Minimum-race-time Energy Allocation Strategies for the Hybrid-electric Formula 1 Power Unit

Pol Duhr[1], Daniele Buccheri[1], Camillo Balerna[1], Alberto Cerofolini[2] and Christopher H. Onder[1]

***Abstract*—The hybrid-electric powertrain currently used in Formula 1 race cars draws its energy from the car's fuel tank and battery. The usable battery size is limited, and refueling during a race is forbidden by the regulations of the Formula 1 race series. From a strategic point of view, lap-by-lap targets for the fuel and battery consumption must be chosen and imposed on the energy management controller of the car. This task is non-trivial due to the influence of the on-board fuel mass on the achievable lap time, as well as the cross-couplings between the electric and the combustion part of the powertrain. A systematic approach is thus required to compute the energy allocation strategy that minimizes the total race time. In this paper, we devise an optimization framework in the form of a non-linear program, yielding the optimal battery and fuel consumption targets for each lap of the race. The approach is based on maps that capture the achievable lap time as a function of car mass and allocated battery and fuel energy. These maps are generated beforehand with a model-based single-lap optimization framework and fitted using artificial neural network techniques. To showcase the approach, we present three case studies: First, we compare the optimal strategy to a heuristic method. The improvement of 2 s over the entire race is substantial, given that the difference only lies in the energy allocation, but not in the overall consumption. It underlines the importance of optimizing the energy allocation. Second, we leverage the framework to compute the optimal fuel load at the beginning of the race. Finally, we apply the developed non-linear program in a shrinking-horizon fashion. Our simulation results show that the resulting model predictive controller correctly reacts to disturbances that frequently occur during a race.**

***Index Terms*—Energy management, Formula 1, hybrid electric, artificial neural networks, non-linear programming.**

I. INTRODUCTION

FORMULA 1 races take place on closed circuits and last for a fixed number of laps, such that a total distance of roughly 300 km is covered. After the starting lights go out, 20 drivers battle for position in order to finish first, piloting cars from different manufacturers. To achieve this goal, not only the driver must display supreme race craft. The manufacturer must also perfectly set up the car's chassis and aerodynamic package, and maximize the potential of its powertrain, which is our focus here. For engineering purposes, car performance over a single lap is characterized by the lap time, whilst the objective of winning the race is commonly translated into minimizing the total time required to cover the race distance.

[1]P. Duhr, D. Buccheri and C. H. Onder are with the Institute for Dynamic Systems and Control ETH Zürich, 8092 Zürich, Switzerland (e-mail: {pduhr,dbuccheri,onder}@ethz.ch)

[2]A. Cerofolini is with the Power Unit Performance and Control Group, Ferrari S.p.A., 41053 Maranello, Italy (e-mail: alberto.cerofolini@ferrari.com)

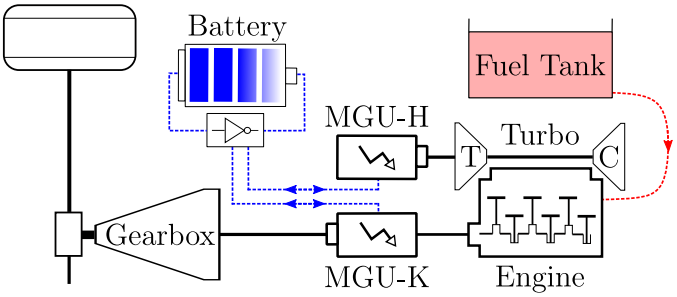

Fig. 1. Schematic of the mechanical and electrical components of the hybrid electric F1 power unit, showing the turbocharged internal combustion engine and the two electric motor-generator units MGU-K and MGU-H. Beside the fuel tank, energy can be drawn from or stored in the battery. Propulsive power is transmitted to the rear wheels via the gearbox.

Formula 1 (F1) cars are hybrid-electric vehicles and feature the Power Unit (PU) depicted in Fig. 1. The turbocharged internal combustion engine (ICE) is compounded by two electric motors: The motor-generator-unit-kinetic (MGU-K) is connected to the engine's crankshaft. Its purpose is kinetic energy recuperation during braking, engine load-point shifting [1] and providing additional power during acceleration. The motor-generator-unit-heat (MGU-H) is coupled to the turbocharger and is mostly used to recuperate excessive power generated by the turbine, but can also be operated in motor mode to speed up the turbocharger and reduce turbo lag. Power transmission takes place via an eight-speed sequential gearbox and a limited-slip differential. There are two on-board sources of energy: the fuel tank for chemical energy, required to run the engine, and the battery for electrical energy, subject to charging and discharging by the two MGUs. To provide an incentive for an efficient design and operation of the powertrain, the technical regulations [2] limit the available energy. The instantaneous fuel mass flow must not exceed 100 kg/h. This makes the peak propulsive power directly depend on the ICE efficiency, since the MGU-K mechanical power is also limited to a maximum of 120 kW. Moreover, the usable battery size is limited to 4 MJ. The fuel is thus the main propellant (a full tank has a chemical energy of about 4800 MJ), but the battery plays an important tactical role as an energy buffer and has a crucial impact on the overall performance. When the regulations for this hybrid PU were first introduced in 2014, fuel consumption for a race was limited to 100 kg [3]. From 2021 onward, the fuel consumption limit was abolished, but re-fueling during the race is still prohibited [4]. Practically, it is not possible to operate the PU at full power whenever desired, otherwise the

driver will run out of fuel and battery energy before finishing the race. This is due to two reasons: First, the race teams want to avoid being too conservative with the fuel load that they carry at the start of the race, since any excess weight penalizes lap time. Choosing the fuel load that ensures the best trade-off between low weight and high energy availability for the PU operation usually means that fuel consumption must be monitored tightly during the race. Second, whilst the battery can be recharged when depleted, this is either very costly in terms of lap time [5], or requires fuel to be converted to battery energy via the MGU-K during part load operation of the PU. Given the strict limits on the energy consumption, the regulations permit the use of a supervisory controller that coordinates the power flows of the PU. Because of the periodic nature of a race, this *energy management* system [6] is generally optimized to achieve minimum lap time over a single lap with a given energy budget [5]. In this paper, we focus on the higher-level task of distributing the available energy over the laps of the race, i.e., defining the single-lap energy budgets, and we refer to this as the *energy allocation* strategy.

In practice, it is implemented via a mix of automatic control and radio communication to the driver. Whilst automatic fuel cut-off at the end of the straights is not permitted for safety reasons, a visual signal on the dashboard display can be provided to the driver, telling him when to lift the throttle pedal to save fuel. In jargon, this technique is called 'lift and coast'. As for the battery consumption target, the driver can set it at the start of each lap with a rotary switch on the steering wheel. The power split between ICE and MGU-K and the recuperation strategy of the MGU-H is then set automatically by the energy management controller of the power unit.

The choice of the energy allocation strategy is a non-trivial task for several reasons. First, the interactions between two energy storages must be considered, which adds a level of complexity compared to battery-electric or engine-only race cars. Second, the fuel consumption over the course of the race entails a substantial variation in vehicle mass, considering that the car only weighs 752 kg. When it becomes lighter, the reduced inertia enables higher cornering speeds and stronger acceleration and deceleration. As a consequence, the lap times tumble. On the other hand, for a given car mass and battery energy allocation, the higher the fuel energy allocation that can be used in a lap, the lower the achievable lap time. Indeed, more fuel is beneficial for the PU operation: The ICE can be operated at full power for longer periods of time, and excess fuel can be used to charge the battery via load-point shifting. This energy can serve to prolong the phases of electric boosting with the MGU-K. It is thus not obvious whether relatively more fuel should be consumed at the beginning of the race to exploit this effect while also quickly decreasing the mass of the car, or rather at the end, when the car is already light. Third, an adaptation of the initial energy allocation strategy might become necessary during the race, caused by events such as pit stops: When the car enters the so-called pit lane for a tire change, it must adhere to a strict speed limit. This can be exploited to recharge the battery, and the energy allocation for the subsequent laps must then be revised to define new lap-by-lap consumption targets. Other examples are safety car or full course yellow phases, during which the cars have to drive at reduced speed [4], or battles for position, which require additional electric boosting.

Against this background, the energy allocation problem must be tackled with a systematic optimization approach. This paper proposes a computationally efficient optimization framework that serves to compute the energy allocation strategy for minimum race time. Its relevance is not confined to F1, given that also other race series have recently moved to hybrid-electric powertrains, e.g., the new Le Mans Hypercars [7].

A. Literature Review

We categorize the relevant research into three different areas. The first pertains to the minimum lap time problem, reflecting the key performance indicator for race cars. One aspect is computing the optimal path on the circuit based on a transient vehicle model, which was first done with nonlinear programming (NLP) in [8], and later improved with curvilinear coordinates [9] and numerical modifications [10]. A noteworthy extension concerns the optimal management of tire degradation over a sequence of laps [11]. A second aspect is the time-optimal energy management strategy, computed on the assumption of a fixed driving path. For hybrid F1 cars, this was done with second-order cone programming [5] and nonconvex NLP [12]. For electric race cars, [5] was modified and extended to evaluate different transmission technologies [13]. Since the power of electric motors can be subject to thermal constraints [14], [15], the optimal thermal management in a racing context was considered in [16], [17]. Inspired by [18], simplified vehicle dynamics models were included to study the torque vectoring of electric race cars [19] and to model the performance envelope of a F1 car in the g-g diagram [20]. Besides, several authors have jointly optimized the energy management and the driving path for hybrid [21], [22] and electric race cars [23], and quite recently, stochastic optimization was proposed to account for the competitors' impact on the energy management [24]. All the cited publications that deal with energy management optimize a single lap.

The second research area consists of minimum-race-time energy management strategies of complex powertrains. In [25], the race is treated as a sequence of spatially discretized laps. However, results are only shown for two laps and not for an entire race, putting a question mark over the computational tractability of such an approach. By contrast, [26] and [27] propose a separation of the race optimization from the single-lap optimization. The race was discretized lap-by-lap and the dependency of lap time on the relevant variables of the considered electric powertrain was captured with maps represented by means of an Artificial Neural Network (ANN), which is a promising technique for control purposes [28] and curve fitting [29]. The race strategy was then determined using Monte Carlo tree search [26] or reinforcement learning [27]. Unfortunately, these methods are devoid of any optimality guarantees, and in one case, the resulting strategy was even shown to be clearly sub-optimal. For the related problem of computing maximum-distance strategies for electrical endurance racing, the optimality issue was addressed by combining this bi-level approach with mixed-integer optimization [30].

The third stream of research deals with race strategy decisions in terms of pit stops, tire changes and overtaking. The importance of related simulation tools for motor sport teams was underlined in [31] for F1 and in [32] for stock car racing. Pit stop strategies for F1 were simulated using a discrete-event approach and a sub-division of each lap into mini-sectors [33], or based on a lap-by-lap discretization [34], also including stochastic effects via Monte Carlo sampling [35]. The aforementioned works also consider the behavior of the competitors. Moving towards decision-making, a data-driven ANN was proposed to decide on the tire strategy [36]. Game-theoretical approaches for planning overtake maneuvers were explored for autonomous race cars [37] and sailing yachts [38]. Methods based on mathematical optimization are very scarce in this field, with one notable example again stemming from the field of sailing yacht racing [39].

B. Research Statement

Whilst a lot of effort was dedicated to the minimum-*lap*-time energy management problem, the minimum-*race*-time energy allocation problem has only been scarcely investigated and has not yet been tackled with numerical optimization algorithms for hybrid-electric race cars. A straightforward approach consists in modeling the race as a sequence of spatially discretized laps and solving the problem all at once with the methods developed for minimum-lap-time optimization. However, as we will show, this is computationally demanding and often hampered by numerical issues, due to the large number of optimization variables involved. The optimization framework presented in this paper addresses these issues. It allows to determine the optimal fuel and battery consumption targets for each lap of the race, and can provide a benchmark solution for heuristic approaches and on-line control, allowing to quantify the sub-optimality. It can also serve to identify the optimal fuel load at the start of the race, that is, to decide whether the fuel tank's size should be fully exploited, or whether it is advantageous to trade some fuel for a lighter car. Furthermore, with such a tool the energy allocation can be adapted during a race in the case of an unexpected disturbance, such as a pit stop or an overtake maneuver.

C. Contributions

Our work comes with two main contributions: First, similar to [26], [27], [30], we discretize the minimum-race-time problem on a *lap-by-lap* basis and separate it from the single-lap energy management problem by means of *lap time maps*. The required data is obtained by solving the minimum-lap-time problem on the given circuit for different energy consumption targets. The maps are then represented with ANNs, given their good fitting capabilities. The resulting NLP yields the optimal energy allocation for each lap and can be solved efficiently with off-the-shelf solvers. In contrast to [26], [27], the solution is guaranteed to be at least a local optimum. Compared to the existing methods, which focus on battery-electric vehicles, our model-based approach also takes into account that the car's mass decreases as fuel is consumed. This is a crucial effect not only in F1 cars, but in every race car with a powertrain that comprises an internal combustion engine. To the best of our knowledge, so far the impact of the fuel consumption on race time has only been included in simulation tools such as [34], in the form of a fixed consumption per lap and an associated lap time sensitivity. In this context, the novelty of our work consists in systematically optimizing the fuel consumption for each lap of the race by solving the minimum-race-time problem, rather than treating it as a pre-defined parameter. Second, we show that our minimum-race-time problem can be easily leveraged for on-line applications, in the form of a model predictive controller (MPC) [40] that updates the energy allocation strategy during the race. As such, it can complement the pit stop strategy simulation and optimization tools already used by most race teams. To validate that the on-line strategies are close-to-optimal, we benchmark the causal MPC strategy against the non-causal optimal solution. MPC has been successfully applied to the energy management of hybrid vehicles, e.g., in [41]–[43]. However, in the context of motor races, the use of MPC for a lap-by-lap update of the energy allocation strategy has not yet been investigated in literature, as far as we know. Specifically, [27] and [30] do not treat the aspect of on-line updates at all, whilst [26] applies a heuristic search algorithm without any optimality guarantees and does not provide a comparison with the optimal solution. Even though this paper focuses on the F1 PU, one of the most technologically advanced hybrid race car powertrains, the method is applicable to any motor sports competition with races over a fixed number of laps and restricted energy usage.

D. Outline

The remainder of the text is structured as follows: In Section II, we present the optimization framework. Next, we discuss its components, namely the minimum-race-time energy allocation problem in Section III, the minimum-lap-time energy management problem in Section IV, and the procedure to fit the lap time maps in Section V. Section VI contains the race strategy optimization results and the case studies. Finally, we conclude the paper in Section VII and give an outlook on future research.

II. RACE OPTIMIZATION FRAMEWORK

The proposed optimization framework to compute the *minimum-race-time energy allocation* strategy is schematically depicted in Fig. 2. The race lasts for N laps of a given circuit, and the fuel mass that is in the car's tank at the start of the race is denoted by m_f. The goal is to compute the fuel energy allocation ΔE_f and the battery energy allocation ΔE_b for each lap $i \in \{1,\dots,N\}$. We observe that the total race time is the sum of all the lap times, and we will thus discretize the race on a lap-by-lap basis. The crucial assumption of our approach is that the achievable lap time T_lap in any lap can be described by a *lap time map* $\mathscr{M}$ as a function of the fuel and battery energy allocation for that lap, as well as the mass of the car m and the battery energy content $E_\mathrm{b,init}$ at the start of the lap.

The choice of these parameters is instigated by the following reasoning: The allocated fuel energy determines whether the engine needs to be turned off already before the car arrives

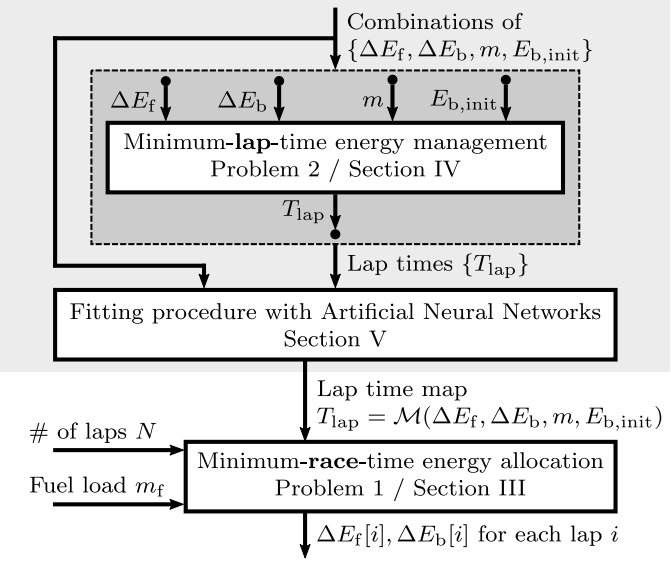

Fig. 2. Schematic representation of the proposed optimization framework. The part shaded in gray is required to generate the lap time maps, which form the basis of our race optimization approach.

at braking zones in order to save fuel, or whether it can be operated at full power along each straight, whilst the allocated battery energy strongly influences the amount of electric boosting that the MGU-K can provide. The energy management strategy in terms of power split, load point shifting and recuperation is also influenced [5]. Hence, the achievable lap time very strongly depends on ΔE_f and ΔE_b. Additionally, the lap time varies with the car mass: A heavier car has a larger inertia in corners, in braking zones and during acceleration and is therefore slower whilst, on the other hand, the normal force on the tires and thus the available tire grip increase [44]. Finally, we have to consider the initial battery energy content $E_\mathrm{b,init}$ at the start of the lap. This is due to the finite size of the battery. Indeed, if the battery is almost fully charged, the recuperation capabilities may be restricted during some parts of the lap. This lack of recuperation leads to a higher lap time, compared to a lap which is not affected by the limits on the battery energy, i.e., a lap with the same ΔE_b but with a value of $E_\mathrm{b,init}$ that is further below the upper bound. Evidence which supports this statement is provided in Section IV-F. Similarly, if the battery is quasi empty, the electric boosting capability is severely restricted by the lower bound, which again has a detrimental effect on lap time.

To synthesize the lap time map, the achievable lap time must be computed for different combinations of fuel and battery energy allocation, car mass and initial battery energy. We do this by solving the *minimum-lap-time energy management problem*, based on a mathematical model of the car's longitudinal and lateral dynamics and its powertrain. This model takes into account the aforementioned effects which have an impact on the lap time. The generated data is then fitted using ANN techniques, which yields a representation of the map that can be integrated into the minimum-race-time optimization problem.

The separation from the single-lap problem renders the race optimization computationally tractable. Moreover, the lap time map has to be generated only once for the circuit under consideration and can then be used for different tasks, as we will show in Section VI. In the following, we will describe in detail the three blocks that form this optimization framework.

III. MINIMUM RACE TIME PROBLEM

In this section, we formulate the optimal control problem that yields the energy allocation strategy for minimum race time. The objective is to minimize the time required to cover the total race distance, which is the sum of the lap times:

$$\min \quad \sum_{i=1}^{N} T_\mathrm{lap}[i]. \tag{1}$$

We assume that the lap time of lap i is described by the lap time map $\mathscr{M}_i$ as a function of the control inputs ΔE_f and ΔE_b, as well as the mass of the car m and the battery energy content E_b at the start of the lap:

$$T_\mathrm{lap}[i] = \mathscr{M}_i(\Delta E_\mathrm{f}[i], \Delta E_\mathrm{b}[i], m[i], E_\mathrm{b}[i]). \tag{2}$$

The map $\mathscr{M}_i$ can be different depending on the lap number i. This allows to model special scenarios such as pit stops for tire changes or (virtual) safety car phases [4], by providing an adapted map for the laps where they occur. We assume that it is known a priori for which laps this is the case.

Next, we model the fuel and battery energy allocation on a lap-by-lap basis. The consumed fuel energy E_f evolves as

$$E_\mathrm{f}[i+1] = E_\mathrm{f}[i] + \Delta E_\mathrm{f}[i] \tag{3}$$

and is monotonically increasing, since fuel cannot be recuperated once consumed, leading to the constraint

$$\Delta E_\mathrm{f}[i] \geq 0. \tag{4}$$

The initial and terminal conditions are

$$E_\mathrm{f}[1] \quad = 0, \tag{5}$$
$$E_\mathrm{f}[N+1] \leq m_\mathrm{f} \cdot H_\mathrm{lhv}, \tag{6}$$

where m_f is the fuel mass that is in the car's fuel tank at the start of the race, and H_lhv is the lower heating value of the fuel. Hence, by (6), the entire fuel may be consumed by the end of the race. Analogously, the battery energy at the start of each lap evolves as

$$E_\mathrm{b}[i+1] = E_\mathrm{b}[i] + \Delta E_\mathrm{b}[i]. \tag{7}$$

The technical regulations [2] stipulate a usable battery size of 4 MJ. We assume that the battery is fully charged at the start of the race, i.e.,

$$E_\mathrm{b}[1] = 4\,\mathrm{MJ}, \tag{8}$$

and that it can be completely depleted by the end of the race:

$$E_\mathrm{b}[N+1] \geq 0\,\mathrm{MJ}. \tag{9}$$

In-between, the battery energy must stay within the bounds

$$E_\mathrm{b}[i] \geq 0\,\mathrm{MJ}, \tag{10}$$
$$E_\mathrm{b}[i] \leq 4\,\mathrm{MJ}. \tag{11}$$

Naturally, the battery energy allocation ΔE_b can be positive or negative, with positive values corresponding to charging and negative values to discharging. Finally, the mass of the car varies according to the consumed fuel mass, captured by the following relationship:

$$m[i+1] = m[i] - \Delta E_\mathrm{f}[i] \cdot \frac{1}{H_\mathrm{lhv}}. \tag{12}$$

The mass at the start of the race corresponds to the mass of the car and driver m_car plus the fuel mass on-board, given by

$$m[1] = m_\mathrm{car} + m_\mathrm{f}. \tag{13}$$

The stated objective and constraints are linear, the only exception being the non-linear and non-convex relationships involved in the ANNs required to model constraint (2) (see Section V). We choose feedforward ANNs with a nonlinear activation function, because they yield very accurate fits at a level of precision that would be difficult to achieve with convex approximations for this particular application. This is important, given that an imprecise lap time map leads to artifacts in the solution of the problem. Hence, the minimum race time problem boils down to the following NLP:

Problem 1. *The optimal energy allocation strategy in terms of fuel and battery target for each race lap is the solution of*

$$\min_{\Delta E_\mathrm{f}, \Delta E_\mathrm{b}} \quad \sum_{i=1}^{N} T_\mathrm{lap}[i]$$

subject to the following constraints $\forall i \in \{1, \ldots, N\}$*:*

Lap time:	(2),
Fuel:	(3), (4), (5), (6),
Battery:	(7), (8), (9), (10), (11),
Car mass:	(12), (13).

Three aspects linked to race strategy that we do not optimize in Problem 1 are the tire choice, the tire degradation [11] and the scheduling of pit stops for tire changes. Optimizing these decisions is beyond the scope of this paper, which focuses on the energy allocation strategy.

IV. MINIMUM LAP TIME PROBLEM

In this section, we formulate an optimization problem for the minimum-lap-time energy management. It will be used in Section V to generate the data that forms the basis of the lap time maps. We work on the model-based convex optimization framework developed in [5], [20], but with several modifications and extensions, leading to an NLP. The optimization approach is based on the assumption that the path around the race circuit is known a priori and not subject to optimization. The so-called 'racing line' was synthesized from velocity and acceleration measurements of one representative lap [8].

Our model features a simple mathematical description of the car's performance envelope for the longitudinal and lateral vehicle dynamics, and of the PU components. We will not discuss the identification of the parameters and the model validation in this paper. They were shown in [5] for the PU and in [20] for the vehicle dynamics performance envelope. We do not include a thermal model of the electric motors, since thermal de-rating is generally not an issue with current-generation F1 cars. The impact of the gearshifts is also neglected, since optimizing the gearshift strategy would lead to a mixed-integer optimization problem [12], [45], which is computationally intractable for the application at hand. In the following, we will summarize the modeling equations.

A. Lap time and longitudinal dynamics

Since the path on the race track is given, we formulate the minimum lap time problem in space domain. The car is assumed to be a point mass m traveling with velocity v along the fixed path, parameterized by the independent path distance variable $s \in [0, S]$, where S denotes the length of the path for a single lap. The objective of the control problem is then to minimize the lap time T, which can be transformed from time domain to space domain as follows:

$$\min \int_0^T \mathrm{d}t = \min \int_0^S \frac{\mathrm{d}t}{\mathrm{d}s}(s) \cdot \mathrm{d}s = \min \int_0^S \frac{1}{v(s)} \cdot \mathrm{d}s. \tag{14}$$

By Newton's second law, the velocity evolves according to the differential equation

$$\frac{\mathrm{d}}{\mathrm{d}s} v(s) = \frac{1}{m \cdot v(s)} \cdot \left(F_\mathrm{p}(s) - F_\mathrm{d}(s) \right), \tag{15}$$

where F_p denotes the propulsive force and F_d the total drag force acting on the car. Note that we assume the car mass m to be constant over the course of a lap, neglecting the overall decrease by ca. 1.5 kg stemming from the fuel consumption. The drag force consists of aerodynamic drag, gravitational forces acting on the car, and rolling friction:

$$F_\mathrm{d}(s) = F_\mathrm{aero}(s) + F_\mathrm{grav}(s) + F_\mathrm{roll}(s). \tag{16}$$

The modeling of these forces in the context of the F1 car was described, identified and validated in [5]. The aerodynamic drag force includes a term that depends on the path curvature $\gamma(s)$ of the racing line and is given by

$$F_\mathrm{aero}(s) = (c_\mathrm{d,0} + c_\mathrm{d,1} \cdot \gamma(s)) \cdot v(s)^2, \tag{17}$$

where $c_\mathrm{d,0}$ and $c_\mathrm{d,1}$ are parameters that have to be identified with measurement data. The hill force is

$$F_\mathrm{grav}(s) = m \cdot g \cdot \sin(\theta(s)), \tag{18}$$

with $g = 9.81\,\mathrm{m/s^2}$ denoting the gravitational constant and $\theta(s)$ the slope. Finally, the rolling friction is given by

$$F_\mathrm{roll}(s) = c_\mathrm{roll} \cdot m \cdot g \cdot \cos(\theta(s)), \tag{19}$$

with the parameter c_roll subject to identification. The propulsive force is related to the propulsive power P_p by

$$F_\mathrm{p}(s) = \frac{P_\mathrm{p}(s)}{v(s)}. \tag{20}$$

Since we are optimizing a single lap embedded in a sequence of laps, we include a velocity periodicity constraint [45],

$$v(0) = v(S). \tag{21}$$

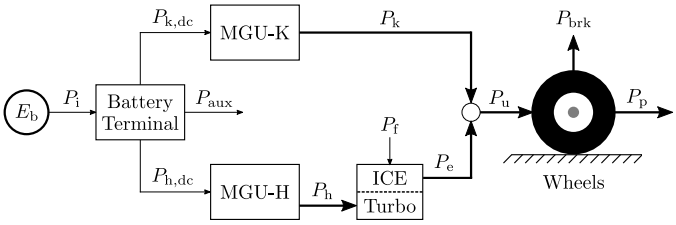

Fig. 3. Schematic representation of the power flows in the F1 powertrain. The arrows indicate the direction of positive power. Thick arrows illustrate mechanical power, whilst thin arrows illustrate electrical power (with the exception of the chemical fuel power).

B. Powertrain model

The power flows that have to be considered within the F1 power unit are depicted schematically in Fig. 3. The effective propulsive power is the sum of the total PU power P_u delivered to the rear wheels and the power $P_\mathrm{brk}^\mathrm{f,fric}$ and $P_\mathrm{brk}^\mathrm{r,fric}$ dissipated in the front and rear friction brakes, respectively:

$$P_\mathrm{p}(s) = c_\mathrm{s} \cdot P_\mathrm{u}(s) - P_\mathrm{brk}^\mathrm{f,fric}(s) - P_\mathrm{brk}^\mathrm{r,fric}(s), \tag{22}$$

where the coefficient $c_\mathrm{s} < 1$ models the wheel slip losses [5]. The friction brakes can only dissipate power, hence

$$P_\mathrm{brk}^\mathrm{f,fric}(s) \geq 0, \quad P_\mathrm{brk}^\mathrm{r,fric}(s) \geq 0 \tag{23}$$

must hold. The engine power P_e and the MGU-K power P_k add up to the total PU power:

$$P_\mathrm{u}(s) = P_\mathrm{e}(s) + P_\mathrm{k}(s). \tag{24}$$

The key input on which the engine power depends is the chemical fuel power P_f, related to the fuel mass flow $\dot{m}_\mathrm{f}$ by

$$P_\mathrm{f}(s) = \dot{m}_\mathrm{f}(s) \cdot H_\mathrm{lhv} \tag{25}$$

and limited by the regulations [2] to

$$P_\mathrm{f}(s) \geq 0, \tag{26}$$
$$P_\mathrm{f}(s) \leq 100\,\mathrm{kg/h} \cdot H_\mathrm{lhv}. \tag{27}$$

We describe the engine power using a Willans model [46] as

$$P_\mathrm{e}(s) = \eta_\mathrm{e}(r_\mathrm{wg}(s)) \cdot P_\mathrm{f}(s) - P_\mathrm{e,0}, \tag{28}$$

where $P_\mathrm{e,0}$ denotes the engine drag power, and where the efficiency η_e depends on the waste-gate position r_wg, which influences the engine back-pressure. The MGU-H is assumed to operate only in generator mode [5]. The power P_h that it recuperates from the turbocharger compound is modeled by

$$P_\mathrm{h}(s) = \eta_\mathrm{h}(r_\mathrm{wg}(s)) \cdot P_\mathrm{f}(s) \leq 0. \tag{29}$$

The waste-gate operation has a large impact on the recuperation efficiency η_h and must therefore be considered for the energy management strategy. The reasons for this are explained in detail in [47].

Regarding the electrical part of the PU, the power at the battery terminal is given by

$$P_\mathrm{b}(s) = P_\mathrm{k,dc}(s) + P_\mathrm{h,dc}(s) + P_\mathrm{aux}, \tag{30}$$

where $P_\mathrm{k,dc}$ and $P_\mathrm{h,dc}$ are the electrical power flows to/from the MGU-K and MGU-H, respectively, and P_aux is a small auxiliary power. A quadratic model parameterized by a coefficient α is used to capture the losses of the MGUs, as well as the relationship between P_b and the internal power P_i of the battery:

$$P_\mathrm{k,dc}(s) = \alpha_\mathrm{k} \cdot P_\mathrm{k}(s)^2 + P_\mathrm{k}(s), \quad \alpha_\mathrm{k} > 0, \tag{31}$$
$$P_\mathrm{h,dc}(s) = \alpha_\mathrm{h} \cdot P_\mathrm{h}(s)^2 + P_\mathrm{h}(s), \quad \alpha_\mathrm{h} > 0, \tag{32}$$
$$P_\mathrm{i}(s) \quad = \alpha_\mathrm{b} \cdot P_\mathrm{b}(s)^2 + P_\mathrm{b}(s), \quad \alpha_\mathrm{b} > 0. \tag{33}$$

Next, we describe the evolution of the fuel energy E_f consumed over the course of the lap:

$$\frac{\mathrm{d}}{\mathrm{d}s} E_\mathrm{f}(s) = \frac{P_\mathrm{f}(s)}{v(s)}. \tag{34}$$

The division by the velocity stems from the fact that we are formulating the differential equation in the space domain rather than in time domain. The initial and terminal conditions on the fuel energy are

$$E_\mathrm{f}(0) = 0, \tag{35}$$
$$E_\mathrm{f}(S) = \Delta E_\mathrm{f}. \tag{36}$$

The parameter ΔE_f is the fuel energy allocation for the lap, as introduced in the minimum race time problem in Section III.

Finally, the battery energy content E_b evolves as

$$\frac{\mathrm{d}}{\mathrm{d}s} E_\mathrm{b}(s) = -\frac{P_\mathrm{i}(s)}{v(s)}. \tag{37}$$

Introducing the parameter $E_\mathrm{b,init}$, the initial and terminal conditions for the battery are

$$E_\mathrm{b}(0) = E_\mathrm{b,init}, \tag{38}$$
$$E_\mathrm{b}(S) = E_\mathrm{b,init} + \Delta E_\mathrm{b}, \tag{39}$$

where ΔE_b is the battery energy allocation for the lap. For $\Delta E_\mathrm{b} > 0$, the battery is charged, while for $\Delta E_\mathrm{b} < 0$, it is discharged. Moreover, there are path constraints on the battery energy, since the difference between the maximum and the minimum state-of-energy of the battery must not exceed 4 MJ while the car is on track [2]:

$$E_\mathrm{b}(s) \geq 0\,\mathrm{MJ}, \tag{40}$$
$$E_\mathrm{b}(s) \leq 4\,\mathrm{MJ}. \tag{41}$$

In the F1 regulations [2], additional limitations on the allowed boosting and recuperation with the MGU-K are specified. For the sake of brevity, these are not discussed here, but nonetheless included in the optimization problem. Interested readers are again referred to [5].

C. Performance envelope

The performance envelope describes the grip limits of the car. It constrains the achievable longitudinal acceleration as a function of the lateral acceleration, and thus has a direct impact on the lap time. The path on the track, assumed to be known, is characterized by its curvature $\gamma(s)$. Hence, following a quasi-steady-state modeling approach [8], the total lateral force F_lat produced by the four tires of the car results from the formula for uniform circular movement:

$$F_\mathrm{lat}(s) = m \cdot v(s)^2 \cdot \gamma(s). \tag{42}$$

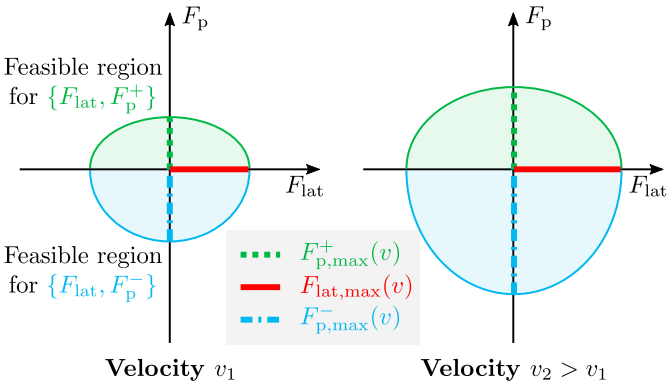

Fig. 4. Schematic representation of the performance envelope model given in (46), (47). As velocity increases, the half-axes that define the upper and lower half-ellipses grow according to (48), (49), (50). This captures the effect of the increasing normal force on the tires due to the aerodynamic downforce.

For the performance envelope, we follow the modeling approach that we previously developed in [20], and which is akin to a tire friction circle, or to the attainable area in the so-called g-g diagram [44]. Specifically, we constrain the total force sustained by the four tires, i.e., the vector combination of propulsive force F_p and lateral force F_lat. We do so by prescribing a bound consisting of two separate half-ellipses for acceleration ($F_\mathrm{p} \geq 0$) and deceleration ($F_\mathrm{p} < 0$), as depicted in Fig. 4. Ultimately, this performance envelope formulation is comparable to a quasi-steady-state approach: The longitudinal dynamics are taken explicitly into account, as described in Section IV-A, while the lateral dynamics are not modeled. Whilst in [20], we constrained the combined propulsive and lateral *acceleration*, here we re-write the model in terms of *forces*. This allows us to capture the effect of different values for the car mass in a simple manner, as outlined below.

First, we introduce two auxiliary variables that serve to constrain acceleration and deceleration phases differently, as described in [20]. Therefore, we split the propulsive force into a positive and a negative component:

$$F_\mathrm{p}(s) = F_\mathrm{p}^+(s) + F_\mathrm{p}^-(s), \tag{43}$$
$$F_\mathrm{p}^+(s) \geq 0, \tag{44}$$
$$F_\mathrm{p}^-(s) \leq 0. \tag{45}$$

We then specify the following two elliptic constraints, which limit F_p^+ and F_p^- depending on F_lat:

$$\frac{F_\mathrm{p}^+(s)^2}{F_{\mathrm{p,max}}^+(v)^2} + \frac{F_\mathrm{lat}(s)^2}{F_{\mathrm{lat,max}}(v)^2} \leq 1, \tag{46}$$
$$\frac{F_\mathrm{p}^-(s)^2}{F_{\mathrm{p,max}}^-(v)^2} + \frac{F_\mathrm{lat}(s)^2}{F_{\mathrm{lat,max}}(v)^2} \leq 1. \tag{47}$$

The variables $F_{\mathrm{p,max}}^+$, $F_{\mathrm{p,max}}^-$ and $F_{\mathrm{lat,max}}$ denote the half-axes of the ellipses, as shown in Fig. 4, and thus the maximum force that can be generated by all four tires in the longitudinal and lateral direction. Regarding the two auxiliary variables, it was shown in [20] that the optimal solution satisfies the complementarity condition $F_\mathrm{p}^+(s) \cdot F_\mathrm{p}^-(s) = 0$ whenever the car is grip-limited, i.e., at the boundary of the upper or lower half-ellipse. The achievable forces are not constant, but increase with the velocity, due to the effect of the aerodynamic downforce generated by the wings, underbody and diffusor of an F1 car [44]. If we assume a simple friction law [48], the maximum total tire force is directly proportional to the normal force acting on the tires, which is the sum of the aerodynamic downforce and the car's weight. The downforce is zero for $v = 0$ and can be modeled as a quadratic function of velocity [44], whilst the car's weight is assumed to be constant during the lap, as previously explained. Combining these considerations, we propose to approximate the velocity dependency of the maximum lateral force as follows:

$$F_{\mathrm{lat,max}}(v) = \underbrace{c_1 \cdot v(s)^2 + c_2 \cdot v(s)}_{\text{aerodynamic}} + \underbrace{c_3 \cdot \frac{m}{m_\mathrm{nom}}}_{\text{weight}}, \tag{48}$$

where c_1, c_2, c_3 are fitting parameters subject to identification. Our simplified approach does not include four individual tire modeling equations, nor does it explicitly model the impact of quantities such as slip angle, longitudinal slip ratio, camber, load transfer, or load sensitivity of the tire friction coefficients [44]. However, following the method proposed in [20], we directly identify the parameters in (48) with measured acceleration data from a representative lap. Specifically, we use Newton's second law and relate force to acceleration to obtain a fit for the maximum lateral force. This procedure implicitly considers the aforementioned quantities and is sufficiently accurate for the purposes of energy management optimization. Multiplying c_3 with m/m_nom, where m_nom is the nominal car mass for which the parameters were identified, allows us to scale the performance envelope depending on the specified mass m without re-identifying the parameters. The scaling is physically meaningful, since it only affects the weight-dependent component of the maximum tire force. We will validate this simple mass sensitivity model in Section V. Finally, the maximum longitudinal forces are modeled identical to (48) by

$$F_{\mathrm{p,max}}^+(v) = d_1^+ \cdot v(s)^2 + d_2^+ \cdot v(s) + d_3^+ \cdot \frac{m}{m_\mathrm{nom}}, \tag{49}$$
$$F_{\mathrm{p,max}}^-(v) = d_1^- \cdot v(s)^2 + d_2^- \cdot v(s) + d_3^- \cdot \frac{m}{m_\mathrm{nom}}, \tag{50}$$

where $d_1^+, d_2^+, d_3^+, d_1^-, d_2^-, d_3^-$ are again parameters subject to identification.

D. Brake-by-wire

Modern F1 cars are equipped with a brake-by-wire system that seamlessly balances the power of the friction brakes with the MGU-K recuperation set by the PU control to enable the driver to brake at the limit [2]. In order to achieve this, a brake balance constraint between the front and rear axles must be respected. The constant $k_\mathrm{bb}^\mathrm{tot}$ prescribes the ratio between the total rear and front braking power as

$$k_\mathrm{bb}^\mathrm{tot} = \frac{-P_\mathrm{brk}^{\mathrm{r,fric}}(s) + P_\mathrm{k}(s) + P_\mathrm{e}(s) + P_{\mathrm{e},0}}{-P_\mathrm{brk}^{\mathrm{f,fric}}(s)}. \tag{51}$$

Indeed, the braking power on the front axle stems only from the friction brakes, whereas for the rear axle the PU power must also be taken into account, corrected by the engine drag $P_{\mathrm{e},0}$. (The latter is a separate tuning parameter for the

car's handling behavior.) The minus sign is due to the fact that the friction brake power is positive, whilst the MGU-K recuperation power is negative. The brake balance ratio is usually set to a value ≈ 1, but can be modified by the driver within certain limits with a switch on the steering wheel, to fine-tune the set-up of the car. We neglect this effect and assume a constant $k_{\mathrm{bb}}^{\mathrm{tot}}$. Since (51) must hold only when the driver presses the brake pedal, i.e., when the friction brakes are activated, we reformulate the constraint to

$$(P_{\mathrm{e}}(s)+P_{\mathrm{e},0})+P_{\mathrm{k}}(s)-P_{\mathrm{brk}}^{\mathrm{r,fric}}(s)=-k_{\mathrm{bb}}^{\mathrm{tot}}\cdot P_{\mathrm{brk}}^{\mathrm{f,fric}}(s)+\varepsilon(s), \quad (52)$$

where we introduce the slack variable

$$\varepsilon(s)\geq 0 \quad (53)$$

subject to the complementarity constraint

$$\left(P_{\mathrm{brk}}^{\mathrm{f,fric}}(s)+P_{\mathrm{brk}}^{\mathrm{r,fric}}(s)\right)\cdot\varepsilon(s)=0. \quad (54)$$

With these provisions, positive PU power can be delivered to the rear axle when the friction brakes power is zero, without violating the brake balance constraint: Indeed, if $P_{\mathrm{brk}}^{\mathrm{f,fric}}=P_{\mathrm{brk}}^{\mathrm{r,fric}}=0$, then $\varepsilon\geq 0$ by (53), (54) and thus (52) yields $P_{\mathrm{e}}+P_{\mathrm{e},0}+P_{\mathrm{k}}=\varepsilon\geq 0$. The addition of the term $P_{\mathrm{e},0}$ in (52) also makes it possible to choose $P_{\mathrm{e}}=-P_{\mathrm{e},0}$ and $P_{\mathrm{k}}=0$, corresponding to the 'lift and coast' operation where neither throttle nor brake pedal are pressed and only the engine drag is acting on the drivetrain. If the friction brakes are active, i.e., $P_{\mathrm{brk}}^{\mathrm{f,fric}}+P_{\mathrm{brk}}^{\mathrm{r,fric}}>0$, then $\varepsilon=0$ by (54) and from (52) the original brake balance constraint (51) is retrieved.

Additionally, there is a mechanical brake balance constraint stemming from the hydraulic brake circuit, defined by the constant $k_{\mathrm{bb}}^{\mathrm{fric}}$ and modeled as

$$P_{\mathrm{brk}}^{\mathrm{f,fric}}(s)\geq k_{\mathrm{bb}}^{\mathrm{fric}}\cdot P_{\mathrm{brk}}^{\mathrm{r,fric}}(s). \quad (55)$$

It is important to include the brake balance constraints, since they limit the MGU-K recuperation power in certain phases of the braking maneuver, but also the amount of fuel that can be burned during braking. Turning on the engine while braking is a technique that can be exploited if one wishes to consume an excess of fuel in a given lap in order to render the car lighter.

E. Minimum lap time control problem

With these preparations, the minimum lap time problem for a single lap can be formulated in the form of an NLP:

Problem 2. *The optimal energy management strategy for a single lap is the solution of*

$$\min_{P_{\mathrm{f}},P_{\mathrm{k}},r_{\mathrm{wg}},P_{\mathrm{brk}}^{\mathrm{f,fric}},P_{\mathrm{brk}}^{\mathrm{r,fric}},\varepsilon}\quad \int_0^S \frac{1}{v(s)}\cdot \mathrm{d}s$$

subject to the following constraints $\forall s\in[0,S]$*:*

Longitudinal dynamics:	(15) – (21),
Powertrain model:	(22) – (24), (26) – (41)
Performance envelope:	(42) – (50)
Brake-by-wire:	(51) – (55).

Special scenarios like, e.g., pit stops can be modeled by slightly modifying Problem 2. We give a brief description of the modifications in Appendix A.

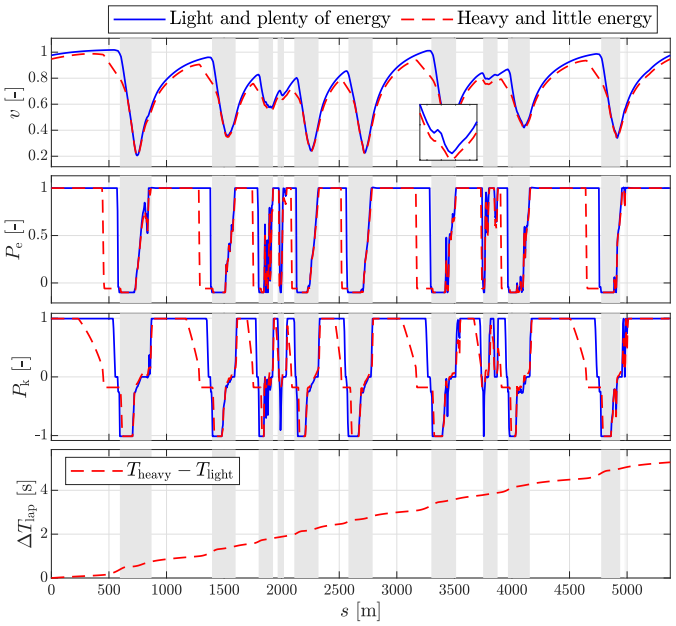

Fig. 5. Solution of Problem 2 for two different scenarios: *'Light and plenty of energy'* with $\{\Delta E_{\mathrm{f}}=78.1\,\mathrm{MJ},\Delta E_{\mathrm{b}}=-1\,\mathrm{MJ},m=760\,\mathrm{kg}\}$, and *'Heavy and little energy'* with $\{\Delta E_{\mathrm{f}}=70\,\mathrm{MJ},\Delta E_{\mathrm{b}}=0\,\mathrm{MJ},m=860\,\mathrm{kg}\}$. The different energy allocation causes drastic changes in the energy management strategy, whilst the difference in vehicle mass results in different cornering speeds and acceleration capabilities, as shown by the zoom-in excerpt. Corners and braking zones are indicated by a gray background, straights by a white one. The bottom subplot shows the accumulated lap time difference.

F. Single-lap solution

The results shown in this paper are all for the Bahrain International Circuit. For reasons of confidentiality, some data has been normalized in the plots. Problem 2 is discretized with the Runge-Kutta-4 method and then implemented in MATLAB as a multiple-shooting problem [49], [50] using the symbolic framework CasADi [51] and solved with the interior-point solver IPOPT [52]. With a spatial discretization step of $\Delta s=5\,\mathrm{m}$, parsing and solving the problem takes around 15 s on a standard consumer laptop with a 4 GHz processor.

As an example, in Fig. 5, we show the solution of Problem 2 for two different combinations of fuel/battery energy allocation and car mass. All other parameters are identical. We can appreciate the effect of constraint (21), which leads to velocity trajectories that are periodic with respect to the start/finish line. Two things become apparent: First, if a smaller quantity of fuel is available, the engine must be turned off at the end of the straights in order to save fuel. Second, if the battery can be discharged, more electric boosting is available on the straights and the MGU-K is operated at full power during larger portions of the lap. Third, if the car is lighter, it gains and sheds speed faster and can also sustain a higher velocity in the corners, due to the fact that $F_{\mathrm{lat,max}}$ in (48) is not directly proportional to m. Clearly, the second scenario with a smaller energy allocation and a heavier car yields a significantly higher lap time - the difference is roughly 5 s (with a lap time of roughly 101 s).

In Fig. 6, we show the solution for three laps with the same battery charge target $\Delta E_{\mathrm{b}}=+1\,\mathrm{MJ}$, but different initial battery

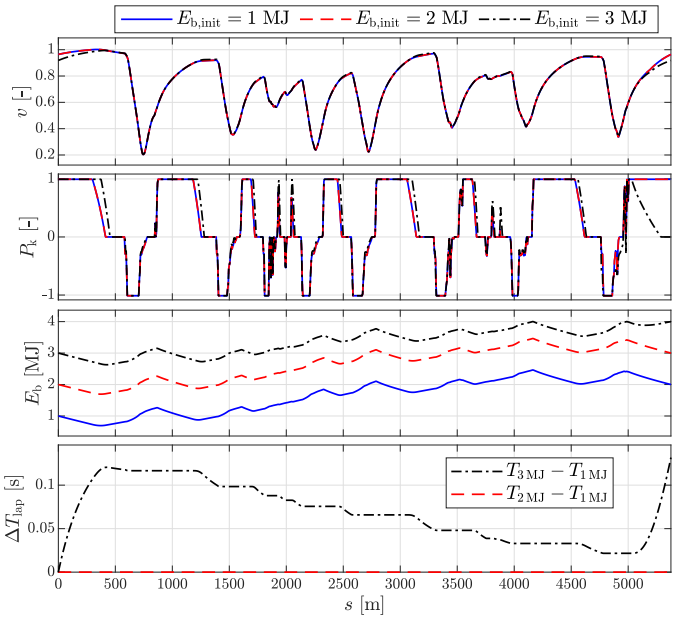

Fig. 6. Solution of Problem 2 for three different values of initial battery energy $E_{\mathrm{b,init}}$. The energy management strategy for the 1 MJ and 2 MJ cases is identical, whilst it differs for the 3 MJ due to the battery path constraint (41) becoming active, as the upper bound is hit. The bottom subplot shows the accumulated lap time difference.

energy levels. For the cases of 1 MJ and 2 MJ, the energy management strategy is literally identical, as the battery energy trajectory is simply shifted[1]. Conversely, with a high initial battery energy content of 3 MJ, the operating strategy of the MGU-K has to be altered as the battery path constraint (41) becomes active three times towards the end of the lap. Hence, less electric boosting is available on the last straight, and lap time increases by almost 0.15 s compared to the other two cases. This is significant in the context of circuit racing in general and F1 in particular, confirming that we must take this effect into account. More information on this can be found in [53], [54].

V. LAP TIME MAPS

Next we describe how to condense the aforementioned trends into lap time maps as a function of the key parameters fuel and battery energy allocation, car mass and initial battery energy. By solving Problem 2 for different values of $\{\Delta E_{\mathrm{f}}, \Delta E_{\mathrm{b}}, m, E_{\mathrm{b,init}}\}$, the optimal lap time for combinations of these features can be obtained. Our goal is to find a suitable fit of this data in the form of a track-specific lap time map $\mathscr{M}$, represented by feedforward ANNs:

$$T_{\mathrm{lap}} = \mathscr{M}(\Delta E_{\mathrm{f}}, \Delta E_{\mathrm{b}}, m, E_{\mathrm{b,init}}) \tag{56}$$

Since the five-dimensional space spanned by lap time as a function of the four features cannot be represented graphically, we show three- and two-dimensional plots of different trends in this section. In total, we generate 17880 data points. The

[1]Note that none of the state dynamics in the optimal control problem depend on E_{b}, hence the absolute level does not make any difference, as long as the 0 MJ and 4 MJ bounds on the battery energy are not hit.

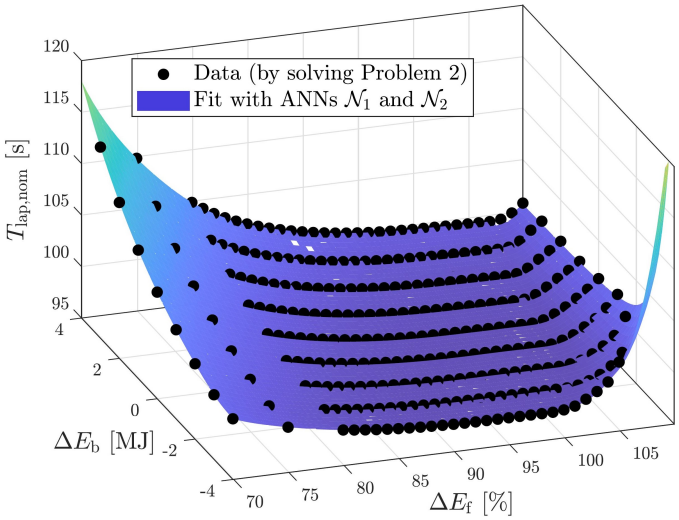

Fig. 7. Map for nominal lap time $T_{\mathrm{lap,nom}}$ as a function of fuel and battery energy allocation, represented for a fixed car mass m. The data is obtained by solving Problem 2 without the battery path constraints (40), (41), i.e., assuming a battery of infinite size. The ANN fit corresponds to (58).

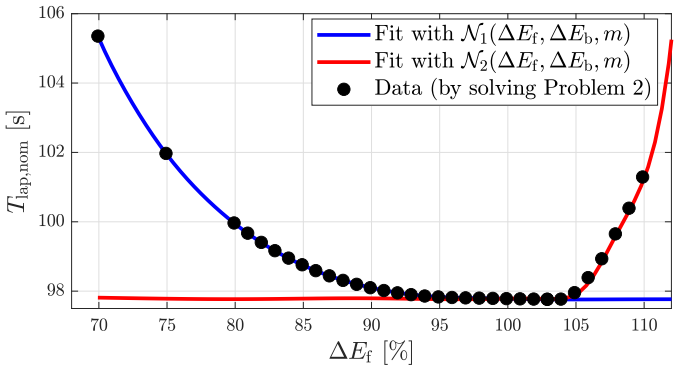

Fig. 8. Detailed view of nominal lap time as a function of fuel allocation, for a fixed car mass m and battery energy allocation ΔE_{b}. The fit with two ANNs $\mathcal{N}_1$ and $\mathcal{N}_2$ as proposed in (58) is shown. The view corresponds to a cross-section of Fig. 7 in the $\Delta E_{\mathrm{f}}/T_{\mathrm{lap,nom}}$ plane at $\Delta E_{\mathrm{b}} = 0$ MJ.

data is densely spaced for feature values which, from experience, are relevant in many practical race scenarios, and more scarcely distributed in other regions. It covers the complete possible range of features. This is important, since ANNs as a regression tool perform well when interpolating, but rather less so when extrapolating beyond the range of features on which they were trained [55].

Starting from (56), we subdivide the lap time into two components: the nominal lap time $T_{\mathrm{lap,nom}}$, fitted to data obtained by solving Problem 2 without the battery path constraints (40), (41), and a correction term $\Delta T_{\mathrm{lap,b}}$ that models the lap time increase when the battery energy bounds are hit during a lap:

$$T_{\mathrm{lap}} = T_{\mathrm{lap,nom}}(\Delta E_{\mathrm{f}}, \Delta E_{\mathrm{b}}, m) + \Delta T_{\mathrm{lap,b}}(E_{\mathrm{b,init}}, \Delta E_{\mathrm{b}}). \tag{57}$$

When directly fitting the lap time to all four features at once, we saw overfitting effects that heavily influenced the solution of Problem 1. Moreover, the penalty incurred due to the battery bounds is almost independent of ΔE_{f} and m. Hence we propose the additive separation approach (57).

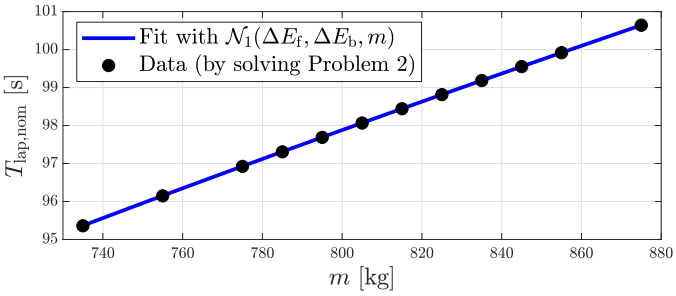

Fig. 9. Nominal lap time as a function of the car mass m, for a fixed battery and fuel energy allocation ΔE_f and ΔE_b.

An excerpt of the generated data for the nominal lap time is displayed in Fig. 7, for a fixed car mass. Generally, the more the battery can be discharged, and the more fuel can be used, the lower the lap time. However, if an overly large quantity of fuel has to be burned, this is not true anymore. We visualize this phenomenon more explicitly in Fig. 8. For fuel allocation values $\Delta E_\mathrm{f} > 105\,\%$, the lap time increases drastically. At this point, the energy management strategy in terms of fuel deployment is already as aggressive as it can be and the quantity of fuel that can be burned by keeping the engine running during braking maneuvers is limited by the brake balance constraints (see Section IV-D). However, the fuel allocation has to be completely used up by the end of the lap, due to the *equality* constraint (36). The optimal solution thus compromises lap time in order to fulfill the fuel constraint, e.g., by delaying the MGU-K power deployment to achieve a lower velocity on the straights, meaning that more time passes during which fuel can be burnt. In practice, besides fully exploiting the brake balance to burn fuel, the engine operation could be altered towards lower efficiency, for example by ignition retardation. The complex mathematical descriptions of the engine that are necessary to capture this effect [12] are computationally intractable for the high-level application at hand (optimizing a single lap takes several hours). We therefore do not consider the effect of ignition retardation here.

The dependency of the nominal lap time on car mass is shown in Fig. 9, for a fixed energy allocation. The trend is almost linear. The averaged mass sensitivity values inferred from the data are given in Table I for the Bahrain circuit and, in addition, for the Spa-Francorchamps circuit, which is by far the longest in the F1 calendar. As one would have expected intuitively, the sensitivity is higher for the longer circuit. Moreover, the obtained value for Bahrain is similar to the ones mentioned in literature for F1 cars (0.040 s/kg in [33], 0.033 s/kg in [34], and 0.030 s/kg in [56]). For the intended purpose, this validates the model described in Section IV.

TABLE I
SENSITIVITY OF LAP TIME W.R.T. CAR MASS

Circuit	Length	Sensitivity
Bahrain	5412 m	0.0351 s/kg
Spa-Francorchamps	7004 m	0.0528 s/kg

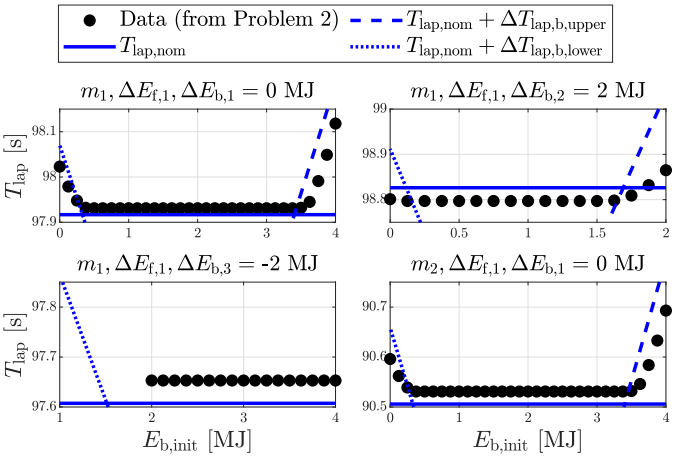

Fig. 10. Lap time as a function of $E_\mathrm{b,init}$, for a fixed car mass m and battery and fuel energy allocation. The plot shows the data obtained by solving Problem 2 with the battery path constraints (40), (41). The fit consists of the nominal lap time augmented by the affine correction described in (59), (60).

We fit the nominal lap time using two ANNs $\mathcal{N}_1$ and $\mathcal{N}_2$ in the following manner:

$$T_\mathrm{lap,nom} = \max\{\mathcal{N}_1(\Delta E_\mathrm{f}, \Delta E_\mathrm{b}, m), \mathcal{N}_2(\Delta E_\mathrm{f}, \Delta E_\mathrm{b}, m)\}. \quad (58)$$

Each ANN consists of 1 layer with 20 neurons and uses the tanh activation function, making it a continuous and differentiable function. The structure of the ANNs is as described in [12]. We use the Deep Learning Toolbox of MATLAB to train and implement the ANNs. The data points are randomly partitioned, with 75 % of the points used for training and 25 % for validation. The resulting fit can be appreciated in Figures 7, 8 and 9. Indeed, the two ANNs serve to fit the two 'branches' of the lap time trend, as is shown in Fig. 8: $\mathcal{N}_1$ is used for the decreasing trend, where more fuel equals to a reduction in lap time, whilst $\mathcal{N}_2$ fits the region where lap time increases because burning all the fuel requires measures that are detrimental to lap time. We propose the split approach (58) for the following reason: When experimenting with a single ANN to fit $T_\mathrm{lap,nom}$, we observed very small 'ripples' in the flat part of the lap time map (values of ΔE_f between 95 % and 105 % in Fig. 8). The solution of Problem 1 then exploited these local lap time minima, which were caused by the fit and did not reflect the data trends. Fitting two separate ANNs resolved this issue.

An excerpt of the data obtained by solving Problem 2 including the battery path constraints is shown in Fig. 10. As can be seen in the top left and bottom right subplots, if the initial battery energy content $E_\mathrm{b,init}$ is close to the upper or lower limit, the lap time deviates substantially from the nominal lap time with the same battery energy allocation $\Delta E_{\mathrm{b},1}$. In the top right subplot, the range of feasible values for $E_\mathrm{b,init}$ only spans up to 2 MJ. For higher $E_\mathrm{b,init}$, the battery energy at the end of the lap would exceed the maximum value of 4 MJ, given that the charge target for these data points is $\Delta E_{\mathrm{b},2} = 2$ MJ. Analogously, regarding the data points in the bottom left subplot, the lowest feasible value is $E_\mathrm{b,init} = 2$ MJ, given that the battery must be discharged by $\Delta E_{\mathrm{b},3} = -2$ MJ. In this scenario and on this particular racetrack, the battery

only becomes empty at the very end of the lap, such that no lap time deterioration is incurred.

We observed that fitting these trends with a regressor resulted in an underestimation of the lap time increase for some of the data points. The solution of Problem 1 then lay too close to the battery bounds in some laps, as the optimizer was 'unaware' of the lap time penalty that its strategy would incur. Further analysis revealed that the data points affected by the battery energy limits form clusters that can be separated from the points with nominal lap time. We therefore propose to use a classifier to capture the impact of the battery energy limits. For the sake of brevity, we only give a short summary of the procedure: We describe the effect of the lower and the upper battery energy limit with two correction terms $\Delta T_{\text{lap,b,lower}}$ and $\Delta T_{\text{lap,b,upper}}$, which are an affine function of $E_{\text{b,init}}$:

$$\Delta T_{\text{lap,b,lower}} = \kappa_{\text{lower}} \cdot E_{\text{b,init}} + \zeta_{\text{lower}}(\Delta E_{\text{b}}), \tag{59}$$

$$\Delta T_{\text{lap,b,upper}} = \kappa_{\text{upper}} \cdot E_{\text{b,init}} + \zeta_{\text{upper}}(\Delta E_{\text{b}}). \tag{60}$$

Whilst κ_{lower}, κ_{upper} are constants, the offset terms ζ_{lower}, ζ_{upper} depend on ΔE_{b}. For reasons of computational tractability, we neglect the small influence of the fuel energy allocation and the car mass on the lap time correction term. The parameters κ_{x} and ζ_{x}, with $\text{x} \in \{\text{lower}, \text{upper}\}$, describe a linear Support Vector Machine (SVM) [57] that separates the data points in the $E_{\text{b,init}}/\Delta T_{\text{lap,b,x}}$ plane into two categories: Those that are affected by the battery energy path constraints, and those that are not, i.e., whose lap time is equal to $T_{\text{lap,nom}}$. The SVMs are derived from a maximum margin formulation and are trained using the hinge loss function [58]. By applying different weights to the data points in the loss function, we can fine-tune the behavior, as will be explained below. The dependency of ζ_{x} on ΔE_{b} is again fitted with two ANNs:

$$\zeta_{\text{x}} = \mathscr{N}_{\text{x}}(\Delta E_{\text{b}}) \quad \text{for x} \in \{\text{lower}, \text{upper}\}. \tag{61}$$

Finally, we combine (59) and (60) to formulate the correction term accounting for the battery path constraints as

$$\begin{aligned} \Delta T_{\text{lap,b}} = \max\{0, &\Delta T_{\text{lap,b,lower}}(E_{\text{b,init}}, \Delta E_{\text{b}}), \\ &\Delta T_{\text{lap,b,upper}}(E_{\text{b,init}}, \Delta E_{\text{b}})\}. \end{aligned} \tag{62}$$

Its effect on the lap time map can be seen in Fig. 10. In the SVM loss function, we weighted the data points affected by the battery energy limits more than those unaffected. The identified parameters for (59), (60) thus result in a very conservative correction term, overestimating the lap time increase. In Problem 1, it provides an incentive to operate the battery at a reasonable distance from its energy bounds. During a race, such a behavior is desirable from a robustness point of view. Regarding the bottom left subplot, we note that the correction term only applies in the infeasible region. It will thus not impact the solution of Problem 1, since these combinations $\{E_{\text{b,init}}, \Delta E_{\text{b,3}}\}$ are inadmissible by virtue of the constraints (10), (11).

In terms of fitting quality, for the nominal lap time component the root mean square error with respect to the data obtained by solving Problem 2 is 0.063 s, or 0.06 % (with respect to the mean lap time in the data). For the complete lap time map including the data points affected by the battery path constraints, the root mean square error is 0.272 s, or 0.28 %. The higher deviation is due to the conservative correction term $\Delta T_{\text{lap,b}}$. Overall, the precision of the fitted lap time map $\mathscr{M}$ is satisfying. Most importantly, it captures the trends correctly, making it suitable for use in Problem 1.

Regarding the implementation of (2) in the NLP framework of Problem 1, we relax the non-smooth expressions to inequality constraints, i.e., for (58), we implement

$$\begin{aligned} T_{\text{lap,nom}} &\geq \mathscr{N}_1(\Delta E_{\text{f}}, \Delta E_{\text{b}}, m), \\ T_{\text{lap,nom}} &\geq \mathscr{N}_2(\Delta E_{\text{f}}, \Delta E_{\text{b}}, m), \end{aligned} \tag{63}$$

and we proceed similarly for (62) by implementing

$$\begin{aligned} &\Delta T_{\text{lap,b}} \geq 0, \\ &\Delta T_{\text{lap,b}} \geq \Delta T_{\text{lap,b,lower}}(E_{\text{b,init}}, \Delta E_{\text{b}}), \\ &\Delta T_{\text{lap,b}} \geq \Delta T_{\text{lap,b,upper}}(E_{\text{b,init}}, \Delta E_{\text{b}}). \end{aligned} \tag{64}$$

Combining (57) with (63) and (64) is identical to writing

$$T_{\text{lap}} \geq \mathscr{M}(\Delta E_{\text{f}}, \Delta E_{\text{b}}, m, E_{\text{b,init}}). \tag{65}$$

Since the objective (1) is to minimize the sum of lap times, (65) will hold with equality in the optimal solution.

VI. RESULTS

In this section, we discuss the results obtained by solving Problem 1 with the maps generated in Section V. First, we validate the approach of separating the minimum-race-time problem from the single-lap energy management problem, followed by a critical discussion of the computational time. Second, we compare the optimal energy allocation solution to a simple heuristic strategy. Third, we show how the optimization framework can be used to determine the optimal fuel load to be put into the car at the start of a race. Fourth, we leverage Problem 1 for online control during a race by implementing it as an MPC.

For all the following comparisons, the lap times T_{lap} associated with the energy allocation strategies are not simply taken from the lap time maps, but are rather computed in post-processing by evaluating Problem 2 for each lap i of the race with the obtained $\{\Delta E_{\text{f}}[i], \Delta E_{\text{b}}[i], m[i], E_{\text{b,init}}[i]\}$. This is important, since a comparison of lap times and total race times calculated from the lap time maps might be falsified by the imprecision of the fits, leading to wrong conclusions. Conversely, evaluating the optimal lap times using Problem 2 enables a fair comparison of different energy allocation strategies.

To compare two different strategies x and y, we define for each lap i the difference in total car mass

$$\Delta m[i] = m_{\text{y}}[i] - m_{\text{x}}[i], \tag{66}$$

(reflecting how the different fuel energy allocation affects the car mass), the difference in lap time

$$\Delta T_{\text{lap}}[i] = T_{\text{lap,y}}[i] - T_{\text{lap,x}}[i], \tag{67}$$

and the accumulated difference in total race time

$$\Delta T_{\text{race}}[i] = \sum_{j=1}^{i-1} T_{\text{lap,y}}[j] - T_{\text{lap,x}}[j]. \tag{68}$$

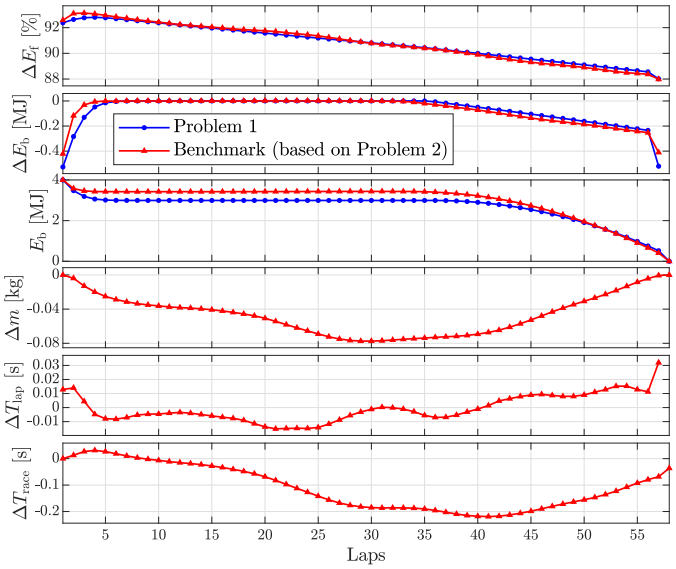

Fig. 11. Validation of the minimum-race-time approach formulated in Problem 1 by comparing it with the benchmark solution obtained by solving Problem 2 not only for a single lap, but for an entire race distance. In the three bottom plots, the differences are defined as $\Delta :=$ Benchmark $-$ Problem 1.

The presented results are for the Bahrain International Circuit, where a F1 race covers $N = 57$ laps and typically lasts for $T_{\text{race}} \approx 1.5$ hours. Problem 1 is implemented in MATLAB using CasADi and solved with IPOPT. For the sake of simplicity, we neglect the effect of the race start in our case studies and use the standard lap time map with the periodicity constraint (21) also in lap 1. A standing start could be considered with a specific lap time map for that scenario.

A. Validation of the optimization approach

Before discussing the strategies obtained by solving Problem 1, we must verify that the lap-by-lap discretization and the reliance on lap time maps that are subject to some fitting imprecision do not affect the optimality of the solution. A straightforward alternative way of computing a benchmark solution consists in solving Problem 2 not only for a single lap, but over an entire race distance, with the path variable $s \in [0, S_{\text{race}}]$ and $S_{\text{race}} = N \cdot S$. Due to the precise modeling equations and the finely resolved spatial discretization, this approach is much more precise, but at the expense of computational tractability. Indeed, for $\Delta s = 5$ m the solver did not even converge. To reduce the number of optimization variables, the step size had to be increased to 10 m, and the brake balance constraints presented in Section IV-D had to be removed. For one carefully chosen scenario, we then managed to obtain a sensible benchmark solution.

The comparison with the solution of Problem 1 is shown in Fig. 11. The two strategies in terms of fuel and battery energy allocation are almost identical, and the resulting difference in car mass is negligible. It can be seen that the solution of Problem 1 settles at a lower battery energy between lap 5 and 40. This is due to the conservative correction term $\Delta T_{\text{lap,b}}$ in the lap time maps, which penalizes an operation close to the battery bounds. In the benchmark problem, the impact of the battery energy bounds is captured exactly, which is exploited in the solution. Despite this effect, the differences in lap time are small (mostly less than 10 ms) and the advantage of the benchmark solution regarding the total race time is only 36 ms. Hence, the lap-by-lap discretization and the maps that form the basis of our minimum-race-time problem do not incur any noteworthy sub-optimality, thereby validating the proposed approach.

B. Discussion of the computational time

The computational time for solving Problem 1 was roughly 1 s on a standard consumer laptop with a 4 GHz processor, whilst computing the benchmark solution took more than one hour. Moreover, solving the problem with the benchmark approach is numerically unstable, generally fails and only works with simplifications for very specific scenarios. However, Problem 1 also comes with a drawback. Whilst the training of the ANNs for the lap time maps takes less than 2 minutes, the generation of the data points is computationally expensive. Solving multiple instances of Problem 2 can be done using parallel computing, but it still took roughly 20 hours on 4 processor cores to generate the 17880 data points required. Thus, our approach shifts the computational bottleneck from solving the minimum-race-time problem to creating the lap time maps. We deem this advantageous for two reasons: First, the method is much more robust and reliable than the benchmark approach discussed in the previous section. Second, once the lap time maps are generated (which can be done off-line before a race weekend), Problem 1 can be solved with great ease to investigate different race scenarios, conduct case studies, such as in Section VI-D, and even for online control during a race, as in Section VI-E. The benchmark approach of Section VI-A is intractable for such applications.

C. Optimal solution vs. heuristic strategy

Next, we compare the solution of Problem 1 to a heuristic strategy, in order to understand the potential in terms of race time and to analyze some characteristic features of the optimal strategy. For the sake of simplicity and to facilitate the interpretation of the results, we do not consider pit stops in this case study. The heuristic strategy is a simple approach that one might choose if no optimization tools whatsoever are at disposal: Both the available fuel and the battery energy are distributed equally over the entire race, i.e.,

$$\begin{aligned} \Delta E_{\text{f}}[i] &= \frac{m_{\text{f}} \cdot H_{\text{lhv}}}{N} \qquad \forall\, i, \\ \Delta E_{\text{b}}[i] &= \frac{4\,\text{MJ}}{N} \qquad \forall\, i. \end{aligned} \tag{69}$$

This means that the allocated energy is the same in each lap.

The comparison is shown in Fig. 12. Regarding the fuel allocation, it is optimal to use relatively more fuel at the beginning of the race and relatively less fuel towards the end. This means that it is advantageous to render the car lighter at an earlier stage of the race. The allocated fuel energy of $\Delta E_{\text{f}} \approx 105\,\%$ in the first five laps corresponds to the maximum amount of fuel that can be consumed without incurring a lap

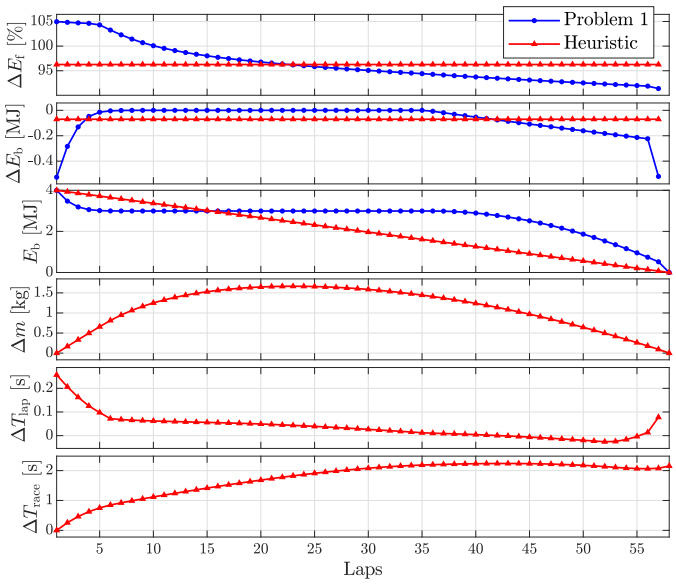

Fig. 12. Comparison between the optimal energy allocation strategy obtained as a solution to Problem 1 and a heuristic strategy. The latter consists in allocating the same battery and fuel energy to every lap of the race. In the three bottom plots, the differences are defined as $\Delta :=$ Heuristic $-$ Problem 1.

time increase, as shown in Fig. 7 and Fig. 8. As one could expect intuitively, it is not optimal to penalize lap time in order to waste fuel. The optimal battery allocation strategy also strongly differs from the heuristic: The battery is quickly depleted to an energy content of roughly 3 MJ in lap 5. At this level, the operation is not affected by the battery energy bounds anymore. From there on, the optimal strategy consists in a charge-sustaining operation until lap 36, and afterwards, it is slowly discharged. The discharge strategy is such that not even the last lap is affected by the battery energy bounds, and thus not compromises lap time. At the beginning of the race, a lot of lap time is lost with the heuristic strategy because the battery energy bounds constrain the power unit operation. After lap 23, lap times are still higher, even though the fuel and battery energy allocated in these laps is larger than in the optimal solution. This occurs because the car is lighter with the optimal fuel allocation strategy, as is evidenced by the difference in car mass, which is more than 1 kg at this stage of the race. Only towards the end of the race does the heuristic strategy gain some time, but it does not make up for the losses incurred previously. The advantage in terms of total race time of the optimal strategy accumulates to over 2 s, or 35 ms per lap on average. Considering that the overall energy consumption is identical and that the PU operation within each lap is optimal in both scenarios, the difference stems solely from a different allocation of the energy over the race. Therefore, the improvement is significant.

D. Determining the optimal fuel load

One important parameter which race strategy engineers have to define before the start of the race is the amount of fuel that is put into the car. There is a trade-off at play: Filling the tank with less fuel renders the car lighter and thus faster, but PU operation might be compromised and the driver might be

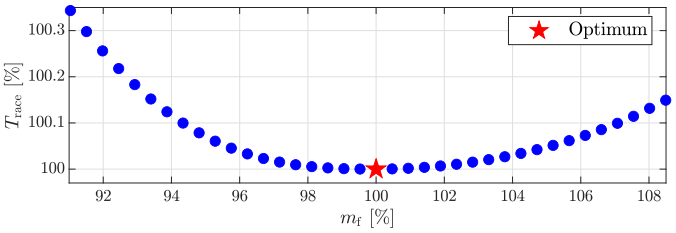

Fig. 13. Total race time computed by solving Problem 1 as a function of the fuel load at the start of the race. The trade-off between mass of the car and restrictions on PU operation is clearly visible. For reasons of confidentiality, m_f and T_{race} have been normalized, with 100 % denoting the optimal values.

forced to save fuel by lifting the throttle pedal at the end of the straights ('lift-and-coast'). Vice-versa, filling the tank with more fuel results in a heavier and thus slower car, but no fuel saving is required. The proposed optimization framework can be used to determine the optimal fuel load and thereby assist the decision process. To this end, Problem 1 is solved for a sweep of different fuel load values m_f. This is done easily, given that the procedure is computationally tractable. Fig. 13 shows the optimal race time as a function of m_f. The trade-off described above is well visible. For this particular car and circuit, deviating from the optimal fuel load by $\pm 5\,\%$ increases total race time by circa 0.05 %, which is in the order of several seconds and thus substantial. Conversely, the trade-off behavior closely around the optimum value is very flat, meaning that a slight mismatch in fuel load does not entail a large performance loss.

E. Online application in the form of an MPC

In our final case study, we show how the presented optimization framework can be leveraged to update the energy allocation strategy online during a race. This is necessary in the case of events such as pit stops or overtake maneuvers. During a pit stop, the driver goes through the pit lane, where he must adhere to a speed limit. Since the pit lane runs parallel to the start-finish straight, this affects the energy management and the achievable lap time in the lap before the pit stop ('in-lap') and in the one after the pit stop ('out-lap'). Thus, special lap time maps $\mathscr{M}_{\text{in-lap}}$ and $\mathscr{M}_{\text{out-lap}}$ have to be used for the concerned laps in the energy allocation optimization. Details on the modeling of the pit stop scenario are given in Appendix A. Another disturbance that can realistically occur during a race is an overtake maneuver. In order to pass an opponent, the driver chooses a higher fuel and battery energy allocation than planned and thereby achieves a higher velocity on the straights, increasing his chances to overtake. If, for instance, the battery is completely depleted afterwards, the energy allocation strategy must be adapted for the remainder of the race.

We slightly modify Problem 1, such that in lap k, it can be solved over the remaining race laps $i \in \{k,\dots,N\}$, starting from the measured battery and fuel energy $E_{\text{f,measured}}$ and $E_{\text{b,measured}}$, and based on the available information regarding the applicable lap time maps. This yields:

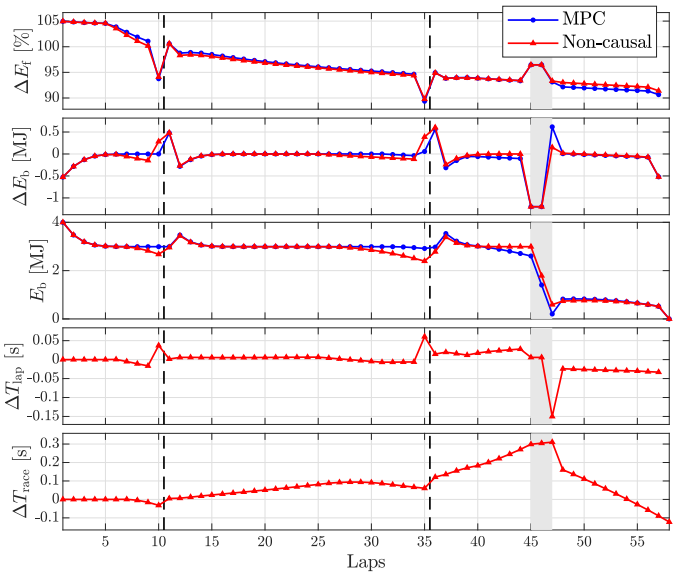

Fig. 14. Online case study for a race with two pit stops (between laps 10-11 and laps 35-36, marked by the dashed line), followed by a battle for position in laps 45-46 (shaded in gray), during which the driver chooses to use more fuel and deplete the battery almost completely. The simulation results with the shrinking-horizon MPC controller are compared to the non-causal solution, which is obtained by solving Problem 1 with perfect knowledge of the pit stops and overtake maneuver. For the two bottom plots, $\Delta :=$ Non-causal $-$ MPC.

Problem 3. *Assuming that at the start of lap k, the total fuel energy consumption $E_{\text{f,measured}}$ and the battery energy level $E_{\text{b,measured}}$ are measured, the optimal energy allocation strategy for the remainder of the race is the solution of*

$$\min_{\Delta E_\text{f}, \Delta E_\text{b}} \quad \sum_{i=k}^{N} T_\text{lap}[i]$$

subject to the following constraints $\forall i \in \{k, \dots, N\}$:

$$\begin{aligned}
&\textit{Lap time:} && (2),\ \mathscr{M}_i = \mathscr{M}_{\text{in/out-lap}} \quad \text{for} \quad i \in \{\text{in/out-lap}\} \\
&\textit{Fuel:} && E_\text{f}[k] = E_{\text{f,measured}},\ (3),(4),(6), \\
&\textit{Battery:} && E_\text{b}[k] = E_{\text{b,measured}},\ (7),(9),(10),(11), \\
&\textit{Car mass:} && m[k] = m_\text{car} + m_\text{f} - \frac{E_{\text{f,measured}}}{H_\text{lhv}},\ (12),(13).
\end{aligned}$$

Solving Problem 3 in each lap $k \in \{1, \dots, N\}$ of the race and applying the energy allocations $\Delta E_\text{f}[k]$ and $\Delta E_\text{b}[k]$ obtained for the first lap of the control horizon leads to a so-called shrinking-horizon MPC [59], [60].

Fig. 14 shows the simulation results with the proposed causal MPC implementation and compares them to the optimal, non-causal strategy obtained by solving Problem 1. The total energy consumption is the same in both cases. Our case study features two pit stops occurring between laps 10-11 and laps 35-36, respectively, as well as a disturbance in the form of an overtake maneuver in laps 45-46. The non-causal solution is obtained with perfect a priori knowledge of the laps affected by the pit stops and the overtake maneuver. The MPC, on the other hand, is informed of the pit stops only at the start of the respective in-lap and is unaware of the overtake maneuver, during which its outputs for ΔE_f and ΔE_b are overwritten. It only 'detects' the overtake a posteriori via the measurements $E_{\text{f,measured}}$, $E_{\text{b,measured}}$ when standard MPC operation is resumed at the start of lap 47, and it reacts by recharging the battery during that lap.

It is apparent that the causal and the non-causal strategy are very similar. The difference in total race time only amounts to about 0.1 s in favor of the non-causal solution. We observe that the laps affected by the pit stops require a very particular energy allocation: The fuel consumption in the in-lap is markedly lower than normal, as the car slows down at the end of the lap and drives through the pit lane at the prescribed speed limit. By contrast, in the out-lap, more fuel has to be consumed, because the car has to accelerate towards the first corner from a very low velocity after exiting the pit lane. The battery is substantially recharged while traversing the pit lane at constant speed, via load-point shifting of the engine. This is visible both in the in-lap and in the out-lap, since the pit lane passage affects both laps. The small differences between the MPC and the non-causal strategy stem from the fact that the non-causal strategy can 'prepare' for the pit stops. This is best explained by looking at the battery energy E_b in laps 30-35: Given the knowledge that a pit stop will occur, it is optimal to deplete the battery somewhat, since it can be recharged in the pit lane. The MPC cannot plan this, since it does not know that a pit stop is coming. Instead, it prescribes a charge-sustaining trajectory, which would be optimal in a race without pit stops (compare Fig. 12). During the overtake maneuver, a similar effect is noticed. The causal MPC has started to gradually discharge the battery, whereas the non-causal strategy, knowing that an overtake with strong battery depletion will occur, keeps the battery energy at a higher level. The MPC strategy therefore recharges the battery by 0.5 MJ after the overtake, which brings the battery energy back to the optimal level for the remaining laps, but causes a lap time loss of 0.15 s compared to the non-causal strategy in lap 47.

Overall, the sub-optimality incurred by reacting to such events in a causal manner is very small, highlighting that the proposed race optimization works well in an MPC context and reacts adequately to disturbances. The results also indicate that it is important to implement an energy allocation strategy that captures the general trends observed in Fig. 12. An optimization-based re-planning can be carried out in the case of unforeseen events, e.g., with the proposed MPC, without markedly penalizing the total race time.

VII. CONCLUSION

In this paper, we have presented an optimization framework that solves the minimum-race-time problem for the hybrid-electric F1 race car based on a lap-by-lap discretization, by separating it from the single-lap optimization problem through the use of lap time maps. We benchmarked the obtained solution with the strategy directly computed based on the detailed model, without relying on lap time maps. This highlighted the precision of our ANN-based map fitting approach, as well as the computational efficiency of the lap-by-lap optimization framework. A comparison to a heuristic strategy underlined the importance of carefully choosing the

energy allocation during a race: Just by optimizing the amount of fuel and battery energy used in each lap, the total race time could be improved by 2 s. Whilst this gain was found for a simplified race scenario without any pit stops or competitor interaction, the comparison clearly highlighted the trade-offs at play and demonstrated the potential of the approach. Two different case studies displayed possible applications of the proposed framework. First, it can be used to determine the optimal fuel load at the start of the race. We found that closely around the optimal value, the sensitivity of total race time towards the fuel load is small. However, slightly larger deviations of $\pm 5\,\%$ from the optimal fuel load increase total race time by several seconds, underlining the importance of optimizing this strategic parameter. Second, due to its computational tractability, the framework can be used as an MPC with only minor modifications. Our results showed that it reacts in a close-to-optimal fashion to common disturbances that can occur during a race.

Further research could focus on integrating a suitable tire degradation model in the optimization framework. Given that the tire degradation also depends on the vehicle mass, this might yield interesting cross-couplings and amplify the impact of the fuel allocation. Moreover, the MPC formulation could be extended towards stochastic models for, e.g., the probability of competitor interaction or a pit stop event. Lastly, it could be of interest for racing teams to combine the framework with pit stop and tire strategy simulation and optimization tools.

ACKNOWLEDGMENT

We thank Ferrari S.p.A. for supporting this project. Moreover, we are grateful to Dr. Ilse New for the proofreading, and to Mr. Tammo Zobel for his helpful comments.

APPENDIX

A. Modeling of pit stops

By making a pit stop during a race, a driver can get new tires fitted [4]. The driver enters the pit lane, which normally runs parallel to the main straight of the track, and stops in front of her/his team's garage. The crew members then change all four tires in roughly 2.5 s and the driver departs again. A pit stop therefore consists of the so-called 'in-lap', towards the end of which the car enters the pit lane, and the 'out-lap', whose first part is not driven on the main straight, but still in the pit lane, until regaining the race track. In the pit lane, the driver must respect a speed limit $v_{\max,\text{pit}}$, usually 80 km/h. For the proof-of-concept of our approach, and due to a lack of data, we do not model the different path taken to enter the pit garage, and we also neglect the fact that the car comes to a halt and remains briefly stationary. We only take into account the speed limit by imposing it as an additional constraint to Problem 2 on the part of the main straight that runs parallel to the pit lane:

$$\text{In-lap:} \quad v(s) \leq v_{\max,\text{pit}} \quad \forall s \in [s_{\text{pit,entry}}, S]. \tag{70}$$

$$\text{Out-lap:} \quad v(s) \leq v_{\max,\text{pit}} \quad \forall s \in [0, s_{\text{pit,exit}}]. \tag{71}$$

With the corresponding additional constraint, Problem 2 can be solved for the in- or out-lap case, allowing to create specific lap time maps $\mathscr{M}_{\text{in-lap}}$ and $\mathscr{M}_{\text{out-lap}}$ for the pit stop scenario by applying the method outlined in Section V. These can then be used in the race optimization given by Problem 1 or in the MPC approach described in Problem 3 for the laps where a pit stop is scheduled.

REFERENCES

[1] A. Sciarretta and L. Guzzella, *Vehicle Propulsion Systems*, 3rd ed. Springer, 2013.
[2] FIA, "2021 Formula One Technical Regulations," Fédération Internationale de l'Automobile, Tech. Rep., 2021.
[3] ——, "2014 Formula One Sporting Regulations," Fédération Internationale de l'Automobile, Tech. Rep., 2014.
[4] ——, "2021 Formula One Sporting Regulations," Fédération Internationale de l'Automobile, Tech. Rep., 2021.
[5] S. Ebbesen, M. Salazar, P. Elbert, C. Bussi, and C. H. Onder, "Time-optimal control strategies for a hybrid electric race car," *IEEE Transactions on Control Systems Technology*, vol. 26, no. 1, pp. 233–247, 2018.
[6] A. Sciarretta and L. Guzzella, "Control of hybrid electric vehicles," *IEEE Control systems*, vol. 27, no. 2, pp. 60–70, 2007.
[7] FIA, "2021 Le Mans Hypercar Technical Regulations," Fédération Internationale de l'Automobile, Tech. Rep., 2021.
[8] D. Casanova, "On minimum time vehicle manoeuvring: The theoretical optimal lap," Ph.D. dissertation, Cranfield University, 2000.
[9] G. Perantoni and D. J. Limebeer, "Optimal control for a Formula One car with variable parameters," *Vehicle System Dynamics*, vol. 52, no. 5, pp. 653–678, 2014.
[10] A. Rucco, G. Notarstefano, and J. Hauser, "An efficient minimum-time trajectory generation strategy for two-track car vehicles," *IEEE Transactions on Control Systems Technology*, vol. 23, no. 4, pp. 1505–1519, 2015.
[11] W. West and D. Limebeer, "Optimal tyre management for a high-performance race car," *Vehicle System Dynamics*, pp. 1–19, 2020.
[12] C. Balerna, M.-P. Neumann, N. Robuschi, P. Duhr, A. Cerofolini, V. Ravaglioli, and C. Onder, "Time-optimal low-level control and gearshift strategies for the Formula 1 hybrid electric powertrain," *Energies*, vol. 14, p. 171, 2021.
[13] O. Borsboom, C. A. Fahdzyana, and M. Salazar, "Time-optimal control strategies for electric race cars with different transmission technologies," in *2020 IEEE Vehicle Power and Propulsion Conference (VPPC)*. IEEE, 2020, pp. 1–5.
[14] C. Yang, M. Zha, W. Wang, L. Yang, S. You, and C. Xiang, "Motor-temperature-aware predictive energy management strategy for plug-in hybrid electric vehicles using rolling game optimization," *IEEE Transactions on Transportation Electrification*, vol. 7, no. 4, pp. 2209–2223, 2021.
[15] D. T. Machacek, K. Barhoumi, J. M. Ritzmann, T. Huber, and C. H. Onder, "Multi-level model predictive control for the energy management of hybrid electric vehicles including thermal derating," *IEEE Transactions on Vehicular Technology*, 2022.
[16] A. Locatello, M. Konda, O. Borsboom, T. Hofman, and M. Salazar, "Time-optimal Control of Electric Race Cars under Thermal Constraints," in *2021 European Control Conference (ECC)*. IEEE, 2021, pp. 905–912.
[17] X. Liu, A. Fotouhi, and D. J. Auger, "Optimal energy management for formula-E cars with regulatory limits and thermal constraints," *Applied Energy*, vol. 279, p. 115805, 2020.
[18] T. Lipp and S. Boyd, "Minimum-time speed optimisation over a fixed path," *International Journal of Control*, vol. 87, no. 6, pp. 1297–1311, 2014.
[19] S. Broere, J. Van Kampen, and M. Salazar, "Minimum-lap-time control strategies for all-wheel drive electric race cars via convex optimization," in *2022 European Control Conference (ECC)*. IEEE, 2022, pp. 1204–1211.
[20] P. Duhr, A. Sandeep, A. Cerofolini, and C. H. Onder, "Convex performance envelope for minimum lap time energy management of race cars," *IEEE Transactions on Vehicular Technology*, vol. 71, no. 8, pp. 8280–8295, 2022.
[21] R. Lot and S. A. Evangelou, "Lap time optimization of a sports series hybrid electric vehicle," in *2013 World Congress on Engineering*, 2013, pp. 1–6.

[22] D. J. Limebeer, G. Perantoni, and A. V. Rao, "Optimal control of Formula One car energy recovery systems," *International Journal of Control*, vol. 87, no. 10, pp. 2065–2080, 2014.

[23] T. Herrmann, F. Christ, J. Betz, and M. Lienkamp, "Energy management strategy for an autonomous electric racecar using optimal control," in *2019 IEEE Intelligent Transportation Systems Conference (ITSC)*. IEEE, 2019, pp. 720–725.

[24] L. Paparusso, M. Riani, F. Ruggeri, and F. Braghin, "Competitors-aware stochastic lap strategy optimisation for race hybrid vehicles," *IEEE Transactions on Vehicular Technology*, 2022.

[25] T. Herrmann, F. Passigato, J. Betz, and M. Lienkamp, "Minimum race-time planning-strategy for an autonomous electric racecar," in *2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, 2020, pp. 1–6.

[26] X. Liu and A. Fotouhi, "Formula-E race strategy development using artificial neural networks and monte carlo tree search," *Neural Computing and Applications*, vol. 32, no. 18, pp. 15 191–15 207, 2020.

[27] X. Liu, A. Fotouhi, and D. J. Auger, "Formula-e race strategy development using distributed policy gradient reinforcement learning," *Knowledge-Based Systems*, vol. 216, p. 106781, 2021.

[28] K. J. Hunt, D. Sbarbaro, R. Żbikowski, and P. J. Gawthrop, "Neural networks for control systems—a survey," *Automatica*, vol. 28, no. 6, pp. 1083–1112, 1992.

[29] C. M. Bishop and C. Roach, "Fast curve fitting using neural networks," *Review of scientific instruments*, vol. 63, no. 10, pp. 4450–4456, 1992.

[30] J. van Kampen, T. Herrmann, and M. Salazar, "Maximum-distance race strategies for a fully electric endurance race car," *European Journal of Control*, p. 100679, 2022.

[31] P. Aversa, L. Cabantous, and S. Haefliger, "When decision support systems fail: Insights for strategic information systems from Formula 1," *The Journal of Strategic Information Systems*, vol. 27, no. 3, pp. 221–236, 2018.

[32] T. Tulabandhula and C. Rudin, "Tire changes, fresh air, and yellow flags: challenges in predictive analytics for professional racing," *Big data*, vol. 2, no. 2, pp. 97–112, 2014.

[33] J. Bekker and W. Lotz, "Planning Formula One race strategies using discrete-event simulation," *Journal of the Operational Research Society*, vol. 60, no. 7, pp. 952–961, 2009.

[34] A. Heilmeier, M. Graf, and M. Lienkamp, "A race simulation for strategy decisions in circuit motorsports," in *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, 2018, pp. 2986–2993.

[35] A. Heilmeier, M. Graf, J. Betz, and M. Lienkamp, "Application of Monte Carlo methods to consider probabilistic effects in a race simulation for circuit motorsport," *Applied Sciences*, vol. 10, no. 12, p. 4229, 2020.

[36] A. Heilmeier, A. Thomaser, M. Graf, and J. Betz, "Virtual strategy engineer: Using artificial neural networks for making race strategy decisions in circuit motorsport," *Applied Sciences*, vol. 10, no. 21, p. 7805, 2020.

[37] A. Liniger and J. Lygeros, "A noncooperative game approach to autonomous racing," *IEEE Transactions on Control Systems Technology*, vol. 28, no. 3, pp. 884–897, 2019.

[38] S. Cacace, R. Ferretti, and A. Festa, "Stochastic hybrid differential games and match race problems," *Applied Mathematics and Computation*, vol. 372, p. 124966, 2020.

[39] F. Tagliaferri and I. M. Viola, "A real-time strategy-decision program for sailing yacht races," *Ocean Engineering*, vol. 134, pp. 129–139, 2017.

[40] E. F. Camacho and C. B. Alba, *Model predictive control*. Springer science & business media, 2013.

[41] H. A. Borhan, A. Vahidi, A. M. Phillips, M. L. Kuang, and I. V. Kolmanovsky, "Predictive energy management of a power-split hybrid electric vehicle," in *2009 American Control Conference*. IEEE, 2009, pp. 3970–3976.

[42] H. Borhan, A. Vahidi, A. M. Phillips, M. L. Kuang, I. V. Kolmanovsky, and S. Di Cairano, "MPC-based energy management of a power-split hybrid electric vehicle," *IEEE Transactions on Control Systems Technology*, vol. 20, no. 3, pp. 593–603, 2012.

[43] X. Li, L. Han, H. Liu, W. Wang, and C. Xiang, "Real-time optimal energy management strategy for a dual-mode power-split hybrid electric vehicle based on an explicit model predictive control algorithm," *Energy*, vol. 172, pp. 1161–1178, 2019.

[44] W. F. Milliken, D. L. Milliken *et al.*, *Race car vehicle dynamics*. Society of Automotive Engineers Warrendale, PA, 1995, vol. 400.

[45] P. Duhr, G. Christodoulou, C. Balerna, M. Salazar, A. Cerofolini, and C. H. Onder, "Time-optimal gearshift and energy management strategies for a hybrid electric race car," *Applied Energy*, vol. 282, p. 115980, 2020.

[46] L. Guzzella and C. H. Onder, *Introduction to Modeling and Control of Internal Combustion Engine Systems*. Springer-Verlag, 2010.

[47] M. Salazar, P. Duhr, C. Balerna, L. Arzilli, and C. H. Onder, "Minimum lap time control of hybrid electric race cars in qualifying scenarios," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 8, pp. 7296–7308, 2019.

[48] K. Kritayakirana and J. C. Gerdes, "Autonomous cornering at the limits: Maximizing a "gg" diagram by using feedforward trail-braking and throttle-on-exit," *IFAC Proceedings Volumes*, vol. 43, no. 7, pp. 548–553, 2010.

[49] H. G. Bock and K.-J. Plitt, "A multiple shooting algorithm for direct solution of optimal control problems," *IFAC Proceedings Volumes*, vol. 17, no. 2, pp. 1603–1608, 1984.

[50] T. Albin Rajasingham, "Nonlinear model predictive control," in *Nonlinear Model Predictive Control of Combustion Engines*. Springer, 2021, pp. 101–137.

[51] J. A. Andersson, J. Gillis, G. Horn, J. B. Rawlings, and M. Diehl, "Casadi: a software framework for nonlinear optimization and optimal control," *Mathematical Programming Computation*, vol. 11, no. 1, pp. 1–36, 2019.

[52] A. Wächter and L. T. Biegler, "On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming," *Mathematical programming*, vol. 106, no. 1, pp. 25–57, 2006.

[53] P. Duhr, M. Schaller, L. Arzilli, A. Cerofolini, and C. H. Onder, "Time-optimal energy management of the Formula 1 power unit with active battery path constraints," in *2021 European Control Conference (ECC)*. IEEE, 2021, pp. 913–920.

[54] P. Duhr, M. Schaller, L. Arzilli, A. Cerofolini, and C. Onder, "Analysis of optimal energy management strategies for the hybrid electric Formula 1 car under consideration of the finite battery size," in *FISITA 2021 World Congress (virtual)*, 2021.

[55] E. Barnard and L. Wessels, "Extrapolation and interpolation in neural network classifiers," *IEEE Control Systems Magazine*, vol. 12, no. 5, pp. 50–53, 1992.

[56] McLaren-Racing-Limited, "Formula One Race Strategy," Royal Academy of Engineering, Tech. Rep., visited on 15/02/2022. [Online]. Available: https://www.raeng.org.uk/publications/other/14-car-racing

[57] T. Hastie, R. Tibshirani, and J. H. Friedman, *The elements of statistical learning: data mining, inference, and prediction*. Springer, 2009, vol. 2.

[58] M. P. Deisenroth, A. A. Faisal, and C. S. Ong, *Mathematics for machine learning*. Cambridge University Press, 2020.

[59] M. M. Thomas, J. Kardos, and B. Joseph, "Shrinking horizon model predictive control applied to autoclave curing of composite laminate materials," in *Proceedings of 1994 American Control Conference-ACC'94*, vol. 1. IEEE, 1994, pp. 505–509.

[60] J. Skaf, S. Boyd, and A. Zeevi, "Shrinking-horizon dynamic programming," *International Journal of Robust and Nonlinear Control*, vol. 20, no. 17, pp. 1993–2002, 2010.