Decomposition methods for multi-horizon stochastic programming

Hongyu Zhang[a,*], Èric Mor Domènech[b], Ignacio E. Grossmann[c], Asgeir Tomasgard[a]

[a]*Department of Industrial Economics and Technology Management, Norwegian University of Science and Technology, Høgskoleringen 1, 7491, Trondheim, Norway*
[b]*Department of Computer Science, Universitat Politècnica de Catalunya, Carrer de Jordi Girona, 31, 08034 Barcelona, Spain*
[c]*Department of Chemical Engineering, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA 15213, USA*

Abstract

Multi-horizon stochastic programming includes short-term and long-term uncertainty in investment planning problems more efficiently than traditional multi-stage stochastic programming. In this paper, we exploit the special structure of multi-horizon stochastic linear programming, and establish that it can be decomposed by Benders decomposition and Lagrangean decomposition. In addition, we propose parallel Lagrangean decomposition with primal reduction that, (1) solves the scenario subproblems in parallel, (2) reduces the primal problem by keeping one copy for each scenario group at each stage, and (3) solves the reduced primal problem in parallel. We compare the parallel Lagrangean decomposition with primal reduction with the standard Lagrangean decomposition and standard Benders decomposition on a stochastic energy system investment planning problem. The computational results show that: (a) the Lagrangean type decomposition algorithms have better convergence at the first iterations to Benders decomposition, and (b) parallel Lagrangean decomposition with primal reduction is up to 9.2 times faster than standard Benders decomposition for a 1% convergence. Based on the computational results, the choice of algorithms for multi-horizon stochastic programming is discussed.

Keywords: Stochastic programming, Multi-horizon stochastic programming, Lagrangean decomposition, Benders decomposition

1. Introduction

Multi-horizon stochastic programming (MHSP) is a powerful modelling approach that can include long-term and short-term uncertainty for long-term investment planning problems with much smaller model size than traditional multi-stage stochastic programming (Kaut et al., 2014). MHSP was further formalised in (Escudero & Monge, 2018). In addition, the bounds and formulation of MHSP have been studied (Maggioni et al., 2020). The literature on MHSP is much more sparse compared with multi-stage stochastic programming. Existing literature mainly centre around the

*Corresponding author
Email addresses: hongyu.zhang@ntnu.no (Hongyu Zhang), ericmord23@gmail.com (Èric Mor Domènech), grossmann@cmu.edu (Ignacio E. Grossmann), asgeir.tomasgard@ntnu.no (Asgeir Tomasgard)

application of MHSP for long-term investment planning problems, especially in energy system planning (Zhang et al., 2022; Backe et al., 2022). The applications show that although MHSP reduces the problem size, the monolithic model can still be hard to solve. Therefore, in this paper, we extend the literature by proposing and comparing decomposition algorithms for MHSP.

Fewer decomposition methods for MHSP have been proposed. (Mazzi et al., 2020) proposed adaptive Benders decomposition to solve large scale optimisation problems with column bounded block-diagonal structure, where subproblems differ in right-hand side and cost coefficients. MHSP belongs to this class of optimisation problems if it is formulated using a node formulation and the operational scenarios are identical in all nodes. They apply the adaptive Benders decomposition to solve a stochastic investment planning problem, and show that the computational time reduces significantly. The limitation of (Mazzi et al., 2020) is that adaptive Benders cannot solve problems where operational scenarios are different in each node. (Zhang et al., 2022) proposed stabilised adaptive Benders decomposition to solve MHSP, and apply it to solve a large-scale power system planning problem. Furthermore, (Zhang et al., 2023) proposed centre point stabilised adaptive Benders decomposition for solving large scale problem with integer variables. The existing literature only has focused on developing Benders type decomposition utilising the node formulation of MHSP. Decomposition algorithms that utilise the scenario formulation of MHSP are missing in the literature.

In this paper, we propose parallel Lagrangean decomposition with Primal Reduction (PLPR) to solve linear programming based MHSP with a scenario formulation. In addition, we show that scenario based MHSP can be decomposed by Lagrangean decomposition. Compared with standard Lagrangean decomposition, the PLPR solves the scenario subproblem in parallel, and reduces the primal problem by keeping one copy in each scenario group at each stage, and solves the primal problem in parallel. The choice of Lagrangean type decomposition and Benders type decomposition is not clear for MHSP. Therefore, we compare Lagrangean decomposition, Benders decomposition and PLPR to provide some computational insights.

The following assumptions are made in this paper: (1) each operational problem can be solved using commercial linear programming solvers, (2) the operational problem in each strategic node has several scenarios but not a multi-stage stochastic programming problem itself, and (3) the problem has relatively complete recourse at every stage.

We apply the proposed algorithms to solve the REORIENT model (Zhang et al., 2023). The REORIENT model is an MHSP proposed for integrated energy system planning considering investment, retrofit and abandonment. In this paper, we turn off the retrofit and abandonment options. Therefore, the problem instances only have continuous variables.

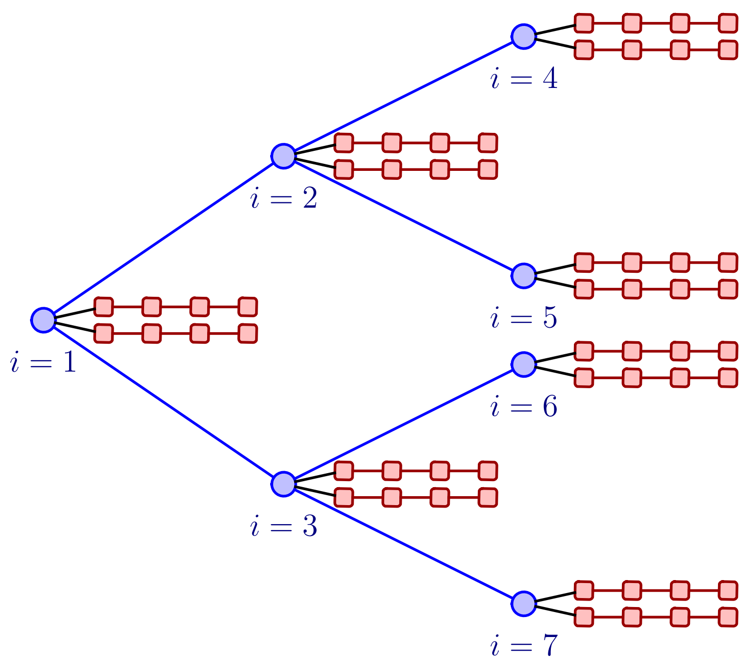
The contributions of this paper are the following: (1) it is the first paper formalising and proposing decomposition methods based on node formulation and scenario formulation of MHSP, (2) PLPR is proposed to utilise the special structure of MHSP to potentially reduce computational time significantly, and (3) the performance of PLPR Benders decomposition, Lagrangean decomposition is compared and analysed.

The outline of the paper is as follows: Section 2 first introduces a node formulation of MHSP and formalises that it can be decomposed by Benders type algorithms. Section 3 provides a scenario formulation of MHSP, and shows that MHSP can be decomposed by Lagrangean decomposition. Section 4 proposes the PLPR algorithm. Section 5 presents the stochastic investment planning model used in the case study. Section 6 reports the computational results and numerical analysis. Section 7 discusses the implications of the method and results and summaries the limitations of the research. Section 8 concludes the paper and suggests further research.

2. Benders decomposition

In this section, we first describe a general node formulation of MHSP, and show that it can be decomposed using Benders decomposition. We then explain that due to the special structure of MHSP, standard Benders can be directly applied for solving multi-stage stochastic programming. Traditionally, multi-stage stochastic programming is usually solved using nested Benders decomposition (Birge, 1985).

Figure 1: Illustration of node formulation for MHSP (blue circles: strategic nodes, red squares: operational periods)



When formulating MHSP using a node formulation, the non-anticipativity constraints are not expressed explicitly. We denote the strategic decision nodes by $i \in \mathcal{I}$, and the set of strategic decision nodes j that are ancestors to a decision node i by $\mathcal{I}_i$. The $\mathcal{S}_i$ denotes the set of operational scenarios that are embedded in strategic node i. The set of operational stages is represented by $\mathcal{T}_i$. The superscripts indicate the type of nodes that vectors and matrices belong to. The subscripts are the indices. The x_i are the strategic decision variables, and y_{its} are the operational variables. The deterministic equivalent of the linear programming MHSP is defined as a full master problem given by Equations (1).

$$\min_{x_i, y_{its}} \sum_{i \in \mathcal{I}} \pi_i \left(c_i^\top x_i + \sum_{s \in \mathcal{S}_i} \omega_{is} \sum_{t \in \mathcal{T}_i} q_{its}^\top Q_{its} y_{its} \right) \tag{1a}$$

$$\text{s.t.} \quad T_j^I x_j + W_i^I x_i \leq h_i^I, \qquad i \in \mathcal{I} \setminus \{1\}, j \in \mathcal{I}_i, \tag{1b}$$

$$T^0 x_i \leq h^0, \qquad i = 1, \tag{1c}$$

$$T_{its}^O x_i + W_{its}^O y_{its} \leq h_{its}^O, \qquad i \in \mathcal{I}, t = 1, s \in \mathcal{S}_i, \tag{1d}$$

$$T_{its}^O y_{i(t-1)s} + W_{its}^O y_{its} \leq h_{its}^O, \qquad i \in \mathcal{I}, t \in \mathcal{T}_i \setminus \{1\}, s \in \mathcal{S}_i, \tag{1e}$$

$$x \geq 0, y \geq 0, \tag{1f}$$

where x and y include all variables x_i and y_{its}, and where π_i is the probability of strategic node i, sum of π_i in each strategic stage is equal to 1, $c_i \in \mathbb{R}^{n_i}$, $h_i^I \in \mathbb{R}^{m_i}$, $W_i^I \in \mathbb{R}^{m_i \times n_i}$, are vectors and matrices at strategic node $i \in \mathcal{I}$, and $T_j^I \in \mathbb{R}^{m_i \times n_j}$ is the matrix for its ancestor nodes $j \in \mathcal{I}_i$. We assume that if $i = 1$, $T_j^I = T^0$, $W_i^I = 0$, and $h_i^I = h^0$. The probability of operational scenario s that is embedded in strategic node i is denoted by ω_{is}, and $\sum_{s \in \mathcal{S}_i} \omega_{is} = 1$. Operational vectors and matrices at operational node i, in operational scenario s, operational stage t are given by $T_{its}^O \in \mathbb{R}^{m_{it} \times n_{it}}$, $W_{its}^O \in \mathbb{R}^{m_{it} \times n_{it}}$, $q_{its} \in \mathbb{R}^{n_{it}}$, $h_{its}^O \in \mathbb{R}^{n_{it}}$. For operational stage $t = 1$, we have $T_{i1s}^O \in \mathbb{R}^{m_{i1} \times n_i}$. Equations (1) provide a general mathematical formulation for MHSP.

By fixing the complicating variable x_i, we can decompose the full size problem using Benders decomposition. The Benders reduced master problem is as follows,

$$\min_{x_i} \sum_{i \in \mathcal{I}} \pi_i (c_i^\top x_i + \beta_i) \tag{2a}$$

$$\text{s.t.} \quad T_j^I x_j + W_i^I x_i \leq h_i^I, \qquad i \in \mathcal{I} \setminus \{1\}, j \in \mathcal{I}_i, \tag{2b}$$

$$T^0 x_i \leq h^0, \qquad i = 1, \tag{2c}$$

$$\beta_i \geq \theta + \lambda^\top (x_i - x), \qquad (x, \theta, \lambda) \in \mathcal{F}_{i(j-1)}, i \in \mathcal{I}, \tag{2d}$$

$$x \geq 0, \tag{2e}$$

where Constraint (2e) are the projected cuts added to the Benders reduced master problem until iteration $j-1$, β_i is a variable for the approximated cost of the operational problem that is embedded in strategic node i. The set of cutting planes associated with subproblem i built up to iteration $j-1$ is denoted by $\mathcal{F}_{i(j-1)}$. The θ collects the optimal objective value of each subproblem i until iteration $j-1$. The subgradient w.r.t. x_{ij} until iteration $j-1$ is collected by λ. The sampled points until iteration $j-1$ are denoted by x.

For a given node i, the Benders subproblem is formulated as

$$\min_{y_{its}} \pi_i \sum_{s \in \mathcal{S}_i} \omega_{is} q_{its}^\top Q_{its} y_{its} \tag{3a}$$

$$
\begin{aligned}
\text{s.t. } & T_{its}^{O} x_i + W_{its}^{O} y_{its} \leq h_{its}^{O}, && i \in \mathcal{I}, t = 1, s \in \mathcal{S}_i, && \text{(3b)}\\
& T_{its}^{O} y_{i(t-1)s} + W_{its}^{O} y_{its} \leq h_{its}^{O}, && i \in \mathcal{I}, t \in \mathcal{T}_i \setminus \{1\}, s \in \mathcal{S}_i, && \text{(3c)}\\
& y \geq 0, && && \text{(3d)}
\end{aligned}
$$

and the Benders subproblems can be solved in parallel.

Traditionally, a stochastic linear program with multiple stages is formulated as a multi-stage stochastic program (Birge & Louveaux, 2011), and then such a problem can be decomposed and solved using nested Benders decomposition. Here we show that by exploiting the special structure of MHSP, we can decompose the problem using classic Benders decomposition (Benders, 1962) to solve multi-stage stochastic programs. In the Benders reduced master problem, we solve for all strategic nodes, and the operational problems are the Benders subproblems. In addition, if W_{ist}^{O} is the same in all nodes, and the operational problem has certain properties, one can improve Benders decomposition by avoiding solving all operational problems at each iteration, such as the adaptive Benders decomposition (Mazzi et al., 2020; Zhang et al., 2022). These approaches also utilise the property of MHSP that Benders subproblems are independent.

The standard Benders decomposition is presented in Algorithm 1.

Algorithm 1 Benders decomposition

```
1: choose ϵ (convergence tolerance), U*_0 := +∞ (initial upper bound), j := 0, F_i0 := {(β_i0, 0, 0)}
   for each i ∈ I;
2: repeat
3:     j := j + 1;
4:     solve the Benders reduced master problem and obtain β_ij and x_ij^RMP; L*_j :=
       Σ_{i∈I} π_i (c_i^⊤ x_ij^RMP + β_ij);
5:     for i ∈ I do
6:         solve Benders subproblem i at (x_ij^RMP, c_i) and obtain θ_ij and λ_ij;
7:     end for
8:     U*_j := min(U*_{j−1}, Σ_{i∈I} π_i (c_i^⊤ x_ij^RMP + θ_ij));
9:     for i ∈ I do
10:        F_ij := F_i(j−1) ∪ {(x_ij^RMP, θ_ij, λ_ij)};
11:    end for
12: until U*_j − L*_j ≤ ϵ.
```

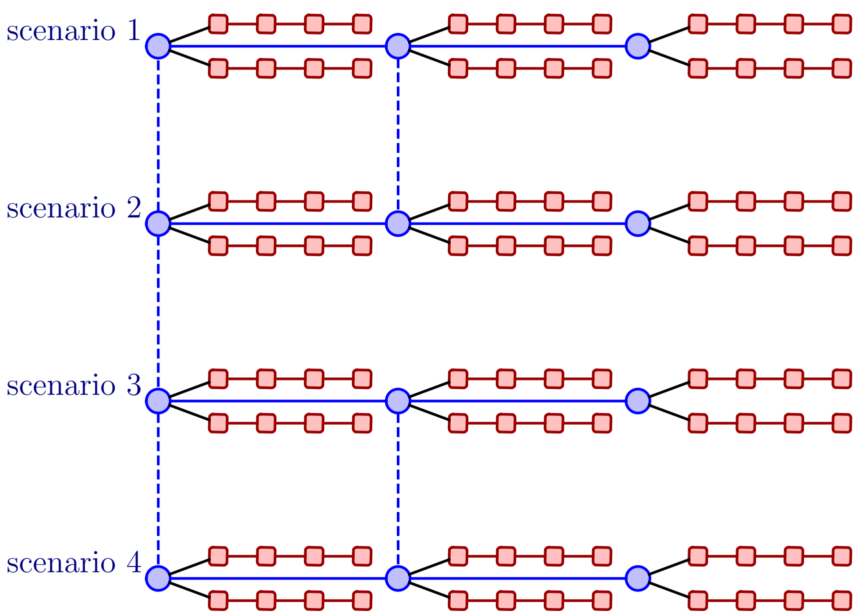
Remark 1. *The special structure of MHSP enables to the application of standard two-stage Benders to solve a multi-stage stochastic programming problem.*

3. Lagrangean decomposition

MHSP can also be formulated in a scenario based formulation. It can then be decomposed by Lagrangean decomposition. When the problem is large, Benders decomposition may have a larger and more ill-conditioned master problem and be hard to converge. In such a case, Lagrangean decomposition may be preferred.

3.1. Scenario formulation for MHSP

Figure 2: Illustration of scenario formulation for MHSP (blue circles: strategic nodes, red squares: operational periods). The blue dashed lines represent the NAC.



Here, we present a scenario formulation for MHSP. We denote the set of strategic stages by $h \in \mathcal{H}$, and the set of strategic scenarios by $i \in \mathcal{S}^I$. The set of operational scenarios is denoted by $s \in \mathcal{S}^O_{hi}$, and the set of operational stages is denoted by $t \in \mathcal{T}^O_{hi}$. We define set $\mathcal{J} := \{(h,i,i') : h \in \mathcal{H}, i, i' \in \mathcal{S}^I, i \text{ and } i' \text{ are indistinguishable in strategic stage } h\}$ for formulating the Non-Anticipativity Constraint (NAC). Variables x_{hi} and y_{hits} are the investment and operational variables respectively. The mathematical formulation of the full size problem is given as follows,

$$\min_{x_{hi}, y_{hits}} \sum_{i \in \mathcal{I}} \pi_i \left(\sum_{h \in \mathcal{H}} \left(c_{hi}^\top x_{hi} + \sum_{s \in \mathcal{S}^O_{hi}} \omega_{his} \sum_{t \in \mathcal{T}^O_{hi}} q_{hits}^\top Q_{hits} y_{hits} \right) \right) \tag{4a}$$

$$\text{s.t.} \quad T^I_{(h-1)i} x_{(h-1)i} + W^I_{hi} x_{hi} \le h^I_{hi}, \qquad h \in \mathcal{H} \setminus \{1\}, i \in \mathcal{S}^I, \tag{4b}$$

$$T^0 x_{hi} \le h^0, \qquad h = 1, i \in \mathcal{S}^I, \tag{4c}$$

$$T^O_{hits} x_i + W^O_{hits} y_{hits} \le h^O_{hits}, \qquad h \in \mathcal{H}, i \in \mathcal{S}^I, t = 1, s \in \mathcal{S}^O_{hi}, \tag{4d}$$

$$T^O_{hits} y_{hi(t-1)s} + W^O_{hits} y_{hits} \le h^O_{hits}, \qquad h \in \mathcal{H}, i \in \mathcal{S}^I, t \in \mathcal{T}^O_{hi} \setminus \{1\}, s \in \mathcal{S}^O_{hi}, \tag{4e}$$

$$x_{hi} = x_{hi'}, \qquad (h,i,i') \in \mathcal{J}, \tag{4f}$$

$$x \ge 0, y \ge 0, \tag{4g}$$

where x and y include all variables x_{hi} and y_{hits}, and where π_i is the probability of strategic scenario

i, sum of π_i is equal to 1, $c_{hi} \in \mathbb{R}^{n_{hi}}$, $h^I_{hi} \in \mathbb{R}^{m_{hi}}$, $W^I_i \in \mathbb{R}^{m_{hi} \times n_{hi}}$, are vectors and matrices for strategic stage $h \in \mathcal{H}$, scenario $i \in \mathcal{S}^I$, and $T^I_{(h-1)i} \in \mathbb{R}^{m_{hi} \times n_{hj}}$ is the matrix for its previous stage. We assume that if $h = 1$, $T^I_{hi} = T^0$, $W^I_{hi} = 0$, and $h^I_{hi} = h^0$. The probability of operational scenario s is denoted by ω_{his}, and $\sum_{s \in \mathcal{S}^O_{hi}} \omega_{his} = 1$. Operational vectors and matrices at operational node i, in operational scenario s, operational stage t are given by $T^O_{hits} \in \mathbb{R}^{m_{it} \times n_{it}}$, $W^O_{hits} \in \mathbb{R}^{m_{hits} \times n_{it}}$, $q_{hits} \in \mathbb{R}^{n_{it}}$, $h^O_{hits} \in \mathbb{R}^{n_{it}}$. For operational stage $t = 1$, we have $T^O_{hi1s} \in \mathbb{R}^{m_{i1} \times n_i}$. Equations (4) correspond to a general scenario based mathematical formulation for MHSP. Equation (4f) is the NAC. Note that due to the properties of the MHSP, operational decisions are independent of future strategic scenarios, and the operational decision variables are embedded in the strategic node. Therefore, NAC is not needed for operational decisions. We denote the full size scenario based formulation, Equations (4), by Lagrangean master problem. The NACs, Equation (4f), are the complicating constraints that link the scenarios.

By relaxing Equation (4f), one can obtain the Lagrangean dual. The problem, given by the Equations (4), is then decomposed by scenarios. The Lagrangean dual is as follows,

$$\min_{x_{hi}, y_{hits}} \sum_{i \in \mathcal{I}} \pi_i \left(\sum_{h \in \mathcal{H}} \left(c_{hi}^\top x_{hi} + \sum_{s \in \mathcal{S}^O_{hi}} \omega_{his} \sum_{t \in \mathcal{T}^O_{hi}} q_{hits}^\top Q_{hits} y_{hits}) \right) \right) - \sum_{(h,i,i') \in \mathcal{J}} \lambda_{hii'}^\top (x_{hi} - x_{hi'}) \tag{5a}$$

$$\text{s.t.} \quad T^I_{(h-1)i} x_{(h-1)i} + W^I_{hi} x_{hi} \leq h^I_{hi}, \qquad h \in \mathcal{H} \setminus \{1\}, i \in \mathcal{S}^I, \tag{5b}$$

$$T^0 x_{hi} \leq h^0, \qquad h = 1, i \in \mathcal{S}^I, \tag{5c}$$

$$T^O_{hits} x_i + W^O_{hits} y_{hits} \leq h^O_{hits}, \qquad h \in \mathcal{H}, i \in \mathcal{S}^I, t = 1, s \in \mathcal{S}^O_{hi}, \tag{5d}$$

$$T^O_{hist} y_{hi(t-1)s} + W^O_{hits} y_{hits} \leq h^O_{hits}, \qquad h \in \mathcal{H}, i \in \mathcal{S}^I, t \in \mathcal{T}^O_{hi} \setminus \{1\}, s \in \mathcal{S}^O_{hi}, \tag{5e}$$

$$x \geq 0, y \geq 0, \tag{5f}$$

where $\lambda_{hii'}$ is the Lagrangean multiplier. The Lagrangean dual can be solved per scenario $i \in \mathcal{S}^I$ in parallel. We use the subgradient method to update the Lagrangean multiplier. The Lagrangean decomposition algorithm is presented in Algorithm 2. For simplifying the notation, we denote the objective value of each Lagrangean subproblem i at iteration j as θ_{ij}.

4. PLPR

This section proposes PLPR, where the subproblems are solved in parallel, and describes the steps that are different from standard Lagrangean decomposition.

First, we use parallel computing when the Lagrangean subproblems are solved. Scenario subproblems are equally distributed to the available processes. If the number of scenarios cannot be divided by the number of processes, some processes will have more scenario subproblems.

Second, we propose a primal reduction step, which is to reduce the size of Equations (4) and parallelise the solution process. In Lagrangean decomposition, to obtain an upper bound, one needs

Algorithm 2 Standard Lagrangean decomposition

choose ϵ (convergence tolerance), $\delta_0 \in (0, 2]$ (initial correction term), $\underline{\gamma} \in (0, 1)$, $\gamma_0 \in (\underline{\gamma}, 1)$, $\overline{\gamma} \in (1, +\infty)$, $U_0^* := +\infty$ (initial upper bound), $L_0^* := -\infty$ (initial lower bound), $j := 0$, $\lambda_{hii'0=0}$;

repeat

$j := j + 1$;

for $i \in \mathcal{S}^I$ **do**

solve the Lagrangean subproblem i, and obtain θ_{ij} and x_{hij};

end for

$L_j := \sum_{i \in \mathcal{S}^I} \theta_{ij}$ and $L_j^* := \max\left(L_{j-1}^*, L_j\right)$;

use bottleneck strategy to construct a feasible x_{hij}^*;

$x_{hij} := x_{hij}$, solve the Lagrangean master problem and obtain U_j;

$U_j^* := \min(U_{j-1}^*, U_j)$;

if $L_j \leq L_{j-1}^*$ **then** $\delta_j := \underline{\gamma}\theta_{j-1}$

else if $x_{hij} \cdot x_{hi(j-1)} < 0$ **then** $\delta_j := \gamma_0\theta_{j-1}$

else $\delta_j := \overline{\gamma}\theta_{j-1}$

end if

for $(h, i, i') \in \mathcal{J}$ **do**

$\lambda_{hii'(j+1)} := \lambda_{hii'j} + \delta_j \frac{\left(U_j^* - L_j^*\right)}{\left\|x_{hij} - x_{hi'j}\right\|^2} \left(x_{hij} - x_{hi'j}\right)$;

end for

until $U_j^* - L_j^* \leq \epsilon$.

to construct a feasible solution from the relaxed solution and solve the original problem, Equations (4) with that fixed investment solution x_{hi}. When the original problem is large, the full size problem may still be hard to solve after fixing some variables.

Once the variable x_{hi} is fixed, the investment-related costs can be directly calculated. Furthermore, the operational problems become parallelisable once x_{hi} is fixed. In addition, not all operational problems need to be solved because once the NAC is restored, some operational problems become exactly equivalent to each other. Therefore, the number of operational problems that need to be solved is theoretically reduced to $|\mathcal{I}|$, where $\mathcal{I}$ is the set of nodes in the node formulation.

Assuming a feasible strategic solution x_{hij} is given as parameters, the primal problem becomes a group of independent operational problems. Each problem is indexed by strategic stage $h \in \mathcal{H}$ and strategic scenario $i \in \mathcal{S}^{RI}$, where $\mathcal{S}^{RI}$ is the reduced set of scenarios. For a given problem $h \in \mathcal{H}, i \in \mathcal{S}^{RI}$, we define the corresponding subproblem as a node subproblem. The formulation of the node subproblem is given as follows,

$$\min_{y_{hits} \in \mathcal{Y}} \pi_i \left(c_{hi}^\top x_{hi} + \sum_{s \in \mathcal{S}_{hi}^O} \omega_{his} \sum_{t \in \mathcal{T}_{hi}^O} q_{hits}^\top Q_{hits} y_{hits} \right) \tag{6a}$$

$$\text{s.t.} \quad T_{hits}^O x_i + W_{hits}^O y_{hits} \leq h_{hits}^O, \qquad h \in \mathcal{H}, i \in \mathcal{S}^I, t = 1, s \in \mathcal{S}_{hi}^O, \tag{6b}$$

$$T_{hits}^O y_{hi(t-1)s} + W_{hits}^O y_{hits} \leq h_{hits}^O, \qquad h \in \mathcal{H}, i \in \mathcal{S}^I, t \in \mathcal{T}_{hi}^O \setminus \{1\}, s \in \mathcal{S}_{hi}^O, \tag{6c}$$

$$y \geq 0. \tag{6d}$$

A Lagrangean upper bound can be obtained after solving all the node subproblems. Here, $c_{hi}^{\top} x_{hi}$ becomes a constant in the objective function. This reduction can produce an exact upper bound because the structure of MHSP makes the operational problem only depend on its investment decisions. The computational time can be significantly reduced by reducing the size of the primal problem and parallelising the solving process. The PLPR is presented in Algorithm 3.

Algorithm 3 PLPR

choose ϵ (convergence tolerance), $\delta_0 \in (0, 2]$ (initial correction term), $\underline{\gamma} \in (0, 1)$, $\gamma_0 \in (\underline{\gamma}, 1)$, $\overline{\gamma} \in (1, +\infty)$, $U_0^* := +\infty$ (initial upper bound), $L_0^* := -\infty$ (initial lower bound), $j := 0$, $\lambda_{hii'0}$=0;
repeat
  $j := j + 1$;
  for $i \in \mathcal{S}^I$ **do**
    assign Lagrangean subproblem i to a computer node, solve it, and obtain θ_{ij} and x_{hij};
  end for
  use bottleneck strategy to construct a feasible x_{hij}^*;
  $L_j := \sum_{i \in \mathcal{S}^I} \theta_{ij}$ and $L_j^* := \max\left(L_{j-1}^*, L_j\right)$;
  for $h \in \mathcal{H}, i \in \mathcal{S}^{RI}$ **do**
    $x_{hij} := x_{hij}^*$, assign the node subproblem hi to a computer node and solve it;
  end for
  obtain $U_j := \sum_{i \in \mathcal{I}} \pi_i \sum_{h \in \mathcal{H}} c_{hi}^{\top} x_{hi} + \sum_{h \in \mathcal{H}, i \in \mathcal{S}^I} \theta_{hij}^{SSP}$;
  set $U_j^* := \min(U_{j-1}^*, U_j)$;
  if $L_j \leq L_{j-1}^*$ **then** $\delta_j := \underline{\gamma} \theta_{j-1}$
  else if $x_{hij} \cdot x_{hi(j-1)} < 0$ **then** $\delta_j := \gamma_0 \theta_{j-1}$
  else $\delta_j := \overline{\gamma} \theta_{j-1}$
  end if
  for $(h, i, i') \in \mathcal{J}$ **do**
    $\lambda_{hii'(j+1)} := \lambda_{hii'j} + \delta_j \frac{(U_j^* - L_j^*)}{\left\| x_{hij} - x_{hi'j} \right\|^2} (x_{hij} - x_{hi'j})$;
  end for
until $U_j^* - L_j^* \leq \epsilon$.

Remark 2. *There is no dedicated NAC constraint for operational decision variables in MHSP because the operational decision variables are embedded in the investment node. This leads to fewer Lagrangean multipliers and a simpler search for a feasible solution.*

5. Mathematical model

This section presents an MHSP model for a power system investment and operational planning problem adapted from (Zhang et al., 2023, 2022). Here, we focus on how the proposed algorithms fit the mathematical model. We use the conventions that calligraphic capitalised Roman letters denote sets, upper case Roman and lower case Greek letters denote parameters, and lower case Roman letters denote variables. The indices are subscripts, and name extensions are superscripts. The names of variables, parameters, sets and indices are single symbols.

5.1. Nomenclature

Investment planning model sets

$\mathcal{P}$	set of technologies, p
$\mathcal{I}$	set of operational nodes, i
$\mathcal{I}_0$	set of investment nodes, i_0
$\mathcal{I}_i$	set of investment nodes i_0 ($i_0 \in \mathcal{I}_0$) ancestor to operational node i ($i \in \mathcal{I}$)
$\mathcal{H}$	set of investment stages, h
$\mathcal{H}_h$	set of all earlier investment stages of stage h ($h \in \mathcal{H}$), h
$\mathcal{S}^I$	set of investment scenarios, i

Operational model sets

$\mathcal{N}$	set of time slices, n
$\mathcal{T}$	set of hours in all time slices, t
$\mathcal{T}_n$	set of hours in time slice n
$\mathcal{G}$	set of thermal generators, g
$\mathcal{S}$	set of electricity storage, s
$\mathcal{R}$	set of renewable generations, r

Investment planning model parameters

$C^{Inv}_{pi/phi}$	unitary investment cost of device p in investment node i/ in stage h scenario i ($p \in \mathcal{P}, i \in \mathcal{I}_0, h \in \mathcal{H}, i \in \mathcal{S}^I$) [€/MW]
$C^{Fix}_{pi/phi}$	unitary fix operational and maintenance cost of device p in node i/ in stage h scenario i ($p \in \mathcal{P}, i \in \mathcal{I}, h \in \mathcal{H}, i \in \mathcal{S}^I$) [€/MW]
$X^{Hist}_{pi/ph}$	historical capacity of device p in investment node i/ in stage h ($p \in \mathcal{P}, i \in \mathcal{I}, h \in \mathcal{H}$) [MW]
X^{Max}_p	maximum installed capacity of device p ($p \in \mathcal{P}$) [MW]
κ	scaling effect depending on the number of operation years between investment nodes/ investment stages
$\delta^{I_0}_i/\delta^I_i/\delta^H_h$	discount factor of investment node i ($i_0 \in \mathcal{I}_0$)/ operational node i ($i \in \mathcal{I}$)/ stage ($h \in \mathcal{H}$)
$\pi^{I_0}_i/\pi^I_i/\pi^{SI}_i$	probability of investment node i ($i_0 \in \mathcal{I}_0$)/ operational node i ($i \in \mathcal{I}$)/ scenario i ($i \in \mathcal{S}^I$)
H^P_p	life time of technology p ($p \in \mathcal{P}$)
x_i	right hand side coefficients of the Benders operational subproblem
c_i	cost coefficients of the Benders operational subproblem
$\mu^E_{i/hi}$	CO_2 budget at operational node i/ in stage h scenario i ($i \in \mathcal{I}, h \in \mathcal{H}, i \in \mathcal{S}^I$)
$\mu^{DP}_{i/hi}$	scaling factor on power demand at operational node i/ in stage h scenario i ($i \in \mathcal{I}, h \in \mathcal{H}, i \in \mathcal{S}^I$)
I^L_i	planning stage of a node i ($i \in \mathcal{I} \cup \mathcal{I}_0$)
$C^{CO2}_{i/hi}$	CO_2 emission price at operational node i/ in stage h scenario i ($i \in \mathcal{I}, h \in \mathcal{H}, i \in \mathcal{S}^I$)

Operational model parameters

W_t	probability multiplied weight of operation period t ($t \in \mathcal{T}$)
H_t	number of hour(s) in one operational period t ($t \in \mathcal{T}$)

α_g^G	maximum ramp rate of gas turbines ($g \in \mathcal{G}$) [MW/MW]
R_{rt}^R	capacity factor of renewable unit r in period t ($r \in \mathcal{R}, t \in \mathcal{T}$)
η_s^{SE}	efficiency of electricity store s ($s \in \mathcal{S}$)
E_g^G	emission factor of gas turbine g ($g \in \mathcal{G}$) [tonne/MWh]
C_g^G/C_s^{SE}	total operational cost of a generator g/ a storage facility s ($g \in \mathcal{G}$/ $s \in \mathcal{S}$) [€/MW]
C^{ShedP}	load shed penalty cost [€/MWh]
P_t^{DP}	power demand in period t ($z \in \mathcal{Z}, t \in \mathcal{T}$) [MW]

Investment planning model variables

c_i^{OPE}	estimated operational cost in operational node i ($i \in \mathcal{I}$) (€)
$x_{pi/phi}^{Acc}$	accumulated capacity of device p in operational node i/ in stage h scenario i ($p \in \mathcal{P}, i \in \mathcal{I}, h \in \mathcal{H}, i \in \mathcal{S}^I$) [MW]
$x_{pi/phi}^{Inst}$	newly invested capacity of device p in investment node i_0/ in stage h scenario i ($p \in \mathcal{P}, i \in \mathcal{I}_0, h \in \mathcal{H}, i \in \mathcal{S}^I$) [MW]
c^{INV}	total expected investment cost (€)

Operational model variables

p_{gi}^{AccG}	accumulated capacity of gas turbine g in operational node i ($g \in \mathcal{G}, i \in \mathcal{I}$) [MW]
p_{ri}^{AccR}	accumulated capacity of renewable unit r in operational node i ($r \in \mathcal{R}, i \in \mathcal{I}$) [MW]
p_{si}^{AccSE}	accumulated charging/discharging capacity of electricity store s in operational node i ($s \in \mathcal{S}, i \in \mathcal{I}$) [MW]
$p_{git/ghit}^{G}$	power generation of gas turbine g in operational node i in period t/ in stage h scenario i period t ($g \in \mathcal{G}, i \in \mathcal{I}, h \in \mathcal{H}, i \in \mathcal{S}^I, t \in \mathcal{T}$) [MW]
$p_{sit/shit}^{SE+}$	charge power of electricity store s in operational node i/ in stage h scenario i period t ($s \in \mathcal{S}, i \in \mathcal{I}, h \in \mathcal{H}, i \in \mathcal{S}^I, t \in \mathcal{T}$) [MW]
$p_{sit/shit}^{SE-}$	discharge power of electricity store s in operational node i/ stage h scenario i in period t($s \in \mathcal{S}, i \in \mathcal{I}, h \in \mathcal{H}, i \in \mathcal{S}^I, t \in \mathcal{T}$) [MW]
$p_{it/hit}^{GShedP}$	generation shed in operational node i/ in stage h scenario i period t ($i \in \mathcal{I}, h \in \mathcal{H}, i \in \mathcal{S}^I, t \in \mathcal{T}$) [MW]
$q_{sit/shit}^{SE}$	energy level of electricity store s in operational node i/ in stage h scenario i at the start of period t ($s \in \mathcal{S}, i \in \mathcal{I}, h \in \mathcal{H}, i \in \mathcal{S}^I, t \in \mathcal{T}$) [MWh]
$p_{it/hit}^{ShedP}$	load shed in operational node i/ in stage h scenario i in period t ($i \in \mathcal{I}, h \in \mathcal{H}, i \in \mathcal{S}^I, t \in \mathcal{T}$) [MW]

5.2. Investment planning model (Benders master problem)

$$\min c^{INV} + \kappa \sum_{i \in \mathcal{I}} \pi_i^I c_i^{OPE} \tag{7a}$$

$$\text{s.t.} \quad c^{INV} = \sum_{i \in \mathcal{I}_0} \delta_i^{I_0} \pi_i^{I_0} \sum_{p \in \mathcal{P}} C_{pi}^{Inv} x_{pi}^{Inst} + \kappa \sum_{i \in \mathcal{I}} \delta_i^I \pi_i^I \sum_{p \in \mathcal{P}} C_{pi}^{Fix} x_{pi}^{Acc} \tag{7b}$$

$$x_{pi}^{Acc} = X_{pi}^{Hist} + \sum_{i_0 \in \mathcal{I}_i | \kappa(I_i^L - I_{i_0}^L) \leq H_p^P} x_{pi}^{Inst}, \qquad p \in \mathcal{P}, i \in \mathcal{I}, \tag{7c}$$

$$x_{pi}^{Acc} \leq X_p^{Max}, \qquad p \in \mathcal{P}, i \in \mathcal{I}, \tag{7d}$$

$$c_i^{OPE} \geq \theta + \lambda^\top (x_i - x) \qquad (x, \theta, \lambda) \in \mathcal{F}_{i(j-1)}, i \in \mathcal{I}, \tag{7e}$$

$$x_{pi}^{Inst}, x_{pi}^{Acc} \in \mathbb{R}_0^+. \tag{7f}$$

The total cost for investment planning, Equation (7a), consists of actual discounted investment costs and discounted fixed operating and maintenance costs c^{INV}, as well as the expected operational cost of the system over the time horizon $\kappa \sum_{i \in \mathcal{I}} \pi_i^I c^{OPE}$. Here, κ is a scaling factor that depends on the time step between two successive investment nodes. Constraint (7c) states that the accumulated capacity of a technology x_{pi}^{Acc} in an operational node equals the sum of the historical capacity X_p^{Hist} and newly invested capacities x_{pi}^{Inst} in its ancestor investment nodes $\mathcal{I}_i$. The parameter X_p^{Max} denotes the maximum accumulated capacity of technologies. We define $x_i = \left(\{x_{pi}^{Acc}, p \in \mathcal{P}\}, \mu_i^{DP}, \mu_i^E\right), i \in \mathcal{I}$ that collects all right hand side coefficients, and will be fixed in the Benders subproblem (8) through vector x_i. Also, $c_i = \left(C_i^{CO_2}\right), i \in \mathcal{I}$ collects all the cost coefficients into vector c_i. The investment planning model Equations (7) corresponds to the Benders reduced master problem Equations (2).

5.3. Operational model (Benders subproblem)

We now compute the operational cost $c^{OPE}(x_i, c_i)$ at one operational node $i \in \mathcal{I}$ by solving Benders subproblem (8) given the decisions x_i and c_i determined in the master problem (7).

$$\min \quad \sum_{t \in \mathcal{T}} W_t H_t \left(\sum_{g \in \mathcal{G}} C_g^G p_{git}^G + \sum_{s \in \mathcal{S}} C_s^{SE} p_{sit}^{SE+} + C^{ShedP} p_{it}^{ShedP} \right) \tag{8a}$$

$$\text{s.t.} \quad p_{git}^G \leq p_{gi}^{AccG}, \qquad g \in \mathcal{G}, t \in \mathcal{T}, \tag{8b}$$

$$p_{sit}^{SE+} \leq p_{si}^{AccSE}, \qquad s \in \mathcal{S}, t \in \mathcal{T}, \tag{8c}$$

$$p_{sit}^{SE-} \leq p_{si}^{AccSE}, \qquad s \in \mathcal{S}, t \in \mathcal{T}, \tag{8d}$$

$$q_{sit}^{SE} \leq \gamma_s^{SE} p_{si}^{AccSE}, \qquad s \in \mathcal{S}, t \in \mathcal{T}, \tag{8e}$$

$$-\alpha_g^G p_{gi}^{AccG} \leq p_{git}^G - p_{gi(t-1)}^G \leq \alpha_g^G p_{gi}^{AccG}, \qquad g \in \mathcal{G}, n \in \mathcal{N}, t \in \mathcal{T}_n, \tag{8f}$$

$$\sum_{g \in \mathcal{G}} p_{git}^G + \sum_{s \in \mathcal{S}} p_{sit}^{SE-} + \sum_{r \in \mathcal{R}} R_{rit}^R p_r^{Acc} + p_{it}^{ShedP} = \mu_i^{DP} P_t^{DP} + \sum_{s \in \mathcal{S}_z} p_{sit}^{SE+} + p_{zit}^{GShedP}, \qquad t \in \mathcal{T}, \tag{8g}$$

$$q_{si(t+1)}^{SE} = q_{sit}^{SE} + H_t(\eta_s^{SE} p_{sit}^{SE+} - p_{sit}^{SE-}), \qquad s \in \mathcal{S}, n \in \mathcal{N}, t \in \mathcal{T}_n, \tag{8h}$$

$$\sum_{t \in \mathcal{T}} \sum_{g \in \mathcal{G}} W_t H_t E_g^G p_{git}^G \leq \mu_i^E, \tag{8i}$$

$$p_{git}^G, p_{gi}^{AccG}, p_{it}^{ShedP}, p_{sit}^{SE+}, p_{sit}^{SE-}, p_{si}^{AccSE}, q_{sit}^{SE}, p_{ri}^{AccR}, p_{it}^{GShedP} \in \mathbb{R}_0^+. \tag{8j}$$

The operational subproblem corresponds to Benders subproblem Equations (3). The operational cost includes total operating costs of all generators and storage facilities $C_g^G p_{gt}^G + C_s^S p_{sit}^{SE+}$ and load shedding costs $C^{ShedP} p_{it}^{ShedP}$. The parameters C_g^G and C_s^{SE} include the variable operational cost of generators and storage. For thermal generators, C_g^G also includes the fuel cost and the CO_2 tax charged for the emissions of generators. Constraint (8f) captures how fast thermal generators can ramp up or down their power output, respectively. The parameter α_g^G is the maximum ramp rate of thermal generators. The power balance, Constraint (8g), ensures that in one operational period t, the sum of total power generation of thermal generators p_{git}^G, power discharged from all the electricity storage p_{sit}^{SE-}, renewable generation $R_t^R p_{ri}^{AccR}$, and load shed p_{it}^{ShedP} equals the sum of power demand $\mu_i^{DP} P_t^{DP}$, and power generation shed p_{it}^{GShedP}. The parameter R_{rt}^R is the capacity factor of a renewable unit that is a fraction of the nameplate capacity p_i^{AccR}. Constraint (8h) states that the state of charge q_{sit}^{SE} in period $t+1$ depends on the previous state of charge q_{sit}^{SE}, the charged power p_{sit}^{SE+} and discharged power p_{sit}^{SE-}. The parameter η_s^{SE} represent the charging efficiency. Constraint (8i) limits the total emission. The parameter H_t is the length of the period t. The symbol E_g^G is the emission factor per unit of power generated. The capacities p_{gi}^{AccG}, p_{si}^{AccSE}, scaling factor of demand μ_i^{DP} and CO_2 budget μ_i^E are passed from the master problem (7) via vector x_i and CO_2 tax that is included in cost coefficient C_g^G is passed from master problem (7) via vector c_i.

5.4. Lagrangean subproblem

In Lagrangean decomposition, the subproblem corresponds to each scenario without the NAC constraint. Unlike Benders decomposition, which projects the operational decision onto the investment space, the Lagrangean subproblem keeps a copy of a part of the original problem. The Lagrangean subproblem for scenario $i \in \mathcal{S}^I$ is as follows:

$$\min \sum_{h\in\mathcal{H}} \delta_h^H \pi_i^{SI} \left(\sum_{p\in\mathcal{P}} \left(C_{phi}^{Inv} x_{pi}^{Inst} + \kappa C_{phi}^{Fix} x_{phi}^{Acc} \right) + \kappa \sum_{t\in\mathcal{T}} W_t H_t \left(\sum_{g\in\mathcal{G}} C_g^G p_{ghit}^G + \sum_{s\in\mathcal{S}} C_s^{SE} p_{shit}^{SE+} + C^{ShedP} p_{hit}^{ShedP} \right) \right) + \sum_{p\in\mathcal{P}} \sum_{h\in\mathcal{H}} \lambda_{phi} x_{phi}^{Inst} \tag{9a}$$

$$\text{s.t.} \quad x_{phi}^{Acc} = X_{phi}^{Hist} + \sum_{h_0\in\mathcal{H}_h | \kappa(h-h_0)\le H_p^P} x_{phi}^{Inst}, \qquad p\in\mathcal{P}, h\in\mathcal{H}, i\in\mathcal{S}^I, \tag{9b}$$

$$x_{phi}^{Acc} \le X_p^{Max}, \qquad p\in\mathcal{P}, h\in\mathcal{H}, \tag{9c}$$

$$p_{ghit}^G \le x_{ghi}^{Acc}, \qquad g\in\mathcal{G}, h\in\mathcal{H}, t\in\mathcal{T}, \tag{9d}$$

$$p_{shit}^{SE+} \le x_{shi}^{Acc}, \qquad s\in\mathcal{S}, h\in\mathcal{H}, t\in\mathcal{T}, \tag{9e}$$

$$p_{shit}^{SE-} \le x_{shi}^{Ac}, \qquad s\in\mathcal{S}, h\in\mathcal{H}, t\in\mathcal{T}, \tag{9f}$$

$$q_{shit}^{SE} \le \gamma_s^{SE} x_{shi}^{Acc}, \qquad s\in\mathcal{S}, h\in\mathcal{H}, t\in\mathcal{T}, \tag{9g}$$

$$-\alpha_g^G p_{ghi}^{AccG} \le p_{ghit}^G - p_{ghi(t-1)}^G \le \alpha_g^G p_{ghi}^{AccG}, \qquad g\in\mathcal{G}, h\in\mathcal{H}, n\in\mathcal{N}, t\in\mathcal{T}_n, \tag{9h}$$

$$\sum_{g\in\mathcal{G}} p_{ghit}^G + \sum_{s\in\mathcal{S}} p_{shit}^{SE-} + \sum_{r\in\mathcal{R}} R_{rt}^R p_{rhi}^{Acc} + p_{hit}^{ShedP} = \mu_{hi}^{DP} P_t^{DP} + \sum_{s\in\mathcal{S}} p_{shit}^{SE+} + p_{hit}^{GShedP}, \qquad h\in\mathcal{H}, t\in\mathcal{T}, \tag{9i}$$

$$q_{shi(t+1)}^{SE} = q_{shit}^{SE} + H_t(\eta_s^{SE} p_{shit}^{SE+} - p_{shit}^{SE-}), \qquad s\in\mathcal{S}, h\in\mathcal{H}, n\in\mathcal{N}, t\in\mathcal{T}_n, \tag{9j}$$

$$\sum_{t\in\mathcal{T}}\sum_{g\in\mathcal{G}} W_t H_t E_g^G p_{ghit}^G \le \mu_{hi}^E, \qquad h\in\mathcal{H}, \tag{9k}$$

$$x_{phti}^{Inst}, x_{phi}^{Acc}, p_{ghit}^G, p_{hit}^{ShedP}, p_{shit}^{SE+}, p_{shit}^{SE-}, q_{shit}^{SE}, p_{hit}^{GShedP} \in \mathbb{R}_0^+. \tag{9l}$$

Equations (9) correspond to the Lagrangean dual Equations (5). The objective function of the Lagrangean subproblem consists of the total investment and operational costs in all stages in scenario i, and the penalty term $\sum_{p\in\mathcal{P}}\sum_{h\in\mathcal{H}}\lambda_{phi}x_{phi}^{Inst}$ for the deviation from the NAC constraint, where λ_{phi} are the Lagrangean multipliers. The investment-related constraints Equations (9b)-(9c) are similar to Equations (7c)-(7d) in the Benders reduced master problem. Also, the operational constraints Equations (9d)-(9k) are similar to Equations (8b)-(8i) in the Benders subproblem. Therefore, we omit to explain all the constraints here.

The Lagrangean master problem in standard Lagrangean decomposition is simply the original full size problem Equations (4) with fixed investment decisions x_{hi}.

5.5. Lagrangean node subproblem

$$\min \kappa \sum_{t\in\mathcal{T}} W_t H_t \left(\sum_{g\in\mathcal{G}} C_g^G p_{ghit}^G + \sum_{s\in\mathcal{S}} C_s^{SE} p_{shit}^{SE+} + C^{ShedP} p_{hit}^{ShedP} \right) \tag{10a}$$

$$\text{s.t.}\quad p_{ghit}^G \le p_{ghi}^{Acc}, \qquad g\in\mathcal{G}, h\in\mathcal{H}, t\in\mathcal{T}, \tag{10b}$$

$$p_{shit}^{SE+} \le p_{shi}^{Acc}, \qquad s\in\mathcal{S}, h\in\mathcal{H}, t\in\mathcal{T}, \tag{10c}$$

$$p_{shit}^{SE-} \le p_{shi}^{Acc}, \qquad s\in\mathcal{S}, h\in\mathcal{H}, t\in\mathcal{T}, \tag{10d}$$

$$q_{shit}^{SE} \le \gamma_s^{SE} p_{shi}^{Acc}, \qquad s\in\mathcal{S}, h\in\mathcal{H}, t\in\mathcal{T}, \tag{10e}$$

$$-\alpha_g^G p_{ghi}^{AccG} \le p_{ghit}^G - p_{ghi(t-1)}^G \le \alpha_g^G p_{ghi}^{AccG}, \qquad g\in\mathcal{G}, h\in\mathcal{H}, n\in\mathcal{N}, t\in\mathcal{T}_n, \tag{10f}$$

$$\sum_{g\in\mathcal{G}} p_{ghit}^G + \sum_{s\in\mathcal{S}} p_{shit}^{SE-} + \sum_{r\in\mathcal{R}} R_{rt}^R p_{rhi}^{Acc} + p_{hit}^{ShedP} = \mu_{hi}^{DP} P_t^{DP} + \sum_{s\in\mathcal{S}_z} p_{shit}^{SE+} + p_{hit}^{GShedP}, \qquad h\in\mathcal{H}, t\in\mathcal{T}, \tag{10g}$$

$$q_{shi(t+1)}^{SE} = q_{shit}^{SE} + H_t(\eta_s^{SE} p_{shit}^{SE+} - p_{shit}^{SE-}), \qquad s\in\mathcal{S}, h\in\mathcal{H}, n\in\mathcal{N}, t\in\mathcal{T}_n, \tag{10h}$$

$$\sum_{t\in\mathcal{T}}\sum_{g\in\mathcal{G}} W_t H_t E_g^G p_{ghit}^G \le \mu_{hi}^E, \qquad h\in\mathcal{H}, \tag{10i}$$

$$x_{phti}^{Inst}, x_{phi}^{Acc}, p_{ghit}^G, p_{zhit}^{ShedP}, p_{shit}^{SE+}, p_{shit}^{SE-}, q_{shit}^{SE}, p_{hit}^{GShedP} \in \mathbb{R}_0^+. \tag{10j}$$

The Lagrangean node subproblem corresponds to the node subproblem Equations (6) in the general formulation of the PLPR. The Lagrangean node subproblem is similar to the Benders subproblem. Once a feasible investment solution is obtained, the investment-related costs can be directly calculated. Then a series of Lagrangean node subproblems are solved in parallel to obtain the total operational costs. Eventually, an upper bound can be calculated.

6. Results

In this section, we provide the case study and the computational results. We use the REORIENT model (Zhang et al., 2023) to solve a single region investment planning problem, and apply standard Benders, standard Lagrangean and PLPR to solve the problem instances. The performance of the methods is compared.

6.1. Case study

In the case study, we use the model to solve a UK power system expansion problem. The data can be found in (Zhang et al., 2023, 2022). We implemented the algorithms and model in Julia 1.8.2 using JuMP (Dunning et al., 2017), and solved with Gurobi (Gurobi Optimization, LLC, 2022). We ran the code on a computer cluster (25 computer nodes) with a 2x 3.6GHz 8 core Intel Xeon Gold 6244 CPU and 384 GB of RAM, running on CentOS Linux 7.9.2009. The cluster was shared by other users and had no resource allocation and queuing systems. Therefore, solution times may have been affected by interfering traffic during program executions.

6.1.1. Computational results

An overview of the case study tested in this paper is presented in Table 1. The different cases vary in the number of investment stages, long-term uncertainty, operational scenarios, and representative hours in the operational problem.

Table 1: Overview of the cases used in the computational study.

	Operational periods per short-term scenario	Short-term scenarios	Long-term scenarios	Number of decision nodes: Present	In 5 years	In 10 years	Total	Problem size (undecomposed): Variables	Constraints
Case 1	168	4	9	1	3	9	13	1.7×10^5	5.0×10^5
Case 2	720	4	9	1	3	9	13	7.3×10^5	2.1×10^6
Case 3	720	8	9	1	3	9	13	1.5×10^6	4.3×10^6
Case 4	2190	4	9	1	3	9	13	2.2×10^6	6.5×10^6
Case 5	168	4	25	1	5	25	31	4.7×10^5	1.4×10^6
Case 6	720	4	25	1	5	25	31	1.6×10^6	4.8×10^6
Case 7	720	8	25	1	5	25	31	4.0×10^6	1.2×10^7
Case 8	2190	4	25	1	5	25	31	6.1×10^6	1.8×10^7

Table 2: Comparative results for standard Benders, standard Lagrangean, and PLPR.

	Standard Benders		Standard Lagrangean		PLPR	
	Iters	Time (s)	Iters	Time (s)	Iters	Time (s)
Case 1	8	10	9	81	8	6
Case 2	9	91	18	1408	8	75
Case 3	9	268	-	-	3	80
Case 4	9	496	-	-	2	54
Case 5	8	25	-	-	16	11
Case 6	9	144	-	-	5	107
Case 7	9	728	-	-	6	149
Case 8	9	1484	-	-	5	281

-: the algorithm cannot solve the test instance in a reasonable amount of time.

The computational time is reported in Table 2. We can see that standard Lagrangean decomposition is much worse than standard Benders decomposition and PLPR. This is because, in standard Lagrangean, both the Lagrangean subproblems and the Lagrangean master problem are solved in series. In addition, the original full size problem after fixing investment decisions is still large to solve. A full size problem is solved at each iteration, which leads to poor performance. This suggests that for MHSP, standard Lagrangean decomposition without parallel computing is not a suitable approach. However, standard Benders performs well and in some cases outperforms the full space problem with Gurobi. The proposed PLPR yields the best performance across all test instances. This is because the Lagrangean subproblem is solved in parallel, so the computational time almost does not increase with the number of scenarios given enough computer nodes. The value of PLPR is more significant for larger instances. A drawback of Lagrangean decomposition and PLPR is that their convergence is highly dependent on the adjustment of step sizes. Extensive tests have been conducted to find suitable parameters for the adjustment. In contrast, Benders decomposition requires no effort in choosing parameters, which makes it more robust.

We note that Lagrangean decomposition has smaller gaps obtained in the initial iterations than Benders decomposition. This is because Benders decomposition requires a sufficiently large number of cutting planes to approximate accurately the objective function, whereas Lagrangean decomposition only needs to find the optimal multipliers. This would suggest that Lagrangean may be the preferred method if the underlying problem does not have to be solved to a very tight tolerance. This may be the case when dealing with huge investment planning problems where a very tight convergence tolerance is not meaningful.

An analysis of computational times is presented in Table 3. We can see that as the number of scenarios increases, the time spent on solving scenario subproblems increases significantly in standard Lagrangean decomposition compared with PLPR. In addition, due to the primal reduction in PLPR and parallel computing, the time spent on solving the primal problem is much less in PLPR than in standard Lagrangean decomposition.

Table 3: Comparative analysis of the computational time (in seconds) of standard Lagrangean and PLPR.

	Dual problem		Primal problem		Multiplier update	
	SL	PLPR	SL	PLPR	SL	PLPR
Case 1	59.1	4.2	21.8	4.5	0.1	0.2
Case 2	933.3	51.2	475.5	23.4	0.2	0.4
Case 3	-	47.7	-	31.8	-	0.4
Case 4	-	33.7	-	19.9	-	0.4
Case 5	-	5.3	-	5.4	-	0.3
Case 6	-	82.9	-	23.5	-	0.6
Case 7	-	88.1	-	60.5	-	0.5
Case 8	-	191.6	-	98.9	-	0.5

6.1.2. Power system investment decisions

This section presents the optimal investment decisions in the first stage for Cases 1-8. We can see from Table 4 that by including operational uncertainty, the investment decisions are considerably different. The differences in long-term and short-term uncertainty in Cases 1-8 are presented in Table 1. Cases 1-4 only differ in operational uncertainty, we can see from Table 4 that the investments in CoalCCS and OnWind are significantly different. It is the same case for Cases 5-8. Cases 1 and 5 differ only in long-term uncertainty, we can see that the investment in OnWind in Case 1 is 79.83 GW compared with 76.63 GW in Case 5. The difference can also be observed by comparing Cases 2 and 6, Cases 3 and 7, and Cases 4 and 8. From this, we can see that both long-term and short-term uncertainty can affect investment decisions significantly. This shows the value of including short-term and long-term uncertainty in a long-term stochastic investment planning problem.

Table 4: Investment decisions in the first stage in the cases.

	New installed capacity (GW)							
	Case 1	Case 2	Case 3	Case 4	Case 5	Case 6	Case 7	Case 8
Coal	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
CoalCCS	11.70	10.17	8.77	9.48	11.68	8.33	8.76	9.29
OCGT	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
CCGT	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Diesel	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Nuclear	7.50	7.50	7.50	7.50	7.50	7.50	7.50	7.50
PHES	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Lithium	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
OnWind	79.83	69.50	77.97	93.97	76.63	72.68	77.02	92.37
OffWind	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
PvSolar	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

7. Discussion

In this paper, we have proposed the PLPR (Parallel Lagrangean decomposition with primal reduction) algorithm, and formalised Benders decomposition and Lagrangean decomposition for MHSP. We tested the proposed methods on a UK power system expansion problem using the REORIENT model.

Through computational tests, we found that PLPR is a very efficient decomposition method for MHSP that utilises the scenario structure of the MHSP. The computational time does not scale much as the problem instance grows due to the use of parallel computing. Despite parallel computing and primal reduction, PLPR inherits the advantages and the disadvantages of Lagrangean decomposition.

We found that Lagrangean decomposition can obtain a good convergence gap in the initial iterations. However, one limitation of Lagrangean type decomposition is that it is sensitive to parameter tuning. Although MHSP reduces the number of multipliers, finding good ones can still be hard. In addition, we notice that Lagrangean decomposition requires substantially more memory than Benders decomposition. Lagrangean decomposition duplicates the variables, which leads to a larger model size. However, Lagrangean decomposition can solve more classes of problems such as the ones with integer operational variables, which were not addressed in this paper. Benders decomposition can only solve linear programming or mixed-integer linear programming with integer variables in the reduced master problem.

Benders decomposition is more robust than Lagrangean decomposition because its convergence does not depend on parameter tuning. However, the drawback of Benders decomposition is that once the scenario tree is large, the master problem may become harder to solve. This is because a standard two-stage Benders solves a multi-stage stochastic program. Therefore, the reduced master problem includes all investment nodes. Once there are a number of investment nodes, or there are integer variables in the reduced master problem, the speed of Benders decomposition may be affected significantly.

For very large problems, combining Lagrangean decomposition with Benders decomposition may be beneficial. For example, use Benders decomposition to solve the Lagrangean subproblem. It is also possible to utilise adaptive oracles (Mazzi et al., 2020) in Lagrangean decomposition.

8. Conclusions and future work

In this paper, we first proposed, formalised and compared decomposition algorithms for linear programming MHSP. We formalised the node and scenario based formulations of MHSP. Decomposition methods including standard Benders, standard Lagrangean and PLPR were proposed based on the special structure of MHSP. Some properties based on the structure of MHSP were presented. By comparing Benders decomposition and Lagrangean decomposition, we found that: (1) PLPR outperforms the other algorithms across all test instances and is up to 9.2 times faster than Benders decomposition for a 1% convergence tolerance; (2) standard Lagrangean is not an efficient method for MHSP; (3) Benders decomposition is more robust in terms of parameter tuning. The choice of algorithms for MHSP was discussed based on the computational tests.

This is the first paper that has systematically studied decomposition methods for MHSP. Future work may include (1) developing algorithms that further exploit the special structure of MHSP, such as Benders decomposition with cut sharing or combined Lagrangean decomposition and Benders decomposition algorithm, and (2) extending the algorithm to solve mixed integer linear programming MHSP.

CRediT author statement

Hongyu Zhang: Conceptualisation, Methodology, Software, Validation, Formal analysis, Investigation, Visualisation, Data curation, Writing - original draft, Writing - review & editing. **Èric Mor Domènech:** Conceptualisation, Methodology, Software, Writing - review & editing. **Ignacio E. Grossmann:** Conceptualisation, Methodology, Supervision, Writing - review & editing. **Asgeir**

Tomasgard: Conceptualisation, Methodology, Supervision, Writing - review & editing, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported by the Research Council of Norway through PETROSENTER LowEmission [project code 296207].

References

Backe, S., Skar, C., del Granado, P. C., Turgut, O., & Tomasgard, A. (2022). EMPIRE: an open-source model based on multi-horizon programming for energy transition analyses. *SoftwareX*, *17*, 100877. doi:10.1016/j.softx.2021.100877.

Benders, J. F. (1962). Partitioning procedures for solving mixed-variables programming problems. *Numerische Mathematik*, *4*, 238–252. doi:10.1007/BF01386316.

Birge, J. R. (1985). Decomposition and partitioning methods for multistage stochastic linear programs. *Operations Research*, *33*, 989–1007. doi:10.1287/OPRE.33.5.989.

Birge, J. R., & Louveaux, F. (2011). *Introduction to stochastic programming*. Springer Science & Business Media. doi:10.1007/978-1-4614-0237-4.

Dunning, I., Huchette, J., & Lubin, M. (2017). JuMP: a modeling language for mathematical optimization. *SIAM Review*, *59*, 295–320. doi:10.1137/15M1020575.

Escudero, L. F., & Monge, J. F. (2018). On capacity expansion planning under strategic and operational uncertainties based on stochastic dominance risk averse management. *Computational Management Science*, *15*, 479–500. doi:10.1007/S10287-018-0318-9/FIGURES/4.

Gurobi Optimization, LLC (2022). Gurobi optimizer reference manual. URL: https://www.gurobi.com.

Kaut, M., Midthun, K. T., Werner, A. S., Tomasgard, A., Hellemo, L., & Fodstad, M. (2014). Multi-horizon stochastic programming. *Computational Management Science*, *11*, 179–193. doi:10.1007/s10287-013-0182-6.

Maggioni, F., Allevi, E., & Tomasgard, A. (2020). Bounds in multi-horizon stochastic programs. *Annals of Operations Research*, *292*, 605–625. doi:10.1007/S10479-019-03244-9/TABLES/5.

Mazzi, N., Grothey, A., McKinnon, K., & Sugishita, N. (2020). Benders decomposition with adaptive oracles for large scale optimization. *Mathematical Programming Computation*, . doi:10.1007/s12532-020-00197-0.

Zhang, H., Grossmann, I. E., Knudsen, B. R., McKinnon, K., Nava, R. G., & Tomasgard, A. (2023). Integrated investment, retrofit and abandonment planning of energy systems with short-term and long-term uncertainty using enhanced Benders decomposition. *arXiv*, . doi:10.48550/arXiv.2303.09927.

Zhang, H., Mazzi, N., Mckinnon, K., Nava, R. G., & Tomasgard, A. (2022). A stabilised Benders decomposition with adaptive oracles applied to investment planning of multi-region power systems with short-term and long-term uncertainty. *arXiv*, . doi:10.48550/arXiv.2209.03471.