{P}\mathbf{y} \leq \mathbf{g} \tag{40}$$

$$\mathbf{R}\mathbf{h}^{(j)} \leq \mathbf{q} - \mathbf{T}\mathbf{u}^{*(j)} - \mathbf{H}\mathbf{y} - \mathbf{W}\boldsymbol{\delta}^{(j)} - \mathbf{X}\boldsymbol{\mu}^{(j)} - \boldsymbol{\mu}^{(j)T}\mathbf{D}\boldsymbol{\delta}^{(j)}, \ \forall\, \mathbf{u}^{*(j)} \in \mathcal{A}_{part} \tag{41}$$

   where Constraint (40) is the matrix compact form of Constraints (25)-(27).
3. Solve the outer subproblem using the inner layer algorithm with $\widehat{\mathbf{y}}^{(j)}$=$\mathbf{y}^*$. Obtain an optimal solution $\rho^{(j)}$ and an optimal attack plan $\widehat{\mathbf{u}}^{(j)}$. Set $LB_{out} = \rho^{(j)}$.
4. If $(UB_{out} - LB_{out})/UB_{out} < 10^{-5}$, $\widehat{\mathbf{y}}^{(j)}$ is the optimal coupling interface and the algorithm is terminated. Otherwise, $\mathcal{A}_{part} = \mathcal{A}_{part} \cup \widehat{\mathbf{u}}^{(j)}$, set $j \leftarrow j + 1$ and return to step 2.

4. Case-study

We applied our proposed DAD problem to the IPGNs used in [3]. This case-study consists of a power network, containing five gas-fired power plant, supplied by a gas network, which needs electricity from the power network. The power network is built upon the topology of the IEEE 14-bus system [18], while for the gas network the topology of the IEEE 9-bus system is used [19]. The geographical position of the IPGNs is shown in Figure 2. For more details on the case-study, the reader is referred to [3].

We assume that the power and gas network are already coupled with an existing coupling interface. Three different existing coupling interface designs, based on network metrics, are tested:

- distance-based coupling interface (also referred to as *Euclidean* coupling interface [3]): each node in the gas network presents an interdependency link coming from the closest node in the power network. Each node in the power network containing a gas-fired power plant presents an interdependency link coming from the closest node in the gas network;

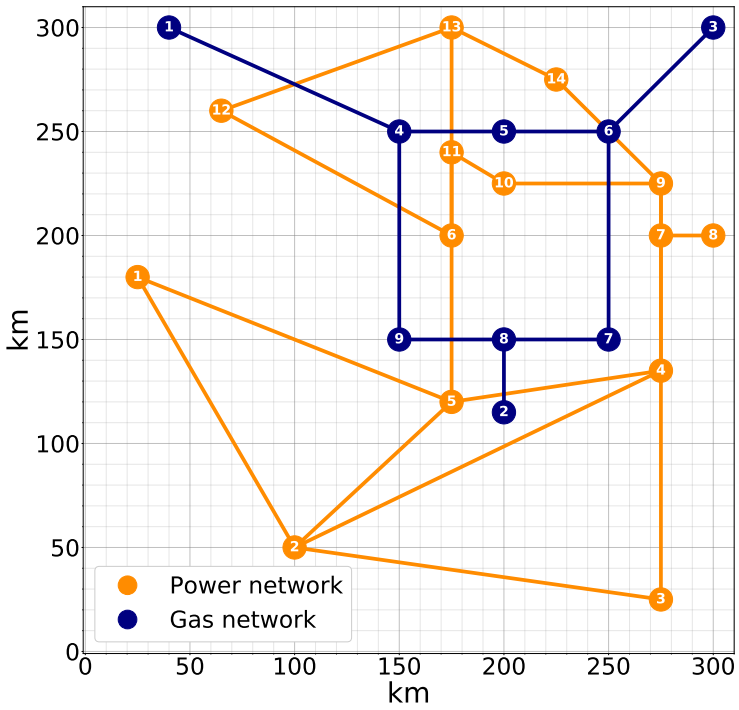

Figure 2: Interdependent power and gas networks. Figure from [3].

- degree-based coupling interface: the node with the k^{th} highest degree in the gas network presents an interdependency link coming from the node with the k^{th} highest degree in the power network. The node with the k^{th} highest degree containing a gas-fired power plant in the power network presents an interdependency link coming from the node with the k^{th} highest degree in the gas network;

- betweenness-based coupling interface: the node with the k^{th} highest betweenness in the gas network presents an interdependency link coming from the node with the k^{th} highest betweenness in the power network. The node with the k^{th} highest betweenness containing a gas-fired power plant in the power network presents an interdependency link coming from the node with the k^{th} highest betweenness in the gas network.

The optimal allocation of redundant interdependency links is investigated with K_{att} from 1 to 5. Four different budgets B_{ci} are used: \$100,\$200, \$300, and \$400. The allocation costs $c_{km}^{g\leftarrow p}$ and $c_{km}^{p\leftarrow g}$ are 1 \$/km. The NC&CG algorithm is solved using Gurobi [20] with an i7-8700@3.20GHz processor and 32 GB RAM.

5. Results and discussion

The results of the optimal allocation of redundant interdependency links in the distance-based, degree-based, and betweenness-based coupling interface are shown in Figures 3, 5, and 7, and Tables 1, 2, and 3, respectively. For comparison, the results of the optimal allocation are compared to the results obtained with a random allocation strategy, shown in Figures 4, 6, and 8. The random allocation strategy is described by the algorithm in Appendix A. For each combination of budget B_{ci} and failed lines K_{att}, 50 simulations of random allocation of redundant interdependency links are performed. The results in Figures 4, 6, and 8 represents the average values and their 95% confidence intervals in the distance-based, degree-based, and betweenness-based coupling interface, respectively.

By graphically comparing Figures 3, 5, and 7 with Figures 4, 6, and 8, we can immediately notice how the optimal allocation of redundant interdependency links clearly outperforms the random allocation strategy, even accounting for the upper bound of the 95% confidence interval.

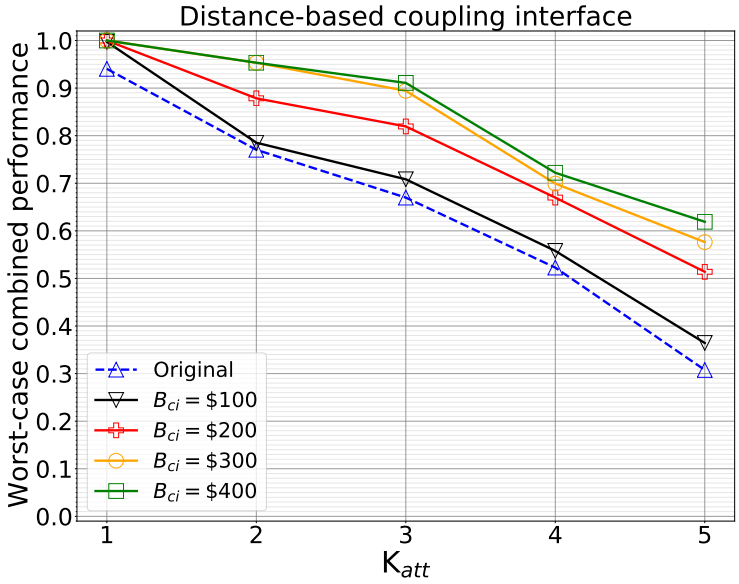

Figure 3: Results of the optimal allocation of redundant interdependency links with a pre-existing distance-based (Euclidean) coupling interface.

For the distance-based coupling interface, as shown in Figure 3 and Table 1, the relative increase of combined performance from the original coupling interface reaches a percentage equal to 101.21% for the case $K_{att} = 5$ and $B_{ci} = \$400$, passing from 0.308 to 0.619. Moreover, as shown in Figure 4, the random allocation strategy does not lead to good results if compared to the optimal allocation, even accounting for the confindence intervals. For example, for the case $K_{att} = 5$ and $B_{ci} = \$400$, the random allocation strategy leads to combined performance equal to 0.338 ± 0.070.

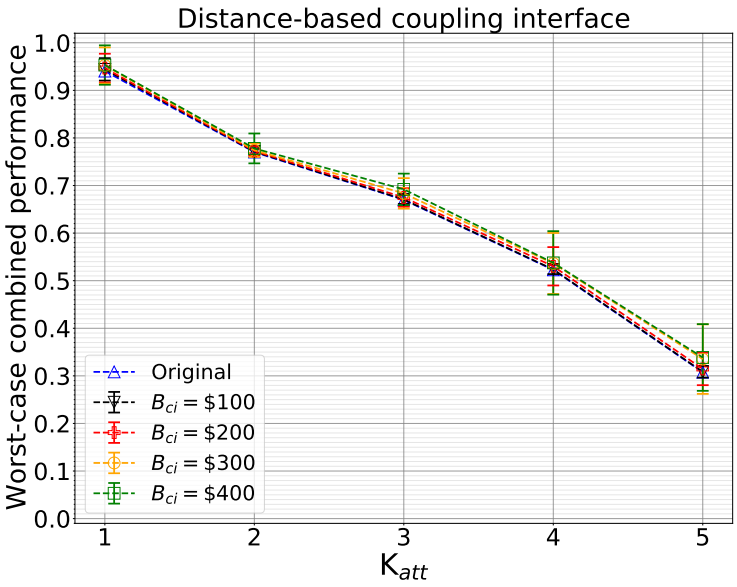

Figure 4: Results of the random allocation of redundant interdependency links with a pre-existing distance-based (Euclidean) coupling interface.

Table 1: Results in terms of worst-case combined performance and relative increase from the original case for the distance-based (Euclidean) coupling interface.

B_{ci}	K_{att}=1	K_{att}=2	K_{att}=3	K_{att}=4	K_{att}=5
Original	0.940	0.770	0.670	0.523	0.308
\$100	0.997 (+6.04%)	0.785 (+1.93%)	0.708 (+5.74%)	0.558 (+6.71%)	0.364 (+18.43%)
\$200	1.0 (+6.33%)	0.879 (+14.03%)	0.819 (+22.33%)	0.670 (+28.13%)	0.514 (+67.04%)
\$300	1.0 (+6.33%)	0.953 (+23.74%)	0.894 (+33.49%)	0.699 (+33.77%)	0.577 (+87.33%)
\$400	1.0 (+6.33%)	0.953 (+23.74%)	0.911 (+35.98%)	0.722 (+38.12%)	0.619 (+101.21%)

For the degree-based coupling interface, as shown in Figure 5 and Table 2, the relative increases of combined performance are even greater than the previous case. For example, for the case $K_{att} = 3$ and $B_{ci} = \$400$, the increase of combined performance from the original coupling interface is 347.55% (from 0.204 to 0.913). Moreover, in the original coupling interface, with $K_{att} = 4$ and $K_{att} = 5$, the worst-case combined performance are equal to 0, and the allocation of redundant

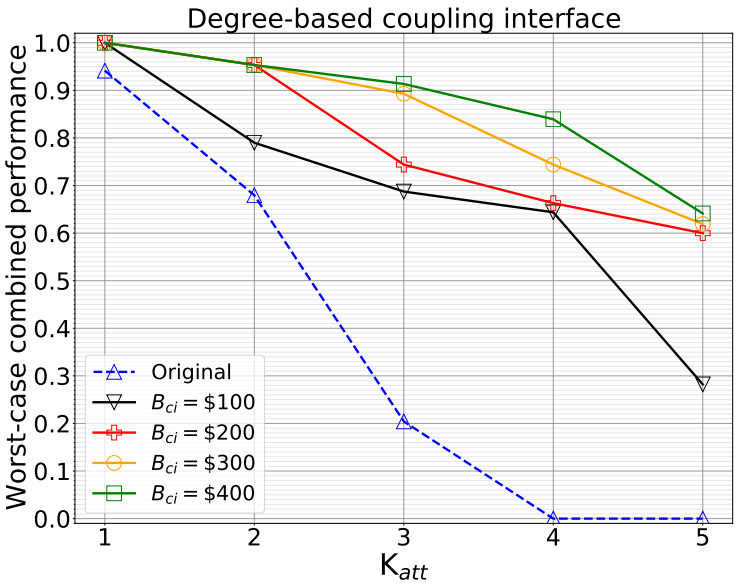

Figure 5: Results of the optimal allocation of redundant interdependency links with a pre-existing degree-based coupling interface.

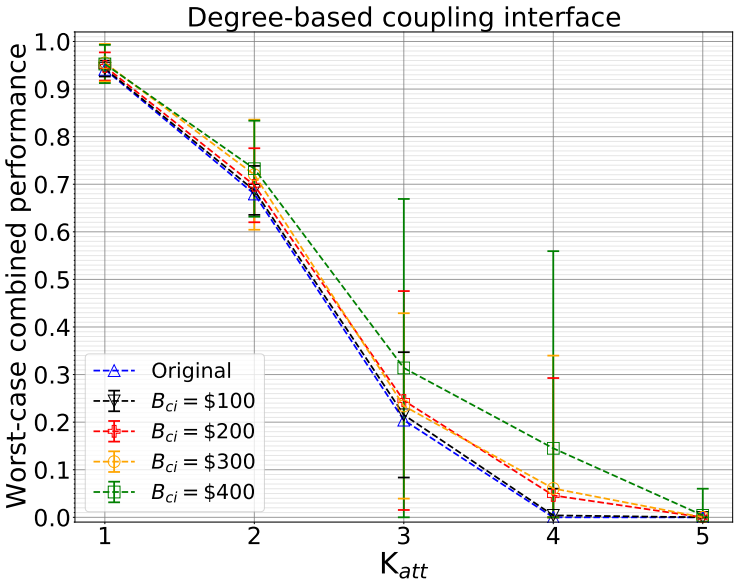

Figure 6: Results of the random allocation of redundant interdependency links with a pre-existing degree-based coupling interface.

interdependency links allows to considerably improve the worst-case combined performance. For example, for the case $K_{att} = 5$ and $B_{ci} = \$400$, we can achieve worst-case combined performance equal to 0.641. The results of the random allocation strategy, shown in Figure 6, are considerably lower if compared to the optimal allocation strategy, even considering the confidence intervals.

Table 2: Results in terms of worst-case combined performance and relative increase from the original case for the degree-based coupling interface.

B_{ci}	K_{att}=1	K_{att}=2	K_{att}=3	K_{att}=4	K_{att}=5
Original	0.941	0.679	0.204	0.0	0.0
\$100	1.0 (+6.27%)	0.790 (+16.35%)	0.687 (+236.76%)	0.644 -	0.282 -
\$200	1.0 (+6.27%)	0.953 (+40.35%)	0.744 (+264.71%)	0.663 -	0.600 -
\$300	1.0 (+6.27%)	0.953 (+40.35%)	0.893 (+337.75%)	0.744 -	0.619 -
\$400	1.0 (+6.27%)	0.953 (+40.35%)	0.913 (+347.55%)	0.839 -	0.641 -

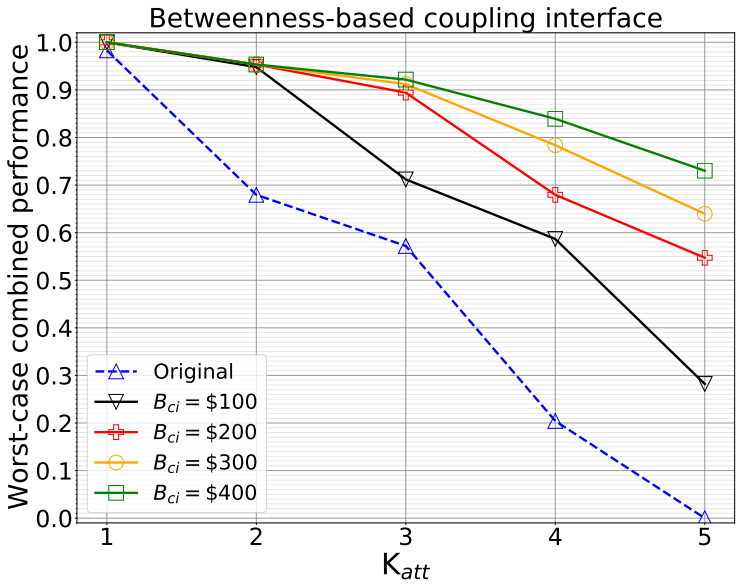

Figure 7: Results of the optimal allocation of redundant interdependency links with a pre-existing betweenness-based coupling interface.

For the betweenness-based coupling interface, as shown in Figure 7 and Table 3, the relative increases of combined performance are similar to the previous case. For example, for the case $K_{att} = 4$ and $B_{ci} = \$400$, the increase of combined performance from the original coupling interface is 311.27% (from 0.204 to 0.839). Moreover, similarly to the previous case, in the original coupling interface, with $K_{att} = 5$, the worst-case combined performance are equal to 0, and the

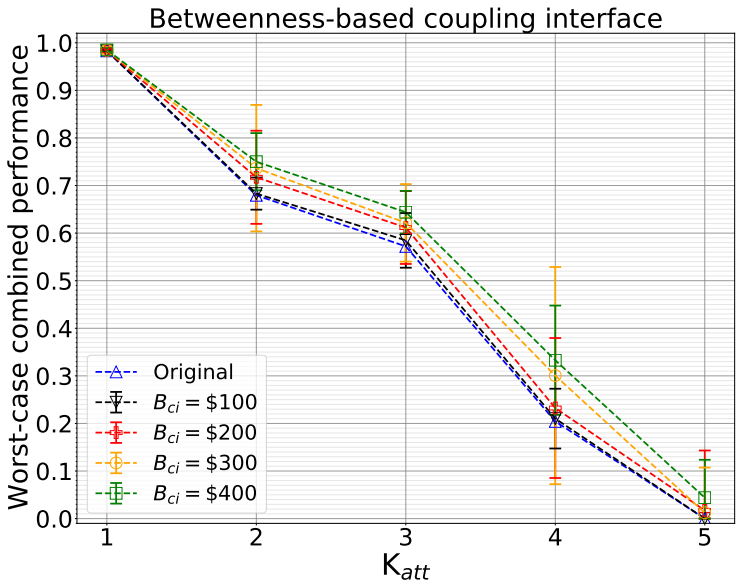

Figure 8: Results of the random allocation of redundant interdependency links with a pre-existing betweenness-based coupling interface.

Table 3: Results in terms of worst-case combined performance and relative increase from the original case for the betweenness-based coupling interface.

B_{ci}	K_{att}=1	K_{att}=2	K_{att}=3	K_{att}=4	K_{att}=5
Original	0.983	0.679	0.572	0.204	0.0
\$100	1.0 (+1.73%)	0.948 (+39.62%)	0.712 (+24.48%)	0.587 (+187.75%)	0.282 -
\$200	1.0 (+1.73%)	0.953 (+40.35%)	0.894 (+56.29%)	0.679 (+232.84%)	0.547 -
\$300	1.0 (+1.73%)	0.953 (+40.35%)	0.912 (+59.44%)	0.784 (+284.31%)	0.640 -
\$400	1.0 (+1.73%)	0.953 (+40.35%)	0.922 (+61.19%)	0.839 (+311.27%)	0.730 -

allocation of redundant interdependency links allows to considerably improve the worst-case combined performance. For example, for the case $K_{att} = 5$ and $B_{ci} = \$400$, we can achieve worst-case combined performance equal to 0.730. The results of the random allocation strategy in Figure 8 are again considerably lower if compared to the optimal allocation strategy, even considering the confidence intervals.

Regarding the computational cost of the proposed approach, we refer to the considerations already mentioned in [3], which we briefly recall here:

- in this work, the computational time is acceptable, as for each combination of K_{att} and B_{ci}, the computational time is less than 10 minutes;
- for larger case-studies, the computational time might increase considerably. However, this fact does not represent an problematic issue, as in design problems long computational times are generally not an obstacle;
- the computational cost of the problem can be reduced by limiting the number of binary variables (e.g. limiting the choice of links which is possible to allocate).

6. Conclusion

In this work, we presented a mathematical programming approach for the optimal allocation of redundant interdependency links in interdependent CIs. A DAD model is used, in order to maximize the combined performance of the interdependent CIs under the worst-case failure scenario. Using IPGNs as illustrative case-study, the results have demonstrated the great potential, in terms of resilience enhancement, of allocating redundancies within the coupling interface of interdependent CIs. Further developments of this work include different operational model of the IPGNs.

Appendix A. Random allocation strategy

Algorithm 1 Random allocation of interdependency links

Initialize $\overline{\mathbf{y}}^{g\leftarrow p}$, $\overline{\mathbf{y}}^{p\leftarrow g}$, B_{ci}, `i`=1, `cost`=0
while `i`=1 **do**
    Define `r`=random binary number
    if `r`=0 **then**
        Initialize random nodes $i \in V_{GN}$ and $j \in V_{PN}$
        if $y_{ij}^{g\leftarrow p}$=0 **then**:
            Set `cost` $\leftarrow$ `cost` $+ d_{ij}^{km} c_{km}^{g\leftarrow p}$
            if `cost` $\leq B_{ci}$ **then**:
                Set $y_{ij}^{g\leftarrow p}$=1
            else if `cost` $> B_{ci}$ **then**
                Set `i`=0 and stop the algorithm
    else if `r`=1 **then**
        Initialize random nodes $i \in V_{PN}$ and $j \in V_{GN}$
        if $y_{ij}^{p\leftarrow g}$=0 **then**:
            Set `cost` $\leftarrow$ `cost` $+ d_{ji}^{km} c_{km}^{p\leftarrow g}$
            if `cost` $\leq B_{ci}$ **then**:
                Set $y_{ij}^{p\leftarrow g}$=1
            else if `cost` $> B_{ci}$ **then**
                Set `i`=0 and stop the algorithm

References

[1] S. M. Rinaldi, J. P. Peerenboom, and T. K. Kelly, "Identifying, understanding, and analyzing critical infrastructure interdependencies," *IEEE control systems magazine*, vol. 21, no. 6, pp. 11–25, 2001.

[2] M. Ouyang, "Review on modeling and simulation of interdependent critical infrastructure systems," *Reliability engineering & System safety*, vol. 121, pp. 43–60, 2014.

[3] A. Bellè, A. F. Abdin, Z. Zeng, Y.-P. Fang, and A. Barros, "A resilience-based framework for the optimal coupling of interdependent critical infrastructures," *Reliability Engineering & System Safety*, under review, 2022.

[4] J. Winkler, L. Dueñas-Osorio, R. Stein, and D. Subramanian, "Interface network models for complex urban infrastructure systems," *Journal of Infrastructure Systems*, vol. 17, no. 4, pp. 138–150, 2011.

[5] M. Ouyang and L. Dueñas-Osorio, "An approach to design interface topologies across interdependent urban infrastructure systems," *Reliability Engineering & System Safety*, vol. 96, no. 11, pp. 1462–1473, 2011.

[6] Y.-P. Fang and E. Zio, "An adaptive robust framework for the optimization of the resilience of interdependent infrastructures under natural hazards," *European Journal of Operational Research*, vol. 276, no. 3, pp. 1119–1136, 2019.

[7] X. Liu, Y.-P. Fang, and E. Zio, "A hierarchical resilience enhancement framework for interdependent critical infrastructures," *Reliability Engineering & System Safety*, vol. 215, p. 107 868, 2021.

[8] M. Ouyang and Z. Wang, "Resilience assessment of interdependent infrastructure systems: With a focus on joint restoration modeling and analysis," *Reliability Engineering & System Safety*, vol. 141, pp. 74–82, 2015.

[9] M. Ouyang, "Critical location identification and vulnerability analysis of interdependent infrastructure systems under spatially localized attacks," *Reliability Engineering & System Safety*, vol. 154, pp. 106–116, 2016.

[10] J. Kong, C. Zhang, and S. P. Simonovic, "Optimizing the resilience of interdependent infrastructures to regional natural hazards with combined improvement measures," *Reliability Engineering & System Safety*, vol. 210, p. 107 538, 2021.

[11] Y. Almoghathawi, K. Barker, and L. A. Albert, "Resilience-driven restoration model for interdependent infrastructure networks," *Reliability Engineering & System Safety*, vol. 185, pp. 12–23, 2019.

[12] K. S. Trivedi, V. Jindal, and S. Dharmaraja, "Stochastic modeling techniques for secure and survivable systems," *Information assurance: Dependability and security in networked systems. Morgan Kaufmann*, pp. 171–207, 2008.

[13] J. Johansson, H. Hassel, and A. Cedergren, "Vulnerability analysis of interdependent critical infrastructures: Case study of the swedish railway system," *International journal of critical infrastructures*, vol. 7, no. 4, pp. 289–316, 2011.

[14] Z. Wang, T. Chen, K. Tang, and X. Yao, "A multi-objective approach to redundancy allocation problem in parallel-series systems," in *2009 IEEE Congress on Evolutionary Computation*, 2009, pp. 582–589. DOI: 10.1109/CEC.2009.4982998.

[15] M. Panteli, C. Pickering, S. Wilkinson, R. Dawson, and P. Mancarella, "Power system resilience to extreme weather: Fragility modeling, probabilistic impact assessment, and adaptation measures," *IEEE Transactions on Power Systems*, vol. 32, no. 5, pp. 3747–3757, 2016.

[16] B. Zeng and L. Zhao, "Solving two-stage robust optimization problems using a column-and-constraint generation method," *Operations Research Letters*, vol. 41, no. 5, pp. 457–461, 2013. DOI: https://doi.org/10.1016/j.orl.2013.05.003.

[17] L. Zhao and B. Zeng, "Vulnerability analysis of power grids with line switching," *IEEE Transactions on Power Systems*, vol. 28, no. 3, pp. 2727–2736, 2013.

[18] Accessed: 30/03/2022. [Online]. Available: https://icseg.iti.illinois.edu/ieee-14-bus-system/.

[19] Accessed: 30/03/2022. [Online]. Available: https://icseg.iti.illinois.edu/wscc-9-bus-system/.

[20] L. Gurobi Optimization, *Gurobi optimizer reference manual*, 2021. [Online]. Available: http://www.gurobi.com.

Paper VI

A. Bellè, Z. Zeng, and A. Barros, "Vulnerability analysis of interdependent energy infrastructures with centralized and decentralized operator models," *Proceedings of the 32nd European Safety and Reliability Conference*, accepted, 2022.

Vulnerability analysis of interdependent energy infrastructures with centralized and decentralized operator models

Andrea Bellè

Chair on Risk and Resilience of Complex Systems, Laboratoire Génie Industriel, CentraleSupélec, Université Paris-Saclay, France. E-mail: andrea.belle@centralesupelec.fr

Zhiguo Zeng

Chair on Risk and Resilience of Complex Systems, Laboratoire Génie Industriel, CentraleSupélec, Université Paris-Saclay, France. E-mail: zhiguo.zeng@centralesupelec.fr

Anne Barros

Chair on Risk and Resilience of Complex Systems, Laboratoire Génie Industriel, CentraleSupélec, Université Paris-Saclay, France. E-mail: anne.barros@centralesupelec.fr

Energy infrastructures (EIs) are large systems which provide essential energy commodities, such as electricity, gas, or heat, to people. As EIs are often interdependent on each other, integrated analysis and optimization are needed. When performing analysis and optimization of interdependent EIs, the behaviour of independent operators should be taken into account. Independent operators might display a decentralized and competitive behaviour, when they interact through the prices of energy commodities in a market-based environment, or a centralized and collaborative behaviour, when they aim at maximizing their combined performance. In this paper, we investigate the impact of centralized and decentralized operators models in the vulnerability analysis of interdependent EIs. Using interdependent power and heat networks (IPHNs), we show that these two classes of models lead to different results in terms of cost and performance. These preliminary results represent the first step in defining a decision-making framework which accounts for the two different behaviours of independent operators: decentralized in normal conditions, and centralized in conditions of disruptions.

Keywords: Energy infrastructures, vulnerability, independent operators, interdependent networks, Stackelberg game, optimization.

1. Introduction

Energy infrastructures (EIs) consist of different systems and technologies, such as power networks, gas networks and heat networks. These infrastructures are often interconnected to, and interdependent on each other. For example, heat networks are often dependent on power networks for the production of heat through electric boilers, and on gas networks for the production of heat through gas boilers (Wang et al., 2020). Similarly, power networks can be dependent on gas networks for the production of electricity in gas-fired power plants, and gas networks need electricity from power networks for running pumps, compressors, valves and other equipment (Fang and Zio, 2019). These infrastructures are essential for maintaining vital societal functions, and analyzing their behaviours in different conditions and ensuring their resilience against disruptive events are key tasks.

Despite being interdependent on each other, energy infrastructures are often operated by separate and independent operators. When performing analysis and optimization of energy infrastructures, the behaviour of independent operators should be taken into account. The actions of energy operators are often modeled as the optimization of a function that can represent cost, profit, or dispatch of energy commodities. In the existing literature, the behaviour of independent EIs operators is usually modeled using two different approaches: centralized and decentralized.

In centralized approaches, a unique centralized operator controls the operations of the interdependent EIs, dispatching the corresponding energy

Proceedings of the 32nd European Safety and Reliability Conference.
Edited by Maria Chiara Leva, Edoardo Patelli, Luca Podofillini and Simon Wilson

ISBN: 981-973-0000-00-0 :: doi: 10.3850/981-973-0000-00-0_output

commodities in order to maximize profit and/or supplied energy, or minimize cost and/or energy not supplied. Centralized models have been often applied in the context of resilience assessment and optimization. For example, centralized operator models have been used to optimize the resilience of interdependent infrastructures via combined improvements (Kong et al., 2021), for the resilience enhancement of interdependent power and gas networks against natural hazards with a robust approach (Fang and Zio, 2019), and for the resilience enhancement of interdependent power and water networks under spatially-localized attacks (Ouyang, 2017). Centralized operator models have also been used for optimizing the joint restoration of disrupted interdependent infrastructures (Almoghathawi et al., 2019; Ouyang and Wang, 2015; Lee II et al., 2007). In the aforementioned works, the operators are centralized, as they can control simultaneously the ensemble of interdependent infrastructures in order to optimize an objective function, which is usually a weighted sum of the performance or cost of each individual infrastructure.

In decentralized models, operators are considered to be independent, and they interact within a market-based environment. For example, operators of interdependent power and gas networks interact by purchasing from and selling to each other energy commodities. The purpose of each operator is usually to optimize their own profit or cost, and the interaction between multiple operators is often modeled through game-theoretic approaches. For example, decentralized models have been considered for the expansion planning of interdependent power and gas networks (Rad et al., 2019; Conejo et al., 2020; Qiu et al., 2014), the optimization of urban energy networks (Jing et al., 2018), security-constrained operations of integrated wind and hydrogen systems (Mirzaei et al., 2019) and the risk assessment of interdependent power and heat networks (Wang et al., 2020).

In general, centralized and decentralized approaches lead to different results, but they are both useful in terms of modeling and decision-making.

Centralized models are often used for modeling resilience assessment and enhancement, and decision-making in conditions of disruption. In fact, in conditions of large disruption, independent operators might fully collaborate in a centralized way in order to minimize negative consequences on the general population. In this case, the action of energy operators are usually modeled as the weighted sum of performance metrics of each individual infrastructure.

Decentralized models are useful to model normal conditions, where energy operators behave independently in a market-based environment, in order to optimize their own profit or cost. In this case, the action of energy operators are usually modeled through game-theoretic approaches.

Decision-makers should take these considerations into account. For example, when performing some joint decision-making, such as a joint transmission and/or generation expansion, decision-makers should consider that operators might display decentralized behaviours in normal conditions, by interacting through prices and demands of energy commodities, and centralized behaviours in conditions of disruption, by aiming at maximizing their combined performance.

In this preliminary work, we investigate the impact of decentralized and centralized models by performing a vulnerability analysis of interdependent power and heat networks.

The rest of the paper is organized as follows: in Section 2, the problem formulation is presented; in Section 3, the illustrative numerical example is presented; in Section 4, the preliminary results are shown and briefly analyzed; in Section 5, some preliminary conclusions and future directions are detailed.

2. Problem formulation

We consider the case-study of interdependent power and heat networks (IPHNs) proposed in Wang et al. (2020). Particularly, we consider a power network PN which supplies electricity to a heat network HN. The heat network is equipped with electric boilers, which need a power supply from the power network, and gas boilers. In this work, gas boilers are assumed to have a reliable gas supply.

Each infrastructure is modeled with a network-

based approach, where a network is described by a graph G=(V, E), where V is the set of M nodes and E is the set of N edges. The origin and destination nodes of each edge k are defined as $O(k)$ and $D(k)$, respectively. In power networks, nodes represent buses and edges represent power lines; in heat networks, nodes represent hubs and edges represent pipelines.

The operations of both power and heat network are modeled with a linear maximal flow approach. The models of both the networks consist of the minimization of an objective function g, which depends on the specific operator model (centralized or decentralized), as shown in (1).

$$\min_{\mathbf{p},\mathbf{d},\mathbf{f}} \; g(\mathbf{p},\mathbf{d},\mathbf{f}) \tag{1}$$

The specific objective functions are detailed in the next sections.

The power network is subject to Constraints (2)-(5).

$$0 \leq p_i \leq \overline{p}_i, \quad \forall i \in V_{PN} \tag{2}$$

$$0 \leq d_i \leq \overline{d}_i \quad \forall i \in V_{PN} \tag{3}$$

$$-u_k^* \overline{f}_k \leq f_k \leq u_k^* \overline{f}_k, \quad \forall k \in E_{PN} \tag{4}$$

$$p_i - (\overline{d}_i - d_i) - \sum_{j \in V_{HN}} d_{j,i}^{h \leftarrow p} + \\ + \sum_{k|D(k)=i} f_k - \sum_{k|O(k)=i} f_k = 0, \;\; \forall i \in V_{PN} \tag{5}$$

Constraints (2) and (3) defines the limit of power production and shedding in each node. The terms $\overline{p}_i$ and $\overline{d}_i$ are the production capacity and the power requested demand in each node, respectively. Constraint (4) bounds the flow of power in each line k, in absolute value, within its capacity $\overline{f}_k$, if the line is functional (u_k^*=1). If the line is failed (u_k^*=0), the flow is enforced to be 0. Constraint (5) ensures that the net power balance in each node is 0. The term $d_{j,i}^{h \leftarrow p}$ represent the power, supplied to node $i \in V_{PN}$, necessary to produce heat in the electric boiler of node $j \in V_{HN}$ (which, thus, depends on node $i \in V_{PN}$).

The heat network is subject to Constraints (6)-(10):

$$0 \leq p_i \leq \overline{p}_i, \quad \forall i \in V_{HN} \tag{6}$$

$$0 \leq d_i \leq \overline{d}_i \quad \forall i \in V_{HN} \tag{7}$$

$$-u_k^* \overline{f}_k \leq f_k \leq u_k^* \overline{f}_k, \quad \forall k \in E_{HN} \tag{8}$$

$$p_i - (\overline{d}_i - d_i) + \sum_{k|D(k)=i} f_k - \\ - \sum_{k|O(k)=i} f_k = 0, \;\; \forall i \in V_{PN} \tag{9}$$

$$p_i = \eta d_{i,j}^{h \leftarrow p}, \quad \forall i \in V_{HN}, j \in V_{PN} \tag{10}$$

Constraints (6)-(9), similarly to Constraints (2)-(5) of the power network, represent limits of heat production and shedding, heat flow, and heat nodal balance. Moreover, if node $i \in V_{HN}$, dependent on node $j \in V_{PN}$, is provided with an electric boiler, its heat production p_i is equal to the power supply $d_{i,j}^{h \leftarrow p}$ multiplied by the power-to-heat conversion efficiency η, as shown in Constraint (10). We assume that each node in the heat network equipped with an electric boiler is dependent on one node of the power network for the electricity supply.

2.1. *Load shedding model*

For the centralized model, we consider an approach which aims at minimizing the combined load shedding of power and heat. We refer to this model as the *load shedding* (LS) model. It simply consists of the objective function in (11):

$$\min_{\mathbf{p},\mathbf{d},\mathbf{f}} \sum_{i \in V_{PN}} \frac{d_i}{\overline{d}_{PN}} + \sum_{i \in V_{HN}} \frac{d_i}{\overline{d}_{HN}} \tag{11}$$

subject to Constraints (2)-(10). The terms $\overline{d}_{PN}$ and $\overline{d}_{HN}$ represent the total requested demand of power and gas, respectively. As it is clearly visible, the single objective functions for power and heat network (minimization of fractions of power and heat shedding, respectively) are simply aggregated within a single linear programming model. With this model, the two independent operators act in a centralized way in order to minimize the combined fraction of power and heat shedding.

2.2. *Game-theoretic cost model*

For the decentralized model, we rely on the game-theoretic approach proposed by Wang et al. (2020). We refer to this model as the *game-theoretic cost* (GTC) model. In this case, operators aims at minimizing their own costs, and they interact by selling and purchasing electricity. In fact, heat operators need to purchase electricity from the power network in order to produce heat in the electric boilers. The operators display a competitive behaviour: power operators aim at selling their electricity at a high price, while the heat operators aim at purchasing it at a low price.

In this case, the cost-based objective functions of power and heat network can not be directly aggregated, but they need to be treated with a game-theoretic approach.

The model of the power network consists of the objective function in (12) subject to (2)-(5) and (13):

$$\min_{\mathbf{p},\mathbf{d},\mathbf{f}} \sum_{i \in V_{PN}} \gamma_i^{power} p_i + \sum_{i \in V_{PN}} \alpha_i^{power} d_i - \sum_{i \in V_{PN}} \sum_{j \in V_{HN}} \beta_i d_{j,i}^{h \leftarrow p} \quad (12)$$

$$\beta^{min} \leq \beta_i \leq \beta^{max}, \quad \forall i \in V_{PN} \quad (13)$$

The cost function consists of three terms:

- the power production cost, where γ_i^{power} is the cost per unit of power production in each node i;
- the penalty cost associated with power load shedding, where α_i^{power} is the penalty per unit of power associated to power load shedding in each node i;
- the profit associated to selling electricity to the heat network, where β_i is the selling price per unit of power produced in each node i.

The electricity selling price is a variable which is bounded between a minimum and maximum value, as shown in Constraint (13).

Similarly, the cost model of the heat network consists of the objective function in (14) subject to (6)-(10).

$$\min_{\mathbf{p},\mathbf{d},\mathbf{f}} \sum_{i \in V_{HN}^{gb}} \gamma_i^{heat} p_i + \sum_{i \in V_{HN}} \alpha_i^{heat} d_i + \sum_{i \in V_{PN}} \sum_{j \in V_{HN}} \beta_i d_{j,i}^{h \leftarrow p} \quad (14)$$

The cost function in consists of three terms, respectively:

- the power production cost in gas boilers, where γ_i^{heat} is the cost per unit of heat in each node i;
- the cost associated with heat load shedding, where α_i^{heat} is the penalty per unit of heat associated to heat load shedding in each node i;
- the cost associated to purchasing electricity from the power network, where β_i is the selling price per MW of power, which is seen as a constant by the model.

The two problems are interconnected through the electricity demand of electric boilers in the heat network, given by the terms $d_{j,i}^{h \leftarrow p}$, and the electricity prices β_i.

Power and heat operators are competing for the electricity prices: power operators aim at increasing prices β_i in order to maximize their profit, while heat operators aim at decreasing the electricity price for minimizing their operational costs. If the price is too high, the heat operators will not purchase electricity; if the price is too low, the power operators will reach a sub-optimal solution in terms of cost minimization. This situation can be modeled using game-theory. In this paper, we assume that the power network operators "move" first by setting the price of electricity, and, subsequently, the heat network operator adjust the electricity demands $d_{j,i}^{h \leftarrow p}$. This situation leads to a Stackelberg game between the power and heat network operators. Under this assumption, as proposed in Wang et al. (2020), the two optimization problems can be recast into a single one by replacing the model of the heat network with its Karush-Kuhn-Tucker (KKT) conditions. For simplicity, we rely directly on a compact matrix formulation. For more details on the single-level reformulation

of the Stackelberg game, the reader is referred to Wang et al. (2020).

The single-level Stackelberg game between power and heat operators can, thus, be formulated as the objective function in (15), subject to (13) and (16)-(20).

$$\min_{\mathbf{h}} \mathbf{c}^T \mathbf{h} \tag{15}$$

subject to:

$$\mathbf{R}^{in}\mathbf{h} - \mathbf{b}^{in} \leq 0 \tag{16}$$

$$\mathbf{R}^{eq}\mathbf{h} - \mathbf{b}^{eq} = 0 \tag{17}$$

$$\nabla \mathbf{c}^T \mathbf{h} + \boldsymbol{\lambda} \nabla \mathbf{R}^{in} \mathbf{h} + \boldsymbol{\mu} \nabla \mathbf{R}^{eq} \mathbf{h} = 0 \tag{18}$$

$$\boldsymbol{\lambda}(\mathbf{R}^{in}\mathbf{h} - \mathbf{b}^{in}) = 0 \tag{19}$$

$$\boldsymbol{\lambda} \geq 0 \tag{20}$$

Equation (15) represents the objective function in (12). Constraints (16) and (17) represents the primary constraints of power and heat networks, previously shown in Equations (2)-(10). Constraints (18)-(20) represents stationarity conditions, complementary slackness, and non-negativity of the KKT conditions of the heat network model (Equation (14) subject to Constraints (6)(10)).

The term $\mathbf{h}$ represents the continuous primary variables $\mathbf{p}$, $\mathbf{d}$ and $\mathbf{f}$, while $\boldsymbol{\lambda}$ and $\boldsymbol{\mu}$ represent the dual variables of inequality and equality constraints of the heat network, respectively. The vector $\mathbf{c}$ contains the coefficients of the objective function, while the matrices $\mathbf{R}^{in}$ and $\mathbf{R}^{eq}$ contain the coefficients of inequality and equality constraints of both power and heat networks. The vectors $\mathbf{b}^{in}$ and $\mathbf{b}^{eq}$ contain the constraints parameters.

The solution of this optimization problem leads to the minimized equilibrium costs for power and heat networks under the assumption of a Stackelberg game.

3. Numerical example

The case study is built upon the one proposed in Wang et al. (2020). The reader is referred to Figure 5 of Wang et al. (2020) for a graphical visualization. We assume that the heat network is composed by 32 nodes connected of 32 edges (Liu et al., 2016). Each node is characterized by a requested demand of heat supply that ranges from 0 to 0.145 MW, for a total demand of 2.164 MW. The heat network is provided with one 1 gas boiler and 3 electric boilers, each of them with a production capacity of 0.8 MW. Each edge has a flow capacity of 1 MW.

The three electric boilers need electricity, and they are supplied by three nodes of the power network, composed of 33 nodes and 32 edges (Baran and Wu, 1989). Each node is characterized by a requested power demand, ranging from 0 to 0.121 MW, for a total demand of 3.655 MW. The power network is equipped with a generator with capacity of 3.5 MW, and four additional generators with a capacity of 0.5 MW. Each edge has a power flow capacity of 3.5 MW.

The electricity price boundaries are set as β^{min}=\$200 and β^{max}=\$1000. The penalty values α_i^{power} and α_i^{heat} for each node are set between β^{min} and β^{max}. The power and heat production cost γ_i^{power} and γ_i^{heat} are set to \$50. The power-to-heat efficiency η is set to 0.8.

We perform a preliminary vulnerability analysis by solving the LS model and the GTC model for every possible combinations of N-1, N-2, and N-3 line contingencies in the power network. The results are evaluated for both the models in terms of average power and heat load shedding and average cost. The cost in the LS model is computed using the equilibrium electricity prices β_i identified by the GTC model under the same contingency scenario.

This illustrative case-study represents a numerical example to analyze the results of the two models and draw some preliminary considerations.

The computations are implemented in Gurobi 9.1 Gurobi Optimization (2021) and performed on a laptop with a 2.60 GHz CPU and 16 GB RAM.

4. Preliminary results

The results in normal conditions (no failed lines), in terms of load shedding and cost, are shown in Tables 1 and 2 .

Table 1. Load shedding in normal conditions with LS model and GTC model.

Model	LS PN	LS HN	Average LS
LS	0.00	0.00	0.00
GTC	0.00	0.186	0.093

Table 2. Costs in normal conditions with LS model and GTC model.

Model	Cost PN	Cost HN	Average cost
LS	-345.80$	653.80$	154.00$
GTC	-189.83$	591.01$	200.59$

We can immediately notice how the two models lead to different results, even in normal conditions. In terms of shedding, the LS model, as expected, leads to no load shedding of power or heat. However, the GTC model leads to a fraction of heat shedding equal to 0.186. A reduction of the heat shedding fraction would imply an electricity purchase with a price above the equilibrium price and, thus, an increase of cost. In addition, the LS model leads to a lower average cost. In fact, the heat operators, aiming at minimizing their load shedding, purchase more electricity from the power network. This contributes to increase the cost of the heat network and decrease the cost of the power network. However, despite the average cost being lower, these values do not represent an equilibrium solution.

The results of the preliminary vulnerability analysis are shown in Figures 1-6. The bar plots denotes the average values of load shedding fraction and cost for both models and N-1, N-2, and N-3 power line contingencies.

The results in terms of load shedding for the combined networks, the power network, and heat network are shown in Figures 1, 2, and 3. As it was expected, in Figure 1, we can notice that the LS model leads to lower level of average combined load shedding in all the three contingency scenarios evaluated. However, as it can be seen in Figures 2, and 3, the same consideration is not valid for the individual load shedding in power and heat networks. In fact, the GTC model leads to fraction of load shedding lower than the LS model in the power network, and considerably higher than the LS model in the heat network. However, overall, the GTC model leads to higher levels of combined load shedding.

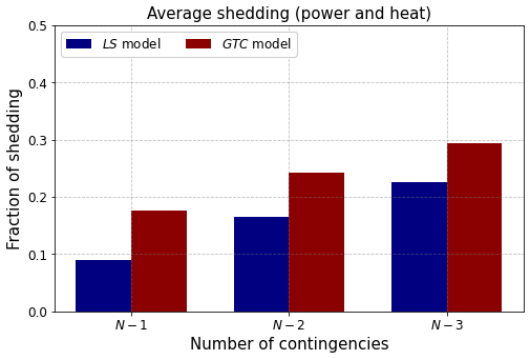

Fig. 1. Average load shedding in the IPHNs.

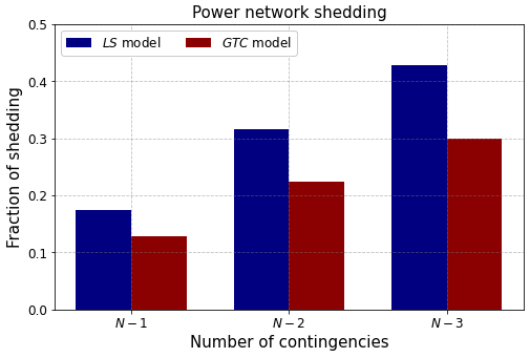

Fig. 2. Average load shedding in the power network.

The results in terms of cost for the combined networks, the power network, and heat network are shown in Figures 4, 5, and 6, respectively. Firstly, looking at the total cost of power and heat network in Figure 4, we notice that for the cases

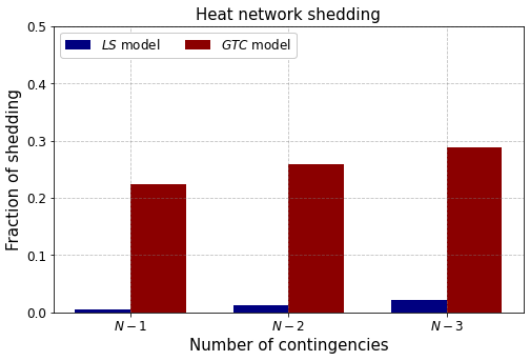

Fig. 3. Average load shedding in the heat network.

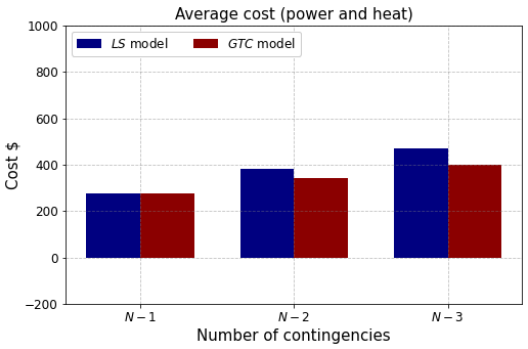

Fig. 4. Average cost in the IPHNs.

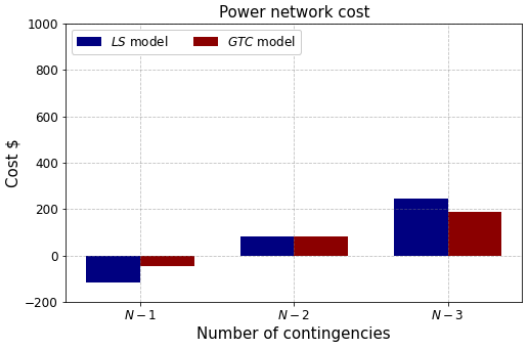

Fig. 5. Average cost in the power network.

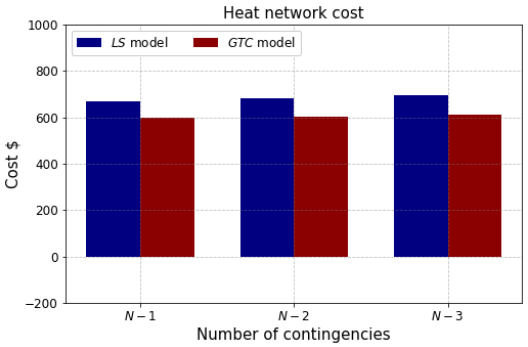

Fig. 6. Average cost in the heat network.

N-2, and N-3, the GTC model leads to lower average costs; however, for the case N-1, the LS model leads to a slightly lower average cost (the difference is small, less than \$1). For understanding this behaviour, the individual costs of power and heat network, in Figures 5 and 6, need to be analyzed. In fact, if we look at the power network cost in Figure 5, we can notice for the case N-1 the LS model leads to a considerably lower cost. However, if we look at the heat network cost in Figure 6, we see that the LS model leads to a higher cost. These are the results of the minimization of load shedding: in fact, the heat operators, when modeled with the LS model, aiming at minimizing the heat load shedding, purchase more electricity that they would with the price-aware GTC model. As a consequence, the power network sells more electricity, which is purchased by the heat network. In conclusion, in terms of cost, similarly to the case in Tables 1 and 2, the LS model leads to solutions far from the optimal equilibrium of the GTC model.

5. Conclusion

With these preliminary results, we show how centralized and decentralized models lead to different results, in terms of cost and shedding, both in normal and disrupted situations. Decentralized models, like the one proposed in Wang et al. (2020), imply a market-based competition between the operators of different infrastructures, and they are useful for modeling normal situations, when operators aim at minimizing their costs. Centralized models imply a collaborative behaviour between operators of different infrastructures, and they are useful in disrupted conditions, when the objective, rather than minimizing a cost, is to restore the infrastructures functionality as rapidly as possible.

In conclusion, centralized and decentralized models lead to different solutions in terms of energy dispatch, as they are based on different ob-

jective functions and assumptions. However, both these classes of models are useful, as they represent operators behaviours in different situations.

When performing optimization and analysis of interdependent EIs, these behaviours in different situations should be taken into account. Further developments of this work will investigate the possibility of a decision-making framework for interdependent EIs which account for decentralized operators behaviours in normal conditions and centralized behaviours in conditions of disruption.

References

Almoghathawi, Y., K. Barker, and L. A. Albert (2019). Resilience-driven restoration model for interdependent infrastructure networks. *Reliability Engineering & System Safety 185*, 12–23.

Baran, M. E. and F. F. Wu (1989). Network reconfiguration in distribution systems for loss reduction and load balancing. *IEEE Power Engineering Review 9*(4), 101–102.

Conejo, A. J., S. Chen, and G. E. Constante (2020). Operations and long-term expansion planning of natural-gas and power systems: A market perspective. *Proceedings of the IEEE 108*(9), 1541–1557.

Fang, Y.-P. and E. Zio (2019). An adaptive robust framework for the optimization of the resilience of interdependent infrastructures under natural hazards. *European Journal of Operational Research 276*(3), 1119–1136.

Gurobi Optimization, L. (2021). Gurobi optimizer reference manual.

Jing, R., M. Wang, H. Liang, X. Wang, N. Li, N. Shah, and Y. Zhao (2018). Multi-objective optimization of a neighborhood-level urban energy network: Considering game-theory inspired multi-benefit allocation constraints. *Applied energy 231*, 534–548.

Kong, J., C. Zhang, and S. P. Simonovic (2021). Optimizing the resilience of interdependent infrastructures to regional natural hazards with combined improvement measures. *Reliability Engineering & System Safety 210*, 107538.

Lee II, E. E., J. E. Mitchell, and W. A. Wallace (2007). Restoration of services in interdependent infrastructure systems: A network flows approach. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews) 37*(6), 1303–1317.

Liu, X., J. Wu, N. Jenkins, and A. Bagdanavicius (2016). Combined analysis of electricity and heat networks. *Applied Energy 162*, 1238–1250.

Mirzaei, M. A., A. S. Yazdankhah, and B. Mohammadi-Ivatloo (2019). Stochastic security-constrained operation of wind and hydrogen energy storage systems integrated with price-based demand response. *International Journal of Hydrogen Energy 44*(27), 14217–14227.

Ouyang, M. (2017). A mathematical framework to optimize resilience of interdependent critical infrastructure systems under spatially localized attacks. *European Journal of Operational Research 262*(3), 1072–1084.

Ouyang, M. and Z. Wang (2015). Resilience assessment of interdependent infrastructure systems: With a focus on joint restoration modeling and analysis. *Reliability Engineering & System Safety 141*, 74–82.

Qiu, J., Z. Y. Dong, J. H. Zhao, Y. Xu, Y. Zheng, C. Li, and K. P. Wong (2014). Multi-stage flexible expansion co-planning under uncertainties in a combined electricity and gas market. *IEEE Transactions on Power Systems 30*(4), 2119–2129.

Rad, V. Z., S. A. Torabi, and H. Shakouri (2019). Joint electricity generation and transmission expansion planning under integrated gas and power system. *Energy 167*, 523–537.

Wang, C., C. Yan, G. Li, S. Liu, and Z. Bie (2020). Risk assessment of integrated electricity and heat system with independent energy operators based on stackelberg game. *Energy 198*, 117349.

Paper VII

A. Bellè, Y.-P. Fang, Z. Zeng, and A. Barros, "A distributionally robust approach for the optimal protection of power networks with endogenous uncertainty," *Proceedings of the 18th IFAC Workshop on Control Application of Optimization*, accepted, 2022.

A distributionally robust approach for the optimal protection of power networks with endogenous uncertainty

Andrea Bellè * **Yi-Ping Fang** * **Zhiguo Zeng** * **Anne Barros** *

* *Chair on Risk and Resilience of Complex Systems, Laboratoire Génie Industriel, CentraleSupélec, Université Paris-Saclay, France.*

Abstract:
Power networks are among the most important infrastructures in any society, and protecting them from a large variety of disruptive events is an essential task. Finding an optimal protection plan often takes the form of a multilevel optimization problem. Recently, approaches based on distributionally robust optimization (DRO) have gained the attention of many scholars. In fact, DRO allows protecting the decision-maker from the ambiguity arising from the imprecisely identified probability distributions of the failure scenarios. In this framework, the probability distribution of the failure scenarios is assumed to be known ambiguously, and contained in an ambiguity set defined by moment-based conditions. In practice, some decisions taken in the protection plan affect the conditions of the ambiguity set. We refer to this situation as DRO with endogenous uncertainty. In this paper, we study the impact of the endogenous uncertainty on the optimal protection plan of a power network, using a numerical example built upon the IEEE 14-bus system with a traditional moment-based ambiguity set.

Keywords: Power network, optimal protection, distributionally robust optmization, moment-based ambiguity set, critical infrastructures.

1. INTRODUCTION

Power networks are among the most important infrastructures for any society, as they provide private and public customers with electricity. Failures and disruption within power networks can lead to considerable negative consequences (Garcia Tapia et al., 2019). The protection of power networks components, i.e., to reinforce them and make them invulnerable to external attacks and hazards, is an effective measure to decrease the vulnerability of power networks and increase the performance in disruptive conditions (Yuan et al., 2014). Due to technical and economical constraints, often, only a limited number of components can be protected. The set of components to protect, which constitutes the protection plan, must then be optimized.

Various works have proposed frameworks for the optimization of protection plans, relying on the traditional *defender-attacker-defender* (DAD) model, adaptive robust optimization and distributionally robust optimization.

In the DAD framework, the decision maker seeks to identify the optimal protection plan against the worst-case scenario within a set of feasible failure/attack scenarios (in this work, the words failure/failed and attack/attacked are used interchangeably). For example, in Ouyang (2017), the author exploits a DAD approach to identify the optimal protection plan against spatially-localized attacks in interdependent power and water networks; similarly, in Ouyang and Fang (2017) and Yuan et al. (2014), the authors provide the optimal protection plan for a power network against attacks and failures.

When the decision maker has information on the failure probability of specific components, an adaptive robust approach can be used. For example, in Fang and Zio (2019), the optimal protection plan for interdependent power and gas networks against typhoons is computed with an adaptive robust approach, accounting for the fragility curves of the power network components and the confidence level of the decision maker.

If distributional information on the failure probability of components is available, a distributionally robust approach can also be applied. For example, in Zhang et al. (2020), the optimal protection plan is identified with a distributionally robust approach, where the probability distribution of feasible failure scenarios is ambiguous. Within this framework, the probability of each scenario is unknown and ambiguous, and contained within an ambiguity set in which moment information of the probability distribution are encoded. A similar approach is applied in Bagheri and Zhao (2019), where the optimal protection plan is identified with a distributionally robust approach under a $N-k$ security criterion.

Distributionally robust approaches have recently attracted the attention of several researchers, thanks to their ability to protect decision maker from the ambiguity which often affects the probability distribution of failure/attack scenarios (Rahimian and Mehrotra, 2019). Within the framework of power network optimization, various recent works have applied distributionally robust approaches. For example, in Babaei et al. (2020), the authors propose a DRO approach for the optimal power network configura-

tion under random failures. In Zhao and Jiang (2017), a contingency-constrained unit commitment model is proposed using DRO. In Alvarado et al. (2018), a transmission expansion planning with distributionally robust security is proposed.

In general, in a distributionally robust framework, a decision maker seeks to identify the decision plan $\mathbf{z}$ which is robust against the worst probability distribution in the ambiguity set. However, first-stage decisions $\mathbf{z}$ can also affect the probability distributions contained within the ambiguity set, which is then dependent on $\mathbf{z}$ (decision-dependent ambiguity set). We refer to this situation as distributionally robust optimization with endogenous uncertainty. The reader is referred to Luo and Mehrotra (2020) and Noyan et al. (2018) for a theoretical framework, Doan (2022) for an application to retrofitting of transportation networks after natural disasters, and Noyan et al. (2021) for a unified modeling framework and an illustrative application to machine scheduling and humanitarian logistics.

When a distributionally robust approach, such as that in Zhang et al. (2020) and Bagheri and Zhao (2019), is used for the optimal protection plan of a power network, the ambiguous distribution is related to the occurrence probability of the feasible failure/attack scenarios. When we protect an element, such as a transmission line, its failure probability is reduced. Consequently, the occurrence probability of each scenario that implies the failure of a protected line changes. Thus, the ambiguity set is decision-dependent, as it can be modified by the protection plan.

To the best of our knowledge, optimal protection plans of power networks with a distributionally robust approach which accounts for this endogenous uncertainty through a decision-dependent ambiguity set have not been investigated.

In this work, we propose a distributionally robust framework for identifying optimal protection plans in power networks with endogenous uncertainty. In our preliminary results, we highlight the difference in results between the cases with and without endogenous uncertainty.

2. PROBLEM FORMULATION

2.1 Power network modeling

A network is a mathematical construct described by a graph $G = (V, E)$, with the set V containing N nodes and the set E containing M edges. Each edge k is directed, and it is defined by an origin node $O(k)$ and a destination node $D(k)$. Each node i is characterized by a production capacity $\overline{p}_i$ and a requested demand $\overline{d}_i$, while each edge k is defined by a flow capacity $\overline{f}_k$ (Bellè et al., 2021).

Power networks are usually modeled using power flow models, such as the DC optimal power flow (Fang and Sansavini, 2017). For the sake of simplicity, in this work, the power network is modeled using a linear maximal flow model, which is an approximation for flow-based infrastructures (González et al., 2016). The purpose of power network operators is to maximize the performance of the power network. The performance is expressed in terms of fraction of requested power demand which is supplied, as shown in Equation (1):

$$\max_{\mathbf{p},\mathbf{d},\mathbf{f}} \frac{\sum_{i\in V} d_i}{\sum_{i\in V} \overline{d}_i} \tag{1}$$

where d_i and $\overline{d}_i$ represents the supplied and requested power demand at bus i, respectively. The optimization in (1) is subject to the constraints of the maximal flow model, shown in Equations (2)-(6):

$$0 \le p_i \le \overline{p}_i, \quad \forall i \in V \tag{2}$$

$$0 \le d_i \le \overline{d}_i, \quad \forall i \in V \tag{3}$$

$$f_k \ge -\overline{f}_k((z_k) + u_k(1 - z_k)), \quad \forall k \in E \tag{4}$$

$$f_k \le \overline{f}_k((z_k) + u_k(1 - z_k)), \quad \forall k \in E \tag{5}$$

$$p_i - d_i + \sum_{D(k)=i} f_k - \sum_{O(k)=i} f_k = 0, \quad \forall i \in V \tag{6}$$

Constraint (2) bounds the power production in each node p_i between 0 and the maximum production capacity $\overline{p}_i$. Similarly, the supplied power demand d_i is bounded between 0 and the requested demand $\overline{d}_i$, as shown in Constraint (3).

Constraints (4) and (5) limit the flow in each line, in absolute value, to the maximum capacity $\overline{f}_k$. The term $((z_k) + u_k(1 - z_k))$ governs the functionality of line k. If a line k is protected, the variable $z_k = 1$ and the line is considered to be invulnerable, similarly to other works in the existing literature, e.g. Ouyang and Fang (2017); otherwise, the variable $z_k = 0$. If a line is attacked (in this work, the word attacked and failed are used interchangeably), the variable $u_k = 0$; otherwise, the variable $u_k = 1$. In this way, if line k is unprotected ($z_k = 0$) and attacked/failed ($u_k = 0$), its flow f_k is forced to be 0; otherwise, the flow f_k is bounded between $-\overline{f}_k$ and $\overline{f}_k$. We highlight that the binary variables z_k and u_k do not appear under the maximization in Equation (1). This is because these are variables of the full DRO formulation. More details are available in the following subsections.

Constraint (6) ensures that the net power balance in each node is equal to 0.

For the sake of clarity, we can express the optimization problem in (1)-(6) with its compact matrix formulation, shown in Equations (7)-(8)

$$\max_{\mathbf{h}} \mathbf{b}^T\mathbf{h} \tag{7}$$

subject to:

$$\mathbf{Rh} \le \mathbf{q} - \mathbf{Tu} - \mathbf{Hz} - \mathbf{z}^T\mathbf{Du} \tag{8}$$

where $\mathbf{h} = \{\mathbf{p}, \mathbf{d}, \mathbf{f}\}$ are the continuous variables, $\mathbf{z}$ and $\mathbf{u}$ are the binary variables, $\mathbf{b}$ is the vector with the objective function coefficients, $\mathbf{R}$, $\mathbf{T}$, $\mathbf{H}$, and $\mathbf{D}$ are the matrices with the constraints coefficients, and $\mathbf{q}$ is the vector with the constraints constants. Equation (7) represents Equation (1), while Constraint (8) represents Constraints (2)-(6).

2.2 Ambiguity set without endogenous uncertainty

In this work, the uncertainty represents the conditional occurrence probability of the feasible failure scenarios contained in the set of feasible failure scenarios $\mathcal{F}$, where the condition is that at least one component has failed. We assume that only edges can fail and we rely on a set of feasible attack/failure scenarios based on the maximum

number of attacked/failed edges K_{att}. The set $\mathcal{F}$ is then defined as in Equation (9):

$$\mathcal{F} = \{\mathbf{u} \mid \{0,1\}^M, ||\mathbf{u}||_1 \geq M - K_{att}\} \tag{9}$$

where K_{att} is the maximum number of lines that can fail and $\mathbf{u}$ is a vector containing the functional state of each line. The k^{th} element of $\mathbf{u}$ is $u_k = 1$ if line k functional, and $u_k = 0$ if it is attacked/failed. Assuming that the scenarios in $\mathcal{F}$ are mutually exclusive, and given the condition that at least one of the lines is attacked/failed, the conditional probability of each scenario to have occurred is defined by a multinomial distribution.

In a DRO approach, the probability distribution of the uncertainty is unknown and ambiguous, and it is contained within an ambiguity set $\mathcal{D}$, which contains all the distribution that can describe the uncertainty and respect some given conditions. Similarly to other existing works, such as Babaei et al. (2020), we rely on a moment-based ambiguity set which restricts the marginal probability of each line k to be attacked/failed between 0 and an upper bound π_k^{max}. The ambiguity set without endogenous uncertainty is defined in Equation (10):

$$\mathcal{D} = \{\mathbb{P} \in \mathcal{P}(\mathcal{F}) : 0 \leq \mathbb{E}_{\mathbb{P}}[1 - \mathbf{u}] \leq \boldsymbol{\pi^{max}}\} \tag{10}$$

where $\mathcal{P}(\mathcal{F})$ defines the set of all probability distributions on a σ-algebra of $\mathcal{F}$, and $\boldsymbol{\pi}^{max}$ is the vector containing the upper bounds of the conditional marginal failure probabilities of each power transmission line. We highlight that the upper bounds $\boldsymbol{\pi}^{max}$ can be estimated from historical data and reliability analysis (Babaei et al., 2020).

2.3 Ambiguity set with endogenous uncertainty

When we protect a power line, we assume that it becomes invulnerable. As a consequence, we can assume that the conditional marginal failure probability of each protected line is 0. The protection plan can, thus, modify the ambiguity set and it represents a source of endogenous uncertainty. The decision-dependent ambiguity set $\mathcal{D}(\mathbf{z})$ is defined as in Equation (11):

$$\mathcal{D} = \{\mathbb{P} \in \mathcal{P}(\mathcal{F}) : 0 \leq \mathbb{E}_{\mathbb{P}}[1 - \mathbf{u}] \leq \boldsymbol{\pi^{max}}(\mathbf{z})\} \tag{11}$$

where the decision-dependent upper bounds are defined as in Equation (12):

$$\boldsymbol{\pi^{max}}(\mathbf{z}) = \frac{M\left((1-\mathbf{z})^T \boldsymbol{\pi^{max}}\right)^T}{M - K_{def}} \tag{12}$$

where M is the number of edges, $\mathbf{z}$ is a vector containing the protection plan ($z_k = 1$ if line k is protected, $z_k = 0$ otherwise) and K_{def} is the parameter, set by the decision-maker, that defines the number of protected lines. The term $M/(M - K_{def})$ is a scaling factor for each upper bound π_k^{max}. In fact, if line k is protected, all the scenarios implying line k to be failed can not occur, and the conditional probability of the other scenarios increase. Consequently, the upper bounds $\boldsymbol{\pi^{max}}$ need to be increased.

2.4 DRO formulations

We define two DRO formulations, one without endogenous uncertainty, and one with endogenous uncertainty, where the ambiguity set is dependent on the protection plan $\mathbf{z}$.

The DRO without endogenous uncertainty corresponds to Equation (13):

$$\max_{\mathbf{z}} \min_{\mathbb{P} \in \mathcal{D}} \mathbb{E}_{\mathbb{P}}\left[Q(\mathbf{z}, \xi)\right] \tag{13}$$

subject to (10) and (14):

$$\sum_{k \in E} z_k = K_{def} \tag{14}$$

and where the recourse function $Q(\mathbf{z}, \xi)$ corresponds to the optimization in (7) subject to (8).

The DRO model with endogenous uncertainty is defined as in (15):

$$\max_{\mathbf{z}} \min_{\mathbb{P} \in \mathcal{D}(\mathbf{z})} \mathbb{E}_{\mathbb{P}}\left[Q(\mathbf{z}, \xi)\right] \tag{15}$$

subject to (11) and (14).

In both the formulations, a decision maker seeks to identify the optimal protection plan $\mathbf{z}$ that maximizes the expected performance of the power network in disrupted conditions. The expected performance are computed considering the worst multinomial distribution contained in the ambiguity set $\mathcal{D}$ or $\mathcal{D}(\mathbf{z})$. In this work, we investigate the impact of considering the endogenous uncertainty within the optimization of the protection plan.

3. SOLUTION PROCEDURE

3.1 Reformulation

We recast the problem into a tractable form, based on the reformulation presented in Babaei et al. (2020), and we solve it with a column and constraint generation (C&CG) approach (Zeng and Zhao, 2013). For simplicity, we derive the reformulation only for the case with endogenous uncertainty. The reformulation for the case without endogenous uncertainty can be derive in a similar way.

For a fixed protection plan $\mathbf{z}^*$, the inner stage of (13) is equivalent to Equation (16) subject to (17) and (18):

$$\min_{\mathbb{P}} \int_{\mathcal{F}} Q(\mathbf{z}, \xi) \mathrm{d}\mathbb{P} \tag{16}$$

$$\int_{\mathcal{F}} \mathrm{d}\mathbb{P} = 1 \tag{17}$$

$$\int_{\mathcal{F}} (1 - u_k) \mathrm{d}\mathbb{P} \leq \pi_k^{max}(z_k), \quad \forall k \in E \tag{18}$$

where $\pi_k^{max}(z_k)$ is the k^{th} element of $\boldsymbol{\pi^{max}}(\mathbf{z})$, defined as in (19):

$$\pi_k^{max}(z_k) = \frac{M(1 - z_k)\pi_k^{max}}{M - K_{def}}. \tag{19}$$

Formulation(16)-(18) is convex in $\mathbb{P}$ and it satisfies Slater's conditions. The dual form of the inner level of (13) is shown in (20):

$$\max_{\alpha, \boldsymbol{\beta} \geq 0} \; -\alpha - \sum_{k \in E} \beta_k \pi_k^{max}(z_k) \tag{20}$$

subject to:

$$\alpha + \sum_{k \in E} \beta_k (1 - u_k^{(i)}) \geq -Q(\mathbf{z}, \xi) \quad \forall \mathbf{u}^{(i)} \in \mathcal{F} \tag{21}$$

and where α and $\boldsymbol{\beta}$ are the dual variables. The optimal value $\widehat{-\alpha}$ corresponds to (22):

$$\widehat{-\alpha} = \min_{\mathbf{u}^{(i)} \in \mathcal{F}} Q(\mathbf{z}, \xi) + \sum_{k \in E} \beta_k (1 - u_k^{(i)}). \tag{22}$$

By substituting (22) in (20) and explicitly expressing $Q(\mathbf{z}, \xi)$, we obtain the reformulation in (23):

$$\max_{\boldsymbol{\beta} \geq 0} \min_{\mathbf{u} \in \{0,1\}} \max_{\mathbf{h}} \mathbf{b}^T \mathbf{h} + \sum_{k \in E} \beta_k (1 - u_k - \pi_k^{max}(z_k)) \tag{23}$$

subject to (8) and (24):

$$\sum_{k \in E} u_k \geq M - K_{att} \tag{24}$$

where Constraint (24) defines the condition on the maximum failed lines K_{att} as denoted in the set $\mathcal{F}$ in Equation (9). For a fixed failure scenario $\mathbf{u}^*$ and fixed dual variables $\boldsymbol{\beta}^*$, the inner level of (23) is an LP problem. By taking its dual form and merging it with the outer and middle levels of (23) and the outer level of (15), we obtain the bilevel formulation in (25):

$$\max_{\mathbf{z}, \boldsymbol{\beta}} \min_{\boldsymbol{\lambda}, \mathbf{u}} (\mathbf{q} - \mathbf{T}\mathbf{u} - \mathbf{H}\mathbf{z} - \mathbf{z}^T \mathbf{D}\mathbf{u})^T \boldsymbol{\lambda} + \sum_{k \in E_{PN}} \beta_k (1 - u_k - \pi_k^{max}(z_k)) \tag{25}$$

subject to (14), (24) and (26):

$$\mathbf{R}^T \boldsymbol{\lambda} = \mathbf{b} \tag{26}$$

where $\boldsymbol{\lambda}$ represents the dual variables of (8). For simplicity, we report directly the compact matrix formulation. In this form, the problem can be recast and solved directly with a C&CG algorithm.

The reformulation for the case without endogenous uncertainty is obtained by simply substituting the decision-dependent upper bounds $\boldsymbol{\pi}^{\boldsymbol{max}}(\mathbf{z})$ with the upper bounds $\boldsymbol{\pi}^{\boldsymbol{max}}$.

3.2 Column&Constraint Generation algorithm

We report the main steps of the C&CG approach for the case with endogenous uncertainty. The main steps for the case without endogenous uncertainty can be derived by substituting the decision-dependent upper bounds $\boldsymbol{\pi}^{\boldsymbol{max}}(\mathbf{z})$ with the upper bounds $\boldsymbol{\pi}^{\boldsymbol{max}}$.

The C&CG algorithm is a traditional cutting-plane strategy. It involves the split of the original problem into a master problem and a subproblem which iteratively exchange binary decision variables. The optimal solution is found by following these steps:

(1) Set $i = 0$, upper bound $UB = \infty$, lower bound $LB = 0$ and $\mathcal{F}_{part} = \emptyset$.

(2) Solve the master problem in Equations (27)-(29). Obtain an optimal solution $\widehat{\rho}^{(i)}$, optimal variables $\widehat{\boldsymbol{\beta}}^{(i)}$, and an optimal protection plan $\widehat{\mathbf{z}}^{(i)}$. Update UB=min $\left(UB, \widehat{\rho}^{(i)}\right)$.

$$\max_{\rho, \mathbf{z}, \mathbf{h}, \boldsymbol{\beta}} \rho \tag{27}$$

subject to (14), (28) and (29): :

$$\rho \leq \mathbf{b}^T \mathbf{h}^{(i)} + \sum_{k \in E_{PN}} \beta_k (1 - u_k^{*(i)} - \pi_k^{max}(z_k)) \tag{28}$$

$$\mathbf{R}\mathbf{h}^{(i)} \leq \mathbf{q} - \mathbf{T}\mathbf{u}^{*(i)} - \mathbf{H}\mathbf{z} - \mathbf{z}^T \mathbf{D}\mathbf{u}^{*(i)} \tag{29}$$

where Constraints (28) and (29) are defined $\forall \mathbf{u}^{*(i)} \in \mathcal{F}_{part}$.

(3) Solve the subproblem in Equations (30) with $\widehat{\mathbf{z}}^{(i)}$=$\mathbf{z}^*$ and $\widehat{\boldsymbol{\beta}}^{(i)}$=$\boldsymbol{\beta}^*$. Obtain an optimal solution and an optimal attack plan $\widehat{\mathbf{u}}^{(i)}$. Update the lower bound LB.

$$\min_{\boldsymbol{\lambda}, \mathbf{u} \in \{0,1\}} (\mathbf{q} - \mathbf{T}\mathbf{u} - \mathbf{H}\mathbf{z}^* - \mathbf{z}^{*T} \mathbf{D}\mathbf{u})^T \boldsymbol{\lambda} + \sum_{k \in E_{PN}} \beta_k^* (1 - u_k - \pi_k^{max}(z_k)) \tag{30}$$

subject to (24) and (26).

(4) If $(UB - LB)/UB < 10^{-4}$, the current solution $\widehat{\mathbf{z}}^{(i)}$ corresponds to the optimal protection plan and the algorithm can be terminated. Otherwise, $\mathcal{F}_{part} = \mathcal{F}_{part} \cup \widehat{\mathbf{u}}^{(i)}$. Set $i \leftarrow i + 1$ and return to step 2.

4. NUMERICAL EXAMPLE

We perform a numerical experiment using an illustrative case study based on the topology of the IEEE 14-bus system (Iyambo and Tzoneva, 2007), shown in Figure 1. The power network is, thus, composed of 14 buses connected by 20 transmission lines. Each bus presents a requested power demand of 14 MW. Buses 1, 2, 3, 6, and 8 present a production capacity of 40 MW. The flow capacity of each line is 22 MW.

For each upper bound π_k^{max}, we assume a value of 0.2. We compute the optimal protection plan with and without endogenous uncertainty for values of K_{def} and K_{att} ranging from 1 to 4.

All the computations are performed on a desktop PC with a 3.20 GHz CPU and 32 GB RAM using the Python API of Gurobi 9.1 (Gurobi Optimization, LLC, 2021).

The results for different combinations of K_{def} and K_{att} are shown in Figure 2.

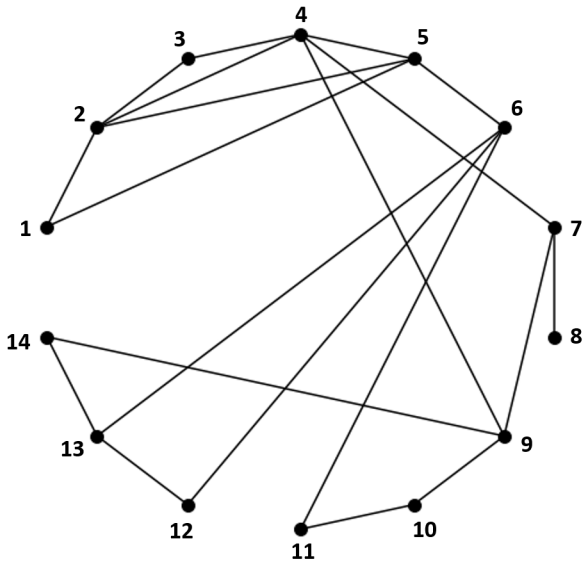

Fig. 1. Topology of the IEEE 14-bus system.

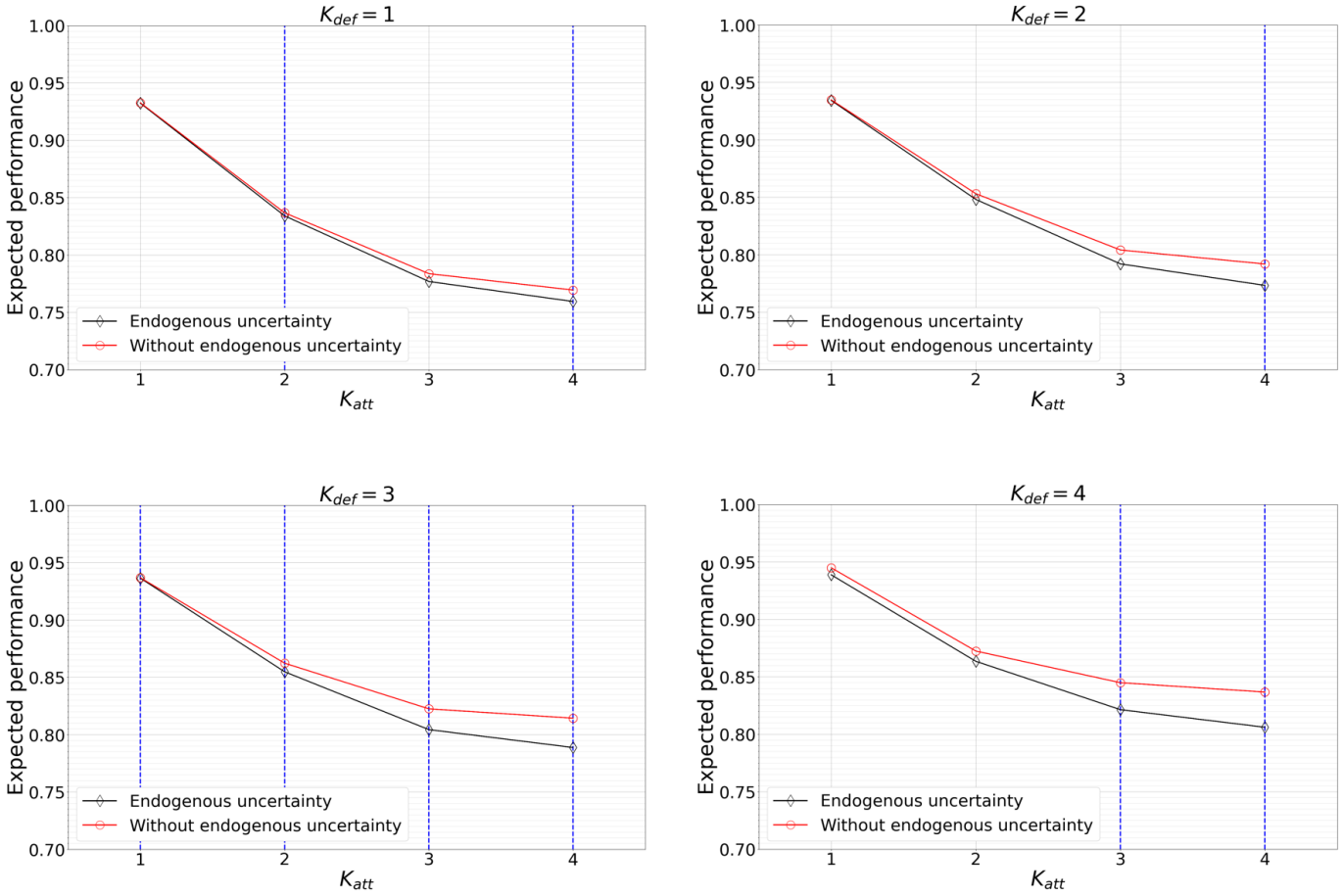

Fig. 2. Expected performance with and without considering endogenous uncertainty within the ambiguity set for different combinations of K_{def} and K_{att}. The blue vertical lines indicate different optimal protection plans between the cases with and without endogenous uncertainty and the same values of K_{def} and K_{att}.

Each plot in Figure 2 represents the results for different values K_{def}, which denote how many lines are protected. The x-axis in each plot represents the K_{att} values, which denote the maximum number of lines which are attacked/failed. The y-axis represents the expected performance of the power network. The expected performance is computed considering the worst distribution in the ambiguity set without endogenous uncertainty $\mathcal{D}$ (red curves), and the decision-dependent ambiguity set with endogenous uncertainty $\mathcal{D}(\mathbf{z})$ (black curves).

From these preliminary results, some important considerations can be drawn.

As it was expected, the expected performance in both the cases - with and without endogenous uncertainty - increases as K_{def} increases and K_{att} decreases.

In addition, we can notice that the expected performance with endogenous uncertainty is always lower than (or equal to) the expected performance without endogenous uncertainty. Moreover, the difference between the two cases increases as K_{def} and K_{att} increase. In fact, the maximum difference is reached for the case with $K_{def} = 4$ and $K_{att} = 4$, where the expected performance with and without endogenous uncertainty take a value of 0.8061 and 0.8374, respectively. In fact, as expected, changing the probability space from which the multinomial distribution is chosen has an impact on the expected performance in disrupted conditions.

Most importantly, we compare the optimal protection plans identified with and without considering the endogenous uncertainty. As it can be seen in Figure 2, in 9 out of the 16 combinations of K_{def} and K_{att} investigated, the protection plans for the two uncertainty cases are different (blue vertical lines). Moreover, the optimal protection plans identified without endogenous uncertainty are not guaranteed to be optimal when considering endogenous uncertainty. For example, with $K_{def} = 4$ and $K_{att} = 4$, an optimal solution for the case without endogenous uncertainty consists in protecting the lines with indices 2, 9, 10, and 14. This protection plan is suboptimal for the case with endogenous uncertainty, as it leads to expected performance of 0.8035, instead of 0.8061, which is attainable, for example, by protecting the lines with indices 2, 6, 9, and 13. In this specific case, the suboptimality leads to limited negative consequences in terms of expected performance in disrupted conditions with endogenous uncertainty. However, the negative impact of suboptimal solutions is strongly dependent on the case-study considered, and a thorough sensitivity analysis on size and parameters of the case-study should be performed. It should also be

highlighted that common optimal solutions, for the cases with and without endogenous uncertainty, can exist.

Including the endogenous uncertainty within the optimization problem has also an impact on the computational performance. In fact, the computational time when including the endogenous uncertainty increases considerably. For example, with $K_{def} = 4$ and $K_{att} = 4$, the case without endogenous uncertainty leads to a computational time of 3.2 seconds, while for the case with endogenous uncertainty the computational time is 218.7 seconds.

5. CONCLUSION

In this work, we proposed a distributionally robust approach with endogenous uncertainty for the protection plan of power networks. Preliminary results have highlighted the importance of including the endogenous uncertainty in optimal protection planning. We showed how including endogenous uncertainty through decision-dependent ambiguity sets impacts the expected performance and the optimal protection plan decision.

Further extensions of this work include: *i)* a more comprehensive sensitivity analysis of the DRO parameters, such as upper bounds $\boldsymbol{\pi}^{\boldsymbol{max}}$, parameters K_{def} and K_{att}, and scaling factor for the decision-dependent upper bounds, *ii)* a thorough comparison between the optimal protection plans identified with and without endogenous uncertainty, in order to quantify their difference in terms of expected performance in the two different probability spaces (with and without endogenous uncertainty), and *iii)* a detailed analysis of the computational performance of the optimization problem.

REFERENCES

Alvarado, D., Moreira, A., Moreno, R., and Strbac, G. (2018). Transmission network investment with distributed energy resources and distributionally robust security. *IEEE Transactions on Power Systems*, 34(6), 5157–5168.

Babaei, S., Jiang, R., and Zhao, C. (2020). Distributionally robust distribution network configuration under random contingency. *IEEE Transactions on Power Systems*, 35(5), 3332–3341.

Bagheri, A. and Zhao, C. (2019). Distributionally robust reliability assessment for transmission system hardening plan under *nk* security criterion. *IEEE Transactions on Reliability*, 68(2), 653–662.

Bellè, A., Abdin, A.F., Zeng, Z., Fang, Y.P., and Barros, A. (2021). A mathematical framework for the optimal coupling of interdependent critical infrastructures. URL `https://hal.archives-ouvertes.fr/hal-03446712`. Working paper or preprint.

Doan, X.V. (2022). Distributionally robust optimization under endogenous uncertainty with an application in retrofitting planning. *European Journal of Operational Research*, 300(1), 73–84.

Fang, Y.P. and Zio, E. (2019). An adaptive robust framework for the optimization of the resilience of interdependent infrastructures under natural hazards. *European Journal of Operational Research*, 276(3), 1119–1136.

Fang, Y. and Sansavini, G. (2017). Optimizing power system investments and resilience against attacks. *Reliability Engineering & System Safety*, 159, 161–173.

Garcia Tapia, A., Suarez, M., Ramirez-Marquez, J.E., and Barker, K. (2019). Evaluating and visualizing the economic impact of commercial districts due to an electric power network disruption. *Risk Analysis*, 39(9), 2032–2053.

González, A.D., Dueñas-Osorio, L., Sánchez-Silva, M., and Medaglia, A.L. (2016). The interdependent network design problem for optimal infrastructure system restoration. *Computer-Aided Civil and Infrastructure Engineering*, 31(5), 334–350.

Gurobi Optimization, LLC (2021). Gurobi Optimizer Reference Manual. URL `https://www.gurobi.com`.

Iyambo, P. and Tzoneva, R. (2007). Transient stability analysis of the ieee 14-bus electric power system. In *AFRICON 2007*, 1–9. IEEE.

Luo, F. and Mehrotra, S. (2020). Distributionally robust optimization with decision dependent ambiguity sets. *Optimization Letters*, 14(8), 2565–2594.

Noyan, N., Rudolf, G., and Lejeune, M. (2018). Distributionally robust optimization with decision-dependent ambiguity set. *Optimization Online*.

Noyan, N., Rudolf, G., and Lejeune, M. (2021). Distributionally robust optimization under a decision-dependent ambiguity set with applications to machine scheduling and humanitarian logistics. *INFORMS Journal on Computing*.

Ouyang, M. (2017). A mathematical framework to optimize resilience of interdependent critical infrastructure systems under spatially localized attacks. *European Journal of Operational Research*, 262(3), 1072–1084.

Ouyang, M. and Fang, Y. (2017). A mathematical framework to optimize critical infrastructure resilience against intentional attacks. *Computer-Aided Civil and Infrastructure Engineering*, 32(11), 909–929.

Rahimian, H. and Mehrotra, S. (2019). Distributionally robust optimization: A review. *arXiv preprint arXiv:1908.05659*.

Yuan, W., Zhao, L., and Zeng, B. (2014). Optimal power grid protection through a defender–attacker–defender model. *Reliability Engineering & System Safety*, 121, 83–89.

Zeng, B. and Zhao, L. (2013). Solving two-stage robust optimization problems using a column-and-constraint generation method. *Operations Research Letters*, 41(5), 457–461.

Zhang, G., Zhang, F., Zhang, X., Wu, Q., and Meng, K. (2020). A multi-disaster-scenario distributionally robust planning model for enhancing the resilience of distribution systems. *International Journal of Electrical Power & Energy Systems*, 122, 106161.

Zhao, C. and Jiang, R. (2017). Distributionally robust contingency-constrained unit commitment. *IEEE Transactions on Power Systems*, 33(1), 94–102.