Energy-efficient Timetables for Railway Traffic: Incorporating DC Power Models *

Robert Burlacu† Patrick Gemander† Tobias Kuen† Jorge Weston‡

September 9, 2024

Abstract

Efficient operation of underground railway systems is critical not only for maintaining punctual service but also for minimizing energy consumption, a key factor in reducing operational costs and environmental impact. To evaluate the energy consumption of the timetables, this paper delves into the development of mathematical models to accurately represent energy dynamics within the underground railway network.

We evaluate the total traction energy consumption in an underground railway network over a specified period, with the analysis discretized to a per-second basis. At each second we evaluate the power flow in the transmission network, a direct current (DC) power grid with fixed powerstation voltages. Quadratic constraints arise when linking power, current, and voltage. To deal with the resulting computational complexity we compare two model formulations, one based on power flow and the other based on current flow. We demonstrate that the current flow model is easier to solve and develop a heuristic to further speed up the solution process.

We integrate the model addressing the power flows in the transmission network with the model that ensures the feasibility of the timetable. Central to our approach is the utilization of Benders row generation to tackle the complexity of the large integrated model. By decomposing the optimization problem into manageable subproblems, we enhance computational efficiency and scalability. To linearize the Benders subproblems we develop relaxations for the non-linear constraints and binary variables. We analyze the performance of the integrated timetabling power flow model on real world data provided by the VAG (Verkehrs-Aktiengesellschaft Nürnberg), the operator of public transport in Nuremberg, Germany. The simulated energy consumption deviates from the actual measurements by only around 1%. The calculated timetable increases energy efficiency by up to 0.8% compared to the previously used model. Further numerical studies demonstrate the effectiveness of the developed solving algorithms.

Keywords: Railway Timetabling, Integer Non-Linear Programming, Benders Decomposition, DC Optimal Power Flow

1 Introduction

The underground railway system is a widely used urban transportation network recognized for its effectiveness and, along with trams, as one of the cleanest forms of public transport (Moreno et al., 2015). The study of energy efficiency in underground railway systems has gained relevance in light of growing global concerns about climate change and the need for clean energy solutions.

*This work was part of the Project EKSSE, funded by the Bundesministerium für Bildung und Forschung (BMBF).

†Fraunhofer-Institut für Integrierte Schaltungen IIS, Gruppe Optimization, Nordostpark 84, 90411 Nürnberg, Germany, (robert.burlacu@iis.fraunhofer.de, patrick.gemander@fau.de, tobias.kuen@iis.fraunhofer.de).

‡Friedrich-Alexander-Universität Erlangen-Nürnberg, Chair Data Science, Cauerstraße 11, 91058 Erlangen, Germany, jorge.weston@fau.de

A critical area of research is reducing energy consumption within DC networks, which represent one of the highest operational costs for underground railway systems. This work specifically focuses on reducing energy consumption in an underground railway system through efficient scheduling and the recovery of braking energy from trains.

Adjusting timetables can significantly impact energy consumption in underground railway systems. Two key factors are relevant in this context: First, the optimal utilization of braking energy, and second, the extension of coasting phases. By synchronizing the braking and accelerating phases of trains, it is possible to capture and reuse braking energy effectively, thereby reducing the overall energy consumption. Additionally, longer running times allow for extended coasting phases, further decreasing energy consumption for individual trains.

However, the efficiency gains from capturing braking energy are tempered by inherent losses in the transmission network, particularly when trains are far apart. The transmission network in underground railway systems operates on a DC power grid, presenting unique challenges for mathematical modeling. Proper optimization techniques are essential for accurately representing the complex dynamics of DC power flows and timetabling.

Energy recovery can be approached in two ways: first, using storage devices to store energy generated during braking, which can then be used to accelerate trains when needed (known as wayside storage); and second, by synchronizing accelerating and braking trains to properly coordinate the energy supplied by braking trains and reduce the amount of energy required from the power network.

Several studies have addressed energy efficiency in underground railway systems through various approaches. For example, Gao and L. Yang, 2019; H. Liu et al., 2018; P. Liu et al., 2018; Su et al., 2013; X. Yang et al., 2015; Zhao et al., 2017 consider the timetable problem alongside train speed profiles to find energy-efficient timetables, but they often overlook the characteristics of DC networks and the related transmission energy losses. These studies focus mainly on the kinetic energy recovery problem and differ in timetable constraints and solution approaches, such as brute force algorithms and various metaheuristics. Similarly, P. Liu et al., 2018 included storage devices in the timetabling with speed profiles model. In S. Yang et al., 2020, the timetabling with speed profiles model is solved alongside a passenger assignment problem via bi-objective optimization using a metaheuristic, later extended in Huang, Liao, and Gao, 2021 to multiple interconnected lines. Finally, Gupta, Van Parys, and Tobin, 2023 present a timetabling with energy recovery, where the energy is approximated via a data-driven approach.

However, these works often do not adequately consider the transmission losses in a DC power network, assuming they can be neglected due to the distance between trains. This assumption can lead to issues. For example, with low-frequency timetables where the distance between trains is greater, the energy recovered from braking might not be sufficient to meet the energy required by accelerating trains. In their study to minimize the energy consumption of a DC railway system via voltage control on substations, Raghunathan et al., 2014 provide a model to represent the transmission losses accurately. Also for timetabling, a joint approach that accounts for power grid losses is essential to prevent overestimating recovered energy and to ensure proper network operation.

Comprehensive reviews of timetabling and energy efficiency in railway systems can be found in Scheepmaker, Goverde, and Kroon, 2017. Arboleya et al., 2020 provide a review of railway feeding infrastructures, and Kang et al., 2024 offer an up-to-date review of underground train timetabling and rescheduling problems.

In this paper, we examine energy-efficient timetabling in underground railway systems while considering transmission network losses. Our model integrates the synchronization of braking and accelerating trains with the power flow dynamics in a DC network, resulting in a non-convex mixed-integer quadratically constrained problem. To address the non-convexity, we develop a heuristic that, through linear approximation and relaxation, converges to a solution that satisfies the physical dynamics of the problem. Our algorithm ensures that the timetabling solutions we

propose are both energy-efficient and practical for real-world implementation.

The remainder of this paper is structured as follows: Section 2 presents the energy-related problem for a fixed timetable. We introduce and compare the current flow problem with the power flow problem in a fixed timetable, introduce the heuristic, and analyze its computational performance. In Section 3, we present the integrated model, considering timetabling and physical dynamics, along with the different solving strategies. Section 4 presents our computational study, and finally, in Section 5, we conclude and suggest possible extensions.

2 Simulation Model for a Fixed Timetable

In this section we focus on the physical model of the electricity flow for a fixed train configuration at a particular moment, i.e., the power flow in the railway network for a given timetable. First, we start by describing in a general way the network that will be considered, and later we compare two formulations: First, Power Flow (PF) and second, Current Flow (CF). Both models are, in their standard form, non-convex quadratic continuous models. Moreover, given our network characteristics it is necessary to include new binary variables and constraints, making the formulations to be mixed-integer non-convex, that are known to be NP-hard in general. Therefore, for the current flow problem (that presented a better computational performance) we develop a heuristic based on a linearization of the model that can tackle this problem in an efficient way and allows us to later integrate timetable planning with the power flow problem.

2.1 Power Network

In this work we study an underground railway network that operates on a DC power grid. In a general way, the network is composed by power stations, the railways connecting them, and a power source for each power station. Additionally, in the network are trains that could be braking or accelerating and that change their location along the planning horizon, making the network dynamic with respect to time. Let us represent the dynamic network by a graph $G(t)$ for t in the time horizon T, and let $A(G(t))$ and $V(G(t))$ be the set of arcs and set of nodes of the graph, respectively. The set $V(G(t))$ is partitioned into:

- Power stations and power station sources, denoted by $V^{PS}(G)$ and $V^{PSS}(G)$, respectively. These nodes do not change over time.
- Braking trains and accelerating trains: The trains position, and behavior, changes over time. These nodes will be denoted by $V^{BT}(G(t))$ and $V^{AT}(G(t))$, respectively.

Finally, let us define by $A^{PS}(G)$ the arcs connecting the power stations with their respective source, this set represent the only time-independent arcs of the graph. Figure 1 represents an example network for a given time. Here, we have 2 accelerating trains depicted in violet and one braking train in yellow driving between 3 stations. The powerstation source nodes are colored orange, the power station nodes teal. In what follows we will omit the time dependency and focus on the network for a particular time, in Section 3 the time dependency will be again considered. Each node $u \in V(G)$, following Watt's law, is characterized by power P_u, current I_u, and voltage U_u. For the particular case of the power station sources, the voltage is a fixed quantity that is known beforehand and will be denoted by $\mathcal{U}$. We also assume that all the power station sources have equal fixed voltages. The power flows along the arcs of the network, for a particular arc $(u, v) \in A(G)$ this can be represented as current flow I_{uv}, or as power flow, going out of node u towards node v, P_{uv}^{from} and arriving at node v from node u, P_{uv}^{to}.

The power flows from higher to lower voltage. However, when the voltage of the power station is higher than the voltage of its respective source a power flow will translate into energy inefficiency for our network. Therefore, to avoid this situation there are gates that only allow

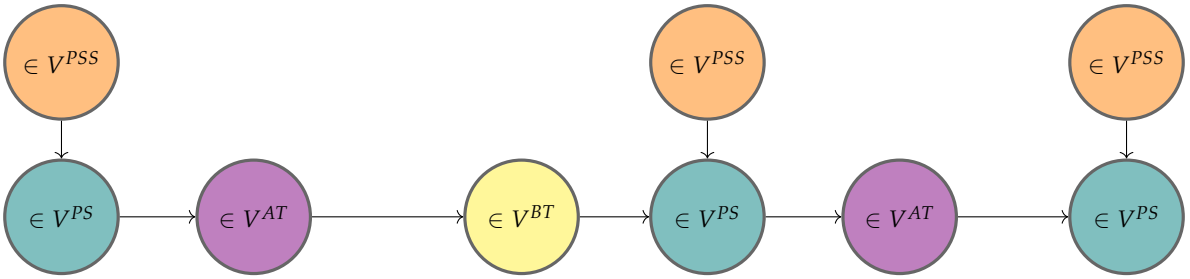


Figure 1: Transmission graph G.

power to flow from power station sources to power stations and prevent flow in the opposite direction. From a modeling point of view, this is achieved by incorporating a binary variable δ_{uv} for all power station arcs $(u, v) \in A^{PS}(G)$ that will activate the arc only if the voltage of the power station source is higher than the voltage in the power station. Finally, the losses along each arc $(u, v) \in A(G)$ are represented with an admittance parameter c_{uv}. The value of this parameter depends on the length of the power line and its material, so we can break it down further to

$$c_{uv} = \frac{\omega}{d_{uv}},$$

where ω is the material constant of the power line and d_{uv} represents its physical length.

2.2 Current and Power Flow Comparison

Now, we are interested in comparing the computational performance of the models PF and CF, and later incorporate the more performant one with the timetable scheduling problem. First, we present the models and later present some computational results to show that the current flow model has a considerably better performance in our experiments.

2.2.1 Current Flow Model CF

Following closely the formulation presented by Hager and Kuen, 2024, and taking into consideration the specific characteristics of our network, the current flow model can be written as:

$$\min_{I,U,\delta} \sum_{(u,v)\in A^{PS}(G)} I_{uv} \tag{CF.1}$$

$$\text{s.t. } I_v = \sum_{w\in\delta^{\text{out}}(v)} I_{vw} - \sum_{u\in\delta^{\text{in}}(v)} I_{uv}, \qquad v \in V^{BT} \cup V^{AT}(G), \tag{CF.2}$$

$$0 = \sum_{w\in\delta^{\text{out}}(v)} I_{vw} - \sum_{u\in\delta^{\text{in}}(v)} I_{uv}, \qquad v \in V^{PS}(G), \tag{CF.3}$$

$$U_u \cdot I_u \leq -p_u, \qquad u \in V^{BT} \cup V^{AT}(G), \tag{CF.4}$$

$$I_{uv} = c_{uv}(U_u - U_v), \qquad (u,v) \in A(G) \setminus A^{PS}(G), \tag{CF.5}$$

$$I_{uv} \geq 0, \qquad (u,v) \in A^{PS}(G), \tag{CF.6}$$

$$\underline{U}_u \leq U_u \leq \overline{U}_u, \qquad u \in V(G) \setminus V^{PSS}(G), \tag{CF.7}$$

$$I_{uv} \geq c_{uv}(\mathcal{U} - U_v), \qquad (u,v) \in A^{PS}(G), \tag{CF.8}$$

$$I_{uv} - M\delta_{uv} \leq c_{uv}(\mathcal{U} - U_v), \qquad (u,v) \in A^{PS}(G), \tag{CF.9}$$

$$I_{uv} \leq M(1-\delta_{uv}), \qquad (u,v) \in A^{PS}(G), \tag{CF.10}$$

$$\mathcal{U} - U_v \leq M(1-\delta_{uv}), \qquad (u,v) \in A^{PS}(G), \tag{CF.11}$$

$$\delta \in \{0,1\}. \tag{CF.12}$$

In here, the objective (CF.1) is to minimize the current injected into the network by the power station sources. Constraints (CF.2) and (CF.2) represent the current flow conservation, i.e., Kirchhoff's first law, for the trains and the power stations, respectively. (CF.4) model the demand (supply) of the accelerating (braking) trains. Ohm's law, for the arcs that are not connected to a power station source, is modeled with constraints (CF.5). Some variables are bounded by constraints (CF.6) and (CF.7). Finally, constraints (CF.8)-(CF.11) model the activation of the arcs connecting the power stations with their respective source. If these gate variables are set to one, Ohm's law is neglected at the associated arc (CF.9), (CF.8), no flow is permitted on the arc (CF.10) and the voltage on the power station node has to be lower than the fixed voltage at the power source node (CF.11).

The complexity of Model (CF) can be summarized by the non-convexity present in constraints (CF.5) and the combinatorial nature of the gate activation constraints. This model serves as the basis for evaluating the energy consumption of the timetables. Comparisons between the traction current data measured and simulated by the VAG (Verkehrs-Aktiengesellschaft Nürnberg), the operator of public transport in Nuremberg, Germany show that the deviation is only around 1%.

2.2.2 Power Flow Model PF

Using Watt's law and following the work of Gan and Low, 2014, we can reformulate Model (CF) obtaining the Power Flow problem:

$$\min_{P,U,\delta} \sum_{(u,v)\in A^{PS}(G)} \frac{P^{TO}_{uv}}{U_v} \tag{PF.1}$$

$$\text{s.t. } p_u \leq \sum_{u\in\delta^{\text{in}}(v)} P^{TO}_{uv} - \sum_{w\in\delta^{\text{out}}(v)} P^{FROM}_{vw}, \qquad v \in V^{BT} \cup V^{AT}(G), \tag{PF.2}$$

$$0 = \sum_{u\in\delta^{\text{in}}(v)} P^{TO}_{uv} - \sum_{w\in\delta^{\text{out}}(v)} P^{FROM}_{vw}, \qquad v \in V^{PS}(G), \tag{PF.3}$$

$$P^{FROM}_{uv} = U_u(U_u - U_v)c_{uv}, \qquad (u,v) \in A(G) \setminus A^{PS}(G), \tag{PF.4}$$

$$P^{TO}_{uv} = U_v(U_u - U_v)c_{uv}, \qquad (u,v) \in A(G) \setminus A^{PS}(G), \tag{PF.5}$$

$$P^{TO}_{uv} \geq 0, \qquad (u,v) \in A^{PS}(G), \tag{PF.6}$$

$$\underline{U}_u \leq U_u \leq \overline{U}_u, \qquad u \in V(G), \tag{PF.7}$$

$$P^{TO}_{uv} \geq U_v(\mathcal{U} - U_v)c_{uv}, \qquad (u,v) \in A^{PS}(G), \tag{PF.8}$$

$$P^{TO}_{uv} - M\delta_{uv} \leq U_v(\mathcal{U} - U_v)c_{uv}, \qquad (u,v) \in A^{PS}(G), \tag{PF.9}$$

$$P^{TO}_{uv} \leq M(1-\delta_{uv}), \qquad (u,v) \in A^{PS}(G), \tag{PF.10}$$

$$\mathcal{U} - U_v \leq M(1-\delta_{uv}), \qquad (u,v) \in A^{PS}(G), \tag{PF.11}$$

$$\delta \in \{0,1\}. \tag{PF.12}$$

In this case, the objective (PF.1) (to minimize the current injected) is non-linear. Kirchhoff's first law is represented by constraints (PF.2) and (PF.3). Constraints (PF.4) and (PF.5) model Ohm's law. Constraints (PF.6) and (PF.7) bound the power flowing into power station sources and the voltage in each node, respectively. Finally, constraints (PF.8)-(PF.11) are, in an equivalent way to Model (CF), the gate activation constraints. In this case, the complexity of Model (PF) lies in the non-convexity present in (PF.1), (PF.4), (PF.5), (PF.8), and (PF.9), and the combinatorial nature of the gate activation constraints.

2.2.3 Computational Performance

Before studying the integrated timetable scheduling and power flow problem, we want to determine which model is computationally more efficient. For this, we compared the runtime of Model (PF) to the runtime of Model (CF) on 5 configurations with 100 randomly created instances each. The configuration names encode the composition of the transmission network G in the scheme $(|V^{PS}(G)|, |V^{AT}(G)|, |V^{BT}(G)|, \max_{u\in V^{BT}\cup V^{AT}(G)} |p_u|)$.

All performance tests in this section were implemented in Python 3.10.13 using Gurobi 11.0.0 with standard parameter setting to solve mixed-integer quadratic problems. We performed the calculations on a laptop with an Intel i7-1165G7 CPU, 32 GB RAM, 4 cores and 2.80 GHz base frequency.

The results in Table 1 show the average runtime of all instances in one configuration in seconds. It gets clear that Model (CF) can be solved significantly more efficiently than Model (PF) with an average speedup of 97%. Therefore, we concentrate in the following on the current flow model.

2.3 Current Model Linearization

The only non-linearity in Model (CF) is in constraints (CF.4). However, we can approximate these constraints by replacing in these inequalities the voltage variable U_u with a constant value

Table 1: Runtime test solving Model (PF) compared to solving Model (CF) using Gurobi.

Config	PF	CF
(10, 5, 5, 2)	1.705	0.055
(15, 7, 7, 2)	7.918	0.152
(15, 12, 7, 2)	21.713	0.155
(10, 5, 5, 10)	1.288	0.069
(15, 7, 7, 10)	4.913	0.090
Average	7.508	0.104

Y_u. Particularly, if $Y_u = \overline{U}_u$ for braking trains and $Y_u = \underline{U}_u$ for accelerating trains we obtain an inner approximation of the problem. With this, we can obtain the following mixed-integer linear problem as an inner approximation of Model (CF)

$$\min_{I,U,\delta} \sum_{u \in V^{PS}(G)} \sum_{v \in \delta^{\text{out}}(u)} I_{uv} \tag{LCF.1}$$

$$\text{s.t. } I_u = \sum_{v \in \delta^{\text{out}}(u)} I_{uv} - \sum_{v \in \delta^{\text{in}}(u)} I_{vu}, \qquad u \in V^{BT} \cup V^{AT}(G), \tag{LCF.2}$$

$$0 = \sum_{v \in \delta^{\text{out}}(u)} I_{uv} - \sum_{v \in \delta^{\text{in}}(u)} I_{vu}, \qquad u \in V^{PS}(G), \tag{LCF.3}$$

$$I_u \leq \frac{-p_u}{Y_u}, \qquad u \in V^{BT} \cup V^{AT}(G), \tag{LCF.4}$$

$$I_{uv} = c_{uv}(U_u - U_v), \qquad (u,v) \in A(G) \setminus A^{PS}(G), \tag{LCF.5}$$

$$I_{uv} \geq 0, \qquad (u,v) \in A^{PS}(G), \tag{LCF.6}$$

$$\underline{U}_u \leq U_u \leq \overline{U}_u, \qquad u \in V(G) \setminus V^{PSS}(G), \tag{LCF.7}$$

$$I_{uv} \geq c_{uv}(\mathcal{U} - U_v), \qquad (u,v) \in A^{PS}(G), \tag{LCF.8}$$

$$I_{uv} - M\delta_{uv} \leq c_{uv}(\mathcal{U} - U_v), \qquad (u,v) \in A^{PS}(G), \tag{LCF.9}$$

$$I_{uv} \leq M(1 - \delta_{uv}), \qquad (u,v) \in A^{PS}(G), \tag{LCF.10}$$

$$\mathcal{U} - U_v \leq M(1 - \delta_{uv}), \qquad (u,v) \in A^{PS}(G), \tag{LCF.11}$$

$$\delta \in \{0,1\}. \tag{LCF.12}$$

Model (LCF) overestimates the real current demand on accelerating trains and underestimates the current supply on braking trains.

2.4 Linear Current Model Heuristic HEU

Even if the linearization used to obtain Model (LCF) is naive and not tight, we can make use of an iterative procedure to improve the approximated constraints and converge to a solution that is feasible for Model (CF). Now, we introduce a heuristic that makes the current demands and supplies approach their actual values, and we also present some remarks regarding the convergence of the heuristic.

As it was previously stated, Model (LCF) is an inner approximation of Model (CF). The heuristic gradually adapts the demand constraints (LCF.4) towards the optimal solution of Model (CF). The heuristic starts by using the voltage's bounds, i.e., $Y_u = \overline{U}_u$ for braking trains and $Y_u = \underline{U}_u$ for accelerating trains. Then, we can solve Model (LCF) and use the solution to update the values of Y_u. This procedure is repeated until the change in Y is smaller than a tolerance ϵ. The heuristic is summarized in Algorithm 1, where $\tilde{I}$ is the optimal solution to Model (CF).

Algorithm 1: Linear Current Model Heuristic HEU

Data: Graph G, Parameters c, p, $\underline{U}$, $\overline{U}$, $\epsilon > 0$
Result: Optimal current flow $I^* \approx \tilde{I}$

$Y_u \leftarrow \underline{U}_u, \quad u \in V^{AT}(G)$;
$Y_u \leftarrow \overline{U}_u, \quad u \in V^{BT}(G)$;
Solve Model LCF;
while $\|U^* - Y\|_2 > \epsilon$ **do**
  $Y_u \leftarrow U^*_u, \quad u \in V^{BT} \cup V^{AT}(G)$;
  (LCF.4) $\leftarrow I_u \leq \frac{-p_u}{Y_u}, \quad u \in V^{BT} \cup V^{AT}(G)$;
  Solve Model LCF

Assuming we can use all the available current from braking trains, and that all accelerating trains use exactly the current they demand, we can calculate the solution for Model (LCF) efficiently in $\mathcal{O}(|A^{PS}(G)|^2 + |A(G)|)$.

Remark 2.1. *For given values for δ_{uv} and I_{uv} $\forall (u,v) \in A^{PS}(G)$ and I_u $\forall u \in V^{BT} \cup V^{AT}(G)$ one can calculate all currents and voltages in the network G in $\mathcal{O}(|A^{PS}(G)|^2 + |A(G)|)$, if a solution to Model (CF) exists.*

To better represent the order of the nodes, we introduce an alternative notation for the nodes in G, see Figure 2. The colors of the nodes are chosen as in Figure 1. The power stations are enumerated from 1 to n, where power station k is adjacent to power station $k-1$ and power station $k+1$ for all $k \in 2, ..., n-1$. For each $k \in [n]$ the corresponding power station source node is named k, the power station node is named k_0. The trains on a path between two neighboring power station nodes k_0 and $k+1_0$ get enumerated from 1 to m_k, where m_k is the number of trains between the two power station nodes. The corresponding train node is named k_l for $l \in [m_k]$ and $k \in [n]$. For all $k \in [n-1]$, we can calculate the voltage of a train node l' from its left-next

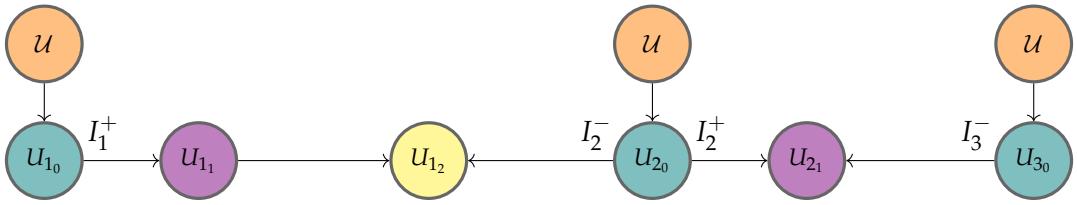


Figure 2: Ordered transmission graph G.

and from its right-next power station node

$$U_{k_0} - \frac{d_{k_0,k_{l'}}}{\omega} I_k^+ - \sum_{l=1}^{l'-1} \frac{d_{k_l,k_{l'}}}{\omega} I_{k_l} = U_{k_{l'}} = U_{k+1_0} - \frac{d_{k_{l'},k+1_0}}{\omega} (-\sum_{l=1}^{m_k} I_{k_l} - I_k^+) - \sum_{l=l'+1}^{m_k} \frac{d_{k_{l'},k_l}}{\omega} I_{k_l}.$$

This allows us to obtain the value for I_k^+ for all $k \in [n-1]$ depending on the voltages at the nearest power station nodes

$$I_k^+ = -\sum_{l=1}^{m_k} \frac{d_{k_l,k+1_0}}{d_{k_0,k+1_0}} I_{k_l} + \frac{\omega}{d_{k_0,k+1_0}} (U_{k_0} - U_{k+1_0}). \tag{1a}$$

If we consider only two power station sources, this reduces to

$$I_k^+ = -\sum_{l=1}^{m_k} \frac{d_{k_l,k+1}}{d_{k,k+1}} I_{k_l}, \tag{1b}$$

since the current from the power station source node equals the current from the power station node. Assume $n \geq 3$. Ohm's law at each power station arc yields

$$\begin{aligned} U_{1_0} &= \mathcal{U} - \frac{d_{1,1_0}}{\omega} I_1^+, \\ U_{k_0} &= \mathcal{U} - \frac{d_{k,k_0}}{\omega}(-\sum_{l=1}^{m_{k-1}} I_{k-1_l} - I_{k-1}^+ + I_k^+), \quad \forall k \in \{2, ..., n-1\}, \\ U_{n_0} &= \mathcal{U} - \frac{d_{n,n_0}}{\omega}(-\sum_{l=1}^{m_{n-1}} I_{n-1_l} - I_{n-1}^+). \end{aligned} \tag{1c}$$

Define

$$\begin{aligned} a_k &:= d_{k,k+1}, && \forall k \in [n-1], \\ b_k &:= -d_{k,k_0}, && \forall k \in [n-1], \\ r_1 &:= -\sum_{l=1}^{m_1} d_{1_l,2} I_{1_l}, \\ r_k &:= -\sum_{l=1}^{m_k} d_{k_l,k+1} I_{k_l} + \sum_{l=1}^{m_{k-1}} d_{k,k_0} I_{k-1_l}, && \forall k \in \{2, ..., n-1\}. \end{aligned}$$

Combining (1a) and (1c) establishes a relationship between the current flows from power station sources given by the following equation system:

$$\underbrace{\begin{pmatrix} a_1 & b_2 & 0 & \cdots & 0 \\ b_2 & a_2 & b_3 & \ddots & \vdots \\ 0 & b_3 & a_3 & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & b_{n-1} \\ 0 & \cdots & 0 & b_{n-1} & a_{n-1} \end{pmatrix}}_{\mathcal{A}} \underbrace{\begin{pmatrix} I_1^+ \\ I_2^+ \\ \vdots \\ I_{n-2}^+ \\ I_{n-1}^+ \end{pmatrix}}_{I^+} = \underbrace{\begin{pmatrix} r_1 \\ r_2 \\ \vdots \\ r_{n-2} \\ r_{n-1} \end{pmatrix}}_{r}. \tag{2}$$

Being a triangular matrix with non-zero entries on the diagonals, $\mathcal{A}$ is invertible and we can calculate I^+ using Thomas algorithm in $\mathcal{O}(n^2)$. I^- is then given by the equations

$$I_k^- = -\sum_{l=1}^{m_{k-1}} I_{k-1_l} - I_{k-1}^+ \quad \forall k \in \{2, ..., n\}$$

and arc currents by Kirchhoff's first law. Lastly, $U_{k_{l'}}$ for $k \in [n-1]$, $l' \in [m_k]$ can be calculated as

$$U_{k_{l'}} = \mathcal{U} - \frac{d_{k,k_0}}{\omega}(I_k^- + I_k^+) - \sum_{l=1}^{l'} \frac{d_{l-1,l}}{\omega} I_k^+ - \sum_{l=1}^{l'-1} \frac{d_{l,l+1}}{\omega} I_{k_l}.$$

Remark 2.2. *If the network only consists of two power stations and accelerating trains, Algorithm 1 converges to the optimal solution of Model* (CF)*, if it exists.*

We use the notation of Remark 2.1. The voltage at each accelerating train node $U_{1_{l'}}$ for all $l \in [m_k]$ can be calculated as follows. From (1b), we already know the current coming from power station 1.

$$I_1^+ = -\sum_{l=1}^{m_1} \frac{d_{1_l,2}}{d_{1,2}} I_{1_l}. \tag{3a}$$

Now we can obtain the voltage at the accelerating train $1_{l'}$.

$$U_{1_{l'}} = \mathcal{U} + \frac{d_{1,1_{l'}}}{\omega} \sum_{l=1}^{m_1} \frac{d_{1_l,2}}{d_{1,2}} I_{1_l} - \sum_{l=1}^{l'-1} \frac{d_{1_l,1_{l'}}}{\omega} I_{1_l}. \tag{3b}$$

To proof that Algorithm 1 converges, we show that the voltage at each train node increases monotonously which implies a monotonously decreasing objective function that is bounded from below by 0.

$$U_{1_{l'}}^{i+1} - U_{1_{l'}}^{i} = -\frac{d_{1,1_{l'}}}{\omega} \sum_{l=l'}^{m_1} \underbrace{\frac{d_{1_l,2}}{d_{1,2}}}_{>0} p_{1_l} \left(\frac{1}{U_{1_l}^{i}} - \frac{1}{U_{1_l}^{i-1}}\right) - \sum_{l=1}^{l'-1} \underbrace{\frac{d_{1,1_{l'}} d_{1_l,2} - d_{1_l,1_{l'}} d_{1,2}}{\omega d_{1,2}}}_{= \frac{d_{1,l} d_{l',2}}{\omega d_{1,2}} > 0} p_{1_l} \left(\frac{1}{U_{1_l}^{i}} - \frac{1}{U_{1_l}^{i-1}}\right).$$

The statement follows by induction and the fact that all voltages increase in the first iteration because they were initialy set to their lower bound. All solution voltages in each iteration lie between their lower bound and the optimal solution of Model (CF) and are therefore feasible.

The algorithm converges to $U_{1_l}^{i+1} - U_{1_l}^{i} \to 0$ and thus towards a solution $(\tilde{U}, \tilde{I})$ that satisfies $\tilde{U}_{1_l} \tilde{I}_{1_l} = -p_{1_l}$ for all train nodes $l \in [m_1]$. This is now a feasible solution for Model (CF) for which all demand constraints are tight. The only way to improve this solution would be by increasing the currents at the train nodes further without increasing the value for $I_{1_l} U_{1_l}$. Using previous transformations, we have for all $l' \in [m_1]$

$$I_{1_{l'}} U_{1_{l'}} = I_{1_{l'}} \left(\mathcal{U} - \frac{d_{1,1_{l'}}}{\omega} \sum_{l=l'}^{m_1} \frac{d_{1_l,2}}{d_{1,2}} I_{1_{l'}} - \sum_{l=1}^{l'-1} \frac{d_{1,l} d_{l',2}}{\omega d_{1,2}} I_{1_{l'}}\right).$$

Taking the derivative of this function with respect to $I_{1_{l'}}$ yields

$$\frac{\partial I_{1_{l'}} U_{1_{l'}}}{\partial I_{1_{l'}}} = \mathcal{U} - \frac{d_{1,1_{l'}}}{\omega} \sum_{l=l'}^{m_1} \frac{d_{1_l,2}}{d_{1,2}} I_{1_l} - \sum_{l=1}^{l'-1} \frac{d_{1,l} d_{l',2}}{\omega d_{1,2}} I_{1_l} - \frac{d_{1,1_{l'}}}{\omega} \frac{d_{1_{l'},2}}{d_{1,2}} I_{1_{l'}} = U_{1_{l'}} - \frac{d_{1,1_{l'}}}{\omega} \frac{d_{1_{l'},2}}{d_{1,2}} I_{1_{l'}},$$

which is greater than zero. Hence, increasing the current at a train node comes with increasing the power at the node and there is no better solution to Model (CF) than $(\tilde{U}, \tilde{I})$ found by Algorithm 1.

In Table 2 we compare the runtime of solving Model (CF) to the runtime of the developed heuristic (HEU). We use the same setup as in Section 2.2.2, the stated runtimes are average times over all instances of one configuration in seconds. The configurations are again named $(|V^{PS}(G)|, |V^{AT}(G)|, |V^{BT}(G)|, \max_{u \in V^{BT} \cup V^{AT}(G)} |p_u|)$. The column $ITER$ represents the average number of steps in the heuristic, so the number of linear models we have to solve to get to an ϵ value of $1e^{-5}$. Note that on average we only have to solve 6 linear models to obtain a close approximation to the optimal solution to the non-linear Model (CF). This, together with the fact that we can utilize a warmstart for the linear models in each steps, yields a significant speedup of 87% in the runtime on average.

Table 2: Runtime test solving Model CF with Gurobi compared to the runtime of Algorithm 1 and number of steps in the heuristic.

Instance	ITER	CF	HEU
(10, 5, 5, 2)	3.3	0.048	0.009
(25, 10, 10, 2)	3.9	0.197	0.020
(50, 25, 25, 2)	4.1	4.322	0.045
(100, 50, 50, 2)	4.5	11.365	0.102
(10, 5, 5, 10)	6.7	0.045	0.017
(25, 10, 10, 10)	7.2	0.137	0.034
(50, 25, 25, 10)	9.7	2.081	0.109
(100, 50, 50, 10)	10.3	6.375	0.245
Average	6.2	3.071	0.072

3 Embedding into Timetabling Model

We now aim to integrate the developed model for the DC power network with the underground train timetable optimization model. By shifting departure times and lengthening running times we can improve both the driving behaviour of the trains and the usage of braking energy in the system to reduce the total energy consumption.

3.1 Modeling a Feasible Timetable

The modeling of a feasible timetable is taken from Gemander, Bärmann, and Martin, 2023. Binary variables x_{kij} are introduced which in the case $x_{kij} = 1$ indicate that the train on leg k departs at time i with running time j. The set X contains all x representing a feasible timetable. The way how X is defined is explained in detail in Gemander, Bärmann, and Martin, 2023. The tuple (k, i, j) is called a departure configuration. Each departure configuration has a power profile assigned to it. This specifies the power demand of (k, i, j) in each second $i, i+1, ..., i+j$. For each $t \in [i, i+j]$ the power demand of the train with departure configuration (k, i, j) is known and given as p_{kijt}. Knowing the departure time, running time and power profile we can also state the location of the train with departure configuration (k, i, j) at time $t \in [i, i+j]$.

3.2 Transmission Graph Adjustment

For each timestamp $t \in T$, all departure configurations (k, i, j) that demand power at time t are grouped into buckets B. Each bucket is assigned to a node in the transmission graph and is denoted by $B(u)$. The potential power demand of the bucket node is defined as the mean power demand of the departure configurations inside the bucket $p_u = \frac{\sum_{(kij) \in B(u)} p_{kijt}}{|B(u)|}$. It is activated if one of the assigned departure configurations is selected.

The departure configurations are not only grouped by their power demand, but also by their location at time t. We term the maximum proximity between two departure configurations in one bucket the *resolution* of the transmission graph, where the proximity of two departure configurations (k_1, i_1, j_1) and (k_2, i_2, j_2) is defined as

$$\left\| \begin{pmatrix} \frac{t-i_1}{2j_1} \\ p_{k_1 i_1 j_1 t} \end{pmatrix} - \begin{pmatrix} \frac{t-i_2}{2j_2} \\ p_{k_2 i_2 j_2 t} \end{pmatrix} \right\|_2$$

for all $t \in T$. We incorporate at least one bucket for each arc that contains a departure configuration.

Instead of train nodes, the transmission network now is filled with bucket nodes located at the center of the departure configurations it contains as shown in Figure 3. This prevents

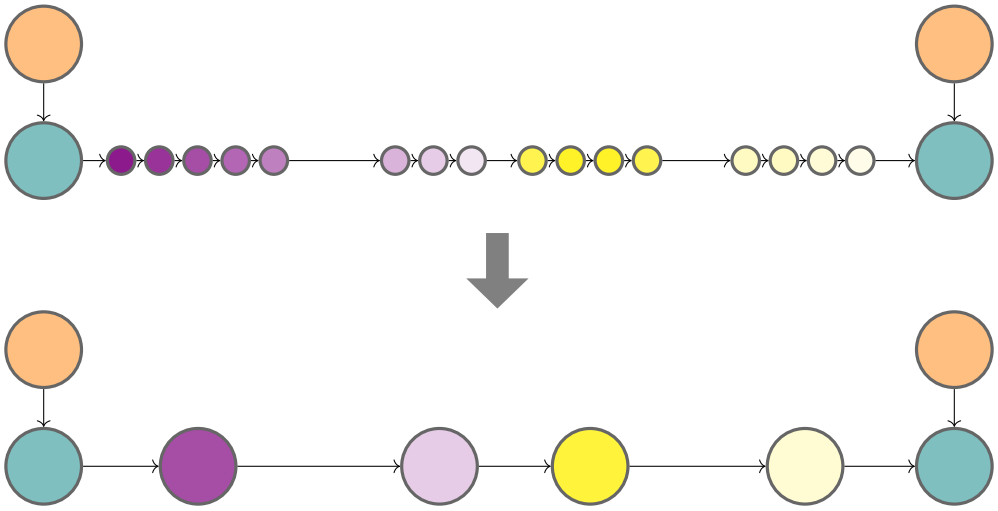


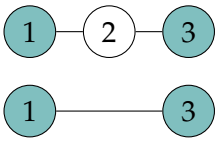
Figure 3: Grouping departure configuration nodes into buckets for transmission graph G_t with timetabling.

large admittance parameters causing numerical issues which would arise if two departure configurations are located too close to each other.

3.3 Admittance Adjustment

When a timetable has been selected, many bucket nodes will have power demand zero, since none of the assigned departure configurations has been chosen. The addition of nodes with zero power demand into the network should have no impact to the overall flow. To archieve this we have to properly determine the admittance parameters. In an example of three nodes where node 2 has zero power demand, we can state the following relationship between the voltages.

$$\begin{aligned}(U_1 - U_2)c_{12} &= (U_2 - U_3)c_{23} \\ \Rightarrow U_2 &= \frac{U_1 c_{12} + U_3 c_{23}}{c_{12} + c_{23}}.\end{aligned}$$



The current should be equal for all arcs $(1,2)$, $(2,3)$ and $(1,3)$. Therefore the admittance parameter when deleting node 2 can be calculated as

$$\begin{aligned}(U_1 - U_3)c_{13} &= (U_1 - U_2)c_{12} \\ \Rightarrow c_{13} &= \frac{c_{12}c_{23}}{c_{12} + c_{23}}.\end{aligned} \tag{4}$$

As we define the admittance on an arc as a constant ω over the length d of the power line the arc represents. This fulfills (4) since

$$d_{uw} = d_{uv} + d_{vw} \quad \Rightarrow \quad \frac{\omega}{d_{uw}} = \frac{\frac{\omega}{d_{uv}}\frac{\omega}{d_{vw}}}{\frac{\omega}{d_{uv}} + \frac{\omega}{d_{vw}}}$$

holds for all $\omega, d_{uw}, d_{uv}, d_{vw} \in \mathbb{R}^+$.

3.4 Timetabling Model BASE

The mixed integer timetabling model stated in Gemander, Bärmann, and Martin, 2023 seeks to minimize the positive part of the sum of all power demands at one timestamp t for all $t \in T$

$$\min_{z,x} \sum_{t \in T} z_t \tag{BASE.1}$$

$$\text{s.t.} \sum_{u \in V^{BT} \cup V^{AT}(G_t)} \sum_{(k,i,j) \in B(u)} p_{kijt} x_{kij} \leq z_t, \qquad t \in T, \tag{BASE.2}$$

$$z \geq 0 \tag{BASE.3}$$

$$x \in X. \tag{BASE.4}$$

The z-variables can be interpreted as the amount of power that is requested from the power-stations, assuming there is no limit on the voltages in the network and no transmission loss on the lines. This model underestimates the actual power demand of the grid and serves as a base model for the study in this paper.

3.5 Integrated Timetabling Current Flow Model CM

The goal is to integrate the constraints implied by physical laws into the (BASE) model. To archieve this, we use Model CF as the evaluation function for the timetables.

$$\min_{I,U,\delta,x} \sum_{t \in T} \sum_{(u,v) \in A^{PS}(G_t)} I_{uv} \tag{CM.1}$$

$$\text{s.t. } I_u = \sum_{v \in \delta^{\text{out}}(u)} I_{uv} - \sum_{v \in \delta^{\text{in}}(u)} I_{vu}, \qquad u \in V^B(G_t),\ t \in T, \tag{CM.2}$$

$$0 = \sum_{v \in \delta^{\text{out}}(u)} I_{uv} - \sum_{v \in \delta^{\text{in}}(u)} I_{vu}, \qquad u \in V^{PS}(G_t),\ t \in T, \tag{CM.3}$$

$$U_u \cdot I_u \leq \sum_{(k,i,j) \in B(u)} -p_{kijt} x_{kij}, \qquad u \in V^B(G_t),\ t \in T, \tag{CM.4}$$

$$I_{uv} = c_{uv}(U_u - U_v), \qquad (u,v) \in A(G_t) \setminus A^{PS}(G_t),\ t \in T, \tag{CM.5}$$

$$I_{uv} \geq 0, \qquad (u,v) \in A^{PS}(G_t),\ t \in T, \tag{CM.6}$$

$$\underline{U}_u \leq U_u \leq \overline{U}_u, \qquad u \in V(G_t) \setminus V^{PSS}(G_t),\ t \in T, \tag{CM.7}$$

$$I_{uv} \geq c_{uv}(\mathcal{U} - U_v), \qquad (u,v) \in A^{PS}(G_t),\ t \in T, \tag{CM.8}$$

$$I_{uv} - M\delta_{uv} \leq c_{uv}(\mathcal{U} - U_v), \qquad (u,v) \in A^{PS}(G_t),\ t \in T, \tag{CM.9}$$

$$I_{uv} \leq M(1 - \delta_{uv}), \qquad (u,v) \in A^{PS}(G_t),\ t \in T, \tag{CM.10}$$

$$\mathcal{U} - U_v \leq M(1 - \delta_{uv}), \qquad (u,v) \in A^{PS}(G_t),\ t \in T \tag{CM.11}$$

$$\delta \in \{0,1\}, \tag{CM.12}$$

$$x \in X. \tag{CM.13}$$

The coupling of the models takes place in the Constraints (CM.4). For each $t \in T$, the demand at each bucket node $u \in V^B(G_t)$ is given by the power demand p_{kijt} of the active departure configuration x_{kij} in side the bucket. The buckets are defined such that there is at most one departure configuration active inside it. Fixing the x variables and therefore commiting to one timetable leads exactly to Model CF which was shown to be solvable efficiently in Section 2.4.

3.6 Solving Strategies

Model CM is a mixed-integer nonlinear problem and is therefore computationally very hard to solve as a whole. This section provides different strategies, relaxations and algorithms, to

deal with this complexity. First, we try to relax the parts of the model, that make it hard to solve, namely the binary part in the gate constraints (CM.10), (CM.11) and the quadratic terms in (CM.4).

3.6.1 Relaxing Gate Constraints

For each $t \in T$, the gate constraints (CM.8), (CM.9), (CM.10), (CM.11) and (CM.12) can be relaxed to

$$I_{uv} \geq c_{uv}(\mathcal{U}_u - U_v), \quad (u,v) \in A^{PS}(G_t), \tag{RG.1}$$

$$I_{uv} \leq c_{uv}(\overline{U}_v - U_v)\frac{\mathcal{U}_u - \underline{U}_v}{\overline{U}_v - \underline{U}_v}, \quad (u,v) \in A^{PS}(G_t) \tag{RG.2}$$

$$\sum_{u \in V^{BT} \cup V^{AT}(G_t)} \sum_{(k,i,j) \in B(u)} p_{kijt}\tilde{x}_{kij} \leq \sum_{(u,v) \in A^{PS}(G_t)} \mathcal{U}_u I_{uv}. \tag{RG.3}$$

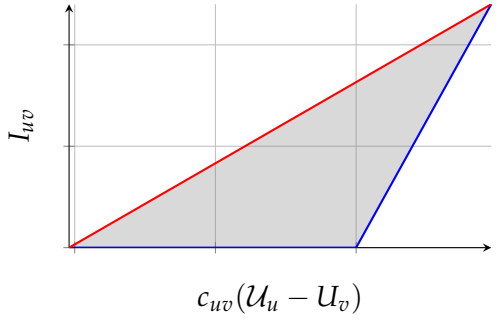


Figure 4: Gate constraint relaxation.

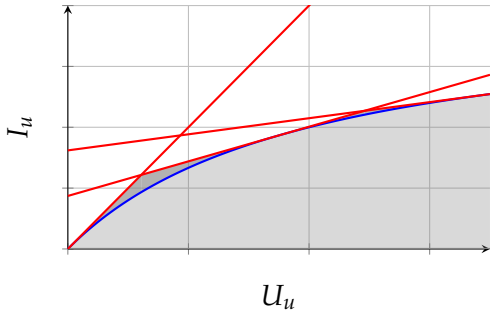


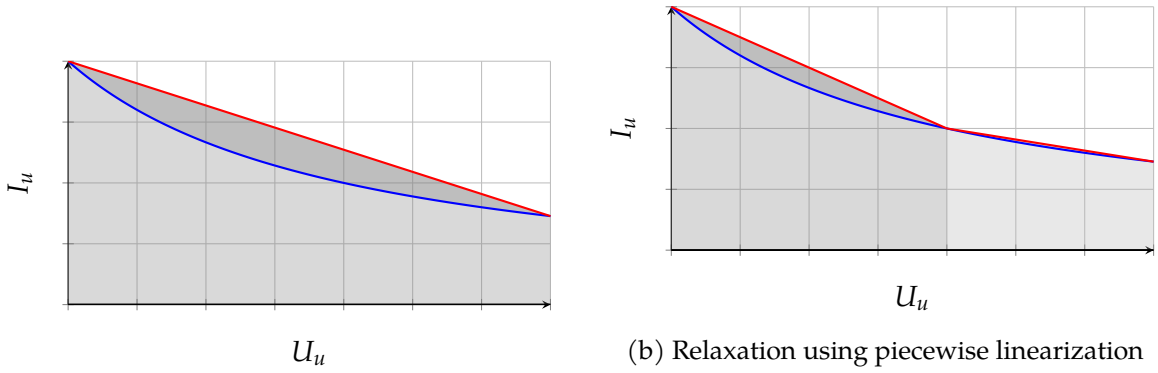
Figure 5: Feasibility set for U_u and I_u with relaxation of demand constraints at accelerating buckets, assuming $P_u = -p_u$.

The blue line in Figure 4 depicts the feasible set defined by (CM.8), (CM.9), (CM.10) and (CM.11). Inequality (RG.1) yields an overestimator for the current shown by the red line in Figure 4. The relaxation allows for some amount of current to flow from a power station source node u to a power station node v although there is a higher voltage on v than on u. In special cases, this could even lead to a higher voltage at the power station source node than at a accelerating train. To prevent the case in which less power is supplied than demanded, we introduce constraints (RG.3).

3.6.2 Relaxing Quadratic Power Balance Constraints

When relaxing the quadratic demand constraints (CM.4) it is helpful to differentiate between buckets that only contain departure configurations with positive power demand, so called accelerating buckets $V^{AB}(G_t)$ and braking buckets $V^{BB}(G_t)$, which only contain configurations with negative power demand. The main difference is that the feasibility set of U_u and I_u for accelerating buckets is convex, while it is concave for braking buckets.

Accelerating Trains The power flowing into a node $u \in V^{AB}(G_t)$ in an optimal solution only takes two possible values. If one of the departure configurations inside the bucket $(i,k,j) \in B(u)$ is selected the value is $-p_{kijt}$, 0 otherwise. We introduce a variable for the power flowing into node u, P_u. For each $t \in T$ for a chosen value $\underline{U}_u \leq \mathcal{V}_u \leq \overline{U}_u$ we introduce a cut to relax (CM.4) that tangents $I_u = \frac{P_u}{U_u}$ at $P_u = -p_u$ and $U_u = \mathcal{V}_u$. Numerical experiments show, that the three support points $\mathcal{V} \in \mathcal{S}_u := \{\underline{U}_u, \frac{\underline{U}_u + \overline{U}_u}{2}, \overline{U}_u\}$, as depicted in Figure 5 are adequate.



(a) Relaxation without using piecewise linearization

(b) Relaxation using piecewise linearization

Figure 6: Feasibility set for U_u and I_u with relaxation of demand constraints at braking buckets.

$$I_u \leq \frac{p_u}{\mathcal{V}^2} U_u + \left(\frac{-\underline{U}_u}{\mathcal{V}^2} + \frac{2}{\mathcal{V}}\right) P_u - p_u \frac{\underline{U}_u}{\mathcal{V}^2}, \quad u \in V^{AB}(G_t),\ \mathcal{V} \in \mathcal{S}_u, \tag{AD.1}$$

$$I_u \leq 0, \quad u \in V^{AB}(G_t), \tag{AD.2}$$

$$P_u \leq \sum_{(k,i,j)\in B(u)} -p_u x_{kij}, \quad u \in V^{AB}(G_t). \tag{AD.3}$$

Constraint (AD.1) is chosen to not cut of the point $I_u = P_u = 0$. This point should be feasible and is enforced by constraints (AD.3) and (AD.2) if no departure configuration is active in the bucket.

Braking Trains To relax the quadratic constraints for each $t \in T$ (CM.4) for braking trains we use

$$I_u \leq \frac{p_u}{\underline{U}_u \overline{U}_u} U_u - p_u \left(\frac{1}{\underline{U}_u} + \frac{1}{\overline{U}_u}\right), \quad u \in V^{BB}(G_t), \tag{BD.1}$$

$$I_u \cdot \underline{U}_u \leq \sum_{(k,i,j)\in B(u)} -p_u x_{kij}, \quad u \in V^{BB}(G_t). \tag{BD.2}$$

Zero or negative current flow if no departure configuration in the bucket is active is enforced by constraints (BD.2). We can increase the tightness of the relaxation using the concept of piecewise linearization. For each breaking train node $u \in V^{BB}(G_t)$ we add a binary variable γ_u indicating whether $U_u \leq \mathcal{V}_u$ or $U_u \geq \mathcal{V}_u$ holds for a chosen support point $\mathcal{V}$. When relaxing quadratic power balance constraints piecewise linear we replace constraints (BD.1) by

$$I_u \leq \frac{p_u}{\underline{U}_u \mathcal{V}} U_u - p_u \left(\frac{1}{\underline{U}_u} + \frac{1}{\mathcal{V}_u}\right) + M\gamma_u, \quad u \in V^{BB}(G_t), \tag{RP.1}$$

$$I_u \leq \frac{p_u}{\mathcal{V} \overline{U}_u} U_u - p_u \left(\frac{1}{\mathcal{V}_u} + \frac{1}{\overline{U}_u}\right) + M(1-\gamma_u), \quad u \in V^{BB}(G_t), \tag{RP.2}$$

$$U_u \leq \mathcal{V}_u + M\gamma_u, \quad u \in V^{BB}(G_t), \tag{RP.3}$$

$$U_u \geq \mathcal{V}_u + M(1-\gamma_u), \quad u \in V^{BB}(G_t), \tag{RP.4}$$

for a chosen value $\underline{U}_u \leq \mathcal{V}_u \leq \overline{U}_u$. The feasible sets for (BD.1) or (RP), respectively, is depicted in Figure 6.

The fully relaxed version of Model CM with constraints (RG),(AD) and (BD), we term RLX. The most accurate linearization of Model CM using only the relaxed constraints (RP) is termed SB.

3.6.3 Adaptive piecewise linearization

Normally, adaptivity is a crucial factor for the performance of approaches based on piecewise linear relaxations. However, in our particular case, adding support points for the piecewise linear relaxation of braking bucket demand constraints is not advantageous. This is due to the fact that the underlying quadratic problem has an immensly symmetric structure, such that essentially all nonlinear parts must be approximated with the same accuracy. Otherwise, the coarsest approximation in the relaxation can be exploited to obtain solutions with a significantly better objective value while maximizing the approximation error.

3.6.4 Benders Cut Generation

The structure of Model CM is very suitable for benders decomposition. If it weren't for the timetable variables x, the model could be solved for each timestep $t \in T$ separately. Hence, we divide the problem into a master problem and multiple subproblems. In the master problem, we choose a feasible timetable. This is done using Model BASE. The power consumption of the timetable is then evaluated by one subproblem for each timestep. The solution of these evaluation subproblems is then used to add a benders cut to the master problem. The subproblems are the relaxed (RG) + (AD) + (BD) versions of Model CM with fixed timetable variables $\tilde{x}, \tilde{z}$.

Benders Subproblem for $t \in T$

$$\min_{I,U} \sum_{(u,v)\in A^{PS}(G_t)} I_{uv} \tag{BS.1}$$

$$\text{s.t. } -\tilde{z}_t \leq -\sum_{(u,v)\in A^{PS}(G_t)} I_{uv}, \qquad (\alpha^1) \tag{BS.2}$$

$$0 = -I_u + \sum_{v\in\delta^{\text{out}}(u)} I_{uv} - \sum_{v\in\delta^{\text{in}}(u)} I_{vu}, \qquad u \in V^B(G_t), \tag{BS.3}$$

$$0 = \sum_{v\in\delta^{\text{out}}(u)} I_{uv} - \sum_{v\in\delta^{\text{in}}(u)} I_{vu}, \qquad u \in V^{PS}(G_t), \tag{BS.4}$$

$$p_u\left(\frac{1}{\underline{U}_u} + \frac{1}{\overline{U}_u}\right) \leq \frac{p_u}{\underline{U}_u\overline{U}_u}U_u - I_u, \qquad u \in V^{BB}(G_t), \qquad (\alpha_u^2) \tag{BS.5}$$

$$\sum_{(k,i,j)\in B(u)} p_{kijt}\tilde{x}_{kij} \leq -I_u \cdot \underline{U}_u, \qquad u \in V^{BB}(G_t), \qquad (\alpha_u^3) \tag{BS.6}$$

$$p_u\frac{\underline{U}_u}{\mathcal{V}^2} \leq \frac{p_u}{\mathcal{V}^2}U_u + \left(\frac{-\underline{U}_u}{\mathcal{V}^2} + \frac{2}{\mathcal{V}}\right)P_u - I_u, \qquad u \in V^{AB}(G_t),\ \mathcal{V} \in \mathcal{S}_u, \qquad (\alpha_{u\mathcal{V}}^4) \tag{BS.7}$$

$$0 \leq -I_u, \qquad u \in V^{AB}(G_t), \tag{BS.8}$$

$$\sum_{(k,i,j)\in B(u)} p_{kijt}\tilde{x}_{kij} \leq -P_u, \qquad u \in V^{AB}(G_t), \qquad (\alpha_u^5) \tag{BS.9}$$

$$c_{uv}\mathcal{U} \leq I_{uv} + c_{uv}U_v, \qquad (u,v) \in A^{PS}(G_t), \qquad (\alpha_{uv}^6) \tag{BS.10}$$

$$-c_{uv}\overline{U}_v\frac{\mathcal{U}-\underline{U}_v}{\overline{U}_v-\underline{U}_v} \leq -I_{uv} - c_{uv}U_v\frac{\mathcal{U}-\underline{U}_v}{\overline{U}_v-\underline{U}_v}, \qquad (u,v) \in A^{PS}(G_t), \qquad (\alpha_{uv}^7) \tag{BS.11}$$

$$\sum_{u\in V^{BT}\cup V^{AT}(G_t)}\ \sum_{(k,i,j)\in B(u)} p_{kijt}\tilde{x}_{kij} \leq \sum_{(u,v)\in A^{PS}(G_t)} \mathcal{U}I_{uv}, \qquad (\alpha^8) \tag{BS.12}$$

$$0 = -I_{uv} + c_{uv}(U_u - U_v), \qquad (u,v) \in A(G_t) \setminus A^{PS}(G_t), \tag{BS.13}$$

$$0 \leq I_{uv}, \qquad (u,v) \in A^{PS}(G_t), \tag{BS.14}$$

$$\underline{U}_u \leq U_u, \qquad u \in V(G_t), \qquad (\alpha_u^9) \tag{BS.15}$$

$$-\overline{U}_u \leq -U_u, \qquad u \in V(G_t). \qquad (\alpha_u^{10}) \tag{BS.16}$$

Additionaly, (BS.7) connects the variables representing the total energy consumption at time t, z_t from Model BASE with the total consumption calculated in the subproblem. As long as the total consumption is underestimated by Model BASE, the respective subproblem will be infeasible. For an extreme ray $\tilde{\alpha}$ of the dual of Model 11, we can introduce the following feasibility cut for each $t \in T$

$$
\begin{aligned}
0 \leq & -\tilde{\alpha}^1 z_t \\
& + \sum_{u \in V^{BB}(G_t)} \tilde{\alpha}^2_u p_u \left(\frac{1}{\underline{U}_u} + \frac{1}{\overline{U}_u}\right) \\
& + \sum_{u \in V^{BB}(G_t)} \tilde{\alpha}^3_u \sum_{(k,i,j) \in B(u)} p_{kijt} x_{kij} \\
& + \sum_{u \in V^{AB}(G_t)} \sum_{\mathcal{V} \in \mathcal{S}_u} \tilde{\alpha}^4_{u\mathcal{V}} p_u \frac{\underline{U}_u}{\mathcal{V}^2} \\
& + \sum_{u \in V^{AB}(G_t)} \tilde{\alpha}^5_u \sum_{(k,i,j) \in B(u)} p_{kijt} x_{kij} \\
& + \sum_{(u,v) \in A^{PS}(G_t)} \tilde{\alpha}^6_{uv} c_{uv} \mathcal{U}_u \\
& + \sum_{(u,v) \in A^{PS}(G_t)} -\tilde{\alpha}^7_{uv} c_{uv} \overline{U}_v \frac{\mathcal{U}_u - \underline{U}_v}{\overline{U}_v - \underline{U}_v} \\
& + \tilde{\alpha}^8 \sum_{u \in V^{BT} \cup V^{AT}(G_t)} \sum_{(k,i,j) \in B(u)} p_{kijt} x_{kij} \\
& + \sum_{u \in V(G_t)} \tilde{\alpha}^9_u \underline{U}_u \\
& + \sum_{u \in V(G_t)} -\tilde{\alpha}^{10}_u \overline{U}_u
\end{aligned}
$$

to Model BASE. This cut can be interpreted as a lower bound on the total power consumption at time t, represented by z_t in Model BASE when selecting timetable x.

Implementation of Benders Decomposition We implemented and tested the presented benders approach in two ways. Whenever we added the cuts from the subproblems in one iteration, we can chose to solve the resulting master problem to optimality or we can generate the cuts using a callback function at each node in the branch-and-bound tree of the master problem. Both approaches have their own advantages. When solving the master problem to optimality in each iteration, we at the same time obtain a lower bound for the integrated model. Using the callback method, the master problem does not have to be solved to optimality in each iteration, therefore the cuts can be added more efficiently. The disadvantage of this method is a slow improvement of the lower bound for the integrated model. We term *BCB* the callback approach. *BS* represents solving the master problem to optimality in each iteration.

4 Computational Results

In this section, we present the computational result sfor the developed models and solving strategies for optimizing timetables of underground train network considering power flows. All algorithms were implemented in Python 3.10.13 using Gurobi 11.0.0 with standard parameter setting to solve mixed-integer problems. We performed the calculations on a server with an Intel Xeon Gold 6326 CPU, 256 GB RAM, 2x16 cores and 2.9 GHz base frequency. The Gurobi solver was limited to 8 threads to avoid memory issues.

4.1 Test Instances

All tests were performed on real-world deployed timetables provided by our project partners VAG, the operator of the underground train system in Nuremberg, Germany. We consider two different timetables, one for weekdays and one for sundays. They differ in frequency of the schedule, there drive fewer trains per hour on sundays than on weekdays. In order to be able to use the timetable optimization in real operation, we plan to repeatedly optimize shorter sections of around 10 minutes of the timetable during a day in a rolling horizon approach. Therefore the instances analyzed in this section are segments of the timetable between 100 and 700 seconds. The transmission network for the DC power simulation is also designed to match the real state of the underground train system in Nürnberg, including informations about the track lengths and admittance parameters on the power lines. The section we are looking at consists of 6 separate sub-grids with a total of 18 substations.

4.2 Resolution Sensitivity Analysis

The first test investigates the impact of varying the resolution of the transmission graph as defined in Section 3.3. The number of bucket nodes is adjusted for each track, and train nodes are clustered with a specified maximum clustering distance. For each test configuration we analyzed 10 instances optimizing different 100 second intervals of the sunday timetable and present the mean value of the results. Table 3 displays the percentage reduction in the number of buckets for different resolutions compared to using one bucket for each departure configuration. The column ∞ indicates that we have exactly one bucket for each track. Additionally, Table 4

Table 3: Percentage of bucket nodes reduced by various resolutions.

Instance	0.05	0.1	0.5	1.0	∞
CM	34.40%	48.04%	72.70%	77.97%	81.49%

presents the relative difference in the objective function values compared to the case with one bucket per train node for the quadratic model CM, its most accurate linearization SB and the relaxed model RLX. The results in Table 3 indicate that we can already reduce almost half of the

Table 4: Relative difference in the objective value to the quadratic model CF for different models and resolutions.

Instance	0.05	0.1	0.5	1.0	∞
CM	0.01%	0.02%	0.04%	0.05%	0.02%
RLX	0.04%	0.29%	2.61%	4.67%	5.92%
SB	0.16%	0.05%	3.01%	2.73%	3.45%

train nodes with a resolution of 0.1. Decreasing the number of buckets in the transmission graph leads to a reduction of bucket nodes by up to 81.49% but also comes with significantly higher deviations from the objective value using one bucket for each train node. This has prompted us to use a resolution of 0.1 in the following computational tests regarding the solving strategies.

4.3 Solving Strategy Comparison

The second test compares different solving strategies for instances with 4 different time horizons between 100 and 700 seconds. The strategies include the solving of the models $BASE$, RLX, and SB as well as the different implementations for the developed Benders decomposition approach BS and BCB. Each strategy gets the solution of $BASE$ as a starting solution. For each timetable

and time horizon we solve 5 different timetable segments with a maximum runtime set to 5 hours. Table 5 shows the geometric mean of the resulting runtimes, Table 6 displays the mean of the *real* energy consumptions of the solutions obtained with each strategy, along with the relative improvement (SAV) compared to $BASE$. For us, the real energy consumption of a timetable is the energy consumption of the timetable evaluated with the quadratic current model CF. For this simulation, we use the heuristic HEU presented in Section 2.4. We evaluate each feasible solution that is produced in the solving process of each strategy and display the best solution in Table 6. The time used for evaluation is not included in Table 5. For Model $BASE$ we present two values. Column $BASE$ is the real energy consumption of the optimal solution to model $BASE$. Column $BASE^*$ displays the lowest real energy comsumption evaluated on all solutions found by the Gurobi solver, also suboptimal solutions.

Table 5: Runtimes for different solving strategies.

Instance	BASE	RLX	BS	BCB	SB
sunday_100	0.06	158.37	3.19	23.73	5374.34
sunday_300	0.14	2205.56	14.01	98.79	18000
sunday_500	0.21	4148.25	26.90	238.47	18000
sunday_700	0.24	4760.78	17.32	263.90	18000
weekday_100	0.14	1038.48	15.79	76.60	18000
weekday_300	0.70	14659.51	142.58	1295.00	18000
weekday_500	1.25	18000	245.49	3616.97	18000
weekday_700	1.44	18000	77.37	5401.85	18000
Average	0.50	7871.37	67.83	1376.91	16421.79

Table 6: Power provided by the powerstations in Wh for timetables computed with different solving strategies.

Instance	BASE	BASE*	RLX	BS	BCB	SB	SAV
sunday_100	179.74	179.74	179.38	**179.35**	179.36	179.69	0.22%
sunday_300	575.49	575.49	**574.89**	574.94	574.93	575.49	0.10%
sunday_500	955.31	955.21	**954.69**	954.79	954.77	955.29	0.07%
sunday_700	1360.61	1359.89	**1359.20**	1361.88	1359.83	1360.23	0.10%
weekday_100	302.52	302.46	300.14	**300.12**	300.16	302.52	0.79%
weekday_300	1023.30	1020.21	1017.34	**1016.59**	1017.05	1023.30	0.65%
weekday_500	1714.37	1710.92	1702.37	1700.81	**1700.77**	1712.31	0.79%
weekday_700	2415.97	2413.44	**2408.32**	2412.83	2408.36	2415.97	0.32%

The integrated view of schedule and DC power network optimization yields noticable improvements particularly for the weekday timetables, where there are more trains driving simultaneously in the network than on sundays. Nevertheless, it must be noted that the easy-to-solve $BASE$ model is a good approximation for the complex physical interrelationships. The produced timetables do not have a significantly worse energy consuption than the ones computed with the more complex power flow models. Solving the model that most accurately represents the power flows in the DC network, Model SB, to optimality was not possible in the timelimit of 5 hours in most of the instances. Although the runtimes of solving Model RLX could be significantly reduced through the use of Benders decomposition, the runtimes of the $BASE$ model are advantageous for the real-time application. The most practical approach to getting high quality solutions in real time seems to be solving Model $BASE$, evaluating all solutions obtained in the solving process and use the best of them with regard to the evaluation using Model CF.

5 Conclusions

In conclusion, this paper represents a fusion of two critical areas of research: timetable optimization and DC power grid optimization in the context of underground railway systems. By combining these previously disparate fields, we have developed a novel approach for generating energy-efficient timetables that accounts for the intricate dynamics of the power network. Our work extends beyond conventional timetable optimization by accurately modeling the physical laws and power losses inherent in the DC power grid, thus ensuring a more comprehensive understanding of energy consumption and demand within the railway system.

Our study compared two model formulations—power flow and current flow—to address the quadratic constraints that arise from the relationship between power, current, and voltage in the DC power grid. The results indicate that the current flow model is computationally more tractable, offering a significant advantage in solving large-scale optimization problems. To further accelerate the solution process, we developed a heuristic that enhances the efficiency of the current flow model, thereby making it more suitable for real-time applications.

Empirical validation using real-world data from the VAG (Verkehrs-Aktiengesellschaft Nürnberg) demonstrated the practical effectiveness of our integrated model. The simulated energy consumption closely aligned with actual measurements, deviating by at most 1%, thereby validating the accuracy of our approach. Additionally, the newly calculated timetable achieved an increase in energy efficiency of up to 0.8% compared to the previously used model, highlighting the tangible benefits of our method in operational settings.

Beyond the immediate improvements in energy efficiency and computational performance, this research contributes to the broader field of railway operations by offering a novel approach that integrates physical energy dynamics with timetable optimization. The methodologies developed here pave the way for future research at the intersection of transportation and energy systems, offering potential applications in other domains where energy efficiency and operational optimization are critical. As urban centers continue to expand, the demand for sustainable and efficient public transportation will grow, making the contributions of this paper increasingly relevant for the operation of next-generation railway systems.

References

Arboleya, Pablo et al. (2020). 'A review of railway feeding infrastructures: Mathematical models for planning and operation'. In: *eTransportation* 5, p. 100063. DOI: https://doi.org/10.1016/j.etran.2020.100063.

Berriel, Ryan O et al. (2023). 'Analysis of the Timetable Impact on Energy Consumption of a Subway Line'. In: *IEEE Transactions on Vehicular Technology*. DOI: 10.1109/TVT.2023.3282816.

Gan, Lingwen and Steven H Low (2014). 'Optimal power flow in direct current networks'. In: *IEEE Transactions on Power Systems* 29.6, pp. 2892–2904. DOI: 10.1109/TPWRS.2014.2313514.

Gao, Ziyou and Lixing Yang (2019). 'Energy-saving operation approaches for urban rail transit systems'. In: *Frontiers of Engineering Management* 6.2, pp. 139–151. DOI: https://doi.org/10.1007/s42524-019-0030-7.

Gemander, Patrick, Andreas Bärmann, and Alexander Martin (2023). 'A Stochastic Optimization Approach to Energy-Efficient Underground Timetabling under Uncertain Dwell and Running Times'. In: *Transportation Science* 57.6, pp. 1627–1645. DOI: https://doi.org/10.1287/trsc.2022.0269.

Gupta, Shuvomoy Das, Bart PG Van Parys, and J Kevin Tobin (2023). 'Energy-optimal Timetable Design for Sustainable Metro Railway Networks'. In: *arXiv preprint arXiv:2309.05489*. DOI: https://doi.org/10.48550/arXiv.2309.05489.

Hager, Lukas and Tobias Kuen (2024). 'Optimization of Underground Train Systems'. en. In: *Unlocking Artificial Intelligence: From Theory to Applications*. Ed. by Christopher Mutschler et al. Cham: Springer Nature Switzerland, pp. 303–319. ISBN: 978-3-031-64832-8. DOI: https://doi.org/10.1007/978-3-031-64832-8.

Huang, Kang, Feixiong Liao, and Ziyou Gao (2021). 'An integrated model of energy-efficient timetabling of the urban rail transit system with multiple interconnected lines'. In: *Transportation Research Part C: Emerging Technologies* 129, p. 103171. DOI: https://doi.org/10.1016/j.trc.2021.103171.

Kang, Liujiang et al. (2024). 'A Critical Review of Subway Train Timetabling and Rescheduling Problems'. In: *IEEE Transactions on Intelligent Transportation Systems*. DOI: 10.1109/TITS.2024.3386728.

Liu, Hongjie et al. (2018). 'Timetable optimization for regenerative energy utilization in subway systems'. In: *IEEE Transactions on Intelligent Transportation Systems* 20.9, pp. 3247–3257. DOI: 10.1109/TITS.2018.2873145.

Liu, Pei et al. (2018). 'Energy-efficient train timetable optimization in the subway system with energy storage devices'. In: *IEEE Transactions on Intelligent Transportation Systems* 19.12, pp. 3947–3963. DOI: 10.1109/TITS.2018.2789910.

Moreno, Teresa et al. (2015). 'Urban air quality comparison for bus, tram, subway and pedestrian commutes in Barcelona'. In: *Environmental Research* 142, pp. 495–510. DOI: https://doi.org/10.1016/j.envres.2015.07.022.

Peña-Alcaraz, Maite et al. (2012). 'Optimal underground timetable design based on power flow for maximizing the use of regenerative-braking energy'. In: *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit* 226.4, pp. 397–408. DOI: https://doi.org/10.1177/0954409711429411.

Raghunathan, A. U. et al. (2014). 'Minimizing Energy Consumption In Railways By Voltage Control On Substations'. In: *WIT Transactions on The Built Environment* 135, pp. 697–708. DOI: https://doi.org/10.2495/CR140581.

Scheepmaker, Gerben M, Rob MP Goverde, and Leo G Kroon (2017). 'Review of energy-efficient train control and timetabling'. In: *European Journal of Operational Research* 257.2, pp. 355–376. DOI: https://doi.org/10.1016/j.ejor.2016.09.044.

Su, Shuai et al. (2013). 'A subway train timetable optimization approach based on energy-efficient operation strategy'. In: *IEEE transactions on intelligent transportation systems* 14.2, pp. 883–893. DOI: 10.1109/TITS.2013.2244885.

Yang, Songpo et al. (2020). 'A bi-objective timetable optimization model incorporating energy allocation and passenger assignment in an energy-regenerative metro system'. In: *Transportation Research Part B: Methodological* 133, pp. 85–113. DOI: https://doi.org/10.1016/j.trb.2020.01.001.

Yang, Xin et al. (2015). 'An energy-efficient scheduling approach to improve the utilization of regenerative energy for metro systems'. In: *Transportation Research Part C: Emerging Technologies* 57, pp. 13–29. DOI: https://doi.org/10.1016/j.trc.2015.05.002.

Zhao, Ning et al. (2017). 'An integrated metro operation optimization to minimize energy consumption'. In: *Transportation Research Part C: Emerging Technologies* 75, pp. 168–182. DOI: 10.1109/TITS.2018.2871347.

Zhou, Yuhe et al. (2018). 'Integrated optimization on train control and timetable to minimize net energy consumption of metro lines'. In: *Journal of Advanced Transportation* 2018.1, p. 7905820. DOI: https://doi.org/10.1155/2018/7905820.