Computing necessary conditions for near-optimality in capacity expansion planning problems☆

Antoine Dubois [a],[*], Damien Ernst [a],[b]

[a] *Department of Electrical Engineering and Computer Science, ULiège, Liège, Belgium*
[b] *LTCI, Telecom Paris, Institut Polytechnique de Paris, Paris, France*

ARTICLE INFO

ABSTRACT

In power systems, large-scale optimisation problems are extensively used to plan for capacity expansion at the supra-national level. However, their cost-optimal solutions are often not exploitable by decision-makers who are preferably looking for features of solutions that can accommodate their different requirements. This paper proposes a generic framework for addressing this problem. It is based on the concept of the epsilon-optimal feasible space of a given optimisation problem and the identification of necessary conditions over this space. This framework has been developed in a generic case, and an approach for solving this problem is subsequently described for a specific case where conditions are constrained sums of variables. The approach is tested on a case study about capacity expansion planning of the European electricity network to determine necessary conditions on the minimal investments in transmission, storage and generation capacity.

1. Introduction

In the coming decades, the European power system will have to face the challenges related to the integration of massive amounts of renewable energy sources and a high level of electrification of the heating, transport and industrial sectors.

The size and level of integration of the European electricity network (*i.e.*, at the transmission level, thousands of substations and power lines connecting them) entail a level of complexity in planning this transition that requires using detailed optimisation models. The increased sophistication of these models comes with drawbacks. In particular, these models essentially focus on unique cost-based optimal solutions that are too restrictive and do not encompass the different requirements of many stakeholders intervening in the decision process for new investments in capacity.

In our opinion, it is preferable to provide *necessary conditions* in capacity investments that guarantee a constrained suboptimality and provide a common ground over which decision-makers can settle and create solutions that accommodate their needs. For example, we could compute the minimum required investment in transmission lines per country to ensure a maximum deviation of 10% from the optimum. Alternatively, one might be interested in knowing if a particular technology – for example, Li-Ion battery or some renewable energy source (RES) type – is necessary for a cost-efficient energy transition.

In this paper, a framework is presented to derive necessary conditions for ϵ-optimality and applied to a capacity expansion planning problem. In Section 2, we discuss the literature related to the optimisation concepts that underlie the framework. The optimisation framework itself is presented in Section 3. Section 4 specifies this framework for the case of conditions consisting of constrained sums of variables and provides a fully-defined methodology for computing *non-implied necessary conditions* in such a context. This methodology is afterwards illustrated on an expansion planning problem in Section 5. Section 6 concludes with the description of future research directions. Finally, Appendix gathers more detailed data on the modelling of the network used in the test case.

2. Literature review

Decision-making based on optimisation results is a complex task. Indeed, this exercise lies at the frontier between human intelligence and machine power, whose coupling is challenging [1], sometimes referred to as post-normal science [2]. Decision-making is linked to *deep uncertainties* [3] where, among other topics, the desirability of alternative outcomes corresponding to different policy objectives is subject to disagreement among stakeholders. These uncertainties and disagreements imply that relying on a single cost-based optimum is often not sufficient. Indeed, there is no guarantee that the findings

obtained via this optimum will be robust regarding parameter perturbation, nor that they will satisfy conflicting objectives. Moreover, as shown in [4], cost-optimal scenarios are not adequate to approximate real-world problems, such as those encountered in the context of the energy transition. This problem highlights the need for the "role of optimisation model methods to be re-thought in full recognition of these limitations", as suggested in [5]. Those authors advocate optimisation methods that "should be used to generate planning alternatives, facilitate their evaluation and elaboration, provide insights and serve as catalysts for human creativity".

We argue that such a *re-thinking* can be achieved by orienting the use of optimisation methods in the search for conditions that are respected across multiple feasible solutions and guarantee a constrained level of suboptimality. The advantage of this approach over unique cost-optimal solutions is to provide decision elements that all stakeholders can agree on and build on using their creativity.

Those solutions can be obtained in a variety of ways. One possibility is the use of multi-objective (or multi-criteria) optimisation [6]. In this field, one searches not for a single solution but a set of *efficient* or *Pareto-optimal* solutions, translating some trade-offs between objectives. A notable drawback with these methods is the general assumption of knowing which objectives are at stake in the problem and being able to model in some form those objectives. Objectives that are a priori unknown (either because of lack of knowledge or unconscious biases) or non-modellable are, thus, left apart.

A technique that circumvents this limitation is what some authors refer to as *"Modelling to Generate Alternatives"* [7]. It consists of exploring solutions located in the inferior or suboptimal region of an optimisation problem [8]. The underlying motivation of this approach is that this region might contain solutions that are better in terms of some unmodelled objectives. Several authors, such as [9–12] or [13], exploit this technique. However, their main objective is to show the variety of solutions that can be extracted rather than to systematically compute conditions that are respected by those solutions. In this paper, we present a framework that puts the identification of such conditions at the centre of the optimisation process.

Finally, in the domain of multi-objective optimisation, [14] have surveyed advanced data-driven methods for extracting *commonalities* among Pareto-optimal solutions. Our framework aims at providing the ground for developing such techniques in suboptimal spaces.

3. Problem formulation

Let us consider the following optimisation problem

$$\min_{x \in \mathcal{X}} f(x) \tag{1}$$

with $\mathcal{X}$ being the feasible space and $f : \mathcal{X} \rightarrow \mathbb{R}^+$ the objective function. Let x^* be an optimal solution. We define an ϵ-*optimal space* as follows:

$$\mathcal{X}^\epsilon = \{x \in \mathcal{X} \mid f(x) \leq (1 + \epsilon)f(x^*), \epsilon \geq 0\}.$$

The set $\mathcal{X}^\epsilon$, depicted in Fig. 1, contains only the feasible solutions with an objective value no greater than $(1 + \epsilon)f(x^*)$. We define ϵ as the *suboptimality coefficient* of such a space, *i.e.* specifying by how much the objective values of the solutions in the space deviate at most from the optimal objective value.

Let us define *conditions* as functions $\phi : \mathcal{X} \rightarrow \{0, 1\}$. Our goal is to identify, among a set Φ of conditions, the ones which are true for any solutions in $\mathcal{X}^\epsilon$. These conditions are called *necessary conditions* for ϵ-optimality, where the parameter ϵ allows one to monitor the level of suboptimality of those necessary conditions. Mathematically,

$$\Phi^{\mathcal{X}^\epsilon} = \{\phi \in \Phi \mid \forall x \in \mathcal{X}^\epsilon : \phi(x) = 1\}$$

is the set of necessary conditions for a given feasible space $\mathcal{X}$, set of conditions Φ and suboptimality coefficient ϵ.
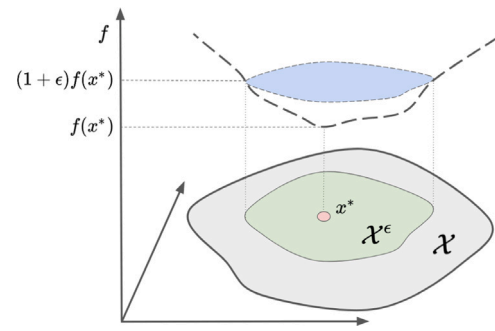

Fig. 1. Three-dimensional representation of an ϵ-optimal space. In the horizontal plane, the space of feasible solutions $\mathcal{X} \subset \mathbb{R}^2$ is depicted while the vertical axis represents the objective function. The red dot is the optimal solution x^* corresponding to the minimal value of the objective function over the feasible space. The blue space coincides with the hyper-plane $f(x) = (1 + \epsilon)f(x^*)$ allowing one to determine the ϵ-optimal space $\mathcal{X}^\epsilon$ shown in green.

3.1. Non-implied necessary conditions on sets of parametric conditions

The goal of the methodology presented in this paper is to support decision-makers in their decision process. However, as explained in the next paragraph, even for a single set of conditions, an infinite number of necessary conditions can be derived. Such quantity of information cannot be used efficiently to take decisions. In this section, the concept of *non-implied necessary condition* is introduced as a solution to this problem.

Let consider the feasible space $\mathcal{X} = \mathbb{R}$ and a set of parametric conditions of the type

$$\Phi = \{\phi_c(x) := x \geq c \mid c \in \mathbb{R}\}.$$

This set contains an infinite number of conditions and can lead to identifying an infinite number of necessary conditions, with which decision-makers might find it cumbersome to deal.

For instance, let $\phi_1(x) := x > 1$ be a necessary condition for ϵ-optimality (*i.e.* $\forall x \in \mathcal{X}^\epsilon : \phi_1(x) = 1$). This automatically *implies* that all ϕ_c where $c < 1$ are necessary conditions. Indeed $\forall x \in \mathcal{X}^\epsilon : x > 1 \Rightarrow x > c$. The only condition that cannot be implied to be a necessary condition from the knowledge of other necessary conditions is the necessary condition ϕ_c with the largest value of c.

This necessary condition is what constitutes a *non-implied necessary condition*. This is a condition that cannot be *implied* to be a necessary condition from the sole knowledge of other conditions that constitute necessary conditions. To minimise the number of necessary conditions that need to be identified and presented to decision-makers, the focus should be placed on the identification of non-implied necessary conditions.

The notion of implication can be formalised by defining the space over which a condition ϕ is true,

$$\mathcal{I}_\phi = \{x \in \mathcal{X} \mid \phi(x) = 1\}.$$

A condition ϕ_2 implies ϕ_1 if $\mathcal{I}_{\phi_2} \subset \mathcal{I}_{\phi_1}$, *i.e.* ϕ_1 is true for all $x \in \mathcal{X}$ over which ϕ_2 is true. Considering sets of parametric conditions, $\mathcal{I}_{\phi_2} = \mathcal{I}_{\phi_1}$ happens only if both conditions are equal. Using this notion, conditions can be defined to be necessary conditions if the space over which they are true encloses $\mathcal{X}^\epsilon$, and non-implied necessary conditions if this space does not include any of the spaces over which other necessary conditions are true. Mathematically, the set of non-implied necessary conditions for ϵ-optimality is defined as

$$\overline{\Phi}^{\mathcal{X}^\epsilon} = \{\phi \in \Phi^{\mathcal{X}^\epsilon} \mid \forall \phi' \in \Phi^{\mathcal{X}^\epsilon} \setminus \{\phi\} : \mathcal{I}_{\phi'} \not\subset \mathcal{I}_\phi\}.$$

Fig. 2(a) provides an illustration of two necessary conditions ϕ_1 and ϕ_2, with ϕ_2 implying ϕ_1. Considering a set of conditions $\Phi =$
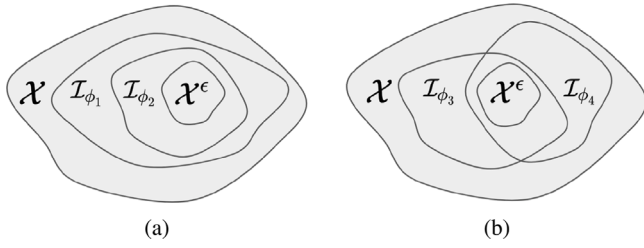
Fig. 2. Graphical illustration of implication using spaces over which conditions are true.

$\{\phi_1, \phi_2\}$ containing uniquely ϕ_1 and ϕ_2, the set of necessary conditions is given by $\Phi^{\mathcal{X}^\epsilon} = \{\phi_1, \phi_2\}$, and the set of non-implied necessary conditions by $\overline{\Phi}^{\mathcal{X}^\epsilon} = \{\phi_2\}$. Fig. 2(b) provides an illustration of two other necessary conditions ϕ_3 and ϕ_4, with no implication. Considering a set of conditions $\Phi = \{\phi_3, \phi_4\}$, then $\overline{\Phi}^{\mathcal{X}^\epsilon} = \Phi^{\mathcal{X}^\epsilon} = \Phi$.

4. Proposed method

In this section, a specific instance of the generic problem specified in Section 3 is considered. This instance is characterised by conditions consisting of parametric constrained sums of variables. We show how for each value of the parameter defining these sums, a unique non-implied necessary condition can be determined.

Theorem 1. *Let $\mathcal{X} \subset \mathbb{R}^n$, $f : \mathcal{X} \to \mathbb{R}^+$ and*

$$\Phi_{\mathbf{d}} = \{\phi_{\mathbf{d}}^c(\mathbf{x}) := \mathbf{d}^T\mathbf{x} \geq c \mid c \in \mathbb{R}\},$$

where $\mathbf{x} \in \mathcal{X}$, be a set of conditions consisting of constrained sums of variables $\mathbf{d}^T\mathbf{x} = \sum_{i=1}^n d_i x_i$ defined by $\mathbf{d} \in \{0, 1\}^n$.

Let $c^* = \min_{\mathbf{x} \in \mathcal{X}^\epsilon} \mathbf{d}^T\mathbf{x}$ then

$$\phi_{\mathbf{d}}^{c^*} := \mathbf{d}^T\mathbf{x} \geq c^*$$

is the only element in the set of non-implied necessary conditions $\overline{\Phi}_{\mathbf{d}}^{\mathcal{X}^\epsilon}$.

Proof. Let us first show that the set of necessary conditions is equal to

$$\Phi_{\mathbf{d}}^{\mathcal{X}^\epsilon} = \{\phi_{\mathbf{d}}^c | c \leq c^*\}.$$

By definition, $c^* = \min_{\mathbf{x} \in \mathcal{X}^\epsilon} \mathbf{d}^T\mathbf{x}$ is the smallest value that $\mathbf{d}^T\mathbf{x}$ can take over $\mathcal{X}^\epsilon$. This implies that

$$\phi_{\mathbf{d}}^{c^*}(\mathbf{x}) := \mathbf{d}^T\mathbf{x} \geq c^*$$

is true for all $\mathbf{x} \in \mathcal{X}^\epsilon$.

Similarly, if $c < c^*$, we know that $\mathbf{d}^T\mathbf{x} \geq c^* > c$ is true for all $\mathbf{x} \in \mathcal{X}^\epsilon$. Thus, all conditions $\phi_{\mathbf{d}}^c$ such that $c \leq c^*$ are necessary conditions.

For $c > c^*$ however, at the optimum $\mathbf{x}_\epsilon^* = \arg\min_{\mathcal{X}^\epsilon} \mathbf{d}^T\mathbf{x}$, we have $\mathbf{d}^T\mathbf{x}_\epsilon^* = c^* < c$ which implies that the condition

$$\phi_{\mathbf{d}}^c(\mathbf{x}) := \mathbf{d}^T\mathbf{x} > c$$

is not true for all $\mathbf{x}$ in $\mathcal{X}^\epsilon$, as $\mathbf{x}_\epsilon^* \in \mathcal{X}^\epsilon$. Therefore, all conditions $\phi_{\mathbf{d}}^c$ such that $c > c^*$ are not necessary conditions.

Now let us prove $\overline{\Phi}_{\mathbf{d}}^{\mathcal{X}^\epsilon} = \{\phi_{\mathbf{d}}^{c^*}\}$. This means that all $\phi_{\mathbf{d}}^c$ with $c < c^*$ are implied by and do not imply $\phi_{\mathbf{d}}^{c^*}$. This can be shown by proving that, for any $c < c^*$,

$$\mathcal{I}_{\phi_{\mathbf{d}}^{c^*}} \subset \mathcal{I}_{\phi_{\mathbf{d}}^c} \text{ and } \mathcal{I}_{\phi_{\mathbf{d}}^c} \not\subset \mathcal{I}_{\phi_{\mathbf{d}}^{c^*}}.$$

We have $\mathcal{I}_{\phi_{\mathbf{d}}^{c^*}} \subset \mathcal{I}_{\phi_{\mathbf{d}}^c}$ because, as shown before, for any $\mathbf{x}$, if $\phi_{\mathbf{d}}^{c^*}(\mathbf{x})$ is true, $\phi_{\mathbf{d}}^c(\mathbf{x})$ with $c < c^*$ is also true. Moreover, $\mathcal{I}_{\phi_{\mathbf{d}}^c} \not\subset \mathcal{I}_{\phi_{\mathbf{d}}^{c^*}}$. Indeed, the element $\mathbf{x}$ such that $\mathbf{d}^T\mathbf{x} = c$ is an element of $\mathcal{I}_{\phi_{\mathbf{d}}^c}$ but not of $\mathcal{I}_{\phi_{\mathbf{d}}^{c^*}}$. $\square$

5. Test case

This methodology will now be applied to a specific test case. The test case is articulated around the problem of capacity expansion planning of the European electricity grid within the objective of the European Union to be carbon-neutral by 2050. Typically, the objective of this problem is to determine capacity investments in transmission, generation and storage assets as well as operation of those assets to satisfy electrical demand while minimising capital and marginal costs.

Decision-makers might be interested in knowing the necessary conditions on the required amount of capacity to be invested in each of those technologies at the European and national levels to ensure that they do not experience more than a well-specified level of cost-suboptimality.

Our methodology will be applied to this problem for computing non-implied necessary conditions for achieving ϵ-optimality on five technologies. More specifically, required minimum investments are first computed for groups of lines at the European, national and individual-line levels. Then, necessary conditions are determined for storage and RES generation, including onshore wind, offshore wind, and utility-scale PV, over the whole network.

In the following section, a short contextualisation of the test case is presented. The test case is then defined following the terms of the methodology presented above. It is followed by a short analysis of the optimal solution to the expansion planning problem before describing the necessary conditions.

5.1. Context

The geographical scope of the expansion planning problem is set to Europe. All countries in the European continent are included, except for Russia, Iceland and some small countries such as Cyprus, Malta and Liechtenstein. In this problem, the network is represented as a grid made up of nodes and lines. When applied at the European level, nodes are generally clustered by country, while lines correspond to aggregations of pre-existing or planned transmission lines between these countries. Fig. 3(a) shows the nodes and lines forming the network. In addition, generators and storage devices are attached to each of those nodes. The temporal scope of the problem is set to one full year, corresponding to the year 2050. More details on the modelling of the network can be found in Appendix.

5.2. Optimisation problem

The expansion planning problem is solved using linear optimisation via the open-source tool PyPSA [15]. In this context, the elements composing problem (1) are described briefly below.

Objective function f. The objective of the problem is to minimise the total annual system cost. To be more specific, the objective function f corresponds to the sum of annualised capital fixed costs and variable costs of generation, storage and transmission across the network.

Feasible space $\mathcal{X}$. The feasible space can be modelled as $\mathcal{X} = \{\mathbf{x} \in \mathbb{R}^n | A\mathbf{x} \geq \mathbf{b}, A \in \mathbb{R}^{m \times n}, \mathbf{b} \in \mathbb{R}^m\}$ with $m \in \mathbb{N}$ and $n \in \mathbb{N}_0$. The variables $\mathbf{x}$ correspond to investment (*i.e.* how much capacity must be added where and to what technology) and operational variables (*e.g.* which quantity of energy each generator must produce at each time step). All variables are continuous, as investments are continuous and unit commitment is not modelled. The bounds on those variables are composed via technical and physical constraints modelled as linear constraints. In addition, a constraint imposing a 99% reduction on CO_2 emissions compared to 1990 levels is added. This value is set Europe-wise and is not set to 100% to ensure the feasibility of the problem.
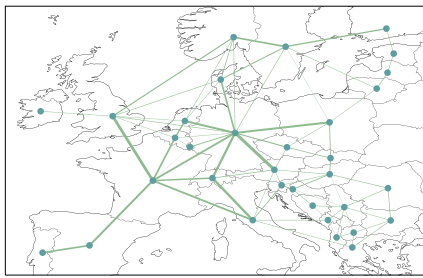
(a) Initial capacity.
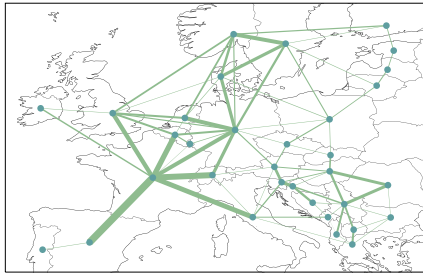
(b) Additional capacity to be cost-optimal.

Fig. 3. European electricity grid representation with the width of lines proportional to capacity in GW.

Set of conditions Φ. As in Section 4, to compute non-implied necessary conditions corresponding to minimum capacity investments, sets of conditions of form

$$\Phi_{\mathbf{d}} = \{\mathbf{d}^T \mathbf{x}_I \geq c \mid c \in \mathbb{R}\}$$

are used, with $\mathbf{x}_I$ a vector of size $|I|$ that collects the different investment variables and $\mathbf{d} \in \{0,1\}^{|I|}$. Depending on the type of investment variables for which a non-implied necessary condition is desired, it suffices to define an appropriate $\mathbf{d}$, *i.e.* whose elements corresponding to the required variables are set to 1, and others to 0.

Suboptimality coefficient ϵ. As mentioned above, necessary conditions are valid for a given value of the suboptimality coefficient ϵ. In this study, necessary conditions are computed for different values of ϵ - ranging from 0% (*i.e.* optimality) to 20% - to see how minimal capacity investment evolves with the suboptimality coefficient.

Computation of non-implied necessary conditions. As a reminder, the computation consists in the following steps:

(1) Compute an optimal solution $\mathbf{x}^*$ for problem (1).
(2) For a given suboptimal coefficient ϵ, compute an ϵ-optimal space $\mathcal{X}^\epsilon$ using this solution.
(3) For this ϵ and a value of $\mathbf{d}$, extract a non-implied necessary condition by solving $\min_{\mathbf{x} \in \mathcal{X}^\epsilon} \mathbf{d}^T \mathbf{x}$.
(4) Repeat step (2) and (3) to obtain non-implied necessary conditions for different values of ϵ and $\mathbf{d}$.

5.3. Optimal solution

Fig. 3 shows the initial topology – with the widths of lines proportional to their capacity in GW – and how much capacity should be added to be optimal (*i.e.* obtained from $\mathbf{x}^*$). Table 1 lists the optimal capacities for the technologies that are expanded in the problem. Capacities of lines are expressed in TWkm, *i.e.* the power capacity installed over a given distance.

Table 1
Optimal capacities.

TWkm			GW					
AC	DC	AC+DC	Onshore wind	Offshore wind	Utility PV	CCGT	OCGT	Li-Ion
128	90	218	168	327	367	49	0	249

For transmission, with the addition of 146 TWkm of lines, the initial 72 TWkm is tripled. Capacity is added to almost all lines, but major investments are made in some countries like Germany and France. In these particular cases, the sums of the optimal capacities of the lines connected to the country are equal to 79 TWkm (from an initial capacity of 21 TWkm) and 57 TWkm (from 23 TWkm).

The capacities of RES, including onshore wind, offshore wind and utility PV, are massive, reaching 862 GW across the three technologies. Gas plants are also deployed, though on a smaller scale due to the constraint on CO_2 emissions. Finally, a substantial amount of Li-Ion batteries is built all around Europe, with the main hub being in Spain, where a lot of PV is also deployed. Note that the initial capacity of these technologies (*i.e.*, RES, storage and gas) are null.

From this optimal solution, the ϵ-optimal spaces corresponding to the values of ϵ listed in the previous section are defined and necessary conditions for transmission, Li-Ion storage and RES generation are computed.

5.4. Necessary conditions on transmission capacity

In this case, the variables of interest correspond to transmission capacity variables, *i.e.* how much capacity – in TWkm – should be added to each line. We minimise capacity in TWkm (*i.e.* the power capacity of the line multiplied by its length) because this value is a good representative of both the physical and economic investments in transmission assets. Necessary conditions are first computed for the sum of capacities of all lines in the network. The analysis is then refined by looking at the sum of capacities of lines connected to a given country and, finally, at the capacities of unique lines.

Whole network. The first non-implied necessary conditions to be ϵ-optimal consist of the required minimum capacity to be added to the whole network. These are obtained by setting to one all values of $\mathbf{d}$ corresponding to transmission capacity investment variables. The other values of $\mathbf{d}$ are set to zero.

Fig. 4(a) shows the values of the required new capacity to sustain a certain ϵ-optimality for different values of ϵ. The required new capacity drops rapidly for small values of ϵ and, for a suboptimality coefficient ϵ of 10%, it has already been divided by a factor two compared to the optimum. However, for larger ϵ, the decrease slows down, and the necessary conditions start reaching a slowly decreasing plateau around 60 TWkm.

National lines. These first results show how much capacity should be added to the initial network to keep costs above a certain ϵ-optimality. National transmission system operators could also be interested in the minimum capacity required to connect their country to the rest of the system. It can be identified by setting all values in $\mathbf{d}$ to zero except for those corresponding to the lines connected to a given country, in which case they are set to one.

Fig. 4(b) focuses on the four countries with the highest added incumbent capacity (in TWkm) in the optimal solution. In decreasing-capacity order, those are France, Germany, Great Britain and Norway. For the two last ones, the necessary additional capacity converges to 0 TWkm for a suboptimality coefficient ϵ of less than 5% even though they had around 20 TWkm of new installed capacity in the optimal solution. France and Germany start from an even higher capacity level in the optimal solution, with 58 and 34 TWkm of added incumbent capacity. When the coefficient ϵ reaches 20%, they still have a non-zero

(a) Sum of the capacities of all lines.

(b) Sum of the capacities of country lines.

(c) Capacity of individual lines.

(d) Sum Li-Ion batteries capacities.

(e) Sum of renewable energy generators capacities (onshore wind, offshore wind, utility PV and sum of the three).
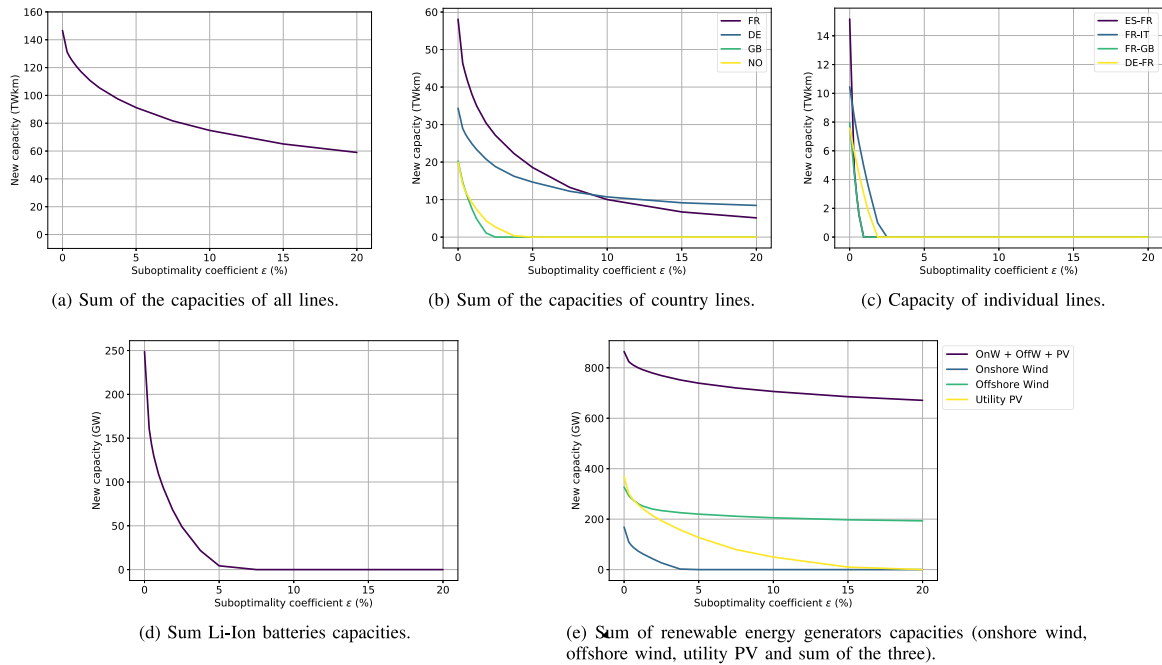
Fig. 4. Non-implied necessary conditions on new capacity for different levels of suboptimality.

additional capacity, with respectively 5 and 8 TWkm, but the decrease is greater than 90% for France and 75% for Germany.

Individual lines. Necessary conditions can be used to identify critical lines in the network. In this case, only one value in **d** is set to 1. Fig. 4(c) shows the value of the necessary conditions for the four lines with the largest capacity increase (in TWkm) in the optimal solution: ES-FR, FR-IT, FR-GB, and DE-FR. The main conclusion drawn from this graph is that no individually line needs to be necessarily expanded to avoid a suboptimality greater than 2.5%.

5.5. Necessary conditions on storage capacity

In the model, there is no pre-existing storage capacity and investment can be made at each bus. In the optimal set-up, 249 GW of Li-ion batteries are built and store 98 TWh over the simulated year. However, Fig. 4(d) shows the necessary condition reaches 0 GW for a suboptimality coefficient ϵ as small as 5%.

5.6. Necessary conditions on RES capacity

Finally, investments in renewable energy sources, including onshore and offshore wind turbines and utility-scale PV panels, are analysed. Four types of necessary conditions are computed: one per technology corresponding to the required minimum in new capacity for that technology and one for the required minimum in the sum of capacities in the three technologies. As for storage, a greenfield approach is used and the new capacity is equal to the total capacity that is installed.

Fig. 4(e) shows that investments in renewable energies are essential as the minimum capacity required to be ϵ-optimal does not drop below 600 GW even as the suboptimality coefficient ϵ rises to 20%. However, this is less clear for each RES technology individually. While the minimum requirement for offshore wind stays consistent with increasing values of the suboptimality coefficient ϵ, the necessary conditions for onshore wind and utility PV converge to 0 GW.

6. Conclusion and future work

In this paper, a framework offering a change of focus for optimisation model methods was presented and applied for capacity expansion planning. Deviating from cost-optimal focused studies, we advocate for the search of non-implied necessary conditions for ϵ-optimality to inform decision-makers efficiently.

The concepts required to define this search in a generic case were formalised. A methodology was then presented to derive necessary conditions in the specific context where conditions consist of constrained sums. Finally, to illustrate the framework, this methodology was applied to a test case related to capacity expansion planning at the European level, focusing on the minimum investments in transmission, storage and generation required for ϵ-optimality.

This work sets the ground for further developments of the presented framework. First, the framework was specified for a fixed set of parameters that define the shape of the feasible space $\mathcal{X}$ and of the objective function f. Changing the value of the parameters could thus impact $\mathcal{X}$ and f, and in turn, the optimal solution and the ϵ-optimal spaces $\mathcal{X}^\epsilon$. As a result, there is no guarantee that necessary conditions found for a fixed set of parameters would remain the same for a different set of parameters. The concept of necessary conditions could thus be extended to overcome this limitation by defining sets of *meta*-necessary conditions valid for different sets of parameters. Determining such necessary conditions and providing guarantees on implications would require more advanced techniques than the one presented in this article.

Second, in this paper, we only presented an algorithmic solution for computing non-implied necessary conditions in the context where conditions consist of constrained sums of variables. It would be interesting to propose algorithmic solutions for other types of conditions.

Third, while we focused on cost-based ϵ-optimality, this concept and the one of necessary conditions can naturally be extended to other objectives.

Finally, it would be interesting to investigate whether other fields than capacity expansion planning could benefit from the framework introduced in this paper.

CRediT authorship contribution statement

Antoine Dubois: Term, Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing – original draft, Writing – review & editing, Visualization. **Damien Ernst:** Term, Conceptualization, Writing – review & editing, Supervision.

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. Network model

Modelling and optimising the network is done using PyPSA [15] in conjunction with REplan [16].

A.1. Topology - Buses and lines

The initial topology of the network is based on the TYNDP18 2027 reference grid developed by ENTSO-E [17]. It consists of bi-directional net transfer capacities (NTC) between countries or regions inside countries. To obtain a one-node-per-country topology, nodes are clustered per country, outgoing lines capacities are summed, and intra-country lines are removed.

Connections are modelled as bi-directional links using a transportation model. The initial capacity of each line is set to the maximum of both NTCs. Lines crossing seas are considered to be HVDC cables, while other lines are represented as HVAC lines. For simulating the N-1 stability constraint, the maximum power flow across any line is set to 70% of its installed capacity (as suggested in [18]).

For keeping the expansion realistic, an upper bound is fixed on the maximum capacity per line. This upper bound is set based on the NTCs of the 'Global Climate Action 2040' scenario of TYNDP2018. However, to provide some slack to the model, this capacity is multiplied by a pre-defined factor of 3. Note that for a multiplication factor equal to 2, around 5% of the load was shed in the optimal solution.

A.2. Load

The model is solved at a 2-hourly resolution. At each time step, the load must be satisfied or, is shed for a cost of 3k€/MWh. Hourly load series per country are extracted from the Open Power System Data project [19]. The reference year used in this model is 2018.

A.3. Generation and storage technologies

The model contains generation and storage technologies. For each technology, one representative plant is used per node where the pre-existing capacity or capacity expansion potential is not null. As detailed below, some of these technologies are expandable and others are not.

Technologies with expandable capacity. Dispatchable capacity can be deployed in the form of CCGT and OCGT. They are the only technologies that produce CO_2 emissions when generating electricity.

Short-term storage can be built as Li-Ion batteries. Those batteries are characterised by two elements: their peak power capacity and the maximum duration during which they can discharge this power. In the test case, this second element is fixed to 4 h and multiplying by the peak power gives the storage capacity of the battery.

For these three technologies, no initial capacity and no upper limit on the amount of new capacity are considered.

Three types of renewable energy sources are added to the model: onshore wind generators, offshore wind generators and utility-scale PV power plants. The per-country capacity factors profiles are obtained

through Renewables.ninja, presented in [20,21], while expansion potential are computed via GLAES [22]. GLAES is parametrised such that, on a cumulative basis, a maximum of 447 GW can be built for onshore wind, 1077 GW for offshore wind and 1150 GW for PV. No initial capacity is considered. The energy produced by those generators can be curtailed without incurring any supplementary cost.

Technologies with fixed capacity. New investments in nuclear power are not considered. Generators in Belgium and Germany, and those commissioned before 1980, are removed from the model. Using the JRC Open Power Plants database [23], this leads to 94 GW of capacity which is in line with the projections made in the 2016 EU Reference Scenario [24].

Hydro-power is modelled through the addition of pre-existing run-of-river generators with a capacity of 34 GW, reservoirs with 105 GW and pumped-hydro storage with 55 GW. Capacities and locations around Europe are extracted from the JRC Hydro-power plants database [25]. For more information on the modelling of input flows, the interested reader can refer to the supplementary material of [26].

A.4. Input parameters and data

All experiments can be reproduced using the code which is available at [27]. The input data used for generating the results presented in this paper can be retrieved in [28] and is preprocessed using the open-source tool EPIPPy [29]. The repository also contains the output of the PyPSA runs and a document describing the techno-economic parameters used in the model and the sources from which they were determined.

References

[1] E.D. Brill, J.M. Flach, L.D. Hopkins, S. Ranjithan, MGA: A Decision support system for complex, incompletely defined problems, IEEE Trans. Syst. Man Cybern. 20 (4) (1990) 745–757.

[2] I. Ravetz, What is post-normal science, Futures-J. Forecast. Plan. Policy 31 (7) (1999) 647–654.

[3] X. Yue, S. Pye, J. DeCarolis, F.G. Li, F. Rogan, B.O. Gallachóir, A review of approaches to uncertainty assessment in energy system optimization models, Energy Strateg. Rev. 21 (2018) 204–217.

[4] E. Trutnevyte, Does cost optimization approximate the real-world energy transition? Energy 106 (2016) 182–193.

[5] E.D. Brill Jr., The use of optimization models in public-sector planning, Manage. Sci. 25 (5) (1979) 413–422.

[6] M. Ehrgott, Multicriteria Optimization, Vol. 491, Springer Science & Business Media, 2005.

[7] D. Brill, S.-Y. Chang, L. Hopkins, Modeling to generate alternatives: The HSJ approach and an illustration using a problem in land use planning, Manage. Sci. 28 (3) (1982) 221–235.

[8] J.F. DeCarolis, Using modeling to generate alternatives (MGA) to expand our thinking on energy futures, Energy Econ. 33 (2) (2011) 145–152.

[9] P. James, K. Ilkka, Modelling to generate alternatives: A technique to explore uncertainty in energy-environment-economy models, Appl. Energy 195 (2017) 356–369.

[10] F.G. Li, E. Trutnevyte, Investment appraisal of cost-optimal and near-optimal pathways for the UK electricity sector transition to 2050, Appl. Energy 189 (2017) 89–109.

[11] L. Nacken, F. Krebs, T. Fischer, C. Hoffmann, Integrated renewable energy systems for Germany–A model-based exploration of the decision space, in: 2019 16th International Conference on the European Energy Market (EEM), IEEE, 2019, pp. 1–8.

[12] J.-P. Sasse, E. Trutnevyte, Distributional trade-offs between regionally equitable and cost-efficient allocation of renewable electricity generation, Appl. Energy 254 (2019) 113724.

[13] F. Neumann, T. Brown, The near-optimal feasible space of a renewable power system model, Electr. Power Syst. Res. 190 (2021) 106690.

[14] S. Bandaru, A.H. Ng, K. Deb, Data mining methods for knowledge discovery in multi-objective optimization: Part A-survey, Expert Syst. Appl. 70 (2017) 139–159.

[15] T. Brown, J. Hörsch, D. Schlachtberger, PyPSA: PYthon for power system analysis, J. Open Res. Softw. 6 (4) (2018).

[16] A. Dubois, D. Radu, REplan, https://github.com/montefesp/REplan.

[17] ENTSO-E, TYNDP18. Joint Scenarios Data. Input Data, [Online]. Available:https://tyndp.entsoe.eu/maps-data.

[18] J. Hoersch, F. Hofmann, D. Schlachtberger, T. Brown, PyPSA-Eur: AN open optimisation model of the European transmission system, Energy Strateg. Rev. 22 (2018) 207–215.

[19] Open Power System Data, Data package time series. version 2019-06-05, 2019, [Online]. Available: https://data.open-power-system-data.org/time_series/2019-06-05.

[20] S. Pfenninger, I. Staffell, Long-term patterns of European PV output using 30 years of validated hourly reanalysis and satellite data, Energy 114 (2016) 1251–1265.

[21] I. Staffell, S. Pfenninger, Using bias-corrected reanalysis to simulate current and future wind power output, Energy 114 (2016) 1224–1239.

[22] D. Ryberg, M. Robinius, D. Stolten, Evaluating land eligibility constraints of renewable energy sources in europe, Energies 11 (5) (2018) 1246.

[23] A. Bocin, I. Hidalgo Gonzalez, K. Kanellopoulos, M. De Felice, JRC Open power plants database (JRC-PPDB-open). European commission, joint research centre (JRC), 2019, [Online]. Available: http://data.europa.eu/89h/9810feeb-f062-49cd-8e76-8d8cfd488a05.

[24] European Commission, EU Reference Scenario 2016 - Energy, transport and GHG emissions Trends to 2050, 68, [Online]. Available: https://ec.europa.eu/energy/sites/ener/files/documents/20160713%20draft_publication_REF2016_v13.pdf.

[25] M. De Felice, G. Peronato, K. Kavvadias, Energy-modelling-toolkit/hydro-power-database: JRC hydro-power database - release 10 (version v10), 2021, [Online]. Available: https://doi.org/10.5281/zenodo.5215920.

[26] D. Radu, M. Berger, A. Dubois, R. Fonteneau, H. Pandžić, Y. Dvorkin, Q. Louveaux, D. Ernst, Assessing the impact of offshore wind siting strategies on the design of the European power system, Appl. Energy 305 (2022) 117700.

[27] A. Dubois, D. Radu, Computing necessary conditions for near-optimality in capacity expansion planning problems - code, 2021, https://zenodo.org/record/6107173.

[28] A. Dubois, D. Radu, M. Berger, Computing necessary conditions for near-optimality in capacity expansion planning problems - dataset, 2021, https://zenodo.org/record/5519081.

[29] A. Dubois, D. Radu, EPIPPy - Expansion planning input preprocessing in python, https://github.com/montefesp/EPIPPy.