Transmission and Generation Investment in Electricity Markets: The Effects of Market Splitting and Network Fee Regimes

Veronika Grimm[a], Alexander Martin[b], Martin Schmidt[b,c], Martin Weibelzahl[b], Gregor Zöttl[d]

[a]*Friedrich-Alexander-Universität Erlangen-Nürnberg, Chair of Economic Theory, Lange Gasse 20, 90403 Nürnberg, Germany*
[b]*Friedrich-Alexander-Universität Erlangen-Nürnberg, Discrete Optimization, Cauerstr. 11, 91058 Erlangen, Germany*
[c]*Energie Campus Nürnberg, Fürther Str. 250, 90429 Nürnberg, Germany*
[d]*Friedrich-Alexander-Universität Erlangen-Nürnberg, Chair of Regulation and Energy Markets, Lange Gasse 20, 90403 Nürnberg, Germany*

Abstract

We propose a equilibrium model that allows to analyze the long-run impact of the regulatory environment on transmission line expansion by the regulator and investment in generation capacity by private firms in liberalized electricity markets. The model incorporates investment decisions of the transmission operator and private firms in expectation of an energy-only market and cost-based redispatch. In different specifications we consider the cases of one vs. multiple price zones (market splitting) and analyze different approaches to recover network cost—in particular lump sum, generation capacity based, and energy based fees. In order to compare the outcomes of our multistage market model with a first best benchmark, we also solve the corresponding integrated planner problem. Using two test networks we illustrate that energy-only markets can lead to suboptimal locational decisions for generation capacity and thus imply excessive network expansion. Market splitting heals these problems only partially. These results are valid for all considered types of network tariffs, although investment slightly differs across those regimes.

Keywords: Equilibrium Models, Mixed-Integer Nonlinear Optimization, Multilevel Programming, Electricity Markets, Network Expansion, Transmission Management
2010 MSC: 90B10, 90C11, 90C33, 90C35, 90C90, 91-08, 91B26, 91B74

1. Introduction

Following the British privatization in the 1980s, various countries around the world liberalized their electricity sectors. Today, in most industrialized countries only the transmission network remains regulated while private firms decide on investment in generation capacities and trade electricity on markets. This structure challenges the planning of transmission and generation capacity expansion. While an entirely regulated electricity sector allows for simultaneous transmission and generation expansion planning, in a liberalized market, investment decisions in transmission and generation capacities are taken by different agents. Investment in generation capacities is typically made by firms and private investors based on their expectations concerning the future regulatory environment, taking into account available network facilities. Network expansion, however, is decided on by regulated firms (or even the regulator), in anticipation of capacity investments by private firms. Traditional optimization approaches, which only consider integrated transmission and generation expansion planning, reveal the optimal expansion plan for transmission and generation but do not offer valuable information on how to achieve those goals in a market environment. In a liberalized market, incentives induced by the interplay of the market environment and regulation determine whether firms make the appropriate investment choices. As our results clearly reveal, the proper design of market rules providing adequate incentives in those markets crucially

Email addresses: veronika.grimm@fau.de (Veronika Grimm), alexander.martin@fau.de (Alexander Martin), mar.schmidt@fau.de (Martin Schmidt), martin.weibelzahl@fau.de (Martin Weibelzahl), gregor.zoettl@fau.de (Gregor Zöttl)

matters. Liberalized electricity markets thus call for new tools to inform the various agents involved: regulators, electricity firms, investors, and other stakeholders.

In this paper we propose a model that allows to analyze investment decisions by the regulator and private firms in liberalized electricity markets. We model energy-only markets and a regulated transmission operator (TO) who uses cost-based redispatch to deal with transmission constraints. In a multistage analysis we study transmission expansion decisions by the regulated TO in anticipation of capacity expansion by private firms. In different instantiations of our model we analyze the effects of market splitting (one vs. multiple price zones) as well as different approaches to recover network cost—in particular a lump sum, a generation capacity based, and an energy based fee. In order to compare the outcomes to a first best benchmark we also solve the integrated planner problem. For the computational studies we restrict ourselves to solving stylized test cases to illustrate the applicability of our framework. The results demonstrate that investment choices in a market environment substantially differ from a first best solution. In our numerical examples the absence of proper locational investment incentives for firms clearly affects investment choices of generators, which, in turn, leads to excessive line investment. This shows that our model allows to compare different network management regimes and to quantify their effects on long-run investment decisions. Our approach is, thus, an important extension of various studies that have mainly considered the short-run properties of different transmission management regimes; see the literature review below. As we show, transmission management has also important implications in the long-run when generation and transmission expansion are taken into account.

Let us emphasize that our approach allows to assess the long-run impact of different transmission management regimes adopted in liberalized electricity markets around the world. Especially in Europe spot market trading does not fully account for transmission constraints. In contrast, capacities are shut down and called by the TO in case that the spot market solution is not technically feasible. Under cost-based redispatch (as it is used in Austria, Switzerland, or Germany) firms called into operation are just compensated for their variable production cost. Consequently, redispatch operations cannot contribute to the recovery of investment cost.[1] Other liberalized electricity markets adopted a system of nodal prices (see, e.g., [2]), where spot market prices directly reflect transmission constraints (e.g., United States, Canada, Australia, or New Zealand), which induces more adequate incentives for generation capacities by private firms. To at least partially overcome the lack of locational signals provided by spot market prices, several countries that rely on a system of redispatch have introduced price zones (e.g., Sweden and Italy). Since the first best solution coincides with the outcome obtained under nodal pricing in our framework, our approach also allows to assess the long-run benefits of a change to this transmission management system.

1.1. Literature Review

Prior to the liberalization of electricity sectors around the world, vertically integrated monopolists (either regulated or directly state owned) were responsible for generation and transmission expansion. Such monopolists needed insights on the cost minimal configuration of the system. As a consequence, traditionally most of the contributions proposed frameworks and techniques to determine overall optimal expansion for generation and transmission facilities; see, e.g., [3], [4], [5], or [6].

In liberalized electricity markets, however, we observe vertical unbundling of transmission and generation facilities. Thus, in addition to insights on the global optimum of an integrated monopolist, research is needed on how the market environment affects decisions of different stakeholders. By now a large literature has developed which analyzes incentives for private and potentially strategic firms to invest in generation facilities. However, these studies typically assume unlimited transmission capacity; examples are [7], [8], [9], [10], [11], [12], or [13].

[1] In contrast, market-based redispatch (used, e.g., in Belgium, Finland, France, or Sweden) yields rents for firms that are called at the redispatch stage and thus induces incentives to build plants at locations with systematic underprovision. Note that this holds only true for the case of cost-based redispatch as analyzed in the present article. However, market-based redispatch is plagued by severe gaming problems, which obtain even for perfectly competitive markets. In the literature this is also often referred to as the inc-dec game. For a discussion of these issues see, e.g., [1].

Another recent strand of literature explicitly models both, generation and transmission investment, typically by making use of bilevel models. [14, 15] are among the first to model investment incentives of generators and transmission network expansion in a such a way. In their contribution they quantify the impact of whether transmission investment anticipates resulting investment of strategic generation companies or not. [16] propose a simulation framework to analyze investment of competitive generation companies and competitive merchant transmission companies. [17] generalize this framework to also include a transmission system operator as a further agent. [18] provide a two-stage model of transmission expansion facing uncertain investment in renewable generation.

[19] and, in an extension, [20] analyze expansion of electricity generation and transmission capacities together with the expansion of a fuel transportation network. For electricity markets, [21] propose a clear-cut bilevel framework which considers optimal network expansion by the transmission company, anticipating investment of competitive generation companies. Also based on a bilevel approach, [22] propose an auction mechanism to implement optimal investment incentives by transmission companies. Those approaches, however, do not explicitly take into account the specific structure of the transmission management regime, but implicitly assume optimal management of the transmission network which is implemented by a regime of locational marginal prices; see [23]. While this models incentives in markets where indeed a system of locational marginal prices is implemented (as, e.g., in the US or Canada), it limits insights with respect to other markets, which might rely on a system of market splitting or coupling and redispatch, which is not captured by the approaches mentioned above. It is the purpose of this article to explicitly analyze the impact of specific design features of transmission management regimes as market splitting and redispatch on the generation and transmission investment incentives.

Let us finally note that, in recent years, an extensive literature has been developed, which analyzes the impact of specific rules of the transmission management regime on short-run market outcomes, i.e., for fixed generation and transmission facilities. Prominent articles include [24], [25], [26], or [27] who compare the short-run implications of zonal systems with redispatch to the system of nodal pricing. Several articles analyze the incentives of different agents that are able to exercise market power under different transmission management regimes; see, e.g., [28], [29], [30], [31], [32], or [33]. Recently, [34], [35], [36], and [37] compared different transmission management regimes based on market coupling or splitting with redispatch. All those articles consider only the short-run perspective, while it is our purpose to consider the long-run effects on investment incentives.

1.2. Outline of the Paper

This paper is organized as follows. Sect. 2 presents the basic economical and technical quantities that are used throughout the paper. In Sect. 3 we introduce the integrated planner approach, while Sect. 4 presents the trilevel model with a cost-based redispatch system and market splitting. The trilevel model is reformulated using novel ideas in Sect. 5 and tailored solution strategies for the reformulated model are given in Sect. 6. The last part of our paper, Sect. 7, presents computational results for test networks that illustrate the applicability of our model. Finally, Sect. 8 concludes.

2. Basic Economical and Technical Setup

In this section we present the basic notation that is used throughout the paper. For the sake of completeness, we present all quantities used in our models in Appendix AppendixA.

2.1. Network Model

We consider an electricity transmission network $\mathcal{G} = (N, L^{\text{ex}})$ with a set of nodes $N = \{n_1, \dots, n_{|N|}\}$ and a set of existing transmission lines L^{ex}. By $\mathcal{L}$ we denote different line types that are characterized by their capacity $\bar{f}$ and their susceptance B. Given different line types $\ell \in \mathcal{L}$, the network operator decides on an optimal network expansion plan, i.e., on the construction of candidate power lines $l \in L^{\text{new}} = \bigcup_{\ell \in \mathcal{L}} L^{\text{new}}_{\ell}$, and on the degradation of existing lines $l \in L^{\text{ex}} = \bigcup_{\ell \in \mathcal{L}} L^{\text{ex}}_{\ell}$. The set of all lines, i.e., existing and candidate transmission lines, is denoted by $L := L^{\text{ex}} \cup L^{\text{new}}$. Throughout the paper we make use of the standard δ-notation, i.e., the set of in- and outgoing edges of a node set $\tilde{N}$ is given by $\delta^{\text{in}}_{\tilde{N}}(\tilde{L})$ or $\delta^{\text{out}}_{\tilde{N}}(\tilde{L})$, respectively, where the line set $\tilde{L}$ is used to denote the set of considered arcs. More formally, we have

$$\delta^{\text{in}}_{\tilde{N}}(\tilde{L}) := \{l \in \tilde{L} \colon l = (n,m) \text{ with } m \in \tilde{N}\}, \quad \delta^{\text{out}}_{\tilde{N}}(\tilde{L}) := \{l \in \tilde{L} \colon l = (n,m) \text{ with } n \in \tilde{N}\}.$$

Our trilevel power market model allows us to account for multiple price zones Z that are given as parts of a partition $N = Z_1 \cup \cdots \cup Z_{|Z|}$ of the node set. In our context, the parts of this partition refer to different price zones in the electricity market. By $L^{\text{inter}}_{\ell} \subseteq L^{\text{ex}}_{\ell} \cup L^{\text{new}}_{\ell}$ we denote inter-zone lines of type ℓ connecting nodes that belong to different zones. As before, $L^{\text{inter}} = \bigcup_{\ell \in \mathcal{L}} L^{\text{inter}}_{\ell}$ denotes the set of inter-zone lines of all types.

Line investment cost are denoted by c^{inv}_l and line degradation cost by c^{del}_l. When making line investment decisions, the TO faces physical network constraints known as Kirchhoff's first and second law. Throughout the paper we use a linear approximation of real power flows known as the lossless direct current (DC) power flow approximation; see, e.g., [38].

2.2. Electricity Demand and Time Horizon

The set of demand nodes is denoted by $N^{\text{dem}} \subseteq N$. Consumers are located exclusively at these nodes, i.e., demand is zero at any other node $n \in N \setminus N^{\text{dem}}$. We further assume a given equidistantly discretized time horizon $T = \{t_1, \dots, t_{|T|}\}$ with time steps $\tau = t_{k+1} - t_k$ for all $k = 1, \dots, |T| - 1$. Elastic demand at demand node $n \in N^{\text{dem}}$ in time period $t \in T$ is modeled by a continuous, strictly decreasing function

$$p_{t,n}(d_{t,n}) : [0, \bar{d}_{t,n}] \to \mathbb{R}^+.$$

Here and in what follows, $d_{t,n}$ denotes demand and $p_{t,n}(d_{t,n})$ is the resulting market price. The saturation point $\bar{d}_{t,n}$ is the unique positive root of $p_{t,n}$. Note that the gross consumer surplus, which is the integral of $p_{t,n}$ over $[0, \bar{d}_{t,n}]$, is a concave function in our case.

2.3. Investment, Production, and Supply

For a given network node $n \in N$, G^{all}_n denotes a finite set of existing technologies and candidate technologies that firms can invest in. We use the set G^{ex}_n for already existing generation technologies and the set G^{new}_n for candidate generation technologies. Thus, $G^{\text{all}}_n = G^{\text{new}}_n \cup G^{\text{ex}}_n$ holds. To account for the characteristics of different production technologies, we allow for so-called equivalent availabilities $\alpha_g \in [0, 1]$ for every $g \in G^{\text{all}}_n$ and $n \in N$. Investment cost of building new generation capacity is denoted by $c^{\text{inv}}_g \in \mathbb{R}^+$.

We assume that all firms act in a competitive environment without any type of market power and act as price takers; see also Sect. 4.2. Variable production cost is denoted by $c^{\text{var}}_g \in \mathbb{R}^+$ for all $g \in G^{\text{all}}_n$ and $n \in N$. In addition, we assume that all variable cost c^{var}_g are pairwise distinct since this is needed to ensure unique spot market solutions; see Sect. 5.

2.4. Network Fees

In our power market model the TO has to collect network fees in order to cover expenses arising from line investment and redispatch. We denote the revenues from these network fees by R and consider three different types of network fee regimes:

- We denote by φ^{ls} the lump sum fee that is paid by the consumers. The corresponding revenues are given by $R^{\text{ls}} = \varphi^{\text{ls}}$.
- We denote by φ^{eb} the energy based fee, i.e., a per unit fee charged for each unit of electricity traded on the spot market. Let $d^{\text{spot}}_{t,n}$ describe spot market demand at node n in time period t. Then, the corresponding revenue is given by

$$R^{\text{eb}} = \varphi^{\text{eb}} \sum_{n \in N^{\text{dem}}} \sum_{t \in T} d^{\text{spot}}_{t,n}.$$

- We denote by φ^{gcb} the generation capacity based fee, i.e., a per unit fee charged for each unit of generation capacity connected to the network. If $\bar{y}_g^{\mathrm{ex}}$ denotes the capacity of an existing generator and $\bar{y}_g^{\mathrm{new}}$ the capacity of a newly installed generator, the corresponding revenues are

$$R^{\mathrm{gcb}} = \varphi^{\mathrm{gcb}} \sum_{n \in N} \left(\sum_{g \in G_n^{\mathrm{new}}} \bar{y}_g^{\mathrm{new}} + \sum_{g \in G_n^{\mathrm{ex}}} \bar{y}_g^{\mathrm{ex}} \right).$$

3. The Integrated Planner Approach as a First Best Benchmark

As a first best benchmark we consider the integrated planner approach where an integrated generation and transmission company (IGTC) simultaneously decides on transmission and generation capacity expansion and chooses welfare maximizing production at the spot markets. There are several formulations in the literature that analyze integrated planner solutions, which might serve as a benchmark in our setting. The formulation chosen here is closely related to the clear-cut formulation of the integrated planner model in [21]. The IGTC maximizes total social welfare which is defined as the difference of gross consumer surplus (aggregated over all demand scenarios) and line investment cost, line degradation cost as well as generation capacity investment cost and variable cost of production. Thus, the objective function of the IGTC reads

$$\begin{aligned} \psi_{\mathrm{IGTC}} := \sum_{n \in N^{\mathrm{dem}}} \sum_{t \in T} \int_0^{d_{t,n}} p_{t,n}(\xi) \,\mathrm{d}\xi - \sum_{l \in L^{\mathrm{new}}} c_l^{\mathrm{inv}} z_l - \sum_{l \in L^{\mathrm{ex}}} c_l^{\mathrm{del}} z_l \\ - \sum_{n \in N} \left(\sum_{g \in G_n^{\mathrm{new}}} c_g^{\mathrm{inv}} \bar{y}_g^{\mathrm{new}} + \sum_{g \in G_n^{\mathrm{all}}} \sum_{t \in T} c_g^{\mathrm{var}} y_{t,g} \right), \end{aligned}$$

where z_l are binary decision variables that decide whether a new line is build or an existing line is degraded. Moreover, $y_{t,g}$ is the actual production of generation technology g at node n. It can be shown that this approach yields the same investment and production outcomes as an idealized nodal pricing system.[2] In what follows, we introduce the constraints that the IGTC is facing. Kirchhoff's first law ensures power balance at any node $n \in N$ in the electricity network. Thus, for every time period $t \in T$ the total power flow out of node $n \in N$ and into that node have to be balanced with respect to total production and demand and that node:

$$d_{t,n} = \sum_{g \in G_n^{\mathrm{all}}} y_{t,g} + \sum_{l \in \delta_n^{\mathrm{in}}(L)} f_{t,l} - \sum_{l \in \delta_n^{\mathrm{out}}(L)} f_{t,l} \quad \text{for all } n \in N, t \in T. \tag{1}$$

Note that summing up (1) for all network nodes yields the market clearing condition

$$\sum_{n \in N^{\mathrm{dem}}} d_{t,n} = \sum_{n \in N} \sum_{g \in G_n^{\mathrm{all}}} y_{t,g} \quad \text{for all } t \in T. \tag{2}$$

Kirchhoff's second law determines the voltage angles in the network:

$$-M_l z_l \le f_{t,l} - B_l(\theta_{t,n} - \theta_{t,m}) \le M_l z_l \quad \text{for all } l = (n,m) \in L^{\mathrm{ex}}, t \in T, \tag{3a}$$

$$-M_l(1 - z_l) \le f_{t,l} - B_l(\theta_{t,n} - \theta_{t,m}) \le M_l(1 - z_l) \quad \text{for all } l = (n,m) \in L^{\mathrm{new}}, t \in T. \tag{3b}$$

[2] It is well known that the integrated planner approach yields the same outcome as a nodal price system in the short-run; see, e.g., [23]. Analogously, our results in Sect. 5 imply that the solution of the long-run integrated planner approach is equivalent to the outcome of a trilevel nodal pricing model which also accounts for investments in transmission and generation capacity. In such a trilevel nodal pricing model, a regulated TO decides on transmission expansion at the first stage. At the second stage, competitive firms decide on generation expansion investment and spot market bids. At the third stage, the TO decides on the welfare maximizing feasible allocation and implements nodal prices.

Here and in what follows, M denotes a sufficiently large constant. In order to obtain physically unique solutions, we have to fix the voltage angle at an arbitrary reference node $n_1 \in N$ in every time period:

$$\theta_{t,n_1} = 0 \quad \text{for all } t \in T. \tag{4}$$

Furthermore, transmission flows are limited by lower and upper bounds. Here, line investment is explicitly taken into account:

$$-(1 - z_l)\bar{f}_l \leq f_{t,l} \leq (1 - z_l)\bar{f}_l \quad \text{for all } l \in L^{\text{ex}}, t \in T, \tag{5a}$$

$$-z_l\bar{f}_l \leq f_{t,l} \leq z_l\bar{f}_l \quad \text{for all } l \in L^{\text{new}}, t \in T. \tag{5b}$$

The next group of constraints ensures that electricity production does not exceed installed generation capacity with respect to the equivalent availabilities of the considered technologies:

$$y_{t,g} \leq \alpha_g \tau \bar{y}_g^{\text{ex}} \quad \text{for all } n \in N, g \in G_n^{\text{ex}}, t \in T, \tag{6a}$$

$$y_{t,g} \leq \alpha_g \tau \bar{y}_g^{\text{new}} \quad \text{for all } n \in N, g \in G_n^{\text{new}}, t \in T. \tag{6b}$$

Finally, we have to impose simple bounds on the variables of the ITGC model:

$$\bar{y}_g^{\text{new}} \geq 0 \quad \text{for all } n \in N, g \in G_n^{\text{new}}, \tag{7a}$$

$$y_{t,g} \geq 0 \quad \text{for all } n \in N, g \in G_n^{\text{all}}, t \in T, \tag{7b}$$

$$d_{t,n} \geq 0 \quad \text{for all } n \in N^{\text{dem}}, t \in T, \tag{7c}$$

$$z_l \in \{0, 1\} \quad \text{for all } l \in L^{\text{ex}} \cup L^{\text{new}}. \tag{7d}$$

In summary, the IGTC faces the following inter-temporal mixed-integer maximization problem with linear constraints and a concave objective:

$$\begin{aligned}
\max \quad & \psi_{\text{IGTC}} && \text{(8a)}\\
\text{s.t.} \quad & \text{Kirchhoff's first and second law: (1), (3),} && \text{(8b)}\\
& \text{voltage phase angle at reference node: (4),} && \text{(8c)}\\
& \text{transmission flow limits: (5),} && \text{(8d)}\\
& \text{generation capacity limits: (6),} && \text{(8e)}\\
& \text{variable restrictions: (7).} && \text{(8f)}
\end{aligned}$$

We finally note that it might be also possible to incorporate additional security of supply constraints, both for line and generation investment. However, we neglect this aspect here and in what follows for the ease of presentation.

4. The Trilevel Power Market Model

In most European countries spot market trading does not fully account for network constraints (if at all). Therefore, congestion management mechanisms and the way network fees are collected play a crucial role for investment incentives—both in network and generation capacity expansion. It is most likely that energy-only markets with redispatch do not result in optimal incentive structures. Thus, one is faced with two main questions: How large is the difference to the first best solution and do there exist alternative mechanisms that have the capability to improve the situation substantially?

In our power market model we consider an energy-only market with cost-based redispatch. Electricity trading and redispatch is organized as follows in such a system: Firms trade electricity day ahead (and possibly intra-day) at a power exchange that does not account for any transmission constraints—or at least not for all of them as in the case of market splitting. After closing the market, the TO checks feasibility of the resulting

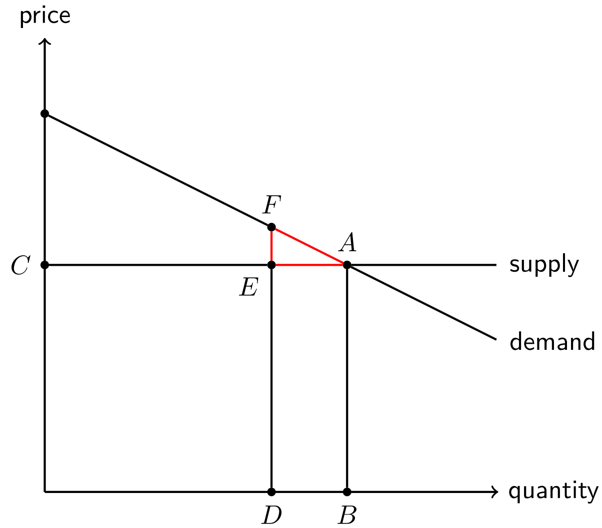

Figure 1: Economic equilibrium of cost-based redispatch

transmission flows. If the allocation is feasible, nothing is changed and electricity is generated and consumed as traded. If transmission is infeasible, the TO redispatches plants and consumers in the cheapest possible way that ensures feasibility of transmission flows. To that aim the TO obliges some producers to (partially) shut down production, while others are called to step in instead, or asks consumers to modify their demand. Plants that are shut down have to pay their variable cost, which they save due to the shutdown, to the TO. Plants that are called to step in receive their variable production cost. Production cost of the called plants are necessarily higher than production cost of the plants that are shut down, since otherwise they would have been successful on the spot market. The resulting cost is collected by the TO through network fees. While in the early days of liberalized electricity markets redispatch was a rare event, nowadays the phenomenon becomes more and more important. See, e.g., [39] for the case of Germany, where the decision to shut down nuclear power plants and to increase the generation capacity of renewables implies a much more uncertain and regionally dispersed supply structure. Figure 1 illustrates the cost-based redispatch mechanism on a stylized two-node example. Here, linear demand is located at one node and production is hosted at the other node. We further assume that there is only one generator with constant variable production cost. As illustrated in Fig. 1, in the case of unlimited transmission capacities the equilibrium quantity B will be produced. However, given physical transmission constraints with D describing the maximum amount of electricity that can be transmitted between the two nodes, the TO obliges consumers to step back from their contracts. As a compensation, the TO pays these consumers the amount equal to the area of the polygon $ABDF$. Additionally, the TO asks firms (typically those with the highest variable cost) to shut down their production. The respective firms pay an amount equal to the area of the rectangle $ABDE$ to the TO, which corresponds exactly to the variable cost that do not arise due to redispatch. Obviously, the redispatch mechanism has no impact on the firms' profits when they are asked to decrease their production. In contrast, the TO faces additional cost equal to the area of the triangle AEF. These additional cost must be collected from the market participants through network fees.

Let us now briefly sketch the structure of our model before we introduce the details at all levels. We consider a trilevel problem where the TO decides about investment in network expansion anticipating an energy-only market with cost-based redispatch. The timing of this game is illustrated in Fig. 2. Note that investment choices are taken once (sequentially by the TO and competitive firms), followed by multiple periods of spot market trading and redispatch (in the case of congestion). We translate this game into a trilevel model as follows: At the first level, the TO decides to invest in line expansion or degradation, anticipating the outcomes at all subsequent levels. The objective of the TO is to maximize welfare. At the second level, we model investment decisions

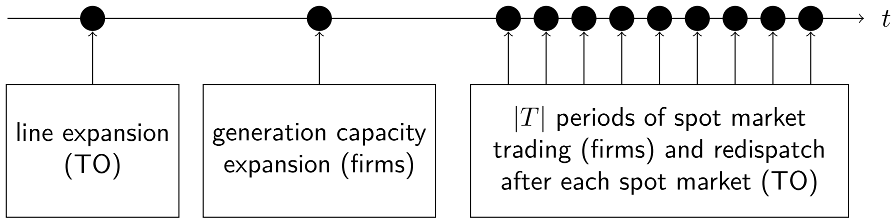

Figure 2: Timing of the multistage game

of private firms in generation capacity as well as trading at a sequence of $|T|$ spot markets with fluctuating demand. In contrast to the TO who is driven by welfare concerns, firms take investment and supply decisions as to maximize profits. We assume that spot market rules yield no price signals within price zones and assume competitive spot markets (see Sect. 4.2 for a detailed discussion of the latter assumption). The redispatch of the $|T|$ spot markets is modeled in the third level and anticipated by firms when they decide on investment and supply at level two. Redispatch occurs whenever traded quantities turn out to be infeasible subject to transmission constraints. Note that consideration of redispatch in a separate third stage is possible since, once network and generation capacities have been chosen, redispatch in time period t cannot have any impact on supply decisions at any later point in time.

We point out that all levels of our power market model are interconnected: the investment of firms and optimal spot market behaviour as well as the redispatch market are part of the TO's constraints at the first level.

Let us finally emphasize that our power market model allows to account for multiple price zones Z, which enables us to investigate the effect of market splitting under cost-based redispatch. Under market splitting, spot market trading takes network constraints between, but not within, zones into account.

In order to formally state the models, we have to introduce some more notation. In what follows, a superindex "spot" indicates quantities after spot market trading and a superindex "redi" denotes quantities after redispatch. Lastly, Δ's denote redispatch quantities, i.e., the difference of the quantity after redispatch and after spot market trading. For instance, $d_{t,n}^{\text{redi}} = d_{t,n}^{\text{spot}} + \Delta d_{t,n}$ specifies the relationship between demand after spot market trading and the actual demand after redispatch.

4.1. First-Level Problem: Optimal Line Expansion

At the first level, the TO decides on a line expansion plan as to maximize welfare, which is given as the difference of gross consumer surplus from all markets and investment and generation costs:

$$\begin{aligned}\psi_1 := \sum_{n \in N^{\text{dem}}} \sum_{t \in T} \int_0^{d_{t,n}^{\text{redi}}} p_{t,n}(\xi)\,\mathrm{d}\xi - \sum_{l \in L^{\text{new}}} c_l^{\text{inv}} z_l - \sum_{l \in L^{\text{ex}}} c_l^{\text{del}} z_l \\ - \sum_{n \in N} \left(\sum_{g \in G_n^{\text{new}}} c_g^{\text{inv}} \bar{y}_g^{\text{new}} + \sum_{g \in G_n^{\text{all}}} \sum_{t \in T} c_g^{\text{var}} y_{t,g}^{\text{redi}} \right).\end{aligned} \tag{9}$$

The TO is restricted by the budget constraint $E = R$, requiring that expenses

$$E = \sum_{n \in N^{\text{dem}}} \sum_{t \in T} \int_{d_{t,n}^{\text{redi}}}^{d_{t,n}^{\text{spot}}} p_{t,n}(\xi)\,\mathrm{d}\xi + \sum_{n \in N} \sum_{g \in G_n^{\text{all}}} \sum_{t \in T} c_g^{\text{var}} \Delta y_{t,g} + \sum_{l \in L^{\text{new}}} c_l^{\text{inv}} z_l + \sum_{l \in L^{\text{ex}}} c_l^{\text{del}} z_l \tag{10}$$

for network expansion and degradation (third and fourth term) and redispatch (first and second term) are covered by revenues R from network fees φ. Thus, revenues R are a function of the collected fee φ, which makes the latter a first-level variable. Note that redispatch cost is composed of remuneration of consumers that cannot be supplied (first term) as well as transfers to (or from) plants that are redispatched (second term). As the TO can decide to reduce spot market production or demand, $\Delta y_{t,g}$ and $\Delta d_{t,n}$ may also become negative. All together, we obtain the following first-level problem:

$$\max \quad \psi_1 \quad \text{s.t.} \quad E = R, \quad z_l \in \{0, 1\} \quad \text{for all } l \in L^{\text{ex}} \cup L^{\text{new}}.$$

4.2. Second-Level Problem: Generation Investment and Spot Market Behavior

At the second level we model the behavior of firms with respect to generation investment and spot market trading. The wholesale electricity market is assumed to be perfectly competitive, i.e., no firm can directly affect prices by strategic investment or supply decisions. It has been shown in the literature that in the absence of transmission constraints a perfectly competitive environment yields welfare maximizing investment and production decisions in our context; see, e.g., [13]. We are aware of the issue that the assumption of perfect competition may not be adequate for power systems in general. However, this assumption is necessary in order to keep the multilevel problem tractable, both theoretically and computationally: For the case of strategic firms investing in several technologies, it has been shown that uniqueness of equilibria does not hold for a reasonable set of assumptions; see [12]. As a consequence, this is assumption has been established as a standard in the literature; see, e.g., [40] or [41].

When making investment and supply decisions, firms only consider physical constraints for which they receive price signals. If the market is not divided into zones, firms receive no signals concerning network capacities and thus, will not account for them. If the market is divided into two or more zones, firms consider those physical constraints which are expressed in price differences due to market splitting: Between any pair of zones, electricity flow cannot exceed the maximum capacity of the respective inter-zone network links—and congestion implies price differences across zones. This is modeled by the following zonal version of Kirchhoff's first law

$$\sum_{n \in N^{\text{dem}} \cap Z_k} d_{t,n}^{\text{spot}} = \sum_{n \in N \cap Z_k} \sum_{g \in G_n^{\text{all}}} y_{t,g}^{\text{spot}} + \sum_{l \in \delta_{Z_k}^{\text{in}}(L)} f_{t,l}^{\text{spot}} - \sum_{l \in \delta_{Z_k}^{\text{out}}(L)} f_{t,l}^{\text{spot}} \tag{11}$$

for all $t \in T, Z_k \in Z$, and market splitting flow restrictions

$$-(1 - z_l)\bar{f}_l \leq f_{t,l}^{\text{spot}} \leq (1 - z_l)\bar{f}_l \quad \text{for all } l \in L^{\text{inter}} \cap L^{\text{ex}}, t \in T, \tag{12a}$$

$$-z_l \bar{f}_l \leq f_{t,l}^{\text{spot}} \leq z_l \bar{f}_l \quad \text{for all } l \in L^{\text{inter}} \cap L^{\text{new}}, t \in T. \tag{12b}$$

Other physical restrictions like Kirchhoff's second law are not considered.[3] In addition, we have variable restrictions in analogy to (7), i.e.,

$$\bar{y}_g^{\text{new}} \geq 0 \quad \text{for all } n \in N, g \in G_n^{\text{new}}, \tag{13a}$$

$$y_{t,g}^{\text{spot}} \geq 0 \quad \text{for all } n \in N, g \in G_n^{\text{all}}, t \in T, \tag{13b}$$

$$d_{t,n}^{\text{spot}} \geq 0 \quad \text{for all } n \in N^{\text{dem}}, t \in T. \tag{13c}$$

To summarize, at level two we consider welfare maximizing generation investment and supply decisions, i.e.,

$$\psi_2 := \sum_{n \in N^{\text{dem}}} \sum_{t \in T} \int_0^{d_{t,n}^{\text{spot}}} p_{t,n}(\xi) \,\mathrm{d}\xi - \sum_{n \in N} \left(\sum_{g \in G_n^{\text{new}}} c_g^{\text{inv}} \bar{y}_g^{\text{new}} + \sum_{g \in G_n^{\text{all}}} \sum_{t \in T} c_g^{\text{var}} y_{t,g}^{\text{spot}} \right) - R, \tag{14}$$

where supply is constrained by generation capacities installed, and transmission constraints across zones. Note that in (14), R again denotes revenues of the TO, which affect demand and generation decisions on the spot

[3] Note that analogous to flow-based market coupling, we could also consider Kirchhoff's second law on the reduced network among zones. Our framework would be perfectly suited for such an analysis. However, due to the overall length of this article we refrain from this aspect.

Table 1: Variables of the trilevel model

first-level variables:	z_l, φ
second-level variables:	$d_{t,n}^{\text{spot}}, y_{t,g}^{\text{spot}}, f_{t,l}^{\text{spot}}, \bar{y}_g^{\text{new}}$
third-level variables:	$d_{t,n}^{\text{redi}}, y_{t,g}^{\text{redi}}, f_{t,l}^{\text{redi}}, \theta_{t,l}^{\text{redi}}, \Delta y_{t,g}, \Delta d_{t,n}, \Delta f_{t,l}$
multilevel variables:	E, R

market. Thus, the second-level problem reads

$$\begin{aligned} \max \quad & \psi_2 \\ \text{s.t.} \quad & \text{generation capacity limits: (6)}, \\ & \text{zonal version of Kirchhoff's first law: (11)}, \\ & \text{market splitting flow restrictions: (12)}, \\ & \text{variable restrictions: (13)}, \end{aligned}$$

where we replaced $y_{t,g}$ by $y_{t,g}^{\text{spot}}$ in (6).

We finally point to the following observations regarding the zonal version of Kirchhoff's first law: Summing up (11) for all zones we obtain the market clearing condition (2). In a model with redispatch and only one zone, (11) coincides with a standard market clearing condition. In the case where every zone consists of exactly one network node, (11) coincides with Kirchhoff's first law ensuring power balance at every network node. Intermediate cases require the market to clear within each zone, accounting for possible transmission constraints across zones through differences in the respective market clearing prices.

4.3. Third-Level Problem: Optimal Redispatch

At the third level, the TO simultaneously decides on redispatch for all $|T|$ spot markets. Reallocation of spot market outcomes is realized in a way that ensures feasibility with respect to transmission constraints at lowest costs. These costs are given by

$$\psi_3 := \sum_{n \in N^{\text{dem}}} \sum_{t \in T} \int_{d_{t,n}^{\text{redi}}}^{d_{t,n}^{\text{spot}}} p_{t,n}(\xi) \, \mathrm{d}\xi + \sum_{n \in N} \sum_{g \in G_n^{\text{all}}} \sum_{t \in T} c_g^{\text{var}} \Delta y_{t,g}$$

and the redispatch decision has to account for all physical transmission constraints and generation capacity limits:

$$\begin{aligned} \min \quad & \psi_3 \\ \text{s.t.} \quad & (1), (3)\text{–}(6), \\ & d_{t,n}^{\text{redi}} = d_{t,n}^{\text{spot}} + \Delta d_{t,n} \quad \text{for all } n \in N^{\text{dem}}, t \in T, \\ & y_{t,g}^{\text{redi}} = y_{t,g}^{\text{spot}} + \Delta y_{t,g} \quad \text{for all } n \in N, g \in G_n^{\text{all}}, t \in T, \\ & f_{t,l}^{\text{redi}} = f_{t,l}^{\text{spot}} + \Delta f_{t,l} \quad \text{for all } l \in L^{\text{inter}}, t \in T, \\ & d_{t,n}^{\text{redi}} \geq 0 \quad \text{for all } n \in N^{\text{dem}}, t \in T, \\ & y_{t,g}^{\text{redi}} \geq 0 \quad \text{for all } n \in N, g \in G_n^{\text{all}}, t \in T. \end{aligned}$$

Here, we replaced $f_{t,l}$ by $f_{t,l}^{\text{redi}}$, $d_{t,n}$ by $d_{t,n}^{\text{redi}}$, and $\theta_{t,n}$ by $\theta_{t,n}^{\text{redi}}$ in the Kirchhoff constraints (1) and (3), in the voltage angle reference constraint (4), in the transmission flow limit constraints (5), and $y_{t,g}$ by $y_{t,g}^{\text{redi}}$ in the generation capacity limit constraints (6). Table 1 gives an overview of the variables of all three stages.

Finally, observe that the trilevel problem yields a different solution than the integrated planner problem. This is mainly driven by the fact that firms at the second level choose generation capacities, which are not optimal from an overall welfare maximizing perspective, since their choice ignores network congestion. Welfare obtained for the solution of the integrated planner problem is thus larger than welfare obtained for the solution of the trilevel market problem.

Table 2: Reformulated trilevel problem: sub- and master problem

	Subproblem (generation investment & spot market)		Master problem (line expansion & redispatch)
max	firms' profits/social welfare	max	social welfare
s.t.	a) generation investment	s.t.	a) line investment
	b) production constraints		b) production constraints
	c) Kirchhoff's 1st law (inter-zonal)		c) power flow constraints
	d) flow restrictions (inter-zonal)		d) revenue/network fee
	e) revenue/network fee		

5. Reformulation of the Trilevel Model

The trilevel problem developed in the last section is a special instance of general multilevel optimization problems, which is a very hard class of optimization problems; cf. [42], who showed that even solving linear bilevel problems is NP-hard. Most algorithmic approaches for bilevel problems make use of first-order optimality conditions and solve the resulting mathematical program with equilibrium constraints (MPEC); see [43], [21], or [20] for applied studies. However, such problem reformulations explicitly rely on nonconvex optimization problems and additionally have the drawback that standard constraint qualifications like the Mangasarian–Fromovitz constraint qualification are violated at every feasible point. Therefore, problem-tailored constraint qualifications and stationarity concepts have been developed for MPECs; see, e.g., [44]. For more information on the topic of multilevel programming and MPECs we refer to [45], [46], and [47].

In this section we present a new reformulation approach that allows us to find global optimal solutions in our specific case without using first-order optimality conditions. Before we present our formal reformulation approach, we point out that the trilevel problem always has a finite global optimal solution, since the feasible region is bounded.

The main reason that makes a reduction of stages of the trilevel model possible is that the objective functions all point "into the same direction". This aspect and the way how we exploit it is discussed in detail in the following. Our reformulation approach builds on a detailed two-step analysis of the connection between the three problem levels. In a first step, we fix all non-second-level variables in the second-level problem. This allows us to iteratively solve independent single-level subproblems, i.e., second-level problems with respect to line investment, and corresponding bilevel master problems consisting of the original first-level and third-level models. In a second step, we show that instead of solving a bilevel master problem, we can solve a single-level master problem. Combining the results of both steps, we iteratively solve single-level sub- and master problems arriving at a global optimal solution to the original trilevel problem. Table 2 depicts the reformulation of the trilevel problem into single-level sub- and master problems.

5.1. From the Trilevel Problem to a Single-Level Sub- and Bilevel Master Problem

First note that the second-level problem (generation investment and spot market behaviour) includes only first- and second-level variables. The reason is that for cost-based redispatch firms never receive additional rents from the third stage. Consequently, although the second-level problem is connected to the third level (redispatch) via the line investment problem at the first level, there is no direct interconnection. To be more precise, the interconnection of the stages is purely driven by inter-zonal line investment variables and the transmission fee variable from the first-level problem. Thus, we can fix these variables and solve the second-level market subproblem for every possible realization of these variables. We then fix the respective second-level variables in the first-level and the third-level problem to the values of this solution. This reduces the problem to a bilevel master problem with a reduced number of variables. Note that the approach to substitute the optimal second-level values requires uniqueness of the second-level solution. This will be shown in the upcoming publication [48] under the assumptions of pairwise distinct variable cost and firms that interact on a fixed network structure.

5.2. From a Bilevel to a Single-Level Master Problem

We now show how the bilevel master problem consisting of line investment and redispatch can be solved efficiently. The key ingredient is given in the following proposition, which states that the third-level objective ψ_3 is an affine transformation of the objective ψ_1 of the first level.

Proposition 1. *Let ψ_1 and ψ_3 be the objective functions of the original first and third level. Then, $\psi_3 = -\psi_1 + b$ holds, where*

$$\begin{aligned} b = &\sum_{n \in N^{\mathrm{dem}}} \sum_{t \in T} \int_0^{d_{t,n}^{\mathrm{spot}}} p_{t,n}(\xi)\, \mathrm{d}\xi - \sum_{l \in L^{\mathrm{new}}} c_l^{\mathrm{inv}} z_l - \sum_{l \in L^{\mathrm{ex}}} c_l^{\mathrm{del}} z_l \\ &- \sum_{n \in N} \sum_{g \in G_n^{\mathrm{new}}} c_g^{\mathrm{inv}} \bar{y}_g^{\mathrm{new}} - \sum_{n \in N} \sum_{g \in G_n^{\mathrm{all}}} \sum_{t \in T} c_g^{\mathrm{var}} y_{t,g}^{\mathrm{spot}} \end{aligned}$$

only depends on spot market and line investment variables.

This insight reveals that the first- and third-level problem have affine-equivalent objective functions and, thus, identical optimization directions. We now exploit the fact that in order to solve a general bilevel problem with affine-equivalent objective functions, it is possible to solve an easier single-level problem, which is equivalent.

Proposition 2. *Consider the bilevel problem*

$$\min_{x_1, x_2} \quad \psi(x_1, x_2) \quad \text{s.t.} \quad g(x_1, x_2) \geq 0, \quad x_2 = \arg\min_z \{\psi(x_1, z) \colon h(x_1, z) \geq 0\}$$

with equivalent objective functions in the upper and lower level and denote the set of optimal solutions by $\mathcal{S}^{\mathrm{bl}}$. Moreover, let

$$\min_{x_1, x_2} \quad \psi(x_1, x_2) \quad \text{s.t.} \quad g(x_1, x_2) \geq 0, \quad h(x_1, x_2) \geq 0$$

be the corresponding single-level problem with solution set $\mathcal{S}^{\mathrm{sl}}$. Then, $\mathcal{S}^{\mathrm{sl}} = \mathcal{S}^{\mathrm{bl}}$ holds.

Proof. Since a solution of the bilevel problem is feasible for the single-level problem and vice versa, the same objective function in the upper and lower level directly implies the result. □

Note that the above result can easily be generalized to p-level programming problems for $p > 2$. For the sake of completeness let us also state the following proposition, which is an immediate consequence.

Proposition 3. *Assume a p-level minimization problem is given with ψ_i denoting the objective function of the i-th problem stage. If there exist affine linear transformations $a_i\psi_i + b_i = \psi_1$ and $a_i > 0$ for all $i \in \{2, \dots, p\}$, then the multilevel model can be solved as a single-level model.*

6. Solution Strategy

In this section we describe how we solve the reformulated problem to global optimality. Since we started from a mixed-integer nonlinear trilevel problem it is clear that this is computationally a very hard task. However, the reformulation discussed in Sect. 5, together with the results given in this section, allow us to set up a binary search strategy to solve the coupled master and subproblems; see Sect. 6.1. In Sect. 6.2, we describe how binary search upper bounds for the network fee are determined and, finally, in Sect. 6.3, we describe a technique for reducing the complexity of the proposed algorithm.

Algorithm 1: Binary search for the reformulated trilevel model

Input : Parameters for the trilevel power market model, upper bound $\bar{\varphi}$ of the network fee
Output : Optimal solution x^* and optimal objective value ψ^* of the trilevel problem

```
1  Set ψ* = 0 and x* = 0.
2  for all network configurations ∅ ⊆ X ⊆ L^ex ∪ L^new do
3      Set a ← 0 and b ← φ̄.
4      while a ≤ b do
5          Set φ ← (a + b)/2 and solve the subproblem for fixed network configuration X and fixed network
           fee φ. Let y*_{X,φ} be the solution.
6          Solve the master problem for fixed network configuration X, fixed network fee φ, and fixed values
           y*_{X,φ} of the subproblem. Let x*_{X,φ} be the solution.
7          if |B(x*_{X,φ})| = 0 then go to line 9
8          if B(x*_{X,φ}) < 0 then set a ← m else set b ← m.
9      if ψ(x*_{X,φ}) > ψ* then set x* ← x*_{X,φ} and ψ* ← ψ(x*).
10 return x* and ψ*.
```

6.1. Binary Search Algorithm

We now describe how the trilevel problem can be solved in an efficient way using a problem-tailored binary search strategy. This approach is based on the trilevel program reformulated as a single-level sub- and master problem as described in Sect. 5. The key insight is that the master and subproblem are only coupled by transmission line investment and network fees. Thus, fixing these values yields decoupled models that can be solved separately. For this purpose, we iterate over all possible network configurations $\mathcal{X}$ with $\emptyset \subseteq \mathcal{X} \subseteq L^{\text{ex}} \cup L^{\text{new}}$ and we additionally neglect the budget constraint $B = R - E = 0$, for which we compute the unique feasible solution via binary search. The resulting algorithm is depicted in Alg. 1.

In order to proof correctness of the algorithm, we have to show that the budget function $B = B(\varphi)$ for fixed line investment is strictly increasing in the network fee φ and that $\operatorname{sign}(B(0)) \neq \operatorname{sign}(B(\bar{\varphi}))$. Note that correctness of the binary search algorithm only requires monotony of the budget function but strict monotonicity is required in order to obtain global optimal solutions. Here and in what follows, $\bar{\varphi}$ denotes the upper bound on the network fee that is used in the binary search. Our discussion in Sect. 4 shows that cost from redispatch are always nonnegative. Moreover, they can only be zero in the case of no line investment. Since $R(0) = 0$ for every fee type, this shows $B(0) \leq 0$. Thus, it is sufficient to prove strict monotonicity of B since then $B(\bar{\varphi}) > 0$ follows for sufficiently large $\bar{\varphi}$. Obviously, there is nothing to show for the case of a lump sum fee. In the next section, we prove a criterion for strict monotonicity for the generation capacity based fee $\varphi = \varphi^{\text{gcb}}$ and focus on the case of the energy-based fee $\varphi = \varphi^{\text{eb}}$ in Sect. 6.1.2.

6.1.1. Monotonicity of the TO's Budget Function for the Generation Capacity Based Fee

In the following, quantities without subindex for network elements or time periods denote the corresponding vector, e.g., $\bar{y}^{\text{new}} = (\bar{y}_n^{\text{new}})_{n \in N}$ is the vector of all newly installed generation capacities. Using this notation, the budget function $B : \mathbb{R} \to \mathbb{R}$ is defined as

$$B(\varphi) := R(\varphi, \bar{y}^{\text{new}}(\varphi)) - E(\bar{y}^{\text{new}}(\varphi)).$$

For the ease of exposition, we consider the slightly simplified spot market model without market splitting

$$\max_{d,y,\bar{y}^{\text{new}}} \quad \psi_{\text{spot}} := \sum_{n\in N^{\text{dem}}} \sum_{t\in T} \int_0^{d_{t,n}} p_{t,n}(\xi)\,\mathrm{d}\xi - \sum_{n\in N} \left(c_n^{\text{inv}} \bar{y}_n^{\text{new}} + \varphi \bar{y}_n^{\text{new}} + \sum_{t\in T} c_n^{\text{var}} y_{t,n} \right) \tag{15a}$$

$$\text{s.t.} \quad 0 \le y_{t,n} \le \bar{y}_n^{\text{new}} \quad \text{for all } n \in N, t \in T, \tag{15b}$$

$$\sum_{n\in N^{\text{dem}}} d_{t,n} - \sum_{n\in N} y_{t,n} = 0 \quad \text{for all } t \in T, \tag{15c}$$

in which we (w.l.o.g.) assume that it can only be invested in one technology per node, that all equivalent availabilities are 1 and that the time steps τ are 1 h. The corresponding redispatch model is given by

$$\max_{d,y} \quad \psi_{\text{redi}} := \sum_{t\in T} \left(\sum_{n\in N^{\text{dem}}} \int_0^{d_{t,n}} p_{t,n}(\xi)\,\mathrm{d}\xi - \sum_{n\in N} c_n^{\text{var}} y_{t,n} \right) \tag{16a}$$

$$\text{s.t.} \quad 0 \le y_{t,n} \le \bar{y}_n^{\text{new}} \quad \text{for all } n \in N, t \in T, \tag{16b}$$

$$\sum_{n\in N^{\text{dem}}} d_{t,n} - \sum_{n\in N} y_{t,n} = 0 \quad \text{for all } t \in T, \tag{16c}$$

$$-\bar{f} \le A(d-y) \le \bar{f}, \tag{16d}$$

where A is the well-known PDTF matrix that relates demand and production at nodes of an electricity network to power flows on lines. Note that we have to extend the demand vector d by additional 0's for all nodes $n \in N \setminus N^{\text{dem}}$. With these two models and the definition (10) of the TO's expenses, the budget function can be rewritten as

$$B(\varphi) = \psi^*_{\text{redi}}(\bar{y}^{\text{new}}(\varphi)) - \psi^*_{\text{spot}}(\varphi) - \sum_{n\in N} c_n^{\text{inv}} \bar{y}_n^{\text{new}}(\varphi),$$

where only those dependencies on parameters are given that do not vanish in the first derivative w.r.t. φ. Strict monotonicity in φ then is equivalent to

$$0 < \frac{\mathrm{d}B}{\mathrm{d}\varphi} = (\nabla_{\bar{y}^{\text{new}}} \psi^*_{\text{redi}})^T \nabla_\varphi \bar{y}^{\text{new}} - \sum_{n\in N} c_n^{\text{inv}} \frac{\partial \bar{y}_n^{\text{new}}}{\partial \varphi} + \sum_{n\in N} \bar{y}_n^{\text{new}}. \tag{17}$$

Thus, we have to compute the change in generation investment with respect to the network fee $\nabla_\varphi \bar{y}^{\text{new}}$ and the change in the optimal redispatch value with respect to the upper bounds on generation $\nabla_{\bar{y}^{\text{new}}} \psi^*_{\text{redi}}$. We start with the former and make use of the first-order optimality conditions of the spot market model (15), which are also sufficient in our case. These conditions comprise dual feasibility

$$p_{t,n}(d_{t,n}) + \lambda_t = 0 \quad \text{for all } n \in N^{\text{dem}}, t \in T, \tag{18a}$$

$$-c_n^{\text{var}} + \alpha_{t,n} - \beta_{t,n} - \lambda_t = 0 \quad \text{for all } n \in N, t \in T, \tag{18b}$$

$$-c_n^{\text{inv}} - \varphi + \sum_{t\in T} \beta_{t,n} = 0 \quad \text{for all } n \in N, \tag{18c}$$

primal feasibility (15b) and (15c) as well as complementarity and nonnegativity of dual variables of inequality constraints (15b). That is, the lower bounds $0 \le y_{t,n}$ are equipped with dual variables $\alpha_{t,n}$, upper bounds $y_{t,n} \le \bar{y}_n$ with $\beta_{t,n}$ and the market clearing condition (15c) with λ_t. Note that a price equilibrium directly follows from (18a), i.e.,

$$p_{t,n_i}(d_{t,n_i}) = p_{t,n_j}(d_{t,n_j}) \quad \text{for all } n_i, n_j \in N^{\text{dem}} \tag{19}$$

holds for all $t \in T$. Thus, it is reasonable to introduce the notation p_t (independent of a demand node n) for the spot market price in time period t. Combining (18a) and (18b) then yields

$$\beta_{t,n} = p_t - c_n^{\text{var}} + \alpha_{t,n} \quad \text{for all } n \in N, t \in T, \tag{20}$$

where we can replace p_t by $p_{t,n} = p_{t,n}(d_{t,n})$ for any demand node $n \in N^{\text{dem}}$. From now on, we assume that all variable costs $c_n^{\text{var}}, n \in N = \{n_1, \dots, n_{|N|}\}$, are positive and pairwise distinct and that the nodes are ordered such that

$$0 < c_{n_1}^{\text{var}} < c_{n_2}^{\text{var}} < \cdots < c_{n_{|N|}}^{\text{var}}$$

holds. This ordering allows us to formulate the following lemma, which follows directly from the optimality of spot market solutions:

Lemma 1. *For every time step $t \in T$ there exists a node index i such that*

$$\bar{y}_{n_j}^{\text{new}} = y_{t,n_j} \quad \text{for all } j < i, \qquad 0 = y_{t,n_j} \quad \text{for all } j > i, \qquad 0 \le y_{t,n_i} < \bar{y}_{n_i}^{\text{new}},$$

holds for every optimal solution d, y and $\bar{y}^{\text{new}}$ of the spot market model.

In what follows, we use the index sets $\mathcal{T}_n := \{t \in T \colon y_{t,n} = \bar{y}_n^{\text{new}}\}$. In order to compute $\nabla_\varphi \bar{y}^{\text{new}}$ we now consider the KKT condition (18c). By subtracting the equation for node n_{i+1} from the one for n_i and using (20), we obtain

$$\gamma_{i+1,i} + \sum_{t \in \mathcal{T}_{n_i} \setminus \mathcal{T}_{n_{i+1}}} p_t = 0, \quad \gamma_{i+1,i} = c_{n_{i+1}}^{\text{inv}} - c_{n_i}^{\text{inv}} + |\mathcal{T}_{n_{i+1}}| c_{n_{i+1}}^{\text{var}} - |\mathcal{T}_{n_i}| c_{n_i}^{\text{var}}.$$

Here, we used the price equilibrium (19), condition (20), Lemma 1, and the relation $\mathcal{T}_{n_{i+1}} \subseteq \mathcal{T}_{n_i}$. We now apply Gaussian elimination steps in which we replace the equation of node n_i by the difference of equation for node n_i and of node n_{i+1} for $i = 1, \dots, |N| - 1$. The equation of the last node stays untouched. This gives

$$\gamma_{i+1,i} + \sum_{t \in \mathcal{T}_{n_i} \setminus \mathcal{T}_{n_{i+1}}} p_t = 0, \quad \text{for all } i = 1, \dots, |N| - 1, \tag{21a}$$

$$-c_{n_{|N|}}^{\text{inv}} - \varphi + \sum_{t \in \mathcal{T}_{n_{|N|}}} (p_t - c_{n_{|N|}}^{\text{var}}) = 0. \tag{21b}$$

With the additional assumption (which is in line with our case studies) that the price functions are linear, i.e. $p_{t,n}(d_{t,n}) = a_{t,n} - s_n d_{t,n}$, where $a_{t,n} > 0$ is the intercept with the price axis and $-s_n < 0$ is the slope. Note that fluctuation is only modeled by shifted intercepts whereas the slope of the price functions stays the same. Furthermore, we need the well-known notion of so-called inverse market demand functions $P_t(D_t)$, which represent the price functions for the aggregated market demand at time step t. They are given as $P_t(D_t) = A_t + S D_t$, with

$$D_t = \sum_{n \in N^{\text{dem}}} d_{t,n}, \quad A_t = \frac{\sum_{n \in N^{\text{dem}}} a_{t,n}/s_n}{\sum_{n \in N^{\text{dem}}} 1/s_n}, \quad S = -\frac{1}{\sum_{n \in N^{\text{dem}}} 1/s_n}. \tag{22}$$

The inverse demand functions can be derived as follows: First solve the single demand functions for $d_{t,n}$, yielding $d_{t,n} = (a_{t,n} - p_{t,n})/s_n$. Since all prices are equal for all nodes, we have

$$D_t = \sum_{n \in N^{\text{dem}}} d_{t,n} = \sum_{n \in N^{\text{dem}}} (a_{t,n} - p_{t,n})/s_n = \sum_{n \in N^{\text{dem}}} a_{t,n}/s_n - P_t \sum_{n \in N^{\text{dem}}} 1/s_n,$$

which gives (22). This notation now allows to rewrite (21) as

$$\gamma_{i+1,i} + \sum_{t \in \mathcal{T}_{n_i} \setminus \mathcal{T}_{n_{i+1}}} P_t \left(\sum_{k=1}^{i} \bar{y}_{n_k}^{\text{new}} \right) = 0, \quad \text{for all } i = 1, \dots, |N| - 1, \tag{23a}$$

$$-c_{n_{|N|}}^{\text{inv}} - \varphi + \sum_{t \in \mathcal{T}_{n_{|N|}}} \left(P_t \left(\sum_{n \in N} \bar{y}_n^{\text{new}} \right) - c_{n_{|N|}}^{\text{var}} \right) = 0. \tag{23b}$$

Since the equations (23a) uniquely define $\bar{y}_{n_i}^{\text{new}}$ for $i = 1, \dots, |N| - 1$ and do not depend on the fee φ, we have

$$\frac{\partial \bar{y}_{n_i}^{\text{new}}}{\partial \varphi} = 0 \quad \text{for all } i = 1, \dots, |N| - 1. \tag{24}$$

In addition, the remaining partial derivative is given by the following lemma.

Lemma 2. *It holds*

$$\frac{\partial \bar{y}^{\text{new}}_{n_{|N|}}}{\partial \varphi} = (|\mathcal{T}_{n_{|N|}}|S)^{-1} =: \Phi < 0.$$

Proof. Equation (23b) is an equation of the form $F(\bar{y}^{\text{new}}_{n_1}(\varphi), \dots, \bar{y}^{\text{new}}_{n_{|N|}}(\varphi); \varphi) = 0$. Thus, (24) and the implicit function theorem imply

$$\frac{\partial F}{\partial \bar{y}^{\text{new}}_{n_{|N|}}} \frac{\partial \bar{y}^{\text{new}}_{n_{|N|}}}{\partial \varphi} = -\frac{\partial F}{\partial \varphi}.$$

The claim follows since $\partial_\varphi F = -1$ and

$$\frac{\partial F}{\partial \bar{y}^{\text{new}}_{n_{|N|}}} = \sum_{t \in \mathcal{T}_{n_{|N|}}} S. \qquad \square$$

It can be easily shown that $|\mathcal{T}_n| > 0$ for all $n \in N$, justifying the definition of Φ. Using the previous lemma, we finally obtain the following monotonicity criterion:

Theorem 1. *Let the spot market and redispatch model with linear demand functions and network fee φ be given as in* (15) *and* (16)*. Furthermore, let ψ^*_{redi} be the optimal value of the redispatch model, $\bar{y}^{\text{new}}_n, n \in N$, be the spot market optimal electricity generation and Φ as defined in Lemma 2. Then, the budget function is strictly increasing in* $[0, \varphi]$ *if*

$$\frac{\mathrm{d}B}{\mathrm{d}\varphi} > 0 \iff \Theta^{\text{gcb}} := \frac{\partial \psi^*_{\text{redi}}}{\partial \bar{y}^{\text{new}}_{n_{|N|}}} + \Phi^{-1} \sum_{n \in N} \bar{y}^{\text{new}}_n - c^{\text{inv}}_{n_{|N|}} < 0 \tag{25}$$

holds.

Proof. The definition of Θ^{gcb} follows directly from the preceding computations. Additionally, Θ^{gcb} is strictly decreasing in $\bar{y}^{\text{new}}_n, n \in N$: The second derivative of the optimal value function of the redispatch model (first term) vanishes due to the envelope theorem and because the capacities appear only linearly in the Lagrangian. The second term also vanishes after differentiating, and the derivative of the third term is Φ^{-1}, which is also negative. Moreover, Θ^{gcb} is increasing in the fee φ because $\bar{y}^{\text{new}}_n, n \in N$, are decreasing in φ. $\square$

As a consequence, the correctness of the binary search thus follows if (25) holds for the generation capacity corresponding to the upper bound $\bar{\varphi}$.

In order to check (25) in practice, we finally note that the first term can be expressed by the dual variables of the redispatch model. From standard sensitivity analysis of convex optimization (cf., e.g., [49]), it follows that

$$\frac{\partial \psi^*_{\text{redi}}}{\partial \bar{y}^{\text{new}}_{n_{|N|}}} = \sum_{t \in T} \beta_{t, n_{|N|}}$$

holds, where $\beta_{t,n_{|N|}}, t \in T$, are the optimal dual variables of the upper bounds in (16b).

6.1.2. Monotonicity of the TO's Budget Function for the Energy Based Fee

In order to show the correctness of the binary search algorithm for the case of the energy based fee $\varphi = \varphi^{\text{eb}}$, we consider the slightly modified spot market model

$$\begin{aligned}
\max_{d, y, \bar{y}^{\text{new}}} \quad & \psi_{\text{spot}} := \sum_{n \in N^{\text{dem}}} \sum_{t \in T} \int_0^{d_{t,n}} p_{t,n}(\xi)\,\mathrm{d}\xi - \sum_{n \in N} \left(c^{\text{inv}}_n \bar{y}^{\text{new}}_n + \sum_{t \in T} c^{\text{var}}_n y_{t,n} \right) - \sum_{n \in N^{\text{dem}}} \sum_{t \in T} \varphi d_{t,n} \\
\text{s.t.} \quad & 0 \le y_{t,n} \le \bar{y}^{\text{new}}_n \qquad \text{for all } n \in N, t \in T, \\
& \sum_{n \in N^{\text{dem}}} d_{t,n} - \sum_{n \in N} y_{t,n} = 0 \quad \text{for all } t \in T.
\end{aligned}$$

The redispatch model (16) stays the same. With these models at hand, we can rewrite the budget function as

$$B(\varphi) = \psi^*_{\text{redi}}(\bar{y}^{\text{new}}(\varphi)) - \psi^*_{\text{spot}}(\varphi) - \sum_{n\in N} c_n^{\text{inv}} \bar{y}_n^{\text{new}}(\varphi).$$

Thus, the budget function is strictly increasing in φ if

$$\frac{\mathrm{d}B}{\mathrm{d}\varphi} = (\nabla_{\bar{y}^{\text{new}}} \psi^*_{\text{redi}})^T \nabla_\varphi \bar{y}^{\text{new}} + \sum_{n\in N^{\text{dem}}} \sum_{t\in T} d_{t,n} - \sum_{n\in N} c_n^{\text{inv}} \frac{\partial \bar{y}_n^{\text{new}}}{\partial \varphi} > 0 \tag{27}$$

holds. As before,

$$\frac{\partial \psi^*_{\text{redi}}}{\partial \bar{y}_n^{\text{new}}} = \sum_{t\in T} \beta_{t,n}$$

is known from standard sensitivity analysis. Thus, we have to compute $\nabla_\varphi \bar{y}^{\text{new}}$. The main strategy is the same as for the generation capacity based fee. Dual feasibility conditions of the spot market model read

$$p_{t,n}(d_{t,n}) - \varphi + \lambda_t = 0 \quad \text{for all } n \in N^{\text{dem}}, t \in T, \tag{28a}$$

$$-c_n^{\text{var}} + \alpha_{t,n} - \beta_{t,n} - \lambda_t = 0 \quad \text{for all } n \in N, t \in T, \tag{28b}$$

$$-c_n^{\text{inv}} + \sum_{t\in T} \beta_{t,n} = 0 \quad \text{for all } n \in N. \tag{28c}$$

As before, the first condition implies the equilibrium of spot market prices (19) for each time step $t \in T$. Thus, λ_t can be eliminated in the first two conditions, yielding

$$\beta_{t,n} = p_t - c_n^{\text{var}} - \varphi + \alpha_{t,n} \quad \text{for all } n \in N, t \in T. \tag{29}$$

Finally, we can eliminate the β's by combining (29) and (28c):

$$\sum_{t\in \mathcal{T}_n} (p_t - c_n^{\text{var}} - \varphi) - c_n^{\text{inv}} = 0 \quad \text{for all } n \in N.$$

Using inverse market functions and the implicit function theorem (as in Sect. 6.1.1) to this system of equations gives

$$\frac{\partial}{\partial \varphi} \left(\sum_{k=1}^{i} \bar{y}_{n_k}^{\text{new}} \right) = \frac{1}{S} \quad \text{for all } i = 1, \dots, |N|,$$

which readily implies

$$\frac{\partial \bar{y}_{n_1}^{\text{new}}}{\partial \varphi} = \frac{1}{S}, \quad \frac{\partial \bar{y}_{n_i}^{\text{new}}}{\partial \varphi} = 0 \quad \text{for all } i = 2, \dots, |N|.$$

This gives us the following theorem:

Theorem 2. *Let the spot market and redispatch model with linear demand functions and network fee φ be given as in* (26) *and* (16). *Furthermore, let ψ^*_{redi} be the optimal value of the redispatch model and $\bar{y}_n^{\text{new}}, n \in N$, be the spot market optimal electricity generation. Then, the budget function is strictly increasing in* $[0, \varphi]$ *if*

$$\frac{\mathrm{d}B}{\mathrm{d}\varphi} > 0 \iff \Theta^{\text{eb}} := \frac{\partial \psi^*_{\text{redi}}}{\partial \bar{y}_{n_1}^{\text{new}}} + S \sum_{n\in N^{\text{dem}}} \sum_{t\in T} d_{t,n} - c_{n_1}^{\text{inv}} < 0. \tag{30}$$

holds.

Proof. The definition of Θ^{eb} again follows from the preceding computations. By the same arguments as before, it can be shown that the demands $d_{t,n}$ are decreasing in φ. Since $S < 0$, this implies that the second term is increasing in φ. Thus, Θ^{eb} is increasing in φ. □

In other words, the last theorem shows the correctness of the binary search algorithm for energy based fees and values of $\bar{\varphi}$ that fulfill criterion (30).

6.2. Computing Upper Network Fee Bounds

Given the importance of appropriate transmission fee bounds, we take the following strategy. In all transmission fee regimes we use zero as a lower bound excluding negative transmission fees. Then, we compute monopoly solutions and derive the respective monopoly markups. Based on the implicit assumption that the TO does not charge fees that are larger than monopoly markups, we use these markups as upper bounds. In the case of the energy based fee, we use the average monopoly markup across all time periods and market zones as an upper bound. In the case of the generation capacity based fee, we use the cumulative monopoly markup across all time periods, where total generation capacity is binding in a given zone. To obtain the upper bound we then take averages over all market zones. As it can be seen easily, the computed upper bound for the generation capacity based fee is in general much larger than for the energy based fee. In our computational study (cf. Sect. 7), the obtained upper bounds fulfill the criteria developed in the last section and thus lead to a well-defined search strategy.

6.3. Reduction Strategy for the Set of Network Configurations

An ingenuous way of incorporating all possible network configurations by simply introducing binary variables for every extension and degradation typically yields very large mixed-integer problems that are hard to solve in practice. In order to reduce the number of combinatorial choices, we model network extensions and degradations in the following way: By $L_{jk}^{\text{new}} \subseteq L^{\text{inter}} \cap L^{\text{new}}$ we denote the set of candidate transmission lines and by L_{jk}^{ex} the set of existing lines between zone Z_j and zone Z_k. Moreover, let Γ_{jk} denote the set of transmission capacities between the two zones that can be realized by an investment in or a degradation of a (sub)set of lines between the zones, i.e.,

$$\Gamma_{jk} := \left\{ \bar{f}_{jk} = \sum_{l \in \mathcal{X}} \bar{f}_l : \emptyset \subseteq \mathcal{X} \subseteq (L_{jk}^{\text{new}} \cup L_{jk}^{\text{ex}}) \right\}.$$

Since intra-zone transmission capacities do not affect spot market trading, intra-zone network modifications do not have to be considered in the subproblem. This allows us to iterate only over the set of inter-zone network modifications in the outer loop of Alg. 1. Moreover, we do not iterate over the specific network configurations but over the set of relevant capacities, i.e., over all elements of

$$\prod_{Z_j, Z_k \in Z, j \neq k} \Gamma_{jk}.$$

By doing so, the constraint for $\bar{f}_{jk} \in \Gamma_{jk}$ of a zonal version of the market splitting flow restrictions (12) now reads

$$-\bar{f}_{jk} \leq \sum_{l \in L_{jk}^{\text{inter}}} f_{t,l} \leq \bar{f}_{jk} \quad \text{for all } t \in T.$$

On the other hand, the master problem determines both a welfare maximizing intra-zone network extension and a welfare maximizing inter-zone network extension that yields the predefined inter-zone transmission capacity of the given network extension and degradation plan. Thus, the master problem is additionally restricted by the following constraint:

$$\sum_{l \in L_{jk}^{\text{ex}}} (1 - z_l) \bar{f}_l + \sum_{l \in L_{jk}^{\text{new}}} z_l \bar{f}_l = \bar{f}_{jk} \quad \text{for all } Z_j, Z_k \text{ with } j \neq k.$$

Note that this reduction technique is only possible given that physical line characteristics other than thermal capacities do not play a role on the spot market.

Taken all together, the fixation of a network extension yields a concave-quadratic maximization over a polyhedral set for the subproblem and additionally relaxing the budget constraint yields a concave-quadratic mixed-integer maximization with respect to linear constraints for the master problem.

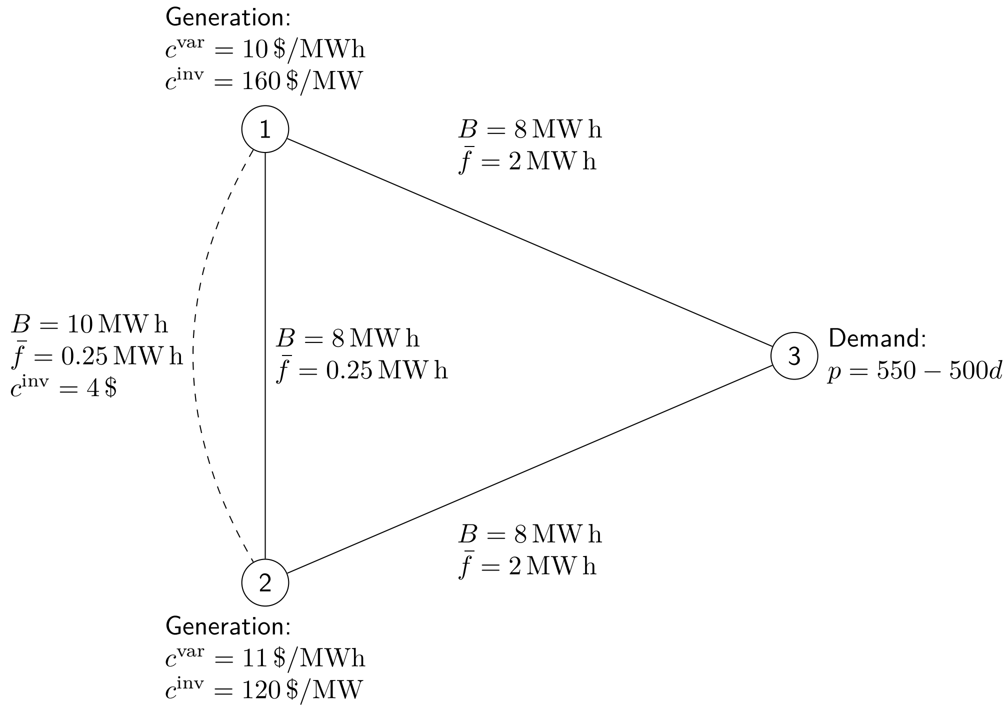

Figure 3: Three-node test network

7. Case Studies

In this section we discuss computational results for two prominent test networks from the literature. The first example is taken from [21] and consists of three nodes. The second test network is a six-node example by [50] that has been widely used in the energy market literature. For both test cases (consisting of the network structure as well as demand and cost parameters) we compute

1. the welfare optimum of the integrated planner model;
2. the market outcomes in the case of a redispatch model without market splitting;
3. the market outcomes in the case of a redispatch model with market splitting.

For the redispatch models we consider all three different network fee regimes, i.e., the lump sum, the energy based, and the generation capacity based fee. Note that we transformed all parameters into equivalent annual hourly values.

We implemented the redispatch model as well as the integrated planner model in Zimpl (see [51]) and used SCIP (see [52]) to generate corresponding mps files. Finally, we solved the problems using CPLEX 12.6; see [53]. All experiments were performed on a 12 core computer equipped with two AMD Opteron(tm) 2435 Processors and 64 GB DDR2-RAM. Using the binary search algorithm described in the last section, all models are solved in a few minutes. This indicates that our method also seems quite promising for larger test instances.

7.1. The Three-Node Test Network

Before we present our main results, we briefly review the input data that is directly taken from [21]. The graph consists of three nodes, three existing lines and one candidate line as depicted in Fig. 3, where all physical and economical data of the network are given. We allow for both investment in the transmission network and in generation capacity. We consider four time periods and assume that at the beginning of the planning horizon there are no existing generators but explicitly model generation investment. Node 1 and node 2 are generation

Table 3: Computational results for the three-node example

	integrated planner	1 price zone φ^{ls}	φ^{eb}	φ^{gcb}	2 price zones φ^{ls}	φ^{eb}	φ^{gcb}
normalized welfare (%)	100	97.0	96.4	96.0	97.0	96.4	96.0
candidate line	yes	yes	yes	yes	yes	yes	yes
investment node 1 (MW)	0.11	0	0	0	0	0	0
investment node 2 (MW)	1.49	1.66	1.63	1.58	1.66	1.63	1.58

nodes. Technology 1 can be built at node 1 and technology 2 can be built at node 2. The equivalent availability of capacity is 80 %.

Consumers are located at node 3. We assume that the demand function is $p = 550 - 500d$. To give the four periods some simple interpretation, we refer to them as the four seasons spring, summer, autumn, and winter. For this reason, we multiply the intercept of the above demand function by 1, 0.5, 1, and 1.5 in time period 1 to 4, respectively.

7.1.1. Discussion

The integrated planner approach yields a social welfare of 1083 \$. This is also the welfare level under a nodal pricing model. A redispatch model with no market splitting, i.e., one price zone, implies a welfare loss as compared to the first best solution in all three network fee regimes. In the case of a generation capacity based fee, an energy based fee, and a lump sum fee the realized welfare levels are 1050 \$, 1044 \$, and 1040 \$, respectively. This welfare loss is mainly driven by a distortion of generation investment. As Table 3 illustrates, the solution to the integrated planner problem implies positive investment in both technologies, where investment in technology 2—which has lower investment cost—is much larger. On the contrary, the trilevel power market model yields generation investment only in technology 2, as technology 1 would earn too little on the spot market to recover investment cost. In addition, in our trilevel power market model generation investment in technology 2 is higher than in the welfare optimum implying an overinvestment at node 2 for all network fee regimes—lowest for the generation capacity based fee, highest for the lump sum fee. Even though generation investment is inefficient in the solution to our market model as compared to the first best solution, line investment is efficient in all three scenarios: The candidate line connecting the nodes 1 and 2 is build in the trilevel power market model as well as in the welfare optimum. Interestingly, introducing a second price zone does not change the presented results. This can be explained by the fact that no matter how the two price zones are designed, there is always at least one non-congested line that connects the two zones.

These results suggest that generation investment can be inefficient in a market environment as spot markets can induce wrong price signals. In addition, market splitting is not always a useful measure as it can leave the welfare level unaffected. Apart from that, the trilevel power market model allows to measure the efficiency of the energy system in general, e.g., to answer the question of how large the difference is between the market model and the first best solution.

7.2. The Six-Node Test Network

We have seen that there is no difference between different network fee regimes and price zone configurations for the small three-node network. This also shows that it might be necessary to consider more complicated networks for discussing the effects of different market configurations and that it is not sufficient to discuss different designs only on minimal test networks. In order to illustrate the effects of market splitting with our model, we thus consider an adapted 52-period test example based on [50]. The network of this example has extensively been analyzed in the literature to illustrate different energy market models. We extend the original test example by several aspects that are important for our approach such as fluctuating demand, generation investment, network investment, and a cost-based redispatch mechanism.

As Fig. 4 illustrates, the network consists of six nodes connected by eight existing transmission lines (solid). Northern nodes (nodes 1 to 3) and southern nodes (nodes 4 to 6) are interconnected by lines with unlimited

Table 4: Computational results for the six-node example

	integrated planner	1 price zone			2 price zones		
		φ^{ls}	φ^{eb}	φ^{gcb}	φ^{ls}	φ^{eb}	φ^{gcb}
normalized welfare (%)	100	62.1	81.8	81.8	94.5	94.2	94.0
candidate line 1	no	yes	yes	yes	yes	yes	yes
candidate line 2	no	no	yes	yes	no	no	no
investment node 1 (MW)	252	2009	1423	1423	1317	1206	1204
investment node 2 (MW)	794	0	0	0	0	0	0
investment node 4 (MW)	620	0	0	0	370	264	251

capacities. The "northern zone" and the "southern zone" are interconnected by only two lines with limited capacities. We consider a situation where two different candidate transmission lines (dashed) can be build to alleviate congestion problems. The technical line data as well as all other data is given in Fig. 4. The equivalent availability of capacity is 80 %.

We assume that there are no existing generators but explicitly model generation investment. Generation units can be build at nodes 1, 2, and 4. Demand is located at nodes 3, 5, and 6. The basic demand functions are $p_3 = 37.5 - 0.05\,d_3$, $p_5 = 75 - 0.1\,d_5$, $p_6 = 80 - 0.1\,d_6$. Assuming these three demand functions taken from [50] we use real-world data to induce demand fluctuation across the 52 periods at all nodes.[4] In particular, we use factors derived from the German 2011 demand realizations for shifting the demand functions; see [54]. Demand levels across nodes in a given period are always shifted by the same factor, which accounts for the fact that the level of demand in a given period is typically correlated across demand nodes.

7.2.1. Discussion

Table 4 shows the welfare levels in the case of one and two price zones as compared to the welfare optimum. Market splitting significantly increases welfare but does not achieve the maximum. While welfare levels under different network regimes are rather similar in the case of two price zones, we find substantial welfare differences for different network regimes in the case of a single price zone, i.e., without market splitting.

Table 4 also illustrates investment decisions in transmission lines and generation capacities for the different scenarios. In the welfare optimum the social planner refrains from transmission line expansion with positive generation capacities being installed at all nodes. Obviously, from a welfare perspective investment in generation capacity at node 4 is preferable to any line investment, although variable production cost of the available technology is substantially higher than variable cost of technologies available at northern nodes. Furthermore, due to the fact that demand satisfied by northern generators is relatively low in the absence of line investment, more generation capacity is built at node 2, where investment cost is lower than at node 1.

Turning to the case of a power market without market splitting, we find that both candidate lines are built, except for candidate line 2 in the case of the lump sum fee, and that generation capacity is exclusively installed at node 1. The intuition is straightforward: Recall that we have no price zones and that transmission constraints are not accounted for at the spot market. In this case the generator with lowest production cost will trade at the spot market while other generators will predominantly be called for at the redispatch stage. The fact that redispatch reimbursement covers only variable cost immediately implies that those generators cannot operate profitably. Consequently, all capacity investment occurs at the node where the production cost is lowest, which is at node 1. Anticipating this decision, the network planner prefers to install both candidate lines—except for the case of a lump sum network fee, where only line 1 is built.

Finally, consider the scenario with two price zones: north and south. Obviously, market splitting allows a generator in the south to earn rents whenever it is impossible to satisfy demand in the south by production from northern generators due to transmission congestion. As a consequence, firms have an incentive to install capacity

[4] Alternatively, one could choose random draws from a distribution that reflect the nature of demand fluctuation in electricity markets.

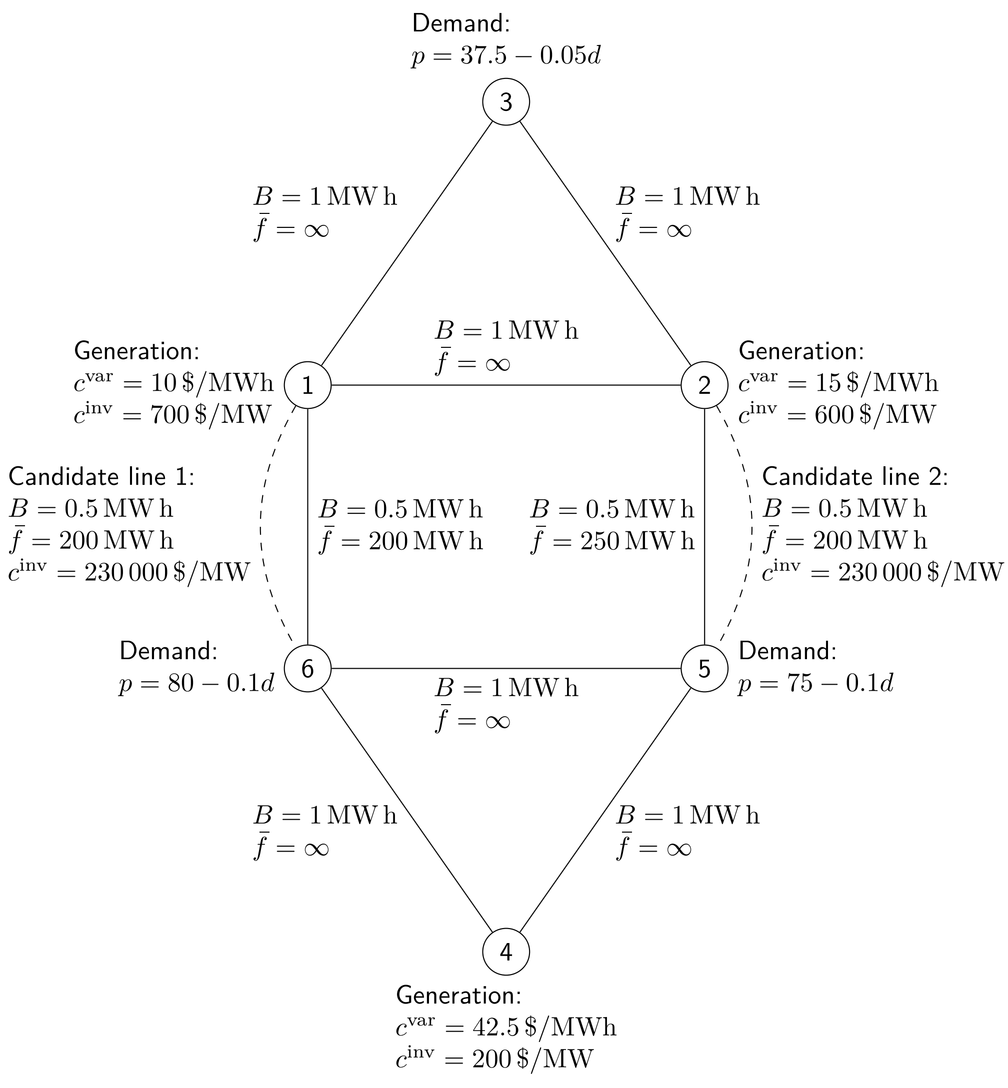

Figure 4: Six-node test network

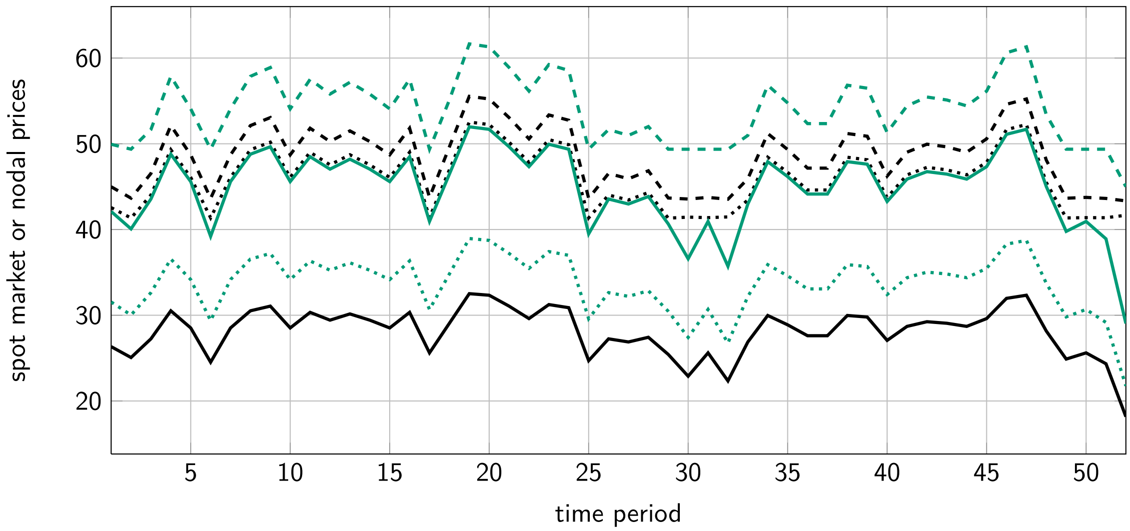

Figure 5: Spot market and nodal prices for the energy based fee. Nodal prices are plotted in black (node 3: solid, node 5: dotted, node 6: dashed) and spot market prices are plotted in green (one zone spot market price: solid, two zone spot market price at northern node: dotted, two zone spot market price at southern node: dashed).

at node 4, which in turn decreases the incentives of the network planner to install additional transmission capacity. Consequently, anticipation of generation investment at node 4 leads to less transmission investment by the TO: only one line is built. As a consequence, market splitting moves the whole system closer to the welfare optimal scenario.

Prices for the case of an energy based fee can be seen in Fig. 5 and support the above intuition. In the welfare optimal solution, prices at the consumption nodes are low in the north (node 3) and high in the south (nodes 5 and 6). With one price zone the spot market price is relatively high. However, generators with high unit production cost have no chance to contract their supply at the spot market. Two price zones also imply a north-south spread which allows the southern generators to recover their investment cost.

8. Conclusion and Outlook

This paper analyzes the long-run impact of different transmission management regimes on investment incentives of generating firms in a market environment with a regulated TO. We propose a trilevel optimization approach to model an electricity market with cost-based redispatch both with and without market splitting. As a first best benchmark we also solve the corresponding integrated planner problem, in which a central planner controls both the grid and generation units. In order to solve our trilevel problem computationally, we present a reformulation that relies on a detailed analysis of the interconnection between the three problem levels. We apply our approach to two simple test instances in order to demonstrate the capabilities to analyze transmission and capacity expansion in a market environment.

Our results clearly show that in a market environment investment choices by the TO and private firms can substantially differ from welfare optimal choices. Obviously, investment in generation units is driven by the incentives for private investors induced by the particular market environment. Absence of proper incentives affects locational decisions of generators and this, in turn, can have substantial effects on optimal line investment. In two examples we demonstrate that welfare optimal line investment is affected significantly by the investment in generation capacity anticipated by the TO. We demonstrate that our model allows to compare different network management regimes and assess their effects on long-run investment decisions. Our approach is, thus, an important extension of various studies that up to now have mainly considered the short-run properties of different transmission management regimes. As we show, transmission management has also important implications in the long-run when generation and transmission expansion are taken into account.

Acknowledgements

This research was performed as part of the Energie Campus Nürnberg and supported by funding through the "Aufbruch Bayern (Bavaria on the move)" initiative of the state of Bavaria. We would like to thank Robert Bixby, Katja Breitmoser, Andreas Ehrenmann, Johannes Müller, and Yves Smeers for valuable comments and discussions. The authors thank the Deutsche Forschungsgemeinschaft for their support within Projekt A05 and B07 in the Sonderforschungsbereich/Transregio 154 "Mathematical Modelling, Simulation and Optimization using the Example of Gas Networks".

References

[1] K. Neuhoff, B. F. Hobbs, D. M. G. Newbery, Congestion management in european power networks: criteria to assess the available options, Tech. rep., DIW Berlin (2011).

[2] P. Joskow, Lessons learned from electricity market liberalization, The Energy Journal 29 (2) (2008) 9–42.

[3] R. A. Gallego, A. Monticelli, R. Romero, Transmission system expansion planning by an extended genetic algorithm, IEEE Proceedings - Generation, Transmission and Distribution 145 (3) (1998) 329–335.

[4] S. Binato, M. V. F. Pereira, S. Granville, A new benders decomposition approach to solve power transmission network design problems, IEEE Transactions on Power Systems 16 (2) (2001) 235–240.

[5] N. Alguacil, A. L. Motto, A. J. Conejo, Transmission expansion planning: a mixed-integer lp approach, IEEE Transactions on Power Systems 18 (3) (2003) 1070–1077.

[6] E. J. de Oliveira, I. C. da Silva Jr., J. L. R. Pereira, S. C. Jr., Transmission system expansion planning using a sigmoid function to handle integer investment variables, IEEE Transactions on Power Systems 20 (3) (2005) 1616–1621.

[7] J. J. Gabszewicz, S. Poddar, Demand fluctuations and capacity utilization under duopoly, Economic Theory 10 (1) (1997) 131–146.

[8] P. Joskow, J. Tirole, Reliability and competitive electricity markets, The RAND Journal of Economics 38 (1) (2007) 60–84.

[9] N. Fabra, N.-H. M. von der Fehr, M.-Á. de Frutos, Market design and investment incentives, The Economic Journal 121 (557) (2011) 1340–1360.

[10] M.-Á. de Frutos, N. Fabra, Endogenous capacities and price competition: The role of demand uncertainty, International Journal of Industrial Organization 29 (4) (2011) 399–411.

[11] F. H. Murphy, Y. Smeers, Generation capacity expansion in imperfectly competitive restructured electricity markets, Operations Research 53 (4) (2005) 646–661.

[12] G. Zöttl, A framework of peak load pricing with strategic firms, Operations Research 58 (6) (2010) 1637–1649.

[13] V. Grimm, G. Zöttl, Investment incentives and electricity spot market competition, Journal of Economics & Management Strategy 22 (4) (2013) 832–851.

[14] E. E. Sauma, S. S. Oren, Proactive planning and valuation of transmission investments in restructured electricity markets, Journal of Regulatory Economics 30 (3) (2006) 261–290.

[15] E. E. Sauma, S. S. Oren, Do generation firms in restructured electricity markets have incentives to support social-welfare-improving transmission investments?, Energy Economics 31 (5) (2009) 676–689.

[16] J. H. Roh, M. Shahidehpour, Y. Fu, Market-based coordination of transmission and generation capacity planning, IEEE Transactions on Power Systems 22 (4) (2007) 1406–1419.

[17] J. H. Roh, M. Shahidehpour, L. Wu, Market-based generation and transmission planning with uncertainties, IEEE Transactions on Power Systems 24 (3) (2009) 1587–1598.

[18] A. H. van der Weijde, B. F. Hobbs, The economics of planning electricity transmission to accommodate renewables: Using two-stage optimisation to evaluate flexibility and the cost of disregarding uncertainty, Energy Economics 34 (6) (2012) 2089–2101.

[19] S. M. Ryan, A. Downward, A. B. Philpott, G. Zakeri, Welfare effects of expansions in equilibrium models of an electricity market with fuel network, IEEE Transactions on Power Systems 25 (3) (2010) 1337–1349.

[20] S. Jin, S. M. Ryan, Capacity expansion in the integrated supply network for an electricity market, IEEE Transactions on Power Systems 26 (4) (2011) 2275–2284.

[21] M. Jenabi, S. M. T. F. Ghomi, Y. Smeers, Bi-level game approaches for coordination of generation and transmission expansion planning within a market environment, IEEE Transactions on Power Systems 28 (3) (2013) 2639–2650.

[22] R. P. O'Neill, E. A. Krall, K. W. Hedman, S. S. Oren, A model and approach to the challenge posed by optimal power systems planning, Mathematical Programming B 140 (2) (2013) 239–266.

[23] W. Hogan, Financial transmission right formulations, Report, Center for Business and Government, John F. Kennedy School of Government, Harvard University, Cambridge, MA.

[24] W. Hogan, Transmission congestion: the nodal-zonal debate revisited, Harvard University, John F. Kennedy School of Government, Center for Business and Government. Retrieved August 29.

[25] A. Ehrenmann, Y. Smeers, Inefficiencies in european congestion management proposals, Utilities policy 13 (2) (2005) 135–152.

[26] K. Neuhoff, J. Barquin, M. G. Boots, A. Ehrenmann, B. F. Hobbs, F. A. M. Rijkers, M. Vazquez, Network-constrained cournot models of liberalized electricity markets: the devil is in the details, Energy Economics 27 (3) (2005) 495–525.

[27] R. Green, Nodal pricing of electricity: how much does it cost to get it wrong?, Journal of Regulatory Economics 31 (2) (2007) 125–149.

[28] S. S. Oren, Economic inefficiency of passive transmission rights in congested electricity systems with competitive generation, The Energy Journal (1997) 63–83.

[29] W. Jing-Yuan, Y. Smeers, Spatial oligopolistic electricity models with cournot generators and regulated transmission prices, Operations Research 47 (1) (1999) 102–112.

[30] S. Borenstein, J. Bushnell, J. Stoft, The competitive effects of transmission capacity in a deregulated electricity industry, Tech. rep., National Bureau of Economic Research (1997).

[31] C. Metzler, B. F. Hobbs, J.-S. Pang, Nash-cournot equilibria in power markets on a linearized dc network with arbitrage: Formulations and properties, Networks and Spatial Economics 3 (2) (2003) 123–150.

[32] R. Gilbert, K. Neuhoff, D. Newbery, Allocating transmission to mitigate market power in electricity networks, The RAND Journal of Economics (2004) 691–709.

[33] X. Hu, D. Ralph, Using epecs to model bilevel games in restructured electricity markets with locational prices, Operations Research 55 (5) (2007) 809–827.

[34] P. Holmberg, E. Lazarczyk, Congestion management in electricity networks: Nodal, zonal and discriminatory pricing, Tech. rep., Research Institute of Industrial Economics (IFN) (2012).

[35] G. Oggioni, Y. Smeers, Market failures of market coupling and counter-trading in europe: An illustrative model based discussion, Energy Economics 35 (2013) 74–87.

[36] G. Oggioni, Y. Smeers, E. Allevi, S. Schaible, A generalized nash equilibrium model of market coupling in the european power system, Networks and Spatial Economics 12 (4) (2012) 503–560.

[37] M. Perninge, L. Söder, Irreversible investments with delayed reaction: an application to generation re-dispatch in power system operation, Mathematical Methods of Operations Research 79 (2) (2014) 195–224.

[38] D. Kirschen, G. Strbac, Transmission networks and electricity markets, Fundamentals of Power System Economics (2004) 141–204.

[39] ENTSO-E, 10-year network development plan 2012 (2012).

[40] J. Boucher, Y. Smeers, Alternative models of restructured electricity systems, part 1: No market power, Operations Research 49 (6) (2001) 821–838.

[41] O. Daxhelet, Y. Smeers, The EU regulation on cross-border trade of electricity: A two-stage equilibrium model, European Journal of Operational Research 181 (3) (2007) 1396–1412. doi:http://dx.doi.org/10.1016/j.ejor.2005.12.040.

[42] R. G. Jeroslow, The polynomial hierarchy and a simple model for competitive analysis, Mathematical programming 32 (2) (1985) 146–164.

[43] B. F. Hobbs, Lcp models of nash-cournot competition in bilateral and poolco-based power markets, in: IEEE Power Engineering Society 1999 Winter Meeting, Vol. 1, IEEE, 1998, pp. 303–308.

[44] H. Scheel, S. Scholtes, Mathematical programs with complementarity constraints: Stationarity, optimality and sensitivity, Mathematics of Operations Research 25 (2000) 1–22.

[45] R. W. Cottle, J.-S. Pang, R. E. Stone, The linear complementarity problem, Vol. 60, SIAM, 2009.

[46] S. Dempe, Foundations of bilevel programming, Springer, 2002.

[47] Z.-Q. Luo, J.-S. Pang, D. Ralph, Mathematical programs with equilibrium constraints, Cambridge University Press, 1996.

[48] V. Grimm, A. Martin, M. Weibelzahl, G. Zöttl, Peak load pricing on a network, working Paper (2015).

[49] S. Boyd, L. Vandenberghe, Convex Optimization, Cambridge University Press, Cambridge, 2004.

[50] H.-P. Chao, S. C. Peck, Reliability management in competitive electricity markets, Journal of Regulatory Economics 14 (2) (1998) 189–200.

[51] T. Koch, Rapid mathematical programming, Ph.D. thesis, Technische Universität Berlin (2004).
URL http://opus.kobv.de/zib/volltexte/2005/834/

[52] T. Achterberg, SCIP: Solving Constraint Integer Programs, Mathematical Programming Computation 1 (1) (2009) 1–41.

[53] CPLEX, User's Manual for CPLEX, IBM Corporation, Armonk, USA, 12th Edition (2013).

[54] ENTSO-E, Electricity price data, https://www.entsoe.eu/Pages/default.aspx, accessed: December 2012 (2012).

AppendixA. Notations and Symbols

Symbol	Description	Unit
$\mathcal{G}$	Transmission network	—
N	Set of nodes of the transmission network	—
N^{dem}	Set of demand nodes	—
T	Set of time periods	—
Z	Set of zones for market splitting	—
G_n^{all}	Set of all generation technologies at node $n \in N$	—
G_n^{ex}	Set of existing generation technologies at node $n \in N$	—
G_n^{new}	Set of candidate generation technologies at node $n \in N$	—
$\mathcal{L}$	Set of line types	—
L_ℓ^{ex}	Set of existing transmission lines of type $\ell \in \mathcal{L}$	—
L_ℓ^{new}	Set of candidate transmission lines of type $\ell \in \mathcal{L}$	—
L_ℓ^{inter}	Set of inter-zone transmission lines of type $\ell \in \mathcal{L}$	—
L^{ex}	Set of all existing transmission lines (set of arcs of graph $\mathcal{G}$)	—
L^{new}	Set of all candidate transmission lines	—
L^{inter}	Set of all inter-zone transmission lines	—
$d_{t,n}$	Demand at demand node $n \in N^{\text{dem}}$ in time period t	MW h
$p_{t,n}$	Price function at demand node $n \in N^{\text{dem}}$ at time period t	\$/MWh
s_n	Slope of $p_{t,n}$ at demand node $n \in N^{\text{dem}}$	\$/MWh^2
$a_{t,n}$	Intercept of $p_{t,n}$ at demand node $n \in N^{\text{dem}}$ in time period t	\$/MWh
c_g^{inv}	Investment cost of candidate generation technology $g \in G_n^{\text{new}}$	\$/MW
c_g^{var}	Variable cost of generation technology $g \in G_n^{\text{all}}$	\$/MWh
c_l^{inv}	Investment cost of candidate transmission line $l \in L^{\text{new}}$	\$
c_l^{del}	Degradation cost of existing transmission line $l \in L^{\text{ex}}$	\$
B_l	DC-power-flow-scaled susceptance of line $l \in L^{\text{ex}} \cup L^{\text{new}}$	MW h
$\theta_{t,n}$	Voltage angle in node $n \in N$ at time period t	rad
$f_{t,l}$	Power flow on line $l \in L^{\text{ex}} \cup L^{\text{new}}$ in time period t	MW h
$\bar{f}_l$	Maximum power flow on line $l \in L^{\text{ex}} \cup L^{\text{new}}$	MW h
α_g	Equivalent availability of generation technology $g \in G_n^{\text{all}}$	1
$y_{t,g}$	Power generation of generator $g \in G_n^{\text{all}}$	MW h
$\bar{y}_g^{\text{new}}$	New power generation capacity installed of generator $g \in G_n^{\text{new}}$	MW
$\bar{y}_g^{\text{ex}}$	Maximum power generation capacity of generator $g \in G_n^{\text{ex}}$	MW
E	Expenses of the transmission network operator	\$
R	Revenues of the transmission network operator	\$
φ^{ls}	Lump sum fee	\$
φ^{eb}	Energy based per unit fee	\$/MWh
φ^{gcb}	Generation capacity based per unit fee	\$/MW
z_l^{ex}	Decision variable for existing line $l \in L^{\text{ex}}$	—
z_l^{new}	Decision variable for candidate line $l \in L^{\text{new}}$	—