Railway Engineering Science (2024) 32(3):257–285
https://doi.org/10.1007/s40534-023-00322-4

Increasing realism in modelling energy losses in railway vehicles and their impact to energy-efficient train control

Michael Nold[1] · **Francesco Corman**[1]

Received: 18 March 2023 / Revised: 17 September 2023 / Accepted: 28 September 2023 / Published online: 5 January 2024

Abstract

The reduction of energy consumption is an increasingly important topic of the railway system. Energy-efficient train control (EETC) is one solution, which refers to mathematically computing when to accelerate, which cruising speed to hold, how long one should coast over a suitable space, and when to brake. Most approaches in literature and industry greatly simplify a lot of nonlinear effects, such that they ignore mostly the losses due to energy conversion in traction components and auxiliaries. To fill this research gap, a series of increasingly detailed nonlinear losses is described and modelled. We categorize an increasing detail in this representation as four levels. We study the impact of those levels of detail on the energy optimal speed trajectory. To do this, a standard approach based on dynamic programming is used, given constraints on total travel time. This evaluation of multiple test cases highlights the influence of the dynamic losses and the power consumption of auxiliary components on railway trajectories, also compared to multiple benchmarks. The results show how the losses can make up 50% of the total energy consumption for an exemplary trip. Ignoring them would though result in consistent but limited errors in the optimal trajectory. Overall, more complex trajectories can result in less energy consumption when including the complexity of nonlinear losses than when a simpler model is considered. Those effects are stronger when the trajectory includes many acceleration and braking phases.

Keywords Train trajectory optimization · Energy-efficient train control (EETC) · Dynamic efficiency · Power losses in railway vehicles

List of symbols

ATO	Automatic train operation
DAS	Driver advisory system
EETC	Energy-efficient train control
GPS	Global positioning system
GTO	Gate-turn-off thyristor
IGBT	Insulated-gate bipolar transistor
NOREG	No regenerative braking
REG	Regenerative braking
RMS	Reduced maximum speed
PMSM	Permanent magnet synchronous motors
E	Total energy
$E_{\mathrm{L.Pdep}}$	Energy losses dependent from the mechanical power
$E_{\mathrm{L.Pind}}$	Energy losses independent from the mechanical power
E_{R}	Energy for the resistance
$E_{\mathrm{L.nonl}}$	Energy losses due to nonlinearities
E_{IC}	Energy irreversibly consumed
E_{Kin}	Kinetic energy
E_{Pot}	Potential energy
η	Efficiency
F	Mechanical traction force at wheel
P_{A}	Power for accelerating
P_{BD}	Power consumed by the dissipative brakes
P_{BR}	Power in the regenerative brakes
P_{In}	Power input
P_{IC}	Total power irreversibly consumed
P_{L}	Power losses
$P_{\mathrm{L.Pdep}}$	Power losses dependent from the mechanical power
$P_{\mathrm{L.Pind}}$	Power losses independent from the mechanical power
$P_{\mathrm{L.nonl}}$	Power losses due to nonlinearities
P_{Mech}	Power output mechanical at wheel
P_{R}	Power for the resistance
ε	Nonlinearities and error term

✉ Michael Nold
michael.nold@ivt.baug.ethz.ch

[1] Institute for Transport Planning and Systems, ETH Zürich, Zürich, Switzerland

c_0	Constant power-independent losses
c_1, c_2, c_3	Constants for the Davis formula
s	Space
t	Time
$t+$	Travel time buffer
v	Speed

1 Introduction

The reduction of energy consumption is an increasingly important topic for mobility. According to the energy strategy of the Swiss Federal Council, the entire transport sector should reduce its energy consumption by around 50% until 2050 [1]. The railway system already has a high energy efficiency compared to private mobility; nevertheless, it must also contribute to this. In order to approach this goal, one direction is to implement vehicle improvements [2–5]. A different direction is to improve operations and usage of the current vehicles. The energy-efficient train control (EETC) is one of cases in the latter direction. The EETC aims to reduce the energy consumption on a trip from an origin to a destination in a given travel time, from a starting speed to an ending speed, by driving a specific speed profile. One key solution is to use the available additional running time which is planned in the timetable (travel time buffer), compared to the minimum technically possible, driving slower and consequently using less energy [6, 7]. The two objectives of running time and energy are conflicting, and we target minimize energy, given a travel time.

It is known that when energy is lost only to accelerate and overcome driving resistance (i.e. no regenerative braking, no losses due to energy conversion, no consideration of energy not related to traction), the optimal solution to the problem of determining the EETC trajectory consists of the sequence of four modes: maximum acceleration, cruising, coasting and maximum braking [6, 7]. In reality, train vehicles are using routinely regenerative braking; and there are losses due to the energy conversion in the vehicle components (e.g. motors), which have complex, nonlinear characteristics. Those losses with increasing realism and detail have been considered in very few publications which consider detailed vehicle models. The available state of the art shows that the optimal speed profile, when those factors are considered, might deviate from the theoretical solution [8–10].

This paper contributes to the topic with the following research contributions in this direction, answering:

- How to *model* the energy usage for realistic railway vehicles at system level, balancing precision, possibility for a mathematical description computational ease. This includes energy used for traction (to accelerate and to overcome resistance), energy lost in energy conversion at multiple levels and in multiple vehicle components (both factors which are dependent on the power exchanged and factors which are independent of the power exchanged) and the nonlinear non-constant efficiency in the energy conversion flow. We propose a categorization of increasing modelling realism and complexity, which comes at a higher computational costs and required data. This allows us to discuss what is the added value of complexity.
- How to make a *detailed analysis* and *numerical estimation* of the difference in the energy optimal trajectories given a target travel time when an increasing amount of realism is included. Different modelling levels are identified and discussed (e.g. traditional EETC trajectory with only running resistance, up to the inclusion of nonlinear non-constant efficiency) in their relative improvement towards energy reduction. We study the usefulness of the increasing realism, from the point of view of actionability, and cross-differences in control actions suggested; estimation errors; and implementability, linked to the task of actually following a complex trajectory by a vehicle and a driver.

Those goals are accomplished by considering multiple energy usages at once. Moreover, we consider variable efficiency (i.e. the ratio between power output and power input) using a look-up table (which can represent non-continuous and nonlinear characteristics) with high resolution in speed and force, fitted to the industrial specification of the engines and conversion technologies. With this information, we can set up a standard discretized network of space–time speed, which we use to find the set of Pareto-optimal trajectories, minimizing energy, given a bound on the travel time. Overall, the different levels of details in modelling losses have a very large range of estimated energy consumption, which directly impacts the dimensioning of energy systems. The usage of the travel time buffer saves much energy; the relative differences in the trajectories, which are optimal for different levels of detail in modelling, are also substantial. Furthermore, technological improvements in the electrical components might additionally save energy. The complexity of implementing energy-efficient train speed profiles and taking care of nonlinearities and highly detailed models asks for advanced driver advisory systems or automated train operations. Only those are able to track the complex speed trajectories computed.

The paper is organized as follows: Sect. 2 is a literature review. Section 3 describes the problem, from the energy flow in a rail vehicle to the identification of energy losses and proposes increasing levels of realism in modelling those effects. Section 4 reports on the algorithm used to determine the optimal trajectories under increasing realism and the associated results in complete trajectories. Section 5 summarizes the results.

2 Literature review

The following literature review starts with an overview of the EETC before going in depth to the papers closer to this paper's focus.

In 1953 a simple mathematical, non-computer-based method for using the travel time buffer to reduce energy consumption was described; however, it cannot guarantee to find the optimal solution [11]. Ichikawa described 1968 probably the first computer-based energy optimization approach able to determine an EETC [12] (as reported by [10, 13]). In the years that followed until today, there is much research on different approaches about the EETC.

One problem is to mathematically determine the shape of the solutions to the EETC problem, which several approaches have studied [13, 14]. For example, analytical approaches [15], semi-analytical Lagrange multipliers [16], numerical evolutionary algorithms [17, 18] or dynamic programming [8, 9] have been used to achieve this goal. The solution of the EETC problem is also dependent on (external) environmental influences, such as slope influences [19], uncertainties in the timetable [20] or uncertainties in the driving resistance force due to different strength of the wind [21]. In the last decades, the EETC has been studied comprehensively in terms of mathematical approaches and influences from the outside. Nevertheless, the description of the EETC problem used in most academic approaches uses strongly simplified vehicle models. The frequently used assumption is to consider the traction energy only, i.e. assuming that energy is only spent to overcome driving resistance and kinetic energy is lost when braking. This means regenerative braking, losses in the energy conversion, traction components and auxiliaries are very often ignored, apart from a few exceptions, some of which are discussed later.

Regenerative braking is used in the Swiss electric rail systems since more than 100 years [22, 23] to prevent overheating of the braking system during downhill driving of mountain railways. In many other countries, large-scale use only began in the last few decades. Regenerative braking complicates the EETC problem by the need to consider also positive energy flows from the vehicle back to the power source. In the current state of the art, regenerative braking is often ignored [20, 21], or otherwise very simplified, such as a constant negative traction force. In reality, the maximum force of regenerative braking is not constant but speed-dependent due to technological factors such as the brake power hyperbola, the pull-out torque of the induction motor, aspects from the control technology, and limitations from electrical circuits [7, 10, 24–26]. When those factors are considered, the electrical braking force can change, over the speed spectrum, more than a factor 2 or disappear completely in certain speed ranges [7, 24]. This fact is ignored in most research papers [7].

Losses in the energy conversion chain and traction components arise in the traction components needed for driving (e.g. the motor, the power inverter). On electric rail vehicles, these traction components enable the conversion of the electrical energy from the overhead line into mechanical energy for the traction force at the wheel [27]. Energy conversion causes energy losses [27]. For one electric 6100 kW vehicle, these losses can be up to around 1000 kW [28]. These losses can be identified, for example, by waste heat. Usually, these losses are simplified and described with a constant efficiency between total power input and power output [7, 10, 29].

In reality, the efficiency is not constant but rather variable; in other words, dynamic. The current speed and traction force are the most important factors for describing it. Given those two factors, moreover, the functional form is nonlinear and non-continuous (changes stepwise, based on different engine regimes). Only very few papers consider a variable efficiency dependent on speed and traction force by a pre-specified continuous function [8–10, 29–31]. Specifically, Ghaviha et al. [10] have considered the dynamic efficiency for EETC and reported that only Franke et al. [8, 9] are known to have considered such a level of detail before.

Auxiliary power consumption arises because each rail vehicle needs energy for auxiliary components necessary for safe driving and to keep the traction equipment in the running regime (e.g. the motor cooling fan). The energy consumption of the auxiliary components can be, for the 6100 kW vehicle mentioned above, up to 150 kW [28]. This is strongly influenced by the operating time and can be significantly reduced during a standstill due to component shutdowns [3, 10, 28–30]. To the best of our knowledge, this relatively small energy consumption, which can be estimated at around 2.5% of the maximum mechanical power at wheel, is usually ignored in the current EETC models. As a noteworthy exception, Ghaviha et al. [10] have considered it, but in a simplified manner.

We call *vehicle component losses* the sum of losses due to traction components and those due to the auxiliary components. The magnitude was quantified empirically on an analysis of energy counting of 49 regular Intercity train runs between the Swiss cities Zürich and Luzern (60 km away, less than an hour travel time, 2 intermediate stops) [32]. After deducting the regenerative energy, the breakdown of total energy consumption in its sources was as follows: the vehicle components losses consumed 32.5%, the driving resistance 24.3%, the non-regenerative brakes absorbed 15.1% and passenger comfort systems (e.g. light and heating) consumed 28.1% of the total energy required for the trip. It should be emphasized that these analyses only represent an order of magnitude because they strongly depend on the driving profile and various other influencing variables

[29, 30]. However, these measurements show that vehicle component losses significantly influence energy consumption. The present paper tackles the challenge of exploiting the potential to reduce the total energy consumption once vehicle component losses are modelled in detail.

One possible way to reduce the energy losses in the vehicle components is to improve the vehicles by means of innovative technology, design, engineering and construction. For instance, switch off motors [33, 34], switch off auxiliaries at standstill [3], or use highly efficient power inverter [4, 5]. A different stream is to reduce energy consumption by developing a model, able to solve the EETC problem, considering with sufficient detail a highly detailed vehicle model. In this direction, we discuss more in detail the work of Franke et al. [8, 9], Ghaviha et al. [10], Xiao et al. [35], and Scheepmaker and Goverde [7].

Franke et al. [8, 9] compared two trajectories determined by optimal train control approaches under different assumptions. The first case considers the standard EETC assumptions, resulting in the optimal trajectory following maximum acceleration, cruising, coasting and maximum braking. The second case considers dynamic efficiency, depending on traction force and speed. To find the optimal trajectory in both cases, they simplified the problem by a coordinate transformation into an equivalent kinetic energy per mass. They used dynamic programming and a lookup table to determine the stepwise contributions of speed changes towards the total energy and travel time. The total travel time is given. When comparing with the trajectory of minimum running time, the results show that total energy consumption is reduced by 14% with standard EETC strategy and by 25% when the dynamic efficiency is considered. This latter trajectory does not use coasting as much as the former one. The authors thus conclude that the coasting strategy is not optimal in case of dynamic efficiency. However, the generality and transferability of these values are limited because the two strategies use different kinds of braking. The EETC uses both regenerative and non-regenerative (dissipative) braking, while their dynamic efficiency approach uses only the regenerative brakes, which is bounded by a nonlinear function. Due to this difference, it is difficult to separate the effects of the consideration of dynamic efficiency (compared to the standard EETC) from the effects of complete usage of regenerative brake usage (compared to using dissipative brakes). Moreover, they do not describe which specific factor, considering the dynamic efficiency, should result in which difference for the optimal trajectory. The paper does not go deep into explanations or analyses and the fact that the dynamic efficiency approach does not use maximum acceleration at higher speeds.

Ghaviha et al. [10] considered dynamic losses focusing on implementing a driver advisory system (DAS) for Android devices. Several compromises were necessary to handle limited device performance, including pre-computing many data. A simplified function for modelling the dynamic efficiency dependent on power and speed is used, based on fitting a bi-dimensional polynomial to some industrially measured values. This system gives traction force recommendations only at discrete moments (each kilometer point) and not continuously. This results in issues in handling precise positioning, including GPS reception problems. Moreover, the driver receives recommendations to adjust the traction force only section-wise. Their optimized trajectory diagrams show that considering constant efficiency contains maximum acceleration and coasting; considering dynamic efficiency does not use the maximum acceleration and does not use coasting. In general, they describe many details about the algorithm and implementation of the Android DAS; however, not many are about the description of the resulting optimal trajectory.

Xiao et al. [35] addressed dynamic efficiency when optimizing the trajectory to result in minimal energy consumption. They reported a comparison between a current approach and an automatic train operation (ATO) strategy, but neither was described in detail. For the comparison, only one route was considered, with different downhill gradients during the acceleration phase and different uphill gradients during the deceleration phase. Due to this, it is not possible to identify whether a change in the trajectory is due to a switching point or a change in gradient. Neither the scenario nor the description is suitable to quantify the dynamic efficiency impact on the acceleration and coasting behaviour.

Scheepmaker and Goverde [7] used a train model extended by considering nonlinear bounded regenerative braking. It came to the result that this braking system influences the optimal cruising speed and the speed at the beginning of the braking phase. However, their train model remains still simplified as it considers constant efficiency. Even though they did not use dynamic efficiency, they suggested that future research consider when more realistic calculations are required.

Approaches to reduce power consumption in railway operations have also reached industrial applications. A driver advisory systems DAS system was developed and introduced for the Swiss Federal Railway [36]. Depending on the current situation on the railway network, the drivers might get a recommendation to reduce the speed. On the one hand, this enables saving energy due to the reduced speed. On the other hand, energy-intensive braking and accelerating on operational signal stops can be avoided. This approach

only recommends a speed profile with a reduced maximum speed (RMS). Coasting or a variable speed profile are not included. Due to this, the federal office of Transport funded the research of this paper to determine the potential for further development in trajectory optimization.

To summarize, there is much research of high mathematical abstraction and quality, using more or less simplified vehicle models, which confirms that the optimal EETC trajectory consists of the four modes: maximum acceleration, cruising, coasting and maximum braking [6, 17, 18, 21]. However, there is very little research which uses increasingly detailed complex models of vehicles and their traction components. In this latter case, the trajectories show differences from the EETC. This can be due to the specific multiple aspects of characterising losses, the confidentiality of data, and complex mathematical modelling, to result in practical scientific insights. However, the only common point that can be proven based on the published state of the art, and only qualitatively, is that maximum acceleration is not used at higher speeds. In this paper, we report some results from an industrial project [29, 30, 37], extended and put in relation to the current state of the art. This justifies the aim of this paper to describe and investigate the influence of increasingly detailed models of vehicles towards determining the energy-optimal trajectory when losses due to vehicle components, dynamic efficiency, and the possibility to use mechanical and regenerative braking are included.

3 Rail vehicle energy flow and energy consumption

3.1 Vehicle energy flow

We here describe the energy flow of rail vehicles, which occurs on a journey from a point A to B. This enables understanding where and why energy is consumed in a rail vehicle and losses occur. We later show how to model those losses mathematically. We call *energy flow* the physical energy (e.g. mechanical energy and electrical energy), which is transferred or converted in the considered technical system. During a train operation, there is an energy flow, which means energy is converted between different kinds of energy. This energy flow, which starts with the supplied energy, is needed to generate mechanical energy (which results in the movement) and finally causes that energy is being consumed and energy losses occur.

Figure 1 describes the simplified energy flow in a typical rail vehicle during a rail operation for a given trip between an origin and a destination. We follow the top line from left to right, which is the most common part. To reach the destination, energy is required, which is usually supplied from outside. This energy can be electrical, fuel, coal or other kinds of energy. Directly from the supplied energy or via the traction components, there is an energy flow to the passenger comfort systems to enable heating, air conditioning, light

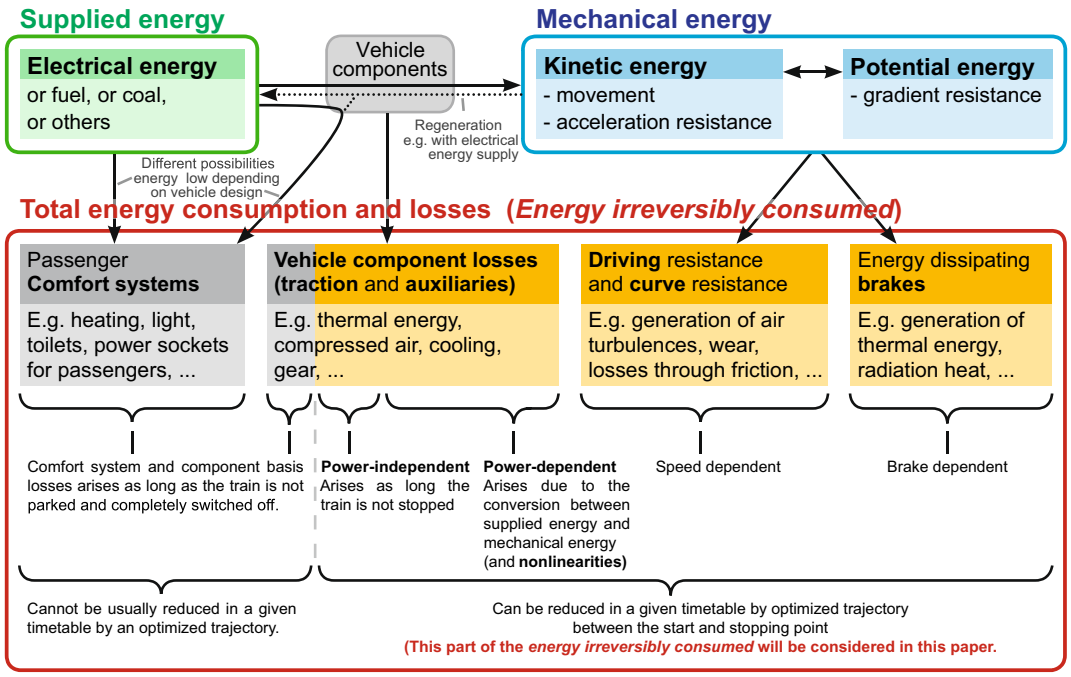

Fig. 1 Simplified visualization of the energy flow during a usual rail operation for a train run between points A and B (based on Refs. [6, 27, 29–32, 39, 41, 58])

and other systems. For locomotion, there is an energy flow from the supplied energy via some traction components, which can convert the energy to mechanical energy. During this conversion, a part is lost due to the limited efficiency of the traction components (e.g. in thermal energy). The bottom part of Fig. 1 reports the types of losses identified in this paper.

The mechanical energy contains kinetic energy (for locomotion) and potential energy (to overcome the gradient resistance in the case of slopes). Due to energy conservation, there is a direct exchange and energy flow between kinetic and potential energy. To overcome the driving resistance and curve resistance there is an energy flow from the mechanical energy. During braking or downhill trips, energy flows reduce the mechanical energy; for instance, there can be an energy flow from the mechanical energy to the energy dissipating brakes or back to the supplied energy if regenerative brakes are used.

3.2 Overview of the energy consumption and losses

The energy flow shows that some energy is transformed during rail operation, but a part of the supplied energy is consumed and/or losses are generated. Mechanical energy is required for the movement. The kinetic and potential energy can vary along the trip and can be reversibly generated. For instance, regenerative braking can feed back energy which does not count towards the total consumption. However, kinetic energy can also be lost, for instance, during braking with energy-dissipating brakes. We are interested in the total energy consumed, i.e. what is not available any more after the operations have been performed, and not in the energy stored temporarily as kinetic energy, which can be regenerated later on. We rather talk not about the energy consumption but rather the energy lost after the operations. The energy which is lost is transformed in an irreversible manner (for instance, into heat, noise) and is not available for further rail operational usage. This is the focus of the paper; we therefore call this *energy irreversibly consumed* (short E_{IC}), and describe the following aspects, which generate energy losses (compare also Fig. 1, bottom boxes):

- *Dissipated energy due to the use of energy-dissipating brakes*: This energy irreversibly consumed is due to the usage of energy-dissipating brakes, which do not convert energy back via regeneration. Brakes are needed, e.g. to reduce the speed of a vehicle or to avoid acceleration on a downhill trip. This can be, e.g. mechanical pad brakes, track brakes, electrical resistor brakes, counter-pressure brakes [25, 26, 38].
- *Energy consumption to overcome the train resistance:* This energy irreversibly consumed overcomes the resistance of the train to movement. Typically, such resistances are modelled by the quadratic speed-dependent Davis formula, possibly with extra factors for the tunnel or curve resistance [27, 31, 39, 40].
- *Energy losses in traction components and auxiliaries*: This energy irreversibly consumed is due to energy losses in the traction components (e.g. motors, power converter, and gear) and the auxiliaries (e.g. motor fan). Auxiliaries consume energy and enable the operation of the traction components. The power losses in the traction components (e. g. motors) arise due to physical reasons. In general, they can be modelled as a total efficiency of energy conversion η, being less than one ($\eta < 1$). In other terms, to result in a given mechanical power output at wheel P_{Mech}, a total of $(1/\eta)P_{\mathrm{Mech}}$ must be supplied as power input P_{In}. The difference between P_{Mech} and P_{In} identifies the total power losses P_{L}. To put it differently, the P_{L} are equal to $\left(\frac{1}{\eta} - 1\right)P_{\mathrm{Mech}}$. The efficiency as a scalar constant number is a simple way to describe the entire vehicle system and traction chain.
- *Energy consumption due to comfort systems or for the preservation of transport goods*: This energy irreversibly consumed primarily arises due to the requirements of the customer. Comfort systems are required on passenger trains, e.g. light, heating, air-conditioning and laptop power sockets. On freight trains, energy is required to preserve the transport goods, e.g. cooling systems during food transport. Reducing the energy losses of comfort systems or preserving transport goods are not the focus of this optimization and the EETC.

Overall, many of those factors have many dependencies. Resistances change on, e.g. atmospheric conditions (time and space specific); driving in curves (infrastructure specific); driving in a tunnel or not (infrastructure specific); design of bogies, wheel and track (vehicle-type specific); revision/refit/improvement of some vehicle (specific of vehicle family and revision status); condition of the specific bogies, traction chain component in a specific vehicle (compared to a typical train-template) (vehicle specific) [27, 31, 39–41].

Some of those phenomena can be included, e.g. in extended kinematic formulas, such as the Davis formula, when made specific to the vehicle and infrastructure. In the following, we assume that all effects of resistances are constant and can be described by the Davis formula so that we can highlight and focus on the energy conversion aspects, which represent the core of this paper. We are also able to study two cases of the same trains (Re460) but of different revision families and show the impact of new electric components in the same vehicle and the same resistance.

In general, the total energy E consumed can be computed by Eq. (1) as the integral of the mechanical traction

forces F at the wheel over space (i.e. distance) s, or the integral of P_{Mech} over time t:

$$E = \int_{s_1}^{s_2} F(s)\mathrm{d}s = \int_{t_1}^{t_2} P_{\text{Mech}}(t)\mathrm{d}t, \tag{1}$$

where s_1 and s_2 are respectively start and ending location of the given trip; t_1 and t_2, the starting and ending time, and similarly v_1 and v_2, are the starting and ending speed; all those are normally given in the EETC problem. The variation potential energy ΔE_{Pot} for both of those locations is also given. The variation in kinetic energy ΔE_{Kin} depends on the starting and ending speed. The $F(s)$ can be further divided in its positive component (when accelerating) and negative component (when braking with regenerative brakes). We label P_{A} for the power for accelerating, P_{BD} for the power consumed by the dissipative brakes and P_{BR} for the power in the regenerative brakes. Overall, at any time step Δt there is

$$\frac{\Delta E_{\text{Kin}}}{\Delta t} + \frac{\Delta E_{\text{Pot}}}{\Delta t} + P_{\text{Mech}} + P_{\text{BD}} + P_{\text{L}} = P_{\text{In}}, \tag{2}$$

where P_{Mech} and P_{R} are described, respectively, as follows:

$$P_{\text{Mech}} = F \cdot v = P_{\text{A}} + P_{\text{BR}} + P_{\text{R}}, \tag{3}$$

$$P_{\text{R}} = v\left(c_1 + c_2 \cdot v + c_3 \cdot v^2\right). \tag{4}$$

The total power irreversibly consumed P_{IC}, with regenerative braking only ($P_{\text{BR}} = 0$), can then be modeled by Eq. (5) as the sum of P_{L} and the power needed to overcome the resistance P_{R}. The paper will focus mostly on those two sources of power irreversibly consumed. We label three components, namely $P_{\text{L.Pdep}}$ for the power-dependent losses, which are constant proportion to the mechanical power; $P_{\text{L.Pind}}$ for the mechanical power-independent losses, which are a constant value; $P_{\text{L.nonl}}$ for the collected power nonlinearities and further modeling errors.

$$P_{\text{IC}} = P_{\text{L}} + P_{\text{R}} = \underbrace{\left(\frac{1}{\eta} - 1\right)}_{P_{\text{L.Pdep}}} P_{\text{M}} + \underbrace{c_0}_{P_{\text{L.Pind}}} + \underbrace{\varepsilon(F, v)}_{P_{\text{L.nonl}}} + \underbrace{v\left(c_1 + c_2 \cdot v + c_3 \cdot v^2\right)}_{P_{\text{R}}}. \tag{5}$$

The EETC is confronted with problems where $\frac{\Delta E_{\text{Kin}}}{\Delta t}$ and $\frac{\Delta E_{\text{Pot}}}{\Delta t}$ both are constant, and can therefore be ignored. We can finally define the EETC problem as the trajectory $v(t)$ and traction force $F(t)$, given $s_1, s_2, t_1$and t_2, that satisfy $\underset{v(t),\,F(t)}{\arg\min} P_{\text{IC}}$. We will then formally consider either the P_{IC} or the energy irreversibly consumed E_{IC} in our study. The EETC problem can be formally described by Eq. (6):

$$\underset{v(t),\,F(t)}{\arg\min} \int_{t_1}^{t_2} P_{\text{In}}(t)\mathrm{d}t = \underset{v(t),\,F(t)}{\arg\min} E_{\text{IC}}. \tag{6}$$

3.3 Energy losses in traction components and auxiliaries: vehicle component losses

Out of the previous section's list, the paper focuses on energy losses in traction components and auxiliaries; we denote them together as vehicle component losses. We will speak of energy and power losses; to derive the energy losses, we need to integrate power losses over the total trip time.

As shown in Fig. 1, the vehicle component losses can be divided into different parts. There is a part which cannot be influenced by the trajectory optimization in a given timetable and remains as long as the train is not fully switched off at a depot or parking facility (e.g. power for control computers, train protection system, and driver cabin heating). The current trend is that this part is reduced by the usage of retrofitted or new vehicles, which can switch off vehicle components at a standstill. Here, we identify two concepts: standstill and switched off. *Standstill* corresponds to any moment when the speed is null. This includes, e.g. commercial stops and waiting in front of signals. Auxiliaries, in this case, are typically still running. Instead, the train is *switched off* when parked longer at a parking location, not providing service to passengers (auxiliaries are typically switched off). Due to this, those losses cannot be reduced, so these types of losses are not considered in the current paper. The focus is instead on the part of the vehicle component losses, which can be reduced by trajectory optimization in a given timetable (also assuming a variable travel time buffer).

The vehicle component losses are quantitatively reported in Fig. 2, which is a real-life base for our analysis. Figure 2a reports the real-life measured losses as a function of speed and traction force, while Fig. 2b shows the resulting efficiency of the total traction chain. Both plots describe the same real-life vehicle type of the Swiss Federal Railways.

Figure 2a shows a surface describing the power losses P_{L} (vertical axis), given speed v and traction force F. We will call "force–speed" diagrams in what follows. The surface increases with both speed and traction force and is bounded on the horizontal domain by the power hyperbola (the engine

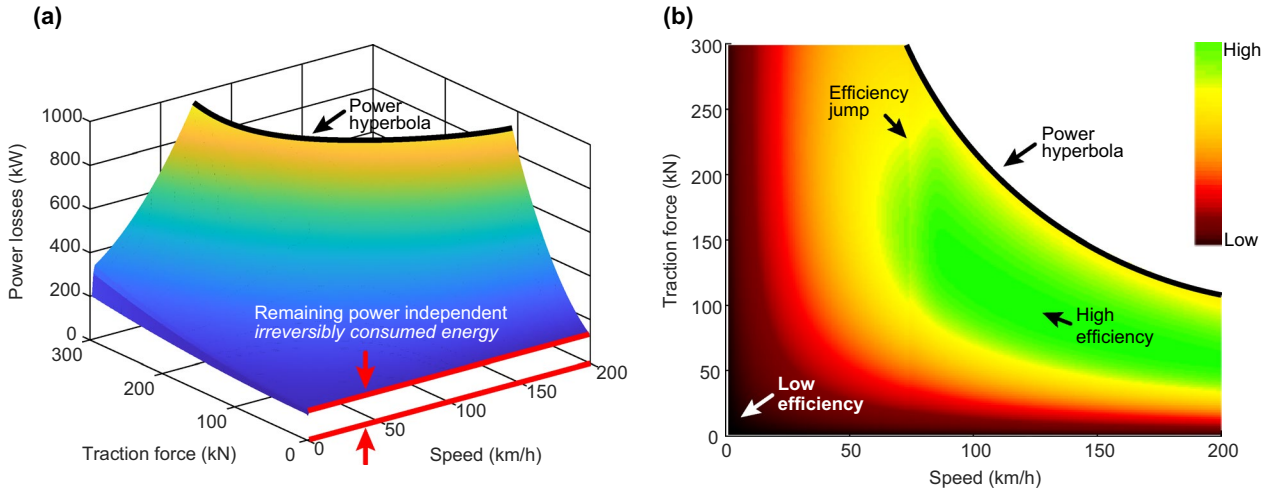

Fig. 2 Diagrams of losses and resulting traction chain efficiency of the traction chain of the Swiss Federal Railway locomotive series Re 460 (using a 3-point-GTO-power inverter), as depending on the engine operating point, delivered from the manufacturer: **a** losses of the entire traction chain, dependent on the speed and traction force ($P_{\text{Losses}}(F, v)$); **b** vehicle efficiency $\eta(F, v)$. The green area has the best efficiency, yellow and red intermediate values, black the lowest efficiency. The resolution of these diagrams is 1 km/h and 1 kN. For generating these maps, the entire traction chain was considered. For instance, in **b** a jump in efficiency due to the power inverter clocking frequency change can be observed. Numerical values are not reported due to confidentiality. The traction chain schema can be found in the Appendix (see Fig. 11)

has a physical power limitation). The surface shows that there is an approximate dependency between the power losses and mechanical power. The mechanical power P_{Mech} is the product of the speed v and traction force F, be this latter positive (accelerating) or negative (braking by means of a regenerative brake). When using only dissipative brakes, the used mechanical power from the engine is 0. We call the power losses that can be described well by this dependency as *power-dependent losses*. Further, Fig. 2a shows that a certain amount of power losses remains independent from the mechanical power. See in Fig. 2a; when $v > 0$ and $F = 0$, $P_{\text{Mech}} = 0$, but $P_{\text{L}} > 0$. These losses can only vanish if the components are completely switched off; this can be done both at a standstill and when the train is switched off. We call this additional part, which can be approximated by some constant, as *power-independent losses.* Due to the power-independent losses, the total efficiency of the traction chain (and consequently of the vehicle), which is the ratio between the mechanical power at the wheel and the power in input, is very low at low speeds and low traction forces.

This effect is described specifically in Fig. 2b, which reports the measured efficiency of the entire traction chain, again depending on speed and traction force; low efficiency is black, high efficiency is green. Due to further complex nonlinear influence, the best efficiency is not on the maximum power (which would be the upper-right boundary corresponding to the power hyperbola), but it lies somehow more in the middle, see the green area. Figure 2b also shows nonlinearities and discontinuities (jumps) in real-life efficiency. In other words, the efficiency changes with both the vehicle speed and traction force; we call this combination the *engine operating point*. In what follows, we investigate how to optimize trajectories by ensuring the engine operating points which describe the entire trajectory result in overall smallest losses.

Figure 2, which is equivalent to diagrams already published in [31, 42, 43], shows that the efficiency and power losses are nonlinear and complex in real vehicles. The complexity and the nonlinearities are the results of the following aspects, which describe how every single component of a traction chain has specific nonlinear complex losses:

- On electric rail vehicles, the electricity results in heat losses which depend on the square of the electric current [44].
- Electric motors and transformers have hysteresis losses and eddy current losses, which are nonlinear [45, 46] and can be described by variable efficiency charts or maps [31, 47–49]. For transformers and asynchronous motors, the losses can be divided into constant *power-independent losses* and additive *power-dependent losses* [50, 51]. Those losses depend on technology and construction: in the case of motors and transformers, this is caused due to basic magnetization, which requires energy independently from the output [44, 51]. In order to reduce these losses, e.g. unnecessary motors are increasingly being switched off if it is technically possible [33].
- Mechanical components like gears and bearings have a nonlinear characteristic, dependent on force and speed [29, 52].

- The power inverter has losses depending on the clocking frequency, which changes with the engine speed. This generates losses due to the basic clocking and is dependent on the power and speed, with some discontinuities and jumps [5, 10]. The discontinuity can also be observed in Fig. 2b and arises due to clocking frequency changes.

Note that not only electric traction components have such nonlinear losses. They occur also on diesel engines [31, 53–56].

Also, note that a part of the power-independent losses, which are included in Fig. 2, arises due to several auxiliary components. In [10], the losses due to those auxiliaries are described with a constant value. They can disappear when the vehicle is at a standstill or switched off [3, 28–30, 43].

3.4 A classification of increasing level of complexity in modelling energy flows

Due to this high complexity, it is understandable how the research has been obliged to simplify those effects in mathematical models. We propose a classification to compare the different simplification steps of the energy irreversible consumed E_{IC}. We will use the suffix of a level to classify the complexity in modelling energy flows. Thereby the complexity increases with the increasing level. We will call it long level of complexity in modelling energy flows or just 'level' for short. Table 1 and the following paragraphs give an overview of the levels of complexity in modelling energy flows. In subsequent denotations, L0, L1, L2 and L3 stand for Levels 0, 1, 2 and 3, respectively.

3.4.1 Level 0: No vehicle components losses

This is the simplest case for modelling the energy flow. Most EETC optimization approaches consider a constant, unitary efficiency $\eta = 1$ and ignore all losses and power consumption in the vehicle components. In case of non-regenerative braking, the optimization would match the usual EETC approach (see e.g. [19–21]). This approach considers the total energy irreversibly consumed $E_{\mathrm{IC.L0.NOREG}}$ as sum of resistance and energy lost in dissipative braking:

$$E_{\mathrm{IC.L0.NOREG}} = \underbrace{\int_{t_1}^{t_2} P_{\mathrm{R}}(t) + P_{\mathrm{BD}}(t)\mathrm{d}t}_{\text{Contains losses until the current time point}} + \Delta E_{\mathrm{Kin}}. \tag{7}$$

Equation (7) also considers that some energy might be stored in the system as kinetic energy, which will be reduced later point to reach a target speed. If no regeneration is considered, this variation in kinetic energy can be done only by dissipative brakes. We prefer to use Eq. (8) as equivalent expression where this kinetic energy does not need to be identified and instead consider only the integral of positive mechanical traction forces at the wheel or mechanical power at the wheel:

$$E_{\mathrm{IC.L0.NOREG}} = \underbrace{\int_{t_1}^{t_2} \max\left\{P_{\mathrm{Mech}}(t), 0\right\}\mathrm{d}t = \underbrace{\int_{s_1}^{s_2} \max\left\{F(s), 0\right\}\mathrm{d}s.}_{\text{Usual EETC approach, here used for L0.NOREG}}}_{\text{Contains additional dissipative brakes losses before occurring}} \tag{8}$$

In the case of regenerative braking, the mechanical power and the traction force will also be integrated over their negative values, and dissipative brakes are not used. Only the driving resistances losses must be minimized, and the energy irreversibly consumed can be evaluated (see Eq. (9)):

$$E_{\mathrm{IC.L0.REG}} = \int_{t_1}^{t_2} P_{\mathrm{R}}(t)\mathrm{d}t. \tag{9}$$

3.4.2 Level 1: Vehicle component losses as a constant proportion of mechanical power

At this level, the total vehicle efficiency $\eta < 1$ is described by a single constant value $\overline{\eta}$ for all engine operating points. This simplification is justified by the physical fact that an engine's higher mechanical power output generates more power losses. This is the most used case if the EETC considers traction component losses (e.g. used by [7]). For determining the energy optimal trajectory with this level, Eq. (10) can be used for describing the energy irreversibly consumed $E_{\mathrm{IC.L1}}$:

$$\begin{aligned} E_{\mathrm{IC.L1}} &= \int_{t_1}^{t_2} \left(P_{\mathrm{R}}(t) + \underbrace{P_{\mathrm{BD}}(t)}_{\substack{\text{In the following zero}\\ \text{because of}\\ \text{100 \% regeneration}}} + \underbrace{P_{\mathrm{L.Pdep}}(t)}_{\substack{\text{Constant proportion}\\ \text{of mechanical power}}} \right) \mathrm{d}t \\ &= \int_{t_1}^{t_2} \left(P_{\mathrm{R}}(t) + \left(\frac{1}{\overline{\eta}} - 1\right) P_{\mathrm{Mech}}(t) \right) \mathrm{d}t. \end{aligned} \tag{10}$$

3.4.3 Level 2: Vehicle component losses as the sum of a constant proportion of mechanical power plus a constant value independent of mechanical power

By considering a power independent part, the previous level of detail is extended. The physical justification for this form comes from the previously described power-independent losses, which arise in the traction chain. It is a sufficiently complex simplification of the reality. We could find no application of this level in the literature. However, it is useful to introduce it, as it is a very simple model with a perceivable effect on the trajectories. To determine the energy optimal trajectory with this level, Eq. (11) can be used for describing the energy irreversibly consumed $E_{\mathrm{IC.L2}}$:

$$E_{\mathrm{IC.L2}} = \int_{t_1}^{t_2} \left(P_{\mathrm{R}}(t) + \underbrace{P_{\mathrm{BD}}(t)}_{\substack{\text{In the following zero}\\ \text{because of}\\ \text{100 \% regeneration}}} + \underbrace{P_{\mathrm{L.Pdep}}(t)}_{\substack{\text{Constant proportion}\\ \text{of mechanical power}}} + \underbrace{P_{\mathrm{L.Pind}}}_{\text{Constant value}} \right) \mathrm{d}t$$
$$= \int_{t_1}^{t_2} \left(P_{\mathrm{R}}(t) + \left(\frac{1}{\eta} - 1 \right) P_{\mathrm{Mech}}(t) + c_0 \right) \mathrm{d}t. \tag{11}$$

3.4.4 Level 3: Description of vehicle component losses with high detail, considering dynamic efficiency

The engines in rail vehicles can be operated at different engine operating points. A change in speed or traction force leads to a change in the engine operating point. The traction component's efficiency is not constant but depends on the engine operating point. This level describes the losses with high quality and can arbitrarily consider the nonlinearities and discontinuity jumpings. It was used with maps by [8, 9] and simplified with nonlinear functions (and no discontinuity jumps) by [10]. In general, engine data for this level can be precisely described by the usage of high-resolution loss- and efficiency-maps [29–31, 43, 48, 49, 54, 55]. The only approximation error of this class of models is in the possible effects, which do not depend only on force and speed, but for instance, on the accumulated temperature on the electrical components; or depending on other variables which cannot be expressed as dependent on speed, force, acceleration or mechanical power. For the energy optimal trajectory with this level, the Eq. (12) can be used:

$$E_{\mathrm{IC.L3}} = \int_{t_1}^{t_2} \left(P_{\mathrm{R}}(t) + \underbrace{P_{\mathrm{BD}}(t)}_{\substack{\text{In the following zero}\\ \text{because of}\\ \text{100 \% regeneration}}} + \underbrace{P_{\mathrm{L}}(F(t), v(t))}_{\substack{\text{Include const value,}\\ \text{constant proportion}\\ \text{of mechanical power}\\ \text{and nonlinearities}}} \right) \mathrm{d}t$$
$$= \int_{t_1}^{t_2} \left(P_{\mathrm{R}}(t) + \left(\frac{1}{\eta} - 1 \right) P_{\mathrm{Mech}}(t) + c_0 + \varepsilon(F, v) \right) \mathrm{d}t. \tag{12}$$

4 Evaluation of the different classes of models

The goal of this section is to compare trajectories, which are derived when optimizing for the different levels introduced, to describe the losses from vehicle components. To enable this, we use an optimization algorithm, which is described in Sect. 4.1. Then, we focus on specific influences towards different driving modes like cruising, acceleration and deceleration. Finally, we compare full trajectories.

4.1 Goal and structure of the optimization algorithm

We are dealing with the movement of trains; our solution space includes train trajectories (i.e., when and how much to accelerate, which cruising speed to hold, how long should one possibly coast over a suitable space, and when and how much to brake). We solve the traditional EETC problem, that is we want to determine the series of control actions (acceleration and braking over time), and consequently speed and space over time, such that the total energy consumed for a task of going from a location s_1 at a speed v_1 to a location s_2 at a speed v_2. The solution trajectory is the kinematic description of those control actions in terms of speed, acceleration/braking. The total energy of a trajectory is the total *irreversibly consumed energy* evaluated at the end of the trajectory.

As standard, we start by determining the trajectory of minimum time. Compared to this, we consider extra travel time $t+$ and enforce that the train must arrive at the end location within $t+$ of the fastest trajectory. We represent $t+$ as a percentage of the minimum running time, as is usual in timetable planning. To more comprehensively study this problem, we study the entire Pareto front of trajectories; i.e. we determine all trajectories that minimize total energy, given a set of $t+$. For the sake of simplicity, we call those

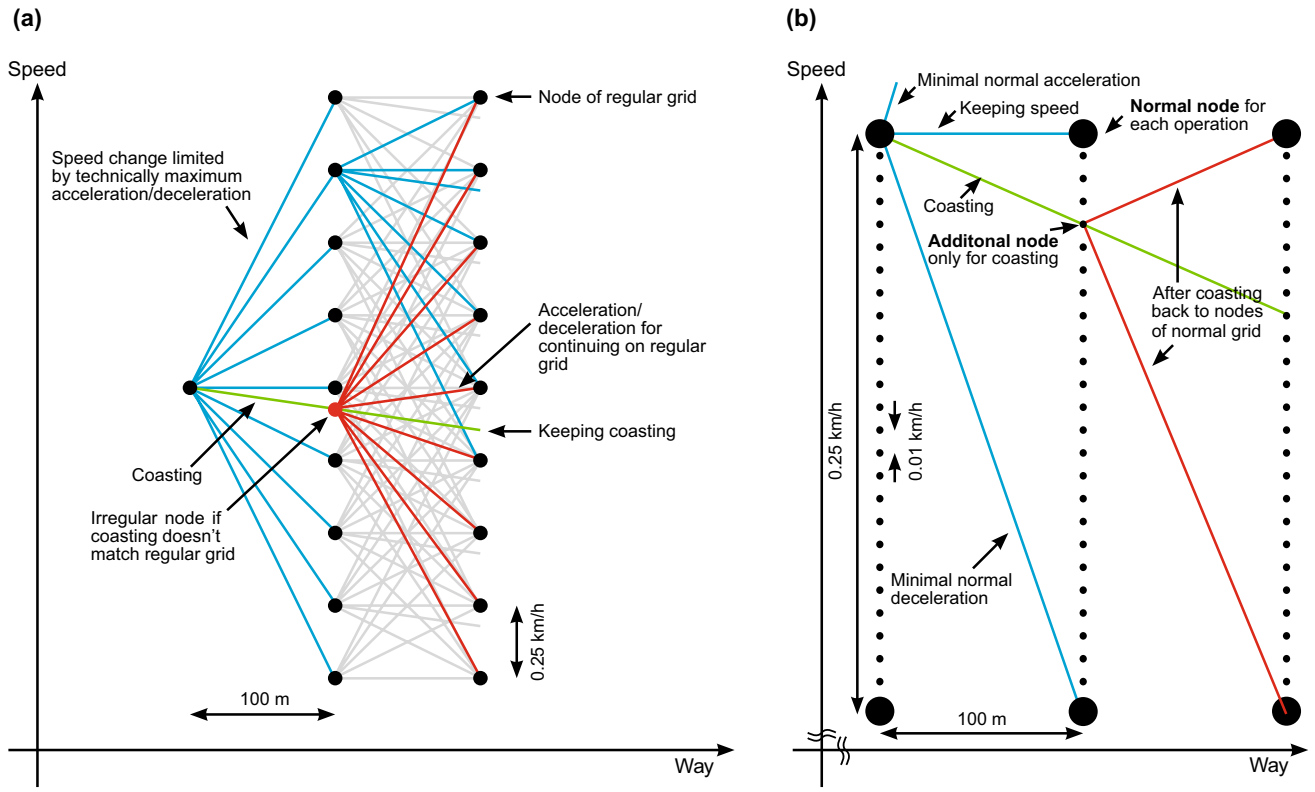

Fig. 3 Schematic visualization of the used grid: **a** macroscopic view of the possible acceleration and deceleration; **b** detail enlargement, which shows standard nodes, and the additional nodes only used to model coasting

Table 1 Overview of the level of complexity in modelling energy flows

Level of complexity in modelling energy flows	Power-dependent losses	Power-independent losses	Nonlinearities and discontinuity jumps	Usage of regenerative brake	Detail level of vehicle component losses and comment
0	–	–	–	No	Component losses are completely ignored; constant efficiency ($\eta(F, v) = 1$)
				Yes	The case, which does not use the regenerative brake is called "Level 0 NOREG"; the case with regenerative brake "Level 0 REG"
1	Included (proportional to P_{Mech})	–	–	Yes	Simple approximation of losses proportional to P_{Mech}, constant efficiency $\eta(F, v) = \overline{\eta}$
2	Included (proportional to P_{Mech})	Included (constant c_0)	–	Yes	Improved approximation of losses
3	Included (proportional to P_{Mech})	Included (constant c_0)	Included as nonlinear factor $\epsilon(F, v)$	Yes	Highly detailed description, very close to the reality

Note: These levels aim to describe the component losses; each of them can be handled with and without regenerative braking—here we use only Level 0, the no regenerative energy dissipating brake

"energy-optimal trajectories given a travel time $t+$" just as energy-optimal trajectories. We try to keep this Pareto analysis for most of the paper and focus on a single $t+$ value for a single analysis only. In reality, $t+$ is often normatively designed to be 5% or 7%. Such a $t+$ is then considered in the timetable design. We assume trains will arrive punctually and do not tradeoff a further energy-reduction with delayed arrival.

The problem of finding those energy-optimal trajectories can be solved by a variety of methods which have been proposed in the literature. As a background, we describe the approach used here, with no claim that it substantially contributes to the state of the art, as it is dynamic programming on a discretized solution space. We remark that given the nonlinearities and complex expressions, indirect methods cannot help, and direct methods computing constructively the trajectory must be used. Moreover, due to the various components' interplay, there are few to no global pruning or heuristics that one can use for all levels. Hence, the resulting algorithm is relatively simple.

We use a speed–space discretized grid which identifies states, represented as nodes. This grid is built with uniformly sampled space values of 100 m, as shown in Fig. 3. We consider uniformly sampled speed values at two levels of detail, as explained later. We can only consider trajectories which connect nodes of this two-level grid at successive spaces, depending on the applied traction force or braking. For any possible force or braking, the resulting speed and travel time and energy can be computed. By summing up all travel times and all energies, we can compute the total travel time and total energy consumption.

The grid size is typically determined as a tradeoff between computation requirements and approximation. If the grid would be too large, fewer operation points in the traction force diagram can be chosen for accelerating and decelerating. This would not result in noticeable difference when looking at a speed diagram. However, in the resulting path through the traction force diagram, it can be observed that only operation points above or below the highest efficiency with a zigzag can be reached, and no stable continuous path. Finally, the trajectory would not be along the lowest energy consumption, above a certain threshold for the grid size. This threshold depends on the calculation and train parameters, and some preliminary testing was done. For our parameters, we determined that a grid size must be smaller than approximately 1 km/h by 100 m or 0.5 km/h by 200 m. We use a maximum step size of 0.25 km/h to be much below the threshold.

A double-level grid is considered as follows. For the normal acceleration, deceleration and cruising, we have a coarser grid, uniformly sampling speed with a step size of 100 m and 0.25 km/h (see Fig. 3a). For coasting, we have an additional finer grid (see Fig. 3b), with the possibility to consider speeds at the level of detail of 0.01 km/h. This finer grid is not complete; i.e. only one arc is considered, which matches the coasting action, depending on the current speed (i.e. on the previous node). The grids are such that the more fine-grained nodes can be reached only by means of a coasting action while applying any other type of speed change will have to be matched to a node of the coarser, regular grid.

We remark following aspects for our simulation: No rules are used to limit the calculation amount (e.g. by limiting the speed range or considering only 4 actions: maximum acceleration, cruising, coasting and maximum braking). We also do not consider intermediate target constraints on speed (e.g. the train path envelope of [20], or a pre-specified cruising speed) but optimize from the beginning to the end. We use a very small threshold for determining dominance of a trajectory over another one, compatible with the machine resolution enabled by our environment (Matlab). Thereby dominated trajectories are removed in both time and energy; non-dominated one are kept; in case multiple trajectories have the same time and energy one is kept randomly. For very low speed we consider a rougher detail, and provide later a justification for this. The maximum speed change depends on the resolution of the grid and of course, on the technical possibility of the vehicle. It is further should be noted that below 20 km/h, nodes are used only to model acceleration/braking from/to a standstill, and cruising at this slow speed is not considered optimal.

The optimization process works iteratively with a loop starting from the first waypoint and continues over each waypoint until the last; the number of *trajectory-variants* increases by adding *trajectory-segments* during each iteration (we call trajectory-segments to describe the segments of a trajectory between two waypoints, and call trajectory-variants to describe the number of various trajectories containing more than one trajectory-segment). The Appendix (Figs. 12 and 13) shows more in detail the simulation structure as a series of detailed flowcharts. The optimization loop can be divided into four parts.

The first three parts add all technically possible trajectory-segments until the next waypoint. The first one is for the regular trajectory-segments, which describe acceleration, deceleration and keeping speed in the coarse grid (blue in Fig. 3). The second adds all technically possible coasting trajectory-segments and transfers them to the finer grid (green in Fig. 3). The third adds the after-coasting trajectory-segments, which are used to transfer the trajectory-variants back after coasting from the finer to the coarser grid (red in Fig. 3).

In the final fourth part of the loop, the number of trajectory-variants is reduced on each node by considering Pareto dominance. Each trajectory-variant is discarded if there is another trajectory-variant which is faster and simultaneously has a smaller total energy consumption.

The final result is a list of non-dominated trajectories on the Pareto front between travel time and energy. This enables us to determine the optimal one, given any target travel time. These four parts are the same for each level introduced and described in this paper. We only change the equations (mentioned in Sect. 3) and parameters for calculating the trajectory-segments depending on the level.

The number of trajectory-variants depends on the level. L0 REG generates the highest amount. At any (space, speed) node, there can be more than 100,000 trajectory-variants. Depending on the level, grid solution, length and amount of saving controlling data, the storage demand is between approximately 5 and 100 GB RAM for the simulation, by calculation time between single minutes up to 72 h. Due to the high resolution, at L0 REG we further refine very small numerical instabilities by filtering trajectories which alternate coasting with accelerating over successive nodes, i.e. those where the coasting phase is broken into two successive intervals, separated by just a node spent non-coasting.

Note that the particularly long calculation time results from the many variants; we consider a very high resolution, using no pruning based on constraints, to reduce variants or computation. Further, it is a simple algorithm not exploiting parallel computing. This was necessary to be sure to avoid possible faults due to a complex algorithm and avoid removing the most energy-efficient variant. With a slightly larger grid and a much more complex algorithm based on constraints or heuristics, and parallel computing, the calculation time can be reduced a lot. Computation times of seconds are possible for the industrial application (see [30, 37]). For the potential and description to solve this EETC by parallel computing, we refer to our industry report [30].

We do not use a mix of both regenerative and dissipative brakes but only one of the two. We only study in detail the dissipative brakes for L0. To allow for a fair comparison, the maximum acceleration and braking forces are the same throughout the levels. In other words, in Level 0, either regenerative or non-regenerative (i.e. dissipative braking), we assume that the deceleration rate is the same (as limited by the power hyperbola). The only difference between the two is that in case of no regenerative brake, the braking energy is lost.

We consider as benchmarks also two typical solutions, namely the one with the minimal travel time and the trajectory called reduced maximum speed (RMS), also discussed in [7], which prescribes only three phases, acceleration with

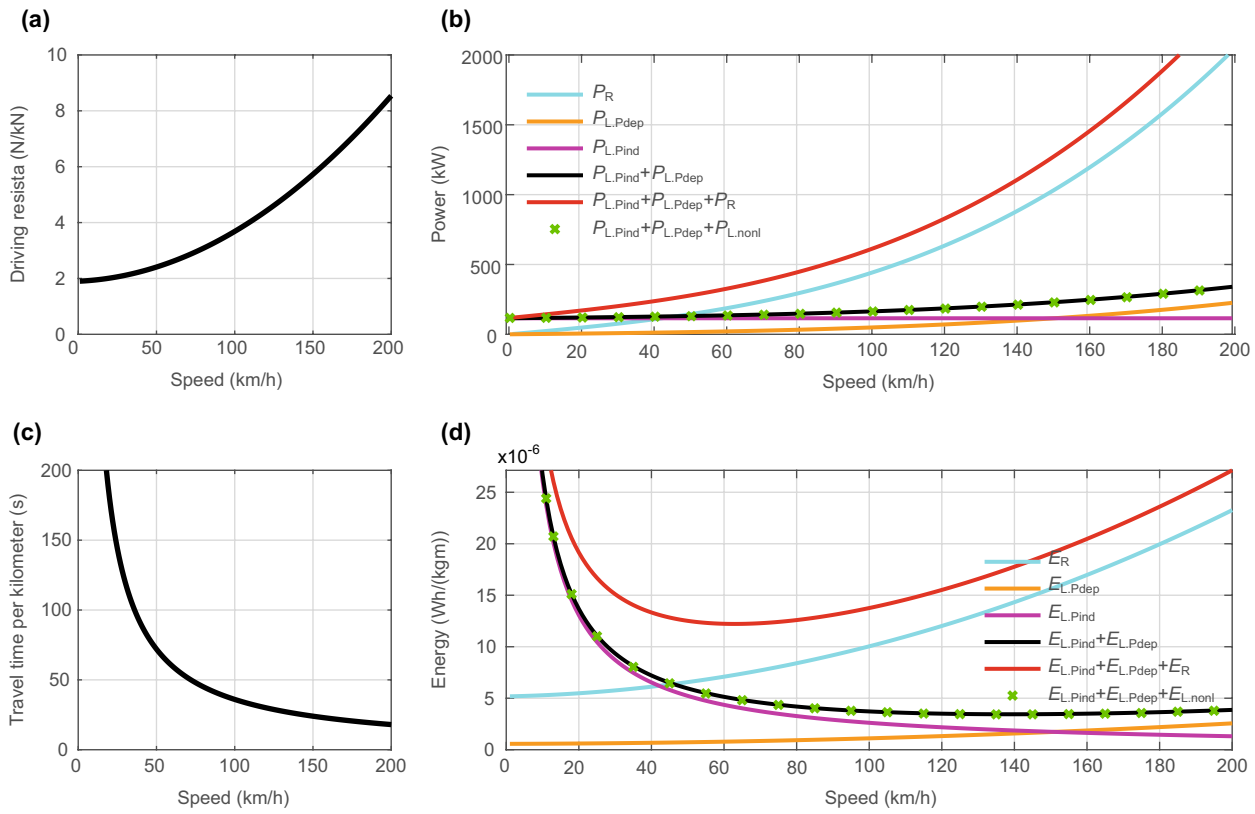

Fig. 4 Quantification of the values to drive different constant speeds: **a** specific driving resistance; **b** different kinds of power versus speed; **c** travel time per kilometer versus speed; **d** Pareto front of specific energy consumption to drive 1 km

maximum acceleration, cruising and braking with maximum braking. In this case, the only variable is the cruising speed, which is chosen to match the total travel time given.

4.2 Cruising: energy-efficient minimum speed

We start our analysis by considering the effect of constant, power-independent energy losses on the relation between travel time and energy. This effect is not relevant to the task of finding the optimal trajectory for standard EETC as far as it assumes a constant, given travel time (and therefore a constant amount of losses depending on it). We anyway show that this factor can have impacts if the target travel time is sufficiently long. The inclusion of the constant term results in a non-monotone relation between speed and energy, with very low speed resulting in higher energy consumption than faster speeds. We show this effect by assuming a vehicle driving at a constant speed between two given locations and varying this speed, and in the meantime neglect slopes, curves, acceleration and deceleration. Given that the speed is constant, we can easily shift to consider the specific energy per distance; we thus analyse the *energy irreversibly consumed* per distance s (e.g. specific energy per meter or kilometer) for driving with a constant speed v, varying the speed and therefore the total travel time.

The energy irreversibly consumed during driving E_{IC} over a fixed space can be computed as follows. Mechanical energy is in this case only needed to overcome the driving resistance (no increase in kinetic energy nor potential energy). This is described with the quadratic Davis formula, dependent on the speed $v \in [0, v_{\max}]$ and three constants $c_1, c_2, c_3 \geq 0$ [31, 40, 41]. With power proportional losses with a constant $\eta = \overline{\eta}$, Eq. (13) results:

$$E_{\mathrm{IC}}(v) = \underbrace{\left(c_1 + c_2 \cdot v + c_3 \cdot v^2\right)s}_{\text{From } P_{\mathrm{R}}} + \underbrace{\left(\frac{1}{\overline{\eta}} - 1\right)\left(c_1 + c_2 \cdot v + c_3 \cdot v^2\right)s}_{\text{From } P_{\mathrm{L.Pdep}}} + \underbrace{c_0 \cdot \frac{s}{v}}_{\text{From } P_{\mathrm{L.Pdep}}} + \varepsilon(F, v) \cdot \frac{s}{v}. \tag{13}$$

When neglecting the nonlinear term with $\varepsilon(F, v)$, Eq. (14) results:

$$E_{\mathrm{IC}}(v) = \underbrace{\underbrace{s \cdot \frac{1}{\eta}\underbrace{\left(c_1 + c_2 \cdot v + c_3 \cdot v^2\right)}_{\text{Quadratic}}}_{\text{L0 and L1}} + \underbrace{s \cdot \frac{c_0}{v}}_{\text{Hyperbolic}}}_{\text{L2 and L3}} \tag{14}$$

Equation (14) shows how the *energy irreversibly consumed* is for Levels 0 and 1 quadratic in the speed, so that lower speed reduces the energy consumption. In this case, the optimal solution is always to drive as slow as possible. From Level 2 onwards, a hyperbolic part is added, which is the result of the constant *power-independent losses*. This part increases for low values of v; driving at very low speed requires so much time that the energy would increase too much. This also justifies the common assumption to avoid considering very low cruising speeds.

In the following, we quantify those insights on an example, modelled with a description at Levels 2 and 3. As a vehicle, we consider a ~ 200-m train with one electric locomotive of type Re460, (84,000 kg, 6.1 MW) and 7 standard UIC passenger coaches, with a total weight of 440,000 kg, including passengers (such trains are routinely used in Switzerland and many European countries). We consider an exemplar distance of 1000 m. Using data from SBB vehicles, we describe this train by the following resistance parameter (c_1, c_2 and c_3) which are calculated according to Sauthoff [27, 41]:

$$\begin{aligned} c_1 &= 8.2\,\mathrm{kN}\left(\frac{\mathrm{m}}{\mathrm{s}}\right)^0, \\ c_2 &= 0.0388\,\mathrm{kN}\left(\frac{\mathrm{m}}{\mathrm{s}}\right)^{-1}, \\ c_3 &= 0.0086\,\mathrm{kN}\left(\frac{\mathrm{m}}{\mathrm{s}}\right)^{-2}. \end{aligned} \tag{15}$$

Level 3 parameters equivalent to those reported in Fig. 2 and parameters for a Level 2 description from a fitting procedure (less than 3% fitting error overall engine operating points) are taken as follows:

$$\begin{aligned} \overline{\eta} &= 0.89, \\ c_0 &= 115\,\mathrm{kN}\left(\frac{\mathrm{m}}{\mathrm{s}}\right)^1. \end{aligned} \tag{16}$$

The value of c_0 (Re460 GTO) describes the *power-independent losses*: a power consumption ≥ 115 kW is due at any moment the train is moving and not at standstill, for the vehicle components. At standstill this part of the consumption vanishes completely. The value $\overline{\eta}$ describes the efficiency.

Figure 4 analyses the trajectories when driving at different cruising speeds on a straight line without curves and slopes. Figure 4a shows the specific driving resistance of the Davis formula. Figure 4b visualizes the different considered kinds of power. P_{R}, the power loss to overcome the driving resistance, is the product of the driving resistance and speed. It starts at zero and increases progressively

with increasing speed. $P_{\text{L.Pdep}}$ arises through the demand of mechanical power at the wheel (here to overcome the driving resistances). It is calculated with the constant efficiency, and increases with increasing speed. $P_{\text{L.Pind}}$ is a constant value because it describes the *power-independent losses*, which arise as long as the train is not stopped. The green crosses describe the vehicle component losses in traction components and auxiliaries from the loss map of the real SBB vehicle, which includes also the nonlinear factor $\varepsilon(F, v)$. The crosses match with small deviations (0.5% on average) the black line, which describes the sum of $P_{\text{L.Pind}}$ and $P_{\text{L.Pdep}}$. The red line describes P_{IC} which is the sum all irreversibly consumed powers, namely P_{R}, $P_{\text{L.Pdep}}$ and $P_{\text{L.Pind}}$. Figure 4c visualizes the travel time, which is needed to drive 1 km. With the product of the power and the travel time to drive 1 km, the energy for 1 km can be calculated. This is shown in Fig. 4d, which also identifies the different kinds of specific energy versus speed (namely energy for the resistance E_{R}, energy losses dependent from the mechanical power $E_{\text{L.Pdep}}$, energy losses independent from the mechanical power $P_{\text{L.Pind}}$, and energy losses due to nonlinearities $E_{\text{L.nonl}}$). The values are the result of the product of the power and travel time.

Overall, Fig. 4 shows that the energy consumption per kilometer decreases with decreasing speed until a certain point is reached (for this vehicle, around 63 km/h by considering the black continuous line). Below this speed, the consumption increases again. Those speeds are though reached only for very large travel time buffers (normal values would be in the range of 5%–7%, see [57]). This effect is due to the speed independent losses and the energy consumers included in Level 2. Going at a lower speed, the speed independent consumers consume energy for a longer time compared to driving the same distance at a higher speed. A theoretical solution would be switching off the entire traction components, but this is not completely possible. For instance, even during coasting, when power exchanged is 0, there might be some residual consumption in the engines. The increasingly used, highly efficient permanent-magnet-synchronous-motors (PMSM) cannot achieve a consumption of 0 (i.e. be completely switched off) as far as speed is different than 0, i.e. during driving.

As long as a vehicle has power independent energy consumption, which can be switched off only at a standstill, there would be an energy-efficient minimum speed, which is vehicle specific. This provides a technically justified reason to reduce the speed range for trajectory optimization variants by avoiding excessively lower speeds.

4.3 Acceleration: reduced acceleration can save energy

We here consider the case of a train accelerating from a standstill to a target speed, given a maximum travel time which we vary. We report the trajectory as a function of the travel time buffer $t+$, which is the extra travel time compared to the technically minimum one. The comparison will be made using different levels, with their different energy functions, which cause different trajectories and energy savings.

For understanding, we have prepared Fig. 14 in the Appendix, which is a graphical description of the encoding of speed, acceleration, space and efficiency using two diagrams with colors and markers. This allows to show the trajectory as it moves through different engine operating points ('speed over space' and 'force over speed'). Namely, Fig. 14a shows the speed–space diagram of a usual train trip. Figure 14b enables showing the path through the force–speed diagram. We consider both positive and negative forces in Fig. 14b. Crosses are spaced regularly in space. The two diagrams are indexed by color and crosses, which report the space development, to match the same trajectory (and engine operating points).

Figure 5 gives an overview of the optimal trajectories, given the different levels. We consider an acceleration from a standstill to 160 km/h (we are considering 160 km/h because most lines in Switzerland are limited to 160 km/h or less) over a distance of 4 km. Figure 5a, b reports 6 travel time buffers, varying from 0 (green) to infinite (black) via the colors blue, purple, red and yellow. To ease the visual match between the top two rows (Fig. 5a, b), the color of the trajectories (in speed–space diagrams) and the paths (in force–speed diagrams) are both changing saturation from brighter (beginning) to grayish (end). We also use crosses to identify points equidistant in space. Thus, it is possible to understand to which part of the trajectory a specific point in the force–speed diagram corresponds. As an example: in the leftmost column, middle row, the path in brown ($t+ = 25\%$) starts from the top of the force–speed diagram, follows the boundary of the power hyperbola, goes down to zero force, goes up again to the power hyperbola, and goes finally down at the end of the trajectory.

Given the plots in Fig. 5a, we analyse how the shape of the trajectory changes for increasing travel time buffer for each level. In Level 0, no regeneration, which corresponds to the traditional EETC problem, there is a starting phase accelerating with maximum acceleration. When the travel time buffer increases, this is split into two acceleration phases (in Fig. 5a, one at 0 km and another at 2 km), with a mixture of coasting and cruising between the two, to use the travel time buffer. The trajectories of Levels 1 and 2 are similar to Level 0, with minimal differences, which arise due to the different energy functions. The main difference is, Levels 0 and 1 do not discard trajectories with low speed. They allow trajectories with the maximum travel time buffer ($t+ = \max$) until the bottom limit of the simulation (20 km/h). However, Level 2, for this $t+$, discards trajectories with a too-low speed because this too-low speed is energetically inefficient

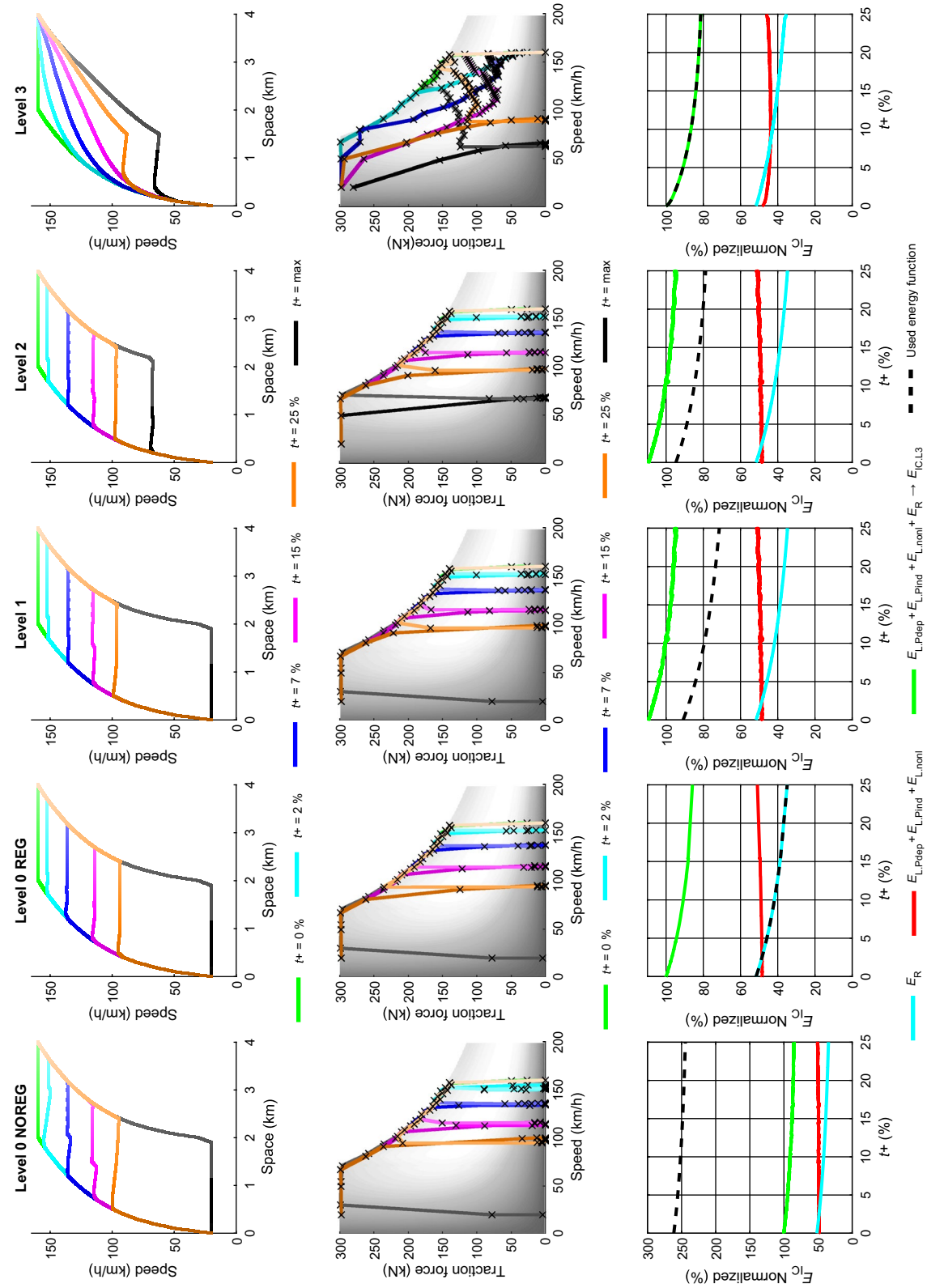

Fig. 5 Acceleration with the different levels: **a** trajectories of the different levels for different amounts of travel time buffer $t+$; **b** paths associated to the trajectories, through the efficiency map of the force–speed diagram (the bright areas within the diagram have a higher efficiency than the darker ones); **c** relative energy consumption dependent on the used travel time buffer, evaluated with the energy function of Level 3. Note: The energy function of Level 0 has a very high energy because it includes the entire kinetic energy of the train and not only the losses (as the kinetic energy is not regenerated). At Level 3 the sum is matching with the optimization function

(compare Sect. 4.2). Due to this, in Level 2, the trajectory with the maximum buffer ($t+ = \max$) always runs well above 20 km/h, actually at the energy-efficient minimum speed.

The necessity of considering the minimum speed can be observed and explained in Fig. 5c, which reports the E_{IC} of the trajectory, normalized by value computed with maximum speed and considering Level 3. The black dotted lines represent the used optimization function; the green lines represent the evaluation function, calculated by the most detailed Level 3 for all cases. For all levels, exploiting a larger travel time buffer decreases the driving resistance losses, but can increase the vehicle component losses (i.e. the sum of the *power-independent losses*, the *power-dependent losses*, and the nonlinear terms). Due to this, the energy consumption reduces for increasing travel time buffer and later increases in Levels 0 and 1. As the energy function of Levels 0 and 1 does not consider these influences, the trajectories with such low speed are not discarded; this happens instead with Level 2 upwards.

Figure 5b reports the paths in the force–speed diagram overlaid on the efficiency map of the traction chain. The crosses are equidistant in space; thus one can have a feeling on how much of a trajectory is spent, at which engine operating point. Overall the engine operating points visited by Levels 0 to 2 are rather regular, oscillating between the upper boundary (power hyperbola) and the bottom horizontal axis. This denotes the standard modes: maximum acceleration, cruising or coasting. Doing this, they spend much time in areas of lower efficiency, just crossing many times areas of higher efficiency in the middle of the diagram, but not spending much time there. See, for instance, how few crosses are reported in the middle areas of the diagram, which denote a very fast transition between other engine operating points. Level 3, instead, has more complex paths which aim to spend more time in areas of higher efficiency (many more crosses happen to be along the path in the intermediate areas of the diagram). The more efficient engine operating points in the middle of the force–speed diagram are used systematically. This results in a new driving mode: *reduced acceleration*. It can be observed that the very same algorithm from Levels 0, 1 and 2 generates the following two types of trajectories by using the energy description function of Level 3:

- The first type occurs with a short and medium travel time buffer. Here the trajectories start with a maximum acceleration and change to a *reduced acceleration* at a higher speed to operate the engine on more energy-efficient engine operating points. They use the increased travel time buffer for energy saving (compare trajectories with 2%, 7% and 15% $t+$ of Fig. 5a.
- The second type occurs with a higher travel time buffer (here 25% $t+$ in this example of Fig. 5). The trajectories start with a maximum acceleration, continue with a *reduced acceleration*, and have an intermediate coasting to use the travel time buffer. After the coasting, they continue with a *reduced acceleration* and a final short maximum acceleration. This intermediate coasting results from the aspect that a too much *reduced acceleration* is also disadvantageous because a too-low traction force is inefficient.

The switching point between these two trajectory types depends on the available travel time buffer, the target space, and the energy function, which is described by the vehicle-specific loss map. A comparison between the levels shows how the engine operating points which Level 3 uses are within the energy-efficient areas (compare crosses of Fig. 5b) and cause a *reduced acceleration*, with resulting energy saving. The energy functions of the lower levels instead use the engine operating points of the maximum acceleration because they ignore the fact that the highest energy efficiency is not matching the maximum acceleration.

Figure 6 analyses further the results of Fig. 5 and puts on the same plot the energy consumption (as evaluated by Level 3) shown in Fig. 5c. One can appreciate how energy consumption generally decreases with an increasing travel time buffer. Nevertheless, other phenomena are visible: the optimization strategy of Level 3 saves more energy with the same travel time buffer as the other ones, which are performing rather similar. This improvement can be quantified in 2.5% for a travel time buffer of 2%, and up to 4% for a travel time buffer of 15% and higher. The fluctuating lines of Levels 0, 1 and 2 results from the aspect that we report in Fig. 6 the energy consumption inclusive nonlinearities, i.e. computed optimizing Levels 0, 1 or 2, but evaluated considering Level 3. These nonlinearities are not included in the optimization function of Levels 0, 1 and 2. The optimization function considered to determine the optimal trajectory for Levels 0, 1 and 2 are non-fluctuating and can be seen in Fig. 5c, dotted lines.

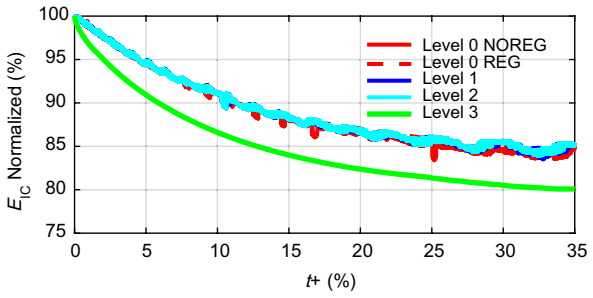

Fig. 6 Percentage reduction of energy consumption, as dependent on the travel time buffer available (x-axis), for an acceleration, considering the different levels, evaluated with the energy function of Level 3

4.4 Deceleration: using buffer for reduced braking reduces energy consumption

We describe the influence of the levels for the deceleration, again with a variable time to perform such a deceleration to a standstill. This case arises, for instance, when stopping at a station just after a previous train has to leave. Also here, the comparison will be made using different levels, with their different energy function, which causes different trajectories and energy savings. A total space of 4 km is considered, with a starting speed of 160 km/h. We report the results similar to the previous section. For the force–speed diagrams, we report only the quadrant with negative traction forces.

Figure 7 gives an overview of the different levels, similar to Fig. 6. We analyse the influence of increasing the travel time buffer for each level. In Level 0, without regenerative braking, the only way to exploit the travel time buffer is to coast; when abundant travel time is available, coasting starts at the beginning; but in any case, it requires full braking at the end, as coasting is not able to reduce the speed enough. For this reason, the leftmost column does not have points beyond 3% $t+$. Here, a strategy like initial braking and cruising or coasting is not advantageous for this energy function, which considers trains with non-regenerative brakes. In fact, all the energy that is in the system would be dissipated until the end by driving resistance, or by the non-regenerative brakes. Due to this, the only saving that can be achieved is to avoid using any traction force, by coasting as soon as possible. Many trajectories with a different mix of braking and coasting might end up with a similar final consumption.

In the case of regenerative braking, for Level 0, the trajectory is different: for an increasing travel time buffer, the cruising will be made stepwise with lower cruising speed and partially very short coasting, and finally, maximum braking. Compared to Level 0, the trajectories of Levels 1 and 2 change, because the description of the losses is more detailed. This results in coasting at intermediate speeds. These two levels have the following trajectory types:

- For a short travel time buffer (here up to 2% $t+$) the initial speed is kept. Following, there is coasting and maximum braking.
- For larger travel time buffer (here more than 2% $t+$), there is an initial maximum braking, then a coasting which reduces the speed, and finally, maximum braking. Compared to the Level 0 non-regenerative, the energy which is saved, due to the lower speed and longer coasting, can be regenerated back.

The main difference between Levels 1 and 2 is the different minimum speed (compare Sect. 4.2). Level 1 does not include in its optimization function the *power-independent losses*, which are included by Level 2. These *power-independent losses* result in the optimizer choosing not to drive too slow, because then the total *power-independent losses* are larger and would increase the trip energy consumption.

The force–speed diagrams describe regenerative braking with bounded values by a specular power hyperbola as the accelerations (like Ref. [13]). Similar to the accelerations, it can be seen that Levels 0, 1 and 2 are most of the time on the boundary of the power hyperbola (maximum deceleration) or coasting (no traction force). The area inside the force–speed diagram, which has higher efficiency, is only traversed between those modes. Instead, Level 3, with the highest detail, also has two other different trajectory types, which use engine operating points inside the force–speed diagrams:

- For a short travel time buffer, there is an initial cruising phase. After this, the engines are switched off, coasting. And finally, there is a maximum deceleration with regeneration.
- For a middle and long travel time buffer, there is a combination of *reduced braking* (similar to the *reduced acceleration*) to use the more energy-efficient paths through the force–speed diagram, with short coasting.

Figure 7c reports the energy consumption for increasing the travel time buffer. Apart from the black dotted line, which is the *energy irreversibly consumed* that the algorithm expects (which varies with the amount of factors considered), the other lines (which are evaluated under the same conditions of Level 3) have a similar shape. For instance, the red line (describing the losses) has a sharp decrease until a travel time buffer $t+$ of 3%, and then a slow increase as the travel time buffer increases. The total (green) has a sharp decrease and then a less sharp decrease with increasing $t+$.

To better highlight those differences, Fig. 8 puts together the normalized energy consumption (with 100% corresponding to no travel time buffer, evaluated in Level 3) as a function of the travel time buffer for all the levels considered. Apart from level 0 REG, for which coasting or cruising results in comparable energy, for short travel time buffers, up to 2%, all levels save the same amount of energy, as the only useful action to do is to coast. After that, the curves in the diagram separate with Level 0 NOREG, Level 1 and Level 2 consuming more energy than Level 3. In fact, Level 0 NOREG, Level 1 and Level 2 save energy by braking first and then coasting at a lower speed, where resistances are lower; and finally regenerating the residual energy. Level 3 instead saves also energy by driving at a lower speed; moreover, Level 3 drives at energy-efficient engine operating points and avoids the maximum braking. The gap between Level 3 and the other levels is

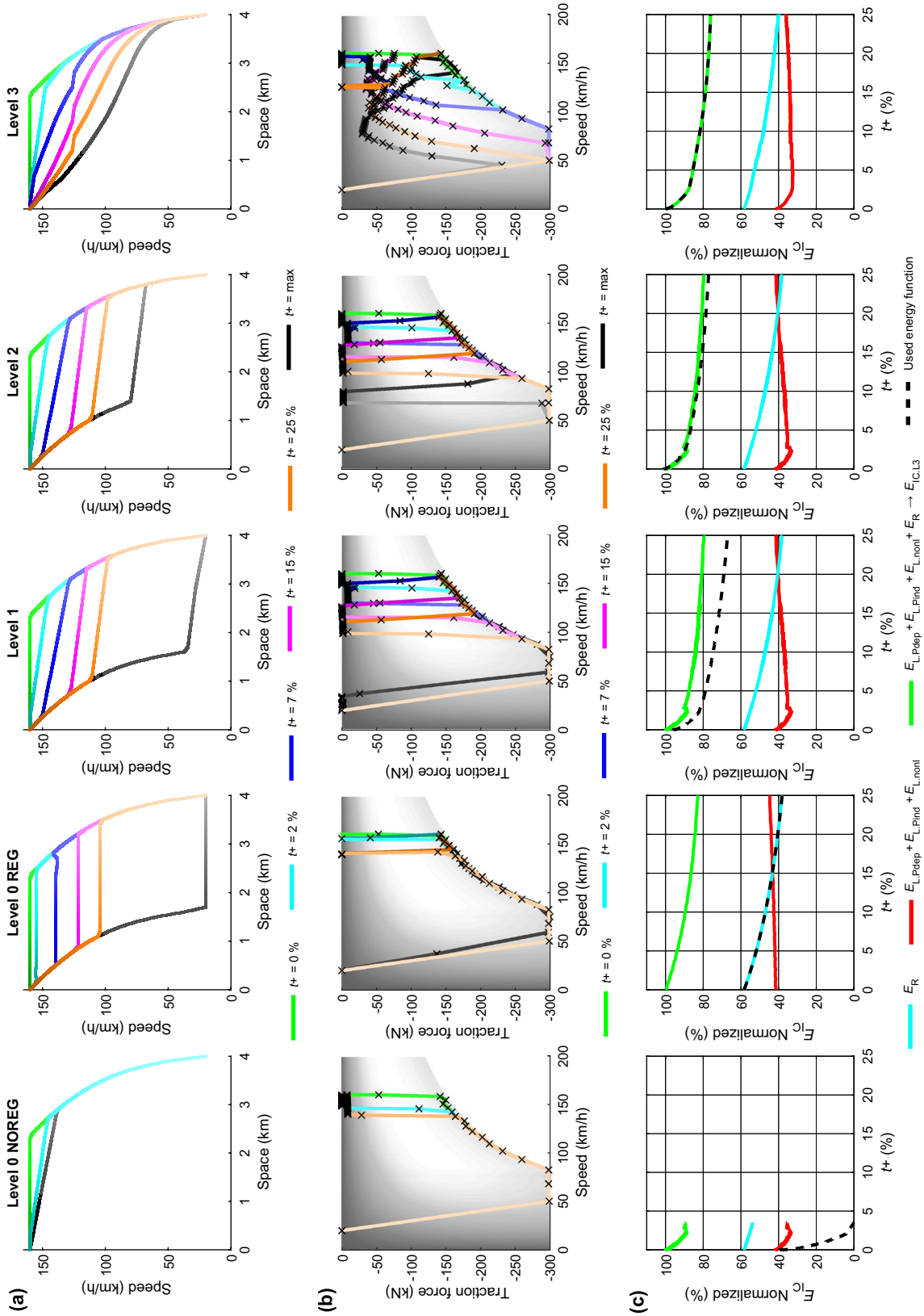

Fig. 7 Deceleration with the different levels: **a** trajectories of the different levels for different amounts of travel time buffer $t+$; **b** paths associated to the trajectories, through the efficiency map of the force–speed diagram (the bright areas within the diagram have a higher energy efficiency than the darker ones); **c** relative energy consumption dependent on the used travel time buffer, evaluated with the energy function of Level 3

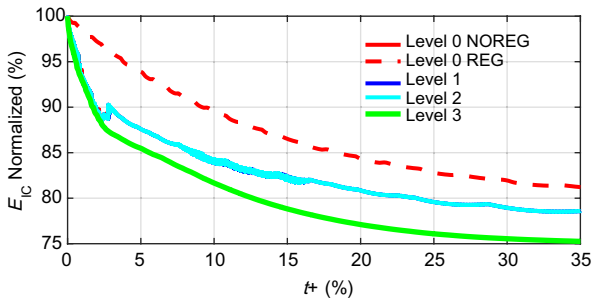

Fig. 8 Percentage reduction of energy consumption, as dependent on the travel time buffer available $t+$ (x-axis), for a deceleration, considering the different levels, evaluated with the energy function of Level 3. Note: The fluctuating lines of Levels 0, 1 and 2 result from the aspect that this diagram shows the energy consumption inclusive nonlinearities. These nonlinearities are not in the optimization function of Levels 0, 1 and 2. Non-fluctuating optimization function of Levels 0, 1 and 2 can be seen in Fig. 7

around 2% for a travel time buffer of 5%, increasing up to 4% for very large travel time buffers.

In Level 3, the coasting is used for what possible. When the travel time buffer is so large that the coasting potential is exhausted, the optimal energy-efficient trajectory is the one that keeps the engine at the most efficient engine operating points. In this sense, we can conclude that the positive effect of coasting is stronger than the positive effect of driving at energy-efficient engine operating points.

Compared to Level 0 NOREG, Level 0 REG might look strange, because Level 0 REG has here the highest energy consumption. However, for this diagram, all trajectories are evaluated with Level 3, which uses regeneration. Trajectories which assume they can use regeneration in combination with cruising at a lower speed, and expect no losses from this choice, have a higher energy consumption (Level 0 REG) compared to trajectories which use coasting (Level 0 NOREG, Level 1, and Level 2). Trajectories which use additional *reduced braking* have the lowest energy consumption (Level 3). This positive effect of coasting also remains for longer travel time buffers.

4.5 Complete trajectories: changes by the dynamic losses of Level 3

We could so far analyse how a higher detail in modelling component losses has an influence in the shape of the trajectory. A take-home message is that during the acceleration phase, the highest potential for energy reduction is available when taking care of the complex nonlinearities of the component losses. We now study a complete trajectory, starting and ending at a standstill, over a distance of 6 km and a maximum speed of 160 km/h. We report the analysis in Fig. 9, in an analogous manner to the previous Figs. 5 and 7. In this case, the force–speed diagram covers both negative and positive traction forces, as both the acceleration and deceleration phases are considered. Compared to before, here the algorithm also decides on how much $t+$ should be spent in the acceleration, cruising or deceleration phase.

Overall, Levels 0, 1 and 2 result in very similar trajectories, which also match the results of the traditional EETC literature, with the 4 modes of maximum acceleration, cruising, coasting and maximum braking.

In Level 2, trajectories which use a low speed are not considered optimal (see the black trajectory with infinite travel time buffer), because a too-low speed increases the energy consumption; this has effects only for very large travel time buffers. A different change arises when comparing Level 3 with the others. Its shape considers the maximum acceleration and coasting only for the short travel time buffer. If more travel time buffer is available before reaching the top speed, the trajectory uses *reduced acceleration* (see the smooth end of the acceleration phase for trajectories with $t+$ bigger than 0) and coasting. Similar to Level 2, trajectories with low speed are discarded.

Overall, trajectories optimal for Level 3 exploit much more both the positive and negative part inside the force–speed diagram, to spend as much time as possible at efficient engine operating points. The crosses along the path in the force–speed diagram (which gives an idea of the time spent at which speed and traction force) are few for Level 0, and even less for Levels 1 and 2. Level 3 instead has elaborated paths, which are smooth as they spend much time at intermediate locations of the diagram, especially for the positive traction force (acceleration).

Figure 9c reports energies for increasing the travel time buffer, where the most visible difference is again that the dotted black lines changes position (i.e. different optimization functions are used), but the overall shape varies to a much smaller extent. In this sense, the optimal trajectory for a level is not too far away from the optimal trajectory of another level. Minor differences pertain to nonlinear effects at intermediate travel time buffers, as well as a different gap between the light blue and the red line. Those effects are analysed in more detail later in the paper.

4.6 Heterogeneity of vehicles: energy-efficient trajectories depend on the specific vehicle

In this last analysis, we show that other vehicles, with other loss maps, result in differences on the Level 3 optimal trajectories. Until now, we have used the loss map of a vehicle with a 3-point GTO power inverter. However, the locomotives considered are being retrofitted with new, more energy-efficient, 3-point IGBT power inverters. We remark that this refit is only a change of electrical components, while resistance, weight and other characteristics remain the same. The

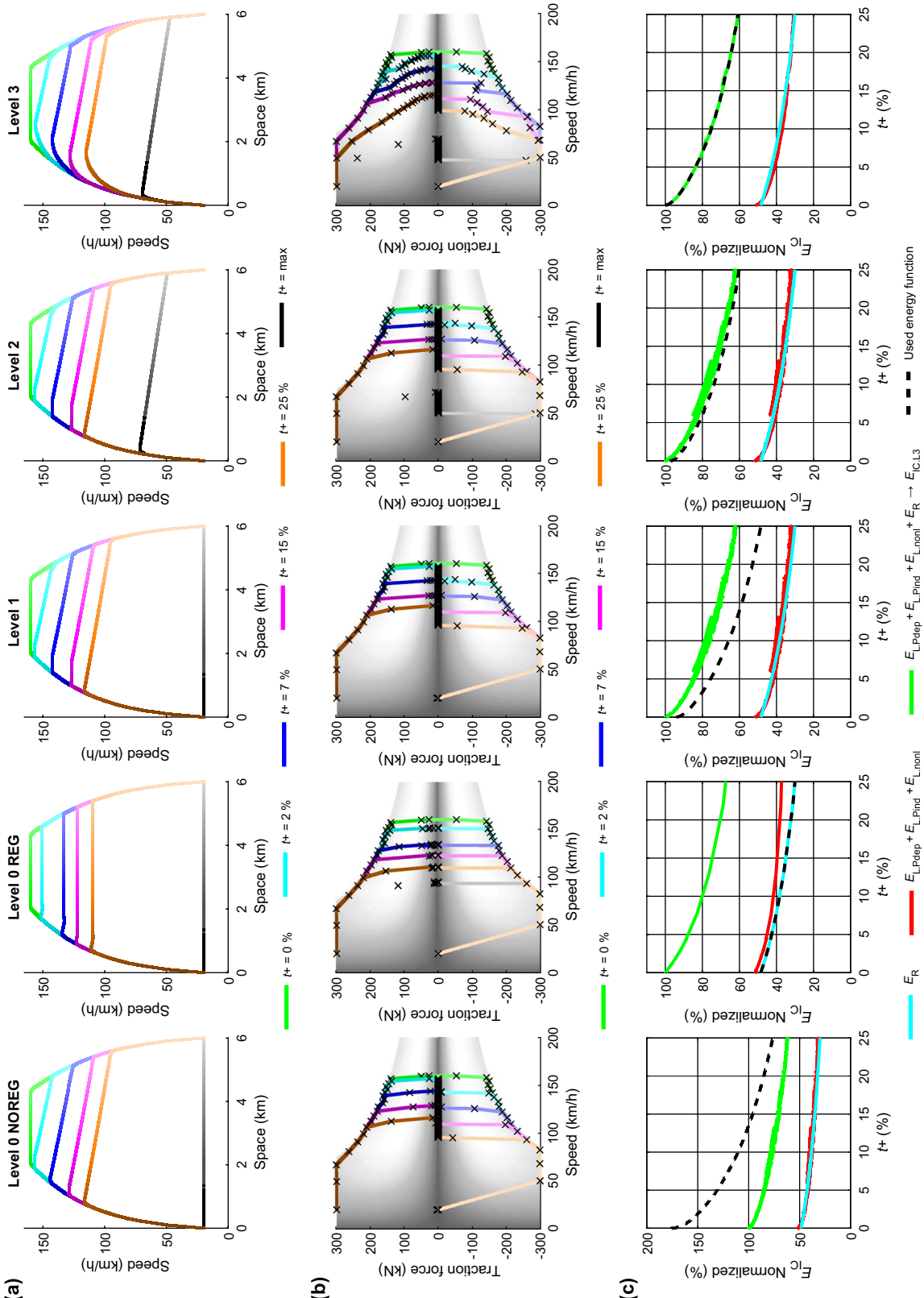

Fig. 9 Complete trajectory with the different levels: **a** trajectories of the different levels for different amounts of travel time buffer $t+$; **b** paths associated to the trajectories, through the efficiency map of the force–speed diagram (the bright areas within the diagram have a higher energy efficiency than the darker ones); **c** relative energy consumption dependent on the used travel time buffer, evaluated with the energy function of Level 3. Note: Level 0 has a very high energy because it includes the entire kinetic energy of the train. For Level 3 the sum is matching with the optimization function

new 3-point IGBT power inverter is a very energy-efficient device, whose design has been optimized for small losses throughout each engine operating point. In contrast to older engines and components of the traction chain, which are primarily optimized for the maximum power, the new engine has also an increased efficiency at middle and low power. Such a specific design causes more nonlinear characteristic, because they are close to the physical component properties. Moreover, the electrical brake has been improved. We exploit in this section the special situation that we have the same vehicle, with differences only in the electrical components.

Figure 10 shows the optimal trajectories over the same test case of Fig. 9 (i.e. 6 km, maximum speed of 160 km/h, for a fixed travel time buffer of 7%), comparing the benchmark of fastest speed, the Level 0 and the two different Level 3, one considering the loss map of the (older) GTO vehicle, one considering the loss map of the (newer) IGBT vehicle. We note that for Level 0, there is no influence of the technology, as Level 0, which corresponds to the standard EETC, does not consider the nonlinear characteristic of these different engines. For Level 3, instead, differences in the trajectories can be observed. IGBT has slightly different switching points and, in general, has *reduced acceleration* and *reduced braking* phase, which are longer in time, and less strong, compared to the Level 3 GTO. This is especially visible during acceleration at high speed (from 1 to 2 km) and deceleration at high speed (around 5 km). This is the result of the different losses and different efficiency. This shows in a different way how an optimal trajectory, considering a Level 3 detail, should result in the two additional driving modes of *reduced acceleration* and *reduced braking*. The electrical components are among the factors determining how much the optimal acceleration and braking should be.

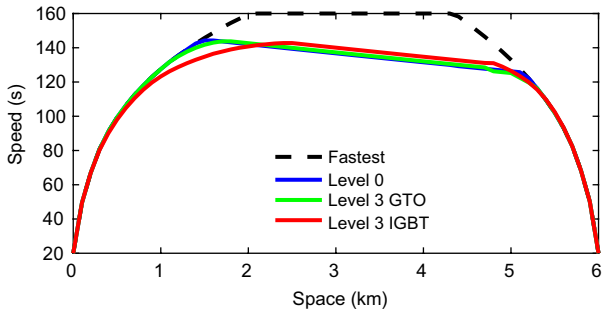

Fig. 10 Trajectory of different trains with Level 3 and the usage of 7% travel time buffer (Note: a visualization of the IGBT loss map is not possible due to confidentiality)

4.7 Quantitative overview

We now report on a quantitative overview of the energy consumption of the optimal trajectories for the different levels. For this, the optimal trajectory by the usage of a travel time buffer of 7% is calculated for each level. The test case is the same 6 km straight line, starting from a standstill and ending at a standstill, with a maximum speed of 160 km/h. The trajectories computed are cross-evaluated by the energy function of each level in Table 2. We here consider the additional benchmark of the reduced maximum speed (RMS), which has not been analysed in detail so far as it has a simple, prescribed trajectory shape (maximum acceleration, cruising, and maximum deceleration). A discussion of this strategy is, for instance, reported by [7]. We use a normalized evaluation by reporting the ratio of energy to the energy of the fastest trajectory, as evaluated by Level 3 GTO. This is reported in Table 2 in percentage values. The missing values are not reported due to confidentiality.

Table 2 Normalized cross evaluation of the energy consumption of optimal trajectories for the different levels, considering the energy functions of the different levels. Travel time buffer is fixed at 7%

Trajectory	Evaluated in Level 0 NOREG	Evaluated in Level 0 REG	Evaluated in Level 1	Evaluated in Level 2	Evaluated in Level 3 with GTO	Evaluated in Level 3 with IGBT
Fastest	176.16	48.51	94.08	97.33	100.00	92.23
RMS	133.81	**40.74**	74.73	81.15	84.07	76.16
Optimal for Level 0 NOREG	**121.97**	40.80	71.28	78.41	80.86	75.35
Optimal for Level 0 REG	134.29	**40.74**	74.92	81.31	84.16	76.34
Optimal for Level 1	121.98	40.80	**71.27**	**78.40**	81.71	74.59
Optimal for Level 2	121.98	40.80	**71.27**	**78.40**	81.71	74.59
Optimal for Level 3 GTO	123.77	40.79	71.80	78.83	**79.99**	–
Optimal for Level 3 IGBT	125.29	40.86	72.06	79.03	–	**73.50**

The bold values are the best performing ones for each column

On the first column of Table 2, it is shown under which level of detail (respectively, with which energy function), the trajectory is optimized. On the top row, it is shown by which level of detail (respectively energy function) the trajectory is evaluated. The diagonal thus shows the best values per column, i.e. the optimization function and the evaluation functions are the same. The values out of the diagonal report how much of an approximation is. It is indeed the case that some trajectories are marginally different and would result anyway in good energy performance, even though they are not optimal.

At first, it can be observed that each column has a very different value range: we start from 120% when evaluated in Level 0 NOREG, to around 40% for Level 0 REG, and 70%, 78% and 80% for, respectively, Level 1, 2 and 3. This underlines how much of the total energy consumption can be when varying the level. Such variations are important for dimensioning of the electrical systems, expected energy consumption and peaks. Going through the first row, one can see that losses can make up more than half (51.5%) of the total energy consumption, comparing Level 0 REG with Level 3. Already an approximated description like Level 1 is able to estimate the consumption with a much smaller error of 6%. Increasing the level of detail reduces this approximation gap further.

This aspect is relevant for dimensioning future electrical railway systems when one takes into account the energy saving potential expected from energy-efficient driving. By using the standard EETC approach (Level 0 NOREG), it can be expected that trajectory optimization can save more than 30% (=1–121.97/176.16, see the Level 0 NOREG). By evaluating the same trajectory with Level 3, this improvement is less than 20% (=1–80.86/100). By applying and evaluating Level 0 NOREG, the railway operating company might think that they need 30% less power plant output, but in reality, it would only be 20% less (for the traction alone).

The variations in the Levels have a different impact on the shape of the trajectory, i.e., what to actually do in terms of speed and acceleration/deceleration actions, when one wants to reduce energy consumption. For this goal, the cross-evaluation is particularly insightful.

The Level 0 NOREG obviously has a much higher value estimated when used as an evaluation function, as it ignores the possibility of regeneration of energy. Also, many different trajectories are possible, which have the same consumption under this level, given the limited descriptive power of losses. Its cross-evaluation shows though that the trajectories evaluated perform reasonably well under all levels. The good results of Level 0 NOREG are mostly due to the extensive use of coasting, which reduces engine power consumption.

The evaluation in Level 0 REG puts all trajectories in a particularly small range of energy consumption. In this sense it has little discriminative power, and the dominance filtering might be misled in the search process for the optimal solutions. This is a problem that the much simplified approaches in the state of the art, do not have, and is a possibly unavoidable drawback of the additional complexity of considering losses. It also shows that the evaluation of Level 0 REG is relatively flat, with many trajectories having an energy consumption very close to each other.

RMS is almost always the worst performing trajectory; this is a result of the very simple approach. Driving slower targets directly the losses due to the driving resistances; however, many other sources of losses exist, as this paper categorized. Those are not reduced and sometimes might even be larger when considering RMS. Its strongest benefit is instead the very easy implementability, which is indeed a fact we will discuss more in detail later on.

For any given evaluation function, there is no difference between the trajectories optimized considering Level 1 and Level 2, by this travel time buffer. In fact, the constant losses, multiplied by a given amount of travel time, have no influence on the shape of the trajectory. They have influence of course on the total energy lost (the columns have variations between each other) and discard low-speed trajectories.

The optimization for Level 3 can improve the energy consumption by a factor of 1% or 2% compared to the second best. This factor, as analysed in the previous sections, depends mostly on the engine operating point in acceleration and deceleration. Those results depend, of course, on the chosen test case, but it can be assumed that the consideration of losses at Level 3 gives the strongest potential for energy saving (compared to lower, less detailed levels) for a trajectory with many accelerations and decelerations, compared to one with very long cruising (for instance the ones discussed in [7]).

The new technology IGBT reaches a saving in total energy of about 6% smaller consumption, comparing the last two columns. This is due to its further optimization in the components, which is much more energy efficient than the old one. In fact, old power inverters have usually been optimized on the engine operating points of the maximum power and traction force. New inverters are optimized on many more engine operating points (e.g. by a specific design of the intermediate circuit voltage). This optimization causes the physical existing nonlinearities to appear more and more.

This also suggests that the requirement for advanced optimization techniques for trajectories is stronger. The IGBT trajectory accelerates less powerful, and decelerates less powerful; in other terms it is connected to acceleration and braking which are more reduced. Due to the component design, the engine operation points in the middle of the force–speed diagram are more efficient.

In absolute terms, the gap between the lower levels and Level 3 is increased further when considering IGBT. Again, here the added value of considering the nonlinear aspects of Level 3 are about 1%–2% better consumption compared to the second best. As a comparison, already driving slower, i.e. exploiting the travel time buffer, reaches a consistent amount of 20%–30% (up to a very optimistic 60%, when considering the rough evaluation by Level 0 REG). Thus, a clear strategy is to start with the implementation of any system able to control the trajectory of trains going slower. The improvement of technology (IGBT vs GTO) comes at a much larger initial investment costs than the improvement given by considering trajectories optimized for Level 3. The improvement for technology allows though to keep a simple task for the driver, while an optimized trajectory for Level 3 might require further effort in driving style. Finally, both improvements are independent and can be combined if required.

5 Results, conclusion and outlook

This paper investigated how to model realistic energy losses in rail vehicles, for the purpose of solving the energy-efficient train control problem (EETC), which determines the trajectories which minimize total energy, given a target travel time. We proposed 4 levels of increasing realism, and evaluated their impact towards the solutions that they would be described as optimal in a travel time/energy multi-objective problem.

The simplest models are common in the literature. Only a handful of papers approach a higher degree of realism. We believe that approximated description, such as the constant plus proportional model of Level 2 can provide a sufficient approximation of the Level 3. This later level is the most complex we consider. Compared to Ref. [10] which smooth those effects into polynomial of degree 2 and 4, we consider the highest resolution possible, which can be expressed as an efficiency and loss map into a force–speed diagram.

Overall, we found that the different levels might have evaluations for a same trajectory differing by as much as 50% of the total consumption. This has strong impacts in the dimensioning of the electrical system for future supply demands, and its peaks. The cross evaluation showed that many approaches, even though they do not correctly model the nonlinear losses, might end up with decently good trajectories, a few percentages away from the best possible.

Considering the more complex levels might break assumptions which allow for simpler resolution used in the optimization of EETC, and requires either approximation, pre-computation, or new algorithmic approaches. In general, the effects of nonlinear losses are evident in avoiding large accelerations, which appear to have low overall efficiency in the traction chain. This contrasts directly with the suggestion to use the maximum acceleration, which is classic in the EETC literature. A few cases showed how a more realistic description of energy conversion results in different trajectories which are energy optimal, by the combined effects of avoiding very low speeds and avoiding very high accelerations.

The different shapes of different trajectories have a direct impact on the possibility of implementing them in real operations. For instance, we compared the RMS approach, which results in lower energy saving. The worst performing approach is the RMS, which might suggest a slow cruising speed, and avoid coasting. Considering the implementation RMS is very simple and a single value can be communicated to the driver. The lower levels (0, 1, 2) must only communicate the driver the switching points between maximum acceleration, cruising, coasting, maximum braking. This is a set of parameters such as cruising speed, timing of the coasting point, and the braking point. We argue that driving these four modes is easy for a driver, as the driver needs to drive at maximum force, or no force, or following a speed controller. Level 3, as far as it specifies a *reduced acceleration* and *reduced braking*, must instead communicate and implement a complete vector of speed and traction force (at a sufficiently high sampling rate, probably in the range of seconds; not in the range of kilometers or minutes). We thus expect that from the implementation point of view, the precise trajectory computed optimizing for losses described at Level 3 is more difficult to track, compared to the lower levels. The paths of *reduced acceleration* and *reduced braking* through the force–speed diagram depend on many factors, including the travel time buffer and slopes. Moreover, the actual travel time buffer depends on the departure delay and can change during the trip. For having

those values available in real-time on a small device, a precomputation of a large amount of trajectories must be performed.

Any change in the driving conditions or the target parameters would result in a change of the trajectory and path in the force–speed diagram. Handling those continuously adjusted, dynamic targets in speed and force can be extremely difficult and particularly tiring, for a human without support (consider already resistances to accept the driver advisory system (DAS) support reported in the literature). A system like DAS or a kind of automatic train operation is necessary, to accurately drive an optimal trajectory specified by Level 3.

From the maintenance and passenger comfort point of view, Level 3 can be advantageous. Compared to the lower levels, Level 3 prescribes reduced acceleration and deceleration and enables a smoother driving. For the passengers, a smoother driving with lower acceleration and lower jerks (change in acceleration) is more comfortable. The same happens to be the case also for the vehicle, where less forces and less strong force changes reduce the wear. There are also advantages for the electric railway power supply. Accelerating and decelerating with the maximum power (at the limit of the power hyperbola) causes a large power load in the electric network, and contributes to strong power peaks. Compared to the lower levels, Level 3 is during the *reduced acceleration* and *reduced braking* not so close to the power hyperbola, therefore resulting in smaller power supply required. Including all those aspects in a single trajectory optimization approach would require multi objective approaches able to include not only energy and travel time in the optimization (as in the Pareto front currently analysed), but also maintenance requirements, passenger comfort and power supply. This would also result in political issues (in quantifying the monetary value of energy, travel time saving, and elusive aspects such as passenger comfort).

Based on the current results, we can identify many potentials for future research, which especially focus on the complex description of losses at Level 2 or 3. We believe the following topics are interesting follow-up works:

- Determining the energy saving potential under further realistic parameters (e.g. different vehicles, slopes, tunnels, varying atmospheric conditions, etc).
- Study of human factors to support acceptability by human drivers in case they need to follow complex trajectories in the force–speed diagram.
- Integration of optimization of train traffic (which determines the separation of trains, and thus the available travel time buffer) with train control (determining the trajectories and thus the travel time).
- Faster computation of solutions by more advanced algorithms, especially on devices with limited computation performance, and interconnection with real system. This might resort on data-driven approaches, surrogate models, or pre-computation.
- Determination of (sufficiently) correct values of even more parameters, and robustness to incorrect parameters also based on online calibration and data-driven approaches. Such approaches can be standard calibration approaches or also self-learning approaches, which are using vehicles energy measurements or vehicle diagnose data.
- Usage of the higher accurate description levels for dimensioning of the power grid and the railway power consumption; and determining especially the peak reduction effect due to *reduced acceleration* and *reduced braking*.

Appendix

See Figs. 11, 12, 13 and 14.

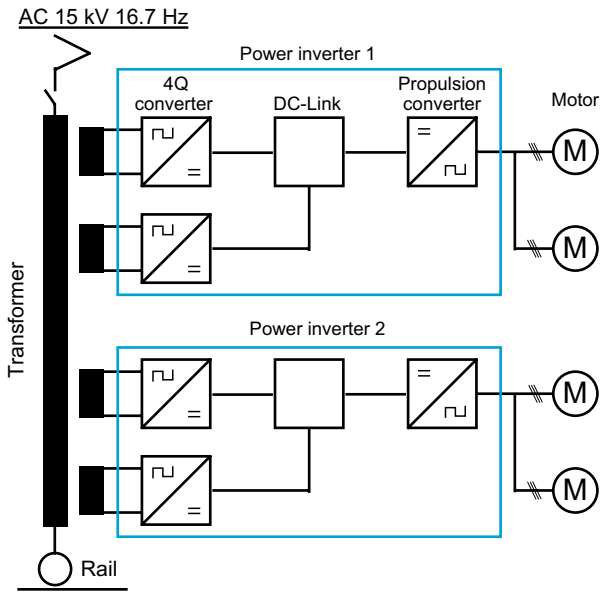

Fig. 11 Simplified visualization of the Re 460 traction chain for two of the four motors based on Refs. [4, 28, 59]. Note: In this simplification grade, the GTO and IGBT power inverters are comparable

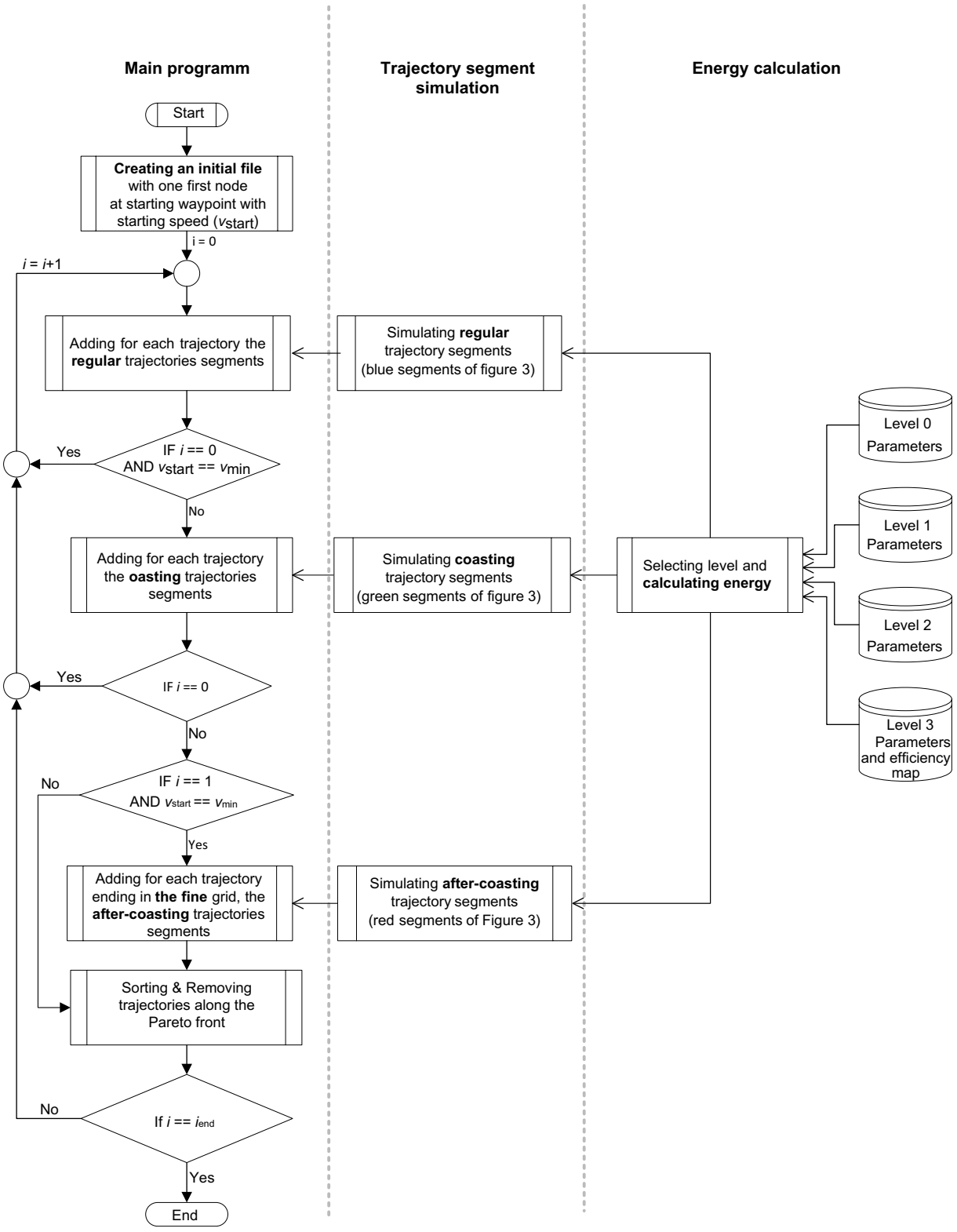

Fig. 12 Flowchart of the general program structure (*i* represents the number of the loop pass)

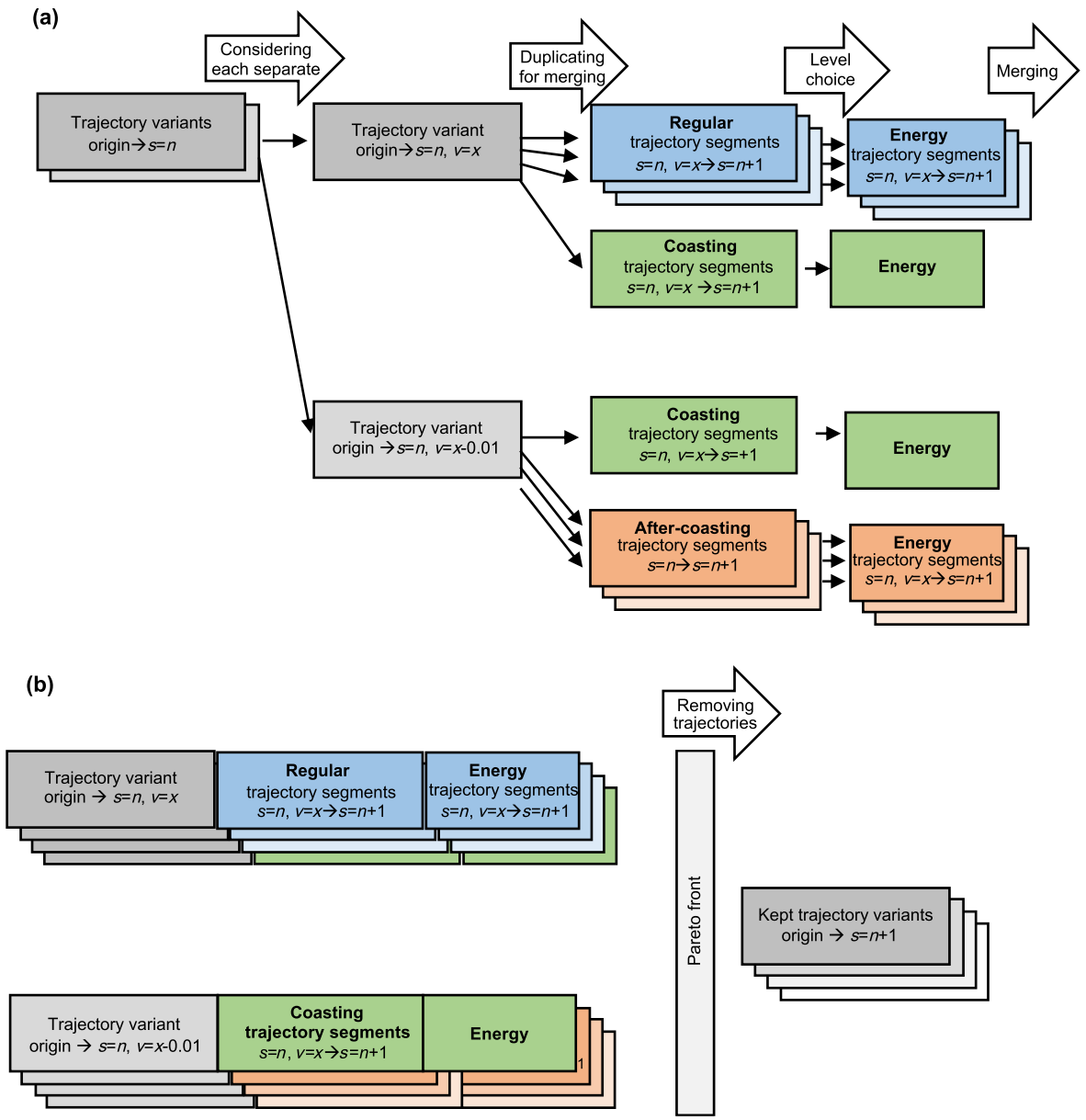

Fig. 13 Simplified visualization of handling a trajectory: **a** adding new segments; **b** reducing variants by dominance, to pareto front. In an iterative scheme, given the Pareto optimal variants from origin to s, we compute and store the trajectory variants until $s+1$, and repeat until destination (n is the number one variant s; x is the value of a speed v)

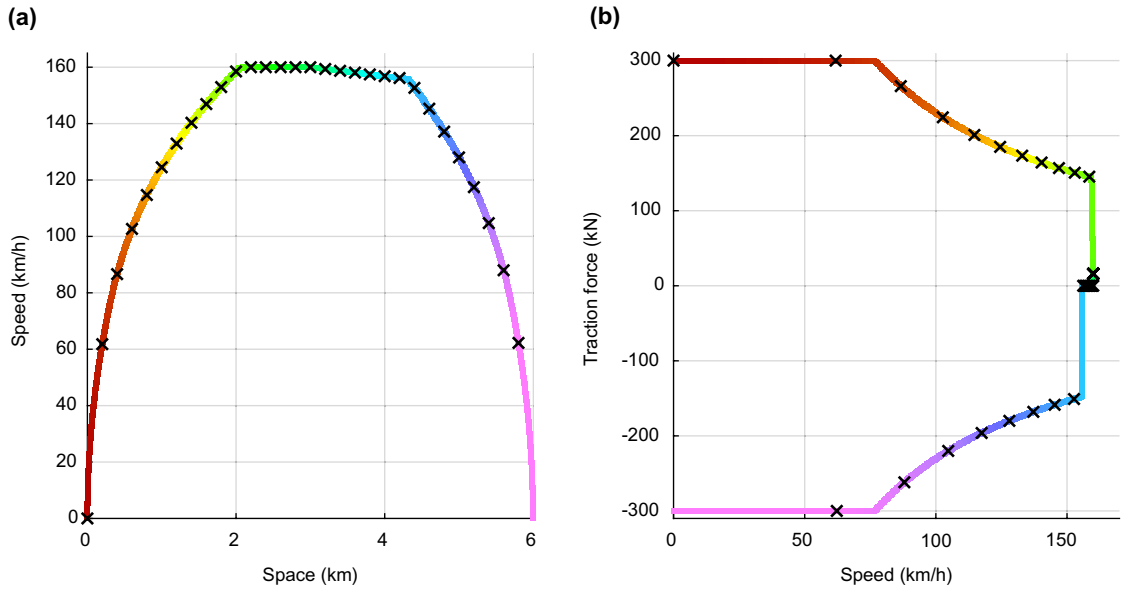

Fig. 14 Explanation of the paths through the force–speed diagram: **a** speed–space diagram of a train run; **b** path through the force–speed diagram. In the diagrams, the crosses are uniformly spaced over the entire distance traveled. In **a**, the train accelerated between 0 and 2 km, cruises between 2 and 3 km, coasts between 3 and 4.2 km, and brakes between 4.2 and 6 km; the line color is changing from dark red to yellow, green, blue and pink with the travel space. In **b**, the line color changes similarly matching to **a** from dark red to yellow, green, blue and pink with the travel space

Acknowledgements This work was supported by Swiss Federal Office of Transport, the ETH foundation and via the grant RAILPOWER.

References

1. BAV (2020) Umsetzung der Energiestrategie 2050 im öffentlichen Verkehr (ESöV 2050). Eidgenössisches Departement für Umwelt, Verkehr, Energie und Kommunikation UVEK - Bundesamt für Verkehr BAV, Bern
2. Garbe R (1924) Die zeitgemäße Heißdampflokomotive. Springer, Berlin Heidelberg, Berlin, Heidelberg
3. Nold M (2018) Energieeinsparung durch Optimierung der Ventilationssteuerung. Bundesamt für Verkehr (BAV), Bern
4. Huggenberger T, Baden (2016) Modernisierung der Lokomotiven Re460 mit IGBT-Stromrichtern. Elektrische Bahnen 485–489
5. Huggenberger T (2017) IGBT-Stromrichter verrlängern die Lebensdauer der Lokomotiven Re460. ABB Review, pp 65–69
6. Howlett PG, Pudney PJ (1995) Energy-efficient train control. Springer, London
7. Scheepmaker GM, Goverde RMP (2020) Energy-efficient train control using nonlinear bounded regenerative braking. Transport Res C: Emerg Technol 121:102852.
8. Franke R, Terwiesch P, Meyer M (2002) An algorithm for the optimal control of the driving of trains. In: Proceedings of the 39th IEEE conference on decision and control (Cat. No.00CH37187), December 12–15, 2000, Sydney, NSW, Australia. IEEE. New York, pp 2123–2128
9. Franke R, Meyer M, Terwiesch P (2002) Optimal control of the driving of trains (optimale steuerung der fahrweise von zügen). Automatisierungstechnik 50(12):606
10. Ghaviha N, Bohlin M, Holmberg C et al (2017) A driver advisory system with dynamic losses for passenger electric multiple units. Transp Res Part C Emerg Technol 85:111–130
11. Müller W (1953) Eisenbahnanlagen und Fahrdynamik. Springer Verlag, Berlin, Heidelberg
12. Ichikawa K (1968) Application of optimization theory for bounded state variable problems to the operation of train. Bull JSME 11(47):857–865
13. Scheepmaker GM, Goverde RMP, Kroon LG (2017) Review of energy-efficient train control and timetabling. Eur J Oper Res 257(2):355–376
14. Bomhauer-Beins A (2019) Energy saving potentials in railway operations under systemic perspectives. Dissertations, ETH Zurich
15. Phil H (1990) An optimal strategy for the control of a train. J Aust Math Soc Series B, Appl Math 31(4):454–471
16. Rodrigo E, Tapia S, Mera JM et al (2013) Optimizing electric rail energy consumption using the Lagrange multiplier technique. J Transp Eng 139(3):321–329
17. Albrecht A, Howlett P, Pudney P et al (2016) The key principles of optimal train control—part 1:formulation of the model,

strategies of optimal type, evolutionary lines, location of optimal switching points. Transp Res Part B Methodol 94:482–508
18. Albrecht A, Howlett P, Pudney P et al (2016) The key principles of optimal train control—part 2:existence of an optimal strategy, the local energy minimization principle, uniqueness, computational techniques. Transp Res Part B Methodol 94:509–538
19. Howlett PG, Milroy IP, Pudney PJ (1994) Energy-efficient train control. Contr Eng Pract 2(2):193–200
20. Wang P, Trivella A, Goverde RMP et al (2020) Train trajectory optimization for improved on-time arrival under parametric uncertainty. Transp Res Part C Emerg Technol 119:102680
21. Trivella A, Wang P, Corman F (2021) The impact of wind on energy-efficient train control. EURO J Transp Logist 10:100013
22. Kummer W (1919) Die Schaltung der Maschinenfabrik Oerlikon zur Energierückgewinnung auf Einphasenbahnen. Schweizerische Bauzeitung 73/74. https://doi.org/10.5169/SEALS-35562
23. SBZ (1920) 1C+C1 Güterzug-Lokomotiven für die Gotthardlinie der S.B.B. https://doi.org/10.5169/SEALS-36464
24. Nold M (2018) Rhaetian Railway's Gmf 4/4 II infrastructure diesel locomotives. Railway Update 2018:118–122
25. Nold M (2019) 60 Jahre Lokomotiven Ge 6/6 II der Rhätischen Bahn (Teil 1). Schweizer Eisenbahn-Revue 2019:106–109
26. Nold M (2019) 60 Jahre Lokomotiven Ge 6/6 II der Rhätischen Bahn (Teil 2). Schweizer Eisenbahn-Revue 2019:164–166
27. Steimel A (2006) Elektrische Triebfahrzeuge und ihre Energieversorgung: Grundlagen und Praxis. Oldenbourg-Industrieverl.
28. Gerber R, Drabek E, Müller R (1991) Die Lokomotiven 2000 – Serie Re 460 – der Schweizerischen Bundesbahn. Schweizer Eisenbahn Revue 321–365
29. Nold M, Corman F (2020) eco 4.0. Vorstudie zur: Traktionsbasierten energieorientierten Echtzeitfahrplanoptimierung. Bundesamt für Verkehr (BAV), Bern
30. Nold M, Corman F (2022) Modelling realistic energy losses from variable efficiency and vehicle systems, in determining energy efficient train control
31. Ihme J (2019) Schienenfahrzeugtechnik, 2, überarbeitete und, erweiterte. Springer Vieweg, Wiesbaden
32. Meyer M, Roth M, Schaller B (2000) Einfluss der Fahrweise und der Betriebssituation auf den Energieverbrauch von Reisezügen. Schweizer Eisenbahn-Revue 360–365
33. Meyer M, Heck A, Walch G et al (2016) Reduktion des Traktionsenergiebedarfs der Allegra-Triebwagen der RhB. Schweizer Eisenbahn-Revue 69–71
34. Zarifyan A, Zarifyan A, Grebennikov N et al (2019) Increasing the energy efficiency of rail vehicles equipped with a multi-motor electrical traction drive. In: 2019 26th International Workshop on Electric Drives: Improvement in Efficiency of Electric Drives (IWED), January 30–February 2, 2019. IEEE, Moscow, pp 1–5
35. Xiao Z, Bai B, Chen N, et al (2019) Energy-efficient control for metros with dynamic losses of traction power system. In: 2019 IEEE Intelligent Transportation Systems Conference (ITSC), October 27–30, 2019, Auckland, New Zealand. IEEE, pp 357–362
36. Graffagnino T, Schäfer R, Tuchschmid M et al (2019) Energy savings with enhanced static timetable information for train drivers. In: 8th International Conference on Railway Operations Modelling and Analysis (ICROMA), Norrköping, Sweden, 17– 20 June 2019
37. Nold M, Thomas H, Corman F (2022) Der Einfluss der Verluste in den Traktionskomponenten auf den Energieverbrauch von Zugfahrten. Schweizer Eisenbahn Revue 84–89
38. Frech W, Nold M (2019) Die HG 4/4 704, ein attraktiver Lokomotiv-Zuwachs auf der Furka-Bergstrecke. Schweizer Eisenbahn-Revue 2019:332–335
39. Filipović Ž (2015) Elektrische Bahnen: Grundlagen, Triebfahrzeuge, Stromversorgung, 5. überarb. Aufl. Springer-Vieweg, Berlin, Heidelberg
40. Peters J-L (1990) Bestimmung des aerodynamischen Widerstandes des ICE/V im Tunnel und auf freier Strecke durch Auslaufversuche. Eisenbahntechnische Rundschau 559–564
41. Wende D (2003) Fahrdynamik des Schienenverkehrs: mit 83 Tabellen und 83 Berechnungsbeispielen, 1st edn. Teubner, Stuttgart Leipzig Wiesbaden
42. Sachs K (1973) Elektrische Triebfahrzeuge: ein Handbuch für die Praxis sowie für Studierende in drei Bänden. Springer, Wien
43. Meyer M, Menth S, Lerjen M (2007) Potentialermittlung Energieeffizienz Traktion bei den SBB. Swiss Federal Office of Transport, Bern
44. Spring E (2009) Elektrische Maschinen: Eine Einführung. Springer, Berlin Heidelberg, Berlin, Heidelberg
45. Veigel M (2018) Ein neues Modell zur Berechnung der fertigungsabhängigen Ummagnetisierungsverluste in Synchronmaschinen. Dissertation, Elektrotechnisches Institut (ETI). https://doi.org/10.5445/IR/1000084543
46. Bauer D (2019) Verlustanalyse bei elektrischen Maschinen für Elektro- und Hybridfahrzeuge zur Weiterverarbeitung in thermischen Netzwerkmodellen. Springer Fachmedien Wiesbaden, Wiesbaden
47. Blesl M, Kessler A (2013) Energieeffizienz in der Industrie. Springer Vieweg Berlin, Heidelberg
48. Mahmoudi A, Soong WL, Pellegrino G, et al (2015) Efficiency maps of electrical machines. In: 2015 IEEE Energy Conversion Congress and Exposition (ECCE), 24–25 September 2015, Montreal, QC, Canada. IEEE, pp 2791–2799
49. Bojoi R, Armando E, Pastorelli M, et al (2016) Efficiency and loss mapping of AC motors using advanced testing tools. In: 2016 XXII International Conference on Electrical Machines (ICEM), 4–7 September 2016, Lausanne, Switzerland. IEEE, New York, pp 1043–1049
50. Siemens (2016) Transformatorauswahl in Abhängigkeit von Belastungsprofilen. Siemens AG, Erlangen
51. Binder A (2017) Elektrische Maschinen und Antriebe: Grundlagen, Betriebsverhalten. Springer, Berlin, Heidelberg
52. Moll T, Hellenbroich G, Herold K et al (2021) Co-simulation von wirkungsgrad, beölung und Wärmehaushalt für getriebe mit integrierter elektromaschine. ATZ Automobiltech Z 123(4):60–64
53. Eberhard F (1996) SBB Gasturbinenlokomotive: Am 4/6 1101. Fachpresse Goldach, Goldach
54. Golloch R (2005) Downsizing bei Verbrennungsmotoren. Springer, Berlin
55. Haake B (2015) Kennfelder. In: Van Basshuysen R, Schäfer F (eds) Handbuch Verbrennungsmotor. Springer Fachmedien Wiesbaden, Wiesbaden, pp 26–33
56. Borchert U (2017) Gasturbine mit verstellbaren Leiteinrichtungen zur Verringerung des Brennstoffverbrauches im untere Teillastbereich und im Leerlauf. FAU University Press, Erlangen
57. De Martinis V, Corman F (2018) Data-driven perspectives for energy efficient operations in railway systems: current practices and future opportunities. Transp Res Part C Emerg Technol 95:679–697
58. Tipler PA, Mosca G (2015) Physik: für Wissenschaftler und Ingenieure, 7, dt. Springer Spektrum, Berlin
59. ABB Switzerland (2018) Compact Converter Borderline. ABB Switzerland, Turgi