Transmission Network Investment Using Incentive Regulation: A Disjunctive Programming Approach.

**D. Khastieva · M. R. Hesamzadeh ·
I. Vogelsang · J. Rosellón.**

Abstract A well-planned electric transmission infrastructure is the foundation of a reliable and efficient power system, especially in the presence of large scale renewable generation. However, the current electricity market designs lack incentive mechanisms which can guarantee optimal transmission investments and ensure reliable integration of renewable generation such as wind. This paper first proposes a stochastic bilevel disjunctive program for optimal transmission investment based on the newly proposed theoretical H-R-G-V incentive mechanism. The upper level is a profit-maximization problem of an independent transmission company (Transco), while the lower level is a welfare maximization problem. The revenue of the Transco is bounded by a regulatory constraint set by the regulator in order to induce socially optimal investments. The application of the H-R-G-V mechanism allows the regulator to ensure social maximum transmission investments and helps to reduce transmission congestion and wind power spillage. The transmission investment under the H-R-G-V mechanism is modeled as a stochastic bilevel disjunctive program. To solve the developed mathematical model we first propose a series of linearization and reformulation techniques to recast the original model as a stochastic mixed integer linear problem (MILP). We exploit the disjunctive nature of the reformulated stochastic MILP model and further propose a Bean decomposition algorithm to efficiently solve the stochastic MILP model. The proposed decomposition algorithm is also modified and accelerated to improve the computational performance. The computational performance of our MILP modeling approach and modified and accelerated Bean decomposition algorithm is studied through several examples in detail. The simulation results confirm a promising performance of both the modeling approach and its solution algorithm.

Keywords Transmission network investments · Incentive regulation · Bean decomposition algorithm · Disjunctive programming.

Dina Khastieva
KTH, Teknikringen 33, 114 28 Stockholm
Tel.: +46 760 744 605
E-mail: dinak@kth.se

1 Introduction

Optimal investment in transmission networks is a major concern in electricity markets around the world, largely due to growing electricity demand and increasing penetration of variable renewable generation. While the generation and retail sectors are effectively managed through competitive market rules, the transmission sector remains a natural monopoly and in many states is managed through independent profit-maximizing transmission companies (Transco). A Transco is responsible for the transmission lines' maintenance and the investments in additional transmission capacity. The reimbursement of transmission investment costs depends on the regulation and the incentive mechanisms adopted by the regulatory entity of the state. Currently the transmission sector is experiencing a lack of investment incentives (Rosellón et al 2018), (Hesamzadeh et al 2018). With the growth of renewables the absence of adequate incentive mechanisms will result in increased transmission congestion costs which will negatively affect social welfare (European Commission 2003), (Dyer 2003).

Various incentive mechanisms were proposed to tackle the incentive problem. They can be divided into two major groups, subsidy mechanisms and constraint mechanisms. Subsidy mechanisms were initially introduced by (Loeb and Magat 1979) and further developed by (Sappington and Sibley 1988) where an incremental surplus subsidy scheme (ISS) was proposed. The mechanism then was applied to transmission pricing and investments in (Gans and King 2000). On the other hand, constraint mechanisms were proposed by (Vogelsang 2001) and (Tanaka 2007), who use price-cap constraints to incentivize optimal transmission investment by a Transco. Under certain conditions, these mechanisms lead to transmission investment decisions which maximize social welfare (Hogan et al 2010). (Joskow and Tirole 2002) propose a reward/penalty mechanism. In this mechanism, the regulator rewards the Transco when the transmission network is expanded and the merchandising surplus as well as network congestion are decreased. (Léautier 2000) proposes an out-turn mechanism. The out-turn is defined as the difference between actual electricity prices and prices without transmission congestion. The Transco is responsible for total out-turn cost and any transmission losses. Variations of such incentive mechanisms were widely applied on energy infrastructure investments (Neumann et al 2015). More recently, an alternative incentive mechanism for transmission investments has been proposed by (Hesamzadeh et al 2018) following the incentive mechanisms in (Sappington and Sibley 1988) and (Hogan et al 2010). The H-R-G-V (Hesamzadeh-Rosellon-Gabriel-Vogelsang) mechanism envisages a dynamic interaction between a profit-maximizing Transco, the regulator and an Independent System Operator (ISO). The H-R-G-V mechanism combines the price cap approach with the ISS scheme. A regulatory entity (regulator) sets a charge for transmission network users (generators and loads) with variable fees corresponding to nodal price differences that are related to the merchandising surplus. In addition, generators and loads are charged with a fixed fee. This mechanism induces social welfare maximizing investments and efficient nodal prices for all planning periods. The comparison of advantages of aforementioned inventive mechanisms is presented in Table 1.

Theoretical justifications that the H-R-G-V mechanism can be effectively used for transmission investments have been presented in (Vogelsang 2018). Mathematically, the transmission investment problem under the H-R-G-V incentive mecha-

Table 1 Comparison of different incentive mechanisms.

Advantages:	Cost-Plus	ISS	H-R-V	H-R-G-V
Does not involve subsidies	yes	no	yes	**yes**
Guarantees socially optimal investments	no	yes	no	**yes**
Based on market information	no	yes	yes	**yes**
Promotes competitive behavior	no	no	yes	**yes**
Simple to model	yes	yes	no	**yes**
Convergence to global solution is guaranteed	yes	yes	no	**yes**

nism is formulated as a non-linear disjunctive program. The mathematical formulation and the solution algorithm previously proposed for modeling the H-R-G-V regulation are applied only to a small example system and they are not practical for real-size applications (Hesamzadeh et al 2018). Moreover, the application of the H-R-G-V mechanism in (Hesamzadeh et al 2018) and in (Vogelsang 2018) was considered only under deterministic input parameters. The growth of variable renewable generation will result in increased stochasticity in the system which should be taken into account in transmission investments as well. Thus, in this paper we extend the application of the mechanism to a stochastic framework and propose a solution algorithm which will mathematically guarantee a globally optimal solution.

The resulting initial model is the deterministic equivalent of a stochastic bilevel disjunctive program with integer variables. The solution of the formulated problem is complicated through several non-linear terms and disjunctive constraints. Thus, through a series of proposed linearization techniques, the initial bilevel program is transformed into a mixed-integer linear program (MILP) with disjunctive constraints. The disjunctive constraints are also complicating the solution process so that state-of the art solvers cannot be applied directly with a guarantee of finding a global solution.

A disjunctive program can be reformulated into a mixed-integer program using several existing techniques including the convex hull, cutting planes and disjunctive constraint linearization techniques[1]. All these techniques provide a reformulation of the original feasible sets and there are limitations specific for each technique. Convex hull methods are proved to provide a tight reformulation. Nevertheless, the approach requires additional variables and constraints which considerably increase the size of the problem for transmission investment models. On the other hand, the disjunctive constraint linearization technique does not affect the size of the problem. However, the disjunctive parameters involved in the reformulation create computational issues for the solver. A disjunctive parameter that is not tuned affects the convergence of the problem (Hooker 2011). The literature provides several methodologies for tuning the disjunctive parameter. The methodologies for tuning the disjunctive parameter can be found in (Trespalacios and Grossmann 2015) and (Hooker 2011). These methodologies are proved to provide good approximations of the disjunctive parameters under certain conditions but additional large scale optimization problems need to be solved for each case and optimality of the outcome still cannot be guaranteed. The problem of the disjunc-

---

[1] The disjunctive constraint linearization technique is also called the big-M reformulation techniques in some literature. Please see references (Hooker 2011) and (Trespalacios and Grossmann 2015)

tive parameter tuning becomes especially hard when the reformulation involves variables without physical upper or lower bounds which is the case in our proposed model.

Various solution algorithms can be considered to solve the reformulated disjunctive program. In (Garces et al 2009) and (Maurovich-Horvat et al 2015) state-of-the-art solvers such as CPLEX are used to solve the transmission investment model. Such methodology guarantees convergence to an optimal solution, however the application to our proposed stochastic MILP will require additional tuning of the disjunctive parameter. In addition, the tractability of the solution is low for large-scale problems (Bertsimas et al 2018). In the large-scale problems, the number of variables and constraints might exceed the level which can be handled by state-of-the-art solvers such as CPLEX. Accordingly, solvers such as CPLEX might not be able to find the optimal solution efficiently (tractability issue). In these cases, it is suggested to decompose the large optimization problem into smaller ones with reduced number of variables and constraints. In this way, the optimal solution can be found more efficiently (improved tractability). One of the most widely used decomposition algorithms is Benders decomposition algorithm. In (Conejo et al 2006) and (McCusker and Hobbs 2003), the Benders decomposition algorithm is used to improve computational tractability considering the uncertainty in the system. The Benders decomposition algorithm proves to be an effective tool and it reduces the computational complexity substantially. While the Benders decomposition algorithm is proved to be effective and can be applied to transmission investment models it does not solve the complications arising from incorporating the incentive mechanism into the transmission investment model. Moreover, the presence of a hard-to-tune disjunctive parameter will prevent a direct application of the Benders decomposition algorithm to our proposed MILP. A sub-optimally tuned disjunctive parameter will result in weak Benders cuts and will cause reduced tractability of the problem (Hooker 2011),(Codato and Fischetti 2004).Additional Gomory cuts were proposed to tackle the problem of disjunctive parameters in (Binato et al 2001), however the proposed approach achieves only an approximate disjunctive parameter without guaranteeing that it was tuned to optimality. Thus, the application of the Benders decomposition algorithm to disjunctive problems has challenges associated with disjunctive parameters.

Employing the disjunctive nature of our reformulated MILP model, we propose an algorithm based on the Bean decomposition algorithm. The Bean decomposition algorithm follows the Benders decomposition algorithm and these two algorithms can be directly compared. The Bean decomposition algorithm directly exploits the disjunctive nature of our proposed MILP model and it has a better computational tractability than Benders decomposition algorithm. We also modify the Bean decomposition algorithm such that disjunctive parameters are completely removed from solution algorithm. Accordingly, they are not needed to be optimally tuned. Moreover the modified Bean decomposition algorithm is accelerated by using some additional constraints. The benefits of our proposed algorithm in comparison to existing solution methodologies are presented in Table 2.

Accordingly, the main contributions of the current paper can be highlighted through the following points:

Table 2 Comparison of solution methodologies. Disj.: Disjunctive

	Global solution	Tuning of disj. parameter is required	Disj. parameter optimality guaranteed	Improved tractability
Direct application of state-of-the-art solvers	yes	yes	no	no
Standard Benders decomposition algorithm	yes	yes	no	no
Benders decomposition with disjunctive parameter tuning	yes	no	no	no
Standard Bean decomposition algorithm	yes	yes	no	no
Our proposed decomposition algorithm	**yes**	**no**	**yes**	**yes**

– The current paper presents an extension of the theoretical H-R-G-V incentive mechanism to the stochastic modeling framework. Wind generation uncertainty is considered in the proposed stochastic framework. The resulting model is a stochastic bilevel disjunctive program which is hard to solve. Accordingly, a series of reformulation and linearization techniques are proposed to recast the original model into an easier-to-solve stochastic MILP with disjunctive constraints.

– The paper then proposes a specialized decomposition algorithm based on the Bean decomposition algorithm to solve the derived MILP model. The proposed decomposition algorithm is aimed at guaranteeing the convergence to the globally optimal solution with good computational tractability and to avoid tuning of the disjunctive parameters. Moreover, it is accelerated by using some additional constraints.

– Several case studies of different size are presented to illustrate the performance of the incentive mechanism and solution methodology. We have demonstrated that the H-R-G-V incentive mechanism can be effectively used for transmission investment. The numerical results show that the proposed decomposition algorithm outperforms the standard Benders decomposition algorithm and it can be used effectively to obtain a globally optimal solution with better computational tractability.

The current paper is organized as follows. Section II describes the incentive-based regulated transmission investment mechanism and formulates the mathematical model. Section III then introduces a stochastic bilevel disjunctive program for the transmission investment planning problem based on the H-R-G-V mechanism. In Section IV the stochastic MILP reformulation of the proposed model is presented. Section V contains a short description of the standard Benders decomposition algorithm followed by our proposed decomposition algorithm. An illustrative example, case studies and the results are presented in Section VI. Section VII provides conclusions. Finally, a short discussion of reliability issues as well as a description of scenario generation methodology are presented in Appendix A.1 and Appendix A.2.

2 Incentive-based regulation: introduction to the H-R-G-V regulatory constraint

The incentive-based regulated transmission investment assumes that three independent organizations are involved in the investment planning and operation of a power system. The first organization is the welfare maximizing independent system operator (ISO) which operates dispatchable conventional and renewable generation assets and calculates the merchandizing surplus. The second organization is a regulated independent transmission company (Transco) which owns a transmission network and is responsible for transmission investment planning and for setting a fixed fee for loads and generators for transmission investments expenses. Finally, the third organization is a regulator which is responsible for providing proper regulatory mechanisms to ensure socially optimal investment decisions, meaning that the regulator has a social welfare maximizing objective. The interaction between ISO, regulator and Transco for transmission investments is illustrated in Figure 1.
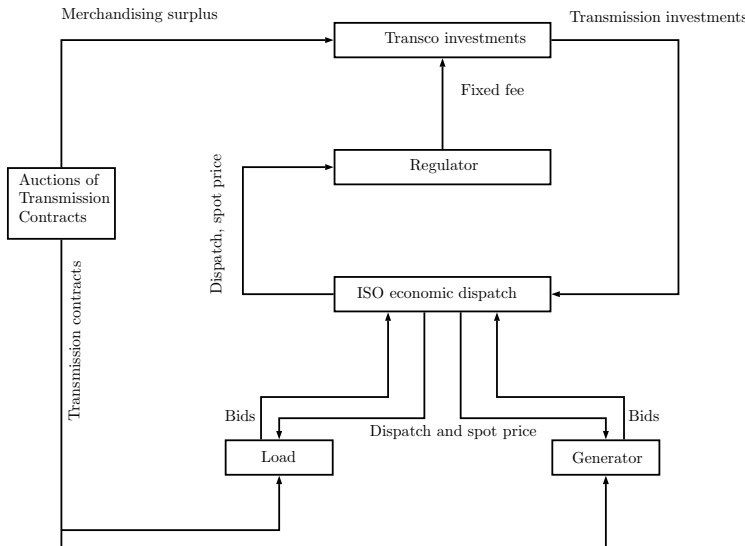

Fig. 1 Incentive-based regulation for transmission investments

The Transco maximizes its profit by expanding its transmission network while considering a fixed fee calculated by the regulator. The Transco communicates transmission investment decisions to the ISO. The ISO dispatches the system and communicates the required information to the regulator. The regulator calculates the fixed fee and communicates the fixed fee to the Transco. The Trasco also receives the merchandising surplus through auctioning transmission contracts.

Under H-R-G-V incentive mechanism the regulatory constraint links the fixed fee to the generator surplus and load surplus. These generator and load surplus

are influenced by transmission investment decisions of Transco. The fixed fee is closely related to the social welfare increase resulting from the network investments. The fixed fee paid by the network users could be seen as a substitute to subsidy payments by the government or the regulator. In order to make a decision on new line investments and transmission tariffs, the Transco has to take into account expected generation dispatch, electricity demand and nodal prices which are the outputs of the problem of the welfare maximizing ISO. In order to provide the expected dispatch of the system, the ISO collects the bids from generators and loads. Under the assumption of perfect competition between generators and between loads the interaction of ISO, generators and loads can be modeled as cost-minimizing dispatch of a power system. Thus, the regulated transmission investments can be formulated as an interaction between ISO, Transco and regulator. From the modeling perspective, the simulation of the interaction between regulator and Transco can be merged. Under the H-R-G-V incentive mechanism the regulator can be represented effectively through the regulatory constraint applied to the operation and investment planning of the Transco. This will lead us to bilevel formulation of the regulated transmission investment problem.

In the upper-level problem, the Transco maximizes its profit which consists of the sum of its network merchandising surplus and a fixed fee to transmission network users (which is limited by the regulatory constraint) minus total investment costs. In the lower-level problem, the ISO takes into account the investment decisions made by the Transco and dispatches generation and loads by maximizing social welfare. The optimal dispatch is used to calculate the merchandising surplus, load surplus and generation surplus. The interaction between upper-level problem and lower-level problem is illustrated in Figure (2).
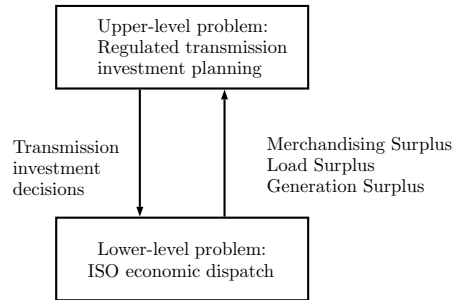

Fig. 2 Relation between upper-level problem and lower-level problem under H-R-G-V incentive regulation

The following assumptions are considered in this paper:

- The Transco and regulator have all information about economic dispatch model of the ISO. The fixed fee is calculated by the regulator and considered by the Transco. The transmission investment costs are only known by the Transco.
- We assume that the maximum demand is expected to grow by a certain percentage each of the following investment planning periods. Moreover, we assume

that demand is horizontal until it reaches a maximum. The maximum demand
level is defined separately for each power system bus.

– Uncertainty comes only from wind generation and it is modeled using a moment-
  matching technique. The demand uncertainty is considered to be relatively low
  and negligible. A short discussion of incorporating additional uncertainties can
  be found in the Appendix A.2.
– We assume that the dispatch is made in a merit order. This means perfect
  competition in generation and load sector.
– The Transco is assumed to be a risk-neutral regulated profit-maximizing entity.
– The regulator has the power to affect the profit of the Transco by setting a
  limit on the fixed fee. The main objective of the regulator is to maximize social
  welfare.
– The transmission investment decision in a particular line corridor is performed
  only once throughout the planning horizon and the investment decision is not
  reversible (Xifan and McDonald 1994). The capacities of the candidate lines
  considered for investment are fixed and known in advance.

The H-R-G-V regulatory constraint is based on the overall change of generation
and demand surplus and is formulated as:

$$\Delta\Phi_t = \Delta\,\mathbb{E}[\pi_t^G] + \Delta\,\mathbb{E}[\pi_t^L] \tag{1}$$

Where Φ_t is a fixed fee at investment period t to be set by the regulator and paid
to the Transco and $\Delta\Phi_t = \Phi_t - \Phi_{t-1}$. $\mathbb{E}[\pi_t^G]$ and $\mathbb{E}[\pi_t^L]$ are the expected generation
and demand surplus at investment period t while $\Delta\,\mathbb{E}[\pi_t^G] = \mathbb{E}[\pi_t^G] - \mathbb{E}[\pi_{t-1}^G]$ and
$\Delta\,\mathbb{E}[\pi_t^L] = \mathbb{E}[\pi_t^L] - \mathbb{E}[\pi_{t-1}^L]$. Mathematically, the bilevel regulated transmission
investment problem can be expressed as:

$$Maximize \sum_{t\in\mathcal{T}}(\mathbb{E}[\pi_t^T] + \Phi_t - \overline{C}_t^T) \tag{2a}$$

$Subject\ to$:

$\qquad$ Transmission investment decision constraint $\;\forall t\in\{\mathcal{T}\backslash t1\}$ $\qquad$ (2b)

$$\Delta\Phi_t = \Delta\,\mathbb{E}[\pi_t^G] + \Delta\,\mathbb{E}[\pi_t^L]\;\;\forall t\in\{\mathcal{T}\backslash t1\} \tag{2c}$$

$$\Phi_{t1} = 0 \tag{2d}$$

$$Maximize \sum_{t\in\mathcal{T}}(\mathbb{E}[\pi_t^G] + \mathbb{E}[\pi_t^T] + \mathbb{E}[\pi_t^L])$$

$\qquad Subject\ to$: system's technical constraints. $\qquad\qquad\qquad\qquad$ (2e)

Here $\mathbb{E}[\pi_t^T]$ is the merchandising surplus earned by the Transco in the spot market.
$\overline{C}_t^T$ and Φ_t are the total transmission investment cost and the fixed fee respectively.
In addition, we assume that the first investment planning period $t1$ is a status quo
period with no investment decisions and no fixed fee.

$\qquad$ Social Welfare (SW_t) in the context of this paper is defined for each investment
planning period as

$$SW_t = \sum_{t\in\mathcal{T}}(\mathbb{E}[\pi_t^G] + \mathbb{E}[\pi_t^T] + \mathbb{E}[\pi_t^L] - \overline{C}_t^T) \tag{3}$$

If we reformulate the objective function (2a) by replacing the fixed fee with the regulatory constraint we obtain (Hesamzadeh et al 2018):

$$\sum_{t\in\mathcal{T}}(\mathbb{E}[\pi_t^T] + \Phi_t - \overline{C}_t^T) = \sum_{t\in\{\mathcal{T}\setminus t1\}}(\mathbb{E}[\pi_t^G] + \mathbb{E}[\pi_t^T] + \mathbb{E}[\pi_t^L] - \overline{C}_t^T) + \mathbb{E}[\pi_{t=t1}^T] +$$
$$(|\mathcal{T}| - 1)(\mathbb{E}[\pi_{t=t1}^G] + \mathbb{E}[\pi_{t=t1}^L]) \qquad (4)$$

The reformulated objective function (4) shows that the regulated objective function of the Transco is equivalent to the social welfare objective (sum of generator, load and transmission surplus minus investment cost) at investment planning periods $t \in \{\mathcal{T}\setminus t1\}$. Transco does not perform investments in period $t1 \in \mathcal{T}$ and the term $(|\mathcal{T}| - 1)(\mathbb{E}[\pi_{t=t1}^G] + \mathbb{E}[\pi_{t=t1}^L])$ will not affect the investment results. Thus, the H-R-G-V mechanism promotes efficient cost allocation of the transmission investments and social maximizing investment decisions.

In (2) the ISO objective function is formulated as maximization of total generator, load and transmission surplus in the spot market. This maximization is equivalent to total load utility minus the total generation cost in the spot market. We assume perfect competition between generators and between loads. Thus, the mathematical formulation of regulated transmission investment with the H-R-G-V incentive mechanism can be presented in general terms as:

$$\textit{Maximize}\quad \textit{Merchandizing surplus + Fixed fee}$$
$$\textit{- Total transmission investment cost} \qquad (5a)$$
$$\textit{Subject to:}$$
$$\textit{H-R-G-V regulatory constraint for each planning period} \qquad (5b)$$
$$\textit{Linearized transmission investment constraints} \qquad (5c)$$
$$\textit{Maximize}\quad \textit{(Total load utility - total generation cost)} \qquad (5d)$$
$$\textit{Subject to:} \qquad (5e)$$
$$\textit{Energy balance constraint} \qquad (5f)$$
$$\textit{Power flow constraints} \qquad (5g)$$
$$\textit{Upper and lower operation bounds} \qquad (5h)$$

2.1 Nomenclature

Indices and Sets

$t, \hat{t} \in \mathcal{T}$	Investment planning periods;
$k \in \mathcal{K}$	Operation sub-periods;
$n \in \mathcal{N}$	Buses ;
$i \in \mathcal{D}$	Demand;
$j \in \mathcal{G}$	Generation;
$w \in \mathcal{W}$	Wind generation;
$l \in \mathcal{L}$	Existing transmission lines;
$m \in \mathcal{M}$	Candidate transmission lines;
$s \in \mathcal{S}$	Wind power generation scenarios;

Parameters

$\Xi^{(m)}$	Disjunctive parameters used to linearize power flow constraints of candidate lines
$\Xi^{(T)}$	Disjunctive parameters used to linearize terms T1 and T2
$\overline{C}_{m,t}$	Investment cost of new line m (\$);
c_j	Marginal cost of generator j, (\$/MWh);
α_i	Intercept of linear utility i, (\$/MWh);
β	Expected annual price escalation rate;
γ	Expected periodic rate of return;
π_s	Probability of scenario s;
$I_{ni}^{(n)}$	Elements of incidence matrix I which shows the relation between set $\mathcal{N}$ and $\mathcal{D}$;
$J_{nj}^{(n)}$	Elements of incidence matrix J which shows the relation between set $\mathcal{N}$ and $\mathcal{G}$;
$W_{nw}^{(n)}$	Elements of incidence matrix W which shows the relation between set $\mathcal{N}$ and $\mathcal{W}$;
$S_{nl}^{(n)}, R_{nl}^{(n)}$	Elements of incident matrices S and $\overline{S}$ which shows the relation between set $\mathcal{N}$ and $\mathcal{L}$;
$R_{n,l}, \overline{R}_{n,m}$	Elements of incident matrices R and $\overline{R}$ which shows the relation between set $\mathcal{N}$ and $\mathcal{M}$;
F_l	Maximum capacity of line l, (MWh);
$\widehat{F}_m$	Maximum capacity of line m, (MWh);
X_l	Reactance of line l, $(p.u.)$;
X_m	Reactance of line m, $(p.u.)$;
G_j	Maximum production of generator j , (MWh);
$\widehat{G}_{wks}$	Maximum production of wind generator w in scenario s in operation sub-period k, (MWh);
D_{it}	Maximum demand i at planning period t, (MWh);
Θ	Voltage angle limits, $(p.u)$;
Ψ	Number of operation sub-periods in one year;

Binary variables

z_{mt}, y_{mt}	Investment variables for line m at investment period t;

Continuous variables

d_{itks}	Consumption of demand i at investment period t, sub-period k and scenario s, (MWh);
g_{jtks}	Production of generator j at investment period t, sub-period k and scenario s, (MWh);
$\widehat{g}_{wtks}$	Production of wind generator w at investment period t, sub-period k and scenario s, (MWh);
f_{ltks}	Flow of line l at investment period t, sub-period k and scenario s, (MWh);
$\widehat{f}_{mtks}$	Flow of line m at investment period t, sub-period k and scenario s, (MWh);
θ_{ntks}	Voltage angle at bus n, investment period t, sub-period k and scenario s, $(p.u.)$;
Φ_t	Fixed fee of Transco at investment period t, (\$);

2.2 The welfare-maximizing ISO

The ISO performs economic dispatch by maximizing the total load utility minus the generation cost given the available transmission lines. The power system consists of buses, which represent demand or generation or both. The buses are connected through existing (l) or newly built (m) transmission lines. Incident matrices $S, R, \overline{S}$ and $\overline{R}$ are used to link sending and receiving buses to existing and candidate lines. The elements of the matrices $S_{nl}^{(n)}, R_{nl}^{(n)}, \overline{R}_{nm}^{(n)}$ and $\overline{S}_{nm}^{(n)}$ will be equal to one if bus n is a sending or receiving bus for line l or m. By analogy incident matrices I, J and W with elements $I_{ni}^{(n)}, J_{nj}^{(n)}$ and $W_{nw}^{(n)}$ are used to map generators and loads to buses of the power system. The economic dispatch problem is presented by a linear optimization problem described in (6) that models the dispatch for a given investment planning period t, an operation sub-period k and a scenario realization s. The assumed scenarios are the result of fluctuating wind power generation which is modeled using a moment-matching scenario generation technique (see Appendix A.2). The detailed description of power system modeling approaches can be found in (Leuthold et al 2012).

The objective of the optimization problem (6) is to maximize the utility of demand minus the cost of generation. $\Omega_s = \{d_{itks}, g_{jtks}, \widehat{g}_{wtks}, f_{ltks}, \widehat{f}_{mtks}, \theta_{ntks}\}$ is the set of decision variables of (6).

$$Maximize_{\Omega_s} \sum_{i \in \mathcal{D}} \alpha_i d_{itks} - \sum_{j \in \mathcal{G}} c_j \, g_{jtks} \tag{6a}$$

$Subject\ to$:

$$\sum_{j \in \mathcal{G}} J_{nj}^{(n)} g_{jtks} + \sum_{w \in \mathcal{W}} W_{nw}^{(n)} \widehat{g}_{wtks} - \sum_{i \in \mathcal{D}} I_{ni}^{(n)} d_{itks} - \sum_{l \in \mathcal{L}} S_{nl}^{(n)} f_{ltks} + \sum_{l \in \mathcal{L}} R_{nl}^{(n)} f_{ltks} -$$

$$\sum_{m \in \mathcal{M}} \overline{S}_{nm}^{(n)} \widehat{f}_{mtks} + \sum_{m \in \mathcal{M}} \overline{R}_{nm}^{(n)} \widehat{f}_{mtks} = 0 \ \forall n \in \mathcal{N}, t \in \mathcal{T}, k \in \mathcal{K}, s \in \mathcal{S} \tag{6b}$$

$$-\frac{100}{X_l} \left(\sum_{n \in \mathcal{N}} S_{nl}^{(n)} \theta_{ntks} - \sum_{n \in \mathcal{N}} R_{nl}^{(n)} \theta_{ntks} \right) + f_{ltks} = 0 \ \forall l \in \mathcal{L}, t \in \mathcal{T}, k \in \mathcal{K}, s \in \mathcal{S} \tag{6c}$$

$$\begin{bmatrix} \widehat{f}_{mtks} = 0 \\ z_{mt} = 0 \end{bmatrix} \vee \begin{bmatrix} \widehat{f}_{mtks} - \frac{100}{X_m} (\sum_{n \in \mathcal{N}} \overline{S}_{nm}^{(n)} \theta_{ntks} - \sum_{n \in \mathcal{N}} \overline{R}_{nm}^{(n)} \theta_{ntks}) = 0 \\ z_{mt} = 1 \end{bmatrix}$$

$\forall m \in \mathcal{M}, t \in \mathcal{T}$ \hfill (6d)

$$-\widehat{F}_m \le \widehat{f}_{mtks} \le \widehat{F}_m \ \forall m \in \mathcal{M}, t \in \mathcal{T}, k \in \mathcal{K}, s \in \mathcal{S} \tag{6e}$$

$$-F_l \le f_{ltks} \le F_l \ \forall l \in \mathcal{L}, t \in \mathcal{T}, k \in \mathcal{K}, s \in \mathcal{S} \tag{6f}$$

$$0 \le g_{jtks} \le G_j \ \forall j \in \mathcal{G}, s \in \mathcal{S}, t \in \mathcal{T}, k \tag{6g}$$

$$0 \le \widehat{g}_{wtks} \le \widehat{G}_{wks} \ \forall w \in \mathcal{W}, t \in \mathcal{T}, k \in \mathcal{K}, s \in \mathcal{S} \tag{6h}$$

$$0 \le d_{itks} \le D_{it} \ \forall i \in \mathcal{D}, t \in \mathcal{T}, k \in \mathcal{K}, s \in \mathcal{S} \tag{6i}$$

$$\theta_{n=n_1 tks} = 0 \ \forall t \in \mathcal{T}, k \in \mathcal{K}, s \in \mathcal{S} \tag{6j}$$

$$d_{itks}, g_{jtks}, \widehat{g}_{wtks}, f_{ltks}, \widehat{f}_{mtks}, \theta_{ntks} \in \mathbb{R} \tag{6k}$$

Following the general formulation in (5) the energy balance constraints (5f) at each bus are formulated in (6b). Power flows constraints (5g) are modeled as in (6c) for existing transmission lines and through disjunctive constraints (6d) for

candidate lines. Integer variables z_{mt} in optimization problem (6) are considered to be parameters and are decision variables of the upper-level problem which will be discussed in the following subsection. Thermal limits of the lines are set in (6f) and (6e). Finally, upper and lower limits on generation, demand and voltage angles are enforced through (6g)-(??). The maximum demand is assumed to change for each investment planning period to reflect the growing trend of the demand over the years. The maximum wind generation is assumed to be a stochastic parameter and varies for each investment planning period t, operation sub-period k and scenario realization s. Constraint (6j) sets bus $n = n1$ as the reference bus.

2.3 The profit-maximizing Transco

The profit maximization model of the Transco is presented through the stochastic bilevel disjunctive program in (7) .

$$
\underset{z_{mt},y_{mt},\Phi_t}{Maximize} \sum_{t\in\mathcal{T}} \frac{(1+\beta)^t}{(1+\gamma)^t} \langle \Phi_t + \Psi \sum_s \pi_s (\sum_{n,i,k} I_{ni}^{(n)} \lambda_{ntks} d_{itks} - \sum_{n,j,k} J_{nj}^{(n)} \lambda_{ntks} g_{jtks} -
$$

$$
\sum_{n\in\mathcal{N},w\in\mathcal{W}} W_{nw}^{(n)} \lambda_{ntks} \widehat{g}_{wtks}) - \sum_{m\in\mathcal{M}} \overline{C}_{m,t} y_{mt} \rangle \tag{7a}
$$

$Subject\ to\ :$

$$
z_{m,t=t1} = 0 \ \forall m \in \mathcal{M} \tag{7b}
$$

$$
z_{mt} = \sum_{\widehat{t}\in\{t1,..t\}} y_{m,\widehat{t}} \ \forall m\in\mathcal{M},t\in\mathcal{T} \tag{7c}
$$

$$
\sum_{t\in\mathcal{T}} y_{mt} \leq 1 \ \forall m\in\mathcal{M},t\in\mathcal{T} \tag{7d}
$$

$$
\Phi_{t=1} = 0 \tag{7e}
$$

$$
\mathbb{E}[\pi_t^L] = \sum_s \pi_s (\sum_{i,k} \alpha_i d_{itks} - \sum_{n,i,k} I_{ni}^{(n)} \lambda_{ntks} d_{itks}) \ \ \forall t\in\mathcal{T} \tag{7f}
$$

$$
\mathbb{E}[\pi_t^G] = \sum_s \pi_s (\sum_{n,j,k} (J_{nj}^{(n)} \lambda_{ntks} g_{jtks} - \sum_{j,k} c_j g_{jtks}) + \sum_{n,w,k} W_{nw}^{(n)} \lambda_{ntks} \widehat{g}_{wtks}) \ \ \forall t\in\mathcal{T}
$$
$$\tag{7g}$$

$$
\Phi_t - \Phi_{t-1} = \Psi(\mathbb{E}[\pi_t^L] - \mathbb{E}[\pi_{s,t-1}^L] + \mathbb{E}[\pi_t^G] - \mathbb{E}[\pi_{s,t-1}^G]) \ \ \forall t\geq t2\in\mathcal{T} \tag{7h}
$$

$$
z_{mt} \in \{0,1\} \tag{7i}
$$

$Where\ \{d_{itks}, g_{jtks}, \widehat{g}_{wtks}, \lambda_{ntks}\} \in$

$$
\{\arg \underset{\Omega_s}{Maximize} \sum_{i\in\mathcal{D}} \alpha_i d_{itks} - \sum_{j\in\mathcal{G}} c_j \ g_{jtks} \tag{7j}
$$

$Subject\ to\ :$

$$
(6b) - (6k)\} \ \ \ \forall t\in\mathcal{T},k\in\mathcal{K},s\in\mathcal{S} \tag{7k}
$$

Optimization problem (7) simulates the bilevel program of the Transco under the H-R-G-V incentive mechanism. The optimization problem (6) is the lower-level program of the optimization problem (7). The objective of the Transco (7a) is

to maximize its expected profit over the investment planning period (t). The expected revenue of the Transco is the expected merchandising surplus. The profit of the Transco is modeled as an expected value due to the stochasticity of wind generation and consequently the stochasticity of ISO economic dispatch. Each investment period (t) consists of operation sub-periods (k) which take into account representative hours of each year. In order to reflect the time value of the investment decision the total profit of the Transco for each investment planning period t is discounted using the present value factor calculated as $\frac{(1+\beta)^t}{(1+\gamma)^t}$. The present value factor is based on the expected rate of price escalation between investment periods β and the expected periodic rate of return γ. The parameter (ψ) is used to match short-term operation costs incurred in each operation sub-period k with long term investment costs incurred in each investment planning period t. For daily operation sub-periods ($\mathcal{K} = \{k1, ...k24\}$) and yearly investment periods, the parameter Ψ will be set to 365 while for hourly operation and yearly investment periods parameter Ψ will be set to 3870. The total Transco profit is the sum of the network merchandising surplus and a fixed fee to loads and generators (Φ_t) minus the investment costs ($\sum_{m \in \mathcal{M}} \overline{C}_{m,t} y_{mt}$). The fixed fee ($\Phi_t$) is set by the regulatory constraint (7h) which is modeled according to the formulation provided in (1). Generation surplus is calculated as a difference in generation revenue from operating in the spot market and generation operating costs as (7g). Similarly, demand surplus is calculated as a difference between the total benefit of consumption calculated using a linear utility function and the cost of energy purchased in the spot market. Generation surplus and demand surplus are affected by the stochastic nature of wind generation and thus should be included as expected values to consider different possible scenarios of economic dispatch. Earlier we have assumed that the first investment planning period $t1$ is a status quo period with neither investments nor a fixed fee. Thus, additional constraints are introduced to set the investment decision y_{mt} and Φ_t to zero at the initial investment planning period $t1$. We assume linear investment costs. The investment decision is taken through binary variable y_{mt} which is equal to 1 if an investment at period t is made and 0 otherwise. Constraint (7d) ensures that the decision to invest in line m is taken only once and it is irreversible. At the same time an additional variable z_{mt} is introduced to capture whether or not a candidate line m exists in any given planning period t. The variable z_{mt} is introduced to simplify the formulation of the bilevel problem and avoids the need to use $\sum_{t \in \mathcal{T}} y_{mt}$ in the lower-level program which may potentially complicate the solution process. The investment decision problem of the Transco as well as the regulatory constraint are subject to the solution of the lower-level programs which are modeled based on formulation (6). It represents the spot market clearance by the ISO in the considered power system. The lower-level program is a set of ISO economic dispatch models (6) which are solved simultaneously for each investment planning period t, operation sub-period k and scenario s.

The proposed transmission investment model is a stochastic bilevel disjunctive program. In the following section we present steps which will lead to an equivalent single-level stochastic MILP formulation with disjunctive constraints.

2.4 The linearization of disjunctive constraints (6d)

Disjunctive constraints can be linearized using a set of disjunctive parameters. This type of reformulation (also known as big-M technique) was well studied in (Lee and Grossmann 2000) and (Trespalacios and Grossmann 2015). The choice of disjunctive parameters is critical for linear reformulation of disjunctive constraints. The parameters should be chosen big enough that the original feasibility set does not change and not too big that the reformulated constraints are as tight as possible. If the disjunctive parameter is chosen carefully then the reformulated problem will be equivalent to the original one. Using this technique, the disjunctive constraints (6d) in the lower-level program of the Transco's problem (7) can be reformulated as linear constraints in (8).

$$
\widehat{f}_{mtks} - \frac{100}{X_m}\Big(\sum_{n\in\mathcal{N}} \overline{S}_{nm}^{(n)}\theta_{ntks} - \sum_{n\in\mathcal{N}} \overline{R}_{nm}^{(n)}\theta_{ntks}\Big) \le
$$

$$
\Xi^{(m)}(1-z_{mt}) \ :(\overline{\sigma}_{mtks}) \ \forall m\in\mathcal{M}, t\in\mathcal{T}, k\in\mathcal{K}, s\in\mathcal{S} \tag{8a}
$$

$$
\widehat{f}_{mtks} - \frac{100}{X_m}\Big(\sum_{n\in\mathcal{N}} \overline{S}_{nm}^{(n)}\theta_{ntks} - \sum_{n\in\mathcal{N}} \overline{R}_{nm}^{(n)}\theta_{ntks}\Big) \ge
$$

$$
- \ \Xi^{(m)}(1-z_{mt}) \ :(\underline{\sigma}_{mtks}) \ \forall m\in\mathcal{M}, t\in\mathcal{T}, k\in\mathcal{K}, s\in\mathcal{S} \tag{8b}
$$

$$
- z_{mt}\Xi^{(m)} \le \widehat{f}_{mtks} \le z_{mt}\Xi^{(m)} : (\underline{\zeta}_{mtks}, \overline{\zeta}_{mtks}) \ \ \forall m\in\mathcal{M}, t\in\mathcal{T}, k\in\mathcal{K}, s\in\mathcal{S} \tag{8c}
$$

2.5 Reformulation using LP duality theorem

The optimization problem (7) after performing the reformulation described in subsection 2.4 becomes a bilevel mixed integer program. The lower-level problem is a linear program, for which LP duality theorem can be applied (Bertsekas 1999). Thus, the lower-level program can be equivalently described by its primal and dual lower level variables which satisfy primal and dual feasibility constraints and strong duality condition. The dual lower level feasibility constraints and strong duality conditions are derived in (9) and (10), respectively.

2.6 Nomenclature

Lagrange multiplier variables of (7)

λ_{ntks} Price at bus n, investment period t,
sub-period k and scenario s, ($\$/MWh$);
sub-period k and scenario s, ($\$/p.u$);

$\overline{\mu}_{ltks}, \underline{\mu}_{ltks}$ Lagrange multipliers for line l upper and lower limit
constraints (6f) at investment period t,
sub-period k and scenario s, ($\$/MWh$);

σ_{ltks} Lagrange multipliers for power flow constraints of line l
constraints (6c) at investment period t,
sub-period k and scenario s, ($\$/MWh$);

$\overline{\gamma}_{mtks}, \underline{\gamma}_{mtks}$	Lagrange multipliers for line m upper and lower limit constraints (6e) at investment period t, sub-period k and scenario s, ($\$/MWh$)
$\overline{\sigma}_{mtks}, \underline{\sigma}_{mtks}$	Lagrange multipliers for line m disjunctive relaxation constraints (8a)-(8b) at investment period t, sub-period k and scenario s, ($\$/MWh$)
$\overline{\zeta}_{mtks}, \underline{\zeta}_{mtks}$	Lagrange multipliers for line m disjunctive relaxation constraints (8c) at investment period t, sub-period k and scenario s, ($\$/MWh$);

2.6.1 Dual constraints

Dual constraints of the linear program (6) are derived as in (9).

$$\alpha_i - \sum_{n \in \mathcal{N}} I_{ni}^{(n)} \lambda_{ntks} + \underline{\omega}_{itks} - \overline{\omega}_{itks} = 0 \quad \forall i \in \mathcal{D}, t \in \mathcal{T}, k \in \mathcal{K}, s \in \mathcal{S} \tag{9a}$$

$$-c_j + \sum_{n \in \mathcal{N}} J_{nj}^{(n)} \lambda_{ntks} + \underline{\nu}_{jtks} - \overline{\nu}_{jtks} = 0 \quad \forall j \in \mathcal{G}, t \in \mathcal{T}, k \in \mathcal{K}, s \in \mathcal{S} \tag{9b}$$

$$\sum_{n \in \mathcal{N}} W_{nw}^{(n)} \lambda_{ntks} + \underline{\eta}_{wtks} - \overline{\eta}_{wtks} = 0 \quad \forall w \in \mathcal{W}, t \in \mathcal{T}, k \in \mathcal{K}, s \in \mathcal{S} \tag{9c}$$

$$-\sum_{n \in \mathcal{N}} S_{nl}^{(n)} \lambda_{ntks} + \sum_{n \in \mathcal{N}} R_{nl}^{(n)} \lambda_{ntks} + \sigma_{ltks} + \underline{\mu}_{ltks} - \overline{\mu}_{ltks} = 0 \quad \forall l \in \mathcal{L}, t \in \mathcal{T}, k \in \mathcal{K}, s \in \mathcal{S} \tag{9d}$$

$$-\sum_{n \in \mathcal{N}} \overline{S}_{nm}^{(n)} \lambda_{ntks} + \sum_{n \in \mathcal{N}} \overline{R}_{nm}^{(n)} \lambda_{ntks} + \underline{\sigma}_{mtks} - \overline{\sigma}_{mtks} + \underline{\gamma}_{mtks} - \overline{\gamma}_{mtks} +$$

$$\underline{\zeta}_{mtks} - \overline{\zeta}_{mtks} = 0 \quad \forall m \in \mathcal{M}, t \in \mathcal{T}, k \in \mathcal{K}, s \in \mathcal{S} \tag{9e}$$

$$-\frac{100}{X_l} \sum_{l \in \mathcal{L}} S_{nl}^{(n)} \sigma_{ltks} + \frac{100}{X_l} \sum_{l} R_{nl}^{(n)} \sigma_{ltks} + \xi_{tks} if(n = n1) -$$

$$\frac{100}{X_m} \sum_{m \in \mathcal{M}} \overline{S}_{nm}^{(n)} \underline{\sigma}_{mtks} + \frac{100}{X_m} \sum_{m \in \mathcal{M}} \overline{R}_{nm}^{(n)} \underline{\sigma}_{mtks} + \frac{100}{X_m} \sum_{m \in \mathcal{M}} \overline{S}_{nm}^{(n)} \overline{\sigma}_{mtks} -$$

$$\frac{100}{X_m} \sum_{m \in \mathcal{M}} \overline{R}_{nm}^{(n)} \overline{\sigma}_{mtks} = 0 \quad \forall n \in \mathcal{N}, t \in \mathcal{T}, k \in \mathcal{K}, s \in \mathcal{S} \tag{9f}$$

2.6.2 Strong duality condition

The strong duality condition of linear program (7) is formulated as in (10).

$$\sum_{i \in \mathcal{D}} \alpha_i d_{itks} - \sum_{j \in \mathcal{G}} c_j \, g_{jtks} = \sum_{t \in \mathcal{T}} \langle \sum_{i \in \mathcal{D}} D_{it} \overline{\omega}_{itks} + \sum_{j \in \mathcal{G}} G_j \overline{\nu}_{jtks} +$$

$$\sum_{l \in \mathcal{L}} F_l (\underline{\mu}_{ltks} + \overline{\mu}_{ltks}) + \sum_{m \in \mathcal{M}} \widehat{F}_m (\underline{\gamma}_{mtks} + \overline{\gamma}_{mtks}) +$$

$$\sum_{m \in \mathcal{M}} \underbrace{z_{mt} \Xi^{(m)} (\underline{\zeta}_{mtks} + \overline{\zeta}_{mtks})}_{T1} +$$

$$\sum_{m\in\mathcal{M}} \underbrace{\Xi^{(m)}(1-z_{mt})(\overline{\sigma}_{mtks}+\underline{\sigma}_{mtks})}_{T2} \rangle \; \forall t\in\mathcal{T},k\in\mathcal{K},s\in\mathcal{S} \tag{10}$$

Lemma 1 shows that terms T1 and T2 are zero at optimal solution.

Lemma 1 *If the disjunctive parameter $\Xi^{(m)}$ is tuned properly and optimization problem (6) is solved to optimality, then terms T1 and T2 in (10) are always equal to zero. Thus, we can reformulate the strong duality constraint as a combination of the strong duality constraint without terms T1 and T2 and the enforcing constraints T1=0 and T2=0, separately.*

Proof For term T1, if $z_{mt}=0$ then $T1=0$. For the case where $z_{mt}=1$, since the disjunctive parameter $\Xi^{(m)}$ is tuned properly and we solve the problem to optimality, then the complimentary conditions are satisfied and constraints (8c) are not active. This, means $(\overline{\zeta}_{mtks}+\underline{\zeta}_{mtks})=0$ or $T1=0$. By analogy we can show that T2 is also always equal to zero when the KarushKuhnTucker (KKT) conditions are satisfied: i.e an optimal solution is reached.

By ensuring constraints T1=0 and T2=0, we can drop T1 and T2 from (10) and rewrite the strong duality condition as:

$$\sum_{i\in\mathcal{D}} \alpha_i d_{itks} - \sum_{j\in\mathcal{G}} c_j\; g_{jtks} =$$

$$\sum_{t\in\mathcal{T}} \langle \sum_{i\in\mathcal{D}} D_{it}\overline{\omega}_{itks} + \sum_{j\in\mathcal{G}} G_j\overline{\nu}_{jtks} + \sum_{l\in\mathcal{L}} F_l(\underline{\mu}_{ltks}+\overline{\mu}_{ltks})+$$

$$\sum_{m\in\mathcal{M}} \widehat{F}_m(\underline{\gamma}_{mtks}+\overline{\gamma}_{mtks}) \; \forall t\in\mathcal{T},k\in\mathcal{K},s\in\mathcal{S} \tag{11a}$$

$$\sum_{m\in\mathcal{M}} z_{mt}\Xi^{(m)}(\underline{\zeta}_{mtks}+\overline{\zeta}_{mtks}) \; \forall m\in\mathcal{M},t\in\mathcal{T},k\in\mathcal{K},s\in\mathcal{S} \tag{11b}$$

$$\sum_{m\in\mathcal{M}} \Xi^{(m)}(1-z_{mt})(\overline{\sigma}_{mtks}+\underline{\sigma}_{mtks}) \; \forall m\in\mathcal{M},t\in\mathcal{T},k\in\mathcal{K},s\in\mathcal{S} \tag{11c}$$

The terms T1 and T2 in (11) are complicated terms due to their non-linear structure. Those terms can be linearized as in (12).

$$-\Xi^{(T)}(1-z_{mt}) \le \overline{\sigma}_{mtks}+\underline{\sigma}_{mtks} \le \Xi^{(T)}(1-z_{mt}) \;\; \forall m\in\mathcal{M},t\in\mathcal{T},k\in\mathcal{K},s\in\mathcal{S} \tag{12a}$$

$$-\Xi^{(T)}z_{mt} \le \overline{\zeta}_{mtks}+\underline{\zeta}_{mtks} \le \Xi^{(T)}z_{mt} \;\; \forall m\in\mathcal{M},t\in\mathcal{T},k\in\mathcal{K},s\in\mathcal{S} \tag{12b}$$

2.6.3 The problem of tuning the disjunctive parameter

The disjunctive parameters $\Xi^{(m)}$ and $\Xi^{(T)}$ used in constraints (8) and (12) typically create some problems in the solution process (Hooker 2011) and (Trespalacios and Grossmann 2015).

First of all, a too large or a too small disjunctive parameter creates numerical errors and rounding errors for the optimizer. Thus, referring to constraints, the parameter $\Xi^{(m)}$ should be chosen small enough not to create numerical complications but big enough to guarantee that the constraints (8) are not binding when $z_{mt}=0$ and the original feasibility set does not change. Such disjunctive parameter can be found by analyzing the power flow limits of the candidate lines and voltage

angle limits. However, when it comes to the second disjunctive parameter, $\Xi^{(T)}$, used in (12) the same methodology cannot be applied. Constraints in (12) contain Lagrange multipliers which do not have natural upper bounds (in case of non-negative Lagrange multipliers). Thus, it will be hard to guarantee the optimality of the tuned disjunctive parameter. Most of the existing methodologies designed to tune the disjunctive parameter focus on models with constraints similar to (8), which have a natural upper bound and thus can be tuned effectively. However, no methodology can guarantee the optimal choice of the disjunctive parameter for constraints without natural bounds as in (12).

Fig. 3 The impact of disjunctive parameter tuning on the relaxed feasible region. The region corresponds to the area inside the dashed or solid lines.
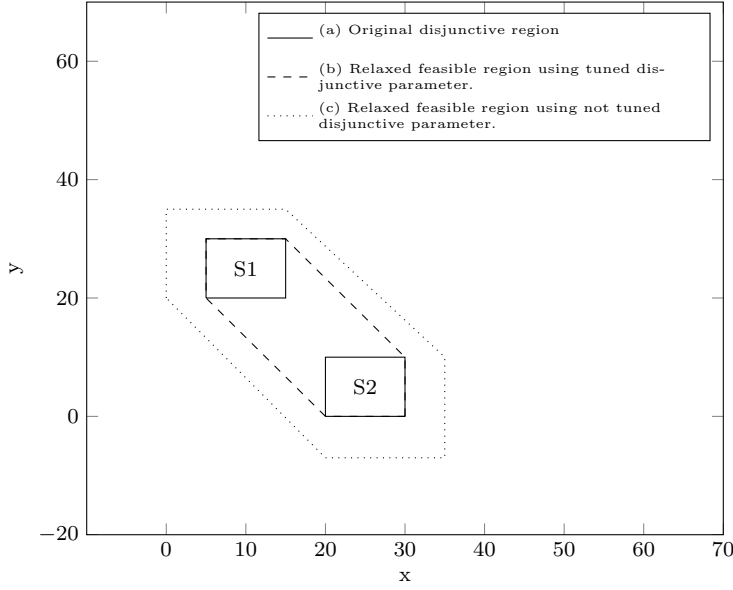
The impact of disjunctive parameter tuning on the relaxed feasible region is conceptually illustrated in Figure 3. The $S1$ and $S2$ are original feasible regions with disjunctive property (either $S1$ or $S2$ is the feasible region). Relaxed feasible regions using optimally tuned disjunctive parameters and the case when disjunctive parameters are not optimally tuned are demonstrated as region (b) and region (c), respectively. As we can see region (b) is a tighter relaxation.

2.7 Linearization

The bilinear terms $\lambda_{ntks}d_{itks}$, $\lambda_{ntks}g_{jtks}$ and $\lambda_{ntks}\widehat{g}_{wtks}$ in the Transco profit function and in the H-R-G-V regulatory constraint can be linearized as in (13) and (14).

$$\sum_{n\in\mathcal{N},i\in\mathcal{D}} I_{ni}^{(n)}\lambda_{ntks}d_{itks} - \sum_{n\in\mathcal{N},j\in\mathcal{G}} J_{nj}^{(n)}\lambda_{ntks}g_{jtks} - \sum_{n\in\mathcal{N},w\in\mathcal{W}} W_{nw}^{(n)}\lambda_{ntks}\widehat{g}_{wtks} =$$

$$\sum_{n\in\mathcal{N}}\lambda_{ntks}(\sum_{i\in\mathcal{D}}I_{ni}^{(n)}d_{itks}-\sum_{j\in\mathcal{G}}J_{nj}^{(n)}g_{jtks}-\sum_{w}W_{nw}^{(n)}\widehat{g}_{wtks})\stackrel{(6b)}{=}$$

$$\sum_{l\in\mathcal{L}}f_{ltks}(-\sum_{n\in\mathcal{N}}S_{nl}^{(n)}\lambda_{ntks}+\sum_{n\in\mathcal{N}}R_{nl}^{(n)}\lambda_{ntks})+\sum_{m\in\mathcal{M}}\widehat{f}_{mtks}(-\sum_{n\in\mathcal{N}}\overline{S}_{nm}^{(n)}\lambda_{ntks}+$$

$$\sum_{n}\overline{R}_{nm}^{(n)}\lambda_{ntks})\stackrel{(9d),(9e)}{=}\sum_{l\in\mathcal{L}}f_{ltks}(-\sigma_{ltks}-\underline{\mu}_{ltks}+\overline{\mu}_{ltks})+$$

$$\sum_{m\in\mathcal{M}}\widehat{f}_{mtks}(-\underline{\sigma}_{mtks}-+\overline{\sigma}_{mtks}-\underline{\gamma}_{mtks}+\overline{\gamma}_{mtks}-\underline{\zeta}_{mtks}+\overline{\zeta}_{mtks})\stackrel{(*)}{=}$$

$$\sum_{l\in\mathcal{L}}F_l(\overline{\mu}_{ltks}+\underline{\mu}_{ltks})+\sum_{m\in\mathcal{M}}\widehat{F}_m(\overline{\gamma}_{mtks}+\underline{\gamma}_{mtks})+T1+T2\stackrel{Lemma(1)}{=}$$

$$\sum_{l\in\mathcal{L}}F_l(\overline{\mu}_{ltks}+\underline{\mu}_{ltks})+\sum_{m\in\mathcal{M}}\widehat{F}_m(\overline{\gamma}_{mtks}+\underline{\gamma}_{mtks}) \quad \forall t\in\mathcal{T},k\in\mathcal{K},s\in\mathcal{S} \tag{13}$$

Where $(*)$ comes from complementary slackness conditions for equations (6f), (6e) and (8a),(8b),(8c).

For bilinear terms in H-R-G-V regulatory constraint we have:

$$\langle\sum_{i\in\mathcal{D}}\alpha_i d_{itks}-\sum_{n\in\mathcal{N},i\in\mathcal{D}}I_{ni}^{(n)}\lambda_{ntks}d_{itks}\rangle+\langle\sum_{n\in\mathcal{N},j\in\mathcal{G}}J_{nj}^{(n)}\lambda_{ntks}g_{jtks}-\sum_{j\in\mathcal{G}}c_j g_{jtks}\rangle+$$

$$\sum_{w}W_{nw}^{(n)}\lambda_{ntks}\widehat{g}_{wtks}\stackrel{(9a)-(9c)}{=}\sum_{i\in\mathcal{D}}d_{itks}(\underline{\omega}_{itks}-\overline{\omega}_{itks})+$$

$$\sum_{j\in\mathcal{G}}g_{jtks}(\underline{\nu}_{jtks}-\overline{\nu}_{jtks})+\sum_{w}\widehat{g}_{wtks}(\underline{\eta}_{wtks}-\overline{\eta}_{wtks})\stackrel{(**)}{=}$$

$$\sum_{i\in\mathcal{D}}D_{it}\overline{\omega}_{itks}+\sum_{j\in\mathcal{G}}G_j\overline{\nu}_{jtks}+\sum_{w}\widehat{G}_{wks}\overline{\eta}_{wtks} \quad \forall t\in\mathcal{T},k\in\mathcal{K},s\in\mathcal{S} \tag{14}$$

Where $(**)$ results from complementary slackness conditions for (6g),(6h) and (6i).

2.8 The reformulated stochastic MILP with linearized disjunctive constraints

The resulting model after performed linearizations and reformulations is presented in (15).

$$\underset{z_{mt},y_{mt},\Omega_p}{Maximize}:\sum_{t\in\mathcal{T}}\frac{(1+\beta)^t}{(1+\gamma)^t}\langle\sum_{s}\pi_s\psi(\sum_{l,k}F_l(\overline{\mu}_{ltks}+\underline{\mu}_{ltks})+\sum_{m,k}\widehat{F}_m(\overline{\mu}_{ltks}+\underline{\mu}_{ltks})+$$

$$\Phi_t-\sum_{m\in\mathcal{M}}\overline{C}_{m,t}y_{mt}\rangle \tag{15a}$$

$Subject\ to:$

$$(7b)-(7e) \tag{15b}$$

$$\Phi_t-\Psi\sum_{s}\pi_s(\sum_{i,k}D_{it}\overline{\omega}_{itks}+\sum_{j,k}G_j\overline{\nu}_{jtks}+\sum_{w,k}\widehat{G}_{wks}\overline{\eta}_{wtks})=$$

$$\Phi_{t-1} - \Psi \sum_s \pi_s \left(\sum_{i,k} D_{it-1} \, \overline{\omega}_{it-1ks} + \sum_{j,k} G_j \overline{\nu}_{jt-1ks} + \sum_{w,k} \widehat{G}_{wks} \overline{\eta}_{wt-1ks} \right) \quad \forall t \in \{\mathcal{T} \setminus t1\}$$

$$\text{(15c)}$$

$$(6b) - (6f), \ (8a) - (8c), \ (6e) - (6k), \ (9a) - (9f), \ (12a) - (12b) \qquad \text{(15d)}$$

$$\sum_{i \in \mathcal{D}} D_{it} \overline{\omega}_{itks} + \sum_{j \in \mathcal{G}} G_j \overline{\nu}_{jtks} + \sum_{l \in \mathcal{L}} F_l (\overline{\mu}_{ltks} + \underline{\mu}_{ltks}) +$$

$$\sum_m \widehat{f}_{mtks} (\overline{\gamma}_{mtks} + \underline{\gamma}_{mtks}) + \sum_w \widehat{G}_{wks} \overline{\eta}_{wtks} =$$

$$\sum_{i \in \mathcal{D}} \alpha_i d_{itks} - \sum_{j \in \mathcal{G}} c_j g_{jtks} \quad \forall t \in \mathcal{T}, k \in \mathcal{K}, s \in \mathcal{S} \qquad \text{(15e)}$$

$$\overline{\omega}_{itks}, \underline{\omega}_{itks}, \overline{\nu}_{jtks}, \underline{\nu}_{jtks}, \overline{\eta}_{wtks}, \underline{\eta}_{wtks}, \overline{\mu}_{ltks}, \underline{\mu}_{ltks}, \sigma_{ltks}, \Phi_t, \overline{\gamma}_{mtks}, \underline{\gamma}_{mtks}, \overline{\sigma}_{mtks},$$

$$\underline{\sigma}_{mtks}, \overline{\zeta}_{mtks}, \underline{\zeta}_{mtks} \geq 0 \qquad \text{(15f)}$$

$$z_{mt}, y_{mt} \in \{0, 1\} \qquad \text{(15g)}$$

Here $\Omega_p = \Omega_s \cup \{\Phi_t, f_{ltks}, \widehat{f}_{mtks}, \theta_{ntks}, d_{itks}, g_{jtks}, \widehat{g}_{wtks}, \overline{\omega}_{itks}, \underline{\omega}_{itks}, \overline{\nu}_{jtks}, \underline{\nu}_{jtks}, \overline{\eta}_{wtks}, \underline{\eta}_{wtks}, \overline{\mu}_{ltks}, \underline{\mu}_{ltks}, \sigma_{ltks}, \overline{\gamma}_{mtks}, \underline{\gamma}_{mtks}, \overline{\sigma}_{mtks}, \underline{\sigma}_{mtks}, \overline{\zeta}_{mtks}, \underline{\zeta}_{mtks}\}$ is the set of decision variables of optimization problem (15).

3 The proposed decomposition algorithm

The stochastic MILP model (15) has a special decomposing structure. Such structure allows us to decompose the problem into a number of independent optimization problems by separating the variables into two vectors. The first vector consists of continuous variables and the second one consist of integer variables. One of the decomposition algorithms for such types of problems is the Benders decomposition algorithm (Lumbreras and Ramos 2013). However, the Benders decomposition algorithm might be ineffective for disjunctive programs, especially when variables in the disjunctive constraint do not have natural upper bounds and the disjunctive parameter cannot be tuned optimally. When the disjunctive parameter is not optimal, the Benders cuts are proved to be weak and the convergence of the Benders decomposition algorithm cannot be guaranteed (Hooker 2011). This is the case for our proposed stochastic MILP model (15). Thus, we propose a specialized decomposition algorithm based on the Bean decomposition algorithm proposed in (Bean et al 1992). In (Bean et al 1992), the authors propose the cuts which are identical to Benders cuts, however, they allow one to eliminate the disjunctive parameter by exploiting the properties of the extreme points. Moreover Bean's cuts are especially applicable for the problems with similar decomposable properties as the proposed problem (15): i.e. for problems that can be decomposed over integer variables which appear in bilinear terms and were linearized using the big-M reformulation. Also, the master problem formulated using Bean's cuts has similar properties as sets partitioning problems and it further improves the tractability of the solution.

The Bean decomposition algorithm is an iterative solution algorithm and has two separate steps at each iteration. First, duality theory is used to determine upper bounds through fixing complicating integer variables (assuming a minimization

program). The second step is to find a lower bound by solving the relaxed problem. The iteration between upper- and lower-bound programs is performed until the upper and lower bounds are close enough and the optimal solution is found. The special structure of the Bean decomposition algorithm allows one to create a tighter lower bound and a computationally more tractable master problem.

3.1 Nomenclature

To present the Bean decomposition algorithm the followings are defined.

Indices and Sets

$a, a' \in \mathcal{A}$	Bean decomposition iterations;
$\Omega_{sa}^{(1)}, \Omega_{sa}^{(2)}$	Sets of extreme points;

Binary variables

u_a, u_0	Auxiliary variables;

Parameters

K_a	Optimal objective value of subproblem;
L_a	Upper-bound constant for a step size;
Υ_a	A penalty factor;
H	Suitably large constants;
$\widehat{z}_{mt}^a, \widehat{y}_{mt}^a$	Fixed investment decisions of line m at period t;

Variables of the subproblem

$\widehat{\lambda}_{itks}^{(9a)}$	Lagrange multipliers for dual constraints(9a),(MWh);
$\widehat{\lambda}_{jtks}^{(9b)}$	Lagrange multipliers for dual constraints(9b),(MWh);
$\widehat{\lambda}_{wtks}^{(9c)}$	Lagrange multipliers for dual constraints (9c), (MWh);
$\widehat{\lambda}_{ltks}^{(9d)}$	Lagrange multipliers for dual constraints (9d), (MWh);
$\widehat{\lambda}_{mtks}^{(9e)}$	Lagrange multipliers for dual constraints (9e), (MWh);
$\widehat{\lambda}_{ntks}^{(9f)}$	Lagrange multipliers for dual constraints (9f), ($p.u$);
$\widehat{\lambda}_{tks}^{(15e)}$	Lagrange multipliers for strong duality conditions (15e);
$\widehat{\lambda}_{t}^{(15c)}$	Lagrange multipliers for regulatory constraint (15c);
$\widehat{\lambda}_{ntks}$	Lagrange multipliers for energy balance constraints (6b) at bus n in period t,k for scenario s, ($\$/MWh$);
$\overline{\mu}_{ltks}^{(6f)}, \underline{\mu}_{ltks}^{(6f)}$	Lagrange multipliers for line l upper and lower limit constraints (6f), ($\$/MWh$);
$\sigma_{ltks}^{(6c)}$	Lagrange multipliers for power flow constraints of line l constraints (6c), ($\$/MWh$);
$\overline{\gamma}_{mtks}^{(6e)}, \underline{\gamma}_{mtks}^{(6e)}$	Lagrange multipliers for line m upper and lower limit constraints (6e), ($\$/MWh$)
$\overline{\sigma}_{mtks}^{(8a)}, \underline{\sigma}_{mtks}^{(8b)}$	Lagrange multipliers for line m disjunctive relaxation constraints (8a)-(8b), ($\$/MWh$)

$\overline{\zeta}_{mtks}^{(8c)}, \underline{\zeta}_{mtks}^{(8c)}$	Lagrange multipliers for line m disjunctive relaxation constraints (8c), ($/MWh$);
$\underline{\vartheta}_{mtks}^{(12a)}, \overline{\vartheta}_{mtks}^{(12a)}$	Lagrange multipliers for T1 and T2 linearization constraints (12a), (MWh);
$\underline{\kappa}_{mtks}^{(12b)}, \overline{\kappa}_{mtks}^{(12b)}$	Lagrange multipliers for T1 and T2 linearization constraints (12b), (MWh);

3.2 The Bean decomposition algorithm

The Bean decomposition algorithm includes a master problem and a sub-problem and can be applied to two-stage stochastic programs with disjunctive constraints. The stochastic master problem is formulated as in (16) while a stochastic sub-problem is formulated as in (17). The sets $\Omega_{sa}^{(1)}$ and $\Omega_{sa}^{(2)}$ represent index sets of extreme points and correspond to Lagrange multipliers of disjunctive constraints (8) and (12). The set $\Omega_{sa}^{(1)}$ corresponds to extreme points of constraints (8a),(8b) and (12a) while the set $\Omega_{sa}^{(2)}$ corresponds to extreme points of constraints (8c) and (12b). The Bean decomposition algorithm iteratively solves master problem and sub-problem. At each iteration the master problem is updated with additional feasibility cut. Iterations are indexed using $a \in \mathcal{A}$. The set $\mathcal{A}$ becomes larger after each iteration if the algorithm does not satisfy the stopping criterion.

The master problem of the standard Bean decomposition algorithm is formulated as in (16).

$$\underset{u_a, z_{mt}, y_{mt}}{Maximize} \sum_{a \in \mathcal{A}} K_a u_a + H u_0 \tag{16a}$$

$Subject\ to$:

$$\sum_{mts \in \Omega_{sa}^{(1)}} z_{mt} + \sum_{mts \in \Omega_{sa}^{(2)}} (1 - z_{mt}) \le |\Omega_{sa}^{(1)}| + |\Omega_{sa}^{(2)}| - 1 + \sum_{a' if (K_{a'} \ge K_a)} u_{a'} \ \forall a \in \mathcal{A} \tag{16b}$$

$$\sum_a u_a = 1 \tag{16c}$$

$$z_{mt} = \sum_{\widehat{t} \in \{t1,..t\}} y_{m\widehat{t}} \ \forall m \in \mathcal{M}, t \in \mathcal{T} \tag{16d}$$

$$\sum_{t \in \mathcal{T}} y_{mt} \le 1 \ \ \forall m \in \mathcal{M} \tag{16e}$$

$$z_{mt}, y_{mt} \in \{0, 1\} \tag{16f}$$

The master problem (16) is a stochastic mixed-integer linear program with sets partitioning characteristics. All the constraints of the master problem contain only binary variables. Variables u_a are introduced for each iteration of the decomposition algorithm and are used to activate the decomposition cuts corresponding to the solution of the sub-problem. Variables z_{mt} and y_{mt} are the transmission investment variables. Variable u_0 is an auxiliary variable used to prevent the formulation to be unbounded from above. The objective function of the master problem consists of the objective function values of the sub-problems K_a at each iteration

multiplied by the corresponding auxiliary variables u_a. The objective function also includes auxiliary variable u_0 multiplied by a large enough parameter H which sets the upper boundary for the master problem and this prevents the master problem from being unbounded. Feasibility cuts based on Bean's cuts are introduced in (16b) and they are based on extreme point sets of the subproblems. The extreme point sets are used to determine whether to include integer variables z_{mt} or not. The cut is activated using auxiliary variables u_a. The cuts in (16b) are equivalent to Benders cuts with optimal disjunctive parameter but the cuts in (16b) do not contain any disjunctive parameter in the formulation (Bean 1992). An additional constraint is introduced in (16c) to ensure that only one cut is activated at each iteration. The rest of constraints (16d)-(16f) are introduced to ensure that the solution of the master problem satisfies line investment constraints (7b)-(7d). Once the master problem is solved, the solution of the investment decision variables z_{mt} and y_{mt} are used to formulate the sub-problem

In order to formulate the sub-problem of the standard Bean decomposition algorithm, we first fix line investment decision variables z_{mt} and y_{mt} in problem (15) and treat them as constants. The investment decision parameters used in the sub-problem are modeled through $\widehat{z}_{mt}^a$ and $\widehat{y}_{mt}^a$. Once the integer variables are fixed and treated as parameters the problem (15) becomes a linear program. The sub-problem is the dual of the optimization problem (15a),(15c)-(15g) where investment decisions are fixed to the values obtained from master problem. The sub-problem of the standard Bean decomposition algorithm is formulated as in (17).

$$
\underset{\Omega_p^d}{Minimize} : \sum_{t\in\mathcal{T}} \langle \sum_{i\in\mathcal{D}} D_{it}\overline{\omega}_{itks}^{(6i)} + \sum_{l\in\mathcal{L}} F_l(\overline{\mu}_{ltks}^{(6f)} + \underline{\mu}_{ltks}^{(6f)}) + \sum_{j\in\mathcal{G}} G_j\overline{\nu}_{jtks}^{(6g)} +
$$

$$
\sum_{m\in\mathcal{M}} \widehat{F}_m(\overline{\gamma}_{mtks}^{(6e)} + \underline{\gamma}_{mtks}^{(6e)}) + \sum_w \widehat{G}_{wks}\overline{\eta}_{wtks}^{(6h)} +
$$

$$
\alpha_i\widehat{\lambda}_{itks}^{(9a)} - c_j\widehat{\lambda}_{jtks}^{(9b)} + \sum_{m\in\mathcal{M}} \underbrace{\widehat{z}_{mt}^a \Xi^{(m)}(\overline{\sigma}_{mtks}^{(8a)} + \underline{\sigma}_{mtks}^{(8b)})}_{T5} +
$$

$$
\sum_{m\in\mathcal{M}} \underbrace{(1 - \widehat{z}_{mt}^a)\Xi^{(m)}(\overline{\zeta}_{mtks}^{(8c)} + \underline{\zeta}_{mtks}^{(8c)})}_{T6} + \sum_{m\in\mathcal{M}} \underbrace{\Xi^{(T)}\widehat{z}_{mt}^a(\overline{\vartheta}_{mtks}^{(12a)} + \underline{\vartheta}_{mtks}^{(12a)})}_{T7} +
$$

$$
\sum_{m\in\mathcal{M}} \underbrace{\Xi^{(T)}(1 - \widehat{z}_{mt}^a)(\overline{\kappa}_{mtks}^{(12b)} + \underline{\kappa}_{mtks}^{(12b)})}_{T8} \rangle \tag{17a}
$$

Subject to :

$$
\alpha_i\widehat{\lambda}_{tks}^{(15e)} - \sum_{n\in\mathcal{N}} I_{ni}^{(n)}\widehat{\lambda}_{ntks} + \underline{\omega}_{itks}^{(6i)} - \overline{\omega}_{itks}^{(6i)} = 0 \quad \forall i\in\mathcal{D},t\in\mathcal{T},k\in\mathcal{K},s\in\mathcal{S} \tag{17b}
$$

$$
-c_j\widehat{\lambda}_{tks}^{(15e)} + \sum_{n\in\mathcal{N}} J_{nj}^{(n)}\widehat{\lambda}_{ntks} + \underline{\nu}_{jtks}^{(6g)} - \overline{\nu}_{jtks}^{(6g)} = 0 \quad \forall j\in\mathcal{G},t\in\mathcal{T},k\in\mathcal{K},s\in\mathcal{S} \tag{17c}
$$

$$
\sum_{n\in\mathcal{N}} W_{nw}^{(n)}\widehat{\lambda}_{ntks} + \underline{\eta}_{wtks}^{(6h)} - \overline{\eta}_{wtks}^{(6h)} = 0 \quad \forall w\in\mathcal{W},t\in\mathcal{T},k\in\mathcal{K},s\in\mathcal{S} \tag{17d}
$$

$$
\sum_{n\in\mathcal{N}} R_{nl}^{(n)}\widehat{\lambda}_{ntks} - \sum_{n\in\mathcal{N}} S_{nl}^{(n)}\widehat{\lambda}_{ntks} + \sigma_{ltks}^{(6c)} + \underline{\mu}_{ltks}^{(6f)} - \overline{\mu}_{ltks}^{(6f)} = 0 \quad \forall l\in\mathcal{L},t\in\mathcal{T},k\in\mathcal{K},s\in\mathcal{S}
$$

$$
\tag{17e}
$$

$$\sum_{n \in \mathcal{N}} \overline{R}_{nm}^{(n)} \widehat{\lambda}_{ntks} - \sum_{n \in \mathcal{N}} \overline{S}_{nm}^{(n)} \widehat{\lambda}_{ntks} + \underline{\gamma}_{mtks}^{(6e)} - \overline{\gamma}_{mtks}^{(6e)} + \underline{\sigma}_{mtks} - \overline{\sigma}_{mtks}^{(8a)} + \underline{\zeta}_{mtks}^{(8c)} -$$

$$\overline{\zeta}_{mtks}^{(8c)} = 0 \forall m \in \mathcal{M}, t \in \mathcal{T}, k \in \mathcal{K}, s \in \mathcal{S} \tag{17f}$$

$$- \frac{100}{X_l} \sum_{l \in \mathcal{L}} S_{nl}^{(n)} \sigma_{ltks}^{(6c)} + \frac{100}{X_l} \sum_{l \in \mathcal{L}} R_{nl}^{(n)} \sigma_{ltks}^{(6c)} + \xi_{ntks}^{(6j)} if(n = n1) -$$

$$\frac{100}{X_m} \sum_{m \in \mathcal{M}} \overline{S}_{nm}^{(n)} \underline{\sigma}_{mtks}^{(8b)} + \frac{100}{X_m} \sum_{m \in \mathcal{M}} \overline{R}_{nm}^{(n)} \underline{\sigma}_{mtks}^{(8b)} + \frac{100}{X_m} \sum_{m \in \mathcal{M}} \overline{S}_{nm}^{(n)} \overline{\sigma}_{mtks}^{(8a)}$$

$$- \frac{100}{X_m} \sum_{m \in \mathcal{M}} \overline{R}_{nm}^{(n)} \overline{\sigma}_{mtks}^{(8a)} = 0 \quad \forall n \in \mathcal{N}, t \in \mathcal{T}, k \in \mathcal{K}, s \in \mathcal{S} \tag{17g}$$

$$\frac{(1+\beta)^t}{(1+\gamma)^t} + \widehat{\lambda}_t^{(15c)} - \widehat{\lambda}_{t+1}^{(15d)} = 0 \forall t \in \mathcal{T} \tag{17h}$$

$$- \sum_{i \in \mathcal{D}} I_{ni}^{(n)} \widehat{\lambda}_{itks}^{(9a)} + \sum_{j \in \mathcal{G}} J_{nj}^{(n)} \widehat{\lambda}_{jtks}^{(9b)} + \sum_w W_{nw}^{(n)} \widehat{\lambda}_{wtks}^{(9c)} - \sum_{l \in \mathcal{L}} S_{nl}^{(n)} \widehat{\lambda}_{ltks}^{(9d)} +$$

$$\sum_{l \in \mathcal{L}} R_{nl}^{(n)} \widehat{\lambda}_{ltks}^{(9d)} - \sum_{m \in \mathcal{M}} \overline{R}_{nm}^{(n)} \widehat{\lambda}_{mtks}^{(9e)} + \sum_{m \in \mathcal{M}} \overline{R}_{nm}^{(n)} \widehat{\lambda}_{mtks}^{(9e)} = 0 \forall n \in \mathcal{N}, t \in \mathcal{T}, k \in \mathcal{K}, s \in \mathcal{S} \tag{17i}$$

$$\widehat{\lambda}_{itks}^{(9a)} \leq 0 \quad \forall i \in \mathcal{D}, t \in \mathcal{T}, k \in \mathcal{K}, s \in \mathcal{S} \tag{17j}$$

$$- D_{it} \widehat{\lambda}_{tks}^{(15e)} - \Psi \pi_s (D_{it} \widehat{\lambda}_t^{(15c)} + D_{it-1} \widehat{\lambda}_{t+1}^{(15d)}) - \widehat{\lambda}_{itks}^{(9a)} \leq 0 \quad \forall i \in \mathcal{D}, t \in \mathcal{T}, k \in \mathcal{K}, s \in \mathcal{S} \tag{17k}$$

$$\widehat{\lambda}_{jtks}^{(9b)} \leq 0 \forall j \in \mathcal{G}, t \in \mathcal{T}, k \in \mathcal{K}, s \in \mathcal{S} \tag{17l}$$

$$- G_j \widehat{\lambda}_{tks}^{(15e)} - \Psi \pi_s (G_j \widehat{\lambda}_t^{(15c)} + G_j \widehat{\lambda}_{t+1}^{(15d)}) - \widehat{\lambda}_{jtks}^{(9b)} \leq 0 \quad \forall j \in \mathcal{G}, t \in \mathcal{T}, k \in \mathcal{K}, s \in \mathcal{S} \tag{17m}$$

$$\widehat{\lambda}_{wtks}^{(9c)} \leq 0 \quad \forall w \in \mathcal{W}, t, k, s \tag{17n}$$

$$- \widehat{G}_{wks} \widehat{\lambda}_{tks}^{(15e)} - \Psi \pi_s (\widehat{G}_{wks} \widehat{\lambda}_t^{(15c)} + \widehat{G}_{wks} \widehat{\lambda}_{t+1}^{(15d)}) - \widehat{\lambda}_{wtks}^{(9c)} \leq 0 \quad \forall w \in \mathcal{W}, t \in \mathcal{T}, k \in \mathcal{K}, s \in \mathcal{S} \tag{17o}$$

$$\widehat{\lambda}_{ltks}^{(9d)} - \frac{100}{X_l} \sum_{n \in \mathcal{N}} S_{nl}^{(n)} \widehat{\lambda}_{ntks}^{(9f)} + \frac{100}{X_l} \sum_{n \in \mathcal{N}} R_{nl}^{(n)} \widehat{\lambda}_{ntks}^{(9f)} = 0 \quad \forall l \in \mathcal{L}, t \in \mathcal{T}, k \in \mathcal{K}, s \in \mathcal{S} \tag{17p}$$

$$\frac{(1+\beta)^t}{(1+\gamma)^t} \pi_s \Psi F_l - F_l \widehat{\lambda}_{tks}^{(15e)} + \widehat{\lambda}_{ltks}^{(9d)} \leq 0 \quad \forall l \in \mathcal{L}, t \in \mathcal{T}, k \in \mathcal{K}, s \in \mathcal{S} \tag{17q}$$

$$\frac{(1+\beta)^t}{(1+\gamma)^t} \pi_s \Psi F_l - F_l \widehat{\lambda}_{tks}^{(15e)} - \widehat{\lambda}_{ltks}^{(9d)} \leq 0 \quad \forall l \in \mathcal{L}, t \in \mathcal{T}, k \in \mathcal{K}, s \in \mathcal{S} \tag{17r}$$

$$\frac{(1+\beta)^t}{(1+\gamma)^t} \pi_s \Psi \widehat{F}_m - \widehat{F}_m \widehat{\lambda}_{tks}^{(15e)} + \widehat{\lambda}_{mtks}^{(9e)} \leq 0 \quad \forall m \in \mathcal{M}, t \in \mathcal{T}, k \in \mathcal{K}, s \in \mathcal{S} \tag{17s}$$

$$\frac{(1+\beta)^t}{(1+\gamma)^t} \pi_s \Psi \widehat{F}_m - \widehat{F}_m \widehat{\lambda}_{tks}^{(15e)} - \widehat{\lambda}_{mtks}^{(9e)} \leq 0 \quad \forall m \in \mathcal{M}, t \in \mathcal{T}, k \in \mathcal{K}, s \in \mathcal{S} \tag{17t}$$

$$\widehat{\lambda}_{mtks}^{(9e)} + \underline{\vartheta}_{mtks}^{(12a)} - \overline{\vartheta}_{mtks}^{(12a)} - \frac{100}{X_m} \sum_{n \in \mathcal{N}} \overline{S}_{nm}^{(n)} \widehat{\lambda}_{ntks}^{(9f)} +$$

$$\frac{100}{X_m} \sum_{n \in \mathcal{N}} \overline{R}_{nm}^{(n)} \widehat{\lambda}_{ntks}^{(9f)} \leq 0 \quad \forall m \in \mathcal{M}, t \in \mathcal{T}, k \in \mathcal{K}, s \in \mathcal{S} \tag{17u}$$

$$- \widehat{\lambda}_{mtks}^{(9e)} + \underline{\vartheta}_{mtks}^{(12a)} - \overline{\vartheta}_{mtks}^{(12a)} + \frac{100}{X_m} \sum_{n \in \mathcal{N}} \overline{S}_{nm}^{(n)} \widehat{\lambda}_{ntks}^{(9f)} -$$

$$\frac{100}{X_m} \sum_{n \in \mathcal{N}} \overline{R}_{nm}^{(n)} \widehat{\lambda}_{ntks}^{(9f)} \leq 0 \quad \forall m \in \mathcal{M}, t \in \mathcal{T}, k \in \mathcal{K}, s \in \mathcal{S} \tag{17v}$$

$$\widehat{\lambda}_{mtks}^{(9e)} + \underline{\kappa}_{mtks}^{(12b)} - \overline{\kappa}_{mtks}^{(12b)} \leq 0 \quad \forall m \in \mathcal{M}, t \in \mathcal{T}, k \in \mathcal{K}, s \in \mathcal{S} \tag{17w}$$

$$- \widehat{\lambda}_{mtks}^{(9e)} + \underline{\kappa}_{mtks}^{(12b)} - \overline{\kappa}_{mtks}^{(12b)} \leq 0 \quad \forall m \in \mathcal{M}, t \in \mathcal{T}, k \in \mathcal{K}, s \in \mathcal{S} \tag{17x}$$

$$\widehat{\lambda}_{ntks}^{(9f)} = 0 \quad \forall n \in \mathcal{N}, t \in \mathcal{T}, k \in \mathcal{K}, s \in \mathcal{S} \tag{17y}$$

$$\overline{\mu}_{ltks}^{(6f)}, \underline{\mu}_{ltks}^{(6f)}, \overline{\gamma}_{mtks}^{(6e)}, \underline{\gamma}_{mtks}^{(6e)}, \overline{\sigma}_{mtks}^{(8a)}, \underline{\sigma}_{mtks}^{(8b)}, \overline{\zeta}_{mtks}^{(8c)}, \underline{\zeta}_{mtks}^{(8c)}, \underline{\vartheta}_{mtks}^{(12a)}, \overline{\vartheta}_{mtks}^{(12a)},$$

$$\underline{\kappa}_{mtks}^{(12b)}, \overline{\kappa}_{mtks}^{(12b)} \geq 0 \tag{17z}$$

Where $\Omega_p^d = \{\widehat{\lambda}_{itks}^{(9a)}, \widehat{\lambda}_{jtks}^{(9b)}, \widehat{\lambda}_{wtks}^{(9c)}, \widehat{\lambda}_{ltks}^{(9d)}, \widehat{\lambda}_{mtks}^{(9e)}, \widehat{\lambda}_{ntks}^{(9f)}, \widehat{\lambda}_{tks}^{(15e)}, \widehat{\lambda}_{t}^{(15c)}, \widehat{\lambda}_{ntks}, \overline{\mu}_{ltks}^{(6f)},$ $\underline{\mu}_{ltks}^{(6f)}, \overline{\gamma}_{mtks}^{(6e)}, \underline{\gamma}_{mtks}^{(6e)}, \overline{\sigma}_{mtks}^{(8a)}, \underline{\sigma}_{mtks}^{(8b)}, \overline{\zeta}_{mtks}^{(8c)}, \underline{\zeta}_{mtks}^{(8c)}, \underline{\vartheta}_{mtks}^{(12a)}, \overline{\vartheta}_{mtks}^{(12a)}, \underline{\kappa}_{mtks}^{(12b)}, \overline{\kappa}_{mtks}^{(12b)}\}$ is the set of decision variables of the problem (17).

4 Modification and acceleration

The sub-problem (17) contains the disjunctive parameters. These disjunctive parameters are the terms $T5$ to $T8$ of the sub-problem. However, using Lemma 2 one can show that these terms will be equal to zero at the optimal solution.

Lemma 2 *If problem (17) is solved to optimality and disjunctive parameters are chosen optimally, then terms T5 to T8 in (17a) are always equal to zero. Thus, we can reformulate the objective function as a combination of the objective function without terms T5 to T8 as in (18a) and with additional constraints (18c)-(18f) which ensure that T5 to T8 are equal to zero*

Proof The proof is by analogy to Lemma 1. Terms T5 to T8 in dual problem (17) correspond to complementary slackness conditions of the primal problem (15a),(15c)-(15g). If the optimal solution is reached then the complementary slackness conditions are satisfied and T5=0, T6=0, T7=0 and T8=0

Therefore, the sub-problem (17) can be reformulated as (18) without any disjunctive parameter.

$$\underset{\Omega_p^d}{Minimize} \sum_{t \in \mathcal{T}} \langle \sum_{i \in \mathcal{D}} D_{it} \overline{\omega}_{itks}^{(6i)} + \sum_{j \in \mathcal{G}} G_j \overline{\nu}_{jtks}^{(6g)} + \sum_{w} \widehat{G}_{wks} \overline{\eta}_{wtks}^{(6h)} +$$

$$\sum_{l \in \mathcal{L}} F_l (\overline{\mu}_{ltks}^{(6f)} + \underline{\mu}_{ltks}^{(6f)}) + \sum_{m \in \mathcal{M}} \widehat{F}_m (\overline{\gamma}_{mtks}^{(6e)} + \underline{\gamma}_{mtks}^{(6e)}) +$$

$$+ \alpha_i \widehat{\lambda}_{itks}^{(9a)} - c_j \widehat{\lambda}_{jtks}^{(9b)} \tag{18a}$$

$$(17b) - (17z) \tag{18b}$$

$$\widehat{z}_{mt}^a (\overline{\sigma}_{mtks}^{(8a)} + \underline{\sigma}_{mtks}^{(8b)}) = 0 \quad \forall m \in \mathcal{M}, t \in \mathcal{T}, k \in \mathcal{K}, s \in \mathcal{S} \tag{18c}$$

$$\widehat{z}_{mt}^a (\overline{\vartheta}_{mtks}^{(12a)} + \underline{\vartheta}_{mtks}^{(12a)}) = 0 \quad \forall m \in \mathcal{M}, t \in \mathcal{T}, k \in \mathcal{K}, s \in \mathcal{S} \tag{18d}$$

$$(1 - \widehat{z}_{mt}^a)(\overline{\zeta}_{mtks}^{(8c)} + \underline{\zeta}_{mtks}^{(8c)}) = 0 \quad \forall m \in \mathcal{M}, t \in \mathcal{T}, k \in \mathcal{K}, s \in \mathcal{S} \tag{18e}$$

$$(1 - \widehat{z}_{mt}^a)(\overline{\kappa}_{mtks}^{(12b)} + \underline{\kappa}_{mtks}^{(12b)}) = 0 \quad \forall m \in \mathcal{M}, t \in \mathcal{T}, k \in \mathcal{K}, s \in \mathcal{S} \tag{18f}$$

The Bean master problem (16) is formulated such that each feasibility cut takes into account all scenarios. In order to create tighter cuts we can reformulate the master problem (16) and replace the feasibility cut (16b) with feasibility cuts for each scenario separately as it is shown in (19). The convergence improvement is shown in (Santoso et al 2005).

$$\sum_{mt \in \Omega_{sa}^{(1)}} z_{mt} + \sum_{mt \in \Omega_{sa}^{(2)}} (1 - z_{mt}) \leq |\Omega_{sa}^{(1)}| + |\Omega_{sa}^{(2)}| - 1 + \sum_{a' \, if(K_{a'} \geq K_a)} u_{a'} \; \forall a \in \mathcal{A}, s \in \mathcal{S}$$

(19)

Furthermore, authors in (Santoso et al 2005) propose a technique to strengthen Benders cuts. We apply a similar procedure to the Bean decomposition algorithm which results in the additional constraint (20) in the master problem formulation (16)

$$\sum_{mt \; if(\widehat{z}_{mt}^a = 1)} (1 - z_{mt}) + \sum_{mt \; if(\widehat{z}_{mt}^a = 0)} z_{mt} \leq L_a \; \forall a \in \mathcal{A}$$

(20)

The intuition behind (20) is based on the need to prevent wide steps between different iterations of the master problem. At the initial iterations of the decomposition algorithm, the solution space of the master problem is vast which means at each iteration the master problem may provide a solution that is very different from the previous iteration solution. By introducing constraint (20), we can prevent the solution from changing too much between different itarations and thereby reduce the necessary number of iterations. The constraint (20) improves the convergence of the Bean decomposition algorithm. However, the parameter L_a depends on the starting point and on the iteration number, and it is hard to identify. The parameter L_a should be manually tuned for each case study. We propose to penalize large steps at each iteration in objective function using a penalty factor Υ_a. The resulting modified Bean master problem is shown in (21).

$$\underset{u_a, z_{mt}, y_{mt}}{Maximize} \sum_a K_a u_a + H u_0 - \Upsilon_a \Big(\sum_{mt \; if(\widehat{z}_{mt}^a = 1)} (1 - z_{mt}) + \sum_{mt \; if(\widehat{z}_{mt}^a = 0)} z_{mt} \Big)$$

(21a)

$Subject \; to \; :$

$$\sum_{mts \in \Omega_{sa}^{(1)}} z_{mt} + \sum_{mts \in \Omega_{sa}^{(2)}} (1 - z_{mt}) \leq |\Omega_{sa}^{(1)}| + |\Omega_{sa}^{(2)}| - 1 + \sum_{a' \, if(K_{a'} \geq K_a)} u_{a'} \; \forall a \in \mathcal{A}$$

(21b)

$$\sum_a u_a = 1$$

(21c)

$$z_{mt} = \sum_{\widehat{t} \in \{t1,..,t\}} y_{m\widehat{t}} \; \forall m \in \mathcal{M}, t \in \mathcal{T}$$

(21d)

$$\sum_{t \in \mathcal{T}} y_{mt} \leq 1 \; \forall m \in \mathcal{M}$$

(21e)

$$z_{mt}, y_{mt} \in \{0, 1\}$$

(21f)

Our proposed solution algorithm solves the modified sub-problem (18) and the modified master problem (21) while increasing the number of iterations until the
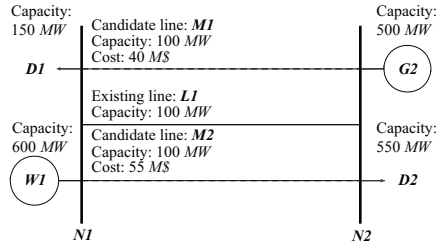
Fig. 4 The single-line diagram of the two-bus system.

optimality gap is satisfied. The proposed decomposition algorithm is presented in Algorithm 1. By using this algorithm the numerical instability problem caused by the disjunctive parameters is removed and whole solution procedure is accelerated.

Algorithm 1 Modified and accelerated Bean decomposition algorithm

---

1: **procedure** BEAN DECOMPOSITION
2:      z_{mt} =initial feasible solution; UB=∞; LB=$-\infty$
3:      **while** UP>LB **do**
4:            Solve subproblem (18)
5:            Update $\Omega_{sa}^{(1)}$ and $\Omega_{sa}^{(2)}$
6:            Append constraints (21b) and solve master problem (21)
7:      **end while**
8:      **return** Optimal solution z_{mt} and y_{mt}
9: **end procedure**

---

5 Illustrative examples

As an illustrative example, the two-bus system presented in Figure 4 is studied in detail. The system contains wind generation which is solely located in bus $N1$. The initial transmission system is congested. We assume five planning periods for transmission investment ($|\mathcal{T}|$=5). Each investment planning period represents one year and includes 8760 hours of operation assuming one representative hour for each investment planning period (ψ=8760, $|\mathcal{K}|$=1). The Transco can build additional transmission lines $M1$ and $M2$. The maximum demand for the first period is set as 550 MW with a 10% rate of increase for each next planning year. Wind is also considered as a dispatchable source of energy with zero marginal cost. The present worth factor is calculated using an annual price escalation rate (β) equal to 2% and expected rate of return (γ) of 10%. When the H-R-G-V incentive mechanism is used, the Transco invests in $M1$ and $M2$ and sets a fixed fee of 4.76 $M\$$ at $t2$ which then further decreases in each of the next periods. At the same time demand reaches 550 MW at bus $N2$ and 150 MW at bus $N1$ which corresponds to their maximum values. When the H-R-G-V incentive mechanism is not used, the Transco will invest only in one line, $M1$, and the merchandising surplus will rise by 33% in comparison with H-R-G-V case. Moreover, demand will not reach the maximum value for $t4$ and $t5$ and the system is still congested. The

Table 3 Investment results in the 2-bus system. Tran. Inv: Transmission Investment. IR: H-R-G-V regulation applied. NR: No regulation applied. [*]Wind spillage$= \frac{\widehat{G}_{wks} - \widehat{g}_{wtks}}{\widehat{G}_{wks}}$

Approach	Line label [from,to]	t1	t2	t3	t4	t5
IR	M1 [1,2]	0	**1**	**1**	**1**	**1**
	M2 [1,2]	0	**1**	**1**	**1**	**1**
NR	M1 [1,2]	0	**1**	**1**	**1**	**1**
	M2 [1,2]	0	0	0	0	0
IR	Fixed Fee (M$)	0	4.76	4.38	4.015	3.67
NR	Fixed Fee (M$)	-	-	-	-	-
IR	Merchandising surplus (M$)	4.69	6.19	5.90	5.62	5.35
NR	Merchandising surplus (M$)	4.69	8.19	7.60	7.05	6.54
Difference		**0**	**-2**	**-1.50**	**-1.43**	**-1.19**
IR	Generator and Load Surplus (M$)	3.25	8.01	7.63	7.27	6.92
NR	Generator and Load Surplus (M$)	3.25	3.41	3.54	3.65	3.75
Difference		**0**	**4.6**	**4.09**	**3.62**	**3.17**
IR	Average price N1($)	27.22	98.01	98.01	98.01	98.01
	Average price N2($)	179.68	179.68	179.68	179.68	179.68
NR	Average price N1($)	27.22	54.45	54.45	54.45	54.45
	Average price N2($)	179.68	179.68	179.68	179.68	179.68
IR	Wind spillage (%)[*]	10	0	0	0	0
NR	Wind spillage (%)[*]	10	2	2	2	2
Difference		**0**	**-2**	**-2**	**-2**	**-2**
IR	Tran. Inv. Cost (M$)	0	1.571	0	0	0
NR	Tran. Inv. Cost (M$)	0	0.762	0	0	0
Difference		**0**	**0.809**	**0**	**0**	**0**

introduction of the H-R-G-V incentive mechanism results in optimal investments in transmission lines. This in turn leads to much lower merchandising surplus and lower overall transmission congestion. The results for the two-bus system are presented in Table 3. They were also compared to welfare-maximizing transmission investment which has served as a benchmark. The results support the conclusions obtained in (Hesamzadeh et al 2018). The H-R-G-V mechanism provides sufficient incentives to the profit maximizing Transco to expand the transmission network such that social welfare is maximized and the congestion in the transmission lines and the wind power spillage are reduced.

As another illustrative example a five-bus system studied. The initial input data for the 5-bus system can be found in Table 4, Table 5 and Table 6. The introduction of the incentive mechanism results in optimal investments in transmission lines which leads to less wind spillage and less transmission congestion. The results for both, two-bus and five-bus case systems are presented in Table 3 and in Table 7, respectively.

6 Numerical results for large test systems

We apply the proposed stochastic MILP formulation presented in (15) to the IEEE 30-bus, 118-bus and 300-bus test systems. Data for the IEEE test systems are taken from data files of Matpower software, (Zimmerman et al 2011). The additional data used for simulations can be found in Table 8. We solve optimization

Table 4 Data of loads in the 5-bus system

Load	Short-run Marginal Utility ($/MWh)	Capacity (MW) at $t1$
d1	300	191
d2	300	196
d3	400	156

Table 5 Data of generators in the 5-bus system

Generator	Bus	Short-run Marginal Utility ($/MWh)	Capacity (MW)
g1	n1	14	40
g2	n1	15	170
g3	n3	30	520
g4	n4	40	200
g5	n5	10	600

Table 6 Data of transmission lines in the 5-bus system

Line label	[from,to]	Reactance (p.u.)	Capacity (MW)	Investment Cost ($/Cct.)
l1	[1,2]	0.4	100	-
l2	[1,4]	0.6	100	-
l3	[1,5]	0.2	100	-
l4	[2,3]	0.2	100	-
l5	[3,4]	0.4	100	-
l6	[4,5]	0.2	100	-
m1	[1,2]	0.2	100	200,000
m2	[3,5]	0.48	100	400,000
m3	[1,4]	0.63	100	310,000
m4	[5,2]	0.3	100	300,000
m5	[3,1]	0.3	100	380,000

problem (15) using the Benders decomposition algorithm, a commercially available state-of-the-art CPLEX solver[2] and our proposed decomposition algorithm. Both decomposition algorithm were implemented in GAMS software.The CPLEX solver is used to solve the MILP master problem and the sub-problem of each decomposition algorithm with the relative gap parameter set to zero. [3] The simulations are run on a computer with two processors and 128 GB of RAM. Wind power scenarios were simulated using the moment-matching technique explained in the Appendix A.2 of this paper (Rubasheuski et al 2014).

The performance of our proposed decomposition algorithm as compared to the Benders decomposition algorithm and the CPLEX solver is presented in Table

---

[2] The disjunctive parameters included in the formulation which is solved by the CPLEX solver are tuned using an iterative method where disjunctive parameters were increased till the point where the further change in the disjunctive parameters did not affect the solution of the problem. It should be noted that we cannot guarantee that disjunctive parameters were tuned to optimality. We are not aware of any methodology which allows one to tune the disjunctive parameters without known upper bound to optimality.

[3] This setting can be relaxed to allow for a small relative gap for both Bean and Benders decomposition algorithms. However, one should keep in mind that the strength of the cuts might be compromised. This is especially the case for Benders decomposition algorithm.

Table 7 Investment results in the 5-bus system.Tran. Inv: Transmission Investment. IR: H-R-G-V regulation applied. NR: No regulation applied. [*]Wind spillage= $\dfrac{\widehat{G}_{wks}-\widehat{g}_{wtks}}{\widehat{G}_{wks}}$

Approach	Line label [from,to]	t1	t2	t3	t4	t5
IR	m1 [1,2]	0	**1**	**1**	**1**	**1**
	m2 [3,5]	0	**1**	**1**	**1**	**1**
	m3 [1,4]	0	**1**	**1**	**1**	**1**
	m4 [5,2]	0	**1**	**1**	**1**	**1**
	m5 [3,1]	0	0	0	0	0
NR	m1 [1,2]	0	**1**	**1**	**1**	**1**
	m2 [3,5]	0	0	0	0	0
	m3 [1,4]	0	**1**	**1**	**1**	**1**
	m4 [5,2]	0	**1**	**1**	**1**	**1**
	m5 [3,1]	0	0	0	0	0
IR	Fixed Fee (M$)	0	13.53	15.10	17.13	10.70
NR	Fixed Fee (M$)	-	-	-	-	-
IR	Merchandising surplus (M$)	46.52	38.71	38.14	36.44	42.01
NR	Merchandising surplus (M$)	46.52	51.74	52.69	53.37	53.06
Difference		**0**	**-13.03**	**-14.55**	**-16.93**	**-11.05**
IR	Generator and Load Surplus (M$)	85.64	99.17	100.74	102.77	96.36
NR	Generator and Load Surplus (M$)	85.64	86.03	86.09	85.75	85.22
Difference		**0**	**13.14**	**14.65**	**17.02**	**11.13**
IR	Wind spillage (%)[*]	7	0	0	0	0
NR	Wind spillage (%)[*]	7	1.43	1.43	1.43	1.43
Difference		**0**	**-1.43**	**-1.43**	**-1.43**	**-1.43**
IReg	Tran. Inv. Cost (k$)	0	11523.8	0	0	0
NReg	Tran. Inv. Cost (k$)	0	7714.3	0	0	0
Difference		**0**	**3809**	**0**	**0**	**0**

Table 8 Input data for case studies.

	IEEE 30-bus	IEEE 118-bus	IEEE 300-bus
Number of candidate lines	20	30	60
Number of existing lines	30	175	411
Conventional Generation,(MWh)	335	4300	20678
Wind Generation, (MWh)	450	2500	12000
Scenarios, (N)	20	20	20
Operation subperiods, (N)	24	105	72
Maximum Load, (MWh)	600	4242	23526
Number of periods	10	10	15

Table 9 Results for IEEE 30-bus case study.

	Number of New Lines	Objective Function, ($)	Computation Time, (h)	Iterations
CPLEX solver	4	145.15	0.485	-
Benders decomposition algorithm	4	145.15	1.48	584
Proposed decomposition algorithm	4	145.15	0.456	152

Table 10 Results for IEEE 118-bus case study. "∗": no solution after 21 hours of simulation.

	Number of New Lines	Objective Function, ($)	Computation Time, (h)	iterations
CPLEX solver	∗	∗	∗	∗
Benders decomposition algorithm	∗	∗	∗	∗
Proposed decomposition algorithm	23	3859	10.14	2510

Table 11 Results for IEEE 300-bus case study. "∗": no solution after 21 hours of simulation

	Number of New Lines	Objective Function, ($)	Computation Time, (h)	iterations
CPLEX solver	∗	∗	∗	∗
Benders decomposition algorithm	∗	∗	∗	∗
Proposed decomposition algorithm	15	10159	14.75	3192

9, Table 10 and Table 11. It should be noted that both the standard Benders decomposition algorithm and the CPLEX solver could not find an optimal solution after 21 hours of simulation.

7 Conclusion

This paper presents a stochastic bilevel disjunctive program for transmission investment planning. The Transco is subject to a proposed H-R-G-V regulatory constraint set by the regulator. The model takes into account uncertain wind generation using a moment matching technique. First, the stochastic bilevel disjunctive program is transformed to a stochastic MILP with linearized disjunctive constraints. A series of linearizations and reformulation techniques are introduced to arrive at the final stochastic MILP with linearized disjunctive constraints. To solve the reformulated MILP model, a specialized decomposition algorithm is developed employing the disjunctive nature of optimization problem. The proposed decomposition algorithm is based on the Bean decomposition algorithm. The scenario-separated feasability cuts and a penalization technique are used. Besides, we show that the proposed decomposition algorithm does not require any tuning of disjunctive parameters.

The stochastic MILP reformulation and proposed decomposition algorithm were applied to case studies of different size. In each case, the proposed H-R-G-V mechanism effectively dealt with congested power systems with integrated stochastic wind generation. The H-R-G-V mechanism incentivizes the Transco to produce welfare-maximum outcomes resulting in much lower congestion cost in comparison to the case where no regulation is present. Welfare-maximum transmission investments not only reduce the congestion cost but also support renewable generation and result in reduced wind power spillage. The computational performance of the proposed decomposition algorithm was tested further on IEEE 30-bus, 118-bus and 300-bus test systems with stochastic wind generation. The numerical results show that the proposed decomposition algorithm helps us to avoid the effect of the

disjunctive parameter on finding the optimal solution and to improve the computational tractability of the problem. Therefore, the proposed H-R-G-V incentive mechanism, which is reformulated as MILP model, and proposed decomposition algorithm may be used as an efficient tool for transmission investment in electric power systems with wind generation.

This work could be extended by including reliability criteria and other sources of uncertainties in the transmission investment model under the H-R-G-V mechanism.

References

Bean J (1992) A Bender's approach to disjunctive programming. Tech. Rep. 92-26, University of Michigan, Ann Arbor, MI, USA

Bean JC, Hopp WJ, Duenyas I (1992) A stopping rule for forecast horizons in nonhomogeneous Markov decision processes. Operations Research 40(6):1188–1199

Bertsekas DP (1999) Nonlinear programming. Athena scientific Belmont

Bertsimas D, Jaillet P, Martin S (2018) Online vehicle routing: The edge of optimization in large-scale applications

Binato S, Pereira MVF, Granville S (2001) A new Benders decomposition approach to solve power transmission network design problems. IEEE Transactions on Power Systems 16(2):235–240, DOI 10.1109/59.918292

Codato G, Fischetti M (2004) Combinatorial Benders' cuts. In: Bienstock D, Nemhauser G (eds) Integer Programming and Combinatorial Optimization, Springer Berlin Heidelberg, Berlin, Heidelberg, pp 178–195

European Commission (2003) Towards a common co-ordinated regional congestion management in Europe. Tech. rep., Consented study, Aachen, Germany, URL http://www.en.consentec.de/publications/studies

Conejo AJ, Castillo E, Mínguez R, García-Bertrand R (2006) Decomposition in linear programming: complicating variables. Decomposition Techniques in Mathematical Programming: Engineering and Science Applications pp 107–139

Dyer J (2003) US department of energy transmission bottleneck project report. Tech. rep., Consortium for Electric Reliability Technology Solution (CERTS), URL "https://certs.lbl.gov/publications/us-department-energy-transmission"

Gans JS, King SP (2000) Options for electricity transmission regulation in Australia. Australian Economic Review 33(2):145–160, DOI 10.1111/1467-8462.00144

Garces LP, Conejo AJ, Garcia-Bertrand R, Romero R (2009) A bilevel approach to transmission expansion planning within a market environment. IEEE Transactions on Power Systems 24(3):1513–1522, DOI 10.1109/TPWRS.2009.2021230

Hesamzadeh M, Rosellón J, Gabriel S, Vogelsang I (2018) A simple regulatory incentive mechanism applied to electricity transmis-

sion pricing and investment. Energy Economics 75:423 – 439, DOI https://doi.org/10.1016/j.eneco.2018.08.033

Hogan W, Rosellon J, Vogelsang I (2010) Toward a combined merchant-regulatory mechanism for electricity transmission expansion. Journal of Regulatory Economics 38(2):113–143, DOI 10.1007/s11149-010-9123-2

Hooker J (2011) Logic-based methods for optimization: combining optimization and constraint satisfaction, vol 2. John Wiley & Sons

Joskow P, Tirole J (2002) Transmission investment: Alternative institutional frameworks. In: Wholesale Markets for Electricity, Toulous, France, pp 22–23

Léautier TO (2000) Regulation of an electric power transmission company. The Energy Journal 21(4):pp. 61–92

Lee S, Grossmann IE (2000) New algorithms for nonlinear generalized disjunctive programming. Computers and Chemical Engineering 24(9):2125–2141

Leuthold FU, Weigt H, von Hirschhausen C (2012) A large-scale spatial optimization model of the European electricity market. Networks and Spatial Economics 12(1):75–107, DOI 10.1007/s11067-010-9148-1, URL https://doi.org/10.1007/s11067-010-9148-1

Loeb M, Magat WA (1979) A decentralized method for utility regulation. The Journal of Law & Economics 22(2):399–404

Lumbreras S, Ramos A (2013) Transmission expansion planning using an efficient version of Benders' decomposition. A case study. In: 2013 IEEE Grenoble Conference, pp 1–7, DOI 10.1109/PTC.2013.6652091

Maurovich-Horvat L, Boomsma TK, Siddiqui AS (2015) Transmission and wind investment in a deregulated electricity industry. IEEE Transactions on Power Systems 30(3):1633–1643, DOI 10.1109/TPWRS.2014.2367107

McCusker S, Hobbs BF (2003) A nested Benders decomposition approach to locating distributed generation in a multiarea power system. Networks and Spatial Economics 3(2):197–223, DOI 10.1023/A:1023912019269

Neumann A, Rosellón J, Weigt H (2015) Removing cross-border capacity bottlenecks in the european natural gas market—a proposed merchant-regulatory mechanism. Networks and Spatial Economics 15(1):149–181, DOI 10.1007/s11067-014-9273-3, URL https://doi.org/10.1007/s11067-014-9273-3

Rosellón J, et al (2018) Electricity (and natural gas) transmission under transformation-an introduction. Economics of Energy & Environmental Policy 7(1)

Rubasheuski U, Oppen J, Woodruff DL (2014) Multi-stage scenario generation by the combined moment matching and scenario reduction method. Operations Research Letters 42(5):374 – 377, DOI http://dx.doi.org/10.1016/j.orl.2014.06.006

Santoso T, Ahmed S, Goetschalckx M, Shapiro A (2005) A stochastic programming approach for supply chain network design under uncertainty. European Journal of Operational Research 167(1):96–115, DOI https://doi.org/10.1016/j.ejor.2004.01.046

Sappington DEM, Sibley DS (1988) Regulating without cost information: The incremental surplus subsidy scheme. International Economic Review 29(2):pp. 297–306

Tanaka M (2007) Extended price cap mechanism for efficient transmission expansion under nodal pricing. Networks and Spatial Economics 7(3):257–275, DOI 10.1007/s11067-006-9016-1

Trespalacios F, Grossmann IE (2015) Improved Big-M reformulation for generalized disjunctive programs. Computers & Chemical Engineering 76:98–103

Vogelsang I (2001) Price regulation for independent transmission companies. Journal of Regulatory Economics 20(2):141–165

Vogelsang I (2018) Can simple regulatory mechanisms realistically be used for electricity transmission investment? The case of H-R-G-V. Economics of Energy & Environmental Policy 7:63–87, DOI 10.5547/2160-5890.7.1.ivog

Xifan W, McDonald J (1994) Modern power system planning. London (UK): Mcgraw-HillBook Company

Zimmerman RD, Murillo-Sánchez CE, Thomas RJ (2011) Matpower: Steady-state operations, planning, and analysis tools for power systems research and education. IEEE Transactions on power systems 26(1):12–19

A Appendix

A.1 Discussion of reliability issues and of economic risks for transmission investment planning

Transmission investment planning is subject to various uncertainties. Renewable generation and load uncertainties can congest the transmission system, especially, when the penetration is high. Furthermore, outages and other malfunctions of the equipment are hard to predict and therefore can affect the reliable operation of a power system. Therefore, these uncertainties have to be taken into account while the decisions on the transmission lines investments are made. In this section we provide a brief discussion of reliability criteria and risks related to the economic uncertainties.

Reliability standards vary for each system and are customarily adapted based on changing characteristics of the system such as, for example, the generation mix. The reliability standards can be incorporated in any transmission planning by enforcing additional technical constraints on transmission operation and planning. Under centralized planning, the reliability standards can be seen as the main criteria for the investments. However, under market based transmission planning the objective of a transmission investor is aimed at profit maximization. Yet, the reliability criteria can be still enforced by a regulator or a system operator (ISO). This applies to transmission investment planning under the H-R-G-V mechanism as well. The H-R-G-V incentive mechanism promotes socially optimal investments and it is aimed at reducing the congestion cost of the system. Reduced congestion, itself, will result in more reliable operation of a power system. Moreover, based on the mathematical models presented in (6) and (7) the system operator can still enforce additional technical reliability constraints which have to be met for the secure operation of the system. For example, the N-1 criterion can be modeled as an extra constraint in the lower-level program. This approach will lead to socially optimal investments while reliability criteria are satisfied. Under a different approach the regulator could promote reliability criteria by assigning monetary value to each criterion. Thus, the reliability of the power system will become a part of the reward structure of the Transco (Vogelsang 2018).

Apart from having to fulfill reliability criteria, the Transco can be subject to additional economic risks. In this paper, we assume that the Transco is a risk neutral entity. However, this assumption can be easily dropped and additional constraints on the risk tolerance of the Transco can be added to the model. However, the complexity of the problem should be reconsidered and therefore additional research is required to properly address the risk attitude of the Transco.

Other uncertainties such as changes in demand, solar power generation and any other renewable generation uncertainties can be incorporated to the proposed model using the same, moment-matching technique, presented in this paper. The description of the moment-matching scenario generation technique can be found in Appendix A.2.

A.2 The moment-matching method for generating wind scenarios

The moment-matching scenario generation technique is based on historical data. The method does not require any knowledge of the distribution of uncertain parameters, instead, it exploits statistical properties of the sampled historical data such as the mean (M_n), standard deviation (SD_n), skewness (SK_n), kurtosis (KRT_n) and correlation ($Corr_{n,m}$). These measurements are also known as moments. In (22) we present the equations used to calculate each of the aforementioned statistical property. Once the moments of the historical sampled data are known, we can generate a number of scenarios with the matching properties. The generation of the scenarios is performed using the mathematical model presented in (23) by minimizing the mismatch for each moment.

$$M_n = \frac{1}{K} \sum_k n_k \tag{22a}$$

$$SD_n = \sqrt{\frac{1}{K} \sum_k (M_n - n_k)^2} \tag{22b}$$

$$SK_n = \frac{1}{K} \frac{\sum_k (M_n - n_k)^3}{SD_n^3} \tag{22c}$$

$$KRT_n = \frac{1}{K} \frac{\sum_k (M_n - n_k)^4}{SD_n^4} \tag{22d}$$

$$Corr_{n,m} = \frac{\sum_k (M_n - n_k)(M_m - m_k)}{\sqrt{\sum_k (M_n - n_k)^2 \sum_k (M_m - m_k)^2}} \tag{22e}$$

Where n and m are the sets of uncertain parameters and k is the considered historical data set with total of K elements.

$$\underset{p_s}{Minimize} : \sum_n ((f^M(n,p) - M_n)^2 + (f^{SD}(n,p) - SD_n)^2 + (f^{SK}(n,p) - SK_n)^2$$
$$+ (f^{KRT}(n,p) - KRT_n)^2) + \sum_{n,m} (f^{Corr}(n,m,p) - Corr_{n,m})^2 \tag{23a}$$

$Subject\ to :$

$$f^M(n,p) = \sum_s n_s p_s \tag{23b}$$

$$f^{SD}(n,p) = \sum_s (f^M(n,p) - n_s)^2 p_s \tag{23c}$$

$$f^{SK}(n,p) = \sum_s \frac{(f^M(n,p) - n_s)^3 p_s}{(f^{SD}(n,p))_n^3} \tag{23d}$$

$$f^{KRT}(n,p) = \sum_s \frac{(f^M(n,p) - n_s)^4 p_s}{(f^{SD}(n,p))^4} \tag{23e}$$

$$f^{Corr}(n,m,p) =$$
$$\sum_s \frac{(f^M(n,p) - n_s)(f^M(m,p) - m_s) p_s}{\sqrt{(f^M(n,p) - n_s)^2 (f^M(m,p) - m_k)^2}} \tag{23f}$$

$$\sum_s p_s = 1 \tag{23g}$$

Where in (23), s is the index for scenarios and p_s is the probability corresponding to each scenario.