


# Electric power system security: the case for an integrated cyber-physical risk management framework

Efthymios Karangelos<sup>a,\*</sup>, Louis Wehenkel<sup>b</sup> 

<sup>a</sup> School of Electrical and Electronic Engineering, University College Dublin, Dublin, Ireland

<sup>b</sup> Department of Electrical Engineering and Computer Science, Université de Liège, Liège, Belgium

## ARTICLE INFO

### Keywords:

Cyber-physical systems  
Electric power system risk management  
Cyber-physical co-optimization  
Tri-level optimization under uncertainty

## ABSTRACT

This paper concerns the security of the electric power transmission grid facing the threat of malicious cyber-physical attackers. We posit that there is no such thing as perfectly effective cyber-security. Rather, any cyber-security measure comes with the possibility that a highly skilled attacker could (eventually find a way to) bypass it. On these grounds, we explicitly represent the exogenous uncertainty regarding the ability of an external adversary to overcome the implemented cyber-security measures. We further formulate a tri-level decision making problem seeking to co-optimize preventive physical and cyber-security measures under the aforementioned uncertainty. Preventive physical security measures refer to the *ex-ante* procurement of reserve capacity, which translates into ramping restrictions in real-time. Cyber-security measures refer to updating the firewall rules so as to impede an intruder from taking over the cyber infrastructure of the grid and disconnecting power generators and transmission branches. We adopt standard assumptions to formalize the inner optimization problems corresponding to the cyber-physical attacker and power grid operator and focus on uncertainty management at the uppermost level of the problem. Our findings establish that physical- and cyber-security measures are non-exchangeable complements in keeping the power grid operation secure, underscoring the potential benefits of selecting them in an integrated manner.

## 1. Introduction

Advanced digital technology is vital for the functioning of today's electric power system. Indeed, the modern power grid relies on its cyber sub-system for most monitoring, protection, control and optimization functionalities and operates as a complex *cyber-physical* system. This essential technology however may also pose a severe vulnerability. The risk that a malicious adversary may intentionally exploit the power grid's cyber sub-system so as to disrupt electricity supply has unfortunately been validated in practice [1]. As a result electric power system cyber-security is nowadays recognized as critical and there is a growing development of standards, rules and regulations [2].

### 1.1. Related literature

The question of securing the grid operation while facing the threat of an intelligent cyber-physical attacker is actively pursued by the research community. Most research works rely on multi-level optimization to explicitly model the interrelated decisions of malicious cyber-physical

attackers and power grid so-called “defenders” (*i.e.*, planners and operators) [3]. Bi-level optimization problems are cast from the perspective of a cyber-physical attacker seeking to optimize her attack vector while explicitly accounting for the optimal reaction of the power grid operator. Such models can be used in an assessment context, so as to identify which vulnerabilities a cyber-physical attacker could target to exploit. Tri-level problems are cast from the (preceding) perspective of a preventive planner. The goal here is to identify preventive measures to counteract a cyber-physical attacker's optimal strategy, in turn maximizing her own objective while anticipating the grid operator's reaction.

Indicatively, He et al. [4] consider the problem of facing a cyber-physical attacker capable of both injecting false data and physically disconnecting transmission lines. An innovative feature of this work is the allocation of Distributed Generators (DGs) in the post-attack stage, translating mathematically into binary variables at the lowest-level optimization problem modeling power system operation. The considered preventive measures only act on the physical infrastructure of the grid

\* Corresponding author.

Email addresses: [efthymios.karangelos@ucd.ie](mailto:efthymios.karangelos@ucd.ie) (E. Karangelos), [l.wehenkel@uliege.be](mailto:l.wehenkel@uliege.be) (L. Wehenkel).

Nomenclature		Variables	
<b>Sets</b>		$p_g$	Active power dispatch of generator $g \in \mathcal{G}$ in the pre-attack state – by the planner
$\mathcal{A}$	Cyber-physical attackers with advanced capabilities ( <i>i.e.</i> , able to by-pass cyber-security measures)	$r_g$	procured reserve capacity from generator $g \in \mathcal{G}$ – by the planner
$\mathcal{B}$	cyber-physical attackers with basic capabilities ( <i>i.e.</i> , unable to by-pass cyber-security measures)	$z_n$	binary variable, taking the value of 1 to indicate the update of the firewall rules at bus $n \in \mathcal{N}$ – by the planner
$\mathcal{G}$	generating units	$w_{n,i}$	binary variable, taking the value of 0 to indicate successful intrusion into the cyber sub-system of bus $n \in \mathcal{N}$ – by cyber-physical attacker $i \in \mathcal{A} \cup \mathcal{B}$
$\mathcal{N}$	buses	$u_{t,i}$	binary variable, taking the value of 0 to indicate the disconnection of transmission branch $t \in \mathcal{T}$ – by cyber-physical attacker $i \in \mathcal{A} \cup \mathcal{B}$
$\mathcal{T}$	transmission branches	$y_{g,i}$	binary variable, taking the value of 0 to indicate the disconnection of generator $g \in \mathcal{G}$ – by cyber-physical attacker $i \in \mathcal{A} \cup \mathcal{B}$
<b>Parameters</b>		$up_{g,i}$	upward active power redispatch of generator $g \in \mathcal{G}$ in the post-attack state corresponding to attacker $i \in \mathcal{A} \cup \mathcal{B}$ – by the power grid operator
$C_g^D$	Dispatch cost of generator $g \in \mathcal{G}$ – in monetary units/MWh (mu/MWh)	$dp_{g,i}$	downward active power redispatch of generator $g \in \mathcal{G}$ in the post-attack state corresponding to attacker $i \in \mathcal{A} \cup \mathcal{B}$ – by the power grid operator
$C_g^R$	reserve capacity cost of generator $g \in \mathcal{G}$ – in mu/MW	$l_{n,i}$	involuntary load shedding at node $n \in \mathcal{N}$ in the post-attack state corresponding to attacker $i \in \mathcal{A} \cup \mathcal{B}$ – by the power grid operator
$C_g^{RD}$	reserve deployment ( <i>i.e.</i> , redispatch) cost of generator $g \in \mathcal{G}$ – in mu/MWh	$f_{t,i}$	active power flow through branch $t \in \mathcal{T}$ in the post-attack state corresponding to attacker $i \in \mathcal{A} \cup \mathcal{B}$
$C^L$	value of lost load – in mu/MWh	$\theta_{n,i}$	voltage angle at node $n \in \mathcal{N}$ in the post-attack state corresponding to attacker $i \in \mathcal{A} \cup \mathcal{B}$
$C^Z$	cost of updating firewall rules – in mu/device		
$D_n$	active power demand at bus $n \in \mathcal{N}$		
$F_t$	capacity of transmission branch $t \in \mathcal{T}$		
$P_g$	active power generation capacity of generator $g \in \mathcal{G}$		
$W_i$	resource limit of cyber-physical attacker $i \in \mathcal{A} \cup \mathcal{B}$		
$X_t$	reactance of transmission branch $t \in \mathcal{T}$		
$Z$	resource limit of cyber-physical planner		
$\gamma_{g,n}$	auxiliary parameter, taking the value of 1 to indicate the bus at which generator $g \in \mathcal{G}$ is connected		
$\pi_i$	manifestation probability of cyber-physical attacker $i \in \mathcal{A} \cup \mathcal{B}$		
$\tau_{t,n}$	auxiliary parameter taking the value of 1(–1) to indicate the sending(receiving) buses of branch $t \in \mathcal{T}$		

by hardening transmission components. In contrast, Lai et al. [5] focus on choosing cyber-security preventive measures only. These authors address the problem of placing firewalls so as to block unauthorized access to protection relays that could be exploited by a cyber-physical attacker seeking to cause a cascading failure. Qin et al. [6] also consider the placement of firewalls to impede unauthorized access to the grid's relays. The novel feature of this work is the consideration of detailed substation topologies.

Both groups of authors in [5,6] adopt the assumption that firewalls are perfectly effective countermeasures, blocking any cyber-physical attacker's unauthorized access with certainty. While such an assumption is quite common in the power systems literature, it may be regarded as questionable in the cyber-security context. In practice, the capabilities of cyber-physical attackers continuously evolve, cyber-security measures have specific protection scopes, while yet undiscovered (a.k.a., zero-day) vulnerabilities may always exist. In line with this, Zhang et al. [7] consider the problem of allocating cyber-security devices that only reduce rather than completely eliminate the probability of a successful cyber-physical attack against the grid's transmission branches. To the best of our knowledge, this is the only research work acknowledging the uncertainty of the effectiveness of cyber-security measures in the context of electric power system security.

## 1.2. Paper scope & contributions

In this paper we also take the pragmatic viewpoint that cyber-security measures should not be considered as 100 % effective. We introduce a novel model for such uncertainty as a probability distribution of exogenous cyber-physical attackers differentiated in terms of their capability to overcome or not the cyber-security measures

protecting the power grid. On this basis, we investigate the stakes for an integrated *cyber-physical risk management* framework wherein (i) preventive cyber-security measures are realistically considered to reduce rather than eliminate the probability of a successful cyber-physical attack, and (ii) these are chosen while acknowledging the synergy with physical-security measures. To do so, we formulate a novel tri-level optimization problem seeking to choose in preventive mode **both** physical- and cyber-security measures under uncertainty regarding the effectiveness of the latter.

Without loss of generality, we focus on the particular threat posed by cyber-physical attackers seeking to intrude into the power grid's substations and disconnect power generators and transmission branches (*i.e.*, lines and transformers) so as to maximize operational costs, including the cost associated with involuntary load shedding. Anticipating such threat, a preventive planner may choose to update the security rules of the firewalls protecting the grid's substations and/or to procure reserves that can be deployed in the aftermath of a successful attack. The overarching framework introduced here, which incorporates uncertainty modeling of attacker capabilities and the integrated selection of cyber- and physical-security measures, is not only relevant for this particular threat but indeed also transferable to alternative threat instances.

To solve the resulting problem, under standard assumptions, we rely on a modified version of the classic *Column & Constraint Generation* (C&CG) algorithm by Ref. [8]. The essential modification to the standard algorithm is an adaptation to the stochastic problem at hand and building upon the work of Xiang and Wang [9]. Results obtained on two benchmark systems, namely the 24-bus version of the IEEE Reliability Test System [10] and the IEEE 118-bus system [11], showcase the complex interrelations between physical- and cyber-security measures. In particular, we find that the consideration of an adversary with the ability



Fig. 1. Cyber-physical power grid security actors.

to overcome cyber-security measures as a credible alternative calls for stronger physical-security measures which in turn reorient the priorities for implementing cyber-security measures.

## 2. Assumptions, models & methods

### 2.1. Problem description

We adopt a planning perspective and consider the problem of protecting the electric power system's operation against the threat of malicious cyber-physical attacks. Fig. 1 illustrates the independent decision-making actors involved in this problem.

The central actor is the *cyber-physical attacker* (a.k.a. *attacker* for brevity) who may seek to compromise the functionalities of the power grid's cyber sub-system so as to disrupt electricity supply. The left-most actor is the so-called *cyber-physical planner* (a.k.a. *planner* for brevity). Anticipating the threat posed by the cyber-physical attacker, she may act on (i) the power grid's cyber sub-system to ensure that it cannot be breached in a way causing an unacceptable electricity supply and/or (ii) the physical power grid so as to ensure that it can maintain an acceptable electricity supply service in spite of the maloperation of its cyber sub-system. Finally, the right-most actor is the *power grid operator* (a.k.a. *operator* for brevity) who, in the aftermath of any cyber-physical attack, must operate the (possibly compromised) electric power system.

We explicitly acknowledge that cyber-security measures have specific protection scopes and are also susceptible to unknown vulnerabilities (e.g., zero-day vulnerabilities), while the capabilities of cyber-physical attackers are constantly evolving. In short, with any cyber-security measure comes the risk that a sufficiently capable cyber-physical attacker may still find a way to compromise the protected functionality. To account for this feature, we consider a discrete set of alternative cyber-physical attackers differentiated in terms of available resources (i.e., number of cyber sub-system components an attacker can compromise) and capabilities to overcome the applicable cyber-security measures. We further associate a probability measure to this set so as to represent the threat landscape as a probability distribution over credible cyber-physical attackers.

Any credible cyber-physical attacker is assumed to be an independent decision-maker, choosing her attack vector upon observing the active cyber- and physical- security measures and while anticipating the optimal reaction of the power grid operator. The corresponding decision-making problem can be cast as a bi-level optimization problem. Similarly, the cyber-physical planner is considered to be an independent decision-maker choosing her risk management strategy while anticipating the optimal reaction of each alternative cyber-physical attacker. Under information symmetry assumptions, the respective decision-making problem can be cast as a tri-level optimization problem under uncertainty.

We consider a set-up wherein cyber-physical attackers may seek to intrude into digital substations<sup>1</sup> and disconnect power generators and

<sup>1</sup> For ease of exposition, and without loss of generality, we henceforth assume that any single bus corresponds to a distinct digital substation and use the terms bus and substation interchangeably to describe the assumed threat.

transmission lines so as to maximize the power grid operational costs. These latter costs include generation redispatching costs and costs of involuntary load shedding and are determined as per the optimal reaction of the power grid operator to the realized grid status. A cyber-physical attacker's strategy is constrained by a resource limit, specifying the number of substations she may simultaneously gain unauthorized access to. Once infiltrating a substation, the cyber-physical attacker can disconnect generators and transmission branches.

Anticipating this uncertain threat, a cyber-physical planner may choose to update the rules of cyber-security devices blocking unauthorized access (a.k.a., firewalls) to the grid's substations. Additionally, this actor should also choose the system reserve capacity which constrains the redispatching of generating units in real-time operation. The objective of the cyber-physical planner is to jointly minimize the cost of base-case operation (i.e., generation dispatch) and cyber-physical security measures (i.e., updating firewall rules and reserve procurement) along with the expected cost of the power grid operation (i.e., expected generation redispatching and load shedding).

### 2.2. Mathematical formulation

Relying on the DC power flow approximation to keep the model of the grid physics linear, we express the trilevel optimization problem of the cyber-physical planner as in (1)–(22).

The top row of the planner's objective (1) includes all 1st-stage costs, specifically the cost of the pre-attack generation dispatch, the cost of reserve procurement, and the cost of mobilizing resources at the physical locations of the different buses of the system to update the firewall rules (3rd term). These 1st-stage costs are to be jointly minimized along with the expectation of the power grid operation cost over the post-attack states while facing the threat described as the postulated set of cyber-physical attackers. The probability weighted summation of these costs, as shown in the bottom row of (1), includes scenario-specific redispatching and load shedding costs.

Inequality constraint (2) is a resource constraint limiting the total number of firewalls that the cyber-physical planner can update with new security rules in the considered lead-time. Constraints (3) and 4) define the feasible region of the planner's generation variables, referring to the pre-attack dispatch and reserve capacity. Constraint (5) models the choice of where to update firewall rules as a binary decision. The group of constraints (6)–(8) enforces that the pre-attack dispatch implies no loss of load and conforms to the branch thermal ratings as per the DC power flow approximation. While the focus is on the considered cyber-physical threat, this group of constraints is included to ensure that reserve procurement is not unrealistic with respect to the pre-attack operation of the system.

For any cyber-physical attacker  $i \in \mathcal{A} \cup \mathcal{B}$  the objective function to be maximized is the cost of power grid operation, including attack-induced reserve deployment and load shedding (9). Constraint (10) defines an attacker's resources in terms of the number of buses she may simultaneously infiltrate. Inequality constraints (11) are only valid for so-called cyber-physical attackers with *basic capabilities*. Attackers with basic capabilities are assumed to be unable to overcome the implemented cyber-security measures.<sup>2</sup> Therefore inequalities (11) enforce that any such actor may not infiltrate any bus that is protected by means of updated firewall rules. The choice to intrude into the cyber sub-system of a specific bus is modeled through a binary decision variable (12). Inequality constraints (13) link the status of transmission branches and the potential intrusion of their associated ("sending/receiving") end-buses. Specifically, these constraints enforce that a branch can only be disconnected if the cyber-physical attacker has managed to successfully intrude either its "sending" or "receiving" bus. We use binary decision

<sup>2</sup> The complementary type of cyber-physical attacker that can overcome cyber-physical measures is referred to as an attacker with *advanced capabilities*.

variables to model branch disconnection decisions (14). Equality constraints (15) model the disconnection of the generating units connected at a breached bus.

Expressions (16)–(22) model the operator’s reaction to the realized attack and, under information symmetry assumptions, the impact of the attack on the power grid operation cost. Constraints (17) and (18) limit the reserve energy to be used by the operator according to the capacity procured by the cyber-physical planner and, of course, the connection status of the respective unit. Equality constraints (19) express the nodal power balance for the post-attack system state, featuring the additional decision variables corresponding to the deployment of reserve energy and the use of involuntary load shedding. Notice here the product of the pre-attack dispatch decided by the planner and the attacker’s decision to compromise the concerned bus. In (20), the non-negative load shedding variable is upper bounded by the respective nodal load demand. Equality constraints (21) define the post-attack power flow as the product of the branch disconnection variable and the fraction of the voltage angle difference divided by the branch reactance. Finally, inequalities (22) impose the branch thermal ratings.

$$\min_{\mathbf{p}, \mathbf{r}, \mathbf{z}} \left\{ \sum_{g \in \mathcal{G}} (C_g^R \cdot r_g + C_g^D \cdot p_g) + C^Z \cdot \sum_{n \in \mathcal{N}} z_n + \sum_{i \in \mathcal{A} \cup \mathcal{B}} \pi_i \cdot \left( \sum_{g \in \mathcal{G}} C_g^{RD} \cdot up_{g,i} + C^L \cdot \sum_{n \in \mathcal{N}} l_{n,i} \right) \right\}, \quad (1)$$

$$\sum_{n \in \mathcal{N}} z_n \leq Z, \quad (2)$$

for all generators  $g \in \mathcal{G}$  :

$$p_g + r_g \cdot \Delta T \leq P_g \cdot \Delta T, \quad (3)$$

$$p_g, r_g \geq 0, \quad (4)$$

for all buses  $n \in \mathcal{N}$  :

$$z_n \in \{0, 1\}, \quad (5)$$

$$\sum_{g \in \mathcal{G}} \gamma_{g,n} \cdot p_g = \sum_{i \in \mathcal{T}} \tau_{i,n} \cdot f_{t,0} + D_n, \quad (6)$$

for all branches  $t \in \mathcal{T}$  :

$$f_{t,0} = (1/X_t) \sum_{n \in \mathcal{N}} \tau_{t,n} \theta_{n,0}, \forall t \in \mathcal{T}, \quad (7)$$

$$-F_t \leq f_{t,0} \leq F_t, \quad (8)$$

for all cyber-physical attackers  $i \in \mathcal{A} \cup \mathcal{B}$  :

$$\max_{\mathbf{v}, \mathbf{w}, \mathbf{y}} \left\{ \sum_{g \in \mathcal{G}} C_g^{RD} \cdot up_{g,i} + C^L \cdot \sum_{n \in \mathcal{N}} l_{n,i} \right\}, \quad (9)$$

$$\sum_{n \in \mathcal{N}} (1 - w_{n,i}) \leq W_i, \quad (10)$$

for all buses  $n \in \mathcal{N}$  :

$$w_{n,i} \geq z_n, \text{ iff } i \in \mathcal{B}, \quad (11)$$

$$w_{n,i} \in \{0, 1\}, \quad (12)$$

for all branches  $t \in \mathcal{T}$  :

$$v_{t,i} \geq \sum_{n \in \mathcal{N}} |\tau_{t,n}| \cdot w_{n,i} - 1, \quad (13)$$

$$v_{t,i} \in \{0, 1\}, \quad (14)$$

for all generators  $g \in \mathcal{G}$  :

$$y_{g,i} = \sum_{n \in \mathcal{N}} \gamma_{g,n} \cdot w_{n,i}, \quad (15)$$

$$\min_{\mathbf{up}, \mathbf{dp}, \mathbf{l}} \left\{ \sum_{g \in \mathcal{G}} C_g^{RD} \cdot up_{g,i} + C^L \cdot \sum_{n \in \mathcal{N}} l_{n,i} \right\}, \quad (16)$$

for all generators  $g \in \mathcal{G}$  :

$$0 \leq up_{g,i} \leq r_g \cdot \Delta T \cdot y_{g,i}, \quad (17)$$

$$0 \leq dp_{g,i} \leq p_g, \quad (18)$$

for all nodes  $n \in \mathcal{N}$  :

$$\sum_{g \in \mathcal{G}} \gamma_{g,n} (p_g \cdot w_{n,i} + up_{g,i} - dp_{g,i}) = \sum_{i \in \mathcal{T}} \tau_{i,n} f_{t,i} + (D_n - l_{n,i}), \quad (19)$$

$$0 \leq l_{n,i} \leq D_n, \quad (20)$$

for all branches  $t \in \mathcal{T}$  :

$$f_{t,i} = (v_{t,i}/X_{t,i}) \sum_{n \in \mathcal{N}} \tau_{t,n} \theta_{n,i}, \quad (21)$$

$$-F_t \leq f_{t,i} \leq F_t, \forall t \in \mathcal{T}, \quad (22)$$

where  $\Delta T = 1 \text{ hr}$ .

### 2.3. Algorithmic solution approach

We rely on a modified version of the *Column & Constraint Generation* (C&CG) introduced in [8] to solve tri-level optimization problem (1)–(22). To the best of our knowledge, Yuan et al. [12] first applied the C&CG algorithm in the context of electric power systems security management to solve the *deterministic* optimization problem of choosing which transmission branches to protect when facing an attacker with the ability to physically disconnect transmission branches. Xiang and Wang [9] used a modified version of this algorithm in a stochastic setting, considering the resources of the attacker as uncertain.

In order to briefly explain the C&CG algorithm, let us first notice that constraints (10) and (12)–(15) *implicitly* form a finite *attack surface* consisting of distinct attack vector scenarios (*i.e.*, substations to intrude and associated branches & generators to disconnect) per cyber-physical attacker  $i \in \mathcal{A} \cup \mathcal{B}$ . We will henceforth denote the attack surface for any single cyber-physical attacker as  $S_i$ . In problem (9)–(22), any cyber-physical attacker essentially chooses the attack vector with the maximal impact on the power grid operation cost out of her corresponding finite set  $S_i$ . This choice of course depends on the cyber-physical planner’s decisions protecting the power grid operation. These latter decisions in turn depend on the attack vectors that the cyber-physical planner anticipates the alternative attackers to potentially launch. The idea then is to iteratively update the planner’s decisions while generating for each alternative cyber-physical attacker  $i$  a growing sequence  $S_i^k$  of subsets containing maximal-impact scenarios chosen from the attack vectors  $S_i$  that can be launched by this attacker. In other words, *explicitly* form finite subsets of attack vector scenarios that would allow the planner to make “similar” decisions as the decisions that would be made while facing the complete attack surface  $\cup_{i \in \mathcal{A} \cup \mathcal{B}} S_i$ . By way of “similar” decisions, we refer here to near-optimal decisions that imply practically equal total costs (*i.e.*, base case dispatch costs, reserve procurement costs and expected impact of the cyber-physical attack on the power grid operation).

Fig. 2 presents the flowchart of the implemented CC&G algorithm. In every iteration  $k$ , given the current finite subsets  $S_i^k$ , a *Single-level Master Problem* is solved to update the decisions of the cyber-physical planner regarding the updating of firewall rules ( $\mathbf{z}^k$ ) as well as the dispatch and reserve capacity of generating units ( $\mathbf{p}^k, \mathbf{r}^k$ ). The solution to this problem also provides an updated lower bound for the total cost which is optimized while taking into consideration subsets  $S_i^k$ . Given the updated decisions of the cyber-physical planner ( $\mathbf{z}^k, \mathbf{p}^k, \mathbf{r}^k$ ), we then solve a *Bi-level Sub-Problem* per cyber-physical attacker  $i \in \mathcal{A} \cup \mathcal{B}$  to update her attack vector ( $\mathbf{v}_i^k, \mathbf{w}_i^k, \mathbf{y}_i^k$ ). Along with the updated attack vectors, the resolution of all bi-level problems also allows us to compute an updated estimation for the 2nd stage costs expressing the expected impact of the potential attacks. We add this updated estimation to the cost of the updated 1st stage decisions ( $\mathbf{z}^k, \mathbf{p}^k, \mathbf{r}^k$ ) to update an upper bound for the total cost. While the difference between the upper and lower bounds is not within the specified tolerance, the attack surface of any alternative cyber-physical attacker is updated as in  $S_i^{k+1} \leftarrow S_i^k \cup \{(\mathbf{v}_i^k, \mathbf{w}_i^k, \mathbf{y}_i^k)\}$ .

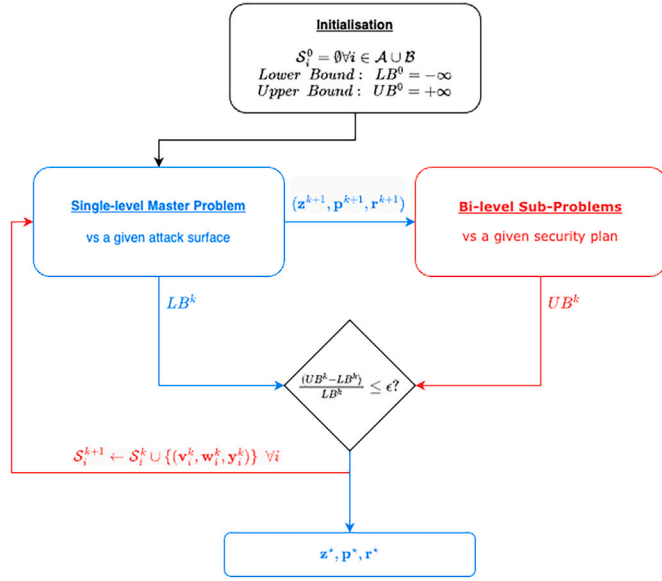


Fig. 2. Column &amp; constraint generation flowchart.

### 2.3.1. Single-level master problem: vs a given attack surface

At every iteration of the algorithm, and given an updated attack surface  $\cup_{i \in \mathcal{A} \cup \mathcal{B}} S_i^k$ , we formulate a single-level scenario sub-problem that corresponds to tri-level problem (1)–(22). To do so, we first introduce attack-vector-specific 2nd stage variables  $up_{g,i}^s$  and  $I_{n,i}^s$  per cyber-physical attacker  $i$  and attack vector  $s$  and then replace the 2nd row of (1) with:

$$+ \sum_{i \in \mathcal{A} \cup \mathcal{B}} \pi_i \cdot \eta_i \left. \right\},$$

and the maximization operation (9) with:

$$\eta_i \geq \sum_{g \in \mathcal{G}} C_g^{RD} \cdot up_{g,i}^s + C^L \cdot \sum_{n \in \mathcal{N}} I_{n,i}^s \quad \forall i \in \{\mathcal{A} \cup \mathcal{B}\}, \forall s \in S_i^k.$$

Constraints (10)–(12), (15), (17)–(20), (22) can be directly re-written per cyber-physical attacker and attack vector. Further, the operator's objective function (16) can be omitted since this actor and the cyber-physical planner are cooperating towards the same interest. The only noteworthy aspect is the treatment of transmission branch constraints (13) and (14) along with (21). We recall here that (13) and (14) allow a cyber-physical attacker to disconnect a transmission branch provided that she has intruded into a substation including any one of its end-buses. In the single-level reformulation, and while considering the finite attack vector scenario subsets  $\cup_{i \in \mathcal{A} \cup \mathcal{B}} S_i^k$ , we must enforce that any transmission branch under attack will be disconnected if any one of its end-buses is attacked and not successfully protected (either by relying on outdated firewall rules or by facing an attacker that can by-pass the updated rules). To do so, we use the following expressions:

For all cyber-physical attackers with **basic** capabilities  $i \in \mathcal{B}$ , attack vector scenarios  $s \in S_i^k$  and **attacked** transmission branches  $t \in \mathcal{T}: v_{t,i}^s = 0$ :

$$f_{t,i}^s \leq M \cdot w_{n_s(t),i}^s \quad (23)$$

$$f_{t,i}^s \geq -M \cdot w_{n_s(t),i}^s \quad (24)$$

$$f_{t,i}^s \leq M \cdot w_{n_r(t),i}^s \quad (25)$$

$$f_{t,i}^s \geq -M \cdot w_{n_r(t),i}^s \quad (26)$$

$$f_{t,i}^s \leq \left( v_{t,i}^s / X_{t,i} \right) \sum_{n \in \mathcal{N}} \tau_{t,n} \theta_{n,i}^s + M \cdot \left( 2 - w_{n_s(t),i}^s + w_{n_r(t),i}^s \right), \quad (27)$$

$$f_{t,i}^s \geq \left( v_{t,i}^s / X_{t,i} \right) \sum_{n \in \mathcal{N}} \tau_{t,n} \theta_{n,i}^s + M \cdot \left( w_{n_s(t),i}^s + w_{n_r(t),i}^s - 2 \right), \quad (28)$$

for all cyber-physical attackers with **advanced** capabilities  $i \in \mathcal{A}$ , attack vector scenarios  $s \in S_i^k$  and **attacked** transmission branches  $t \in \mathcal{T}: v_{t,i}^s = 0$ :

$$f_{t,i}^s = 0, \quad (29)$$

for all cyber-physical attackers  $i \in \mathcal{A} \cup \mathcal{B}$ , attack vector scenarios  $s \in S_i^k$  and **non-attacked** transmission branches  $t \in \mathcal{T}: v_{t,i}^s = 1$ :

$$f_{t,i}^s = (1/X_{t,i}) \sum_{n \in \mathcal{N}} \tau_{t,n} \theta_{n,i}^s. \quad (30)$$

In (23)–(28) indices  $n_s(t)$ ,  $n_r(t)$  denote the sending and receiving buses of transmission branch  $t$ . Accordingly, inequality constraints (23)–(26) set the power flow through a branch to zero if at least one of its end-buses is successfully attacked. Inequalities (27)–(28) enforce the power flow definition from equality (21) only for branches whose end-buses are both successfully protected. Equality constraint (29) only holds for the sub-class of cyber-physical attackers with advanced capabilities. Since these actors can by-pass the cyber-security measures, the power flow of any attacked branch will in any case be equal to zero. Finally, the standard DC power flow definition equality is imposed for any transmission branch that is not under attack (30).

### 2.3.2. Bi-level sub-problems: vs a given cyber-physical security plan

At every iteration of the algorithm, and given updated decisions for the cyber-physical planner  $(z^k, p^k, r^k)$ , we formulate per cyber-physical attacker  $i \in \mathcal{A} \cup \mathcal{B}$  a bi-level problem as in (9)–(22). To solve these problems we rely on the fact that the lowest-level minimization problem (16)–(22) is a convex optimization problem. Exploiting this property, we remove the minimization operator of (16), introduce the dual feasibility constraints of (16)–(22) and enforce the strong duality condition [13].

## 3. Results

We consider a set of case studies defined on two well known benchmarks, namely the IEEE 24-bus and 118-bus systems. Figs. 3 and 4 plot the respective single-line diagrams. Our motivation for using two different benchmarks is to derive observations that do not depend on the properties of the considered system. We use all original system data as documented by Ref. [11] and supplement these with:

1. An assumed *value of lost load* parameter, set at 5000 monetary units per MWh;
2. an assumed *cost of updating firewall rules* parameter, set at 5.55 monetary units per device;
3. An assumed *reserve capacity cost* parameter, set per generator at 25 % of the respective dispatch cost parameter.

To put results into context, we will henceforth report all cost data as a percentage of the base-case dispatch cost of the respective benchmark. Such cost is computed while considering an empty set of potential cyber-physical attackers and solving optimization problem (1)–(8). It amounts to 41,904.11 monetary units for the IEEE 24-bus benchmark and 109,826.08 monetary units for the IEEE 118-bus benchmark. Table 1 collects the test cases considered and introduces the order in which results will be presented. As can be seen, apart from different benchmark systems, we differentiate test cases in terms of the cyber-security resources available to the planner (3rd column), the resources available to a cyber-physical attacker with basic (BSC) and/or advanced (ADV) capabilities (4th and 5th columns respectively) as well as the probabilities<sup>3</sup> associated with these actors (6th and 7th columns respectively).

<sup>3</sup> All probability values used in our numerical analysis are arbitrary. In real-life, Transmission System Operators would have to rely on Threat Intelligence Agencies for estimating such probabilities.

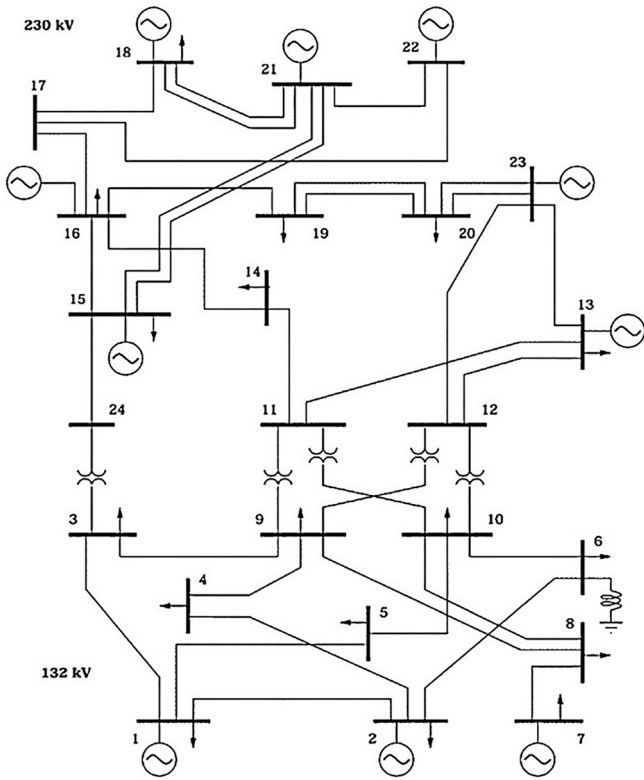


Fig. 3. Single-line diagram of the IEEE 24-bus benchmark [14].

Table 1  
Summary of test case parameters.

Case	System	Resources			Attack Probability	
		Planner	Attackers		BSC	ADV
			BSC	ADV		
A	24-bus	24	2	0	0.01	0
B	24-bus	24	0	2	0.0	0.01
C	24-bus	24	2	2	0.005	0.005
D	24-bus	3	2	0	0.01	0
E	24-bus	3	0	2	0.0	0.01
F	24-bus	3	2	2	0.005	0.005
G	118-bus	3	2	0	0.01	0
H	118-bus	3	0	2	0.0	0.01
I	118-bus	3	2	2	0.005	0.005

3.1. Unlimited cyber-security resources

We start by considering a set of test-cases on the 24-bus system and assume that the cyber-physical planner has unlimited cyber-security resources, which would allow her to update the firewall rules for all buses of the system. We first consider in case A that there exists only a single credible cyber-physical attacker and this actor does not have the capability to overcome the applicable cyber-security measures. In complementary case B we consider that the same cyber-physical attacker can by-pass firewalls, even under the updated rules. In both cases, the probability associated with the single credible attacker is set at 0.01 and her resource limit is set to 2, meaning that she has the ability to intrude at most 2 buses. Table 2 includes the results for these two test cases.

Starting from the data in the top row of Table 2, if cyber-security measures are fully effective, under the assumed conditions it makes (economic) sense to fully rely on these so as to completely neutralize the threat posed by the cyber-physical attacker. Indeed, the optimal strategy here is to pay a negligible additional cost of 0.29 % at the planning

Table 2  
Solution overview for test cases A–C.

Case	Buses Secured	Costs (%)		
		Total	Dispatch + Reserve	Exp. (2nd-stage)
A	1–10, 12–23	100.29	100	0
B	–	200.54	117.17	83.37
C	1–10, 13–23	157.35	107.20	49.88

stage so as to ensure that a potential cyber-physical attack may not have any impact on the power grid operation. Notice that this additional 0.29 % is only spent on updating the firewall rules, since the optimal costs of generation dispatch and reserve procurement are identical to the aforementioned base-case results (computed while assuming no credible threat). It should also be noted that the optimal planning strategy is to deploy the considered cyber-security measure in 22 out of the 24 system buses. Recalling here that the 24-bus benchmark is relatively lightly loaded and has ample transmission capacity, in the considered steady-state, linear model it could withstand the outage of two transmission (*i.e.*, no load or generation) buses.

The middle row of Table 2 concerns the case wherein the cyber-physical attacker would have the advanced capability to by-pass the cyber-security measures under consideration. It is easy to understand that, when facing only such a threat actor, there is no point in using cyber-security measures. Rather, the only form of security can come from physical measures. Under the assumed conditions, it makes (economic) sense to spend an additional cost of 17.17 % in choosing the generation dispatch and procuring reserve capacity (as seen in the 4th column). Still, this would not completely neutralize the credible threat which carries a substantial expected impact. It would be equal to 83.37 % of the base-case dispatch cost. Our detailed results indicate that the cyber-physical attacker’s strategy would be to intrude on buses 15 (load and significant generation capacity) and 23 (largest generation capacity) thus provoking both reserve deployment and involuntary load shedding.

Let us now move on to test case C, which as seen in Table 1 concerns two alternative credible cyber-physical attackers differentiated only in terms of their ability to overcome or not the updated firewall rules. Both these actors have the same resources as in cases A and B and share equally the total probability associated with the attackers in cases A or B. The bottom row of Table 2 reports the corresponding solution overview. As evidenced by these results, acknowledging that a cyber-physical attacker may not be impeded by cyber-security measures results in using a blend of cyber- and physical- security measures. Indeed, the optimal solution includes both updating the firewall rules for 21 system buses and paying an additional 7.2 % in generation dispatch and reserve capacity. What is most interesting here is the synergy between the latter and former measures. The physical system is more strongly prepared, by way of the generation dispatch and reserve capacity, since it is possibly facing an attacker that could by-pass cyber-security measures. This strength also carries over to possibly facing the alternative attacker that would not be able to by-pass the firewall rules. The available reserve capacity, along with the reduced probability of the cyber-physical attacker with basic capabilities, implies that she is expected to only cause a minor impact on the power grid operation. As a result, the security of the cyber sub-system against this threat can be slightly relaxed. Compared to case A, one less bus is chosen to be protected by cyber-security measures in case C.

3.2. Limited cyber-security resources

We continue with a set of test-cases on the 24-bus and 118-bus benchmarks in the setting where the cyber-security resources of the planner are limited. We believe that considering such limitations is very relevant, since in practice deploying cyber-security measures may involve mobilizing specialized workforce to access the concerned systems locally. For

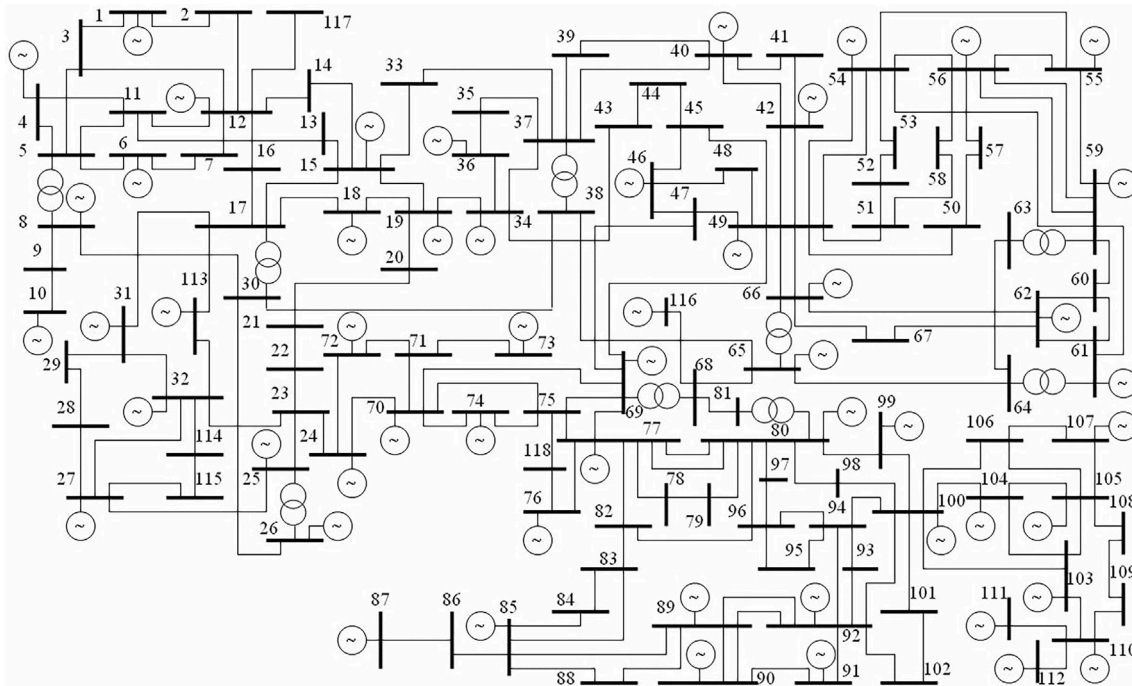


Fig. 4. Single-line diagram of the IEEE 118-bus benchmark [15].

Table 3

Solution overview for test cases D–F.

Case	Buses	Costs (%)			
		Secured	Total	Dispatch + Reserve	Exp. (2nd-stage)
D	15,18,23		172.47	117.54	54.89
E	–		200.54	117.17	83.37
F	15,18,23		186.91	117.73	69.13

both benchmarks, we limit the cyber-physical planner’s resources to 3 while maintaining the assumption that any attacker could at most intrude into 2 buses. We form alternative test cases by varying the sets of credible cyber-physical attackers in the same style as previously. Specifically, for any benchmark, we consider the alternative hypotheses of facing (i,ii) a single credible attacker with either basic or advanced capabilities as well as (iii) two credible attackers with basic and advanced capabilities respectively.

### 3.2.1. 24-bus benchmark

Table 3 concerns the test cases on the 24-bus benchmark. The results for case D, compared to those for case A, confirm that the limit on the available cyber-security resources is indeed restrictive. This restriction is partly counterbalanced by the procurement of reserve capacity at an additional cost of 17.54 %. Once again we can notice here how cyber- and physical- security measures can complement one another if chosen in an integrated manner. It would however be impossible to completely neutralize the impact of the considered cyber-physical attack on the power grid operation cost, since the deployment of the procured reserve also comes at a cost. It is also interesting to acknowledge that the cyber-physical planner’s optimal decision making includes updating the firewall rules for buses 15 and 23. We recall from case B that these would form the optimal attack vector for the attacker with advanced capabilities. The results for case E are, as anticipated, identical to those of case B. Since both these cases concern an attacker with advanced capabilities only, the cyber-security measures are ineffective hence the respective resource limit is immaterial. Comparing results for cases F and D, we

Table 4

Solution overview for test cases G–I.

Case	Buses	Costs (%)			
		Secured	Total	Dispatch + Reserve	Exp. (2nd-stage)
G	26,59,80		129.05	108.20	20.83
H	–		136.78	106	30.78
I	26,59,69		133.99	107.52	26.45

find that, for this specific benchmark, the consideration that a cyber-physical attacker may find a way to overcome the applied cyber-security measures has a moderate effect on the optimal decisions of the planner regarding physical security measures and no effect on the optimal decisions regarding cyber-security measures. Indeed, the only difference is in the optimal reserve & dispatch cost which is only increased by 0.19 %. It appears preferable, to tolerate a larger expected impact of a potential attack on the power grid operation cost, and precisely increased by 14.24 %.

### 3.2.2. 118-bus benchmark

Table 4 concerns the test cases on the 118-bus benchmark. We make an interesting observation by comparing the outcomes for cases G and I. Recall that case G only acknowledges the possibility of a cyber-physical attacker with basic capabilities while case I also acknowledges the possibility of an attacker with advanced capabilities to overcome cyber-security measures. As per our results, the optimal choice of which buses to protect by means of updated firewall rules is different between these two cases. Again, the reason is the synergy between cyber- and physical-security measures. The hypothesis in case I calls for different physical-security measures, serving as the optimal way to counterbalance the threat posed by the cyber-physical attacker with advanced capabilities. These physical measures would also be effective for the alternative of the attacker with basic capabilities, providing a form of security and leading to a different choice of (complementary) cyber-security measures.

Table 5 details the evolution of the problem’s optimal solution as per the C&CG algorithm for test case I. Per iteration, the 2nd column of this

**Table 5**  
CC&G algorithm evolution for test case I.

Iteration	Buses		Cost (%)			
	Secured	Attacked		1st-stage		Total
		BSC	ADV			
1		69,80	69,80	100	100	
2	69,80	10,26	66,80	108.39	116.87	
3	10,26,80	69,66	69,66	108.69	117.32	
4	66,10,80	59,49	59,49	108.76	119.36	
5	69,59,10	66,100	69,100	106	124.46	
6	69,59,10	26,9	69,80	108.61	129.24	
7	69,59,26	89,80	69,80	108.61	129.24	
8	59,26,80	69,89	69,80	109.39	129.41	
9	59,26,80	100,89	59,49	109.33	132.34	
10	26,49,80	66,59	59,49	109.82	133.35	
11	59,26,80	49,8	59,49	109.82	133.35	
12	69,59,26	68,66	69,80	107.16	133.57	
13	59,26,89	100,80	59,49	108.04	133.68	
14	69,59,26	10,80	69,80	107.52	133.99	

table holds the indices of the buses wherein the cyber-physical planner would decide to update the firewall rules. The 3rd and 4th columns contain the attack vectors launched by the cyber-physical attackers with basic and advanced capabilities in response to the planner's updated decisions. The last two columns show the cost of the optimal physical security measures and the total cost of the cyber-physical planner (*i.e.*, her objective).

At the 1st iteration of the algorithm, the cyber-physical planner considers no credible threat and thus implements no security measure. Both attackers would in such case launch the same attack vector since none of them is impeded by any barrier. Acknowledging this possibility, at the 2nd iteration, the cyber-physical planner chooses to both (i) completely neutralize the attack vector of the attacker with basic capabilities (by updating firewall rules) and (ii) mitigate the potential impact of the alternative attacker with advanced capabilities (by changing the generation dispatch and procuring reserve capacity). It is also interesting to comment on the solution for iteration 4. Since the planner's resource limit is set at 3, she may not completely neutralize the subset of attack vectors for the attacker with basic capabilities. This subset already includes 3 alternative pairs of buses and therefore requires 5 resources to be rendered ineffectual. Notice that bus 69 is common in two of these attack vectors. The planner's optimal decision is to spread her resources to partially mitigate each one of these 3 alternative vectors rather than completely neutralize one of them. By doing so, the planner is seeking to ensure that there is no attack vector with considerably high impact that would also push the expectation term and the overall objective higher. It is also interesting that bus 69 is not one of the buses to be protected by the available cyber-security resources. We found that this bus would also be targeted by the cyber-physical attacker with advanced capabilities and infer that physical security measures have been chosen to mitigate this threat.

It is also obvious from Table 5 that, as anticipated, the cyber-physical attacker with advanced capabilities only reacts to an update of the physical security measures. Indeed, between iterations where these measures are not updated this actor would repeatedly launch the same vector (*e.g.*, between iterations 6 and 7). We should explain here that it is not necessary for the planner to update both cyber- and physical- security measures at every iteration, since seeking to co-optimize her planning cost along with the *joint* expectation of the attack impact over the set of credible attackers. We also notice that the cyber-physical attacker with advanced capabilities would frequently return to the same attack vector (buses 69,80) even when the physical- security measures seem to change. The updated dispatching and reserve decisions reduce the expected impact of this vector, however it still remains the best choice for this attacker relatively to her other options. We finally remark that while the total cost for the cyber-physical planner can only increase between

**Table 6**  
Computational cost.

Case	Execution time (s)				Iterations
	Total	Master	Sub-problems		
			Basic	Advanced	
A	8.38	1.1	7.28	0	14
B	2.06	0.61	0	1.45	3
C	13.52	1.24	7.40	4.87	13
D	8.41	1.62	6.79	0	16
E	2.06	0.61	0	1.45	3
F	10.57	1.14	4.61	4.82	11
G	1409.24	9.24	1400	0	15
H	605.4	7.56	0	597.84	6
I	3307.55	20.75	1600.80	1686	14

**Table 7**  
Sensitivity of the computational cost to the attacker resources.

Attacker resources	Execution time/iteration (s)			Iterations
	Master	Sub-problems		
		Basic	Advanced	
2	0.10	0.42	0.44	11
4	0.17	0.99	1.37	15
6	1.01	1.78	1.65	30

different iterations, the optimal mix between 1st and 2nd stage costs is not necessarily constant. This explains why the reserve & dispatch cost appears to decrease between some successive iterations.

### 3.3. Computational cost

Our implementation of the C&CG algorithm for the resolution of tri-level optimization problem (1)–(22) was written in Julia [16], using the JuMP modeling language [17]. We used the PowerModels.jl package [18] for parsing the benchmark data and solved all optimization problem instances with the CPLEX solver [19]. All test cases were solved using a MacBook with an 8-core M2 chip and 24 GB of RAM.

Table 6 reports the computational cost of all test cases in terms of the total execution time and number of iterations. The effect of the size of the considered system is obvious. We may also notice by comparing the data for cases A–C to those for cases D–F that the assumed resource constraint of the cyber-physical planner has a marginal effect on computational cost. On the other hand, we have found that the assumed value for a cyber-physical attacker's resource constraint has a multi-fold effect on the computational cost. Firstly, it generally increases the number of iterations required to achieve convergence. As the algorithm requires more iterations to converge to the optimal solution, the *Master Problem* also requires more execution time per iteration to be solved. Finally, marginally increasing the resource limit also increases the execution time required to solve each instance of the *Bi-level Sub-problem* (9)–(22). Table 7 demonstrates these effects by documenting the computational cost of variants of case C with additional resources available to both alternative cyber-physical attackers.

## 4. Conclusions

In this paper we focused on the threat posed by (malicious) cyber-physical attackers, who may use the cyber sub-system of the power grid as the medium to disrupt the continuous supply of electricity. The topic is rather timely not only due to the increasing role of digital monitoring, control, communications and computation infrastructures in the operation of the modern power grid but, unfortunately, also due to geo-political developments. The salient feature of this work is the acknowledgment that even though cyber-security measures can be effective, a sophisticated threat actor may eventually find a (new) way to circumvent these. We therefore investigated the stakes for an integrated

risk management approach choosing both cyber- and physical- security measures under uncertainty regarding the effectiveness of the former.

To do so, we formalized an optimization problem from the perspective of a cyber-physical planner seeking to co-optimize her costs along with the expected impact of a cyber-physical attack on the power grid operation cost. The latter includes a cost of energy not supplied component, penalizing the event of an attack leading to electricity supply interruption. The planner faces a distribution of alternative credible cyber-physical attackers, including attackers that can overcome cyber-security measures. Our main finding is a dual synergy between cyber- and physical security measures. The use of physical-security measures was shown to be primarily associated with the potential ineffectiveness of cyber-security measures. Put simply, an attacker who may overcome cyber-security measures would be counteracted by preparing the physical grid to withstand her attack vector. As a secondary effect, taking physical-security measures against this credible threat makes the grid physically stronger, thus more able to withstand any credible threat. In fine, this physically stronger grid may have different vulnerabilities and different priorities/needs for cyber-security. There is therefore scope for choosing cyber- and physical- security measures in an integrated manner.

It is relevant to mention here the distinctive features of the considered threat with respect to the ‘classical’, *physical-only* threats to the security of the electric power grid, such as component outages (a.k.a. contingencies) and power injection forecasting errors. While physical-only threats are purely exogenous and show relatively favorable statistical properties, cyber-physical threats are adaptive to the security management strategy and intentionally designed for the maximum attainable impact. To represent these features mathematically, we adopted the framework of tri-level optimization (under uncertainty). Admittedly, the solvability of our tri-level optimization framework comes with limitations that we have imposed on the representation of the grid physics and of its real-time control problem statement (steady-state and linear power flow model, with only continuous real-time controls and no infeasibility). These limitations however do not seem critical for the conceptual question we address here and the qualitative conclusion we reach. It is nevertheless desirable to limitations these restrictions in order to develop more realistic solution strategies for more advanced versions of the integrated cyber-physical power grid security management problem.

We remark here that these solvability limitations mostly refer to the lowest-level optimization problem (16)–(22) used by a cyber-physical attacker so as to anticipate the power grid operator’s optimal reaction. One may argue that in practice cyber-physical attackers are equally bound by similar solvability limitations and would have to resort to similar simplifying assumptions in order to identify and launch near optimal attack vectors. In this spirit, a possible next research step would be to differentiate between the physical model used in (16)–(22) to determine the optimal attack vector(s) of the considered cyber-physical attacker(s) and the physical model used to assess and control the impact of such attack vector(s) on the power grid. Indeed, there is no *a priori* restriction on using a more detailed physical model either in the so-called *Master Problem* or in an additional post-processing step. This additional step would serve to evaluate more precisely an *upper bound* given the resolution of all *Bilevel Sub-problem(s)* and the respective attack vector(s).

The computational cost of the tri-level optimization framework is clearly also a limiting issue that requires dedicated attention and further research. In spite of such cost, we relied on this framework since it offers the transparency/interpretability required for the analysis we sought to perform here. Several improvements in algorithmic efficiency are conceivable on top of the basic implementation that we used, including for instance the resolution of Bi-level Sub-problem(s) in parallel over the credible set of cyber-physical attackers. We wish to also acknowledge the potential interest in moving away from approaches purely based on analytical optimization techniques. The particular alternative approach that we aim to explore in future research is *Multi-agent*

*Reinforcement Learning*. We refer the interested reader to [20] for a comprehensive exposition of the relevance of using Reinforcement Learning for cyber-physical system security management.

As a concluding point, we return to the original question and the integration of cyber- and physical- resources for electric power system security management. We have shown here that the integration of such resources is indeed well-justifiable, especially given the risk that malicious adversaries may develop the skill to overcome cyber-security measures. We showed this while focusing on a single instance of power grid operation setting and on cyber- and physical- measures that could conceivably be chosen on similar time-frames. Addressing the electric-power system cyber-physical security management problem as a multi-stage stochastic decision making problem is the end goal of this research stream.

#### CRediT authorship contribution statement

**Efthymios Karangelos:** Writing – review & editing, Writing – original draft, Validation, Software, Project administration, Methodology, Investigation, Conceptualization. **Louis Wehenkel:** Writing – review & editing, Resources, Investigation, Funding acquisition.

#### Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered potential competing interests:

Efthymios Karangelos reports that financial support was provided by Belgian Energy Transition Funds. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

#### Acknowledgment

This work has been prepared with the support of the Belgian Energy Transition Fund, project CYPRESS (<https://cypress-project.be/>).

#### Data availability

All data used are publicly available and we have cited the sources.

#### References

- [1] J.E. Sullivan, D. Kamensky, How cyber-attacks in Ukraine show the vulnerability of the us power grid, *Electr. J.* 30 (3) (2017) 30–35.
- [2] M.S. Alvarez-Alvarado, C. Apolo-Tinoco, M.J. Ramirez-Prado, F.E. Alban-Chacón, N. Pico, J. Aviles-Cedeno, A.A. Recalde, F. Moncayo-Rea, W. Velasquez, J. Rengifo, Cyber-physical power systems: a comprehensive review about technologies drivers, standards, and future perspectives, *Comput. Electr. Eng.* 116 (2024) 109149.
- [3] S. Karamdel, X. Liang, S.O. Faried, M. Mitolo, Optimization models in cyber-physical power systems: a review, *IEEE Access* 10 (2022) 130469–130486.
- [4] H. He, S. Huang, Y. Liu, T. Zhang, A tri-level optimization model for power grid defense with the consideration of post-allocated dgs against coordinated cyber-physical attacks, *Int. J. Electr. Power Energy Syst.* 130 (2021) 106903.
- [5] K. Lai, M. Illindala, K. Subramaniam, A tri-level optimization model to mitigate coordinated attacks on electric power systems in a cyber-physical environment, *Appl. Energy* 235 (2019) 204–218.
- [6] C. Qin, C. Zhong, B. Sun, X. Jin, Y. Zeng, A tri-level optimal defense method against coordinated cyber-physical attacks considering full substation topology, *Appl. Energy* 339 (2023) 120961.
- [7] Z. Zhang, S. Huang, Y. Chen, B. Li, S. Mei, Cyber-physical coordinated risk mitigation in smart grids based on attack-defense game, *IEEE Trans. Power Syst.* 37 (1) (2022) 530–542.
- [8] B. Zeng, L. Zhao, Solving two-stage robust optimization problems using a column-and-constraint generation method, *Oper. Res. Lett.* 41 (5) (2013) 457–461.
- [9] Y. Xiang, L. Wang, An improved defender–attacker–defender model for transmission line defense considering offensive resource uncertainties, *IEEE Trans. Smart Grid* 10 (3) (2018) 2534–2546.
- [10] C. Grigg, P. Wong, P. Albrecht, R. Allan, M. Bhavaraju, R. Billinton, Q. Chen, C. Fong, S. Haddad, S. Kuruganty, et al, The IEEE reliability test system-1996. A report prepared by the reliability test system task force of the application of probability methods subcommittee, *IEEE Trans. Power Syst.* 14 (3) (1999) 1010–1020.
- [11] C. Coffrin, D. Gordon, P. Scott, Nesta, the NICTA energy system test case archive, 2019. [Online]. Available: <https://arxiv.org/abs/1411.0359>.
- [12] W. Yuan, L. Zhao, B. Zeng, Optimal power grid protection through a defender–attacker–defender model, *Rel. Eng. Syst. Saf.* 121 (2014) 83–89.

- [13] T. Kleinert, M. Labbé, F. Plein, M. Schmidt, Closing the gap in linear bilevel optimization: a new valid primal-dual inequality, *Optim. Lett.* 15 (4) (2021) 1027–1040.
- [14] K. Sun, D.-Z. Zheng, Q. Lu, A simulation study of obdd-based proper splitting strategies for power systems under consideration of transient stability, *IEEE Trans. Power Syst.* 20, 03 2005) 389–399.
- [15] A.R. Al-Roomi, Power Flow Test Systems Repository, Halifax, Nova Scotia, Canada, 2015. [Online]. Available: <https://al-roomi.org/power-flow>.
- [16] J. Bezanson, A. Edelman, S. Karpinski, V. Shah, Julia: a fresh approach to numerical computing, *SIAM Rev.* 59 (1) (2017). [Online]. Available: 65–98, <https://doi.org/10.1137/141000671>.
- [17] I. Dunning, J. Huchette, M. Lubin, JuMP: a modeling language for mathematical optimization, *SIAM Rev.* 59 (2) (2017) 295–320. [Online]. Available: <https://doi.org/10.1137/15M1020575>.
- [18] C. Coffrin, R. Bent, K. Sundar, Y. Ng, M. Lubin, PowerModels.JL: an open-source framework for exploring power formulations, in: 2018 Power Systems Computation Conference (PSCC), June 2018.
- [19] Cplex, IBM ILOG, V12. 1: User's manual for CPLEX, *Int. Bus. Mach. Corp.* 46 (53) (2009) 157.
- [20] C. Li, M. Qiu, Reinforcement Learning for Cyber-Physical Systems: With Cybersecurity Case Studies, Chapman and Hall/CRC, 2019.