Technische Universität München

Fakultät für Mathematik

Lehrstuhl für Angewandte Geometrie und Diskrete Mathematik

A Polyhedral Analysis of Start-up Process Models in Unit Commitment Problems

Matthias Silbernagl

Vollständiger Abdruck der von der Fakultät für Mathematik der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitzender:	Prof. Dr. Michael Ulbrich
Prüfer der Dissertation:	1. Prof. Dr. Peter Gritzmann
	2. Prof. Dr. Alexander Martin
	3. Prof. Dr. Thomas Hamacher

Die Dissertation wurde am 11.07.2016 bei der Technischen Universität München eingereicht und durch die Fakultät für Mathematik am 30.10.2016 angenommen.

Abstract

This work investigates mixed-integer start-up models in Unit Commitment problems. We consider the costs and the power trajectories caused by a start-up.

After a polyhedral analysis of the start-up cost function, we present two families of models based on different approaches: The first type of formulations explicitly captures the temperature of a power plant and models the typical inversely exponential start-up costs. The second type of formulations classifies each start-up by binary variables, allowing arbitrary costs and power trajectories. The practical relevance of both models is experimentally demonstrated using data of the German power system.

Zusammenfassung

Diese Arbeit untersucht die gemischt-ganzzahlige Modellierung von Startvorgängen in der Kraftwerkseinsatzplanung. Betrachtet werden dabei die durch einen Start verursachten Kosten und Verläufe der Stromproduktion.

Nach einer polyedrischen Analyse der Startkostenfunktion stellen wir zwei Familien von Modellen basierend auf unterschiedlichen Ansätzen vor: Die erste Art von Formulierungen modelliert den Temperaturverlauf eines Kraftwerks explizit und bildet damit die typischen invers exponentiellen Startkosten ab. Die zweite Art von Formulierungen typisiert Startvorgänge durch Binärvariablen und erlaubt dadurch beliebige Kosten und Produktionsverläufe. Die praktische Relevanz beider Modelle wird durch Experimente anhand von Daten des deutschen Kraftwerkparks nachgewiesen.

Contents

Chapter 1

Introduction

The end of the 19th century saw the birth of the first local power systems. Over mere 50 years, these local systems evolved into networks spanning whole geographic regions. They became and remain indispensable to modern society [Hug93].

To keep a power system in balance, its power plants are coordinated to constantly meet the electricity demand, at substantial costs. Due to inherent strengths and weaknesses of different electricity production technologies, highly heterogeneous power plant fleets have been developed. The arising complexity provides ample opportunity for optimization and leads to the natural question "how should a power system be operated?" Part of this question—the day-to-day scheduling of electricity generating units—is commonly known as the family of *Unit Commitment (UC)* problems.

Despite more than a century of research, this problem family is far from being solved. Moreover, the structure of the problem has recently been shifted by the introduction of renewable energy sources in significant quantities. The production of the prevalent renewable energy sources, wind and solar energy, is inherently volatile, aggravating the task of keeping electricity production and demand in balance. As a consequence, conventional units need to change production level and start-up more often, resulting in a higher share of start-up costs among the total costs. Hence, accurate models of start-ups, and in particular of start-up costs, become a priority.

The main results presented in this work are two novel families of start-up cost models,

- the temperature formulations represented by P^{temp} (joint work with René Brandenberg and Matthias Huber [SHB16; BHS]), and
- the start-up type formulations culminating in P^{δ} (based on the start-up types introduced in [Muc66]).

As demonstrated by numerical experiments on real-world models of the German Power system (see Chapter 5), these new formulations significantly improve both solution time and integrality gap compared to the state-of-the-art approaches in [CA06] (denoted by P^{t}_{ex}) and in [SBB10] (denoted by P^{δ}_{ex}).

The new formulations are capable of solving problems about twice as large as P^{δ}_{ex} in the same time to the same optimality tolerance (see Fig. 1.1). The gap to P^{t}_{ex} is even more pronounced, which was already shown in [MELR13b]. The computational

performance of all formulations except P^{temp} depends strongly on the approximation tolerance of the start-up costs - in the case of Fig. 1.1 a tolerance of $\text{CU}_{\text{tol}} = 10\%$ is used. In contrast, the temperature formulation P^{temp} by design models the typically assumed exponential start-up cost function (1.3.1) without approximation.

The superior computational performance of the new formulations can be explained by their much lower integrality gap. In our experiments, the median integrality gap is decreased down to one third of that of P^{δ}_{ex} (see Fig. 1.2). Note that if unapproximated start-up costs are modeled (i. e. $\text{CU}_{\text{tol}} = 0\%$), then the integrality gaps of P^{δ} and P^{temp} are equal.

These improvements stem from our thorough polyhedral analysis of the epigraphs of the start-up cost functions and the start-up types. Apart from the computational progress, our results analytically compare the tightness of all formulations regarding the modeled start-up costs (see Fig. 1.3). In particular, we prove the conjecture of [MELR13b] that P^{δ}_{ex} models a stronger lower bound on the start-up costs than P^{t}_{ex}.

Our contribution is organized in four chapters which, after introducing the necessary background in the following sections, is laid out in detail in Section 1.4. We give a short overview here:

Chapter 2 starts by analyzing the start-up cost epigraphs, which leads to

- a linear, irredundant $\mathcal{H}$-representation of the epigraph of the start-up cost function in a single period with a linear separation algorithm in Section 2.1 (joint work with René Brandenberg and Matthias Huber [BS14; SHB16]),
- an exponential $\mathcal{H}$-representation of the epigraph of the start-up cost function summed over all periods, which is irredundant in the general case and possesses

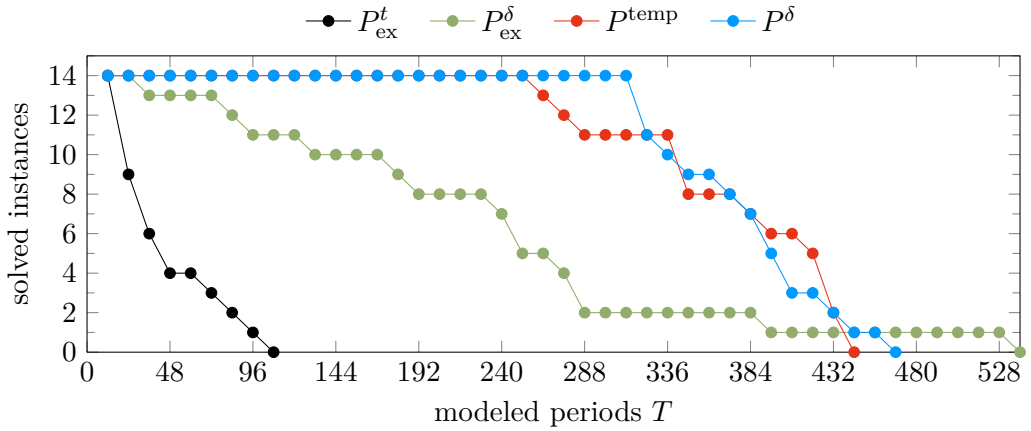


Figure 1.1: Number of solved instances in 15 minutes to an optimality tolerance of 1%, showing that the state-of-the-art models P^{t}_{ex} and P^{δ}_{ex} are dominated by the newly introduced P^{δ} and P^{temp} in most cases. P^{t}_{ex}, P^{δ}_{ex}, P^{δ} approximate the start-up costs with a tolerance of 10%.

a linear separation algorithm in Section 2.2 (joint work with René Brandenberg and Matthias Huber [BHS16]), and
- an exponential class of facets of the epigraph of the start-up cost function combined for all periods in Section 2.3.

By explicitly modeling the temperature of a unit, **Chapter 3** provides the extended formulation P^{temp} of the epigraph of the summed start-up costs with $\mathcal{O}(T)$ variables and $\mathcal{O}(T^2)$ inequalities. Only $\mathcal{O}(T)$ of these inequalities are necessary to model the start-up costs correctly, while the remaining $\mathcal{O}(T^2)$ inequalities tighten the linear relaxation and can be separated in $\mathcal{O}(T)$ (joint work with René Brandenberg and Matthias Huber [SHB16; BHS]). As shown in [HS15], the temperature formulations can be extended to incorporate the start-up process (see Section 6.2).

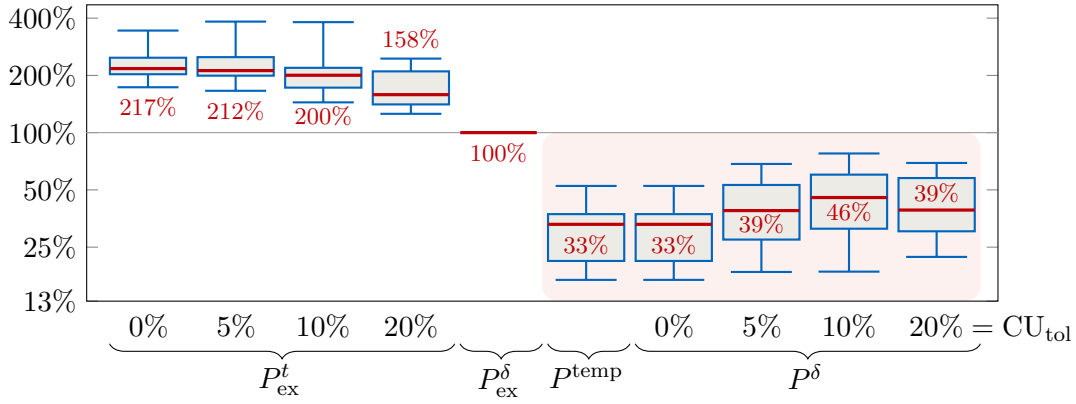


Figure 1.2: Integrality gap relative to the state-of-the-art P^{δ}_{ex} with start-up costs approximated to different tolerances CU_{tol} for P^{t}_{ex} and P^{δ}. The newly introduced models P^{temp} and P^{δ} (red) outclass the existing formulations P^{t}_{ex} and P^{δ}_{ex}.

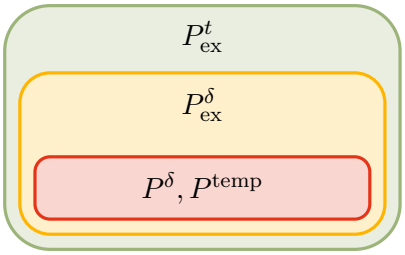


Figure 1.3: Venn diagram comparing the tightness of the formulations regarding the sum of the start-up costs. P^{δ} and P^{temp} model the same bound on the total start-up costs, which is higher than the bound of P^{δ}_{ex} and P^{t}_{ex}. All subset relationships are strict in general.

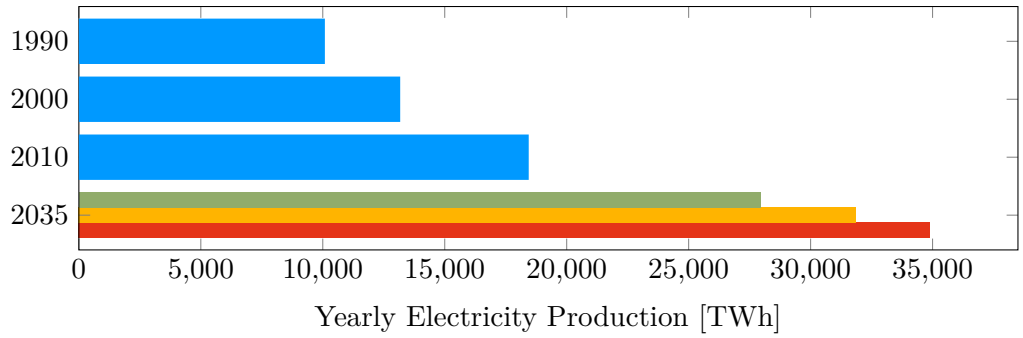


Figure 1.4: Historical and projected electricity production assuming different government policies (blue: historical, red: current policies, yellow: expected future policies, green: policies limiting global temperature increase to 2°C) [Int12] [1]

State-of-the-art start-up process models classify each start-up as one of S types (cf. [SBB10]). By interpreting these types as flows in a particular network, **Chapter 4** significantly tightens and generalizes these models such that the resulting polyhedron P^{δ} with $\mathcal{O}(ST)$ variables and $\mathcal{O}(T^2)$ inequalities is an extended formulation of the epigraph of the start-up costs in all periods. The start-up costs are modeled correctly by $\mathcal{O}(ST)$ of these inequalities, and the remaining inequalities tighten the linear relaxation and can be separated in $\mathcal{O}(T^2)$. The start-up process models from [SBB10] still apply to these start-up types.

Chapter 5 compares our models experimentally to the state-of-the-art using data of the German Power system. At the hands of scenarios of the years 2011 and of 2025, it substantiates the earlier theoretical results: The new extended formulations P^{temp}, P^{δ} outperform the existing formulations both in terms of solution time and integrality gap.

1.1 Technical and Economical Background

From 1990 to 2010, the world-wide electricity production has increased by 80%. Depending on government policies, the production is expected to grow by another 50% to 90% until 2035, mainly driven by the development of newly industrialized countries and the electrification of transport (see Fig. 1.4). As of 2010, electricity supplies 24% of the final energy demand, with increasing tendency. So, efficient power plant operation is highly relevant from both an economical and environmental point of view [Int12].

[1] Based on IEA data from "Renewable Energy Medium-Term Market Report 2013" © OECD/IEA 2013, IEA Publishing, and from "World Energy Outlook 2012" © OECD/IEA 2012, IEA Publishing; modified by Matthias Silbernagl. License: `http://www.iea.org/t&c/termsandconditions/`

The last decade has seen an unprecedented increase in electricity production from renewable energy sources. Globally, the production has increased from 3.5 TWh in 2006 to 4.9 TWh in 2012, and is projected to reach 6.9 TWh in 2018. Approximately half of the growth can be attributed to volatile energy sources, i. e. mainly wind and solar energy (see Fig. 1.5) [Int13].

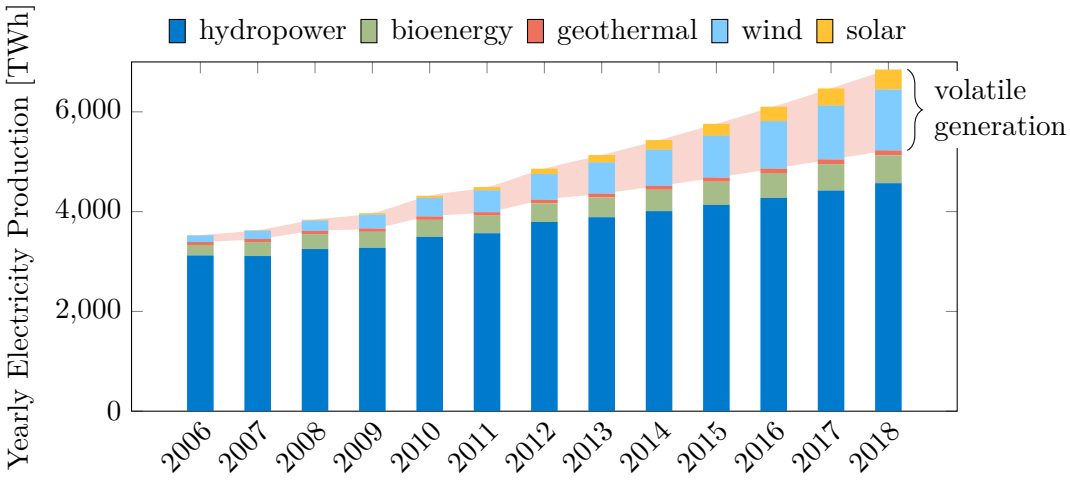


Figure 1.5: World renewable electricity production [Int13] [1]

When electricity from volatile energy sources is available, it is given preference due to policy (e. g. in the European Union) and due to low production costs. Consequentially, the remaining electricity producers need to keep the power network stable by providing the residual demand, i. e. the difference between demand and renewable production. As observable in the example of Germany (see Fig. 1.6), the relative volatility of the residual demand increases with increasing production from wind and solar. Therefore, the minimal residual demand, which is called the *base load*, decreases drastically.

Due to these changes in base load, many thermal units, i. e. nuclear, lignite, coal, gas, and biomass units, which have been designed for essentially uninterrupted operation are forced to shut down and start up regularly [KSH13]. Depending on the time that such a thermal unit has been offline before a start-up, the *downtime*, the unit executes a well-defined sequence to start up, following a specific power trajectory and incurring substantial costs, mainly due to the needed fuel and wear&tear caused by thermal stress [SBB10]. Given its significant impact on the operation of units and the resulting costs, it is essential to model the time-dependent start-up process and costs accurately and efficiently.

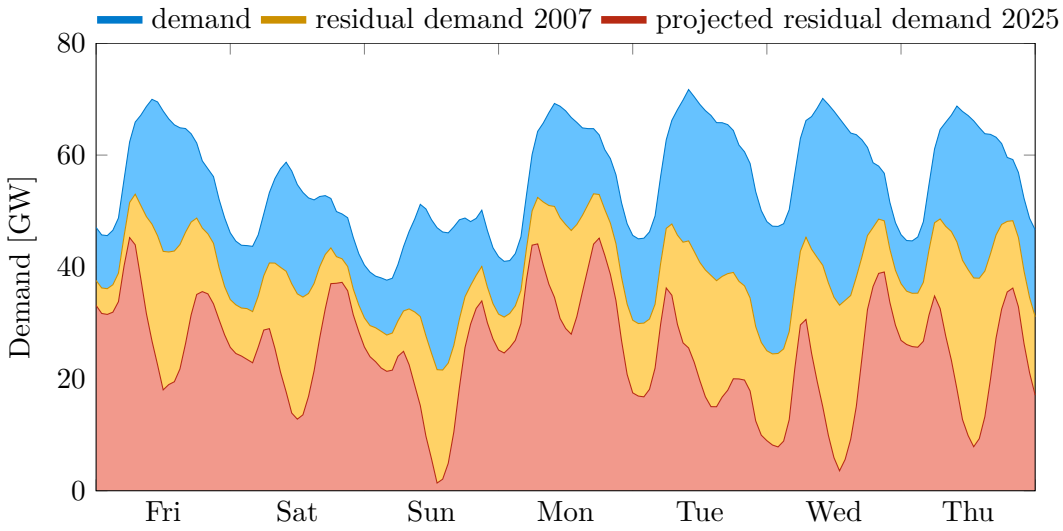


Figure 1.6: Total and residual demand in an exemplary week in Germany. Demand data for 2007 is provided by ENTSO-E [ENT]. Electricity production from solar and wind energy is extrapolated from installed capacity and production in 2014, projected capacity in 2025 and NASA weather profiles [Rie+11]. For comparability, all values are scaled to a yearly consumption of 520 TWh.

1.2 Mathematical Background

The roots of the Unit Commitment (UC) problem lie in the *Economic Dispatch problem*, where a certain demand needs to be distributed among a predefined set of online units at minimal cost, subject to operational constraints. Due to its complexity and non-linearity, this problem is initially solved by heuristic methods. As early as 1934 however, the so-called approach of "equal incremental rates" is presented in [SS34], which introduces an optimality condition representing a special case of the (yet to be published) Karush-Kuhn-Tucker conditions.

Enabled by the recent breakthrough in mixed integer programming (MIP) due to Gomory [Gom58], in 1962 [Gar62] extended the Economic Dispatch problem by the question which power plants should be online in which periods, thereby introducing the *Unit Commitment* problem (sometimes also referred to as the *Unit Commitment and Economic Dispatch* problem). To this end, the Economic Dispatch problem is simplified and linearized, loosing considerable detail.

The subsequent research saw the restoration of most of this detail, accompanied by a notable increase in modeled facets made possible by the inclusion of the commitment decisions. These results have been achieved using solution techniques ranging from (Meta-)Heuristics to Dynamic Programming, Lagrangian Relaxation, and MIP (see

[SF94; Pad04]). Due to its modeling flexibility and its ability to control the quality of the solution, MIP has emerged as the favored approach in recent years (see [SPO05; Ott10]).

The operational and economic constraints used in the literature vary heavily depending on the focus of the respective publication. We utilize a commonly used subset consisting of

- minimal and maximal production constraints,
- ramping constraints, i. e. constraints on the change of the production level,
- minimum up- and downtime constraints,
- affine linear production costs,
- start-up costs, and
- start-up production.

After summarizing the MIP, Subsection 1.2.2 gives a detailed account of the chosen constraints. The start-up costs and production, which are the main research interest of this work, are discussed separately in Section 1.3.

1.2.1 Nomenclature

The modeled time range is split into T consecutive periods of non-negative lengths $\mathrm{L}^1, \ldots, \mathrm{L}^T$, indexed by the set $\{1, \ldots, T\}$. To accommodate adaptive discretizations with period lengths based on the required level of detail, e. g. with shorter periods during working hours and longer periods during night, the period lengths may differ. While all proposed formulations explicitly consider non-homogeneous period lengths, the existing approaches summarized in this section are presented assuming $\mathrm{L}^1 = \ldots = \mathrm{L}^T = 1$ for the sake of simplicity.

The set $\mathcal{I}$ denotes electricity generating units which can be controlled. Virtually non-controllable units like wind and solar units are excluded from the optimization model: they merely change the demand by a constant term, and possibly introduce uncertainty. Each of the controllable units has an individual set of parameters (see Table 1.1) and is modeled using the variables in Table 1.2.

Regarding the notation, the term $[a..b]$ denotes $[a, b] \cap \mathbb{Z}$, and the shortcut $[b]$ denotes $[1..b]$. Note that we generally use superscripts as in "x^t" for the period index t and subscripts as in "x_i" for all other indices.

Parameter	Domain	Scale unit	Description
D^t	$\mathbb{Q}_{\geq 0}$	MW	Demand
L^t	$\mathbb{Q}_{>0}$	h	Period length
$\overline{\mathrm{P}}_i$	$\mathbb{Q}_{\geq 0}$	MW	Maximal production in online state
$\underline{\mathrm{P}}_i$	$\mathbb{Q}_{\geq 0}$	MW	Minimal production in online state
RU_i	$\mathbb{Q}_{\geq 0}$	MW/h	Maximal production increase in online state
RD_i	$\mathbb{Q}_{\geq 0}$	MW/h	Maximal production decrease in online state
SU_i	$\mathbb{Q}_{\geq 0}$	MW	Maximal production increase at start-up
SD_i	$\mathbb{Q}_{\geq 0}$	MW	Maximal production decrease at shutdown
A_i	$\mathbb{Q}_{\geq 0}$	costs/MWh	Cost increase per MWh of production
B_i	$\mathbb{Q}_{\geq 0}$	costs/h	Fixed costs in online state
CD_i	$\mathbb{Q}_{\geq 0}$	costs	Shutdown costs
UT_i	$\mathbb{N}$	periods	Minimal uptime
DT_i	$\mathbb{N}$	periods	Minimal downtime
POT_i	$\mathbb{Q}_{\geq 0}$	periods	Uptime before start of modeled time range
PDT_i	$\mathbb{Q}_{\geq 0}$	periods	Downtime before start of time range
$\mathrm{PU}_i^t(s, t')$	$\mathbb{Q}_{\geq 0}$	MW	Start-up production, type-dependent
$\mathrm{CU}_i^t(l)$	$\mathbb{Q}_{\geq 0}$	costs	Start-up costs, downtime dependent

Table 1.1: Parameters of the Unit Commitment problem

Variable	Domain	Scale unit	Description
v_i^t	$\{0, 1\}$		Operational state
y_i^t	$\{0, 1\}$		Start-up indicator
z_i^t	$\{0, 1\}$		Shutdown indicator
p_i^t	$\mathbb{R}_{\geq 0}$	MW	Production

Table 1.2: Variables of the Unit Commitment problem

1.2.2 Unit Commitment Reference

Objective function:

$$\min \sum_{i \in \mathcal{I}} \sum_{t \in [T]} \left(\mathrm{A}_i p_i^t + \mathrm{B}_i v_i^t + \mathrm{CD}_i \mathrm{z}_i^t \right) + \text{Start-up costs}$$

Operational State and Indicators:

$$\begin{aligned} v_i^t, \mathrm{y}_i^t, \mathrm{z}_i^t &\in \{0,1\}, && i \in \mathcal{I}, t \in [T] \\ \mathrm{y}_i^t - \mathrm{z}_i^t &= v_i^t - v_i^{t-1}, && i \in \mathcal{I}, t \in [2\,..\,T] \\ \mathrm{y}_i^1 - \mathrm{z}_i^1 &= v_i^1 - \begin{cases} 0 & \text{if } \mathrm{PDT}_i > 0, \\ 1 & \text{else,} \end{cases} && i \in \mathcal{I} \end{aligned}$$

Production and Production Limits:

$$\begin{aligned} p_i^t &\in \mathbb{R}_{\geq 0}, && i \in \mathcal{I}, t \in [T] \\ \underline{\mathrm{P}}_i v_i^t \leq p_i^t &\leq \overline{\mathrm{P}}_i v_i^t, && i \in \mathcal{I}, t \in [T] \\ p_i^t &\leq p_i^{t-1} + \mathrm{RU}_i v_i^{t-1} + \mathrm{SU}_i \mathrm{y}_i^t, && i \in \mathcal{I}, t \in [2..T] \\ p_i^t &\geq p_i^{t-1} - \mathrm{RD}_i v_i^t - \mathrm{SD}_i \mathrm{z}_i^t, && i \in \mathcal{I}, t \in [2..T] \end{aligned}$$

Minimal Up- and Downtime:

$$\begin{aligned} \sum_{l=0}^{\mathrm{UT}_i - 1} \mathrm{y}_i^{t-l} &\leq v_i^t, && i \in \mathcal{I}, t \in [\mathrm{UT}_i..T] \\ \sum_{l=0}^{\mathrm{DT}_i - 1} \mathrm{z}_i^{t-l} &\leq 1 - v_i^t, && i \in \mathcal{I}, t \in [\mathrm{DT}_i..T] \\ v_i^1 = \ldots = v_i^{\mathrm{UT}_i - \mathrm{POT}_i} &= 1, && i \in \mathcal{I} \text{ with } \mathrm{POT}_i > 0 \\ v_i^1 = \ldots = v_i^{\mathrm{DT}_i - \mathrm{PDT}_i} &= 0, && i \in \mathcal{I} \text{ with } \mathrm{PDT}_i > 0 \end{aligned}$$

Demand:

$$\sum_{i \in \mathcal{I}} p_i^t + \mathrm{SP} = \mathrm{D}^t . t \in [T]$$

Operational State and Indicators

The defining feature of the Unit Commitment problem is the operational state of each unit,

$$\forall\, i \in \mathcal{I}, t \in [T]: \qquad v_i^t \in \{0,1\},$$

where $v_i^t = 1$ iff unit i is committed to be online in period t. Already in [Gar62], the operational state is paired with the start-up indicator y_i^t and the shutdown indicator z_i^t,

$$\forall\, i \in \mathcal{I}, t \in [T]: \qquad \mathrm{y}_i^t, \mathrm{z}_i^t \in \{0,1\}, \tag{1.2.1}$$

with $\mathrm{y}_i^t = 1$ iff unit i starts up in period t and $\mathrm{z}_i^t = 1$ iff unit i shuts down in period t.

These indicators are derived from the operational state by

$$\forall\, i \in \mathcal{I}, t \in [2\,..\,T]: \qquad \mathrm{y}_i^t - \mathrm{z}_i^t = v_i^t - v_i^{t-1}, \tag{1.2.2}$$

$$\forall\, i \in \mathcal{I}: \qquad \mathrm{y}_i^1 - \mathrm{z}_i^1 = v_i^1 - \begin{cases} 1 & \text{if } \mathrm{PDT}_i = 0, \\ 0 & \text{else.} \end{cases} \tag{1.2.3}$$

By themselves, these equations allow solutions with $v_i^{t-1} = v_i^t$ and $\mathrm{y}_i^t = \mathrm{z}_i^t = 1$. This is prevented by the inequalities

$$\forall\, i \in \mathcal{I}, t \in [T]: \qquad \mathrm{y}_i^t \leq v_i^t,$$
$$\forall\, i \in \mathcal{I}, t \in [T]: \qquad \mathrm{z}_i^t \leq 1 - v_i^t.$$

which are dominated by the minimum up-/downtime inequalities, see (1.2.6) and (1.2.7) below.

In [CA06], the UC problem is modeled without y_i and z_i, resulting in improved solution times due to a smaller number of binary variables. This conclusion is negated in [OAV12], where a similar model including the indicator variables is found to be superior, which is partially explained by a possible improvement in the used MIP solver CPLEX. Furthermore, the minimum up-/downtime and ramping inequalities can be tightened using y_i and z_i.

We use the indicator variables

- in the extended formulation of the minimum up-/downtime (1.2.6) and (1.2.7),
- in the ramping inequalities (1.2.4) and (1.2.5), and
- in the time-dependent start-up costs presented in Chapter 3, Section 4.4, and Subsection 4.5.2.

Production and Production Limits

The electricity production of a unit is modeled by the non-negative variable p_i^t,

$$\forall\, i \in \mathcal{I}, t \in [T]: \qquad p_i^t \in \mathbb{R}_{\geq 0}.$$

Each unit has a certain online capacity $\overline{\mathrm{P}}_i$, which is generally modeled as

$$\forall\, i \in \mathcal{I}, t \in [T]: \qquad p_i^t \leq \overline{\mathrm{P}}_i v_i^t.$$

Note that we consider neither a varying maximal production due to environmental factors (e. g. temperature, humidity) nor the possibility of exceeding the maximal production for short time periods.

Most units have a minimal online production $\underline{\mathrm{P}}_i$, typically in the range of $0.4\overline{\mathrm{P}}_i$ to $0.8\overline{\mathrm{P}}_i$, which is enforced by

$$\forall\, i \in \mathcal{I}, t \in [T]: \qquad p_i^t \geq \underline{\mathrm{P}}_i v_i^t.$$

Units which do not have this restriction, e. g. hydro units, are modeled using $\underline{P}_i = 0$.

The production level of a unit may not change arbitrarily. The parameters RU_i and RD_i denote the maximal speed when increasing / decreasing the production level of an operating unit. For example, a unit with $\mathrm{RU}_i = 50\mathrm{MW/h}$ is able to increase it's production level from 120MW to at most 170MW in one hour. Two additional parameters SU_i and SD_i denote the maximal production level at start-up and before shutdown, respectively.

In [AC00], these constraints are modeled as

$$\forall\, i \in \mathcal{I}, t \in [T]: \qquad p_i^t \le p_i^{t-1} + \mathrm{RU}_i v_i^{t-1} + \mathrm{SU}_i \mathrm{y}_i^t, \tag{1.2.4}$$

$$\forall\, i \in \mathcal{I}, t \in [T]: \qquad p_i^t \ge p_i^{t-1} - \mathrm{RD}_i v_i^t - \mathrm{SD}_i \mathrm{z}_i^t. \tag{1.2.5}$$

Minimal Up- and Downtime

At start-up and shutdown, a fossil-fuel unit produces more exhaust gas than allowed during regular operation. To limit the extra pollution, such units are restricted from starting up and shutting down too frequently. Typically, this is enforced by a minimum uptime UT_i and minimum downtime DT_i, meaning that unit i may not be online for less than UT_i consecutive periods and offline for less than DT_i consecutive periods.

For each unit i, the resulting polyhedron of feasible operational schedules

$$P_i^{\mathrm{UDT}} := \mathrm{conv}\Big\{ v_i \in \{0,1\}^T \text{ fulfilling the minimum up- and downtime} \Big\},$$

may have an exponential number of facets, which are described by the alternating up/down inequalities presented in [LLM04].

We use an extended formulation of P_i^{UDT} presented in [RT05], which models the min up-/downtime constraints by the linear class of turn-on and turn-off inequalities,

$$\forall\, i \in \mathcal{I}, t \in [\mathrm{UT}_i..T]: \qquad \sum_{l=0}^{\mathrm{UT}_i - 1} \mathrm{y}_i^{t-l} \le v_i^t, \tag{1.2.6}$$

$$\forall\, i \in \mathcal{I}, t \in [\mathrm{DT}_i..T]: \qquad \sum_{l=0}^{\mathrm{DT}_i - 1} \mathrm{z}_i^{t-l} \le 1 - v_i^t. \tag{1.2.7}$$

The pre-model uptime POT_i and downtime PDT_i are respected by introducing the trivial inequalities

$$\forall\, i \in \mathcal{I} \text{ with } \mathrm{POT}_i > 0, t \in [\mathrm{UT}_i - \mathrm{POT}_i]: \qquad v_i^t = 1,$$

$$\forall\, i \in \mathcal{I} \text{ with } \mathrm{PDT}_i > 0, t \in [\mathrm{DT}_i - \mathrm{PDT}_i]: \qquad v_i^t = 0.$$

Demand

The electricity produced by each unit is transported to the electricity consumers via a power network. The interconnections of such a network have limited capacities and can restrict the set of feasible production schedules of units. However, current electricity markets do not take the power network into consideration when determining the market price and selecting successful bids.

Following recent publications regarding the start-up process [CA06; MELR13b; MELR13a; SBB10], we do not model the power network and enforce solely that demand and production meet,

$$\forall\, t \in [T]: \qquad \sum_{i \in \mathcal{I}} p_i^t + \mathrm{SP} = \mathrm{D}^t. \tag{1.2.8}$$

SP is a placeholder for the start-up production in some models, which is considered in Subsection 1.3.3. Models without start-up production use $\mathrm{SP} = 0$.

Production and Shutdown Costs

The overall costs consist of production costs cp_i^t, shutdown costs cd_i^t, and start-up costs.

$$\sum_{i \in \mathcal{I}} \sum_{t \in [T]} \left(\mathrm{cp}_i^t + \mathrm{cd}_i^t \right) + \text{Start-up Costs}.$$

The start-up costs are represented by the place-holder "Start-up Costs", and discussed in Section 1.3.

We model the production costs as in [Muc66], where each unit incurs fixed costs B_i while online and variable costs A_i per unit of production,

$$\forall\, i \in \mathcal{I}, t \in [T]: \qquad \mathrm{cp}_i^t := \mathrm{A}_i p_i^t + \mathrm{B}_i v_i^t.$$

Due to the affine linearity, the total costs per unit of production decrease with increasing production (see Fig. 1.7).

The shutdown costs are generally modeled as constant costs CD_i incurred at each shutdown,

$$\forall\, i \in \mathcal{I}, t \in [T]: \qquad \mathrm{cd}_i^t := \mathrm{CD}_i \mathrm{z}_i^t.$$

Both production and shutdown costs are linear expressions of the variables p_i^t, v_i^t, and cd_i^t, and their definition is substituted in the objective function,

$$\sum_{i \in \mathcal{I}} \sum_{t \in [T]} \left(\mathrm{A}_i p_i^t + \mathrm{B}_i v_i^t + \mathrm{CD}_i \mathrm{z}_i^t \right) + \text{Start-up Costs}.$$

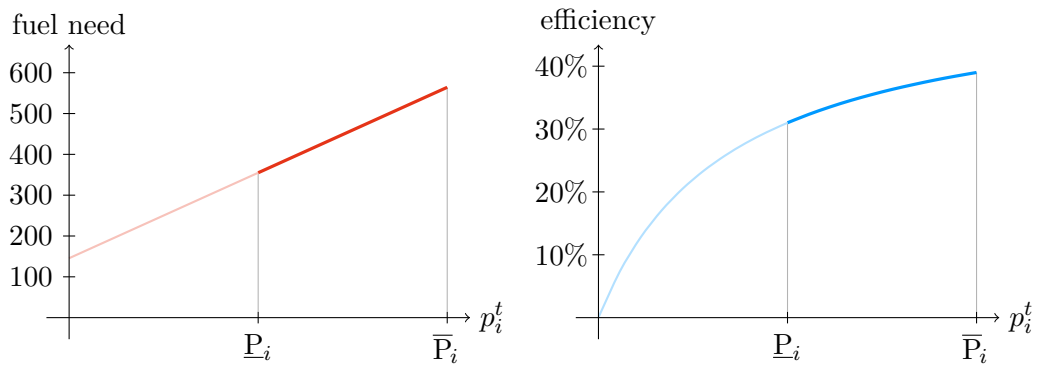


Figure 1.7: Fuel need and efficiency of a typical thermal unit. Note the non-zero y intercept of the fuel need function resulting in the curvature of the efficiency function. The unit has a maximal production $\overline{P}_{i,t} = 220$ MW, a minimal production $\underline{P}_{i,t} = 110$ MW (50%), and respective efficiencies of 39% and 31%.

1.2.3 Computational Complexity

The Unit Commitment problem is a well-known $\mathbb{NP}$-complete problem in the weak sense (see e. g. [Tse96; GZP03]) and, to the best of our knowledge, no proof of strong $\mathbb{NP}$-completeness has been published. The weak $\mathbb{NP}$-completeness is typically proved by using only a single period and units with a single feasible production level, i. e. $\underline{P}_i = \overline{P}_i$, effectively yielding the $\mathbb{NP}$-complete partition problem. Units with $\underline{P}_i = \overline{P}_i$ however do not occur in real-world instances. So, one may restrict the UC problem such that $\underline{P}_i \leq (1-\epsilon)\overline{P}_i$ without hurting its practical applicability, but invalidating this proof.

The complexity of the UC problem does not stem only from the difficulty of matching the demand, but from the difficulty of matching the demand *with maximal efficiency*. The efficiency of a hydrothermal unit depends both directly on the production level and indirectly on the preceding start-up costs. We give a proof relying on a varying production efficiency.

Proposition 1.1 *For UC problems, it is weakly $\mathbb{NP}$-complete to decide whether a solution with costs of at most $C \in \mathbb{Q}$ exists.*

Proof. Let $n, e_1, \ldots, e_n \in \mathbb{N}$ be given. The $\mathbb{NP}$-complete partition problem is to decide whether there exists $E \subset [n]$ such that

$$\sum_{i \in E} e_i = \sum_{i \in [n] \setminus E} e_i.$$

The existing proof constructs a UC problem with demand $D = \frac{1}{2} \sum_{i \in [n]} e_i$ and $\underline{P}_i = \overline{P}_i = e_i$, such that the problem is feasible iff such a set E exists. In the following,

we relax the minimal production to $\underline{\mathrm{P}}_i = 0$, but choose the production efficiency such that production costs of $C := \sum_{i\in[n]} e_i$ are only possible iff such an E exists.

We model the units $\mathcal{I} := [n]$ with parameters

- production limits $\underline{\mathrm{P}}_i := 0$, $\overline{\mathrm{P}}_i := e_i$,
- variable production costs $\mathrm{A}_i := 1$, and
- fixed production costs $\mathrm{B}_i := \overline{\mathrm{P}}_i$.

The number of periods is $T = 1$ and we hence drop the period index t. We define the demand as $\mathrm{D} := 1/2 \sum_{i\in[n]} e_i$ and $C := \sum_{i\in[n]} e_i = 2\mathrm{D}$.

Since only a single period is considered, ramping and minimum up-/down times are not relevant, and the UC problem degenerates to

$$\begin{aligned}
\min \sum_{i\in[n]} (\mathrm{A}_i p_i + \mathrm{B}_i v_i) & \\
p_i \le \overline{\mathrm{P}}_i v_i, \qquad & i \in [n] \\
\sum_{i\in\mathcal{I}} p_i = \mathrm{D}, & \\
v_i \in \{0,1\}, \qquad & i \in [n] \\
p_i \in \mathbb{R}_{\ge 0}, \qquad & i \in [n].
\end{aligned}$$

Note that the production costs are bounded by

$$\mathrm{A}_i p_i + \mathrm{B}_i v_i = p_i + \overline{\mathrm{P}}_i v_i \ge 2p_i,$$

and that $\mathrm{A}_i p_i + \mathrm{B}_i v_i = 2p_i$ iff $p_i = \overline{\mathrm{P}}_i v_i$. Summing over all units, the demand induces costs of at least

$$\sum_{i\in[n]} (\mathrm{A}_i p_i + \mathrm{B}_i v_i) \ge \sum_{i\in[n]} 2p_i = 2\mathrm{D} = C,$$

and a solution of the UC problem has costs C iff $p_i = \overline{\mathrm{P}}_i v_i$ for each unit i.

The remainder of this proof follows the existing approach closely: The optimal solution of a UC problem has costs C iff

$$\frac{1}{2} \sum_{i\in[n]} e_i = \mathrm{D} = \sum_{i\in[n]} p_i = \sum_{i\in[n]} \overline{\mathrm{P}}_i v_i = \sum_{i\in E} \overline{\mathrm{P}}_i = \sum_{i\in E} e_i$$

with $E := \{i \in [n] \mid v_i = 1\}$. Since this is equivalent to

$$\sum_{i\in E} e_i = \sum_{i\in[n]\setminus E} e_i,$$

E is a solution to the partition problem. So, the UC problem is at least as hard as the partition problem, and since it clearly lies in $\mathbb{NP}$, it is $\mathbb{NP}$-complete. □

1.3 Foundations of Start-up Modeling

As in the majority of the existing literature, we allow two operational states for a unit, online and offline, and do not consider intermediate states such as synchronization, soaking, desynchronization, banking, as well as transient interruptions. Hence, we define a start-up as the transition from offline to online state. Its main characteristics are the incurred costs and the preceding production, both of which depend on the downtime (cf. Fig. 1.8). Note that [SBB10] shows how to introduce the additional states in the start-up type model presented in Chapter 4.

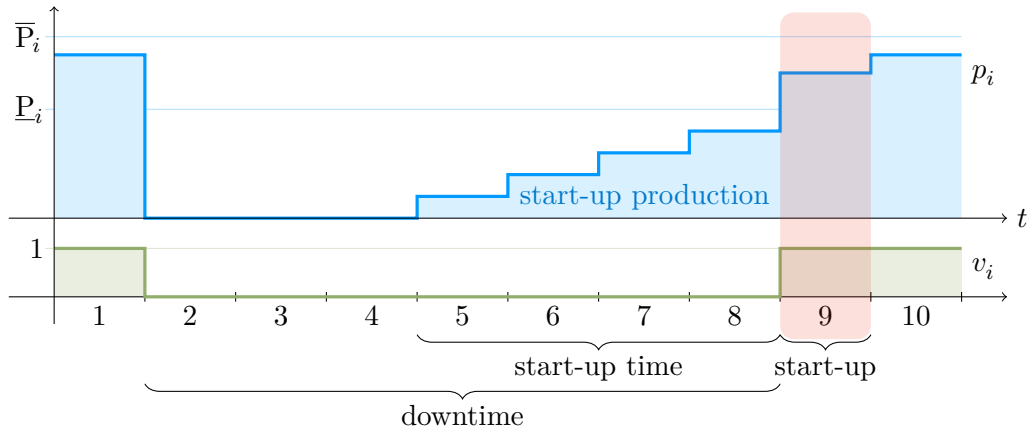


Figure 1.8: A start-up in period 9 after 7 offline periods, preceded by a typical start-up production spanning periods 5–8.

As shown in the numerical experiments in Chapter 5, the integrality gap and the computational performance of a UC formulation strongly depend on the quality of the lower bound on the start-up costs. Moreover, virtually every UC formulation since [Gar62] models the start-up costs, while the start-up production has only been introduced recently, e. g. in [SBB10]. Hence, we concentrate our theoretical analysis on the start-up costs, and demonstrate the effectiveness of our models with start-up production experimentally.

In this section, we

- give the basic notation and parameters of start-ups in Subsection 1.3.1,
- establish the start-up cost epigraphs and their importance to start-up cost modeling in Subsection 1.3.2,
- describe how the start-up production is modeled in the literature in Subsection 1.3.3, and
- summarize our contribution to start-up models in Section 1.4.

1.3.1 Start-up Notation

Start-up Cost Notation

When starting up, a unit incurs costs which increase with the downtime. These start-up costs are usually characterized as "exponential" (see e. g. [WW96; ZG88; MHV12; EHG11; RR91; PE10; AC00; MELR13b]) and defined as

$$\mathrm{CU}_i(L) = \begin{cases} \mathrm{CU}_i^{\mathrm{fixed}} + (1 - \mathrm{e}^{-\lambda_i L})\mathrm{CU}_i^{\mathrm{var}} & \text{if } L > 0, \\ 0 & \text{else,} \end{cases} \tag{1.3.1}$$

where $\mathrm{CU}_i^{\mathrm{fixed}}$ denotes the fixed costs, $\mathrm{CU}_i^{\mathrm{var}}$ denotes the maximal variable costs, and λ_i determines how fast the start-up costs converge to $\mathrm{CU}_i^{\mathrm{fixed}} + \mathrm{CU}_i^{\mathrm{var}}$ with increasing downtime L (cf. Fig. 1.9).

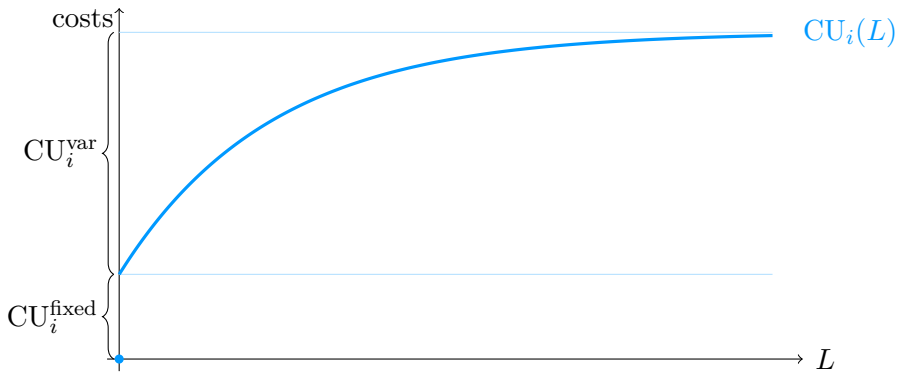


Figure 1.9: Typical exponential start-up cost function of a thermal unit

The term "exponential" refers to the variable start-up costs $(1 - \mathrm{e}^{-\lambda_i L})\mathrm{CU}_i^{\mathrm{var}}$, which increase inversely exponential towards the maximum value $\mathrm{CU}_i^{\mathrm{var}}$. Their characteristic development is caused by the exponential decay of the unit's temperature, which needs to be offset by heating.

Note that this definition of the start-up costs is an idealization and neglects such factors as the ambient temperature and air humidity. For this reason and to include non-exponential start-up costs, we restrict the class of admissible functions as little as possible. Specifically, we consider increasing functions in Sections 2.1 and 2.3 on single/all start-up costs, concave functions in Section 2.2 on the summed start-up costs, and strictly increasing functions in Chapter 4 on start-up types (i. e. on categorizing start-ups by the preceding downtime). Only the temperature model in Chapter 3 is limited to exponential start-up costs.

We expect the variable start-up costs of a unit $i \in \mathcal{I}$ to be given as a non-negative, increasing function CU_i depending on the non-negative downtime length L, with $\mathrm{CU}_i(0) = 0$ and $\mathrm{CU}_i(L) > 0$ for all $L > 0$:

$$\forall\, i \in \mathcal{I}: \quad \mathrm{CU}_i: \ \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}, \quad L \mapsto \text{start-up costs after downtime } L. \tag{1.3.2}$$

To match the discretization of the modeled time range into T periods of lengths $\mathrm{L}^1, \ldots, \mathrm{L}^T$, we also discretize the start-up cost functions such that $\mathrm{CU}_i^t(l)$ denotes the costs of a start-up in period t after l offline periods. To this end, we define the function $\mathrm{OL}^t(l)$ as the downtime in that situation, that is

$$\forall\, i \in \mathcal{I}, t \in [T],\, l \in [0\,..\,t{-}1]: \quad \mathrm{OL}^t(l) := \sum_{j=1}^{l} \mathrm{L}^{t-j} + \begin{cases} \mathrm{PDT}_i & \text{if } l = t-1, \\ 0 & \text{else.} \end{cases} \tag{1.3.3}$$

The case distinction differentiates between the case where the downtime lies completely within the modeled time range and the case where it stretches to include the pre-model downtime PDT_i. The corresponding start-up costs $\mathrm{CU}_i^t(l)$ are

$$\forall\, i \in \mathcal{I}, t \in [T], l \in [0\,..\,t{-}1]: \qquad \mathrm{CU}_i^t(l) := \mathrm{CU}_i(\mathrm{OL}^t(l)). \tag{1.3.4}$$

Start-up Production Notation

In preparation of a start-up, a unit increases its production following a specific power trajectory (cf. Fig. 1.8). In thermal units, the main purpose of this procedure is to increase the temperature in a controlled manner while limiting the amount of material tensions due to temperature gradients. Hence, the power trajectory, and in particular its duration, depend on the preceding downtime analogously to the start-up costs.

While the start-up time $\mathrm{TU}_i(L)$ is typically assumed to increase inversely exponential (cf. Fig. 1.10), the production during start-up is modeled as either constant or linearly increasing in the literature (see e. g. [SBB10; MELR13a]).

We expect the start-up time to be given by a non-negative function TU_i

$$\mathrm{TU}_i: \qquad \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}, \quad L \mapsto \begin{cases} \text{start-up time} & \text{if } L > 0, \\ 0 & \text{else.} \end{cases}$$

In the discretized time range of our model, the start-up time in period t must be expressed as a number of preceding offline periods s. We define the start-up time $\mathrm{TU}_i^t(l)$ in period t after l preceding offline periods such that the discretization error $\left|\mathrm{OL}^t(\mathrm{TU}_i^t(l)) - \mathrm{TU}_i(\mathrm{OL}^t(l))\right|$ is minimal,

$$\forall\, t \in [T], l \in [t-1]: \quad \mathrm{TU}_i^t(l) := \min\Big\{ s \in [1\,..\,t{-}1] \;\Big|\; \mathrm{OL}^t(s) \geq \mathrm{TU}_i(\mathrm{OL}^t(l)) - 0.5 \Big\}.$$

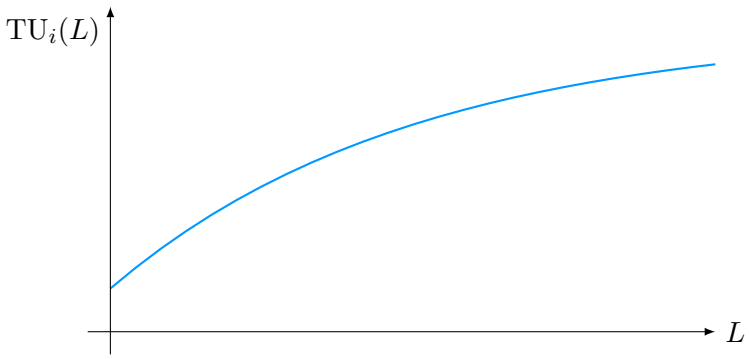


Figure 1.10: Exemplary start-up time $\mathrm{TU}_i(L)$ of a thermal unit, which increases inversely exponential with the preceding downtime L.

The start-up production in period t caused by a start-up in period t' after l preceding offline periods is denoted by $\mathrm{PU}_i^t(t', l)$. Following the discretization of the start-up time, we model it as a linear function with values from $\underline{\mathrm{PU}}_i$ to $\overline{\mathrm{PU}}_i$,

$$\mathrm{PU}_i^t(t', l) = \begin{cases} \left(\overline{\mathrm{PU}}_i - \underline{\mathrm{PU}}_i\right) \dfrac{\mathrm{L}^{t'-\mathrm{TU}_i^{t'}(l)} + \ldots + \mathrm{L}^{t}}{\mathrm{L}^{t'-\mathrm{TU}_i^{t'}(l)} + \ldots + \mathrm{L}^{t'-1}} + \underline{\mathrm{PU}}_i & \text{if } t' - \mathrm{TU}_i^{t'}(l) \leq t < t', \\ 0 & \text{else,} \end{cases} \tag{1.3.5}$$

where the term

$$\frac{\mathrm{L}^{t'-\mathrm{TU}_i^{t'}(l)} + \ldots + \mathrm{L}^{t}}{\mathrm{L}^{t'-\mathrm{TU}_i^{t'}(l)} + \ldots + \mathrm{L}^{t'-1}}$$

is chosen to increase linearly from 0 in period $t = t' - \mathrm{TU}_i^{t'}(l) - 1$ (which is not part of the start-up time) to 1 in period $t' - 1$.

1.3.2 Epigraphs of Start-up Costs Functions

As noted, the operational state of a unit is modeled by the variables $v_i^t \in \{0, 1\}$, where $v_i^t = 1$ iff unit i is online in period t. Deriving the start-up costs in each period t from these variables yields the *discrete start-up cost functions*

$$\forall\, i \in \mathcal{I}, t \in [T]: \quad \mathrm{DCU}_i^t: \quad \{0,1\}^T \to \mathbb{R}_{\geq 0}, \quad v_i \mapsto \begin{cases} \mathrm{CU}_i^t(\mathrm{ol}^t(v_i)) & \text{if } v_i^t = 1, \\ 0 & \text{else,} \end{cases} \tag{1.3.6}$$

where

$$\mathrm{ol}^t(v_i) := \max\{l \in [t-1] \mid v_i^{t-1} = \ldots = v_i^{t-l} = 0\}. \tag{1.3.7}$$

In Fig. 1.11 this relationship is depicted for an exemplary operational schedule. Note that ol and OL follow the notational convention of l and L to denote numbers of periods in lower case and time lengths in upper case.

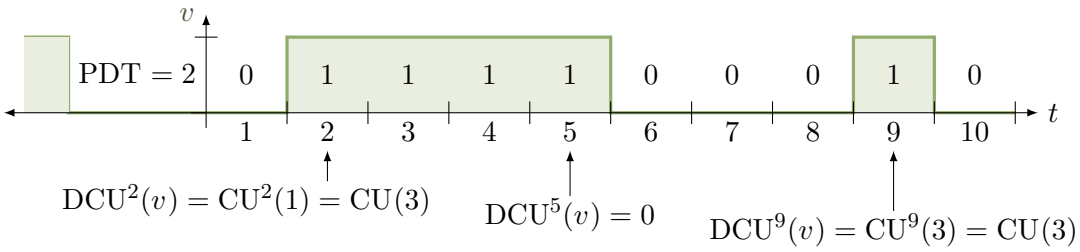


Figure 1.11: Start-up costs incurred for an exemplary operational schedule $v \in \{0,1\}^T$ (green) with $T = 10$ uniform periods ($\mathrm{L}^1 = \ldots = \mathrm{L}^T = 1$). The first start-up takes place in period 2, with costs $\mathrm{DCU}^2(v)$ of $\mathrm{CU}(3)$ due to the preceding offline period with length $\mathrm{L}^1 = 1$ and the pre-model downtime $\mathrm{PDT} = 2$. The only other start-up costs are incurred in period 9, after three offline periods with a total length of $\mathrm{L}^6 + \mathrm{L}^7 + \mathrm{L}^8 = 3$.

Using DCU_i^t, any Unit Commitment problem with start-up costs can be written as

$$\min\left\{ c(v_1,\ldots,v_{|\mathcal{I}|},d) + \sum_{i\in\mathcal{I}} \sum_{t\in[T]} \mathrm{DCU}_i^t(v_i) \;\middle|\; (v_1,\ldots,v_{|\mathcal{I}|},d) \in \mathcal{F} \right\}, \tag{1.3.8}$$

with additional variables $d \in \mathbb{R}^n$, feasible set $\mathcal{F} \subset \{0,1\}^{T|\mathcal{I}|} \times \mathbb{R}^n$, and objective function c. In the presented Unit Commitment problem, d is a placeholder for the electricity production p, startup/shutdown indicators y, z and cost variables cp, cd; $\mathcal{F}$ models production limits, ramping limits, minimum up-/downtime, demand, and costs. Recognize that even if the objective c is linear, the extended objective is non-linear due to DCU_i^t.

Since the domain $\{0,1\}^T$ of DCU_i^t is finite, the convex hull of its epigraph,

$$\begin{aligned} \mathrm{conv}(\mathrm{epi}(\mathrm{DCU}_i^t)) &= \mathrm{conv}\Big(\Big\{ (v_i, \mathrm{cu}_i^t) \;\Big|\; v_i \in \{0,1\}^T, \mathrm{cu}_i^t \geq \mathrm{DCU}_i^t(v_i) \Big\}\Big) \\ &= \mathrm{conv}(V^t) + \mathrm{pos}(u_{T+1}), \end{aligned} \tag{1.3.9}$$

is a polyhedron with recession cone $\mathrm{pos}(u_{T+1})$ (where u_{T+1} denotes the unit vector in the direction of the last coordinate cu_i^t), and with vertices

$$V^t = \{ (v_i, \mathrm{DCU}_i^t(v_i)) \mid v_i \in \{0,1\}^T \}.$$

So, using the variables cu_i^t, the minimization problem (1.3.8) may be rewritten as

$$\min\left\{\begin{array}{c} c(v_1,...,v_{|\mathcal{I}|},d)+\sum_{i\in\mathcal{I}}\sum_{t\in[T]}\mathrm{cu}_i^t: \\ \\ (v_i,\mathrm{cu}_i^t)\in\mathrm{conv}(\mathrm{epi}(\mathrm{DCU}_i^t)), \qquad i\in\mathcal{I},t\in[T] \\ (v_1,...,v_{|\mathcal{I}|},d)\in\mathcal{F} \end{array}\right\}. \tag{1.3.10}$$

If c is affine linear and $\mathcal{F}$ is expressible in a mixed integer program (MIP), then the minimization problem (1.3.10) is a MIP as well, since the polyhedron $\mathrm{conv}(\mathrm{epi}(\mathrm{DCU}_i^t))$ may be modeled using its $\mathcal{V}$-representation in (1.3.9) (see [DW60]). Note that due to the 2^T vertices of $\mathrm{conv}(\mathrm{epi}(\mathrm{DCU}_i^t))$, this representation is exponential.

In addition to modeling the start-up costs, $\mathrm{conv}(\mathrm{epi}(\mathrm{DCU}_i^t))$ also acts as a reference formulation: It models the best-possible lower bound on the start-up costs cu_i^t which can be achieved using inequalities of the shape

$$\mathrm{cu}_i^t \geq \sum_{t\in[T]}\alpha^t v_i^t+\beta \qquad \text{with } \alpha^t,\beta\in\mathbb{R}, \tag{1.3.11}$$

if the set of feasible solutions $\mathcal{F}$ meets the condition (1.3.14) given in the following.

In fact, by definition of the epigraph, the constraint "$(v_i,\mathrm{cu}_i^t)\in\mathrm{conv}(\mathrm{epi}(\mathrm{DCU}_i^t))$" dominates each inequality (1.3.11) that fulfills

$$\forall\, v_i\in\{0,1\}^T: \qquad \mathrm{DCU}_i^t(v_i)\geq\sum_{t\in[T]}\alpha^t v_i^t+\beta. \tag{1.3.12}$$

An inequality (1.3.11) that is valid for the minimization problem (1.3.10) satisfies

$$\forall\,(v_1,...,v_{|\mathcal{I}|},d)\in\mathcal{F}: \qquad \mathrm{DCU}_i^t(v_i)\geq\sum_{t\in[T]}\alpha^t v_i^t+\beta. \tag{1.3.13}$$

Condition (1.3.12) is stronger than (1.3.13), since $\mathcal{F}$ may exclude some operational schedules $v_i\in\{0,1\}^T$ for unit i (cf. Fig. 1.12). If we however assume that

$$\forall\,\tilde{v}_i\in\{0,1\}^T \;\; \exists\,(v_1,...,v_{|\mathcal{I}|},d)\in\mathcal{F} \text{ with } v_i=\tilde{v}_i, \tag{1.3.14}$$

then both conditions are equivalent, and so the constraint "$(v_i,\mathrm{cu}_i^t)\in\mathrm{conv}(\mathrm{epi}(\mathrm{DCU}_i^t))$" dominates all valid inequalities of shape (1.3.11).

The functions DCU_i^t are not mutually independent, e. g. $\mathrm{DCU}_i^{t-1}(v_i)>0$ implies $\mathrm{DCU}_i^t(v_i)=0$. So, stronger lower bounds may be deduced on combinations of multiple start-up costs cu_i^t. Using the *discrete start-up cost functions of all periods*,

$$\mathrm{DCU}_i: \quad \{0,1\}^T\to\mathbb{R}^T_{\geq 0}, \quad v_i\mapsto(\mathrm{DCU}_i^1(v_i),\ldots,\mathrm{DCU}_i^T(v_i)),$$

we may replace the condition

$$(v_i, \mathrm{cu}_i^t) \in \mathrm{conv}(\mathrm{epi}(\mathrm{DCU}_i^t)), \qquad i \in \mathcal{I}, t \in [T]$$

in (1.3.10) by

$$(v_i, \mathrm{cu}_i) \in \mathrm{conv}(\mathrm{epi}(\mathrm{DCU}_i)), \qquad i \in \mathcal{I},$$

thus tightening the lower bound on the start-up costs. Analogous to $\mathrm{conv}(\mathrm{epi}(\mathrm{DCU}_i^t))$, "$(v_i, \mathrm{cu}_i) \in \mathrm{conv}(\mathrm{epi}(\mathrm{DCU}_i))$" dominates each inequality

$$\sum_{t\in[T]} \gamma^t \mathrm{cu}_i^t \geq \sum_{t\in[T]} \alpha^t v_i^t + \beta$$

which is valid for (1.3.10).

Virtually all Unit Commitment models in the literature minimize the (unweighted) sum of the start-up costs. So, in place of DCU_i, it suffices to consider the *summed start-up cost functions*

$$\mathrm{DCU}_i^\Sigma : \quad \{0,1\}^T \to \mathbb{R}^T_{\geq 0}, \quad v_i \mapsto \sum_{t\in[T]} \mathrm{DCU}_i^t(v_i).$$

Replacing

$$(v_i, \mathrm{cu}_i) \in \mathrm{conv}(\mathrm{epi}(\mathrm{DCU}_i)), \qquad i \in \mathcal{I},$$

in (1.3.10) by

$$(v_i, \mathrm{cu}_i^\Sigma) \in \mathrm{conv}(\mathrm{epi}(\mathrm{DCU}_i^\Sigma)), \qquad i \in \mathcal{I},$$

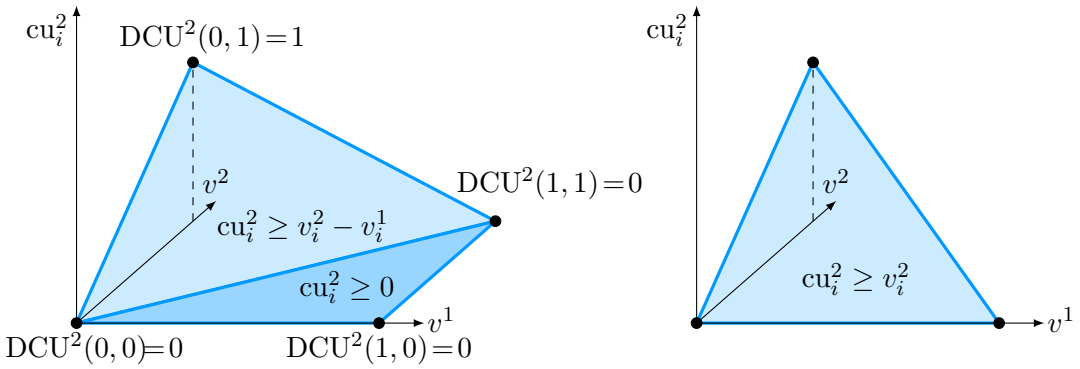


Figure 1.12: The lower boundary of $\mathrm{conv}(\mathrm{epi}(\mathrm{DCU}^2))$ (left) and of $\mathrm{conv}(\mathrm{epi}(\mathrm{DCU}^2|_{\mathcal{F}}))$, i. e. restricted to the schedules $\mathcal{F} = \{(0,0),(0,1),(1,0)\}$ (right). Note that due to the restriction to $\mathcal{F}$, the two inequalities describing the lower boundary in the left figure are replaced by a single tighter inequality in the right figure.

and therefore minimizing over cu_i^{Σ} yields the same lower bound on the sum of the start-up costs as $\min \sum_{t\in[T]} \mathrm{cu}_i^t$ in $\mathrm{conv}(\mathrm{epi}(\mathrm{DCU}_i))$. Yet $\mathrm{conv}(\mathrm{epi}(\mathrm{DCU}_i))$ remains interesting due to its extended formulation presented in Chapter 4 based on start-up types (cf. Subsection 1.3.3).

While condition (1.3.14) is not met in typical Unit Commitment problems, most notably due to minimum up- and downtime, $\mathrm{conv}(\mathrm{epi}(\mathrm{DCU}_i^t))$, $\mathrm{conv}(\mathrm{epi}(\mathrm{DCU}_i^{\Sigma}))$, and $\mathrm{conv}(\mathrm{epi}(\mathrm{DCU}_i))$ nonetheless provide goalposts against which every start-up model should be compared. Several of the proposed formulations achieve the same bounds as the start-up cost epigraphs, in contrast to all existing formulations. Moreover, Section 6.4 points out how to tighten some of these models if a minimum downtime is enforced.

1.3.3 Start-up Types

To model the characteristics of the start-up process, [SBB10] extends the so-called *start-up types* introduced in [Muc66]: the start-ups of each unit are subdivided in a set of types, e. g. hot, warm, and cold start-up type, and each type possesses a specific start-up time and a power trajectory.

This subdivision is performed by partitioning the possible offline lengths $[0 \,..\, t-1]$ in each period t into intervals $\mathcal{L}_0^t, \dots, \mathcal{L}_{S^t}^t$, such that all offline lengths $l \in \mathcal{L}_s^t$ belong to the same start-up type $s \in [S^t]$. Each type s is represented by a variable $\delta_s^t \in \{0,1\}$, where $\delta_s^t = 1$ iff the unit starts up in period t after $l \in \mathcal{L}_s^t$ offline periods. Using these variables, modeling the properties of each start-up becomes straightforward. In particular, the start-up costs equal

$$\sum_{t\in[T],s\in[0..S^t]} \mathrm{CU}_i^t\Big(\min \mathcal{L}_s^t\Big)\delta_s^t,$$

where $\mathrm{CU}_i^t\big(\min \mathcal{L}_s^t\big)$ represents the costs of start-up type s.

Furthermore, the start-up production in period t can be expressed as

$$\sum_{t'\in[t+1..T],s\in[S^t]} \widetilde{\mathrm{PU}}_i^t(t',s)\delta_s^{t'},$$

where $\widetilde{\mathrm{PU}}_i^t(t',s)$ denotes the start-up production in period t due to a start-up in period t' of type s, and approximates the values $\mathrm{PU}_i^t(t',l)$ with $l \in \mathcal{L}_s^{t'}$. Thus, the power trajectories at start-up are introduced in our UC model from Section 1.2 by redefining the demand equation (1.2.8) as

$$\forall\, t \in [T]: \qquad \sum_{i\in\mathcal{I}} p_i^t + \sum_{i\in\mathcal{I}} \sum_{\substack{t'\in[t+1\,..\,T]\\ s\in[S^{t'}]}} \widetilde{\mathrm{PU}}_i^t(t',s)\delta_{i,s}^{t'} = \mathrm{D}^t. \tag{1.3.15}$$

While the start-up types are able to model all characteristics of a start-up, they are beneficial even when only modeling start-up costs. In comparison to the state-of-the-art model not based on start-up types, [MELR13b] demonstrates experimentally that the start-up type model is computationally more efficient due to a superior integrality gap.

1.4 Contribution

Subsection 1.3.2 shows that the start-up costs in a UC problem can be modeled using the $\mathcal{V}$-representations of the start-up cost epigraphs $\operatorname{conv}(\operatorname{epi}(\mathrm{DCU}^t))$, $\operatorname{conv}(\operatorname{epi}(\mathrm{DCU}^\Sigma))$, or $\operatorname{conv}(\operatorname{epi}(\mathrm{DCU}))$. Unfortunately, the $\mathcal{V}$-representations are of exponential size, since each $v \in \{0,1\}^T$ induces a vertex $(v, \mathrm{DCU}^t(v))$, $(v, \mathrm{DCU}(v))$, or $(v, \mathrm{DCU}^\Sigma(v))$ of the respective epigraph. Therefore, the principal aim of our work lies on finding (extended) formulations of the start-up cost epigraphs. In particular, we improve the start-up type formulation introduced in [Muc66], which is a promising approach to model start-up costs and production.

As we consider the units individually, we drop the index i in the following.

The main results are laid out in four chapters:

- Chapter 2 derives a linear $\mathcal{H}$-representation of $\operatorname{conv}(\operatorname{epi}(\mathrm{DCU}^t))$, an exponential $\mathcal{H}$-representation of $\operatorname{conv}(\operatorname{epi}(\mathrm{DCU}^\Sigma))$, and an exponential class of facets of $\operatorname{conv}(\operatorname{epi}(\mathrm{DCU}))$ (joint work with René Brandenberg and Matthias Huber, partially published in [BS14; SHB16; BHS16]).
- Chapter 3 introduces the temperature model, an extended formulation of the epigraph $\operatorname{conv}(\operatorname{epi}(\mathrm{DCU}^\Sigma))$ with $\mathcal{O}(T)$ variables and $\mathcal{O}(T^2)$ inequalities. It is based on explicitly modeling the temperature of a unit and can be complemented to include a start-up production (joint work with René Brandenberg and Matthias Huber, partially published in [SHB16; BHS]).
- Chapter 4 presents the start-up type model, an extended formulation of the epigraph $\operatorname{conv}(\operatorname{epi}(\mathrm{DCU}))$ with $\mathcal{O}(T^2)$ variables and $\mathcal{O}(T^2)$ inequalities which significantly tightens the approach in [SBB10].
- Chapter 5 demonstrates numerically that the above models outperform the existing formulations both in terms of integrality gap and solution time.

After elaborating on each of these chapters in the following, we summarize our contribution in Subsection 1.4.1 by giving an overview of the existing and newly introduced formulations, and the separation algorithms and formulation relationships presented in this work.

Chapter 2: Epigraphs of Start-up Cost Functions

Chapter 2 exploits the structure of the discrete start-up cost function DCU^t to analyze the epigraphs $\mathrm{conv}(\mathrm{epi}(\mathrm{DCU}^t))$, $\mathrm{conv}(\mathrm{epi}(\mathrm{DCU}^\Sigma))$, and $\mathrm{conv}(\mathrm{epi}(\mathrm{DCU}))$.

In [CA06], the enclosing polyhedron P^t_{ex} of $\mathrm{conv}(\mathrm{epi}(\mathrm{DCU}^t))$ with $\mathcal{O}(T)$ inequalities is proposed, which models the start-up costs correctly for integral operational schedules $v \in \{0,1\}^T$. This suffices since $v \in \{0,1\}^T$ is enforced by the integrality constraints of the MIP, but only provides a weak bound on the start-up costs for fractional $v \in [0,1]^T$. Hence, the linear relaxations considered in the MIP solution process provide only poor lower bounds, causing slow convergence of the optimality gap. We tighten the inequalities of [CA06] in Section 2.1, introducing a $\mathcal{H}$-representation of $\mathrm{conv}(\mathrm{epi}(\mathrm{DCU}^t))$ with $\mathcal{O}(T)$ inequalities (see Theorem 2.8) which decreases the integrality gap significantly.

In practice, the number of inequalities in P^t_{ex} is usually reduced by modeling an approximation of the start-up cost function. We apply this approach to $\mathrm{conv}(\mathrm{epi}(\mathrm{DCU}^t))$ and present an algorithm which derives an approximation leading to a model of minimal size for any tolerance $\mathrm{CU}_{\mathrm{tol}} > 0$ (see Proposition 2.14).

Section 2.2 applies an iterative lifting to $\mathrm{conv}(\mathrm{epi}(\mathrm{DCU}^\Sigma))$ starting from the trivial inequality $\mathrm{cu}^\Sigma \geq 0$, deriving the resulting inequalities analytically. It shows that in general, the facets on the lower boundary of $\mathrm{conv}(\mathrm{epi}(\mathrm{DCU}^\Sigma))$ have a one-to-one relationship to binary trees of size T (see Theorem 2.37): for each $v \in [0,1]^T$, it pin-points the *Cartesian tree* (cf. [Vui80]) of v as the binary tree corresponding to the facet which defines the value $\mathrm{LCU}^\Sigma(v)$ (see Corollary 2.35).

Based on this characterization, we derive a generally irredundant but exponential $\mathcal{H}$-representation of $\mathrm{conv}(\mathrm{epi}(\mathrm{DCU}^\Sigma))$ whose inequalities can be separated in linear time using the Cartesian tree algorithm given in [GBT84] (see Proposition 2.39). While modeling $\mathrm{conv}(\mathrm{epi}(\mathrm{DCU}^\Sigma))$ in a cutting plane approach is impractical due to slow convergence and numerical problems, this separation algorithm is the basis for the separation of the practically relevant formulation P^{temp} presented in Chapter 3.

Section 2.3 generalizes the inequalities of $\mathrm{conv}(\mathrm{epi}(\mathrm{DCU}^t))$ to an exponential class of facet-inducing inequalities of $\mathrm{conv}(\mathrm{epi}(\mathrm{DCU}))$, showing that in general any $\mathcal{H}$-representation of $\mathrm{conv}(\mathrm{epi}(\mathrm{DCU}))$ is exponential (see Proposition 2.49). While Section 2.3 does not derive an $\mathcal{H}$-representation of $\mathrm{conv}(\mathrm{epi}(\mathrm{DCU}))$ itself, Chapter 4 presents an extended formulation of $\mathrm{conv}(\mathrm{epi}(\mathrm{DCU}))$ which provides a starting point to find the facets of $\mathrm{conv}(\mathrm{epi}(\mathrm{DCU}))$ (see Section 6.3).

Using their polynomial separation algorithms, the Ellipsoid method solves optimization problems over $\mathrm{conv}(\mathrm{epi}(\mathrm{DCU}^t))$ and $\mathrm{conv}(\mathrm{epi}(\mathrm{DCU}^\Sigma))$ in polynomial time (see e. g. [Sch98]). Remember however that we are interested in the Unit Commitment problem, including for example the constraints reviewed in Section 1.2. Therefore, the following sections study the formulations P^{temp} and P^δ which lead to better computational

performance than conv(epi(DCU^t)) and conv(epi(DCU)) when used as part of a UC problem.

Chapter 3: The Temperature Model

As described in Subsection 1.3.1, the start-up costs are typically modeled by an exponential start-up cost function $\mathrm{CU}(t) = A\mathrm{e}^{-\lambda t} + B$, which is caused by the exponential decay of the temperature of a unit during downtime. By explicitly modeling a unit's temperature and heating, Chapter 3 introduces two new formulations:

- $\hat{P}^{\mathrm{temp}}$ models the start-up costs correctly for $v \in \{0,1\}^T$ using $\mathcal{O}(T)$ variables and $\mathcal{O}(T)$ inequalities (see Theorem 3.13). Its integrality gap is typically considerably lower than that of the the state-of-the-art formulations leading to a superior computational performance (see Sections 5.3 and 5.4).
- The additional $\mathcal{O}(T^2)$ *residual temperature inequalities* defined in (3.3.1) further decrease the integrality gap, yielding P^{temp}. Building on the results on conv(epi(DCU^{Σ})) in Section 2.2, we show that P^{temp} is an extended formulation of conv(epi(DCU^{Σ})) and derive a linear-time separation algorithm for P^{temp} (see Theorems 3.26 and 3.33), which proves to be computationally competitive.

Section 3.4 generalizes the temperature approach from a linear formulation to a convex formulation and thereby extends the representable start-up cost functions by a class of non-exponential functions. At the same time however it shows that the temperature approach only results in linear formulations for exponential start-up cost functions $\mathrm{CU}(t) = A\mathrm{e}^{-\lambda t} + B$ (see Proposition 3.35).

Finally, Section 6.2 (based on [HS15]) points out that these formulations can be extended to model the underlying heating limitations of each unit, therefore including specific start-up time functions and power trajectories.

Chapter 4: Start-up Types

As introduced in Subsection 1.3.3, the current state-of-the-art model P^{δ}_{ex} from [SBB10] models start-ups by classifying each start-up as one of S start-up types. Chapter 4 characterizes these start-up types as optimal flows in a special network flow model (see Proposition 4.3), allowing us to tighten P^{δ}_{ex}:

- By interpreting the inequalities of P^{δ}_{ex} as combinations of flow inequalities, tighter inequalities are derived. The resulting formulation $\hat{P}^{\delta}$ (see Eq. (4.3.8)) is of the same size as P^{δ}_{ex} ($\mathcal{O}(ST)$ variables and $\mathcal{O}(ST)$ inequalities) but models a significantly tighter lower bound on the start-up costs.
- Generalizing the inequalities of $\hat{P}^{\delta}$ leads to the even tighter formulation P^{δ} in (4.3.10) with $\mathcal{O}(ST)$ variables and $\mathcal{O}(T^2)$ inequalities. These inequalities can be

separated in time $\mathcal{O}(T^2)$ (see Theorem 4.13), effectively evaluating each inequality in $\mathcal{O}(1)$ despite its $\mathcal{O}(ST)$ non-zero coefficients. Furthermore, we show that in the edge case $S = T$, P^δ degenerates into the flow polyhedron P^f in (4.2.5) with $\mathcal{O}(T^2)$ variables but only $\mathcal{O}(T)$ inequalities.

As demonstrated by the numerical experiments in Chapter 5, $\hat{P}^\delta$ and P^δ are computationally superior to P^δ_{ex}, even when modeling the start-up production.

The network flow interpretation furthermore allows to transfer fundamental results of combinatorial optimization:

- By applying the total unimodularity of network flow programs to the polyhedron P^f, we prove that P^f is an extended formulation of conv(epi(DCU)) (see Theorem 4.5), and
- by using the Min-Cut/Max-Flow theorem, we show that P^f and P^δ are equivalent under projection (see Theorem 4.10), and thereby that P^δ is an extended formulation of conv(epi(DCU)), too.

Even surpassing the tightness of conv(epi(DCU)), Section 6.4 details how the formulations P^δ_{ex}, $\hat{P}^\delta_{\text{yz}}$, P^δ, and P^f can be further tightened in the presence of minimal up- and downtimes inequalities.

Finally, Section 4.5 considers the step-wise models P^t_{ex} and conv(epi(DCU^t)) in the framework of network flows. Using the start-up and shutdown indicators, we derive the formulation P^{yz} in Eq. (4.5.4) with the same modeling approach as conv(epi(DCU^t)) but the same solution structure as P^δ_{ex}. This allows us to highlight the differences between conv(epi(DCU^t)), P^δ_{ex}, and P^δ in the network flow framework (see Fig. 4.8). In particular, we prove that P^δ_{ex} models a tighter bound on the start-up costs than P^t_{ex} (see Proposition 4.19, Proposition 4.18), which has already been conjectured in [MELR13b] based on numerical experiments.

Chapter 5: Numerical Experiments

Chapter 5 compares the presented start-up formulations numerically by modeling the German power system, both with and without start-up production. The experiments corroborate the proved relationships: The novel formulation $\hat{P}^{\text{temp}}$ and $\hat{P}^\delta$ in general have a considerable lower integrality gap than the existing P^t_{ex} and P^δ_{ex}, and are in turn consistently outperformed by P^{temp} and P^δ.

This advantage translates directly into a better computational performance, as is demonstrated by examining the number of solved instances in a certain time frame. Again, $\hat{P}^{\text{temp}}$, P^{temp}, $\hat{P}^\delta$, and P^δ outperform the existing models significantly. Interestingly, the gap between $\hat{P}^{\text{temp}}$, P^{temp} and between $\hat{P}^\delta$, P^δ, respectively, is minor, with $\hat{P}^{\text{temp}}$ even dominating P^{temp} in some cases.

1.4.1 Summary

Figure 1.13 gives an overview of the novel formulations (red), the existing state-of-the-art formulations (blue), and their relationships uncovered in this work (red).

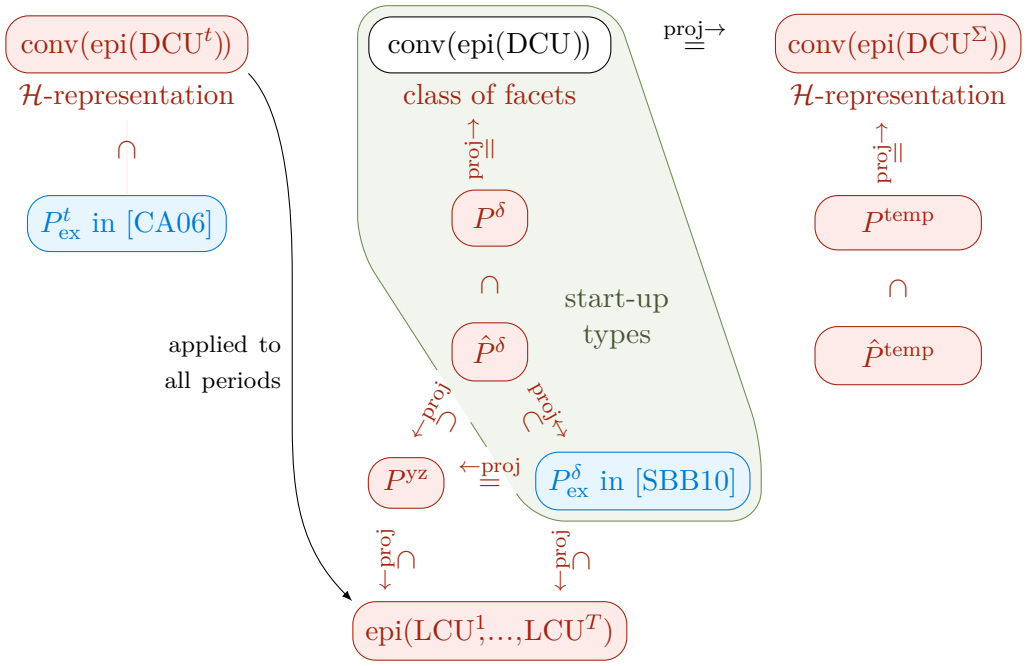


Figure 1.13: Comparison of the start-up cost models, partially under projection ("proj"). Blue nodes denote state-of-the art models, nodes and set relationships in red denote our contribution presented in this thesis. Green marks formulations with start-up types, which can be used to model the start-up process. The combination of $(v, \mathrm{cu}^t) \in \mathrm{conv}(\mathrm{epi}(\mathrm{DCU}^t))$ for all periods t is denoted by $\mathrm{epi}(\mathrm{LCU}^1, \ldots, \mathrm{LCU}^T)$. Note that the proposed formulations both dominate the state-of-the-art and model all three start-up cost epigraphs.

As shown, for each of the existing formulations we present a superior formulation, and model all three start-up cost epigraphs,

- either directly in the case of $\mathrm{conv}(\mathrm{epi}(\mathrm{DCU}^t))$,
- by the extended formulation P^δ in the case of $\mathrm{conv}(\mathrm{epi}(\mathrm{DCU}))$,
- or both directly and by the extended formulation P^{temp} for $\mathrm{conv}(\mathrm{epi}(\mathrm{DCU}^\Sigma))$.

A list of all formulations, their source, and their number of variables and inequalities is given in Table 1.3. It should be highlighted that we have tightened each of the existing formulations without increasing their respective number of variables and inequalities: $\mathrm{conv}(\mathrm{epi}(\mathrm{DCU}^t))$ tightens P^t_{ex} and $\hat{P}^\delta$ tightens P^δ_{ex}.

Formulation	Source	Variables	Inequalities
P^f	(4.2.5)	$\mathcal{O}(T^2)$	$\mathcal{O}(T)$
P^δ	(4.3.10)	$\mathcal{O}(ST)$	$\mathcal{O}(T^2)$
$\hat{P}^\delta$	(4.3.8)	$\mathcal{O}(ST)$	$\mathcal{O}(ST)$
P^δ_{ex}	[Muc66]	$\mathcal{O}(ST)$	$\mathcal{O}(ST)$
P^{yz}	(4.5.4)	$\mathcal{O}(ST)$	$\mathcal{O}(ST)$
$\text{epi}(\text{LCU}^1, \dots, \text{LCU}^T)$	Theorem 2.8	$\mathcal{O}(T)$	$\mathcal{O}(ST)$
P^t_{ex}	[CA06]	$\mathcal{O}(T)$	$\mathcal{O}(ST)$
$\text{epi}(\text{LCU}^\Sigma)$	Theorem 2.36	$\mathcal{O}(T)$	$\mathcal{O}(2^T)$
P^{temp}	Definition 3.17	$\mathcal{O}(T)$	$\mathcal{O}(T^2)$
$\hat{P}^{\text{temp}}$	Definition 3.4	$\mathcal{O}(T)$	$\mathcal{O}(T)$

Table 1.3: All start-up cost formulations considered in this thesis with source and number of variables and inequalities. The formulations are split in groups concerning $\text{epi}(\text{LCU}^t)/\,\text{conv}(\text{epi}(\text{DCU}))$ and $\text{epi}(\text{LCU}^\Sigma)$, ordered by decreasing tightness.

We analyze all relationships between these formulations, as represented by the Venn diagram in Fig. 1.14. For each of the set relationships, we provide an example where the subset is proper. The relationships are listed with a reference to proof and example in Table 1.4.

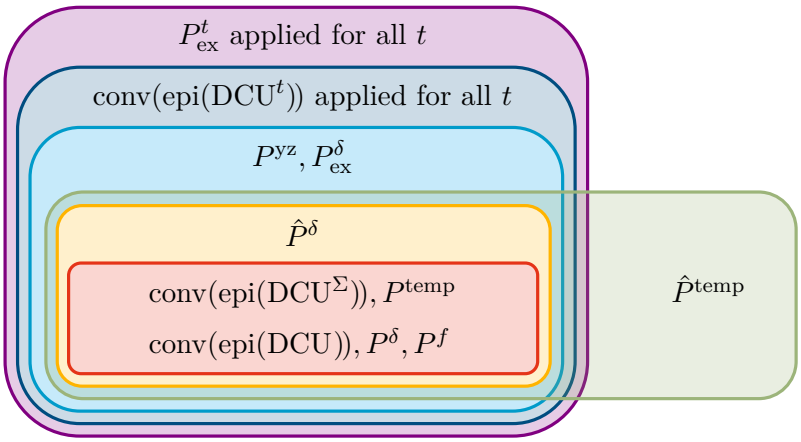


Figure 1.14: Venn diagram of the projections of all formulations to $\text{conv}(\text{epi}(\text{DCU}^\Sigma))$. All subset relationships are strict in general. Note that since $\hat{P}^{\text{temp}}$ is only defined for exponential start-up costs and since in that case $\hat{P}^\delta = P^\delta$, $\hat{P}^{\text{temp}}$ is shown as containing $\hat{P}^\delta$.

Relationship	Source
$\operatorname{conv}(\operatorname{epi}(\mathrm{DCU})) \overset{\text{proj}}{=} P^\delta$	Corollary 4.11
$P^\delta \subsetneq \hat{P}^\delta$	$P^\delta \subsetneq \hat{P}^\delta$ by $P^\delta = \pi^{f\delta}(P^f)$ and Eq. (4.3.9)
$\hat{P}^\delta \subsetneq P^\delta_{\text{ex}}$	$\hat{P}^\delta \subset P^\delta_{\text{ex}}$ by Proposition 4.16, $\hat{P}^\delta \neq P^\delta_{\text{ex}}$ by the example in Fig. 4.7
$P^\delta_{\text{ex}} \overset{\text{proj}}{=} P^{\text{yz}}$	Proposition 4.18
$P^{\text{yz}} \subsetneq \operatorname{epi}(\mathrm{LCU}^1, \dots, \mathrm{LCU}^T)$	$P^{\text{yz}} \subsetneq \operatorname{epi}(\mathrm{LCU}^1, \dots, \mathrm{LCU}^T)$ by Proposition 4.19, $P^{\text{yz}} \neq \operatorname{epi}(\mathrm{LCU}^1, \dots, \mathrm{LCU}^T)$ by the example in Fig. 4.8
$\operatorname{epi}(\mathrm{LCU}^1, \dots, \mathrm{LCU}^T) \subsetneq P^t_{\text{ex}}$	$\operatorname{epi}(\mathrm{LCU}^1, \dots, \mathrm{LCU}^T) \subset P^t_{\text{ex}}$ by Eq. (2.1.3), $\operatorname{epi}(\mathrm{LCU}^1, \dots, \mathrm{LCU}^T) \neq P^t_{\text{ex}}$ by the example in Fig. 2.2
$\operatorname{epi}(\mathrm{LCU}^\Sigma) \overset{\text{proj}}{=} P^{\text{temp}}$	Theorem 3.26
$P^{\text{temp}} \subsetneq \hat{P}^{\text{temp}}$	$P^{\text{temp}} \subset \hat{P}^{\text{temp}}$ by definition, $P^{\text{temp}} \neq \hat{P}^{\text{temp}}$ by Proposition 3.14
$\hat{P}^{\text{temp}} \overset{\text{proj}}{\not\subset} P^t_{\text{ex}}$	(4.5.9)
$P^\delta_{\text{ex}} \overset{\text{proj}}{\not\subset} \hat{P}^{\text{temp}}$	(4.5.10)

Table 1.4: List of relationships between formulations and references to their proofs. The strict relationships $\subsetneq$ hold in general, but may be reduced to $\subset$ in special cases.

Accompanying each of the tightened formulations, we provide a separation algorithm which improves over the naive algorithm by enumeration (see Table 1.5). To the best of our knowledge, not faster separation algorithms were presented for the existing formulations.

Formulation	Inequalities	Naive separation	Separation
P^δ	$\mathcal{O}(ST)$	$\mathcal{O}(S^2T^2)$	$\mathcal{O}(T^2)$ (Theorem 4.13)
$\operatorname{epi}(\mathrm{LCU}^1, \dots, \mathrm{LCU}^T)$	$\mathcal{O}(ST)$	$\mathcal{O}(ST^2)$	$\mathcal{O}(T^2)$ (Proposition 2.9)
$\operatorname{epi}(\mathrm{LCU}^\Sigma)$	$\mathcal{O}(2^T)$	$\mathcal{O}(T2^T)$	$\mathcal{O}(T)$ (Proposition 2.39)
P^{temp}	$\mathcal{O}(T^2)$	$\mathcal{O}(T^2)$	$\mathcal{O}(T)$ (Theorem 3.33)

Table 1.5: Start-up cost formulations with number of inequalities and running times of the separation by enumeration and the separation algorithm provided in this work.

1.5 Acknowledgments

I thank my co-authors Dr. René Brandenberg and Matthias Huber for the enjoyable and effective collaboration on our joint publications [BS14; SHB16; HS15; BHS16], which are part of Chapters 2 and 3 and Section 6.2. I express my sincere thanks to Dr. René Brandenberg for countless hours of fruitful discussion, his experienced guidance, and invaluable comments on drafts.

I am gratefully indebted to my supervisor Prof. Gritzmann for his stimulating comments and insights which have greatly improved this manuscript, and for his trust and support in the past six year. His excellent lectures have sparked my interest in integer optimization.

I express my sincere gratitude to my co-author Matthias Huber and his supervisor Prof. Hamacher. Any impact of this work in the engineering community is owed to their contribution to experiments, modeling, and data fueled by their experience with real-world power modeling. I am grateful to the *International Graduate School for Science and Engineering (IGSSE)* for sponsoring the cooperation between our chairs.

I take this opportunity to thank Statoil ASA for their collaboration in a two-year joint research project, and in particular my project colleagues Patrick Blomquist, Paul Andrew Holtom, Øystein Klokk, and Christophe Roth Maingourd for sharing their insight into power markets.

I would like to thank my colleagues at the faculty of mathematics for creating an open and inspiring atmosphere. I am especially grateful to Paul Stursberg, Dr. Philipp Ahlhaus, and Dr. Felix Schmiedl for many motivating exchanges on our research.

Finally, I thank my wife Jasmin for her unending patience and support, my parents and my brother for their support and encouragement, and my sons for putting everything into perspective.

Chapter 2

$\mathcal{H}$-Representations of the Epigraphs of Start-up Cost Functions

The volatility of renewable electricity production forces thermal units to start-up more often. This is true not only for gas units, which traditionally have the task to cover peak load, but also for coal units which were designed to supply base load. In the light of the increasing penetration of renewables, it is necessary to model start-up costs more accurately.

This chapter is dedicated to start-up cost functions of a single unit $i \in \mathcal{I}$. In the context of the Unit Commitment problem, start-up costs need to be minimized. We therefore analyze the epigraphs of the functions of

- the start-up costs incurred in an individual period (Section 2.1, joint work with René Brandenberg and Matthias Huber partially published in [BS14; SHB16]),
- the summed start-up costs of all modeled periods (Section 2.2, joint work with René Brandenberg and Matthias Huber published in [BHS16]), and
- the function of all individual start-up costs (Section 2.3).

One of the prevalent models in the literature was presented in [CA06], and based on a piece-wise linear description of the start-up costs proposed in [NR00]. Building on this formulation, we derive a linear-size $\mathcal{H}$-representation of each of these individual costs, thereby improving our understanding of the start-up cost function (Section 2.1).

However, the individual start-up costs are typically used to derive the sum of the start-up costs, which is then minimized as part of the objective function. By considering these summed costs, the integrality gap can be reduced substantially. In Section 2.2, we present a generally irredundant, exponential-size $\mathcal{H}$-representation of the epigraph of this summed start-up cost function and give a linear, exact separation algorithm. This result provides the theoretical framework for the *temperature model* in Chapter 3, which is equivalent to the epigraph of the summed start-up costs and relies on its separation algorithm.

Section 2.3 focuses on the epigraph of the function of all individual start-up costs, for which it identifies an exponential-size class of facet-inducing inequalities.

2.1 The Start-up Costs in a Single Period

This section presents joint work with René Brandenberg and Matthias Huber [BS14; SHB16]. We consider the *discrete start-up cost function* DCU^t, defined in (1.3.6) as the start-up costs incurred in period t of a given operational schedule v,

$$\forall\, t \in [T]: \quad \mathrm{DCU}^t: \ \{0,1\}^T \to \mathbb{R}_{\geq 0}, \ \ v \mapsto \begin{cases} \mathrm{CU}^t(\mathrm{ol}^t(v)) & \text{if } v^t = 1, \\ 0 & \text{else,} \end{cases} \tag{2.1.1}$$

where $\mathrm{ol}^t(v)$ denotes the number of offline periods prior to period t,

$$\mathrm{ol}^t(v) := \max\{l \in [t-1] \mid v^{t-1} = \ldots = v^{t-l} = 0\}.$$

By tightening an existing formulation, we gain a generally irredundant, linear-size $\mathcal{H}$-representation of the convex hull of its epigraph, and an exact separation algorithm with time complexity $\mathcal{O}(t)$.

2.1.1 The Existing Step-wise Start-up Cost Model

Based on a piece-wise linear representation of the start-up cost function presented in [NR00], [CA06] introduces a formulation to model start-up costs which depend on the preceding downtime. The necessary inequalities emerge naturally when explicitly modeling the epigraph of DCU.

As seen in the introduction, in a minimization problem the start-up costs of unit i can be modeled as

$$\forall\, t \in [T]: \qquad (v, \mathrm{cu}^t) \in \mathrm{epi}(\mathrm{DCU}^t).$$

This constraint on v and cu^t is equivalent to the nonlinear inequality

$$\forall\, t \in [T]: \qquad \mathrm{cu}^t \geq \mathrm{DCU}^t(v),$$

which by definition is again equivalent to

$$\forall\, t \in [T] \text{ with } v^t = 1: \quad \mathrm{cu}^t \geq \mathrm{CU}^t(\mathrm{ol}^t(v)).$$

Since the start-up cost coefficients $\mathrm{CU}^t(l)$ are increasing, the bound on cu^t remains unchanged if offline lengths $l \leq \mathrm{ol}^t(v)$ are considered as well,

$$\forall\, t \in [T], l \in [0\,..\,t-1] \text{ with } v^t = 1, v^{t-l} = \ldots = v^{t-1} = 0: \quad \mathrm{cu}^t \geq \mathrm{CU}^t(l).$$

Finally, by canonically rewriting the binary condition on the operational schedule v, we get the generalized constraints as presented in [CA06],

$$\forall\, t \in [T],\, l \in [0\,..\,t-1]: \qquad \mathrm{cu}^t \geq \mathrm{CU}^t(l)\Big(v^t - \sum_{j=1}^{l} v^{t-j}\Big). \tag{2.1.2}$$

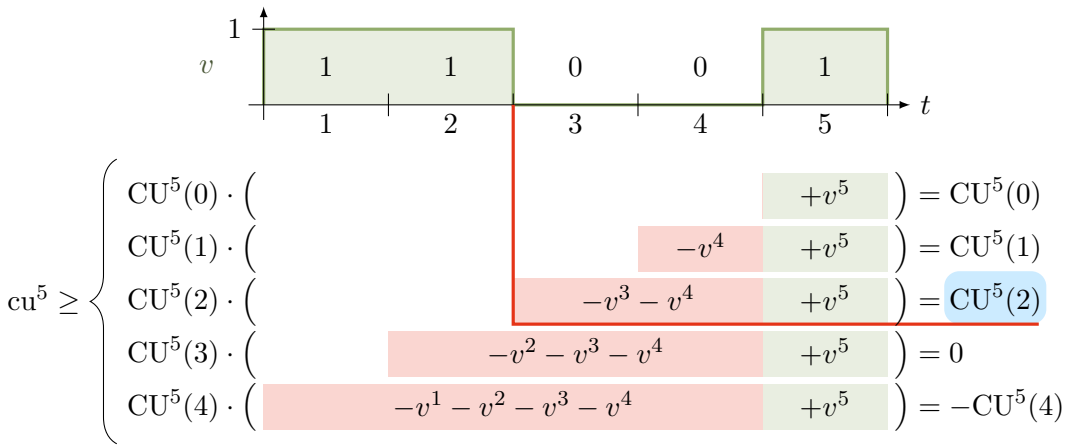


Figure 2.1: Start-up cost inequalities proposed in [CA06] and based on [NR00]. The third constraint provides the correct bound on the start-up costs cu^5 (marked blue). The bounds by the first two constraints are lower due to their lower start-up cost coefficients $\mathrm{CU}^5(1)$ and $\mathrm{CU}^5(2)$. Since $v^2 = 1$, the sums marked in red that include "$-v^2$" are greater than or equal v^5, and hence the bounds by the last two constraints are non-positive.

The expression in parenthesis, $v^t - \sum_{j=1}^{l} v^{t-j}$, is 1 iff the considered unit starts up in period t after being offline in periods $[t-l \,..\, t-1]$. Otherwise, the term is less or equal to 0. Figure 2.1 visualizes the idea behind these constraints.

So, the *existing step-wise start-up cost polyhedron* P^t_{ex},

$$P^t_{\mathrm{ex}} := \Big\{(v, \mathrm{cu}) \in [0,1]^T \times \mathbb{R}^T \;\Big|\; (v, \mathrm{cu}^t) \text{ fulfills (2.1.2) for } t\Big\},$$

correctly models the epigraph of the start-up cost function DCU^t for integral $v \in \{0,1\}^T$,

$$\mathrm{epi}(\mathrm{DCU}^t) = P^t_{\mathrm{ex}} \cap (\mathbb{Z}^T \times \mathbb{R}) = \Big\{(v, \mathrm{cu}) \in \{0,1\}^T \times \mathbb{R}^T \;\Big|\; (v, \mathrm{cu}^t) \text{ fulfills (2.1.2) for } t\Big\}.$$

If we relax the integrality of v, this is no longer true in general. More precisely, if the start-up cost function DCU^t is not constant, then the convex extension LCU^t of DCU^t is not described accurately,

$$\begin{aligned} &\forall\, t \in [T] \text{ with } \mathrm{DCU}^t(t-1) > \mathrm{DCU}^t(1) > 0 : \\ &\qquad \mathrm{epi}(\mathrm{LCU}^t) = \mathrm{conv}(\mathrm{epi}(\mathrm{DCU}^t)) \subsetneq P^t_{\mathrm{ex}}. \end{aligned} \tag{2.1.3}$$

For a given period $t \in [T]$ with $\mathrm{DCU}^t(t-1) > \mathrm{DCU}^t(1) > 0$, which implies $t \geq 3$, we can see this by considering the point $v \in [0,1]^T$ (cf. Fig. 2.2) with

$$v := (0, \ldots, \underbrace{1/2}_{t-2}, 0, \underbrace{1}_{t}, 0, \ldots, 0)$$

which is uniquely representable as the convex combination $1/2(v' + v'')$ with points

$$v' := (0, \dots, \ 0 \ , 0, \ 1 \ , 0, \dots, 0)$$
$$v'' := (0, \dots, \underbrace{1}_{t-2}, 0, \underbrace{1}_{t}, 0, \dots, 0)$$

Thus, the smallest value cu^t such that $(v, \mathrm{cu}^t) \in \mathrm{conv}(\mathrm{epi}(\mathrm{DCU}^t))$ is

$$\mathrm{cu}^t = \frac{1}{2}(\mathrm{CU}(t-1) + \mathrm{CU}(1)) > \max\left\{\mathrm{CU}(1), \frac{1}{2}\mathrm{CU}(t-1)\right\},$$

but on the other hand $(v, \max\{\mathrm{CU}(1), \frac{1}{2}\mathrm{CU}(t-1)\})$ fulfills (2.1.2) for each $t \in [T]$.

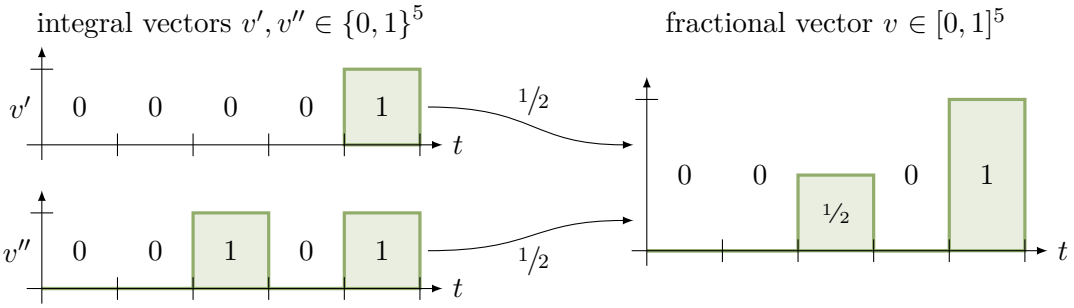


Figure 2.2: Convex combination $v = 1/2v' + 1/2v''$ of $v' = (0,0,0,0,1)$ and $v'' = (0,0,1,0,1)$, leading to $\mathrm{LCU}^5(v) = 1/2\mathrm{DCU}^5(v') + 1/2\mathrm{DCU}^5(v'') = 1/2\mathrm{CU}^5(4) + 1/2\mathrm{CU}^5(1)$. P^t_{ex} only models a start-up cost of $\max\{1/2\mathrm{CU}^5(1), 1/2\mathrm{CU}^5(1)\}$.

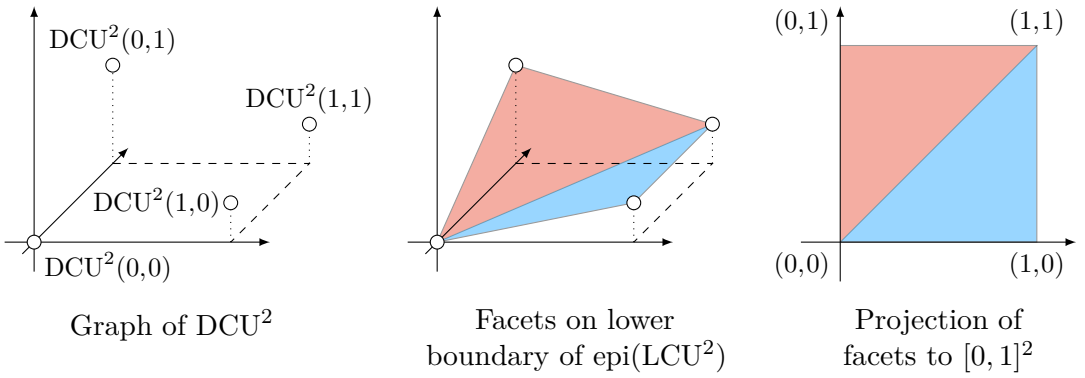
In the following, we tighten the model P^t_{ex} to an $\mathcal{H}$-representation of $\mathrm{epi}(\mathrm{LCU}^t)$.

2.1.2 A Geometric Interpretation

Our goal of finding an $\mathcal{H}$-representation of $\mathrm{epi}(\mathrm{LCU}^t)$ amounts to deducing the facets on the lower boundary of $\mathrm{epi}(\mathrm{LCU}^t)$. If we project these facets to the convex hull $[0,1]^T$ of the domain of DCU^t, we gain a partition of $[0,1]^T$ which fully describes the facets (see Fig. 2.3).

This observation reveals a geometric interpretation of the subsequent results in Subsection 2.1.4, which lead to an $\mathcal{H}$-representation of $\mathrm{epi}(\mathrm{LCU}^t)$:

1. Proposition 2.2 lists the vertices $v \in \{0,1\}^T$ contained in the projection to $[0,1]^T$ of each facet.
2. Corollary 2.3 expands on the above proposition by describing a subset of the edges of each facet.
3. Lemma 2.7 finally gives the facet appertaining to each v in $[0,1]^T$.



Graph of DCU^2 | Facets on lower boundary of $\mathrm{epi}(\mathrm{LCU}^2)$ | Projection of facets to $[0,1]^2$

Figure 2.3: By projecting the facets of the convex hull of the epigraph of $\mathrm{DCU}^2 : \{0,1\}^2 \to \mathbb{R}$ to the convex hull of its domain $[0,1]^2$, we gain a partition of $[0,1]^2$.

For a function $f\colon \{0,1\}^T \to \mathbb{R}$, one would expect that in general the partition of $[0,1]^T$ induced by the facets of $\mathrm{conv}(\mathrm{epi}(f))$ consists of simplices. A simplex contained in $[0,1]^T$ has a volume of at most $\frac{n^{n/2}}{n!2^n}$ (see e. g. [GK94]): therefore $\Omega(2^n)$ simplices are required to partition $[0,1]^T$. Hence, $\mathrm{conv}(\mathrm{epi}(f))$ is expected to have at least an exponential number of facets.

As we show in the following however, $\mathrm{epi}(\mathrm{LCU}^t)$ in general has only T facets on its lower boundary, which are induced by T inequalities with parameters $l \in [0\,..\,T-1]$ (see (2.1.2) and Theorem 2.8). This is explained by the structure of DCU, which results in partitions of $[0,1]^T$ where the volume of the subset corresponding to a given parameter $l \in \{0,\dots,t-1\}$ is independent of the dimension $T \geq l+2$ (see Fig. 2.4).

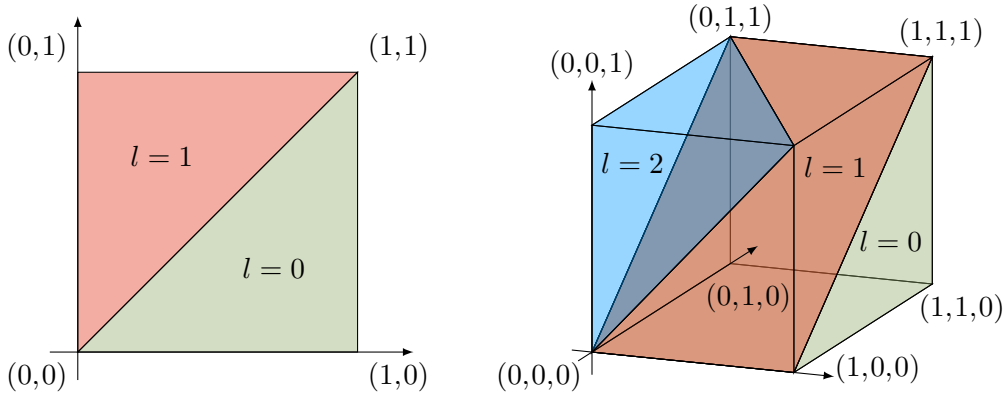


Figure 2.4: Partition of the cube $[0,1]^T$ corresponding to the facets of $\mathrm{epi}(\mathrm{LCU}^t)$ for parameters $t = T = 2$ (left) and $t = T = 3$ (right). Note that the volume of the facet for $l = 0$ (green) equals $\frac{1}{2}$ for both $T = 2$ and $T = 3$. Furthermore, the facet for $l = T-1$ is a simplex with orthogonal corner $(0,\dots,0,1)$ in both cases.

The reason hereof is that the value $\mathrm{DCU}^t(v)$ only depends a subset of the coordinates of $v \in \{0,1\}^T$. If we have $v^t = v^{t-l-1} = 1$ for some $l \in [0 \,..\, t-2]$, then all $\tilde{v} \in \{0,1\}^T$ with

$$\forall\, t' \in [t-l-1 \,..\, t]: \quad \tilde{v}^{t'} = v^{t'}$$

have the same start-up costs, $\mathrm{DCU}^t(v) = \mathrm{DCU}^t(\tilde{v})$.

Hence, the number of vertices $(v, \mathrm{DCU}^t(v))$ with the same costs $\mathrm{DCU}^t(v) = \mathrm{CU}^t(l)$ increases exponentially with $T - l$. For example for parameters $T = 8$, $t = 7$, and $l = 3$, Fig. 2.5 lists 8 schedules with the same start-up costs which, as we see in Proposition 2.2, lie on the same facet.

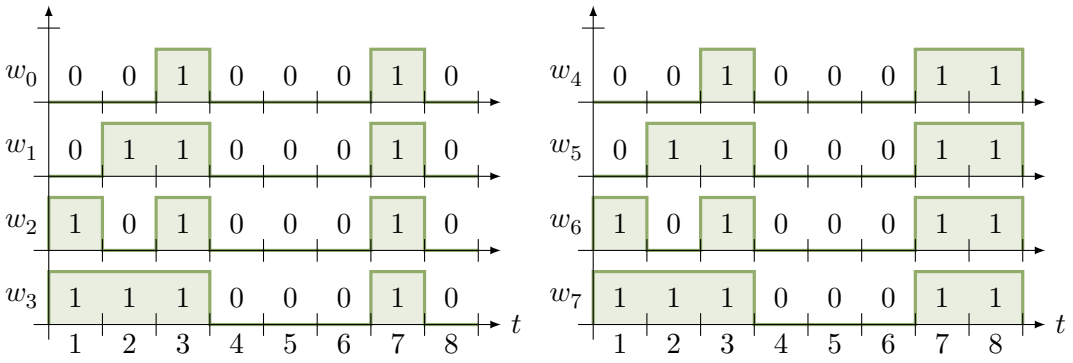


Figure 2.5: Operational schedules w_j of vertices $(w_j, \mathrm{DCU}^t(w_j))$ with the same start-up costs $\mathrm{DCU}^7(w_j) = \mathrm{CU}^7(3)$ and lie on the same facet.

2.1.3 Lifted Start-up Cost Inequalities

By rewriting the constraints (2.1.2) of the existing start-up cost model as

$$\forall\, t \in [T],\, l \in [0 \,..\, t-1]: \quad \mathrm{cu}^t \geq \mathrm{CU}^t(l) v^t - \sum_{j=1}^{l} \mathrm{CU}^t(l) v^{t-j},$$

an alternative interpretation suggests itself. The right-hand side of each of these constraints reaches its maximum for $v^t = 1$ and $v^{t-1} = \ldots = v^{t-l} = 0$, i. e. if the unit starts up in period t after at least l offline periods. If the unit is online in a single period $t - j$ with $j \in [l]$, the start-up costs are "taken back", i. e. the right-hand side is reduced to 0 or less.

This is evident in Fig. 2.1, where the right-hand side of the constraint for $l = 3$ is 0 since $v^2 = 1$. However, the unit does incur start-up costs, such that the constraint would still be valid with a right-hand side of $\mathrm{CU}^5(2)$. Hence, the coefficient of v^2 could

be lifted, i. e. increased such that the resulting constraint dominates the original one and remains valid. This idea leads to the inequalities

$$\forall\, t \in [T],\, l \in [0\,..\,t-1]: \quad \mathrm{cu}^t \geq \mathrm{CU}^t(l)v^t - \sum_{j=1}^{l} \Big(\mathrm{CU}^t(l) - \mathrm{CU}^t(j-1)\Big)v^{t-j}, \quad (2.1.4)$$

which we call *lifted start-up cost inequalities*. In the remainder of this section, we prove that they are valid and induce all non-trivial facets of the epigraph of the start-up cost functions LCU^t.

Proposition 2.1 *For each $t \in [T]$, the lifted start-up cost inequality for t is valid for* $\mathrm{epi}(\mathrm{LCU}^t)$.

Proof. Since $\mathrm{epi}(\mathrm{LCU}^t) = \mathrm{conv}(\mathrm{epi}(\mathrm{DCU}^t))$, it suffices to show that the inequality is valid for each point $(v, \mathrm{cu}^t) \in \mathrm{epi}(\mathrm{DCU}^t)$. If $v^t = 0$, no start-up costs are incurred in period t. Thus,

$$\mathrm{cu}^t \geq \mathrm{DCU}^t(v^t) = 0 = \underbrace{\mathrm{CU}^t(l)v^t}_{=0} - \sum_{j=1}^{l} \underbrace{\big(\mathrm{CU}^t(l) - \mathrm{CU}^t(j-1)\big)}_{\geq 0} v^{t-j}.$$

If $v^t = 1$, then the unit was offline for the preceding $\mathrm{ol}^t(v)$ periods, and the start-up costs by definition equals $\mathrm{CU}^t(\mathrm{ol}^t(v))$. Since $\mathrm{CU}^t(l)$ increases with l, for each $l \in [0\,..\,\mathrm{ol}^t(v)]$ we have

$$\mathrm{cu}^t \geq \mathrm{DCU}^t(v) = \mathrm{CU}^t(\mathrm{ol}^t(v)) \geq \mathrm{CU}^t(l)v^t - \sum_{j=1}^{l} \Big(\mathrm{CU}^t(l) - \mathrm{CU}^t(j-1)\Big) \underbrace{v^{t-j}}_{=0},$$

and for each $l \in [\mathrm{ol}^t(v)+1\,..\,t-1]$,

$$\begin{aligned}\mathrm{cu}^t &\geq \mathrm{DCU}^t(v) = \mathrm{CU}^t(\mathrm{ol}^t(v)) = \mathrm{CU}^t(l)v^t - \Big(\mathrm{CU}^t(l) - \mathrm{CU}^t(\mathrm{ol}^t(v))\Big)\underbrace{v^{t-\mathrm{ol}^t(v)-1}}_{=1} \\ &\geq \mathrm{CU}^t(l)v^t - \sum_{j=1}^{l} \Big(\mathrm{CU}^t(l) - \mathrm{CU}^t(j-1)\Big)v^{t-j}.\end{aligned}$$

□

By considering the inequalities in the last proof, we can readily identify the points which fulfill a certain lifted start-up cost inequality with equality, and thus lie on the face of $\mathrm{epi}(\mathrm{LCU}^t)$ induced by this inequality. We start with the vertices of $\mathrm{epi}(\mathrm{LCU}^t)$ in the following proposition.

Proposition 2.2 *For each $t \in [T]$, $l \in [0\,..\,t-1]$, a vertex $(v, \mathrm{DCU}^t(v))$ of* $\mathrm{epi}(\mathrm{DCU}^t)$ *fulfills the lifted start-up cost inequality for t and l with equality, if one of these cases applies:*

a) $v^t = 0$ *and* $v^{t-1} = \ldots = v^{t-l} = 0$.

b) $v^t = 1$ *and* $\exists\, \mathrm{ol}^t \in [0\,..\,l-1]$ *with* $v^{t-j} = \begin{cases} 1 & \text{if } j = \mathrm{ol}^t + 1 \\ 0 & \text{else} \end{cases}$ *for* $j \in [l]$.

c) $v^t = 1$, $v^{t-1} = \ldots = v^{t-l} = 0$, *and* $l = t-1$ *or* $v^{t-l-1} = 1$.

Proof. For each vertex $(v, \mathrm{DCU}^t(v))$ of $\mathrm{epi}(\mathrm{DCU}^t)$ to which on of the above cases applies, it holds that

- $\mathrm{DCU}^t(v) = 0$ if the point matches case a),
- $\mathrm{DCU}^t(v) = \mathrm{CU}^t(\mathrm{ol}^t)$ if case b) applies, and
- $\mathrm{DCU}^t(v) = \mathrm{CU}^t(l)$ if the point corresponds to case c).

If the point $(v, \mathrm{cu}^t) \in \mathrm{epi}(\mathrm{DCU}^t)$ matches cases a) and c), substituting its values of $v^{t-l}, \ldots, v^{t-1}$ into the right-hand side of the lifted start-up cost inequality results in $\mathrm{cu}^t \geq \mathrm{CU}^t(l)v^t$, which is fulfilled with equality.

If the point $(v, \mathrm{cu}^t) \in \mathrm{epi}(\mathrm{DCU}^t)$ matches case b), substituting gives

$$\mathrm{CU}^t(l)v^t - \Big(\mathrm{CU}^t(l) - \mathrm{CU}^t(\mathrm{ol}^t)\Big)v^{t-\mathrm{ol}^t-1} = \mathrm{CU}^t(\mathrm{ol}^t) = \mathrm{cu}^t. \qquad \square$$

The integrality of the coordinates $v^{t'}$ with $t' \in [t-l-1] \cup [t+1 \,..\, T]$ is not necessary for the correctness of the last proof. We can therefore extend the last result to describe certain edges of the facets In fact, the convex combinations of all vertices of $\mathrm{epi}(\mathrm{DCU}^t)$ which correspond to the same case in Proposition 2.2 and have the same start-up costs spans all such partially fractional points. Naturally, these points still lie on the same face of $\mathrm{epi}(\mathrm{LCU}^t)$.

Corollary 2.3 *For each $t \in [T]$, $l \in [0 \,..\, t-1]$, a fractional point $(v, \mathrm{cu}^t) \in [0,1]^T \times \mathbb{R}$ lies on the face of $\mathrm{epi}(\mathrm{LCU}^t)$ induced by the start-up cost inequality for t and T, if one of these cases applies:*

a) $v^t = 0$, $v^{t-1} = \ldots = v^{t-l} = 0$ *and* $\mathrm{cu}^t = 0$.

b) $v^t = 1$, $\exists\, \mathrm{ol}^t \in [0 \,..\, l-1]$ *with* $v^{t-j} = \begin{cases} 1 & \text{if } j = \mathrm{ol}^t + 1 \\ 0 & \text{else} \end{cases}$ *for* $j \in [l]$, $\mathrm{cu}^t = \mathrm{CU}^t(\mathrm{ol}^t)$.

c) $v^t = 1$, $v^{t-1} = \ldots = v^{t-l} = 0$, $l = t-1$ *or* $v^{t-l-1} = 1$, *and* $\mathrm{cu}^t = \mathrm{CU}^t(l)$.

Of course, all convex combinations of these points again lie on the same face.

The requirements laid out in Proposition 2.2 can be visualized in form of a matrix, such that a point $(v, \mathrm{cu}^t) \in [0,1]^T \times \mathbb{R}$ fulfills the requirements iff they match a row of this matrix (cf. Fig. 2.6). Stars match any value.

It is easy to check that the set of all points $(v, \mathrm{cu}^t) \in [0,1]^T \times \mathbb{R}$ has an affine dimension of T.

Corollary 2.4 *For each $t \in [T]$, $l \in [0 \,..\, t-1]$, the lifted start-up cost inequality for t and l induces a facet of* $\mathrm{epi}(\mathrm{LCU}^t)$.

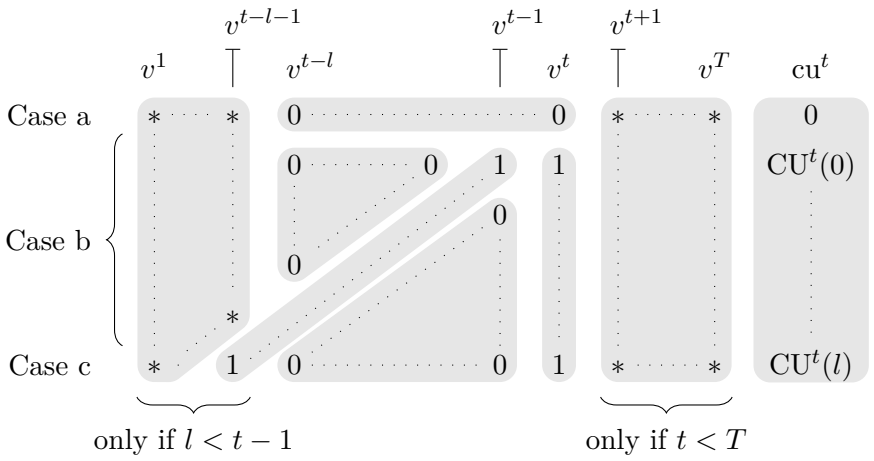


Figure 2.6: Patterns of the points fulfilling the lifted start-up cost inequality for t and l with equality.

2.1.4 An $\mathcal{H}$-Representation

Having proved that the lifted start-up cost inequalities induce facets, we proceed by showing that these inequalities (and the trivial inequality $0 \leq v^t \leq 1$) already suffice to separate all points outside $\mathrm{epi}(\mathrm{LCU}^t)$ from $\mathrm{epi}(\mathrm{LCU}^t)$. To this end, we bring the right-hand sides of these inequalities into relation.

Proposition 2.5 *For each $t \in [T]$ and $l \in [t-1]$, it holds that*

$$rhs^t(l) = rhs^t(l-1) + \Big(\mathrm{CU}^t(l) - \mathrm{CU}^t(l-1)\Big)\left(v^t - \sum_{j=1}^{l} v^{t-j}\right),$$

where $rhs^t(l)$ denotes the right-hand side of the lifted start-up cost inequality with parameter l.

Proof. For each $t \in [T]$ and $l \in [t-1]$, the right-hand side of the respective lifted start-up cost inequality can be reorganized to

$$\mathrm{rhs}^t(l) = \mathrm{CU}^t(l)\left(v^t - \sum_{j=1}^{l} v^{t-j}\right) + \sum_{j=1}^{l} \mathrm{CU}^t(j-1)v^{t-j}.$$

The claimed equality holds since

$$\begin{aligned}
\mathrm{rhs}^t(l) &= \mathrm{CU}^t(l)\left(v^t - \sum_{j=1}^{l-1} v^{t-j}\right) + \sum_{j=1}^{l-1} \mathrm{CU}^t(j-1)v^{t-j} - \Big(\mathrm{CU}^t(l) - \mathrm{CU}^t(l-1)\Big)v^{t-l} \\
&= \left.\begin{aligned}&\mathrm{CU}^t(l-1)\left(v^t - \sum_{j=1}^{l-1} v^{t-j}\right) + \sum_{j=1}^{l-1} \mathrm{CU}^t(j-1)v^{t-j} \\ &\qquad + \Big(\mathrm{CU}^t(l) - \mathrm{CU}^t(l-1)\Big)\left(v^t - \sum_{j=1}^{l-1} v^{t-j}\right)\end{aligned}\right\} \mathrm{rhs}^t(l-1) \\
&\qquad - \Big(\mathrm{CU}^t(l) - \mathrm{CU}^t(l-1)\Big)v^{t-l} \\
&= \mathrm{rhs}^t(l-1) + \Big(\mathrm{CU}^t(l) - \mathrm{CU}^t(l-1)\Big)\left(v^t - \sum_{j=1}^{l} v^{t-j}\right). \qquad \square
\end{aligned}$$

Starting from $\mathrm{rhs}^t(0) = \mathrm{CU}^t(0)v^t$, this recursive relation of $\mathrm{rhs}^t(l)$ leads to an alternative formulation of the lifted start-up cost inequalities.

Corollary 2.6 *For each $t \in [T]$, $l \in [0 \,..\, t-1]$, the lifted start-up cost inequality* (2.1.4) *equals*

$$\mathrm{cu}^t \geq \sum_{j=0}^{l} \Big(\mathrm{CU}^t(j) - \mathrm{CU}^t(j-1)\Big)\left(v^t - \sum_{k=1}^{j} v^{t-k}\right), \tag{2.1.5}$$

defining $\mathrm{CU}^t(-1) = 0$ for convenience.

We use the relationship of the right-hand sides to analyze how the bound on the start-up costs for a certain vector $v \in [0,1]^T$ changes in dependence of l. Since the start-up cost function CU^t is increasing, the term $\mathrm{CU}^t(l) - \mathrm{CU}^t(l-1)$ is non-negative for all $l \in [0 \,..\, t-1]$. The sign of the change of the right-hand side for increasing l thus depends solely on the term $v^t - \sum_{j=1}^{l} v^{t-j}$,

$$\begin{aligned}
\mathrm{rhs}^t(l) &= \mathrm{rhs}^t(l-1) + \Big(\mathrm{CU}^t(l) - \mathrm{CU}^t(l-1)\Big)\left(v^t - \sum_{j=1}^{l} v^{t-j}\right) \\
&\begin{cases}
\geq \mathrm{rhs}^t(l-1) & \text{if } v^t - \sum_{j=1}^{l} v^{t-j} \geq 0, \\
\leq \mathrm{rhs}^t(l-1) & \text{if } v^t - \sum_{j=1}^{l} v^{t-j} \leq 0.
\end{cases}
\end{aligned}$$

The term $v^t - \sum_{j=1}^{l} v^{t-j}$ decreases with increasing l. So, there exists a pivotal value of l until which the right-hand sides increase, and from which on the right-hand sides decrease (not necessarily strictly). The lifted start-up cost inequality with this pivotal l thus dominates the other inequalities for this particular $v \in [0,1]^T$. This idea leads to the following lemma.

Lemma 2.7 *For each $t \in [T]$ and $(v, \mathrm{cu}^t) \in [0,1]^T \times \mathbb{R}$, choose $\overline{\mathrm{ol}} \in [0 \,..\, t-1]$ s.t.*

$$\sum_{j=1}^{\overline{\mathrm{ol}}} v^{t-j} \leq v^t \leq \sum_{j=1}^{\overline{\mathrm{ol}}+1} v^{t-j}, \tag{2.1.6}$$

defining $v^0 = 1$ for notational convenience. Then (v, cu^t) lies in $\mathrm{epi}(\mathrm{LCU}^t)$ iff the lifted start-up cost inequality for t and $\overline{\mathrm{ol}}$ is fulfilled.

Proof. Before starting with the main proof, it should be noted that such a length $\overline{\mathrm{ol}}$ exists, but is not necessarily unique. The variable v^0 is defined such that the condition

$$v^t \leq \sum_{j=1}^{\overline{\mathrm{ol}}+1} v^{t-j}$$

is certainly true for $\overline{\mathrm{ol}} = t-1$. At the same time, the condition

$$\sum_{j=1}^{\overline{\mathrm{ol}}} v^{t-j} \leq v^t \tag{2.1.7}$$

is trivially fulfilled for $\overline{\mathrm{ol}} = 0$. So, the last $\overline{\mathrm{ol}} \in [0 \,..\, t-1]$ for which (2.1.7) holds must also fulfill (2.1.6) and is thus a feasible choice.

We prove the statement by determining the unique $\tilde{\mathrm{cu}}^t \in \mathbb{R}$ such that $(v, \tilde{\mathrm{cu}}^t)$ lies on the boundary of $\mathrm{epi}(\mathrm{LCU}^t)$, and comparing cu^t and $\tilde{\mathrm{cu}}^t$. To this end we decompose $(v, \tilde{\mathrm{cu}}^t)$ into a convex combination of points $(w_j, \mathrm{DCU}^t(w_j))$

- lying on the facet of $\mathrm{epi}(\mathrm{LCU}^t)$ induced by the lifted start-up cost inequality for t and $\overline{\mathrm{ol}}$ and
- with start-up costs known from Corollary 2.3.

The idea of this decomposition is visualized in Fig. 2.7.

For $\overline{\mathrm{ol}} < t-1$, we require the technical coefficient μ fulfilling

$$\mu\Big(\sum_{j=1}^{\overline{\mathrm{ol}}} v^{t-j} - v^t + 1\Big) + v^t - \sum_{j=1}^{\overline{\mathrm{ol}}+1} v^{t-j} = v^{t-\overline{\mathrm{ol}}-1}.$$

If $v^t = 1$ and $\sum_{j=1}^{\overline{\mathrm{ol}}} v^{t-j} = 0$, this is trivially fulfilled by $\mu := 0$, and else by

$$\mu := \frac{\sum_{j=1}^{\overline{\mathrm{ol}}+1} v^{t-j} - v^t}{\sum_{j=1}^{\overline{\mathrm{ol}}} v^{t-j} - v^t + 1}.$$

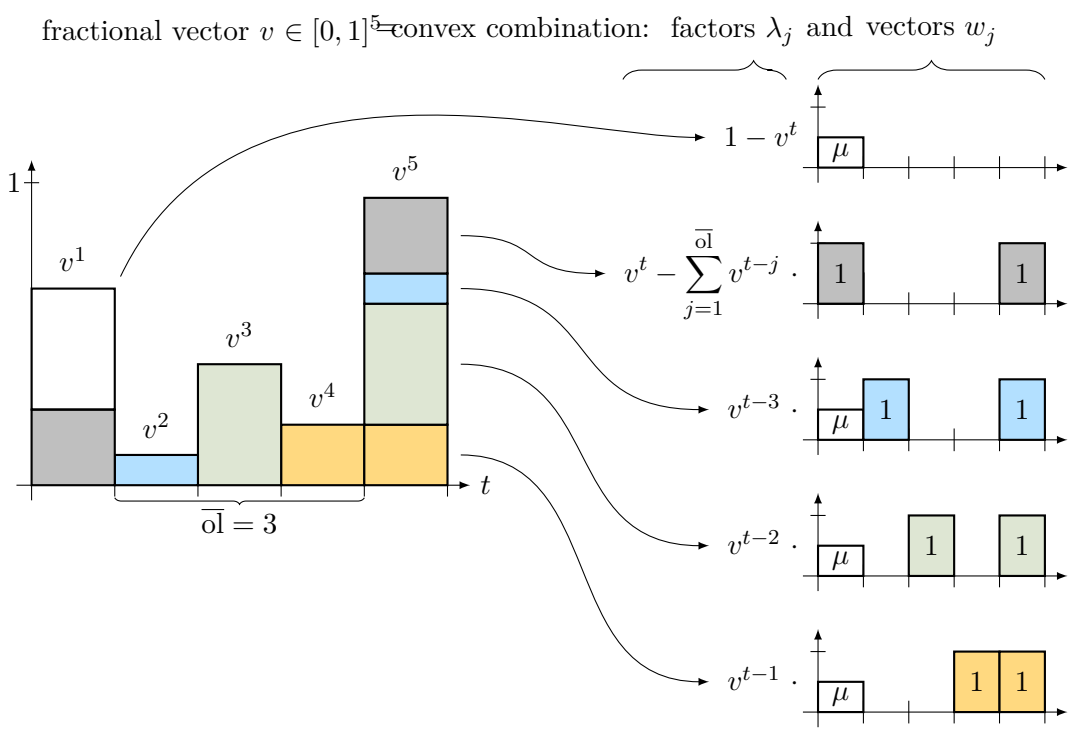


Figure 2.7: Decomposition of a fractional vector $v \in [0,1]^5$ into a convex combination of vectors w_j known to lie on a single facet (with appropriate start-up costs cu^t). The facet is induced by the lifted start-up cost inequality with $t = 5$ and $\overline{\text{ol}} = 3$. The term μ is a technical term as stated in the proof of Lemma 2.7.

Because of (2.1.6), μ lies in $[0,1]$.

Using this coefficient, define the vectors $w_0, \ldots, w_{\overline{\text{ol}}+1} \in [0,1]^T$ as

$$w_0^{l'} = \begin{cases} 0 & \text{if } l' \in [t-\overline{\text{ol}} \,..\, t], \\ \mu & \text{if } l' = t - \overline{\text{ol}} - 1, \\ v^{l'} & \text{else.} \end{cases}$$

$$w_{\overline{\text{ol}}+1}^{l'} = \begin{cases} 0 & \text{if } l' \in [t-\overline{\text{ol}} \,..\, t-1], \\ 1 & \text{if } l' \in \{t, t - \overline{\text{ol}} - 1\}, \\ v^{l'} & \text{else.} \end{cases}$$

$$\forall\, j \in [\overline{\text{ol}}]: \quad w_j^{l'} = \begin{cases} 0 & \text{if } l' \in [t-\overline{\text{ol}} \,..\, t-1] \setminus \{t-j\}, \\ 1 & \text{if } l' \in \{t, t-j\}, \\ \mu & \text{if } l' = t - \overline{\text{ol}} - 1, \\ w^{l'} & \text{else.} \end{cases}$$

The respective convex combination coefficients are $\lambda_0, \ldots, \lambda_{\overline{\text{ol}}+1}$ with

$$\lambda_0 := 1 - v^t, \qquad \lambda_j := v^{t-j}, \qquad \lambda_{\overline{\text{ol}}+1} := v^t - \sum_{j=1}^{\overline{\text{ol}}} v^{t-j}.$$

It is easy to verify that the coefficients λ_j lie in $[0,1]$ (using (2.1.6)) and sum up to 1.

We prove that v is the convex combination of the above vectors with the above coefficients by coordinate-wise comparison, again following the idea shown in Fig. 2.7:

$$\forall\, t' \in [T] \setminus [t-\overline{\mathrm{ol}}-1 \,..\, t]: \quad \sum_{j=0}^{\overline{\mathrm{ol}}+1} \lambda_j w_j^{t'} = \sum_{j=0}^{\overline{\mathrm{ol}}+1} \lambda_j v^{t'} = v^{t'},$$

$$\forall\, t' \in [t-\overline{\mathrm{ol}} \,..\, t-1]: \quad \sum_{j=0}^{\overline{\mathrm{ol}}+1} \lambda_j w_j^{t'} = \lambda_{t-t'} = v^{t'}$$

$$\sum_{j=0}^{\overline{\mathrm{ol}}+1} \lambda_j w_j^{t} = \sum_{j=1}^{\overline{\mathrm{ol}}+1} \lambda_j w_j^{t} = \sum_{j=1}^{\overline{\mathrm{ol}}+1} \lambda_j = v^t,$$

$$\sum_{j=0}^{\overline{\mathrm{ol}}+1} \lambda_j w_j^{t-\overline{\mathrm{ol}}-1} = \mu \sum_{j=0}^{\overline{\mathrm{ol}}} \lambda_j + \lambda_{\overline{\mathrm{ol}}+1} = \mu\Big(\sum_{j=1}^{\overline{\mathrm{ol}}} v^{t-j} - v^t + 1\Big) + v^t - \sum_{j=1}^{\overline{\mathrm{ol}}} v^{t-j} = v^{t-\overline{\mathrm{ol}}-1}.$$

The start-up costs for these vectors are $\mathrm{DCU}^t(w_0) = 0$ and $\mathrm{DCU}^t(w_j) = \mathrm{CU}^t(j-1)$ for all $j \in [\overline{\mathrm{ol}}+1]$. So, the analogous convex combination of the start-up costs equals

$$\tilde{\mathrm{cu}}^t = \sum_{j=1}^{\overline{\mathrm{ol}}+1} \lambda_j \mathrm{CU}^t(j-1).$$

By Corollary 2.3 the points $(w_0, 0)$, $(w_1, \mathrm{CU}^t(0))$, …, $(w_{\overline{\mathrm{ol}}+1}, \mathrm{CU}^t(\overline{\mathrm{ol}}))$ lie on the facet of $\mathrm{epi}(\mathrm{LCU}^t)$ induced by the lifted start-up cost inequality for t and $\overline{\mathrm{ol}}$. Hence, their convex combination $(v, \tilde{\mathrm{cu}}^t)$ also lies on the same facet. By definition of the epigraph,

$$\mathrm{LCU}^t(v) = \tilde{\mathrm{cu}}^t = \sum_{j=1}^{\overline{\mathrm{ol}}+1} \lambda_j \mathrm{CU}^t(j-1).$$

In conclusion, $(v, \mathrm{cu}^t) \in \mathrm{epi}(\mathrm{LCU}^t)$ iff $\mathrm{cu}^t \geq \mathrm{LCU}^t(v) = \tilde{\mathrm{cu}}^t$, which is true iff (v, cu^t) fulfills the lifted start-up cost inequality for t and $\overline{\mathrm{ol}}$. □

It should be noted that the length $\overline{\mathrm{ol}}$ as used in the last lemma is the generalization of $\mathrm{ol}^t(v)$ as defined in (1.3.7) for the fractional situation. If v is integral, then condition (2.1.6) is always fulfilled for the number of offline periods $\mathrm{ol}^t(v)$ preceding period t.

Lemma 2.7 implies that each point $(v, \mathrm{cu}^t) \in [0,1]^T \times \mathbb{R}$ which fulfills all lifted start-up cost inequalities is contained in $\mathrm{epi}(\mathrm{LCU}^t)$. Since these inequalities are also

valid for epi(LCU^t) by Proposition 2.1, they form an $\mathcal{H}$-representation of epi(LCU^t) (together with the trivial inequalities).

Theorem 2.8

$$\operatorname{epi}(\mathrm{LCU}^t) = \left\{ \begin{array}{ll} (v, \mathrm{cu}^t) \in \mathbb{R}^{T+1} : & \\ \quad \mathrm{cu}^t \geq \mathrm{CU}^t(l) v^t - \sum_{j=1}^{l} \Big(\mathrm{CU}^t(l) - \mathrm{CU}^t(j-1) \Big) v^{t-j}, & l \in [0 \,..\, t-1] \\ \quad 0 \leq v^t \leq 1, & t \in [T] \end{array} \right\}.$$

As a result, we can separate epi(LCU^t) by considering the lifted start-up cost inequalities. Since there are t such inequalities with $\Theta(t)$ non-zero coefficients, the canonical algorithm has a running time of $\Theta(t^2)$. By exploiting the relationship of the right-hand sides (Proposition 2.5) this can be improved to the linear-time separation algorithm 2.1.1.

Algorithm 2.1.1: Separating a given point from epi(LCU^t)

Input : $(v, \mathrm{cu}^t) \in [0,1]^T \times \mathbb{R}$
Output : $(v, \mathrm{cu}^t) \in \operatorname{epi}(\mathrm{LCU}^t)$, or a separating inequality

$l \leftarrow 0$;
$v^\Sigma \leftarrow v^t$;
rhs $\leftarrow \mathrm{CU}^t(0) v^t$;
while $l < t-1$ *and* $v^\Sigma > v^{t-l-1}$ **do**
  $l \leftarrow l + 1$;
  $v^\Sigma \leftarrow v^\Sigma - v^{t-l}$;
  rhs $\leftarrow$ rhs $+ (\mathrm{CU}^t(l) - \mathrm{CU}^t(l-1)) v^\Sigma$;
if $\mathrm{cu}^t \geq$ *rhs* **then**
  return $(v, \mathrm{cu}^t) \in \operatorname{epi}(\mathrm{LCU}^t)$;
else
  return (v, cu^t) *can be separated from* epi(LCU^t) *by*

$$\mathrm{cu}^t \geq \mathrm{CU}^t(l) v^t - \sum_{j=1}^{l} \Big(\mathrm{CU}^t(l) - \mathrm{CU}^t(j-1) \Big) v^{t-j};$$

Proposition 2.9 *For each $t \in [T]$ and each $(v, \mathrm{cu}^t) \in [0,1]^T \times \mathbb{R}$, Algorithm 2.1.1 either confirms that $(v, \mathrm{cu}^t) \in \operatorname{epi}(\mathrm{LCU}^t)$, or finds a lifted start-up cost inequality separating (v, cu^t) from* epi(LCU^t) *in $\mathcal{O}(t)$.*

Proof. The correctness of the algorithm can be affirmed by noting that before and after each iteration of the `while`-loop on lines 4–7, it holds that

- $v^{\Sigma} = v^t - \sum_{j=1}^{l} v^{t-j}$, and
- rhs equals the right hand side of lifted start-up cost inequality for t and l.

The `while`-loop is executed until either $l = t-1$ or $v^t \leq \sum_{j=1}^{l+1} v^{t-j}$. Thus, after this loop, l is maximal in $[0\,..\,t-1]$ with $v^t > \sum_{j=1}^{l} v^{t-j}$, fulfilling the requirements of Lemma 2.7.

So, if the condition $\mathrm{cu}^t \geq$ rhs on line 8 is true, then the point (v, cu^t) fulfills the lifted start-up cost inequality for t and l, and by Lemma 2.7 lies in the epigraph $\mathrm{epi}(\mathrm{LCU}^t)$. Otherwise, this constraint separates (v, cu^t) from $\mathrm{epi}(\mathrm{LCU}^t)$.

The `while`-loop is executed at most $t-1$ times and the separating inequality on line 11 can also be constructed in $\mathcal{O}(t)$. In sum, the running time is $\mathcal{O}(t)$. □

2.1.5 The Convex Extension LCU^t of DCU^t

Knowing an $\mathcal{H}$-representation of $\mathrm{epi}(\mathrm{LCU}^t)$, we can deduce an explicit (non-linear) representation of the convex extension LCU^t of DCU^t. Again, this definition is based on the alternative formulation of the lifted start-up cost inequalities in (2.1.5).

Proposition 2.10 *For each $t \in [T]$ and $v \in [0,1]^T$,*

$$\mathrm{LCU}^t(v) = \sum_{l=0}^{t-1} \Big(\mathrm{CU}^t(l) - \mathrm{CU}^t(l-1)\Big) \max\left\{0, v^t - \sum_{j=1}^{l} v^{t-j}\right\}, \tag{2.1.8}$$

defining $\mathrm{CU}^t(-1) = 0$ for notational convenience.

Proof. By definition of the epigraph, the value of $\mathrm{LCU}^t(v)$ is the smallest value cu^t such that $(v, \mathrm{cu}^t) \in \mathrm{epi}(\mathrm{LCU}^t)$. Due to Lemma 2.7, $(v, \mathrm{cu}^t) \in \mathrm{epi}(\mathrm{LCU}^t)$ iff (v, cu^t) fulfills all lifted start-up cost inequalities. Using the alternative representation (2.1.5), the minimal such cu^t is

$$\mathrm{LCU}^t(v) = \mathrm{cu}^t = \max_{k \in [0\,..\,t-1]} \left\{ \sum_{l=0}^{k} \Big(\mathrm{CU}^t(l) - \mathrm{CU}^t(l-1)\Big) \left(v^t - \sum_{j=1}^{l} v^{t-j}\right) \right\} \tag{2.1.9}$$

In Lemma 2.7 we have seen that the maximum of the right-hand sides is attained by an $\overline{\mathrm{ol}} \in [0\,..\,t-1]$ with

$$\sum_{j=1}^{\overline{\mathrm{ol}}} v^{t-j} \leq v^t \leq \sum_{j=1}^{\overline{\mathrm{ol}}+1} v^{t-j} \qquad \text{(defining } v^0 = 1\text{)}.$$

The right-hand side of Eq. (2.1.8) corresponds exactly to the right hand side of Eq. (2.1.9) for $k = \overline{\mathrm{ol}}$. □

Since all lifted start-up cost inequalities (2.1.4) are homogeneous, LCU^t is homogeneous as well - a fact which can also be seen by examining (2.1.8).

Corollary 2.11 *The start-up cost function* LCU^t *is homogeneous, i. e. it holds that*

$$\forall\, v \in [0,1]^T, \lambda \geq 0: \qquad \mathrm{LCU}^t(\lambda v) = \lambda\, \mathrm{LCU}^t(v)$$

2.1.6 Redundancy and Approximations of Start-up Costs

Modeling $\mathrm{epi}(\mathrm{LCU}^t)$ for all $t \in [T]$ requires ${}^{T(T+1)}/_2$ lifted start-up cost inequalities and T inequalities $0 \leq v^t \leq 1$, which proves to cause a high computational burden. Traditionally, the number of necessary inequalities (2.1.2) is reduced by replacing the costs $\mathrm{CU}^t(l)$ by approximations $\tilde{\mathrm{CU}}^t(l)$ with a small number of steps, i. e. different values [Gar62; Muc66; SPO05; CA06]. We transfer this reduction to $\mathrm{epi}(\mathrm{LCU}^t)$ by showing that for this approximation $\tilde{\mathrm{CU}}^t(l)$, some inequalities (2.1.4) may be redundant.

Furthermore, to the best of our knowledge, the approximation $\tilde{\mathrm{CU}}^t(l)$ has always been regarded part of the input in the existing publications, and its choice has not been discussed. We provide an elementary algorithm which, given a certain error tolerance, finds an approximation of the $\mathrm{CU}^t(l)$ with a minimal number of steps.

Since the lifted start-up cost inequalities induce facets, the $\mathcal{H}$-representation in Theorem 2.8 is only redundant if two of them are equivalent. Using the alternative formulation in Corollary 2.6, it is straightforward to derive under which conditions this occurs.

Proposition 2.12 *For each* $t \in [T]$, $l_1, l_2 \in [t-1]$, *the two lifted start-up cost inequalities for* t *and* l_1 *or* l_2 *are equal iff* $\mathrm{CU}^t(l_1) = \mathrm{CU}^t(l_2)$.

Proof. Let t, l_1, and l_2 with $l_1 \leq l_2$.

By looking at the coefficient of v^t in (2.1.4), we see that two such constraints may only be equal if $\mathrm{CU}^t(l_1) = \mathrm{CU}^t(l_2)$.

Now, assume that $\mathrm{CU}^t(l_1) = \mathrm{CU}^t(l_2)$. Since the start-up costs increase with l, this implies $\mathrm{CU}^t(l_1) = \mathrm{CU}^t(l) = \mathrm{CU}^t(l_2)$ for all $l \in [l_1..l_2]$. Using (2.1.5), we obtain

$$\mathrm{rhs}^t(l_2) - \mathrm{rhs}^t(l_1) = \sum_{j=l_1}^{l_2} \Big(\underbrace{\mathrm{CU}^t(j) - \mathrm{CU}^t(j-1)}_{=0} \Big) \left(v^t - \sum_{k=1}^{j} v^{t-k} \right) = 0,$$

implying that equality of the two constraints. □

Due to the monotonicity of the $\mathrm{CU}^t(l)$, Proposition 2.12 implies that the lifted start-up cost inequalities for t and $l \in [l_1..l_2]$ are equal.

An irredundant $\mathcal{H}$-representation is gained by simply removing duplicate inequalities.

Corollary 2.13 *The $\mathcal{H}$-representation*

$$\operatorname{epi}(\mathrm{LCU}^t) = \left\{ \begin{array}{ll} (v, \mathrm{cu}^t) \in \mathbb{R}^{T+1} : & \\ \quad \mathrm{cu}^t \geq \mathrm{CU}^t(l)v^t - \sum_{j=1}^{l} \Big(\mathrm{CU}^t(l) - \mathrm{CU}^t(j-1)\Big)v^{t-j}, & \\ \qquad\qquad l \in [0\,..\,t-1] \ \textit{with}\ \mathrm{CU}^t(l) > \mathrm{CU}^t(l-1) & \\ \quad 0 \leq v^t \leq 1, \qquad t \in [T] & \end{array} \right\}.$$

is irredundant.

Thus, the $\mathcal{H}$-representation given in 2.8 is irredundant for strictly increasing start-up costs $\mathrm{CU}^t(l)$. On the other hand, the number of irredundant inequalities decreases for each $l \in [t-1]$ with $\mathrm{CU}^t(l) = \mathrm{CU}^t(l+1)$. This suggest a natural question: Given certain costs $\mathrm{CU}^t(l)$, how can we find an approximation $\tilde{\mathrm{CU}}^t(l)$ that lies within a given tolerance and minimizes the number of inequalities? Figure 2.8 highlights that such an approach can be effective even at low error tolerances.

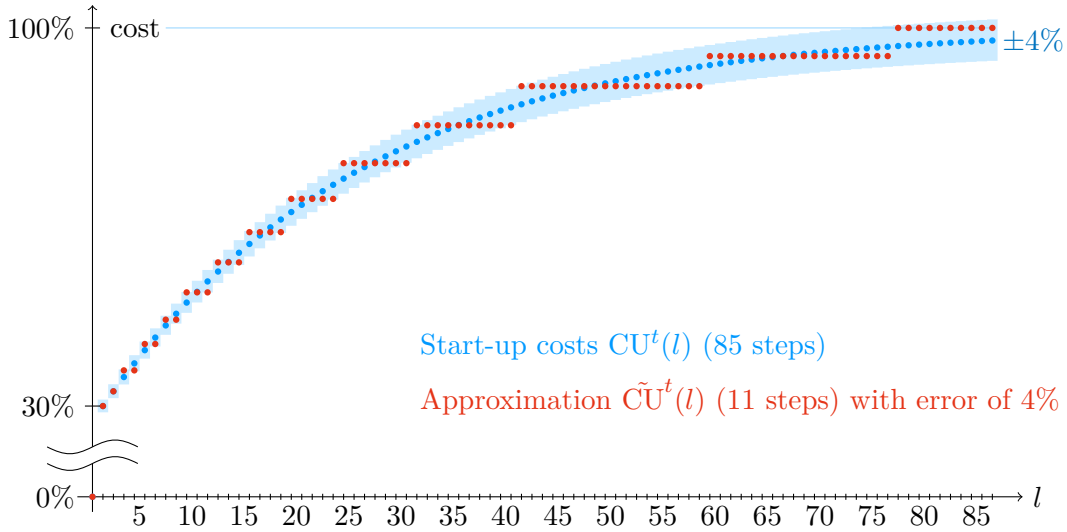


Figure 2.8: Approximation of a start-up cost function with a tolerance of 4%
The start-up cost function of a thermal unit (blue) with 30% fixed costs and 70% variable costs (absolute values are not relevant) is approximated in the first 85 periods by a function with 11 steps to a relative error of 4%. The error margin is marked in light blue.

For the sake of a simpler notation, we exclude the trivial case $t = 1$. Given an error tolerance $\mathrm{CU}_{\mathrm{tol}}$, the problem of finding an approximation $\tilde{\mathrm{CU}}^t(l)$ with a minimal number of steps can be stated as

$$\min\left\{\begin{array}{cc} \left|\bigcup\limits_{l\in[t-1]}\left\{\tilde{\mathrm{CU}}^t(l)\right\}\right| : & \\ \left|\tilde{\mathrm{CU}}^t(l) - \mathrm{CU}^t(l)\right| \leq \mathrm{CU}_{\mathrm{tol}}\,\mathrm{CU}^t(l), & l \in [t-1] \\ \tilde{\mathrm{CU}}^t(1) \leq \tilde{\mathrm{CU}}^t(2) \leq \ldots \leq \tilde{\mathrm{CU}}^t(t-1), & \\ \tilde{\mathrm{CU}}^t(l) \in \mathbb{R}_{\geq 0} & l \in [t-1] \end{array}\right\}. \tag{2.1.10}$$

If the minimal number of different start-up cost values is n, there exist period counts $1 = l_1 < l_2 < \ldots < l_n < l_{n+1} = t$ such that the start-up costs in each of the intervals $[l_j, l_{j+1}-1]$ for $j \in [n]$ are equal. For any such interval $[l_j, l_{j+1}-1]$, it is straightforward to check that the minimal local relative error is

$$\delta(l_j, l_{j+1}-1) := \frac{\mathrm{CU}^t(l_{j+1}-1) - \mathrm{CU}^t(l_j)}{\mathrm{CU}^t(l_{j+1}-1) + \mathrm{CU}^t(l_j)},$$

which is attained by the start-up costs

$$\tilde{\mathrm{CU}}^t(l_j) = \ldots = \tilde{\mathrm{CU}}^t(l_{j+1}-1) = \gamma(l_j, l_{j+1}-1) := \frac{2\,\mathrm{CU}^t(l_j)\,\mathrm{CU}^t(l_{j+1}-1)}{\mathrm{CU}^t(l_{j+1}-1) + \mathrm{CU}^t(l_j)},$$

except if $\mathrm{CU}^t(l_{j+1}-1) = 0$, where the minimal local error of 0 is attained by $\tilde{\mathrm{CU}}^t(l_j) = \ldots = \tilde{\mathrm{CU}}^t(l_{j+1}-1) = 0$. Thus, the minimization problem 2.1.10 is reduced to

$$\min\left\{\begin{array}{cc} n : & \\ \delta(l_j, l_{j+1}-1) \leq \mathrm{CU}_{\mathrm{tol}}, & j \in [n] \\ l_1 = 1, & \\ l_{n+1} = t, & \\ l_i < l_{j+1}, & j \in [n] \\ n \in [t-1], & \\ l_j \in \mathbb{N} & j \in [n+1] \end{array}\right\}. \tag{2.1.11}$$

The function δ is decreasing in its first parameter, and increasing in its second parameter. Thus, we can solve (2.1.11) by starting with $l_1 := 1$, and greedily choosing the maximal feasible value for each l_j, in turns.

Algorithm 2.1.2: Finding optimal step boundaries l_j

```
1 l_1 ← 1
2 n ← 0
3 while l_{n+1} < t do
4 |   n ← n + 1
5 |   l_{n+1} ← l_n + 1
6 |   while l_{n+1} < t ∧ bestError(l_n, l_{n+1}) ≤ CU_tol do
7 |   |   l_{n+1} ← l_{n+1} + 1
```

Proposition 2.14 *Algorithm 2.1.2 solves* (2.1.11) *in* $\mathcal{O}(t)$ *steps.*

Proof. It is straightforward to check that $1 = l_1 < l_2 < \ldots < l_n < l_{n+1} = t$ is always fulfilled (remember that we assumed $t \geq 2$). After line 5, it trivially holds that $\delta(l_n, l_{n+1} - 1) = 0$. So, by design of the loop termination condition on line 6, after each iteration of the loop it still holds that $\delta(l_n, l_{n+1} - 1) \leq \mathrm{CU}_{\mathrm{tol}}$. Thus, the computed $l_1, \ldots, l_{n+1}$ are a feasible solution to our minimization problem 2.1.11.

Suppose that there exists a better solution $1 = \tilde{l}_1 < \tilde{l}_2 < \ldots < \tilde{l}_m < \tilde{l}_{m+1} = t$ with $m < n$. Since $\tilde{l}_{m+1} = t = l_{n+1} > l_{m+1}$, there exist $j \in [2 \,..\, m{+}1]$ such that $l_j < \tilde{l}_j$. Denote the minimal such index as j^*. Trivially, we have $l_{j^*} < \tilde{l}_{j^*} \leq t$. Thus, the loop on lines 6-7 terminated due to $\delta(l_{j^*-1}, l_{j^*}) > \mathrm{CU}_{\mathrm{tol}}$, implying

$$\delta(\tilde{l}_{j^*-1}, \tilde{l}_{j^*} - 1) \geq \delta(\tilde{l}_{j^*-1}, l_{j^*}) = \delta(l_{j^*-1}, l_{j^*}) > \mathrm{CU}_{\mathrm{tol}},$$

a contradiction to the feasibility of the $\tilde{l}_1, \ldots, \tilde{l}_{m+1}$.

Finally, the running time of the algorithm follows from the increment of l_{n+1} in each loop. □

We transformed the minimization problem 2.1.10 into 2.1.11 by reducing the search for approximate start-up costs to the search for a feasible partitioning of $[t-1]$. Analogously, the minimization problem 2.1.10 can be reduced to a search for a feasible partitioning of $[\mathrm{CU}^t(1), \mathrm{CU}^t(t-1)]$. This provides an upper bound on the minimal number n of start-up cost values, which only depends on the interval $[\mathrm{CU}^t(1), \mathrm{CU}^t(t-1)]$ and on the tolerance $\mathrm{CU}_{\mathrm{tol}}$.

Proposition 2.15 *The objective function of the minimization problem 2.1.10 is bounded from above by*

$$\left\lceil \frac{\log(\mathrm{CU}^t(t-1)) - \log(\mathrm{CU}^t(1))}{\log(1 + \mathrm{CU}_{\mathrm{tol}}) - \log(1 - \mathrm{CU}_{\mathrm{tol}})} \right\rceil$$

for any $0 \leq \mathrm{CU}_{\mathrm{tol}} < 1$.

Proof. The statement is proved by providing a feasible solution to the minimization problem 2.1.11. Define $\overline{n}$ as the upper bound

$$\overline{n} := \left\lceil \frac{\log(\mathrm{CU}^t(t-1)) - \log(\mathrm{CU}^t(1))}{\log(1+\mathrm{CU}_{\mathrm{tol}}) - \log(1-\mathrm{CU}_{\mathrm{tol}})} \right\rceil.$$

We partition the continuous interval $[\mathrm{CU}^t(1), \mathrm{CU}^t(t-1)]$ in $\overline{n}$ intervals with boundaries c^j defined as

$$\forall\, j \in [\overline{n}+1]: \qquad c_j := \left(\frac{1+\mathrm{CU}_{\mathrm{tol}}}{1-\mathrm{CU}_{\mathrm{tol}}}\right)^{j-1} \mathrm{CU}^t(1).$$

By the choice of $\overline{n}$ and c^j, we have $c_1 = \mathrm{CU}^t(1)$ and $c_{\overline{n}} < \mathrm{CU}^t(t-1) \leq c_{\overline{n}+1}$. Thus, the intervals $[c_1, c_2), [c_2, c_3), \ldots, [c_{\overline{n}}, c_{\overline{n}+1}]$ are a partitioning of a superset of $[\mathrm{CU}^t(1), \mathrm{CU}^t(t-1)]$. The size of the intervals is chosen such that

$$\tilde{\delta}(c_{j-1}, c_j) := \frac{c_j - c_{j-1}}{c_j + c_{j-1}} = \frac{c_{j-1}\left(\frac{c_j}{c_{j-1}} - 1\right)}{c_{j-1}\left(\frac{c_j}{c_{j-1}} + 1\right)} = \frac{\frac{1+\mathrm{CU}_{\mathrm{tol}}}{1-\mathrm{CU}_{\mathrm{tol}}} - 1}{\frac{1+\mathrm{CU}_{\mathrm{tol}}}{1-\mathrm{CU}_{\mathrm{tol}}} + 1} = \mathrm{CU}_{\mathrm{tol}}.$$

We define the period counts $l_1, \ldots, l_{\overline{n}}$ such that the start-up cost values for downtimes $l \in [l_j, l_{j+1} - 1]$ coincide with the above $[c_j, c_{j+1}]$ intervals,

$$\forall\, l \in [\overline{n}+1]: \qquad l_j := \begin{cases} \min\{l \in [t-1] \mid \mathrm{CU}^t(l) \geq c^j\} & \text{if } j \in [\overline{n}] \\ t & \text{else } (j = \overline{n}+1). \end{cases}$$

The min-expression is well-defined because of $c^{\overline{n}} < \mathrm{CU}^t(t-1)$. Since the c^j are increasing, the l_j are increasing as well. If two indices l_j and l_{j+1} are equal, one of them can be "dropped", decreasing $\overline{n}$ by one. Thus we can assume that the l_j increase strictly. By definition, it holds that $l_{\overline{n}+1} = t$, and $c^1 = \mathrm{CU}^t(1)$ implies $l_1 = 1$.

Finally, for each $j \in [\overline{n}]$, by definition of the l_j we have

$$\mathrm{CU}^t(l_j) \geq c^j \qquad \text{and} \qquad \mathrm{CU}^t(l_{j+1} - 1) \leq c^{j+1},$$

which by choice of the c^j implies that

$$\delta(l_j, l_{j+1} - 1) \leq \tilde{\delta}(c_{j-1}, c_j) \leq \mathrm{CU}_{\mathrm{tol}}.$$

Thus, the values $l_1, \ldots, l_{\overline{n}+1}$ are a feasible solution to our minimization problem 2.1.11. □

The bound on the maximal number of needed inequalities given in Proposition 2.15 does not depend on the number of periods t, but only on the error tolerance $\mathrm{CU}_{\mathrm{tol}}$ and the ratio of the initial start-up costs $\mathrm{CU}^t(1)$ and the maximal start-up costs $\mathrm{CU}^t(t-1)$. This period-independent bound is shown in Fig. 2.9 for different ratios and error tolerances.

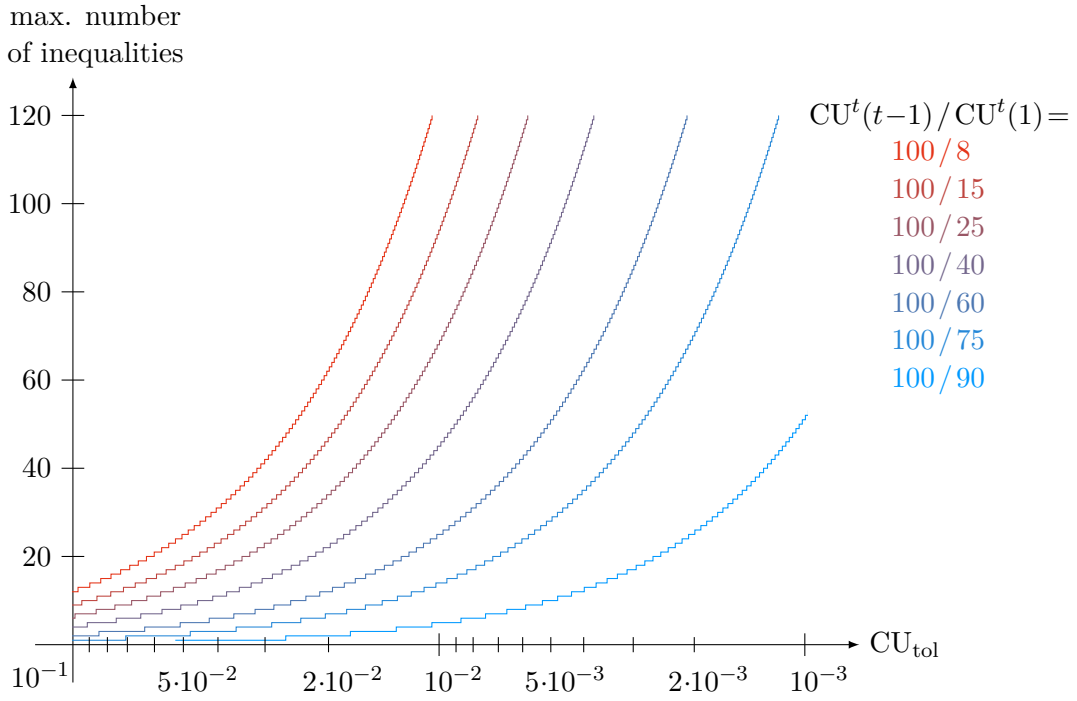


Figure 2.9: The bound on the number of irredundant inequalities due to Proposition 2.15, for varying parts of fixed and variable start-up costs, depending on $\mathrm{CU}_{\mathrm{tol}}$.

Lastly, Algorithm 2.1.2 is easily varied to find period indices l_j, such that the start-up costs for any chosen downtime l are not approximated, i. e. such that the resulting costs $\tilde{\mathrm{CU}}^t(l)$ equal the original costs $\mathrm{CU}^t(l)$. In our model, we choose to accurately model the start-up costs at $\mathrm{CU}^t(1)$ and the asymptotic final start-up costs.

2.2 The Summed Start-up Costs

This section presents joint work with René Brandenberg and Matthias Huber [BHS16].

We consider the sum of the incurred start-up costs of a single unit. Our major result is a correspondence between the facets of its epigraph and some binary trees for concave start-up cost functions CU, which is bijective if CU is strictly concave. We derive an exponential $\mathcal{H}$-representation of the convex hull of this epigraph, and provide an exact linear separation algorithm. These results significantly decrease the integrality gap of the mixed integer formulation of a Unit Commitment problem compared to current literature. Moreover, this section provides the basis for the temperature formulation in Section 3.3.

The *discrete summed start-up cost function* DCU^Σ is defined for operational schedules $v \in \{0,1\}^T$ based on the start-up costs DCU^t in individual periods,

$$\mathrm{DCU}^\Sigma : \quad \{0,1\}^T \to \mathbb{R}_{\geq 0}, \quad v \mapsto \sum_{t=1}^{T} \mathrm{DCU}^t(v), \qquad \text{and}$$

the *(linearized) summed start-up cost function* LCU^Σ is the convex hull of DCU^Σ on the set $[0,1]^T$, i. e. the unique function such that $\mathrm{epi}(\mathrm{LCU}^\Sigma) = \mathrm{conv}(\mathrm{epi}(\mathrm{DCU}^\Sigma))$.

Since the domain of DCU^Σ is $\{0,1\}^T$, the epigraph $\mathrm{epi}(\mathrm{LCU}^\Sigma)$ is a polyhedron, which means that LCU^Σ is piece-wise linear, and $\mathrm{DCU}^\Sigma = \mathrm{LCU}^\Sigma\big|_{\{0,1\}^T}$. Summarizing, $\mathrm{epi}(\mathrm{LCU}^\Sigma)$ possesses the following irredundant $\mathcal{V}$-representation.

Corollary 2.16 *The vertices of the polyhedron* $\mathrm{epi}(\mathrm{LCU}^\Sigma)$ *are*

$$V^\Sigma := \Big\{(v, \mathrm{cu}^\Sigma) \in \{0,1\}^T \times \mathbb{R}_{\geq 0} \;\Big|\; \mathrm{cu}^\Sigma = \mathrm{DCU}^\Sigma(v)\Big\},$$

and it holds that

$$\mathrm{epi}(\mathrm{LCU}^\Sigma) = \mathrm{conv}(\mathrm{epi}(\mathrm{DCU}^\Sigma)) = \mathrm{conv}(V^\Sigma) + \mathrm{pos}(u_{T+1}),$$

where u_{T+1} *denotes the unit vector in the direction of the last coordinate* cu^Σ.

The epigraph $\mathrm{epi}(\mathrm{LCU}^\Sigma)$ inherits the trivial facets of its domain $[0,1]^T$, induced by $0 \leq v^t \leq 1$).

At its core, our contribution results from four observations:

1. Iteratively lifting a trivial inequality variable by variable results in facets of $\mathrm{epi}(\mathrm{LCU}^\Sigma)$ and in coefficients which can be derived explicitly from the lifting order.

2. Different lifting orders result in the same facet. Each lifted facet is already uniquely described by a special partial order, which can be expressed as a binary tree, warranting the name *binary tree inequality* (BTI) for the lifted inequalities.

3. The binary tree associated with a lifted facet readily identifies the points on this facet. Furthermore, these facets describe the complete lower boundary (with respect to the last coordinate cu^Σ) of the epigraph, proving that it does not have further non-trivial facets.

4. For a fractional point, a suitable binary tree and the coefficients of the respective lifted facet can be determined efficiently.

These observations are presented as follows:

- in Subsection 2.2.1 the result of lifting a trivial inequality is analyzed,
- in Subsection 2.2.2 special notation for binary trees is introduced,
- in Subsection 2.2.3 the so-called *binary tree inequalities* are derived and shown to be facet-inducing,
- in Subsection 2.2.4 these inequalities are shown to complete an $\mathcal{H}$-representation of the epigraph of the summed start-up costs, and finally
- in Subsection 2.2.5 an exact separation algorithm is presented.

2.2.1 Lifting Inequalities

In this subsection, we observe that by lifting the trivial inequality $\mathrm{cu}^\Sigma \geq 0$, we gain facet-inducing inequalities with explicitly derivable coefficients.

By Corollary 2.11 the functions LCU^t are homogeneous, i. e. $\lambda\mathrm{LCU}^t(v) = \mathrm{LCU}^t(\lambda v)$. Hence LCU^Σ is homogeneous and it holds for all $(x, y) \in [0,1]^T \times \mathbb{R}, \lambda \in (0,1]$ that

$$\begin{aligned}(v, \mathrm{cu}^\Sigma) \in \mathrm{epi}(\mathrm{LCU}^\Sigma) &\Leftrightarrow \mathrm{cu}^\Sigma \geq \mathrm{LCU}^\Sigma(v) \Leftrightarrow \lambda\mathrm{cu}^\Sigma \geq \lambda\mathrm{LCU}^\Sigma(v) = \mathrm{LCU}^\Sigma(\lambda v)\\ &\Leftrightarrow \lambda(v, \mathrm{cu}^\Sigma) \in \mathrm{epi}(\mathrm{LCU}^\Sigma).\end{aligned}$$

This means, $\mathrm{epi}(\mathrm{LCU}^\Sigma)$ is the intersection of a cone with $[0,1]^T \times \mathbb{R}$, and thus all facets of $\mathrm{epi}(\mathrm{LCU}^\Sigma)$, except those induced by $0 \leq v^t \leq 1$, must also be facets of this cone. Hence, the facet-inducing inequalities are homogeneous, i. e. without constant term, and since cu^Σ is not bounded from above in $\mathrm{epi}(\mathrm{LCU}^\Sigma)$, of the kind

$$\mathrm{cu}^\Sigma \geq \sum_{t\in[T]} \alpha_t v^t.$$

Starting from $\mathrm{cu}^\Sigma \geq 0$, such facets can be identified using the standard sequential lifting method (see [Pad73]). For each order $\sigma : [T] \to [T]$, this method determines the coefficients $\alpha_{\sigma(j)}$ consecutively by considering the start-up costs for operational schedules v with coordinates $\sigma(j+1), \ldots, \sigma(T)$ fixed to 0,

$$\forall\, j \in [0..T]: \qquad \mathcal{F}^j(\sigma) := \Big\{ v \in \{0,1\}^T \;\Big|\; \forall\, k \in [j{+}1 \,..\, T] : w_j^{\sigma(k)} = 0 \Big\}.$$

Here, each $\mathcal{F}^j(\sigma)$ extends $\mathcal{F}^{j-1}(\sigma)$ by all v with $w_j^{\sigma(j)} = 1$,

$$\forall\, j \in [T]: \qquad \mathcal{F}_1^j(\sigma) := \mathcal{F}^j(\sigma) \setminus \mathcal{F}^{j-1}(\sigma) = \{v \in \mathcal{F}^j(\sigma), v^{\sigma(j)} = 1\}.$$

Being fulfilled by $0 \in \mathrm{epi}(\mathrm{LCU}^\Sigma)$ with equality, $\mathrm{cu}^\Sigma \geq 0$ induces a face of $\mathrm{epi}(\mathrm{LCU}^\Sigma)$ of at least dimension 0. By determining the coefficients $\alpha_{\sigma(j)}$ as

$$\forall\, j \in [T]: \qquad \alpha_{\sigma(j)} := \min\Big\{\mathrm{DCU}^\Sigma(v) - \sum_{k=1}^{j-1} \alpha_{\sigma(k)} v^{\sigma(k)} \;\Big|\; v \in \mathcal{F}_1^j(\sigma)\Big\},$$

the sequential lifting method produces a series of inequalities defining faces of $\mathrm{epi}(\mathrm{LCU}^\Sigma)$ of at least dimension j, culminating in a facet of $\mathrm{epi}(\mathrm{LCU}^\Sigma)$.

Both the number of vertices and facets of the polyhedrons $\mathcal{F}_1^j(\sigma)$ grow exponentially in general, rendering the above calculation of $\alpha_{\sigma(j)}$ computationally infeasible. In the case of $\mathrm{DCU}^\Sigma(v)$ however, we claim that these coefficients can be derived as

$$\begin{aligned} \forall\, j \in [T]: \qquad \alpha_{\sigma(j)} &= \mathrm{DCU}^\Sigma(w_j(\sigma)) - \mathrm{DCU}^\Sigma(w_{j-1}(\sigma)) \\ &= \mathrm{DCU}^\Sigma(w_{j-1}(\sigma) + u_{\sigma(j)}) - \mathrm{DCU}^\Sigma(w_{j-1}(\sigma)), \end{aligned} \tag{2.2.1}$$

with vectors $w_j(\sigma) \in \mathcal{F}_1^j(\sigma)$ defined as

$$\begin{aligned} \forall\; j \in [T]: \qquad & w_j(\sigma) := w_{j-1}(\sigma) + u_{\sigma(j)}, \\ & w_0(\sigma) := (0, \ldots, 0). \end{aligned} \tag{2.2.2}$$

To prove this, a closer look at the change of the summed start-up costs that defines $\alpha_{\sigma(j)}$ is necessary. This change depends on $w_{j-1}(\sigma)$ and $\sigma(j)$, and not on the relative lifting order of $\sigma(1), \ldots, \sigma(j-1)$ and $\sigma(j+1), \ldots, \sigma(T)$. We express this by considering $\tilde{v}, v \in \{0,1\}^T$ and $t \in [T]$ with $\tilde{v} = v - u_t$ in place of $w_{j-1}(\sigma)$, $w_j(\sigma)$, and $\sigma(j)$.

As depicted in Fig. 2.10, there are at most two indices $t' \in [T]$ such that the start-up costs $\mathrm{CU}^{t'}(v)$ and $\mathrm{CU}^{t'}(\tilde{v})$ differ,

1. the index t itself, and
2. the minimal index $q \in [t{+}1 \,..\, T]$ with $\tilde{v}^q = 1$, if such an index exists.

These start-up costs depend on the number of offline periods immediately before and after period t. The number of offline periods preceding t is given by $\mathrm{ol}^t(v)$ (see (1.3.3)). For each $\tilde{v} \in \{0,1\}^T$, $t \in [T]$, we define a similar function for the offline periods succeeding t as

$$\mathrm{or}^t(\tilde{v}) := \max\{j \in [0 \,..\, T{-}t] \mid \tilde{v}^{t+1} = \tilde{v}^{t+2} = \ldots = \tilde{v}^{t+j} = 0\}. \tag{2.2.3}$$

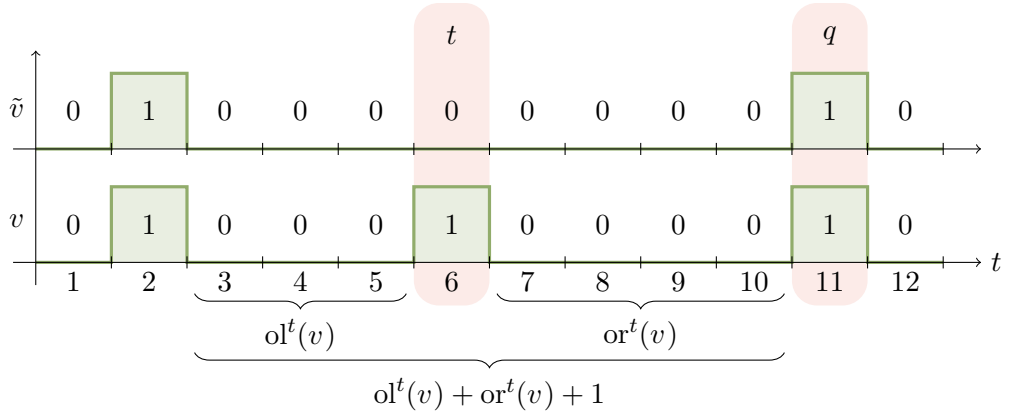


Figure 2.10: A step of the sequential lifting method, from the vector $\tilde{v}$ to the vector v. By setting $w_j^t = 1$, the downtime of length $\mathrm{ol}^t(v) + \mathrm{or}^t(v) + 1$ is split into two downtimes of lengths $\mathrm{ol}^t(v)$ and $\mathrm{or}^t(v)$, thereby changing the summed start-up costs.

So, for two operational schedules $v, \tilde{v} \in \{0,1\}^T$ which differ solely in period t, the start-up costs can be different only in periods t and $t + \mathrm{or}^t(v)$. Abbreviating $l := \mathrm{ol}^t(v)$ and $r := \mathrm{or}^t(v)$, we obtain

$$\mathrm{DCU}^\Sigma(v) - \mathrm{DCU}^\Sigma(\tilde{v}) = \begin{cases} \mathrm{DCU}^t(v) + \mathrm{DCU}^{t+r+1}(v) - \mathrm{DCU}^{t+r+1}(\tilde{v}) & \text{if } t + r < T, \\ \mathrm{DCU}^t(v) & \text{else.} \end{cases}$$

Thus, the change in the summed start-up costs depends only on the offline periods prior and after t, and can be further simplified.

Proposition 2.17 *For each $v, \tilde{v} \in \{0,1\}^T$ and $t \in [T]$ with $\tilde{v} = v - u_t$, it holds*

$$\mathrm{DCU}^\Sigma(v) - \mathrm{DCU}^\Sigma(\tilde{v}) = \delta^t(\mathrm{ol}^t(v), \mathrm{or}^t(v)),$$

where

$$\begin{aligned} \delta^t : \quad & [0 \,..\, t-1] \times [0 \,..\, T-t] \to \mathbb{R} \\ & (l, r) \mapsto \begin{cases} \mathrm{CU}^t(l) + \mathrm{CU}^{t+r+1}(r) - \mathrm{CU}^{t+r+1}(l + r + 1) & \textit{if } t + r < T, \\ \mathrm{CU}^t(l) & \textit{else.} \end{cases} \end{aligned}$$

The following lemma shows where the concavity of the start-up cost function CU is used.

Lemma 2.18 *For each $t \in [T]$, δ^t is increasing in l and r if* CU *is concave, and strictly increasing if* CU *is strictly concave.*

Proof. Let $t \in [T]$, $l, \tilde{l} \in [0 \,..\, t-1]$ with $l < \tilde{l}$ and $r, \tilde{r} \in [0 \,..\, T-t]$ with $r < \tilde{r}$ be given. Denote the period indices

$$q := t + r + 1 \quad \text{and} \quad \tilde{q} := t + \tilde{r} + 1,$$

implying $q < \tilde{q}$ (see Fig. 2.10). Recall that by definition, $\mathrm{CU}^t(l) = \mathrm{CU}(\mathrm{OL}^t(l))$, where $\mathrm{OL}^t(l)$ denotes the offline length corresponding to l offline periods prior to period t (see (1.3.4)). In Fig. 2.10 for example, l corresponds to $\mathrm{ol}^t(v)$ and r corresponds to $\mathrm{or}^t(v)$. So, $\mathrm{OL}^q(l + r + 1)$ is the downtime prior to q in $\tilde{v}$, and $\mathrm{OL}^q(r)$ and $\mathrm{OL}^t(l)$ are the downtimes prior to q and t in v, respectively. Hence

$$\mathrm{OL}^q(l + r + 1) = \mathrm{OL}^q(r + 1) + \mathrm{OL}^t(l).$$

We start by proving that δ^t is increasing in l. For $t + r = T$, δ^t increases in l since CU^t increases,

$$\delta^t(\tilde{l}, r) - \delta^t(l, r) = \mathrm{CU}^t(\tilde{l}) - \mathrm{CU}^t(l) \geq 0.$$

For $t + r < T$, we obtain

$$\begin{aligned}
\delta^t(\tilde{l},r) - \delta^t(l,r) &= \mathrm{CU}^q(r) + \mathrm{CU}^t(\tilde{l}) - \mathrm{CU}^q(\tilde{l} + r + 1) \\
&\quad - \Big(\mathrm{CU}^q(r) + \mathrm{CU}^t(l) - \mathrm{CU}^q(l + r + 1)\Big) \\
&= \Big(\mathrm{CU}^t(\tilde{l}) - \mathrm{CU}^t(l)\Big) - \Big(\mathrm{CU}^q(\tilde{l} + r + 1) - \mathrm{CU}^1(l + r + 1)\Big) \\
&= \Big(\mathrm{CU}(\mathrm{OL}^t(\tilde{l})) - \mathrm{CU}(\mathrm{OL}^t(l))\Big) - \Big(\mathrm{CU}(\mathrm{OL}^q(\tilde{l} + r + 1)) - \mathrm{CU}(\mathrm{OL}^q(l + r + 1))\Big) \\
&= \Big(\mathrm{CU}(\mathrm{OL}^t(l) + \mathrm{OL}^{t-l-1}(\tilde{l} - l)) - \mathrm{CU}(\mathrm{OL}^t(l))\Big) \\
&\quad - \Big(\mathrm{CU}(\mathrm{OL}^q(l + r + 1) + \mathrm{OL}^{t-l-1}(\tilde{l} - l)) - \mathrm{CU}(\mathrm{OL}^q(l + r + 1))\Big)
\end{aligned}$$

which, when abbreviating $x := \mathrm{OL}^t(l)$, $y := \mathrm{OL}^q(l + r + 1)$, $s := \mathrm{OL}^{t-l-1}(\tilde{l} - l)$, equals

$$\delta^t(\tilde{l}, r) - \delta^t(l, r) = \Big(\mathrm{CU}(x + s) - \mathrm{CU}(x)\Big) - \Big(\mathrm{CU}(y + s) - \mathrm{CU}(y)\Big).$$

The non-negativity of the final term follows from the characterization of concave functions by subdifferentials, using that $x < y$ and $s > 0$.

The statement that δ^t is increasing in r follows analogously in the case of $t + \tilde{r} < T$. If $t + \tilde{r} = T$, then

$$\begin{aligned}
\delta^t(l, \tilde{r}) - \delta^t(l, r) &= \mathrm{CU}^t(l) - \Big(\mathrm{CU}^t(l) + \mathrm{CU}^q(r) - \mathrm{CU}^q(l + r + 1)\Big) \\
&= \mathrm{CU}^q(l + r + 1) - \mathrm{CU}^q(r) \geq 0.
\end{aligned}$$

Finally, in the case of a strictly concave start-up cost function CU, all of the above inequalities are also strict. □

Proposition 2.17 shows that we may restate our claim regarding the lifted coefficients $\alpha_{\sigma(j)}$ in (2.2.1) as

$$\forall\, j \in [T]: \quad \alpha_{\sigma(j)} = \mathrm{DCU}^{\Sigma}(w_j(\sigma)) - \mathrm{DCU}^{\Sigma}(w_{j-1}(\sigma)) = \delta^{\sigma(j)}\big(\mathrm{ol}^{\sigma(j)}(w_j), \mathrm{or}^{\sigma(j)}(w_j)\big). \tag{2.2.4}$$

Based on Lemma 2.18, this could be proved by induction over j. Furthermore, [Pad73] shows that, as $\mathrm{cu}^{\Sigma} \geq 0$ is valid for $\mathrm{epi}(\mathrm{LCU}^{\Sigma})$, these inequalities define facets of $\mathrm{epi}(\mathrm{LCU}^{\Sigma})$. Since the current subsection is intended to be purely motivational, we prefer to re-introduce the lifted inequalities in Section 2.2.3, and prove that they induce facets without referring to the sequential lifting method, see Theorem 2.31.

The necessary steps are essentially the same:

- proving that the lifted inequalities are valid requires the same arguments as showing $\alpha_{\sigma(j)} \leq \delta^{\sigma(j)}\big(\mathrm{ol}^{\sigma(j)}(w_j), \mathrm{or}^{\sigma(j)}(w_j)\big)$, and
- proving that these inequalities induce facets requires the same arguments as showing $\alpha_{\sigma(j)} \geq \delta^{\sigma(j)}\big(\mathrm{ol}^{\sigma(j)}(w_j), \mathrm{or}^{\sigma(j)}(w_j)\big)$.

2.2.2 Notation for Binary Trees

Describing the lifted inequalities of the last subsection necessitates non-standard notation for binary trees, which is presented in this section. A binary tree is defined as an undirected, rooted tree, where each node t has at most two child nodes: a left child $\mathrm{llink}(t)$ and a right child $\mathrm{rlink}(t)$ (cf. Fig. 2.11).

Basic notation includes:

- Each binary tree B has a *root* denoted by $\mathrm{root}(B)$.
- The number of edges on the unique path from a node t to the root is called the *depth* $d(t)$ of t. The nodes of this path are the *ancestors* of t (including t itself). Vice versa, t is called a descendant of each of its ancestors.
- The path from a node t to the root is unique. The number $d(t)$ of its edges is called the *depth* of t, and its nodes are the *ancestors* of t (including t itself). Vice versa, t is called a descendant of each of its ancestors.
- The first ancestor, i. e. the node immediately succeeding t on the path to the root, is called the *parent* $p(t)$ of t. Conversely, t is said to be a *child* of $p(t)$.
- The subtree $S(t)$ comprising all descendants of a node t (including t) is the *principal subtree in* t. If the left/right child of t exists, then its principal subtree is the *left subtree* $L(t)$/*right subtree* $R(t)$ of t. All of these subtrees are binary trees as well.
- The number of nodes in these subtrees in t are denoted by $s(t) := |S(t)|$, $l(t) := |L(t)|$ and $r(t) := |R(t)|$.

Proposition 2.19 *For each binary tree B on nodes V, there exists a unique mapping* $\operatorname{rank} : V \to \{1, \dots, |V|\}$ *such that*

$$\forall\, t \in V \setminus \{\operatorname{root}(B)\} : \qquad \operatorname{rank}(t) \begin{cases} < \operatorname{rank}(p(t)) & \text{if } t \text{ is the left child of } p(t), \\ > \operatorname{rank}(p(t)) & \text{if } t \text{ is the right child of } p(t), \end{cases}$$

and such that the nodes of each principal subtree $S(t)$ are a mapped to a contiguous interval. This mapping is called the rank function of B.

Proof. We prove existence and uniqueness of the rank function by induction over the number $n := |V|$ of nodes of B. For $n = 1$, the unique rank function is clearly given by $\operatorname{rank} : V \to \{1\}$ with $\operatorname{rank}(t) = 1$ for the single node $t \in V$.

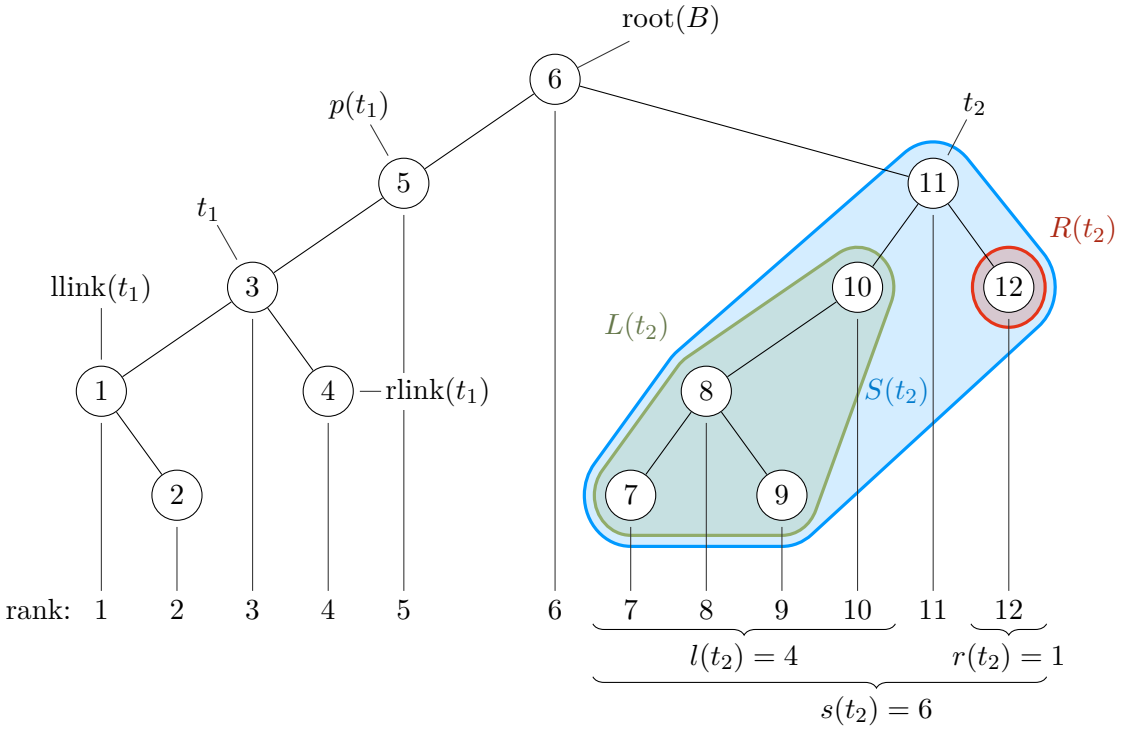


Figure 2.11: A binary tree with nodes labeled by ranks from 1 to 12, and $\operatorname{root}(B) = 6$. The parent of node $t_1 = 3$ is $p(t_1) = 5$, its left child is $\operatorname{llink}(t_1) = 1$ and its right child is $\operatorname{rlink}(t_1) = 4$. The principal subtree $S(t_2)$ in $t_2 = 11$ is marked in blue, its left subtree $L(t_2)$ in green and its right subtree $R(t_2)$ in red. The respective subtree sizes are $s(t_2) = 6$, $l(t_2) = 4$ and $r(t_2) = 1$.

For $n \geq 2$, let $r := \mathrm{root}(B)$. By induction, there exist rank functions for the subtrees $L(r)$ and $R(r)$, $\mathrm{rank}_l : L(r) \to \{1, \ldots, l(r)\}$ and $\mathrm{rank}_r : R(r) \to \{1, \ldots, r(r)\}$, respectively. Since $l(r) + 1 + r(r) = s(r) = n$, the mapping

$$\begin{aligned} \mathrm{rank}: \quad & V \to \{1, \ldots, n\} \\ & t \mapsto = \begin{cases} \mathrm{rank}_l(t) & \text{if } t \in L(t), \\ l(r) + 1 & \text{if } v = r, \\ l(r) + 1 + \mathrm{rank}_r(t) & \text{else,} \end{cases} \end{aligned}$$

is a rank function for B.

Furthermore, this rank function is unique: Choose any mapping rank' fulfilling the properties of a rank function for B. Since $\mathrm{rank}'(l(r)) < \mathrm{rank}'(r)$ and since a rank function must map the nodes of each $L(r)$ to a contiguous interval, $\mathrm{rank}'(t) < \mathrm{rank}'(r)$ for all $t \in L(r)$. Analogously, $\mathrm{rank}'(r) < \mathrm{rank}'(t)$ for all $t \in R(r)$. As rank' is a mapping to $\{1, \ldots, n\}$, we have $l(r) < \mathrm{rank}'(r) \leq n - r(r)$, which given that $r(r) = n - l(r) - 1$ implies $\mathrm{rank}'(r) = l(r) + 1 = \mathrm{rank}(r)$.

The restriction of rank' to $L(r)$ is a rank function for $L(r)$ and thereby unique, implying $\mathrm{rank}'|_{L(r)} = \mathrm{rank}|_{L(r)}$. Analogously, $\mathrm{rank}'|_{R(r)} = \mathrm{rank}|_{R(r)}$. In sum, $\mathrm{rank}' = \mathrm{rank}$, proving the uniqueness of the rank function. □

The rank function has a straightforward interpretation: When drawing a binary tree such that the left (right) subtree of each node t is located entirely left (right) of t, the rank numbers the nodes from left to right (see Fig. 2.11). Several basic properties can be derived from this observation:

- For each binary tree B and $t \in B$, it holds that

$$\{\mathrm{rank}(t') \mid t' \in S(t)\} = \{\mathrm{rank}(t') \mid t' \in L(t)\} \cup \{\mathrm{rank}(t)\} \cup \{\mathrm{rank}(t') \mid t' \in R(t)\}.$$

- Since by definition the ranks of nodes in a subtree are contiguous, it holds that

$$\begin{aligned} \{\mathrm{rank}(t') \mid t' \in L(t)\} &= [\mathrm{rank}(t) - l(t) \,..\, \mathrm{rank}(t) - 1], \\ \{\mathrm{rank}(t') \mid t' \in R(t)\} &= [\mathrm{rank}(t) + 1 \,..\, \mathrm{rank}(t) + r(t)], \text{ and} \\ \{\mathrm{rank}(t') \mid t' \in S(t)\} &= [\mathrm{rank}(t) - l(t) \,..\, \mathrm{rank}(t) + r(t)]. \end{aligned} \tag{2.2.5}$$

- Finally, since $L(t) = S(\mathrm{llink}(t))$ and $R(t) = S(\mathrm{rlink}(t))$, we have

$$\mathrm{rank}(\mathrm{llink}(t)) + r(\mathrm{llink}(t)) + 1 = \mathrm{rank}(t) = \mathrm{rank}(\mathrm{rlink}(t)) - l(\mathrm{rlink}(t)) - 1. \tag{2.2.6}$$

For example, in the binary tree shown in Fig. 2.11 it holds that

$$\{\mathrm{rank}(t') \mid t' \in S(t_2)\} = \underbrace{[7..10]}_{\{\mathrm{rank}(t')|t'\in L(t_2)\}} \cup \underbrace{\{11\}}_{\mathrm{rank}(t_2)} \cup \underbrace{\{12\}}_{\{\mathrm{rank}(t')|t'\in R(t_2)\}} \quad \text{and}$$

$$\underbrace{1}_{\mathrm{rank}(\mathrm{llink}(t_1))} + \underbrace{1}_{r(\mathrm{llink}(t_1))} + 1 = \underbrace{3}_{\mathrm{rank}(t_1)} = \underbrace{4}_{\mathrm{rank}(\mathrm{rlink}(t_1))} - \underbrace{0}_{l(\mathrm{rlink}(t_1))} - 1.$$

The special nature of the coefficients of the facets of $\mathrm{epi}(\mathrm{LCU}^{\Sigma})$ is best characterized by some non-standard notation, the *top-left* and *top-right* nodes (see Fig. 2.12). These nodes are defined recursively, with the root being the first top-left and top-right node. Left children of top-left nodes are also top-left nodes, and right children of top-right nodes are top-right nodes too. The last top-left node has rank 1, and the last top-right node has rank n, where n is the number of nodes in the binary tree.

Definition 2.20 For each binary tree B on n nodes, define $\mu_1 := \eta_1 := \mathrm{root}(B)$,

$$\forall\, j \in [d(1)]: \quad \mu_{j+1} := \mathrm{llink}(\mu_j), \qquad \forall\, j \in [d(n)]: \quad \eta_{j+1} := \mathrm{rlink}(\eta_j),$$

where μ_j denotes the j-th top-left node and η_j denotes the j-th top-right node.

Conversely, ancestors of top-left (top-right) nodes must be top-left (top-right) nodes as well.

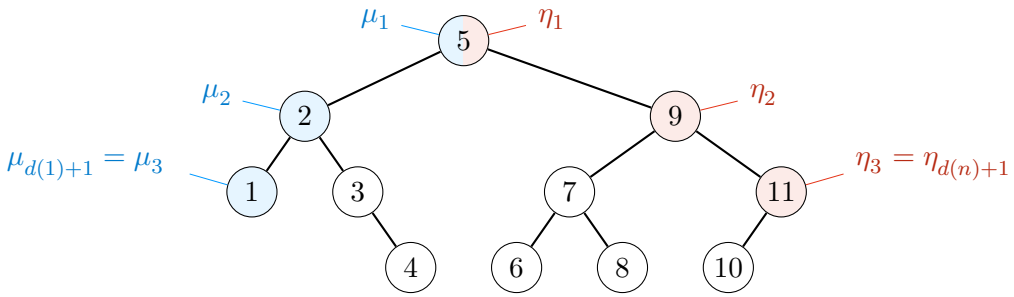


Figure 2.12: A binary tree with top-left nodes μ_i and top-right nodes η_i. Both μ_1 and η_1 always equal $\mathrm{root}(B)$. The last top-left node $\mu_{d(1)+1}$ is always the node with rank 1, and the last top-right node $\eta_{d(n)+1}$ is always the node with maximal rank.

We continue by proving basic facts regarding top-right nodes, and transfer them to top-left nodes by "mirroring" the binary trees.

In a graph with n nodes, the node n is contained in the principal subtree of each top-right node η_j, and by definition possesses the maximal rank. Since the principal subtree of η_j spans the ranks $[\mathrm{rank}(\eta_j)-l(\eta_j) \,..\, \mathrm{rank}(\eta_j)+r(\eta_j)]$ (see (2.2.5)), it must

hold that $\mathrm{rank}(\eta_j) + r(\eta_j) = n$. On the other hand, if the rank of a node t fulfills $\mathrm{rank}(t) + r(t) = n$, then n lies in the right subtree of t. So, t is an ancestor of n, and thus a top-right node.

Proposition 2.21 *For each binary tree B on n nodes, a node t is a top-right node in B iff* $\mathrm{rank}(t) + r(t) = n$.

This result may be extended to subtrees as well:

Lemma 2.22 *For each binary tree B on n nodes, a node t is a top-right node in the left subtree of a node t' iff* $\mathrm{rank}(t) + r(t) + 1 = \mathrm{rank}(t')$.

Proof. By (2.2.5), the left subtree $L(t')$ of t' consists of the nodes with ranks in the interval $[\mathrm{rank}(t') - l(t') \,..\, \mathrm{rank}(t') - 1]$. Thus, the rank function $\mathrm{rank}^{L(t')}$ of $L(t')$, which must keep the same order as rank and ranges from 1 to $l(t')$ equates

$$\mathrm{rank}^{L(t')}(t) = \mathrm{rank}(t) - \mathrm{rank}(t') + l(t') + 1.$$

So, each node $t \in L(t')$ is a top-right node in $L(t')$ iff $\mathrm{rank}^{L(t')}(t) + r(t) = |L(t')| = l(t')$ (see Proposition 2.21), which may be rewritten as

$$\mathrm{rank}(t) + r(t) + 1 = \mathrm{rank}^{L(t')}(t) + \mathrm{rank}(t') - l(t') - 1 + r(t) + 1 = \mathrm{rank}(t').$$

Assume that there exists a node t of B with $t \notin L(t')$ but $\mathrm{rank}(t) = \mathrm{rank}(t') - r(t) - 1$. Then $\mathrm{rank}(t) < \mathrm{rank}(t')$, and $t \notin S(t')$. On the other hand, $\mathrm{rank}(t') > \mathrm{rank}(t) + r(t)$ and hence $t' \notin S(t)$.

Choose the first common ancestor s of t and t', which due to $t \notin S(t')$ and $t' \notin S(t)$ equals neither t nor t'. By the choice of s it holds that $t \in L(s)$ and $t' \in R(s)$, and thus

$$\mathrm{rank}(t) + r(t) < \mathrm{rank}(s) < \mathrm{rank}(t') - l(t'),$$

a contradiction to $\mathrm{rank}(t) + r(t) + 1 = \mathrm{rank}(t')$. □

Note that since $\mathrm{rank}(t) + r(t) + 1 \in [\mathrm{rank}(t) + 1 \,..\, n + 1]$, t must be either a top-right node or a top-right node in the left subtree of some node t'.

Corollary 2.23 *For each binary tree on n nodes, a node t is*

- *a top-right node iff* $\mathrm{rank}(t) + r(t) = n$, *and*
- *a top-right node in the left subtree of the node t' with* $\mathrm{rank}(t') = \mathrm{rank}(t) + r(t) + 1$ *iff* $\mathrm{rank}(t) + r(t) < n$.

Let $\overline{B}$ denote the "mirrored" version of a binary tree B of size n, i. e. the binary tree that results from exchanging the left and right children of each node. The depth, parent and subtree size functions remain unchanged ($d^{\overline{B}} = d^B$, $p^{\overline{B}} = p^B$, $s^{\overline{B}} = s^B$), the left and right subtree size functions are exchanged ($r^{\overline{B}} = l^B$, $l^{\overline{B}} = r^B$), and the rank function is mirrored ($\mathrm{rank}^{\overline{B}}(t) = n + 1 - \mathrm{rank}^B$). By applying this mirroring to the properties of the top-right nodes, we can derive equivalent properties of the top-left nodes.

Corollary 2.24 *For each binary tree, a node t is*

- *a top-left node iff* $\mathrm{rank}(t) = l(t) + 1$, *and*
- *a top-left node in the right subtree of the node $\tilde{l}$ with* $\mathrm{rank}(\tilde{l}) = \mathrm{rank}(t) - l(t) - 1$ *iff* $\mathrm{rank}(t) > l(t) + 1$.

The root of a binary tree on n nodes is both a top-right and top-left node, and thus its rank equals

$$\mathrm{rank}(\mathrm{root}(B)) = n - r(r) = l(r) + 1.$$

2.2.3 The Binary Tree Inequalities

In this subsection, we show that all lifted inequalities correspond in a one-to-one way to binary trees, which motivates naming them *binary tree inequalities (BTIs)*. Together with $0 \leq v^t \leq 1$, these inequalities induce all facets of $\mathrm{epi}(\mathrm{LCU}^\Sigma)$.

We start with an example where α_4 and α_9 are lifted in both possible orders,

- α_4 before α_9 with intermediate vector w_j, and
- α_9 before α_4 with intermediate vector $\tilde{v}_j$ (cf. Fig. 2.13).

If the coefficient α_6 is already lifted, the relative order in which α_4 and α_9 are lifted does not influence the period counts,

$$\begin{aligned} \mathrm{ol}^4(w_j) = \mathrm{ol}^4(w_{j+1}) \qquad &\text{and} \qquad \mathrm{or}^4(w_j) = \mathrm{or}^4(w_{j+1}), \\ \mathrm{ol}^9(\tilde{v}_j) = \mathrm{ol}^9(w_{j+1}) \qquad &\text{and} \qquad \mathrm{ol}^9(\tilde{v}_j) = \mathrm{ol}^9(w_{j+1}). \end{aligned}$$

Thereby, the values of the lifted coefficients α_4 and α_9 are equal for both cases. This holds in general: As soon as a coefficient α_t has been lifted, for each subsequently lifted coefficient $\alpha_{t'}$ with $t' < t$ we have $\mathrm{or}^{t'}(v) < t - t'$. Thus, the period counts, and by extension the coefficient $\alpha_{t'}$, do not depend on the lifting order of coefficients $\alpha_{\tilde{t}}$ with $\tilde{t} > t$. Analogously if $t' > t$, the coefficient $\alpha_{t'}$ does not depend on the lifting order of coefficients $\alpha_{\tilde{t}}$ with $\tilde{t} < t$.

A lifting order σ corresponds to a linear order $\preccurlyeq_\sigma$ on $[T]$ with

$$\forall\, t_1, t_2 \in [T]: \qquad t_1 \preccurlyeq_\sigma t_2 \quad :\Leftrightarrow \quad \sigma^{-1}(t_1) \leq \sigma^{-1}(t_2).$$

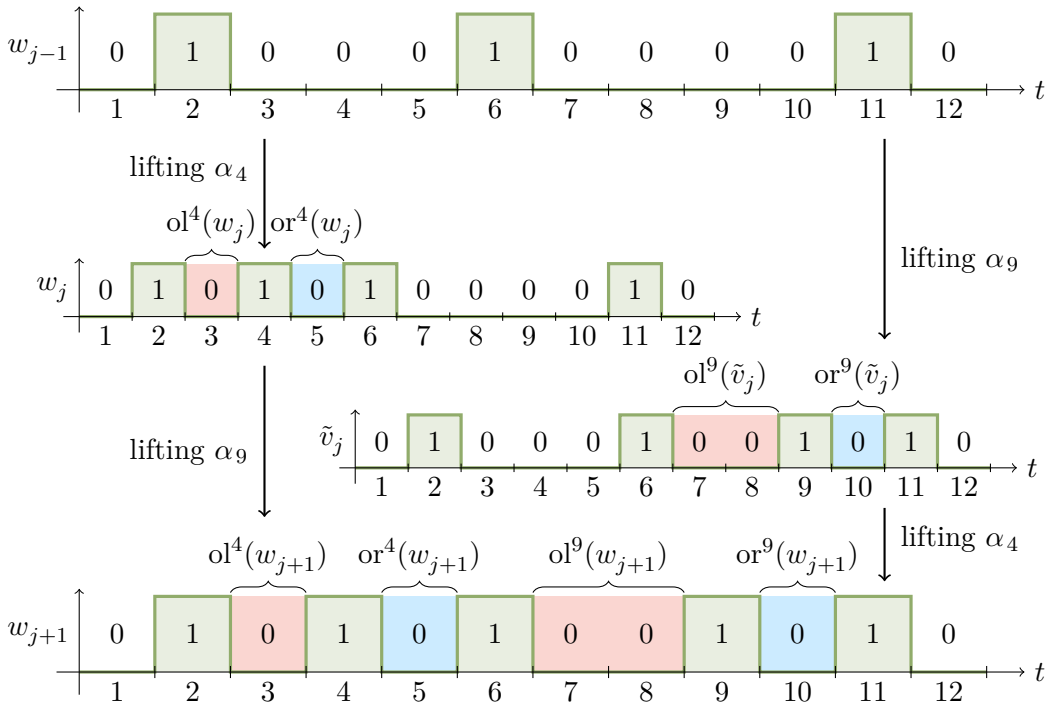


Figure 2.13: Lifting coefficients α_4 and α_9 in both orders, with intermediate vectors.

As argued above, the lifted coefficients $\alpha_{t_1}, \alpha_{t_2}$ do not depend on whether $t_1 \preccurlyeq_\sigma t_2$ or $t_2 \preccurlyeq_\sigma t_1$ if

$$\exists\, t_3 \in [T]: \quad \min\{t_1, t_2\} < t_3 < \max\{t_1, t_2\}, \quad t_3 \preccurlyeq_\sigma t_1, \quad t_3 \preccurlyeq_\sigma t_2.$$

Eliminating such relationships from $\preccurlyeq_\sigma$ yields the partial order $\preccurlyeq'_\sigma$ defined as

$$\forall\, t_1, t_2 \in [T]: \qquad t_1 \preccurlyeq'_\sigma t_2 \quad :\Leftrightarrow \quad t_1 \preccurlyeq_\sigma t_2 \text{ and } t_2 \in D_\sigma(t_1), \tag{2.2.7}$$

where

$$D_\sigma(t) := \Big[\max\Big\{t' \in [t-1] \;\Big|\; t' \preccurlyeq_\sigma t\Big\} + 1 \;..\; \min\Big\{t' \in [t{+}1 \;..\; T] \;\Big|\; t' \preccurlyeq_\sigma t\Big\} - 1\Big].$$

Proposition 2.25 *For each ordering $\sigma : [T] \to [T]$, $\preccurlyeq'_\sigma$ as defined in (2.2.7) is a partial order.*

Proof. By definition, for each $t_1 \in [T]$ we have $t_1 \in D_\sigma(t_1)$. Moreover, for each t_2 with $t_1 \preccurlyeq'_\sigma t_2$, we have $t_2 \in D_\sigma(t_1)$ and thus

$$\max\{t' \in [t_1 - 1] \mid t' \preccurlyeq_\sigma t_1\} + 1 \le t_2 \le \min\{t' \in [t_1{+}1 \;..\; T] \mid t' \preccurlyeq_\sigma t_1\} - 1,$$

implying

$$\begin{aligned}\max\{t' \in [t_1 - 1] \mid t' \preccurlyeq_\sigma t_1\} &= \max\{t' \in [t_2 - 1] \mid t' \preccurlyeq_\sigma t_1\} \\ &\leq \max\{t' \in [t_2 - 1] \mid t' \preccurlyeq_\sigma t_2\}.\end{aligned}$$

Analogously, $\min\{t' \in [t_1{+}1 \,..\, T] \mid t' \preccurlyeq_\sigma t_1\} \geq \min\{t' \in [t_2{+}1 \,..\, T] \mid t' \preccurlyeq_\sigma t_2\}$ and hence $D_\sigma(t_2) \subset D_\sigma(t_1)$ (more specifically, since $t_1 \notin D_\sigma(t_2)$, $D_\sigma(t_2) \subsetneq D_\sigma(t_1)$).

The reflexivity and antisymmetry of $\preccurlyeq_\sigma$ is directly inherited by $\preccurlyeq'_\sigma$ since $t \in D_\sigma(t)$. For each $t_1, t_2, t_3 \in [T]$ with $t_1 \preccurlyeq'_\sigma t_2$ and $t_2 \preccurlyeq'_\sigma t_3$, we have

- $t_1 \preccurlyeq_\sigma t_2$ and $t_2 \preccurlyeq_\sigma t_3$, implying $t_1 \preccurlyeq_\sigma t_3$, and
- $t_3 \in D_\sigma(t_2) \subset D_\sigma(t_1)$.

So $t_1 \preccurlyeq'_\sigma t_3$, proving that $\preccurlyeq'_\sigma$ is transitive. In conclusion, $\preccurlyeq'_\sigma$ is a partial order. □

The lifted inequality is fully determined by $\preccurlyeq'_\sigma$, i. e. each linearization of this partial order leads to the same lifted inequality. Fig. 2.14 shows an exemplary partial order (twice, as a Hasse diagram) and two possible linearizations $\preccurlyeq_\sigma$ and $\preccurlyeq_{\overline{\sigma}}$ represented by the permutations σ and $\overline{\sigma}$.

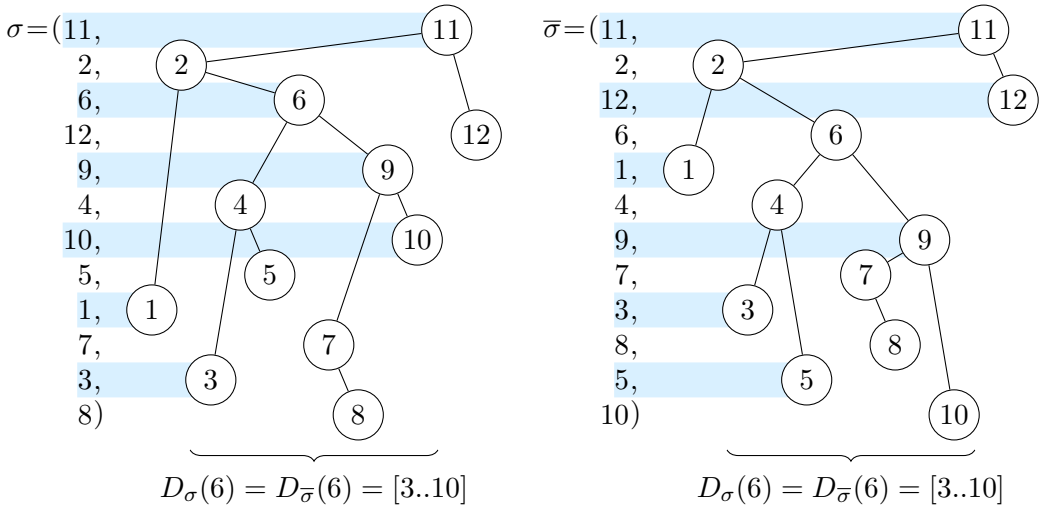


Figure 2.14: A partial order determining the lifted coefficients (twice, as a Hasse diagram) and two possible linearizations (as permutations σ and $\overline{\sigma}$) leading to the same lifted inequality.

As Fig. 2.14 suggests, the Hasse diagrams of a partial order is a binary tree with root node $\sigma^{-1}(1)$ and, for each node t, with

- the node t_l with $D_\sigma(t_l) = D_\sigma(t) \cap [t - 1]$ as the left child of t (if it exists), and
- the node t_r with $D_\sigma(t_r) = D_\sigma(t) \cap [t{+}1 \,..\, T]$ as the right child of t (if it exists).

Each coefficient α_t corresponds to a node t and the partial order then prescribes that each coefficient α_t must be lifted before the coefficients associated with the descendants of node t.

We claim in (2.2.2) that the vectors w_j encountered in the lifting process with order σ are

$$w_0 := (0, \dots, 0) \quad \text{and} \quad \forall\, j \in [T]: \quad w_j := w_{j-1} + u_{\sigma(j)}.$$

Observe that using the corresponding linear order $\preccurlyeq_\sigma$, these vectors fulfill

$$\forall\, j \in [T], t \in [T]: \qquad w_j^t = \begin{cases} 1 & \text{if } t \preccurlyeq_\sigma \sigma(j), \\ 0 & \text{else.} \end{cases}$$

Moreover, we claim in (2.2.4) that the lifting process results in an inequality with coefficients $\alpha_{\sigma(j)} = \delta^{\sigma(j)}(\mathrm{ol}^{\sigma(j)}(w_j), \mathrm{or}^{\sigma(j)}(w_j))$. Using the above representation of w_j, the involved offline lengths $\mathrm{ol}^{\sigma(j)}(w_j)$ and $\mathrm{or}^{\sigma(j)}(w_j)$ equate to

$$\begin{aligned} \mathrm{ol}^{\sigma(j)}(w_j) &= \max\{l \in [0 \,..\, \sigma(j)-1] \mid \forall\, t \in [\sigma(j)-l \,..\, \sigma(j)-1] : t \preccurlyeq_\sigma \sigma(j)\}, \text{ and} \\ \mathrm{or}^{\sigma(j)}(w_j) &= \max\{l \in [0 \,..\, T-\sigma(j)] \mid \forall\, t \in [\sigma(j)+1 \,..\, \sigma(j)+l] : t \preccurlyeq_\sigma \sigma(j)\}. \end{aligned}$$

As argued, these lengths remain unchanged when replacing $\preccurlyeq_\sigma$ by the corresponding partial order $\preccurlyeq'_\sigma$. Furthermore, their above representation shows that they can be derived from the Hasse diagram of $\preccurlyeq'_\sigma$: The size $l(\sigma(j))$ of the left subtree of $\sigma(j)$ equals $\mathrm{ol}^{\sigma(j)}(w_j)$ and the size $r(\sigma(j))$ of the right subtree of $\sigma(j)$ equals $\mathrm{or}^{\sigma(j)}(w_j)$ (see Fig. 2.15).

The claimed coefficients α_t thus solely depend on the partial order $\preccurlyeq'_\sigma$ and may be expressed as

$$\alpha_t = \delta^t(l(t), r(t)),$$

where $t = \sigma(j)$, and $l(t), r(t)$ are determined by the partial order.

Each of the start-up cost terms in $\delta^t(l(t),r(t))$ corresponds to the cost incurred when starting up after being offline during the periods contained in either the left subtree $L(t)$ of t, the right subtree $R(t)$ of t, or the principal subtree $S(t)$ of t (see (2.2.5)),

$$\delta^t(l(t), r(t)) = \underbrace{\mathrm{CU}^t(l(t))}_{\text{offline in } L(t)} + \underbrace{\mathrm{CU}^{t+r(t)+1}(r(t))}_{\text{offline in } R(t)} - \underbrace{\mathrm{CU}^{t+r(t)+1}(l(t)+1+r(t))}_{\text{offline in } S(t)}.$$

To simplify the notation, we generalize the definition (see (1.3.4)) of the offline length OL: for each $B \in \mathcal{B}$ and $t \in [T]$, we abbreviate

$$\mathrm{OL}(S(t)) := \mathrm{OL}^{t+r(t)+1}(s(t)), \tag{2.2.8}$$

and consequently, by (2.2.5)

$$\mathrm{OL}(L(t)) = \mathrm{OL}^t(l(t)) \quad \text{and} \quad \mathrm{OL}(R(t)) = \mathrm{OL}^{t+r(t)+1}(r(t)). \tag{2.2.9}$$

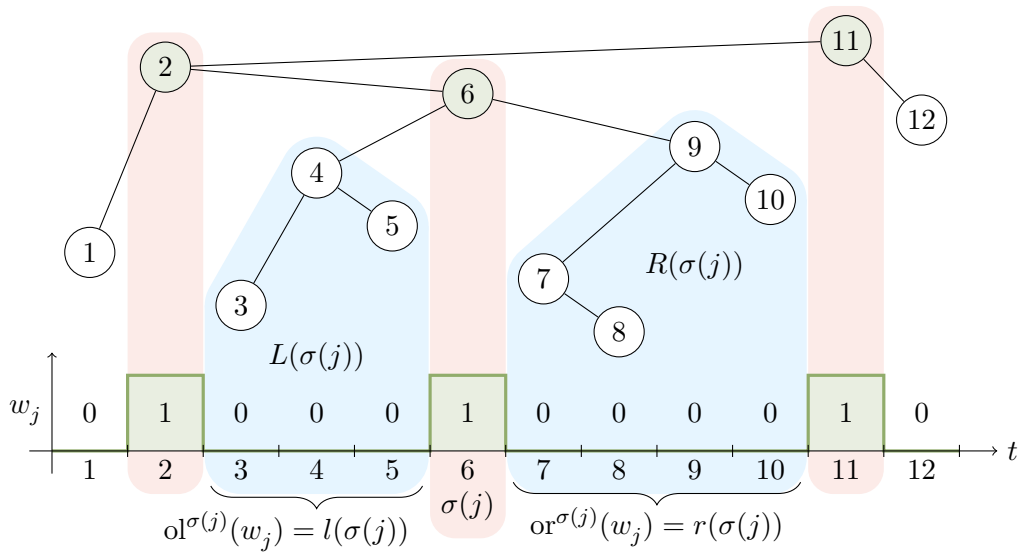


Figure 2.15: The lifting step as shown in Fig. 2.10, with a possible lifting order represented by a binary tree B. The offline lengths $\mathrm{ol}^{\sigma(j)}(w_j)$ and $\mathrm{or}^{\sigma(j)}(w_j)$ adjacent to period $\sigma(j)$ are equal to the sizes of the left and right subtree of node $\sigma(j)$ in B.

Since by definition $\mathrm{CU}^t(l) = \mathrm{CU}(\mathrm{OL}^t(l))$, it follows that

$$\delta^t(l(t), r(t)) = \mathrm{CU}(\mathrm{OL}(L(t))) + \mathrm{CU}(\mathrm{OL}(R(t))) - \mathrm{CU}(\mathrm{OL}(S(t))). \tag{2.2.10}$$

Hence we may write all lifted inequalities as follows.

Definition 2.26 A *rank-labeled* binary tree is a binary tree B on nodes $[n]$, where $n \in \mathbb{N}$, and $\mathrm{rank}(i) = i$ for all $i \in [n]$. Let $\mathcal{B}$ denote the family of all rank-labeled binary trees on $[T]$. For each $B \in \mathcal{B}$, we define the *binary tree inequality (BTI)* as

$$\mathrm{cu}^\Sigma \geq \sum_{t \in [T]} \delta^t(l(t), r(t))\, v^t,$$

using the sizes $l(t)$ or $r(t)$ of the left or right subtrees of t, respectively, and δ^t as defined in Proposition 2.17.

In the following, we confirm that the BTIs, together with the trivial inequalities $0 \leq v^t \leq 1$, define all non-trivial facets of $\mathrm{epi}(\mathrm{LCU}^\Sigma)$ by proving that

- they are valid (Lemma 2.28),
- they are fulfilled with equality by all points $(w_j, \mathrm{DCU}^\Sigma(w_j))$ encountered during the lifting process (Lemma 2.29), and
- these points are linearly independent (Theorem 2.31).

Moreover, we will show that all points not in epi(LCU^Σ) can be separated by the BTIs or the trivial inequalities $0 \le v^t \le 1$ in $\mathcal{O}(T)$.

In the following, we need to put the vertices $(v, \mathrm{DCU}^\Sigma(v))$ of the epigraph epi(LCU^Σ) into relation with the binary trees $B \in \mathcal{B}$.

Definition 2.27 For each $B \in \mathcal{B}$ with edges E and $v \in \{0,1\}^T$, define $B(v)$ as the subgraph of B induced by the 1-entries of v,

$$B(v) := (V_S, E_S) \quad \text{where} \quad V_S := \{t \in [T] \mid v^t = 1\} \quad \text{and} \quad E_S := \{e \in E \mid e \subset V_S\}.$$

Note that in general, $B(v)$ is not connected.

The next lemma proves that the binary tree inequalities are valid using the concavity of CU, which is exploited through the monotonicity of δ^t (Lemma 2.18).

Lemma 2.28 *For each $B \in \mathcal{B}$, the BTI induced by B is valid for* epi(LCU^Σ).

Proof. Due to the convexity of epi(LCU^Σ) and since a binary tree inequality bounds cu^Σ only from below, it suffices to prove that all vertices $(v, \mathrm{DCU}^\Sigma(v)) \in V^\Sigma$ of epi(LCU^Σ) fulfill all BTIs. To do so, for each $B \in \mathcal{B}$, we prove that its induced BTI is valid by induction over the number of nodes n in its induced subtree $B(v)$.

For $n = 0$, the only such point is $0 \in V^\Sigma$. Since each BTI is homogeneous, 0 fulfills all of them with equality. For $n \ge 1$, choose a leaf t in $B(v)$, i. e. a node such that the subtree of t in B does not contain any other nodes from $B(v)$. Define the vector $\tilde{v}$ as

$$\tilde{v}^{t'} := \begin{cases} v^{t'} & \text{if } t' \neq t, \\ 0 & \text{if } t' = t, \end{cases},$$

differing from v only in period t. By Proposition 2.17, we get

$$\mathrm{DCU}^\Sigma(v) = \mathrm{DCU}^\Sigma(\tilde{v}) + \delta^t(\mathrm{ol}^t(v), \mathrm{or}^t(v)).$$

These vectors, the induced subtrees, and the offline lengths $\mathrm{ol}^t(v)$, $\mathrm{or}^t(v)$ are shown in Fig. 2.16.

By the choice of t, its left subtree, which contains the nodes $t-l(t), \ldots, t-1$, does not contain any nodes from $B(v)$. Hence, $v^{t-l(t)} = \ldots = v^{t-1} = 0$, implying $\mathrm{ol}^t(v) \ge l(t)$. Since the right subtree of t does not contain any nodes from $B(v)$ either, we analogously obtain $\mathrm{or}^t(v) \ge r(t)$.

Thus, using the monotonicity of δ^t, it holds that

$$\begin{aligned} \mathrm{DCU}^\Sigma(v) &= \mathrm{DCU}^\Sigma(\tilde{v}) + \delta^t(\mathrm{ol}^t(v), \mathrm{or}^t(v)) \ge \mathrm{DCU}^\Sigma(\tilde{v}) + \delta^t(l(t), r(t)) \\ &\overset{\text{ind.hyp.}}{\ge} \sum_{t' \in [T]} \delta^{t'}(l(t'), r(t'))\, \tilde{v}^{t'} + \delta^t(l(t), r(t)) = \sum_{t' \in [T]} \delta^{t'}(l(t'), r(t'))\, v^{t'}. \end{aligned}$$ □

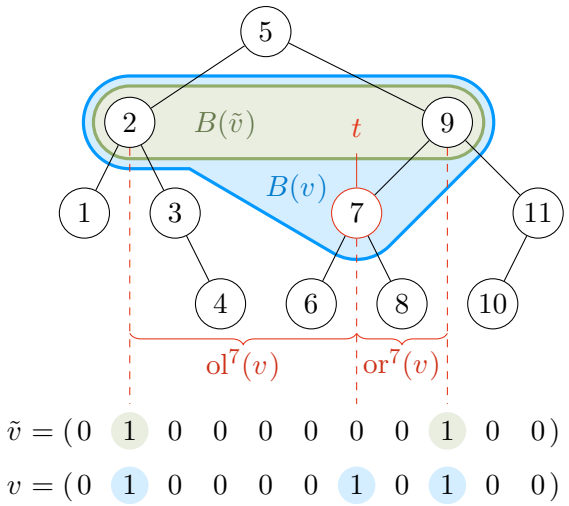


Figure 2.16: Removing the leaf $t = 7$ from the induced subgraph $B(v)$ results in the subgraph $B(\tilde{v})$ induced by $\tilde{v}$. The lengths $\mathrm{ol}^7(v)$ and $\mathrm{or}^7(v)$ denote the offline lengths before and after period 7 (see (1.3.7) and (2.2.3)).

The central argument of the proof of the last lemma is that in each step of the induction, the inequality

$$\delta^t(\mathrm{ol}^t(v), \mathrm{or}^t(v)) \geq \delta^t(l(t), r(t))$$

holds due to $\mathrm{ol}^t(v) \geq l(t)$ and $\mathrm{or}^t(v) \geq r(t)$. Assume that, for a given binary tree B and a vertex of $\mathrm{epi}(\mathrm{LCU}^\Sigma)$, this inequality is fulfilled with equality in each induction step. Then the vertex also fulfills the BTI induced by B with equality. We characterize such vertices in the next lemma.

Lemma 2.29 *For each $B \in \mathcal{B}$, $(v, \mathrm{DCU}^\Sigma(v)) \in V^\Sigma$, if the induced subgraph $B(v)$ is a tree containing $\mathrm{root}(B)$, then $(v, \mathrm{DCU}^\Sigma(v))$ fulfills the BTI induced by B with equality.*

Proof. The condition that the induced subtree $B(v)$ contains $\mathrm{root}(B)$ is fulfilled iff for each node $t \in B(v)$ all ancestors of t also lie in $B(v)$.

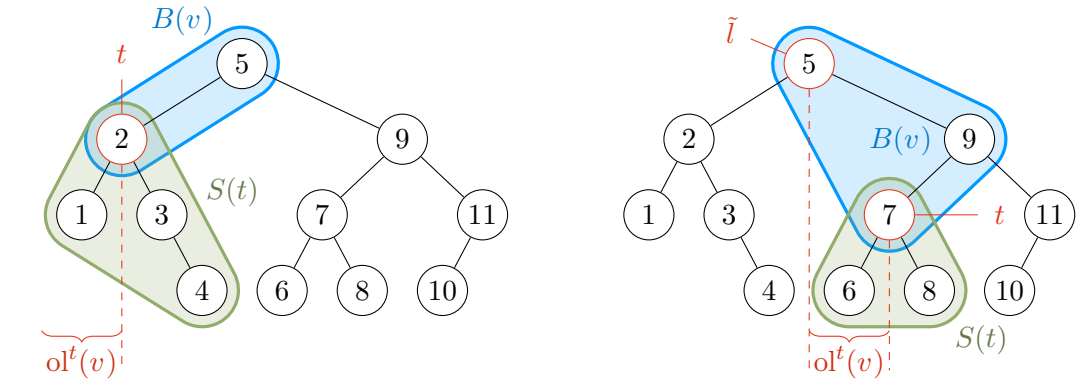
Analogously to the proof of Lemma 2.28, we show that $(v, \mathrm{DCU}^\Sigma(v))$ fulfills the binary tree inequality induced by $B \in \mathcal{B}$ by induction over the number of nodes n in the subgraph $B(v)$.

For $n = 1$, the only such vertex is $v \in V^\Sigma$ with the single non-zero coordinate $v^r = 1$, where $r := \mathrm{root}(B)$. Clearly,

$$\mathrm{DCU}(v) = \mathrm{CU}^r(r-1) = \delta^r(r-1, T-r) = \delta^r(l(r), r(r)) = \sum_{t \in [T]} \delta^t(l(t), r(t)) v^t.$$

For $n \geq 2$, let t be a leaf of $B(v)$ except the root, and define $\tilde{v} := v - u_t$ as in the proof of Lemma 2.28. Note that since $B(v)$ is a tree and $\text{root}(B) \in B(v)$, the subgraph $B(\tilde{v})$ is also a tree with $n-1$ nodes, which contains $\text{root}(B)$.

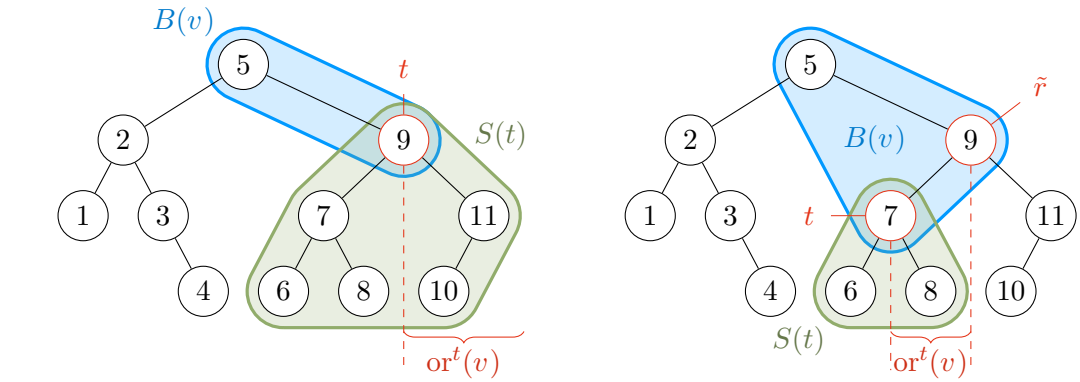
Since no nodes in the principal subtree of t, except t itself, are in $B(v)$, analogously to the proof of Lemma 2.28, it holds that $\text{ol}^t(v) \geq l(t)$ and $\text{or}^t(v) \geq r(t)$. We show that these bounds are sharp by examining two cases for $\text{ol}^t(v) = l(t)$ (see Fig. 2.17) and two cases for $\text{or}^t(v) = r(t)$ (see Fig. 2.18).



$v = (0\ 1\ 0\ 0\ 1\ 0\ 0\ 0\ 0\ 0\ 0)$ — a) t is a top-left node

$v = (0\ 0\ 0\ 0\ 1\ 0\ 1\ 0\ 1\ 0\ 0)$ — b) t is a top-left node in $R(\tilde{l})$

Figure 2.17: Number of offline periods $\text{ol}^t(v)$ in v before period t, which is either bounded by the start of the model or by the first left ancestor $\tilde{l}$ of node t



$v = (0\ 0\ 0\ 0\ 1\ 0\ 0\ 0\ 1\ 0\ 0)$ — a) t is a top-right node

$v = (0\ 0\ 0\ 0\ 1\ 0\ 1\ 1\ 1\ 0\ 0)$ — b) t is a top-right node in $L(\tilde{r})$

Figure 2.18: Number of offline periods $\text{or}^t(v)$ in vector v after period t, which is either bounded by the end of the model or by the first right ancestor $\tilde{r}$ of node t

If t is a top-left node in B, then $t = l(t) + 1$ (see Corollary 2.24), implying

$$l(t) \leq \mathrm{ol}^t(v) \leq t - 1 = l(t),$$

and thus $\mathrm{ol}^t(v) = l(t)$. Otherwise, there exists a node $\tilde{l}$ such that t is a top-left node in the right subtree $R(\tilde{l})$ of $\tilde{l}$. Since $\tilde{l}$ is an ancestor of t, we have $\tilde{l} \in B(v)$. Hence,

$$l(t) \leq \mathrm{ol}^t(v) \leq t - \tilde{l} - 1 \overset{\text{Corollary 2.24}}{=} \left(\tilde{l} + l(t) + 1\right) - \tilde{l} - 1 = l(t).$$

If t is a top-right node in B, then $t = T - r(t)$, implying $r(t) \leq \mathrm{or}^t(v) \leq T - t = r(t)$. Else, t is a top-right node in the left subtree of a node $\tilde{r}$ (see Corollary 2.23). Since $\tilde{r}$ is an ancestor of t, we have $\tilde{r} \in B(v)$, and thus again

$$r(t) \leq \mathrm{or}^t(v) \leq \tilde{r} - t - 1 \overset{\text{Lemma 2.22}}{=} r(t).$$

In conclusion, $\mathrm{ol}^t(v) = l(t)$ and $\mathrm{or}^t(v) = r(t)$ holds. Analogous to Lemma 2.28,

$$\begin{aligned}\mathrm{DCU}^\Sigma(v) &= \mathrm{DCU}^\Sigma(\tilde{v}) + \delta^t(\mathrm{ol}^t(v), \mathrm{or}^t(v)) = \mathrm{DCU}^\Sigma(\tilde{v}) + \delta^t(l(t), r(t)) \\ &\overset{\text{ind.hyp.}}{=} \sum_{t' \in [T]} \delta^{t'}(l(t'), r(t'))\, \tilde{v}^{t'} + \delta^t(l(t), r(t)) = \sum_{t' \in [T]} \delta^{t'}(l(t'), r(t'))\, v^{t'}. \qquad \square\end{aligned}$$

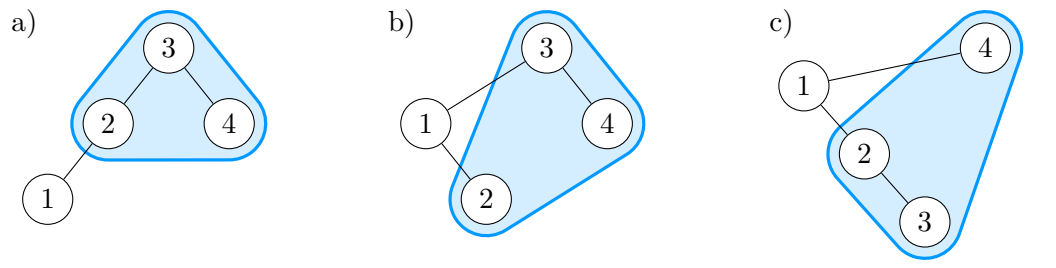
In general, a vertex $(v, \mathrm{DCU}^\Sigma(v))$ may lie on a facet induced by a binary tree B which does not meet the requirements of the preceding lemma, e. g. for a linear start-up cost function. The next result shows that this does not happen for strictly concave start-up cost functions CU, which by Lemma 2.18 lead to strictly increasing functions $\delta^t(l, r)$.

Consider the model with parameters $\mathrm{L}^1 = \ldots = \mathrm{L}^4 = 1$, $\mathrm{PDT} = 0$, and the linear start-up cost function $\mathrm{CU}(l) = l$. This leads to the BTI coefficients

$$\delta^t(l(t), r(t)) = \begin{cases} -1 & \text{if } t \text{ is not a top-right node,} \\ l(t) & \text{else.} \end{cases}$$

The point $(v, \mathrm{DCU}^\Sigma(v))$ with $v = (0, 1, 1, 1)$, $\mathrm{DCU}^\Sigma(v) = l$ lies on the facets induced by all three binary trees shown in Fig. 2.19, since it fulfills their BTIs with equality. Despite this, its induced subtree $B(v)$ is not a tree in the case of the binary trees b) and c). Note moreover that the binary trees a) and b) are not equal but still induce the same facet.

Lemma 2.30 *Let* CU *be strictly concave,* $B \in \mathcal{B}$*, and* $(v, \mathrm{DCU}^\Sigma(v)) \in V^\Sigma$*. If* $(v, \mathrm{DCU}^\Sigma(v))$ *fulfills the BTI induced by* B *with equality, then either* $v = 0$ *or the induced subgraph* $B(v)$ *is a tree containing* $\mathrm{root}(B)$*.*



$\mathrm{cu}^\Sigma \geq 2v^3 - v^1 - v^2$ $\mathrm{cu}^\Sigma \geq 2v^3 - v^1 - v^2$ $\mathrm{cu}^\Sigma \geq 3v^4 - v^1 - v^2 - v^3$

Figure 2.19: Three binary trees with respective BTI. Note that for $v = (0, 1, 1, 1)$, the induced subtree $B(v)$ (highlighted in blue) is a tree in case a), but not in cases b) and c), while v fulfills all three BTIs with equality.

Proof. Again we prove the statement by induction over the number of nodes n in the subgraph $B(v)$. The case $n = 0$ is fulfilled trivially.

For $n \geq 1$, analogously to the proof of Lemma 2.28, choose a node t such that the subtree of t in B does not contain any other nodes from $B(v)$. This choice implies $\mathrm{ol}^t(v) \geq l(t)$ and $\mathrm{or}^t(v) \geq r(t)$. Lemma 2.18 shows that $\delta^t(l, r)$ is strictly increasing in l and r if CU is strictly concave, and thus

$$\delta^t(\mathrm{ol}^t(v), \mathrm{or}^t(v)) \geq \delta^t(l(t), r(t)).$$

For $\tilde{v} := v - u_t$, Proposition 2.17 yields

$$\mathrm{DCU}^\Sigma(v) - \delta^t(\mathrm{ol}^t(v), \mathrm{or}^t(v)) = \mathrm{DCU}^\Sigma(\tilde{v})$$

which, since $(\tilde{v}, \mathrm{DCU}^\Sigma(\tilde{v})) \in \mathrm{epi}(\mathrm{LCU}^\Sigma)$, may be bounded by

$$\mathrm{DCU}^\Sigma(\tilde{v}) \geq \sum_{t' \in [T]} \delta^{t'}(l(t'), r(t'))\tilde{v}^{t'} = \sum_{t' \in [T]} \delta^{t'}(l(t'), r(t'))v^{t'} - \delta^t(l(t), r(t))$$

which, by the choice of v and t, equals

$$\mathrm{DCU}^\Sigma(\tilde{v}) = \mathrm{DCU}^\Sigma(v) - \delta^t(l(t), r(t)) \geq \mathrm{DCU}^\Sigma(v) - \delta^t(\mathrm{ol}^t(v), \mathrm{or}^t(v)) = \mathrm{DCU}^\Sigma(\tilde{v}).$$

Therefore, the above inequality is fulfilled with equality, and we conclude

$$\delta^t(\mathrm{ol}^t(v), \mathrm{or}^t(v)) = \delta^t(l(t), r(t)) \quad \text{and} \quad \mathrm{DCU}^\Sigma(\tilde{v}) = \sum_{t' \in [T]} \delta^{t'}(l(t'), r(t'))\tilde{v}^{t'}.$$

Firstly, as CU is assumed to be strictly concave, δ^t is strictly increasing in l and r (see Proposition 2.17). Recalling that $\mathrm{ol}^t(v) \geq l(t)$ and $\mathrm{or}^t(v) \geq r(t)$ by the choice of t, we infer $\mathrm{ol}^t(v) = l(t)$ and $\mathrm{or}^t(v) = r(t)$.

Secondly, since $|B(\tilde{v})| = |B(v)| - 1$, the induction hypothesis states that either $\tilde{v} = 0$ or $B(\tilde{v})$ is a tree containing root(B). If $\tilde{v} = 0$, then

$$s(t) = l(t) + 1 + r(t) = \mathrm{ol}^t(v) + 1 + \mathrm{or}^t(v) = t - 1 + 1 + (T - t) = T,$$

hence t is the root of B and $B(v)$ is the single-noded tree consisting of $t = \mathrm{root}(B)$.

Else $\tilde{v} \neq 0$, implying that $B(\tilde{v})$ is a tree containing root(B). Therefore $t \neq \mathrm{root}(B)$, meaning that t has a parent $p(t)$ in $B(\tilde{v})$. By (2.2.6), if t is a right child, then

$$p(t) = t - l(t) - 1 = t - \mathrm{ol}^t(v) - 1,$$

and if t is a left child, then

$$p(t) = t + r(t) + 1 = t + \mathrm{or}^t(v) + 1.$$

In both cases, $v^{p(t)} = 1$ by definition of $\mathrm{ol}^t(v)$ or $\mathrm{or}^t(v)$ ((1.3.7) and (2.2.3)), and thus $p(t) \in B(v)$. Given that $B(\tilde{v})$ is a tree containing root(B), we obtain that $B(v)$ must be a tree containing root(B) too. □

The vectors v encountered during a lifting process gain a non-zero entry in every step, and are thus linearly independent.

Theorem 2.31 *All binary tree inequalities induce facets of* epi(LCU^Σ).

Proof. For each binary tree $B \in \mathcal{B}$, the induced BTI is valid for epi(LCU^Σ) (see Lemma 2.28).

Choose a permutation σ of $[T]$ such that the nodes $\sigma(t)$ are ordered by their depth $d(\sigma(t))$ in B, i. e. such that for each $t \in [T-1]$ it holds that $d(\sigma(t)) \leq d(\sigma(t+1))$. For each $j \in [T]$, define the vector

$$\forall\, t \in [T]: \qquad w_j^t := \begin{cases} 1 & \text{if } \sigma^{-1}(t) \leq j, \\ 0 & \text{else,} \end{cases}$$

which induces the subgraph $B(w_j)$ containing the nodes $\{\sigma(1), \ldots, \sigma(j)\}$. $B(w_j)$ fulfills the requirements of Lemma 2.29, and thus the vertex $(w_j, \mathrm{DCU}^\Sigma(w_j))$ fulfills the BTI with equality.

Denoting the permutation matrix associated to σ with Π_σ, we obtain

$$\Pi_\sigma \cdot \begin{pmatrix} w_1 & w_2 & \ldots & w_T \end{pmatrix} = \begin{pmatrix} 1 & \cdots & \cdots & 1 \\ 0 & 1 & \cdots & 1 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 \end{pmatrix},$$

and thus $0, (w_1, \mathrm{DCU}^\Sigma(w_1))$, …, $(w_T, \mathrm{DCU}^\Sigma(w_T))$ are affine linearly independent. Since they amount to $T+1$ points on the face F of epi(LCU^Σ) induced by the BTI, F must be a facet. □

2.2.4 Sufficiency of the BTIs

In this subsection, we prove that the binary tree inequalities (Definition 2.26), together with the trivial facets $0 \leq v^t \leq 1$, are sufficient for an $\mathcal{H}$-representation of epi(LCU^Σ). To do so we show that the facets induced by all the BTIs fully describe the lower boundary of epi(LCU^Σ). To this end, we extend Lemma 2.29 on the facets containing a vertex $(v, \mathrm{DCU}^\Sigma(v)) \in V^\Sigma$ lies to points with $v \in [0,1]^T$.

Lemma 2.29 provides the sufficient (but not necessary) condition for the vertices of epi(LCU^Σ) lying on a certain facet: "A vertex $(v, \mathrm{DCU}^\Sigma(v))$ lies on the facet corresponding to a binary tree $B \in \mathcal{B}$, if the induced subgraph $B(v)$ is a tree containing root(B)." In other words, each node $t \in B(v)$ needs to be connected to root(B) within $B(v)$. The unique path P_t from t to root(B) consists of its ancestors,

$$P_t = t \to p(t) \to p(p(t)) \to p(p(p(t))) \to \ldots \to \mathrm{root}(B).$$

By definition, $t \in B(v) \Leftrightarrow v^t = 1$. So, $B(v)$ contains the paths P_t for all $t \in B(v)$ iff

$$v^t = 1 \Rightarrow v^{p(t)} = 1 \quad \text{for all } t \in [T] \setminus \{\mathrm{root}(B)\},$$

which, since $v \in \{0,1\}^T$, is equivalent to

$$v^t \leq v^{p(t)} \quad \text{for all } t \in [T] \setminus \{\mathrm{root}(B)\}. \tag{2.2.11}$$

Coincidentally, this condition is also important when searching efficiently in a point set in the Cartesian plane, and is denoted by "B is a Cartesian tree for v" in [Vui80]. We use a definition adapted to our purposes, equivalent to the recursive definition of Cartesian trees in [GBT84]:

Definition 2.32 For each $v \in \mathbb{R}^T$, a binary tree $B \in \mathcal{B}$ with

$$v^t \leq v^{p(t)} \quad \text{for all } t \in [T] \setminus \{\mathrm{root}(B)\}$$

is called a *Cartesian tree* for v.

For a rank-labeled binary tree $B \in \mathcal{B}$, the condition $x_{\mathrm{rank}(v)} \leq x_{\mathrm{rank}(p(v))}$ simplifies to $x_v \leq x_{p(v)}$.

The following lemma shows that condition (2.2.11) applies to any $v \in [0,1]^T$ as well.

Lemma 2.33 *If $v \in [0,1]^T$ and $B \in \mathcal{B}$ is a Cartesian tree for v, then $(v, \mathrm{LCU}^\Sigma(v))$ lies on the facet of* epi(LCU^Σ) *induced by B.*

Proof. For each $v \in [0,1]^T$ and each Cartesian tree $B \in \mathcal{B}$ for v, we give a set of vertices $(\tilde{v}_j, \mathrm{DCU}^\Sigma(\tilde{v}_j))$ of epi(LCU^Σ) which lie on the facet of epi(LCU^Σ) induced by

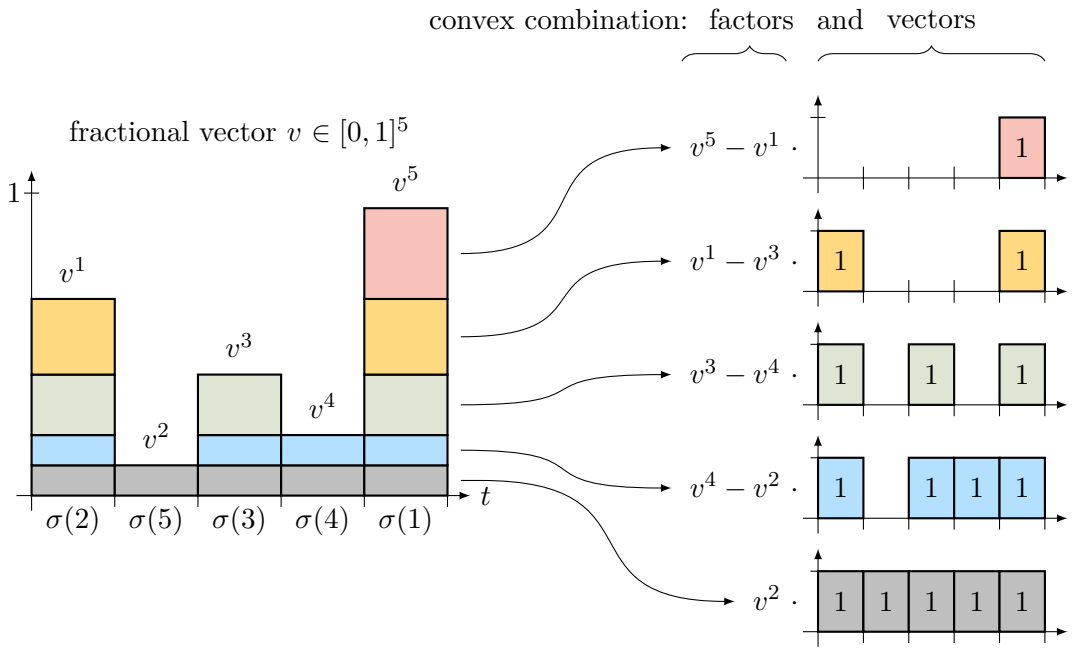


Figure 2.20: Decomposition of a fractional vector $v \in [0,1]^5$ into binary vectors $w_j \in \{0,1\}^5$, such that all points $(w_j, \mathrm{LCU}^{\Sigma}(w_j))$ lie on the same facet.

B, and whose vectors $\tilde{v}_j$ can be convexly combined to v. This convex combination is depicted schematically in Fig. 2.20.

Choose a permutation σ which orders the indices $t \in [T]$ by decreasing value of v^t, and increasing depth $d(t)$ in B if the values v^t are equal, i. e. such that

$$\forall\, j_1, j_2 \in [T], j_1 < j_2 : \qquad \begin{aligned} & v^{\sigma(j_1)} \geq v^{\sigma(j_2)} \quad \text{and} \\ & v^{\sigma(j_1)} = v^{\sigma(j_2)} \quad \Rightarrow \quad d(\sigma(j_1)) \leq d(\sigma(j_2)). \end{aligned}$$

For all $j \in [T]$, define the vectors $\tilde{v}_j \in \{0,1\}^T$ and the coefficients λ_j as

$$\tilde{v}_j^t := \begin{cases} 1 & \text{if } \sigma^{-1}(t) \leq j, \\ 0 & \text{else,} \end{cases} \qquad \text{and} \qquad \lambda_j := \begin{cases} v^{\sigma(j)} - v^{\sigma(j+1)} & \text{for } j \leq T-1, \\ v^{\sigma(T)} & \text{else.} \end{cases}$$

By definition of σ, all λ_j are non-negative, and obviously they sum up to $v^{\sigma(1)}$. Hence, for each $t \in [T]$, it holds that

$$\sum_{j=1}^{T} \lambda_j \tilde{v}_j^t = \sum_{j=\sigma^{-1}(t)}^{T} \lambda_j = \sum_{j=\sigma^{-1}(t)}^{T-1} (v^{\sigma(j)} - v^{\sigma(j+1)}) + v^{\sigma(T)} = v^{\sigma(\sigma^{-1}(t))} = v^t.$$

Therefore, the vector v indeed is a convex combination of the vectors $\tilde{v}_j$ and 0,

$$v = \sum_{j=1}^{T} \lambda_j \tilde{v}_j + (1 - v^{\sigma(1)})\, 0.$$

We proceed by showing that the vertices $(\tilde{v}_j, \mathrm{DCU}^{\Sigma}(\tilde{v}_j))$ lie on the facet induced by B. For each $j \in [T]$, the induced subgraph $B(\tilde{v}_j)$ is the induced subgraph of B on nodes $\{t \in [T] \mid \sigma^{-1}(t) \leq j\}$. Consider an arbitrary node $t \in B(\tilde{v}_j)$ and its parent $p(t)$. Due to the definition of B, both

$$v^t \leq v^{p(t)} \quad \text{and} \quad d(t) > d(p(t)) \quad \text{hold.}$$

So, by choice of the permutation σ, we have

$$\sigma^{-1}(p(t)) < \sigma^{-1}(t) \leq j,$$

and therefore $p(t) \in B(\tilde{v}_j)$ too. As noted in the motivation preceding this lemma, this condition is equivalent to "$B(\tilde{v}_j)$ contains the root of B and is a tree", and thus Lemma 2.29 implies that $(\tilde{v}_j, \mathrm{DCU}^{\Sigma}(\tilde{v}_j))$ lies on the facet induced by B.

Being a convex combination of such points, the point $(v, \sum_{j=1}^{T} \lambda_j \mathrm{DCU}^{\Sigma}(\tilde{v}_j))$ lies on the same facet, implying $\mathrm{LCU}^{\Sigma}(v) = \sum_{j=1}^{T} \lambda_j \mathrm{DCU}^{\Sigma}(\tilde{v}_j)$ and hence that $(v, \mathrm{LCU}^{\Sigma}(v))$ lies on the facet of $\mathrm{epi}(\mathrm{LCU}^{\Sigma})$ induced by B. □

Note that Lemma 2.33 is a generalization of Lemma 2.29, since for discrete vectors $v \in \{0,1\}^T$ the induced subgraph $B(v)$ is a Cartesian tree for v iff it is empty or a tree containing $\mathrm{root}(B)$.

By Lemma 2.33, each BTI defines a part of the lower boundary of $\mathrm{epi}(\mathrm{LCU}^{\Sigma})$, and thereby also a part of LCU^{Σ}.

Corollary 2.34 *If $B \in \mathcal{B}$ is a Cartesian tree for $v \in [0,1]^T$, then*

$$\mathrm{LCU}^{\Sigma}(v) = \sum_{t \in [T]} \delta^t(l(t), r(t))\, v^t.$$

Since epigraphs are characterized by their lower boundary, this is equivalent to:

Corollary 2.35 *If $(v, \mathrm{cu}^{\Sigma}) \in [0,1]^T \times \mathbb{R}$ and $B \in \mathcal{B}$ is a Cartesian tree for v, then $(v, \mathrm{cu}^{\Sigma})$ lies in $\mathrm{epi}(\mathrm{LCU}^{\Sigma})$ iff it fulfills the BTI induced by B.*

Proof.

$$(v, \mathrm{cu}^{\Sigma}) \in \mathrm{epi}(\mathrm{LCU}^{\Sigma}) \Leftrightarrow \mathrm{cu}^{\Sigma} \geq \mathrm{LCU}^{\Sigma}(v) = \sum_{t \in [T]} \delta^t(l(t), r(t)) v^t$$
$$\Leftrightarrow (v, \mathrm{cu}^{\Sigma}) \text{ fulfills the BTI induced by } B.$$ □

Finally, since there exists a rank-labeled Cartesian tree for each $v \in [0,1]^T$ ([GBT84]), the BTIs and the trivial inequalities $0 \leq v^t \leq 1$ completely describe $\mathrm{epi}(\mathrm{LCU}^\Sigma)$.

Theorem 2.36

$$\mathrm{epi}(\mathrm{LCU}^\Sigma) = \left\{ \begin{array}{ll} (v, \mathrm{cu}^\Sigma) \in \mathbb{R}^{T+1} : & \\ \qquad \mathrm{cu}^\Sigma \geq \sum\limits_{t \in [T]} \delta^t(l(t), r(t))\, v^t, & B \in \mathcal{B} \\ \qquad 0 \leq v^t \leq 1, & t \in [T] \end{array} \right\}. \tag{2.2.12}$$

It remains to discuss whether this $\mathcal{H}$-representation is irredundant. As noted in the paragraph before Lemma 2.30, if the start-up cost function CU is not strictly concave, e. g. if CU is linear, then a vertex $(v, \mathrm{DCU}^\Sigma(v))$ may fulfill a BTI induced by a binary tree $B \in \mathcal{B}$ which is not a Cartesian tree for v. This leads to multiple binary trees inducing the same facet of $\mathrm{epi}(\mathrm{LCU}^\Sigma)$, rendering the above $\mathcal{H}$-representation redundant.

However, if CU is strictly concave, Lemma 2.30 describes all vertices $(v, \mathrm{DCU}^\Sigma(v))$ on a facet. By showing that different binary trees induce facets with different vertices, we prove that the given $\mathcal{H}$-representation is irredundant:

Theorem 2.37 *If* CU *is strictly concave, then the $\mathcal{H}$-representation of* $\mathrm{epi}(\mathrm{LCU}^\Sigma)$ *given in* (2.2.12) *is irredundant.*

Proof. For each $B_1, B_2 \in \mathcal{B}$ with $B_1 \neq B_2$, we construct a vertex $(v, \mathrm{DCU}^\Sigma(v)) \in V^\Sigma$ such that B_1 is a Cartesian tree for v, but B_2 is not a Cartesian tree for v. Then, by Lemma 2.29, $(v, \mathrm{DCU}^\Sigma(v))$ lies on the facet induced by B_1, and by Lemma 2.30, $(v, \mathrm{DCU}^\Sigma(v))$ does not lie on the facet induced by B_2, proving that the induced facets are not equal.

If $\mathrm{root}(B_1) \neq \mathrm{root}(B_2)$, then $(u_{\mathrm{root}(B_1)}, \mathrm{DCU}^\Sigma(u_{\mathrm{root}(B_1)}))$ clearly is such a vertex. Otherwise, we have $r := \mathrm{root}(B_1) = \mathrm{root}(B_2)$, and therefore the edge sets E_1 of B_1 and E_2 of B_2 differ. Each edge in B_1 connects a node t to its parent $p^{B_1}(t)$. Choose $e = \{t, p^{B_1}(t)\} \in E_1 \setminus E_2$ with minimal $d(t)$. Since $e \notin E_2$ and $t \neq r$, t has different parents $p_1 := p^{B_1}(t)$ and $p_2 := p^{B_2}(t)$ in B_1 and B_2. Now, denote the path from r to p_1 in B_1 by P, and define $v \in \{0,1\}^T$ as

$$\forall\, t' \in [T]: \qquad v^{t'} := \begin{cases} 1 & \text{if } t' = t \text{ or } t' \in P, \\ 0 & \text{else.} \end{cases}$$

By definition, B_1 is a Cartesian tree for v.

We conclude this proof by showing that $p_2 \notin P$, which implies $v^{p_2} = 0$, and thereby that B_2 is not a Cartesian tree for v: Assume $p_2 \in P$. Then, since choosing e such that $d(t)$ is minimal, we have $P \subset B_2$, and therefore p_1 is a descendant of p_2 in both B_1 and B_2. Since t is a child of p_2 in B_2 and $t \notin P$, it holds that t and p_1 are

separated to the subtrees $L^{B_2}(p_2)$ and $R^{B_2}(p_2)$ of p_2, and thus either $t < p_2 < p_1$ or $p_1 < p_2 < t$. However, t is a child of p_1 in B_1, and thus both, p_1 and t, either lie in the left subtree $L^{B_1}(p_2)$ or right subtree $R^{B_1}(p_2)$ of p_2 in B_1. Since this implies either $t, p_1 < p_2$ or $p_2 < t, p_1$, we obtain a contradiction, proving $p_2 \notin P$. □

The number of different binary trees on T nodes is given by the T-th *Catalan number* C_T, which is asymptotically equivalent to ${}^{4^T}/_{(T^{\frac{3}{2}}\sqrt{\pi})}$ (see [Ros00]). This allows us to count the number of facets of epi(LCU$^\Sigma$):

Corollary 2.38 *If* CU *is strictly concave, the number of facets of* epi(LCU$^\Sigma$) *is*

$$C_T + 2T \sim \frac{4^T}{T^{\frac{3}{2}}\sqrt{\pi}},$$

where C_T denotes the T-th Catalan number.

2.2.5 Separation

Generally, the $\mathcal{H}$-representation of epi(LCU$^\Sigma$) given in the last subsection is of exponential size, and thus not (directly) suitable for computational purposes. This is overcome by a cutting plane approach based on an exact separation algorithm for epi(LCU$^\Sigma$) presented in this section. Assuming that the start-up cost function CU can be evaluated in $\mathcal{O}(1)$, e. g. for the generally used exponential start-up cost function $\mathrm{CU}(l) = A\mathrm{e}^{-\lambda l} + B$, we show that this separation algorithm has a running time of $\mathcal{O}(T)$. Otherwise, the running time would change proportionally.

Lemma 2.33 states that a point $(v, \mathrm{cu}^\Sigma) \in [0,1]^T \times \mathbb{R}$ lies in epi(LCU$^\Sigma$) iff the BTI induced by the Cartesian tree for v is fulfilled. Thus, the idea of the separation algorithm for epi(LCU$^\Sigma$) is to find a Cartesian tree for the vector v, and construct its induced BTI.

A linear-time algorithm for finding a Cartesian tree for v has already been given in [GBT84], and is revisited as Algorithm A.0.1 in Appendix A, denoted as *FindCartesianTree*. In summary, this algorithm starts with a tree with a single node 1, and iteratively adds the remaining nodes $t \in [2..T]$. The key observation is that the node t in each iteration must be added such that it results to be

- the last top-right node (as to receive the correct rank t), and to be
- beneath all top-right nodes t' with $v^{t'} > v^t$ and above all top-right nodes t' with $v^{t'} < v^t$.

The algorithm represents the resulting Cartesian tree by its left and right children $\mathrm{llink}(t)$ and $\mathrm{rlink}(t)$ of each node $t \in [T]$. To construct the induced BTI, its coefficients

$$\alpha_t = \delta^t(l(t), r(t)) = \mathrm{CU}(\mathrm{OL}(L(t))) + \mathrm{CU}(\mathrm{OL}(R(t))) - \mathrm{CU}(\mathrm{OL}(S(t))) \qquad (2.2.13)$$

need to be computed.

This requires the subtree sizes $l(t)$, $r(t)$ and $s(t)$, which due to (2.2.5) and (2.2.6) equal

$$l(t) = s(\text{llink}(t)) = l(\text{llink}(t)) + 1 + r(\text{llink}(t)) = l(\text{llink}(t)) + t - \text{llink}(t), \tag{2.2.14}$$
$$r(t) = s(\text{rlink}(t)) = l(\text{rlink}(t)) + 1 + r(\text{rlink}(t)) = \text{rlink}(t) - t + r(\text{rlink}(t)), \tag{2.2.15}$$
$$s(t) = l(t) + 1 + r(t). \tag{2.2.16}$$

By the definition of the offline lengths $\text{OL}^t(l)$ (see (1.3.3)), we have

$$\forall\, t \in [T], l_1 \in [0\,..\,t-1], l_2 \in [t - l_1 - 1] : \quad \text{OL}^t(l_1 + l_2) = \text{OL}^t(l_1) + \text{OL}^{t-l_1}(l_2),$$

and the desired offline lengths $\text{OL}(L(t))$, $\text{OL}(R(t))$ and $\text{OL}(S(t))$ may be derived from (2.2.8),(2.2.9) for each $t \in [T]$ as

$$\text{OL}(L(t)) = \text{OL}^t(l(t)) = \text{OL}^t(t-1) - \begin{cases} \text{OL}^{t-l(t)}(t - l(t) - 1) & \text{if } l(t) < t - 1, \\ 0 & \text{else.} \end{cases} \tag{2.2.17}$$
$$\text{OL}(R(t)) = \text{OL}^{t+r(t)+1}(r(t)) = \text{OL}^{t+r(t)+1}(t + r(t)) - \text{OL}^{t+1}(t), \tag{2.2.18}$$
$$\text{OL}(S(t)) = \text{OL}^{t+r(t)+1}(s(t)) = \text{OL}(L(t)) + \text{OL}(R(t)) + \text{L}^t. \tag{2.2.19}$$

Using (2.2.13)-(2.2.19), the next result shows that the coefficients α_t of the induced BTI are computable in linear time.

Proposition 2.39 *Algorithm 2.2.1 solves the separation problem for* $\text{epi}(\text{LCU}^{\Sigma})$ *in* $\mathcal{O}(T)$.

Algorithm 2.2.1: SeparateBTI

Input : Point $(v, \mathrm{cu}^{\Sigma}) \in [0,1]^T \times \mathbb{R}$
Output : $(v, \mathrm{cu}^{\Sigma}) \in \mathrm{epi}(\mathrm{LCU}^{\Sigma})$, or a separating inequality.

$B \leftarrow \mathrm{FindCartesianTree}(v)$;

for $t = 1, \ldots, T$ **do**
  $l(t) := \begin{cases} l(\mathrm{llink}(t)) + t - \mathrm{llink}(t) & \text{if } \mathrm{llink}(t) \neq \emptyset \\ 0 & \text{else;} \end{cases}$
for $t = T, \ldots, 1$ **do**
  $r(t) := \begin{cases} \mathrm{rlink}(t) - t + r(\mathrm{rlink}(t)) & \text{if } \mathrm{rlink}(t) \neq \emptyset \\ 0 & \text{else;} \end{cases}$
  $s(t) := r(t) + l(t) + 1$;

$\mathrm{OL}^1(0) := \mathrm{PDT}$;
for $t = 2, \ldots, T$ **do**
  $\mathrm{OL}^t(t-1) := \mathrm{OL}^{t-1}(t-2) + \mathrm{L}^{t-1}$;
for $t = 1, \ldots, T$ **do**
  $\mathrm{OL}(L(t)) := \mathrm{OL}^t(t-1) - \begin{cases} \mathrm{OL}^{t-l(t)}(t-l(t)-1) & \text{if } l(t) < t-1, \\ 0 & \text{else.} \end{cases}$
  $\mathrm{OL}(R(t)) := \mathrm{OL}^{t+r(t)+1}(t+r(t)) - \mathrm{OL}^{t+1}(t)$;
  $\mathrm{OL}(S(t)) := \mathrm{OL}(L(t)) + \mathrm{L}^t + \mathrm{OL}(R(t))$;

for $t \in [T]$ **do**
  $\alpha_t := \begin{cases} \mathrm{CU}(\mathrm{OL}(L(t))) + \mathrm{CU}(\mathrm{OL}(R(t))) - \mathrm{CU}(\mathrm{OL}(S(t))) & \text{if } t + r(t) < T, \\ \mathrm{CU}(\mathrm{OL}(L(t))) & \text{else.} \end{cases}$
if $\mathrm{cu}^{\Sigma} \geq \sum_{t \in [T]} \alpha_t v^t$ **then**
  return $(v, \mathrm{cu}^{\Sigma}) \in \mathrm{epi}(\mathrm{LCU}^{\Sigma})$;
else
  return $(v, \mathrm{cu}^{\Sigma})$ may be separated from $\mathrm{epi}(\mathrm{LCU}^{\Sigma})$ by the BTI with coefficients α_t;

2.3 The Start-up Costs in All Periods

This section investigates the epigraph of the start-up cost function of all periods. Its $\mathcal{H}$-representation is exponential, which we show by identifying an exponential class of facet inducing inequalities.

We assume that $\mathrm{CU}(l) > 0$ for $l > 0$, which holds for virtually all types of units. While the results are straightforward to amend for $\mathrm{CU}(l) = 0$, this would introduce technical details.

The *discrete start-up cost function of all periods* is defined as the combination of the discrete start-up cost functions DCU^t in (2.1.1), which model the start-up costs in each individual period t,

$$\mathrm{DCU}: \quad \{0,1\}^T \to \mathbb{R}^T_{\geq 0}, \quad v \;\mapsto \Big(\mathrm{DCU}^1(v), \ldots, \mathrm{DCU}^T(v)\Big).$$

Analogous to DCU^t and DCU^Σ, the convex hull of the epigraph of DCU is a polyhedron with vertices $(v, \mathrm{DCU}(v))$,

$$\mathrm{conv}(\mathrm{epi}(\mathrm{DCU})) = \mathrm{conv}\left(\left\{(v, \mathrm{DCU}(v)) \;\middle|\; v \in \{0,1\}^T\right\}\right) + \mathrm{pos}\{u_{T+t} \mid t \in [T]\}, \quad (2.3.1)$$

In contrast to DCU^t and DCU^Σ however, in general no convex hull function of DCU exists, i. e. there exists no function LCU such that $\mathrm{epi}(\mathrm{LCU}) = \mathrm{conv}(\mathrm{epi}(\mathrm{DCU}))$, as demonstrated by the following example:

Consider a model with parameters $T = 3$, $\mathrm{L}^1 = \mathrm{L}^2 = \mathrm{L}^3 = 1$, $\mathrm{PDT} = 1$, and $\mathrm{CU}(l) = 16(1 - 2^{-l})$. The operational schedule $v = (1/2, 1/2, 1)$ may be decomposed convexly into $v = 1/2(0,0,1) + 1/2(1,1,1)$ and $v = 1/2(0,1,1) + 1/2(1,0,1)$. Since

$$\begin{aligned}
&\frac{1}{2}\mathrm{DCU}(0,0,1) + \frac{1}{2}\mathrm{DCU}(1,1,1) = \frac{1}{2}(0,0,14) + \frac{1}{2}(8,0,0) = (4,0,7) \quad \text{and} \\
&\frac{1}{2}\mathrm{DCU}(0,1,1) + \frac{1}{2}\mathrm{DCU}(1,0,1) = \frac{1}{2}(0,12,0) + \frac{1}{2}(8,0,8) = (4,6,4),
\end{aligned}$$

$\mathrm{conv}(\mathrm{epi}(\mathrm{DCU}))$ contains both $(v,(4,0,7))$ and $(v,(4,6,4))$. For any function LCU with $\mathrm{epi}(\mathrm{LCU}) = \mathrm{conv}(\mathrm{epi}(\mathrm{DCU}))$, it holds that $\mathrm{LCU}(v) \leq \min\{(4,0,7),(4,6,4)\} = (4,0,4)$, and thus $(v,(4,0,4)) \in \mathrm{epi}(\mathrm{LCU}) = \mathrm{conv}(\mathrm{epi}(\mathrm{DCU}))$. However, it is straight-forward to determine that there is no convex combination of points $(v, \mathrm{DCU}(v))$ equating $(v,(4,0,4))$, a contradiction.

Thus, this section does not investigate epi(LCU) analogously to Section 2.1 and Section 2.2, but instead considers $\mathrm{conv}(\mathrm{epi}(\mathrm{DCU}))$.

In the following, we introduce a class of facets of $\mathrm{conv}(\mathrm{epi}(\mathrm{DCU}))$ by

- generalizing the lifted start-up cost inequalities (2.1.4) defined in Subsection 2.3.1, and
- showing that these inequalities define facets of $\mathrm{conv}(\mathrm{epi}(\mathrm{DCU}))$ in Subsection 2.3.2.

2.3.1 Composite Start-up Cost Inequalities

Recall the lifted start-up cost inequalities defined in (2.1.4), Section 2.1,

$$\forall\, t \in [T], l \in [t-1]: \qquad \mathrm{cu}^t \geq \mathrm{CU}^t(l)v^t - \sum_{j=1}^{t} \big(\mathrm{CU}^t(l) - \mathrm{CU}^t(j-1)\big)v^{t-j},$$

which induce all facets of $\mathrm{conv}(\mathrm{epi}(\mathrm{DCU}^t))$. It is not surprising that these inequalities induce facets of $\mathrm{conv}(\mathrm{epi}(\mathrm{DCU}))$ as well.

Being designed for $\mathrm{conv}(\mathrm{epi}(\mathrm{DCU}^t))$, the inequality with parameters l, t only bounds the single start-up cost variable cu^t. In the context of $\mathrm{conv}(\mathrm{epi}(\mathrm{DCU}))$, this inequality can be extended by including multiple cu^t: For a given set $\mathcal{J} \subset [l-1]$, we lift the coefficients of cu^{t-j} for each $j \in \mathcal{J}$ and call the resulting inequality a *composite start-up cost inequality*.

Definition 2.40 For each $t \in [T]$, $l \in [0\,..\,t-1]$, and $\mathcal{J} \subset [l-1]$, define the composite start-up cost inequality as

$$\mathrm{cu}^t + \sum_{j\in\mathcal{J}} \omega_j(t,l)\mathrm{cu}^{t-j} \geq \mathrm{CU}^t(l)v^t - \sum_{j=1}^{t} \alpha_j(t,l,\mathcal{J})v^{t-j}$$

with the non-negative coefficients

$$\omega_j(t,l) := \min\Big\{\omega \in \mathbb{R}_{\geq 0} \;\Big|\; \forall\, k \in [j{+}1\,..\,l]:\ \omega\,\mathrm{CU}^{t-j}(k-j) \geq \mathrm{CU}^t(k) - \mathrm{CU}^t(j)\Big\}$$

and

$$\alpha_j(t,l,\mathcal{J}) := \begin{cases} \mathrm{CU}^t(j) - \mathrm{CU}^t(j-1) & \text{for } j \in \mathcal{J}, \\ \mathrm{CU}^t(l) - \mathrm{CU}^t(j-1) & \text{for } j \notin \mathcal{J}. \end{cases}$$

If the constraint parameters t, l and $\mathcal{J}$ are evident from the context, we omit them and refer to the coefficients as ω_j and α_j. It is straightforward to check that these coefficients are non-negative, and even positive if $\mathrm{CU}(t)$ is strictly increasing.

The composite start-up cost inequalities generalize the lifted start-up cost inequalities, which constitute the special case $\mathcal{J} = \emptyset$. In the following, we

- show their validity in Lemma 2.41,
- prove that they induce facets in Theorem 2.47 and
- count them in Proposition 2.49.

Lemma 2.41 *All composite start-up cost inequalities are valid for* $\mathrm{conv}(\mathrm{epi}(\mathrm{DCU}))$.

Proof. Since the composite start-up cost inequalities bound the start-up costs from below, it suffices to prove that each vertex $(v, \mathrm{DCU}(v))$ of $\mathrm{conv}(\mathrm{epi}(\mathrm{DCU}))$ satisfies each inequality. Consider the composite start-up cost inequality with parameters $t \in [T]$, $l \in [0 \,..\, t-1]$, and $\mathcal{J} \subset [l-1]$. This inequality is fulfilled by each vertex (v, cu) with $v^t = 0$, since its left-hand side is non-negative and its right-hand side for v is non-positive,

$$\mathrm{cu}^t + \sum_{j \in \mathcal{J}} \omega_j \mathrm{cu}^{t-j} \geq 0 \geq \mathrm{CU}^t(l) \underbrace{v^t}_{=0} - \sum_{j=1}^{l} \alpha_j v^{t-j}.$$

Of the vertices (v, cu) with $v^t = 1$, we only need to examine those with $v^{t'} = 1$ for all $t' \in [t-l-1]$. If such a vertex fulfills the considered inequality, then so do all vertices $(\tilde{v}, \tilde{\mathrm{cu}})$ with $v^{t'} = \tilde{v}^{t'}$ for all $t' \in [t-l \,..\, t]$, since

- different values of $\tilde{v}^{t'}$ with $t' > t$ do not influence the inequality at all, and
- different values of $\tilde{v}^{t'}$ with $t' < t-l$ may only increase the left-hand side of the inequality, and do not influence its right-hand side.

The validity of the inequality for these vertices is proved by induction over the number of online periods in the range $[t-l \,..\, t-1]$,

$$n := |N(v)| \quad \text{with} \quad N(v) := \{j \in [l] \mid v^{t-j} = 1\}.$$

If $n = 0$, then $\mathrm{cu}^t = \mathrm{CU}^t(l)$ and the composite start-up cost inequality is fulfilled with equality.

If $n > 0$, denote the two smallest elements of $N(v) \cup \{l+1\}$ by

$$j^1 := \min N(v), \quad j^2 := \min\Big(N(v) \setminus \{j^1\} \cup \{l+1\}\Big)$$

and define the vertex $(\tilde{v}, \tilde{\mathrm{cu}})$ as

$$\forall\, t' \in [T]: \qquad \tilde{v}^{t'} := \begin{cases} 0 & \text{if } t' = t - j^1, \\ v^{t'} & \text{else,} \end{cases} \quad \text{and} \quad \tilde{\mathrm{cu}} := \mathrm{DCU}(\tilde{v}).$$

By definition $|N(\tilde{v})| = |N(v)| - 1 = n - 1$, and hence the composite start-up cost inequality is satisfied by $(\tilde{v}, \tilde{\mathrm{cu}})$. Its validity for (v, cu) is inferred by comparing the differences between the left and right hand sides for (v, cu) and $(\tilde{v}, \tilde{\mathrm{cu}})$. The start-up costs cu and $\tilde{\mathrm{cu}}$ differ solely in periods t and j^1,

$$\begin{aligned} \mathrm{cu}^t &= \mathrm{CU}^t(j^1 - 1), & \mathrm{cu}^{t-j^1} &= \mathrm{CU}^{t-j^1}(j^2 - j^1 - 1), \\ \tilde{\mathrm{cu}}^t &= \mathrm{CU}^t(j^2 - 1), & \tilde{\mathrm{cu}}^{t-j^1} &= 0. \end{aligned}$$

The difference Δ_{L} of the left hand sides of the inequality for (v, cu) and $(\tilde{v}, \tilde{\mathrm{cu}})$ depends on whether the coefficient for cu^{t-j^1} is lifted, and equals

$$\Delta_{\mathrm{L}} := \mathrm{cu}^t + \sum_{j\in\mathcal{J}} \omega_j \mathrm{cu}^{t-j} - \tilde{\mathrm{cu}}^t - \sum_{j\in\mathcal{J}} \omega_j \tilde{\mathrm{cu}}^{t-j} = \begin{cases} \mathrm{cu}^t - \tilde{\mathrm{cu}}^t + \omega_{j^1} \mathrm{cu}^{t-j^1} & \text{if } j^1 \in \mathcal{J}, \\ \mathrm{cu}^t - \tilde{\mathrm{cu}}^t & \text{if } j^1 \notin \mathcal{J}. \end{cases}$$
$$= \begin{cases} \mathrm{CU}^t(j^1-1) - \mathrm{CU}^t(j^2-1) + \omega_{j^1} \mathrm{CU}^{t-j^1}(j^2-j^1-1) & \text{if } j^1 \in \mathcal{J}, \\ \mathrm{CU}^t(j^1-1) - \mathrm{CU}^t(j^2-1) & \text{if } j^1 \notin \mathcal{J}. \end{cases}$$

The difference Δ_{R} of the right hand sides also depends on whether the coefficient of cu^{t-j^1} is lifted, but implicitly through α_{j^1},

$$\Delta_{\mathrm{L}} := \mathrm{CU}^t(l) v^t - \sum_{j=1}^{t} \alpha_j v^{t-j} - \mathrm{CU}^t(l) \tilde{v}^t + \sum_{j=1}^{t} \alpha_j \tilde{v}^{t-j} = -\alpha_{j^1}.$$

We now show $\Delta_{\mathrm{L}} \geq \Delta_{\mathrm{R}}$. If $j^1 \notin \mathcal{J}$, then

$$\Delta_{\mathrm{L}} = \mathrm{CU}^t(j^1-1) - \mathrm{CU}^t(j^2-1) \geq \mathrm{CU}^t(j^1-1) - \mathrm{CU}^t(l) = -\alpha_{j^1} = \Delta_{\mathrm{R}}.$$

Else, if $j^1 \in \mathcal{J}$ and $j^2 = j^1 + 1$, then

$$\Delta_{\mathrm{L}} = \mathrm{CU}^t(j^1-1) - \mathrm{CU}^t(\underbrace{j^2-1}_{=j^1}) + \omega_{j^1} \mathrm{CU}^{t-j^1}(\underbrace{j^2-j^1-1}_{=0}) = -\alpha_{j^1} = \Delta_{\mathrm{R}}.$$

Finally, if $j^1 \in \mathcal{J}$ and $j^2 > j^1 + 1$, then

$$\Delta_{\mathrm{L}} = \mathrm{CU}^t(j^1-1) - \mathrm{CU}^t(j^2-1) + \omega_{j^1} \mathrm{CU}^{t-j^1}(j^2-j^1-1)$$
$$\overset{\text{def. 2.40}}{\geq} \mathrm{CU}^t(j^1-1) - \mathrm{CU}^t(j^1) = -\alpha_{j^1} = \Delta_{\mathrm{R}}.$$

Thus, the composite start-up cost inequality is fulfilled for (v, cu), and by induction for all vertices $(v, \mathrm{DCU}(v))$. □

2.3.2 Facets

The proof of Lemma 2.41 already suggests how to derive vertices $(v, \mathrm{DCU}(v))$ of the epigraph conv(epi(DCU)) which lie on the face induced by a composite start-up cost inequality: by inductively constructing them with $\Delta_{\mathrm{L}} = \Delta_{\mathrm{R}}$ in each step. We enumerate these vertices, ultimately proving that each composite start-up cost inequality induces a facet of conv(epi(DCU)). Figure 2.21 gives these vertices for an exemplary inequality.

To simplify the notation, we denote the face induced by the composite start-up cost inequality with parameters $t \in [T]$, $l \in [0 \,..\, t-1]$, $\mathcal{J} \subset [l-1]$ by $\mathcal{F}(t, l, \mathcal{J})$.

Firstly, the composite start-up cost inequality is trivially fulfilled with equality for each vertex (v, cu) with $v^{t-l} = \ldots = v^{t-1} = 0$:

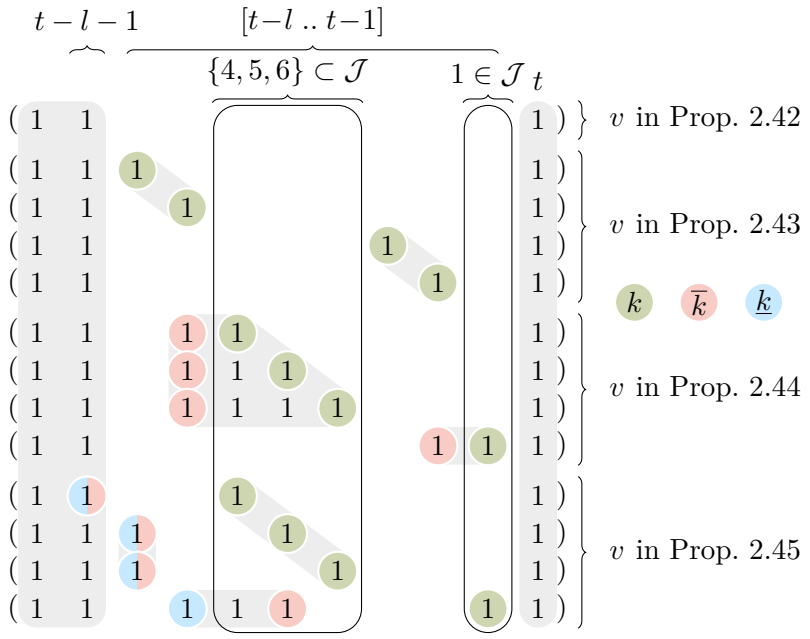


Figure 2.21: Operational schedules v of non-trivial vertices $(v, \mathrm{DCU}(v))$ contained in the face $\mathcal{F}(t, l, \mathcal{J})$ of an exemplary composite start-up cost inequality with parameters $t = T = 11$, $l = 8$, and $\mathcal{J} = \{1, 4, 5, 6\}$ with equality. Only non-zero coefficients are shown. The indices k, $\underline{k}$, and $\overline{k}$ used in the respective results are marked.

Proposition 2.42 *For each $t \in [T]$, $l \in [0\,..\,t-1]$, and $\mathcal{J} \subset [l-1]$, $\mathcal{F}(t, l, \mathcal{J})$ contains $(v, \mathrm{DCU}(v))$ with*

$$v^{t'} = \begin{cases} 1 & \text{if } t' \in [t-l-1] \cup \{t\}, \\ 0 & \text{else.} \end{cases}$$

Proof.

$$\mathrm{cu}^t + \sum_{j \in \mathcal{J}} \omega_j \mathrm{cu}^{t-j} = \mathrm{CU}^t(l) = \mathrm{CU}^t(l) \underbrace{v^t}_{=1} - \sum_{j=1}^{t} \alpha_j \underbrace{v^{t-j}}_{=0}. \qquad \square$$

The induction step of the proof of Lemma 2.41 discerns three cases,

- $j^1 \notin \mathcal{J}$,
- $j^1 \in \mathcal{J}$ and $j^2 = j^1 + 1$, and
- $j^1 \in \mathcal{J}$ and $j^2 > j^1 + 1$.

In the first case, we show that the equality $\Delta_{\mathrm{L}} = \Delta_{\mathrm{R}}$ is attained if $j^2 = l + 1$, which implies that $v^{t-j^1} = 1$ and $v^{t-j} = 0$ for all $j \in [l], j \neq j^1$.

Proposition 2.43 *For each $t \in [T]$, $l \in [0\,..\,t-1]$, $\mathcal{J} \subset [l-1]$, and $k \in [l-1] \setminus \mathcal{J}$, $\mathcal{F}(t,l,\mathcal{J})$ contains $(v, \mathrm{DCU}(v))$ with*

$$v^{t'} = \begin{cases} 1 & \text{for } t' \in [t-l-1] \cup \{t, t-k\}, \\ 0 & \text{else.} \end{cases}$$

Proof.

$$\begin{aligned} \mathrm{cu}^t + \sum_{j\in\mathcal{J}} \omega_j \mathrm{cu}^{t-j} &= \mathrm{CU}^t(k-1) = \mathrm{CU}^t(l) - \Big(\mathrm{CU}^t(l) - \mathrm{CU}^t(k-1)\Big) = \mathrm{CU}^t(l) - \alpha_k \\ &= \mathrm{CU}^t(l) - \alpha_k \underbrace{v^{t-k}}_{=1} - \sum_{j=1}^{t} \alpha_j \underbrace{v^{t-j}}_{=0} = \mathrm{CU}^t(l) v^t - \sum_{j=1}^{t} \alpha_j v^{t-j}. \end{aligned}$$ □

In the second case, i. e. if $j^1 \in \mathcal{J}$ and $j^2 = j^1 + 1$, $\Delta_{\mathrm{L}} = \Delta_{\mathrm{R}}$ is always fulfilled. We construct a vertex $(v, \mathrm{DCU}(v))$ such that either the first or second case would apply in each induction step.

Proposition 2.44 *For each $t \in [T]$, $l \in [0\,..\,t-1]$, $\mathcal{J} \subset [l-1]$, and $k \in \mathcal{J}$, define $\overline{k} := \min [k+1\,..\,l] \setminus \mathcal{J}$. Then, $\mathcal{F}(t,l,\mathcal{J})$ contains $(v, \mathrm{DCU}(v))$ with*

$$v^{t'} = \begin{cases} 1 & \text{if } t' \in [t-l-1] \cup \{t\}, \\ 1 & \text{if } t' \in [t-\overline{k}\,..\,t-k], \\ 0 & \text{else.} \end{cases}$$

Proof. By the choice of $\overline{k}$, $\overline{k} \notin \mathcal{J}$ and $[k\,..\,\overline{k}-1] \subset \mathcal{J}$. Hence, the coefficients α_j equate to

$$\alpha_{\overline{k}} = \mathrm{CU}^t(l) - \mathrm{CU}^t(\overline{k}-1) \quad \text{and} \quad \forall\, j \in [k\,..\,\overline{k}-1]: \quad \alpha_j = \mathrm{CU}^t(j) - \mathrm{CU}^t(j-1).$$

Thus, we have

$$\begin{aligned} \mathrm{cu}^t + \sum_{j\in\mathcal{J}} \omega_j \mathrm{cu}^{t-j} &= \mathrm{CU}^t(k-1) \\ &= \mathrm{CU}^t(l) - \Big(\mathrm{CU}^t(l) - \mathrm{CU}^t(\overline{k}-1)\Big) - \Big(\mathrm{CU}^t(\overline{k}-1) - \mathrm{CU}^t(k-1)\Big) \\ &= \mathrm{CU}^t(l) - \Big(\mathrm{CU}^t(l) - \mathrm{CU}^t(\overline{k}-1)\Big) - \sum_{j=k}^{\overline{k}-1} \Big(\mathrm{CU}^t(j) - \mathrm{CU}^t(j-1)\Big) \\ &= \mathrm{CU}^t(l) - \alpha_{\overline{k}} - \sum_{j=k}^{\overline{k}-1} \alpha_j = \mathrm{CU}^t(l) \underbrace{v^t}_{=1} - \sum_{j=k}^{\overline{k}-1} \alpha_j \underbrace{v^{t-j}}_{=1} - \sum_{\substack{j=1 \\ j\notin[k\,..\,\overline{k-1}]}}^{t} \alpha_j \underbrace{v^{t-j}}_{=0} \\ &= \mathrm{CU}^t(l) v^t - \sum_{j=1}^{t} \alpha_j v^{t-j}. \end{aligned}$$ □

In the third case, i. e. if $j^1 \in \mathcal{J}$ and $j^2 > j^1 + 1$, the equality $\Delta_{\mathrm{L}} = \Delta_{\mathrm{R}}$ holds iff the value of the coefficient ω_{j^1} in its definition is attained for $k = j^2$ (see Definition 2.40). We construct a vertex $(v, \mathrm{DCU}(v))$ such that this third case applies in the final induction step, preceded by induction steps of the first and second case.

Proposition 2.45 *For each $t \in [T]$, $l \in [0 \,..\, t-1]$, $\mathcal{J} \subset [l-1]$ and $k \in \mathcal{J}$, define $\underline{k}$ as*

$$\underline{k} := 1 + \max\{j \in [k+1 \,..\, l] \mid \omega_k \mathrm{CU}^{t-k}(j-k) = \mathrm{CU}^t(j) - \mathrm{CU}^t(k)\}$$
$$\text{and} \quad \overline{k} := \min [\underline{k} \,..\, l+1] \setminus \mathcal{J}.$$

Then, $\mathcal{F}(t, l, \mathcal{J})$ contains $(v, \mathrm{DCU}(v))$ with

$$v^{t'} = \begin{cases} 1 & \text{if } t' \in [t-l-1] \cup \{t\}, \\ 1 & \text{if } t' \in \{t-k\} \cup [t-\overline{k} \,..\, t-\underline{k}], \\ 0 & \text{else.} \end{cases}$$

Proof. Firstly, the definition of v implies the start-up costs

$$\mathrm{cu}^{t'} = \begin{cases} \mathrm{CU}^t(k-1) & \text{if } t' = t, \\ \mathrm{CU}^{t-k}(\underline{k}-k-1) & \text{if } t' = t-k, \\ \mathrm{CU}^{t-\overline{k}}(l-\overline{k}-1) & \text{if } t' = t-\overline{k}, \\ 0 & \text{else.} \end{cases}$$

By the choice of $\overline{k}$, $[\underline{k} \,..\, \overline{k}-1] \subset \mathcal{J}$ and $\overline{k} \notin \mathcal{J}$, and hence

$$\mathrm{cu}^t + \sum_{j \in \mathcal{J}} \omega_j \mathrm{cu}^{t-j} = \mathrm{CU}^t(k-1) + \omega_k \mathrm{CU}^{t-k}(\underline{k}-k-1)$$

The index $\underline{k}$ is chosen such that, using Definition 2.40 of ω_k,

$$\omega_k \mathrm{CU}^{t-k}(\underline{k}-k-1) = \mathrm{CU}^t(\underline{k}-1) - \mathrm{CU}^t(k).$$

Therefore,

$$\begin{aligned} \mathrm{cu}^t + \sum_{j \in \mathcal{J}} \omega_j \mathrm{cu}^{t-j} &= \mathrm{CU}^t(k-1) + \mathrm{CU}^t(\underline{k}-1) - \mathrm{CU}^t(k) \\ &= \mathrm{CU}^t(l) - \Big(\mathrm{CU}^t(k) - \mathrm{CU}^t(k-1)\Big) - \Big(\mathrm{CU}^t(\overline{k}-1) - \mathrm{CU}^t(\underline{k}-1)\Big) \\ &\quad - \Big(\mathrm{CU}^t(l) - \mathrm{CU}^t(\overline{k}-1)\Big) \end{aligned}$$

Similar to the proof of Proposition 2.44, we expand the term into a sum,

$$\mathrm{cu}^t + \sum_{j\in\mathcal{J}} \omega_j \mathrm{cu}^{t-j} = \mathrm{CU}^t(l) - \Big(\mathrm{CU}^t(k) - \mathrm{CU}^t(k-1)\Big) - \sum_{j=\underline{k}}^{\overline{k}-1} \mathrm{CU}^t(j) - \mathrm{CU}^t(j-1)$$
$$- \Big(\mathrm{CU}^t(l) - \mathrm{CU}^t(\overline{k}-1)\Big) = \mathrm{CU}^t(l) - \alpha_k - \sum_{j=\underline{k}}^{\overline{k}-1} \alpha_j - \alpha_{\overline{k}}$$

and by definition of v^t get

$$\mathrm{cu}^t + \sum_{j\in\mathcal{J}} \omega_j \mathrm{cu}^{t-j} = \mathrm{CU}^t(l) v^t - \sum_{j=1}^{t} \alpha_j v^{t-j}. \qquad \square$$

Since each composite start-up cost inequality is homogeneous, it is fulfilled with equality by all vertices $(v, \mathrm{DCU}(v))$ which have positive values solely for variables with coefficient 0 in that inequality.

Proposition 2.46 *For $t \in [T]$, $l \in [0\,..\,t-1]$, and $\mathcal{J} \subset [l-1]$, $\mathcal{F}(t,l,\mathcal{J})$ contains*

- *the null vector 0,*
- *$(u_{t'}, \mathrm{DCU}(u_{t'}))$ for $t' \in [T] \setminus [t-l\,..\,t]$ and*
- *$(0, u_{t'})$ for $t' \in [T] \setminus \{t\}$, $t - t' \notin \mathcal{J}$,*

where $u_{t'}$ denotes the t-th unit vector.

Proof. Since $\mathrm{DCU}(0) = 0$, both the null vector 0 and the points $(u_{t'}, \mathrm{DCU}(u_{t'}))$ are vertices of $\mathrm{conv}(\mathrm{epi}(\mathrm{DCU}))$. We noted in (2.3.1) that $\mathrm{conv}(\mathrm{epi}(\mathrm{DCU}))$ contains the rays $\{u_{T+t} \mid t \in [T]\}$,

$$\mathrm{conv}(\mathrm{epi}(\mathrm{DCU})) = \mathrm{conv}\left(\left\{(v, \mathrm{DCU}(v)) \;\middle|\; v \in \{0,1\}^T\right\}\right) + \mathrm{pos}\{u_{T+t} \mid t \in [T]\}.$$

Therefore, $\mathrm{conv}(\mathrm{epi}(\mathrm{DCU}))$ also contains the points $(0, u_{t'})$.

Each of the mentioned points fulfills $v^t = v^{t-1} = \ldots = v^{t-l} = 0$, $\mathrm{cu}^t = 0$, and $\mathrm{cu}^{t-j} = 0$ for all $j \in \mathcal{J}$. So, each point satisfies the composite start-up cost inequality with equality,

$$\mathrm{cu}^t + \sum_{j\in\mathcal{J}} \omega_j(t,l) \mathrm{cu}^{t-j} = 0 = \mathrm{CU}^t(l) v^t - \sum_{j=1}^{t} \alpha_j(t,l,\mathcal{J}) v^{t-j},$$

and lies in $\mathcal{F}(t,l,\mathcal{J})$. $\square$

Although the faces $\mathcal{F}(t,l,\mathcal{J})$ typically contain further vertices $(v, \mathrm{DCU}(v))$, the enumerated points suffice to prove that they are facets:

Theorem 2.47 *For each $t \in [T]$, $l \in [0\,..\,t-1]$, and $\mathcal{J} \subset [l-1]$, $\mathcal{F}(t,l,\mathcal{J})$ is a facet.*

Proof. By definition, the face $\mathcal{F}(t,l,\mathcal{J})$ is a facet iff its affine hull $\operatorname{aff}(\mathcal{F}(t,l,\mathcal{J}))$ has dimension $2T-1$. This is certainly fulfilled if $\varphi(\operatorname{aff}(\mathcal{F}(t,l,\mathcal{J}))) = \mathbb{R}^{2T-1}$ using the projection φ which discards the coordinate cu^t,

$$\varphi: \quad \mathbb{R}^{2T} \to \mathbb{R}^{2T-1}, \quad (v,\mathrm{cu}) \mapsto (v,\mathrm{cu}^1,\ldots,\mathrm{cu}^{t-1},\mathrm{cu}^{t+1},\ldots,\mathrm{cu}^T).$$

By linearity of φ, $\varphi(\operatorname{aff}(\mathcal{F}(t,l,\mathcal{J}))) = \operatorname{aff}(\varphi(\mathcal{F}(t,l,\mathcal{J})))$, and since $\mathcal{F}(t,l,\mathcal{J})$ contains the null vector (cf. Proposition 2.46), the affine hull $\operatorname{aff}(\varphi(\mathcal{F}(t,l,\mathcal{J})))$ equals the linear hull $\operatorname{span}(\varphi(\mathcal{F}(t,l,\mathcal{J})))$.

In the following, we consider linear hulls $L_0,\ldots,L_4$ of the points proved to lie in $\mathcal{F}(t,l,\mathcal{J})$ in Propositions 2.42 to 2.46. These linear hull L_i span the projection of an increasing set of these points,

- L_0 of the points in Proposition 2.46,
- L_1 of the points in Propositions 2.42 and 2.46,
- L_2 of the points in Propositions 2.42, 2.43 and 2.46,
- L_3 of the points in Propositions 2.42 to 2.44 and 2.46, and
- L_4 of the points in Propositions 2.42 to 2.46.

The order of considering these points is chosen to allow a straight-forward characterization of the L_i: As we will show, each L_i has an orthonormal basis consisting of unit vectors, and each added point increases this basis by a further unit vector.

The linear hull of (the projection of) the trivial points enumerated in Proposition 2.46 equals

$$\begin{aligned} L_0 &= \operatorname{span}\Big(\varphi\Big(\{u_{t'} \mid t' \in [T] \setminus [t-l \mathinner{..} t]\} \cup \{u_{T+t'} \mid t' \in [T] \setminus \{t\} \text{ and } t-t' \notin \mathcal{J}\}\Big)\Big) \\ &= \operatorname{span}\Big(\{u_{t'} \mid t' \in [T] \setminus [t-l \mathinner{..} t]\} \cup \{u_{T+t'} \mid t' \in [T-1] \text{ and } t-t' \notin \mathcal{J}\}\Big). \end{aligned}$$

The point $(v,\mathrm{cu}) \in \mathcal{F}(t,l,\mathcal{J})$ in Proposition 2.42 has non-zero entries in coordinates $v^1,\ldots,v^{t-l-1},v^t,\mathrm{cu}^1,\mathrm{cu}^t$. Of these coordinates, only v^t does not correspond to a unit vector in the basis of L_0. Thus, $L_1 = \operatorname{span}(L_0 \cup \varphi((v,\mathrm{cu})))$ possesses the orthonormal basis

$$L_1 = \operatorname{span}\Big(\{u_{t'} \mid t' \in [T] \setminus [t-l \mathinner{..} t-1]\} \cup \{u_{T+t'} \mid t' \in [T-1] \text{ and } t-t' \notin \mathcal{J}\}\Big).$$

For each $k \in [l-1] \setminus \mathcal{J}$, Proposition 2.43 lists a point $(v,\mathrm{cu}) \in \mathcal{F}(t,l,\mathcal{J})$ with non-zero entries in coordinates $v^1,\ldots,v^{t-l-1},v^{t-k},v^t,\mathrm{cu}^1,\mathrm{cu}^{t-k},\mathrm{cu}^t$. Since $k \notin \mathcal{J}$, the linear hull L_2 of L_1 and the points in Proposition 2.43 has the orthonormal basis

$$L_2 = \operatorname{span}\Big(\{u_{t'} \mid t' \in [T] \text{ and } t-t' \notin \mathcal{J}\} \cup \{u_{T+t'} \mid t' \in [T-1] \text{ and } t-t' \notin \mathcal{J}\}\Big).$$

For each $k \in \mathcal{J}$, Proposition 2.44 defines $\overline{k} \notin \mathcal{J}$ and gives the point $(v, \mathrm{cu}) \in \mathcal{F}(t, l, \mathcal{J})$ with non-zero entries in $v^1, \ldots, v^{t-l-1}, v^{t-\overline{k}}, \ldots, v^{t-k}, v^t, \mathrm{cu}^1, \mathrm{cu}^{t-\overline{k}}, \mathrm{cu}^t$. Hence, the linear hull L_3 of L_2 and the points in Proposition 2.44 has the orthonormal basis

$$L_3 = \operatorname{span}\Big(\{u_{t'} \mid t' \in [T]\} \cup \{u_{T+t'} \mid t' \in [T-1] \text{ and } t - t' \notin \mathcal{J}\}\Big)$$

Finally, for each $k \in \mathcal{J}$ Proposition 2.45 lists a point $(v, \mathrm{cu}) \in \mathcal{F}(t, l, \mathcal{J})$ with non-zero entries $v^1, \ldots, v^{t-l-1}, v^{t-\overline{k}}, \ldots, v^{t-\underline{k}}, v^{t-k}, v^t, \mathrm{cu}^1, \mathrm{cu}^{t-\overline{k}}, \mathrm{cu}^{t-k}, \mathrm{cu}^t$. As $k \in \mathcal{J}$, this yields the linear hull $L_4 = \mathbb{R}^{2T-1}$.

In conclusion $\dim(\operatorname{aff}(\mathcal{F}(t, l, \mathcal{J}))) \geq \dim(L_4) = 2T - 1$ and thus $\mathcal{F}(t, l, \mathcal{J})$ is a facet of $\operatorname{conv}(\operatorname{epi}(\mathrm{DCU}))$. □

Having introduced the class of facet-inducing composite start-up cost inequalities, we now determine their number. For strictly increasing CU, the coefficients ω_j and α_j are positive, and thus the parameters $t \in [T]$, $l \in [0 \,..\, t-1]$, $\mathcal{J} \subset [l-1]$ can be derived uniquely from the non-zero coefficients of a given inequality. Since each inequality uniquely induces a facet, all inequalities are irredundant.

Proposition 2.48 *If* CU *is strictly increasing, then composite start-up cost inequalities are irredundant.*

Therefore, we can count the composite start-up cost inequalities simply by counting the number of possible parameters.

Proposition 2.49 *If* CU *is strictly increasing, then there exist* $2^T - 1$ *different composite start-up cost inequalities.*

Proof.

$$\begin{aligned}\sum_{t\in[T]} \sum_{l=0}^{t-1} \Big|\{\mathcal{J} : \mathcal{J} \subset [l-1]\}\Big| &= \sum_{t\in[T]} \left(1 + \sum_{l=1}^{t-1} 2^{l-1}\right) = \sum_{t\in[T]} \left(1 + \sum_{l=0}^{t-2} 2^{l}\right) \\ &= \sum_{t\in[T]} 2^{t-1} = 2^T - 1.\end{aligned}$$

□

Chapter 3

The Temperature Model

This chapter presents joint work with René Brandenberg and Matthias Huber of which parts have been published in [SHB16; BHS]. While Section 2.1 considers general concave start-up cost functions, it proposes a model specifically designed for the exponential start-up cost function as defined in (1.3.1),

$$\mathrm{CU}(l) = \mathrm{CU}^{\mathrm{fixed}} + \mathrm{CU}^{\mathrm{var}}\left(1 - \mathrm{e}^{-\lambda l}\right) \tag{3.0.1}$$

with fixed costs $\mathrm{CU}^{\mathrm{fixed}} > 0$, maximal variable costs $\mathrm{CU}^{\mathrm{var}} > 0$, and a parameter $\lambda \in (0, 1)$ as depicted in Fig. 3.1.

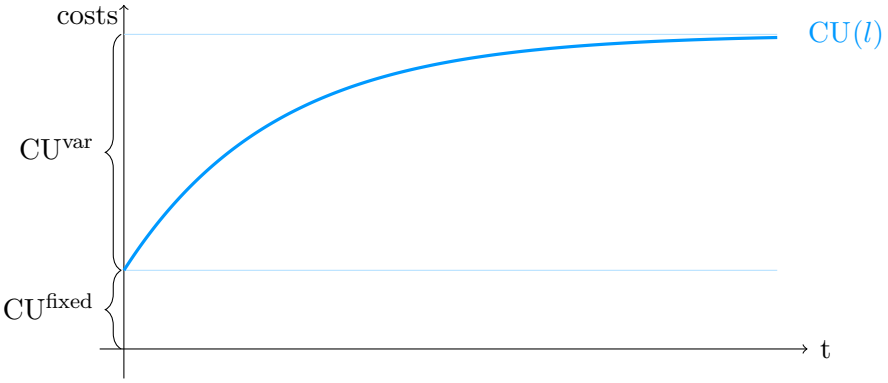


Figure 3.1: Exponential start-up cost function with $\lambda = 0.04$, fixed costs $\mathrm{CU}^{\mathrm{fixed}}$ and maximal variable costs $\mathrm{CU}^{\mathrm{var}}$ depending on the downtime l.

By introducing new variables temp^t and h^t for the temperature and heating of a unit in period $t \in [T]$, we model the start-up costs with significantly fewer inequalities than the models $\mathrm{conv}(\mathrm{epi}(\mathrm{DCU}^t))$ and $\mathrm{conv}(\mathrm{epi}(\mathrm{DCU}^\Sigma))$ given in Chapter 2, resulting in substantially improved solution times. Using $2T$ extra variables, $\mathcal{O}(T^2)$ inequalities yield an extended formulation of the epigraph $\mathrm{conv}(\mathrm{epi}(\mathrm{DCU}^\Sigma))$ of the summed start-up costs, and only $\mathcal{O}(T)$ inequalities are required to model the start-up cost function accurately. These inequalities can be separated in linear time, enabling an efficient cutting plane algorithm.

The results are presented as follows:

1. Based on the physical interpretation of the exponential start-up cost function (3.0.1) in Subsection 1.3.1, Section 3.1 derives a continuous-time model for the temperature of a unit. Discretizing this model yields the temperature and heating values for each operational schedule $v \in \{0,1\}^T$.
2. Using this physical interpretation, Section 3.2 introduces the polyhedron $\hat{P}^{\text{temp}}$ composed of $\mathcal{O}(T)$ inequalities which models temperature and heating, and thereby also the start-up costs for $v \in \{0,1\}^T$.
3. Section 3.3 gives $\mathcal{O}(T^2)$ additional inequalities, yielding the extended formulation P^{temp} of $\text{conv}(\text{epi}(\text{DCU}^{\Sigma}))$, which hence models the summed start-up costs for fractional $v \in [0,1]^T$. A linear-time separation algorithm for these inequalities is given in Subsection 3.3.2.

3.1 A Physical Interpretation of the Start-up Costs

The temperature model is based on a physical interpretation of the exponential start-up costs. The variable part of these costs stem (mostly) from the boiler start-up costs, i. e. the costs of reheating the boiler back to operating temperature, which is proportional to the heat loss suffered during the downtime of the unit.

We assume that the heat is mainly lost to the environment, which means that we exclude combined heat&power units that are able to actively extract heat. For such units, the proposed temperature formulation is straightforward to amend, but results in different start-up cost functions.

Furthermore, we assume that the time required for re-heating is negligible, which appears to be an underlying assumption of the exponential start-up costs (3.0.1). We discuss the case of non-negligible heating times in Section 6.2, where the heating speed is bounded. This however results in additional heat losses while starting, and thereby in costs which differ from the exponential start-up costs.

We start by considering the temperature $\widetilde{\text{temp}}(t)$ of a unit as a continuous function, where $\widetilde{\text{temp}}(t)$ is normalized such that $\widetilde{\text{temp}}(t) = 0$ is the ambient temperature and $\widetilde{\text{temp}}(t) = 1$ is the operational temperature. In a unit without heat extraction, we model the heat loss using the law of heat conduction, which states that the density of the heat flux from the unit to the ambient space is proportional to the temperature difference,

$$\frac{\mathrm{d}\widetilde{\text{temp}}}{\mathrm{d}t}(t) = -\lambda \widetilde{\text{temp}}(t),$$

where the *heat loss coefficient* λ is characteristic to the unit. Solving this linear homogeneous differential equation yields a temperature of

$$\widetilde{\text{temp}}(l) = \mathrm{e}^{-\lambda l} \tag{3.1.1}$$

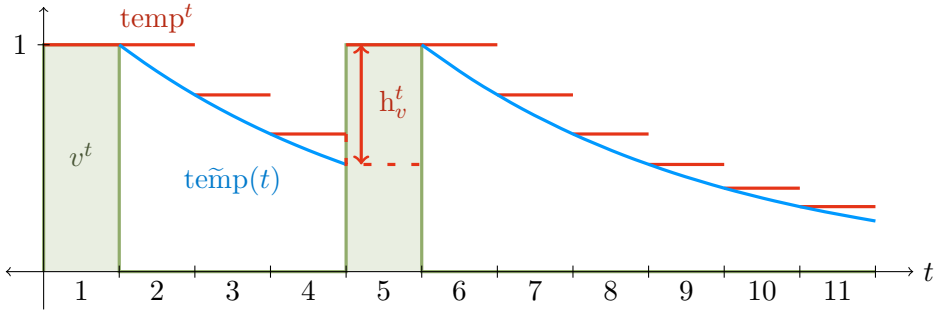


Figure 3.2: Temperature development and heating in the continuous and discretized model.

after a downtime l (cf. Fig. 3.2).

Thus, the heat loss after l offline periods amounts to $(1-\mathrm{e}^{-\lambda l})$, and the total start-up costs result to

$$\mathrm{CU}(l) = \mathrm{CU}^{\text{fixed}} + \mathrm{CU}^{\text{var}}\left(1-\mathrm{e}^{-\lambda l}\right),$$

where $\mathrm{CU}^{\text{fixed}}$, $\mathrm{CU}^{\text{var}} \geq 0$, and $\lambda \in (0,1)$ are appropriate constants specific to each unit. This matches the typically used definition of the start-up costs in (3.0.1).

To apply this temperature model to the Unit Commitment problem, it needs to be discretized. For an operational schedule $v \in \{0,1\}^T$, we denote the discretization of the temperature as temp_v, where temp_v^t is defined as the continuous temperature at the beginning of period t (c.f. Fig. 3.2). Analogously, the heating needed when starting up in period t is denoted by h_v^{t-1}.

As expected, the exponential decay of the physical temperature during offline periods carries over to the discretized temperature temp_v. Recall that $\mathrm{ol}^t(v)$ counts the number of offline periods preceding period t (see (1.3.7)),

$$\mathrm{ol}^t(v) = \max\{l \in [t-1] \mid v^{t-1} = \ldots = v^{t-l} = 0\}.$$

Then, $\mathrm{OL}^t(\mathrm{ol}^t(v))$ equals the total offline length (see (1.3.3)),

$$\mathrm{OL}^t(l) = \sum_{j=1}^{l} \mathrm{L}^{t-j} + \begin{cases} \text{PDT} & \text{if } l = t-1, \\ 0 & \text{else,} \end{cases}$$

where PDT denotes the pre-model downtime. Using $\mathrm{OL}^t(l)$, the temperature can be expressed as

$$\forall\, t \in [T]: \qquad \text{temp}_v^t := \begin{cases} 1 & \text{if } v^t = 1, \\ \mathrm{e}^{-\lambda \mathrm{OL}^t(\mathrm{ol}^t(v))} & \text{else.} \end{cases} \tag{3.1.2}$$

The decay of the temperature results in an inverse exponential increase of the heating after the downtime,

$$\forall\, t \in [T]: \qquad \mathrm{h}_v^{t-1} := \begin{cases} 1 - \mathrm{e}^{-\lambda \mathrm{OL}^t(\mathrm{ol}^t(v))} & \text{if } v^t = 1, \\ 0 & \text{else.} \end{cases} \tag{3.1.3}$$

Note that the indices of h_v lie in $[0\,..\,T-1]$, as the heating for a start-up in period t is performed in period $t-1$.

To model the fixed start-up costs, we use the start-up and shutdown indicators $\mathrm{y}^t, \mathrm{z}^t$ due to [Gar62], which are equal to 1 iff the unit starts up/shuts down in period t.

$$\forall\, t \in [T]: \qquad \mathrm{y}_v^t := \begin{cases} 1 & \text{if } v^t = 1 \text{ and } \mathrm{OL}^t(\mathrm{ol}^t(v)) > 0, \\ 0 & \text{else.} \end{cases} \tag{3.1.4}$$

$$\forall\, t \in [T]: \qquad \mathrm{z}_v^t := \begin{cases} 1 & \text{if } v^t = 0 \text{ and } \mathrm{OL}^t(\mathrm{ol}^t(v)) = 0, \\ 0 & \text{else.} \end{cases} \tag{3.1.5}$$

As shown in the following, y^t and h^{t-1} suffice to derive the start-up costs. Still, we require z_v and temp_v to model the extended formulation of $\mathrm{conv}(\mathrm{epi}(\mathrm{DCU}^{\Sigma}))$ with $\mathcal{O}(T^2)$ inequalities (see Theorem 3.26).

Proposition 3.1 *For each $v \in \{0,1\}^T$, using the orthogonal projection*

$$\begin{aligned} \pi^{\mathrm{temp}}: \quad \mathbb{R}^{5T} \to \mathbb{R}^{2T}, \quad (v, \mathrm{y}_v, \mathrm{z}_v, \mathrm{temp}_v, \mathrm{h}_v) \mapsto (v, \mathrm{cu}), \\ \mathrm{cu}^t := \mathrm{CU}^{\mathrm{fixed}}\,\mathrm{y}^t + \mathrm{CU}^{\mathrm{var}}\,\mathrm{h}^{t-1}, \end{aligned} \tag{3.1.6}$$

it holds that $\pi^{\mathrm{temp}}(v, \mathrm{y}_v, \mathrm{z}_v, \mathrm{temp}_v, \mathrm{h}_v) = (v, \mathrm{DCU}(v))$.

Proof. Substituting the definitions of h_v and y_v into (3.1.6) yields

$$\begin{aligned} \mathrm{cu}^t &= \mathrm{CU}^{\mathrm{fixed}}\mathrm{y}_v^t + \mathrm{CU}^{\mathrm{var}}\mathrm{h}_v^{t-1} = \mathrm{CU}^{\mathrm{fixed}}\mathrm{y}_v^t + \mathrm{CU}^{\mathrm{var}}\Big(1 - \mathrm{e}^{-\lambda \mathrm{OL}^t(\mathrm{ol}^t(l))}\Big) \\ &= \begin{cases} \mathrm{CU}^{\mathrm{fixed}} + \mathrm{CU}^{\mathrm{var}}\Big(1 - \mathrm{e}^{-\lambda \mathrm{OL}^t(\mathrm{ol}^t(l))}\Big) & \text{if } \mathrm{OL}^t(\mathrm{ol}^t(l)) > 0 \text{ and } v^t = 1, \\ 0 & \text{else,} \end{cases} \end{aligned}$$

which equals the definition of the start-up costs (see (1.3.6), (1.3.4),(1.3.2)) for the exponential start-up cost function in (3.0.1). □

Thus, the set V^{temp} of all such points,

$$V^{\mathrm{temp}} := \big\{(v, \mathrm{y}_v, \mathrm{z}_v, \mathrm{temp}_v, \mathrm{h}_v) \mid v \in \{0,1\}^T\big\}.$$

projects on the vertices V of the epigraph conv(epi(DCU)) of the start-up costs in all periods, i. e. $\pi^{\text{temp}}(V^{\text{temp}}) = V$. This motivates Section 3.2 and Section 3.3, which focus on modeling V^{temp} by enclosing polyhedra.

However, the defining equations (3.1.2) and (3.1.3) of temp^t and h^t are non-linear. In preparation of the linear inequalities presented in the next section, we derive recursive linear representations of temp^t and h^t in the following. These representations are based on the solution (3.1.1) of the differential equation describing the heat loss, which suggests the recursion

$$\text{temp}_v^{t+1} = \text{e}^{-\lambda \text{L}^t} \text{temp}_v^t$$

for the temperature during offline periods.

Taking into account the temperature during online periods and the downtime PDT prior to the modeled time range, we extend this relationship to all cases:

Proposition 3.2 *For each $v \in \{0,1\}^T$,*

$$\forall\, t \in [T]: \quad \text{temp}_v^t = \begin{cases} 1 & \text{if } v^t = 1, \\ e^{-\lambda \text{PDT}} & \text{if } v^t = 0 \text{ and } t = 1, \\ 1 & \text{if } v^t = 0,\ t > 1 \text{ and } v^{t-1} = 1, \\ e^{-\lambda \text{L}^{t-1}} \text{temp}_v^{t-1}(v) & \text{if } v^t = 0,\ t > 1 \text{ and } v^{t-1} = 0. \end{cases}$$

Proof. For $v^t = 1$, $\text{temp}_v^t = 1$ by definition. For $v^t = 0$, we have $\text{temp}_v^t = \text{e}^{-\lambda \text{OL}^t(\text{ol}^t(v))}$, and the equivalency follows directly from the properties of the number of preceding offline periods $\text{ol}^t(v)$,

$$\text{ol}^t(v) = \begin{cases} 0 & \text{if } t = 1 \text{ or } v^{t-1} = 1, \\ \text{ol}^{t-1}(v) + 1 & \text{else,} \end{cases}$$

and the properties of the offline length $\text{OL}^t(\text{ol}^t(v))$,

$$\text{OL}^t(\text{ol}^t(v)) = \begin{cases} \text{PDT} & \text{if } t = 1, \\ 0 & \text{if } t > 1 \text{ and } v^{t-1} = 1, \\ \text{OL}^{t-1}(\text{ol}^{t-1}(v)) + \text{L}^{t-1} & \text{if } t > 1 \text{ and } v^{t-1} = 0. \end{cases}$$ □

The recursion for the heating h_v is derived analogously:

Proposition 3.3 *For each $v \in \{0,1\}^T$,*

$$\forall\, t \in [T]: \quad \text{h}_v^{t-1} = \begin{cases} 0 & \text{if } v^t = 0, \\ 1 - e^{-\lambda \text{PDT}} & \text{if } v^t = 1 \text{ and } t = 1, \\ 0 & \text{if } v^t = 1,\ t > 1 \text{ and } v^{t-1} = 1, \\ 1 - e^{-\lambda \text{L}^{t-1}} \text{temp}_v^{t-1} & \text{if } v^t = 1,\ t > 1 \text{ and } v^{t-1} = 0. \end{cases}$$

3.2 An $\mathcal{H}$-Representation for Integral Operational Schedules

Based on Propositions 3.2 and 3.3, this section presents $\mathcal{O}(T)$ feasible inequalities for V^{temp}. The resulting polyhedron $\hat{P}^{\text{temp}}$ correctly models the start-up costs for integral operational schedules $v \in \{0,1\}^T$ when minimizing their sum.

The requirement to minimize the sum of the start-up costs stems from the need to prevent premature heating and can be relaxed to minimizing certain *weighted* sums of the start-up costs. We point out its necessity in Subsection 3.2.1 and specify the exact requirement in Definition 3.7.

The operational state v, the start-up indicators y^t and the shutdown indicators z^t are part of the Unit Commitment problem presented in Section 1.2. For convenience, we repeat their defining inequalities (1.2.2),(1.2.3):

$$\forall\, t \in [T]: \qquad 0 \leq v^t \leq 1 \tag{3.2.1}$$

$$\forall\, t \in [2..T]: \qquad \mathrm{y}^t - \mathrm{z}^t = v^t - v^{t-1}, \tag{3.2.2}$$

$$\mathrm{y}^1 - \mathrm{z}^1 = v^1 - \begin{cases} 1 & \text{if PDT} = 0, \\ 0 & \text{else.} \end{cases} \tag{3.2.3}$$

$$\forall\, t \in [T]: \qquad \mathrm{y}^t, \mathrm{z}^t \geq 0 \tag{3.2.4}$$

The operational temperature is enforced by

$$\forall\, t \in [T]: \qquad \mathrm{temp}^t \geq v^t. \tag{3.2.5}$$

A lower bound suffices, as higher temperatures never lead to an optimal solution (cf. Lemma 3.10).

During an offline period t, the temperature decays by a factor of $\mathrm{e}^{-\lambda \mathrm{L}^t}$. On the other hand, the temperature must remain constant at 1 during an online period. Keeping in mind the possible heating, this results in the temperature development equations

$$\mathrm{temp}^1 = \mathrm{e}^{-\lambda \mathrm{PDT}} + \mathrm{h}^0, \tag{3.2.6}$$

$$\forall\, t \in [T-1]: \qquad \mathrm{temp}^{t+1} = \mathrm{e}^{-\lambda \mathrm{L}^t} \mathrm{temp}^t + (1 - \mathrm{e}^{-\lambda \mathrm{L}^t}) v^t + \mathrm{h}^t. \tag{3.2.7}$$

Note that without the heating h^t, substituting either $v^t = 0$ or $v^t = 1$ yields the case distinction in equation (3.1.2) for $\mathrm{temp}_v^{i,t}$.

The linear representation of the temperature development (3.2.7) in temp^{t+1}, temp^t, v^t, and h^t is crucial for the efficiency of the temperature model. Section 3.4 generalizes (3.2.7) for different temperature decays and points out that only an exponential temperature decay results in a linear temperature development.

Since we assume that the unit can not actively extract heat, the heating variables are non-negative,

$$\forall\, t \in [T]: \qquad \mathrm{h}^{t-1} \geq 0. \tag{3.2.8}$$

Noting that heating in period 1 is applied iff the unit is online and was previously offline, i. e. $v^1 = 1$ and $\mathrm{PDT} > 0$, we find the additional valid equation

$$\mathrm{h}^0 = (1 - \mathrm{e}^{-\lambda \mathrm{PDT}})v^1. \tag{3.2.9}$$

Together, inequalities (3.2.2)-(3.2.9) define the polyhedron $\hat{P}^{\mathrm{temp}}$.

Definition 3.4

$$\hat{P}^{\mathrm{temp}} := \Big\{(v, \mathrm{y}, \mathrm{z}, \mathrm{temp}, \mathrm{h}) \in \mathbb{R}^{5T} \text{ fulfilling (3.2.1)-(3.2.9)}\Big\}.$$

The next lemma points out that this polyhedron contains the feasible solutions of the temperature model V^{temp}.

Lemma 3.5

$$V^{\mathrm{temp}} \subset \hat{P}^{\mathrm{temp}} \subset \mathbb{R}^{5T}_{\geq 0}$$

Proof. It is straightforward to check that all variables of $\hat{P}^{\mathrm{temp}}$ must be non-negative, and thus $\hat{P}^{\mathrm{temp}} \subset \mathbb{R}^{5T}_{\geq 0}$.

We show that each point $(v, \mathrm{y}_v, \mathrm{z}_v, \mathrm{temp}_v, \mathrm{h}_v) \in V^{\mathrm{temp}}$ fulfills all inequalities defining $\hat{P}^{\mathrm{temp}}$ in order of their introduction.

The bounds $0 \leq v^t \leq 1$ are trivially fulfilled by definition of V^{temp}. The validity of inequalities (3.2.2)-(3.2.4) for the start-up and shutdown indicators has been shown by [Gar62]: For $t > 1$, $\mathrm{OL}^t(\mathrm{ol}^t(v^t)) > 0$ iff $v^{t-1} = 0$, and therefore (3.1.4) and (3.1.5) yield

$$\mathrm{y}^t_v - \mathrm{z}^t_v = \begin{cases} 1 & \text{if } v^t = 1 \text{ and } v^{t-1} = 0 \\ -1 & \text{if } v^t = 0 \text{ and } v^{t-1} = 1 \\ 0 & \text{else} \end{cases} = v^t - v^{t-1},$$

validating (3.2.2). For $t = 1$, $\mathrm{OL}^t(\mathrm{ol}^t(v^t)) > 0$ iff $\mathrm{PDT} > 0$, and thus (3.1.4) and (3.1.5) yield

$$\mathrm{y}^1_v - \mathrm{z}^1_v = \begin{cases} 1 & \text{if } v^1 = 1 \text{ and } \mathrm{PDT} > 0 \\ -1 & \text{if } v^1 = 0 \text{ and } \mathrm{PDT} = 0 \\ 0 & \text{else} \end{cases} = \begin{cases} v^1 & \text{if } \mathrm{PDT} > 0 \\ v^1 - 1 & \text{if } \mathrm{PDT} = 0 \end{cases}.$$

Hence, (3.2.3) is valid as well. Moreover, the indicators are non-negative by definition, and so (3.2.4) holds.

The first case in (3.1.2) guarantees that inequality (3.2.5) enforcing the operational temperature holds. To verify inequality (3.2.6), we consider two cases. If $v^1 = 0$, then

$$\mathrm{temp}^1_v \overset{(3.1.2)}{=} \mathrm{e}^{-\lambda \mathrm{PDT}} \overset{(3.1.3)}{=} \mathrm{e}^{-\lambda \mathrm{PDT}} + \mathrm{h}^0_v,$$

and if $v^1 = 1$, then

$$\mathrm{temp}_v^1 \overset{(3.1.2)}{=} 1 = \mathrm{e}^{-\lambda \mathrm{PDT}} + (1 - \mathrm{e}^{-\lambda \mathrm{PDT}}) \overset{(3.1.3)}{=} \mathrm{e}^{-\lambda \mathrm{PDT}} + \mathrm{h}_v^0.$$

For inequality (3.2.7) governing the temperature loss in periods $t \geq 1$, three cases are required. If $v^t = 1$, then by Proposition 3.2 and 3.3

$$\mathrm{temp}_v^{t+1} = 1 = \mathrm{e}^{-\lambda \mathrm{L}^t} \underbrace{\mathrm{temp}_v^t}_{=1} + \Big(1 - \mathrm{e}^{-\lambda \mathrm{L}^t}\Big) \underbrace{v^t}_{=1} + \underbrace{\mathrm{h}_v^t}_{=0},$$

if $v^t = 0$ and $v^{t+1} = 0$, then

$$\mathrm{temp}_v^{t+1} = \mathrm{e}^{-\lambda \mathrm{L}^t} \mathrm{temp}_v^t = \mathrm{e}^{-\lambda \mathrm{L}^t} \mathrm{temp}_v^t + \Big(1 - \mathrm{e}^{-\lambda \mathrm{L}^t}\Big) \underbrace{v^t}_{=0} + \underbrace{\mathrm{h}_v^t}_{=0},$$

and if $v^t = 0$ and $v^{t+1} = 1$, then

$$\mathrm{temp}_v^{t+1} = 1 = \mathrm{e}^{-\lambda \mathrm{L}^t} \mathrm{temp}_v^t + \Big(1 - \mathrm{e}^{-\lambda \mathrm{L}^t} \mathrm{temp}_v^t\Big) \underbrace{v^t}_{=0} + \underbrace{\mathrm{h}_v^t}_{=1-\mathrm{e}^{-\lambda \mathrm{L}^t} \mathrm{temp}_v^t}.$$

Since the temperature temp_v is bounded from above by 1, it is easy to see that the heating h_v is non-negative by checking every case of (3.1.3), confirming (3.2.8).

Finally, note that $\mathrm{OL}^1(\mathrm{ol}^1(v)) = \mathrm{OL}^1(0) = \mathrm{PDT}$, (3.2.9) holds by definition of h_v^{t-1} in (3.1.3), since

$$\mathrm{h}_v^0 = \begin{Bmatrix} 1 - \mathrm{e}^{-\lambda \mathrm{PDT}} & \text{if } v^1 = 1, \\ 0 & \text{else.} \end{Bmatrix} = (1 - \mathrm{e}^{-\lambda \mathrm{PDT}}) v^1. \qquad \square$$

By Proposition 3.1, $\pi^{\mathrm{temp}}(V^{\mathrm{temp}})$ equals the vertices of conv(epi(DCU)), which are therefore contained in $\pi^{\mathrm{temp}}(\hat{P}^{\mathrm{temp}})$. Furthermore, temperature and heating are only bounded from below in $\hat{P}^{\mathrm{temp}}$, and hence the start-up costs are not bounded from above. So, $V^{\mathrm{temp}} \subset \hat{P}^{\mathrm{temp}}$ implies

Corollary 3.6

$$\mathrm{conv}(\mathrm{epi}(\mathrm{DCU})) \subset \pi^{\mathrm{temp}}(\hat{P}^{\mathrm{temp}}).$$

3.2.1 Correctness for Integral Operational Schedules

Having established that solutions of the temperature model are contained in $\hat{P}^{\text{temp}}$, we spend the remainder of this section determining under which conditions these solutions are optimal and the start-up costs are modeled correctly.

Note that in $\hat{P}^{\text{temp}}$, a unit may heat even if it is not starting. However, since the temperature losses are proportional to the difference between temperature and operational state, premature heating leads to extra heat loss (see Fig. 3.3).

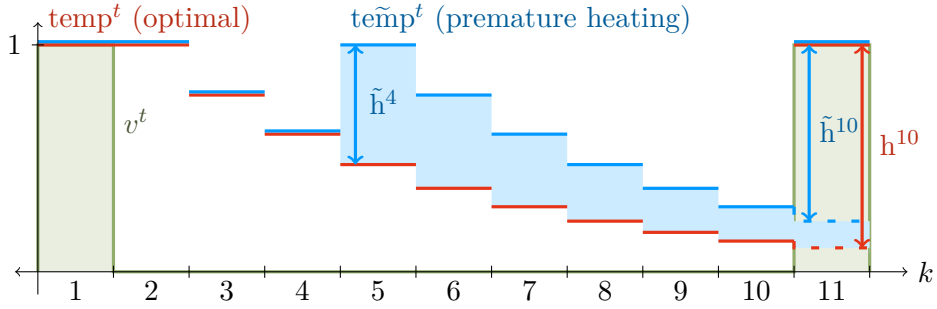


Figure 3.3: Temperature for premature heating $\tilde{\text{h}}$ compared to minimal heating h. The premature heating $\tilde{\text{h}}^4$ is strictly greater than the saved heating in period 10.

For example, consider a point $(v, \text{y}, \text{z}, \text{temp}, \text{h}) \in \hat{P}^{\text{temp}}$ with heating in period $t-1$ such that $\text{temp}^t > v^t$. Postponing a part $0 < \Delta \leq \min\{\text{h}^{t-1}, \text{temp}^t - v^t\}$ of the heating to period t results in

$$\tilde{\text{temp}}^{t'} = \begin{cases} \text{temp}^{t'} - \Delta & \text{if } t' = t, \\ \text{temp}^{t'} & \text{else,} \end{cases} \qquad \tilde{\text{h}}^{t'} = \begin{cases} \text{h}^{t'} - \Delta & \text{if } t' = t-1, \\ \text{h}^{t'} + \text{e}^{-\lambda \text{L}^{t-1}}\Delta & \text{if } t' = t, \\ \text{h}^{t'} & \text{else.} \end{cases} \tag{3.2.10}$$

Since $\tilde{\text{temp}}^t \geq v^t$, the point $(v, \text{y}, \text{z}, \tilde{\text{temp}}, \tilde{\text{h}})$ still lies in $\hat{P}^{\text{temp}}$. Its heating, and by extension its associated start-up costs, are shifted from period $t-1$ to period t, and decreased by a factor of $\text{e}^{-\lambda \text{L}^{t-1}} < 1$.

As we show in the following, if $p \in \hat{P}^{\text{temp}}$ minimizes the total heating for an operational schedule $v \in \{0,1\}$, then there is no premature heating.

Definition 3.7 A positive objective function vector $a \in \mathbb{R}^T_{>0}$ with

$$\forall\, t \in [T-1]: \qquad a_t > \text{e}^{-\lambda \text{L}^t} a_{t+1},$$

is called *heating-minimizing*, and for each set P with $V^{\text{temp}} \subset P \subset \hat{P}^{\text{temp}}$ and each $v \in [0,1]^T$, a solution of

$$\min\Big\{ a^T \text{cu} \;\Big|\; (v, \text{y}, \text{z}, \text{temp}, \text{h}) \in P, (v, \text{cu}) = \pi^{\text{temp}}(v, \text{y}, \text{z}, \text{temp}, \text{h}) \Big\} \tag{3.2.11}$$

is called a *heating-minimal solution for v in P*.

The class of heating-minimizing objective function vectors in particular includes the one vector $a = (1, \dots, 1)^T$, resulting in minimal summed start-up costs. In some cases, for example when modeling variable fuel prices, one might want to choose a different objective function vector a. Note that for a rapid increase in fuel prices, a may not be heating-minimizing and premature heating may be optimal. While the resulting start-up costs do not conform to the exponential start-up cost function, they may be closer to the real behavior of a unit in such a situation.

We show in the following that heating-minimizing objective functions prevent premature heating, even for fractional operational schedules.

Proposition 3.8 *For each $v \in [0,1]^T$, each heating-minimal solution $(v,\mathrm{y},\mathrm{z},\mathrm{temp},\mathrm{h})$ for v in $\hat{P}^{\mathrm{temp}}$ fulfills*

$$\forall\, t \in [T]: \quad \mathrm{h}^{t-1} > 0 \;\Rightarrow\; \mathrm{temp}^t = v^t.$$

Proof. Assume there exists a heating-minimal solution $(v, \mathrm{y}, \mathrm{z}, \mathrm{temp}, \mathrm{h})$ for v in $\hat{P}^{\mathrm{temp}}$ with premature heating, i. e.

$$\exists\, t \in [T]: \qquad \mathrm{h}^{t-1} > 0 \;\text{ and }\; \mathrm{temp}^t > v^t.$$

Choose $\Delta := \max\{\mathrm{h}^{t-1}, \mathrm{temp}^t - v^t\} > 0$, and denote by $(v, \mathrm{y}, \mathrm{z}, \tilde{\mathrm{temp}}, \tilde{\mathrm{h}})$ the point with postponed heating as in (3.2.10). Let cu and $\tilde{\mathrm{cu}}$ be the start-up costs associated with $(v, \mathrm{y}, \mathrm{z}, \mathrm{temp}, \mathrm{h})$ and $(v, \mathrm{y}, \mathrm{z}, \tilde{\mathrm{temp}}, \tilde{\mathrm{h}})$, i. e. $(v, \mathrm{cu}) = \pi^{\mathrm{temp}}(v, \mathrm{y}, \mathrm{z}, \mathrm{temp}, \mathrm{h})$ and $(v, \tilde{\mathrm{cu}}) = \pi^{\mathrm{temp}}(v, \mathrm{y}, \tilde{\mathrm{temp}}, \tilde{\mathrm{h}})$. Since both points have the same start-up indicators y, we have

$$a^T \mathrm{cu} - a^T \tilde{\mathrm{cu}} = \mathrm{CU}^{\mathrm{var}}\Big(a^T \mathrm{h} - a^T \tilde{\mathrm{h}}\Big),$$

which, noting that h and $\tilde{\mathrm{h}}$ differ by Δ in periods t and possibly $t+1$, is bounded by

$$a^T \mathrm{cu} - a^T \tilde{\mathrm{cu}} \geq \begin{cases} \mathrm{CU}^{\mathrm{var}} \Delta (a_t - a_{t+1} \mathrm{e}^{-\lambda \mathrm{L}^t}) > 0 & \text{if } t < T, \\ \mathrm{CU}^{\mathrm{var}} a_t \Delta > 0 & \text{if } t = T, \end{cases}$$

a contradiction to the optimality of $(v, \mathrm{y}, \mathrm{z}, \mathrm{temp}, \mathrm{h})$. □

As a direct consequence, heating prior to an offline period is never heating-minimal:

Proposition 3.9 *For each $v \in [0,1]^T$, each heating-minimal solution $(v,\mathrm{y},\mathrm{z},\mathrm{temp},\mathrm{h})$ for v in $\hat{P}^{\mathrm{temp}}$ fulfills*

$$\forall\, t \in [T]: \quad v^t = 0 \;\Rightarrow\; \mathrm{h}^{t-1} = 0.$$

Proof. For $t = 1$, this holds due to (3.2.9). For $t > 1$, assume there exists $t \in [T]$ with $v^t = 0$ and $\mathrm{h}^{t-1} > 0$. By Proposition 3.8, this implies $\mathrm{temp}^t = v^t = 0$, and thereby

$$0 = \mathrm{temp}^t \overset{(3.2.7)}{=} \mathrm{e}^{-\lambda \mathrm{L}^{t-1}} \mathrm{temp}^{t-1} + (1 - \mathrm{e}^{-\lambda \mathrm{L}^{t-1}}) v^{t-1} + \mathrm{h}^{t-1} \geq \mathrm{h}^{t-1} > 0,$$

a contradiction. □

In the following, we show that every heating-minimal solution $(v, \mathrm{y}, \mathrm{z}, \mathrm{temp}, \mathrm{h})$ in $\hat{P}^{\mathrm{temp}}$ with $v \in \{0,1\}^T$ lies in the set V^{temp}, i. e. that

$$(v, \mathrm{y}, \mathrm{z}, \mathrm{temp}, \mathrm{h}) = (v, \mathrm{y}_v, \mathrm{z}_v, \mathrm{temp}_v, \mathrm{h}_v).$$

Hence, $\hat{P}^{\mathrm{temp}}$ correctly models temperature, heating, and start-up/shutdown indicators when minimizing using a heating-minimizing objective function vector.

Lemma 3.10 *For each $v \in \{0,1\}^T$, each heating-minimal solution $(v, \mathrm{y}, \mathrm{z}, \mathrm{temp}, \mathrm{h})$ for v in $\hat{P}^{\mathrm{temp}}$ fulfills $\mathrm{temp} = \mathrm{temp}_v$ and $\mathrm{h} = \mathrm{h}_v$.*

Proof. For each $t \in [T]$, Proposition 3.8 proves that $\mathrm{h}^{t-1} > 0$ implies $\mathrm{temp}^t = v^t$. Since the temperature may only rise when heating is applied and since $v^t \leq 1$, it follows that $\mathrm{temp}^t \leq 1$. On the other hand, $v^t \leq \mathrm{temp}^t$, and thus

$$\forall\, t \in [T] \text{ with } v^t = 1: \qquad \mathrm{temp}^t = 1 = \mathrm{temp}_v^t.$$

If $v^1 = 0$, then by (3.2.9) $\mathrm{h}^0 = 0$ and hence

$$\mathrm{temp}^1 \overset{(3.2.6)}{=} \mathrm{e}^{-\lambda \mathrm{PDT}} + \mathrm{h}^0 = \mathrm{e}^{-\lambda \mathrm{PDT}} = \mathrm{temp}_v^1.$$

Finally, assume that there exists $t \in [2..T]$ with $v^t = 0$ and $\mathrm{temp}^t \neq \mathrm{temp}_v^t$, and choose the first such period t^*. By the choice of t^* we have $\mathrm{temp}^{t-1} = \mathrm{temp}_v^{t-1}$, and Proposition 3.9 implies $\mathrm{h}^{t^*-1} = 0$. If $v^{t^*-1} = 0$, then

$$\begin{aligned}
\mathrm{temp}^{t^*} &\overset{(3.2.7)}{=} \mathrm{e}^{-\lambda \mathrm{L}^{t^*-1}} \mathrm{temp}^{t^*-1} + (1 - \mathrm{e}^{-\lambda \mathrm{L}^{t^*-1}}) v^{t^*-1} + \mathrm{h}^{t^*-1} \\
&= \mathrm{e}^{-\lambda \mathrm{L}^{t^*-1}} \mathrm{temp}_v^{t^*-1} + (1 - \mathrm{e}^{-\lambda \mathrm{L}^{t^*-1}}) v^{t^*-1} \\
&\overset{(3.2)}{=} \left\{ \begin{array}{ll} \mathrm{e}^{-\lambda \mathrm{L}^{t^*-1}} \mathrm{temp}^{t^*-1} & \text{if } v^{t^*-1} = 0, \\ 1 & \text{if } v^{t^*-1} = 1, \end{array} \right\} \overset{(3.2)}{=} \mathrm{temp}_v^{t^*},
\end{aligned}$$

again a contradiction. Hence, $\mathrm{temp} = \mathrm{temp}_v$.

We proceed by showing that $\mathrm{h} = \mathrm{h}_v$. Since no premature heating is applied due to Proposition 3.8, we know that

$$\forall\, t \in [T], v^t = 0: \quad \Rightarrow \quad \mathrm{h}^{t-1} = 0 = \mathrm{h}_v^{t-1}.$$

If $v^1 = 1$, the temperature development equation (3.2.6) implies

$$\mathrm{h}^0 = \mathrm{temp}^1 - \mathrm{e}^{-\lambda \mathrm{PDT}} = 1 - \mathrm{e}^{-\lambda \mathrm{PDT}} = \mathrm{h}_v^0.$$

Analogously, for each $t \geq 2$ with $v^t = 1$, we have by (3.2.7)

$$\begin{aligned} \mathrm{h}^{t-1} &= 1 - \mathrm{e}^{-\lambda \mathrm{L}^{t^*-1}} \mathrm{temp}^{t-1} - (1 - \mathrm{e}^{-\lambda \mathrm{L}^{t^*-1}}) v^{t-1} \\ &= \begin{Bmatrix} 1 - \mathrm{e}^{-\lambda \mathrm{L}^{t^*-1}} \mathrm{temp}^{t^*-1} & \text{if } v^{t^*-1} = 0, \\ 0 & \text{if } v^{t^*-1} = 1. \end{Bmatrix} = \mathrm{h}_v^{t^*-1}. \end{aligned}$$ □

For non-zero fixed start-up costs, [Gar62] shows that inequalities (3.2.2)-(3.2.4) properly model the start-up indicators.

Proposition 3.11 *For each $v \in \{0,1\}^T$, each heating-minimal solution* $(v, \mathrm{y}, \mathrm{z}, \mathrm{temp}, \mathrm{h})$ *for v in $\hat{P}^{\mathrm{temp}}$ fulfills* $\mathrm{y} = \mathrm{y}_v$ *and* $\mathrm{z} = \mathrm{z}_v$.

So, for each integral operational schedule $v \in \{0,1\}^T$, the unique heating-minimal solution is $(v, \mathrm{y}_v, \mathrm{z}_v, \mathrm{temp}_v, \mathrm{h}_v) \in V^{\mathrm{temp}}$, which models the start-up costs correctly by Proposition 3.1.

Corollary 3.12 *For each $v \in \{0,1\}^T$, the unique heating-minimal solution for v in $\hat{P}^{\mathrm{temp}}$ is $(v, \mathrm{y}_v, \mathrm{z}_v, \mathrm{temp}_v, \mathrm{h}_v) \in V^{\mathrm{temp}}$.*

By extension, the optimal solution of the heating-minimizing optimization problem in (3.2.11) is independent of the feasible set P with $V^{\mathrm{temp}} \subset P \subset \hat{P}^{\mathrm{temp}}$. We use this fact in the next section, where the formulation $\hat{P}^{\mathrm{temp}}$ is tightened using further valid inequalities for V^{temp}.

Theorem 3.13 *For each $v \in \{0,1\}^T$ and each P with $V^{\mathrm{temp}} \subset P \subset \hat{P}^{\mathrm{temp}}$, the unique heating-minimal solution for v in P is $(v, \mathrm{y}_v, \mathrm{z}_v, \mathrm{temp}_v, \mathrm{h}_v) \in V^{\mathrm{temp}}$.*

3.3 The Temperature Polyhedron

This section introduces a class of $\mathcal{O}(T^2)$ inequalities valid for V^{temp}, which can be separated in $\mathcal{O}(T)$. Adding these inequalities to the $\mathcal{H}$-representation of $\hat{P}^{\mathrm{temp}}$ results in an extended formulation of the summed start-up cost epigraph $\mathrm{conv}(\mathrm{epi}(\mathrm{DCU}^{\Sigma}))$.

Consider the operational schedules $v' = (1,0,0,0,0)$ and $v'' = (1,0,0,0,1)$ and their convex combination $v = (1-\mu)v' + \mu v'' = (1,0,0,0,\mu)$ in Fig. 3.4. Since this decomposition of v is unique, there exists exactly one point $(v, \mathrm{y}, \mathrm{z}, \mathrm{temp}, \mathrm{h}) \in \mathrm{conv}(V^{\mathrm{temp}})$ with $v = (1,0,0,0,\mu)$. Combining the temperatures of v' and v'' yields

$$\mathrm{temp}^5 = (1-\mu)\mathrm{temp}_{v_1}^5 + \mu \mathrm{temp}_{v_2}^5 = (1-v^5)\mathrm{e}^{-3\lambda} + v^5 = \mathrm{e}^{-3\lambda} \underbrace{\mathrm{temp}^2}_{=1} + (1 - \mathrm{e}^{-3\lambda}) v^5.$$

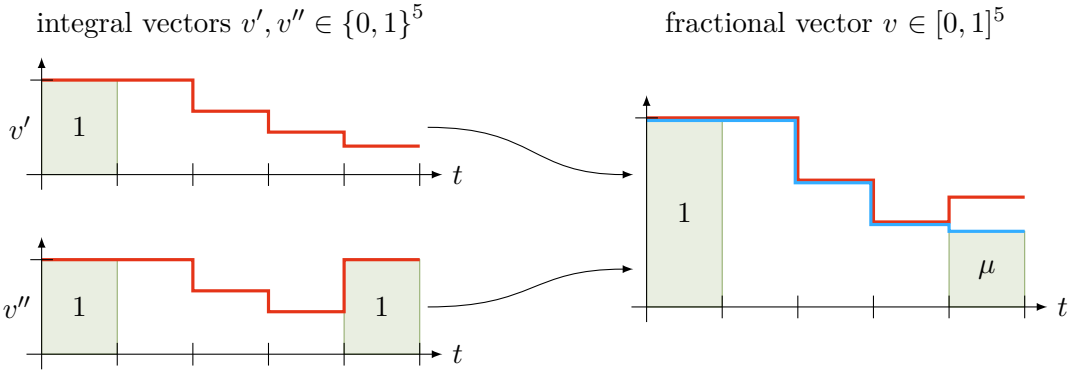


Figure 3.4: Convex combination of the vectors $v' = (1,0,0,0,0)$, $v'' = (1,0,0,0,1)$ and their respective temperatures (red) to $v = (1,0,0,0,\mu)$, in contrast to the minimal feasible temperature for v in $\hat{P}^{\text{temp}}$ (blue).

Generalizing the above example yields the *residual temperature inequalities (RTIs)*

$$\forall\, t \in [T], l \in [t-1]:$$
$$\text{temp}^t \geq \begin{cases} \mathrm{e}^{-\lambda \text{OL}^t(l)}\, \text{temp}^{t-l} + \left(1 - \mathrm{e}^{-\lambda \text{OL}^t(l)}\right) v^t & \text{if } l < t-1, \\ \mathrm{e}^{-\lambda \text{OL}^t(t-1)} + \left(1 - \mathrm{e}^{-\lambda \text{OL}^t(t-1)}\right) v^t & \text{if } l = t-1. \end{cases} \tag{3.3.1}$$

We denote the RTI with parameters t and l as RTI(t,l). Using the same approach as exemplified in Fig. 3.4, we show no RTI(t,l) is satisfied for all points in $\hat{P}^{\text{temp}}$.

Proposition 3.14 *For each $t \in [T]$, $l \in [t-1]$, there exists $(v, \mathrm{y}, \mathrm{z}, \text{temp}, \mathrm{h}) \in \hat{P}^{\text{temp}}$ violating RTI(t,l).*

Proof. It is straight-forward to check that the point $(v, \mathrm{y}, \mathrm{z}, \text{temp}, \mathrm{h})$ with

$$\forall\, t' \in [T]: \quad v^{t'} = \mathrm{y}^{t'} = \begin{cases} \mathrm{e}^{-\lambda \text{OL}^t(t-1)} & \text{if } t' = t, \\ 0 & \text{else,} \end{cases} \qquad \begin{aligned} \text{temp}^{t'} &= \mathrm{e}^{-\lambda \text{OL}^{t'}(t'-1)}, \\ \mathrm{h}^{t'} &= 0, \end{aligned}$$
$$\mathrm{z}^{t'} = \begin{cases} 1 & \text{if } t' = 1,\ \text{PDT} = 0, \\ \mathrm{e}^{-\lambda \text{OL}^t(t-1)} & \text{if } t' = t+1, \\ 0 & \text{else,} \end{cases}$$

fulfills the inequalities (3.2.1)-(3.2.9) and thereby lies in $\hat{P}^{\text{temp}}$.

On the other hand, this point does not satisfy RTI(t,l): For $l = t-1$, since $v^t = \mathrm{e}^{-\lambda \text{OL}^t(t-1)} > 0$ and $\mathrm{e}^{-\lambda \text{OL}^t(t-1)} < 1$, we have

$$\text{temp}^t = \mathrm{e}^{-\lambda \text{OL}^t(t-1)} < \mathrm{e}^{-\lambda \text{OL}^t(t-1)} + \left(1 - \mathrm{e}^{-\lambda \text{OL}^t(t-1)}\right) v^t.$$

and analogously for $l < t-1$

$$\mathrm{temp}^t = \mathrm{e}^{-\lambda \mathrm{OL}^t(t-1)} = \mathrm{e}^{-\lambda \mathrm{OL}^t(l)} \mathrm{e}^{-\lambda \mathrm{OL}^{t-l}(t-l-1)} < \mathrm{e}^{-\lambda \mathrm{OL}^t(l)}\, \mathrm{temp}^{t-l} + \left(1 - \mathrm{e}^{-\lambda \mathrm{OL}^t(l)}\right) v^t. \qquad \square$$

The terms $\mathrm{e}^{-\lambda \mathrm{OL}^t(l)} \mathrm{temp}^{t-l}$ and $\mathrm{e}^{-\lambda \mathrm{OL}^t(t-1)}$ represent the residual temperature left in period t, if no heating is applied during periods $t-l$ to $t-1$, motivating their name. The alternative formulation of the RTIs

$$\forall\, t \in [T], l \in [t-1]: \quad \mathrm{temp}^t - v^t \geq \begin{cases} \mathrm{e}^{-\lambda \mathrm{OL}^t(l)} (\mathrm{temp}^{t-l} - v^t) & \text{if } l < t-1, \\ \mathrm{e}^{-\lambda \mathrm{OL}^t(l)} (1 - v^t) & \text{if } l = t-1, \end{cases} \tag{3.3.2}$$

highlights their dominance over $\mathrm{temp}^t \geq v^t$ and is well-suited for proving its validity for V^{temp}.

Lemma 3.15 *All RTIs are valid for V^{temp}.*

Proof. Consider $(v, \mathrm{y}_v, \mathrm{z}_v, \mathrm{temp}_v, \mathrm{h}_v) \in V^{\mathrm{temp}}$ and $\mathrm{RTI}(t,l)$, $t \in [T]$, $l \in [t-1]$. If $v^t = 1$, then the left hand side of (3.3.2) equals 0 by (3.2.5) and the right hand side is not positive since $\mathrm{temp}_v \leq 1$.

Recall that $\mathrm{ol}^t(v)$ denotes the number of offline periods prior to t, and therefore $v^{t-l} = 0$ for $l \leq \mathrm{ol}^t(v)$. If $v^t = 0$, then for $l \leq \mathrm{ol}^t(v)$ by definition (3.1.2)

$$\mathrm{temp}_v^t = \mathrm{e}^{-\lambda \mathrm{OL}^t(\mathrm{ol}^t(v))} \quad \text{and} \quad \mathrm{temp}_v^{t-l} = \mathrm{e}^{-\lambda \mathrm{OL}^{t-l}(\mathrm{ol}^t(v)-l)}.$$

This leads to

$$\mathrm{temp}_v^t = \mathrm{e}^{-\lambda \mathrm{OL}^t(\mathrm{ol}^t(v))} \begin{cases} \geq \mathrm{e}^{-\lambda \mathrm{OL}^t(l)} \geq \mathrm{e}^{-\lambda \mathrm{OL}^t(l)} \mathrm{temp}_v^{t-l} & \text{if } l > \mathrm{ol}^t(v), \\ = \mathrm{e}^{-\lambda \mathrm{OL}^t(l)} = \mathrm{e}^{-\lambda \mathrm{OL}^t(l)} \mathrm{temp}_v^{t-l} & \text{if } l = \mathrm{ol}^t(v), \\ = \mathrm{e}^{-\lambda \mathrm{OL}^t(l)} \mathrm{e}^{-\lambda \mathrm{OL}^{t-l}(\mathrm{ol}^t(v)-l)} = \mathrm{e}^{-\lambda \mathrm{OL}^t(l)} \mathrm{temp}_v^{t-l} & \text{else.} \end{cases}$$

For $l < t-1$, we thus have

$$\mathrm{temp}_v^t - v^t = \mathrm{temp}_v^t \geq \mathrm{e}^{-\lambda \mathrm{OL}^t(l)} \mathrm{temp}_v^{t-l} = \mathrm{e}^{-\lambda \mathrm{OL}^t(l)} (\mathrm{temp}_v^{t-l} - v^t).$$

For $l = t-1$, we know that $l \geq \mathrm{ol}^t(v)$, and hence

$$\mathrm{temp}_v^t - v^t = \mathrm{temp}_v^t \geq \mathrm{e}^{-\lambda \mathrm{OL}^t(l)} = \mathrm{e}^{-\lambda \mathrm{OL}^t(l)} (1 - v^t). \qquad \square$$

The proof of the last statement already hints at the points in V^{temp} which fulfill an RTI with equality.

Proposition 3.16 *A vertex* $(v, \mathrm{y}_v, \mathrm{z}_v, \mathrm{temp}_v, \mathrm{h}_v) \in V^{\mathrm{temp}}$ *fulfills RTI*(t, l) *with* $t \in [T]$, $l \in [t-1]$ *with equality if*

- $v^{t-l} = \ldots = v^t = 0$,
- $l < t-1$ *and* $v^{t-l-1} = v^t = 1$, *or*
- $l = t-1$ *and* $v^t = 1$.

Proof. We first consider the case $v^{t-l} = \ldots = v^t = 0$, which implies $l \leq \mathrm{ol}^t(v)$. Note that $l = t-1$ necessitates $l = \mathrm{ol}^t(v)$. Similarly to the proof of Lemma 3.15, (3.1.2) and $v^t = 0$ yield

$$\mathrm{temp}_v^t = \mathrm{e}^{-\lambda \mathrm{OL}^t(\mathrm{ol}^t(v))} = \begin{cases} \mathrm{e}^{-\lambda \mathrm{OL}^t(l)} \mathrm{temp}_v^{t-l} + (1 - \mathrm{e}^{-\lambda \mathrm{OL}^t(l)}) v^t & \text{if } l < t-1, \\ \mathrm{e}^{-\lambda \mathrm{OL}^t(l)} + (1 - \mathrm{e}^{-\lambda \mathrm{OL}^t(l)}) v^t. & \text{if } l = t-1. \end{cases}$$

On the other hand, if $l < t-1$ and $v^{t-l-1} = v^t = 1$, then $\mathrm{temp}_v^{t-l} = \mathrm{temp}_v^t = 1$, and so

$$\mathrm{temp}_v^t = 1 = \mathrm{e}^{-\lambda \mathrm{OL}^t(l)} \mathrm{temp}_v^{t-l} + (1 - \mathrm{e}^{-\lambda \mathrm{OL}^t(l)}) v^t.$$

Analogously, if $l = t-1$ and $v^t = 1$, then

$$\mathrm{temp}_v^t = 1 = \mathrm{e}^{-\lambda \mathrm{OL}^t(l)} + (1 - \mathrm{e}^{-\lambda \mathrm{OL}^t(l)}) v^t. \qquad \square$$

In the following, we consider the polyhedron $\hat{P}^{\mathrm{temp}}$ tightened by the RTIs, which we call the temperature polyhedron P^{temp}.

Definition 3.17

$$P^{\mathrm{temp}} := \Big\{ (v, \mathrm{y}, \mathrm{z}, \mathrm{temp}, \mathrm{h}) \in \hat{P}^{\mathrm{temp}} \text{ fulfilling (3.3.1)} \Big\}.$$

Note that $P^{\mathrm{temp}} \subsetneq \hat{P}^{\mathrm{temp}}$, as Proposition 3.14 constructs a point for each $t \in [T]$, $l \in [t-1]$ which lies in $\hat{P}^{\mathrm{temp}} \setminus P^{\mathrm{temp}}$.

Since $V^{\mathrm{temp}} \subset P^{\mathrm{temp}} \subset \hat{P}^{\mathrm{temp}}$, Theorem 3.13 shows that each heating-minimal solution in P^{temp} for integral operational schedule $v \in \{0,1\}^T$ lies in V^{temp} and has correct start-up costs. Furthermore, the RTIs do not bound the resulting start-up costs from above, and so analogous to Corollary 3.6:

Corollary 3.18

$$\mathrm{conv}(\mathrm{epi}(\mathrm{DCU})) \subset \pi^{\mathrm{temp}}(P^{\mathrm{temp}}).$$

3.3.1 Equivalency to the Summed Start-up Cost Epigraph

Section 2.2 introduces the summed start-up cost DCU^Σ and its convex extension LCU^Σ, and analyzes the epigraph $\mathrm{epi}(\mathrm{LCU}^\Sigma) = \mathrm{conv}(\mathrm{epi}(\mathrm{DCU}^\Sigma))$. In this subsection, we show that P^{temp} is an extended formulation of $\mathrm{epi}(\mathrm{LCU}^\Sigma)$, i. e. that using the canonical orthogonal projection

$$\pi^\Sigma : \quad \mathbb{R}^{5T} \to \mathbb{R}^{2T}, \quad (v, \mathrm{y}, \mathrm{z}, \mathrm{temp}, \mathrm{h}) \mapsto (v, \mathrm{cu}^\Sigma)$$

with

$$\mathrm{cu}^\Sigma := \sum_{t\in[T]} \Big(\mathrm{CU}^{\mathrm{fixed}}\,\mathrm{y}^t + \mathrm{CU}^{\mathrm{var}}\,\mathrm{h}^{t-1}\Big),$$

we have $\pi^\Sigma(P^{\mathrm{temp}}) = \mathrm{epi}(\mathrm{LCU}^\Sigma)$.

Theorem 2.36 gives an $\mathcal{H}$-representation of $\mathrm{epi}(\mathrm{LCU}^\Sigma)$ composed of the binary tree inequalities (BTIs, cf. Definition 2.26) and the trivial inequalities $0 \le v \le 1$. We prove our claim by showing that each BTI is dominated by a set of inequalities of P^{temp} under the projection π^Σ.

Each BTI corresponds to a rank-labeled binary tree $B \in \mathcal{B}$ with nodes $[T]$ (see Definition 2.26). Recall that in such a binary tree,

- $S(t)$, $L(t)$, and $R(t)$ denote the principal subtree, the left subtree, and the right subtree of node t, respectively, and
- $s(t)$, $l(t)$, and $r(t)$ are defined as the number of nodes of the respective subtrees (cf. Subsection 2.2.2).

Furthermore, the definition of $\mathrm{OL}^t(l)$ is extended in (2.2.8) such that $\mathrm{OL}(S(t))$ denotes the downtime of a start-up after being offline during the periods $t \in S(t)$. Since $S(t)$ comprises the nodes $[t-l(t) \,..\, t+r(t)]$,

$$\mathrm{OL}(S(t)) := \mathrm{OL}^{t+r(t)+1}(s(t)).$$

Analogously, $\mathrm{OL}(L(t)) = \mathrm{OL}^t(l(t))$ and $\mathrm{OL}(R(t)) = \mathrm{OL}^{t+r(t)+1}(r(t))$.

Using this notation, the BTI are defined as

$$\forall\, B \in \mathcal{B}: \qquad \mathrm{cu}^\Sigma \ge \sum_{t\in[T]} \delta^t(l(t), r(t))\, v^t$$

with coefficients δ^t derived in (2.2.10) as

$$\delta^t(l(t), r(t)) = \begin{cases} \mathrm{CU}(\mathrm{OL}(L(t))) + \mathrm{CU}(\mathrm{OL}(R(t))) - \mathrm{CU}(\mathrm{OL}(S(t))) & \text{if } t + r(t) < T, \\ \mathrm{CU}(\mathrm{OL}(L(t))) & \text{else,} \end{cases}$$

where, the condition $t+r(t) < T$ holds iff t is not a top-right node (cf. Proposition 2.21).

We showed in Proposition 2.17 that these coefficients δ^t correspond to the change in the summed start-up costs cu^Σ associated with v when changing v^t from 0 to 1 (cf.

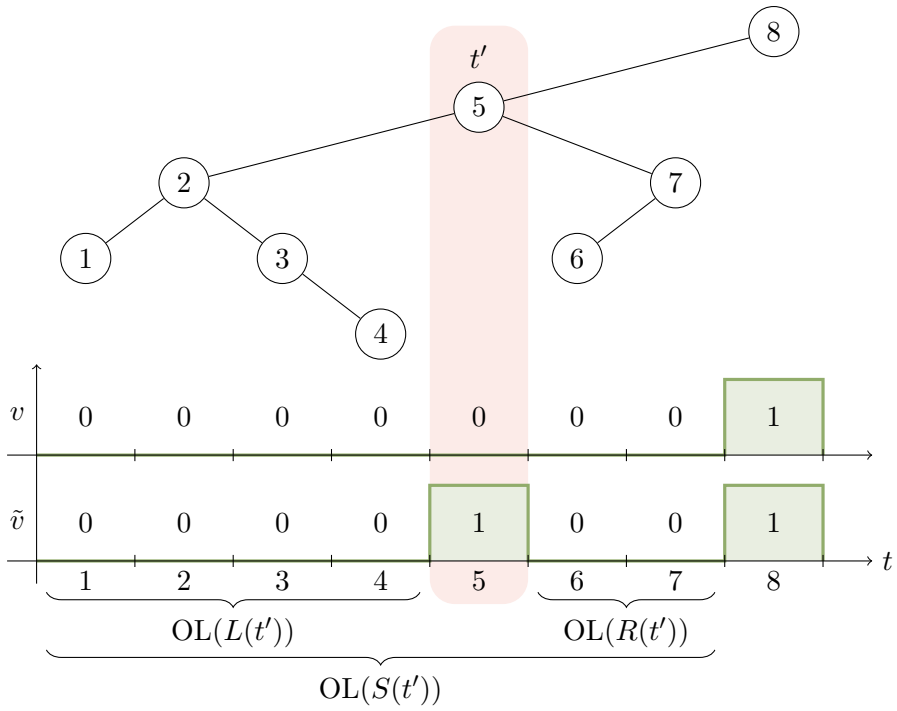


Figure 3.5: Change from $v^{t'} = 0$ to $v^{t'} = 1$ which defines the coefficient $\delta^{t'}$.

Fig. 3.5). If $t + r(t) = T$, the start-up costs $\mathrm{CU}(\mathrm{OL}(R(t)))$, $\mathrm{CU}(\mathrm{OL}(S(t)))$ are dropped from δ^t, since both refer to a start-up in period $T + 1$ which lies outside the modeled time range.

For the exponential start-up cost function, the start-up costs $\mathrm{CU}(\mathrm{OL}(L(t)))$ in the coefficients δ^t of the BTIs equate to

$$\mathrm{CU}(\mathrm{OL}(L(t))) = \mathrm{CU}^{\mathrm{var}}\left(1 - \mathrm{e}^{-\lambda \mathrm{OL}(L(t))}\right) + \begin{cases} \mathrm{CU}^{\mathrm{fixed}} & \text{if } \mathrm{OL}(L(t)) > 0, \\ 0 & \text{else.} \end{cases}$$

For $t + r(t) < T$, using that $\mathrm{OL}(S(t)) \geq \mathrm{L}^t > 0$ and $(\mathrm{OL}(R(t)) > 0 \Leftrightarrow r(t) > 0)$, the remaining start-up costs in δ^t equate to

$$\mathrm{CU}(\mathrm{OL}(R(t))) = \mathrm{CU}^{\mathrm{var}}\left(1 - \mathrm{e}^{-\lambda \mathrm{OL}(R(t))}\right) + \begin{cases} \mathrm{CU}^{\mathrm{fixed}} & \text{if } r(t) > 0, \\ 0 & \text{else,} \end{cases}$$

$$\mathrm{CU}(\mathrm{OL}(S(t))) = \mathrm{CU}^{\mathrm{var}}\left(1 - \mathrm{e}^{-\lambda \mathrm{OL}(S(t))}\right) + \mathrm{CU}^{\mathrm{fixed}}.$$

These start-up costs are composed by a fixed part with factor $\mathrm{CU}^{\mathrm{fixed}}$ and a variable part with factor $\mathrm{CU}^{\mathrm{var}}$. Substituting the start-up costs in the definition of δ^t shows that δ^t can be decomposed analogously.

Proposition 3.19 *For each $B \in \mathcal{B}$ and each $t \in [T]$, defining*

$$\beta^t = \left\{ \begin{cases} 1 & \text{if } \mathrm{OL}(L(t)) > 0, \\ 0 & \text{else,} \end{cases} \right\} - \left\{ \begin{cases} 1 & \text{if } r(t) = 0 \text{ and } t < T, \\ 0 & \text{else,} \end{cases} \right\},$$

$$\gamma^t = \begin{cases} 1 + e^{-\lambda \mathrm{OL}(S(t))} - e^{-\lambda \mathrm{OL}(L(t))} - e^{-\lambda \mathrm{OL}(R(t))} & \text{if } t + r(t) < T, \\ 1 - e^{-\lambda \mathrm{OL}(L(t))} & \text{else.} \end{cases}$$

each binary tree inequality can be written as

$$\mathrm{cu}^\Sigma \geq \sum_{t \in [T]} \delta^t(l(t), r(t)) v^t = \mathrm{CU}^{\mathrm{fixed}} \sum_{t \in [T]} \beta^t v^t + \mathrm{CU}^{\mathrm{var}} \sum_{t \in [T]} \gamma^t v^t.$$

The division in fixed and variable costs is reflected in the projection π^Σ, where the sum of the start-up costs is defined as

$$\mathrm{cu}^\Sigma = \mathrm{CU}^{\mathrm{fixed}} \sum_{t \in [T]} \mathrm{y}^t + \mathrm{CU}^{\mathrm{var}} \sum_{t \in [T]} \mathrm{h}^{t-1}.$$

We prove the dominance of the inequalities of P^{temp} over each BTI by showing that for each $(v, \mathrm{y}, \mathrm{z}, \mathrm{temp}, \mathrm{h}) \in P^{\mathrm{temp}}$

$$\sum_{t \in [T]} \mathrm{y}^t \geq \sum_{t \in [T]} \beta^t v^t \quad \text{and} \quad \sum_{t \in [T]} \mathrm{h}^{t-1} \geq \sum_{t \in [T]} \gamma^t v^t,$$

which implies that for $(v, \mathrm{cu}^\Sigma) = \pi^\Sigma(v, \mathrm{y}, \mathrm{z}, \mathrm{temp}, \mathrm{h})$,

$$\mathrm{cu}^\Sigma = \mathrm{CU}^{\mathrm{fixed}} \sum_{t \in [T]} \mathrm{y}^t + \mathrm{CU}^{\mathrm{var}} \sum_{t \in [T]} \mathrm{h}^{t-1} \geq \mathrm{CU}^{\mathrm{fixed}} \sum_{t \in [T]} \beta^t v^t + \mathrm{CU}^{\mathrm{var}} \sum_{t \in [T]} \gamma^t v^t. \quad (3.3.3)$$

We first consider the fixed costs: There, as we will see, the lower bounds on y^t as given by (3.2.2) of $\hat{P}^{\mathrm{temp}}$,

$$\mathrm{y}^t = v^t - v^{t-1} + \mathrm{z}^t \geq v^t - v^{t-1} \qquad \text{and} \qquad \mathrm{y}^t \geq 0,$$

suffice.

Lemma 3.20 *For each $(v, \mathrm{y}, \mathrm{z}, \mathrm{temp}, \mathrm{h}) \in P^{\mathrm{temp}}$ and each $B \in \mathcal{B}$,*

$$\sum_{t \in [T]} \mathrm{y}^t \geq \sum_{t \in [T]} \beta^t v^t.$$

Proof. The sum of the start-up indicator variables is bounded by

$$\sum_{t \in [T]} \mathrm{y}^t \geq \sum_{\substack{t \in [T] \\ \mathrm{OL}(L(t)) > 0}} \mathrm{y}^t \geq \sum_{\substack{t \in [T] \\ \mathrm{OL}(L(t)) > 0}} v^t - \sum_{\substack{t \in [2..T] \\ \mathrm{OL}(L(t)) > 0}} v^{t-1},$$

since we have $\mathrm{y}^t \geq v^t - v^{t-1}$ for $t \geq 2$, and since, if $\mathrm{OL}^1(0) > 0$, we also have $\mathrm{PDT} > 0$ and thus $\mathrm{y}^1 = v^1$.

For $t \geq 2$, $\mathrm{OL}(L(t)) > 0$ is equivalent to $l(t) > 0$, which is equivalent to $r(t-1) = 0$. Thus, we have

$$\sum_{\substack{t \in [2..T] \\ \mathrm{OL}(L(t))>0}} v^{t-1} = \sum_{\substack{t \in [T-1] \\ r(t)=0}} v^t,$$

and therefore the bound on the start-up indicators can be reformulated as

$$\sum_{t \in [T]} \mathrm{y}^t \geq \sum_{\substack{t \in [T] \\ \mathrm{OL}(L(t))>0}} v^t - \sum_{\substack{t \in [T-1] \\ r(t)=0}} v^t = \sum_{t \in [T]} \beta^t v^t. \qquad \square$$

Subsection 2.2.1 highlights that the binary tree structure underlying the BTIs is caused by the fact that for an operational schedule $v \in \{0,1\}^T$ with $v^t = 1$, the start-up costs in periods prior and after t are independent. This property is intrinsic to the temperature: $v^t = 1$ implies $\mathrm{temp}^t = 1$, and hence the temperatures prior and after t are independent.

Furthermore, recall the cost-optimal decomposition of a point (v, cu^Σ) performed in the proof of Lemma 2.33 and represented in Fig. 2.20. Considering a different example, Fig. 3.6 shows that this decomposition fits all $\mathrm{RTI}(t, l(t))$.

We prove in the remainder of this section that the variable part of each BTI is formed by combining the RTIs with parameters t and $l = l(t)$, using the temperature development inequalities (3.2.6) and (3.2.7), which amounts to showing

$$\sum_{t \in [0\,..\,T-1]} \mathrm{h}^t \geq \sum_{t \in [T]} \gamma^t v^t.$$

The proof considers subtrees of B of increasing size:

- Proposition 3.22 starts with nodes t without a left child $\mathrm{llink}(t)$,
- Lemma 3.23 continues with principal subtrees $S(t)$ of nodes t that are not top-right, i. e. with $t + r(t) < 1$,
- Lemma 3.24 proceeds with left subtrees $L(t)$ of top-right nodes t, and
- Lemma 3.25 concludes with the complete tree.

To clearly identify the RTIs which are necessary to dominate a given BTI, we summarize the contribution of the RTIs in the following technical lemma. Note that for a given $B \in \mathcal{B}$, only inequalities $\mathrm{RTI}(t, l)$ with $l = l(t)$ are required in this proof and in the following results.

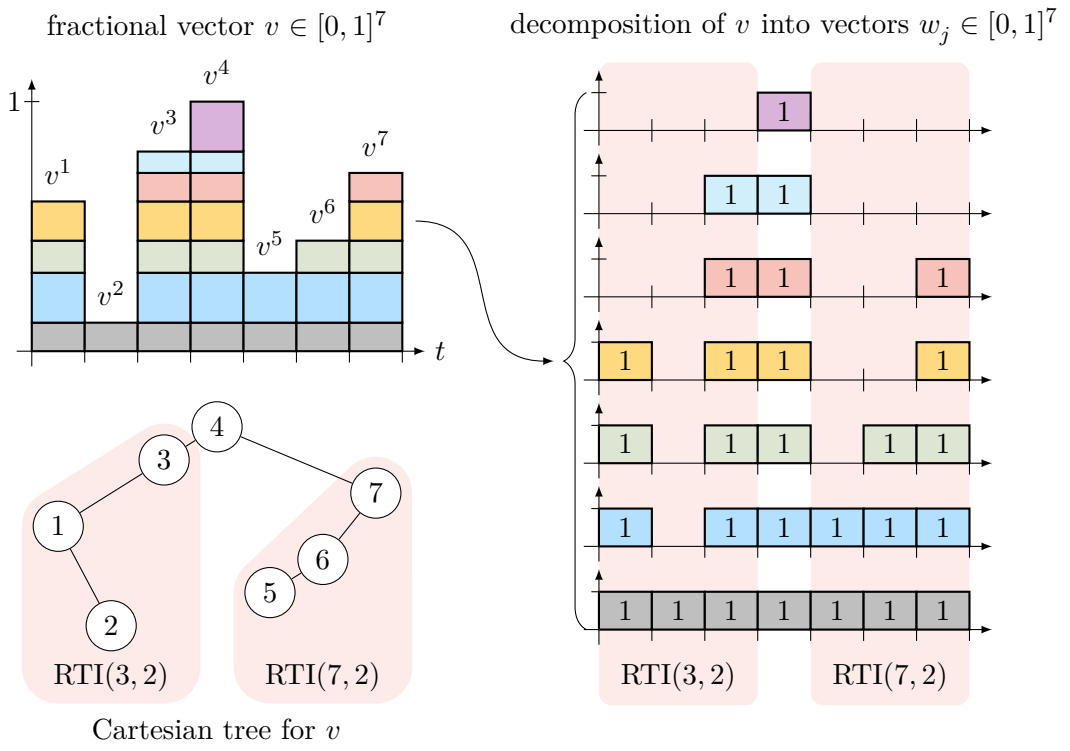


Figure 3.6: The summed start-up costs $\mathrm{LCU}^{\Sigma}(v)$ of the fractional $v \in [0,1]^7$ (top left) are given by the BTI defined by the Cartesian tree for v (bottom left). Each point $(w_j, \mathrm{DCU}^{\Sigma}(w_j))$ in the decomposition according to Lemma 2.33 of $(v, \mathrm{LCU}^{\Sigma}(v))$ (right) fulfills this BTI with equality. Proposition 3.16 shows that each w_j and its induced temperature also fulfills $\mathrm{RTI}(t, l(t))$.

Lemma 3.21 *Let* $(v, \mathrm{y}, \mathrm{z}, \mathrm{temp}, \mathrm{h}) \in P^{\mathrm{temp}}$, $B \in \mathcal{B}$, *and* $t \in B$ *be given. If* $t + r(t) < T$ *then*

$$\begin{aligned} \mathrm{h}^t \geq\ & \gamma^t v^t + e^{-\lambda \mathrm{OL}(R(t))}\mathrm{temp}^{t+1} - \mathrm{temp}^t \\ & + \begin{cases} (e^{-\lambda \mathrm{OL}(L(t))} - e^{-\lambda \mathrm{OL}(S(t))})\mathrm{temp}^{t-l(t)} & \text{if } l(t) < t-1, \\ (e^{-\lambda \mathrm{OL}(L(t))} - e^{-\lambda \mathrm{OL}(S(t))}) & \text{else,} \end{cases} \end{aligned} \tag{3.3.4}$$

and if $t + r(t) = T$ *then*

$$\mathrm{temp}^t - \left\{ \begin{matrix} e^{-\lambda \mathrm{OL}(L(t))}\mathrm{temp}^{t-l(t)} & \text{if } l(t) < t-1 \\ e^{-\lambda \mathrm{OL}(L(t))} & \text{else} \end{matrix} \right\} \geq \gamma^t v^t. \tag{3.3.5}$$

Proof. If $t + r(t) = T$, the statement follows directly from the definition of $\mathrm{RTI}(t, l(t))$ in (3.3.1),

$$\mathrm{temp}^t - \left.\begin{cases} \mathrm{e}^{-\lambda \mathrm{OL}(L(t))}\mathrm{temp}^{t-l(t)} & \text{if } l(t) < t-1 \\ \mathrm{e}^{-\lambda \mathrm{OL}(L(t))} & \text{else} \end{cases}\right\} \geq (1 - \mathrm{e}^{-\lambda \mathrm{OL}(L(t))})v^t = \gamma^t v^t.$$

Else, if $t + r(t) < T$, then the temperature development equation (3.2.7) and $\mathrm{h}^t \geq 0$ imply

$$\begin{aligned} \mathrm{h}^t &\geq \mathrm{e}^{-\lambda \mathrm{OL}(R(t))}\mathrm{h}^t \\ &= \mathrm{e}^{-\lambda \mathrm{OL}(R(t))}\mathrm{temp}^{t+1} - \mathrm{e}^{-\lambda(\mathrm{OL}(R(t))+\mathrm{L}^t)}\mathrm{temp}^t - \mathrm{e}^{-\lambda \mathrm{OL}(R(t))}(1 - \mathrm{e}^{-\lambda \mathrm{L}^t})v^t \\ &= \mathrm{e}^{-\lambda \mathrm{OL}(R(t))}\mathrm{temp}^{t+1} - \mathrm{temp}^t + (1 - \mathrm{e}^{-\lambda \mathrm{OL}(R(t))})v^t \\ &\quad + (1 - \mathrm{e}^{-\lambda(\mathrm{OL}(R(t))+\mathrm{L}^t)})(\mathrm{temp}^t - v^t). \end{aligned} \tag{3.3.6}$$

If $l(t) < t-1$, then $\mathrm{RTI}(t, l(t))$ as given in (3.3.2) and multiplied by the non-negative factor $(1 - \mathrm{e}^{-\lambda(\mathrm{OL}(R(t))+\mathrm{L}^t)})$ equals

$$\begin{aligned} (1 - \mathrm{e}^{-\lambda(\mathrm{OL}(R(t))+\mathrm{L}^t)})(\mathrm{temp}^t - v^t) &\geq (1 - \mathrm{e}^{-\lambda(\mathrm{OL}(R(t))+\mathrm{L}^t)})\mathrm{e}^{-\lambda \mathrm{OL}(L(t))}(\mathrm{temp}^{t-l(t)} - v^t) \\ &\overset{(2.2.19)}{=} \left(\mathrm{e}^{-\lambda \mathrm{OL}(L(t))} - \mathrm{e}^{-\lambda \mathrm{OL}(S(t))}\right)(\mathrm{temp}^{t-l(t)} - v^t). \end{aligned}$$

Applying this inequality to (3.3.6) yields

$$\begin{aligned} \mathrm{h}^t \geq\ & \mathrm{e}^{-\lambda \mathrm{OL}(R(t))}\mathrm{temp}^{t+1} - \mathrm{temp}^t + \left(\mathrm{e}^{-\lambda \mathrm{OL}(L(t))} - \mathrm{e}^{-\lambda \mathrm{OL}(S(t))}\right)\mathrm{temp}^{t-l(t)} \\ &+ \Big(\underbrace{1 - \mathrm{e}^{-\lambda \mathrm{OL}(L(t))} - \mathrm{e}^{-\lambda \mathrm{OL}(R(t))} + \mathrm{e}^{-\lambda \mathrm{OL}(S(t))}}_{=\gamma^t}\Big)v^t. \end{aligned}$$

If $l(t) = t-1$, the statement follows analogously without "$\mathrm{temp}^{t-l(t)}$", due to the case distinction in the definition of the RTIs (3.3.1). □

Proposition 3.22 *Let $(v, \mathrm{y}, \mathrm{z}, \mathrm{temp}, \mathrm{h}) \in P^{\mathrm{temp}}$, $B \in \mathcal{B}$, and $t \in B$ with $l(t) = 0$ be given. If $l(t) < t-1$, then*

$$0 = \mathrm{temp}^t - e^{-\lambda \mathrm{OL}(L(t))}\mathrm{temp}^{t-l(t)},$$

and if $l(t) = t-1$, then

$$\mathrm{h}^0 = \mathrm{temp}^t - e^{-\lambda \mathrm{OL}(L(t))}.$$

Proof. If $l(t) < t-1$, then $\mathrm{OL}(L(t)) = \mathrm{OL}^t(l(t)) = \mathrm{OL}^t(0) = 0$, implying

$$0 = \mathrm{temp}^t - \mathrm{e}^{-\lambda \mathrm{OL}(L(t))}\mathrm{temp}^t = \mathrm{temp}^t - \mathrm{e}^{-\lambda \mathrm{OL}(L(t))}\mathrm{temp}^{t-l(t)}.$$

If $l(t) = t-1$, then $t = 1$ and $\mathrm{e}^{-\lambda \mathrm{OL}(L(t))} = \mathrm{e}^{-\lambda \mathrm{PDT}}$. The equation for the initial temperature (3.2.6) implies

$$\mathrm{h}^0 = \mathrm{temp}^1 - \mathrm{e}^{-\lambda \mathrm{PDT}} = \mathrm{temp}^t - \mathrm{e}^{-\lambda \mathrm{OL}(L(t))}.$$ □

Lemma 3.23 *Let* $(v, \mathrm{y}, \mathrm{z}, \mathrm{temp}, \mathrm{h}) \in P^{\mathrm{temp}}$, $B \in \mathcal{B}$, *and* $t \in B$ *with* $t + r(t) < T$ *be given. If* $l(t) < t - 1$, *then*

$$\sum_{t' \in S(t)} \mathrm{h}^{t'} \geq \mathrm{temp}^{t+r(t)+1} - e^{-\lambda \mathrm{OL}(S(t))} \mathrm{temp}^{t-l(t)} + \sum_{t' \in S(t)} \gamma^{t'} v^{t'},$$

and if $l(t) = l - 1$, *then*

$$\mathrm{h}^0 + \sum_{t' \in S(t)} \mathrm{h}^{t'} \geq \mathrm{temp}^{t+r(t)+1} - e^{-\lambda \mathrm{OL}(S(t))} + \sum_{t' \in S(t)} \gamma^{t'} v^{t'}.$$

Proof. The above inequalities (in particular $\mathrm{h}^{t'}$ and $\mathrm{temp}_t^{t+r(t)+1}$) are well-defined since $t + r(t) < T$, and thus $T \notin S(t)$.

We start by proving the claim for nodes $t \in B$ with $l(t) < t - 1$ by induction over the size $s(t)$ of $S(t)$. For the induction start, i. e. for $t \in B$ with $s(t) = 1$, Lemma 3.23 is implied Lemma 3.21. We consider $t \in B$ with $s(t) > 1$ in the following.

If t has a left child t_l, it holds that

$$\begin{aligned} \sum_{t' \in L(t)} \mathrm{h}^{t'} = \sum_{t' \in S(t_l)} \mathrm{h}^{t} &\overset{\text{ind. hyp.}}{\geq} \mathrm{temp}^{t_l + r(t_l)+1} - \mathrm{e}^{-\lambda \mathrm{OL}(S(t_l))} \mathrm{temp}^{t_l - l(t_l)} + \sum_{t' \in S(t_l)} \gamma^{t'} v^{t'} \\ &= \mathrm{temp}^{t} - \mathrm{e}^{-\lambda \mathrm{OL}(L(t))} \mathrm{temp}^{t-l(t)} + \sum_{t' \in L(t)} \gamma^{t'} v^{t'}. \end{aligned} \tag{3.3.7}$$

If t has no left child, (3.3.7) holds because of Proposition 3.22. Analogously, if t has a right child t_r, the induction yields

$$\begin{aligned} \sum_{t' \in R(t)} \mathrm{h}^{t'} = \sum_{t' \in S(t_r)} \mathrm{h}^{t} &\overset{\text{ind. hyp.}}{\geq} \mathrm{temp}^{t_r + r(t_r)+1} - \mathrm{e}^{-\lambda \mathrm{OL}(S(t_r))} \mathrm{temp}^{t_r - l(t_r)} + \sum_{t' \in S(t_r)} \gamma^{t'} v^{t'} \\ &= \mathrm{temp}^{t+r(t)+1} - \mathrm{e}^{-\lambda \mathrm{OL}(R(t))} \mathrm{temp}^{t+1} + \sum_{t' \in R(t)} \gamma^{t'} v^{t'}. \end{aligned} \tag{3.3.8}$$

If t has no right child, (3.3.8) holds since

$$\mathrm{temp}_t^{t+r(t)+1} = \mathrm{temp}_t^{t+1} = \mathrm{e}^{-\lambda \mathrm{OL}(R(t))} \mathrm{temp}^{t+1}.$$

By combining the inequalities (3.3.4), (3.3.7), and (3.3.8), the sum of the heating variables can be bounded by

$$\begin{aligned}\sum_{t'\in S(t)} \mathrm{h}^{t'} &= \mathrm{h}^t + \sum_{t'\in L(t)} \mathrm{h}^{t'} + \sum_{t'\in R(t)} \mathrm{h}^{t'} \\ &\geq \mathrm{e}^{-\lambda \mathrm{OL}(R(t))}\mathrm{temp}^{t+1} - \mathrm{temp}^t + (\mathrm{e}^{-\lambda \mathrm{OL}(L(t))} - \mathrm{e}^{-\lambda \mathrm{OL}(S(t))})\mathrm{temp}^{t-l(t)} + \gamma^t v^t \\ &\quad + \mathrm{temp}^t - \mathrm{e}^{-\lambda \mathrm{OL}(L(t))}\mathrm{temp}^{t-l(t)} + \sum_{t'\in L(t)} \gamma^{t'} v^{t'} \\ &\quad + \mathrm{temp}^{t+r(t)+1} - \mathrm{e}^{-\lambda \mathrm{OL}(R(t))}\mathrm{temp}^{t+1} + \sum_{t'\in R(t)} \gamma^{t'} v^{t'} \\ &= \mathrm{temp}^{t+r(t)+1} - \mathrm{e}^{-\lambda \mathrm{OL}(S(t))}\mathrm{temp}^{t-l(t)} + \sum_{t'\in S(t)} \gamma^{t'} v^{t'}.\end{aligned}$$

For nodes $t \in B$ with $l(t) = t - 1$, the statement follows analogously □

Lemma 3.24 *Let $(v, \mathrm{y}, \mathrm{z}, \mathrm{temp}, \mathrm{h}) \in P^{\mathrm{temp}}$, $B \in \mathcal{B}$, and $t \in B$ with $t + r(t) = T$ be given. If $l(t) < t - 1$, then*

$$\sum_{t'\in L(t)} \mathrm{h}^{t'} \geq \sum_{t'\in L(t)} \gamma^{t'} v^{t'} + \gamma^t v^t,$$

and if $l(t) = t - 1$,

$$\mathrm{h}^0 + \sum_{t'\in L(t)} \mathrm{h}^{t'} \geq \sum_{t'\in L(t)} \gamma^{t'} v^{t'} + \gamma^t v^t.$$

Proof. For each node $t \in B$ with $t + r(t) = T$ and $l(t) < t - 1$, we have by Proposition 3.22 (if t has no left child) or by Lemma 3.23 (applied to left child $\mathrm{llink}(t)$ of t) that

$$\sum_{t'\in L(t)} \mathrm{h}^{t'} \geq \mathrm{temp}^t - \mathrm{e}^{-\lambda \mathrm{OL}(L(t))}\mathrm{temp}^{t-l(t)} + \sum_{t'\in L(t)} \gamma^{t'} v^{t'},$$

which, by (3.3.5) may be bounded by

$$\sum_{t'\in L(t)} \mathrm{h}^{t'} \geq \sum_{t'\in L(t)} \gamma^{t'} v^{t'} + \gamma^t v^t.$$

For the root node $t = \mathrm{root}(B)$, which is the unique node with $t + r(t) = T$ and $l(t) = t - 1$, we analogously have

$$\mathrm{h}^0 + \sum_{t'\in L(t)} \mathrm{h}^{t'} \geq \mathrm{temp}^t - \mathrm{e}^{-\lambda \mathrm{OL}(L(t))} + \sum_{t'\in L(t)} \gamma^{t'} v^{t'} \geq \sum_{t'\in L(t)} \gamma^{t'} v^{t'} + \gamma^t v^t$$

due to Proposition 3.22, Lemma 3.23, and (3.3.5). □

Lemma 3.25 *For each* $(v, \mathrm{y}, \mathrm{z}, \mathrm{temp}, \mathrm{h}) \in P^{\mathrm{temp}}$ *and* $B \in \mathcal{B}$ *it holds that*

$$\sum_{t\in[T]} \mathrm{h}^{t-1} \geq \sum_{t\in[T]} \gamma^t v^t.$$

Proof. The top-right nodes of B (see Definition 2.20) are defined recursively: the first top-right node η_1 is the root, and each subsequent top-right node η_i is the right child of η_{i-1}. We repeat basic facts on top-right nodes from Subsection 2.2.2:

- the depth d of each top-right node η_i is $i-1$, i. e. $d(\eta_i) = i+1$, and the last top-right node is $\eta_{d(T)+1} = T$ (see Definition 2.20),
- each node of B is either a top-right node or belongs to the left subtree of a top-right node (see Corollary 2.23), and
- for each top-right node t it holds that $t + r(t) = T$ (see Proposition 2.21).

Thus, the heating variables h^t can be split into top-right nodes and into left subtrees of top-right nodes as

$$\sum_{t\in[T]} \mathrm{h}^{t-1} = \mathrm{h}^0 + \sum_{r=1}^{d(T)+1} \sum_{t\in L(\eta_r)} \mathrm{h}^t + \sum_{r=1}^{d(T)} \mathrm{h}^{\eta_i} \geq \mathrm{h}^0 + \sum_{t\in L(\eta_1)} \mathrm{h}^t + \sum_{r=2}^{d(T)+1} \sum_{t\in L(\eta_r)} \mathrm{h}^t,$$

and, using Lemma 3.23 and Lemma 3.24, bounded by

$$\sum_{t\in[T]} \mathrm{h}^{t-1} \geq \sum_{r=1}^{d(T)+1} \left(\sum_{t\in L(\eta_r)} \gamma^t v^t + \gamma^{\eta_r} v^{\eta_r} \right) = \sum_{t\in[T]} \gamma^t v^t. \qquad \square$$

As pointed out in (3.3.3), the statements of Lemma 3.20 and Lemma 3.25 imply that each BTI is dominated by inequalities of P^{temp}, as

$$\sum_{t\in[T]} \left(\mathrm{CU}^{\mathrm{fixed}}\, \mathrm{y}^t + \mathrm{CU}^{\mathrm{var}}\, \mathrm{h}^{t-1} \right) \geq \mathrm{CU}^{\mathrm{fixed}} \sum_{t\in[T]} \beta^t v^t + \mathrm{CU}^{\mathrm{var}} \sum_{t\in[T]} \gamma^t v^t = \sum_{t\in[T]} \delta^t v^t.$$

Therefore, the projection $\pi^{\Sigma}(p)$ of each point $p \in P^{\mathrm{temp}}$ fulfills all BTIs and lies in $\mathrm{epi}(\mathrm{LCU}^{\Sigma})$. On the other hand, $\mathrm{conv}(\mathrm{epi}(\mathrm{DCU})) \subset \pi^{\mathrm{temp}}(P^{\mathrm{temp}})$ (see Corollary 3.18), which by definition of π^{Σ} implies $\mathrm{epi}(\mathrm{LCU}^{\Sigma}) \subset \pi^{\Sigma}(P^{\mathrm{temp}})$.

Theorem 3.26

$$\pi^{\Sigma}(P^{\mathrm{temp}}) = \mathrm{epi}(\mathrm{LCU}^{\Sigma}).$$

3.3.2 Separation

Since each of the $\mathcal{O}(T^2)$ RTI has 3 non-zero coefficients, an exhaustive search separates the RTI in $\mathcal{O}(T^2)$. In this section, we use Lemma 3.21 to identify T crucial RTIs for each $v \in [0,1]^T$, based on which we present an $\mathcal{O}(T)$ algorithm that

- separates the projection $\pi^\Sigma(p)$ of each point $p \in \hat{P}^{\text{temp}}$ from $\text{epi}(\text{LCU}^\Sigma)$, and
- separates each heating-minimal point $p \in \hat{P}^{\text{temp}}$ from P^{temp}.

This separation algorithm may be extended canonically to arbitrary points $p \in \mathbb{R}^{5T}$ while preserving the running time, as each of the $\mathcal{O}(T)$ inequalities in the $\mathcal{H}$-representation of $\hat{P}^{\text{temp}}$ has at most 4 non-zero coefficients.

The proof of $\pi^\Sigma(P^{\text{temp}}) = \text{epi}(\text{LCU}^\Sigma)$ in the last section shows that each of the BTIs, which define $\text{epi}(\text{LCU}^\Sigma)$, is dominated by a projection of the inequalities of $\hat{P}^{\text{temp}}$ and the RTIs. Moreover, the RTIs used in this projection are identified in the proof of Lemma 3.21, and depend solely on B:

Corollary 3.27 *For each $B \in \mathcal{B}$ and $p \in \hat{P}^{\text{temp}}$, if p fulfills RTI$(t, l(t))$ for each $t \in [T]$ with $l(t) > 0$, then $\pi^\Sigma(p)$ fulfills the BTI induced by B.*

By Corollary 2.35, a point $(v, \text{cu}^\Sigma) \in [0,1]^T \times \mathbb{R}$ lies in $\text{epi}(\text{LCU}^\Sigma)$ iff the BTI induced by a Cartesian tree B for v (cf. Definition 2.32) is fulfilled.

Corollary 3.28 *For each $p = (v, \text{y}, \text{z}, \text{temp}, \text{h}) \in \hat{P}^{\text{temp}}$ with Cartesian tree B for v, if p fulfills RTI$(t, l(t))$ for each $t \in [T]$ with $l(t) > 0$, then $\pi^\Sigma(p) \in \text{epi}(\text{LCU}^\Sigma)$.*

Thus, to separate the projection $\pi^\Sigma(p)$ of a point $p \in \hat{P}^{\text{temp}}$, it suffices to check all $\text{RTI}(t, l)$ with $l = l(t) > 0$:

- if these inequalities are fulfilled then $\pi^\Sigma(p) \in \text{epi}(\text{LCU}^\Sigma)$, and
- else the violated RTIs separate p from P^{temp}.

To simplify the algorithm, we note that the definition of the RTI can be extended canonically to $l = 0$,

$$\forall\, t \in [T], l \in [0 \,..\, t-1]:$$
$$\text{temp}^t \geq \begin{cases} \mathrm{e}^{-\lambda \text{OL}^t(l)}\, \text{temp}^{t-l} + \left(1 - \mathrm{e}^{-\lambda \text{OL}^t(l)}\right) v^t & \text{if } l < t-1, \\ \mathrm{e}^{-\lambda \text{OL}^t(t-1)} + \left(1 - \mathrm{e}^{-\lambda \text{OL}^t(t-1)}\right) v^t & \text{if } l = t-1, \end{cases}$$

such that $\text{RTI}(t, l)$ with $l = 0$ are always fulfilled:

Proposition 3.29 *For each point $(v, \text{y}, \text{z}, \text{temp}, \text{h}) \in \hat{P}^{\text{temp}}$ and each $t \in [T]$, the inequality RTI$(t, 0)$ holds with equality.*

Proof. The claim holds for $t \in [2..T]$ since $\mathrm{OL}^t(0) = 0$. For $t = 1$, we have $\mathrm{OL}^1(0) = \mathrm{PDT}$, and therefore (3.2.6) and (3.2.9) imply

$$\mathrm{temp}^1 = \mathrm{e}^{-\lambda \mathrm{PDT}} + \mathrm{h}^0 = \mathrm{e}^{-\lambda \mathrm{PDT}} + (1 - \mathrm{e}^{-\lambda \mathrm{PDT}})v^1 = \mathrm{e}^{-\lambda \mathrm{OL}^1(0)} + (1 - \mathrm{e}^{-\lambda \mathrm{OL}^1(0)})v^1. \quad \square$$

We give the separation algorithm without explicitly calculating the subtree sizes $l(t)$ and the downtimes $\mathrm{OL}^t(l(t))$. See Algorithm 2.2.1 for how these values can be derived in $\mathcal{O}(T)$ for a binary tree $B \in \mathcal{B}$.

Algorithm 3.3.1: SeparateRTI

Input : Point $(v, \mathrm{y}, \mathrm{z}, \mathrm{temp}, \mathrm{h}) \in \hat{P}^{\mathrm{temp}}$
Output : $\pi^{\Sigma}(v, \mathrm{y}, \mathrm{z}, \mathrm{temp}, \mathrm{h})$ lies in $\mathrm{epi}(\mathrm{LCU}^{\Sigma})$ or a violated RTI.

```
B ← FindCartesianTree (v) (represented by root, llink, and rlink);
for t ∈ [T] do
    if l(t) < t − 1 then
        if temp^t < e^{−λOL^t(l(t))} temp^{t−l(t)} + (1 − e^{−λOL^t(l(t))}) v^t then
            stop (v, y, z, temp, h) violates RTI(t, l(t));
    else
        if temp^t < e^{−λOL^t(l(t))} + (1 − e^{−λOL^t(l(t))}) v^t then
            stop (v, y, z, temp, h) violates RTI(t, l(t));
stop π^Σ(v, y, z, temp, h) lies in epi(LCU^Σ);
```

Proposition 3.30 *For each $p \in \hat{P}^{\mathrm{temp}}$, Algorithm 3.3.1 confirms $\pi^{\Sigma}(p) \in \mathrm{epi}(\mathrm{LCU}^{\Sigma})$ or stops with an RTI which separates p from P^{temp}, in $\mathcal{O}(T)$.*

By using Algorithm 3.3.1 in a cutting plane approach, we can model the constraint $\pi^{\Sigma}(v, \mathrm{y}, \mathrm{z}, \mathrm{temp}, \mathrm{h}) \in \mathrm{epi}(\mathrm{LCU}^{\Sigma})$ and thereby reach the same bound on the summed start-up costs as in $\mathrm{epi}(\mathrm{LCU}^{\Sigma})$.

Still, Algorithm 3.3.1 does not separate $p \in P^{\mathrm{temp}}$ from $\hat{P}^{\mathrm{temp}}$ as there exist points $p \in \hat{P}^{\mathrm{temp}} \setminus P^{\mathrm{temp}}$ with $\pi^{\Sigma}(p) \in \mathrm{epi}(\mathrm{LCU}^{\Sigma})$. In the remainder of this subsection, we demonstrate that such points are not heating-minimal. In other words, we prove that for a heating-minimal solution p, Algorithm 3.3.1 either finds a violated RTI or confirms that p fulfills all RTIs. Moreover, the proof highlights the connection between the RTIs and Cartesian trees.

A Cartesian tree of a vector $v \in [0, 1]^T$ is given in (2.2.11) as a rank-labeled binary tree on nodes $[T]$ such that for each node $t \in [T]$ and its parent $p(t)$, it holds that

$$v^t \leq v^{p(t)}.$$

Such a tree has two important properties:

- since for each $t \in [T]$ with $l(t) < t-1$, the node t is a descendant of the node $t - l(t) - 1$ (cf. Corollary 2.24),

$$\forall\, t \in [T] \text{ with } l(t) < t-1 : \qquad v^t \leq v^{t-l(t)-1}, \tag{3.3.9}$$

- and since the nodes $[t-l(t) \,..\, t-1]$ lie in the left subtree of t (cf. (2.2.5)),

$$\forall\, t \in [T], t' \in [t-l(t) \,..\, t-1] : \qquad v^t \geq v^{t'}. \tag{3.3.10}$$

As the next lemma demonstrates, property (3.3.9) implies that $\mathrm{RTI}(t, \tilde{l})$ with $\tilde{l} > l(t)$ is dominated by the combination of $\mathrm{RTI}(t, l(t))$ and $\mathrm{RTI}(t - l(t) - 1, \tilde{l} - l(t) - 1)$.

Lemma 3.31 *For each $p \in \hat{P}^{\mathrm{temp}}$ with Cartesian tree B, $t \in [T]$, $\tilde{l} \in [l(t)+1 \,..\, t-1]$, the combination of RTI$(t, l(t))$ and RTI$(t - l, \tilde{l} - l(t) - 1)$ dominate RTI$(t, \tilde{l})$.*

Proof. First, suppose $l(t) < \tilde{l} < t-1$. We start from formulation (3.3.2) of $\mathrm{RTI}(t, l(t))$,

$$\mathrm{temp}^t - v^t \geq \mathrm{e}^{-\lambda \mathrm{OL}^t(l(t))} (\underbrace{\mathrm{temp}^{t-l(t)} - v^t}_{(*)}). \tag{3.3.11}$$

Due to the temperature development equation (3.2.7) and (3.3.9), the term $(*)$ on its right-hand side can be bounded by

$$\begin{aligned} \mathrm{temp}^{t-l(t)} - v^t &= \mathrm{e}^{-\lambda \mathrm{L}^{t-l(t)}} \mathrm{temp}^{t-l(t)-1} + (1 - \mathrm{e}^{-\lambda \mathrm{L}^{t-l(t)-1}}) v^{t-l(t)-1} - v^t \\ &= \mathrm{e}^{-\lambda \mathrm{L}^{t-l(t)-1}} (\mathrm{temp}^{t-l(t)-1} - v^t) + (1 - \mathrm{e}^{-\lambda \mathrm{L}^{t-l(t)-1}}) (\underbrace{v^{t-l(t)-1} - v^t}_{\geq 0}) \\ &\geq \mathrm{e}^{-\lambda \mathrm{L}^{t-l(t)-1}} (\underbrace{\mathrm{temp}^{t-l(t)-1} - v^t}_{(**)}). \end{aligned} \tag{3.3.12}$$

Using $\mathrm{RTI}(t - l(t) - 1, \tilde{l} - l(t) - 1)$, the term $(**)$ in (3.3.12) is bounded by

$$\begin{aligned} \mathrm{temp}^{t-l(t)-1} - v^t &\geq v^{t-l(t)-1} + \mathrm{e}^{-\lambda \mathrm{OL}^{t-l(t)-1}(\tilde{l}-l(t)-1)} (\mathrm{temp}^{t-\tilde{l}} - v^{t-l(t)-1}) - v^t, \\ &= \mathrm{e}^{-\lambda \mathrm{OL}^{t-l(t)-1}(\tilde{l}-l(t)-1)} (\mathrm{temp}^{t-\tilde{l}} - v^t) \\ &\qquad + (1 - \mathrm{e}^{-\lambda \mathrm{OL}^{t-l(t)-1}(\tilde{l}-l(t)-1)}) (\underbrace{v^{t-l(t)-1} - v^t}_{\geq 0}) \\ &\geq \mathrm{e}^{-\lambda \mathrm{OL}^{t-l(t)-1}(\tilde{l}-l(t)-1)} (\mathrm{temp}^{t-\tilde{l}} - v^t). \end{aligned} \tag{3.3.13}$$

Combining (3.3.11), (3.3.12), and (3.3.13) results in

$$\mathrm{temp}^t - v^t \geq \mathrm{e}^{-\lambda \mathrm{OL}^t(l(t))} \mathrm{e}^{-\lambda \mathrm{L}^{t-l(t)-1}} \mathrm{e}^{-\lambda \mathrm{OL}^{t-l(t)-1}(\tilde{l}-l(t)-1)} (\mathrm{temp}^{t-\tilde{l}} - v^t),$$

which equals $\mathrm{RTI}(t,\tilde{l})$, as can be seen by aggregating the heat loss factors to

$$\begin{aligned}\mathrm{temp}^t - v^t &\geq \mathrm{e}^{-\lambda\left(\sum_{t'\in[t-l(t)..t-1]}\mathrm{L}^{t'} \,+\, \mathrm{L}^{l(t)-1} \,+\, \sum_{t'\in[t-\tilde{l}..t-l(t)-2]}\mathrm{L}^{t'}\right)}(\mathrm{temp}^{t-\tilde{l}} - v^t)\\ &= \mathrm{e}^{-\lambda\sum_{t'\in[t-\tilde{l}..t-1]}\mathrm{L}^{t'}}(\mathrm{temp}^{t-\tilde{l}} - v^t) = \mathrm{e}^{-\lambda\mathrm{OL}^t(t-\tilde{l})}(\mathrm{temp}^{t-\tilde{l}} - v^t).\end{aligned}$$

Finally, for $\tilde{l} = t-1$ the statement is proved analogously by replacing $\mathrm{temp}^{t-\tilde{l}}$ with 1.□

The property (3.3.10) is used in the following to show that $\mathrm{RTI}(t,\tilde{l})$ with $\tilde{l} < l(t)$ is dominated by $\mathrm{RTI}(t,l(t))$.

Lemma 3.32 *For each heating-minimal solution* $(v,\mathrm{y},\mathrm{z},\mathrm{temp},\mathrm{h})$ *in* $\hat{P}^{\mathrm{temp}}$ *with Cartesian tree* B *for* v*,* $t\in[T]$*,* $RTI(t,\tilde{l}+1)$ *dominates* $RTI(t,\tilde{l})$*.*

Proof. If $\mathrm{h}^{t-\tilde{l}-1} = 0$, then

$$\mathrm{temp}^{t-\tilde{l}} = \mathrm{e}^{-\lambda\mathrm{L}^{t-\tilde{l}-1}}\mathrm{temp}^{t-\tilde{l}-1} + (1-\mathrm{e}^{-\lambda\mathrm{L}^{t-\tilde{l}-1}})v^{t-\tilde{l}-1}.$$

Using the above equation and (3.3.10), $\mathrm{RTI}(t,\tilde{l}+1)$ implies

$$\begin{aligned}\mathrm{temp}^t &\geq \mathrm{e}^{-\lambda\mathrm{OL}^t(\tilde{l}+1)}\mathrm{temp}^{t-\tilde{l}-1} + (1-\mathrm{e}^{-\lambda\mathrm{OL}^t(\tilde{l}+1)})v^t\\ &= \mathrm{e}^{-\lambda\mathrm{OL}^t(\tilde{l})}\mathrm{temp}^{t-\tilde{l}} + (1-\mathrm{e}^{-\lambda\mathrm{OL}^t(\tilde{l})})v^t + \mathrm{e}^{-\lambda\mathrm{OL}^t(\tilde{l})}(1-\mathrm{e}^{-\lambda\mathrm{L}^{t-\tilde{l}-1}})(\underbrace{v^t - v^{t-\tilde{l}-1}}_{\geq 0})\\ &\geq \mathrm{e}^{-\lambda\mathrm{OL}^t(\tilde{l})}\mathrm{temp}^{t-\tilde{l}} + (1-\mathrm{e}^{-\lambda\mathrm{OL}^t(\tilde{l})})v^t,\end{aligned}$$

and so $\mathrm{RTI}(t,\tilde{l})$ is dominated by $\mathrm{RTI}(t,\tilde{l}+1)$.

If $\mathrm{h}^{t-\tilde{l}-1} > 0$, then by Proposition 3.8 $\mathrm{temp}^{t-\tilde{l}} = v^{t-\tilde{l}}$. Hence, (3.3.10) implies

$$\mathrm{temp}^t - v^{t-\tilde{l}} \geq v^t - v^{t-\tilde{l}} \geq 0 = \mathrm{e}^{-\lambda\mathrm{OL}^t(\tilde{l})}(\mathrm{temp}^{t-\tilde{l}} - v^{t-\tilde{l}}).$$ □

By combining the two previous results, we strengthen Corollary 3.28 for heating-minimal solutions:

Theorem 3.33 *For each heating-minimal solution* p *of* $\hat{P}^{\mathrm{temp}}$*, Algorithm 3.3.1 separates* p *from* P^{temp} *in a running time of* $\mathcal{O}(T)$*.*

Proof. By Corollary 3.28, for each point $p = (v,\mathrm{y},\mathrm{z},\mathrm{temp},\mathrm{h}) \in \hat{P}^{\mathrm{temp}}$ Algorithm 3.3.1 either confirms that $\pi^{\Sigma}(p) \in \mathrm{epi}(\mathrm{LCU}^{\Sigma})$ or stops with a violated RTI in $\mathcal{O}(T)$.

If the algorithm stops with a violated RTI, then $p \notin P^{\mathrm{temp}}$. Else, p fulfills $\mathrm{RTI}(t,l(t))$ for a Cartesian tree B for v and each $t\in[T]$. We prove that the remaining RTIs are fulfilled as well by induction over $t\in[T]$.

For $t = 1$, the single inequality $\mathrm{RTI}(1,0)$ is trivially fulfilled.

For $t \geq 2$, Lemma 3.31 can be applied for $l = l(t)+1$ and each $\tilde{l} \in [l \,..\, t-1]$, showing that $\mathrm{RTI}(t, \tilde{l})$ is fulfilled for $\tilde{l} \in [l(t)+1 \,..\, t-1]$.

Assume there exists $l \in [l(t)-1]$ such that $\mathrm{RTI}(t, l)$ is violated. If $\tilde{l}$ denotes the maximal such l, then $\mathrm{RTI}(t, \tilde{l}+1)$ is fulfilled. Hence, Lemma 3.32 proves that $\mathrm{RTI}(t, \tilde{l})$ is fulfilled, a contradiction. Thus, $\mathrm{RTI}(t, l)$ is fulfilled for all $l \in [t-1]$.

In conclusion, if Algorithm 3.3.1 confirms that $\pi^{\Sigma}(p) \in \mathrm{epi}(\mathrm{LCU}^{\Sigma})$, then p fulfills all RTI and hence $p \in P^{\mathrm{temp}}$. □

3.4 Generalization of Temperature Development

In Section 3.2, we derived the temperature formulation $\hat{P}^{\mathrm{temp}}$ for the exponential start-up cost function

$$\mathrm{CU}(L) = \begin{cases} \mathrm{CU}^{\mathrm{var}}(1 - f(L)) + \mathrm{CU}^{\mathrm{fixed}} & \text{if } L > 0, \\ 0 & \text{if } L = 0, \end{cases}$$

where $f(L) = \mathrm{e}^{-\lambda L}$. This section generalizes the temperature model to functions f satisfying three conditions:

- f is strictly decreasing and continuous,
- $f(0) = 1$ and $f(L) \to 0$ for $L \to \infty$ to retain the meaning of the fixed costs $\mathrm{CU}^{\mathrm{fixed}}$ and variable costs $\mathrm{CU}^{\mathrm{var}}$, and
- f is strictly convex for keeping premature heating inefficient.

As we will show, $f = \mathrm{e}^{-\lambda L}$ is unique with the property of resulting in a affine linear inequalities in Proposition 3.35, and that f may be modeled using convex inequalities iff $f(f^{-1}(x) + y)$ is concave with respect to x for $y > 0$.

The fundamental idea of this chapter is to explicitly model the variable temp by the value $f(\mathrm{L})$,

$$\forall\, t \in [T]: \quad \mathrm{temp}_v^t := \begin{cases} 1 & \text{if } v^t = 1, \\ f(\mathrm{OL}^t(\mathrm{ol}^t(v))) & \text{else,} \end{cases}$$

where $\mathrm{OL}^t(\mathrm{ol}^t(v))$ equals the downtime prior to period t in the operational schedule v.

To model the connection between two consecutive temperatures, we replace the exponential case of the temperature development equations (3.2.6),(3.2.7) by their possibly non-linear generalizations

$$\mathrm{temp}^1 = f(\mathrm{PDT}) + \mathrm{h}^t, \tag{3.4.1}$$

$$\forall\, t \in [T-1]: \qquad \mathrm{temp}^{t+1} \leq f(f^{-1}(\mathrm{temp}^t) + \mathrm{L}^t) + (1 - f(\mathrm{L}^t))v^t + \mathrm{h}^t. \tag{3.4.2}$$

The generalized temperature set $\hat{S}^{\text{temp}}$ is defined similarly as $\hat{P}^{\text{temp}}$ with (3.2.6), (3.2.7) replaced by (3.4.1), (3.4.2):

$$\hat{S}^{\text{temp}} := \Big\{(v, \mathrm{y}, \mathrm{z}, \text{temp}, \mathrm{h}) \in \mathbb{R}^{5T} \text{ fulfilling (3.2.1)–(3.2.5),(3.2.9),(3.4.1),(3.4.2)}\Big\}. \tag{3.4.3}$$

Section 3.2 proves in three steps that $\hat{P}^{\text{temp}}$ models the start-up costs, showing that

- $\hat{P}^{\text{temp}}$ is valid for V^{temp} in Lemma 3.5,
- optimal solutions in $\hat{P}^{\text{temp}}$ for heating-minimal objective function vectors do not heat prematurely in Proposition 3.8, and
- optimal solutions in $\hat{P}^{\text{temp}}$ for heating-minimal objective function vectors model correct start-up costs in Lemma 3.10.

Assuming that premature heating is prevented, Lemma 3.5 and Lemma 3.10 are straightforward to generalize to $\hat{S}^{\text{temp}}$ by noting that

- if $v^t = 1$, then using (3.2.5) $\text{temp}^t = 1$ and therefore

$$\text{temp}_v^{t+1} = 1 = f(\underbrace{f^{-1}(\text{temp}_v^t)}_{=f^{-1}(1)=0} + \mathrm{L}^t) + (1 - f(\mathrm{L}^t)) \underbrace{v^t}_{=1} + \underbrace{\mathrm{h}_v^t}_{=0},$$

- if $v^t = 0$ and $v^{t+1} = 0$, then

$$\text{temp}_v^{t+1} = f(\mathrm{OL}^{t+1}(\mathrm{ol}^{t+1}(v))) = f(\underbrace{f^{-1}(\text{temp}_v^t)}_{=\mathrm{OL}^t(\mathrm{ol}^t(v))} + \mathrm{L}^t) + (1 - f(\mathrm{L}^t)) \underbrace{v^t}_{=0} + \underbrace{\mathrm{h}_v^t}_{=0},$$

- and if $v^t = 0$ and $v^{t+1} = 1$, then

$$\text{temp}_v^{t+1} = 1 = f(\underbrace{f^{-1}(\text{temp}_v^t)}_{=\mathrm{OL}^t(\mathrm{ol}^t(v))} + \mathrm{L}^t) + (1 - f(\mathrm{L}^t)) \underbrace{v^t}_{=0} + \underbrace{\mathrm{h}_v^t}_{=1-f(\mathrm{OL}^{t+1}(\mathrm{ol}^{t+1}(v)))}.$$

Regarding the premature heating, extending Proposition 3.8 to $\hat{S}^{\text{temp}}$ highlights why f must be strictly convex. Due to the complexity of generalizing the definition of heating-minimal objective function vectors in Definition 3.7, we restrict ourselves to the typically used vector $a := (1, \ldots, 1)$.

Proposition 3.34 *For each $v \in [0,1]$, each $(v, \mathrm{y}, \mathrm{z}, \text{temp}, \mathrm{h}) \in \hat{S}^{\text{temp}}$ which minimizes*

$$\sum_{t \in [T]} \mathrm{h}^{t-1}$$

in $\hat{S}^{\text{temp}}$ fulfills

$$\forall\, t \in [T]: \quad \mathrm{h}^{t-1} > 0 \;\Rightarrow\; \text{temp}^t = v^t.$$

Proof. The argumentation of this proof is analogous to the original proof of Proposition 3.8. Assume there exists $t \in [T]$ such that

$$\mathrm{h}^{t-1} > 0 \quad \text{and} \quad \mathrm{temp}^t > v^t.$$

Then, we can postpone an amount $\Delta := \max\{\mathrm{h}^{t-1}, \mathrm{temp}^t - v^t\} > 0$ of heating in period $t-1$ to period t. To keep temp^{t+1} constant, the heating in period t must increase by

$$\Gamma := \underbrace{f(f^{-1}(\mathrm{temp}^t) + \mathrm{L}^t)}_{=\mathrm{temp}^{t+1}} - f(f^{-1}(\underbrace{\mathrm{temp}^t - \Delta}_{\widetilde{\mathrm{temp}}^t}) + \mathrm{L}^t).$$

The resulting point $(v, \mathrm{y}, \mathrm{z}, \widetilde{\mathrm{temp}}, \tilde{\mathrm{h}})$ with

$$\widetilde{\mathrm{temp}}^{t'} = \begin{cases} \mathrm{temp}^{t'} - \Delta & \text{if } t' = t, \\ \mathrm{temp}^{t'} & \text{else,} \end{cases} \qquad \tilde{\mathrm{h}}^{t'} = \begin{cases} \mathrm{h}^{t'} - \Delta & \text{if } t' = t-1, \\ \mathrm{h}^{t'} + \Gamma & \text{if } t' = t, \\ \mathrm{h}^{t'} & \text{else,} \end{cases}$$

lies in $\hat{S}^{\mathrm{temp}}$.

Since f is strictly decreasing and $\Delta > 0$, $f^{-1}(\mathrm{temp}^t) < f^{-1}(\mathrm{temp}^t - \Delta)$. Thus, as f is strictly decreasing and strictly convex, Γ can be bounded by

$$\Gamma < f(f^{-1}(\mathrm{temp}^t)) - f(f^{-1}(\mathrm{temp}^t - \Delta)) = \Delta.$$

Thereby, the amount of heating in $(v, \mathrm{y}, \mathrm{z}, \widetilde{\mathrm{temp}}, \tilde{\mathrm{h}}) \in \hat{S}^{\mathrm{temp}}$ is

$$\sum_{t\in[T]} \tilde{\mathrm{h}}^{t-1} = \sum_{t\in[T]} \mathrm{h}^t + \Gamma - \Delta < \sum_{t\in[T]} \mathrm{h}^{t-1},$$

a contradiction to the minimality of the total heating in $(v, \mathrm{y}, \mathrm{z}, \mathrm{temp}, \mathrm{h})$. □

Having proved that $\hat{S}^{\mathrm{temp}}$ correctly models the start-up costs, we consider the shape of $\hat{S}^{\mathrm{temp}}$. The only inequality in the representation (3.4.3) of $\hat{S}^{\mathrm{temp}}$ which may not be affine linear is (3.4.2), due to the term $f(f^{-1}(\mathrm{temp}) + \mathrm{L}^t)$.

This term determines whether $\hat{S}^{\mathrm{temp}}$ is a polyhedron, a convex set, or a non-convex set:

- inequality (3.4.2) is affine linear iff $f(f^{-1}(\mathrm{temp}) + L)$ is affine linear in temp, and
- inequality (3.4.2) is convex iff $f(f^{-1}(\mathrm{temp}) + L)$ is concave in temp.

We continue by showing that $f(f^{-1}(\mathrm{temp}) + L)$ is affine linear iff $f = \mathrm{e}^{-\lambda L}$, and conclude this section by giving the exemplary class of functions $f = (1+x)^{-p}$ such that $f(f^{-1}(\mathrm{temp}) + L)$ is concave.

Proposition 3.35 *$\hat{S}^{\mathrm{temp}}$ is a polyhedron iff $f(L) = e^{-\lambda L}$ for some $\lambda > 0$.*

Proof. As noted, $\hat{S}^{\text{temp}}$ is a polyhedron iff $f(f^{-1}(\text{temp}) + L)$ is affine linear in temp. If $f(L) = \mathrm{e}^{-\lambda L}$, then

$$f(f^{-1}(\text{temp}) + L) = f\left(\frac{\ln(\text{temp})}{-\lambda} + L\right) = \mathrm{e}^{\ln(\text{temp})-\lambda L} = \mathrm{e}^{-\lambda L}\text{temp},$$

and hence $f(f^{-1}(\text{temp}) + L)$ is linear in temp. On the other hand, assume that there exist functions $a(L)$ and $b(L)$ with

$$f(f^{-1}(\text{temp}) + L) = a(L)\text{temp} + b(L).$$

Since f is strictly decreasing and converges to 0, we have $0 < f(f^{-1}(\text{temp})+L) < \text{temp}$, and so

$$(1 - a(L))\text{temp} > b(L).$$

Hence, considering temp $\to 0$, $b(L) \geq 0$. Summarizing, $b(L)$ is bounded as

$$(1 - a(L))\text{temp} > b(L) \geq 0,$$

which implies $a(L) < 1$ and, again by considering temp $\to 0$, $b(L) = 0$.

Furthermore, it holds for all $L', L'' \in \mathbb{R}_{>0}$ that

$$\begin{aligned} a(L'')a(L')\text{temp} &= a(L'')f(f^{-1}(\text{temp}) + L') = f(f^{-1}(f(f^{-1}(\text{temp}) + L')) + L'') \\ &= f(f^{-1}(\text{temp}) + L' + L'') = a(L' + L'')\text{temp}. \end{aligned}$$

By [Rud76, p. 197, ex. 6], the only continuous function with $a(L')a(L'') = a(L' + L'')$ is $a(L) = \mathrm{e}^{-\lambda L}$ with $\lambda \in \mathbb{R}$, and since $a(L) < 1$, we have $\lambda > 0$. Finally, $f(L)$ can be derived as

$$\forall\, L > 0 : \qquad f(L) = f(\underbrace{f^{-1}(1)}_{=0} + L) = a(L) = \mathrm{e}^{-\lambda L}. \qquad \square$$

Proposition 3.36 *If $f(x) = (1 + x)^{-p}$ with $p > 0$, then $\hat{S}^{\text{temp}}$ is convex.*

Proof. First note that $f(x) = (1 + x)^{-p}$ is admissible, since f is continuous on $(0, 1)$, strictly decreasing, strictly convex, and fulfills $f(0) = 1$ and $f(x) \to 0$ as $x \to \infty$.

The term $f(f^{-1}(\text{temp}) + L)$ evaluates to

$$f(f^{-1}(\text{temp}) + L) = \frac{\text{temp}}{(\sqrt[p]{\text{temp}}L + 1)^p},$$

and its second partial derivative with respect to temp is

$$\frac{\delta^2 f(f^{-1}(\text{temp}) + L)}{\delta \text{temp}^2} = -\frac{p+1}{p}\frac{\sqrt[p]{\text{temp}}L}{\text{temp}(\sqrt[p]{\text{temp}}L + 1)^{p+1}} < 0.$$

Thus, $f(f^{-1}(\text{temp}) + L)$ is concave in temp and $\hat{S}^{\text{temp}}$ is convex. $\square$

Chapter 4

Start-up Types

Start-up type models assign a type s to each start-up using binary variables δ_s^t. This approach was introduced in [Muc66] and enhanced in [SBB10], and is capable of modeling the complete start-up process, including e. g. costs, synchronization times, soak times, and power trajectories (see [SBB10]). Even when only modeling time-dependent start-up costs, the computational experiments in [MELR13b] demonstrate that start-up types are beneficial, since the solution times are shortened considerably by tightening the linear relaxation compared to the start-up cost models presented e. g. [NR00; CA06] and in Section 2.1.

The present chapter provides an interpretation of the start-up types as optimal flows in a special network flow model. The flow conservation inequalities of this network significantly tighten the inequalities in [Muc66; SBB10], yielding a smaller integrality gap and better computational performance. In particular, the resulting model P^δ is an extended formulation of the epigraph of all start-up costs conv(epi(DCU)) (see Section 2.3).

Moreover, Theorem 4.20 uses the interpretation of the start-up types as flows to compare the tightness of the models P_{ex}^t (see [CA06]), epi(LCU^t) (see Section 2.1), P_{ex}^δ (see [Muc66]), and the newly introduced P^δ.

As opposed to the extended formulation P^{temp} of conv(epi(DCU^Σ)) presented in Chapter 3, start-up type models allow any increasing start-up cost function and an arbitrary weighting of the start-up costs in each period. Furthermore, while we only consider start-up costs, the start-up production can be modeled based on the start-up types as shown in [SBB10] (cf. Subsection 1.3.3).

In the following, the model is introduced in five steps:

- Section 4.1 presents the interpretation of a special case of start-up types as optimal solutions in a network flow model.
- Section 4.2 shows that the polyhedron corresponding to this network flow model is an extended formulation of conv(epi(DCU)) with $\mathcal{O}(T^2)$ variables and $\mathcal{O}(T)$ inequalities.
- For piece-wise constant start-up cost functions with S steps (e. g. approximations as in Subsection 2.1.6), Section 4.3 groups the flows into S start-up types, thereby

reducing the number of variables to $\mathcal{O}(ST)$ at the cost of $\mathcal{O}(T^2)$ inequalities. Of these, $\mathcal{O}(T)$ inequalities suffice to model the types correctly for $v \in \{0,1\}^T$, and the remaining inequalities can be separated in $\mathcal{O}(T^2)$.

- Section 4.4 notes a relationship between the types and the widely-used start-up/shutdown indicators, and substitutes some types with indicators.
- Finally, Section 4.5 exploits the flow interpretation to compare our model with the models from [Muc66; SBB10; CA06] and the epigraphs in Sections 2.1 and 2.3.

4.1 The Network Flow Interpretation

We consider the network with nodes

$$\begin{aligned} V &:= \{0, 1_-, \dots, T_-, 1_+, \dots, T_+\} \quad \text{and edges} \\ E &:= \Big\{(t_-, t_+) \,\Big|\, t \in [T]\Big\} \;\cup\; \Big\{(t_+, t'_-) \,\Big|\, t \in [T], t' \in [0 \,..\, t-1]\Big\}. \end{aligned} \tag{4.1.1}$$

In this network, each node pair t_-, t_+ corresponds to a period $t \in [T]$. Their connecting edge (t_-, t_+) represents the operational state v^t and thus has capacity 1. As we see in the following, the edges (t_+, t'_-) between nodes from different periods indicate whether the unit was shut down in period t' and started up in period t. The capacity of these edges is unlimited, but the flow on them is implicitly bounded by 1 due to the capacities of the edges (t_-, t_+) and the flow conservation in nodes t_+.

To represent shutdowns and start-ups outside the modeled time range, we use the sink node 0, and model the nodes t_- as sources, i. e. with outgoing flow at least as high as incoming flow.

Denoting the flow on edge (t_-, t_+) as v^t and the flow on edge (t_+, t'_-) as $f^t_{t-t'-1}$, a flow can be represented by (v, f) (cf. Fig. 4.1).

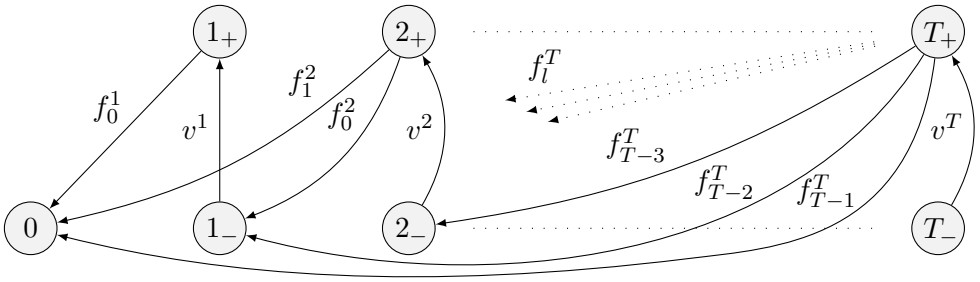


Figure 4.1: Network flow interpretation of the start-up types.

Using the number of offline periods $\mathrm{ol}^t(v)$ preceding period t, we define the canonical flow of an operational schedule $v \in \{0,1\}^T$ as $(v, \hat{f}(v))$ (see Fig. 4.2). For each $t \in [T], l \in [0\,..\,t-1]$, the flow on edge $(t_+, (t-l-1)_-)$ is given by

$$\hat{f}_l^t(v) := \begin{cases} 1 & \text{if } v^t = 1 \text{ and } l = \mathrm{ol}^t(v), \\ 0 & \text{else.} \end{cases} \tag{4.1.2}$$

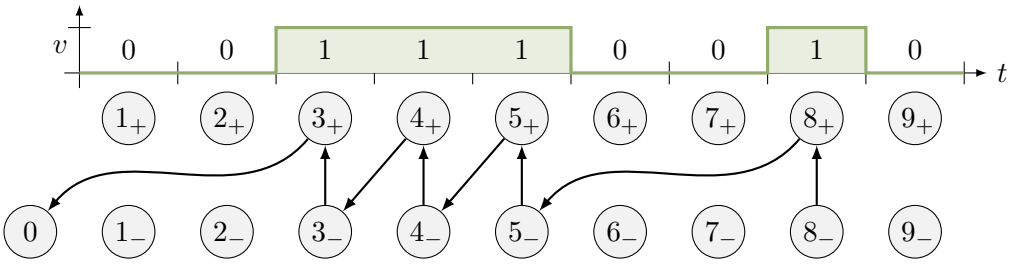


Figure 4.2: Exemplary canonical flow $(v, \hat{f}(v))$ of an operational schedule $v \in \{0,1\}^T$ with $T = 9$. The shown edges have flow 1 and all hidden edges have flow 0.
Note how each start-up is indicated by an edge (t_+, t'_-) with $t' < t-1$ having flow 1: The start-up is performed in period t and the length of the preceding downtime is given by $t - t' - 1$.

In the following, we characterize the feasible flows within the described network, and bring them into relation with the start-up costs. Firstly, the canonical flow of an operational schedule is always feasible:

Proposition 4.1 *For each $v \in \{0,1\}^T$, $(v, \hat{f}(v))$ is a feasible flow.*

Proof. The edge capacities are fulfilled since $0 \leq v^t \leq 1$. Since the sink 0 has only incoming edges, it suffices to check the flow conservation of the nodes t_- and t_+.

For each $t \in [T]$, it holds by definition of $\hat{f}(v)$ that

$$\sum_{l=0}^{t-1} \hat{f}_l^t(v) = \hat{f}^t_{\mathrm{ol}^t(v)}(v) = v^t,$$

fulfilling the flow conservation at nodes t_+. For the nodes t_-, a case distinction is necessary:

- If $v^t = 0$, then for each $t' \in [t+1\,..\,T]$, $\mathrm{ol}^{t'}(v) \neq t' - t - 1$, and thus

$$\sum_{t'=t+1}^{T} \hat{f}^{t'}_{t'-t-1}(v) = 0 = v^t.$$

- For $v^t = 1$, suppose that there exist $t_1 < t_2 \in [T]$ with $v^{t_1} = 1$, $t = t_1 - \mathrm{ol}^{t_1}(v) - 1$, and $t = t_2 - \mathrm{ol}^{t_2}(v) - 1$. Then by definition $\mathrm{ol}^{t_2}(v) \leq t_2 - t_1 - 1 < t_2 - t - 1$, a contradiction. So, there exists at most one $t' \in [T]$ with $t' - \mathrm{ol}^{t'}(v) - 1 = t$, i. e.

$$\sum_{t'=t+1}^{T} \hat{f}^{t'}_{t'-t-1}(v) \leq 1 = v^t. \qquad \square$$

The capacity of 1 on the edges (t_-, t_+) constrains v to $[0,1]^T$, and through the flow conservation on the nodes t_+ also limits f^t_l:

Proposition 4.2 *Each feasible flow (v, f) lies in $[0,1]^{T+T(T+1)/2}$.*

Proof. The number of components of f is $\sum_{t\in[T]} \sum_{l=0}^{t-1} 1 = T(T+1)/2$. For each $t \in [T]$,

$$\max_{l\in[0\,..\,t-1]} f^t_l \leq \sum_{l=0}^{t-1} f^t_l = v^t \leq 1. \qquad \square$$

Figure 4.2 suggests that non-zero flows f^t_l with $l > 0$ indicate offline periods, and may therefore be used to derive the start-up costs incurred by v. We define the projection π^f as

$$\pi^f : \quad [0,1]^T \times [0,1]^{T(T+1)/2} \to [0,1]^T \times \mathbb{R}^T \qquad \text{with} \quad \mathrm{cu}^t := \sum_{l=0}^{t-1} \mathrm{CU}^t(l) f^t_l.$$
$$(v, f) \mapsto (v, \mathrm{cu})$$

Indeed, it follows directly from the definition in (4.1.2) that the costs associated to $(v, \hat{f}(v))$ equal the incurred start-up costs $\mathrm{DCU}(v)$, i. e.

$$\forall\, v \in \{0,1\}^T : \qquad \pi^f(v, \hat{f}(v)) = (v, \mathrm{DCU}(v)). \tag{4.1.3}$$

As we show next, the network is designed such that $(v, \hat{f}(v))$ not only results in correct start-up costs, but moreover minimizes the costs among all (v, f).

Proposition 4.3 *For each $v \in \{0,1\}^T$ and each feasible flow (v, f),*

$$(v, \mathrm{DCU}(v)) = \pi^f(v, \hat{f}(v)) \leq \pi^f(v, f),$$

and if CU *is strictly increasing and $\hat{f}(v) \neq f$, then $\pi^f(v, \hat{f}(v)) < \pi^f(v, f)$.*

Proof. For each $t \in [T]$, the definition of $\mathrm{ol}^t(v)$ implies $v^{t-l} = 0$ for all $l \in [\mathrm{ol}^t(v) - 1]$. Using the flow conservation for the nodes t'_- with $t' \in [t - \mathrm{ol}^t(v) + 1\,..\,t-1]$, we obtain

$$\forall\, l \in [0\,..\,\mathrm{ol}^t(v) - 1] : \qquad f^t_l = 0.$$

Since CU is increasing, $\pi^f(v,f)$ is bounded by

$$\pi^f(v,f) = \sum_{l=0}^{t-1} \mathrm{CU}^t(l) f_l^t = \sum_{l=\mathrm{ol}^t(v)}^{t-1} \mathrm{CU}^t(l) f_l^t \overset{*}{\geq} \mathrm{CU}^t(\mathrm{ol}^t(v)) \sum_{l=\mathrm{ol}^t(v)}^{t-1} f_l^t = \mathrm{CU}^t(\mathrm{ol}^t(v)) \sum_{l=0}^{t-1} f_l^t,$$

which by the flow conservation in nodes t_+ and (4.1.2) implies

$$\pi^f(v,f) \geq \mathrm{CU}^t(\mathrm{ol}^t(v)) \underbrace{\sum_{l=0}^{t-1} f_l^t}_{=v^t} = \mathrm{CU}^t(\mathrm{ol}^t(v)) \underbrace{\sum_{l=0}^{t-1} \hat{f}_l^t(v)}_{=v^t} = \sum_{l=0}^{t-1} \mathrm{CU}^t(l) \hat{f}_l^t(v) = \pi^f(v, \hat{f}(v)).$$

If CU is strictly increasing and $f \neq \hat{f}(v)$, then the inequality above marked by "$*$" is strict for some $t \in [T]$. □

4.2 The Start-up Flow Polyhedron

This section shows that the canonical linear model of the proposed network flow problem is an extended formulation of conv(epi(DCU)).

The capacity 1 of the edges (t_-, t_+) and the unboundedness of the edges (t_+, t'_-) is expressed as

$$\forall\, t \in [T]: \qquad 0 \leq v^t \leq 1, \tag{4.2.1}$$

$$\forall\, t \in [T], l \in [0\,..\,t-1]: \qquad f_l^t \geq 0. \tag{4.2.2}$$

The flow conservation for the nodes t_+ is modeled by

$$\forall\, t \in [T]: \qquad \sum_{l=0}^{t-1} f_l^t = v^t. \tag{4.2.3}$$

For the nodes t_-, which are sources, the outgoing flow must be at least as high as the incoming flow,

$$\forall\, t \in [T-1]: \qquad \sum_{t'=t+1}^{T} f_{t'-t-1}^{t'} \leq v^t. \tag{4.2.4}$$

We call the resulting polyhedron

$$P^f := \left\{(v,f) \in \mathbb{R}^{T+T(T+1)/2} \text{ fulfilling inequalities (4.2.1)-(4.2.4)}\right\} \tag{4.2.5}$$

the start-up flow polyhedron.

As the inequalities (4.2.1)-(4.2.4) model a flow problem on a directed graph with edges of capacity 1 or unlimited capacity, an $\mathcal{H}$-representation of P^f can be given in matrix form as

$$P^f = \left\{ (v,f) \in \mathbb{R}^{T+T(T+1)/2} \;\middle|\; \begin{pmatrix} C_+ \\ -C_+ \\ C_- \\ I_T\ 0 \\ -I \end{pmatrix} (v,f) \leq \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \right\},$$

using the oriented incidence matrices C_+ and C_- of the graph in (4.1.1) restricted to the nodes t_+ and t_-, respectively, the identity matrix I_T of size T, the identity matrix I of size $T + {}^{T(T+1)}/_2$, and the zero matrix 0 of appropriate size. The matrices C_- and C_+ are shown in Fig. 4.3 (using $T = 5$).

$$C_- = \begin{array}{c} \begin{array}{ccccc|ccccccccccccccc} v^1 & v^2 & v^3 & v^4 & v^5 & f_0^1 & f_0^2 & f_1^2 & f_0^3 & f_1^3 & f_2^3 & f_0^4 & f_1^4 & f_2^4 & f_3^4 & f_0^5 & f_1^5 & f_2^5 & f_3^5 & f_4^5 \end{array} \\ \left(\begin{array}{ccccc|ccccccccccccccc} -1 & & & & & & 1 & & & 1 & & & & 1 & & & & & 1 & \\ & -1 & & & & & & & 1 & & & 1 & & & & & 1 & & \\ & & -1 & & & & & & & & & 1 & & & & 1 & & & & \\ & & & -1 & & & & & & & & & & & & 1 & & & & \\ & & & & -1 & & & & & & & & & & & & & & & \end{array} \right) \begin{array}{c} 1_- \\ 2_- \\ 3_- \\ 4_- \\ 5_- \end{array} \end{array}$$

$$C_+ = \begin{array}{c} \left(\begin{array}{ccccc|ccccccccccccccc} 1 & & & & -1 & & & & & & & & & & & & & & & \\ & 1 & & & & -1 & -1 & & & & & & & & & & & & & \\ & & 1 & & & & & & -1 & -1 & -1 & & & & & & & & & \\ & & & 1 & & & & & & & & -1 & -1 & -1 & -1 & & & & & \\ & & & & 1 & & & & & & & & & & & -1 & -1 & -1 & -1 & -1 \end{array} \right) \begin{array}{c} 1_+ \\ 2_+ \\ 3_+ \\ 4_+ \\ 5_+ \end{array} \end{array}$$

Figure 4.3: Incidence matrices C_- and C_+ of the nodes t_- and t_+ in the network flow problem underlying P^f. The edges are identified by their associated flow.

The matrix defining P^f is totally unimodular since it is composed of incidence and identity matrices (see [HK56]). Hence, as the right hand side is integral, each vertex of P^f is integral as well.

As P^f is contained in the cube $[0,1]^{T+T(T+1)/2}$ (see Proposition 4.2), each vertex of P^f is binary and each binary vector in P^f is extremal.

Proposition 4.4 *The set of vertices of P^f is $V^f := P^f \cap \{0,1\}^{T+T(T+1)/2}$.*

In particular, the vertices of P^f include the canonical flows $(v, \hat{f}(v))$, which dominate all other vertices (v, f) in terms of start-up costs (cf. Proposition 4.3). Hence, we have

$$\begin{aligned}\operatorname{conv}\Big\{\pi^f(v,\hat{f}(v)) \,\Big|\, v \in \{0,1\}^T\Big\} \subset \pi^f(P^f) = \operatorname{conv}\Big\{\pi^f(v,f) \,\Big|\, (v,f) \in V^f\Big\} \\ \subset \operatorname{conv}\Big\{\pi^f(v,\hat{f}(v)) \,\Big|\, v \in \{0,1\}^T\Big\} + \operatorname{pos}\{u_{T+t} \mid t \in [T]\},\end{aligned}$$

where u_t denotes the t-th unit vector.

Again by Proposition 4.3, the points $\pi^f(v, \hat{f}(v))$ equal the vertices of the epigraph of the start-up costs conv(epi(DCU)),

$$\operatorname{conv}(\operatorname{epi}(\mathrm{DCU})) = \operatorname{conv}\Big\{(v, \mathrm{DCU}(v)) \,\Big|\, v \in \{0,1\}^T\Big\} + \operatorname{pos}\{u_{T+t} \mid t \in [T]\},$$

and hence the lower boundaries of P^f and conv(epi(DCU)) are equal under projection.

Theorem 4.5

$$\pi^f(P^f) + \operatorname{pos}\{u_{T+t} \mid t \in [T]\} = \operatorname{conv}(\operatorname{epi}(\mathrm{DCU})).$$

4.3 The Start-up Type Polyhedron

[SBB10] presents the idea of grouping start-ups with the same start-up costs, such that multiple offline lengths result in the same *start-up type*. This reduces the number of variables for step-wise constant start-up cost functions, e. g. approximations of start-up cost functions computed with Algorithm 2.1.2.

In this section, we introduce the concept of start-up types to P^f, reducing the number of variables from $\mathcal{O}(T^2)$ to $\mathcal{O}(ST)$, where S is the number start-up types. The resulting extended formulation P^δ of conv(epi(DCU)) has $\mathcal{O}(T^2)$ inequalities, of which $\mathcal{O}(ST)$ suffice to model the start-up costs for integral operational schedules $v \in \{0,1\}^T$.

The inequalities of P^δ have $\mathcal{O}(ST)$ non-zero coefficients, resulting in a separation by exhaustive search with running time $\mathcal{O}(ST^3)$. We present a $\mathcal{O}(T^2)$ separation algorithm which checks each inequality in constant time.

After defining start-up types in the context of P^f, we analyze the resulting formulation

- by deriving valid inequalities in Subsection 4.3.1,
- by showing that a subset of these valid inequalities models the start-up costs correctly for $v \in \{0,1\}^T$ in Subsection 4.3.2,
- by proving that all valid inequalities model the start-up costs correctly for $v \in [0,1]^T$ in Subsection 4.3.3, and
- by providing a separation algorithm in Subsection 4.3.4.

Consider a start-up cost function with $\mathrm{CU}^t(l) = \mathrm{CU}^t(l+n)$ for some t, l, n. Then, the flows $f_l^t, \dots, f_{l+n}^t$ induce the same start-up costs $\mathrm{CU}^t(l)$. The idea of the start-up types is to model these flows by a single variable $\delta_s^t := f_l^t + \cdots f_{l+n}^t$, where s denotes the start-up type with offline lengths $[l..l+n]$.

For a given start-up cost function, we define a grouping of the offline lengths into start-up types for period t as a partition of $[0 \,..\, t-1]$ given by non-empty intervals $\mathcal{L}_0^t, \dots, \mathcal{L}_{S^t}^t$ with $\mathcal{L}_0^t = \{0\}$, such that the lengths in each interval induce the same costs,

$$\forall\, s \in [0..S^t], l_1, l_2 \in \mathcal{L}_s^t : \qquad \mathrm{CU}^t(l_1) = \mathrm{CU}^t(l_2),$$

and such that the lengths increase with s,

$$\forall\, s \in [S^t] : \qquad \max \mathcal{L}_{s-1}^t = \min \mathcal{L}_s^t - 1.$$

If CU is a step function, then each such interval $\mathcal{L}_s^t$ corresponds to a step of CU. We denote the total number of such types by $S^\Sigma := \sum_{t\in[T]}(S^t+1)$.

Combining the flows f_l^t with $l \in \mathcal{L}_s^t$ into the start-up type δ_s^t can be expressed by the projection

$$\pi^{f\delta} : \quad \mathbb{R}^{T+T(T+1)/2} \to \mathbb{R}^{T+S^\Sigma}, \quad (v, f) \mapsto (v, \delta),$$

with

$$\forall\, t \in [T], s \in [0..S^t] : \qquad \delta_s^t := \sum_{l\in\mathcal{L}_s^t} f_l^t.$$

By the choice of $\mathcal{L}_s^t$, the start-up costs induced by these composite flows equate to

$$\forall\, t \in [T] : \qquad \mathrm{cu}^t = \sum_{l=0}^{t-1} \mathrm{CU}^t(l) f_l^t = \sum_{s=0}^{S^t} \mathrm{CU}^t\Big(\min \mathcal{L}_s^t\Big)\delta_s^t.$$

Thus, the start-up types (v, δ) are canonically projected to conv(epi(DCU)) by

$$\pi^\delta : \quad \mathbb{R}^{T+S^\Sigma} \to \mathbb{R}^{2T}, \quad (v, \delta) \mapsto (v, \mathrm{cu})$$

with

$$\forall\, t \in [T] : \qquad \mathrm{cu}^t := \sum_{s=0}^{S^t} \mathrm{CU}^t\Big(\min \mathcal{L}_s^t\Big)\delta_s^t.$$

Note that by definition the start-up costs $\pi^f(v, f)$ associated to (v, f) are equal to the start-up costs $\pi^\delta(v, \delta)$ associated to $(v, \delta) = \pi^{f\delta}(v, f)$.

Proposition 4.6

$$\pi^\delta \circ \pi^{f\delta} = \pi^f.$$

Therefore $\pi^\delta(\pi^{f\delta}(P^f)) = \pi^f(P^f) = \mathrm{conv}(\mathrm{epi}(\mathrm{DCU}))$, and $\pi^{f\delta}(P^f)$ is also an extended formulation of conv(epi(DCU)).

4.3.1 Valid Inequalities

The $\mathcal{H}$-representation of P^f is given by (4.2.1)-(4.2.4).

The edge capacities and the non-negativity of the flow remain valid for $\pi^{f\delta}(P^f)$,

$$\forall\, t \in [T]: \qquad 0 \le v^t \le 1, \tag{4.3.1}$$

$$\forall\, t \in [T], s \in [0\,..\,S^t]: \qquad \delta_s^t \ge 0. \tag{4.3.2}$$

The conservation inequality (4.2.3) on the nodes t_+,

$$\forall\, t \in [T]: \qquad \sum_{l=0}^{t-1} f_l^t = v^t,$$

is easily translated to $\pi^{f\delta}(P^f)$, as each start-up type δ_s^t is the combination of flows f_l^t with the same start node t:

$$\forall\, t \in [T]: \qquad \sum_{s=0}^{S^t} \delta_s^t = v^t. \tag{4.3.3}$$

The flow conservation inequality (4.2.4) at nodes t_-,

$$\forall\, t \in [T-1]: \qquad \sum_{t'=t+1}^{T} f_{t'-t-1}^{t'} \le v^t,$$

however cannot be translated in the same manner, since each start-up type has multiple end nodes. Thus, a start-up type δ_s^t can only be bounded if the inequality regards all end nodes of δ_s^t simultaneously. To simplify the notation, we denote the indices of the end nodes of δ_s^t by $\mathcal{E}_s^t$, i. e.

$$\forall\, t \in [T],\ s \in [0..S^t]: \qquad \mathcal{E}_s^t := \{t - l - 1 \mid l \in \mathcal{L}_s^t\}.$$

For each set $A \subset [T]$, the nodes $\{t_- \mid t \in A\}$ thus comprise the end nodes of all start-up types δ_s^t with

$$(t,s) \ \in\ \mathcal{I}(A) := \Big\{(t',s') \,\Big|\, t' \in [T], s' \in [0..S^{t'}] \ \text{ with } \ \mathcal{E}_{s'}^{t'} \subset A\Big\}. \tag{4.3.4}$$

We call $\mathcal{I}(A)$ the set of *incoming start-up types* of A. By definition, these start-up types correspond to a subset of the incoming start-up flows of $\{t_- \mid t \in A\}$, i. e.

$$\sum_{(t,s)\in\mathcal{I}(A)} \delta_s^t = \sum_{(t,s)\in\mathcal{I}(A)} \overbrace{\sum_{l\in\mathcal{L}_s^t} f_l^t}^{=\delta_s^t} = \sum_{(t,s)\in\mathcal{I}(A)} \sum_{e\in\mathcal{E}_s^t} f_{t-e-1}^t \le \sum_{e\in A} \sum_{t=e+1}^{T} f_{t-e-1}^t.$$

By summing the flow conservation inequalities (4.2.4) of nodes $\{t_- \mid t \in A\}$, we can thereby bound their incoming start-up types,

$$\forall\, A \subset [T]: \qquad \sum_{(t,s)\in\mathcal{I}(A)} \delta_s^t \leq \sum_{e\in A} \sum_{t=e+1}^{T} f_{t-e-1}^t \overset{(4.2.4)}{\leq} \sum_{t\in A} v^t. \tag{4.3.5}$$

This combined flow conservation inequality for A inequality is valid for $\pi^{f\delta}(P^f)$.

Note however that this inequality is not irredundant. For each set $A \subset [T]$ that can be partitioned in two sets $A_1 \dot\cup A_2 = A$ with $\mathcal{I}(A) = \mathcal{I}(A_1) \cup \mathcal{I}(A_2)$, the flow conservation inequalities for A_1 and A_2 imply the inequality for A,

$$\sum_{(t,s)\in\mathcal{I}(A_1\cup A_2)} \delta_s^t = \sum_{(t,s)\in\mathcal{I}(A_1)} \delta_s^t + \sum_{(t,s)\in\mathcal{I}(A_2)} \delta_s^t \leq \sum_{t\in A_1} v^t + \sum_{t\in A_2} v^t = \sum_{t\in A_1\cup A_2} v^t.$$

By definition, $\mathcal{I}(A) = \mathcal{I}(A_1) \cup \mathcal{I}(A_2)$ iff each set of end nodes $\mathcal{E}_s^t \subset [a..b]$ fulfills $\mathcal{E}_s^t \subset A_1$ or $\mathcal{E}_s^t \subset A_2$. Since the sets $\mathcal{E}_s^t$ are intervals, such A_1, A_2 exist for all $A \subset [T]$ which are not intervals. Hence, no irredundant inequalities (4.3.5) are disregarded by only considering

$$\forall\, a \in [T-1], b \in B(a): \qquad \sum_{(t,s)\in\mathcal{I}([a..b])} \delta_s^t \leq \sum_{t=a}^{b} v^t \tag{4.3.6}$$

with $B(a) := \Big\{ b \in [a\,..\,T-1] \;\Big|\; \forall\, b' \in [a\,..\,b-1] : \mathcal{I}([a..b]) \neq \mathcal{I}([a..b']) \cup \mathcal{I}([b'+1\,..\,b]) \Big\}$.

Corollary 4.7 *The set of inequalities* (4.3.5) *dominates the set of inequalities* (4.3.6).

4.3.2 Integral Operational Schedules

The canonical flow $\hat f(v)$ is extended to start-up types by

$$\forall\, t \in [T], s \in [0..S^t]: \qquad \hat\delta_s^t(v) := \begin{cases} 1 & \text{if } v^t = 1 \text{ and } \mathrm{ol}^t(v) \in \mathcal{L}_s^t, \\ 0 & \text{else,} \end{cases}$$

By definition $(v, \hat\delta(v)) = \pi^{f\delta}(v, \hat f(v))$, and hence feasibility and associated start-up costs follow directly from Proposition 4.1 and (4.1.3).

Corollary 4.8 *For each* $v \in \{0,1\}^T$,

$$(v, \hat\delta(v)) \in \pi^{f\delta}(P^f) \qquad \textit{and} \qquad \pi^{\delta}(v, \hat\delta(v)) = (v, \mathrm{DCU}(v)).$$

Proposition 4.3 shows that P^f models the start-up costs correctly for $v \in \{0,1\}^T$ by noting that start-up costs associated with the canonical flow $\hat f(v)$ are minimal.

The proof centers on the flow conservation of the nodes t_-, which enforces $f_l^t = 0$ for $l < \mathrm{ol}^t(v)$.

An analogous argument can be made for the start-up types using the following selection of inequalities (4.3.6):

$$\forall\, [a..b] \in \left\{ \mathcal{E}_s^t \;\middle|\; t \in [T], s \in [0\,..\,S^t] \right\} : \qquad \sum_{(t,s)\,\in\,\mathcal{I}([a..b])} \delta_s^t \leq \sum_{t\,\in\,[a..b]} v^t. \tag{4.3.7}$$

Note that these inequalities are a subset of the inequalities (4.3.6): $[a..b] = \mathcal{E}_s^t$ implies for each $b' \in [a\,..\,b{-}1]$ that $(t,s) \in \mathcal{I}([a..b]) \setminus \mathcal{I}([a..b']) \cup \mathcal{I}([b'{+}1\,..\,b])$, and thus $b \in B(a)$.

We call the resulting polyhedron

$$\hat{P}^\delta := \left\{ (v,\delta) \in \mathbb{R}^{T+S^\Sigma} \text{ fulfilling (4.3.1), (4.3.2),(4.3.3), and (4.3.7)} \right\} \tag{4.3.8}$$

the *start-up type polyhedron* for integers, and apply the approach of Proposition 4.3 to $\hat{P}^\delta$:

Proposition 4.9 *Each $(v,\delta) \in \hat{P}^\delta$ with $v \in \{0,1\}^T$ fulfills $\pi^\delta(v,\delta) \geq (v, \mathrm{DCU}(v))$.*

Proof. We show that $\pi^\delta(v,\delta) \geq (v, \mathrm{DCU}(v))$ by defining the start-up costs cu such that $(v,\mathrm{cu}) = \pi^\delta(v,\delta)$ and examining cu^t for each $t \in [T]$.

Since the intervals $\mathcal{L}_s^t$ partition $[0\,..\,t{-}1]$, there exists a unique $s^* \in [0..S^t]$ with $\mathrm{ol}^t(v) \in \mathcal{L}_{s^*}^t$. By definition of $\mathrm{ol}^t(v)$, we have $v^{t'} = 0$ for each $s \in [0\,..\,s^*{-}1]$, $t' \in \mathcal{E}_s^t$. Therefore by (4.3.7)

$$\delta_s^t \leq \sum_{(t',s')\,\in\,\mathcal{I}(\mathcal{E}_s^t)} \delta_{s'}^{t'} \leq \sum_{t'\in\mathcal{E}_s^t} v^{t'} = 0,$$

implying $\delta_s^t = 0$. Since CU is increasing, it follows from inequality (4.3.3) that

$$\begin{aligned} \mathrm{cu}^t &= \sum_{s=0}^{S^t} \mathrm{CU}^t(\min \mathcal{L}_{s^*}^t)\delta_s^t = \sum_{s=s^*}^{S^t} \mathrm{CU}^t(\min \mathcal{L}_{s^*}^t)\delta_s^t \\ &\geq \mathrm{CU}^t(\min \mathcal{L}_{s^*}^t) \sum_{s=s^*}^{S^t} \delta_s^t = \mathrm{CU}^t(\min \mathcal{L}_{s^*}^t) \sum_{s=0}^{S^t} \delta_s^t = \mathrm{CU}^t(\min \mathcal{L}_{s^*}^t) v^t, \end{aligned}$$

which by the choice of s^* and by definition of the start-up costs in (1.3.6) implies

$$\mathrm{cu}^t \geq \mathrm{CU}^t(\min \mathcal{L}_{s^*}^t) v^t = \mathrm{CU}^t(\mathrm{ol}^t(v)) v^t = \mathrm{DCU}^t(v). \qquad \square$$

4.3.3 Fractional Operational Schedules

In general, $\pi^{f\delta}(P^f)$ is a proper subset of $\hat{P}^\delta$, meaning that $\hat{P}^\delta$ does not bound the start-up costs tightly for fractional operational schedules. Consider the following example with $T = 5$, start-up costs

$$\forall\, t \in [T]: \qquad \mathrm{CU}^t(1) = \mathrm{CU}^t(2) = 2, \quad \mathrm{CU}^t(3) = \mathrm{CU}^t(4) = 3$$

independent of t and start-up types

$$\begin{aligned} &\mathcal{L}^1_0 = \{0\}, \qquad \mathcal{L}^2_0 = \{0\}, \mathcal{L}^2_1 = \{1\}, \qquad \mathcal{L}^4_0 = \{0\}, \mathcal{L}^4_1 = [1..2], \mathcal{L}^4_2 = \{3\}, \\ &\qquad\qquad \mathcal{L}^3_0 = \{0\}, \mathcal{L}^3_1 = [1..2], \qquad \mathcal{L}^5_0 = \{0\}, \mathcal{L}^5_1 = [1..2], \mathcal{L}^5_2 = [3..4]. \end{aligned}$$

For the operational schedule $v = (1, 1/2, 0, 1/2, 1)$, the point $(v, f) \in P^f$ with

$$f^4_1 = 1/2, f^5_3 = 1/2, \text{ and } f^t_l = 0 \text{ for all other } (t, l),$$

results in minimal start-up costs $\mathrm{cu} = (0, 0, 0, 2/2, 3/2)$. Yet, for $(v, \delta) \in \hat{P}^\delta$ the start-up types

$$\delta^4_1 = 1/2, \delta^5_1 = 1/2, \text{ and } \delta^t_s = 0 \text{ for all other } (t, s),$$

are feasible, resulting in strictly lower costs $\tilde{\mathrm{cu}} = (0, 0, 0, 2/2, 2/2)$. Thus, in general

$$\pi^{f\delta}(P^f) \subsetneq \hat{P}^\delta. \tag{4.3.9}$$

This subsection shows that by using all instances of inequality (4.3.6), the so-called *start-up type polyhedron*

$$P^\delta := \left\{(v, \delta) \in \mathbb{R}^{T+S^\Sigma} \text{ fulfilling (4.2.1), (4.3.2),(4.3.3), and (4.3.6)}\right\} \tag{4.3.10}$$

models $\pi^{f\delta}(P^f)$ and is thereby an extended formulation of conv(epi(DCU)).

Since inequalities (4.3.3) and (4.3.6) are valid for $\pi^{f\delta}(P^f)$, it holds that $\pi^{f\delta}(P^f) \subset P^\delta$. We prove $P^\delta \subset \pi^{f\delta}(P^f)$ by showing that each type δ may be "split" into feasible flows f, i. e. that for each $(v, \delta) \in P^\delta$ there exists $(v, f) \in P^f$ such that $\pi^{f\delta}(v, f) = (v, \delta)$.

Theorem 4.10

$$\pi^{f\delta}(P^f) = P^\delta.$$

Proof. The idea of this proof is to model the splitting of δ into f as a network flow problem. This problem is designed such that any maximal flow corresponds to $(v, f) \in P^f$ with $\pi^{f\delta}(v, f) = (v, \delta)$.

The underlying graph $G = (V, E)$ is shown in Fig. 4.4. Its vertices V consist of a source a, a sink b and two sets of nodes X and Y,

$$V := \{a, b\} \cup X \cup Y \text{ with } X := \{x^t_s \,|\, t \in [T], s \in [0..S^t]\}, \quad Y := \{y^t \,|\, t \in [0..T-1]\}.$$

Its directed edges E are subdivided into three categories,

$$E := E_\delta \cup E_f \cup E_v$$

with

$$E_\delta := \left\{ (a, x_s^t) \;\middle|\; t \in [T], s \in [0..S^t] \right\} \subset \{a\} \times X,$$
$$E_f := \left\{ (x_s^t, y^{t'}) \;\middle|\; t \in [T], s \in [0..S^t], t' \in \mathcal{E}_s^t \right\} \subset X \times Y,$$
$$E_v := \left\{ (y^t, b) \;\middle|\; t \in [T-1] \right\} \subset Y \times \{b\}.$$

Each edge (a, x_s^t) corresponds to the start-up type δ_s^t, each edge $(x_s^t, y^{t'})$ corresponds to the start-up flow f_l^t with $l = t - t' - 1 \in \mathcal{L}_s^t$, and each edge (y^t, b) corresponds to v^t.

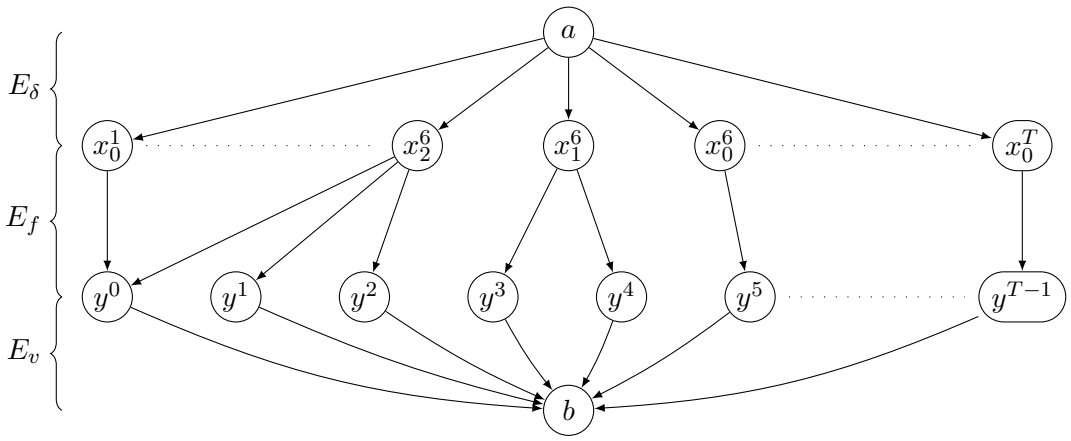


Figure 4.4: Graph modeling the splitting of δ into f. The edges between x_s^t and $y^{t'}$ are shown for x_s^6 with $S^6 = 2$.

The edge capacities of E_δ and E_v are derived from the considered point (v, δ), and the capacities of (y_0, b) and E_f are unlimited,

$$c: \quad E \to \mathbb{R}_{\geq 0}, \quad e \mapsto \begin{cases} \delta_s^t & \text{if } e = (a, x_s^t), \\ v^t & \text{if } e = (y^t, b) \text{ with } t \geq 1, \\ \infty & \text{if } e = (y^0, b), \\ \infty & \text{if } e = (x_s^t, y^{t'}). \end{cases}$$

A flow in this network is uniquely determined by the flow on the edges $(x_s^t, y^{t'})$: Denoting the flow on $(x_s^t, y^{t'})$ by f_l^t with $l = t - t' - 1 \in \mathcal{L}_s^t$, the flow conservation of the nodes x_s^t and y^t state

- that the flow on each edge (a, x_s^t) equals $\sum_{l \in \mathcal{L}_s^t} f_l^t$, and

- that the flow on each edge (y^t, b) equals $\sum_{t'=t+1}^{T} f^{t'}_{t'-t-1}$.

Therefore, the capacities of the edges (y^t, b) imply the inequality (4.2.4) of P^f,

$$\forall\, t \in [T-1]: \qquad \sum_{t'=t+1}^{T} f^{t'}_{t'-t-1} \leq v^t,$$

and the capacities of the edges (a, x^t_s) imply

$$\forall\, t \in [T], s \in [0..S^t]: \qquad \sum_{l \in \mathcal{L}^t_s} f^t_l \leq \delta^t_s,$$

which is equivalent to $\pi^{f\delta}(v, f) \leq (v, \delta)$.

Note that if there exists an a-b-flow with value

$$C := \sum_{t \in [T], s \in [0..S^t]} \delta^t_s \quad ,$$

then the outgoing edges of node a are saturated, and it holds that

$$\forall\, t \in [T], s \in [0..S^t]: \qquad \sum_{l \in \mathcal{L}^t_s} f^t_l = \delta^t_s,$$

which is equivalent to $\pi^{f\delta}(v, f) = (v, \delta)$.

For such a flow, the inequality (4.2.3) of P^f is fulfilled as well since

$$\forall\, t \in [T]: \qquad \sum_{l \in [0..T-1]} f^t_l = \sum_{s \in [0..S^t]} \delta^t_s = v^t,$$

Hence $(v, f) \in P^f$, which proves $P^\delta \subset \pi^{f\delta}(P^f)$. The remainder of this proof shows that an a-b-flow with value C exists by demonstrating that a minimal a-b-cut has weight C.

We claim that the cut (A^*, B^*) with $A^* = \{a\}$ and $B^* := b \cup X \cup Y$ is such a minimal cut. Its cut-set is exactly E_δ, and hence its weight equals

$$\sum_{t \in [T], s \in [0..S^t]} \delta^t_s \quad = C.$$

Assume there exists a cut (A, B) with smaller weight. Since the edges (y^0, b) and E_f have unlimited capacity, it holds that $y^0 \in B$ and that for each $(x^t_s, y^{t'}) \in E_f$ either $\{x^t, y^{t'}\} \subset A$ or $\{x^t, y^{t'}\} \subset B$. By definition of E_f, this implies

$$\forall\, x^t_s \in A, t' \in \mathcal{E}^t_s: \qquad y^{t'} \in A,$$

which, when defining $0 \notin S := \{t \in [T] : y^t \in A\}$, is equivalent to

$$\forall\, x_s^t \in A: \qquad \mathcal{E}_s^t \subset S. \tag{4.3.11}$$

The weight of the cut (A, B) equals

$$\underbrace{\sum_{x_s^t \in X \cap B} \delta_s^t}_{\text{edges } E_\delta \cap A \times B} + \underbrace{\sum_{y^t \in Y \cap A, t \geq 1} v^t}_{\text{edges } E_v \cap A \times B}. \tag{4.3.12}$$

Since $(v, \delta) \in P^\delta$, (v, δ) fulfills inequality (4.3.6) and by Corollary 4.7 also (4.3.5). Hence, we can bound the total capacity of the edges $E_v \cap A \times B$ by

$$\sum_{y^t \in Y \cap A, t \geq 1} v^t = \sum_{t \in S} v^t \overset{(4.3.6)}{\geq} \sum_{(t,s) \in \mathcal{I}(S)} \delta_s^t = \sum_{\substack{t \in [T], s \in [0..S^t] \\ \mathcal{E}_s^t \subset S}} \delta_s^t \overset{(4.3.11)}{\geq} \sum_{x_s^t \in A} \delta_s^t.$$

Thus, the weight (4.3.12) of the cut (A, B) can be bounded from below by

$$\sum_{x_s^t \in X \cap B} \delta_s^t + \sum_{y^t \in Y \cap A, t \geq 1} v^t \geq \sum_{x_s^t \in X \cap B} \delta_s^t + \sum_{x_s^t \in X \cap A} \delta_s^t = \sum_{x_s^t \in X} \delta_s^t = C.$$

Therefore, the weight of a minimum cut is C, implying that the value of a maximum flow is C (cf. [FF56]). As shown, such a maximum flow represented by f fulfills $\pi^{f\delta}(v, f) = (v, \delta)$ and $(v, f) \in P^f$, and hence proves $P^\delta \subset \pi^{f\delta}(P^f)$. □

As noted, since $\pi^f = \pi^\delta \circ \pi^{f\delta}$ (see Proposition 4.6) and $\pi^f(P^f) = \text{conv}(\text{epi}(\text{DCU}))$ (see Theorem 4.5), P^δ is an extended formulation for conv(epi(DCU)).

Corollary 4.11

$$\pi^\delta(P^\delta) + \text{pos}\{u_{T+t} \mid t \in [T]\} = \text{conv}(\text{epi}(\text{DCU})).$$

Finally, both $\hat{P}^\delta$ and P^δ are generalizations of P^f. Specifically, if for each $t \in [T]$, $[0\,..\,t-1]$ is partitioned such that $\mathcal{L}_s^t = \{s\}$, then by definition $\pi^{f\delta}$ degenerates into the identity function, implying $P^f = P^\delta$. Furthermore, since $B(a) = \{a\}$ for each $a \in [T-1]$, inequalities (4.2.4), (4.3.6) and (4.3.7) are equal, implying that the $\mathcal{H}$-representations of P^f, P^δ, and $\hat{P}^\delta$ are equal.

Proposition 4.12 *If $\mathcal{L}_s^t = \{s\}$ for each $t \in [T]$, $s \in [0..S^t]$, then $\pi^{f\delta} = \text{id}$ and the $\mathcal{H}$-representations of P^f, P^δ, and $\hat{P}^\delta$ are equal.*

4.3.4 Separation

In P^δ, the $\mathcal{O}(T)$ flow conservation inequalities (4.2.4) of P^f are replaced by the $\mathcal{O}(T^2)$ inequalities (4.3.6). To reduce the impact on solution times, we present an exact separation algorithm for (4.3.6) with running time $\mathcal{O}(T^2)$. Note that due to the $\mathcal{O}(ST)$ non-zero coefficients of each of the $\mathcal{O}(T^2)$ inequalities, an exhaustive search requires a running time of $\mathcal{O}(ST^3)$ to separate (4.3.6).

The efficiency of our separation algorithm depends on an efficient construction of the sets of incoming start-up types $\mathcal{I}([a..b])$ of the nodes $\{t_- \mid t \in [a..b]\}$. By their definition in (4.3.4), it holds that

$$\mathcal{I}([a..b]) = \Big\{(t,s) \;\Big|\; t \in [T], s \in [0..S^t] \text{ with } \mathcal{E}^t_s \subset [a..b]\Big\} = \bigcup_{a' \in [a..b]} \bigcup_{b' \in [a'..b]} \mathcal{A}([a'..b'])$$

$$\text{with} \qquad \mathcal{A}([a..b]) = \Big\{(t,s) \;\Big|\; t \in [T], s \in [0..S^t] \;\text{ with }\; \mathcal{E}^t_s = [a..b]\Big\}.$$

Determining the sets $\mathcal{A}([a..b])$ is straightforward. The sets $\mathcal{I}([a..b])$ can be computed recursively by first deriving

$$\mathcal{B}([a..b]) := \bigcup_{b' \in [a..b]} \mathcal{A}([a..b']) = \mathcal{B}([a \,..\, b{-}1]) + \mathcal{A}([a..b]) \tag{4.3.13}$$

and then calculating

$$\mathcal{I}([a..b]) = \bigcup_{a' \in [a..b]} \mathcal{B}([a'..b]) = \mathcal{I}([a{+}1 \,..\, b]) \cup \mathcal{B}([a..b]) \tag{4.3.14}$$

When separating a point $(v,\delta) \in \hat{P}^\delta$ however, we do not require the actual sets $\mathcal{I}([a..b])$, but rather the sum of the values of the included start-up types, i. e.

$$\varphi(\mathcal{I}([a..b])) \quad \text{with} \quad \varphi(S) := \sum_{(t,s) \in S} \delta^t_s.$$

Using (4.3.13)–(4.3.14), Algorithm 4.3.1 computes these values in $\mathcal{O}(T^2)$, and checks the resulting inequality. Note that each iteration of the outer loop on lines 6–11 only requires the values $\varphi(\mathcal{I}([a'..b']))$ with $a' \in \{a, a+1\}$. Therefore, a careful implementation of these variables only requires $\mathcal{O}(T)$ space.

Theorem 4.13 *Algorithm 4.3.1 separates P^δ from $\hat{P}^\delta$ by inequality* (4.3.6) *in $\mathcal{O}(T^2)$.*

Proof. The sums $\sum_{t=a}^{b} v^t$ used on line 10 are easily computed in a total time of $\mathcal{O}(T^2)$. Due to its nested loop structure, the algorithm considers each tuple (a,b) with $a \in [T-1]$ and $b \in [a \,..\, T{-}1]$ at most once. Since the operations lines 8–11 have running time $\mathcal{O}(1)$, the total complexity of the algorithm is $\mathcal{O}(T^2)$.

Algorithm 4.3.1: Separate P^δ from $\hat{P}^\delta$ by inequality (4.3.6)

Input : Point $(v, \delta) \in \hat{P}^\delta$
Output : (v, δ) lies in P^δ or a violated inequality (4.3.6).

1 $\varphi(\mathcal{A}([a..b])) \leftarrow 0 \qquad \forall\ [a..b] \subset [T-1]$;
2 **for** $t \in [T], s \in [0\ ..\ S^t{-}1]$ **do**
3 $\quad \varphi(\mathcal{A}(\mathcal{E}^t_s)) \leftarrow \varphi(\mathcal{A}(\mathcal{E}^t_s)) + \delta^t_s$;
4 **for** $a = T-1, \ldots, 1$ **do**
5 $\quad \varphi(\mathcal{B}([a\ ..\ a{-}1])) \leftarrow 0$;
6 $\quad \varphi(\mathcal{I}([a{+}1\ ..\ a])) \leftarrow 0$;
7 $\quad$ **for** $b = a, \ldots, T-1$ **do**
8 $\quad\quad \varphi(\mathcal{B}([a..b])) \leftarrow \varphi(\mathcal{B}([a\ ..\ b{-}1])) + \varphi(\mathcal{A}([a..b]))$;
9 $\quad\quad \varphi(\mathcal{I}([a..b])) \leftarrow \varphi(\mathcal{I}([a{+}1\ ..\ b])) + \varphi(\mathcal{B}([a..b]))$;
10 $\quad\quad$ **if** $\varphi(\mathcal{I}([a..b])) > \sum_{t=a}^{b} v^t$ **then**
11 $\quad\quad\quad$ **stop** (v, δ) violates inequality (4.3.6) with parameters a, b;
12 **return** $(v, \delta) \in P^\delta$;

It is straightforward to check that lines 1–3 compute $\varphi(\mathcal{A}([a..b]))$ by considering each δ^t_s. Furthermore, the calculations on lines 5–9 apply the relationships presented in (4.3.13)–(4.3.14), as well as in Algorithm 4.3.1, and hence the algorithm computes the values $\varphi(\mathcal{I}([a..b]))$ correctly. So, line 10 checks inequality (4.3.5) as desired.

What remains to be proved is that only violated inequalities (4.3.6) with $b \in B(a)$ are reported. Assume the algorithm stops claiming that (v, δ) violates inequality (4.3.6) for parameters $a \in [T-1]$, $b \in [a\ ..\ T{-}1]$ due to

$$\varphi(\mathcal{I}([a..b])) > \sum_{t=a}^{b} v^t. \tag{4.3.15}$$

If $b \notin B(a)$, then there exists $b' \in [a..b]$ such that $\mathcal{I}([a..b]) = \mathcal{I}([a..b']) \cup \mathcal{I}([b'+1..b])$. By design of the loops, prior to stopping the algorithm has already checked that

$$\varphi(\mathcal{I}([a..b'])) \leq \sum_{t=a}^{b'} v^t \quad \text{and} \quad \varphi(\mathcal{I}([b'{+}1\ ..\ b])) \leq \sum_{t=b'+1}^{b} v^t,$$

which contradicts (4.3.15) since $\varphi(\mathcal{I}([a..b])) = \varphi(\mathcal{I}([a..b'])) + \varphi(\mathcal{I}([b'+1..b]))$. Therefore, $b \in B(a)$. □

4.4 Start-up and Shutdown Indicators

While the polyhedron P^δ in the last section models the types δ_l^t based on the operational schedule v, [SBB10] derives the start-up types from the start-up and shutdown indicators y^t and z^t, which are part of most Unit Commitment formulations (see (1.2.1)–(1.2.2)). In this section, we show that $\delta_0^t = v^t - \mathrm{y}^t = v^{t-1} - \mathrm{z}^t$ and use this relationship to substitute δ_0^t in P^δ. This reduces the number of variables by T and simplifies the comparison to [SBB10] performed in Section 4.5.

The start-up and shutdown indicator variables y^t and z^t are introduced in [Gar62], and are widely used in UC formulations, for example to model minimum up-/downtime and ramping limits (see e. g. [AC00; RT05; OAV12; MELR13b; MELR13a]), and exponential start-up costs (see Chapter 3).

The indicators are expressed by two sets of binary variables $\mathrm{y}, \mathrm{z} \in \{0,1\}^T$ where $\mathrm{y}^t = 1$ iff the unit starts up in period t and $\mathrm{z}^t = 1$ iff the unit shuts down in period t. In our basic formulation in Section 1.2, these variables are modeled by the inequalities (1.2.1)–(1.2.3), which we repeat for convenience:

$$\forall\, t \in [T]: \qquad 0 \leq v^t \leq 1, \tag{4.4.1}$$

$$\forall\, t \in [2..T]: \qquad \mathrm{y}^t - \mathrm{z}^t = v^t - v^{t-1}, \tag{4.4.2}$$

$$\mathrm{y}^1 - \mathrm{z}^1 = v^1 - \begin{cases} 1 & \mathrm{PDT} = 0, \\ 0 & \text{else}, \end{cases} \tag{4.4.3}$$

$$\forall\, t \in [T]: \qquad \mathrm{y}^t \leq v^t, \tag{4.4.4}$$

$$\forall\, t \in [T]: \qquad \mathrm{z}^t \leq 1 - v^t, \tag{4.4.5}$$

$$\forall\, t \in [T]: \qquad \mathrm{y}^t, \mathrm{z}^t \in \{0,1\}^T.$$

In this section, we only use the linear relaxation of the above integrality constraint, which can be modeled as

$$\forall\, t \in [T]: \qquad \mathrm{y}^t, \mathrm{z}^t \geq 0, \tag{4.4.6}$$

since the inequality $\mathrm{y}^t, \mathrm{z}^t \leq 1$ is dominated by (4.4.4) and (4.4.5).

The start-up and shutdown indicators can be interpreted as the change of the operational schedule v^t from period $t-1$ to t, where y^t is the increase and z^t is the decrease, i. e.

$$\forall\, t \in [2..T]: \qquad \mathrm{y}^t = \max\{0, v^t - v^{t-1}\}, \qquad \mathrm{z}^t = \max\{0, v^{t-1} - v^t\}$$

On the other hand, the types δ_0^t represent the amount of the operational schedule which remains constant from period $t-1$ to t,

$$\forall\, t \in [2..T]: \qquad \delta_0^t = \min\{v^{t-1}, v^t\},$$

Thus, the operational schedule is split into constant and changing part (cf. Fig. 4.5),

$$\forall\, t \in [2..T]: \qquad v^t = \mathrm{y}^t + \delta_0^t, \qquad v^{t-1} = \mathrm{z}^t + \delta_0^t. \tag{4.4.7}$$

Relationship (4.4.7) is easily verified for integral $v \in \{0,1\}^T$ and its natural types $\hat{\delta}_l^t(v)$: $\mathrm{y}^t = 1$ iff $v^t = 1$ and $v^{t-1} = 0$, and $\hat{\delta}_0^t(v) = 1$ iff $v^t = 1$ and $v^{t-1} = 1$. Hence, $\mathrm{y}^t + \hat{\delta}_0^t(v) = 1$ iff $v^t = 1$. Moreover, $\mathrm{z}^t + \hat{\delta}_0^t(v) = \mathrm{y}^t - v^t + v^{t-1} + \hat{\delta}_0^t(v) = v^{t-1}$.

Since the start-up and shutdown indicators are already part of a typical Unit Commitment formulation, we can reduce the number of variables by substituting the variables δ_0^t using (4.4.7). Taking into account the case $t = 1$, this substitution is expressed by the mapping

$$\pi^{\delta\to\mathrm{z}}: \quad \mathbb{R}^{T+S^\Sigma} \to \mathbb{R}^{3T+(S^\Sigma-T)}, \quad (v,\delta) \mapsto (v, \mathrm{y}, \mathrm{z}, \delta_{\geq 1}) \tag{4.4.8}$$

with

$$\mathrm{y}^t := \begin{cases} v^t - \delta_0^t & \text{if } t \geq 2 \text{ or PDT} = 0, \\ v^t & \text{else,} \end{cases} \qquad \mathrm{z}^t := \mathrm{y}^t - v^t + \begin{cases} v^{t-1} & \text{if } t \geq 2, \\ 1 & \text{if } t = 1 \text{ and PDT} = 0, \\ 0 & \text{else,} \end{cases}$$

where $\delta_{s\geq 1}$ comprises all variables δ_s^t with $t \in [T]$ and $s \in [S^t]$, but excludes δ_0^t.

In the following, we consider the resulting polyhedron

$$P_{\mathrm{yz}}^\delta := \pi^{\delta\to\mathrm{z}}(P^\delta). \tag{4.4.9}$$

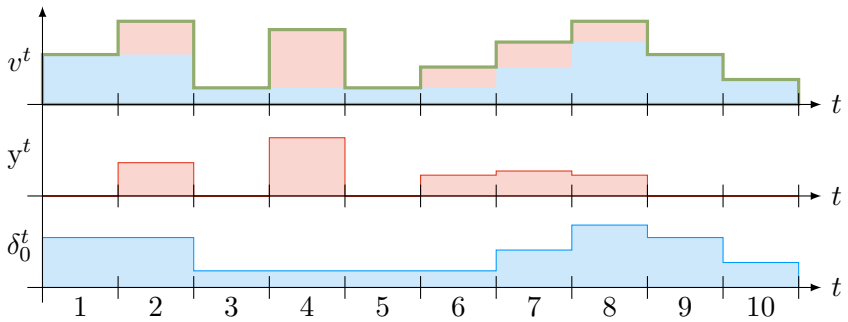


Figure 4.5: Fractional operational schedule v^t with corresponding start-up indicators y^t and start-up types δ_0^t. Note that $v^t = \mathrm{y}^t + \delta_0^t$, $v^{t-1} = \mathrm{z}^t + \delta_0^t$.

In Corollary 4.11, it was shown that if each point $(v, \delta) \in P^\delta$ is associated with the start-up costs

$$\forall\, t \in [T]: \qquad \mathrm{cu}^t = \sum_{s \in [0\,..\,S^t]} \mathrm{CU}^t(\min \mathcal{L}_s^t)\delta_s^t,$$

then P^δ models the lower boundary of conv(epi(DCU)) under projection. Since $\mathrm{CU}^t(0) = 0$ for $t \geq 2$ and $\delta_0^1 = v^1$ by (4.3.3), the same costs are associated with $(v, \mathrm{y}, \mathrm{z}, \delta_{\geq 1}) := \pi^{\delta \to \mathrm{z}}(v, \delta)$ by

$$\pi_{\mathrm{yz}}^\delta: \quad \mathbb{R}^{2T+S^\Sigma} \to \mathbb{R}^{2T}, \quad (v, \mathrm{y}, \mathrm{z}, \delta_{\geq 1}) \mapsto (v, \mathrm{cu})$$

with

$$\forall\, t \in [T]: \qquad \mathrm{cu}^t = \sum_{s \in [S^t]} \mathrm{CU}^t(\min \mathcal{L}_s^t)\delta_s^t + \begin{cases} \mathrm{CU}^1(0)v^1 & \text{if } t = 1, \\ 0 & \text{else.} \end{cases}$$

Therefore, Corollary 4.11 can be extended to P_{yz}^δ:

Corollary 4.14

$$\pi_{\mathrm{yz}}^\delta(P_{\mathrm{yz}}^\delta) + \mathrm{pos}\{u_{T+t} \mid t \in [T]\} = \mathrm{conv}(\mathrm{epi}(\mathrm{DCU})).$$

We now develop an $\mathcal{H}$-representation of P_{yz}^δ from the given $\mathcal{H}$-representation of P^δ, which comprises the inequalities (4.3.3), (4.3.6), and (4.3.2). Recall that by definition of $\pi^{\delta \to \mathrm{z}}$, we have

$$\forall\, t \in [2..T]: \qquad \delta_0^t = v^t - \mathrm{y}^t = v^{t-1} - \mathrm{z}^t$$

for each $(v, \delta) \in \mathbb{R}^{2T+S^\Sigma}$ and its projection $(v, \mathrm{y}, \mathrm{z}, \delta_{\geq 1}) := \pi^{\delta \to \mathrm{z}}(v, \delta)$.

By substituting $\delta_0^t = v^t - \mathrm{y}^t$ in the flow conservation inequality (4.3.3) of the nodes t_+, for $t \geq 2$ we obtain

$$\forall\, t \in [2..T]: \qquad \sum_{s=1}^{S^t} \delta_s^t = \mathrm{y}^t. \tag{4.4.10}$$

Using $\delta_0^t = v^{t-1} - \mathrm{z}^t$, the flow conservation in the nodes t_- (4.3.6) translates to

$$\forall\, a \in [T-1], b \in B(a): \qquad \sum_{\substack{(t,s) \in \mathcal{I}([a..b]) \\ s \geq 1}} \delta_s^t \;\leq\; \sum_{t=a}^{b} \mathrm{z}^{t+1}. \tag{4.4.11}$$

The non-negativity of the start-up types δ_s^t with $s \geq 1$ in (4.3.2) remains unchanged,

$$\forall\, t \in [T], s \in [S^t]: \qquad \delta_s^t \geq 0. \tag{4.4.12}$$

Furthermore, the following result shows that the projection $\pi^{\delta \to \mathrm{z}}$ is designed such that the resulting indicators as modeled by (4.4.1)–(4.4.6).

Proposition 4.15

$$P^{\delta}_{\mathrm{yz}} = \left\{(v, \mathrm{y}, \mathrm{z}, \delta_{\geq 1}) \in \mathbb{R}^{2T+S^{\Sigma}} \text{ fulfilling (4.4.1)–(4.4.6),(4.4.10)–(4.4.12)}\right\}.$$

Proof. We start by proving that each point $(v, \mathrm{y}, \mathrm{z}, \delta_{\geq 1}) \in \mathbb{R}^{2T+S^{\Sigma}}$ which fulfills (4.4.1)–(4.4.6), (4.4.10)–(4.4.12) lies in P^{δ}_{yz}. For such a point, define

$$\delta_0^t = \begin{cases} v^t - \mathrm{y}^t & \text{if } t \geq 2, \\ v^t & \text{else,} \end{cases}$$

resulting in $\pi^{\delta\to\mathrm{z}}(v, \delta) = (v, \mathrm{y}, \mathrm{z}, \delta_{\geq 1})$. This point (v, δ) lies in P^{δ}, since

- (4.4.4) implies $\delta_0^t \geq 0$, which together with (4.4.12) is equivalent to (4.3.2),
- for $t = 1$, (4.3.3) is fulfilled by definition of δ_0^t,
- for $t \geq 2$, (4.3.3) is equivalent to (4.4.10), and
- (4.3.6) is equivalent to (4.4.11).

Therefore, $(v, \mathrm{y}, \mathrm{z}, \delta_{\geq 1}) = \pi^{\delta\to\mathrm{z}}(v, \delta) \in \pi^{\delta\to\mathrm{z}}(P^{\delta}) = P^{\delta}_{\mathrm{yz}}$.

For the converse direction, we show that each $(v,\mathrm{y},\mathrm{z},\delta_{\geq 1}) \in P^{\delta}_{\mathrm{yz}}$ fulfills (4.4.1)–(4.4.6), (4.4.10)–(4.4.12) by considering the point $(v, \delta) \in P^{\delta}$ with $(v, \mathrm{y}, \mathrm{z}, \delta_{\geq 1}) = \pi^{\delta\to\mathrm{z}}(v, \delta)$. Again, inequalities (4.4.10)–(4.4.12) on the variables δ_l^t are implied by (4.3.3)–(4.3.2) by construction.

Moreover, the definition of $\pi^{\delta\to\mathrm{z}}$ in (4.4.8) guarantees that

- (4.4.2), (4.4.3) holds by definition of z^t,
- (4.4.4) is implied by $\delta_0^t \geq 0$,
- $\mathrm{y}^t \geq 0$ in (4.4.6) is fulfilled since $\delta_0^t \leq \sum_{s=0}^{S^t} \delta_s^t = v^t$ (see (4.3.3)), and
- $\mathrm{z}^t \geq 0$ in (4.4.6) holds due to

$$\mathrm{z}^t = \begin{cases} \mathrm{y}^t - v^t + v^{t-1} = v^{t-1} - \delta_0^t \geq v^{t-1} - \sum_{s=0}^{S^t} \delta_s^t \geq 0 & \text{for } t \geq 2, \\ \mathrm{y}^1 - v^1 + 1 \geq 1 - v^1 \geq 0 & \text{for } t = 1, \mathrm{PDT} = 0, \\ \mathrm{y}^1 - v^1 = 0 & \text{for } t = 1, \mathrm{PDT} > 0. \end{cases}$$

Therefore, $(v, \mathrm{y}, \mathrm{z}, \delta_{\geq 1})$ fulfills (4.4.2)–(4.4.6),(4.4.10)-(4.4.12). □

Moreover, the projection $\pi^{\delta\to\mathrm{z}}$ may also be applied to the start-up type polyhedron $\hat{P}^{\delta}$ for integers, which models the start-up costs correctly for $v \in \{0,1\}^T$ (see Proposition 4.9) using only the subset (4.3.7) of the inequalities (4.3.6). Substituting δ_0^t by $v^t - \mathrm{y}^t$ in (4.3.7) yields

$$\forall\, [a..b] \in \left\{\mathcal{E}_s^t \,\middle|\, t \in [T], s \in [S^t]\right\} : \qquad \sum_{\substack{(t,s) \in \mathcal{I}([a..b]) \\ s \geq 1}} \delta_s^t \leq \sum_{t \in [a..b]} \mathrm{z}^{t+1}. \tag{4.4.13}$$

Analogously to Proposition 4.15, it holds for $\hat{P}^{\delta}_{\mathrm{yz}} := \pi^{\delta\to \mathrm{z}}(\hat{P}^{\delta})$ that

$$\hat{P}^{\delta}_{\mathrm{yz}} = \left\{(v, \mathrm{y}, \mathrm{z}, \delta_{\geq 1}) \in \mathbb{R}^{2T+S^{\Sigma}} \text{ fulfilling (4.4.1)–(4.4.6),(4.4.10),(4.4.12),(4.4.13)}\right\}. \tag{4.4.14}$$

Finally, since the structure of inequality (4.3.6) remains unchanged in (4.4.11), adapting the separation algorithm 4.3.1 to P^{δ}_{yz} is straightforward.

4.5 Comparison of Start-up Cost Models

The approach of categorizing start-ups using variables δ^t_s was introduced in [Muc66], and the grouping into start-up types was proposed in [SBB10]. Even if only start-up costs are modeled, the formulation in [SBB10] represents the state of the art (cf. [MELR13b]). This section shows that the inequalities of $\hat{P}^{\delta}$, P^{δ}, both tighten and generalize the inequalities of the existing formulations, and that already $\hat{P}^{\delta}$ improves the bound on the start-up costs.

Moreover, we introduce a formulation based on the start-up and shutdown indicators, which in terms of start-up cost bounds is equivalent to [Muc66; SBB10], but stricter than the epigraph of the start-up costs $\mathrm{conv}(\mathrm{epi}(\mathrm{DCU}^t))$ in a single period and [CA06]. By extension, we are able to compare the tightness of all considered start-up cost formulations in Subsection 4.5.4.

To visualize the difference between these models, we consider their respective cost-minimal point (v, δ) for the fractional operational schedule $v = 1/4(3, 2, 1, 2, 4)$ and with the trivial partition $\mathcal{L}^t_s = \{s\}$ (cf. Fig. 4.6).

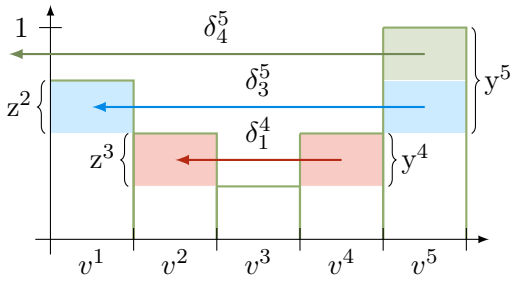


Figure 4.6: A visualization of the point (v, δ) with $v = 1/4(3, 2, 1, 2, 4)$, $\delta^4_1 = \delta^5_3 = \delta^5_4 = 1/4$, $\delta^1_0 = 3/4$, $\delta^2_0 = \delta^5_0 = 1/2$, $\delta^3_0 = \delta^4_0 = 1/4$. The values v^t are represented by bars of corresponding height. The type δ^t_l on the edge (t_+, t'_-) with $t' = t - l - 1$ is depicted by an arrow starting in v^t and ending in $v^{t'}$, and its value is indicated by the part of v^t of the same color. The types δ^t_0, which are not modeled in all compared formulations, and the types δ^t_l with $\delta^t_l = 0$ are not shown.

4.5.1 The Existing Start-up Type Models

The formulation presented in [Muc66] models time-dependent start-up costs without grouping downtimes, i. e. in the same manner as P^f. The parts pertaining to these flows can be summarized as

$$\begin{aligned} &\forall\, t \in [2..T]: && \sum_{l=1}^{t-1} f_l^t = \mathrm{y}^t, \\ &\forall\, t \in [T], l \in [t-1]: && f_l^t \leq \mathrm{z}^{t-l}, \\ &\forall\, t \in [T], l \in [t-1]: && f_l^t \geq 0. \end{aligned}$$

Due to its size, the formulation in [Muc66] is computationally expensive. By grouping the individual start-up variables f to start-up types δ, [SBB10] critically improves the problem size for piece-wise constant start-up cost functions. The resulting start-up type model can be summarized as

$$P_{\mathrm{ex}}^{\delta} := \Big\{(v, \mathrm{y}, \mathrm{z}, \delta_{\geq 1}) \in \mathbb{R}^{2T+S^{\Sigma}} \text{ fulfilling (4.4.1)–(4.4.6), (4.5.1)–(4.5.3)}\Big\}$$

with

$$\forall\, t \in [2..T]: \qquad \sum_{s=1}^{S^t} \delta_s^t = \mathrm{y}^t, \tag{4.5.1}$$

$$\forall\, t \in [T], s \in [S^t - 1]: \qquad \delta_s^t \leq \sum_{l \in \mathcal{L}_s^t} \mathrm{z}^{t-l}, \tag{4.5.2}$$

$$\forall\, t \in [T], s \in [S^t]: \qquad \delta_s^t \geq 0. \tag{4.5.3}$$

When comparing this formulation to the $\mathcal{H}$-representation of $\hat{P}_{\mathrm{yz}}^{\delta}$ in (4.4.14), we note that (4.5.1),(4.5.3) are equivalent to (4.4.10),(4.4.12), and that (4.5.2) is dominated by (4.4.13), since $\mathcal{I}(\mathcal{E}_s^t)$ always contains the tuple (t, s). Therefore, $\hat{P}_{\mathrm{yz}}^{\delta}$ models a stronger bound on the start-up costs than P_{ex}^{δ}, and since $P_{\mathrm{yz}}^{\delta} \subset \hat{P}_{\mathrm{yz}}^{\delta}$, we conclude:

Proposition 4.16

$$P_{\mathrm{yz}}^{\delta} \subset \hat{P}_{\mathrm{yz}}^{\delta} \subset P_{\mathrm{ex}}^{\delta}.$$

The difference between P_{yz}^{δ} and P_{ex}^{δ} is highlighted in Fig. 4.7.

4.5.2 The Epigraph of the Start-up Costs in a Single Period

In Section 2.1, we derived the convex extension LCU^t of the start-up costs DCU^t in period t (see Proposition 2.10) and an $\mathcal{H}$-representation of $\mathrm{epi}(\mathrm{LCU}^t) = \mathrm{conv}(\mathrm{epi}(\mathrm{DCU}^t))$ (see Theorem 2.8),

$$\mathrm{epi}(\mathrm{LCU}^t) = \left\{ \begin{array}{ll} (v, \mathrm{cu}^t) \in \mathbb{R}^{T+1}: & \\ \quad \mathrm{cu}^t \geq \mathrm{CU}^t(l) v^t - \sum_{j=1}^{l} \Big(\mathrm{CU}^t(l) - \mathrm{CU}^t(j-1)\Big) v^{t-j}, & l \in [0\,..\,t-1] \\ \quad 0 \leq v^t \leq 1, & t \in [T] \end{array} \right\}.$$

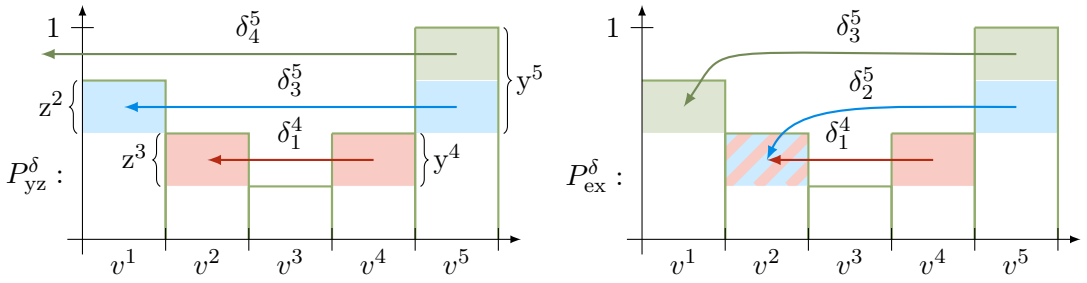


Figure 4.7: Cost-minimizing types δ for $v = {}^1\!/_4(3,2,1,2,4)$ in P_{yz}^{δ} (left) and P_{ex}^{δ} (right), visualized as in Fig. 4.6. While P_{yz}^{δ} bounds $\delta_1^4 + \delta_2^5 \leq z^3$, P_{ex}^{δ} only models $\delta_1^4 \leq z^3$ and $\delta_2^5 \leq z^3$ separately, resulting in a lower bound on the start-up costs. This is indicated by the overlapping rectangles which represent the value of z^3 and z_5^2.

As argued in Section 1.3, epi(LCU^t) models the best possible bound on the start-up costs in period t (if condition (1.3.14) holds). However, since epi(LCU^t) does not consider dependencies between multiple periods, its combination for all periods epi($\mathrm{LCU}^1, \ldots, \mathrm{LCU}^T$) is strictly dominated by conv(epi(DCU)). In this section, we show that epi($\mathrm{LCU}^1, \ldots, \mathrm{LCU}^T$) is also dominated by the projection $\pi^{\delta}(P_{ex}^{\delta})$ of the type polyhedron from [Muc66; SBB10].

To prove this, we present the *start-up cost polyhedron using indicators* P^{yz}, which bounds the start-up costs as strictly as P_{ex}^{δ}, but stricter than epi($\mathrm{LCU}^1, \ldots, \mathrm{LCU}^T$). This polyhedron is defined as

$$P^{yz} := \Big\{(v, y, z, cu) \in \mathbb{R}^{4T} \text{ fulfilling (4.4.1)–(4.4.6) and (4.5.5)}\Big\} \tag{4.5.4}$$

with inequalities (4.4.1)–(4.4.6) describing v, y, z, and inequality

$$\forall\, t \in [T], l \in [0\,..\,t-1]: \qquad cu^t \geq \mathrm{CU}^t(l)\, y^t - \sum_{j=1}^{l-1} (\mathrm{CU}^t(l) - \mathrm{CU}^t(j))\, z^{t-j}. \tag{4.5.5}$$

We start the comparison of P_{ex}^{δ} and P^{yz} by showing that the inequalities of P_{ex}^{δ} dominate (4.5.5) under projection. To this end, we project (v, y, z, cu) to the space of (v, cu),

$$\pi^{yz}: \quad \mathbb{R}^{4T} \to \mathbb{R}^{2T}, \quad (v, y, z, cu) \mapsto (v, cu).$$

Proposition 4.17

$$\pi_{yz}^{\delta}(P_{ex}^{\delta}) + \mathrm{pos}\{u_{T+t} \mid t \in [T]\} \subset \pi^{yz}(P^{yz}).$$

Proof. For each $(v, \mathrm{y}, \mathrm{z}, \delta_{\geq 1}) \in P^{\delta}_{\mathrm{ex}}$, define $(v, \mathrm{cu}) := \pi^{\delta}_{\mathrm{yz}}(v, \mathrm{y}, \mathrm{z}, \delta_{\geq 1})$, i. e.

$$\forall\, t \in [T]: \qquad \mathrm{cu}^t = \sum_{s=1}^{S^t} \mathrm{CU}^t(\min \mathcal{L}^t_s)\delta^t_s + \begin{cases} \mathrm{CU}^1(0)v^1 & \text{if } t = 1, \\ 0 & \text{else.} \end{cases}$$

For $t = 1$, inequality (4.5.5) trivially holds since $\mathrm{cu}^1 = \mathrm{CU}^1(0)v^1 \geq \mathrm{CU}^1(0)\mathrm{y}^1$. For $t \in [2..T]$ and $l \in [t-1]$, choose $s^* \in [S^t]$ such that $l \in \mathcal{L}^t_{s^*}$. Then,

$$\begin{aligned} \mathrm{cu}^t &= \sum_{s=1}^{S^t} \mathrm{CU}^t\Big(\min \mathcal{L}^t_s\Big)\delta^t_s = \mathrm{CU}^t(l) \sum_{s=1}^{S^t} \delta^t_s - \sum_{s=1}^{S^t} \Big(\mathrm{CU}^t(l) - \mathrm{CU}^t\Big(\min \mathcal{L}^t_s\Big)\Big)\delta^t_s \\ &\geq \mathrm{CU}^t(l) \sum_{s=1}^{S^t} \delta^t_s - \sum_{s=1}^{s^*-1} \Big(\mathrm{CU}^t(l) - \mathrm{CU}^t\Big(\min \mathcal{L}^t_s\Big)\Big)\delta^t_s, \end{aligned}$$

due to $\mathrm{CU}^t(l) \leq \mathrm{CU}^t(\min \mathcal{L}^t_s)$ for $s \geq s^*$.

Using inequalities (4.5.1) and (4.5.2), the start-up costs can be bounded by

$$\begin{aligned} \mathrm{cu}^t &\geq \mathrm{CU}^t(l)\mathrm{y}^t - \sum_{s=1}^{s^*-1} \Big(\mathrm{CU}^t(l) - \mathrm{CU}^t\Big(\min \mathcal{L}^t_s\Big)\Big) \sum_{j \in \mathcal{L}^t_s} \mathrm{z}^{t-j} \\ &= \mathrm{CU}^t(l)\mathrm{y}^t - \sum_{s=1}^{s^*-1} \sum_{j \in \mathcal{L}^t_s} (\mathrm{CU}^t(l) - \mathrm{CU}^t(j))\mathrm{z}^{t-j}. \end{aligned}$$

Since the sets $\mathcal{L}^t_1, \ldots, \mathcal{L}^t_{s^*-1}$ partition $[1 \,..\, \max \mathcal{L}^t_{s^*-1}]$ and $\mathrm{CU}^t(l) = \mathrm{CU}^t(j)$ for each $j \in [\max \mathcal{L}^t_{s^*-1}+1 \,..\, l-1] \subset \mathcal{L}^t_{s^*}$, we finally have

$$\mathrm{cu}^t = \mathrm{CU}^t(l)\mathrm{y}^t - \sum_{j=1}^{\max \mathcal{L}^t_{s^*-1}} (\mathrm{CU}^t(l) - \mathrm{CU}^t(j))\,\mathrm{z}^{t-j} = \mathrm{CU}^t(l)\mathrm{y}^t - \sum_{j=1}^{l-1} (\mathrm{CU}^t(l) - \mathrm{CU}^t(j))\,\mathrm{z}^{t-j}.$$

Therefore $(v, \mathrm{y}, \mathrm{z}, \mathrm{cu})$ fulfills (4.5.5). By definition $(v, \mathrm{y}, \mathrm{z}, \mathrm{cu})$ also satisfies (4.4.1)–(4.4.6), and hence $(v, \mathrm{y}, \mathrm{z}, \mathrm{cu}) \in P^{\mathrm{yz}}$.

Since $\pi^{\delta}_{\mathrm{yz}}(v, \mathrm{y}, \mathrm{z}, \delta_{\geq 1}) = (v, \mathrm{cu}) = \pi^{\mathrm{yz}}(v, \mathrm{y}, \mathrm{z}, \mathrm{cu})$, we have $\pi^{\delta}_{\mathrm{yz}}(P^{\delta}_{\mathrm{ex}}) \subset \pi^{\mathrm{yz}}(P^{\mathrm{yz}})$. Given that the start-up costs cu^t are not bounded from above in P^{yz}, this implies $\pi^{\delta}_{\mathrm{yz}}(P^{\delta}_{\mathrm{ex}}) + \mathrm{pos}\{u_{T+t} \mid t \in [T]\} \subset \pi^{\mathrm{yz}}(P^{\mathrm{yz}})$. □

The following proposition shows the converse direction, i. e. that P^{yz} models the same lower boundary on the start-up costs as P^{δ}_{ex}, by explicitly constructing $\delta_{\geq 1}$ with minimal start-up costs for a given $(v, \mathrm{y}, \mathrm{z})$ such that $(v, \mathrm{y}, \mathrm{z}, \delta_{\geq 1}) \in P^{\delta}_{\mathrm{ex}}$.

Proposition 4.18

$$\pi^{\delta}_{\mathrm{yz}}(P^{\delta}_{\mathrm{ex}}) + \mathrm{pos}\{u_{T+t} \mid t \in [T]\} = \pi^{\mathrm{yz}}(P^{\mathrm{yz}}).$$

Proof. For each $(v, \mathrm{y}, \mathrm{z}, \mathrm{cu}) \in P^{\mathrm{yz}}$ define $\delta^t_{\geq 1}$ as

$$\forall\, s \in [S^t - 1]: \qquad \delta^t_s := \min\Big\{ \sum_{l \in \mathcal{L}^t_s} \mathrm{z}^{t-l}, \max\Big\{0, \mathrm{y}^t - \sum_{s'=1}^{s-1} \delta^t_{s'}\Big\}\Big\},$$

$$\delta^t_{S^t} := \max\Big\{0, \mathrm{y}^t - \sum_{s=1}^{S^t-1} \delta^t_s\Big\}.$$

The point $(v, \mathrm{y}, \mathrm{z}, \delta_{\geq 1})$ satisfies the inequalities (4.4.1)–(4.4.6) for v, y, z by definition of P^{yz}, and the flow inequalities (4.5.1)–(4.5.3) of P^{δ}_{ex} by definition of $\delta^t_{\geq 1}$. Hence, $(v, \mathrm{y}, \mathrm{z}, \delta_{\geq 1}) \in P^{\delta}_{\mathrm{ex}}$.

We continue by showing that the start-up costs $(v, \tilde{\mathrm{cu}}) := \pi^{\delta}_{\mathrm{yz}}(v, \mathrm{y}, \mathrm{z}, \delta_{\geq 1})$ induced by $\delta_{\geq 1}$ are at most cu. The start-up costs $\mathrm{CU}^1(0)$ in period 1 are positive iff the unit was offline prior to the model time, i. e. PDT > 0, which implies $\mathrm{y}^1 = v^1$. Thus,

$$\tilde{\mathrm{cu}}^1 = \mathrm{CU}^1(0) v^1 = \mathrm{CU}^1(0) \mathrm{y}^1 \leq \mathrm{cu}^1.$$

For each $t \in [2..T]$, choose $s^* \in [S^t]$ maximal such that $\delta^t_{s^*} > 0$, and $l := \min \mathcal{L}^t_{s^*}$. Then, the start-up costs $\tilde{\mathrm{cu}}^t$ can be bounded by

$$\begin{aligned}\tilde{\mathrm{cu}}^t &= \sum_{s=1}^{S^t} \mathrm{CU}^t\Big(\min \mathcal{L}^t_s\Big) \delta^t_s = \sum_{s=1}^{s^*} \mathrm{CU}^t(\min \mathcal{L}^t_s) \delta^t_s \\ &= \mathrm{CU}^t(l) \sum_{s=1}^{s^*} \delta^t_s - \sum_{s=1}^{s^*-1} \Big(\mathrm{CU}^t(l) - \mathrm{CU}^t\Big(\min \mathcal{L}^t_s\Big)\Big) \delta^t_s\end{aligned}$$

By the choice of s^* and definition of δ^t_s, it holds that

$$\sum_{s=1}^{s^*} \delta^t_s = \mathrm{y}^t \qquad \text{and} \qquad \forall\, s \in [s^* - 1]: \quad \delta^t_s = \sum_{j \in \mathcal{L}^t_s} \mathrm{z}^{t-j}.$$

Hence, the above bound on $\tilde{\mathrm{cu}}^t$ can be expressed as

$$\tilde{\mathrm{cu}}^t = \mathrm{CU}^t(l) \mathrm{y}^t - \sum_{s=1}^{s^*-1} \Big(\mathrm{CU}^t(l) - \mathrm{CU}^t\Big(\min \mathcal{L}^t_s\Big)\Big) \sum_{j \in \mathcal{L}^t_s} \mathrm{z}^{t-j}$$

Since $\mathrm{CU}^t(j) = \mathrm{CU}^t(\min \mathcal{L}^t_s)$ for each $j \in \mathcal{L}^t_s$, and since the sets $\mathcal{L}^t_1, \ldots, \mathcal{L}^t_{s^*-1}$ partition $[1 \,..\, \max \mathcal{L}^t_{s^*-1}]$ and $\max \mathcal{L}^t_{s^*-1} = \min \mathcal{L}^t_{s^*} - 1 = l - 1$, we have

$$\begin{aligned}\tilde{\mathrm{cu}}^t &= \mathrm{CU}^t(l) \mathrm{y}^t - \sum_{s=1}^{s^*-1} \sum_{j \in \mathcal{L}^t_s} (\mathrm{CU}^t(l) - \mathrm{CU}^t(j)) \mathrm{z}^{t-j} \\ &= \mathrm{CU}^t(l) \mathrm{y}^t - \sum_{j=1}^{l-1} (\mathrm{CU}^t(l) - \mathrm{CU}^t(j))\, \mathrm{z}^{t-l} \leq \mathrm{cu}^t.\end{aligned}$$

In conclusion, $\pi^{\delta}_{\mathrm{yz}}(v, \mathrm{y}, \mathrm{z}, \delta_{\geq 1}) = (v, \tilde{\mathrm{cu}}) \leq (v, \mathrm{cu}) = \pi^{\mathrm{yz}}(v, \mathrm{y}, \mathrm{z}, \mathrm{cu})$, and therefore $\pi^{\mathrm{yz}}(P^{\mathrm{yz}}) \subset \pi^{\delta}_{\mathrm{yz}}(P^{\delta}_{\mathrm{ex}}) + \mathrm{pos}\{u_{T+t} \,|\, t \in [T]\}$. □

Having proved that P^{δ}_{ex} and P^{yz} model the same lower bound on the start-up costs, we now compare P^{yz} to the epigraph $\mathrm{epi}(\mathrm{LCU}^t)$ of the individual start-up cost function DCU^t. In Theorem 2.8, we showed that $\mathrm{epi}(\mathrm{LCU}^t)$ equals

$$\mathrm{epi}(\mathrm{LCU}^t) = \{(v, \mathrm{cu}^t) \in \mathbb{R}^{T+1} \text{ fulfilling } v \in [0,1]^T \text{ and } (2.1.4)\},$$

where inequality (2.1.4) is defined as

$$\forall\, l \in [0 \,..\, t-1]: \qquad \mathrm{cu}^t \geq \mathrm{CU}^t(l) v^t - \sum_{j=1}^{l} (\mathrm{CU}^t(l) - \mathrm{CU}^t(j-1))\, v^{t-j}.$$

The epigraphs $\mathrm{epi}(\mathrm{LCU}^t)$ can be combined for all t, yielding a model of the start-up costs in all periods,

$$\begin{aligned}\mathrm{epi}(\mathrm{LCU}^1, \ldots, \mathrm{LCU}^T) &= \Big\{(v, \mathrm{cu}) \in \mathbb{R}^{2T} \;\Big|\; v \in [0,1]^T, \forall\, t \in [T] : \overbrace{\mathrm{cu}^t \geq \mathrm{LCU}^t(v)}^{\Leftrightarrow\ (v, \mathrm{cu}^t) \in \mathrm{epi}(\mathrm{LCU}^t)}\Big\} \\ &= \Big\{(v, \mathrm{cu}) \in \mathbb{R}^{2T} \text{ fulfilling } v \in [0,1]^T \text{ and } (2.1.4) \text{ for } t \in [T]\Big\}.\end{aligned}$$

Due to the similar structure of their inequalities (4.5.5) and (2.1.4), it is straightforward to show that P^{yz} dominates $\mathrm{epi}(\mathrm{LCU}^1, \ldots, \mathrm{LCU}^T)$ in terms of start-up costs:

Proposition 4.19

$$\pi^{\mathrm{yz}}(P^{\mathrm{yz}}) \subset \mathrm{epi}(\mathrm{LCU}^1, \ldots, \mathrm{LCU}^T).$$

Proof. For each $(v, \mathrm{y}, \mathrm{z}, \mathrm{cu}) \in P^{\mathrm{yz}}$, $t \in [T]$, and $l \in [0 \,..\, t-1]$, we have

$$\mathrm{cu}^t \geq \mathrm{CU}^t(l) \mathrm{y}^t - \sum_{j=1}^{l-1} (\mathrm{CU}^t(l) - \mathrm{CU}^t(j))\, \mathrm{z}^{t-j}.$$

Using $\mathrm{y}^t = v^t - v^{t-1} - \mathrm{z}^t \geq v^t - v^{t-1}$ and $\mathrm{z}^t = \mathrm{y}^t - v^t + v^{t-1} \leq v^{t-1}$, these start-up costs cu^t are bounded by

$$\begin{aligned}\mathrm{cu}^t &\geq \mathrm{CU}^t(l)(v^t - v^{t-1}) - \sum_{j=1}^{l-1} (\mathrm{CU}^t(l) - \mathrm{CU}^t(j))\, v^{t-j-1} \\ &= \mathrm{CU}^t(l) v^t - \sum_{j=1}^{l} (\mathrm{CU}^t(l) - \mathrm{CU}^t(j-1))\, v^{t-j}.\end{aligned}$$

Therefore $\pi^{\mathrm{yz}}(v, \mathrm{y}, \mathrm{z}, \mathrm{cu}) = (v, \mathrm{cu}) \in \mathrm{epi}(\mathrm{LCU}^1, \ldots, \mathrm{LCU}^T)$. □

On a side note, analogously to the relationship between P^{yz} and P^{δ}_{ex}, there exists an extended formulation of $\mathrm{epi}(\mathrm{LCU}^1, \dots, \mathrm{LCU}^T)$ using start-up types. The same approach as in the proofs of Propositions 4.17 and 4.18 with z^{t-l} replaced by v^{t-l-1} shows that

$$\begin{aligned}&\mathrm{epi}(\mathrm{LCU}^1, \dots, \mathrm{LCU}^T) = \\ &\quad = \pi^{\delta}_{\mathrm{yz}}\left(\left\{ \begin{array}{l} (v,\mathrm{y},\mathrm{z},\delta_{\geq 1}) \in \mathbb{R}^{3T+(S^{\Sigma}-T)} \\ \text{fulfilling (4.4.1)-(4.4.6),(4.5.6)–(4.5.8)} \end{array} \right\}\right) + \mathrm{pos}\{u_{T+t} \mid t \in [T]\}\end{aligned}$$

with inequalities

$$\forall\, [2..T]: \qquad \sum_{s=1}^{S^t} \delta^t_s = \mathrm{y}^t, \tag{4.5.6}$$

$$\forall\, t \in [T], s \in [S^t - 1]: \qquad \delta^t_s \leq \sum_{l \in \mathcal{L}^t_s} v^{t-l-1}, \tag{4.5.7}$$

$$\forall\, t \in [T], s \in [S^t]: \qquad \delta^t_s \geq 0. \tag{4.5.8}$$

Using this representation, the difference between P^{δ}_{yz}, P^{yz}, and $\mathrm{epi}(\mathrm{LCU}^1, \dots, \mathrm{LCU}^T)$ is highlighted in Fig. 4.8 by considering cost-minimizing types in the three models.

4.5.3 The Temperature Polyhedron for Integers

The formulations presented in Chapter 3, which are based on explicitly modeling the temperature of a unit, follow an entirely different approach than the start-up type formulations. This is reflected by the fact that in general, $\hat{P}^{\mathrm{temp}}$ can be compared to neither P^t_{ex}, $\mathrm{epi}(\mathrm{LCU}^t)$, nor P^{δ}_{ex}. We show this by giving examples of points which have strictly highest start-up costs in $\hat{P}^{\mathrm{temp}}$ and P^{δ}_{ex}, respectively.

Consider the exponential start-up cost function $\mathrm{CU}(t) = (1 - \mathrm{e}^{-\ln(2)t}) + 1$ in a model with uniform period lengths $\mathrm{L}^t = 1$. Then, $\hat{P}^{\mathrm{temp}}$ models the temperature temp^t, heating h^t, and start-up indicators y^t as

$$\mathrm{temp} = (1, 1, 1/2, 1/4), \quad \mathrm{h} = (0, 0, 0, 0), \quad \mathrm{y} = (0, 0, 0, 1/2),$$

for the operational schedule $v = (1, 0, 0, 1/2)$, resulting in the costs $\mathrm{cu} = (0, 0, 0, 1/2)$. For the same schedule however P^t_{ex} yields $\mathrm{cu} = (0, 0, 0, {}^{\mathrm{CU}^4(2)}/_2) = (0, 0, 0, 3/8) > (0, 0, 0, 1/2)$. Therefore, in general

$$\pi^{\mathrm{temp}}(\hat{P}^{\mathrm{temp}}) \not\subset P^t_{\mathrm{ex}}, \tag{4.5.9}$$

which implies $\pi^{\mathrm{temp}}(\hat{P}^{\mathrm{temp}}) \not\subset \mathrm{epi}(\mathrm{LCU}^t)$ and $\pi^{\mathrm{temp}}(\hat{P}^{\mathrm{temp}}) \not\subset P^{\delta}_{\mathrm{ex}}$.

For the operational schedule $v = (1, 1/2, 1/2, 0, 1/2, 1)$, $\hat{P}^{\mathrm{temp}}$ models

$$\mathrm{temp} = (1, 1, 3/4, 5/8, 1/2, 1), \quad \mathrm{h} = (0, 0, 0, 0, 3/16, 1/2), \quad \mathrm{y} = (0, 0, 0, 0, 1/2, 1/2),$$

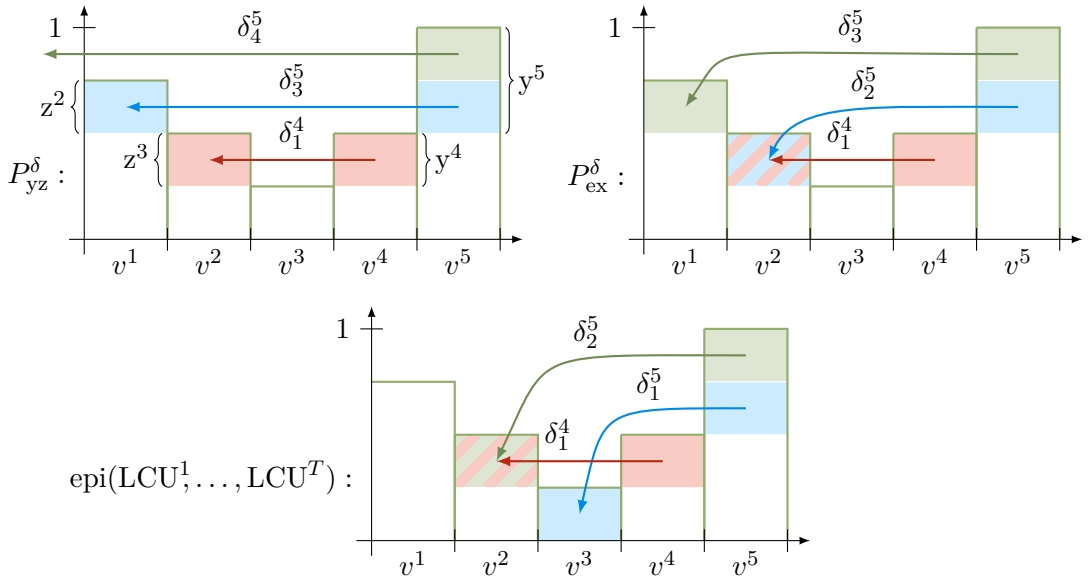


Figure 4.8: Cost-minimizing types for $v = {}^1/_4(3,2,1,2,4)$ in the type model P^δ_{yz} (top left), the existing type model P^δ_{ex} (top right), and in the type model equivalent to epi$(\text{LCU}^1,\dots,\text{LCU}^T)$ (bottom), visualized as in Fig. 4.6. While P^δ_{yz} bounds $\delta_1^4 + \delta_2^5 \leq \text{z}^3$, P^δ_{ex} only models $\delta_1^4 \leq \text{z}^3$ and $\delta_2^5 \leq \text{z}^3$ separately, and epi$(\text{LCU}^1,\dots,\text{LCU}^T)$ only models $\delta_1^4 \leq v^2$ and $\delta_1^5 \leq v^3$, resulting in a lower bound on the start-up costs.

leading to a total start-up cost of $\text{cu}^\Sigma = {}^{27}/_{16}$. The start-up costs in P^δ_{ex} however equate to $\text{cu} = (0,0,0,0,{}^{\text{CU}^5(1)}/_2, {}^{\text{CU}^6(2)}/_2) = (0,0,0,0,)$ such that $\text{cu}^\Sigma = {}^{\text{CU}^5(1)}/_2 + {}^{\text{CU}^6(2)}/_2 = {}^3/_4 + {}^7/_8 = {}^{13}/_8 = {}^{26}/_{16} < {}^{27}/_{16}$. So, in general

$$\pi^\delta(P^\delta_{\text{ex}}) \not\subset \pi^{\text{temp}}(\hat{P}^{\text{temp}}), \tag{4.5.10}$$

which analogously demonstrates $P^t_{\text{ex}} \not\subset \pi^{\text{temp}}(\hat{P}^{\text{temp}})$ and $\text{epi}(\text{LCU}^t) \not\subset \pi^{\text{temp}}(\hat{P}^{\text{temp}})$.

4.5.4 Conclusion

In this chapter and in Section 2.1, we have considered the existing start-up cost models

- P^t_{ex} is the existing step-wise start-up cost polyhedron presented in [CA06] (see Subsection 2.1.1) and
- P^δ_{ex} is the existing start-up type polyhedron presented in [SBB10] (see Subsection 4.5.1),

and have introduced the novel models

- $\text{epi}(\text{LCU}^1, \dots, \text{LCU}^T)$ is the combination of $\text{epi}(\text{LCU}^t)$ for all $t \in [T]$,
- P^δ is the start-up type polyhedron in Section 4.3 and P^δ_{yz} is its variation using indicators in Proposition 4.15,
- $\hat{P}^\delta$ is the start-up type polyhedron for integers in Subsection 4.3.2 and $\hat{P}^\delta_{\text{yz}}$ is its variation using indicators in Proposition 4.15, and
- P^{yz} is the start-up cost polyhedron using indicators in Subsection 4.5.2.

Combining the results in Theorem 4.10, Proposition 4.16, Proposition 4.18, and Proposition 4.19 allows us to compare the tightness of all of these formulations.

Theorem 4.20 *Defining* $R := \text{pos}\{u_{T+t} \mid t \in [T]\}$,

$$\begin{aligned}\text{conv}(\text{epi}(\text{DCU})) &= \pi^\delta(P^\delta) + R = \pi^\delta_{\text{yz}}(P^\delta_{\text{yz}}) + R \\ &\subset \pi^\delta(\hat{P}^\delta) + R = \pi^\delta_{\text{yz}}(\hat{P}^\delta_{\text{yz}}) + R \\ &\subset \pi^\delta_{\text{yz}}(P^\delta_{\text{ex}}) + R = \pi^{\text{yz}}(P^{\text{yz}}) \\ &\subset \text{epi}(\text{LCU}^1, \dots, \text{LCU}^T) \subset P^t_{\text{ex}},\end{aligned}$$

where conv(epi(DCU)) *is the epigraph of the start-up costs in all periods (cf. Section 2.3), and* π^δ, π^δ_{yz}, *and* π^{yz} *denote the respective projections to the space of* conv(epi(DCU)).

Note that except for P^δ and P^δ_{yz}, the newly introduced tighter formulations do not introduce further variables and inequalities, but strengthen the inequalities of the existing models (cf. Table 4.1).

Model	Source	Variables	Inequalities
P^f	(4.2.5)	$\mathcal{O}(T^2)$	$\mathcal{O}(T)$
P^δ, P^δ_{yz}	(4.3.10), (4.4.9)	$\mathcal{O}(ST)$	$\mathcal{O}(T^2)$
$\hat{P}^\delta$, $\hat{P}^\delta_{\text{yz}}$	(4.3.8), (4.4.14)	$\mathcal{O}(ST)$	$\mathcal{O}(ST)$
P^δ_{ex}	[Muc66]	$\mathcal{O}(ST)$	$\mathcal{O}(ST)$
P^{yz}	(4.5.4)	$\mathcal{O}(T)$	$\mathcal{O}(ST)$
$\text{epi}(\text{LCU}^1, \dots, \text{LCU}^T)$	Theorem 2.8	$\mathcal{O}(T)$	$\mathcal{O}(ST)$
P^t_{ex}	[CA06]	$\mathcal{O}(T)$	$\mathcal{O}(ST)$

Table 4.1: Sizes of step-wise start-up cost formulations and start-up type formulations, in order of decreasing tightness.

Chapter 5

Numerical Experiments

This section compares the integrality gap and the computational performance of the proposed and existing start-up cost and start-up process models by embedding them into the Unit Commitment model presented in Subsection 1.2.2.

The main findings can be summarized as:

- The temperature models $\hat{P}^{\text{temp}}$ and P^{temp}, and the start-up type models $\hat{P}^{\delta}$ and P^{δ} considerably outperform the state of the art.
- Approximating the start-up cost function by a step-wise function with tolerance CU_{tol} (cf. Subsection 2.1.6) reduces the number of start-up types and therefore the model size of $\hat{P}^{\delta}$, but increases the integrality gap. Out of $\text{CU}_{\text{tol}} \in \{5\%, 10\%, 20\%\}$, $\hat{P}^{\delta}$ and P^{δ} perform best for $\text{CU}_{\text{tol}} = 10\%$.
- While the RTIs in P^{temp} significantly reduce the integrality gap of $\hat{P}^{\text{temp}}$, their separation is beneficial only in some cases.

All tests have been performed with the FICO Xpress Optimizer (see [Fic]) and with the separation of cutting planes for P^{temp} and P^{δ} implemented in FICO Xpress Mosel.

5.1 The Scenarios

As noted in the introduction, the main cause of the increasing number of start-ups is the volatile production from renewable energy sources. To highlight the computational viability of the proposed start-up models, even for future scenarios with higher renewable energy penetration, we perform the numerical experiments both on real-world data for 2011 and on a forecast for 2025, similar to the scenarios proposed in [SHB16].

All models use a uniform period length of one hour, i. e. $\text{L}^1 = \ldots = \text{L}^T = 1$. The pre-model uptimes POT_i are defined to be equal to the minimum uptime UT_i, such that each unit is able to shut down in the first period. To keep the computational effort reasonable, each problem instance only considers a limited number of periods $T \leq 540$, corresponding to a maximal duration of 22.5 days. We retain the diversity of the years 2011 and 2025 by repeating each experiment for 14 different time ranges starting in the S-th hour of the year with

$$S \in \{624k + 433 : k \in [0..13]\}. \tag{5.1.1}$$

By calculating the start points S using the coefficients 624 and 433,

- the modeled time ranges are distributed uniformly over the years 2011 and 2025 (since $\lfloor {}^{365}/_{14} \rfloor \approx 26 = {}^{624}/_{24}$),
- each time range starts at midnight (since $433 = 18 \cdot 24 + 1$),
- the atypical Christmas and New Year holiday is part of a single time range, and
- two time ranges start on any day of the week, respectively (since ${}^{624}/_{24} = 26$ and 7 are relatively prime).

The power plant data is based on the publication [Ger14] by the German Federal Network Agency for the year 2014, consisting of 228 individually controlled power plants. Observing the plans for a nuclear phase-out in Germany, all 9 remaining nuclear power plants are removed in the 2025 scenario and replaced by 4 combined cycle gas units, resulting in a total of 223 power plants.

As our UC model requires parameters not contained in [Ger14], in particular the minimal production, efficiency, and start-up costs, we complement it based on the results reported in [Kum+12; EUR03; Ege+14]. The start-up costs are defined as in (1.3.1),

$$\forall\, i \in \mathcal{I}, t \in [T], l \in [t-1]: \qquad \mathrm{CU}_i^t(l) := \mathrm{CU}_i^{\mathrm{var}}(1 - \mathrm{e}^{-\lambda_i l}) + \mathrm{CU}_i^{\mathrm{fixed}},$$

with the fixed start-up costs $\mathrm{CU}_i^{\mathrm{fixed}}$, the maximal variable start-up costs $\mathrm{CU}_i^{\mathrm{var}}$ and the heat loss rate λ_i. The heat loss rates depend on the unit type and are chosen as $\lambda_i = 0.03$ for lignite and nuclear units, $\lambda_i = 0.05$ for coal units, $\lambda_i = 0.1$ for combined cycle gas units, and $\lambda_i = 0.3$ for gas units.

In some experiments, these start-up costs are approximated by Algorithm 2.1.2 to a tolerance of $\mathrm{CU}_{\mathrm{tol}} \in \{5\%, 10\%, 20\%\}$, which results in smaller problem sizes in the considered formulations.

The start-up time is defined as

$$\mathrm{TU}_i^t(l) = \max\Big\{1, \Big\lceil A_i(1 - \mathrm{e}^{-\lambda_i l}) - 0.5 \Big\rceil\Big\},$$

with parameters A_i fitted to the start-up times given in [SBB10, Table II]. The start-up production is modeled by power trajectory (1.3.5) with parameters $\underline{\mathrm{PU}}_i = 0$ and $\overline{\mathrm{PU}}_i = \underline{\mathrm{P}}_i$. This yields the linear power trajectory from 0 to $\underline{\mathrm{P}}_i$ at start-up, which is typically used in the literature (see e. g. [SBB10; MELR13a]).

Renewable energy sources which cannot be scheduled, i. e. wind and solar energy, do not need to be modeled as units in the Unit Commitment problem. Given the short time span modeled in the following experiments (at most 10 days), their production is assumed to be known to a sufficient accuracy from weather forecasts and their contribution is included by defining the demand D^t as the difference between the total demand and the production from these renewables. Hourly demand data is provided

by ENTSO-E [ENT] and scaled to a yearly electricity consumption of 520 TWh. Wind and solar electricity generation profiles are computed based on the weather data in the NASA MERRA database [Rie+11],

- on the actual installed capacity in 2011 (29 GW wind, 23 GW solar, 3 GW biomass, 4 GW hydro), and
- on the forecast installed capacity for 2025 (50 GW wind, 50 GW solar, 5.5 GW biomass, 4.5 GW hydro).

Biomass and hydro units contribute relatively little to the fluctuation of the renewable power production. We simplify the model by assuming that these units constantly produce at full capacity.

In a small fraction of the periods, the production from renewable energy sources exceeds the demand. We assume that the renewable production is restricted such that the residual demand D^t fulfills

- $\mathrm{D}^t \geq 0$ MWh in models without start-up production, and
- $\mathrm{D}^t \geq 1000$ MWh in models with start-up production, to ensure that an adequate number of units is able to start up even after an exceptionally low residual demand without causing overproduction.

5.2 The Models

In this thesis, we have examined the two existing start-up cost models,

- the step-wise start-up cost polyhedron P^t_{ex} presented in [CA06] (see Subsection 2.1.1), and
- the start-up type model P^δ_{ex} presented in [SBB10] (see Subsection 4.5.1).

Moreover, we have proposed some new start-up cost models,

- the epigraphs $\mathrm{conv}(\mathrm{epi}(\mathrm{DCU}^t))$ for all $t \in [T]$ in Theorem 2.8, denoted by P^t,
- the temperature polyhedron for integers $\hat{P}^{\mathrm{temp}}$ in Definition 3.4,
- the temperature polyhedron P^{temp} in Definition 3.17,
- the start-up cost polyhedron P^{yz} using indicators in Subsection 4.5.2,
- the start-up type polyhedron for integers $\hat{P}^\delta$ in Subsection 4.3.2, and
- the start-up type polyhedron P^δ in Section 4.3.

We embed these approaches into the Unit Commitment formulation summarized in Subsection 1.2.2, where the placeholders cu_i are substituted by the start-up costs as defined in the respective start-up cost models.

P^{temp} and P^δ are implemented using Algorithm 3.3.1 and Algorithm 4.3.1 in a Branch&Cut approach. These separation algorithms are modified to select the Q most

violated inequalities such that at most Q_i apply to a single unit. We use the values $Q = 0.5|\mathcal{I}|$ and $Q_i = 0.15\,T$ for P^{temp}, and $Q = |\mathcal{I}|$ and $Q_i = 0.15\,T$ for P^{δ}, which lead to an adequate computational performance in the considered test cases.

Subsection 1.3.3 notes that the start-up type variables $\delta^t_{i,s}$ of P^{δ}_{ex}, $\hat{P}^{\delta}$, and P^{δ} provide the basis of a start-up process formulation, and that the start-up production can be modeled by replacing the demand equation (1.2.8)

$$\forall\, t \in [T]: \qquad \sum_{i \in \mathcal{I}} p^t = \mathrm{D}^t$$

of the UC formulation by (1.3.15),

$$\forall\, t \in [T]: \qquad \sum_{i \in \mathcal{I}} \Big(p^t_i + \sum_{t' \in [t+1\,..\,T], s \in [S^{t'}]} \mathrm{PU}^t_i(t', s)\, \delta^{t'}_{i,s} \Big) = \mathrm{D}^t.$$

We assess the impact of the start-up process by considering all start-up type models with and without start-up production.

5.3 Integrality Gap

In practice, Unit Commitment problems are not solved to optimality due to excessive computational effort, as are most Mixed Integer Programs. Instead, state-of-the-art solvers find a feasible commitment with a cost within a given tolerance of the minimal cost in two steps:

- First, they gain a lower bound on the cost by solving the linear relaxation of the UC formulation,
- which is then iteratively refined and complemented with feasible commitments.

The gap of the initial linear relaxation to the optimal solution is called the *integrality gap* and the target gap of the iterative refinement is called *optimality tolerance*. This optimality tolerance controls the trade-off between the solution quality and computation time. On one hand, it limits the possible financial loss due to non-optimality: In our model of Germany 2014, the electricity production in an average week causes a cost of more than 170 million euro. An optimality tolerance of 1% means that the accepted solution may cost as much as 1.7 million euro more than necessary. Hence, this tolerance should be selected as low as possible. On the other hand, increasing the ratio between integrality gap and optimality tolerance sharply increases the computational effort of the iterative refinement.

The integrality gap is thus a major driver of the computational effort and it is essential to use a formulation with a low integrality gap. In the following, we experimentally compare the integrality gaps of the presented models for 14 instances with the start

points S as given in (5.1.1) and $T = 96$ periods (4 days). In the following figures, we highlight the newly introduced models in red.

As proved in Theorem 4.20, all experimental results yield

- equal integrality gaps when using the start-up cost models P^{yz} and P^{δ}_{ex}, which we denote by "P^{δ}_{ex}" in the following charts, and
- equal integrality gaps when using the start-up cost models P^{δ} and P^{temp}, which we denote by "P^{δ}".

Moreover, since we approximate the start-up cost function only in the experiment shown in Fig. 5.4, $\hat{P}^{\delta}$ degenerates into P^{δ} in the remaining figures Fig. 5.1–5.3. Also recall that we denote $\text{epi}(\text{LCU}^1, \ldots, \text{LCU}^T)$ by "P^t".

Figure 5.1 compares the integrality gap in scenario 2011 (left chart) and scenario 2025 (right chart). Due to the higher penetration of renewable energy in 2025, more start-ups and ramping are necessary, leading to a 2–4 times higher integrality gap. The higher number of start-ups increases the spread between lowest and highest median integrality gaps from 1.2 percentage points in 2011 to 2.6 percentage points in 2025. In both cases the significant impact of the start-up cost model is demonstrated.

We consider the relative differences between the formulations in the following. Examining the integrality gaps relative to P^{δ}_{ex} (see Fig. 5.2) confirms Theorem 4.20 and the experimental results of [MELR13b], which state that the existing start-up type polyhedron P^{δ}_{ex} in [SBB10] dominates the step-wise model P^{t}_{ex} in [CA06]. The tightening of P^{t}_{ex} to $P^t = \text{epi}(\text{LCU}^1, \ldots, \text{LCU}^T)$ (inequalities (2.1.2) to (2.1.4)) does not change this assessment.

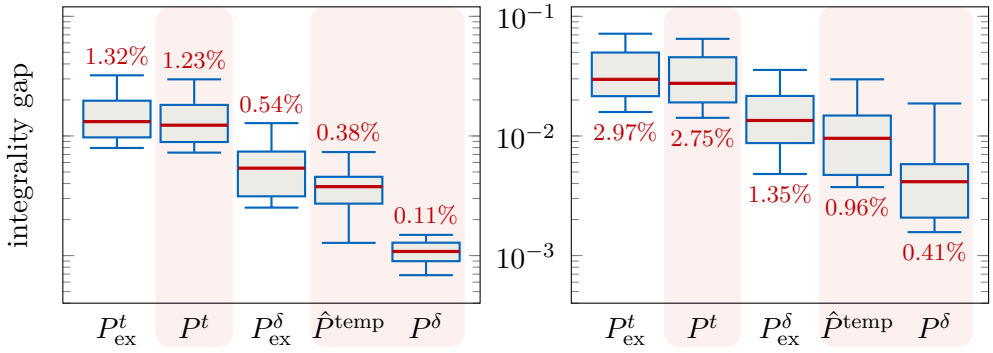


Figure 5.1: Integrality gap in scenarios 2011 (left) and 2025 (right) represented in a box plot: For each model the median is represented by a red line and number, the range from first to third quartile is spanned by a gray box, and the minimum and maximum values are marked by whiskers.
The median integrality gaps in 2025 are higher than in 2011 by a factor of 2–4. P^{δ} always outperforms the other models (see Fig. 5.3).

Still, the state-of-the-art P^δ_{ex} is outperformed by the temperature polyhedron for integers $\hat{P}^{\text{temp}}$, which reduces the median integrality gap by further 24–28%. While $\hat{P}^{\text{temp}}$ is typically tighter than P^δ_{ex}, Fig. 5.3 demonstrates that there are examples were P^δ_{ex} strictly dominates $\hat{P}^{\text{temp}}$. Albeit they did not occur in our numerical experiments, it is easy to construct instances where P^t_{ex} strictly dominates $\hat{P}^{\text{temp}}$. Consider the UC problem with a single unit, $T = 2$, PDT $= 0$, and demand $\mathrm{D} = (0, 1/2\overline{\mathrm{P}}_1)$, which has an optimal fractional solution with $v_1 = (0, 1/2)$. If $\lambda_1 < \ln(1/2)$, then no heating in period 1 is necessary, and the start-up costs in $\hat{P}^{\text{temp}}$ equal $1/2\mathrm{CU}_1^{\text{fixed}}$. On the other hand, the unit incurs start-up costs of $1/2\mathrm{CU}_1^2(1) = 1/2((1 - \mathrm{e}^{-\lambda \mathrm{L}^1})\mathrm{CU}_1^{\text{var}} + \mathrm{CU}_1^{\text{fixed}}) > 1/2\mathrm{CU}_1^{\text{fixed}}$ in P^t_{ex}, strictly dominating $\hat{P}^{\text{temp}}$.

The difference of $\hat{P}^{\text{temp}}$ and P^{temp} (corresponding to "P^δ") caused by the residual temperature inequalities defined in (3.3.1) is considerable. Figure 5.3 points out that the RTIs reduce the median integrality gap by 55–69%.

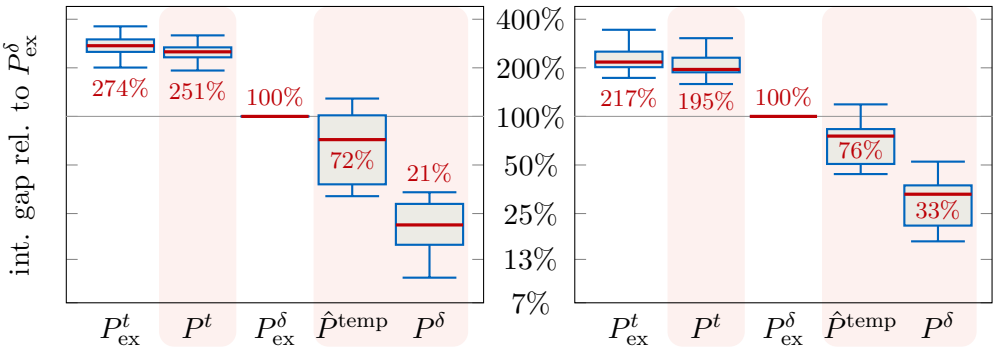


Figure 5.2: Integrality gap relative to P^δ_{ex} in scenarios 2011 (left) and 2025 (right). P^δ_{ex} outclasses P^t_{ex} and $P^t = \text{epi}(\text{LCU}^1,\ldots,\text{LCU}^T)$, but is typically dominated by $\hat{P}^{\text{temp}}$.

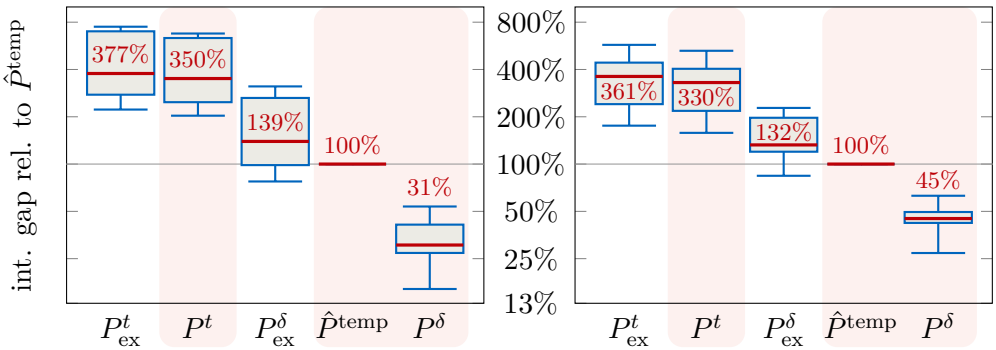


Figure 5.3: Integrality gap relative to $\hat{P}^{\text{temp}}$ in scenarios 2011 (left) and 2025 (right). While $\hat{P}^{\text{temp}}$ typically dominates the state-of-the-art P^δ_{ex}, the distance to P^{temp} and P^δ is considerable.

As the start-up costs increase strictly, $\hat{P}^\delta$ degenerates into P^δ. We therefore contrast the integrality gaps for tolerances $\mathrm{CU}_{\mathrm{tol}} \in \{5\%, 10\%, 20\%\}$ in the scenario 2025 (see Fig. 5.4). The tightened inequalities in $\hat{P}^\delta$ significantly reduce the integrality gap in comparison to P^δ_{ex}. Note however that the ratio between $\hat{P}^\delta$ and P^δ strongly correlates with $\mathrm{CU}_{\mathrm{tol}}$: While for $\mathrm{CU}_{\mathrm{tol}} = 5\%$ only 9% separate P^δ and $\hat{P}^\delta$, the difference grows to 72% for $\mathrm{CU}_{\mathrm{tol}} = 20\%$. At the same time, the advantage of $\hat{P}^\delta$ over P^δ_{ex} diminishes. In contrast, the relative integrality gaps of P^t_{ex} and P^t decrease consistently.

Finally, Fig. 5.5 highlights the impact of the start-up production. While the integrality gaps consistently decrease, their relative difference for P^δ_{ex}, $\hat{P}^\delta$, and P^δ_{yz} stays similar. This is explained by the fact that the start-up production effectively reduces the start-up costs: For each unit which produces during start-up, the other units ramp down and thereby decrease their production costs.

5.4 Computational Performance

This subsection analyzes the computational performance of each model by determining the maximal number of periods T for each start point which the model solves to a certain optimality gap within a given time limit.

We consider two common solver configurations,

1. a time limit of 15 minutes and an optimality tolerance of 1% for scenario 2011, and
2. a time limit of 30 minutes and an optimality tolerance of 0.5% for scenario 2025.

In scenario 2011, all models fail to solve some instances due to poor upper bounds, even for low values of T. The reason is that the FICO Xpress Optimizer separates cutting planes at the root node before applying heuristics, and sometimes does not

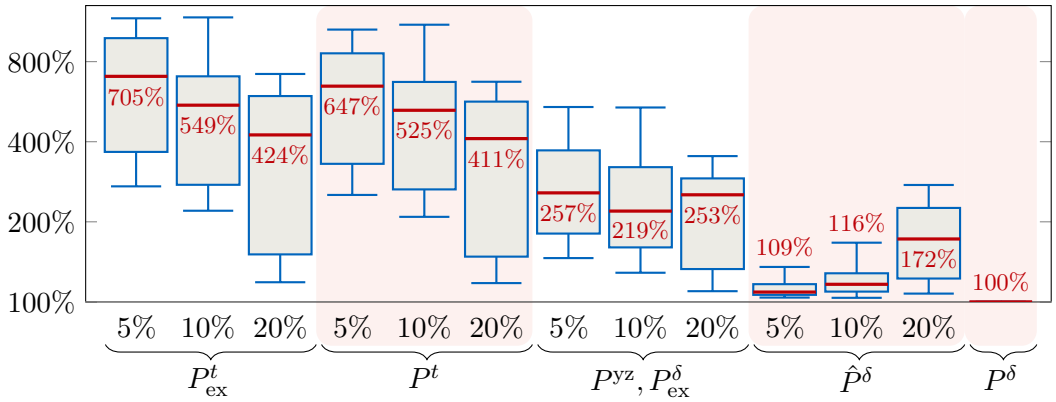


Figure 5.4: Integrality gap relative to P^δ with $\mathrm{CU}_{\mathrm{tol}} \in \{5\%, 10\%, 20\%\}$ in scenario 2025. $\hat{P}^\delta$ is significantly tighter than P^δ_{ex} and nearly matches P^δ at $\mathrm{CU}_{\mathrm{tol}} = 5\%$.

find (good) feasible solutions within the time limit. We amend this by disabling the built-in cutting planes with `XPRS_CUTSTRATEGY=0` and enabling thorough heuristics with `XPRS_HEURSTRATEGY=3` for scenario 2011.

The results are depicted separately in Fig. 5.6 and Fig. 5.7, with P^{δ}_{ex} in both charts for comparability. Figure 5.6 shows that P^{δ}_{ex} outperforms all step-wise start-up cost models P^{t}_{ex}, $P^{t} = \text{epi}(\text{LCU}^1, \dots, \text{LCU}^T)$, and P^{yz}. While the advantage over P^{yz} is only minor, this result is still unexpected, as P^{yz} reduces the model size compared to P^{δ}_{ex}, while retaining its integrality gap. P^{t}_{ex} and P^{t} offer similar performances due to equal model sizes and only slightly differing integrality gap.

Figure 5.7 demonstrates the superiority of the temperature formulations $\hat{P}^{\text{temp}}$ and P^{temp}, and the start-up type formulations $\hat{P}^{\delta}$ and P^{δ} over the state-of-the-art P^{δ}_{ex}. $\hat{P}^{\text{temp}}$ and P^{temp} are matched only by P^{δ} at $\text{CU}_{\text{tol}} = 10\%$, which is notable since $\hat{P}^{\text{temp}}$ and P^{temp} model the exponential start-up costs exactly. Separating the RTIs is not beneficial in this case, as P^{temp} solves slightly fewer instances than $\hat{P}^{\text{temp}}$.

P^{δ} consistently outperforms $\hat{P}^{\delta}$, and both formulations exceed P^{δ}_{ex} for almost all T and CU_{tol}. However, the linear relaxation of P^{δ}_{ex} is solved faster due to the lower number of non-zero coefficients of its inequalities. Hence, P^{δ}_{ex} is successful for some instances where the solution of the linear relaxations of $\hat{P}^{\delta}$ and P^{δ} does not complete within the time limit. Note in particular that $\hat{P}^{\delta}$ and P^{δ} abruptly stop to solve any instances at $T = 192$ (for $\text{CU}_{\text{tol}} = 0\%$) and $T = 288$ (for $\text{CU}_{\text{tol}} = 5\%$).

All models in Fig. 5.6 perform better for a higher approximation tolerance CU_{tol} due to smaller model sizes and smaller integrality gaps for P^{t}_{ex} and P^{t} (see Fig. 5.4). In contrast, the models $\hat{P}^{\delta}$ and P^{δ} in Fig. 5.7 are fastest at $\text{CU}_{\text{tol}} = 10\%$, which offers the best compromise of model size and integrality gap (see Fig. 5.4).

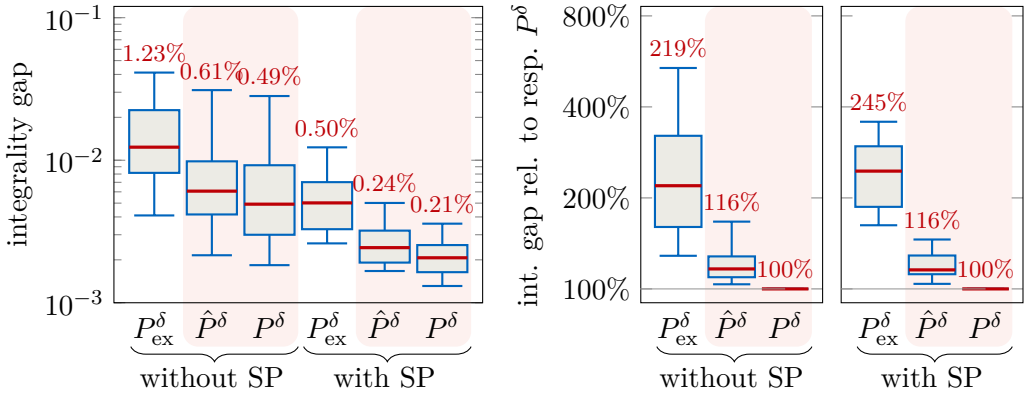


Figure 5.5: Integrality gaps, absolute (left) and relative to P^{δ} (right), in scenario 2025 with and without start-up production. While the integrality gaps decrease when modeling start-up production (left), their ratios between the models remain similar (right).

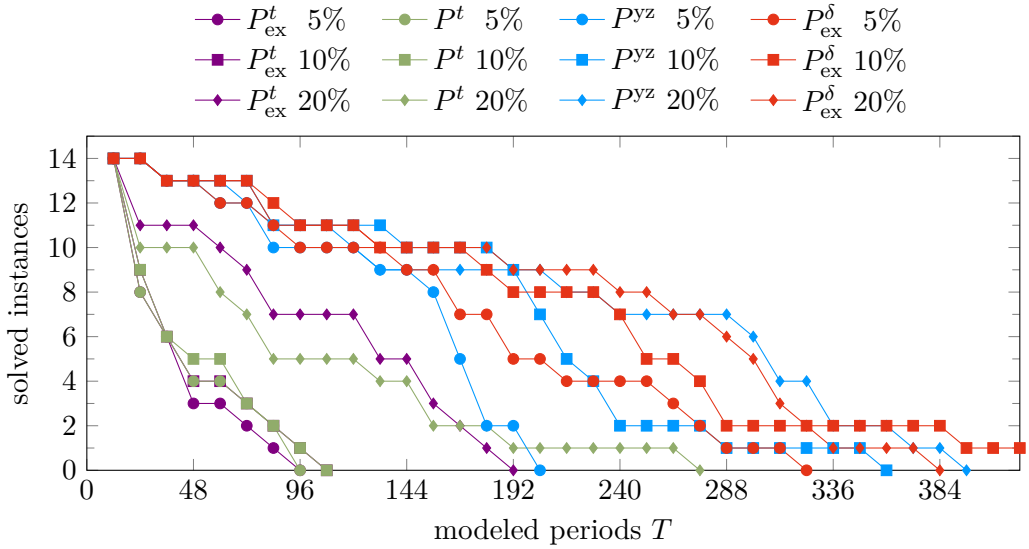


Figure 5.6: Number of instances of scenario 2011 solved to an optimality gap of 1% in 15 minutes. P^δ_{ex} outperforms all step-wise models P^t_{ex}, P^t, P^{yz}. The tightening of P^t_{ex} to $P^t = \text{epi}(\text{LCU}^1, \dots, \text{LCU}^T)$ does not improve the performance.

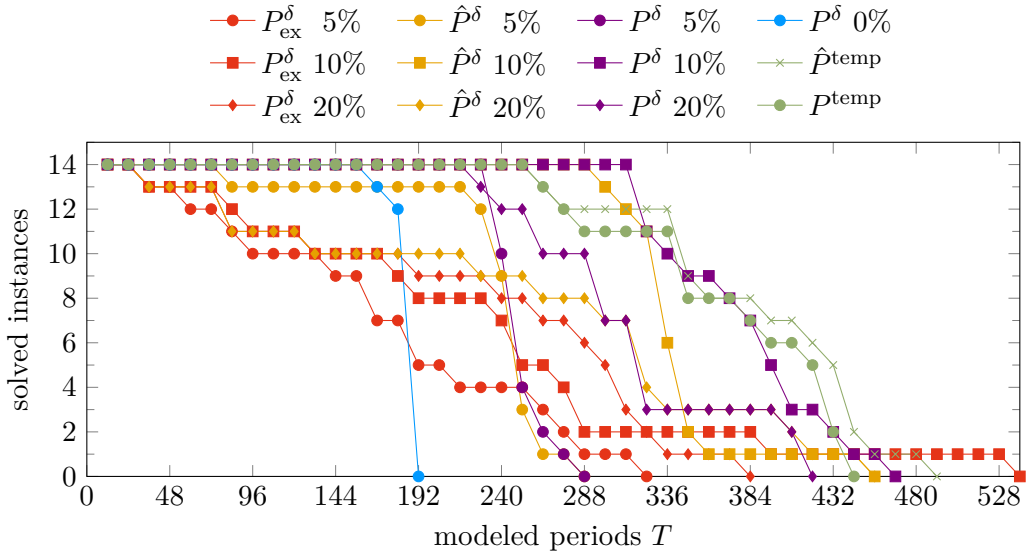


Figure 5.7: Number of instances of scenario 2011 solved to an optimality gap of 1% in 15 minutes. $\hat{P}^{\text{temp}}$, P^{temp}, and P^δ with $\text{CU}_{\text{tol}} = 10\%$ outperform all other models.

The number of variables and inequalities depends approximately linearly on T for all formulations, except for P^δ with $\text{CU}_{\text{tol}} = 0\%$: Here, the number of variables increases quadratically (see Figs. 5.8 and 5.9).

P^t_{ex}, P^t, P^{yz} have the lowest number of variables, followed by $\hat{P}^{\text{temp}}$ and P^{temp}. The number of variables of P^δ_{ex}, $\hat{P}^\delta$, P^δ is equal and decreases with increasing tolerance CU_{tol}. Still, it is higher than in $\hat{P}^{\text{temp}}$ and P^{temp} for $\text{CU}_{\text{tol}} = 20\%$ (see Fig. 5.8).

Regarding inequalities, P^δ with $\text{CU}_{\text{tol}} = 0\%$ has the lowest number followed closely by $\hat{P}^{\text{temp}}$. The other formulations P^t_{ex},P^t,P^{yz},P^δ_{ex},$\hat{P}^\delta$, and P^δ have the same number of inequalities, which increases with decreasing tolerance CU_{tol}.

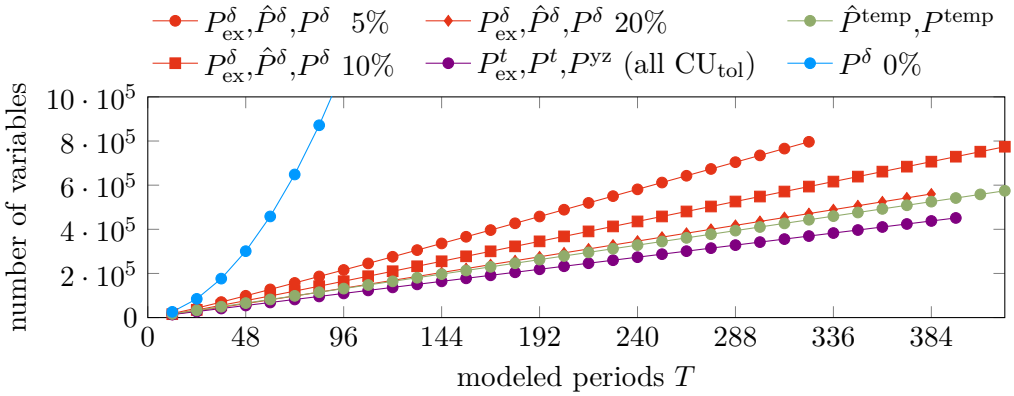


Figure 5.8: The number of variables of the formulations for increasing T. Note that the number of variables depends linearly on T in general, but quadratically on T for P^δ with $\text{CU}_{\text{tol}} = 0\%$. Moreover, the number of variables is independent of CU_{tol} for P^t_{ex}, P^t, P^{yz} and dependent on CU_{tol} for P^δ_{ex}, $\hat{P}^\delta$, P^δ.

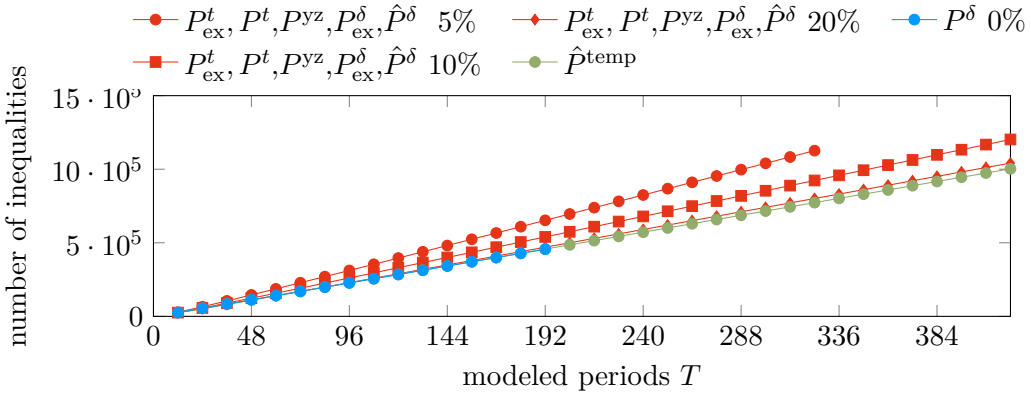


Figure 5.9: The number of inequalities of the formulations for increasing T. Note that all numbers increase linearly with. P^{temp} and P^δ are not included in this chart since the major part of their inequalities is separated.

In scenario 2025, the model size and therefore CU_{tol} is less important (see Fig. 5.10). The results of the models $\hat{P}^\delta$ and P^δ coincide for $\text{CU}_{\text{tol}} \in \{5\%, 10\%\}$, indicating that

the generalized inequalities in P^{δ} are not effective due to the already close integrality gaps. For $\mathrm{CU}_{\mathrm{tol}} = 20\%$ however, P^{δ} outperforms $\hat{P}^{\delta}$.

Figure 5.11 shows all models with their respective optimal $\mathrm{CU}_{\mathrm{tol}}$. The proposed formulations $\hat{P}^{\mathrm{temp}}$, P^{temp}, $\hat{P}^{\delta}$, and P^{δ} are relatively close in performance and dominate P^{t}_{ex}, P^{t}, P^{yz}, and P^{δ}_{ex}. At a higher number of periods, $\hat{P}^{\delta}$ and P^{δ} yield the best results. The models P^{t}_{ex} and P^{t} solve the exact same instances. Similarly, the results for P^{δ}_{ex} and P^{yz} are practically equal.

Comparing $\hat{P}^{\mathrm{temp}}$ and P^{temp} highlights an interference between cutting planes and heuristics similar as in scenario 2011: As long as the problem size is small enough to separate the RTIs and apply heuristics within the time limit, P^{temp} has a higher success rate than $\hat{P}^{\mathrm{temp}}$ (see Fig. 5.11 up to $T = 132$). For larger problems however, the heuristics of P^{temp} are cut short before finding a high-quality solution, and $\hat{P}^{\mathrm{temp}}$ solves more instances due to its less extensive cutting plane stage. To mitigate this, we advise to apply cutting planes and (custom) heuristics in parallel. Nevertheless, P^{temp} solves a total of 9 instances more than $\hat{P}^{\mathrm{temp}}$.

Finally, including the start-up production does not fundamentally change the performance of the start-up type models (cf. Fig. 5.12): $\hat{P}^{\delta}$ and P^{δ} still consistently exceed P^{δ}_{ex}. The importance of the integrality gap is even more pronounced than for scenario 2025, leading to P^{δ} generally improving over $\hat{P}^{\delta}$. While $\mathrm{CU}_{\mathrm{tol}} = 10\%$ again yields the best results for $\hat{P}^{\delta}$, there is no clear optimal $\mathrm{CU}_{\mathrm{tol}}$ for P^{δ}. This is also explained by the stronger dependence on integrality gaps, as P^{δ} possesses nearly the same integrality gap regardless of $\mathrm{CU}_{\mathrm{tol}}$.

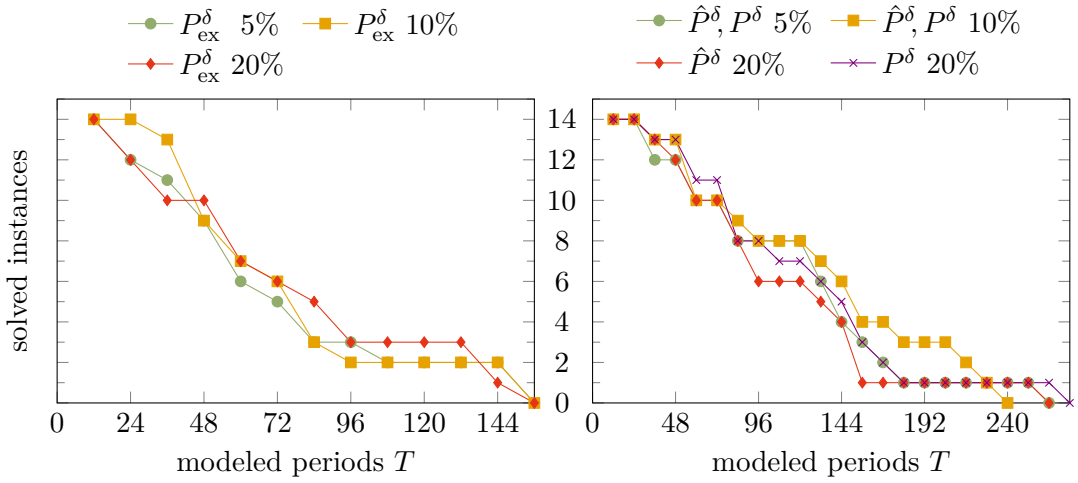


Figure 5.10: Number of instances of scenario 2025 solved to an optimality gap of 0.5% in 30 minutes. P^{δ}_{ex} (left) is typically fastest with $\mathrm{CU}_{\mathrm{tol}} = 20\%$, while $\hat{P}^{\delta}$ and P^{δ} (right) are fastest with $\mathrm{CU}_{\mathrm{tol}} = 10\%$.

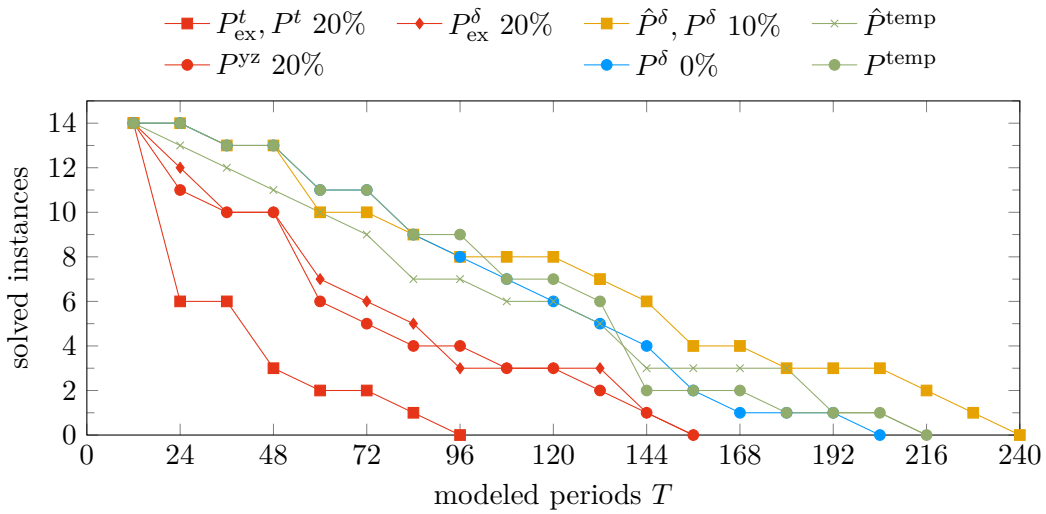


Figure 5.11: Number of instances of scenario 2025 solved to an optimality gap of 0.5% in 30 minutes, showing only the best-performing start-up cost tolerance CU_{tol} for each formulation. P^t_{ex}, P^t and $\hat{P}^\delta$, P^δ yield the same results, respectively. P^δ_{ex} is outperformed by $\hat{P}^{temp}$, $\hat{P}^{temp}$, $\hat{P}^\delta$, P^δ, and P^f.

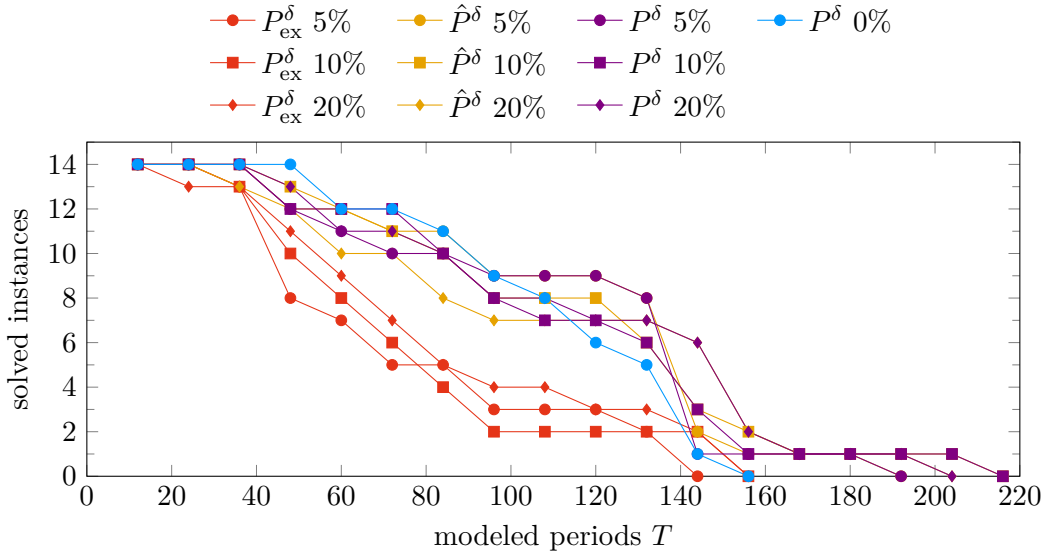


Figure 5.12: Number of instances of scenario 2025 solved to an optimality gap of 0.5% in 30 minutes, including start-up production. The results remain similar, with $\hat{P}^\delta$ and P^δ consistently outperforming P^δ_{ex}.

Chapter 6

Summary and Outlook

6.1 Summary

In this work, we thoroughly examine the start-up process of a electricity producing unit and

- analyze the epigraphs of the start-up cost functions DCU^t, DCU^Σ, and DCU in Chapter 2,
- present the temperature model, an extended formulation of $\mathrm{conv}(\mathrm{epi}(\mathrm{DCU}^\Sigma))$, in Chapter 3, and
- introduce the start-up type model, an extended formulation of $\mathrm{conv}(\mathrm{epi}(\mathrm{DCU}))$, in Chapter 4,

which are interesting both for their theoretical insight and practically applicability.

The analysis of the epigraphs of the start-up cost functions in Chapter 2 yields

- a linear, irredundant $\mathcal{H}$-representation of $\mathrm{conv}(\mathrm{epi}(\mathrm{DCU}^t))$ in Section 2.1,
- an exponential $\mathcal{H}$-representation of $\mathrm{conv}(\mathrm{epi}(\mathrm{DCU}^\Sigma))$ in Section 2.2, which is irredundant in the general case and possesses a linear separation algorithm, and
- an exponential class of facet-inducing inequalities of $\mathrm{conv}(\mathrm{epi}(\mathrm{DCU}))$ in Section 2.3.

Each epigraph dominates a certain important class of inequalities (cf. Subsection 1.3.2), and thus serves as a goalpost for our research. Moreover, these results deepen our understanding of the start-up costs by characterizing the convex extensions LCU^t, LCU^Σ of the discrete start-up cost functions DCU^t, DCU^Σ and by providing a separation algorithm for $\mathrm{conv}(\mathrm{epi}(\mathrm{DCU}^\Sigma))$.

Based on a physical interpretation of the exponential start-up costs, Chapter 3 introduces the temperature and heating variables, and presents

- the temperature model $\hat{P}^{\mathrm{temp}}$ with $\mathcal{O}(T)$ inequalities, and
- the temperature model P^{temp}, an extended formulation of $\mathrm{conv}(\mathrm{epi}(\mathrm{DCU}^\Sigma))$ with $\mathcal{O}(T^2)$ inequalities and linear separation algorithm.

These formulations provide the most compact description of the exponential start-up costs without relying on approximations.

Finally, Chapter 4 identifies a network flow problem which underlies all start-up type models. Based on this interpretation, it introduces

- the start-up type model P^{δ}, an extended formulation of conv(epi(DCU)) with $\mathcal{O}(ST)$ variables, $\mathcal{O}(T^2)$ inequalities, and a $\mathcal{O}(T^2)$ separation algorithm, where S is the number of start-up types,
- the start-up type model $\hat{P}^{\delta}$, which models the start-up costs correctly for $v \in \{0,1\}^T$ using a subset of $\mathcal{O}(ST)$ inequalities of P^{δ}, and
- the model P^{yz}, which provides the same lower bound on the start-up costs as the state-of-the-art start-up type models in [Muc66; SBB10] with $\mathcal{O}(T)$ variables and $\mathcal{O}(ST)$ inequalities.

Using the flow interpretation, Theorem 4.20 derives a total order on the tightness of all proposed start-up cost formulations (under projection) except $\hat{P}^{\text{temp}}$, which is shown in Fig. 6.1. The figure highlights that the models introduced in this work (red) dominate the state-of-the-art (blue) and that we reached our goal of describing each of the start-up cost epigraphs.

Moreover, the temperature formulation $\hat{P}^{\text{temp}}$, which models the start-up costs for integral operational schedules with $\mathcal{O}(T)$ inequalities, is shown to be incomparable to the models $P^{t}_{\text{ex}}, P^{\delta}_{\text{ex}}, P^{\text{yz}}$: The experiments in Chapter 5 indicate that the bound on the start-up costs provided by $\hat{P}^{\text{temp}}$ is typically tighter than P^{yz}, but may be weaker than P^{t}_{ex} in some cases (see Fig. 5.2).

Moreover, Chapter 5 confirms the practical superiority of the presented models, showing that they strongly decrease the integrality gap when modeling start-up costs and processes alike and significantly increase the computational performance.

Still, there remain several interesting open questions, which we address in the following subsections.

6.2 Modeling the Start-up Process in the Temperature Model

This section recaps joint work with Matthias Huber [HS15].

In the temperature formulation in Chapter 3, we model the exponential start-up cost function $\mathrm{CU}(L) = \mathrm{e}^{-\lambda L}$ by introducing variables for the temperature temp^t and heating h^t of a unit. If we restrict ourselves to $\mathrm{L}^1 = \ldots = \mathrm{L}^T = 1$, these variables are connected by the temperature development equation (3.2.7) as

$$\mathrm{temp}^{t+1} = \mathrm{e}^{-\lambda}\mathrm{temp}^t + (1 - \mathrm{e}^{-\lambda})v^t + \mathrm{h}^t.$$

To accurately model the exponential start-up costs, Chapter 3 assumes that the heating time is negligible, effectively leaving h^t unbounded.

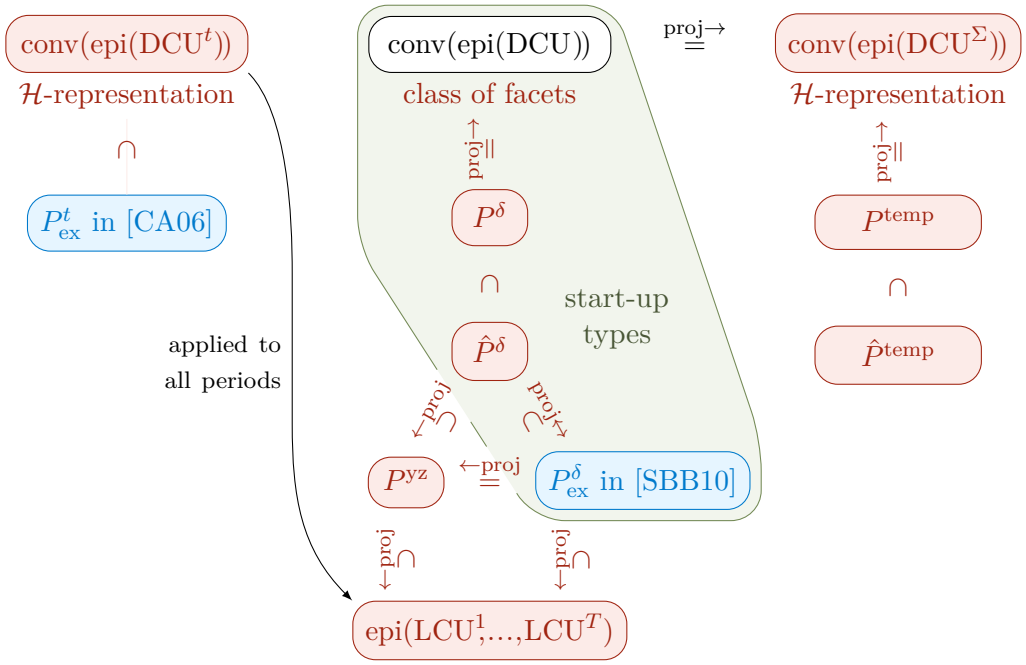


Figure 6.1: Comparison of the tightness of the lower bounds of the start-up cost models, partially under projection ("proj"). Blue nodes denote state-of-the art models, nodes and set relationships in red denote our contribution presented in this thesis. In particular, we prove $P^t_{\text{ex}} \subset P^\delta_{\text{ex}}$, which has been stated in [MELR13b] based on experiments.
The green area marks formulations based on start-up types, which can be used to model the start-up process. The combination of $(v, \text{cu}^t) \in \text{conv}(\text{epi}(\text{DCU}^t))$ for all periods t is denoted by $\text{epi}(\text{LCU}^1, \dots, \text{LCU}^T)$.
Note that the proposed models both dominate the state-of-the-art and describe the start-up cost epigraphs $\text{conv}(\text{epi}(\text{DCU}^t))$ and $\text{conv}(\text{epi}(\text{DCU}^\Sigma))$.

The heating speed of thermal units however is regulated to mitigate material tensions induced by temperature gradients, which necessitate maintenance and shorten the service life of the unit [KSH13]. By limiting the amount of heating performed in each period to $\overline{\text{H}}$,

$$\forall\, t \in [T]: \qquad \text{h}^{t-1} \leq \overline{\text{H}},$$

the total heating at start-up is spread over a time of

$$\text{TU}(L) = \frac{1}{\lambda} \ln\left(\frac{\overline{\text{H}}/\lambda}{\overline{\text{H}}/\lambda - (1 - \text{e}^{-\lambda L})}\right),$$

and therefore over $\lceil \text{TU}(L) \rceil$ periods, depending on the prior downtime L. Hence, this effectively models a time-dependent start-up time (see Fig. 6.2).

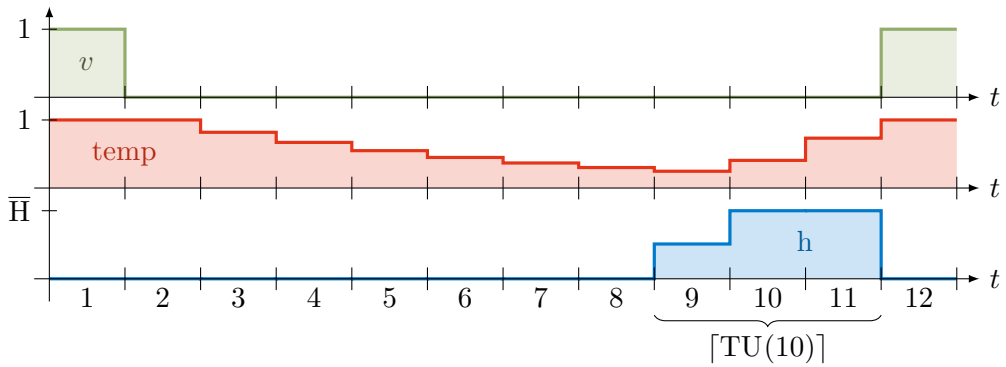


Figure 6.2: Due to the limited heating speed, the shown unit needs to heat up for $\lceil \mathrm{TU}(10) \rceil = 3$ periods before becoming operational in period 12.

Due to the lengthier starting process, a unit looses more heat and requires a total heating of $\overline{\mathrm{H}} \lceil \mathrm{TU}(L) \rceil$, which is interpolated by

$$\frac{\overline{\mathrm{H}}}{\lambda} \ln\left(\frac{\overline{\mathrm{H}}/\lambda}{\overline{\mathrm{H}}/\lambda - (1 - \mathrm{e}^{-\lambda L})}\right) > 1 - \mathrm{e}^{-\lambda L}$$

and for $L \to \infty$ converges to

$$\frac{\overline{\mathrm{H}}}{\lambda} \ln\left(\frac{\overline{\mathrm{H}}/\lambda}{\overline{\mathrm{H}}/\lambda - 1}\right) > 1.$$

Note that both the start-up time and total heating depend on the ratio $\overline{\mathrm{H}}/\lambda$ of the heating speed and the heat loss rate.

To retain the meaning of $\mathrm{CU}^{\mathrm{var}}$ as the maximal variable start-up cost, we define the start-up costs in the case of limited heating as

$$\mathrm{cu}^t := \tilde{\mathrm{CU}}^{\mathrm{var}} \mathrm{h}^{t-1} + \mathrm{CU}^{\mathrm{fixed}} \mathrm{y}^t \quad \text{with} \quad \tilde{\mathrm{CU}}^{\mathrm{var}} := \mathrm{CU}^{\mathrm{var}} \left(\frac{\overline{\mathrm{H}}}{\lambda} \ln\left(\frac{\overline{\mathrm{H}}/\lambda}{\overline{\mathrm{H}}/\lambda - 1}\right)\right)^{-1}.$$

The resulting start-up cost function differs from the exponential start-up costs, albeit only marginally for typical thermal units (cf. Fig. 6.3). Therefore, the temperature model with bounded heating may be applied even if the unit exhibits exponential start-up costs.

Seeing how the start-up time naturally arises in the temperature formulation when respecting a basic limitation of thermal units, we believe that further aspects of the start-up process follow from considering the underlying technical restrictions during a start-up. For example, the start-up production could be tied to either the heating or the temperature while heating, leading to a constant or nearly linear start-up power trajectory as proposed in [SBB10].

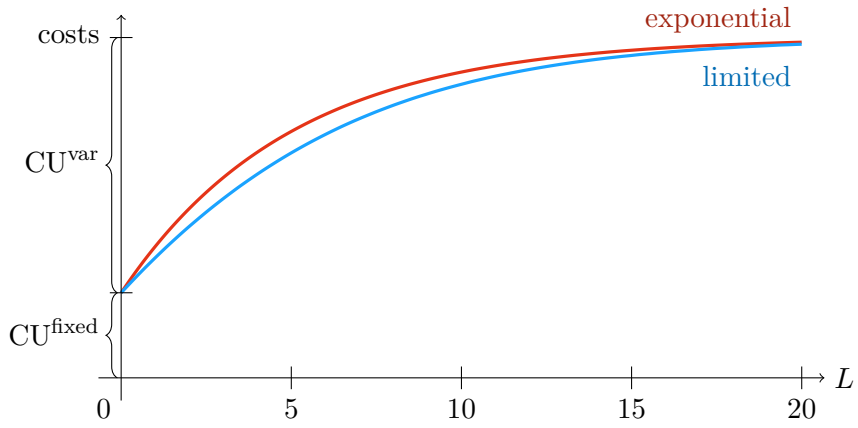


Figure 6.3: Start-up costs with limited heating speed ("limited") compared to exponential start-up costs ("exponential"), depending on the downtime L. Note that the "limited" costs require a smaller coefficient $\tilde{\mathrm{CU}}^{\mathrm{var}} < \mathrm{CU}^{\mathrm{var}}$ than the "exponential" costs to reach the same maximal start-up costs

6.3 The Epigraph of Start-up Costs in All Periods

Section 2.3 presents the composite start-up inequalities, an exponential class of facet-inducing inequalities of conv(epi(DCU)), but leaves the problem of deriving an $\mathcal{H}$-representation of conv(epi(DCU)) open. The start-up flow polyhedron P^f presented in Section 4.2, which is an extended formulation of the lower boundary of conv(epi(DCU)), may provide a viable solution approach.

A fundamental result on extended formulations (see e. g. [Bal05]) considers a projection of the set Q defined as

$$Q := \{(x, u) \in \mathbb{R}^q \times \mathbb{R}^p : Au + Bx \leq b, x \in S\},$$

with matrices $A \in \mathbb{R}^{m \times q}, B \in \mathbb{R}^{m \times p}$, vector $b \in \mathbb{R}^m$, and set $S \subset \mathbb{R}^q$. It shows that projecting Q on its variables x results in

$$\pi_x(Q) = \{x \in \mathbb{R}^q : (v^T B)x \leq v^T b, v \in \mathrm{extr}(W), x \in S\}, \tag{6.3.1}$$

where $\mathrm{extr}(W)$ denotes the extremal rays of the cone W,

$$W := \{v \in \mathbb{R}^m : v^T A = 0, v \geq 0\}.$$

To apply this result, we explicitly incorporate the start-up costs as calculated in the projection π^f from P^f to conv(epi(DCU)), gaining the polyhedron

$$\tilde{P}^f := \left\{(v, \mathrm{cu}, f) \in \mathbb{R}^{2T+T(T+1)/2} \text{ fulfilling inequalities (6.3.2)-(6.3.6)}\right\}$$

with inequalities

$$\sum_{l=0}^{t-1} f_l^t \geq v^t, \qquad t \in [T] \tag{6.3.2}$$

$$\sum_{t'=t+1}^{T} f_{t'-t-1}^{t'} \leq v^t \qquad t \in [T-1] \tag{6.3.3}$$

$$\mathrm{cu}^t \geq \sum_{l=0}^{t-1} \mathrm{CU}^t(l) f_l^t \qquad t \in [T] \tag{6.3.4}$$

$$0 \leq v^t \leq 1, \qquad t \in [T] \tag{6.3.5}$$

$$f_l^t \geq 0, \qquad t \in [T], l \in [0 \,..\, t-1] \tag{6.3.6}$$

and with corresponding orthogonal projection

$$\tilde{\pi}^f : \quad \mathbb{R}^{2T+T(T+1)/2} \to \mathbb{R}^{2T}, \quad (v, \mathrm{cu}, f) \mapsto (v, \mathrm{cu}).$$

Substituting $Q = \tilde{P}^f$ and $\pi_x = \tilde{\pi}^f$ in (6.3.1) yields an $\mathcal{H}$-representation of the projection $\tilde{\pi}^f(\tilde{P}^f) = \mathrm{conv}(\mathrm{epi}(\mathrm{DCU}))$, which can be simplified to

$$\tilde{\pi}^f(\tilde{P}^f) = \left\{ \begin{array}{l} (v, \mathrm{cu}) \in \mathbb{R}^{2T} : \\ \displaystyle\sum_{t \in [T]} \lambda_c^t \mathrm{cu}^t \geq \lambda_s^T v^T + \sum_{t \in [T-1]} (\lambda_s^t - \lambda_e^t) v^t, \quad (\lambda_c, \lambda_s, \lambda_e) \in \mathrm{extr}(W) \\ \qquad 0 \leq v^t \leq 1 \qquad\qquad t \in [T] \end{array} \right\},$$

where λ_c corresponds to the start-up cost bounds (6.3.4), λ_s corresponds to the flow start inequalities (6.3.2), and λ_e corresponds to the flow end inequalities (6.3.3).

The respective cone W is defined as

$$W := \left\{ \begin{array}{l} (\lambda_c, \lambda_s, \lambda_e) \in \mathbb{R}^T \times \mathbb{R}^{T-1} \times \mathbb{R}^T : \\ \qquad \mathrm{CU}^t(l)\lambda_c^t - \lambda_s^t + \lambda_e^{t-l-1} \geq 0, \qquad t \in [T], l \in [0 \,..\, t-2] \\ \qquad \mathrm{CU}^t(t-1)\lambda_c^t - \lambda_s^t \geq 0, \qquad t \in [T] \\ \qquad\qquad \lambda_c^t \geq 0, \qquad t \in [T] \\ \qquad\qquad \lambda_s^t \geq 0, \qquad t \in [T] \\ \qquad\qquad \lambda_e^t \geq 0 \qquad t \in [T-1] \end{array} \right\}.$$

The composite start-up cost inequalities (cf. Definition 2.40) provide some of the rays extr(W). Each inequality with parameters $t \in [T]$, $l \in [0 \,..\, t-1]$, $\mathcal{J} \subset [l-1]$ corresponds to the ray

$$\left\{ q(\lambda_c, \lambda_s, \lambda_e) \,\middle|\, q \in \mathbb{R}_{\geq 0} \right\} \in \mathrm{extr}(W)$$

with coefficients

$$\forall\, t' \in [T]: \qquad \lambda_c^{t'} = \begin{cases} 1 & \text{if } t' = t, \\ \omega_j(t,l) & \text{if } t' = t-j \text{ with } j \in \mathcal{J}, \\ 0 & \text{else,} \end{cases}$$

$$\forall\, t' \in [T]: \qquad \lambda_s^{t'} = \begin{cases} \mathrm{CU}^t(l) & \text{if } t' = t, \\ \mathrm{CU}^t(l) - \mathrm{CU}^t(t-t') & \text{if } t - t' \in \mathcal{J}, \\ 0 & \text{else,} \end{cases}$$

$$\forall\, t' \in [T-1]: \qquad \lambda_e^{t'} = \begin{cases} \mathrm{CU}^t(l) - \mathrm{CU}^t(t-t'-1) & \text{if } t' \in [l], \\ 0 & \text{else.} \end{cases}$$

Deriving the remaining extremal rays of W would lead to an explicit $\mathcal{H}$-representation of conv(epi(DCU)).

6.4 Minimum Downtime and Start-ups

Subsection 1.3.2 establishes the importance of start-up cost epigraphs by showing that if the set of feasible solutions of a Unit Commitment problem fulfills condition (1.3.14), then these epigraphs dominate important classes of inequalities. Specifically, conv(epi(DCU^t)), conv(epi(DCU^Σ)), and conv(epi(DCU)) model the best possible lower bound on the start-up costs with respect to the inequalities

$$\mathrm{cu}^t \geq \sum_{t\in[T]} \alpha_t v^t + \beta, \quad \sum_{t\in[T]} \mathrm{cu}^t \geq \sum_{t\in[T]} \alpha_t v^t + \beta, \quad \text{and} \quad \sum_{t\in[T]} \gamma_t \mathrm{cu}^t \geq \sum_{t\in[T]} \alpha_t v^t + \beta,$$

respectively.

As subsequently noted, condition (1.3.14) does not hold for units with a minimum downtime $\mathrm{DT} > 1$, for which it is straightforward to show that the lower bound on the start-up costs can be improved. An $\mathcal{H}$-representation of conv(epi(DCU^t)) is given by the lifted start-up cost inequalities (2.1.4),

$$\forall\, t \in [T],\, l \in [t-1]: \quad \mathrm{cu}^t \geq \mathrm{CU}^t(l) v^t - \sum_{j=1}^{l} (\mathrm{CU}^t(l) - \mathrm{CU}^t(j-1)) v^{t-j},$$

which can be tightened to

$$\forall\, t \in [T],\, l \in [t-1]: \quad \mathrm{cu}^t \geq \mathrm{CU}^t(l)(v^t - v^{t-1} - \sum_{j=\mathrm{DT}+1}^{l} \Big(\mathrm{CU}^t(l) - \mathrm{CU}^t(j-1)\Big) v^{t-j}.$$

For $v^t = 0$ or $v^{t-1} = 1$, the new inequality is fulfilled since its right-hand side is non-positive. For $v^t = 1$ and $v^{t-1} = 0$, the minimum downtime states that $v^{t-1} = \ldots = v^{t-\mathrm{DT}} = 0$, such that the right-hand sides of the new inequality and the original inequality (2.1.4) are equal.

Analogously, the defining inequality (4.5.5) of the polyhedron P^{yz},

$$\forall\, t \in [T], l \in [0 \,..\, t-1]: \qquad \mathrm{cu}^t \geq \mathrm{CU}^t(l)\mathrm{y}^t - \sum_{j=1}^{l-1} (\mathrm{CU}^t(l) - \mathrm{CU}^t(j))\, \mathrm{z}^{t-j},$$

can be tightened to

$$\forall\, t \in [T], l \in [0 \,..\, t-1]: \qquad \mathrm{cu}^t \geq \mathrm{CU}^t(l)\mathrm{y}^t - \sum_{j=\mathrm{DT}}^{l-1} (\mathrm{CU}^t(l) - \mathrm{CU}^t(j))\, \mathrm{z}^{t-j}.$$

Regarding conv(epi(DCU)), the stronger lower bound is easy to see when considering its extended formulations, the start-up flow polyhedron P^f and the start-up type polyhedron P^δ. A minimum downtime DT effectively renders all start-up flows f_l^t with $l \in [\mathrm{DT} - 1]$ superfluous and P^f can be tightened by setting

$$\forall\, t \in [T], l \in [\mathrm{DT} - 1]: \qquad f_l^t = 0.$$

Equivalently, P^δ can be tightened by choosing the partition $\mathcal{L}_0^t, \ldots, \mathcal{L}_{S^t}^t$ of $[0 \,..\, t-1]$ such that $\mathcal{L}_1^t = [\mathrm{DT} - 1]$ and fixing δ_1^t,

$$\forall\, t \in [T]: \qquad \delta_1^t = 0.$$

Given these straightforward improvements, one cannot help asking:

- Is a further tightening possible? Do additional non-redundant inequalities exist?
- Can the binary tree inequalities or the temperature model be tightened in the case of a minimum downtime?
- Do other typical restrictions on the operational schedule of a unit result in similar opportunities for improvement?

Bibliography

[AC00] J. Arroyo and A. Conejo. “Optimal response of a thermal unit to an electricity spot market”. In: *IEEE Trans. Power Syst.* 15.3 (2000), pp. 1098–1104.

[Bal05] E. Balas. “Projection, lifting and extended formulation in integer and combinatorial optimization”. In: *Ann. Oper. Res.* 140.1 (2005), pp. 125–161.

[BHS] R. Brandenberg, M. Huber, and M. Silbernagl. “An Extended Formulation of the Summed Start-up Costs in a Unit Commitment Problem via Power Plant Temperatures”. In Preparation.

[BHS16] R. Brandenberg, M. Huber, and M. Silbernagl. “The Summed Start-up Costs in a Unit Commitment Problem”. In: (2016). EURO J. Comput. Optim., Early access: `http://www.springer.com/-/2/AVOG6nyoAgfPWjhr_594`.

[BS14] R. Brandenberg and M. Silbernagl. *Implementing a Unit Commitment Power Market Model in FICO Xpress-Mosel*. FICO Xpress Optimization Suite whitepaper. FICO, 2014.

[CA06] M. Carrión and J. Arroyo. “A computationally efficient mixed-integer linear formulation for the thermal unit commitment problem”. In: *IEEE Trans. Power Syst.* 21.3 (2006), pp. 1371–1378.

[DW60] G. B. Dantzig and P. Wolfe. “Decomposition principle for linear programs”. In: *Oper. Res.* 8.1 (1960), pp. 101–111.

[Ege+14] J. Egerer, C. Gerbaulet, R. Ihlenburg, F. Kunz, B. Reinhard, C. v. Hirschhausen, A. Weber, and J. Weibezahn. *Electricity sector data for policy-relevant modeling: Data documentation and applications to the German and European electricity markets*. Data Documentation 72. Berlin: Deutsches Institut für Wirtschaftsforschung (DIW), 2014.

[EHG11] J. Ebrahimi, S. H. Hosseinian, and G. B. Gharehpetian. “Unit commitment problem solution using shuffled frog leaping algorithm”. In: *IEEE Trans. Power Syst.* 26.2 (2011), pp. 573–581.

[FF56] L. R. Ford and D. R. Fulkerson. “Maximum flow through a network”. In: *Canadian J. Math.* 8.3 (1956), pp. 399–404.

[Fic] *FICO Xpress Optimization Suite*. URL: `http://www.fico.com/en/products/fico-xpress-optimization-suite` (visited on 05/27/2015).

[Gar62] L. Garver. "Power Generation Scheduling by Integer Programming-Development of Theory". In: *IEEE Trans. Power App. Syst.* 81.3 (1962), pp. 730–734.

[GBT84] H. N. Gabow, J. L. Bentley, and R. E. Tarjan. "Scaling and Related Techniques for Geometry Problems". In: *Proc. 16th Annu. ACM Symp. Theory Comput.* STOC '84. New York, NY, USA: ACM, 1984, pp. 135–143.

[GK94] P. Gritzmann and V. Klee. "On the complexity of some basic problems in computational convexity: I. Containment problems". In: *Discrete Math.* 136.1 (1994), pp. 129–174.

[Gom58] R. E. Gomory. "Outline of an algorithm for integer solutions to linear programs". In: *B. Am. Math. Soc.* 64.5 (1958), pp. 275–278.

[GZP03] X. Guan, Q. Zhai, and A. Papalexopoulos. "Optimization based methods for unit commitment: Lagrangian relaxation versus general mixed integer programming". In: *2003 IEEE Power Eng. Soc. General Meeting*. Vol. 2. IEEE, 2003.

[HK56] A. Hoffman and J. Kruskal. "Integral boundary points of convex polyhedra". In: *Linear Inequalities and Related Systems*. Ed. by K. Kuhn and A. Tucker. Princeton University Press, Princeton, NJ, 1956, pp. 223–246.

[HS15] M. Huber and M. Silbernagl. "Modeling Start-Up Times in Unit Commitment by Limiting Temperature Increase and Heating". In: *12th Int. Conf. Eur. Energy Market*. Lisbon, 2015.

[Hug93] T. Hughes. *Networks of Power: Electrification in Western Society, 1880-1930*. Softshell Books. Baltimore: Johns Hopkins University Press, 1993.

[KSH13] P. Keatley, A. Shibli, and N. J. Hewitt. "Estimating power plant start costs in cyclic operation". In: *Appl. Energ.* 111 (2013), pp. 550–557.

[Kum+12] N. Kumar, P. M. Besuner, S. A. Lefton, D. D. Agan, and D. A. Hileman. *Power Plant Cycling Costs*. Tech. rep. Research Report, prepared for National Renewabl Energy Laboratory and Western Electricity Coordinating Council. Sunnyvale, California: Intertek APTECH, 2012.

[LLM04] J. Lee, J. Leung, and F. Margot. "Min-up/min-down polytopes". In: *Discrete Optim.* 1.1 (2004), pp. 77–85.

[MELR13a] G. Morales-España, J. Latorre, and A. Ramos. "Tight and Compact MILP Formulation for the Thermal Unit Commitment Problem". In: *IEEE Trans. Power Syst.* 28.4 (2013), pp. 4897–4908.

[MELR13b] G. Morales-España, J. Latorre, and A Ramos. "Tight and Compact MILP Formulation of Start-Up and Shut-Down Ramping in Unit Commitment". In: *IEEE Trans. Power Syst.* 28.2 (2013), pp. 1288–1296.

[MHV12] M. Moghimi Hadji and B. Vahidi. "A solution to the unit commitment problem using imperialistic competition algorithm". In: *IEEE Trans. Power Syst.* 27.1 (2012), pp. 117–124.

[Muc66] J. A. Muckstadt. "Scheduling in power systems". `http://hdl.handle.net/2027.42/6720`. PhD thesis. University of Michigan, 1966.

[NR00] M. P. Nowak and W. Römisch. "Stochastic Lagrangian Relaxation Applied to Power Scheduling in a Hydro-Thermal System under Uncertainty". In: *Ann. Oper. Res.* 100.1-4 (2000), pp. 251–272.

[OAV12] J. Ostrowski, M. Anjos, and A. Vannelli. "Tight Mixed Integer Linear Programming Formulations for the Unit Commitment Problem". In: *IEEE Trans. Power Syst.* 27.1 (2012), pp. 39–46.

[Ott10] A. Ott. "Evolution of computing requirements in the PJM market: Past and future". In: *2010 IEEE Power Energy Soc. General Meeting*. 2010, pp. 1–4.

[Pad04] N. Padhy. "Unit commitment-a bibliographical survey". In: *IEEE Trans. Power Syst.* 19.2 (2004), pp. 1196–1205.

[Pad73] M. W. Padberg. "On the Facial Structure of Set Packing Polyhedra". In: *Mathematical Programming* 5.1 (1973), pp. 199–215.

[PE10] V. S. Pappala and I. Erlich. "A variable-dimension optimization approach to unit commitment problem". In: *IEEE Trans. Power Syst.* 25.3 (2010), pp. 1696–1704.

[Rie+11] M. M. Rienecker et al. "MERRA: NASA's Modern-Era Retrospective Analysis for Research and Applications". In: *J. Climate* 24.14 (2011), pp. 3624–3648.

[Ros00] K. H. Rosen, ed. *Handbook of discrete and combinatorial mathematics*. Discrete Math. Appl. CRC press, 2000.

[RR91] S. Ruzic and N. Rajakovic. "A new approach for solving extended unit commitment problem". In: *IEEE Trans. Power Syst.* 6.1 (1991), pp. 269–277.

[RT05] D. Rajan and S. Takriti. "Minimum Up/Down Polytopes of the Unit Commitment Problem with Start-Up Costs". In: *IBM Res. Rep.* (2005).

[Rud76] W. Rudin. "Principles of mathematical analysis". In: 3d ed. International series in pure and applied mathematics. New York: McGraw-Hill, 1976.

[SBB10] C. Simoglou, P. Biskas, and A. Bakirtzis. "Optimal Self-Scheduling of a Thermal Producer in Short-Term Electricity Markets by MILP". In: *IEEE Trans. Power Syst.* 25.4 (2010), pp. 1965 –1977.

[Sch98] A. Schrijver. *Theory of linear and integer programming*. Chichester: John Wiley & Sons, 1998.

[SF94] G. Sheble and G. Fahd. "Unit commitment literature synopsis". In: *IEEE Trans. Power Syst.* 9.1 (1994), pp. 128 –135.

[SHB16] M. Silbernagl, M. Huber, and R. Brandenberg. "Improving Accuracy and Efficiency of Start-up Cost Formulations in MIP Unit Commitment by Modeling Power Plant Temperatures". In: *IEEE Trans. Power Syst.* 31.4 (2016), pp. 2578–2586.

[SPO05] D. Streiffert, R. Philbrick, and A Ott. "A mixed integer programming solution for market clearing and reliability analysis". In: *2005 IEEE Power Eng. Soc. General Meeting*. 2005, 2724–2731 Vol. 3.

[SS34] M. J. Steinberg and T. H. Smith. "The theory of incremental rates and their practical Application to load division—Part I". In: *Electrical Engineering* 53.3 (1934), pp. 422–445.

[Tse96] C.-L. Tseng. *On power system generation unit commitment problems*. University of California, Berkeley, 1996.

[Vui80] J. Vuillemin. "A unifying look at data structures". In: *Commun. ACM* 23.4 (1980), pp. 229–239.

[WW96] A. Wood and B. Wollenberg. *Power generation, operation and control*. 2nd ed. Wiley, 1996.

[ZG88] F. Zhuang and F. Galiana. "Towards a more rigorous and practical unit commitment by Lagrangian relaxation". In: *IEEE Trans. Power Syst.* 3.2 (1988), pp. 763–773.

[ENT] ENTSO-E, European Network of Transmission System Operators for Electricity. *Hourly Load Values*. URL: `http://www.entsoe.eu/data/data-portal/consumption/` (visited on 06/01/2015).

[EUR03] EURELECTRIC. *Efficiency in Electricity Generation*. Tech. rep. `www.eurelectric.org/Download/Download.aspx?DocumentID=13549`. EURELECTRIC, 2003.

[Ger14] German Federal Network Agency. *List of Power Plants*. 2014. URL: `http://www.bundesnetzagentur.de/cln_1412/DE/Sachgebiete/ElektrizitaetundGas/Unternehmen_Institutionen/Versorgungssicherheit/Erzeugungskapazitaeten/Kraftwerksliste/kraftwerksliste-node.html` (visited on 06/01/2015).

[Int12] International Energy Agency. *World Energy Outlook 2012*. OECD/IEA, 2012.

[Int13] International Energy Agency. *Renewable Energy Medium-Term Market Report*. Tech. rep. OECD/IEA, 2013.

Appendix A

Finding Cartesian Trees

By [GBT84], the Cartesian tree of a vector $x \in \mathbb{R}^n$ can be constructed in $\mathcal{O}(n)$ using an iterative approach. Since Subsection 2.2.5 relies on its applicability, we reprise the algorithm and its proof of correctness in detail.

Algorithm A.0.1: FindCartesianTree

Input : Vector $x \in \mathbb{R}^n$

Output : Cartesian tree B for the vector x, represented by llink and rlink

```
1  for i = 1, ..., n do
2      llink(i) ← ∅, rlink(i) ← ∅;
3  η_1 ← 1;
4  R ← 1;
5  for i = 2, ..., n do
6      if x_{η_1} ≤ x_i then
7          llink(i) ← η_1;
8          η_1 ← i;
9          R ← 1;
10     else
11         j ← R;
12         while x_{η_j} ≤ x_i do
13             j ← j − 1;
14         llink(i) ← rlink(η_j);
15         rlink(η_j) ← i;
16         η_{j+1} ← i;
17         R ← j + 1;
18 return binary tree represented by llink and rlink;
```

Lemma A.1 *For each $x \in \mathbb{R}^n$, Algorithm A.0.1 finds a Cartesian tree for x in $\mathcal{O}(n)$.*

Proof. For $i \in [n]$, let $\text{llink}^i(j)$, $\text{rlink}^i(j)$, η_j^i and R^i denote the values of the variables $\text{llink}(j)$, $\text{rlink}(j)$, η_j and R during the execution of the algorithm,

- before the first execution of the main `for`-loop (lines 5–17) for $i = 1$, and
- after the $(i-1)$-th iteration of the main `for`-loop for $i > 1$.

We proof by induction over $i \in [n]$ that llink^i and rlink^i restricted to $[i]$ represent a Cartesian tree B^i for $(x_1, \ldots, x_i)$ with $\text{rank}^{B^i}(k) = k$ for each node $k \in [i]$ and top-right nodes $\eta_1^i, \ldots, \eta_{R^i}^i$.

The functions llink^1 and rlink^1 restricted to $[1]$ represent the binary tree B^1 with the single node 1, which is the unique Cartesian tree for (x_1). Both R^1 and η_1^1 are correct.

For $i > 1$, consider the case where the `if`-block (lines 6–9) was executed in the last iteration of the main loop: here line 7 appends B^{i-1} as the left subtree of i (see Fig. A.1). Since the root i has no right child, the sole top-right node is $\eta_1^i = i$, which is reflected by the update of the top-right nodes on lines 8–9.

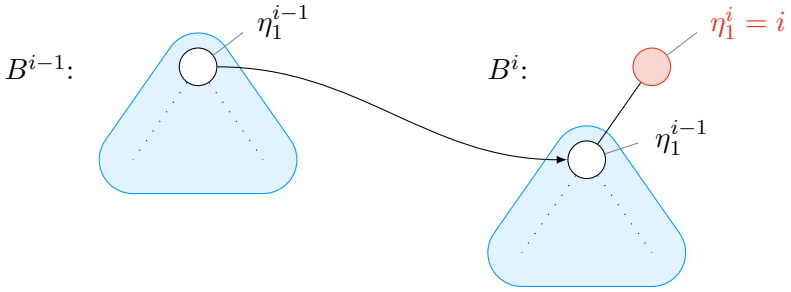


Figure A.1: Adding a node i with $x_{\eta_1^{i-1}} \leq x_i$ to the Cartesian tree B^{i-1}, resulting in the Cartesian tree B^i.

By definition of the rank function, $\text{rank}^{B^1}(i) = i$ and the ranks of the nodes in its left subtree $L^{B^i}(i) = B^{i-1}$ are $[i-1]$. Since their relative order remains unchanged, their ranks remain unchanged, and $\text{rank}^{B^i}(k) = \text{rank}^{B^{i-1}}(k) = k$ for all $k \in [i-1]$.

All nodes $k \in B^i$ except η_1^{i-1} and i have the same parent in B^{i-1} and B^i, and thus fulfill $x_k \leq x_{p(k)}$. Since the `if`-block has been executed, $x_{\eta_1^{i-1}} \leq x_i$ holds as well and B^i is a Cartesian tree for $(x_1, \ldots, x_i)$.

Now, consider the case where the `else`-block (lines 10–17) has been executed in the last iteration of the main loop. Since the statement $x_i < x_{\eta_1^{i-1}}$ holds, the `while`-loop on lines 12–13 stops with $j \geq 1$. Lines 14–15 move the right subtree T of η_j^{i-1} to the left subtree of i, and append i as the right child of η_j^{i-1} (see Fig. A.2). Hence, the top-right nodes $\eta_{j+1}^{i-1}, \ldots, \eta_{R^{i-1}}^{i-1}$ are replaced by $\eta_{j+1}^i = i$, and lines 16–17 update η_{j+1} and R correctly.

As the node $\eta_{R^{i-1}}^{i-1} = i-1$ lies in the moved subtree T, we have $\text{rank}^{B^i}(i) = i$ and T contains the nodes with ranks $[i-s(T) \,..\, i-1]$ both in B^{i-1} and B^i. By extension, the

remainder of the tree B^{i-1} also contains the nodes with the same ranks $[i - s(T) - 1]$ in B^i. So, the ranks in B^{i-1} and B^i are equal, and $\text{rank}^{B^i}(i) = i$.

All nodes $k \in B^i$ except i and η_{j+1}^{i-1} (if $j < R^{i-1}$) have the same parent in B^{i-1} and B^i, and thus fulfill $x_{p(k)} \geq x_k$. Since the `while`-loop terminates with $x_{p(i)} = \eta_j^{i-1} \leq x_i$ the Cartesian tree property is fulfilled for i. If $j < R^{i-1}$, the same holds for η_{j+1}^{i-1}, as we have $x_{p(\eta_{j+1}^{i-1})} = x_i < \eta_{j+1}^{i-1}$. So, B^i is a Cartesian tree for $(x_1, \ldots, x_i)$.

This concludes the induction and proves that the tree resulting from the algorithm is a Cartesian tree for the vector $(x_1, \ldots, x_n)$.

The whole algorithm except the while-loop (lines 12-13) clearly has a running time of $\mathcal{O}(n)$. The number of top-right nodes R is initialized with 1, and is only changed on line 17 to $R = j + 1$. Since the condition on line 12 for $j = 1$ is the same as the condition on line 7, the while-loop stops with $j \geq 1$. Thus, R is always greater than or equal to 1. On the other hand, R is only possibly increased on line 17 by at most 1. Thus, the body of the while-loop is executed at most $n - 1$ times during the complete algorithm. □

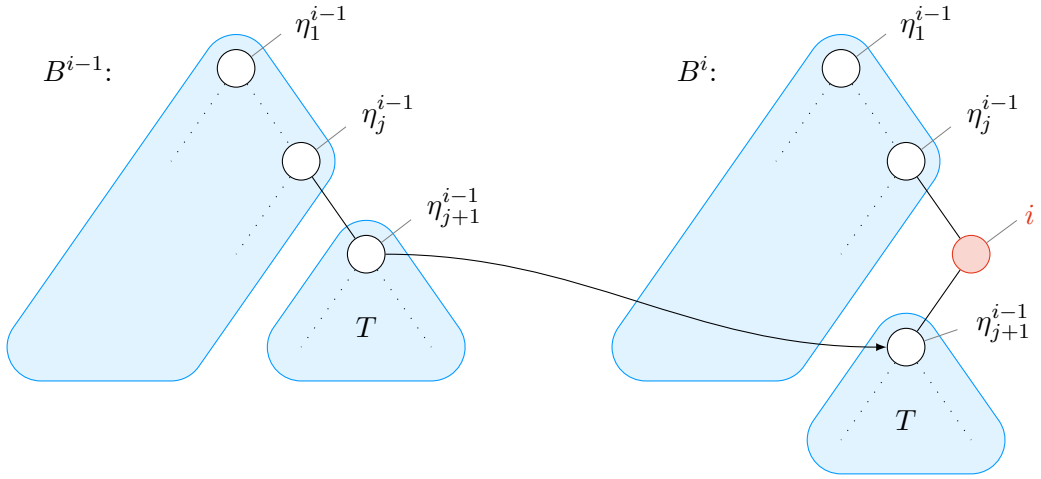


Figure A.2: Adding a node i with $x_{\eta_j^{i-1}} > x_i \geq x_{\eta_{j+1}^{i-1}}$ to the Cartesian tree B^{i-1}, resulting in the Cartesian tree B^i.

Glossary

boiler Component of a unit which is heated by the burning fuel to generate steam. 92

demand, residual Difference between electricity demand and production from other sources; in this publication: difference between demand and renewable energy production. 5, 6, 155

downtime Time during which a unit does not produce electricity, i. e. is offline. 5, 7, 8, 10, 11, 15–19, 22, 25, 26, 32, 50, 51, 55, 56, 91–95, 106, 116, 119, 125, 140, 145, 167, 169, 171, 172

Economic Dispatch Problem of cost-optimally distributing an electricity demand on a set of online units. Does not include the decision which units should be offline and which units should be online. 6

load, base Typical minimal residual demand in a power system. 5, 31

load, peak Typical maximal residual demand in a power system. 31

material tension Force on material; in this publication due to temperature gradients which result in differing material densities. 17, 167

maximal production Maximal sustainable electricity production of a unit. Generally determined by the thermal design. 7, 10, 11, 13

minimal production Minimal electricity production in online state of a unit. Generally defined by law, since lower production levels lead to a higher concentration of carbon monoxide and other toxic gases in the exhaust gas. 7, 13, 14, 154

MWh Megawatt hour, a unit of energy of adequate dimension to describe power plants. The average German citizen used 7 MWh of electricity in 2011. 8, 155

offline State during which a unit does not produce electricity or heat. 11, 15, 17–19, 22, 32, 33, 36, 37, 43, 54–56, 65–69, 78, 93, 95–97, 104, 106, 125, 126, 129, 130, 148, 183, 184

online State during which a unit contributes electricity to the power system. 6, 8–12, 15, 18, 36, 82, 95–97, 183–185

operational schedule Schedule determining the operational state of a unit over time. 11, 19, 20, 24, 32, 52, 53, 55, 92, 93, 96, 99, 100, 102, 105, 109, 119, 125, 129, 134, 140, 141, 144, 172

operational state General state of a unit: online, offline, banking, soaking, synchronizing, desynchronizing. 8–10, 15, 18, 96, 99, 183

power plant Electricity producing factory, typically consisting of multiple units in a single building complex. 1, 4, 6, 154, 183

power trajectory Pre-defined controlled development of a unit's electricity production level over time, mainly in the context of a start-up. 1, 5, 17, 22, 25, 123, 154, 168, 184

production costs Costs of electricity production excluding start-up costs. 5, 7, 12, 14, 159

ramp Change of the electricity production level of a unit during normal operation. 7, 10, 14, 19, 140, 157, 159, 185

shutdown Transition from online to offline status of a unit. 5, 8, 11, 12, 26, 94, 97, 140, 141, 144, 185

start-up Transition from offline to online status of a unit. 1, 4, 5, 8, 11, 15–19, 22, 23, 26, 31, 94, 97, 106, 107, 123, 129, 140, 141, 144, 153–155, 157, 159, 167, 168, 184, 185

start-up costs Costs caused by a start-up, both directly and indirectly. 1, 3, 4, 7, 10, 12, 13, 15–25, 31–33, 36–38, 40–43, 47–49, 51–55, 80, 82, 91–96, 98–100, 102, 105–110, 116, 120, 121, 123, 125, 126, 129, 130, 132–134, 140, 142–152, 154, 155, 158–160, 165, 166, 168–171

start-up process Process to transition a unit from offline to online state, encompassing a pre-defined power trajectory. 4, 5, 12, 22, 27, 123, 153, 156, 165, 167, 168

start-up production Production of a unit while performing a start-up. 7, 8, 12, 15, 18, 22, 23, 26, 123, 154–156, 159, 160, 163, 164, 168

start-up time Time required for to complete the start-up process of a unit. 15, 17, 18, 22, 25, 154, 167, 168

state, banking State in which the unit is kept at operational temperature without producing electricity, e. g. to allow a fast start-up. 15, 183

state, desynchronizing State in which a unit is disconnecting from the power system. 15, 183

state, soaking State during start-up in which the heat in the turbines is allowed to distribute by keeping the unit at a certain intermediate temperature. 15, 123, 183

state, synchronizing State in which a unit is connecting to the power system. 15, 123, 183

thermal stress Deteriorating force due to quick changes in temperature. 5

unit Smallest electricity producing entity which may be started up and shut down independently. 1, 3, 5–7, 10–14, 16–18, 22, 23, 33, 36, 37, 52, 80, 91, 92, 94, 96, 97, 99, 100, 140, 148, 155, 158, 159, 165, 167, 168, 171, 172, 183–185

Unit Commitment Problem of cost-optimally distributing an electricity demand on a set of units, deciding which units need to be online and which units should be offline. 1, 6, 8–10, 13–15, 19, 21–25, 31, 52, 93, 96, 140, 141, 153–156, 158, 171

unit, biomass Thermal unit producing energy by burning organic matter, e. g. wood, (converted) plants, waste. 5, 155

unit, coal Thermal unit producing heat by burning coal. 5, 31, 154

unit, combined cycle gas Thermal unit producing electricity by transforming the heat of burning gas in multiple combined thermodynamic cycles. 154, 185

unit, combined heat&power Unit whose waste heat is used in heating-intensive processes, e. g. chemical processing, drying, building heating. 92

unit, gas Unit producing electricity by transforming the heat of gas in a single thermodynamic cycle. Less efficient than a combined cycle gas unit, but with faster ramping and start-up. 5, 31, 154

unit, hydro Unit producing energy by converting the gravitational force of water. 11, 155

unit, hydrothermal Unit which is either a thermal unit or a hydro; mainly used to address units with controllable production level. 13

unit, lignite Thermal unit producing energy by burning lignite, i. e. brown coal. 5, 154

unit, nuclear Thermal unit producing heat by nuclear fission. 5, 154

unit, solar Unit producing energy by converting solar energy. 7, 155

unit, thermal Unit which produces electricity by generating heat, e. g. by burning fuel and converting the heat to electricity. 5, 13, 16–18, 31, 47, 167, 168

unit, wind Unit producing energy by converting wind energy. 7, 155

uptime Time during which a unit produces electricity, i. e. is online. 7, 8, 10, 11, 14, 19, 22, 26, 140, 153

wear&tear Deterioration of a unit due to production, ramping, start-ups, and shut-downs. 5