BattOpt: Optimal Facility Planning for Electric Vehicle Battery Recycling

Matthew Brun* Xu Andy Sun†

Abstract

The electric vehicle (EV) battery supply chain will face challenges in sourcing scarce, expensive minerals required for manufacturing and in disposing of hazardous retired batteries. Integrating recycling technology into the supply chain has the potential to alleviate these issues; however, players in the battery market must design investment plans for recycling facilities. In this paper, we propose a two-stage stochastic optimization model for computing minimum cost recycling capacity decisions, in which retired batteries are recycled and recovered materials are used to manufacture new batteries. The model is a separable concave minimization subject to linear constraints, a class for which we design a new finitely convergent global optimization algorithm based on piecewise linear approximation that solves up to 10x faster than comparable algorithms. We propose an equivalent reformulation of the model that reduces the total number of variables by introducing integrality constraints. The reformulation can also be solved by our global algorithm with drastically reduced solve times. We detail a cut grouping strategy for Benders' decomposition in the second stage which improves convergence relative to single-cut and multi-cut implementations. To produce a set of second-stage scenarios, we design an approach for generating time-series projections for new battery demand, retired battery supply, and material costs. Analysis of the optimal solutions shows that effective investment in recycling can reduce battery manufacturing costs by 22% and reduce environmental impacts by up to 7%.

Key words: electric vehicle battery recycling, facility planning, global optimization, stochastic optimization

1 Introduction

Electric vehicles (EVs) are a principal technology in the energy transition that will play an important role in reducing emissions. The demand for the lithium-ion batteries that power EVs is expected to grow by over 20% year-over-year through 2030, reaching a market valuation in excess of $360 billion (Campagnol et al. 2022). Manufacture of these batteries requires "critical minerals", materials defined by a vulnerability to supply chain disruption, including lithium (Li), cobalt (Co), nickel (Ni), and manganese (Mn) (U.S. Geological Survey 2022). As demand for batteries grows significantly, so will the demand for these materials; for instance, lithium requirements are expected to grow 6-fold by 2030 under current climate pledges (IEA 2022).

*Operations Research Center, Massachusetts Institute of Technology, Cambridge, MA 02139. Email: brunm@mit.edu

†Sloan School of Management, Massachusetts Institute of Technology, Cambridge, MA 02139. Email: sunx@mit.edu

EV batteries have variable lifespans, but generally are retired after a capacity reduction of 20% with an expected age of 10 years (Marano et al. 2009). As large numbers of batteries are removed from vehicles, attention must be paid to their safe disposal. EV batteries contain toxic and heavy metals, so improper disposal may pose health and environmental risks (Kang et al. 2013). Further, recent European Union regulation sets recycled material composition requirements for new batteries and material recovery targets from retired batteries (Council of the European Union 2023).

In light of these challenges, a promising solution is the recycling of retired batteries, allowing recovery of component materials for reuse. Lithium-ion battery packs are made up of many individual cells, each of which contain a graphite-based anode, lithium-based cathode, electrolytes, binding polymers, and copper and aluminum foils, inside a plastic or steel shell. Common cathode compounds include lithium iron phosphate (LFP), lithium nickel cobalt aluminum oxide (NCA), and lithium nickel manganese cobalt oxide (NMC). Many battery materials can be recovered via recycling, although most valuable are the metals contained in cathode compounds.

Current technologies for lithium-ion battery recycling include pyrometallurgical, hydrometallurgical, and mechanical processes (Harper et al. 2019). *Pyrometallurgical recycling* exposes battery packs to high temperatures, incinerating extraneous battery components and inducing chemical reactions to recover compounds from the cathode material (Makuza et al. 2021). *Hydrometallurgical recycling* reacts cathode materials with a leaching solution, from which compounds are precipitated (Yao et al. 2018). A newer process, referred to as *direct recycling*, uses mechanical processes such as shredding and crushing to separate and recover constituent materials, followed by "relithiation" of cathode material to restore the chemical properties of a new material (Wu et al. 2023).

EverBatt (Dai et al. 2019) is a recently proposed closed-loop model for analyzing the cost and environmental impacts of battery recycling and manufacturing with detailed data for recycling process yields, facility costs, and emissions calculations. In a *closed-loop* model, recycled materials are used for the manufacturing of new products. This differs from *open-loop* models, in which recovered materials are sold or disposed. In the EverBatt model, retired batteries are first recycled. If pyrometallurgical or hydrometallurgical recycling is used, the extracted cathode precursors are remanufactured into cathode material in a step called *cathode production*. Under direct recycling, as cathode material is recovered in its original form, this step is not necessary. Then, new batteries are manufactured using the recycled cathode material and other purchased materials. EverBatt models the cost and environmental impact of recycling, cathode production, and manufacturing facilities under fixed retired battery input and new battery output quantities.

As players in the battery industry look towards incorporation of recycling technologies (Forbes 2022, BAF 2023), important questions surrounding investment and capacity planning will arise. Firms will decide how much recycling capacity to construct, which technologies to invest in, and when to use recycled materials to produce new batteries. In this paper, we provide an optimization-based method to address these questions. We convert the EverBatt framework into a two-stage stochastic concave minimization model. In the first stage, a multi-year capacity decision is made for recycling and cathode production facilities. Then, over a number of second-stage scenarios, operational decisions are made on how many batteries to recycle and how to use recycled materials to manufacture new batteries.

We summarize our contributions as follows:

- We model capacity planning for EV battery recycling as a two-stage separable concave minimization problem subject to linear constraints. To construct second-stage scenarios, we design methodology that leverages recent EV vehicle stock projections and critical mineral price projections to build a variety of retired battery supply, new battery demand, and material cost projections.
- We propose a novel adaptive piecewise linear approximation algorithm (aPWL) for the problem class that converges finitely to a globally optimal solution. In computational tests, we show that the aPWL algorithm compares favorably to benchmark algorithms, solving up to 10x faster than the next best algorithm on some instances.
- We develop an equivalent reformulation of our model that reduces the total number of variables by introducing integrality constraints. The reformulation reduces solve times from hours to minutes.
- We leverage a Benders' decomposition approach to separate the problem over second-stage scenarios. We evaluate cut aggregation strategies within the Benders' decomposition and show that a grouping of similar scenarios outperforms traditional single-cut and multi-cut implementations.
- We run computational studies to analyze the impact of closed loop recycling on new battery manufacturing. Our solutions show that an optimal recycling investment plan reduces the expected cost of battery production by 22%, expected energy consumption by 6%, and expected greenhouse gas emissions by 7% through 2050, relative to a solution that does not utilize any recycling.

The remainder of the paper is organized as follows. In Section 2, we review relevant literature. In Section 3, we introduce our models for EV battery recycling facility investment planning. In Section 4, we propose a novel algorithm for solving linearly constrained separable concave minimization problems and a decomposition scheme over second-stage scenarios. In Section 5, we describe our data and scenario generation, show computational results on a variety of test problems, and analyze representative optimal investment decisions. In Section 6, we conclude the paper.

2 Literature Review

2.1 Battery Recycling Optimization

Capacity planning problems for manufacturing applications have been studied extensively in the literature. For thorough reviews, see Van Mieghem (2003) and Martínez-Costa et al. (2014). Many studies consider a discrete set of feasible capacity decisions. We instead focus on formulations with concave costs and continuous decisions. In this problem class, Lee and Luss (1987) and Rajagopalan et al. (1998) simplify the continuous space to a discrete set of

potentially optimal decisions and solve the problem via dynamic programming. Li and Tirupati (1994) similarly compute a finite set of candidate capacities and propose heuristic algorithms. In this work, we introduce a reformulation by identifying an underlying structure of optimal decisions, but do not enumerate all such solutions.

Previous work has explored the application of capacity and logistics planning to EV battery recycling networks. Hoyer et al. (2015), Tadaros et al. (2022), and Rosenberg et al. (2023) study open-loop battery recycling processes. Hoyer et al. (2015) propose a mixed-integer linear program (MILP) for a deterministic multi-period model that optimizes capacity decisions for battery disassembly, conditioning, and recycling facilities in Germany. Feasible capacity decisions are restricted to a discrete set of values, modeled by the construction of "modules" with fixed capacities. Tadaros et al. (2022) and Rosenberg et al. (2023) also propose multi-period deterministic MILPs for planning of open-loop EV battery recycling, the former for the Swedish supply chain and the latter the German. Tadaros et al. (2022) formulate the problem as a facility location problem with fixed capacities, while Rosenberg et al. (2023) optimize both the planned location of a facility as well as its capacity, selected from a discrete set of possible capacity levels.

Li et al. (2018), Wang et al. (2020), and Li et al. (2023) introduce models for closed-loop battery recycling. Li et al. (2018) propose a multi-period model with stochasticity in the number and age of retired batteries avaiable for recycling, and Wang et al. (2020) investigate a single-period deterministic setting. Both models are solved with genetic algorithms. Li et al. (2023) consider optimization of the recycling network design in a single-period model with stochasticity in material recovery rate, new battery demand, and retired battery supply. This model optimally locates facilities with fixed capacities and is solved as a two-stage MILP with second-stage decomposition.

Our work is unique in its consideration of a multi-period, stochastic, closed-loop battery recycling problem. A multi-period stochastic model motivates new approaches for generating scenarios with parameters that vary across time. Our model further chooses capacities from a continuous range, in contrast with the discrete feasible sets of other works. This approach allows for more granularity in capacity planning and expansion without making a priori assumptions on the size of constructed facilities. However, considering a continuous space of capacity decisions introduces concave cost functions due to economies of scale, making the model a nonconvex program instead of a MILP and increasing the difficulty of identifying globally optimal capacity decisions.

2.2 Linearly Constrained Concave Programming

Many important optimization problems can be formulated as a concave minimization over a polyhedron, including quadratic assignment (Bazaraa and Sherali 1982), binary programming (Raghavachari 1969), and applications with economies of scale (Feldman et al. 1966). Algorithms for globally solving concave programs date back to Tuy's cutting planes (Tuy 1964); subsequently, a variety of branch and bound algorithms have been proposed (Rosen 1983, Benson 1985). For a survey of methods, see Pardalos and Rosen (1986) and Pardalos and Romeijn (2002).

We focus on problems with separable objective functions (i.e., representable as a sum of univariate functions). Shectman and Sahinidis (1998) propose a finitely convergent branch and bound algorithm for separable concave minimization over polyhedra, in which the feasible region is partitioned with rectangular splits and the objective is

underestimated by linear functions, allowing the subproblems to be solved as linear programs. This algorithm is similar to that of Falk and Soland (1969), with improved finite convergence guarantees. D'Ambrosio et al. (2009) propose a global algorithm for separable nonconvex optimization. The algorithm splits each nonconvex function into convex and concave segments and approximates the concave segments by piecewise linear under-approximators. The approximation is solved globally as a mixed-integer convex program, and the under-approximators of the convex pieces are updated at the incumbent iterate. This algorithm converges in the limit to a global optimum. Magnanti and Stratila (2004) show approximate equivalence between MILP and linearly constrained separable concave minimization by generating piecewise linear approximations of the concave cost functions with a number of pieces that scales linearly in the inverse of the approximation error.

Our proposed algorithm solves over a sequence of piecewise linear under-approximations of the objective function, where each approximation is exact at previous iterates. As opposed to Falk and Soland (1969) and Shectman and Sahinidis (1998), we partition on all variables simultaneously at a global optimum of the incumbent approximation. This reduces the total number of iterations, improving compatibility with decompositions over a second stage. Additionally, by representing piecewise functions as MILPs, we leverage the tools, including cuts, heuristics, and branching techniques, incorporated into modern commercial solvers. D'Ambrosio et al. (2009) utilize a similar approximation and update scheme but apply the method to a broader problem class which does not admit a finite termination result.

3 EV Battery Recycling and Manufacturing Model

We model the supply chain for new EV batteries through a sequence of production steps: retired battery recycling (REC), material conversion (MC), cathode production (CP), and new battery manufacturing (NB). Figure 1 depicts the material flow through these steps. Battery recycling extracts materials from retired batteries via physical or chemical methods. If the recycling method recovers cathode powder, the cathode powder can be used directly for new battery manufacturing. Otherwise, the recycled materials are used to produce new cathode powder in the cathode production step. This step manufactures new cathode material using new or recycled precursors. Material conversion is an intermediate step to convert recycled chemicals into the compounds required as inputs for cathode production. Finally, in new battery manufacturing, EV battery cells are produced from cathode powder that is obtained from recycling or purchased from a market.

Recycling and cathode production require facilities, and facility capacities limit throughput. We assume that demand for new batteries is always satisfied, so battery manufacturing facility decisions are not included in the model. The facilities can be located in a number of zones. Zones may contain multiple facilities, and production runs in parallel across zones, with supply and demand for retired and new batteries allocated to each zone. Recycled material and retired batteries in inventory can be transported between zones.

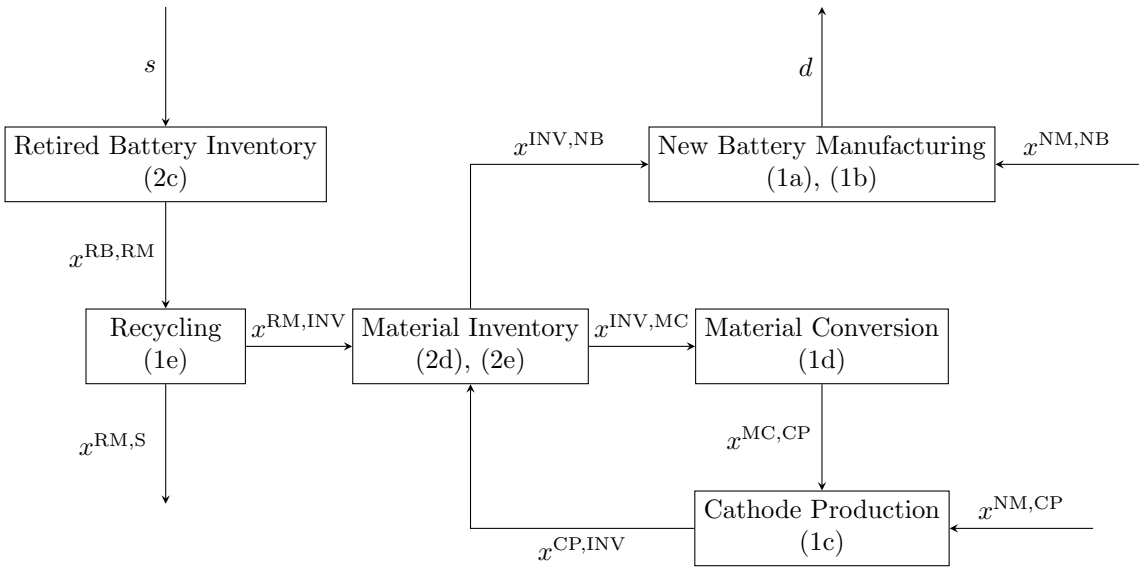


Figure 1: Material flows through the model in a single time period and location. RB represents retired batteries available for recycling, INV materials in inventory, RM recycled materials, S materials sold, and NM new materials purchased.

3.1 Two-Stage Stochastic Battery Recycling Model

The model is a two-stage stochastic optimization problem. A facility capacity decision for each zone is made in the first stage, then uncertainty in parameters for new battery demand, retired battery supply, and material cost is realized and operational decisions are made in the second stage. The set of uncertain scenarios for material cost, new battery demand, and retired battery supply is given by Ω and the probability associated with scenario $\omega \in \Omega$ is denoted p_ω. The model time horizon contains T periods (e.g., years), $\mathcal{T} = [\![T]\!]$, in which operational decisions are made. These periods are grouped into L planning periods (e.g., spans of 5 years), $\mathcal{L} = [\![L]\!]$, in which facility capacity decisions are made. The set $\mathcal{T}_l \subseteq \mathcal{T}$ gives the set of time periods associated with planning period $l \in \mathcal{L}$, and, conversely, $l_t \in \mathcal{L}$ gives the planning period associated with time period $t \in \mathcal{T}$. The set of zones for facility location is given by $\mathcal{Z}$. Variables x are operational decisions, indexed by scenario and time period, while variables y are capacity decisions, indexed by planning period.

Production

Constraints (1) model the changes between materials during new battery manufacturing, material conversion, cathode production, and battery recycling.

$$\sum_{i \in \mathcal{I}} \Delta^{\text{NB}}_{i,k} d_{\omega,z,t,i} = x^{\text{NM,NB}}_{\omega,z,t,k} + x^{\text{INV,NB}}_{\omega,z,t,k} \qquad \forall k \in \mathcal{K}^{\text{CP}},\ t \in \mathcal{T},\ z \in \mathcal{Z},\ \omega \in \Omega \tag{1a}$$

$$\sum_{i \in \mathcal{I}} \Delta^{\text{NB}}_{i,k} d_{\omega,z,t,i} = x^{\text{NM,NB}}_{\omega,z,t,k} \qquad \forall k \in \mathcal{K} \setminus \mathcal{K}^{\text{CP}},\ t \in \mathcal{T},\ z \in \mathcal{Z},\ \omega \in \Omega \tag{1b}$$

$$\sum_{k' \in \mathcal{K}^{\text{CP}}} \Delta^{\text{CP}}_{k',k} x^{\text{CP,INV}}_{\omega,z,t,k'} = x^{\text{NM,CP}}_{\omega,z,t,k} + x^{\text{MC,CP}}_{\omega,z,t,k} \qquad \forall k \in \mathcal{K} \setminus \mathcal{K}^{\text{CP}},\ t \in \mathcal{T},\ z \in \mathcal{Z},\ \omega \in \Omega \tag{1c}$$

$$\sum_{k' \in \mathcal{K} \setminus \mathcal{K}^{\mathrm{CP}}} \Delta^{\mathrm{MC}}_{k',k} x^{\mathrm{MC,CP}}_{\omega,z,t,k'} = x^{\mathrm{INV,MC}}_{\omega,z,t,k} \qquad \forall k \in \mathcal{K} \setminus \mathcal{K}^{\mathrm{CP}},\ t \in \mathcal{T},\ z \in \mathcal{Z},\ \omega \in \Omega \tag{1d}$$

$$x^{\mathrm{RM,INV}}_{\omega,z,t,k} + x^{\mathrm{RM,S}}_{\omega,z,t,k} = \sum_{i \in \mathcal{I},\ j \in \mathcal{J}} \Delta^{\mathrm{REC}}_{k,i,j} x^{\mathrm{RB,RM}}_{\omega,z,t,i,j} \qquad \forall k \in \mathcal{K},\ t \in \mathcal{T},\ z \in \mathcal{Z},\ \omega \in \Omega \tag{1e}$$

We remind that the process dynamics represented in constraints (1) are depicted in Figure 1. The set $\mathcal{K}$ gives the set of materials tracked across the production processes, such as graphite, copper, and carbon black. The set of battery chemistries is given by $\mathcal{I}$, including chemistries of types NMC, NCA, and LFP. Cathode powders are given by the set $\mathcal{K}^{\mathrm{CP}} \subset \mathcal{K}$. New batteries are manufactured from new or recycled cathode powder (1a) and other non-recycled new materials (e.g., graphite, aluminum) (1b) to meet battery demand $d_{\omega,z,t,i}$ for each chemistry and zone. Variables $x^{\mathrm{NM,NB}}_{\omega,z,t,k}$ and $x^{\mathrm{INV,NB}}_{\omega,z,t,k}$ give the mass of new and recycled material k, respectively, used in new battery production, and the parameter $\Delta^{\mathrm{NB}}_{i,k}$ gives the amount of material k needed to manufacture a unit mass of battery with chemistry i. Equation (1c) models the production of cathode powder from new or recycled materials, where the variables $x^{\mathrm{NM,CP}}_{\omega,z,t,k}$ and $x^{\mathrm{MC,CP}}_{\omega,z,t,k}$ give the mass of new and recycled material inputs to cathode production, respectively. Output from cathode production is given by the variable $x^{\mathrm{CP,INV}}_{\omega,z,t,k}$. Recycled materials used in cathode production first undergo material conversion, modeled by (1d). Inputs to material conversion are given by variable $x^{\mathrm{INV,MC}}_{\omega,z,t,k}$ and outputs by $x^{\mathrm{MC,CP}}_{\omega,z,t,k}$. The mass of material k required to manufacture a unit of material k' in cathode production and material conversion are given by parameters $\Delta^{\mathrm{CP}}_{k',k}$ and $\Delta^{\mathrm{MC}}_{k',k}$, respectively. Finally, the battery recycling processes are modeled by (1e). The set $\mathcal{J}$ gives the available recycling processes (pyrometallurgical, hydrometallurgical, and direct), variable $x^{\mathrm{RB,RM}}_{\omega,z,t,i,j}$ tracks the mass of batteries of chemistry i recycled with process j, and $\Delta^{\mathrm{REC}}_{k,i,j}$ gives the amount of material k recovered per unit of recycled battery with chemsitry i when using process j. The output materials can be placed into inventory to be used in manufacturing or sold to the market, modeled by variables $x^{\mathrm{RM,INV}}_{\omega,z,t,k}$ and $x^{\mathrm{RM,S}}_{\omega,z,t,k}$, respectively.

Inventory Balance

Constraints (2) model the flow of retired batteries and recycled materials between inventory and production processes.

$$x^{\mathrm{RB}}_{\omega,z,0,i} = 0 \qquad \forall i \in \mathcal{I},\ z \in \mathcal{Z},\ \omega \in \Omega \tag{2a}$$

$$x^{\mathrm{INV}}_{\omega,z,0,k} = 0 \qquad \forall k \in \mathcal{K},\ z \in \mathcal{Z},\ \omega \in \Omega \tag{2b}$$

$$x^{\mathrm{RB}}_{\omega,z,t,i} = x^{\mathrm{RB}}_{\omega,z,t-1,i} + \sum_{z' \in \mathcal{Z} \setminus \{z\}} \left(x^{\mathrm{TR,RB}}_{\omega,z',z,t,i} - x^{\mathrm{TR,RB}}_{\omega,z,z',t,i} \right) + s_{\omega,z,t,i} - \sum_{j \in \mathcal{J}} x^{\mathrm{RB,RM}}_{\omega,z,t,i,j}$$
$$\forall i \in \mathcal{I},\ t \in \mathcal{T},\ z \in \mathcal{Z},\ \omega \in \Omega \tag{2c}$$

$$x^{\mathrm{INV}}_{\omega,z,t,k} = x^{\mathrm{INV}}_{\omega,z,t-1,k} + \sum_{z' \in \mathcal{Z} \setminus \{z\}} \left(x^{\mathrm{TR,RM}}_{\omega,z',z,t,k} - x^{\mathrm{TR,RM}}_{\omega,z,z',t,k} \right) + x^{\mathrm{RM,INV}}_{\omega,z,t,k} - x^{\mathrm{INV,MC}}_{\omega,z,t,k}$$
$$\forall k \in \mathcal{K} \setminus \mathcal{K}^{\mathrm{CP}},\ t \in \mathcal{T},\ z \in \mathcal{Z},\ \omega \in \Omega \tag{2d}$$

$$x^{\mathrm{INV}}_{\omega,z,t,k} = x^{\mathrm{INV}}_{\omega,z,t-1,k} + \sum_{z' \in \mathcal{Z} \setminus \{z\}} \left(x^{\mathrm{TR,RM}}_{\omega,z',z,t,k} - x^{\mathrm{TR,RM}}_{\omega,z,z',t,k} \right) + x^{\mathrm{RM,INV}}_{\omega,z,t,k} + x^{\mathrm{CP,INV}}_{\omega,z,t,k} - x^{\mathrm{INV,NB}}_{\omega,z,t,k}$$

$$\forall k \in \mathcal{K}^{\mathrm{CP}},\ t \in \mathcal{T},\ z \in \mathcal{Z},\ \omega \in \Omega \tag{2e}$$

The mass of batteries with chemistry i in inventory is given by variable $x^{\mathrm{RB}}_{\omega,z,t,i}$, and inventory for material k is given by $x^{\mathrm{INV}}_{\omega,z,t,k}$. The variables $x^{\mathrm{TR,RM}}_{\omega,z,z',t,k}$ and $x^{\mathrm{TR,RB}}_{\omega,z,z',t,i}$ represent the transport of material k and of batteries with chemistry i, respectively, from zone z to zone z'. Constraints (2a)-(2b) enforce that initial inventory is zero. Retired battery conservation across time periods is modeled by (2c), where the mass of newly retired batteries is given by $s_{\omega,z,t,i}$, and batteries leave inventory to be recycled. Flow balance for materials is modeled by (2d)-(2e), where material enters inventory from battery recycling or cathode production and is removed for material conversion and new battery production.

Capacity

Constraints (3) model recycling and cathode production facility capacities.

$$\sum_{n \in \mathcal{N}^{\mathrm{REC}}_l} y^{\mathrm{REC}}_{z,l,j,n} \geq \sum_{i \in \mathcal{I}} x^{\mathrm{RB,RM}}_{\omega,z,t,i,j} \qquad \forall j \in \mathcal{J},\ l \in \mathcal{L},\ t \in \mathcal{T}_l,\ z \in \mathcal{Z},\ \omega \in \Omega \tag{3a}$$

$$\sum_{n \in \mathcal{N}^{\mathrm{CP}}_{l,k}} y^{\mathrm{CP}}_{z,l,k,n} \geq x^{\mathrm{CP,INV}}_{\omega,z,t,k} \qquad \forall k \in \mathcal{K}^{\mathrm{CP}},\ l \in \mathcal{L},\ t \in \mathcal{T}_l,\ z \in \mathcal{Z},\ \omega \in \Omega \tag{3b}$$

$$\sum_{n \in \mathcal{N}^{\mathrm{REC}}_l} y^{\mathrm{REC}}_{z,l,j,n} \geq \sum_{n \in \mathcal{N}^{\mathrm{REC}}_{l-1}} y^{\mathrm{REC}}_{z,l-1,j,n} \qquad \forall j \in \mathcal{J},\ l \in \mathcal{L} \setminus \{1\},\ z \in \mathcal{Z} \tag{3c}$$

$$\sum_{n \in \mathcal{N}^{\mathrm{CP}}_{l,k}} y^{\mathrm{CP}}_{z,l,k,n} \geq \sum_{n \in \mathcal{N}^{\mathrm{CP}}_{l-1,k}} y^{\mathrm{CP}}_{z,l-1,k,n} \qquad \forall k \in \mathcal{K}^{\mathrm{CP}},\ l \in \mathcal{L} \setminus \{1\},\ z \in \mathcal{Z} \tag{3d}$$

$$y^{\mathrm{REC}}_{z,l,j,n} \leq u^{\mathrm{REC}} \qquad \forall j \in \mathcal{J},\ l \in \mathcal{L},\ z \in \mathcal{Z},\ n \in \mathcal{N}^{\mathrm{REC}}_l \tag{3e}$$

$$y^{\mathrm{CP}}_{z,l,k,n} \leq u^{\mathrm{CP}} \qquad \forall k \in \mathcal{K}^{\mathrm{CP}},\ l \in \mathcal{L},\ z \in \mathcal{Z},\ n \in \mathcal{N}^{\mathrm{CP}}_{l,k} \tag{3f}$$

First-stage capacity decisions are made by planning period, so the facility capacity in planning period l applies to all scenarios and all time periods $\mathcal{T}_l$ within the planning period. Recycling capacity can be constructed at up to N^{REC}_l identical facilities within each zone, where $\mathcal{N}^{\mathrm{REC}}_l = [\![N^{\mathrm{REC}}_l]\!]$ is the facility index set. Variable $y^{\mathrm{REC}}_{z,l,j,n}$ gives the capacity for facility n by recycling process j, zone z, and planning period l. Similarly, cathode production capacity is constructed on up to $N^{\mathrm{CP}}_{l,k}$ lines, where different cathode powders are manufactured on distinct lines. The set $\mathcal{N}^{\mathrm{CP}}_{l,k} = [\![N^{\mathrm{CP}}_{l,k}]\!]$ gives the line indices for each cathode powder type. Variable $y^{\mathrm{CP}}_{z,l,k,n}$ gives the capacity of line n producing cathode powder k. Constraint (3a) limits the throughput of each recycling process by the total capacity across all facilities. Similarly, cathode production throughput is constrained by (3b). Constraints (3c) and (3d) require that total capacity is nondecreasing across planning periods. Finally, recycling facilities and cathode production lines have maximum capacities of u^{REC} and u^{CP}, respectively, enforced by (3e) and (3f).

Cost

Costs related to the production process are modeled by equations (4).

$$C^{\mathrm{OP}}_{\omega,t}(x) = \sum_{z\in\mathcal{Z}}\Bigg(\sum_{k\in\mathcal{K}} c^{\mathrm{NB,NM}}_{\omega,t,k} x^{\mathrm{NM,NB}}_{\omega,z,t,k} + \sum_{k\in\mathcal{K}\setminus\mathcal{K}^{\mathrm{CP}}}\left(c^{\mathrm{CP,NM}}_{\omega,t,k} x^{\mathrm{NM,CP}}_{\omega,z,t,k} + c^{\mathrm{MC}}_{\omega,z,t,k} x^{\mathrm{MC,CP}}_{\omega,z,t,k}\right) + \sum_{k\in\mathcal{K}} v_{\omega,t,k}\cdot\left(\rho x^{\mathrm{INV}}_{\omega,z,t,k} - \eta x^{\mathrm{RM,S}}_{\omega,z,t,k}\right) + \sum_{k\in\mathcal{K}^{\mathrm{CP}}} c^{\mathrm{CP}}_{\omega,z,t,k} x^{\mathrm{CP,INV}}_{\omega,z,t,k} + \sum_{i\in\mathcal{I},\ j\in\mathcal{J}} c^{\mathrm{REC}}_{\omega,z,t,i,j} x^{\mathrm{RB,RM}}_{\omega,z,t,i,j} + \sum_{z'\in\mathcal{Z}\setminus\{z\}}\left(\sum_{k\in\mathcal{K}} c^{\mathrm{TR,RM}}_{z,z'} x^{\mathrm{TR,RM}}_{\omega,z,z',t,k} + \sum_{i\in\mathcal{I}} c^{\mathrm{TR,RB}}_{z,z'} x^{\mathrm{TR,RB}}_{\omega,z,z',t,i}\right)\Bigg) \quad \forall t\in\mathcal{T},\ \omega\in\Omega \tag{4a}$$

$$C^{\mathrm{PL}}_{t}(y) = \sum_{z\in\mathcal{Z}}\left(\sum_{j\in\mathcal{J}}\sum_{n\in\mathcal{N}^{\mathrm{REC}}_{l_t}} f^{\mathrm{REC}}_{z,j}(y^{\mathrm{REC}}_{z,l_t,j,n}) + \sum_{k\in\mathcal{K}^{\mathrm{CP}}}\sum_{n\in\mathcal{N}^{\mathrm{CP}}_{l_t,k}} f^{\mathrm{CP}}_{z,k}(y^{\mathrm{CP}}_{z,l_t,k,n})\right) \quad \forall t\in\mathcal{T} \tag{4b}$$

Operational costs per time period, including material, utility, transportation, and inventory costs, as well as revenue incurred by selling recycled material, are linear functions of second-stage decisions x and are represented by the functions $C^{\mathrm{OP}}_{\omega,t}(x)$, defined in (4a). The parameters $c^{\mathrm{NB,NM}}_{\omega,t,k}$ and $c^{\mathrm{CP,NM}}_{\omega,t,k}$ give the purchase costs per unit of new material k used in new battery manufacturing and cathode production, respectively. The variable costs of producing a unit of material k via material conversion, manufacturing a unit of cathode powder k via cathode production, and recycling a unit of batteries of chemistry i via process j are given by $c^{\mathrm{MC}}_{\omega,z,t,k}$, $c^{\mathrm{CP}}_{\omega,z,t,k}$, and $c^{\mathrm{REC}}_{\omega,z,t,i,j}$, respectively. Per-unit transportation costs from zone z to zone z' are given by $c^{\mathrm{TR,RM}}_{z,z'}$ for materials and $c^{\mathrm{TR,RB}}_{z,z'}$ for batteries. The revenue generated per unit of recycled material k sold is a proportion $\eta \leq 1$ of the value of the material $v_{\omega,t,k}$; selling extraneous recycled materials that are not used in downstream manufacturing helps offset recycling facility costs. Inventory costs are similarly incurred as a proportion ρ of the value of material held in inventory. We note that parameters that vary across scenarios are indexed by ω.

Planning costs include capital and labor costs, and are represented by the functions $C^{\mathrm{PL}}_{t}(y)$, defined in (4b). Capital costs include costs for facility construction and equipment purchases. These costs are concave functions due to economies of scale. Planning costs are defined in terms of functions f, where $f^{\mathrm{REC}}_{z,j}(y^{\mathrm{REC}}_{z,l,j,n})$ gives the long-term costs of facilities by recycling process and zone, and $f^{\mathrm{CP}}_{z,k}(y^{\mathrm{CP}}_{z,l,k,n})$ those of a cathode production line producing material k. The functions $f : \mathbb{R}_+ \to \mathbb{R}$ are concave lower semicontinuous piecewise with the form

$$f(y) = \begin{cases} \sum_i q_i y^{r_i} + w & y > 0 \\ 0 & y = 0. \end{cases} \tag{5}$$

The functions introduce a cost of 0 when no capacity is constructed, and otherwise are the sum of concave power terms ($r_i \in (0, 1]$ and $q_i \geq 0$) and a constant term ($w \geq 0$), and thus are monotonic increasing.

Complete Model

We seek to minimize total expected costs over the model horizon with an annual discount factor of γ, subject to the constraint logic presented in this section:

$$\begin{aligned} \min_{x,y} \quad & \sum_{t\in\mathcal{T}}(1-\gamma)^{t-1}\left(C_t^{\mathrm{PL}}(y)+\sum_{\omega\in\Omega}p_\omega C_{\omega,t}^{\mathrm{OP}}(x)\right) \\ \text{s.t.} \quad & (1),\ (2),\ (3) \\ & x\geq 0,\ y\geq 0. \end{aligned} \tag{P}$$

The model (P) is a minimization of a separable concave function subject to linear constraints.

3.2 Model Properties

In the remainder of this paper, we assume that Assumption 1 holds. The assumption enforces reasonable characteristics on the model data, including nonnegativity of costs, values, supply, demand, material requirements, and capacities. The properties of the functions f hold by their structure (5), and discount factor γ naturally falls on $[0,1)$.

Assumption 1. The parameters d, s, c, v, and Δ are nonnegative, u is positive, $\gamma<1$, and functions f are concave monotonic increasing with $f(0)=0$.

Each second-stage scenario has a relatively complete recourse property, given in Proposition 1. This property ensures that, for any first-stage feasible capacity decision y, there is some feasible operational decision x. We further show in Proposition 2 that the model is feasible and has an optimal solution at an extreme point of its feasible region. These and other original proofs are provided in the appendix.

Proposition 1. *For any $\overline{y}\geq 0$ that satisfies* (3c)-(3f)*, $\{(y,x)\geq 0\ :\ (1),\ (2),\ (3),\ y=\overline{y}\}\neq\emptyset$.*

Proposition 2. *The model* (P) *is feasible and has an optimal solution at an extreme point of its feasible region.*

For convenience of notation, we introduce the operator $\mathcal{C}(\mathcal{X};v)$ which counts the number of elements of a set $\mathcal{X}$ that take the value ν: $\mathcal{C}(\mathcal{X};\nu)=\sum_{x\in\mathcal{X}}\mathbb{I}(x=\nu)$. In Theorem 1, we note that any extreme point of the feasible region of (P) satisfies a specific structure in the first-stage decision y. Specifically, each facility is constructed to its upper bound before adding any capacity to the next facility. For a recycling solution y^{REC} with total capacity $Y_{z,l,j}:=\sum_{n\in\mathcal{N}_l^{\mathrm{REC}}}y_{z,l,j,n}^{\mathrm{REC}}$, the structure is defined by

$$\begin{aligned} \mathcal{C}(\{y_{z,l,j,n}^{\mathrm{REC}}\}_{n\in\mathcal{N}_l^{\mathrm{REC}}};u^{\mathrm{REC}})&\geq\left\lceil\frac{Y_{z,l,j}}{u^{\mathrm{REC}}}\right\rceil-1; \\ \mathcal{C}(\{y_{z,l,j,n}^{\mathrm{REC}}\}_{n\in\mathcal{N}_l^{\mathrm{REC}}};0)&\geq|\mathcal{N}_l^{\mathrm{REC}}|-\left\lceil\frac{Y_{z,l,j}}{u^{\mathrm{REC}}}\right\rceil \quad \forall j\in\mathcal{J},\ l\in\mathcal{L},\ z\in\mathcal{Z}, \end{aligned} \tag{6}$$

with a corresponding structure for $y_{z,l,k,n}^{\mathrm{CP}}$ defined in the appendix. Under this structure, there is at most one facility

with capacity not equal to its upper or lower bound. As (P) has an optimal extreme point, we conclude that the model has an optimal solution with this structure.

Theorem 1. *Let* $(\tilde{y}, \tilde{x})$ *be an extreme point of* $\{(y, x) \geq 0 \ : \ (1),\ (2),\ (3)\}$. *Then,* $\tilde{y}^{REC}$ *satisfies* (6) *and* $\tilde{y}^{CP}$ *satisfies the corresponding property.*

3.3 Reformulating the Model

Under the structure of the optimal solution implied by Theorem 1, we reformulate the capacity decision variables y and associated constraints (3). In the reformulation, integer variables $y^{\text{REC}}_{z,l,j}$ and $y^{\text{CP}}_{z,l,k}$ give the number of facilities with capacities at their upper bound, and continuous variables $y^{\text{REC}}_{z,l,j,+}$ and $y^{\text{CP}}_{z,l,k,+}$ give the capacity of the single facility not at its upper or lower bound.

$$u^{\text{REC}} y^{\text{REC}}_{z,l,j} + y^{\text{REC}}_{z,l,j,+} \geq \sum_{i \in \mathcal{I}} x^{\text{RB,RM}}_{\omega,z,t,i,j} \qquad \forall j \in \mathcal{J},\ l \in \mathcal{L},\ t \in \mathcal{T}_l,\ z \in \mathcal{Z},\ \omega \in \Omega \tag{7a}$$

$$u^{\text{CP}} y^{\text{CP}}_{z,l,k} + y^{\text{CP}}_{z,l,k,+} \geq x^{\text{CP,INV}}_{\omega,z,t,k} \qquad \forall k \in \mathcal{K}^{\text{CP}},\ l \in \mathcal{L},\ t \in \mathcal{T}_l,\ z \in \mathcal{Z},\ \omega \in \Omega \tag{7b}$$

$$u^{\text{REC}} y^{\text{REC}}_{z,l,j} + y^{\text{REC}}_{z,l,j,+} \geq u^{\text{REC}} y^{\text{REC}}_{z,l-1,j} + y^{\text{REC}}_{z,l-1,j,+} \qquad \forall j \in \mathcal{J},\ l \in \mathcal{L} \setminus \{1\},\ z \in \mathcal{Z} \tag{7c}$$

$$u^{\text{CP}} y^{\text{CP}}_{z,l,k} + y^{\text{CP}}_{z,l,k,+} \geq u^{\text{CP}} y^{\text{CP}}_{z,l-1,k} + y^{\text{CP}}_{z,l-1,k,+} \qquad \forall k \in \mathcal{K}^{\text{CP}},\ l \in \mathcal{L} \setminus \{1\},\ z \in \mathcal{Z} \tag{7d}$$

$$y^{\text{REC}}_{z,l,j} \in \{0, \ldots, |\mathcal{N}^{\text{REC}}_l| - 1\} \qquad \forall j \in \mathcal{J},\ l \in \mathcal{L},\ z \in \mathcal{Z} \tag{7e}$$

$$y^{\text{CP}}_{z,l,k} \in \{0, \ldots, |\mathcal{N}^{\text{CP}}_{l,k}| - 1\} \qquad \forall k \in \mathcal{K}^{\text{CP}},\ l \in \mathcal{L},\ z \in \mathcal{Z} \tag{7f}$$

$$y^{\text{REC}}_{z,l,j,+} \leq u^{\text{REC}} \qquad \forall j \in \mathcal{J},\ l \in \mathcal{L},\ z \in \mathcal{Z} \tag{7g}$$

$$y^{\text{CP}}_{z,l,k,+} \leq u^{\text{CP}} \qquad \forall k \in \mathcal{K}^{\text{CP}},\ l \in \mathcal{L},\ z \in \mathcal{Z} \tag{7h}$$

We also introduce an adjusted version of the first-stage objective function:

$$\overline{C}^{\text{PL}}_t(y) = \sum_{z \in \mathcal{Z}} \left(\sum_{j \in \mathcal{J}} \left(f^{\text{REC}}_{z,j}(y^{\text{REC}}_{z,l_t,j,+}) + y^{\text{REC}}_{z,l_t,j} f^{\text{REC}}_{z,j}(u^{\text{REC}}) \right) + \sum_{k \in \mathcal{K}^{\text{CP}}} \left(f^{\text{CP}}_{z,k}(y^{\text{CP}}_{z,l_t,k,+}) + y^{\text{CP}}_{z,l_t,k} f^{\text{CP}}_{k,z}(u^{\text{CP}}) \right) \right) \forall t \in \mathcal{T}. \tag{8}$$

The full reformulated model is then given by

$$\begin{aligned} \min_{x,y} \quad & \sum_{t \in \mathcal{T}} (1-\gamma)^{t-1} \left(\overline{C}^{\text{PL}}_t(y) + \sum_{\omega \in \Omega} p_\omega C^{\text{OP}}_{\omega,t}(x) \right) \\ \text{s.t.} \quad & (1),\ (2),\ (7) \\ & x \geq 0,\ y \geq 0. \end{aligned} \tag{PMI}$$

The new model (PMI) is equivalent to the original formulation (P), shown in Theorem 2. It also retains the linearly constrained separable concave structure of (P) with the addition of integrality constraints. Although these constraints add some complexity, the number of first-stage variables is significantly reduced as they are no longer indexed by

$\mathcal{N}^{\text{REC}}$ and $\mathcal{N}^{\text{CP}}$.

Theorem 2. *The models* (P) *and* (PMI) *have the same optimal objective value.*

4 Algorithmic Approaches

4.1 Adaptive Piecewise Linear Approximation Algorithm

We propose a finitely convergent algorithm to find globally optimal solutions for separable concave minimization problems with linear constraints and mixed-integer variables. The algorithm under-approximates each univariate concave objective function with a piecewise linear function, optimizes over the approximation, then updates the functions by introducing breakpoints at the incumbent solution, making the approximation exact at the current iterate. The subproblem objectives provide lower bounds on the global optimum, and revisiting an iterate yields an optimality certificate.

In this section, we define new notation which is distinct from the previous section. For analysis of our algorithm, we introduce a stylized version of the models (P) and (PMI):

$$\min_{(y,x)\in\mathcal{X}} \quad \sum_{i=1}^{n_y} f_i(y_i) + \sum_{\omega\in\Omega} p_\omega c_\omega^T x_\omega, \tag{SCP}$$

where the feasible region is $\mathcal{X} = \{(y, \{x_\omega\}_{\omega\in\Omega}) \geq 0 \ : \ y \in \mathbb{Z}^{n_{\text{I}}} \times \mathbb{R}^{n_y - n_{\text{I}}},\ Ay = b,\ B_\omega x_\omega + Dy = d_\omega\ \forall \omega \in \Omega\}$. The functions $f_i \ : \ \mathbb{R} \to \mathbb{R}$ are concave lower semicontinuous, $p_\omega \in \mathbb{R}$, $c_\omega \in \mathbb{R}^{n_x}$, $A \in \mathbb{R}^{m_1 \times n_y}$, $b \in \mathbb{R}^{m_1}$, $B_\omega \in \mathbb{R}^{m_2 \times n_x}$, $D \in \mathbb{R}^{m_2 \times n_y}$ and $d_\omega \in \mathbb{R}^{m_2}$. The number of integer first-stage variables is given by n_{I}. We assume $\mathcal{X}$ is nonempty and bounded, so any feasible y satisfies some consistent bounds $\underline{y} \leq y \leq \overline{y}$. This assumption is reasonable as the models have finite optimal solutions by Proposition 2. We further assume a relatively complete recourse property, which holds for our models by Proposition 1:

$$\{x \geq 0 \ : \ B_\omega x = d_\omega - Dy\} \neq \emptyset \quad \forall \omega \in \Omega,\ y \in \mathbb{Z}^{n_{\text{I}}} \times \mathbb{R}^{n_y - n_{\text{I}}} \ : \ y \geq 0,\ Ay = b. \tag{9}$$

We now detail our algorithm. Given a set of k_i breakpoints $\{y_i^{(j)}\}_{j=1}^{k_i}$ for variable y_i ordered so that $y_i^{(j)} < y_i^{(j+1)}$, we construct piecewise linear functions $\overline{f}_i \ : \ [y_i^{(1)}, y_i^{(k_i)}] \to \mathbb{R}$ by

$$\overline{f}_i(y_i) = \begin{cases} \dfrac{f_i(y_i^{(j+1)}) - f_i(y_i^{(j)})}{y_i^{(j+1)} - y_i^{(j)}}(y_i - y_i^{(j)}) + f_i(y_i^{(j)}) & y_i \in [y_i^{(j)}, y_i^{(j+1)}) \quad \forall j \in [\![k_i - 1]\!]. \end{cases} \tag{10}$$

We build a subproblem (SP) that replaces the objective of (SCP) with the piecewise functions $\overline{f}_i$:

$$\min_{(y,x)\in\mathcal{X}} \quad \sum_{i=1}^{n} \overline{f}_i(y_i) + \sum_{\omega\in\Omega} p_\omega c_\omega^T x_\omega. \tag{SP}$$

Algorithm 1: Adaptive Piecewise Linear Approximation

1 Initialize $(y_i^{(1)}, y_i^{(2)}) \leftarrow (\underline{y}_i, \overline{y}_i)$ and $k_i \leftarrow 2 \quad \forall i \in [\![n_y]\!]$. Set $s \leftarrow 1$.

2 For all $i \in [\![n_y]\!]$, construct $\overline{f}_i^s$ by (10) with breakpoints $\{y_i^{(j)}\}_{j=1}^{k_i}$.

3 Solve $\min_{(y,x)\in\mathcal{X}} \quad \sum_{i=1}^{n} \overline{f}_i^s(y_i) + \sum_{\omega\in\Omega} p_\omega c_\omega^T x_\omega$ to optimal solution $(\tilde{y}^s, \{\tilde{x}_\omega^s\}_{\omega\in\Omega})$.

4 If $\overline{f}_i^s(\tilde{y}_i^s) \geq f_i(\tilde{y}_i^s)$ for all i, STOP, $(\tilde{y}^s, \{\tilde{x}_\omega^{(s)}\}_{\omega\in\Omega})$ is optimal for (SCP).

5 Otherwise, for each i where $\overline{f}_i^s(\tilde{y}_i^s) < f_i(\tilde{y}_i^s)$, add breakpoint $y_i^{(k_i+1)} \leftarrow \tilde{y}_i^s$. Set $k_i \leftarrow k_i + 1$. Reindex $\{y_i^{(j)}\}_{j=1}^{k_i}$ so that $y_i^{(j)} < y_i^{(j+1)}$ for all i, j.

6 Set $s \leftarrow s + 1$. Go to 2.

As $\overline{f}_i$ are underestimators of f_i, the subproblems are relaxations of (SCP). The piecewise functions can be modeled with SOS-2 constraints or mixed-integer variables.

Denote the (polyhedral) set formed by fixing the discrete variables of $\mathcal{X}$ to some solution $\overline{y}$ by $\mathcal{X}(\overline{y}) = \{(y, x) \in \mathcal{X} \ : \ y_i = \overline{y}_i \ \forall i \in [\![n_I]\!]\}$. Lemma 1 shows that there is an optimal solution for (SP) that falls at an extreme point of $\mathcal{X}(y^*)$ for some y^*.

Lemma 1. *Let $(y^*, x^*) \in \mathcal{X}$ be an optimal solution to* (SP)*. Then, there is an optimal solution that is an extreme point of $\mathcal{X}(y^*)$.*

Assumption 2. Solving (SP) yields an optimal solution (y^*, x^*) that is an extreme point of $\mathcal{X}(y^*)$.

We propose the adaptive piecewise linear approximation algorithm (aPWL), Algorithm 1, that solves a sequence of subproblems (SP) and updates the breakpoints defining the piecewise functions at every iteration. As the subproblems minimize piecewise linear functions subject to linear and integrality constraints, they can be solved to global optimality via an LP-based branch and bound solver. Assumption 2 establishes that solving (SP) in step 3 yields one of the optimal extreme points described in Lemma 1. A branch and bound solver can be tailored to generate optimal extreme points of the restricted feasible region, satisfying this assumption. Under Assumption 2, the algorithm converges finitely to a global optimum of (SCP), shown in Theorem 3.

Theorem 3. *Algorithm 1 terminates finitely with a global optimum of* (SCP)*.*

4.2 Scenario Decomposition

We further leverage the two-stage stochastic structure of the model by employing a Benders' decomposition scheme (Benders 1962, Birge and Louveaux 2011) in step 3 of Algorithm 1. The subproblems (SP) are solved iteratively by generating cuts from the second-stage problems, which can be solved in parallel as linear programs. Importantly, every second-stage scenario has a relatively complete recourse property (9), so the decomposition scheme does not need to add cuts that enforce feasibility.

For each second-stage problem indexed by ω, we denote the extreme points of the dual feasible region $\{\pi \ : \ p_\omega c_\omega \geq B_\omega^T \pi\}$ by $\pi_j^{(\omega)}$ for $j \in \mathcal{J}_\omega$, where $\mathcal{J}_\omega$ is an index set for the dual extreme points of problem ω. Under property (9),

every second-stage problem has a finite optimal solution for any capacity decision y. Thus, each dual problem has a finite optimum at some extreme point $\pi_j^{(\omega)}$. The dual problems can then be equivalently represented by the convex piecewise linear functions

$$v_\omega(y) = \max_{j \in \mathcal{J}_\omega} (d_\omega - Dy)^T \pi_j^{(\omega)}. \tag{11}$$

The Benders' decomposition approximates v_ω by adding valid cuts at incumbent first-stage solutions y. Importantly, added cuts remain valid between iterations of Algorithm 1, allowing warmstarting of the decomposition between iterations and decreasing the number of new cuts required to converge.

For the sake of computational efficiency, we only partially decompose across scenarios. We construct groups of scenarios and apply the Benders' cut generation approach to each group. This approach is similar to that employed by Adulyasak et al. (2015). Let the groups $\{\mathcal{G}_e\}_{e=1}^{n_g}$ be a partition of the scenarios Ω. Then, we generate cuts for the grouped scenario functions

$$\max_{\{j_\omega\}_{\omega \in \mathcal{G}_e} \in \prod_{\omega \in \mathcal{G}_e} \mathcal{J}_\omega} \sum_{\omega \in \mathcal{G}_e} (d_\omega - Dy)^T \pi_{j_\omega}^{(\omega)}, \tag{12}$$

instead of the individual functions (11). Grouping in such a way reduces the number of convex functions to model but increases the number of linear pieces in each function.

5 Computational Results and Analysis

5.1 Data Generation

We model the recycling and manufacture of six battery chemistries: NMC(111), NMC(532), NMC(622), NMC(811), NCA, and LFP. NMC are lithium nickel manganese cobalt oxides, NCA is a lithium nickel cobalt aluminium oxide, and LFP is lithium iron phosphate. Batteries can be recycled through pyrometallurgical, hydrometallurgical, or direct processes. The time horizon covers 2021 to 2050 with annual granularity, and planning periods are the six 5-year intervals: 2021-2025, 2026-2030, 2031-2035, 2036-2040, 2041-2045, and 2046-2050.

Deterministic parameters are computed using the data and methodology proposed in the EverBatt model (Dai et al. 2019). We compute the parameters Δ, u, f, $c^{\mathrm{TR,RM}}$, and $c^{\mathrm{TR,RB}}$ from extensive low-level formulas and data provided within EverBatt. We take a discount rate $\gamma = 3\%$ to match the Federal Energy Management Program discount rate for 2021 (Kneifel and Webb 2020). Inventory costs are set as $\rho = 25\%$ of material value (Azzi et al. 2014), and we assume that $\eta = 70\%$ of material value is recovered when selling recycled materials to the market.[1.]

The numbers of recycling facilities and cathode production lines are determined by the number of facilities and lines needed to recycle the largest annual retired battery supply and to manufacture cathode powder to satisfy the largest annual new battery demand during each planning period:

$$|\mathcal{N}_l^{\mathrm{REC}}| = \left\lceil \frac{1}{u^{\mathrm{REC}}} \max_{\omega \in \Omega,\ t \in \mathcal{T}_l} \sum_{i \in \mathcal{I},\ z \in \mathcal{Z}} s_{\omega,z,t,i} \right\rceil \quad \text{and} \quad |\mathcal{N}_{l,k}^{\mathrm{CP}}| = \left\lceil \frac{1}{u^{\mathrm{CP}}} \max_{\omega \in \Omega,\ t \in \mathcal{T}_l} \sum_{i \in \mathcal{I},\ z \in \mathcal{Z}} \Delta_{i,k}^{\mathrm{NB}} d_{\omega,z,t,i} \right\rceil,$$

where maximum capacities are $u^{\text{REC}} = 100,000$ and $u^{\text{CP}} = 2,000$ tonnes/year (Dai et al. 2019).

Demand Scenarios

We follow the methodology and data from Xu et al. (2020) for constructing deterministic global EV stock projections under the International Energy Agency's Sustainable Development Scenario (SDS) and Stated Policies Scenario (STEPS) (IEA 2020). The projections give the total number of active EVs (including battery and plug-in hybrid EVs) in each year over the 30 year period from 2021 to 2050. We label these two projections EV_t^{SDS} and EV_t^{STEPS}. Time periods are standardized, so $t = 1$ corresponds to 2021 and $T = 30$ to 2050.

Let EV_t be the random variable that represents the total number of active EVs (EV *stock*) in year t. We define the distribution of EV_t by the truncated multivariate normal distribution

$$EV_t = \frac{1}{2}\left|EV_t^{\text{SDS}} - EV_t^{\text{STEPS}}\right| Z^{\text{EV}} + \frac{1}{2}\left(EV_t^{\text{SDS}} + EV_t^{\text{STEPS}}\right),$$

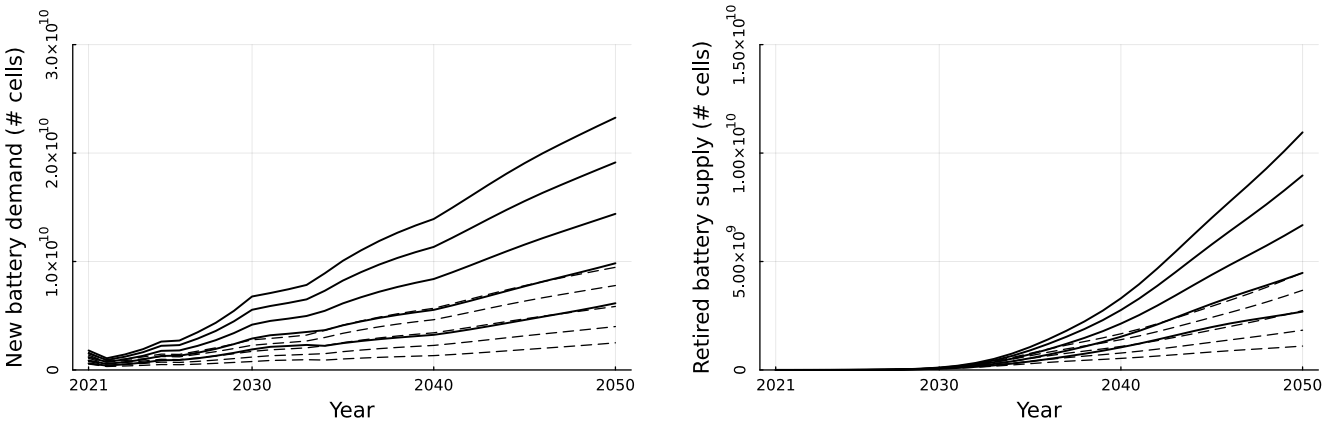
where the randomness is described by a single random variable Z^{EV} that is distributed as a truncated standard normal on the interval $[-2, 2]$. We then generate a discrete distribution by approximating Z^{EV} with n_d values via Gaussian quadrature (Miller III and Rice 1983). For each scenario index $\xi \in [\![n_d]\!]$, Gaussian quadrature produces an observation of Z^{EV}, labeled Z_ξ^{EV}, and a corresponding scenario probability p_ξ^{EV}. Gaussian quadrature guarantees that the first $2n_d - 1$ moments of the discrete distributions match those of the underlying truncated normal distributions. From these observations Z^{EV} we compute EV stock observations, labeled $EV_{\xi,t}$. Under this approach, observations across time within each scenario lie at the same cumulative probability level. Thus, scenarios with high EV stock in some time period will have high EV stock in all time periods.

We next separate the number of active batteries by age. Using data from Xu et al. (2020), we designate LS_a as the percentage of batteries of age a that retire at their current age, with a maximum lifespan of A. Denote by $AB_{\xi,t,a}$ the number of active batteries in year t, scenario ξ that are a years old. The parameter $AB_{\xi,t,a}$ is computed recursively,

$$AB_{\xi,t,a} = \begin{cases} (1 - LS_a) AB_{\xi,t-1,a-1} & a > 0 \\ \left(EV_{\xi,t} - EV_{\xi,t-1} + \sum_{a=1}^{A} LS_a AB_{\xi,t-1,a-1}\right)^+ & a = 0 \end{cases} \qquad \forall \xi \in [\![n_d]\!],\ t \in \mathcal{T},\ a \in \{0, \ldots, A\},$$

where $(\cdot)^+ = \max\{\cdot, 0\}$. By this definition, the number of new ($a = 0$) batteries is the sum of the year-over-year increase in active batteries and the number of retired batteries, and older ($a > 0$) batteries are reduced by the retirement proportion LS_a as they age. The initial state $AB_{\xi,0,a}$ is taken from historical data (IEA 2020, 2023).

We convert from battery counts to mass and split by cathode chemistry, distinguishing between two projections for market shares of battery chemistries. Xu et al. (2020) also provide these projections: in the first (the NCX projection), NMC chemistries are more prominent, and in the second (the LFP projection), LFP chemistries are more prominent. We label these market share projections $MS_{\text{NCX},t,i}$ and $MS_{\text{LFP},t,i}$, representing the proportion of new battery sales



(a) (b)

Figure 2: Ten ($n_d = 5$) new battery demand (a) and recycled battery supply (b) scenarios for chemistry NMC(622) with $\beta = 1$. NCX scenarios are shown with solid lines and LFP scenarios with dottted lines.

that are of chemistry i by year and projection. These parameters include historical data through year $-A$. We now index scenarios by market share projection and EV stock observation: $(\psi, \xi) \in \{\text{NCX}, \text{LFP}\} \times [\![n_d]\!]$. To account for changes in average battery size due to differences in battery EV (BEV) and plug-in hubrid (PHEV) adoption, we compute an additional parameter $m_{\xi,t}$, representing the average mass of a new battery. The computation of $m_{\xi,t}$ is described in the appendix.

Supply and demand for new batteries by chemistry is then given, respectively, by

$$s_{(\psi,\xi),z,t,i} = \beta_z m_{\xi,t} \sum_{a=1}^{A} MS_{\psi,t-a,i} LS_a AB_{\xi,t-1,a-1} \quad \text{and} \quad d_{(\psi,\xi),z,t,i} = \beta_z m_{\xi,t} MS_{\psi,t,i} AB_{\xi,t,0},$$

where $\beta_z \in (0, 1]$ is a scale parameter that gives the proportion of the global EV market allocated to each zone z. Figure 2 shows a set of supply and demand scenarios generated with this method.

Cost Scenarios

We assume that changes in cost coefficients are due to fluctuations in the value of "critical minerals": copper (Cu), cobalt (Co), lithium (Li), manganese (Mn), and nickel (Ni). Let $\mathcal{M} = \{\text{Cu}, \text{Co}, \text{Li}, \text{Mn}, \text{Ni}\}$ be this set of metals. From this assumption, only material costs and values vary by scenario and time; all other costs are static at the values from EverBatt.

To account for correlations between EV demand and metal prices, we construct n_c cost subscenarios for each demand scenario $(\psi, \xi) \in \{\text{NCX}, \text{LFP}\} \times [\![n_d]\!]$. We leverage metal price projections from Boer et al. (2024), who provide an annual median price estimate and a 40% highest posterior density (HPD) region under each of the SDS and STEPS projections from 2020 to 2040.[2.] Projections are extended to 2050 setting the values for years 2041-2050 to the value from 2040. Let $v_{t,m}^{\text{MED,SDS}}$ and $v_{t,m}^{\text{MED,STEPS}}$ be the median estimates under each projection, and $[v_{t,m}^{\text{LB,SDS}}, v_{t,m}^{\text{UB,SDS}}]$ and $[v_{t,m}^{\text{LB,STEPS}}, v_{t,m}^{\text{UB,STEPS}}]$ be the HPD intervals for each metal m.

As demand and cost data are both associated with the same underlying IEA projections, we use the relationship between the demand scenarios and the IEA projections to produce a set of scenario-adjusted cost distribution parameters $v^{\text{MED}}_{\xi,t,m}$, $v^{\text{LB}}_{\xi,t,m}$, and $v^{\text{UB}}_{\xi,t,m}$, e.g.,

$$v^{\text{MED}}_{\xi,t,m} = \frac{1}{2}\left|v^{\text{MED,SDS}}_{\xi,t,m} - v^{\text{MED,STEPS}}_{\xi,t,m}\right| \operatorname*{proj}_{[-1,1]}(Z^{\text{EV}}_{\xi}) + \frac{1}{2}\left(v^{\text{MED,SDS}}_{\xi,t,m} + v^{\text{MED,STEPS}}_{\xi,t,m}\right).$$

Under this formula, for demand scenarios that are closer to the SDS projection than STEPS, the cost distribution parameters are also closer to their SDS counterpart. This introduces reasonable relationships between EV demand and metal prices.

Let $v_{\xi,t,m}$ be the random variable that represents the price of metal m, with the distribution

$$v_{\xi,t,m} = \left(\frac{v^{\text{UB}}_{\xi,t,m} - v^{\text{MED}}_{\xi,t,m}}{\Phi^{-1}(0.7)}\mathbb{I}(Z^{\text{COST}} \geq 0) + \frac{v^{\text{LB}}_{\xi,t,m} - v^{\text{MED}}_{\xi,t,m}}{\Phi^{-1}(0.3)}\mathbb{I}(Z^{\text{COST}} < 0)\right) Z^{\text{COST}} + v^{\text{MED}}_{\xi,t,m},$$

where Φ is the cumulative distribution function for the standard normal distribution. Again, the randomness is described by a single random variable Z^{COST} that is distributed as a truncated standard normal on $[\Phi^{-1}(0.3), \Phi^{-1}(0.7)]$. The marginal distributions for each $v_{\xi,t,m}$ are described by two truncated normal distributions, one for samples above the median and another for those below. The distributions are constructed so that $v^{\text{MED}}_{\xi,t,m}$ matches the median of $v_{\xi,t,m}$ and the HPD interval $[v^{\text{LB}}_{\xi,t,m}, v^{\text{UB}}_{\xi,t,m}]$ defines an equal-tailed 40% credible interval for $v_{\xi,t,m}$. We again generate a discrete distribution via two Gaussian quadratures for Z^{COST}, one on the interval $[\Phi^{-1}(0.3), 0]$ and one on $[0, \Phi^{-1}(0.7)]$, with $\left\lfloor \frac{n_c}{2} \right\rfloor$ samples allocated to each. The quadrature yields values $v_{(\xi,\zeta),t,m}$ and probabilities $p^{\text{COST}}_{\xi,\zeta}$ for each scenario $\zeta \in [\![n_c]\!]$.[3.] This approach only guarantees that the first $n_c - 2$ moments of the discrete distribution match the underlying distribution, but ensures that scenario observations across time lie at the same cumulative probability level of the underlying distributions.

From the metal price observations, we construct scenario-dependent cost parameters. Consider some material cost parameter c (e.g., cost of lithium carbonate, cost of LFP cathode powder), where $\bar{c}$ gives the deterministic value of this parameter from EverBatt. Let PC_m be the proportion by mass of the material composed of the metal m. We compute the base cost $c^{\text{BASE}}_{\xi} = \bar{c} - \sum_{m\in\mathcal{M}} PC_m v^{\text{MED}}_{\xi,0,m}$, which is the cost of the material excluding the value of the metals it contains. In each cost scenario ζ, contributions from metal prices are added back to the base cost, $c_{(\xi,\zeta),t} = c^{\text{BASE}}_{\xi} + \sum_{m\in\mathcal{M}} PC_m v_{(\xi,\zeta),t,m}$. Then, the cost coefficients in (4a) are computed as the sum of the relevant scenario-dependent material cost time series $c_{(\xi,\zeta),t}$ and other utility cost terms. Figure 3 shows a set of cost scenarios generated with this method.

Under this scenario generation approach, the set of scenarios is $\Omega = \{\text{NCX}, \text{LFP}\} \times [\![n_d]\!] \times [\![n_c]\!]$, totaling $|\Omega| = 2n_d n_c$ scenarios. For $\omega = (\psi, \xi, \zeta) \in \Omega$, associated probabilities are given by $p_\omega = \frac{1}{2} p^{\text{EV}}_{\xi} p^{\text{COST}}_{\zeta}$, demand and supply by $d_{\omega,z,t,i} = d_{(\psi,\xi),z,t,i}$ and $s_{\omega,z,t,i} = s_{(\psi,\xi),z,t,i}$, and material costs by $c_{\omega,t} = c_{(\xi,\zeta),t}$. The tree for scenario generation is depicted in Figure 4. Importantly, this structure yields natural groupings for the decomposition discussed in Section 4.2.

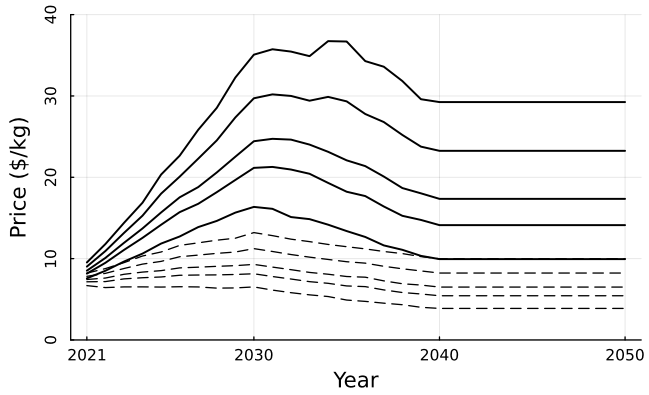


Figure 3: Cost scenarios for value v of lithium carbonate, with five subscenarios ($n_c = n_d = 5$) of demand scenario (NCX, 2) (dotted lines, low EV demand) and five of (NCX, 4) (solid lines, high EV demand).

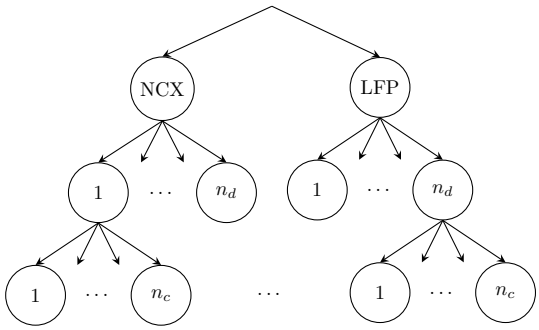


Figure 4: Tree for scenario generation. Within each market share scenario (NCX,LFP), there are n_d demand scenarios, and within each demand scenario there are n_c cost scenarios.

As ζ varies in the set $[\![n_c]\!]$, the scenario supply $s_{\omega,z,t,i}$ and demand $d_{\omega,z,t,i}$ do not change. Thus, a natural grouping is $\mathcal{G}_{\psi,\xi} = \{(\psi, \xi, \zeta)\}_{\zeta=1}^{n_c}$ for $(\psi, \xi) \in \{\text{NCX}, \text{LFP}\} \times [\![n_d]\!]$.

5.2 Computational Performance

Concave Minimization Algorithms

We first evaluate the performance of aPWL (Algorithm 1) and the impact of the reformulated problem in a setting where subproblems are solved in the extensive form, with no second-stage decomposition. These models have a single second stage, where supply and demand are taken from the SDS and NCX projections and material costs are held constant at values from EverBatt. To encourage greater branching on first-stage variables, direct recycling capacity is fixed to 0, encouraging investment in cathode production facilities in the optimal solution. The model contains a single zone, the United States (U.S.).

For the continuous problem (P), we compare aPWL to a finitely convergent spatial branch and bound algorithm (sB&B, Shectman and Sahinidis (1998)) and to the global optimization package BARON (Sahinidis 1996). For the reformulation (PMI), we compare aPWL to BARON. In this setting, we do not evaluate sB&B due to the presence of integrality constraints. To solve over discontinuous objective functions with BARON, we reformulate (5) with binary variables:

$$y \leq uz, \quad f(y) = \sum_i q_i y^{r_i} + wz, \quad z \in \{0, 1\}.$$

Table 1 compares performance in the deterministic setting by algorithm and model across a variety of scales, generated by varying the parameter $\beta_{\text{U.S.}}$. For the original formulation (P), aPWL significantly outperforms all other methods, solving all problems within the time limit. aPWL terminates up to 240x faster than sB&B and 14x faster than BARON for problems that are universally solved within the time limit. The reformulation (PMI) allows for extreme reductions in solve times, allowing global optima to be found in minutes with either aPWL or BARON. For

	Problem (P)						Problem (PMI)			
		aPWL		sB&B		BARON		aPWL		BARON
$\beta_{\text{U.S.}}$	n_y	Time (s)	# Iter	Time (s)	# Iter	Time (s)	n_y	Time (s)	# Iter	Time (s)
0.1	3,709	112.0	2	27,617.2	6,193	1,569.7	96	4.5	2	8.9
0.15	5,553	1,615.0	2	13,943.1	1,110	3,704.9	96	8.7	2	13.0
0.2	7,399	824.3	2	23,370.9	810	8,327.7	96	14.9	2	18.6
0.25	9,235	2,080.5	3	8,727.3	151	19,752.5	96	23.0	1	27.1
0.3	11,081	2,152.8	2	14,579.4	167	32,158.2	96	33.4	1	37.3
0.35	12,925	3,943.9	2	31,200.0	239	—	96	50.2	1	56.1
0.4	14,744	6,032.2	2	—	—	—	96	65.6	1	69.8
0.45	16,613	10,279.5	2	—	—	—	96	80.6	1	87.7
0.5	18,456	25,018.8	3	—	—	—	96	99.6	1	101.8

Table 1: Algorithm and formulation comparison for the extensive form of (P) and (PMI). A — indicates that the model was not solved within 10 hours. n_y gives the number of first-stage variables.

the reformulation, aPWL slightly outperforms BARON on all problems.

Decomposition Approaches

We next evaluate the performance of the Benders' decomposition when solving subproblems within aPWL. These experiments use instances of the reformulated problem (PMI) with multiple second-stage scenarios. We compare different scenario groupings against solving the subproblem in its extensive form, with no decomposition. The scenario groupings considered are the standard single-cut ($\mathcal{G}_1 = \Omega$) and multi-cut ($\mathcal{G}_\omega = \{\omega\}\ \forall \omega \in \Omega$) aggregation strategies, as well as the intermediate grouping discussed previously ($\mathcal{G}_{\psi,\xi} = \{(\psi, \xi, \zeta)\}_{\zeta=1}^{n_c}\ \forall (\psi, \xi) \in \{\text{NCX}, \text{LFP}\} \times [\![n_d]\!]$). The problems have two zones: $\mathcal{Z} = \{\text{U.S.}, \text{China}\}$. A set of test cases are generated by varying the model scale $(\beta_{\text{U.S.}}, \beta_{\text{China}})$ and the number of scenarios $|\Omega| = 2n_d n_c$, with $n_c = n_d$ for all problems. As optimal solutions in this setting have nonzero cathode production capacity, we no longer fix direct recycling capacity to 0.

Table 2 shows the computational results of the aPWL algorithm with the various decomposition approaches. Benders' decompositions outperform the extensive form in all test cases. Further, the grouped-cut aggregation strategy outperforms both single-cut and multi-cut implementations, solving up to 9x faster than the next fastest approach. The grouped-cut strategy effectively balances the tradeoff between increased subproblem size from adding many cuts in each iteration and information loss from aggregating cuts. With grouped-cut aggregation, the decomposition improves solve times by up to 44x over the extensive form. As the scale β increases, solve times decrease. This occurs due to the use of relative optimality tolerances, with objective values increasing at larger scales. However, the size of the reformulated model does not grow accordingly.

Numerical experiments are conducted in Julia 1.9 on MIT SuperCloud (Reuther et al. 2018). The experiments are run on 48 core Intel Xeon Platinum 8260@2.40 GHz processors with 192 GB of RAM. Linear and mixed-integer

$(\beta_{\text{U.S.}}, \beta_{\text{China}})$	$\lvert\Omega\rvert$	Extensive Form Time (s)	Benders' Single-Cut Time (s)	Benders' Single-Cut # Iter	Benders' Grouped-Cut Time (s)	Benders' Grouped-Cut # Iter	Benders' Multi-Cut Time (s)	Benders' Multi-Cut # Iter
(0.01,0.025)	50	4,637.8	—	—	1,810.8	91 (7)	3,732.3	64 (5)
	200	—	31,899.6	846 (10)	3,401.4	87 (9)	—	—
(0.05,0.125)	50	320.6	97.9	263 (5)	45.8	57 (7)	70.3	35 (4)
	200	1,128.3	209.8	222 (4)	107.7	53 (7)	929.7	38 (6)
(0.1,0.25)	50	241.1	90.4	257 (7)	30.8	55 (5)	60.6	39 (3)
	200	980.4	255.6	268 (6)	60.0	43 (4)	234.7	33 (2)
(0.15,0.375)	50	542.8	74.4	241 (5)	21.4	50 (1)	33.1	36 (1)
	200	1,074.3	219.7	249 (5)	64.8	48 (3)	212.9	33 (2)
(0.2,0.5)	50	1,144.6	78.6	242 (3)	25.7	51 (4)	28.6	33 (1)
	200	1,731.8	233.8	254 (4)	65.3	44 (4)	102.6	30 (1)

Table 2: Subproblem decomposition algorithm performance. # Iter counts the total Benders' iterations across all subproblems, with the number of iterations of Algorithm 1 given in parentheses. A — indicates that the model was not solved within 10 hours.

	Case 1 (U.S.)			Case 2 (U.S. + China)		
	R0	DR0	DR1	R0	DR0	DR1
Cost (billion $)	1,797.3	1,421.7	**1,407.8**	6,290.5	4,901.5	**4,882.3**
Energy Consumption (EJ)	25.7	25.3	**24.0**	90.9	89.6	**87.7**
GHG Emissions (Mt)	1,632.7	1,600.2	**1,518.8**	5,912.0	5,833.6	**5,714.0**

Table 3: Comparison of optimal solutions in expected cost and environmental impact over model horizon. R0 indicates the solution with no recycling capacity, DR0 the solution where no direct recycling is allowed, and DR1 the solution where all recycling methods are allowed.

problems are solved with Gurobi 10.0. We use BARON 23.6.23 with Ipopt 3.14.4 and CPLEX 22.1 as the nonlinear and mixed-integer subsolvers. Problems are solved to a relative optimality tolerance of 10^{-4} and a primal feasibility tolerance of 10^{-8}.

5.3 Optimal Investment Strategies

We analyze the optimal solutions of two cases. The first (Case 1) has a single zone for the U.S., with $\beta_{\text{U.S.}} = 0.2$. The second (Case 2) has two zones, one for the U.S. and one for China, with $(\beta_{\text{U.S.}}, \beta_{\text{China}}) = (0.2, 0.5)$. Both cases have 200 second-stage scenarios ($n_c = n_d = 10$). The 20% and 50% of global EV market shares approximately correspond to projections for market share in the U.S. and China, respectively, through 2030 (IEA 2023). As direct recycling is newer and less commercialized than other recycling processes, we analyze two variations of each case: one where direct recycling facilities can be constructed (labeled DR1) and one where direct recycling capacity is fixed to 0 (labeled DR0).

Comparison of Impact

To demonstrate the value of recycling, we compare to a solution where no recycling capacity is constructed (labeled R0) in objective cost, total energy consumption, and greenhouse gas (GHG) emissions. Energy consumption and GHG emissions are evaluated in expectation over second-stage scenarios and aggregated over the model horizon. These metrics are computed as linear functions of throughput with coefficients given by EverBatt, and include the impact from recycling and cathode production processes as well as the upstream contribution from purchased new materials. Improvement in these metrics is due to the replacement of new materials with their recycled counterpart during new battery manufacturing.

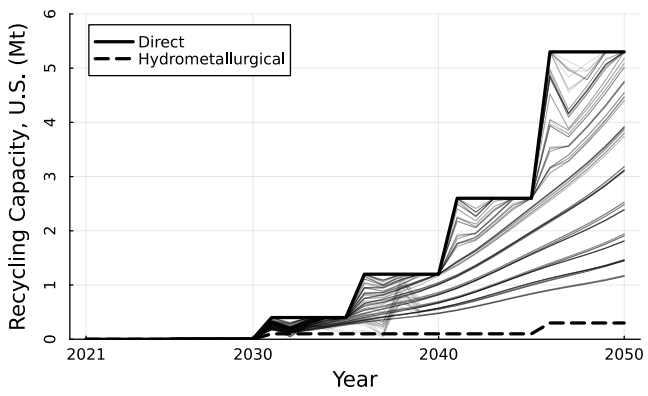
Table 3 compares the metrics across the variations R0, DR0, and DR1. Incorporating recycling into the supply chain reduces costs by about 22% in Case 2, with 0.3% additional reduction when direct recycling is allowed. In Case 1, the cost savings are more significant under direct recycling, increasing from 21% to 22%. Across both cases, energy consumption is reduced marginally ($\sim 1\%$) when non-direct recycling is adopted, with more significant reductions of 6% in Case 1 and 4% in Case 2 under direct recycling. Similarly, GHG emissions are reduced $1-2\%$ under non-direct recycling, and up to 7% under direct recycling. While there are large, comparable cost savings under direct or non-direct recycling, improvements in environmental impact are more significant under direct recycling, especially when only considering production in the U.S.

Optimal Investment Plan

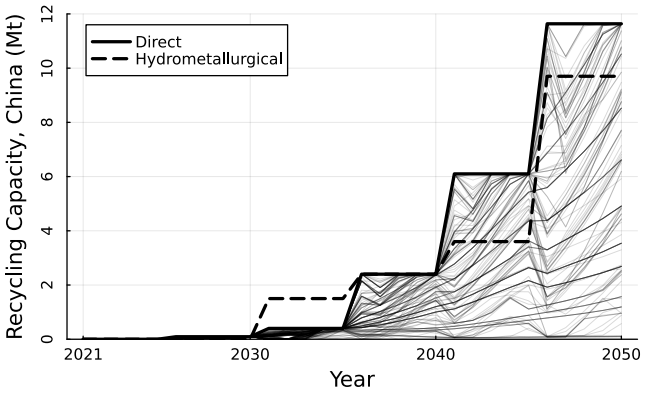
Figure 5 depicts the optimal capacity decisions for the DR1 variations of both cases. Analysis of the DR0 variations is included in the appendix. In Case 1, the U.S. primarily constructs direct recycling capacity starting in 2026, supplemented by a lesser amount of hydrometallurgical capacity. The optimal solution has no pyrometallurgical capacity, and only a single cathode production line, with 2 kt of annual capacity, is constructed for NMC(811), opening in 2031. In Case 2, a mixture of hydrometallurgical and direct recycling capacity is constructed in China, starting in 2026. The U.S. constructs only direct recycling capacity in a lesser amount than in Case 1, starting in 2036. Cathode production capacity for all NMC chemistries is developed in China.[4.] In the second stage, retired batteries are often transported from the U.S. to China to be recycled, then recycled cathode powder is transported back to the U.S.

Closed-Loop Recycled Material Usage

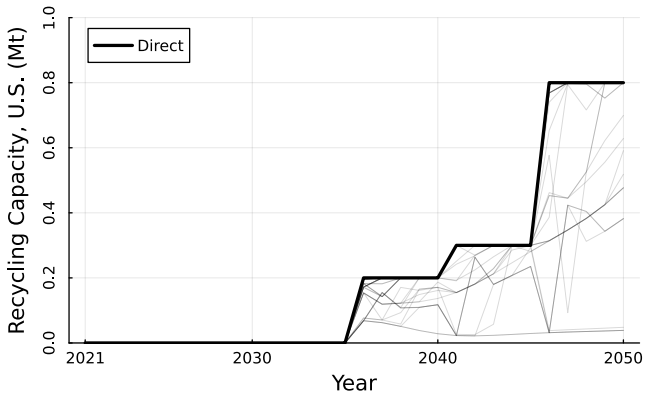
Figure 6 visualizes the proportion of cathode material used in new battery manufacturing that is sourced from recycled batteries as a distribution over second-stage scenarios. This proportion is computed for each scenario and planning period as $\frac{\sum_{z\in\mathcal{Z},\ k\in\mathcal{K}^{\mathrm{CP}},\ t\in\mathcal{T}_l} x^{\mathrm{INV,NB}}_{z,\omega,t,k}}{\sum_{z\in\mathcal{Z},\ k\in\mathcal{K}^{\mathrm{CP}},\ t\in\mathcal{T}_l} \left(x^{\mathrm{INV,NB}}_{z,\omega,t,k}+x^{\mathrm{NM,NB}}_{z,\omega,t,k}\right)}$; this value corresponds to the closed-loop recycling potential defined by Xu et al. (2020). LFP chemistry scenarios have lower recycling rates than NCX scenarios because the common chemistries in LFP scenarios do not contain cobalt and thus are cheaper to acquire. This property gives the distributions a bimodal



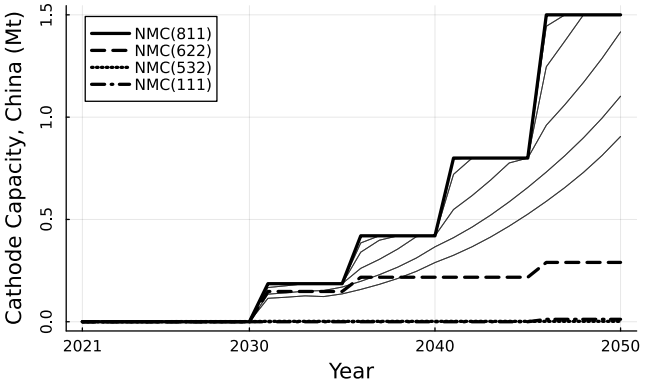
(a) Case 1, U.S. Recycling Capacity



(b) Case 2, China Recycling Capacity

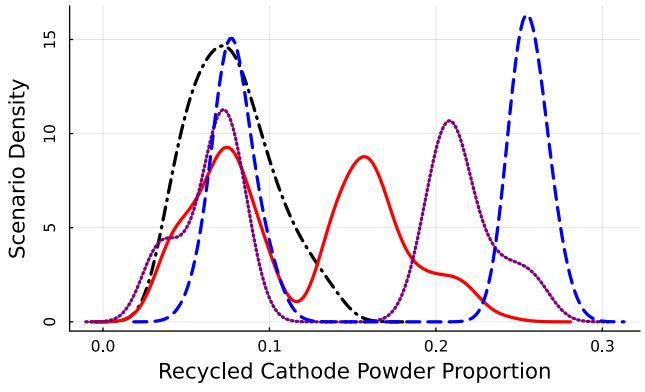


(c) Case 2, U.S. Recycling Capacity

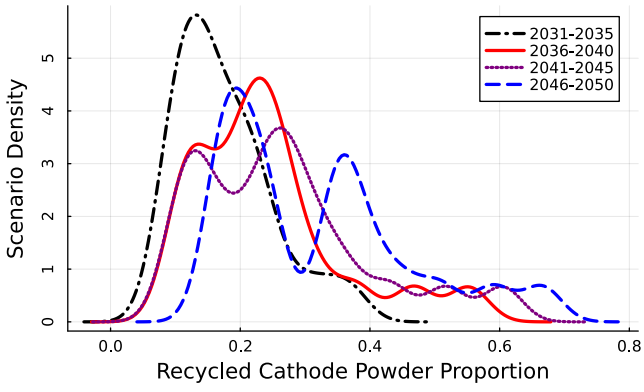


(d) Case 2, China Cathode Production Capacity

Figure 5: Optimal capacity investment decisions for recycling and cathode production facilities by zone. Utilization of direct recycling capacity (a,b,c) and NMC(811) cathode production capacity (d) for each second-stage scenario is shown with thin lines. Cathode production capacity for Case 1 is not shown.



(a) Case 1



(b) Case 2

Figure 6: Density across scenarios of the proportion of cathode powder used in new battery manufacturing that is produced via recycling, indexed by planning period.

structure. The proportion of cathode powder obtained from recycling increases in later planning periods, approaching a maximum of 30% for some scenarios in Case 1 and 70% in Case 2. The expected proportion of cathode powder obtained from recycling across the model horizon is 13% in Case 1 and 23% in Case 2.

Open-Loop Solutions

Our model takes the perspective of closed-loop supply chain, where materials recovered from recycling are used to manufacture new batteries. The open-loop perspective instead sells the recovered materials to a market. Under the assumption that the nonnegativity constraint on $x^{\mathrm{NM,NB}}$ is nonbinding, $x^{\mathrm{NM,NB}}$ can be replaced with its definition from constraint (1a). This assumption holds if material requirements for new battery manufacture cannot be met through recycling alone. Then, the objective contribution from $x^{\mathrm{NM,NB}}$ becomes

$$\sum_{z\in\mathcal{Z},\ \omega\in\Omega,\ t\in\mathcal{T},\ k\in\mathcal{K}^{\mathrm{CP}}} c^{\mathrm{NM,NB}}_{\omega,t,k} x^{\mathrm{NM,NB}}_{z,\omega,t,k} = \sum_{z\in\mathcal{Z},\ \omega\in\Omega,\ t\in\mathcal{T},\ k\in\mathcal{K}^{\mathrm{CP}}} c^{\mathrm{NM,NB}}_{\omega,t,k} \left(\sum_{i\in\mathcal{I}} \Delta^{\mathrm{NB}}_{i,k} d_{\omega,z,t,i} - x^{\mathrm{INV,NB}}_{z,\omega,t,k} \right).$$

This change effectively removes the new battery manufacture portion of the model and adds a revenue associated with "selling" recycled cathode powder at the market price. Thus, the solutions from our model are also representative of optimal decisions under an open-loop recycling model.

6 Conclusion

In this work, we formulate a model to determine optimal investment strategies in EV battery recycling and cathode production facilities, derived from the EverBatt tool. We also propose an equivalent formulation of the model which reduces solve times by introducing integrality constraints. We construct a procedure for generating scenarios for new battery demand, retired battery supply, and material costs. To solve the model, we introduce an algorithm for separable concave minimization over polyhedra that improves performance over existing algorithms and a cut grouping strategy for Benders' decomposition that converges faster than standard approaches. The optimal solutions show large improvements in the cost of manufacturing new batteries when recycling is incorporated into the supply chain, as well as reductions in the environmental impact of battery manufacturing. Overall, the model and solution approaches provide a framework for players in the EV battery market to plan for the integration of recycling technologies into the supply chain.

Endnotes

1. Pourmohammadi et al. (2008) estimate η as 60% for recycled aluminum. Recycled carbon black is similarly discounted by $\eta \in [70\%, 85\%]$ relative to its virgin counterpart (Contec 2023).

2. Boer et al. (2024) do not provide projections for manganese prices. However, we find that manganese and copper

prices are highly correlated. Price data from IMF (2022) between 2013 and 2021 show that manganese and copper prices have a correlation coefficient of 0.91. As validation, a correlation of prices between 1950 and 2010 from U.S. Geological Survey (2013) yields a coefficient of 0.85. Given this high correlation, we predict the median value and HPD bounds for manganese from the corresponding values for copper using a linear regression fitted on data from IMF (2022).

3. The quadrature probabilities are reweighted by $\frac{1}{2}$ to account for $\mathbb{P}(Z^{\text{COST}} \geq 0)$ and $\mathbb{P}(Z^{\text{COST}} < 0)$.

4. EverBatt does not require a cathode production facility to manufacture LFP cathode powder, likely due to the simplicity of its manufacturing process. For this reason, we allow recycled LFP to be manufactured regardless of the amount of constructed cathode production capacity.

References

Adulyasak Y, Cordeau JF, Jans R (2015) Benders decomposition for production routing under demand uncertainty. *Operations Research* 63(4):851–867.

Azzi A, Battini D, Faccio M, Persona A, Sgarbossa F (2014) Inventory holding costs measurement: a multi-case study. *The International Journal of Logistics Management* 25(1):109–132.

BAF (2023) Partnership between BASF and Nanotech Energy will enable production of lithium-ion batteries in North America with locally recycled content and low CO2 footprint. Accessed June 2024, `https://www.basf.com/global/en/media/news-releases/2023/09/p-23-292.html`.

Bazaraa MS, Sherali HD (1982) On the use of exact and heuristic cutting plane methods for the quadratic assignment problem. *Journal of the Operational Research Society* 33(11):991–1003.

Benders JF (1962) Partitioning procedures for solving mixed-variables programming problems. *Numerische Mathematik* 4(1):238–252.

Benson HP (1985) A finite algorithm for concave minimization over a polyhedron. *Naval Research Logistics Quarterly* 32(1):165–177.

Birge JR, Louveaux F (2011) *Introduction to Stochastic Programming* (Springer Science & Business Media).

Boer L, Pescatori A, Stuermer M (2024) Energy transition metals: Bottleneck for net-zero emissions? *Journal of the European Economic Association* 21(1):200–229.

Campagnol N, Pfeiffer A, Tryggestad C (2022) Capturing the battery value-chain opportunity. Accessed October 2023, `https://www.mckinsey.com/industries/electric-power-and-natural-gas/our-insights/capturing-the-battery-value-chain-opportunity`.

Contec (2023) On the rise: The price of carbon black isn't sustainable. Accessed December 2023, `https://contec.tech/price-of-carbon-black-sustainability`.

Council of the European Union (2023) Council adopts new regulation on batteries and waste batteries. Accessed October 2023, `https://www.consilium.europa.eu/en/press/press-releases/2023/07/10/council-adopts-new-regulation-on-batteries-and-waste-batteries/`.

Dai Q, Spangenberger J, Ahmed S, Gaines L, Kelly JC, Wang M (2019) EverBatt: A closed-loop battery recycling cost and environmental impacts model. Technical report, Argonne National Lab.

D'Ambrosio C, Lee J, Wächter A (2009) A global-optimization algorithm for mixed-integer nonlinear programs having separable non-convexity. *European Symposium on Algorithms*, 107–118.

Falk JE, Soland RM (1969) An algorithm for separable nonconvex programming problems. *Management Science* 15(9):550–569.

Feldman E, Lehrer FA, Ray TL (1966) Warehouse location under continuous economies of scale. *Management Science* 12(9):670–684.

Forbes (2022) VW and Audi to recycle EV batteries through Tesla cofounder's company. Accessed June 2024, `https://www.forbes.com/sites/alanohnsman/2022/07/12/vw-and-audi-to-recycle-ev-batteries-through-tesla-cofounders-company`.

Harper G, Sommerville R, Kendrick E, Driscoll L, Slater P, Stolkin R, Walton A, Christensen P, Heidrich O, Lambert S, Abbott A, Ryder K, Gaines L, Anderson P (2019) Recycling lithium-ion batteries from electric vehicles. *Nature* 575:75–86.

Horst R, Pardalos PM, Van Thoai N (2000) *Introduction to Global Optimization* (Springer Science & Business Media).

Hoyer C, Kieckhäfer K, Spengler TS (2015) Technology and capacity planning for the recycling of lithium-ion electric vehicle batteries in Germany. *Journal of Business Economics* 85:505–544.

IEA (2020) Global EV outlook 2020. Technical report, IEA.

IEA (2022) Global supply chains of EV batteries. Technical report, IEA.

IEA (2023) Global EV data explorer. Accessed August 2023, `https://www.iea.org/data-and-statistics/data-tools/global-ev-data-explorer`.

IMF (2022) Primary commodity price system. Accessed December 2022, `https://data.imf.org/commodityprices`.

Kang DHP, Chen M, Ogunseitan OA (2013) Potential environmental and human health impacts of rechargeable lithium batteries in electronic waste. *Environmental Science & Technology* 47(10):5495–5503.

Kneifel J, Webb D (2020) Life cycle costing manual for the federal energy management program. `https://doi.org/10.6028/NIST.HB.135-2020`.

Lee SB, Luss H (1987) Multifacility-type capacity expansion planning: algorithms and complexities. *Operations Research* 35(2):249–253.

Li L, Dababneh F, Zhao J (2018) Cost-effective supply chain for electric vehicle battery remanufacturing. *Applied Energy* 226:277–286.

Li P, Huai Z, Jiang J (2023) Optimization of closed-loop supply chain network design under uncertainty: Considering electric vehicle battery recycling. *Industrial Engineering and Applications*, volume 35 of *Advances in Transdisciplinary Engineering*, 182–190.

Li S, Tirupati D (1994) Dynamic capacity expansion problem with multiple products: Technology selection and timing of capacity additions. *Operations Research* 42(5):958–976.

Magnanti TL, Stratila D (2004) Separable concave optimization approximately equals piecewise linear optimization. *Integer Programming and Combinatorial Optimization*, 234–243.

Makuza B, Tian Q, Guo X, Chattopadhyay K, Yu D (2021) Pyrometallurgical options for recycling spent lithium-ion batteries: A comprehensive review. *Journal of Power Sources* 491:229622.

Marano V, Onori S, Guezennec Y, Rizzoni G, Madella N (2009) Lithium-ion batteries life estimation for plug-in hybrid electric vehicles. *2009 IEEE Vehicle Power and Propulsion Conference*, 536–543.

Martínez-Costa C, Mas-Machuca M, Benedito E, Corominas A (2014) A review of mathematical programming models for strategic capacity planning in manufacturing. *International Journal of Production Economics* 153:66–85.

Miller III AC, Rice TR (1983) Discrete approximations of probability distributions. *Management Science* 29(3):352–362.

Pardalos PM, Romeijn HE (2002) *Handbook of Global Optimization*, volume 2 (Springer).

Pardalos PM, Rosen JB (1986) Methods for global concave minimization: A bibliographic survey. *SIAM Review* 28(3):367–379.

Pourmohammadi H, Rahimi M, Dessouky M (2008) Sustainable reverse logistics for distribution of industrial waste/byproducts: A joint optimization of operation and environmental costs. *Supply Chain Forum: An International Journal*, volume 9, 2–17.

Raghavachari M (1969) On connections between zero-one integer programming and concave programming under linear constraints. *Operations Research* 17(4):680–684.

Rajagopalan S, Singh MR, Morton TE (1998) Capacity expansion and replacement in growing markets with uncertain technological breakthroughs. *Management Science* 44(1):12–30.

Reuther A, Kepner J, Byun C, Samsi S, Arcand W, Bestor D, Bergeron B, Gadepally V, Houle M, Hubbell M, Jones M, Klein A, Milechin L, Mullen J, Prout A, Rosa A, Yee C, Michaleas P (2018) Interactive supercomputing on 40,000 cores for machine learning and data analysis. *IEEE High Performance Extreme Computing Conference*, 1–6.

Rosen JB (1983) Global minimization of a linearly constrained concave function by partition of feasible domain. *Mathematics of Operations Research* 8(2):215–230.

Rosenberg S, Glöser-Chahoud S, Huster S, Schultmann F (2023) A dynamic network design model with capacity expansions for EoL traction battery recycling – A case study of an OEM in Germany. *Waste Management* 160:12–22.

Sahinidis NV (1996) BARON: A general purpose global optimization software package. *Journal of Global Optimization* 8:201–205.

Shectman JP, Sahinidis NV (1998) A finite algorithm for global minimization of separable concave programs. *Journal of Global Optimization* 12:1–36.

Tadaros M, Migdalas A, Samuelsson B, Segerstedt A (2022) Location of facilities and network design for reverse logistics of lithium-ion batteries in Sweden. *Operational Research* 22:895–915.

Tuy H (1964) Concave programming with linear constraints. *Doklady Akademii Nauk* 159(1):32–35.

US Geological Survey (2013) Metal prices in the United States through 2010. Technical report, U.S. Geological Survey, `http://dx.doi.org/10.3133/sir20125188`.

US Geological Survey (2022) 2022 final list of critical minerals. Accessed October 2023, `https://www.usgs.gov/news/national-news-release/us-geological-survey-releases-2022-list-critical-minerals`.

Van Mieghem JA (2003) Capacity management, investment, and hedging: Review and recent developments. *Manufacturing & Service Operations Management* 5(4):269–302.

Wang L, Wang X, Yang W (2020) Optimal design of electric vehicle battery recycling network – From the perspective of electric vehicle manufacturers. *Applied Energy* 275:115328.

Wu J, Zheng M, Liu T, Wang Y, Liu Y, Nai J, Zhang L, Zhang S, Tao X (2023) Direct recovery: A sustainable recycling technology for spent lithium-ion battery. *Energy Storage Materials* 54:120–134.

Xu C, Dai Q, Gaines L, Hu M, Tukker A, Steubing B (2020) Future material demand for automotive lithium-based batteries. *Communications Materials* 1:99.

Yao Y, Zhu M, Zhao Z, Tong B, Fan Y, Hua Z (2018) Hydrometallurgical processes for recycling spent lithium-ion batteries: a critical review. *ACS Sustainable Chemistry & Engineering* 6(11):13611–13627.

A Appendix

This appendix is organized as follows. Section A.1 contains tables that summarize the data and variables used in model construction. Section A.2 describes the computation of the average battery mass parameter used in demand scenario generation. Section A.3 analyzes the optimal solutions of the DR0 variations of Cases 1 and 2. Section A.4 provides proofs of our results.

A.1 Summary of Model Data

Tables A.1, A.2, and A.3 give descriptions of the index sets, constant parameters, and variables used to define the models (P) and (PMI).

Table A.1: Sets

$\mathcal{I}$	Set of battery chemistries
$\mathcal{J}$	Set of recycling processes
$\mathcal{K}$	Set of materials encountered during recycling and manufacturing
$\mathcal{K}^{\mathrm{CP}}$	Set of cathode powders
$\mathcal{L}$	Set of planning periods in the model horizon
$\mathcal{N}_{l,k}^{\mathrm{CP}}$	Set of cathode production line indices used to produce material $k \in \mathcal{K}^{\mathrm{CP}}$ during planning period $l \in \mathcal{L}$
$\mathcal{N}_{l}^{\mathrm{REC}}$	Set of recycling facility indices during planning period $l \in \mathcal{L}$
$\mathcal{T}$	Set of time periods in the model horizon
$\mathcal{T}_l$	Set of time periods associated with planning period $l \in \mathcal{L}$
$\mathcal{Z}$	Set of zones in which production facilities can be located
Ω	Set of second-stage supply, demand and cost scenarios

Table A.2: Parameters

$c_{\omega,z,t,k}^{\mathrm{CP}}$	Variable cost of producing one unit of material $k \in \mathcal{K}^{\mathrm{CP}}$ via cathode production in time period $t \in \mathcal{T}$, zone $z \in \mathcal{Z}$ and scenario $\omega \in \Omega$
$c_{\omega,t,k}^{\mathrm{CP,NM}}$	Cost per unit of material $k \in \mathcal{K} \setminus \mathcal{K}^{\mathrm{CP}}$ purchased new to manufacture cathode powder in time period $t \in \mathcal{T}$ and scenario $\omega \in \Omega$
$c_{\omega,z,t,k}^{\mathrm{MC}}$	Variable cost for generating one unit of material $k \in \mathcal{K} \setminus \mathcal{K}^{\mathrm{CP}}$ via material conversion in time period $t \in \mathcal{T}$, zone $z \in \mathcal{Z}$ and scenario $\omega \in \Omega$
$c_{\omega,t,k}^{\mathrm{NB,NM}}$	Cost per unit of material $k \in \mathcal{K}$ purchased new to manufacture batteries in time period $t \in \mathcal{T}$ and scenario $\omega \in \Omega$
$c_{\omega,z,t,i,j}^{\mathrm{REC}}$	Variable cost per unit mass of battery of chemistry $i \in \mathcal{I}$ recycled with process $j \in \mathcal{J}$ in time period $t \in \mathcal{T}$, zone $z \in \mathcal{Z}$ and scenario $\omega \in \Omega$

$c^{\text{TR,RB}}_{z,z'}$	Cost per unit mass of battery transported from zone $z \in \mathcal{Z}$ to zone $z' \in \mathcal{Z} \setminus \{z\}$
$c^{\text{TR,RM}}_{z,z'}$	Cost to transport a unit of recycled material from zone $z \in \mathcal{Z}$ to zone $z' \in \mathcal{Z} \setminus \{z\}$
$C^{\text{OP}}_{\omega,t}$	Operational costs as a function of operational decisions in time period $t \in \mathcal{T}$ and scenario $\omega \in \Omega$
C^{PL}_{t}	Planning costs as a function of capacity decisions in time period $t \in \mathcal{T}$
$\overline{C}^{\text{PL}}_{t}$	Planning costs as a function of capacity decisions in time period $t \in \mathcal{T}$ for the reformulated capacity variable structure
$d_{\omega,z,t,i}$	Demand for batteries of chemistry $i \in \mathcal{I}$ in time period $t \in \mathcal{T}$, zone $z \in \mathcal{Z}$ and scenario $\omega \in \Omega$
$f^{\text{CP}}_{z,k}$	Concave function mapping cathode production capacity for a line producing material $k \in \mathcal{K}^{\text{CP}}$ in zone $z \in \mathcal{Z}$ to annual labor and capital cost
$f^{\text{REC}}_{z,j}$	Concave function mapping recycling capacity for process $j \in \mathcal{J}$ in zone $z \in \mathcal{Z}$ to annual labor and capital cost
l_t	Planning period associated with time period $t \in \mathcal{T}$
p_ω	Probability associated with scenario $\omega \in \Omega$
$s_{\omega,z,t,i}$	Supply of batteries of chemistry $i \in \mathcal{I}$ in time period $t \in \mathcal{T}$, zone $z \in \mathcal{Z}$ and scenario $\omega \in \Omega$
u^{CP}	Maximum capacity of a cathode production line
u^{REC}	Maximum capacity of a recycling process at each facility
$v_{\omega,t,k}$	Value per unit of material $k \in \mathcal{K}$ in time period $t \in \mathcal{T}$ and scenario $\omega \in \Omega$
$\Delta^{\text{CP}}_{k',k}$	Amount of material $k \in \mathcal{K} \setminus \mathcal{K}^{\text{CP}}$ used to produce one unit of material $k' \in \mathcal{K}^{\text{CP}}$ via cathode production
$\Delta^{\text{MC}}_{k',k}$	Amount of material $k \in \mathcal{K} \setminus \mathcal{K}^{\text{CP}}$ used to produce one unit of material $k' \in \mathcal{K} \setminus \mathcal{K}^{\text{CP}}$ via material conversion
$\Delta^{\text{NB}}_{i,k}$	Amount of material $k \in \mathcal{K}$ used to produce a unit mass of battery of chemistry $i \in \mathcal{I}$
$\Delta^{\text{REC}}_{k,i,j}$	Amount of material $k \in \mathcal{K}$ generated by recycling a unit mass of battery of chemistry $i \in \mathcal{I}$ with process $j \in \mathcal{J}$
γ	Annual discount factor
η	Proportion of material value recovered when selling recycled material to the market
ρ	Proportion of material value incurred as annual inventory cost

Table A.3: Variables

$x^{\text{CP,INV}}_{\omega,z,t,k}$	Amount of output material $k \in \mathcal{K}^{\text{CP}}$ from cathode production put in inventory in time period $t \in \mathcal{T}$, zone $z \in \mathcal{Z}$ and scenario $\omega \in \Omega$
$x^{\text{INV}}_{\omega,z,t,k}$	Amount of recycled material $k \in \mathcal{K}$ in inventory in time period $t \in \mathcal{T}$, zone $z \in \mathcal{Z}$ and scenario $\omega \in \Omega$
$x^{\text{INV,MC}}_{\omega,z,t,k}$	Amount of recycled material $k \in \mathcal{K} \setminus \mathcal{K}^{\text{CP}}$ used in material conversion in time period $t \in \mathcal{T}$, zone $z \in \mathcal{Z}$ and scenario $\omega \in \Omega$
$x^{\text{INV,NB}}_{\omega,z,t,k}$	Amount of recycled material $k \in \mathcal{K}^{\text{CP}}$ used to produce new batteries in time period $t \in \mathcal{T}$, zone $z \in \mathcal{Z}$ and scenario $\omega \in \Omega$

$x^{\text{MC,CP}}_{\omega,z,t,k}$	Amount of output material $k \in \mathcal{K} \setminus \mathcal{K}^{\text{CP}}$ from material conversion used for cathode production in time period $t \in \mathcal{T}$, zone $z \in \mathcal{Z}$ and scenario $\omega \in \Omega$
$x^{\text{NM,CP}}_{\omega,z,t,k}$	Amount of new material $k \in \mathcal{K} \setminus \mathcal{K}^{\text{CP}}$ used in cathode production in time period $t \in \mathcal{T}$, zone $z \in \mathcal{Z}$ and scenario $\omega \in \Omega$
$x^{\text{NM,NB}}_{\omega,z,t,k}$	Amount of new material $k \in \mathcal{K}$ used to produce new batteries in time period $t \in \mathcal{T}$, zone $z \in \mathcal{Z}$ and scenario $\omega \in \Omega$
$x^{\text{RB}}_{\omega,z,t,i}$	Amount of retired battery of chemistry $i \in \mathcal{I}$ in inventory in time period $t \in \mathcal{T}$, zone $z \in \mathcal{Z}$ and scenario $\omega \in \Omega$
$x^{\text{RB,RM}}_{\omega,z,t,i,j}$	Amount of retired battery of chemistry $i \in \mathcal{I}$ recycled with process $j \in \mathcal{J}$ in time period $t \in \mathcal{T}$, zone $z \in \mathcal{Z}$ and scenario $\omega \in \Omega$
$x^{\text{RM,INV}}_{\omega,z,t,k}$	Amount of output material $k \in \mathcal{K}$ from battery recycling put in inventory in time period $t \in \mathcal{T}$, zone $z \in \mathcal{Z}$ and scenario $\omega \in \Omega$
$x^{\text{RM,S}}_{\omega,z,t,k}$	Amount of output material $k \in \mathcal{K}$ from battery recycling sold to the market in time period $t \in \mathcal{T}$, zone $z \in \mathcal{Z}$ and scenario $\omega \in \Omega$
$x^{\text{TR,RB}}_{\omega,z,z',t,i}$	Amount of retired battery with chemistry $i \in \mathcal{I}$ transported from zone $z \in \mathcal{Z}$ to zone $z' \in \mathcal{Z} \setminus \{z\}$ in time period $t \in \mathcal{T}$ and scenario $\omega \in \Omega$
$x^{\text{TR,RM}}_{\omega,z,z',t,k}$	Amount of recycled material $k \in \mathcal{K}$ transported from zone $z \in \mathcal{Z}$ to zone $z' \in \mathcal{Z} \setminus \{z\}$ in time period $t \in \mathcal{T}$ and scenario $\omega \in \Omega$
$y^{\text{CP}}_{z,l,k,n}$	Annual capacity of cathode production line $n \in \mathcal{N}^{\text{CP}}_{l,k}$ producing material $k \in \mathcal{K}^{\text{CP}}$ in zone $z \in \mathcal{Z}$ during planning period $l \in \mathcal{L}$
$y^{\text{REC}}_{z,l,j,n}$	Annual capacity of recycling process $j \in \mathcal{J}$ at facility $n \in \mathcal{N}^{\text{REC}}_{l}$ in zone $z \in \mathcal{Z}$ during planning period $l \in \mathcal{L}$
$y^{\text{CP}}_{z,l,k}$	Number of cathode production lines producing material $k \in \mathcal{K}^{\text{CP}}$ in zone $z \in \mathcal{Z}$ during planning period $l \in \mathcal{L}$ with capacity at the upper bound u^{CP}
$y^{\text{REC}}_{z,l,j}$	Number of facilities in zone $z \in \mathcal{Z}$ during planning period $l \in \mathcal{L}$ with capacity for recycling process $j \in \mathcal{J}$ at the upper bound u^{REC}
$y^{\text{CP}}_{z,l,k,+}$	Annual capacity of the cathode production line producing material $k \in \mathcal{K}^{\text{CP}}$ in zone $z \in \mathcal{Z}$ during planning period $l \in \mathcal{L}$ that does not have capacity at its upper or lower bound
$y^{\text{REC}}_{z,l,j,+}$	Annual capacity of the recycling facility for process $j \in \mathcal{J}$ in zone $z \in \mathcal{Z}$ during planning period $l \in \mathcal{L}$ that does not have capacity at its upper or lower bound

A.2 Battery Mass Parameter Calculation

The parameter $m_{\xi,t}$ gives the average mass of a new battery by year and scenario, accounting for differences in battery size between BEVs and PHEVs. Let BEV_t^{SDS} and BEV_t^{STEPS} represent the proportion of EV stock composed of BEVs

under the SDS and STEPS projections, respectively, using data from Xu et al. (2020). Note that $1 - BEV_t^{\text{SDS}}$ gives the proportion composed of PHEVs. Similarly to EV_t, we compute samples for BEV_t from the scenario quadrature observations Z_ξ^{EV}:

$$BEV_{\xi,t} = \frac{1}{2}|BEV_t^{\text{SDS}} - BEV_t^{\text{STEPS}}|Z_{\xi,t}^{\text{EV}} + \frac{1}{2}(BEV_t^{\text{SDS}} + BEV_t^{\text{STEPS}}).$$

We then convert the proportion of battery stock $BEV_{\xi,t}$ to a proportion of new battery sales $\overline{BEV}_{\xi,t}$ recursively by

$$\overline{BEV}_{\xi,t} = \frac{BEV_{\xi,t}EV_{\xi,t} - BEV_{\xi,t-1}EV_{\xi,t-1} + \sum_{a=1}^{A} \overline{BEV}_{\xi,t-a} LS_a AB_{\xi,t,a}}{AB_{\xi,t,0}}.$$

Note that historical BEV proportions are available through year $-A$. Let m^{BEV} and m^{PHEV} be the average mass of BEV and PHEV batteries (Xu et al. 2020). The average mass of a new battery is then computed

$$m_{\xi,t} = m^{\text{BEV}}\overline{BEV}_{\xi,t} + m^{\text{PHEV}}(1 - \overline{BEV}_{\xi,t}).$$

A.3 Optimal Investment Plan for DR0 Variations

Figure A.1 depicts the optimal capacity decisions for the DR0 variations of Cases 1 and 2. Across both cases, total recycling capacity is comparable to the DR1 variations, with direct recycling capacity replaced with hydrometallurgical. The amount of capacity constructed in the U.S. is reduced in Case 2. In both cases, construction of recycling capacity begins in 2026. In Case 1, cathode production capacity is constructed for NMC(811) and NMC(622), while almost no capacity (a single line) was constructed in the DR1 variation. Case 2 similarly increases the amount of NMC(811) and NMC(622) cathode production capacity and eliminates capacity for NMC(532) and NMC(111).

Figure A.2 visualizes the density of recycled cathode powder proportion by scenario in DR0 variations, using the same methodology as in Figure 6. The proportion of recycled cathode powder increases over time, and the distributions are right-tailed, meaning that a few scenarios see high representation of recycled material in cathode powder consumed for new battery manufacture. These scenarios generally correspond to scenarios with below-average demand for new batteries. In the DR0 variations, the expected proportion of cathode powder obtained from recycling across the model horizon is 26% in Case 1, and is 32% in Case 2. The representation of recycled cathode powder is higher in DR0 variations than DR1. This occurs because pure metals are recovered from hydrometallurgical recycling and can be manufactured into any type of cathode powder, whereas with direct recycling the recovered cathode powder type is always the same as that of the recycled battery. Such flexibility allows for increased use of recycled materials, albeit at a slightly higher total cost.

A.4 Proofs of Results

Proposition 1. *For any $\overline{y} \geq 0$ that satisfies* (3c)-(3f), $\{(y, x) \geq 0 : (1),\ (2),\ (3),\ y = \overline{y}\} \neq \emptyset$.

Proof. Fix $y \geq 0$ so that (3c)-(3f) are satisfied. We construct a corresponding feasible x that only buys new materials

to manufacture batteries and does not recycle. Take $x^{\mathrm{NM,NB}}_{\omega,z,t,k} = \sum_{i\in\mathcal{I}} \Delta^{\mathrm{NB}}_{i,k} d_{\omega,z,t,i}$. Further, set $x^{\mathrm{RB}}_{\omega,z,t,i} = \sum_{t'=1}^{t} s_{\omega,z,t',i}$ and $x^{\mathrm{RB}}_{\omega,z,0,i} = 0$. Set all other components of x to 0. Because battery supply s, demand d, and material requirements Δ are nonnegative under Assumption 1, $x \geq 0$.

The construction of $x^{\mathrm{NM,NB}}$ ensures that (1a) and (1b) are satisfied, and that of x^{RB} ensures that (2c) is satisfied. The remainder of constraints in (1) and (2) only involve components of x that are 0 and are thus satisfied. As $y \geq 0$, the solution x is feasible for (3a) and (3b). Thus, given y, we can construct $x \geq 0$ so that constraints (1)-(3) are satisfied.

□

Lemma A.1 (Horst et al. (2000), Theorem 1.19)**.** *Let $\mathcal{X}$ be a bounded polytope and $f : \mathcal{X} \to \mathbb{R}$ be a concave function. Then, f attains its minimum at an extreme point of $\mathcal{X}$.*

Proof. Let $\{x^{(j)}\}_{j=1}^{J}$ be the extreme points of $\mathcal{X}$. As $\mathcal{X}$ is bounded, $\mathcal{X} = \mathrm{conv}\{x^{(j)}\}_{j=1}^{J}$. Then, for any $x \in \mathcal{X}$, there is some $\lambda \geq 0$ with $\sum_{j=1}^{J} \lambda_j = 1$ and $\sum_{j=1}^{J} \lambda_j x^{(j)} = x$. By concavity of f,

$$f(x) = f\left(\sum_{j=1}^{J} \lambda_j x^{(j)}\right) \geq \sum_{j=1}^{J} \lambda_j f(x^{(j)}) \geq \min_{j\in[\![J]\!]} f(x^{(j)}) \left(\sum_{j=1}^{J} \lambda_j\right) = \min_{j\in[\![J]\!]} f(x^{(j)}).$$

Therefore, f achieves its minimum on $\mathcal{X}$ at some element of $\{x^{(j)}\}_{j=1}^{J}$, which is an extreme point of $\mathcal{X}$.

□

Lemma A.2. *Let f be a concave function and $\mathcal{X} \subseteq \mathbb{R}^p_+$ be a nonempty polyhedron. Then, if $\inf_{x\in\mathcal{X}} f(x) > -\infty$, it attains a finite optimal solution at an extreme point of $\mathcal{X}$.*

Proof. Suppose $\inf_{x\in\mathcal{X}} f(x) > -\infty$. Let $\{x^{(j)}\}_{j=1}^{J}$ and $\{r^{(k)}\}_{k=1}^{K}$ be the extreme points and extreme rays of $\mathcal{X}$, and denote $\epsilon = \inf_{x\in\mathcal{X}} f(x)$, which is finite as the problem is feasible. Note that

$$\mathcal{X} = \mathrm{conv}\{x^{(j)}\}_{j=1}^{J} + \mathrm{cone}\{r^{(k)}\}_{k=1}^{K}.$$

Further, as $\mathcal{X} \subseteq \mathbb{R}^p_+$, it does not contain a line, which implies that it has as least one extreme point.

Select any $x \in \mathrm{conv}\{x^{(j)}\}_{j=1}^{J}$ and $r \in \mathrm{cone}\{r^{(k)}\}_{k=1}^{K}$. To prove by contradiction, suppose that $f(x+r) < f(x)$. Take $\delta = \frac{f(x)-\epsilon+1}{f(x)-f(x+r)}$ and note that $\delta \geq 1$ as $\epsilon \leq f(x+r)$. Then, by concavity of f,

$$f(x+r) = f\left(\frac{1}{\delta}(x+\delta r) + \frac{\delta-1}{\delta} x\right) \geq \frac{1}{\delta} f(x+\delta r) + \frac{\delta-1}{\delta} f(x)$$

and

$$f(x+\delta r) \leq \delta\left(f(x+r) - f(x)\right) + f(x) = \epsilon - 1.$$

As $x + \delta r \in \mathcal{X}$, this contradicts the definition of ϵ, and thus $f(x) \leq f(x + r)$. This shows that

$$\inf_{x \in \mathcal{X}} f(x) \quad = \quad \inf_{x \in \text{conv}\{x^{(j)}\}_{j=1}^{J}} f(x).$$

The latter is a concave minimization over a bounded polytope, and thus attains its optimal value ϵ at an extreme point of $\text{conv}\{x^{(j)}\}_{j=1}^{J}$ by Lemma A.1. Extreme points of the convex hull of a discrete set are elements of the set, and these elements are extreme points of $\mathcal{X}$. Thus, the optimal value ϵ is obtained at an extreme point of $\mathcal{X}$.

□

Proposition 2. *The model* (P) *is feasible and has an optimal solution at an extreme point of its feasible region.*

Proof. We first establish that (P) is feasible. Take $y = 0$. As $u^{\text{REC}} > 0$ and $u^{\text{CP}} > 0$ under Assumption 1, y is feasible for (3c)-(3f). By Proposition 1, there is some $x \geq 0$ so that (y, x) satisfies (1), (2), and (3). Then, (y, x) is feasible for (P).

Next, consider the relaxation of (P):

$$\begin{aligned} \min_{x,y} \quad & \sum_{t \in \mathcal{T}} (1-\gamma)^{t-1} \left(C_t^{\text{PL}}(y) + \sum_{\omega \in \Omega} p_\omega C_{\omega,t}^{\text{OP}}(x) \right) \\ \text{s.t.} \quad & \text{(1e), (2a), (2c)} \\ & x \geq 0,\ y \geq 0. \end{aligned} \tag{13}$$

Any feasible solution for (P) is feasible for (13) with the same objective value. Note that, as constraints (3) are relaxed, there are no constraints coupling variables x and y. Thus, (13) is equivalent to

$$\begin{aligned} \min_{y \geq 0} \quad & \sum_{t \in \mathcal{T}} (1-\gamma)^{t-1} C_t^{\text{PL}}(y) \quad + \quad \min_{x \geq 0} \quad \sum_{t \in \mathcal{T}} (1-\gamma)^{t-1} \sum_{\omega \in \Omega} p_\omega C_{\omega,t}^{\text{OP}}(x) \\ & \qquad\qquad\qquad\qquad\qquad\qquad \text{s.t.} \quad \text{(1e), (2a), (2c).} \end{aligned} \tag{14}$$

As C_t^{PL} is monotonic increasing in each variable y by Assumption 1, $y = 0$ is optimal for the first problem of (14).

Consider any direction $x \geq 0$ in which the feasible region of the second problem of (14) is unbounded. By (2c),

$$x_{\omega,z,t,i}^{\text{RB}} - x_{\omega,z,t-1,i}^{\text{RB}} + \sum_{j \in \mathcal{J}} x_{\omega,z,t,i,j}^{\text{RB,RM}} - \sum_{z' \in \mathcal{Z} \setminus \{z\}} \left(x_{\omega,z',z,t,i}^{\text{TR,RB}} - x_{\omega,z,z',t,i}^{\text{TR,RB}} \right) = 0.$$

Summing over z yields $\sum_{z \in \mathcal{Z}} \left(x_{\omega,z,t,i}^{\text{RB}} - x_{\omega,z,t-1,i}^{\text{RB}} + \sum_{j \in \mathcal{J}} x_{\omega,z,t,i,j}^{\text{RB,RM}} \right) = 0$. As $x_{\omega,z,0,i}^{\text{RB}} = 0$ and $x \geq 0$, by induction $x^{\text{RB}} = 0$ and $x^{\text{RB,RM}} = 0$. Then, (1e) gives $x_{\omega,z,t,k}^{\text{RM,INV}} + x_{\omega,z,t,k}^{\text{RM,S}} - \sum_{i \in \mathcal{I},\ j \in \mathcal{J}} \Delta_{k,i,j}^{\text{REC}} x_{\omega,z,t,i,j}^{\text{RB,RM}} = 0$, so $x^{\text{RM,INV}} = 0$ and $x^{\text{RM,S}} = 0$ as $x \geq 0$.

Then, consider the change in objective in direction x, given by $\sum_{t \in \mathcal{T}} (1-\gamma)^{t-1} \sum_{\omega \in \Omega} p_\omega C_{\omega,t}^{\text{OP}}(x)$. As $\gamma < 1$, $c \geq 0$, $v \geq 0$, $x \geq 0$, and $x^{\text{RM,S}} = 0$, it holds that $\sum_{t \in \mathcal{T}} (1-\gamma)^{t-1} \sum_{\omega \in \Omega} p_\omega C_{\omega,t}^{\text{OP}}(x) \geq 0$, and the direction does not improve

the objective. As the second problem of (14) is a linear program, this proves that its objective is bounded. Thus, neither problem in (14) is unbounded, so (13) and (P) are bounded in objective. As (P) is a concave minimization over a polyhedron contained in the nonnegative orthant and is bounded in objective, it attains its optimal value at an extreme point of its feasible region by Lemma A.2.

□

Theorem 1. *Let $(\tilde{y}, \tilde{x})$ be an extreme point of $\{(y,x) \geq 0 \ : \ (1),\ (2),\ (3)\}$. Then, $\tilde{y}^{REC}$ satisfies (6) and $\tilde{y}^{CP}$ satisfies the corresponding property.*

Proof. The comparable solution structure for $\tilde{y}^{\mathrm{CP}}$ with total capacity $Y_{z,l,k} = \sum_{n \in \mathcal{N}^{\mathrm{CP}}_{l,k}} \tilde{y}^{\mathrm{CP}}_{z,l,k,n}$ is given by

$$\begin{aligned} \mathcal{C}(\{\tilde{y}^{\mathrm{CP}}_{z,l,k,n}\}_{n \in \mathcal{N}^{\mathrm{CP}}_{l,k}}; u^{\mathrm{CP}}) &\geq \left\lceil \frac{Y_{z,l,k}}{u^{\mathrm{CP}}} \right\rceil - 1; \\ \mathcal{C}(\{\tilde{y}^{\mathrm{CP}}_{z,l,k,n}\}_{n \in \mathcal{N}^{\mathrm{CP}}_{l,k}}; 0) &\geq |\mathcal{N}^{\mathrm{CP}}_{l,k}| - \left\lceil \frac{Y_{z,l,k}}{u^{\mathrm{CP}}} \right\rceil \quad \forall k \in \mathcal{K}^{\mathrm{CP}},\ l \in \mathcal{L},\ z \in \mathcal{Z}. \end{aligned} \tag{15}$$

We demonstrate the result only for the property of the recycling capacity solution $\tilde{y}^{\mathrm{REC}}$, as the same arguments apply to the cathode production capacity solution $\tilde{y}^{\mathrm{CP}}$. The result is shown by contraposition. Denote the polyhedron $\mathcal{X} = \{(y,x) \geq 0 \ : \ (1),\ (2),\ (3)\}$ and consider any $(y,x) \in \mathcal{X}$. Fix some z, l, j and let the total recycling capacity be $Y_{z,l,j} = \sum_{n \in \mathcal{N}^{\mathrm{REC}}_l} y^{\mathrm{REC}}_{z,l,j,n}$. Suppose that (6) does not hold for y^{REC}.

First, consider the case where $\mathcal{C}(\{y^{\mathrm{REC}}_{z,l,j,n}\}_{n \in \mathcal{N}^{\mathrm{REC}}_l}; u^{\mathrm{REC}}) \leq \left\lceil \frac{Y_{z,l,j}}{u^{\mathrm{REC}}} \right\rceil - 2$. Suppose for contradiction that there are one or fewer indices n with $u^{\mathrm{REC}} > y^{\mathrm{REC}}_{z,l,j,n} > 0$. Then, the total number of indices corresponding to nonzero elements is bounded by $\left\lceil \frac{Y_{z,l,j}}{u^{\mathrm{REC}}} \right\rceil - 1$, and

$$\sum_{n \in \mathcal{N}^{\mathrm{REC}}_l} y^{\mathrm{REC}}_{z,l,j,n} \leq u^{\mathrm{REC}} \left(\left\lceil \frac{Y_{z,l,j}}{u^{\mathrm{REC}}} \right\rceil - 1 \right) < Y_{z,l,j} = \sum_{n \in \mathcal{N}^{\mathrm{REC}}_l} y^{\mathrm{REC}}_{z,l,j,n}.$$

By this contradiction, there must be at least two indices n_1, n_2 with $u^{\mathrm{REC}} > y^{\mathrm{REC}}_{z,l,j,n} > 0$ for $n \in \{n_1, n_2\}$.

Second, consider the case where $\mathcal{C}(\{y^{\mathrm{REC}}_{z,l,j,n}\}_{n \in \mathcal{N}^{\mathrm{REC}}_l}; 0) \leq |\mathcal{N}^{\mathrm{REC}}_l| - \left\lceil \frac{Y_{z,l,j}}{u^{\mathrm{REC}}} \right\rceil - 1$. Again, suppose for contradiction that there are one or fewer indices n with $u^{\mathrm{REC}} > y^{\mathrm{REC}}_{z,l,j,n} > 0$. Then, there are at least $\left\lceil \frac{Y_{z,l,j}}{u^{\mathrm{REC}}} \right\rceil$ elements at the upper bound u^{REC} and one additional nonzero element (that may also take the value u^{REC}), so

$$\sum_{n \in \mathcal{N}^{\mathrm{REC}}_l} y^{\mathrm{REC}}_{z,l,j,n} > u^{\mathrm{REC}} \left\lceil \frac{Y_{z,l,j}}{u^{\mathrm{REC}}} \right\rceil \geq Y_{z,l,j} = \sum_{n \in \mathcal{N}^{\mathrm{REC}}_l} y^{\mathrm{REC}}_{z,l,j,n}.$$

By this contradiction, there must again be at least two indices n_1, n_2 with $u^{\mathrm{REC}} > y^{\mathrm{REC}}_{z,l,j,n} > 0$ for $n \in \{n_1, n_2\}$.

In either scenario, select such indices n_1, n_2, and denote

$$\delta = \min\{y^{\mathrm{REC}}_{z,l,j,n_1}, y^{\mathrm{REC}}_{z,l,j,n_2}, u^{\mathrm{REC}} - y^{\mathrm{REC}}_{z,l,j,n_1}, u^{\mathrm{REC}} - y^{\mathrm{REC}}_{z,l,j,n_2}\}.$$

By construction, $\delta > 0$. We construct perturbed solutions

$$\overline{y}^{\text{REC}}_{z,l,j,n}(1) = \begin{cases} y^{\text{REC}}_{z,l,j,n} & n \notin \{n_1, n_2\} \\ y^{\text{REC}}_{z,l,j,n} + \delta & n = n_1 \\ y^{\text{REC}}_{z,l,j,n} - \delta & n = n_2 \end{cases}$$

and

$$\overline{y}^{\text{REC}}_{z,l,j,n}(2) = \begin{cases} y^{\text{REC}}_{z,l,j,n} & n \notin \{n_1, n_2\} \\ y^{\text{REC}}_{z,l,j,n} - \delta & n = n_1 \\ y^{\text{REC}}_{z,l,j,n} + \delta & n = n_2. \end{cases}$$

Let $\overline{y}^{\text{CP}}(1) = \overline{y}^{\text{CP}}(2) = y^{\text{CP}}$. By construction, $0 \leq \overline{y}^{\text{REC}}_{z,l,j,n}(1) \leq u^{\text{REC}}$ and $\sum_{n \in \mathcal{N}^{\text{REC}}_l} \overline{y}^{\text{REC}}_{z,l,j,n}(1) = \sum_{n \in \mathcal{N}^{\text{REC}}_l} y^{\text{REC}}_{z,l,j,n}$, so $(\overline{y}(1), x)$ satisfies all constraints of $\mathcal{X}$. Similarly, $(\overline{y}(2), x) \in \mathcal{X}$. Finally, we note that $(y, x) = \frac{1}{2}(\overline{y}(1), x) + \frac{1}{2}(\overline{y}(2), x)$, so (y, x) is a convex combination of other points in $\mathcal{X}$, and thus is not an extreme point.

This shows that, if a solution y^{REC} does not satisfy (6), then the solution (y, x) is not an extreme point of $\mathcal{X}$. Thus, the contrapositive of this result is also true, namely, an extreme point $(\tilde{y}, \tilde{x})$ of $\mathcal{X}$ must satisfy (6). By similar logic, it can be shown that an extreme point must also satisfy (15). □

Theorem 2. *The models* (P) *and* (PMI) *have the same optimal objective value.*

Proof. We first demonstrate that a feasible solution for (PMI) can be mapped to an equivalent feasible solution for (P). Let $(\overline{y}, \overline{x})$ be feasible for (PMI). We define $\tilde{y}$ by

$$\tilde{y}^{\text{REC}}_{z,l,j,n} = \begin{cases} u^{\text{REC}} & n \leq \overline{y}^{\text{REC}}_{z,l,j} \\ \overline{y}^{\text{REC}}_{z,l,j,+} & n = \overline{y}^{\text{REC}}_{z,l,j} + 1 \\ 0 & \text{otherwise,} \end{cases}$$

$$\tilde{y}^{\text{CP}}_{z,l,k,n} = \begin{cases} u^{\text{CP}} & n \leq \overline{y}^{\text{CP}}_{z,l,k} \\ \overline{y}^{\text{CP}}_{z,l,k,+} & n = \overline{y}^{\text{CP}}_{z,l,k} + 1 \\ 0 & \text{otherwise.} \end{cases}$$

By construction, $0 \leq \tilde{y}^{\text{REC}}_{z,l,j,n} \leq u^{\text{REC}}$ and $0 \leq \tilde{y}^{\text{CP}}_{z,l,k,n} \leq u^{\text{CP}}$. Further, $\sum_{n \in \mathcal{N}^{\text{REC}}_l} \tilde{y}^{\text{REC}}_{z,l,j,n} = u^{\text{REC}} \overline{y}^{\text{REC}}_{z,l,j} + \overline{y}^{\text{REC}}_{z,l,j,+}$ and similarly $\sum_{n \in \mathcal{N}^{\text{CP}}_{l,k}} \tilde{y}^{\text{CP}}_{z,l,k,n} = u^{\text{CP}} \overline{y}^{\text{CP}}_{z,l,k} + \overline{y}^{\text{CP}}_{z,l,k,+}$. Due to feasibility of $(\overline{y}, \overline{x})$ for (7), the solution $(\tilde{y}, \overline{x})$ is feasible for

constraints (3). Further, $\overline{x}$ remains feasible for (1) and (2). Thus, $(\tilde{y}, \overline{x})$ is feasible for (P). The objective values satisfy

$$\begin{aligned} C_t^{\mathrm{PL}}(\tilde{y}) &= \sum_{z\in\mathcal{Z}} \left(\sum_{j\in\mathcal{J}} \sum_{n\in\mathcal{N}_{l_t}^{\mathrm{REC}}} f_{z,j}^{\mathrm{REC}}(\tilde{y}_{z,l_t,j,n}^{\mathrm{REC}}) + \sum_{k\in\mathcal{K}^{\mathrm{CP}}} \sum_{n\in\mathcal{N}_{l_t,k}^{\mathrm{CP}}} f_{z,k}^{\mathrm{CP}}(\tilde{y}_{z,l_t,k,n}^{\mathrm{CP}}) \right) \\ &= \sum_{z\in\mathcal{Z}} \left(\sum_{j\in\mathcal{J}} \left(\overline{y}_{z,l_t,j}^{\mathrm{REC}} f_{z,j}^{\mathrm{REC}}(u^{\mathrm{REC}}) + f_{z,j}^{\mathrm{REC}}(\overline{y}_{z,l_t,j,+}^{\mathrm{REC}}) \right) + \sum_{k\in\mathcal{K}^{\mathrm{CP}}} \left(\overline{y}_{z,l_t,k}^{\mathrm{CP}} f_{z,k}^{\mathrm{CP}}(u^{\mathrm{CP}}) + f_{z,k}^{\mathrm{CP}}(\overline{y}_{z,l_t,k,+}^{\mathrm{CP}}) \right) \right) \\ &= \overline{C}_t^{\mathrm{PL}}(\overline{y}), \end{aligned}$$

as $f(0) = 0$. Thus $(\tilde{y}, \overline{x})$ has the same objective value in (P) as $(\overline{y}, \overline{x})$ in (PMI).

In the opposite direction, we consider an optimal solution for (P) that occurs at an extreme point of its feasible region, which exists by Proposition 2. Denote this optimal extreme point $(\tilde{y}, \tilde{x})$. By Theorem 1, this solution satisfies (6) and (15).

As $u^{\mathrm{REC}} > 0$ under Assumption 1, each index $n \in \mathcal{N}_l^{\mathrm{REC}}$ can be counted by at most one of $\mathcal{C}(\{\tilde{y}_{z,l,j,n}^{\mathrm{REC}}\}_{n\in\mathcal{N}_l^{\mathrm{REC}}}; 0)$ and $\mathcal{C}(\{\tilde{y}_{z,l,j,n}^{\mathrm{REC}}\}_{n\in\mathcal{N}_l^{\mathrm{REC}}}; u^{\mathrm{REC}})$. By this property and the structure (6),

$$|\mathcal{N}_l^{\mathrm{REC}}| \geq \mathcal{C}(\{\tilde{y}_{z,l,j,n}^{\mathrm{REC}}\}_{n\in\mathcal{N}_l^{\mathrm{REC}}}; 0) + \mathcal{C}(\{\tilde{y}_{z,l,j,n}^{\mathrm{REC}}\}_{n\in\mathcal{N}_l^{\mathrm{REC}}}; u^{\mathrm{REC}}) \geq |\mathcal{N}^{\mathrm{REC}}| - 1.$$

If the lower bound holds with equality, we can select the unique $\overline{n}_{z,l,j}$ so that $0 < \tilde{y}_{z,l,j,(\overline{n}_{z,l,j})}^{\mathrm{REC}} < u^{\mathrm{REC}}$. Otherwise, we select arbitrary $\overline{n}_{z,l,j}$. The index has the property that $\tilde{y}_{z,l,j,n}^{\mathrm{REC}} \in \{0, u^{\mathrm{REC}}\}$ for any $n \neq \overline{n}_{z,l,j}$. We can similarly select $\overline{n}_{z,l,k}$ for $\tilde{y}^{\mathrm{CP}}$. Now, we define

$$\begin{aligned} &\overline{y}_{z,l,j}^{\mathrm{REC}} = \mathcal{C}(\{\tilde{y}_{z,l,j,n}^{\mathrm{REC}}\}_{n\in\mathcal{N}_l^{\mathrm{REC}}\setminus\{\overline{n}_{z,l,j}\}}; u^{\mathrm{REC}}); \\ &\overline{y}_{z,l,j,+}^{\mathrm{REC}} = \tilde{y}_{z,l,j,\overline{n}_{z,l,j}}^{\mathrm{REC}}; \\ &\overline{y}_{z,l,k}^{\mathrm{CP}} = \mathcal{C}(\{\tilde{y}_{z,l,k,n}^{\mathrm{CP}}\}_{n\in\mathcal{N}_{l,k}^{\mathrm{CP}}\setminus\{\overline{n}_{z,l,k}\}}; u^{\mathrm{REC}}); \text{ and} \\ &\overline{y}_{z,l,k,+}^{\mathrm{CP}} = \tilde{y}_{z,l,k,\overline{n}_{z,l,k}}^{\mathrm{CP}}. \end{aligned}$$

By construction, $\overline{y}_{z,l,j}^{\mathrm{REC}}$ and $\overline{y}_{z,l,k}^{\mathrm{CP}}$ satisfy the integrality constraints (7e) and (7f), and further $\overline{y}_{z,l,j,+}^{\mathrm{REC}}$ and $\overline{y}_{z,l,k,+}^{\mathrm{CP}}$ satisfy nonnegativity and the upper bounds (7g) and (7h) by the feasibility of $\tilde{y}$ for (P). By definition of $\overline{n}$,

$$\sum_{n\in\mathcal{N}_l^{\mathrm{REC}}} \tilde{y}_{z,l,j,n}^{\mathrm{REC}} = u^{\mathrm{REC}} \mathcal{C}(\{\tilde{y}_{z,l,j,n}^{\mathrm{REC}}\}_{n\in\mathcal{N}_l^{\mathrm{REC}}\setminus\{\overline{n}_{z,l,j}\}}; u^{\mathrm{REC}}) + \tilde{y}_{z,l,j,(\overline{n}_{z,l,j})}^{\mathrm{REC}} = u^{\mathrm{REC}} \overline{y}_{z,l,j}^{\mathrm{REC}} + \overline{y}_{z,l,j,+}^{\mathrm{REC}},$$

and similarly

$$\sum_{n\in\mathcal{N}_{l,k}^{\mathrm{CP}}} \tilde{y}_{z,l,k,n}^{\mathrm{CP}} = u^{\mathrm{CP}} \overline{y}_{z,l,k}^{\mathrm{CP}} + \overline{y}_{z,l,k,+}^{\mathrm{CP}}.$$

Thus, feasibility of $(\tilde{y}, \tilde{x})$ for (3) implies that $(\overline{y}, \tilde{x})$ is feasible for (7a)-(7d), and $\tilde{x}$ remains feasible for (1) and (2).

Thus, $(\overline{y}, \tilde{x})$ is feasible for (PMI). Considering the objective,

$$\sum_{n \in \mathcal{N}_l^{\mathrm{REC}}} f_{z,j}^{\mathrm{REC}}(\tilde{y}_{z,l,j,n}^{\mathrm{REC}}) = \mathcal{C}(\{\tilde{y}_{z,l,j,n}^{\mathrm{REC}}\}_{n \in \mathcal{N}_l^{\mathrm{REC}} \setminus \{\overline{n}_{z,l,j}\}}; u^{\mathrm{REC}}) f_{z,j}^{\mathrm{REC}}(u^{\mathrm{REC}}) + f_{z,j}^{\mathrm{REC}}(\tilde{y}_{z,l,j,(\overline{n}_{z,l,j})}^{\mathrm{REC}})$$

$$= \overline{y}_{z,l,j}^{\mathrm{REC}} f_{z,j}^{\mathrm{REC}}(u^{\mathrm{REC}}) + f_{z,j}^{\mathrm{REC}}(\overline{y}_{z,l,j,+}^{\mathrm{REC}}),$$

as $f_{z,j}^{\mathrm{REC}}(0) = 0$ by Assumption 1. Again, a similar result holds for $\tilde{y}^{\mathrm{CP}}$. Therefore, the objectives satisfy

$$\begin{aligned} C_t^{\mathrm{PL}}(\tilde{y}) &= \sum_{z \in \mathcal{Z}} \left(\sum_{j \in \mathcal{J}} \sum_{n \in \mathcal{N}_{l_t}^{\mathrm{REC}}} f_{z,j}^{\mathrm{REC}}(\tilde{y}_{z,l_t,j,n}^{\mathrm{REC}}) + \sum_{k \in \mathcal{K}^{\mathrm{CP}}} \sum_{n \in \mathcal{N}_{l_t,k}^{\mathrm{CP}}} f_{z,k}^{\mathrm{CP}}(\tilde{y}_{z,l_t,k,n}^{\mathrm{CP}}) \right) \\ &= \sum_{z \in \mathcal{Z}} \left(\sum_{j \in \mathcal{J}} \overline{y}_{z,l_t,j}^{\mathrm{REC}} f_{z,j}^{\mathrm{REC}}(u^{\mathrm{REC}}) + f_{z,j}^{\mathrm{REC}}(\overline{y}_{z,l_t,j,+}^{\mathrm{REC}}) + \sum_{k \in \mathcal{K}^{\mathrm{CP}}} \overline{y}_{z,l_t,k}^{\mathrm{CP}} f_{z,k}^{\mathrm{CP}}(u^{\mathrm{CP}}) + f_{z,k}^{\mathrm{CP}}(\overline{y}_{z,l_t,k,+}^{\mathrm{CP}}) \right) \\ &= \overline{C}_t^{\mathrm{PL}}(\overline{y}), \end{aligned}$$

and $(\overline{y}, \tilde{x})$ has the same objective value in (PMI) as $(\tilde{y}, \tilde{x})$ has in (P). We have already shown that any feasible solution for (PMI) has a corresponding feasible solution for (P) with the same objective value, so $(\overline{y}, \tilde{x})$ is optimal for (PMI), as otherwise there must also be a better solution than $(\tilde{y}, \tilde{x})$ for (P). Therefore, as $(\tilde{y}, \tilde{x})$ is optimal for (P) and $(\overline{y}, \tilde{x})$ is optimal for (PMI) with the same objective value, the problems have the same optimal objective.

□

Lemma A.3. *The functions $\overline{f}_i$ defined in (10) have the following properties:*

1. *$\overline{f}_i(y) \leq f_i(y)$ for $y \in [y_i^{(1)}, y_i^{(k_i)}]$;*
2. *$\overline{f}_i$ is concave and continuous; and*
3. *$\overline{f}_i(y_i^{(j)}) = f_i(y_i^{(j)})$ for all $j \in \{1, \ldots, k_i\}$.*

Proof. 1. Select j so that $y \in [y_i^{(j)}, y_i^{(j+1)})$. Then, taking $\lambda = \frac{y - y_i^{(j)}}{y_i^{(j+1)} - y_i^{(j)}} \in [0, 1)$,

$$\begin{aligned} \overline{f}_i(y) &= \frac{y - y_i^{(j)}}{y_i^{(j+1)} - y_i^{(j)}} f_i(y_i^{(j+1)}) + \frac{y_i^{(j+1)} - y}{y_i^{(j+1)} - y_i^{(j)}} f_i(y_i^{(j)}) = \lambda f_i(y_i^{(j+1)}) + (1 - \lambda) f_i(y_i^{(j)}) \\ &\leq f_i(\lambda y_i^{(j+1)} + (1 - \lambda) y_i^{(j)}) = f_i(y) \end{aligned}$$

by the concavity of f_i.

2. Denote $\tilde{f}_i(y) = \min_{j \in \{1, \ldots, k_i - 1\}} \left\{ \frac{f_i(y_i^{(j+1)}) - f_i(y_i^{(j)})}{y_i^{(j+1)} - y_i^{(j)}} (y - y_i^{(j)}) + f_i(y_i^{(j)}) \right\}$. Consider $y \in [y_i^{(1)}, y_i^{(k_i)}]$ and any index j. If $y \geq y_i^{(j+1)}$, let $\lambda = \frac{y_i^{(j+1)} - y_i^{(j)}}{y - y_i^{(j)}} \in (0, 1]$ so that $y_i^{(j+1)} = \lambda y + (1 - \lambda) y_i^{(j)}$. Then, by concavity of f_i,

$$f_i(y_i^{(j+1)}) \geq \lambda f_i(y) + (1 - \lambda) f_i(y_i^{(j)}),$$

and

$$\frac{f_i(y_i^{(j+1)}) - f_i(y_i^{(j)})}{y_i^{(j+1)} - y_i^{(j)}}(y - y_i^{(j)}) + f_i(y_i^{(j)}) \geq f_i(y) \geq \overline{f}_i(y).$$

Otherwise, if $y < y_i^{(j)}$, let $\lambda = \frac{y_i^{(j)} - y}{y_i^{(j+1)} - y} \in (0, 1]$ so that $y_i^{(j)} = \lambda y_i^{(j+1)} + (1 - \lambda)y$. By concavity of f_i,

$$f_i(y_i^{(j)}) \geq \lambda f_i(y_i^{(j+1)}) + (1 - \lambda) f_i(y),$$

and

$$\frac{f_i(y_i^{(j+1)}) - f_i(y_i^{(j)})}{y_i^{(j+1)} - y_i^{(j)}}(y - y_i^{(j)}) + f_i(y_i^{(j)}) \geq f_i(y) \geq \overline{f}_i(y).$$

Finally, if $y \in [y_i^{(j)}, y_i^{(j+1)})$,

$$\frac{f_i(y_i^{(j+1)}) - f_i(y_i^{(j)})}{y_i^{(j+1)} - y_i^{(j)}}(y - y_i^{(j)}) + f_i(y_i^{(j)}) = \overline{f}_i(y).$$

Thus, we have that $\overline{f}_i(y) \leq \tilde{f}_i(y)$. Further, $\overline{f}_i(y) \in \left\{ \frac{f_i(y_i^{(j+1)}) - f_i(y_i^{(j)})}{y_i^{(j+1)} - y_i^{(j)}}(y - y_i^{(j)}) + f_i(y_i^{(j)}) \right\}_{j=1}^{k_i - 1}$, so $\overline{f}_i(y) \geq \tilde{f}_i(y)$, and $\overline{f}_i(y) = \tilde{f}_i(y)$. Therefore, as the function $\overline{f}_i$ is a minimum of linear functions, it is continuous and concave.

3. The equality follows from direct evaluation.

□

Corollary A.1. *Let $\overline{f}_i^1$ be the piecewise linear function associated with breakpoints $\{y_i^{(j)}\}_{j=1}^{k_i}$ and $\overline{f}_i^2$ the function with breakpoints $\{y_i^{(j)}\}_{j=1}^{k_i} \setminus \{y_i^{(r)}\}$ for some $r \notin \{1, k_i\}$. Then, $\overline{f}_i^1(y) \geq \overline{f}_i^2(y)$ for $y \in [y_i^{(1)}, y_i^{(k_i)}]$.*

Proof. By Lemma A.3, $\overline{f}_i^1$ is concave. Let j_1 and j_2 be consecutive breakpoints for $\overline{f}_i^2$, so

$$j_2 = \begin{cases} j_1 + 1 & j_1 \neq r - 1 \\ j_1 + 2 & \text{otherwise.} \end{cases}$$

Then, by Lemma A.3,

$$\frac{f_i(y_i^{(j_2)}) - f_i(y_i^{(j_1)})}{y_i^{(j_2)} - y_i^{(j_1)}}(y_i - y_i^{(j_1)}) + f_i(y_i^{(j_1)}) = \frac{\overline{f}_i^1(y_i^{(j_2)}) - \overline{f}_i^1(y_i^{(j_1)})}{y_i^{(j_2)} - y_i^{(j_1)}}(y_i - y_i^{(j_1)}) + \overline{f}_i^1(y_i^{(j_1)})$$

for any $y_i \in [y_i^{(1)}, y_i^{(k_i)}]$. Thus, $\overline{f}_i^2$ can be equivalently constructed in the form (10) by using the concave function $\overline{f}_i^1$ in place of f_i. It then holds that $\overline{f}_i^2$ is a piecewise under-approximator of $\overline{f}_i^1$ and, by Lemma A.3, $\overline{f}_i^1(y) \geq \overline{f}_i^2(y)$ for $y \in [y_i^{(1)}, y_i^{(k_i)}]$.

□

Lemma 1. *Let $(y^*, x^*) \in \mathcal{X}$ be an optimal solution to* (SP)*. Then, there is an optimal solution that is an extreme point of $\mathcal{X}(y^*)$.*

Proof. Denote the optimal cost for (SP) by $z^{\text{SP}} = \sum_{i=1}^{n} \overline{f}_i(y_i^*) + \sum_{\omega\in\Omega} p_\omega c_\omega^T x_\omega^*$. Consider the restricted problem

$$z^{\text{RP}} = \min_{(y,x)\in\mathcal{X}(y^*)} \quad \sum_{i=1}^{n} \overline{f}_i(y_i) + \sum_{\omega\in\Omega} p_\omega c_\omega^T x_\omega. \tag{RSP}$$

As $\mathcal{X}(y^*) \subseteq \mathcal{X}$, (RSP) is a restriction of (SP), and $z^{\text{SP}} \leq z^{\text{RP}}$. However, (y^*, x^*) is feasible for (RSP), so $z^{\text{RP}} \leq z^{\text{SP}}$. Thus, the optimal values $z^{\text{RP}} = z^{\text{SP}}$, and an optimal solution for (RSP) is also optimal for (SP).

Further,

$$\mathcal{X}(y^*) = \{(y, \{x_\omega\}_{\omega\in\Omega}) \geq 0 \ : \ y_i = y_i^* \ \forall i \in [\![n_{\text{I}}]\!],\ Ay = b,\ B_\omega x_\omega + Dy = d_\omega \ \forall \omega \in \Omega\}$$

is a polytope under the assumption that $\mathcal{X}$ is bounded. Lemma A.3 gives that the functions $\overline{f}_i$ are concave, so the objective function of (RSP) is concave and (RSP) is a concave minimization over a polytope. Thus, (RSP) has an optimal solution at an extreme point of $\mathcal{X}(y^*)$ by Lemma A.1. This solution is also optimal for (SP), and therefore, there is an extreme point of $\mathcal{X}(y^*)$ that is optimal for (SP). □

Theorem 3. *Algorithm 1 terminates finitely with a global optimum of* (SCP).

Proof. Consider the set of feasible realizations of the integer components of y,

$$\mathcal{Y} = \{z \in \mathbb{Z}_+^{n_{\text{I}}} \ : \ \exists (y,x) \in \mathcal{X},\ y_i = z_i \ \forall i \in [\![n_{\text{I}}]\!]\} \subseteq \{z \in \mathbb{Z}_+^{n_{\text{I}}} \ : \ \underline{y}_i \leq z_i \leq \overline{y}_i \ \forall i \in [\![n_{\text{I}}]\!]\}.$$

Clearly, $\mathcal{Y}$ is finite. We construct the set of points that are extreme points of $\mathcal{X}(z)$ for some $z \in \mathcal{Y}$,

$$\mathcal{E} = \bigcup_{z\in\mathcal{Y}} \{(y,x) \ : \ (y,x) \text{ extreme point of } \mathcal{X}(z)\}.$$

This set is also finite, as there are finitely many extreme points of any polytope $\mathcal{X}(z)$ and finitely many such $z \in \mathcal{Y}$. We next construct the set of unique first-stage solutions with a corresponding extreme point:

$$\overline{\mathcal{E}} = \bigcup_{(y,x)\in\mathcal{E}} \{y\}.$$

Then, $|\overline{\mathcal{E}}| \leq |\mathcal{E}|$, so $\overline{\mathcal{E}}$ is finite.

For any y feasible for (SP) or (SCP) and any set of breakpoints $\{y_i^{(j)}\}_{j=1}^{k_i}$ generated by Algorithm 1,

$$y_i^{(1)} = \underline{y}_i \leq y_i \leq \overline{y}_i = y_i^{(k_i)}$$

by the initialization in step 1 of the algorithm, so $y_i \in [y_i^{(1)}, y_i^{(k_i)}]$. As a result, the first property of Lemma A.3 will apply for any feasible solution y to these problems. Further, $y_i^{(1)}$ and $y_i^{(k_i)}$ maintain the same values across iterations by step 5 of the algorithm.

Suppose that Algorithm 1 produces a sequence of iterates $(\tilde{y}^s, \{\tilde{x}_\omega^s\}_{\omega\in\Omega})$ of length at least $|\overline{\mathcal{E}}| + 1$. We note that

(SP) has a finite optimal solution by the boundedness assumption on $\mathcal{X}$ (Lemma A.1). Under Assumption 2, at every iteration s, the iterate $(\tilde{y}^s, \{\tilde{x}^s_\omega\}_{\omega\in\Omega})$ is an extreme point of $\mathcal{X}(\tilde{y}^s)$, so $\tilde{y}^s \in \overline{\mathcal{E}}$ for all s.

As the number of iterates is strictly larger than $|\overline{\mathcal{E}}|$, some first-stage solution must be revisited, so $\tilde{y}^{s_1} = \tilde{y}^{s_2}$ for some $s_1 < s_2 \leq |\overline{\mathcal{E}}| + 1$. Denote the set of breakpoints at the start of iteration s_2 by $\{y_i^{(j)}\}_{j=1}^{k_i}$. Note that the breakpoints which generate $\overline{f}_i^{s_1}$ are a subset of those which generate $\overline{f}_i^{s_2}$ and include $\{y_i^{(1)}, y_i^{(k_i)}\}$, so $\overline{f}_i^{s_2}(\tilde{y}_i^{s_1}) \geq \overline{f}_i^{s_1}(\tilde{y}_i^{s_1})$ as a consequence of Corollary A.1. From step 5 of the algorithm, either $\tilde{y}_i^{s_1} \in \{y_i^{(j)}\}_{j=1}^{k_i}$ or $\overline{f}_i^{s_1}(\tilde{y}_i^{s_1}) \geq f_i(\tilde{y}_i^{s_1})$ for each i. In the former case, $\overline{f}_i^{s_2}(\tilde{y}_i^{s_2}) = f_i(\tilde{y}_i^{s_2})$ by Lemma A.3. In the latter case,

$$\overline{f}_i^{s_2}(\tilde{y}_i^{s_2}) = \overline{f}_i^{s_2}(\tilde{y}_i^{s_1}) \geq \overline{f}_i^{s_1}(\tilde{y}_i^{s_1}) \geq f_i(\tilde{y}_i^{s_1}) = f_i(\tilde{y}_i^{s_2}).$$

Thus, the condition in step 4 of the algorithm is met at iteration s_2, and Algorithm 1 terminates finitely within $|\overline{\mathcal{E}}|+1$ iterations.

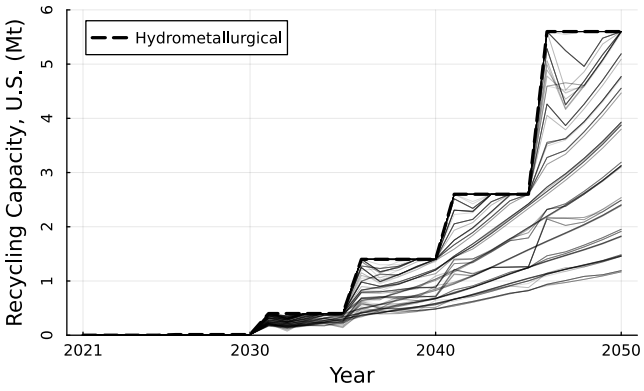
Let the algorithm terminate at iteration s with iterate $(\tilde{y}^s, \{\tilde{x}^s_\omega\}_{\omega\in\Omega})$. Then,

$$\begin{aligned}
\sum_{i=1}^n \overline{f}_i^s(\tilde{y}_i^s) + \sum_{\omega\in\Omega} p_\omega c_\omega^T \tilde{x}^s_\omega &= \min_{(y,\{x_\omega\}_{\omega\in\Omega})\in\mathcal{X}} \sum_{i=1}^n \overline{f}_i^s(y_i) + \sum_{\omega\in\Omega} p_\omega c_\omega^T x_\omega \\
&\leq \min_{(y,\{x_\omega\}_{\omega\in\Omega})\in\mathcal{X}} \sum_{i=1}^n f_i(y_i) + \sum_{\omega\in\Omega} p_\omega c_\omega^T x_\omega \\
&\leq \sum_{i=1}^n f_i(\tilde{y}_i^s) + \sum_{\omega\in\Omega} p_\omega c_\omega^T \tilde{x}^s_\omega \\
&\leq \sum_{i=1}^n \overline{f}_i^s(\tilde{y}_i^s) + \sum_{\omega\in\Omega} p_\omega c_\omega^T \tilde{x}^s_\omega,
\end{aligned}$$

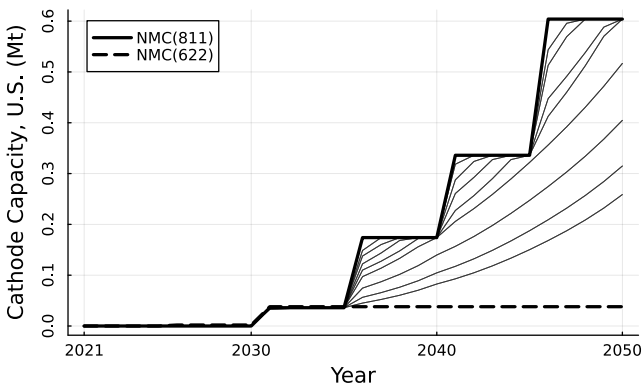
where the first inequality follows from Lemma A.3 and the third from the termination criterion in step 4 of the algorithm. Therefore,

$$\min_{(y,\{x_\omega\}_{\omega\in\Omega})\in\mathcal{X}} \sum_{i=1}^n f_i(y_i) + \sum_{\omega\in\Omega} p_\omega c_\omega^T x_\omega = \sum_{i=1}^n f_i(\tilde{y}_i^s) + \sum_{\omega\in\Omega} p_\omega c_\omega^T \tilde{x}^s_\omega,$$

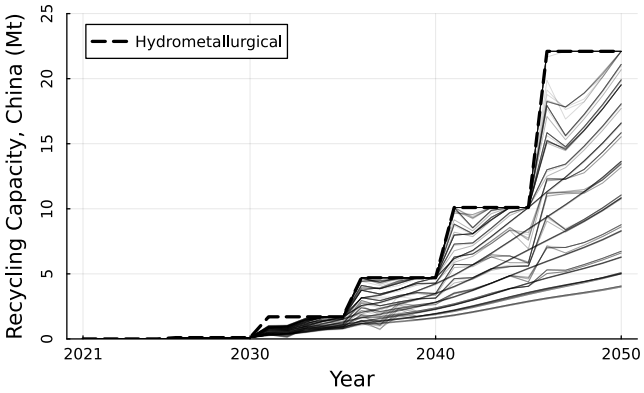
and $(\tilde{y}^s, \{\tilde{x}^s_\omega\}_{\omega\in\Omega})$ is a global optimum for (SCP). □



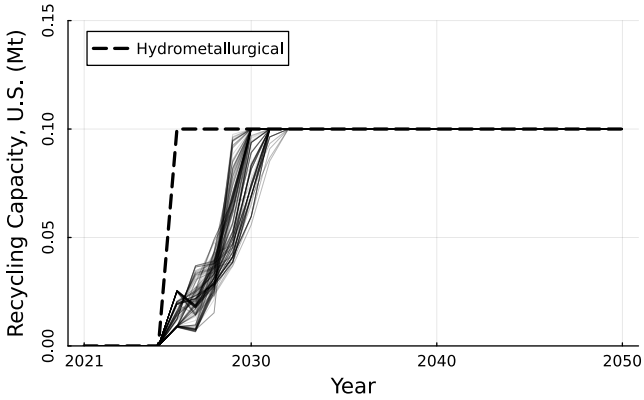
(a) Case 1, U.S. Recycling Capacity



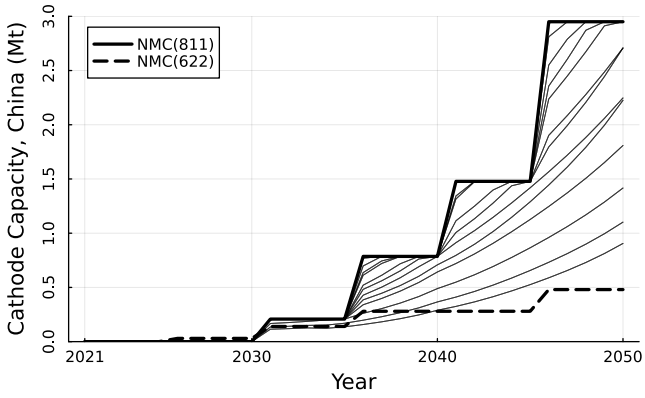
(b) Case 1, U.S. Cathode Production Capacity



(c) Case 2, China Recycling Capacity

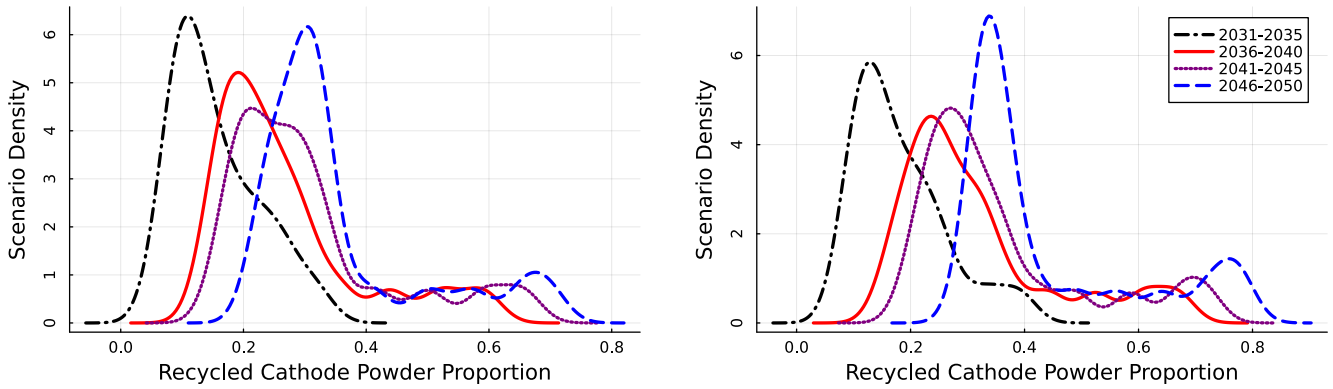


(d) Case 2, U.S. Recycling Capacity



(e) Case 2, China Cathode Production Capacity

Figure A.1: Optimal capacity investment decisions for recycling and cathode production facilities by zone in DR0 variations. Utilization of hydrometallurgical recycling capacity (a,c,d) and NMC(811) cathode production capacity (b,e) for each second-stage scenario is shown with thin lines.



(a) Case 1

(b) Case 2

Figure A.2: Density across scenarios of the proportion of cathode powder used in new battery manufacturing that is produced via recycling, indexed by planning period, in DR0 variations.