An algorithm for the optimal control of the driving of trains

Rüdiger Franke
ABB Corporate Research
Ruediger.Franke@de.abb.com

Peter Terwiesch
ABB Corporate Research
Peter.Terwiesch@ch.abb.com

Markus Meyer
ADtranz (Switzerland)
Markus.Meyer@ch.adtranz.com

Abstract

We discuss an algorithm that optimizes the driving style of a train. The objective is to minimize the electrical energy used for traction subject to constraints such as travel time, speed limits, available traction power, etc. The optimization is based on a nonlinear point-mass model of the train, which includes the equations of motion and which considers the setpoint-dependent efficiency of the propulsion system. Although nonlinear, the equations of motion are formulated in a way that allows their piecewise analytical solution, thus greatly increasing computational efficiency. A discrete dynamic programming algorithm is developed for the deterministic and efficient numerical solution of the nonlinear optimal control problem. Both simulation results and practical measurements indicate energy savings between 10 and 30%, depending on operating conditions. The resulting optimal driving style is qualitatively different from previous solutions obtained with more simplified train models as reported in the literature. The algorithm forms a suitable basis for a nonlinear model-predictive controller operating in hard real time.

1 Introduction

The optimal control of trains has been an active research topic for many years. The general aim is to drive a train in a way that minimizes overall consumption of electrical energy, subject to constraints on time and to physical limitations imposed by the train and the operating environment. In an increasingly competitive transportation market, interest in energy efficiency among railway operators has been a subject of increased interest in recent years, both for retrofit of existing vehicles and the aquisition of new ones.

Literature on energy-efficient operation of trains goes back to the late 1960s, when the optimal control problem for a simplified linear train model was solved analytically by applying Pontryagin's Maximum principle, see e.g. [2]. The resulting optimal driving style basically consists of four sections: maximum acceleration, cruising at constant speed, coasting (zero traction force), and maximum deceleration. Several methods have been proposed to obtain the switching time points and/or the cruise speed, which are the remaining decision variables, see e.g. [6]. Reported practical implementations are often limited to obtaining the time point when coasting should start, e.g. [7], or to additionally calculate the optimal cruise speed, e.g. [8]. However, the analytical solution to the simplified linear problem has a number of limitations. Some authors propose empirical improvements, e.g. to replace the section of constant cruise speed with a coast remotor cycle [8], or to exploit the height profile of the track [6]. Moreover, the number of decision variables increases if the maximal permitted track speed varies during a ride.

The present paper describes a new algorithm for the optimal train control problem. We consider a nonlinear train model that is a closer approximation of reality. This allows us to overcome the limitations of the more simple linear solution, accounting for effects such as air resistance, which cannot be neglected at higher speeds. Most importantly, the new approach takes into account the setpoint-dependent efficiency of the propulsion system, resulting in a solution that is both qualitatively and quantitatively different from the solution to the simplified problem. Using kinetic energy instead of velocity as a dynamic state variable, an analytical solution for the nonlinear equations of motion can be found. Exploiting the numerical efficiency of this analytical solution instead of numerical integration of the nonlinear model, we implement a discrete dynamic programming algorithm that is able to solve the optimal control problem in real time.

2 Train model

The train is commonly modeled as mass point in the related literature, see e.g. [6]. Note that this is not a limitation compared to a distributed mass model, as the distribution of masses over the length of the train can be taken into account by pre-processing the altitude profile of the track before it is used with the point-mass model. Starting from this modeling assumption, we decide to use kinetic energy per mass unit $e_{\text{kin}} = 0.5v^2$ and time t as states, and position s of the train as the independent variable. The use of position s instead of time

t as independent variable simplifies the incorporation of track-related data, such as track slope, speed limits, and tunnel resistance. Depending on the exact local conditions, information on s is obtained in real time by a combination of global positioning signals, track-side equipment, gyro, and model integration, making it known to an accuracy of about 10 meters. Digital track maps with information on gradient, altitude, speed limits, tunnels, etc are available in a number of countries with a resolution of 0.1 to 1 meter. Note that the choice of e_{kin} instead of speed v does not eliminate all model nonlinearities. However, not only does the influence of air resistance on the states become linear, but also the nonlinearity is such that a piecewise analytical solution of the model differential equations becomes possible, thus permitting a significant increase in the speed of numerical solutions. Our equations of motion have the form

$$\begin{aligned} \frac{de_{\text{kin}}}{ds} &= \frac{1}{m\rho}\left[F_{\text{tract}} - mg\sin\alpha - a - Ce_{\text{kin}}\right], \quad (1) \\ & \quad e_{\text{kin}}(0) = e_{\text{kin}0}, \\ \frac{dt}{ds} &= \sqrt{\frac{1}{2e_{\text{kin}}}}, \quad t(0) = t_0. \quad (2) \end{aligned}$$

Here m is the mass of the train and ρ is a factor to consider rotating masses. F_{tract} is the applied traction force, g the gravitation constant, α the slope of the track, and a is a parameter describing constant parts of the running resistance of the train. The term C contains parts of the running resistance of the train that depend quadratically on velocity, especially air resistance. The main advantage of this formulation of the equations of motion is that they can be solved analytically if everything but e_{kin}, t, and s is constant. The solution for e_{kin} is

$$e_{\text{kin}}(s) = \frac{A}{C} + \left[e_{\text{kin}}(0) - \frac{A}{C}\right]\exp(-Cs) \quad (3)$$

with $A = F_{\text{tract}} - mg\sin\alpha - a$ the constant disturbance term of the differential equation (1). As kinetic energy must be nonnegative, the condition

$$A \geq -\frac{Ce_{\text{kin}}(0)}{\exp(Cs) - 1} \quad (4)$$

follows from (3).

Now we can solve (5) for t. We have to distinguish four cases for obtaining $t(s)$ analytically:

$$t(s) = t(0) + \begin{cases} \Delta t_{\text{free}}(s), & A = 0 \\ \Delta t_{\text{econ}}(s), & A > 0, \frac{C}{A}\,e_{\text{kin}}(0) = 1 \\ \Delta t_{\text{apos}}(s), & A > 0, \frac{C}{A}\,e_{\text{kin}}(0) \neq 1 \\ \Delta t_{\text{aneg}}(s), & -\frac{Ce_{\text{kin}}(0)}{\exp(Cs)-1} < A < 0. \end{cases} \quad (5)$$

In the special cases of free motion ($A = 0$) and constant speed, the solutions are

$$\Delta t_{\text{free}}(s) = \frac{s}{\sqrt{2e_{\text{kin}}(0)}} \quad (6)$$

and

$$\Delta t_{\text{econ}}(s) = \frac{2}{C\sqrt{2e_{\text{kin}}(0)}}\left[\exp\left(\frac{C}{2}s\right) - 1\right], \quad (7)$$

respecively. Besides this, we have generally for positive A

$$\Delta t_{\text{apos}}(s) = \frac{2}{\sqrt{2AC}}\,\text{artanh}\left(\frac{\sqrt{q_1} - \sqrt{q_2}}{1 - \sqrt{q_1 q_2}}\right) \quad (8)$$

with

$$\begin{aligned} q_1 &= 1 + \left[\frac{C}{A}\,e_{\text{kin}}(0) - 1\right]\exp(-Cs), \\ q_2 &= \frac{C}{A}\,e_{\text{kin}}(0) \end{aligned}$$

and for negative A

$$\Delta t_{\text{aneg}}(s) = \frac{2}{\sqrt{-2AC}}\,\arctan\left(\frac{\sqrt{-q_2} - \sqrt{-q_1}}{1 + \sqrt{q_1 q_2}}\right). \quad (9)$$

Even though the analytical solution is not trivial for our nonlinear train model, it allows an extremely fast and reliable calculation of possible motions of the train, especially when compared to numerical integration of the nonlinear model, which would need to be used otherwise in order to account for the nonlinear effects described above.

The traction force F_{tract} is our control variable. We distinguish two parts: an electrical part F_{el} for acceleration and recuperator brake, and a mechanical part F_{mech} for the pneumatic brake. The sum of both parts gives the actual traction force ($F_{\text{tract}} = F_{\text{el}} + F_{\text{mech}}$).

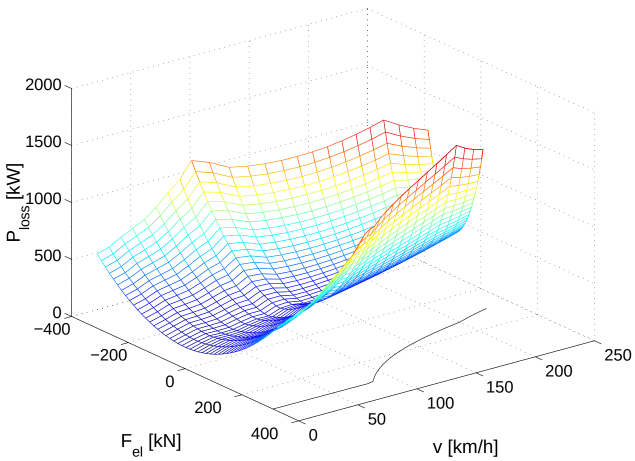

Figure 1: Loss power of a converter locomotive of type Re 460 as function of traction force (electrical part) and velocity. Furthermore the maximum traction force is shown in the F-v plane as function of velocity.

In order to be able to minimize the electrical energy taken from the mains, we need to know the efficiency of the propulsion system. This efficiency is often considered constant in the literature. However, it actually varies greatly with operating conditions [3]. This is why we consider the power losses of the propulsion system P_{loss} as a function of F_{el} and velocity v, introducing a new nonlinear term into our optimization problem. The steady-state loss power function is derived approxiamately from a detailed propulsion model. Figure 1 shows the loss power of a typical converter locomotive.

The objective for a ride between positions s_0 and s_f is minimization of the use of electrical energy E_{el} by means of selection of F_{tract}:

$$E_{\text{el}} = \int_{s_0}^{s_f} F_{\text{el}} v + P_{\text{loss}}(F_{\text{el}}, v)\, ds \rightarrow \min_{F_{\text{tract}}} \qquad (10)$$

Constraints arise, besides from the equations of motion (1) and (2), from the time table, from the permitted track speed, as well as from comfort bounds and physical limits. Figure 1 shows the maximum traction force. Mainly due to limitations of the power of the propulsion system, the maximum traction force is a function of velocity. Furthermore the maximum and minimum traction forces are limited by comfort bounds on the maximum acceleration and deceleration. Thus the bounds on the traction force are also functions of the slope α of the track.

3 Optimization algorithm

The optimization algorithm is the key part of a Nonlinear Model Predictive Controller (NMPC), which can be used to optimize the driving style of the train. We implemented a Discrete Dynamic Programming (DDP) algorithm to solve the optimal control problem numerically.

First the control horizon $[s_0, s_f]$ is split into K stages $k = 0, 1, \ldots, K-1$, so that the track and train parameters in equation (1) as well as the sought traction force can be considered constant in each stage. Stage k covers the control interval $[s^k, s^{k+1})$ with the length $\Delta s^k = s^{k+1} - s^k$. The multistage optimization problem is

$$\sum_k f_0(e^k_{\text{kin}}, A^k, C^k, \Delta s^k) \rightarrow \min_{F^k_{\text{tract}}} \qquad (11)$$

subject to the motion of the train

$$e^{k+1}_{\text{kin}} = f_1(e^k_{\text{kin}}, A^k, C^k, \Delta s^k), \qquad (12)$$

$$t^{k+1} = t^k + f_2(e^k_{\text{kin}}, A^k, C^k, \Delta s^k), \qquad (13)$$

limitations on the traction force

$$F_{\text{min}}(\alpha^k) < F^k_{\text{tract}} < F_{\text{max}}(\alpha^k, e^k_{\text{kin}}, e^{k+1}_{\text{kin}}), \qquad (14)$$

speed limits

$$e^k_{\text{min}} < e^k_{\text{kin}} < e^k_{\text{max}}, \qquad (15)$$

bounds on passage times

$$t^k_{\text{min}} < t^k < t^k_{\text{max}}, \qquad (16)$$

as well as the given initial state $(t_0, e_{\text{kin}0})$ and the scheduled final state $(t_f, e_{\text{kin}f})$

$$(t^0, e^0_{\text{kin}}) = (t_0, e_{\text{kin}0}), \qquad (17)$$

$$(t^K, e^K_{\text{kin}}) = (t_f, e_{\text{kin}f}). \qquad (18)$$

The functions f_1 and f_2 are defined with (3) and (5), respectively. The function f_0 is obtained using a numerical approximation of (10) over one control interval.

This discrete optimal control problem could be solved numerically with a nonlinear programming algorithm, e.g. [1], provided that all constraints and the objective are formulated sufficiently smooth for the solution method. But the result of one such calculation is a control trajectory. Special care has to be taken to account for model errors and disturbances, which result in prediction errors in an on-line application. That is why the optimal control problem is normally solved repeatedly when using nonlinear programming in NMPC. In this way the prediction is updated during the operation. It is required that the dynamics of the controlled process is slow, compared to the time needed to solve one optimal control problem.

In our case the process dynamics are rather fast, i.e. the NMPC should be able to update a prediction within seconds. The main advantage of dynamic programming is that the result of one calculation is a control function, which can be used in a regulatory feedback control. With DDP the state space is discretized and one optimal control value is calculated for each discrete cell, i.e. for each possible state (see e.g. [5] for more details on DDP). As the computational complexity depends exponentially on the dimension of the state space, the applicability of DDP is limited to medium-dimension state and decision spaces. In our case the state space has two dimensions (kinetic energy and time) and the calculation of one state transition is cheap with the analytical solution of our equations of motion. Furthermore we can exploit that f_0, f_1 and f_2 in (11), (12) and (13) do not depend on time. This means that we need to evaluate these functions only once for each possible transition in kinetic energy over a stage. The once obtained values can be reused in the discrete dynamic programming algorithm at all possible initial times of the stage.

A further important advantage of the algorithm is that the time needed to solve one optimal control problem is known in advance. This is because the time for calculating an individual state transition is deterministic, and because the total number of calculated state transitions is given by the number of discrete cells. The deterministic solution time simplifies a real-time application.

The result of the DDP algorithm is not only a control trajectory, but optimal control values are stored for the whole range of possible velocities (kinetic energies) and passage times at discrete positions. This allows to react rapidly on disturbances like time delays or temporary speed limits.

4 Example

We demonstrate the DDP implementation on a realistic example. We compare the energy savings obtained with our numerical solution of the optimal control problem to a coasting strategy. The coasting strategy is a simple application of the analytical solution for the linear train model.

Figure 3 (upper diagram) shows results of a train performance calculation for a passenger train between Zürich main station and Lenzburg, the first stop towards Genève. The train consists of 9 waggons and one converter locomotive of type Re 460. The timetables of most railways include time reserves, which are necessary to handle delays under abnormal traffic conditions. In a normal situation these reserves can be used for an economic driving style. As the train in the example would arrive more than 2 minutes early with the fastest possible driving style, a coasting phase is applied before the final stop. In table 1 it can be seen that in this way about 14 % of the electrical energy can be saved. The main part of this energy saving results from the reduced use of the mechanical brakes, an other part from reduced speed (less air drag) during the final phase of the run.

The lower diagram of figure 3 shows plots for the same example when applying our optimization algorithm. As a result of the consideration of the efficiency of the propulsion system, no maximum acceleration or maximum decelaration is applied at high velocities. This is qualitatively different from the analytical solution for a linear train model. Furthermore the speed is reduced in the tunnel between kilometer 17 and 23, where the air resistance is higher. The traction force and the power trajectories are considerably smoother. Thus with optimal driving about 25 % of the net electrical energy consumed at the pantograph can be saved, compared to fastest driving. This is significantly higher than with the coasting strategy (see table 1).

Table 1: Examplary travel time t and use of electrical energy E_{el} for different driving styles of a passenger train between Zürich and Lenzburg.

Driving style	fastest	coasting	optimal
t [m:s]	16:41	19:00	19:00
E_{el} [kWh]	469	404	352

5 Practical Tests

First measurements during the real operation of a passenger train were performed in winter 1999/2000 (see [4]). The train consisted of a locomotive of type Re 460 and 10 waggons and was operated on the line Zurich – Luzern of Swiss Federal Railways (SBB). In a first phase, the energy consumption and its statistical distribution was measured for a number of trains in scheduled service (between 50 and 100 data sets per section of the line). In a second step, two controlled runs were carried out, where the train driver operated the locomo-

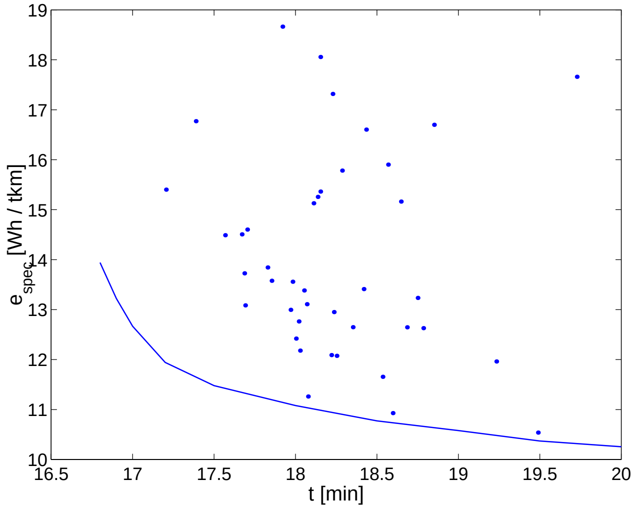

Figure 2: Distribution of energy usage under normal service, compared to optimization results. The use of energy is normalized with transported mass. It is shown as function of travel time. The results were obtained for rides from Luzern to Zug. The section is about 28 km long.

tive exactly according to the precalculated optimization results.

Figure 2 shows a comparison of results measured for normal service and calculated with the optimization. It can be seen that a few very skilled drivers reach the lowest possible energy consumption. This optimum was also obtained by operating the train according to the optimization results. Calculation and measurement results for this case were equal within the given accuracy of the sensors (3 % range).

The tests proved the correctness and practicability of the optimization algorithm for the given application. Compared to today's mean energy consumption values, savings between 10 and 30 % are possible with our algorithm.

6 Conclusions

We discuss a new algorithm for the minimization of the traction energy used by a train. Compared to approaches known from the literature, we have obtained the following improvements.

A different formulation allows us to solve the nonlinear equtions of motion analytically for sections of constant track parameters. Exploiting this, a Discrete Dynamic Programming (DDP) algorithm is implemented, which is suitable for a real-time implementation of a Nonlinear Model Predictive Controller (NMPC).

Furthermore we incorporate the set point dependent efficiency of the propulsion system, in order to minimize the electrical energy taken from the mains. Compared to the common approach, when the efficiency is assumed constant, i.e. basically the mechanical energy on the track is minimized, our algorithm gives a qualitatively new solution for the optimal driving style of a modern converter locomotive.

Simulation results and first practical investigations show that the traction energy savings obtained with our algorithm are between 10 and 30%, compared to trivial optimization strategies and to the current mean value for manual driving. First practical tests showed the correctness and applicability of our results.

Five patents were filed for the train model and its use in the optimization of the driving style.

References

[1] Rüdiger Franke and Eckhard Arnold. Applying new numerical algorithms to the solution of discrete-time optimal control problems. In K. Warwick and M. Kárný, editors, *Computer-Intensive Methods in Control and Signal Processing: The Curse of Dimensionality*, pages 105–118. Birkhäuser, Basel, 1997.

[2] Peter Horn. Über die Anwendung des Maximum-Prinzips von Pontrjagin zur Ermittlung von Algorithmen für eine energieoptimale Zugsteuerung. *Wissenschaftliche Zeitschrift der Hochschule für Verkehrswesen "Friedrich List" in Dresden*, 18(4), 1971.

[3] Markus Meyer and Martin Aeberhard. Vom Gratisstrom zur Energiesparlokomotive – Energieverbrauch bei elektrischen Bahnen. *Eisenbahn-Revue*, pages 28–39, January 1997.

[4] Markus Meyer, Michael Roth, and Beat Schaller. Einfluss der Fahrweise und der Betriebssituation auf den Energieverbrauch von Reisezügen. *Eisenbahn-Revue*, pages 360–365, August 2000.

[5] M. Papageorgiou. *Optimierung*. Oldenburg Verlag, München, 2 edition, 1996.

[6] Eckhard Schüler-Hainisch. *Verfahren zur Optimierung des Energieverbrauchs von Zugfahrten*. Number 157 in Forschritt-Berichte VDI, Reihe 12: Verkehrstechnik/Fahrzeugtechnik. VDI Verlag, Düsseldorf, 1991.

[7] Gerhard Voß and Dirk Sanftleben. Rechnergestützte energiesparende Fahrweise im Regelbetrieb. *ETR – Eisenbahntechnische Rundschau*, 47(1):25–31, 1998.

[8] Shinobu Yasukawa, Shinichiro Fujita, Takehisa Hasebe, and Koichi Sato. Development of an on-board energy-saving train operation system for the Shinkansen electric railcars. *Quarterly Reports of RTRI, Tokyo*, 28(2–4):54–62, 1987.

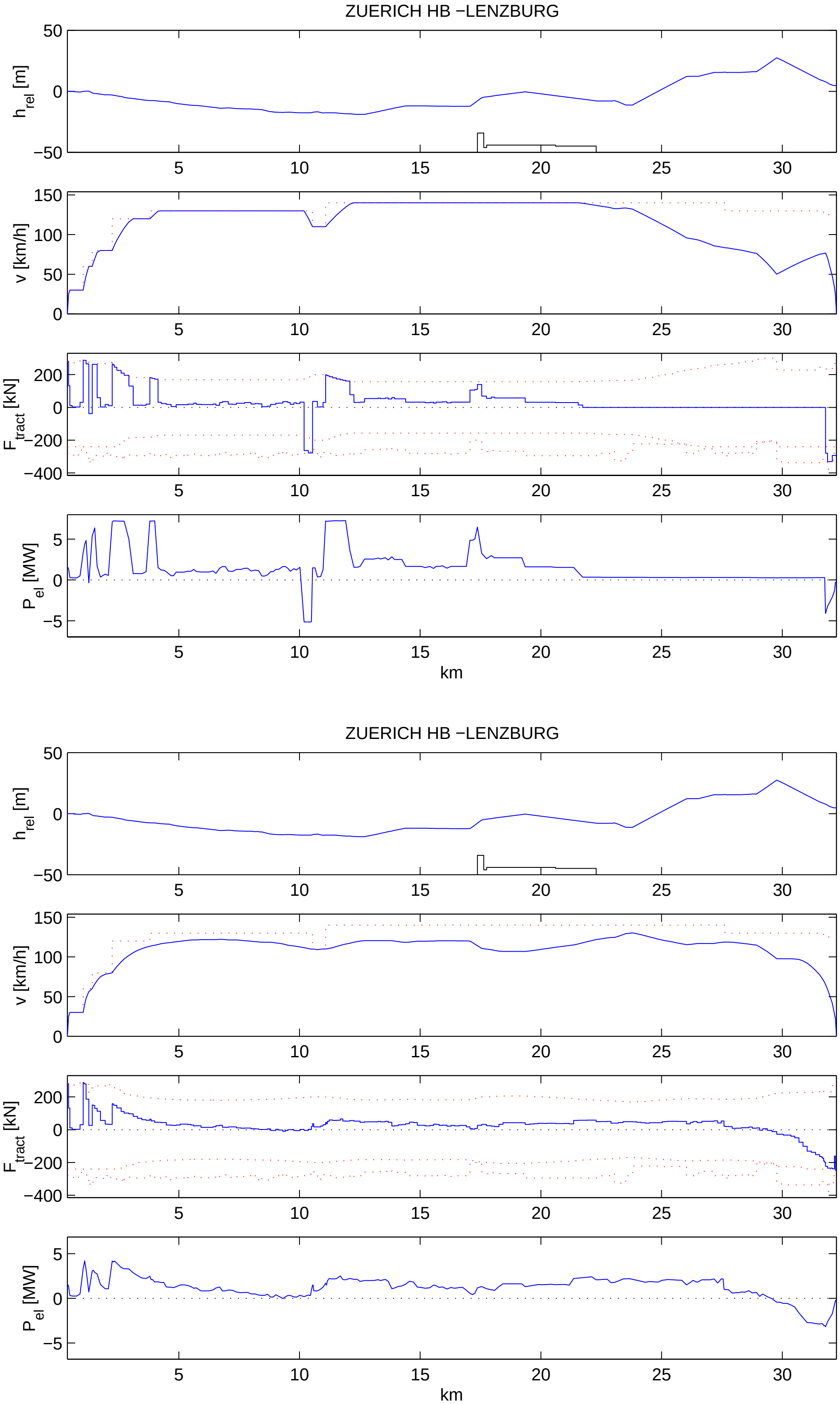

Figure 3: Train performance plots for a passenger train between Zürich and Lenzburg applying a coasting strategy (upper diagram) and optimal driving (lower diagram). The plots show the relative height profile h_{rel} (tunnels are indicated at the km axis), speed v (dotted line for permitted track speed), traction force F_{tract} (dashed lines for maximum traction force, minimum traction force covered by rekuperator brake and total minimum traction force), and electrical power P_{el}.