Application of Novel Nested Decomposition Techniques to Long-Term Planning Problems

P. Falugi, I. Konstantelos, G. Strbac
Department of Electrical and Electronic Engineering, Imperial College London, London, uk
{p.falugi, i.konstantelos, g.strbac}@imperial.ac.uk

***Abstract*—Cost effective, long term planning under uncertainty constitutes a significant challenge since a meaningful description of the planning problem is given by large Mixed Integer Linear Programming (MILP) models which may contain thousands of binary variables and millions of continuous variables. In this paper, a novel multistage decomposition scheme, based on Nested Benders decomposition is applied to the transmission planning problem. The difficulties in using temporal decomposition schemes in the context of planning problems due to the presence of non-sequential investment state equations are highlighted. An efficient and highly-generalizable framework for recasting the temporal constraints of such problems in a structure amenable to nested decomposition methods is presented. The proposed scheme's solution validity and substantial computational benefits are clearly demonstrated through the aid of case studies on the IEEE24-bus test system.**

***Index Terms*—Long-Term Planning, Nested Bender Decomposition, Stochastic programming**

NOMENCLATURE

The symbols used for sets and quantities are as follows

Ω_G	Set of generation units, indexed g.
Ω_E	Set of all stages.
Ω_M	Set of all nodes belonging to the scenario tree.
Ω_{W_ℓ}	Set of expansion options for line ℓ.
Ω_N	Set of system buses.
Ω_{N_s}	Set of system buses for which a storage candidate is available.
$\Omega_{N_{sT}}$	Set of system buses for which a storage device exists or a candidate storage is available.
Ω_L	Set of transmission lines.
Ω_B	Set of demand blocks indexed by b.
Ω_T^b	Set of periods in demand block b indexed by t.
$\varepsilon(m)$	Stage to which the node m belongs to.
$\Phi_k(m)$	Set of all parent nodes of m up to stage $\varepsilon(m)-k$.
$\Phi_0^e(m)$	Parent node of m at stage e.

Given a set $\mathcal{S}$, $|\mathcal{S}|$ denotes the cardinality of the set $\mathcal{S}$.
The input parameters of the optimization problem, summarised by the vector $\boldsymbol{\rho}$, are as follows

$F_{\ell,w}^{\max}$	Maximum capacity provided by expansion option w for line ℓ.
$\bar{p}_{m,g,t}$	Maximum generation for unit g at the operating point (m,t).
F_ℓ^0	Initial capacity for line ℓ.
H_n^0	Existing storage at bus $n \in \Omega_{N_{sT}}$.
$D_{t,n}^b$	Demand at bus n in period t for demand block b (MW).
$B_{n,g}$	Bus-to-generation incidence matrix.
$I_{n,\ell}$	Bus-to-line incidence matrix.
χ_ℓ	Line length in km.
W_b	Weight of demand block b.
X_ℓ	Reactance of transmission line ℓ.
u_ℓ	Sending bus for line ℓ.
v_ℓ	Receiving bus for line ℓ.
$\bar{h}$	Maximum charge/discharge rate of storage device (MW).
$\bar{\eta}$	Energy capacity of storage device (MWh).
τ_t^b	Time duration of demand period t in demand block b.
γ^H	Build time of storage device.
κ^H	Annual capital cost of storage device (£/yr).
$\kappa_{\ell,w}$	Build time for line ℓ and expansion option w.
κ_g^G	Operation cost of generating unit g (£/MWh).
$\gamma_{\ell,w}$	Annualised fixed investment cost for line ℓ, option w (£/(km yr)).
$c_{\ell,w}$	Annualised variable investment cost for line ℓ, option w (£/(MW km yr)).
Γ	System balance penalty constant (£/MWh).
ρ_{ef}	Storage efficiency.
σ_b	First period of demand block b.
T_b	Last period of demand block b.
$r_{\varepsilon(m)}^I$	Cumulative discount factor for investment cost in epoch $\varepsilon(m)$.
$r_{\varepsilon(m)}^O$	Cumulative discount factor for operation cost in epoch $\varepsilon(m)$.

The decision variables of the optimization problem are stacked in the vector $\mathbf{x}$ as follows

$f_{m,\ell,w}^{inv}$	Transmission capacity to be built for line ℓ using option w at node m.
$\beta_{m,\ell,w}$	Binary variable modeling the choice of expansion option w for line ℓ at node m.
$H_{m,n}$	Binary variable modelling the decision at node m of building a storage device at bus n.
$d_{m,t,n}$	Curtailed demand at bus n at operating point (m,t).
$p_{m,t,g}$	Output of generation unit g at operating point (m,t).
$f_{m,t,\ell}$	Power flow on line ℓ at operating point (m,t).
$\theta_{m,t,n}$	Bus angle at node n for operating point (m,t).
$h_{m,t,n}$	Output of storage device at bus n at operating point (m,t).

$\tilde{h}_{m,t,n}$ State of charge of storage device at bus n at operating point (m, t).

I. INTRODUCTION

Achieving the ambitious decarbonization goals set by governments worldwide will entail significant changes to the way electrical energy is generated, transmitted and used. While renewable sources will play a major role in supporting this transition, their cost effective integration within conventional energy systems constitutes a signicant challenge. Substantial transmission investment is required to support the emerging power flows, especially given that low-carbon sources are typically located far away from the load centres. Most importantly, following the ownership unbundling that has taken place in many jurisdictions, planners are facing increased uncertainty regarding the size, type and location of new connections. In view of the difficulty of obtaining permissions for new transmission corridors and the fast pace at which new generators can be built in areas with limited exporting capability, retaining a reactive stance is no longer feasible. Planners are starting to shift to an anticipatory planning paradigm and as a result, require suitable models to identify strategic investment opportunities [1].

Traditionally, the transmission expansion planning problem [2] has been formulated on the basis of identifying the optimal investment plan for a target horizon (e.g. [3], [4]). Recently, it has been shown that stochastic transmission planning based on scenario trees is a framework that enables planners to identify strategic opportunities which can enable the cost-efficient management of long-term uncertainty by taking into account the inter-temporal resolution of uncertainty (e.g. [5], [6], [7]). In this vein, non-network solutions such as storage, Demand-Side Response (DSR) and FACTS devices can possess significant strategic value, enabling planners to manage interim congestion until investment in large-scale interconnection projects is firmly established e.g. [8], [9]. It follows that a comprehensive planning model is required to include a plethora of competing investment options characterized by different building times, thus resulting in scenario trees of high temporal resolution. In addition, considering that the size of a scenario tree grows exponentially with the sources of uncertainty it describes along with the increasing need for extended chronological simulations of time-shifting elements, new planning tools are facing an extremely high computational burden; a meaningful description of the transmission expansion planning problem is given by large MILP models which may contain thousands of binary variables and tens of millions of continuous variables.

Decomposition techniques, based on Benders Decomposition [10] have been broadly applied to power systems [11], [12], [13] to optimize the desired objectives [14]. Transmission and generation planning models exhibit a problem structure which is highly exploitable by Benders decomposition methods [15] when a separation between investment and operation costs is performed. Enhanced multi-cut formulations [16], [9] have been proposed to improve convergence properties of Benders decomposition techniques. However these traditional decomposition techniques, which have been employed in the past, are reaching their limits: in long-term planning problems the master problem can quickly become intractable due to the large number of binary variables.

In this paper, a novel multistage decomposition scheme, based on Nested Benders decomposition [17], [18], is applied to the transmission planning problem. The proposed decomposition is performed along the temporal axis defining a sequence of master problems and subproblems similar to [19], [20] where an averaging of the future cost with respect to the realizations of the uncertainty is performed. The difficulties in using temporal decomposition schemes in the context of planning problems due to the presence of non-sequential investment state equations are highlighted. To this end an efficient and generalizable reformulation is proposed. Another major challenge is the non-convex structure of subproblems due to the presence of binary variables. To tackle this challenge, recently-developed approximation and tight relaxation techniques can be applied to the planning problem. Approximation techniques for multi-stage stochastic optimization problems with mixed-integer variables are reviewed and analyzed in [21], [22]. A first approximation consists in relaxing the mixed integer problems on the backward pass and estimating a lower bound and an upper bound of the cost. Even if the relaxed problem does not converge to the global optimum, the lower bound and the upper bound give useful information on the investment and if the gap is small, the solution given by the upper bound is acceptable. If the solution is not acceptable it is possible to refine the estimation of the bounds. Note that the Nested Benders decomposition is well suited for parallelization ([23] and [24]). The proposed scheme's solution validity and substantial computational benefits are clearly demonstrated through the aid of case studies on test systems. The model takes into account candidate investments on line reinforcements and storage installations and models operation across five typical days. It is of paramount importance to stress that the impressive computational gains achieved with the nested solution strategy are readily extensible to other application such as the stochastic unit commitment problem which also contains binary variables and non-sequential state equations due to minimum up/down times.

II. PROBLEM FORMULATION

The stochastic process $\{\xi\}_{e=1}^{N_E}$, where $N_E \triangleq |\Omega_E|$, affecting the transmission planning problem is approximated by a process forming a scenario tree which is based on a finite set Ω_M of nodes [25]. In the scenario tree the root node $m = 1$ is located at epoch $e = 1$. Since we assume the first epoch is not affected by uncertainty, at $e = 1$ there is only one node which is the root node. Every other node m has a unique predecessor $m^- = \Phi_0^{\varepsilon(m)-1}(m)$ and a set $\mathcal{N}^+(m)$ of successors. A node m belongs to stage $\varepsilon(m)$ if its distance from the root node is $\varepsilon(m) - 1$. A leaf node m is a node at stage N_E satisfying $\mathcal{N}^+(m) = \emptyset$ (it has no successors). Probability that tree node m occurs is denoted by π_m. Using the scenario tree representation of the uncertainty it is possible to formulate the stochastic transmission planning problem as a MILP.

The objective of the transmission planning problem is to minimize the expected system cost under the uncertainty described by a multi-stage scenario tree while satisfying operation and investment constraints. The mathematical formulation presented in the following sections enables the strategic investment in different line reinforcement options as well as storage assets. Note that the presented planning framework can accommodate other technologies such as DSR following straightfroward modifications to the formulation (for an example see [26]).

Two important aspects in transmission expansion models are economies of scale and the delay between investment and asset commissioning. Capturing the economies of scale present in transmission projects is essential in modelling both timing and sizing flexibility. The chosen approach is to express capital costs in terms of fixed and variable components. The planner can choose to incur some fixed costs for the ability to invest in a specific project in the future. On the other hand, incurred variable costs depend on the eventual reinforcement size, with each option limited to a maximum capacity addition. In addition, the presented formulation enables the modeling of delays. A set of investment options with different upgradeability levels and construction times have been included in the formulation to capture available choices present in a realistic setting, where the planner can choose to invest in an anticipatory manner. Our model includes investment candidate options in line capacity and storage devices.

The expected system cost is a probability weighted combination of discounted investment and operation costs, $V_m^I(\boldsymbol{\rho}, \mathbf{x}_m)$ and $V_m^O(\boldsymbol{\rho}, \mathbf{x}_m)$ respectively and it is minimized along a planning horizon

$$V^*(\rho) = \min_{\mathbf{x}} \sum_{m\in\Omega_M} \pi_m \left[r^I_{\varepsilon(m)} V_m^I(\boldsymbol{\rho}, \mathbf{x}) + r^O_{\varepsilon(m)} V_m^O(\boldsymbol{\rho}, \mathbf{x}) \right] \quad \text{subject to } (4),\ (5),\ (6),\ (7),\ (8),\ (9),\ (10),\ (11),\ (12),\ (13) \tag{1}$$

where $\mathbf{x}$, the vector of decision variables, and $\boldsymbol{\rho}$ the parameters vector, are defined in the nomenclature. The investment cost $V_m^I(\boldsymbol{\rho}, \mathbf{x}_m)$ at node m is defined by

$$\begin{aligned} V_m^I(\cdot) &= \textstyle\sum_{\ell\in\Omega_L} \sum_{w\in\Omega_{W_\ell}} (c_{\ell,w} f^{inv}_{m,\ell,w} + \gamma_{\ell,w}\beta_{m,\ell,w})\chi_\ell \\ &+ \textstyle\sum_{n\in\Omega_{N_s}} H_{m,n}\kappa^H \end{aligned} \tag{2}$$

The operation cost $V_m^O(\boldsymbol{\rho}, \mathbf{x}_m)$ at node m is defined by

$$V_m^O(\cdot) \triangleq \sum_{b\in\Omega_B} W_b \sum_{t\in\Omega_T^b} \tau_t^b \left[\sum_{g\in\Omega_G} \kappa_g^G p_{m,t,g} + \sum_{n\in\Omega_N} \Gamma d_{m,t,n} \right]. \tag{3}$$

The aggregate capacity $F^{inv}_{m,\ell}$ is the sum of the investment decisions taken in all nodes belonging to the path $\Phi_0(m)$ subject to the building option time $\kappa_{\ell,w}$

$$F^{inv}_{m,\ell} = \sum_{w\in\Omega_{W_\ell}} \sum_{\phi\in\Phi_{\kappa_{\ell,w}}(m)} f^{inv}_{\phi,\ell,w} \quad \forall \ell\in\Omega_L, m\in\Omega_M. \tag{4}$$

The bound on the capacity that can be built for line ℓ at node m using investment option w is as follows

$$f^{inv}_{m,\ell,w} \leq \sum_{\phi\in\Phi_0(m)} \beta_{\phi,\ell,w} F^{max}_{\ell,w} - \sum_{\phi\in\Phi_1(m)} f^{inv}_{\phi,\ell,w} \tag{5}$$

where $\ell\in\Omega_L$, $m\in\Omega_M$ and $w\in\Omega_{W_\ell}$. Exclusive investment options have been taken into account adding the constraints

$$\sum_{w\in\Omega_{W_\ell}} \sum_{\phi\in\Phi_0(m)} \beta_{\phi,\ell,w} \leq 1 \quad \forall \ell\in\Omega_L \tag{6}$$

The amount of storage $\tilde{H}_{m,n}$ added up at bus n and node m is given by

$$\tilde{H}_{m,n} = \sum_{\phi\in\Phi_{\gamma^H}(m)} H_{\phi,n} \qquad \forall n\in\Omega_{N_s} \tag{7}$$

The constraints (4),(5),(6) and (7) describe the evolution of different forms of investment along the prediction horizon.
At the operational level the limits to the generation dispatch for each operation point (m,t) is

$$0 \leq p_{m,t,g} \leq \bar{p}_{m,g,t} \quad \forall g\in\Omega_G,\ \forall t\in\Omega_T^b. \tag{8}$$

The distribution of the power over the network is described by the Direct Current Optimal Power Flow (DCOPF) formulation as follows

$$f_{m,t,\ell} = \frac{\theta_{m,t,u_\ell} - \theta_{m,t,v_\ell}}{X_\ell} \quad \forall t\in\Omega_T^b,\ \forall \ell\in\Omega_L,\ \forall b\in\Omega_B, \tag{9}$$

The system balance requires to impose that the sum of the injected power in each node is equal to the local demand level

$$\begin{aligned} &\sum_{g\in\Omega_G} B_{n,g} p_{m,t,g} + \sum_{\ell\in\Omega_L} I_{n,\ell} f_{m,t,\ell} + d_{m,t,n} - h_{m,t,n} = D^b_{t,n} \\ &\quad \forall t\in\Omega_T^b,\ \forall b\in\Omega_B,\ \forall n\in\Omega_N\ \forall m\in\Omega_M \end{aligned} \tag{10}$$

where $d_{m,t,n} \geq 0$ and it is used as a slack variable to ensure that operation is feasible even in cases of inadequate generation capacity. On each line ℓ the power flow has to remain inside the available system capacity

$$-(F^{inv}_{m,\ell} + F^0_\ell) \leq f_{m,t,\ell} \leq F^{inv}_{m,\ell} + F^0_\ell \quad \begin{aligned} &\forall t\in\Omega_T^b,\ \forall b\in\Omega_B \\ &\forall \ell\in\Omega_L,\ \forall m\in\Omega_M \end{aligned} \tag{11}$$

The storage operation can be effectively modeled with a first order difference equation as follows

$$\begin{aligned} &\tilde{h}_{m,t,n} = \rho_{ef}\tilde{h}_{m,t-1,n} + \tau_t^b h_{m,t,n} \quad \forall t\in\Omega_T^b\backslash\{\sigma_b\}, \\ &\quad \forall b\in\Omega_B,\ \forall n\in\Omega_{N_{sT}}\ \forall m\in\Omega_M \\ &h_{m,t,n} = 0,\ \tilde{h}_{m,t,n} = 0\ \forall n\notin\Omega_{N_{sT}} \end{aligned} \tag{12}$$

where the initial condition $\tilde{h}_{m,\sigma_b,n}$ for block b can be an input parameter or defined as a decision variable. Moreover it can be required to equate $\tilde{h}_{m,\sigma_b,n}$ to the value of $\tilde{h}_{m,t,n}$ at the last period of the demand block b i.e. $\tilde{h}_{m,\sigma_b,n} = \tilde{h}_{m,T_b,n}$. The operating storage is required to satisfy physical constraints $\forall t \in \Omega_T^b$, $\forall b \in \Omega_B$, $\forall m \in \Omega_M$

$$\begin{array}{ll} |h_{m,t,n}| \leq (\tilde{H}_{m,n} + H_n^0)\bar{h} & \forall n \in \Omega_{N_s} \\ |h_{m,t,n}| \leq H_n^0 \bar{h} & \forall n \in \Omega_{N_{sT}} \backslash \{\Omega_{N_s}\} \\ 0 \leq \tilde{h}_{m,t,n} \leq \bar{\eta} & \forall n \in \Omega_{N_{sT}} \end{array} \tag{13}$$

The transmission planning problem (1), referred to as Undecomposed, requires the solution of a large-scale Stochastic Mixed Integer Programming (SMIP) problem. The SMIP problem becomes computationally intractable using standard solution approaches when a considerable number of uncertainty scenarios is included especially in the presence of a large number of binary variables. In [9] the transmission planning problem has been tackled with a Benders decomposition scheme. The Benders decomposition allows to solve efficiently large scale problems but it is not possible to include a sufficiently significant number of possible scenarios because the size of the master problem increase rapidly with the number of nodes in the scenario tree and its solution become impossible. Here a novel decomposition based on Nested Benders decomposition scheme is used to render the model tractable for systems with multiple scenarios and operating points. The proposed decomposition is performed along the temporal axis defining a sequence of master problems and subproblems similarly to [19]. Each optimization problem can represent both a master problem and a subproblem. A sequence of master problems is solved forward in time to compute trial values of the decision variables. Then a sequence of subproblems is solved backward in time to compute the dual variables to approximate the future cost in each optimization problem. The application of this decomposition, based on Nested Benders, to the dynamic formulation of the transmission planning problem is not straightforward for three reasons: (i) the subproblems contain binary variables, (ii) the investments are required to satisfy dynamic constraints along the scenario tree and (iii) the building time of the additional capacity is affected by delay.

III. Temporal Decomposition and Investment Modeling

A multi-stage temporal decomposition requires that in each stage the constraints are only coupled with decision variables belonging to the previous stage as illustrated in Figure 1.

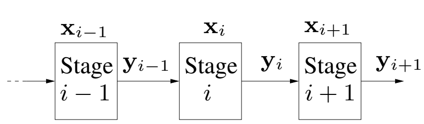

Figure 1: Desired relation among stages.

The vector $\mathbf{x}_i$ denotes the variables involved in stage i; the output $\mathbf{y}_i = C_i \mathbf{x}_i$, where C_i is a matrix of appropriate dimensions, only depends on $\mathbf{x}_i$ and describes the variables of interest to the next stage $i+1$. Since all the nodes belonging to the same stage i are decoupled with each other we can illustrate the optimization model looking at a single node m belonging to stage $i \in \Omega_E$. The structure illustrated in Figure 1 is achieved by introducing additional variables and state equations. Mutually exclusive investment constraints (6) in line capacity can be reformulated, introducing additional decision variables $x_{m,\ell}^{\beta} \in \mathbb{R}$ for all $\ell \in \Omega_L$, as follows

$$\begin{array}{l} x_{m,\ell}^{\beta} = y_{m^-,\ell}^{\beta} \\ x_{m,\ell}^{\beta} + \sum_{w=1}^{N_{W_\ell}} \beta_{m,\ell,w} \leq 1 \quad \forall \ell \in \Omega_L \end{array} \tag{14}$$

where $y_{m,\ell}^{\beta} \triangleq x_{m,\ell}^{\beta} + \sum_{w=1}^{N_{W_\ell}} \beta_{m,\ell,w}$ and $x_{1,\ell}^{\beta} = y_{1^-,\ell}^{\beta} = 0$.
Similarly limits (5) on transmission capacity $f_{m,\ell,w}^{inv}$ for each line $\ell \in \Omega_L$ at node $m \in \Omega_M$ are required to satisfy

$$\begin{array}{l} \mathbf{x}_{m,\ell}^{c} = \mathbf{y}_{m^-,\ell}^{c} \\ \mathbf{x}_{m,\ell}^{c} + \mathcal{F}^{\max,\ell} \boldsymbol{\beta}_m^{\ell} - \mathbf{f}_{m,\ell}^{inv} \geq \mathbf{0} \end{array} \tag{15}$$

where $\mathbf{x}_{m,\ell}^{c} \in \mathbb{R}^{N_{W_\ell}}$ is an additional decision variable. The inequalities in (15) are defined componentwise, $\mathbf{y}_{m,\ell}^{c} \triangleq \mathbf{x}_{m,\ell}^{c} + \mathcal{F}^{\max,\ell} \boldsymbol{\beta}_m^{\ell} - \mathbf{f}_{m,\ell}^{inv}$, $\mathbf{x}_{1,\ell}^{c} = \mathbf{y}_{1^-,\ell}^{c} = \mathbf{0}$, $\boldsymbol{\beta}_m^{\ell} \triangleq [\beta_{m,\ell,1}, \ldots, \beta_{m,\ell,N_{W_\ell}}]'$ $\mathcal{F}^{\max,\ell} \in \mathbb{R}^{N_{W_\ell} \times N_{W_\ell}}$ is a diagonal matrix with the diagonal entries given by $[F_{\ell,1}^{max}, \ldots, F_{\ell,N_{W_\ell}}^{max}]'$ and $\mathbf{f}_{m,\ell}^{inv} \triangleq [f_{m,\ell,1}^{inv}, \ldots, f_{m,\ell,N_{W_\ell}}^{inv}]'$. Note that the term $\mathbf{y}_{m^-,\ell}^{c} \in \mathbb{R}^{N_{W_\ell}}$ describes the transmission capacity that has been decided to be built in the set of all parent nodes of m. Constraints (15) only depend on a subset of variables belonging to the previous stage.

In contrast to $\beta_{m,\ell,w}$ and $f_{m,\ell,w}^{inv}$, describing the investment process, it is fundamental to note that the aggregate capacity $F_{m,\ell}^{inv}$ available for operational use at node m is affected by a delay modeling the building time of the transmission capacity. The choice of reinforcing line ℓ is taken at stage $\varepsilon(m)$ but the transmission capacity will be available at a stage $j \geq \varepsilon(m)$ depending on the features of the expansion options. When a delay larger than one stage is present, the sequential state equations break down. As such, it is not straightforward to get the structure required by Nested Benders' decompositions. The target structure can be achieved by defining an extended state carrying information on when the capacity is going to be available for operational use. Given $\ell \in \Omega_L$ let $\kappa_\ell^{\max} \triangleq \max\{\max_{w \in \Omega_{W_\ell}} \kappa_{\ell,w}, 1\}$. Let $\mathbf{z}_{m^-,\ell}^{F_\ell} \in \mathbb{R}^{\kappa_\ell^{\max}}$ be a state vector whose k-th entry contains the transmission capacity decided by all father nodes of m which is available at stage $\varepsilon(m) + k - 1$ for $k = 1, \ldots, \kappa_\ell^{\max}$. Let $\mathcal{B}^{\ell} \in \mathbb{R}^{\kappa_\ell^{\max} \times N_{W_\ell}}$ and $\mathcal{A}^{\ell} \in \mathbb{R}^{\kappa_\ell^{\max} \times \kappa_\ell^{\max}}$ be matrices whose elements, respectively $b_{j,w}^{\ell}$ and $a_{j,k}^{\ell}$ for $j = 1, \ldots, \kappa_\ell^{\max}$, $k = 1, \ldots, \kappa_\ell^{\max}$ and $w = 1, \ldots, N_{W_\ell}$, are given by

$$b_{j,w}^{\ell} = \begin{cases} 1 & \text{if } j \geq \kappa_{\ell,w} \; w \in \Omega_{W_\ell} \\ 0 & \text{otherwise} \end{cases} \tag{16}$$

$$a_{j,k}^{\ell} = \begin{cases} 1 & \text{if } k = j+1, \text{ or } j = \kappa_\ell^{\max}, \; k = \kappa_\ell^{\max} \\ 0 & \text{otherwise} \end{cases} \tag{17}$$

The aggregate capacity $F_{m,\ell}^{inv}$ for $\ell \in \Omega_L$ can be modeled by

$$\begin{array}{l} \mathbf{x}_{m,\ell}^{F_\ell} = \mathbf{y}_{m^-,\ell}^{F_\ell} \\ F_{m,\ell}^{inv} = x_{m,\ell}^{F_\ell}(1) + \mathcal{B}_0^{\ell} \mathbf{f}_{m,\ell}^{inv} \end{array} \tag{18}$$

where $\mathbf{x}^{F_\ell}_{m,\ell} = [x^{F_\ell}_{m,\ell}(1), \ldots, x^{F_\ell}_{m,\ell}(\kappa^{\max}_\ell)]'$ is a decision vector, $\mathbf{y}^{F_\ell}_{m,\ell} \triangleq \mathcal{A}^\ell \mathbf{x}^{F_\ell}_{m,\ell} + \mathcal{B}^\ell \mathbf{f}^{inv}_{m,\ell}$ with $\mathbf{x}^{F_\ell}_{1,\ell} = 0$ and $\mathcal{B}^\ell_0 \in \mathbb{R}^{1\times N_{W_\ell}}$ is defined as

$$b^{\ell,0}_w = \begin{cases} 1 & \text{if } \kappa_{\ell,w} = 0,\ w \in \Omega_{W_\ell} \\ 0 & \text{otherwise} \end{cases} \tag{19}$$

According to (18), we have successfully rendered investment variables at stage $\varepsilon(m)$ to only depend on$\varepsilon(m)$ and $\varepsilon(m)-1$. Constraints (14),(15) and (18) describe investment and commissioning of transmission capacity. The same approach can be applied to investment options such as storage.

The storage capacity $\tilde{H}_{m,n}$ commissioned up to decision point m, at bus $n \in \Omega_{N_s}$, can be modeled using the decision vector $\mathbf{x}^{h_n}_m$, carrying information on the amount of storage available at node m decided in the previous stages, as follows

$$\begin{aligned} &\mathbf{x}^{h_n}_m = \mathbf{y}^{h_n}_{m^-} \\ &\tilde{H}_{m,n} = x^{h_n}_m(1) + \mathcal{B}^h_0 H_{m,n} \end{aligned} \tag{20}$$

where $\mathcal{B}^h_0$ is a scalar equal to 1 if $\gamma^H = 0$ and 0 otherwise, $\mathbf{y}^{h_n}_m = \mathcal{A}^h \mathbf{x}^{h_n}_{m^-} + \mathcal{B}^h H_{m,n}$ with $\mathbf{x}^{h_n}_1 = \mathbf{0}$. The entries $a^h_{j,k}$ and b^h_j of $\mathcal{A}^h \in \mathbb{R}^{\gamma^{\max}\times\gamma^{\max}}$ and $\mathcal{B}^h \in \mathbb{R}^{\gamma^{\max}\times 1}$ are defined by

$$a^h_{j,k} = \begin{cases} 1 & \text{if } k = j+1, \text{ or } j = \gamma^{\max},\ k = \gamma^{\max} \\ 0 & \text{otherwise} \end{cases} \tag{21}$$

$$b^h_j = \begin{cases} 1 & \text{if } j = \gamma^{\max} \\ 0 & \text{otherwise} \end{cases} \tag{22}$$

where $\gamma^{\max} = \max\{\gamma^H, 1\}$. Using the proposed formulation, constraints (20) at stage $\varepsilon(m)$ only depend on variables at stage $\varepsilon(m)$ and $\varepsilon(m)-1$. Then the maximum number of decision variables added up to each optimization problem is the same for all nodes and equal to $|\Omega_L| + \sum_{\ell\in\Omega_L} |\Omega_{W_\ell}| + \sum_{\ell\in\Omega_L} \kappa^{\max}_\ell + \gamma^{\max}|\Omega_{N_s}|$. Note that all additional variables can be defined to be continuous without affecting the optimal solution.

IV. Nested Decomposition applied to the transmission planning problem

In this section we propose a multistage decomposition scheme exploiting the constraints' structure achieved in section III. The nodes of the scenario tree correspond to decision points and a master problem and a subproblem is associated to each node. In this way the original problem is decomposed into small tractable optimization problems of uniform size. However, the subproblems are not convex since some of the variables have binary restrictions. A convexification of the future value function can be achieved by relaxing the binary decision variables to be continuous. This relaxation allows us to compute Lagrange multipliers in the subproblems and obtain a first approximation of the future value functions. The relaxed formulation does not guarantee convergence to the optimal solution but returns upper and lower bounds. The upper bound gives the cost of actually investing in options that are available whereas the lower bound gives the cost of investing in the minimum options required to meet the system constraints. If the gap between the lower and the upper bound is too large it is possible to refine the approximation arbitrarily by sequential convexification techniques reviewed in [21], [22].

Let $V_m(\boldsymbol{\rho}, \mathbf{x}_m) \triangleq \pi_m \left[r^I_{\varepsilon(m)} V^I_m(\boldsymbol{\rho}, \mathbf{x}_m) + r^O_{\varepsilon(m)} V^O_m(\boldsymbol{\rho}, \mathbf{x}_m) \right]$. The master problems and the subproblems at iteration k for each node m are denoted by $\mathbb{P}^M_m(\boldsymbol{\rho}, y^k_{m^-})$ and $\mathbb{P}^S_m(\boldsymbol{\rho}, y^k_{m^-})$ respectively and, they can be compactly defined by the problem $\mathbb{P}^\zeta_m(\boldsymbol{\rho}, y^k_{m^-})$ as follows

$$\min_{\mathbf{x}_m, \alpha^i_j \geq 0} V_m(\boldsymbol{\rho}, \mathbf{x}_m) + \sum_{j\in\mathcal{N}^+(m)} \alpha^m_j \tag{23}$$

subject to

$$(14),\ (15),\ (18),\ (20) \tag{24}$$

$$(8),\ (9),\ (10) \tag{25}$$

$$(11),\ (12),\ (13) \tag{26}$$

$$\begin{aligned} &\text{If } \varepsilon(m) < N_E, \quad \alpha^m_j \geq V^{S,*}_j(\boldsymbol{\rho}, y^\nu_m) + (\mathbf{x}^I_m - \mathbf{x}^{I,\nu}_m)' F'_s \Lambda^{S,\nu}_j \\ &\quad j \in \mathcal{N}^+(m),\ \nu = 1, \ldots, k-1,\ k > 1 \end{aligned} \tag{27}$$

where $\mathbf{x}^{I,k}_m$ denotes the investment variables contained in the optimal solution $\mathbf{x}^k_m$ to $\mathbb{P}^M_m(\boldsymbol{\rho}, y^k_{m^-})$. When $\zeta = S$, all the decision variables in the optimization problem $\mathbb{P}^S_m(\boldsymbol{\rho}, y^k_{m^-})$ take values over the real numbers i.e the binary variables are relaxed. The complicating variables are only given by investment variables (see section (III)), so we can express the Benders cuts (27) as linear functions of investment variables denoted by $\mathbf{x}^I_m$. The vector $\Lambda^k_{\varepsilon(m)}$ denotes the dual optimal solution to $\mathbb{P}^S_m(\boldsymbol{\rho}, y^k_{m^-})$ and $V^{S,*}_m(\boldsymbol{\rho}, y^k_{m^-})$ its optimal value function. The dual optimal variables associated with the constraints containing terms belonging to $y^k_{m^-}$ (see constraints (14), (15), (18) and (20)) are indicated by $\Lambda^{S,k}_j$. The matrix F_s in (27) is such that $y^k_m = F_s \mathbf{x}^{I,k}_m$ (see Figure 1). Note that in the proposed decomposition F_s is constant and identical in each node so that it can be pre-computed in advance.

The proposed multi-stage decomposition algorithm, illustrated in Figure 2, is an iterative procedure solving a sequence of master problems forward in time and a sequence of subproblems backward in time. The convergence criterion for the iterative procedure is based on estimated lower bounds of the optimal cost $V^*(\boldsymbol{\rho})$ of the transmission planning problem. A convergence criterion typically used is checking if the lower bound does not improve within a certain tolerance ϵ. The lower bound at iteration k is given by

$$Z^k_l = V^{M,*}_1(\boldsymbol{\rho}, y^k_{1^-}). \tag{28}$$

where $V^{M,*}_1(\boldsymbol{\rho}, y^k_{1^-})$ is the optimal cost of the master problem $\mathbb{P}^M_1(\boldsymbol{\rho}, y^k_{1^-})$. The upperbound of $V^*(\boldsymbol{\rho})$ at iteration k is given by

$$Z^k_u = \sum_{m\in\Omega_M} V_m(\boldsymbol{\rho}, \mathbf{x}^k_m). \tag{29}$$

Since the backward pass optimises a continuous relaxed problem, the convergence of Z^k_l and Z^k_u to $V^*(\boldsymbol{\rho})$ is not guaranteed. If a higher solution accuracy is required it is possible to refine the convexifications of the subproblems as described in [21], [22] at the expense of increasing the computational burden.

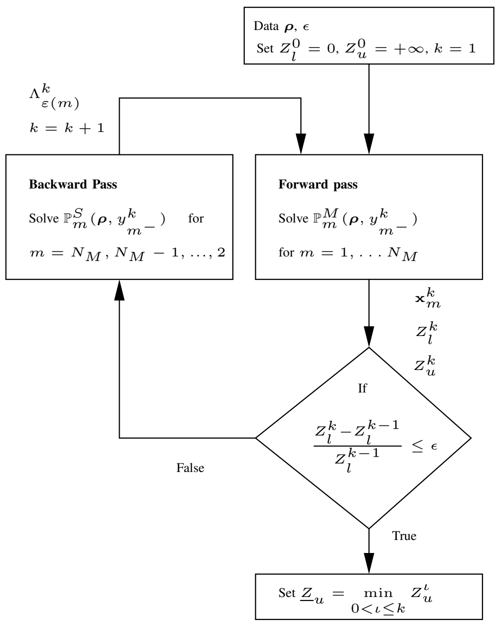

Figure 2: Decomposition Algorithm

The Benders decomposition applied in [9] solves a total of $|\Omega_B||\Omega_M|+1$ optimization problems per iteration. The number of binary variables in the master problem is $(\sum_{\ell\in\Omega_L}|\Omega_{W_\ell}| + |\Omega_{N_s}|)|\Omega_M|$. Note that the number of binary variables grows linearly with the number of nodes $N_M = |\Omega_M|$ in the scenario tree. The proposed decomposition requires the solution of N_M master problems and $N_M - 1$ subproblems. In our formulation the number of binary variables involved in the master problem $\mathbb{P}^M_m(\boldsymbol{\rho}, y^k_{m-})$ is $\sum_{\ell\in\Omega_L}|\Omega_{W_\ell}| + |\Omega_{N_s}|$ and does not depend on the value of N_M. Since mixed integer problems are NP-Hard [27] it is preferable to formulate problems with a small number of discrete variables; in this respect the proposed decomposition is clearly advantageous compared to Benders' decomposition when the value of N_M is large.

V. IEEE-RTS CASE STUDY

In this section we analyse the computational benefits of the proposed decomposition scheme using the IEEE24-bus system (IEEE-RTS) [28] which is extensively used in benchmarking studies worldwide. The network topology is the same as in [28] but some minor changes have been implemented to render the analysis more straightforward. The model consists of 24 buses, 39 lines and 27 generation units. The operating cost of the three generation technologies, nuclear, coal and oil, have been set to 6, 50 and 150 £/MWh respectively. In the first stage the total conventional generation capacity is 3105 MW. The length of all the lines is 50km and a line connecting buses 7 and 8 has been added. The hydro plants have been removed to further increase the need for investment in this system. The peak load is assumed to be 2850 MW along all the stages. The uncertainty of the generation capacity of a prospective wind farm development at bus node 24 is modeled as a scenario tree using the variable input parameter $p^{\max}_{m,g}$. In order to model different operating conditions occurring during a year we have used 5 demand blocks of 24 hours representing the daily trend of the demand. One block for each calendar season and an extra block modeling peak demand during the period close to Christmas are introduced. The blocks representing the calendar seasons are assumed to repeat 87.5 times while the extra block is repeated 15 times. The repetition of the blocks is included by weighting the corresponding part of the cost accordingly. The weekly wind and loading time series defining the input parameters $D^b_{t,n}$ and $\bar{p}_{m,g,t}$ respectively have been extracted from measurements of Great Britain's aggregate wind production and electricity demand in 2012.

In term of investment, the planner can choose to invest on transmission line reinforcement and storage devices and the value of the parameters are reported in Table I.

TABLE I: TRANSMISSION LINE REINFORCEMENT OPTIONS

w	Reinforcement Capacity [MW]	$c_{\ell,w}$ £/(MW km yr)	$\gamma_{\ell,w}$ £/(km yr)	$\kappa_{\ell,w}$
A	200	50	60000	1, $\forall\ell$
B	400	50	80000	1, $\forall\ell$
C	800	50	130000	2, $\forall\ell$

The transmission line reinforcement options have a building time delay $\gamma_{\ell,w}$ of 1 stage. The annualized capital cost of storage devices amounts to $15,000,000$£/yr and their building time is not affected by delay. The energy capacity $\bar{\eta}$ of the storage device is 1600MWh and the maximum charge/discharge rate $\bar{h}$ is 400MW. Load curtailment has been economically penalized using $\Gamma = 30,000$ £/MWh. An annual discount rate of 5% has been chosen to compute the cumulative discount investment and operation costs $r^I_{\varepsilon(m)}$ and $r^O_{\varepsilon(m)}$.

The computational benefits of proposed Nested Benders' decomposition is investigated comparing the performance with the Undecomposed model and the Benders decomposition used in [9]. The comparisons have been performed using a single storage candidate at bus 24 and a binary scenario tree with an increasing number of stages $N_E = 4, \ldots 8$ and nodes N_M.

The optimisation models have been implemented and executed using Matlab with Gurobi 6.0 ([29]) on a Xeon computer with two 3.46GHz processors. The standard Gurobi's settings have been used apart from selecting the Deterministic Concurrent approach. The convergence criterion for the Benders decomposition in [9] is given by $(\underline{Z}_u - Z^k_l)/\underline{Z}_u \leq 10^{-5}$ while the stopping criterion of the Nested Benders decomposition uses an $\epsilon = 10^{-12}$. Table II shows the numerical advantages with respect to Benders Decomposition [9] and the original undecomposed problem when the number of stages and nodes N_M in the scenario tree increases.

TABLE II: COMPUTATIONAL STUDIES

Nodes $\mathrm{N_M}$	**N. of iterations**	**CPU time**	**Cost lower bound £m**	**Cost upper bound £m**
Benders decomposition				
15	12	55s	6346.3	6346.3
31	15	5m 52s	6917.1	6917.1
63	11	3d 19h 39m	7309.5	7328.5
127	6	15d 14h 35m	7559.8	7642.9
255	7	21d 7h	7585.3	7674.4
Nested Benders decomposition				
15	13	1m 58s	6301.3	6348.8
31	14	4m 28s	6840.8	6922.8
63	16	10m 39s	7219.7	7316.9
127	12	16m 13s	7449.2	7612.3
255	12	28m 28s	7464.9	7637.4
Undecomposed Problem				
15	1	9m 3s	6346.3	6346.3
31	1	13h 19m	6917.1	6917.1
63	Timed out after 21 days			

The number of iterations required by the decomposition algorithms is practically uniform with respect to the number of stages but the execution time required by the Benders decomposition considerably increases with the number of scenarios. In all the computational studies reported in table II the total CPU time required by the Nested Benders decomposition is less than half hour while the Benders decomposition requires several days when the number of nodes in the scenario tree increases. Moreover, note that the lower and upper bound of the cost in the Benders decomposition does not achieve the convergence criterion when the number of nodes is large, even if the algorithm ran for several days. Table III shows investment decision for $N_M = 15$ together with the investment (IC), operation (OC) and total costs (TC) for all scenarios. The storage devices are denoted as STOR.

TABLE III: CASE STUDY FOR $N_M = 15$.

	Investment Decisions				Costs (£m)		
	Epochs						
	1	2	3	4	IC	OC	TC
Benders' decomposition							
s_1	STOR	A(6, 7)	A(27)	–	269.1	5225.1	5494.2
s_2	STOR	A(6, 7)	A(27)	-	269.1	5322.2	5591.3
s_3	STOR	A(6, 7)	-	-	251.2	5610.5	5861.7
s_4	STOR	A(6, 7)	-	-	251.2	5830.8	6082.0
s_5	STOR	-	A(7)	-	212.4	6208.7	6421.0
s_6	STOR	-	A(7)	-	212.4	6410.6	6623.0
s_7	STOR	-	-	-	196.3	6863.9	7060.2
s_8	STOR	-	-	-	196.3	7203.3	7399.6
Nested Benders' decomposition							
s_1	STOR	A(7)	A(6, 27)	-	258.4	5249.3	5507.7
s_2	STOR	A(7)	A(6, 27)	-	258.4	5346.2	5604.6
s_3	STOR	A(7)	-	-	224.4	5630.0	5854.4
s_4	STOR	A(7)	-	-	224.4	5832.1	6056.5
s_5	STOR	-	A(7)	-	212.4	6208.7	6421.0
s_6	STOR	-	A(7)	-	212.4	6410.6	6623.0
s_7	STOR	-	-	-	196.3	6863.9	7060.2
s_8	STOR	-	-	-	196.3	7203.3	7399.6

The scenario tree used in this case for the IEEE-RTS case study is reported in Figure 3.

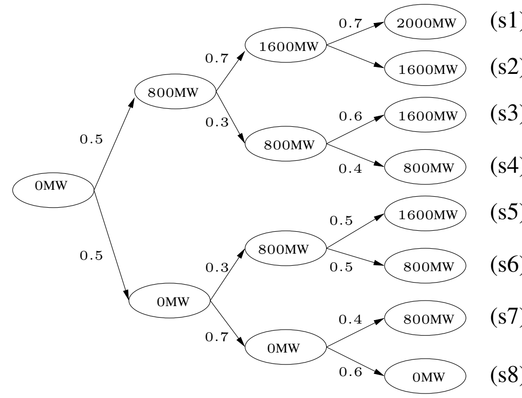

Figure 3: Transition probabilities for $N_M = 15$

VI. CONCLUSIONS

In this paper a novel multistage decomposition scheme is applied to the transmission planning problem. The main difficulty in using temporal decomposition schemes in the context of planning problems is due to the presence of non-sequential investment state equations. A general formulation of non-sequential investment constraints suitable for temporal decompositions is attained. The implemented decomposition is based on the definition of a master problem and a sub-problem for each node in the scenario tree; each iteration of the algorithm solves N_M optimization problems. In this way the original problem is decomposed into small tractable subproblems with an uniform size. This is contrasted to conventional Benders' decomposition approaches where the number of binary variables present in the master problem increases linearly with the number of scenario tree nodes N_M. Through case studies on IEEE-RTS we show the significant computational benefits obtained by applying the temporal decomposition when the size of the scenario tree increases. The use of the problem structure in the proposed decomposition renders the transmission planning problem tractable for long-term cost-benefit studies. The proposed framework is efficient and highly-generalizable to investment and operation problems involving decision delays.

REFERENCES

[1] V. Rious, Y. Perez, and J. Glachant, "Power transmission network investment as an anticipation problem," *Rev. Netw. Econ.*, vol. 10, pp. 1–21, 2011.

[2] R. Hemmati, R. A. Hooshmand, and A. Khodabakhshian, "Comprehensive review of generation and transmission expansion planning," *IET Generation, Transmission & Distribution*, 2013.

[3] E. E. Sauma and S. S. Oren, "Proactive planning and valuation of transmission investments in restructured electricity markets," *J. Regulat. Econ.*, vol. 30, pp. 261–290, 2006.

[4] J. H. Zhao, Z. Y. Dong, P. Lindsay, and K. P. Wong, "Flexible transmission expansion planning with uncertainties in an electricity market," *IEEE Trans. Power Syst.*, vol. 24, pp. 479–488, 2009.

[5] L. G. Epstein, "Decision making and the temporal resolution of uncertainty," *Int. Econ. Rev.*, vol. 21, pp. 269–283, 1980.

[6] B. G. Gorenstin, N. M. Campodonico, J. P. Costa, and M. V. F. Pereira, "Power system expansion planning under uncertainty," *IEEE Trans. Power Syst.*, vol. 8, 1993.

[7] S. Ahmed, A. J. King, and G. Parija, "A multi-stage stochastic integer programming approach for expansion under uncertainty," *J. Global Optimiz.*, vol. 26, pp. 3–24, 2003.

[8] S. Giannelos, I. Konstantelos, and G. Strbac, "Option Value of Soft Open Points in Distribution Networks," in *PowerTech, 2015 IEEE Eindhoven*, 2015.

[9] I. Konstantelos and G. Strbac, "Valuation of flexible transmission investment options under uncertainty," *IEEE Transactions on Power systems*, 2015.

[10] J. Benders, "Partitioning methods for solving mixed variables programming problems," *Numerische Mathematik*, vol. 4, pp. 238–252, 1962.

[11] S. Granville, E. P. R. Institute, and S. U. S. O. Laboratory, *Mathematical Decomposition Techniques for Power System Expansion Planning: Decomposition methods and uses*. The Institute, 1988.

[12] A. Conejo, E. Castillo, R. Minguez, and R. Garcia-Bertrand, *Decomposition techniques in mathematical programming: engineering and science applications*. Springer-Verlag, Berlin, 2006.

[13] C. Lee, S. K. Ng, J. Zhong, and F. F. Wu, "Transmission expansion planning from past to future," in *IEEE PES power systems conference and exposition*, 2006, pp. 257–265.

[14] M. Shahidehpour and F. Yong, "Benders decomposition: Applying Benders decomposition to power systems," *IEEE Power Energy Mag.*, vol. 3, no. 2, pp. 20–21, 2005.

[15] F. D. Munoz, B. F. Hobbs, and J. P. Watson, "New bounding and decomposition approaches for MILP investment problems: Multi-area transmission and generation planning under policy constraints," *European Journal of Operational Research*, 2015.

[16] J. R. Birge and F. V. Louveaux, "A multicut algorithm for two-stage stochastic linear programs," *Eur. J. Oper. Res.*, vol. 34, pp. 384–392, 1988.

[17] R. V. Slyke and R. Wets, "L-shaped linear programs with applications to optimal control and stochastic programming." *SIAM J. Appl. Math.*, vol. 17, pp. 638–663, 1969.

[18] J. R. Birge, "Decomposition and partitioning methods for multistage stochastic programming," *Operations Research*, vol. 33, pp. 989–1007, 1985.

[19] M. Pereira and L. Pinto, "Multi-stage stochastic optimization applied to energy planning," *Mathematical Programming*, vol. 52, pp. 359–375, 1991.

[20] N. Newham, "Transmission investment planning using sddp," in *Power Engineering Conference*, 2007, pp. 1–5.

[21] S. Sen, "Stochastic mixed-integer programming algorithms: Beyond benders' decomposition," in *Encyclopedia of Operations Research and Management Science*. Wiley, 2010.

[22] S. Vigerske, "Decomposition of multistage stochastic programs and a constraint integer programming approach to mixed-integer nonlinear programming," Ph.D. dissertation, Humboldt-University Berlin, 2013.

[23] J. R. Birge, C. J. Donohue, D. F. Holmes, and O. G. Svintsitski, "A parallel implementation of the nested decomposition algorithm for multistage stochastic linear programs," *Mathematical Programming*, vol. 75, pp. 327–352, 1996.

[24] M. A. H. Dempster and R. T. Thompson, "Parallelization and aggregation of nested benders decomposition," *Annals of Operations Research*, vol. 81, pp. 163–188, 1998.

[25] J. Birge and F. Louveaux, *Introduction to Stochastic Programming*. Springer, New York, 1997.

[26] I. Konstantelos, S. Giannelos, and G. Strbac, "Strategic valuation of smart grid technology options in distribution networks," submitted to IEEE Transactions on Power Systems.

[27] M. Garey and D. S. Johnson, *Computers and intractability a guide to the theory of NP-completeness*. W. H. Freeman and Company, 1979.

[28] R. T. S. T. F. of the Application of Probability Methods Subcommittee, "IEEE reliability test system," *IEEE Trans. Power App. Syst.*, vol. 13, pp. 2047–2054, 1979.

[29] I. Gurobi Optimization, "Gurobi optimizer reference manual," 2015. [Online]. Available: http://www.gurobi.com