Multi-Fidelity Benders Decomposition for Generation, Storage, and Transmission Expansion Planning

D. Mitchell[a,*], A. Liebman[a], A.T. Ernst[a], P. Le Bodic[a], S. Dunstall[b]

[a]*Monash University, Clayton, 3800, Melbourne, Australia*
[b]*CSIRO Data 61, Clayton, 3800, Melbourne, Australia*

Abstract

Modern energy grid expansion planning, by necessity, includes timeseries data to accurately model storage and renewable assets. Representative time periods are commonly used as a way to decrease problem size and therefore mitigate the increased complexity from this inclusion. However, there are many choices around these representative periods: length; location in planning horizon; boundary conditions. Each of these can influence what is considered the optimal planning outcome, potentially causing disparity with the original problem's optimal solution. As an alternative to this approach, we present a novel methodology to contend with the entire planning horizon including the embedded timeseries by augmenting Benders decomposition (BD). To improve the speed of generating Benders cuts, the full subproblem can be partitioned into smaller independent subproblems covering the entire horizon. However, this approach ignores the boundary conditions between subproblems, hence they only produce a lower bound for the GSTEP problem. Here we define a *multi-fidelity* BD approach that converges quickly by using small subproblems for approximate objective evaluation and fast cut generation. However, it still guarantees optimality by ultimately solving a few full subproblems. We apply this approach to a variety of modified IEEE test cases and several scenarios of the real-world Australian National Energy Market to show that dynamically adjusting the subproblems' size between multiple fidelities can provide significant speed-up. Although the proposed method is tested in the context of this specific case study, it is generically applicable to other long-term planning problems that use Benders decomposition.

Keywords: Benders decomposition, generation storage and transmission expansion planning, mixed-integer linear programming, multi-cut acceleration, Multiple-Fidelity Benders decomposition.

1. Introduction

The Benders decomposition (BD) framework [1], [2], [3], [4] has been successfully used to solve energy grid Expansion Planning (EEP), specifically transmission and thermal generation expansion planning. BD is a natural decomposition that splits the difficult mixed-integer linear program (MILP) of EEP into two problems: the *master* problem, an integer program which makes investment decisions, and a *subproblem*, a linear program that evaluates operational costs under master-level decisions. Unfortunately, the inclusion of assets with time-dependent variations, such as renewable generation technologies, has created the need to solve the operational subproblem over a planning horizon, dramatically increasing complexity. For example, solar power varies with solar cycles and is subject to seasonal trends, requiring the subproblem to evaluate the operational cost from power flow at each time-step across the planning horizon. This is further complicated by the inclusion of battery storage assets.

In the case where batteries are included in the expansion plan, the problem becomes Generation, Storage and Transmission Expansion Planning (GSTEP). While batteries provide a direct way to mitigate the variability in power from renewable sources to better match demand, their inclusion introduces linking constraints between time-steps that are necessary to allow energy stored by one time period to be used in another. These extra constraints further complicate the model.

A common approach to mitigate this increase in model complexity is to compress the planning horizon into

*Corresponding author
Email address: drew.mitchell1@monash.edu (D. Mitchell)

representative hours, days, or weeks which attempt to capture the dynamics of demand, and renewable variability of the system. This approach has been used within BD frameworks and on non-decomposed, or monolithic, MILP alike [5], [6], [7], [8], [9], [10]. Using this approach, however, the investment decisions obtained are highly dependent on the length of representation chosen [11].

Hence, practitioners using BD, who wish to maintain the information encoded within the full planning horizon, are faced with the following choice: solve the subproblem for the entire planning horizon; or split the planning horizon, but lose the ability to faithfully model the behaviour of batteries between time periods and, without intervention, lose guaranteed BD optimality.

For realistic test cases, solving a single, long, subproblem is impractical, so the planning horizon must be split. The question of how to handle the state-of-charge of batteries between consecutive subproblems remains. The "fixed-charge" strategy consists in setting the state-of-charge of all batteries to a fixed amount at the beginning of each time-step of each subproblem [12]. However, the starting charge level is "free energy", arbitrarily influencing how cost-efficient batteries are to the master problem. Nevertheless, this approach is attractive as it ensures that the subproblems return a lower bound on the operational costs if the batteries start full and the price of energy is never negative. Under these assumptions, the master problem of BD converges to a lower-bound of the optimal GSTEP solution.

In this paper, we extend and improve on this concept of partitioning the planning horizon to create relaxed independent subproblems and utilise decomposition techniques [13], namely adding cuts from these relaxed problems before strengthening the formulation, to create a novel framework.

The contributions presented in this paper are:

1. An algorithm, *multi-fidelity* BD (MFBD) which exploits the BD structure by initially solving a set of short (low-fidelity) independent subproblems for fast cut generation. These low-fidelity subproblems span the entire planning horizon but with relaxed boundary conditions on storage state-of-charge variables. Next we increase fidelity by solving longer timescale subproblems to obtain tighter bounds, keeping the cheap cuts, before eventually recovering the full planning horizon subproblem for guaranteed optimality when solved for all facet defining cuts.
2. Further acceleration of MFBD using an adaptive strategy to decide when to increase fidelity: while it would be possible to solve each fidelity to optimality before switching to a higher fidelity, we instead detect when the convergence of BD slows down for the current fidelity to increase the fidelity.
3. A theorem with accompanying proof for guaranteeing the preservation of the optimal solution when solving, and adding cuts from, low-fidelity subproblems.
4. Case study results demonstrating the computational speed-up over Benders Decomposition obtained using the proposed algorithm.

The remainder of this paper is organised as follows. Section 2 contains a brief literature review of energy grid expansion planning and BD. Section 3 outlines a standard formulation of the GSTEP problem considering generation, transmission, and storage expansion with algorithm. Section 4 presents our novel acceleration technique, justification and modified algorithm. Section 5 highlights the numerical results applied to some modified standard IEEE test cases, as well as test case scenarios of the 2019 Australian National Energy Market (NEM). Section 6 concludes the paper.

2. Literature Review

This literature review is split into three key components. First, we consider a brief overview of the current state of EEP. Research in this area is extremely diverse, and as such we will focus primarily on works directly related to the version of EEP below, in Section 3. Next, we discuss some of the areas and techniques for accelerating Benders decomposition (BD). Finally, we highlight the application of BD to the EEP problem. For more comprehensive literature reviews, we recommend [14, 15].

2.1. Energy Grid Expansion Planning

The problem of EEP was first introduced in 1970 by *Garver* [16, 17], in the form of transmission network expansion planning for energy networks (TNEP). More recently, with advancements in renewable energy technology and political debate about the reliability and cost of electricity, the problem has transformed from transmission network expansion planning into the long-term planning of entire energy grids [18, 6].

Given the broad nature of the research into this topic, it is important we clarify which areas we are specifically investigating. There are many methods researchers use and have used to solve energy expansion planning: from mathematical techniques to meta-heuristics and machine learning [14, 15].

Regardless of the method used to solve this problem, the problem can be abstracted in many ways. In its most realistic representation, the flow of electricity is represented as alternating current (AC), which is non-linear and clearly complicates optimisation approaches. The flow can, of course, be fully relaxed to obtain a transportation model which readily extends to linear programming (LP) techniques at the cost of accuracy. A compromise between these representations uses a direct current (DC) approximation in a disjunctive model [19].

2.2. Benders Decomposition

A popular methodology for dealing with problems containing binary variables is Benders decomposition (BD) [20]. For a holistic investigation of BD, we recommend the review by *Rahmaniani et al.* [21].

The concept of adding multiple non-aggregated cuts into the master from independent subproblems in each iteration has been previously explored, most relevantly in [22]. This paper applies a multiple cut generation approach to the generation and transmission expansion planning problem. The authors demonstrate significant speed-up using a multiple cut approach based on generating cuts for planning years and "operating modes" [22].

There is a limit, however, to how many cuts can be added without creating a master problem that is prohibitively large to solve. The more cuts added each iteration the faster this state is approached, hence either the subproblems must be clustered together somehow, or cuts that no longer serve a purpose must be removed from the master itself [23]. As discussed in [21], further work is required on cut removal strategies. Finally, a recent (2017) work [24] investigating the parallelisation of a "Nested" BD applied to hydrothermal scheduling provides key insight into what may be done when considering problems with inter-time dependent constraints. This "Nested" BD considers a multistage planning horizon, where each stage depends on the previous, hence creating a nested structure. The authors demonstrate the ability to relax constraints between what would otherwise be interlinked subproblems allowing for parallelisation and significant computational speedup [24].

2.3. Benders Decomposition in Grid Expansion Planning

BD has been applied to EEP with positive results since 1985 [1, 25]. The research scope for BD applied to energy grid expansion planning is broad and covers the range of both the static and dynamic (multi-year) version of the problem [8], as well as considering security constraints [26].

Even under the BD framework, expansion planning can be difficult and time-consuming to solve. Some current strategies to accelerate this process partition the planning horizon into smaller time windows, at a single fidelity (or size), and solve: each constituent subproblem but with altered boundary conditions between them i.e. each day of a planning year, in order with cyclic conditions [12]; or some set of subproblems from one part of the planning horizon i.e. the month of June or 4 consecutive days of the year [10, 6]. To speed up convergence, work has also been done to augment the Benders framework. Specifically, a Branch-and-Cut Benders Decomposition (BCBD), where BD is embedded inside the Branch-and-Cut framework has been utilised, to speed up convergence for security constrained TEP problem (SCTEP) [4]. This speed-up is obtained by reducing the number of nodes/solutions needed to be explored via pruning nodes in the Branch-and-Bound tree.

The authors in [8] consider a formulation of the dynamic expansion planning problem where each subsequent investment year is considered as a subproblem of the previous. The system is solved in a two-step process, with a forward pass to obtain a feasible solution and a backward pass to obtain the price and add cuts for investment decisions in each year. This backwards pass is solved using a Lagrangian relaxation of the problem. Further work was done, [27], to present a tailored BD in which the subproblems consider each year in the dynamic planning problem and the master deals with the complicating variables of investment over a multitude of years. This work also introduces a hull formulation [28] to replace the big-M and alternative big-M formulation. This hull formulation has more continuous variables but a tighter LP relaxation. Experimental results indicated a large time saving when using tailored BD.

3. Benders Decomposition for GSTEP

In this section, we first provide a monolithic formulation for GSTEP under DC-OPF: MILP (1), [4, 29]. Table 1 summarises the notation used. The MILP is then decomposed using BD. The master, Section 3.1, and subproblem, Section 3.2, are explained in their respective subsections. The splitting of this subproblem into smaller independent problems is then covered in Section 3.3. Finally, an algorithm for solving the single-fidelity BD is provided in Section 3.4.

$$\min_{x_i,\, p_{t.g},\, q_{t,g}} \sum_{i\in\mathcal{I}} C_i^{\text{inv}} x_i + \sum_{t\in\mathcal{T}} \left(\sum_{g\in\mathcal{G}} C_g^{\text{gen}} p_{t,g} + \sum_{n\in\mathcal{N}} C^q q_{t,n} \right) \tag{1a}$$

$$\text{s.t.} \quad F_l^{\min} x_l \le f_{t,l} \le F_l^{\max} x_l \qquad \forall t\in\mathcal{T}, l\in\mathcal{L}, \tag{1b}$$

$$f_{t,l} - B_l(\theta_{t,\nu(l)} - \theta_{t,\omega(l)}) \le M_l(1-x_l) \qquad \forall t\in\mathcal{T}, l\in\mathcal{L}, \tag{1c}$$

$$-f_{t,l} + B_l(\theta_{t,\nu(l)} - \theta_{t,\omega(l)}) \le M_l(1-x_l) \qquad \forall t\in\mathcal{T}, l\in\mathcal{L}, \tag{1d}$$

$$\theta^{\min} \le \theta_{t,n} \le \theta^{\max} \qquad \forall t\in\mathcal{T}, n\in\mathcal{N}, \tag{1e}$$

$$p_{t,g} \le P_g^{\max} x_g \qquad \forall\, t\in\mathcal{T}, g\in\mathcal{G}, \tag{1f}$$

$$p_{t,r} \le P_{t,r}^{\max} x_r \qquad \forall\, t\in\mathcal{T}, r\in\mathcal{R}, \tag{1g}$$

$$p_{t,e}^{\text{dis}} \le P_e^{\text{dis}} x_e \qquad \forall\, t\in\mathcal{T}, e\in\mathcal{E}, \tag{1h}$$

$$p_{t,e}^{\text{char}} \le P_e^{\text{char}} x_e \qquad \forall\, t\in\mathcal{T}, e\in\mathcal{E}, \tag{1i}$$

$$v_{t,e} \le C_e^{\max} x_e \qquad \forall\, t\in\mathcal{T}\cup\{|\mathcal{T}|+1\}, e\in\mathcal{E}, \tag{1j}$$

$$p_{t,e} = p_{t,e}^{\text{dis}} - p_{t,e}^{\text{char}} \qquad \forall\, t\in\mathcal{T}, e\in\mathcal{E}, \tag{1k}$$

$$p_{t,e}^{\text{dis}} - \epsilon p_{t,e}^{\text{char}} = v_{t,e} - v_{t+1,e} \qquad \forall\, t\in\mathcal{T}, e\in\mathcal{E}, \tag{1l}$$

$$\sum_{i\in\mathcal{I}(n)} p_{t,i} + \sum_{l\in\mathcal{L}|\omega(l)=n} f_{t,l} - \sum_{l\in\mathcal{L}|\nu(l)=n} f_{t,l} + q_{t,n} = d_{t,n} \quad \forall t\in\mathcal{T}, n\in\mathcal{N}, \tag{1m}$$

where,

$$x_i \in \mathbb{Z}_+^n \cup \{0\}^n \qquad \forall i\in\mathcal{I}, \tag{1n}$$

$$f_{t,l} \in \mathbb{R} \qquad \forall t\in\mathcal{T}, l\in\mathcal{L}, \tag{1o}$$

$$\theta_{t,n} \in \mathbb{R} \qquad \forall t\in\mathcal{T}, n\in\mathcal{N}, \tag{1p}$$

$$p_{t,g} \ge 0 \qquad \forall\, t\in\mathcal{T}, g\in\mathcal{G}, \tag{1q}$$

$$p_{t,e}^{\text{dis}} \ge 0 \qquad \forall\, t\in\mathcal{T}, e\in\mathcal{E}, \tag{1r}$$

$$p_{t,e}^{\text{char}} \ge 0 \qquad \forall\, t\in\mathcal{T}, e\in\mathcal{E}, \tag{1s}$$

$$p_{t,r} \ge 0 \qquad \forall\, t\in\mathcal{T}, r\in\mathcal{R}, \tag{1t}$$

$$p_{t,e} \in \mathbb{R} \qquad \forall\, t\in\mathcal{T}, e\in\mathcal{E}, \tag{1u}$$

$$v_{t,e} \ge 0 \qquad \forall\, t\in\mathcal{T}\cup\{|\mathcal{T}|+1\}, e\in\mathcal{E}, \tag{1v}$$

$$q_{t,n} \ge 0 \qquad \forall\, t\in\mathcal{T}, n\in\mathcal{N}. \tag{1w}$$

MILP (1) is then split into a master problem containing the binary (investment decision) variables and a set of subproblems, $\mathcal{S}$, representing DC optimal power flow over a network (DC-OPF) under these investment decisions [4, 29].

3.1. Master (Investment) Mixed Integer Linear Program

The master problem for GSTEP minimises the cost of investing in each asset by assigning each a binary investment decision variable, x_i, over the set of assets $\mathcal{I}$. Each investment variable x_i has a corresponding annualised investment cost C_i^{inv}. The operational costs are provided by the decision variables, α_s, for each subproblem $s\in\mathcal{S}$. These α_s variables are bounded (2b) by the optimality cuts from each subproblem; they pass back information from the subproblems to guide investment decisions in the master problem. Specifically, in (2b), $z_s(\bar{\boldsymbol{x}}^k)$ is the objective value of subproblem, s, in iteration, k, $\lambda_{s,i}^k$ is the dual price of each investment decision $\bar{x}_i$ from iteration k.

Table 1: Notation

Sets	**Explanation**
$\mathcal{T}$	Discrete time-steps (Monolithic MILP)
$\mathcal{S}_\gamma$	Subproblem indices for fidelity $\gamma \in \Gamma$
$\mathcal{T}_s$	Discrete time-steps in subproblem s
$\mathcal{P}_\gamma$	End time-steps for subproblems for fidelity $\gamma \in \Gamma$
Γ	Fidelities: {Daily, Weekly, Monthly, Yearly}
$\mathcal{N}$	Nodes
$\mathcal{L}$	Transmission lines
$\mathcal{G}$	Thermal generators
$\mathcal{R}$	Renewable generators
$\mathcal{E}$	Storage assets
$\mathcal{I}$	$\mathcal{L} \cup \mathcal{G} \cup \mathcal{R} \cup \mathcal{E}$
Parameters	
B_l	Susceptance in line $l \in \mathcal{L}$
$F_l^{\min}$	Minimum power flow in line $l \in \mathcal{L}$
$F_l^{\max}$	Maximum power flow in line $l \in \mathcal{L}$
$\theta^{\min}$	Minimum phase angle
$\theta^{\max}$	Maximum phase angle
$P_g^{\max}$	Maximum generation of thermal generator $g \in \mathcal{G}$
$P_{t,r}^{\max}$	Maximum generation of renewable generator $r \in \mathcal{R}$ at time-step $t \in \mathcal{T}_s$
P_e^{char}	Maximum charge rate to battery $e \in \mathcal{E}$
P_e^{dis}	Maximum discharge rate from battery $e \in \mathcal{E}$
$C_e^{\max}$	Maximum state-of-charge battery $e \in \mathcal{E}$
ϵ	Storage loss parameter
$d_{t,n}$	Demand at node $n \in \mathcal{N}$ at time-step $t \in \mathcal{T}_s$
C_g^{gen}	Generation/fuel cost of generator $g \in \mathcal{G}$
C_i^{inv}	Investment cost of asset $i \in \mathcal{I}$
C^q	Cost of unserved demand
M_l	Big-M constants per line $l \in \mathcal{L}$
Decision Variables	
$f_{t,l}$	Flow in line $l \in \mathcal{L}$ at time-step $t \in \mathcal{T}_s$
$\theta_{t,n}$	Phase angle at node $n \in \mathcal{N}$ at time-step $t \in \mathcal{T}_s$
$p_{t,i}$	Power from asset $i \in \{\mathcal{G} \cup \mathcal{R} \cup \mathcal{E}\}$ at time-step $t \in \mathcal{T}_s$
$v_{t,e}$	state-of-charge of battery $e \in \mathcal{E}$ at time-step $t \in \mathcal{T}_s$
$q_{t,n}$	Unserved demand at node $n \in \mathcal{N}$ at time-step $t \in \mathcal{T}_s$
x_i	Binary investment decision variable for existence of asset $i \in \mathcal{I}$
α_s	Optimality cut variable for subproblem $s \in \mathcal{S}_\gamma$

$$\min_{x_i, \alpha_d} \sum_{i \in \mathcal{I}} C_i^{\text{inv}} x_i + \sum_{s \in \mathcal{S}} \alpha_s, \tag{2a}$$

subject to:

$$\alpha_s \geq z_s^k(\bar{\boldsymbol{x}}) + \sum_{i \in \mathcal{I}} \lambda_{s,i}^k (x_i - \bar{x}_i^k) \qquad \forall s \in \mathcal{S},\ k, \tag{2b}$$

$$\alpha_s \geq 0 \qquad \forall s \in \mathcal{S}, \tag{2c}$$

$$x_i \in \mathbb{Z}_+^n \cup \{0\}^n \qquad \forall i \in \mathcal{I}. \tag{2d}$$

3.2. Linear Subproblem (DC-Optimal Power Flow)

The subproblems consider the operational cost under the set o f investment decisions $\bar{\boldsymbol{x}}$. The objective value $z_s(\bar{\boldsymbol{x}})$ is the minimum cost of power generation plus the cost of unserved demand for subproblem s. The power output for each generator $g \in \mathcal{G}$, at each time-step $t \in \mathcal{T}_s$, is given by the decision variables $p_{t,g}$. The generation cost is measured in LP (3) by multiplying the power generation decision variables by their associated per-unit generation cost, C_g^{gen}, and summing over time-steps and generators. Note that we assume the cost of generation is fixed in time. Ramp-rate constraints and unit commitment are not considered in this model. The inclusion of ramp-rates adds in additional time-linking constraints, which compound with those from state-of-charge variables. However, for the times-scales and asset types considered in this paper, these ramp-rates do not have significant impact. Only coal-fired thermal generators have sufficiently shallow enough ramp-rate to affect dispatch, albeit marginally, as compared to storage limits, which are binding and act across all assets.

Unserved demand is determined at each node n, and each time-step t, with the decision variables $q_{t,n}$. The associated cost of unserved demand, in LP (3), is the multiplication between this variable and the per-unit cost, C^q, added over all time-steps and nodes. This minimisation is subject to power flow, generation, storage and nodal balance constraints, (5),(6),(7),(8a),(8b),(9), explained below. Finally we note that λ_i is the dual variable associated with the constraint fixing the investment decision variables to the values chosen by the master.

$$z_s(\bar{\boldsymbol{x}}) = \min_{p_{t.g},\, q_{t,g}} \sum_{t \in \mathcal{T}_s} \left(\sum_{g \in \mathcal{G}} C_g^{\text{gen}} p_{t,g} + \sum_{n \in \mathcal{N}} C^q q_{t,n} \right) \tag{3a}$$

$$\text{s.t.} \qquad \text{Eqs. (1b) to (1m) and (1o) to (1w)} \quad \forall\ t \in \mathcal{T}_s,$$

$$x_i = \bar{x}_i \qquad \forall i \in \mathcal{I} \qquad : \lambda_i. \tag{3b}$$

We consider, in our testing, both a DC-OPF and a simple transportation model.

In the transportation model, the flow decision variables $f_{t,l}$ in each line l at each time-step t are only bounded by the minimum and maximum thermal limits on each line $F_l^{\max}$ and $F_l^{\min}$ such that,

$$F_l^{\min} x_l \leq f_{t,l} \leq F_l^{\max} x_l \quad \forall t \in \mathcal{T}_s, l \in \mathcal{L}. \tag{4}$$

When considering DC-OPF the power flow constraints along transmission lines in our network are further restricted/governed by Kirchhoff's second voltage law (5a), (5b) in the disjunctive big-M form. Consider

$$f_{t,l} - B_l(\theta_{t,\nu(l)} - \theta_{t,\omega(l)}) \geq -M_l(1 - x_l) \quad \forall t \in \mathcal{T}_s, l \in \mathcal{L}, \tag{5a}$$

and

$$f_{t,l} - B_l(\theta_{t,\nu(l)} - \theta_{t,\omega(l)}) \leq M_l(1 - x_l) \quad \forall t \in \mathcal{T}_s, l \in \mathcal{L}, \tag{5b}$$

where we introduce the phase angle variables $\theta_{t,n}$ at node n, and time-step t, and where $\nu(l)$ and $\omega(l)$ refer to the start and end nodes of line l, respectively. The inclusion of these constraints means the loss of monotonicity in the objective: the addition of candidate lines does not necessarily decrease the subproblem cost. The constant B_l is the susceptance of line l, and M is a large "big-M" constant. The phase angle variables are also bounded by $\theta^{\min}$ and $\theta^{\max}$:

$$\theta^{\min} \leq \theta_{t,n} \leq \theta^{\max} \quad \forall t \in \mathcal{T}_s, n \in \mathcal{N}. \tag{5c}$$

The only limitations placed on thermal generators is that they sit in their operating range, such that, at time-step t, we have the power from thermal generator g, $p_{t,g}$, bounded by

$$0 \leq p_{t,g} \leq P_g^{\max} x_g \quad \forall\ t \in \mathcal{T}_s,\ g \in \mathcal{G}, \tag{6}$$

with investment decisions for generator g, x_g and maximum generation output given by $P_g^{\max}$. There are analogous constraints for the power from renewable generators. For renewable generator r, the decision variable for power output $p_{t,r}$, is bounded by $P_{t,r}^{\max}$, corresponding to the asset's renewable trace at time-step t and the investment decision variable for the corresponding asset, x_r, such that

$$0 \leq p_{t,r} \leq P_{t,r}^{\max} x_r \quad \forall\ t \in \mathcal{T}_s,\ r \in \mathcal{R}. \tag{7}$$

The final set of constraints bounds the dynamics of storage assets: charging, discharging, and charge level. The charging/discharging decision variable $p_{t,e}$ at time-step t, for storage unit e, is bounded by the charging and discharging constants P_e^{char}, P_e^{dis} respectively, with the decision variable x_e.

$$-P_e^{\text{char}} x_e \leq p_{t,e} \leq P_e^{\text{dis}} x_e \quad \forall\ t \in \mathcal{T}_s,\ e \in \mathcal{E}. \tag{8a}$$

The state-of-charge decision variables $v_{t,e}$, for each storage asset e, at time-step t, are bounded by constant $C_e^{\max}$, as the maximum state-of-charge of this asset:

$$0 \leq v_{t,e} \leq C_e^{\max} \quad \forall t \in \mathcal{T}_s \cup \{|\mathcal{T}_s| + 1\},\ e \in \mathcal{E}. \tag{8b}$$

At each time-step t, and asset e the charge and discharge variables $p_{t,e}^{\text{char}}$ and $p_{t,e}^{\text{dis}}$ are governed by the change in charge level on either side of the time-step. To account for round-trip efficiency, a loss factor of ϵ is applied to the charging variable. Simply,

$$p_{t,e}^{\text{dis}} - \epsilon p_{t,e}^{\text{char}} = v_{t,e} - v_{t+1,e} \quad \forall t \in \mathcal{T}_s,\ e \in \mathcal{E}. \tag{8c}$$

Then, the net flow $p_{t,e}$ in or out of the assets can be determined. This is written as,

$$p_{t,e} = p_{t,e}^{\text{dis}} - p_{t,e}^{\text{char}} \quad \forall t \in \mathcal{T}_s,\ e \in \mathcal{E}. \tag{8d}$$

Note that (8c) is the only constraint that acts temporally, linking the current time-step t, to the next time-step, $(t+1)$. As such, when the planning horizon is split into multiple low-fidelity subproblems, this constraint is relaxed at the boundary between subproblems, see Section 3.3. A common alternative approach is to include cyclic boundary conditions where the first and last time-step must have the same state of charge. This, however, does not guarantee a lower-bound cut for the subproblems.

The final constraint measures the unserved energy. At each node n and time-step t, it requires that the sum of power generated, stored and dispatched, along with all power flowing into and out of each node, plus unserved energy $q_{t,n}$, add up to the demand $d_{t,n}$. Explicitly,

$$\sum_{i \in \mathcal{I}(n)} p_{t,i} + \sum_{l \in \mathcal{L} | \omega(l) = n} f_{t,l} - \sum_{l \in \mathcal{L} | \nu(l) = n} f_{t,l} + q_{t,n} = d_{t,n} \quad \forall\ t \in \mathcal{T}_s, n \in \mathcal{N}. \tag{9}$$

The inclusion of unserved demand in the subproblems guarantees that the subproblems are always feasible.

3.3. Subproblem generation

When considering a single subproblem spanning the entire planning horizon, in this case a "Yearly" fidelity with $\mathcal{S}_Y = \{1\}$, the above formulation is a standard BD with one subproblem and a single cut added each iteration. When partitioning the planning horizon into smaller intervals and generating independent subproblems, some treatment for Eq. (8c) must be considered at the boundary when one subproblem ends and another begins. This is the only constraint in the above formulation that links time-steps together. By duplicating the variable corresponding to the state-of-charge at the end of a subproblem and assigning the newly created variable to the next subproblem the two time-steps are no longer connected and independence is gained. To illustrate this, consider a partition where a single year is split into daily subproblems, a "Daily"

fidelity. Then, $\mathcal{S}_D = \{1, 2, \ldots, 365\}$ with each subproblem containing 24 time-steps. Or, explicitly: for $s = 1$ we have $\mathcal{T}_1 = \{1, 2 \ldots, 24\}$, for $s = 2$, $\mathcal{T}_2 = \{25, 26 \ldots, 48\}$, and so on.

Then to separate subproblem 2 from subproblem 1 we can remove the constraints

$$p_{24,e} = v_{24,e} - v_{25,e} \quad \forall\, e \in \mathcal{E}. \tag{10}$$

Then we replace them with two new sets of constraints

$$p_{24,e} = v_{24,e} - \bar{v}_{25,e} \qquad \forall\, e \in \mathcal{E}, \tag{11}$$
$$v_{25,e} = \bar{v}_{25,e} \qquad \forall\, e \in \mathcal{E}. \tag{12}$$

The newly created $\bar{v}_{25,e}$ variables that duplicate the state-of-charge at the end of subproblem 1. By removing Eq. (12), the subproblems become decoupled and can be solved independently. This process can be continued for each day, or more generally applied to any boundary where a new subproblem is generated. To generalise this, we define a set of endpoints of subproblems, for a given fidelity γ, $\mathcal{P}_\gamma \subset \mathcal{T}$ and replace with

$$p^{\text{dis}}_{t,e} - \epsilon p^{\text{char}}_{t,e} = v_{t,e} - \bar{v}_{t+1,e} \qquad \forall\, t \in \mathcal{T}, e \in \mathcal{E}, \tag{13}$$
$$v_{t,e} = \bar{v}_{t,e} \qquad \forall\, t \in \mathcal{T}, e \in \mathcal{E}. \tag{14}$$
$$v_{t,e} = \bar{v}_{t,e} \qquad \forall\, t \in \mathcal{P}_\gamma, e \in \mathcal{E}. \tag{15}$$

The decomposition into independent subproblems then involves relaxing (15). The decomposed subproblems can then be written as

$$z_s(\bar{\boldsymbol{x}}) = \min_{p_{t.g},\, q_{t,g}} \sum_{t \in \mathcal{T}_s} \left(\sum_{g \in \mathcal{G}} C^{\text{gen}}_g p_{t,g} + \sum_{n \in \mathcal{N}} C^q q_{t,n} \right) \tag{16a}$$

$$\text{s.t.} \qquad x_i = \bar{x}_i \qquad \forall i \in \mathcal{I} \qquad : \lambda_i, \tag{16b}$$
$$\text{Eqs. (1b) to (1k) and (1m)} \quad \forall\, t \in \mathcal{T}_s,$$
$$\bar{v}_{t,e} \le C^{\max}_e x_e \qquad \forall\, t \in \mathcal{T}_s \cup \{|\mathcal{T}_s| + 1\}, e \in \mathcal{E}, \tag{16c}$$
$$p^{\text{dis}}_{t,e} - \epsilon p^{\text{char}}_{t,e} = v_{t,e} - \bar{v}_{t+1,e} \qquad \forall\, t \in \mathcal{T}_s, e \in \mathcal{E}, \tag{16d}$$
$$v_{t,e} = \bar{v}_{t,e} \qquad \forall\, t \in \mathcal{T}_s, e \in \mathcal{E}, \tag{16e}$$
$$\text{where,}$$
$$\text{Eqs. (1o) to (1u) and (1w)} \quad \forall\, t \in \mathcal{T}_s,$$
$$v_{t,e} \ge 0 \qquad \forall\, t \in \mathcal{T}_s \cup \{|\mathcal{T}_s| + 1\}, e \in \mathcal{E}, \tag{16f}$$
$$\bar{v}_{t,e} \ge 0 \qquad \forall\, t \in \mathcal{T}_s \cup \{|\mathcal{T}_s| + 1\}, e \in \mathcal{E}. \tag{16g}$$

This formulation leaves the variables $v_{t,e}$ and $\bar{v}_{t,e}$ uncoupled at the final time-step of each subproblem, s, allowing subproblems s and $s + 1$ to be solved independently.

3.4. Algorithms

We provide, in Algorithm 1, the general steps for solving the above formulation [12]. Firstly, the fidelity of the subproblem must be chosen, these can be intuitively picked from cyclic structures such as days, weeks, months or more flexibly chosen as any discretisation of the planning horizon. Once the length of time for each problem is chosen, the subproblems are generated by segmenting the planning horizon into these lengths consecutively. For details on generating independent subproblems see Section 3.3. Then the main loop of Benders decomposition can run. This involves iterating between obtaining a solution of investment variables by solving the master problem, then, utilising this solution to solve the subproblem(s) to obtain optimality cut(s). These cut(s) are added to the master before it is solved for a new solution. This cycle continues until the termination criterion is met. Algorithm 1 does not change fidelities at any point and storage is forced to be full at the beginning of each subproblem. This produces a lower bound on the optimal objective for the full problem. The 'UB' in Algorithm 1 is only an upper bound to the approximate problem. At no point in this algorithm is the true (full horizon) operational cost evaluated. To strengthen this weakness, we present the multi-fidelity approach.

Algorithm 1 Single Fidelity Benders Decomposition

Input: Fidelity γ
Initialise: master problem
Initialise: subproblem end points $\mathcal{P}_\gamma$
Initialise: subproblem(s) **for** $s \in \mathcal{S}_\gamma$
do
  sol, LB $\leftarrow$ SolveMaster()
  for $s \in \mathcal{S}$ **do**
    cut(s) $\leftarrow$ SolveSub(s,sol)
    UpdateMaster(cut(s))
  end for
  Update: UB; $Gap = \min\{Gap, \frac{UB-LB}{UB}\}$
while $Gap \geq tol$

4. Multi-Fidelity Benders Decomposition

Within the model presented in Section 3, the only linking constraints across time correspond to the state-of-charge variables of energy storage assets. When relaxing the linking constraints between subproblems (see Section 3.3), the size of the time windows, or the fidelity, will affect the attractiveness of storage investment. This bias can be mitigated by considering higher fidelity subproblems and thus tightening the formulation and obtaining more discerning Benders cuts.

We therefore present a novel method for accelerating the BD process by solving a series of low-fidelity subproblems and adding their cuts back to the master, before switching to the "true" problem once enough progress has been made. With this method, we present an important change to the master problem for GSTEP. We consider, instead of a single α variable per subproblem, a variable for each subproblem of the lowest-fidelity. Specifically, we use a variable α_d per daily subproblem. When the subproblem size being solved is daily, then the formulation presented above does not change. However, when considering higher fidelity subproblems i.e. weekly, the master problem optimality cut's left-hand-side must adjust to sum over all α_d variables involved in each subproblem. This changes the master problem formulation to:

$$\begin{aligned} \min_{x_i,\, \alpha_d} \quad & \sum_{x_i \in \mathcal{I}} C_i^{\text{inv}} x_i + \sum_{d \in \mathcal{S}_D} \alpha_d \\ \text{s.t.} \quad & \sum_{d \in s} \alpha_d \geq z_s^k + \sum_{i \in \mathcal{I}} \quad \lambda_{s,i}^k (x_i - x_i^k), \\ & \qquad\qquad\qquad\qquad \forall\ s \in \mathcal{S}_\gamma,\ \forall\ k, \\ & \alpha_d \geq 0 \qquad\qquad \forall\ d \in \mathcal{S}_D, \\ & x_i \in \mathbb{Z}_+^n \cup \{0\}^n \quad \forall\ i \in \mathcal{I}. \end{aligned} \tag{17}$$

For example a daily fidelity cut from the first day will take the form,

$$\alpha_1 \geq z_1^k + \sum_{i \in \mathcal{I}} \quad \lambda_{1,i}^k (x_i - x_i^k).$$

Then, when moving from daily to weekly subproblems, a cut from the first week will take the form,

$$\sum_{d \in 1:7} \alpha_d \geq z_1^k + \sum_{i \in \mathcal{I}} \quad \lambda_{1,i}^k (x_i - x_i^k).$$

Algorithm 2 is a high-level generalisation of this process. We define an initial list of fidelities, **L**, from which the subproblems can be generated by the process outlined in Section 3.3. The change condition to move from one fidelity of subproblems to the next is customisable. For example, we test a simple version of Algorithm 2 using a bi-fidelity approach. This means initialising a fidelity list, **L**, with just two entries, e.g. **L** = ["Daily", "Yearly"]. Then in practice, the algorithm solves daily subproblems to generate fast cuts before

Algorithm 2 Multi-Fidelity Benders Decomposition

Input: Fidelity list: **L**
Initialise: master problem
for $\gamma \in$ **L do**
  Initialise: subproblem end points $\mathcal{P}_\gamma$
  Initialise: subproblem(s) **for** $s \in \mathcal{S}_\gamma$
  do
    sol $\leftarrow$ SolveMaster(cuts)
    for $s \in \mathcal{S}_\gamma$ **do**
      cut $\leftarrow$ SolveSub(s, sol)
      UpdateMaster(cut,s)
    end for
  while Change condition not met
end for

changing to the yearly problem when the BD *Gap* tolerance is met. The yearly subproblem is then solved again until the tolerance is met. An extension to this rather simple approach is to lengthen the fidelity list and augment change condition. The adaptive approach we investigate is centred on switching fidelity when the convergence slows down. When significant progress is no longer being made for a given subproblem, the next fidelity is selected and the algorithm continues. Any cuts generated by previous fidelities are kept when a new fidelity is selected. This allows information from previous solutions to be preserved. Once the yearly problem is reached, the master solved until the *Gap* falls below the desired tolerance.

4.1. Optimality guarantee

The multi-fidelity approach is an exact optimisation algorithm. From, Theorem 2 (see Appendix A) we know that the cuts generated at lower fidelity are valid lower bounds and do not eliminate optimal solutions. As the algorithm continues and increases fidelity, the subproblems become larger and their cuts stronger, until we reach a subproblem spanning the entire planning horizon. The final phase where the original (full) subproblem is solved then guarantees that ultimately an optimal solution is found:

Corollary 1. *The optimal solution using MFBD (Algorithm 2) is optimal for the full formulation if the final fidelity consists of a single subproblem.*

The benefit from solving relaxed problems first comes from quickly cutting off poor solutions. Using a single subproblem spanning the entire planning horizon these solutions would still be removed. However, generating these cuts is much slower, as demonstrated in the results below. This theory is extensible to multi-year investment planning horizons without loss of optimality. Specifically, adding constraints to the master problem relating to multi-stage investment decisions do not affect the theorem, proof or corollary.

5. Results

5.1. Test cases and parameters

The results obtained in this paper are on the modified 6, 14 and 24 IEEE bus systems [30] as well as test instances based on the Australian Energy Market Integrated Systems Plan from 2020 [31]. These NEM cases are:

- Central: costs are determined by market forces and current federal and state government policies;
- Slow Change: slower economic growth and emission reductions;
- Step Change: both consumer-led and technology-led transitions occur in the midst of aggressive global decarbonisation;
- High DER: rapid consumer adoption of distributed energy resources (DER), which includes technologies such as rooftop solar, wind and battery storage;

- Fast Change: greater investment in grid-scale technology.

The data used in this paper is available online [32].
Table 2 is a summary of the test cases used. Further, Table 3 is provided for reference of the fidelity of each proposed and solved subproblem. We note that in some fidelities, the number of time-steps over all subproblems does not add up to a full year of hours, 8760. This is resolved by including all remaining time-steps in the final subproblem of a given fidelity. Within these test cases, we use a planning horizon of 365 days with an hourly resolution. We set the cost of unserved demand to be \$1000/MWh, the phase angle bounds to be $\pm\frac{\pi}{2}$ and the storage loss parameter to 0.9. The big-M constants were calculated using

Table 2: Table summary of test cases.

	Nodes	Lines		Generators		Renewables		Storage		Flow	Citation
Test Cases	-	Init	Cand	Init	Cand	Init	Cand	Init	Cand	-	-
6bus	6	7	3	3	4	0	4	0	3	DC	[30]
14bus	14	20	7	5	6	0	6	0	4	DC	[30]
24bus	24	34	6	12	6	0	10	0	6	DC	[30]
Central	5	7	4	36	21	106	106	15	29	T	[31]
Step Change	5	7	4	36	21	106	106	15	29	T	[31]
Slow Change	5	7	4	36	21	106	106	15	29	T	[31]
High DER	5	7	4	36	21	106	106	15	29	T	[31]
Fast Change	5	7	4	36	21	106	106	15	29	T	[31]

In the NEM case we note that candidates for Storage, Renewables and Generators refer to build sites and have much larger corresponding build limits.
Storage here is an umbrella for battery, pumped hydro and regular hydro technologies.
DC refers to DC Optimal power flow. T refers to a transportation model of power flow.

the shortest path method per candidate line [33, 34]. For more information on the exact specifications of the modified IEEE test cases, please refer to [30]. For all test cases the *Gap* tolerance was set to 0.01%. Both the bi-fidelity and the adaptive multi-fidelity approach were only terminated when the *Gap* for the yearly subproblem fell below the given tolerance or the wall time of 10,000 seconds was exceeded. The tested adaptive approach was fixed to move between daily, weekly, monthly, and the full yearly subproblem in the direction of increasing subproblem length. The change condition used, in our adaptive approach, is to check if the improvement in *Gap* over the previous 10 iterations was greater than 0.5%, or that the *Gap* falls below tolerance. All tests were run on a server with specifications: AMD EPYC-Rome Processor, instruction set [SSE2|AVX|AVX2]. Gurobi (12.0.0) was used, with default only parameters with 6 threads. For LP solving, Gurobi used concurrent optimisers: primal simplex, dual simplex, and barrier. Models were written in Julia using the JuMP package.

5.2. Analysis

We see across each IEEE test case in Figs. 1a to 1c, that all multiple fidelity approaches, either a bi-fidelity or the adaptive, are significantly faster than when solving only the full yearly subproblem from the beginning, as in classic Benders Decomposition. We notice a trade-off between time taken to solve the master problem and time taken in solving each subproblem over the course of the algorithm when decreasing the fidelity. As the number of iterations increases, so too does the number of constraints in our master problem. This leads

Table 3: Table summary of subproblem fidelities.

Fidelity	Subproblem notation	Number of time-steps	Number of subproblems
Daily	$\mathcal{S}_D$	24	365
Weekly	$\mathcal{S}_W$	168	52
Monthly	$\mathcal{S}_M$	720	12
Yearly	$\mathcal{S}_Y$	8760	1

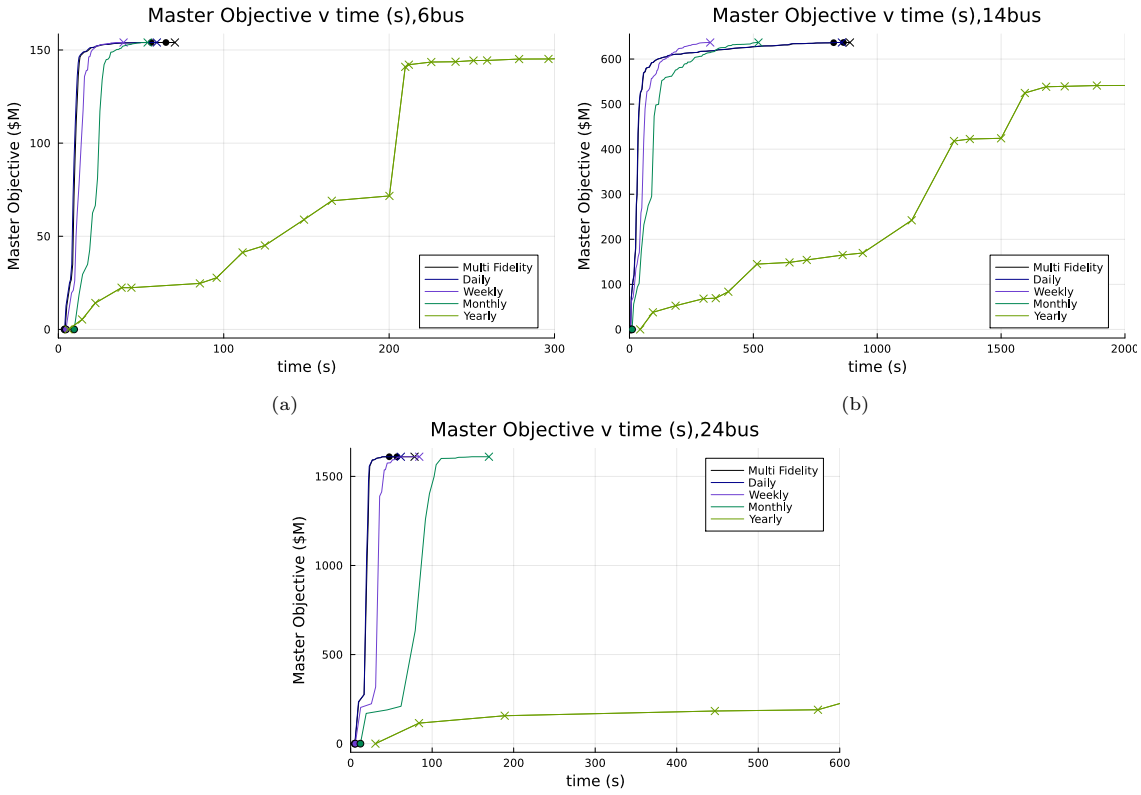

Figure 1: Master objective against time for IEEE scenarios comparing single yearly fidelity, bi-fidelity and adaptive multi-fidelity (Mulit Fidelity) approaches. Crosses indicate an iteration in which the full yearly subproblem was solved, and solid circles indicate where the adaptive approach has switched fidelity. The legend entries indicate which fidelity each approach starts with, apart from Multi Fidelity which starts using daily subproblems.

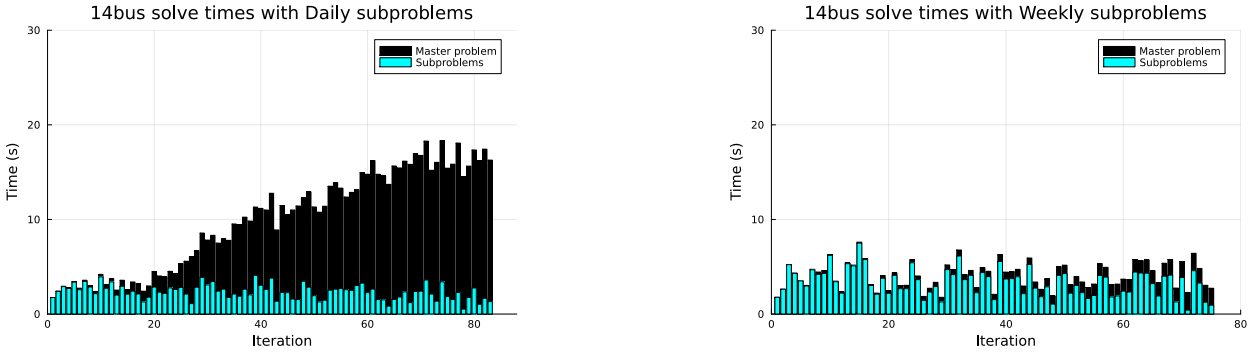

Figure 2: Comparison of time spent solving the master and subproblems of the IEEE 14 bus test case.

to increased complexity and as such increase in solve time, which in some test cases can prove prohibitive. In Fig. 2 we observe this growth in the time spent solving the master problem. Specifically, when solving daily subproblems, Fig. 2a, we observe significant growth in the master solve time, as 365 cuts can be added each iteration. Over many iterations, this growth dominates the computational effort. Comparatively, when using weekly subproblems, Fig. 2b, only 52 cuts are added each iteration, these are stronger and do not take significantly longer to solve and so the optimal solution is obtained before this growth in constraints can become a problem. Combining these observations, we determine that switching based solely on the *Gap*

tolerance, as in the bi-fidelity case, can result in fast growth in the master problem solve times and therefore convergence is heavily reliant on the initial fidelity chosen.

Tables 4 to 6 provide result summaries and solution times for the IEEE and NEM cases. For conciseness, we only provide these results for the adaptive multi-fidelity approach, classic BD, and the monolithic MILP formulation. Across all three IEEE test cases, the adaptive multi-fidelity approach achieves optimality much faster than either the monolithic MILP or the classic BD approach. For the 14 bus test case the monolithic MILP does not find the optimal solution within the wall time. The structure of the IEEE tests cases is

Table 4: Solve times for Classic BD.

	Time (s)			Number of	Cost ($M)	(%)
Test Cases	MP	SP	Total	Iterations	Best LB	Gap
6bus	2.4	581	595	56	154.02	0.0
14bus	5.4	9,504	9,539	97	636.74	0.0
24bus	0.6	1,499	1,513	27	1,610.26	0.0
Central	0.4	12,962	12,977	7	942.03	90.17
Step Change	0.1	12,184	12,197	5	597.10	95.33
Slow Change	0.1	11,947	11,962	7	610.69	87.13
High DER	0.3	10,335	10,352	6	704.05	91.89
Fast Change	0.3	10,248	10,265	7	756.02	92.84

Table 5: Solve times for Adaptive MFBD.

	Time (s)			Number of	Cost ($M)	(%)
Test Cases	MP	SP	Total	Iterations	Best LB	Gap
6bus	38	18	70	40	154.02	0.0
14bus	666	206	890	85	636.74	0.0
24bus	4	57	78	22	1,610.25	0.0
Central	294	9,535	10,008	98	4,335.92	12.5
Step Change	238	10,153	10,524	86	4,246.38	11.9
Slow Change	225	12,103	12,512	89	3,022.15	11.6
High DER	237	10,072	10,479	98	3,817.93	13.7
Fast Change	274	9,779	10,206	89	4,314.81	13.0

Table 6: Monolitic MILP Evolution.

Case	Root relaxation time (s)	Root Solution ($M)	Best Bound ($M)	Gap (%)	Final Time (s)
6 bus	45	146.06	154.02	0	2,511
14 bus	347	576.51	589.04	11.7	10,000
24 bus	304	1,605.68	1,610.26	0	9,423
Central	1,449	4,857.03	4,858.43	6.30	12,570
Step Change	1,525	4,729.12	4,730.53	0.29	10,029
Slow Change	1,605	3,383.87	3,391.39	0.01	10,024
High DER	2,199	4,329.17	4,331.50	6.13	11,116
Fast Change	1,512	4,853.42	4,854.95	6.41	10,677

such that there are no storage assets initially present or as part of the optimal solution. However, during the course of the algorithm, solutions including storage assets are tested. In these test cases, even the multiple fidelity approach can get stuck on low fidelity subproblems, as in the 14 bus case. the switching strategy in this instance is not ideal. In comparison to classic BD solving only a yearly subproblem, however, the

multiple fidelity approach is still much faster. Again, we emphasise that without prior knowledge of how each test case may behave, a generalised approach is preferential.

We then consider the NEM test cases listed in Table 2. The results of which, as shown in Fig. 3, highlight the effectiveness of an adaptive multi-fidelity approach especially as compared to classic BD. In each of the five scenarios, our adaptive approach is comparable to if not faster than the next best methodology at any point in the run time. We notice that the switching points allow the adaptive curve to avoid stalling in convergence and follow the curve of the best fidelity. The inconsistency in performance of bi-fidelity approaches, such as the Daily bi-fidelity curve in Fig. 3d, further highlights the importance of a methodology that does not require pre-knowledge of how well it will perform. Tables 4 and 5 further highlight the significant gain over classic BD, the smaller subproblems from the MFBD approach allow for a much greater number of useful cuts to be added, pushing the lower bound far beyond what is possible with only a single cut each iteration. Each of the approaches in the NEM cases terminated due to the wall time limit of 10,000 seconds being reached. Within this time limit, the Monthly bi-fidelity approach (which is the only method to reach a comparable objective value) does not switch to evaluating yearly subproblems in all but one test case. For expansion planning problems, 10,000 seconds may be considered a short cutoff point, considering the difficulty of the problem and the long planning horizons. However, the MFBD methodology presented can be readily extended to the dynamic expansion planning problem with multi-year time horizon and security constraint expansion planing. Solving an accurate problem down to a reasonable gap quickly is critical for these problems, since they involve many years in the former and many scenarios in the latter case.

Table 6 provides a summary of the key monolithic MILP solving information. Gurobi's barrier solve finds a solution to the root node linear relaxation (first and second columns), which provides a high lower bound, relatively quickly before moving to the branch-and-bound, which then makes little progress. The best bound is the final integer bound Gurobi is able to obtain, and the gap presented is the integrality gap to the incumbent solution. The final bound for the IEEE cases is the optimal solution, as discussed previously, while the NEM cases do not reach the optimal solution. The lower bounds obtained by Gurobi when solving the monolithic MILP, in the NEM cases, are higher than those we obtain using MFBD. However, these are inconsistent across all test cases. Without knowing how well the monolithic MILP will converge on a given problem, there is potential for significant speed up when using our approach.

More detailed result tables can be found in the appendix. These include: problem sizes for each fidelity and the MILP; a breakdown by fidelity of the Adaptive MFBD algorithm solution times; and investment solution summary information of the Adaptive MFBD algorithm for each test case.

Fig. 4 highlights the effect of relaxation on subproblem cost as the algorithm converges. After each fidelity has converged sufficiently, and we switch to the next, there is a clear spike in subproblem cost as the formulation is tightened through the addition of previously relaxed linking constraints. This objective then decreases again as new cuts are added, and alternative investment solutions are tried.

The analogue of this is Fig. 5 showcasing the jump in investment cost as new solutions must be investigated given the tighter formulation. From this we also see that the solutions explored in the lower-fidelity iterations are, for the most part, not revisited once the subproblems have moved to higher-fidelities. Within each fidelity, the addition of the investment cost and subproblem objectives provide an upper bound on the optimal solution cost at that fidelity. When yearly problems are considered, this provides an upper bound to the optimal solution to the monolithic problem and is analogous to the incumbent solution provided by Gurobi. Hence, these graphs provide a higher solution cost than those presented in Fig. 3. When the optimal solution relative to the yearly subproblem is reached, then the subproblem cost and investment cost add to give the same cost as the optimal monolithic problem and the same cost as the optimal master solution (with all facet defining cuts added).

6. Conclusion

We demonstrate on GSTEP a modified BD strategy, MFBD, focused on balancing time between cut generation and solving the master problem. To obtain this scheme, we require small relaxations across temporal linking constraints to generate independent subproblems over multiple fidelities. The generated low-fidelity subproblems are fast to solve and generate more cuts each iteration, these cuts represent a lower bound. Then high-fidelity subproblems, that accurately represent the problem but which can take prohibitively long when used in isolation, are solved. We observe that by utilising a switching mechanism,

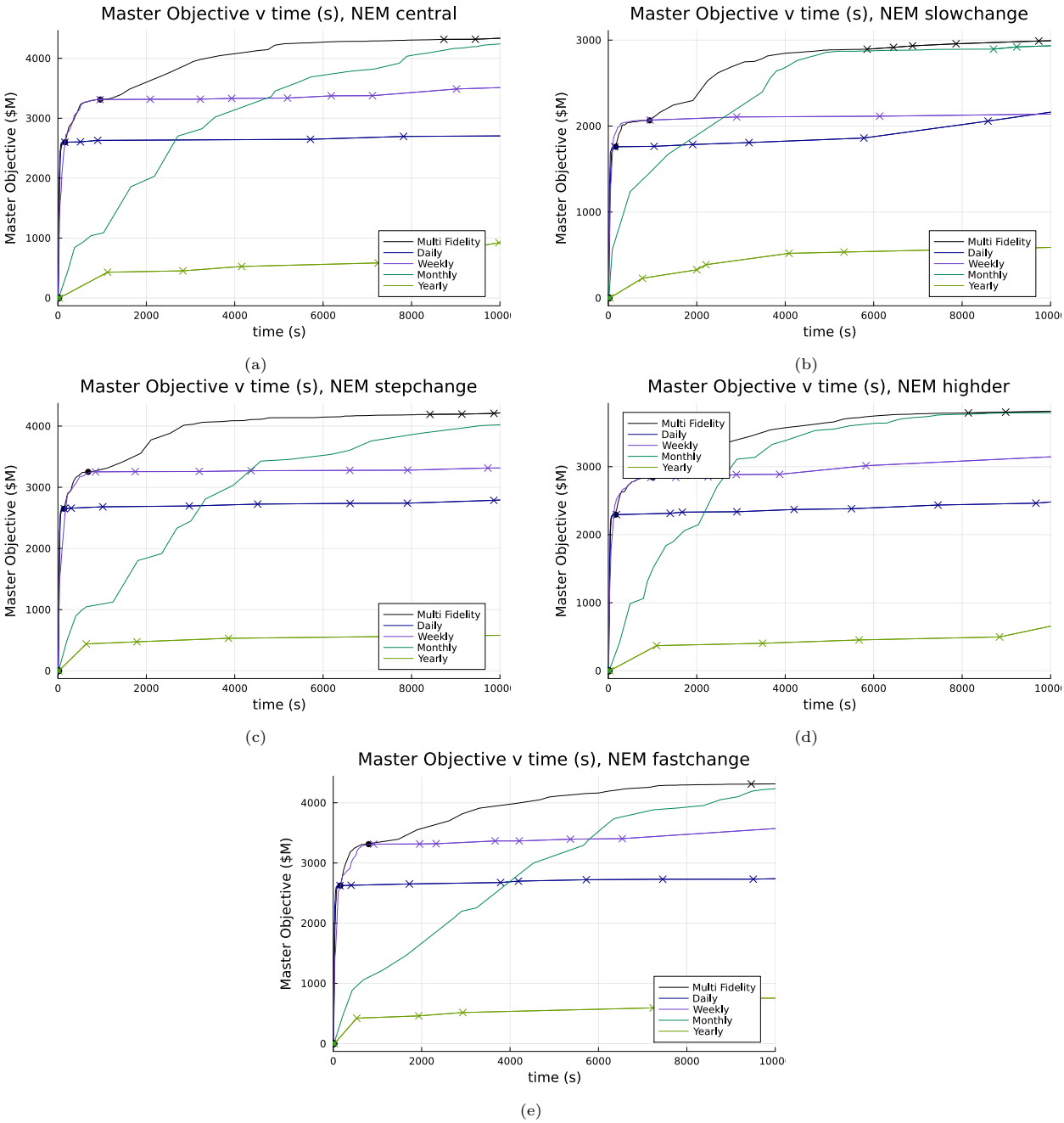

Figure 3: Master objective against time for NEM scenarios comparing single yearly fidelity, bi-fidelity and adaptive multi-fidelity approaches. Crosses indicate an iteration in which the full yearly subproblem was solved, and solid circles indicate where the adaptive approach has switched fidelity. The legend entries indicate which fidelity each approach started with, apart from adaptive which starts using daily subproblems

to adaptively change fidelity of the subproblem being solved, a speed-up can be obtained for larger test cases when compared to classic Benders Decomposition. This can be attributed to the inevitable slow-down caused by adding a multitude of cuts to the master when solving low-fidelity problems. By striking this balance we avoid such slowdown which can, in certain cases, be catastrophic. We also include a theorem with proof that the presented MFBD algorithm will converge to the optimum of GSTEP when the full-time-horizon subproblem is solved for all facet defining cuts. Ultimately, the presented switching strategy is one

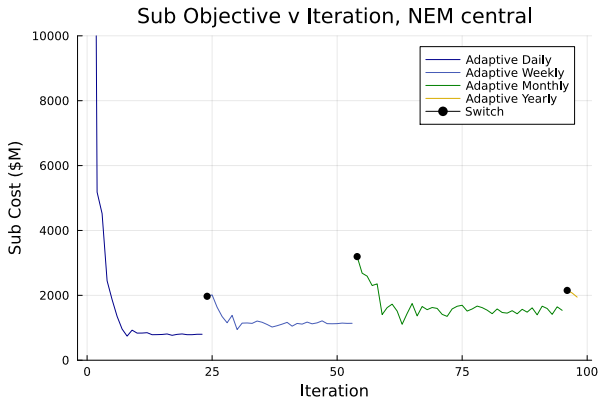

Figure 4: Subproblem objective function against iteration of the adaptive multifidelity approach.

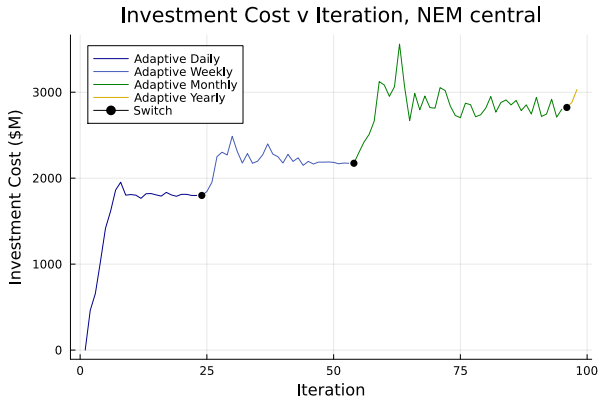

Figure 5: Investment Cost against iteration of the adaptive multi-fidelity approach.

of many possible alternatives, and further work will include refinement of the adaptive-switching rule for best performance across a wider expanse of data sets. Future work will focus on expanding and solidifying a suite of standardised test datasets such that comparison between existing methodologies is possible.

Acknowledgements

The authors would like to pay special tribute to Professor Ariel Liebman (1968—2023). This research is supported in part through a Monash research scholarship (top-up) funded by Russell and Jenny Tait.

Appendix A. Mathematical foundation of optimality using Multi-Fidelity Benders Decomposition

To demonstrate the general validity of the approach proposed here, consider the following MILP that is structurally similar to the problem of GSTEP. With some investment cost c for integer investment variables x, operational cost d for continuous operational variables y (combining f, θ, p, q and v variables from Table 1) and constraint matrices A, B and E.

$$\begin{aligned} \min_{x,y} \quad & c^\top x + d^\top y \\ \text{s.t.} \quad & Ax + By \leq b, \\ & Ey = \mathbf{0}, \\ & y \geq \mathbf{0}, \\ & x \in \mathbb{Z}_2^n \end{aligned} \tag{A.1}$$

Here the $Ey = \mathbf{0}$ consists of the linking constraints (14). The remaining constraints of (5)-(9) are part of the $Ax + By \leq b$ inequalities for suitable A, B and b.

Under Benders Decomposition the master problem for the above is:

$$\begin{aligned} \min_{x,\,\alpha} \; & c^\top x + \alpha \\ \text{s.t.} \quad & \alpha \geq \mathbf{0}, \\ & x \in \mathbb{Z}_2^n \end{aligned} \tag{A.2}$$

with associated subproblem:

$$\begin{aligned} z(\bar{x}) = \min_{y} \; & d^\top y \\ \text{s.t.} \quad & By \leq b - A\bar{x}, \\ & Ey = \mathbf{0}, \\ & y \geq \mathbf{0}. \end{aligned} \tag{A.3}$$

The dual for this is given by:

$$\begin{aligned} k(\bar{x}) = \max_{\pi,\,\lambda} \; & (b - A\bar{x})^\top \pi + \mathbf{0}^\top \lambda \\ \text{s.t.} \quad & B^\top \pi + E^\top \lambda \leq d, \\ & \lambda \text{ unrestricted}, \qquad \pi \leq \mathbf{0} \end{aligned} \tag{A.4}$$

The proof that master problem (A.2) with primal & dual subproblem (A.3) & (A.4) are equivalent to MILP (A.1) is given by M. Conforti *et al.* [35], under the assumption that all Benders cuts obtained from LP (A.4) are added to the master (A.2).

The proposed multi-fidelity scheme begins by solving a relaxed version of the operational GSTEP problem analogous to removing the requirement that $Ey = \mathbf{0}$ in problem (A.3) and thus giving:

$$\begin{aligned} z_r(\bar{x}) = \min_{y} \; & d^\top y \\ \text{s.t.} \quad & By \leq b - A\bar{x}, \\ & y \geq \mathbf{0} \end{aligned} \tag{A.5}$$

Which has an associated dual given by:

$$\begin{aligned} k_r(\bar{x}) = \max_{\pi} \; & (b - A\bar{x})^\top \pi \\ \text{s.t.} \quad & B^\top \pi \leq d, \\ & \pi \leq \mathbf{0} \end{aligned} \tag{A.6}$$

Lemma 1. *If the primal problem is feasible and bounded, the dual problem is bounded.*

Theorem 2. *Consider a bounded, feasible, mixed-integer linear program of the form in problem (A.1) with Benders Decomposition and dual as given by (A.2) and (A.4). The cut generated by solving the relaxed problem (A.5) and its dual (A.6) constitute a valid lower bound for the unrelaxed problem, and as such cannot cut off an optimal solution.*

Proof. As all problems are feasible and bounded, we know there is an optimal solution for each problem. We denote these optimal solutions as $x^{(n)}, y^{(n)}$ for primal problem (n) and $\pi^{(m)}, \lambda^{(m)}$ for dual problem (m). The *optimal* solution for dual problem (A.6), labelled as $\pi^{(A.6)}$, is a *feasible* solution of dual problem (A.4) choosing $\lambda = \mathbf{0}$.

By definition, the *optimal* solution for dual problem (A.4), $(\pi^{(A.4)}, \lambda^{(A.4)})$, will give a higher objective value than the *feasible* solution $(\pi^{(A.6)}, \lambda = \mathbf{0})$.

We therefore have for any *feasible* solution $(\bar{x}, \bar{y})$ of problem (A.1) and some *optimal* solution $y^{(A.3)}$ of

problem (A.3) that

$$\begin{aligned} c^\top \bar{x} + d^\top \bar{y} &\geq c^\top \bar{x} + d^\top y^{(A.3)} \\ &\overset{\text{S.D}}{=} (b - A\bar{x})^\top \pi^{(A.4)} + 0^\top \lambda^{(A.4)} \\ &\geq (b - A\bar{x})^\top \pi^{(A.6)} + 0^\top \mathbf{0} \\ &= (b - A\bar{x})^\top \pi^{(A.6)}. \end{aligned}$$

S.D indicates that the equality holds due to strong duality. The above inequality demonstrates that the cuts generated from solving the relaxed dual problem, (A.6) are a lower bound for problem (A.3). □

Appendix B. Test case sizes

Table B.7: MILP and LP sizes for each test case (before presolve).

		Rows	Continuous	Integer	Total	Non-zeros
6bus	Monolithic MILP	770,910	420,851	24	420,875	2,119,376
	Master	24	365	24	389	24
	Subproblem (Daily)	2,142	1,182	0	1,182	5,846
	Subproblem (Weekly)	14,814	8,094	0	8,094	40,696
	Subproblem (Monthly)	63,390	34,590	0	34,590	174,278
	Subproblem (Yearly)	770,910	420,510	0	420,510	2,119,376
14bus	Monolithic MILP	1,681,976	806,293	48	806,341	4,717,231
	Master	48	365	48	413	48
	Subproblem (Daily)	4,664	2,264	0	2,264	12,994
	Subproblem (Weekly)	32,312	15,512	0	15,512	90,574
	Subproblem (Monthly)	138,296	66,296	0	66,296	387,964
	Subproblem (Yearly)	1,681,976	805,976	0	805,976	4,717,231
24bus	Monolithic MILP	2,601,806	1,279,337	74	1,279,411	7,155,160
	Master	74	365	74	439	74
	Subproblem (Daily)	7,214	3,590	0	3,590	19,710
	Subproblem (Weekly)	49,982	24,614	0	24,614	137,352
	Subproblem (Monthly)	213,926	105,206	0	105,206	588,284
	Subproblem (Yearly)	2,601,806	1,279,046	0	1,279,046	7,155,160
5bus NEM	Monolithic MILP	4,906,325	4,424,253	324	4,424,577	13,717,781
	Master	637	365	324	689	637
	Subproblem (Daily)	13,852	12,532	0	12,532	37,835
	Subproblem (Weekly)	94,492	85,252	0	85,252	262,583
	Subproblem (Monthly)	403,612	364,012	0	364,012	1,123,399
	Subproblem (Yearly)	4,906,012	4,424,212	0	4,424,212	13,717,468

Table B.7 gives a summary of the sizes, when given to Gurobi, in terms of number of rows, columns and non-zero coefficients for the linear and mixed integer programs. All subproblems listed are LPs, the master is a MILP.

Table B.8 provides the algorithm evolution information broken down by fidelity being solved for the Adaptive MFBD approach. For each test case, the time spent solving the master problem, subproblem, and the total time in each fidelity is provided. The number of iterations of each fidelity as well as the final lower bound obtained in each fidelity are also presented. The gap for these lower bounds is relative to the UB found in that specific fidelity. As such, the only globally valid gap and UB are those of the Yearly subproblems.

References

[1] M. V. F. Pereira, L. M. V. G. Pinto, S. H. F. Cunha, G. C. Oliveira, A decomposition approach to automated generation/transmission expansion planning, IEEE Transactions on Power Apparatus and Systems PAS-104 (11) (1985) 3074–83. doi:10.1109/TPAS.1985.318815.

Table B.8: Solve times for Adaptive MFBD by fidelity.

		Time (s)			Number of	Cost ($M)	(%)
Test Cases	Fidelity	MP	SP	Total	Iterations	Best LB	Gap
6bus	Daily	32	15	53	37	154.0	0.02*
	Weekly	2	0.3	3	1	154.02	0.0*
	Monthly	2	0.4	9	1	154.02	0.0*
	Yearly	2	3	5	1	154.02	0.0
14bus	Daily	604	194	807	81	635.91	0.13*
	Weekly	30	2	34	2	636.73	0.0*
	Monthly	16	1	23	1	636.74	0.0*
	Yearly	16	10	26	1	636.74	0.0
24bus	Daily	3	32	44	19	1,610.19	0.0*
	Weekly	0.3	2	3	1	1,610.25	0.0*
	Monthly	0.3	3	9	1	1,610.25	0.0*
	Yearly	0.3	21	22	1	1,610.25	0.0
Central	Daily	21	77	153	23	2,599.07	0.01*
	Weekly	82	663	788	30	3,311.74	0.0*
	Monthly	177	7,496	7,751	32	4,315.61	0.38*
	Yearly	14	1,299	1,316	3	4,335.92	12.5
Step Change	Daily	16	70	133	20	2,694.07	0.0*
	Weekly	61	443	534	25	3,251.36	0.0*
	Monthly	144	7,519	7,713	37	4,190.83	0.6*
	Yearly	17	2,120	2,144	4	4,246.38	11.9
Slow Change	Daily	16	97	163	21	1,759.09	0.0*
	Weekly	92	612	751	33	2065.74	0.2*
	Monthly	95	4,761	4,905	29	2,863.9	0.1*
	Yearly	23	6,633	6,693	6	3,022.15	11.6
High DER	Daily	20	89	155	25	2.294.88	0.0*
	Weekly	87	693	837	36	2.840.35	0.0*
	Monthly	115	6,942	7,116	34	3,788.91	0.4*
	Yearly	15	2,348	2,372	3	3,817.93	13.7
Fast Change	Daily	20	80	154	22	2,624.07	0.0*
	Weekly	82	500	630	27	3,312.55	0.1*
	Monthly	161	8,426	8,636	38	4,311.15	0.0*
	Yearly	11	773	787	2	4,314.81	13.0

*These gaps are **relative** to the fidelity of the problem being solved, not a valid upper bound of the master with yearly subproblem.

Table B.9: 6 bus Investment Decisions

Asset Type	Capacity Investment (MW)			Investment Cost ($M)		
	CBD	MFBD	MILP	CBD	MFBD	MILP
Lines	200	200	200	0.6	0.6	0.6
Generators	200	200	200	33.5	33.5	33.5
Renewables	115	115	115	12.2	12.2	12.2
Storage	0	0	0	0	0	0
Total	515	515	515	46.3	46.3	46.3

[2] S. Binato, M. Pereira, S. Granville, A new benders decomposition approach to solve power transmission network design problems, Power Systems, IEEE Transactions on 16 (2001) 235–40. doi:10.1109/59.918292.

[3] S. Lumbreras, A. Ramos, Transmission expansion planning using an efficient version of benders' decom-

Table B.10: 14 bus Investment Decisions

Asset Type	Capacity Investment (MW)			Investment Cost ($M)		
	CBD	MFBD	MILP	CBD	MFBD	MILP
Lines	570	570	760	2.8	2.8	3.2
Generators	870	870	870	151.6	151.6	151.6
Renewables	270	270	150	32.3	32.3	16.0
Storage	0	0	0	0	0	0
Total	1,710	1,710	1,780	186.7	186.7	170.8

Table B.11: 24 bus Investment Decisions

Asset Type	Capacity Investment (MW)			Investment Cost ($M)		
	CBD	MFBD	MILP	CBD	MFBD	MILP
Lines	825	825	825	4.8	4.8	4.8
Generators	1,850	1,850	1,850	324.8	324.8	324.8
Renewables	545	545	545	58	58	58
Storage	0	0	0	0	0	0
Total	3,220	3,220	3,220	387.7	387.7	387.7

Table B.12: NEM Central Investment Decisions

Asset Type	Capacity Investment (MW)			Investment Cost ($M)		
	CBD	MFBD	MILP	CBD	MFBD	MILP
Lines	0	170	0	0.0	6.6	0.0
Generators	6,600	4,900	6,100	459.4	619.1	872.5
Renewables	2	21,073	18,423	0.1	1,912.9	1,730.5
Storage	64,000	53,200	36,600	482.5	493.3	325.5
Total	70,602	79,343	61,123	942.0	3031.9	2928.4

Table B.13: NEM Step Change Investment Decisions

Asset Type	Capacity Investment (MW)			Investment Cost ($M)		
	CBD	MFBD	MILP	CBD	MFBD	MILP
Lines	2100	170	0	114.8	6.6	0.0
Generators	5,700	5,000	4,850	395.6	466.8	525.1
Renewables	10	23,667	24,603	0.6	1,988.3	2,193.9
Storage	12,200	48,000	51,000	86.1	573.0	454.4
Total	20,010	76,837	80,453	597.1	3,034.7	3,173.4

Table B.14: NEM Slow Change Investment Decisions

Asset Type	Capacity Investment (MW)			Investment Cost ($M)		
	CBD	MFBD	MILP	CBD	MFBD	MILP
Lines	170	170	170	8.5	8.5	8.5
Generators	2,700	4,000	5,050	229.3	473.2	695.7
Renewables	0	5,088	5,218	0.0	527.3	516.8
Storage	35,800	72,000	0	372.9	333.7	0.0
Total	38,670	81,258	10,438	610.7	1,342.8	1,221.0

position. a case study, in: 2013 IEEE Grenoble Conference, 2013, pp. 1–7. doi:10.1109/PTC.2013.6652091.

[4] S. Huang, V. Dinavahi, A branch-and-cut benders decomposition algorithm for transmission expansion

Table B.15: NEM High DER Investment Decisions

Asset Type	Capacity Investment (MW)			Investment Cost ($M)		
	CBD	MFBD	MILP	CBD	MFBD	MILP
Lines	0	170	0	0.0	6.6	0.0
Generators	8,100	3,750	5,700	560.9	427.0	833.8
Renewables	0	19,105	17,122	0.0	1,709.1	1,642.1
Storage	43,800	47,600	28,200	143.1	405.9	249.7
Total	51,900	70,625	51,022	704.0	2,548.5	2,725.6

Table B.16: NEM Fast Change Investment Decisions

Asset Type	Capacity Investment (MW)			Investment Cost ($M)		
	CBD	MFBD	MILP	CBD	MFBD	MILP
Lines	0	800	0	0.0	124.5	0.0
Generators	10,200	4,950	6,400	706.3	680.4	893.3
Renewables	7	17,999	18,069	0.9	1,708.7	1,762.0
Storage	11,400	45,600	36,600	48.8	413.9	325.1
Total	21,607	69,349	61,069	756.0	2,927.5	2,980.4

planning, IEEE Systems Journal 13 (1) (2019) 659–69. doi:10.1109/JSYST.2017.2775610.

[5] M. Larsen, E. Sauma, Economic and emission impacts of energy storage systems on power-system long-term expansion planning when considering multi-stage decision processes, Journal of Energy Storage 33 (2021) 101883. doi:https://doi.org/10.1016/j.est.2020.101883.
URL https://www.sciencedirect.com/science/article/pii/S2352152X20317205

[6] M. Moradi-Sepahvand, T. Amraee, Integrated expansion planning of electric energy generation, transmission, and storage for handling high shares of wind and solar power generation, Applied Energy 298 (2021) 117137. doi:https://doi.org/10.1016/j.apenergy.2021.117137.
URL https://www.sciencedirect.com/science/article/pii/S0306261921005778

[7] D. S. Mallapragada, D. J. Papageorgiou, A. Venkatesh, C. L. Lara, I. E. Grossmann, Impact of model resolution on scenario outcomes for electricity sector system expansion, Energy 163 (2018) 1231–44. doi:https://doi.org/10.1016/j.energy.2018.08.015.
URL https://www.sciencedirect.com/science/article/pii/S0360544218315238

[8] C. L. Lara, D. S. Mallapragada, D. J. Papageorgiou, A. Venkatesh, I. E. Grossmann, Deterministic electric power infrastructure planning: Mixed-integer programming model and nested decomposition algorithm, European Journal of Operational Research 271 (3) (2018) 1037–54. doi:https://doi.org/10.1016/j.ejor.2018.05.039.
URL https://www.sciencedirect.com/science/article/pii/S0377221718304466

[9] K. Yagi, R. Sioshansi, Nested benders's decomposition of capacity-planning problems for electricity systems with hydroelectric and renewable generation, Computational Management Science 21 (01 2024). doi:10.1007/s10287-023-00469-9.

[10] J. Ma, V. Silva, R. Belhomme, D. S. Kirschen, L. F. Ochoa, Evaluating and planning flexibility in sustainable power systems, in: 2013 IEEE Power Energy Society General Meeting, 2013, pp. 1–11. doi:10.1109/PESMG.2013.6672221.

[11] T. Levin, P. L. Blaisdell-Pijuan, J. Kwon, W. N. Mann, High temporal resolution generation expansion planning for the clean energy transition, Renewable and Sustainable Energy Transition 5 (2024) 100072. doi:https://doi.org/10.1016/j.rset.2023.100072.
URL https://www.sciencedirect.com/science/article/pii/S2667095X23000284

[12] S. Wijekoon, A. Liebman, A. Aleti, S. Dunstall, Enhanced generation, energy storage and transmission expansion planning for renewables with operational flexibility through unit commitment, in: 2019 Modern Electric Power Systems (MEPS), 2019, pp. 1–6. doi:10.1109/MEPS46793.2019.9395031.

[13] R. Romero, A. Monticelli, A hierarchical decomposition approach for transmission network expansion planning, IEEE Transactions on Power Systems 9 (1) (1994) 373–80. doi:10.1109/59.317588.

[14] N. G. Ude, H. Yskandar, R. G. Coneth, A comprehensive state-of-the-art survey on the transmission network expansion planning optimization algorithms, IEEE Access 7 (2019) 123158–123181. doi:10.1109/ACCESS.2019.2936682.

[15] M. Mahdavi, C. Sabillon Antunez, M. Ajalli, R. Romero, Transmission expansion planning: Literature review and classification, IEEE Systems Journal 13 (3) (2019) 3129–40. doi:10.1109/JSYST.2018.2871793.

[16] R. Villasana, L. L. Garver, S. J. Salon, Transmission network planning using linear programming, IEEE Transactions on Power Apparatus and Systems PAS-104 (2) (1985) 349–56. doi:doi:10.1109/TPAS.1985.319049.

[17] L. L. Garver, Transmission network estimation using linear programming, IEEE Transactions on Power Apparatus and Systems PAS-89 (7) (1970) 1688–97. doi:doi:10.1109/TPAS.1970.292825.

[18] R. Romero, A. Monticelli, A. Garcia, S. Haffner, Test systems and mathematical models for transmission network expansion planning, IEE Proceedings-Generation, Transmission and Distribution 149 (1) (2002) 27–36.

[19] L. Bahiense, G. C. Oliveira, M. Pereira, S. Granville, A mixed integer disjunctive model for transmission network expansion, IEEE Transactions on Power Systems 16 (3) (2001) 560–5. doi:10.1109/59.932295.

[20] J. F. Benders, Partitioning procedures for solving mixed-variables programming problems, Numerische mathematik 4 (1) (1962) 238–52.

[21] R. Rahmaniani, T. G. Crainic, M. Gendreau, W. Rei, The benders decomposition algorithm: A literature review, European Journal of Operational Research 259 (3) (2017) 801–17.

[22] M. Jenabi, S. Fatemi Ghomi, S. A. Torabi, S. H. Hosseinian, Acceleration strategies of benders decomposition for the security constraints power system expansion planning, Annals of Operations Research 235 (2015) 337–69.

[23] A. Ruszczyński, Decomposition methods, in: Stochastic Programming, Vol. 10 of Handbooks in Operations Research and Management Science, Elsevier, 2003, pp. 141–211. doi:https://doi.org/10.1016/S0927-0507(03)10003-5.
URL https://www.sciencedirect.com/science/article/pii/S0927050703100035

[24] T. N. Santos, A. L. Diniz, C. L. T. Borges, A new nested benders decomposition strategy for parallel processing applied to the hydrothermal scheduling problem, IEEE Transactions on Smart Grid 8 (3) (2017) 1504–12. doi:10.1109/TSG.2016.2593402.

[25] S. Lumbreras, A. Ramos, A benders' decomposition approach for optimizing the electric system of offshore wind farms, in: 2011 IEEE Trondheim PowerTech, 2011, pp. 1–8. doi:10.1109/PTC.2011.6019371.

[26] S. Lumbreras, A. Ramos, S. Cerisola, A progressive contingency incorporation approach for stochastic optimization problems, IEEE Transactions on Power Systems 28 (2) (2013) 1452–60. doi:10.1109/TPWRS.2012.2225077.

[27] C. Li, A. J. Conejo, P. Liu, B. P. Omell, J. D. Siirola, I. E. Grossmann, Mixed-integer linear programming models and algorithms for generation and transmission expansion planning of power systems, European Journal of Operational Research 297 (3) (2022) 1071–82. doi:https://doi.org/10.1016/j.ejor.2021.06.024.
URL https://www.sciencedirect.com/science/article/pii/S0377221721005397

[28] I. E. Grossmann, F. Trespalacios, Systematic modeling of discrete-continuous optimization models through generalized disjunctive programming, AIChE Journal 59 (9) (2013) 3276–3295.

[29] S. Wijekoon, Solution methods for optimal energy system planning with renewable and non-renewable generation (2020). doi:10.26180/5eddb6216e35f.
URL https://bridges.monash.edu/articles/thesis/Solution_Methods_for_Optimal_Energy_System_Planning_with_Renewable_and_Non-Renewable_Generation/12444269/1

[30] S. Wijekoon, A. Liebman, S. Dunstall, Test cases for capacity expansion planning with unit commitment (2019). doi:10.26180/5cdcc2c988948.
URL https://bridges.monash.edu/articles/online_resource/Test_Cases_for_Capacity_Expansion_Planning_with_Unit_Commitment/8135573/1

[31] AEMO, 2020 integrated system plan for the national electricity market, Tech. rep., Australian Energy Market Operator (AEMO) (2020).
URL https://aemo.com.au/en/energy-systems/major-publications/integrated-system-plan-isp/2020-integrated-system-plan-isp

[32] D. Mitchell, Dataset for GSTEP testing (9 2025). doi:10.26180/30227410.v1.
URL https://bridges.monash.edu/articles/dataset/Dataset_for_GSTEP_testing/30227410

[33] S. Binato, M. Pereira, S. Granville, A new benders decomposition approach to solve power transmission network design problems, IEEE Transactions on Power Systems 16 (2) (2001) 235–240. doi:10.1109/59.918292.

[34] L. S. Moulin, M. Poss, C. Sagastizábal, Transmission expansion planning with re-design, Energy systems 1 (2) (2010) 113–139.

[35] M. Conforti, G. Cornuéjols, G. Zambelli, Integer programming models, Springer, 2014, Ch. 8, pp. 321–49.