Supplementary Material for 'Scalable Multi-Level Optimization for Sequentially Cleared Energy Markets with a Case Study on Gas and Carbon Aware Unit Commitment'

Appendix D. Proof of Theorem 3.1

Proof. We denote the unbounded and finite of Problem (i) as $\mathcal{U}_i$ and $\mathcal{F}_i$ for $i \in \{5, 6, 7\}$, respectively. The combination of outcomes of BSP subproblems (6) and (7) are $\mathcal{A} = \{(\mathcal{U}_6, \mathcal{U}_7), (\mathcal{U}_6, \mathcal{F}_7), (\mathcal{F}_6, \mathcal{U}_7), (\mathcal{F}_6, \mathcal{F}_7)\}$. Similarly, the possible outcomes of BSP Problem (5) is $\mathcal{B} = \{\mathcal{U}_5, \mathcal{F}_5\}$. This proof is to show that the solution of BSP Problem (5) can be determined by solving BSP subproblems (6) and (7), i.e., there is a surjective mapping $\mathcal{A} \to \mathcal{B}$ from the outcomes of (6) and (7) $\mathcal{A}$ to the outcomes of BSP Problem (5) $\mathcal{B}$ (Byeon and Van Hentenryck, 2022). For clarity and brevity, we denote the variables $(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n, \boldsymbol{u}_{y1}, \ldots, \boldsymbol{u}_{yn}, \boldsymbol{v}_{y1}, \ldots, \boldsymbol{v}_{yn}, \boldsymbol{u})$ as $(\boldsymbol{x}, \boldsymbol{u}_y, \boldsymbol{v}_y, \boldsymbol{u})$ and $(\boldsymbol{y}_1, \ldots, \boldsymbol{y}_n, \boldsymbol{u}_{x1}, \ldots, \boldsymbol{u}_{xn}, \boldsymbol{w}) = (\boldsymbol{y}, \boldsymbol{u}_x, \boldsymbol{w})$.

1. If $\mathcal{U}_6$: We denote the unbounded ray of Problem (6) as $(\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{u}}_y, \tilde{\boldsymbol{v}}_y, \tilde{\boldsymbol{u}})$ and $\mathcal{U}_6(\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{u}}_y, \tilde{\boldsymbol{v}}_y, \tilde{\boldsymbol{u}}) := \sum_{i=1}^n \big(\gamma_i c_i^T \tilde{\boldsymbol{x}}_i - \gamma_i \tilde{\boldsymbol{u}}_{yi}^T (b_{yi} - H_{yi}\hat{\boldsymbol{z}}) - \gamma_i \tilde{\boldsymbol{v}}_{yi}^T (e_{yi} + E_{yi}\hat{\boldsymbol{z}})\big) - \tilde{\boldsymbol{u}}^T (q_y - R_y \hat{\boldsymbol{z}}) < 0$. Note that $\mathcal{U}_6 = \infty$ and without the loss of generality, we can assume the variable $\boldsymbol{w}$ to be zero in Problem (7). Then,
   (a) If $\mathcal{U}_7$: Let $(\tilde{\boldsymbol{y}}, \tilde{\boldsymbol{u}}_x, 0)$ be the unbounded ray of Problem (7). Then $\tilde{\mu}_1 := (\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{u}}_y, \tilde{\boldsymbol{v}}_y, \tilde{\boldsymbol{u}}, \boldsymbol{0}, \boldsymbol{0}, \boldsymbol{0})$ and $\tilde{\mu}_2 := (\boldsymbol{0}, \boldsymbol{0}, \boldsymbol{0}, \boldsymbol{0}, \tilde{\boldsymbol{y}}_n, \tilde{\boldsymbol{u}}_x, 0)$ are respective an unbounded ray of Problem (5). In this case, Problem (5) is unbounded.
   (b) Else if $\mathcal{F}_7$: Let $(\hat{\boldsymbol{y}}, \hat{\boldsymbol{u}}_x, 0)$ be the optimal solution of Problem (7) with optimal objective value as $\mathfrak{D}_7$. Then for any $\alpha > 0$, $(\boldsymbol{0}, \boldsymbol{0}, \boldsymbol{0}, \boldsymbol{0}, \hat{\boldsymbol{y}}, \hat{\boldsymbol{u}}_x, 0) + \alpha(\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{u}}_y, \tilde{\boldsymbol{v}}_y, \tilde{\boldsymbol{u}}, \boldsymbol{0}, \boldsymbol{0}, \boldsymbol{0})$ is a feasible solution to Problem (5) with objective value as $\mathfrak{D}_7 - \alpha\mathcal{U}_6$ which increases as α increases. Therefore, $\tilde{\mu}_1 := (\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{u}}_y, \tilde{\boldsymbol{v}}_y, \tilde{\boldsymbol{u}}, \boldsymbol{0}, \boldsymbol{0}, \boldsymbol{0})$ is also an unbounded ray for Problem (5).
2. Else If $\mathcal{F}_6$: We denote the optimal solution of Problem (6) as $(\hat{\boldsymbol{x}}, \hat{\boldsymbol{u}}_y, \hat{\boldsymbol{v}}_y, \hat{\boldsymbol{u}})$ and let $\mathfrak{D}_6(\hat{\boldsymbol{x}}, \hat{\boldsymbol{u}}_y, \hat{\boldsymbol{v}}_y, \hat{\boldsymbol{u}}) := \sum_{i=1}^n \big(\gamma_i c_i^T \hat{\boldsymbol{x}}_i - \gamma_i \hat{\boldsymbol{u}}_{yi}^T (b_{yi} - H_{yi}\hat{\boldsymbol{z}}) - \gamma_i \hat{\boldsymbol{v}}_{yi}^T (e_{yi} + E_{yi}\hat{\boldsymbol{z}})\big) - \hat{\boldsymbol{u}}^T (q_y - R_y \hat{\boldsymbol{z}})$ be the optimal objective value. Then,
   (a) If $\mathcal{U}_7$: Let $(\tilde{\boldsymbol{y}}, \tilde{\boldsymbol{u}}_x, \tilde{\boldsymbol{w}})$ be the unbounded ray of Problem (7). Then $\tilde{\mu}_3 := (\hat{\boldsymbol{x}}\tilde{\boldsymbol{w}}, \hat{\boldsymbol{u}}_y\tilde{\boldsymbol{w}}, \hat{\boldsymbol{v}}_y\tilde{\boldsymbol{w}}, \hat{\boldsymbol{u}}\tilde{\boldsymbol{w}}, \tilde{\boldsymbol{y}}, \tilde{\boldsymbol{u}}_x, \tilde{\boldsymbol{w}})$ is a feasible solution of Problem (5) and the objective value $\sum_{i=1}^n \tilde{\boldsymbol{y}}_i^T (b_i - D_i \hat{\boldsymbol{z}}) + \tilde{\boldsymbol{u}}_{xi}^T (b_{xi} - H_{xi}\hat{\boldsymbol{z}}) - \tilde{\boldsymbol{w}}\mathfrak{D}_6$ is positive because $\tilde{\mu}_3$ yields a positive objective value indicates an unbounded direction in Problem (5). Therefore, $\tilde{\mu}_3$ is an unbounded ray of Problem (5) and Problem (5) is unbounded.
   (b) Else if $\mathcal{F}_7$: Let $(\hat{\boldsymbol{y}}, \hat{\boldsymbol{u}}_x, \hat{\boldsymbol{w}})$ be the optimal solution of Problem (7) with optimal objective value as $\mathfrak{D}_7$. First, note that $\hat{\mu}_4 := (\hat{\boldsymbol{x}}\hat{\boldsymbol{w}}, \hat{\boldsymbol{u}}_y\hat{\boldsymbol{w}}, \hat{\boldsymbol{v}}_y\hat{\boldsymbol{w}}, \hat{\boldsymbol{u}}\hat{\boldsymbol{w}}, \hat{\boldsymbol{y}}, \hat{\boldsymbol{u}}_x, \hat{\boldsymbol{w}})$ is a feasible solution of Problem (5). Suppose $\boldsymbol{w} > 0$, we introduce auxiliary variables $(\boldsymbol{x}', \boldsymbol{u}'_y, \boldsymbol{v}'_y, \boldsymbol{u}') = (\frac{\boldsymbol{x}}{\boldsymbol{w}}, \frac{\boldsymbol{u}_y}{\boldsymbol{w}}, \frac{\boldsymbol{v}_y}{\boldsymbol{w}}, \frac{\boldsymbol{u}}{\boldsymbol{w}})$. Then Problem (5) can be rewritten as

$$\max_{\boldsymbol{w}>0} \mathfrak{D}(\boldsymbol{w}) \tag{B:1}$$

   where

$$\mathfrak{D}(w) := \max_{\boldsymbol{y}\geq 0, \boldsymbol{u}_x \geq 0} \sum_{i=1}^n \boldsymbol{y}_i^T (b_i - D_i \hat{\boldsymbol{z}}) + \boldsymbol{u}_{xi}^T (b_{xi} - H_{xi}\hat{\boldsymbol{z}}) - \boldsymbol{w}\mathfrak{D}_6$$
$$\text{s.t.} \quad \text{Eq. (7b)–Eq. (7c)}$$

   Note that Problem $\max_{w>0} \mathfrak{D}(w)$ (B:1) is equivalent to Problem (7) where the nonnegativity constraint for $\boldsymbol{w}$ is restricted by strict inequality. Therefore, we have $\max_{w>0} \mathfrak{D}(w) \leq \mathfrak{D}_7$. When $\boldsymbol{w} = 0$, Problem (5) can be decomposed into Problem (7) with a restriction $\boldsymbol{w} = 0$ (denoted as Problem (7)') and Problem (6) with right-hand side be replaced by 0 (denoted as

Problem (6)'). Taking the dual of Problem (6)', we obtain a dual objective function of $\max 0$. Note that Problem (6)' is either unbounded or zero at optimality, as an easy solution is all variables at zero. Therefore the optimum must be zero otherwise problem (6) is unbounded. This implied also that Problem (5) when $\boldsymbol{w} = 0$ is also bounded above by $\mathfrak{D}_7$, which proves that $\hat{\mu}_4 := (\hat{\boldsymbol{x}}, \hat{\boldsymbol{w}}, \hat{\boldsymbol{u}}_y\hat{\boldsymbol{w}}, \hat{\boldsymbol{v}}_y\hat{\boldsymbol{w}}, \hat{\boldsymbol{u}}\hat{\boldsymbol{w}}, \hat{\boldsymbol{y}}, \hat{\boldsymbol{u}}_x, \hat{\boldsymbol{w}})$ is the optimal solution of Problem (5).

□

Appendix E. Corollary 1

Corollary 1. *Let $\mathcal{P}_6$ and $\mathcal{R}_6$ be the set of all extreme points and rays of BSP1 (6) and $\mathcal{P}_7$ and $\mathcal{R}_7$ be the set of all extreme points and rays of BSP2 (7), respectively. With $\mathcal{P}_6$, $\mathcal{P}_7$, $\mathcal{R}_6$ and $\mathcal{R}_7$, the single-level problem (2) is equivalent to the following problem:*

$$\min_{\boldsymbol{z}\in\{0,1\}^m, \boldsymbol{z}\in\mathcal{Z}} \quad \gamma c_z^T \boldsymbol{z} + t \tag{E.1a}$$

s.t.

$$t \geq \sum_{i=1}^{n} \left(\hat{\boldsymbol{y}}_i^T(b_i - D_i\boldsymbol{z}) + \hat{\boldsymbol{u}}_{xi}^T(b_{xi} - H_{xi}\boldsymbol{z})\right) - \hat{\boldsymbol{w}}\Bigg[\sum_{i=1}^{n}\left(\gamma_i c_i^T\hat{\boldsymbol{x}}_i - \gamma_i\hat{\boldsymbol{u}}_{yi}^T(b_{yi} - H_{yi}\boldsymbol{z}) - \gamma_i\hat{\boldsymbol{v}}_{yi}^T(e_{yi} + E_{yi}\boldsymbol{z})\right) \tag{E.1b}$$

$$- \hat{\boldsymbol{u}}^T(q_y - R_y\boldsymbol{z})\Bigg], \quad \forall(\hat{\boldsymbol{x}}, \hat{\boldsymbol{u}}_y, \hat{\boldsymbol{v}}_y, \hat{\boldsymbol{u}}, \hat{\boldsymbol{y}}, \hat{\boldsymbol{u}}_x, \hat{\boldsymbol{w}}) \in \mathcal{P}_6 \times \mathcal{P}_7,$$

$$\sum_{i=1}^{n}\left(\gamma_i c_i^T\tilde{\boldsymbol{x}}_i - \gamma_i\tilde{\boldsymbol{u}}_{yi}^T(b_{yi} - H_{yi}\boldsymbol{z}) - \gamma_i\tilde{\boldsymbol{v}}_{yi}^T(e_{yi} + E_{yi}\boldsymbol{z})\right) - \tilde{\boldsymbol{u}}^T(q_y - R_y\boldsymbol{z}) \geq 0, \quad \forall(\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{u}}_y, \tilde{\boldsymbol{v}}_y, \tilde{\boldsymbol{u}}) \in \mathcal{R}_6, \tag{E.1c}$$

$$0 \geq \sum_{i=1}^{n}\left(\tilde{\boldsymbol{y}}_i^T(b_i - D_i\boldsymbol{z}) + \tilde{\boldsymbol{u}}_{xi}^T(b_{xi} - H_{xi}\boldsymbol{z})\right), \quad \forall(\tilde{\boldsymbol{y}}, \tilde{\boldsymbol{u}}_x, 0) \in \mathcal{R}_7, \tag{E.1d}$$

$$0 \geq \sum_{i=1}^{n} \left(\tilde{\boldsymbol{y}}_i^T(b_i - D_i\boldsymbol{z}) + \tilde{\boldsymbol{u}}_{xi}^T(b_{xi} - H_{xi}\boldsymbol{z})\right) - \hat{\boldsymbol{w}}\Bigg[\sum_{i=1}^{n}\left(\gamma_i c_i^T\hat{\boldsymbol{x}}_i - \gamma_i\hat{\boldsymbol{u}}_{yi}^T(b_{yi} - H_{yi}\boldsymbol{z}) - \gamma_i\hat{\boldsymbol{v}}_{yi}^T(e_{yi} + E_{yi}\boldsymbol{z})\right) \tag{E.1e}$$

$$- \hat{\boldsymbol{u}}^T(q_y - R_y\boldsymbol{z})\Bigg], \quad \forall(\hat{\boldsymbol{x}}, \hat{\boldsymbol{u}}_y, \hat{\boldsymbol{v}}_y, \hat{\boldsymbol{u}}) \in \mathcal{P}_6, (\tilde{\boldsymbol{y}}, \tilde{\boldsymbol{u}}_x, \tilde{\boldsymbol{w}}) \in \mathcal{R}_7 : \tilde{\boldsymbol{w}} > 0.$$

Proof. As suggested in the proof of Theorem 3.1, $\hat{\mu}_4 := (\hat{\boldsymbol{x}}, \hat{\boldsymbol{w}}, \hat{\boldsymbol{u}}_y\hat{\boldsymbol{w}}, \hat{\boldsymbol{v}}_y\hat{\boldsymbol{w}}, \hat{\boldsymbol{u}}\hat{\boldsymbol{w}}, \hat{\boldsymbol{y}}, \hat{\boldsymbol{u}}_x, \hat{\boldsymbol{w}})$ is an extreme points of Problem (7) *iif* $(\hat{\boldsymbol{x}}, \hat{\boldsymbol{u}}_y, \hat{\boldsymbol{v}}_y, \hat{\boldsymbol{u}}, \hat{\boldsymbol{y}}, \hat{\boldsymbol{u}}_x, \hat{\boldsymbol{w}}) \in \mathcal{P}_6 \times \mathcal{P}_7$, which is shown in Eq. (E.1b). The proof of Theorem 3.1 also suggests that there are three extreme rays:

1. $\tilde{\mu}_1 := (\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{u}}_y, \tilde{\boldsymbol{v}}_y, \tilde{\boldsymbol{u}}, \boldsymbol{0}, \boldsymbol{0}, \boldsymbol{0}), \quad \forall(\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{u}}_y, \tilde{\boldsymbol{v}}_y, \tilde{\boldsymbol{u}}) \in \mathcal{R}_6$
2. $\tilde{\mu}_2 := (\boldsymbol{0}, \boldsymbol{0}, \boldsymbol{0}, \boldsymbol{0}, \tilde{\boldsymbol{y}}_n, \tilde{\boldsymbol{u}}_x, 0), \quad \forall(\tilde{\boldsymbol{y}}, \tilde{\boldsymbol{u}}_x, 0) \in \mathcal{R}_7$
3. $\tilde{\mu}_3 := (\hat{\boldsymbol{x}}\tilde{\boldsymbol{w}}, \hat{\boldsymbol{u}}_y\tilde{\boldsymbol{w}}, \hat{\boldsymbol{v}}_y\tilde{\boldsymbol{w}}, \hat{\boldsymbol{u}}\tilde{\boldsymbol{w}}, \tilde{\boldsymbol{y}}, \tilde{\boldsymbol{u}}_x, \tilde{\boldsymbol{w}}), \forall(\hat{\boldsymbol{x}}, \hat{\boldsymbol{u}}_y, \hat{\boldsymbol{v}}_y, \hat{\boldsymbol{u}}) \in \mathcal{P}_6, (\tilde{\boldsymbol{y}}, \tilde{\boldsymbol{u}}_x, \tilde{\boldsymbol{w}}) \in \mathcal{R}_7 : \tilde{\boldsymbol{w}} > 0$

Thus, Eqs. (E.1c)–(E.1e) hold and are equivalent to the projection of the feasible region of the linear relaxation of Problem (2) onto the space of $\boldsymbol{z}$. □

Appendix F. Proof of Theorem 3.2

Proof. Built upon Theorem 3.1, it suffices to show that solving Problem (8) is sufficient to obtain the optimal solution or unbounded ray of Problem (7) with $c_{xi}^T = \gamma_i c_i^T$ conditions (i.e., Problem (F.1)). First, Problem

Algorithm 1: Benders Separation Algorithm for Theorem 3.1

Input: $\hat{\boldsymbol{z}} \in \{0,1\}^m, \hat{\boldsymbol{z}} \in \mathcal{Z}$

Solve Problem (6);
if *Problem (6) is unbounded with an unbounded ray* $(\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{u}}_y, \tilde{\boldsymbol{v}}_y, \tilde{\boldsymbol{u}}) \in \mathcal{R}_6$ **then**
  Solve Problem (7) with $\mathfrak{D} = \infty$ (i.e., by fixing $\boldsymbol{w} = 0$);
  if *Problem (7) is unbounded with an unbounded ray* $(\tilde{\boldsymbol{y}}, \tilde{\boldsymbol{u}}_x, 0) \in \mathcal{R}_7$ **then**
    Add the feasibility cuts (E.1c) and (E.1d) to the RMP;
  else
    Add the feasibility cut (E.1c) to the RMP;
else
  Obtain its optimal solution $(\hat{\boldsymbol{x}}, \hat{\boldsymbol{u}}_y, \hat{\boldsymbol{v}}_y) \in \mathcal{P}_6$ and let $\mathfrak{D}_6$ be its optimal objective value;
  Solve Problem (7) with $\mathfrak{D}_6$;
  if *Problem (7) is unbounded with an unbounded ray* $(\tilde{\boldsymbol{y}}, \tilde{\boldsymbol{u}}_x, \tilde{\boldsymbol{w}}) \in \mathcal{R}_7$ **then**
    Add the feasibility cut (E.1e);
  else
    Obtain its optimal solution $(\hat{\boldsymbol{y}}, \hat{\boldsymbol{u}}_x, \hat{\boldsymbol{w}}) \in \mathcal{P}_7$;
    Add the optimality cut (E.1b);
    Update the best primal bound with the obtained feasible solution;

(7) with $c_{xi}^T = \gamma_i c_i^T$ conditions is:

$$\max_{\substack{\boldsymbol{y}\geq 0,\\ \boldsymbol{u}_x \geq 0,\\ \boldsymbol{w}\geq 0}} \quad \sum_{i=1}^{n} \left(\boldsymbol{y}_i^T (b_i - D_i \hat{\boldsymbol{z}}) + \boldsymbol{u}_{xi}^T (b_{xi} - H_{xi} \hat{\boldsymbol{z}}) \right) - \boldsymbol{w} \mathfrak{D}_6 \tag{F.1a}$$

$$\text{s.t.} \quad \boldsymbol{y}_i^T A_i + \boldsymbol{y}_{i+1}^T B_{i+1} + G_{xi}^T \boldsymbol{u}_{xi} \leq \gamma_i c_i^T (\boldsymbol{w} + 1), \quad \forall i \in [n-1], \tag{F.1b}$$

$$\boldsymbol{y}_n^T A_n + G_{xn}^T \boldsymbol{u}_{xn} \leq \gamma_n c_n^T (\boldsymbol{w} + 1). \tag{F.1c}$$

We introduce auxiliary variables as $(\boldsymbol{y}', \boldsymbol{u}_x') = (\frac{\boldsymbol{y}}{\boldsymbol{w}+1}, \frac{\boldsymbol{u}_x}{\boldsymbol{w}+1})$ and Problem (F.1) can be rewritten as

$$\max_{\substack{\boldsymbol{y}'\geq 0, \boldsymbol{u}_x' \geq 0\\ \boldsymbol{w}\geq 0}} \sum_{i=1}^{n} \left(\boldsymbol{y}_i^{'T} (b_i - D_i \hat{\boldsymbol{z}}) + \boldsymbol{u}_{xi}^{'T} (b_{xi} - H_{xi} \hat{\boldsymbol{z}}) \right) + \boldsymbol{w} \left[\sum_{i=1}^{n} \boldsymbol{y}_i^{'T} (b_i - D_i \hat{\boldsymbol{z}}) + \boldsymbol{u}_{xi}^{'T} (b_{xi} - H_{xi} \hat{\boldsymbol{z}}) - \mathfrak{D}_6 \right] \tag{F.2a}$$

$$\text{s.t.} \quad \boldsymbol{y}_i^{'T} A_i + \boldsymbol{y}_{i+1}^{'T} B_{i+1} + G_{xi}^T \boldsymbol{u}_{xi}^{'} \leq \gamma_i c_i^T, \quad \forall i \in [n-1], \tag{F.2b}$$

$$\boldsymbol{y}_n^{'T} A_n + G_{xn}^T \boldsymbol{u}_{xn}^{'} \leq \gamma_n c_n^T. \tag{F.2c}$$

Suppose Problem (8) has a finite optimum $\mathfrak{D}_8$ at $(\hat{\boldsymbol{y}}, \hat{\boldsymbol{u}}_x)$ and $\mathfrak{D}_8 > \mathfrak{D}_6$. Then, for any $\alpha > 0$, $(\boldsymbol{y}', \boldsymbol{u}_x', \boldsymbol{w}) = (\hat{\boldsymbol{y}}, \hat{\boldsymbol{u}}_x, \alpha)$ is feasible to Problem (F.2) and its objective value increases as α increases, and thus Problem (F.2) is unbounded, so is Problem (F.1). Note that by converting $(\boldsymbol{y}', \boldsymbol{u}_x', \boldsymbol{w})$ to the solution of Problem (F.1) using $(\boldsymbol{y}', \boldsymbol{u}_x') = (\frac{\boldsymbol{y}}{\boldsymbol{w}+1}, \frac{\boldsymbol{u}_x}{\boldsymbol{w}+1})$, we can see that $(\hat{\boldsymbol{y}}, \hat{\boldsymbol{u}}_x, 1)$ is an unbounded ray of Problem (F.1). When $\mathfrak{D}_8 \leq \mathfrak{D}_6$, the term multiplied by $\boldsymbol{w}$ in the objective function of (F.2) can be disregarded, thus Problem (F.2) and Problem (8) have a finite optimum $\mathfrak{D}_{(F.2)}$ at $(\hat{\boldsymbol{y}}, \hat{\boldsymbol{u}}_x, 0)$. Otherwise, when Problem (8) is unbounded with an unbounded ray of $(\tilde{\boldsymbol{y}}, \tilde{\boldsymbol{u}}_x)$, Problem (F.1) is unbounded by $(\tilde{\boldsymbol{y}}, \tilde{\boldsymbol{u}}_x, 0)$. □

Appendix G. Corollary 2

Corollary 2. *Let $\mathcal{P}_6$ and $\mathcal{R}_6$ be the set of all extreme points and rays of Problem (6) and $\mathcal{P}_8$ and $\mathcal{R}_8$ be the set of all extreme points and rays of Problem (8), respectively. With $\mathcal{P}_6$, $\mathcal{P}_6$, $\mathcal{R}_8$, and $\mathcal{R}_8$, the single-level* *problem (2) is equivalent to the following problem:*

$$\min_{\boldsymbol{z}\in\{0,1\}^m,\boldsymbol{z}\in\mathcal{Z}} \quad \gamma c_z^T \boldsymbol{z} + t \tag{G.1a}$$

s.t.

$$t \geq \sum_{i=1}^{n} \left(\hat{\boldsymbol{y}}_i^T (b_i - D_i\boldsymbol{z}) + \hat{\boldsymbol{u}}_{xi}^T (b_{xi} - H_{xi}\boldsymbol{z}) \right), \quad \forall (\hat{\boldsymbol{x}}, \hat{\boldsymbol{u}}_y, \hat{\boldsymbol{v}}_y, \hat{\boldsymbol{u}}, \hat{\boldsymbol{y}}, \hat{\boldsymbol{u}}_x) \in \mathcal{P}_6 \times \mathcal{P}_8,\ \textit{\textbf{if}}\ \mathfrak{D}_6 \geq \mathfrak{D}_8, \tag{G.1b}$$

$$\sum_{i=1}^{n} \left(\gamma_i c_i^T \tilde{\boldsymbol{x}}_i - \gamma_i \tilde{\boldsymbol{u}}_{yi}^T (b_{yi} - H_{yi}\boldsymbol{z}) - \gamma_i \tilde{\boldsymbol{v}}_{yi}^T (e_{yi} + E_{yi}\boldsymbol{z}) \right) - \tilde{\boldsymbol{u}}^T (q_y - R_y \boldsymbol{z}) \geq 0, \quad \forall (\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{u}}_y, \tilde{\boldsymbol{v}}_y, \tilde{\boldsymbol{u}}) \in \mathcal{R}_6, \tag{G.1c}$$

$$0 \geq \sum_{i=1}^{n} \left(\tilde{\boldsymbol{y}}_i^T (b_i - D_i\boldsymbol{z}) + \tilde{\boldsymbol{u}}_{xi}^T (b_{xi} - H_{xi}\boldsymbol{z}) \right), \quad \forall (\tilde{\boldsymbol{y}}, \tilde{\boldsymbol{u}}_x, 0) \in \mathcal{R}_8, \tag{G.1d}$$

$$\begin{aligned} \sum_{i=1}^{n} \left(\hat{\boldsymbol{y}}_i^T (b_i - D_i\boldsymbol{z}) + \hat{\boldsymbol{u}}_{xi}^T (b_{xi} - H_{xi}\boldsymbol{z}) \right) \leq \sum_{i=1}^{n} &\left(\gamma_i c_i^T \hat{\boldsymbol{x}}_i - \gamma_i \hat{\boldsymbol{u}}_{yi}^T (b_{yi} - H_{yi}\boldsymbol{z}) - \gamma_i \hat{\boldsymbol{v}}_{yi}^T (e_{yi} + E_{yi}\boldsymbol{z}) \right) \\ &- \hat{\boldsymbol{u}}^T (q_y - R_y \boldsymbol{z}), \quad \forall (\hat{\boldsymbol{x}}, \hat{\boldsymbol{u}}_y, \hat{\boldsymbol{v}}_y, \hat{\boldsymbol{u}}, \hat{\boldsymbol{y}}, \hat{\boldsymbol{u}}_x, \hat{\boldsymbol{w}}) \in \mathcal{P}_6 \times \mathcal{P}_8,\ \textit{\textbf{if}}\ \mathfrak{D}_6 < \mathfrak{D}_8. \end{aligned} \tag{G.1e}$$

Proof. As suggested in the proof of Theorem 3.2, $\hat{\mu}_4 := (\hat{\boldsymbol{x}}, \hat{\boldsymbol{w}}, \hat{\boldsymbol{u}}_y\hat{\boldsymbol{w}}, \hat{\boldsymbol{v}}_y\hat{\boldsymbol{w}}, \hat{\boldsymbol{u}}\hat{\boldsymbol{w}}, \hat{\boldsymbol{y}}, \hat{\boldsymbol{u}}_x, \hat{\boldsymbol{w}})$ is an extreme points of Problem (14) *iif* $(\hat{\boldsymbol{x}}, \hat{\boldsymbol{u}}_y, \hat{\boldsymbol{v}}_y, \hat{\boldsymbol{u}}, \hat{\boldsymbol{y}}, \hat{\boldsymbol{u}}_x, \hat{\boldsymbol{w}}) \in \mathcal{P}_6 \times \mathcal{P}_8$ and $\mathfrak{D}_6 \geq \mathfrak{D}_8$, which is shown in Eq. (G.1b). The proof of Theorem 3.1 also suggests that there are three extreme rays:

1. $\tilde{\mu}_1 := (\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{u}}_y, \tilde{\boldsymbol{v}}_y, \tilde{\boldsymbol{u}}, \boldsymbol{0}, \boldsymbol{0}, \boldsymbol{0}), \quad \forall (\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{u}}_y, \tilde{\boldsymbol{v}}_y, \tilde{\boldsymbol{u}}) \in \mathcal{R}_6$
2. $\tilde{\mu}_2 := (\boldsymbol{0}, \boldsymbol{0}, \boldsymbol{0}, \boldsymbol{0}, \tilde{\boldsymbol{y}}_n, \tilde{\boldsymbol{u}}_x, 0), \quad \forall (\tilde{\boldsymbol{y}}, \tilde{\boldsymbol{u}}_x, 0) \in \mathcal{R}_8$
3. $\tilde{\mu}_3 := (\hat{\boldsymbol{x}}, \hat{\boldsymbol{u}}_y, \hat{\boldsymbol{v}}_y, \hat{\boldsymbol{u}}, \hat{\boldsymbol{y}}, \hat{\boldsymbol{u}}_x, \hat{\boldsymbol{w}}), \quad \forall (\hat{\boldsymbol{x}}, \hat{\boldsymbol{u}}_y, \hat{\boldsymbol{v}}_y, \hat{\boldsymbol{u}}, \hat{\boldsymbol{y}}, \hat{\boldsymbol{u}}_x, \hat{\boldsymbol{w}}) \in \mathcal{P}_6 \times \mathcal{P}_8 :$ **if** $\mathfrak{D}_6 < \mathfrak{D}_8$

Thus, Eqs. (G.1c)–(G.1e) hold and are equivalent to the projection of the feasible region of the linear relaxation of Problem (2) onto the space of $\boldsymbol{z}$. □

Algorithm 2: Benders Separation Algorithm for Theorem 3.2

Input: $\hat{\boldsymbol{z}} \in \{0,1\}^m, \hat{\boldsymbol{z}} \in \mathcal{Z}$

Solve Problem (6) and (8) independently and let $\mathfrak{D}_6$ and $\mathfrak{D}_8$ denote their respective objective value;
if $\mathfrak{D}_8 = \infty$ *with an unbounded ray* $(\tilde{\boldsymbol{y}}, \tilde{\boldsymbol{u}}_x) \in \mathcal{R}_8$ **then**
    Add feasibility cuts (G.1d) to the RMP;
else
    Obtain the optimal solution of Problem (8) $(\hat{\boldsymbol{y}}, \hat{\boldsymbol{u}}_x) \in \mathcal{P}_8$;
    if $\mathfrak{D}_6 = \infty$ *with an unbounded ray* $(\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{u}}_y, \tilde{\boldsymbol{v}}_y, \tilde{\boldsymbol{u}}) \in \mathcal{R}_6$ **then**
        Add the feasibility cut (G.1c);
    else
        Obtain the optimal solution $(\hat{\boldsymbol{x}}, \hat{\boldsymbol{u}}_y, \hat{\boldsymbol{v}}_y, \hat{\boldsymbol{u}}) \in \mathcal{P}_6$;
        if $\mathfrak{D}_6 < \mathfrak{D}_8$ **then**
            Add the feasibility cut (G.1e) to the RMP;
        else
            Add the optimality cut (G.1b) to the RMP;
            Update the best primal bound with the obtained feasible solution;

Appendix H. Proof of Theorem 3.3

Proof. For BSP1 (Problem (6)), if without the *upper dual complicating constraints* (1e), the separated Benders subproblem (6) becomes

$$\min_{\boldsymbol{x}\geq 0,\boldsymbol{u}_y\geq 0,\boldsymbol{v}_y\geq 0} \quad \sum_{i=1}^{n} \gamma_i c_i^T \boldsymbol{x}_i - \gamma_i \boldsymbol{u}_{yi}^T (b_{yi} - H_{yi}\hat{\boldsymbol{z}}) - \gamma_i \boldsymbol{v}_{yi}^T (e_{yi} + E_{yi}\hat{\boldsymbol{z}}) \tag{H.1a}$$

$$\text{s.t.} \quad A_1\boldsymbol{x}_1 - G_{y1}^T\boldsymbol{u}_{y1} - Y_{y1}^T\boldsymbol{v}_{y1} \geq b_1, \tag{H.1b}$$

$$A_i\boldsymbol{x}_i + B_i\boldsymbol{x}_{i-1} - G_{yi}^T\boldsymbol{u}_{yi} - Y_{yi}^T\boldsymbol{v}_{yi} \geq b_i, \quad \forall i \in [n]^+, \tag{H.1c}$$

$$K_{yi}^T\boldsymbol{v}_{yi} \leq \mathbf{1}, \quad i \in [n]. \tag{H.1d}$$

We observe that the objective for the i^{th} lower level problems are scaled by γ_i. As discussed in the proof of Theorem 2.1 in Appendix A, this scaling reflects the sequence of computations. Specifically, the $(i-1)^{th}$ problem is computed first, and the results of $\boldsymbol{x}_{i-1}$ are subsequently treated as parameters in the i^{th} problem to compute $(\boldsymbol{x}_i, \boldsymbol{u}_{yi}, \boldsymbol{v}_{yi})$, $\forall i \in [n]^+$. Therefore, for each $i \in [n]$, the i^{th} subproblem takes the following form:

$$\min_{\boldsymbol{x}\geq 0,\boldsymbol{u}_y\geq 0,\boldsymbol{v}_y\geq 0} \quad \gamma_i c_i^T \boldsymbol{x}_i - \gamma_i \boldsymbol{u}_{yi}^T (b_{yi} - H_{yi}\hat{\boldsymbol{z}}) - \gamma_i \boldsymbol{v}_{yi}^T (e_{yi} + E_{yi}\hat{\boldsymbol{z}}) \tag{H.2a}$$

$$\text{s.t.} \quad A_i\boldsymbol{x}_i + B_i\hat{\boldsymbol{x}}_{i-1} - G_{yi}^T\boldsymbol{u}_{yi} - Y_{yi}^T\boldsymbol{v}_{yi} \geq b_i, \tag{H.2b}$$

$$K_{yi}^T\boldsymbol{v}_{yi} \leq \mathbf{1}. \tag{H.2c}$$

Here, $\hat{\boldsymbol{x}}_{i-1}$ is treated as a fixed parameter obtained from solving the $(i-1)^{\text{th}}$ subproblem. The term $B_i\hat{\boldsymbol{x}}_{i-1}$ is included only for $i \geq 2$ and omitted when $i = 1$. Consequently, the impact of the scaling problem can be mitigated through this sequential computation process.

For BSP2 (Problem (8)), we replace n dual variables $\boldsymbol{y}'$ with $\boldsymbol{y}'_i = \boldsymbol{y}_i/\gamma_i$ and $\boldsymbol{u}'_x$ with $\boldsymbol{u}'_{xi} = \boldsymbol{u}_{xi}/\gamma_i$, then Problem (8) becomes

$$\max_{\boldsymbol{y}'\geq 0,\boldsymbol{u}'_x\geq 0} \quad \sum_{i=1}^{n} \gamma_i \boldsymbol{y}_i^{'T} (b_i - D_i\hat{\boldsymbol{z}}) + \gamma_i \boldsymbol{u}_{xi}^{'T} (h_{xi} - H_{xi}\hat{\boldsymbol{z}}) \tag{H.3a}$$

$$\text{s.t.} \quad \boldsymbol{y}_i^{'T} A_i + (1-\gamma)\boldsymbol{y}_{i+1}^{'T} B_{i+1} + G_{xi}^T\boldsymbol{u}_{xi}^{'} \leq c_i^T, \quad \forall i \in [n-1], \tag{H.3b}$$

$$\boldsymbol{y}_n^{'T} A_n + G_{xn}^T\boldsymbol{u}_{xn}^{'} \leq c_n^T. \tag{H.3c}$$

As γ approaches to 1, the term $(1-\gamma)\boldsymbol{y}_{i+1}^{'T} B_{i+1}$ approaches to 0 and can be disregarded from the problem. Therefore, for γ sufficiently close to 1, Problem (H.3) can be decomposed into n independent subproblems, one for each $i \in [n]$, of the following form:

$$\max_{\boldsymbol{y}\geq 0,\boldsymbol{u}_x\geq 0} \quad \boldsymbol{y}_i^T (b_i - D_i\hat{\boldsymbol{z}}) + \boldsymbol{u}_{xi}^T (h_{xi} - H_{xi}\hat{\boldsymbol{z}}) \tag{H.4a}$$

$$\text{s.t.} \quad \boldsymbol{y}_i^T A_i + G_{xi}^T\boldsymbol{u}_{xi} \leq c_i^T. \tag{H.4b}$$

These subproblems can be solved independently in parallel, with no coupling terms remaining when γ approaches to 1.

For a given $\hat{\boldsymbol{z}}$, valid cuts can be generated using Corollary 2, after solving the decomposed subproblems (H.2) and (H.4) for all $i \in [n]$, involving the variables $\boldsymbol{x}_i$, $\boldsymbol{u}_{yi}$, $\boldsymbol{v}_{yi}$, $\boldsymbol{y}_i$, and $\boldsymbol{u}_{xi}$. □

References

Byeon, G. and Van Hentenryck, P. (2022). Benders subproblem decomposition for bilevel problems with convex follower. *INFORMS Journal on Computing*, 34(3):1749–1767.