Optimal control of a turbocharged direct injection diesel engine by direct method optimization

Jose Manuel Luján, Carlos Guardiola, Benjamín Pla and Alberto Reig
e-mail: {jlujan,carguaga,benplamo,alreiber}@mot.upv.es

CMT-Motores Térmicos, Universitat Politècnica de València, Spain

Abstract

This work studies the effect and performance of an optimal control strategy on engine fuel efficiency and pollutant emissions. An accurate mean-value control-oriented engine model has been developed and experimental validation on a wide range of operating conditions was carried out. A direct optimization method based on Euler's collocation scheme is used in combination with the above model in order to address the optimal control of the engine. This optimization method provides the optimal trajectories of engine controls (fueling rate, exhaust gases recirculation valve position, variable turbine geometry position and start of injection) to reproduce a predefined route (speed trajectory including variable road grade), minimizing fuel consumption with limited NO_x emissions and a low soot stamp. This optimization procedure is performed for a set of different NO_x emission limits in order to analyze the tradeoff between optimal fuel consumption and minimum emissions.

Optimal control strategies are validated in an engine test bench and compared against engine factory calibration. Experimental results show that significant improvements in both fuel efficiency and emissions reduction can be achieved with optimal control strategy. Fuel savings at about 4% and less than half of the factory NO_x emissions were measured in the actual engine, while soot generation was still low. Experimental results and optimal control trajectories are thoroughly analyzed, identifying the different strategies that allowed those performance improvements.

This work was supported by Ministerio de Economía y Competitividad through Project TRA2016-78717-R.

1 Introduction

Internal Combustion Engines (ICEs) are experiencing a significant increase in complexity, especially regarding pollutant reduction devices and technologies. Nowadays, vehicles incorporate a large number of different systems aimed at reducing fuel consumption while fulfilling pollutant emissions regulations. Double Exhaust Gases Recirculation (EGR) loop, Variable Geometry Turbine (VGT), double or triple forced induction devices, Variable Valve Timing (VVT), Direct Injection (DI) systems with multiple injection events, Diesel Oxidation Catalyst (DOC), Diesel Particulate Filter (DPF) or Selective Catalytic Reduction (SCR) devices are some examples [1].

All these systems work coupled together, involving a challenging control problem, where the behavior of a single device influences the rest of the chain. Usually, actuations aiming for both objectives (minimizing fuel and pollutant emissions) go in opposite directions [2], so at the end it is a tradeoff. The control of such a complex nonlinear system requires of advanced algorithms and, eventually, a large batch of test cell experiments.

On the other hand, Optimal Control (OC) theory has become an attractive methodology to approach complex problems thanks to the advances in computational power of the last decades and the access to state-of-the-art computers for the general public. OC theory is a type of model-based control where different mathematical methods are used to calculate the controls that must be applied to a system to minimize a cost index [3]. It is already used in many different engineering fields such as aviation [4], train scheduling [5] or aerospace trajectory planning [6].

Despite of its advantages, OC is not yet established in the automotive industry, where traditional approaches, PID controllers and heuristic rules are the most common. However, several works may be found in literature where OC theory is used with successful results: [7] applies the OC theory to calculate the optimal behavior of a VGT during load transients finding that pressure may be built faster and smoother than with traditional PID controllers; [8] approaches the combined control of fueling rate, EGR and VGT opening to minimize fuel consumption for a prescribed driving cycle; [9] studies the possibility to optimally switch between different calibrations to fulfill emission limits along a route, minimizing fuel consumption compared to a single calibration; [10] calculates the advantage of optimally switching between low and high pressure EGR loops in a diesel engine analyzing the frequency and number of switches; [11] finds the optimal heat release law by applying Dynamic Programming (DP) to an in-cylinder pressure model to minimize fuel and pollutant emissions; Pontryagin Minimum Principle (PMP) is very popular among energy management and power-split in Hybrid Electric Vehicle (HEV) works with a wide variety of approaches such as [12–14].

This paper presents an OC strategy to manage controls in a DI turbocharged diesel engine, namely fueling rate, EGR valve position, VGT opening and Start of Injection (SOI), for a given driving cycle, minimizing fuel consumption with limited NO_x emissions and low soot generation. The main contributions of this work compared to the current state of the art are as follows:

- The application of OC to the simultaneous control of multiple engine actuators, removing the need of maps or ECU calibrations on the control scheme in favour of a powertrain physical model. Opposedly to other efficient and complex control algorithms in literature, an OC approach, such as the one in this study, provides mathematical guarantee of minimum fuel consumption and/or pollutant emissions for a given powertrain model and a set of constraints.

- The experimental benchmark for a long real driving cycle (25 minutes), covering multiple engine operating areas and transient conditions. Existing studies usually limit the application of OC to short transients (generally few seconds) and reduced engine operating areas.

- The use of Direct Method (DM) in combination with Euler's collocation scheme as the Ordinary Differential Equation (ODE) transcription method. Typically, higher order collocation schemes are used for DM but Euler's proved to be a better choice in terms of computational efficiency, being able to handle such a huge Optimal Control Problem (OCP) with reduced computation time while keeping a fairly good accuracy.

In order to reach the above targets, first, an experimental test campaign is performed with a diesel engine (see section 2) in order to fit a zero-dimensional Mean Value Engine Model (MVEM) (see section 3). A DM based on Euler's collocation scheme in combination with an Nonlinear Programming (NLP) solver is used to address the corresponding optimization problem, obtaining a set of trajectories to be applied to each of the engine controls (see section 4). An experimental validation is carried out on the same test bench with different NO_x emission limits (see section 5) in order to find the optimal tradeoff between fuel consumption and NO_x generation. Finally, after analyzing the behavior of the optimal trajectories, the main conclusions of the paper are drawn, showing that DM offers an approach to solve OCPs that are too large or too complex for DP or PMP, and that the application of OC significantly improves the fuel economy and emissions reduction during real driving conditions (see section 6).

2 Experimental setup

The OCP presented in this paper is experimentally benchmarked in an engine test bench to evaluate the performance of the proposed methods for the control of a real engine. The engine testing facility, a real driving cycle and a testing campaign are described in the following points.

2.1 Testing facility

These control works are performed in a Euro 5 turbocharged diesel engine equipped with VGT and high pressure EGR. Table 1 summarizes the main specifications of this engine.

Configuration	Euro 5 diesel, 4 in line, 16 valve, common rail, VGT, EGR
Displacement	2 liters
Max torque	340 Nm @ 2000 rpm
Max power	120 kW @ 3750 rpm
After-treatment	DOC and DPF

Table 1: Main specifications of the engine.

Control inputs	u_f	Fueling rate
	u_{egr}	EGR valve position
	u_{vgt}	VGT opening
	$u_{\delta soi}$	SOI offset
Model states	p_2	Intake manifold pressure
	X_{Oi}	Intake manifold oxygen fraction
	p_3	Exhaust manifold pressure
	X_{Oe}	Exhaust manifold oxygen fraction
	ω_{tc}	Turbocharger speed
Disturbances	N	Engine speed

Table 2: Summary of controls, states and disturbances to the OCP.

The test bench dyno is controlled by a Horiba SPARC/STARS system. The ECU features an ETK port that allows monitoring and accessing to the ECU volatile memory on real time. The control variables of interest (listed in table 2) are bypassed through the ETK port. The engine is instrumented with several thermocouples and pressure sensors at different spots of the airpath as a supplement to the factory sensors. A turbocharger speed sensor and two commercial NO_x probes are also fitted to the engine. A Horiba MEXA-7170DEGR exhaust gas analyzer and an AVL 439 opacimeter measure raw pollutant emissions (before after-treatment), intake CO_2 (in order to have accurate steady state EGR ratio readings) and exhaust gas opacity.

All the above measurements are ultimately sent to a dSPACE MicroAutoBox II real time device that manages sensor readings and OC trajectories to be injected into the ECU.

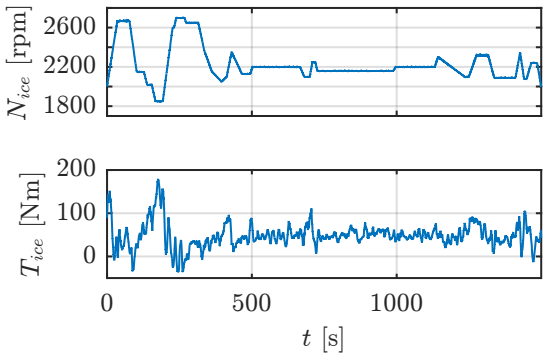

Figure 1: Real driving cycle engine speed and torque trajectories.

2.2 Driving cycle

The benchmarking driving cycle is a daily commute between two cities, recorded from actual driving with GPS and OBD devices. The engine speed is captured from the vehicle sensors while torque trajectory is reconstructed according to the vehicle speed and road height profile as collected from Google Maps API. This trip features positive and negative slopes during highway operation on a 40 km and 25 minutes long route. Speed and torque profiles are shown in figure 1. Note that the sharpness of the engine speed is due to test bench limitations which prevented to reproduce the exact actual trajectory.

2.3 Test campaign

In order to fit the model to the actual engine, a set of experimental measurements has been carried out. These tests explore multiple control combinations as well as the dynamic response of the engine. To this effect, three tests have been performed:

- Test 1: parametric tests at 20 different engine operating points that cover the typical operating points on the benchmarking driving cycle. Fueling rate, EGR, VGT and SOI are explored at each operating point. This provides steady state information at a wide range of engine conditions, namely 980 different combinations of controls and operating points (4 hours).
- Test 2: WLTC cycle with factory calibration.

This test provides transient information at a dynamic cycle (30 minutes).

- Test 3: benchmarking driving cycle trajectories with factory calibration. This test is only used as a validation to check that the model has an acceptable accuracy (25 minutes).

In addition to the above model fitting tests, a set of experiments have also been executed to validate the proposed OC strategy. These experiments consist of 12 optimal trajectories (calculated with the presented methodology) applied to the engine controls and an additional experiment with factory calibration. These experiments will be introduced in section 5.

3 Model

A zero-dimensional MVEM [15] has been built to calculate the OC trajectories. This engine model is inspired by other existing MVEMs such as [16]; however, an effort has been done to produce a smooth and fully continuous model. It consists of intake and exhaust manifolds, EGR loop, turbocharger and cylinder submodels. It has been fitted to the experimental data described above in section 2.3.

3.1 Intake and exhaust manifolds

Manifolds are modeled as reservoirs [17]. The equilibrium between the incoming and outgoing gases produces a variation on the reservoir pressure, energy and composition. These quantities are solved using the mass and energy conservation principles:

$$\dot{m}(t) = \dot{m}_{us}(t) - \dot{m}_{ds}(t) \tag{1}$$

$$\dot{U}(t) = \dot{H}_{us}(t) - \dot{H}_{ds}(t) + \dot{Q}(t) \tag{2}$$

with m the mass of gas in the reservoir, ds and us subindices refer to downstream and upstream flows, U its internal energy, H the enthalpy of incoming and outgoing flows, and $\dot{Q}$ the heat flow exchanged with the environment, which is null assuming an adiabatic manifold.

The energy associated to the reservoir gas temperature is:

$$U = c_v m\theta = \frac{1}{\kappa - 1} pV \tag{3}$$

where c_v is the heat capacity at constant volume and θ the temperature inside the reservoir. Note that the right hand side expression is obtained by introducing the ideal gas equation substituting mass and temperature. The energy flow at upstream and downstream flows are:

$$\dot{H}_{us} = c_p \dot{m}_{us} \theta_{us} \tag{4}$$

$$\dot{H}_{ds} = c_p \dot{m}_{ds} \theta_{ds} \tag{5}$$

with c_p the heat capacity and constant pressure. The introduction of eqs. (3) to (5) in the energy conservation relation (2) enables to express the variation of the reservoir pressure as:

$$\dot{p} = \frac{\kappa R}{V} \left[\dot{m}_{us}\theta_{us} - \dot{m}_{ds}\theta_{ds}\right] = \frac{R\theta}{V} \left[\dot{m}_{us} - \dot{m}_{ds}\right] \tag{6}$$

The right hand side expression assumes an isothermal reservoir. This may be justified because the temperature variations in the manifold are significantly slower than other quantities such as pressure or gas composition. This simplification is essential to develop a light enough control-oriented model, avoiding temperature as an additional state to the problem.

The mass conservation principle (1) allows to track the composition of the reservoir gas. Particularly, the oxygen mass m_O in this system verifies:

$$\dot{m}_O = \dot{m}_{Ous} - \dot{m}_{Ods} \tag{7}$$

Oxygen mass flow can also be expressed as a function of the oxygen fraction X_O for convenience:

$$\dot{m}_O = \dot{m} X_O + m \dot{X}_O \tag{8}$$

$$\dot{m}_{Ous} = \dot{m}_{us} X_{Ous} \tag{9}$$

$$\dot{m}_{Ods} = \dot{m}_{ds} X_{Ods} \tag{10}$$

The above expressions can be seamlessly introduced in (7). If m is substituted by the ideal gas law, and $\dot{m}$ for the reservoir mass conservation principle in (1), the variation of the oxygen fraction at the reservoir results:

$$\dot{X}_O = \frac{R\theta}{pV} \dot{m}_{us} (X_{Ous} - X_O) \tag{11}$$

Expressions for pressure and oxygen fraction variations can be particularized for both intake and exhaust manifolds. The intake manifold (p_2 and X_{Oi}) receives fresh air from the intercooler ($\dot{m}_c$ and X_{Oc}) and the EGR gases ($\dot{m}_{egr}$ and X_{Oegr}), while this mixture goes into the cylinder inlet port ($\dot{m}_{ip}$):

$$\dot{p}_2 = \frac{R\theta_{im}}{V_{im}} \left[\dot{m}_c + \dot{m}_{egr} - \dot{m}_{ip}\right] \tag{12}$$

$$\begin{aligned}\dot{X}_{Oi} = \frac{R\theta_{im}}{p_2 V_{im}} &\left[\dot{m}_c(X_{Oc} - X_{Oi})\right. \\ &\left. + \dot{m}_{egr}(X_{Oegr} - X_{Oi})\right]\end{aligned} \tag{13}$$

At the exhaust manifold (p_3 and X_{Oe}), the gases are diverted from the cylinder exhaust port ($\dot{m}_{ep}$ and X_{Oep}) to the EGR valve and the exhaust line ($\dot{m}_t$):

$$\dot{p}_3 = \frac{R\theta_{em}}{V_{em}} \left[\dot{m}_{ep} - \dot{m}_{egr} - \dot{m}_t\right] \tag{14}$$

$$\dot{X}_{Oe} = \frac{R\theta_{em}}{p_3 V_{em}} \dot{m}_{ep} \left[X_{Oep} - X_{Oe}\right] \tag{15}$$

3.2 Turbocharger

The turbocharger is composed of a compressor, a turbine and a mechanical link between both. Models for these three subsystems are described below.

3.2.1 Compressor

The compressor model is based on its operating maps [16]. Contrary to the common representation of compressor mass flow and efficiency as a function of corrected rotational speed $\widehat{\omega}_{tc}$ and pressure ratio Π_c, this model replaces the pressure ratio for a normalized pressure ratio to account for the surge limit. The normalized pressure ratio is defined as:

$$\widehat{\Pi}_c = \frac{\Pi_c - 1}{\Pi_{srg} - 1} \tag{16}$$

The compressor surge happens at pressure ratios over Π_{srg}, so $\widehat{\Pi}_c$ must stay within the $[0, 1]$ range to avoid unstable operation.

According to experimental readings, the surge limit Π_{srg} can be approximated to the following empirical equation:

$$\Pi_{srg} = c_{10} \exp(c_{11}\widehat{\omega}_{tc}) + c_{20} \exp(c_{21}\widehat{\omega}_{tc}) \tag{17}$$

where c are tunable coefficients. Similarly, compressor mass flow and efficiency have been also approximated to the empirical expressions below:

$$\dot{m}_c = c_{00} + c_{10}\widehat{\omega}_{tc} + c_{01}\widehat{\Pi}_c + c_{11}\widehat{\omega}_{tc}\widehat{\Pi}_c + c_{02}\widehat{\Pi}_c^2 \tag{18}$$

$$\begin{aligned}\eta_c = &\left[c_{00} + c_{10}\widehat{\omega}_{tc} + c_{01}\widehat{\Pi}_c + c_{20}\widehat{\omega}_{tc}^2 + c_{11}\widehat{\omega}_{tc}\widehat{\Pi}_c\right. \\ &\left. + c_{02}\widehat{\Pi}_c^2\right] \left[1 - \exp(c_A\widehat{\Pi}_c + c_B)\right]\end{aligned} \tag{19}$$

3.2.2 Turbine

Following the compressor model philosophy, the turbine is represented with its operating maps [16]. As long as turbines are not affected by surge, corrected mass flow and efficiency are modeled with an analytical approximation to the experimental maps as a function of the pressure ratio Π_t and VGT opening u_{vgt}:

$$\begin{aligned}\dot{m}_{t,cor} = &c_{00} + c_{10}\Pi_t + c_{01}u_{vgt} + c_{20}\Pi_t^2 \\ &+ c_{11}\Pi_t u_{vgt} + c_{02}u_{vgt}^2\end{aligned} \tag{20}$$

$$\begin{aligned}\eta_t = &\left[1 - \exp(c_A(\Pi_t - 1))\right] \left[c_2 u_{vgt}^2 + c_1 u_{vgt} + c_0\right] \\ &+ c_B(\Pi_t - 1)\end{aligned} \tag{21}$$

3.2.3 Mechanical link

Compressor and turbine wheels are linked with a shaft so there is an energy transfer between exhaust gases and the compressor mass flow. The power generated by the turbine can be estimated assuming an adiabatic expansion with isentropic efficiency η_t:

$$P_t = \eta_t \frac{\gamma R}{\gamma - 1} \dot{m}_t \theta_{em} \left(1 - \Pi_t^{\frac{1-\gamma}{\gamma}}\right) \tag{22}$$

Similarly, the power required to perform an adiabatic compression at the compressor is:

$$P_c = \frac{1}{\eta_c} \frac{\gamma R}{\gamma - 1} \dot{m}_c \theta_{ci} \left(\Pi_c^{\frac{\gamma-1}{\gamma}} - 1\right) \tag{23}$$

The balance between those two power values above determines the acceleration of the turbocharger shaft:

$$\dot{\omega}_{tc} = \frac{P_t - P_c}{I_{tc}\omega_{tc}} \tag{24}$$

where I_{tc} is the moment of inertia of the rotating masses.

3.3 EGR loop

The EGR valve is modeled as an isothermal orifice [17] with the following assumptions: (i) the flow is compressible, (ii) it accelerates up to the narrowest point of the orifice without losses trading some pressure for kinetic energy, and (iii) after the orifice, the flow is completely turbulent and kinetic energy is dissipated without pressure recuperation. Then, the mass flow through the valve is:

$$\dot{m}_{egr} = AC_d \frac{p_3}{\sqrt{R\theta_{em}}}\Psi \tag{25}$$

where A is the effective opening of the valve, C_d the discharge coefficient and Ψ the expansion factor which is a function of the expansion ratio $\Pi_{egr} = p_3/p_2$. The expression for Ψ is a piecewise function that discriminates the conditions when the flow is chocked from those when it is not [18]. For the sake of simplicity, that expression is approximated to the continuous function below:

$$\begin{aligned}\Psi = &\left[\frac{1}{2} + \frac{1}{2}\mathrm{erf}(c_1\Pi_{egr} + c_0)\right] \\ &\cdot \left[c_2 \exp(c_3\Pi_{egr}) + c_4\right]\end{aligned} \tag{26}$$

The effective opening A is a function of the valve position. The following empirical expression for it is based on experimental observations:

$$A_{egr} = c_2\left[\frac{1}{2}\mathrm{erf}(c_1 u_{egr} + c_0) - \frac{1}{2}\mathrm{erf}(c_0)\right] \tag{27}$$

3.4 Cylinder

The events taking part at the cylinder are complex and highly nonlinear [2]. In order to avoid heavy equations that limit the control capabilities of the model–and as long as no intra-cycle control is to be performed–the cylinder is represented as a combination of physically meaningful expressions and black-box models whose inputs are the thermodynamic conditions of the intake gas and the control variables [17]. The interesting outputs are the indicated efficiency η_{ind}, the exhaust temperature θ_{ep} and $\mathrm{NO_x}$ emissions X_{nox}:

$$\begin{aligned}\eta_{ind} &= \eta_{ind}(N, u_f, u_{\delta soi}, X_{Oi}, p_2) \\ \theta_{ep} &= \theta_{ep}(N, u_f, u_{\delta soi}, X_{Oi}, p_2) \\ X_{nox} &= X_{nox}(N, u_f, u_{\delta soi}, X_{Oi}, p_2)\end{aligned} \tag{28}$$

with N the engine speed and $u_{\delta soi}$ the SOI offset from factory calibration (a positive value means an earlier SOI). This model entails an important simplification but it is a reasonable hypothesis since the characteristic time of the airpath is several orders of magnitude over the in-cylinder phenomena.

The gas mass flow that gets into the cylinders can be calculated assuming ideal gas behavior and a volumetric efficiency η_v [2]:

$$\dot{m}_{ip} = \eta_v \frac{p_2 N V_d}{120 R\theta_{im}} \tag{29}$$

with V_d the engine displacement. There are several empirical expressions available in the literature to approach the volumetric efficiency [2, 16, 17, 19]. However, based on experimental observations of the current engine, the following expression is proposed:

$$\eta_v = c_A \exp\left(\frac{c_2}{c_1 p_2 + c_0}\right) + c_B u_{vgt} + c_C \tag{30}$$

The presence of u_{vgt} in the above equation is due to purely empirical reasons, but enables to account for the effect of the VGT backpressure.

Since no reservoirs are considered in the cylinder, the exhaust mass flow is:

$$\dot{m}_{ep} = \dot{m}_{ip} + \dot{m}_f \tag{31}$$

with the fuel mass flow a consequence of the fueling rate u_f:

$$\dot{m}_f = u_f \frac{n_{cyl} N}{120 \cdot 10^6} \tag{32}$$

with n_{cyl} the number of cylinders.

Assuming perfect burning of fuel, the oxygen fraction at the exhaust gas is a function of the stoichiometric Air to Fuel Ratio (AFR):

$$X_{Oep} = \frac{X_{Oi}\dot{m}_{ip} - \text{AFR}X_{Oc}\dot{m}_f}{\dot{m}_{ep}} \tag{33}$$

Similarly, engine λ can also be calculated from the AFR:

$$\lambda = \frac{\dot{m}_{ip}X_{Oi}}{\text{AFR}\dot{m}_f X_{Oc}} \tag{34}$$

The indicated torque generation is a consequence of the fuel energy liberation during the combustion process:

$$T_i = \eta_{ind}\frac{u_f H_f n_{cyl}}{4 \cdot 10^6 \pi} \tag{35}$$

where H_f is the fuel heating power. The torque losses can be represented by a linear function of the engine speed, which may be fitted with a motoring experiment:

$$T_l = c_0 + c_1 N \tag{36}$$

Then, the effective torque at the crankshaft is:

$$T_e = T_i - T_l \tag{37}$$

3.5 Experimental validation

The accuracy of effective torque and NO_x emissions are critical for the OCP as long as they are part of the constraints and the variables that define the performance of the engine–to be precise, the fueling rate is another critical variable but, since it is a control that is forced at the experiments, no model is in between. In figure 2 it may be appreciated that torque and NO_x emissions are well correlated to experiments, especially the total NO_x emissions. Note that, in order to provide a fair model validation, this data corresponds to the validation cycle which the model is not fitted to (Test 3 in section 2.3). On the bottom plots of this figure, the dynamic response of intake pressure and turbocharger speed (states of the OCP) is shown. Despite a minor offset at $t = [245, 255]$ s, settling time is quite similar in simulations (red) and experiments (blue).

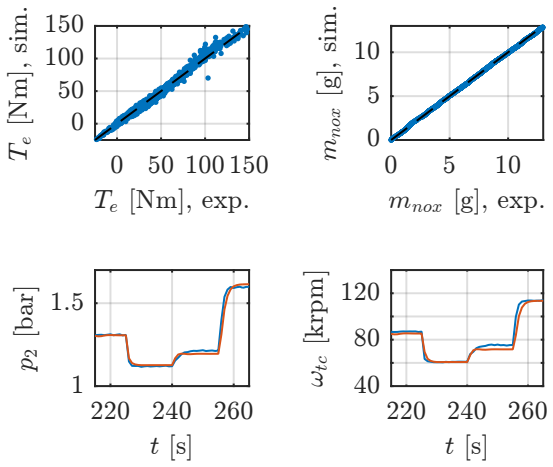

Figure 2: Model validation results. At top plots, correlation between experiments and simulations are shown for effective torque and accumulated NO_x emissions. On bottom plots, dynamics of the model are represented for two model states: intake manifold pressure and turbocharger speed. Blue corresponds to experiments and red to simulations.

4 Optimal control methodology

4.1 Problem description

The aim of this work is to find an OC strategy that minimizes fuel consumption with a specific limit on NO_x emissions while keeping a low soot stamp for a real driving cycle. This target can be translated into an OCP as the minimization of the cost index:

$$J = \int_0^{\mathcal{T}} \dot{m}_f(x, u, t)\mathrm{d}t \tag{38}$$

varying the controls trajectory. The controls u and states x are listed in table 2 for the current problem. The NO_x limit may be specified with an integral constraint that bounds the total emissions:

$$\int_0^{\mathcal{T}} \dot{m}_{nox}(x, u, t)\mathrm{d}t - \widehat{m}_{nox} \leq 0 \tag{39}$$

where $\widehat{m}_{nox}$ is the maximum amount of permissible NO_x generation during the cycle.

Since the target of this work is to calculate the appropriate control trajectories to reproduce a pre-

scribed driving cycle, its engine speed and torque trajectories must be followed. Engine speed is a disturbance to the problem and, therefore, is always fulfilled. However, the engine torque is a consequence of the controls. Then, an additional constraint is required to fulfill the driving cycle torque request $\widetilde{T}$:

$$\widetilde{T} - T_e(x, u, t) = 0 \tag{40}$$

Note that in the above formulation hat accents ($\widehat{\cdot}$) are used to denote limit values while tildes ($\widetilde{\cdot}$) refer to setpoint trajectories.

In order to avoid excessive smoke generation and to keep a low soot stamp, an opacity constraint should be included. Unfortunately, opacity models may be too expensive to be included in a control-oriented engine model that already has 5 states, and quasi-steady approaches show a poor correlation with experimental values. However, in figure 3 it may be appreciated that there is a relation between opacity and λ, as the less fresh air that takes part in the combustion process, the chances to generate more smoke increase. The red line overlaid in this same figure is the Pareto front, which can be used as a rough correlation between opacity and λ, enabling to limit smoke generation by limiting λ to the appropriate value. Of course, this is just a qualitative question as long as a stronger limitation in λ does not guarantee any particular opacity level, but will surely limit the smoke generation at some point with a low computational impact. Then, a λ constraint is included in the OCP:

$$\widehat{\lambda}(t) - \lambda(x, u, t) \leq 0 \tag{41}$$

where $\widehat{\lambda}$ is the boundary limit calculated according to the relation from figure 3 with an arbitrary 2% higher opacity level than factory calibration in order to allow some room for improvement.

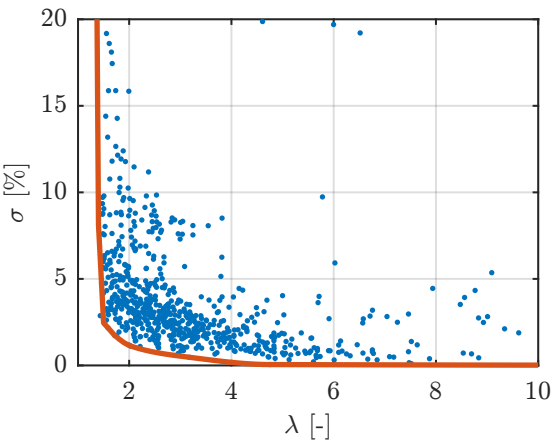

Figure 3: Relation between λ and opacity. Blue dots correspond to steady state measurements (Test 1 according to the test campaign description in section 2.3). The red line is the Pareto front for this tradeoff.

4.2 Optimization algorithm

The above OCP features complex nonlinear equations with 5 states and 4 controls. At a time step of $\Delta t = 0.05$ seconds, the current 25 minutes long driving cycle entails about 120000 unknowns for the control trajectories.

There are three main families of OC methods to approach OCPs: (i) DP [3, 20, 21], (ii) PMP [22–24], and (iii) DMs [25–27]. DP can effectively address complex problems but suffers the *curse of dimensionality* [28], making a problem like this technically untractable with current computers. PMP is not affected by those size-related issues but when several states are involved, the ODEs are difficult to address and tend to be ill-conditioned resulting in the impossibility to reach a numerical solution [3, 29].

On the contrary, DMs do not suffer any particular issue when facing complex nonlinear OCPs of any dimension. The main drawback is that the solution is a local optimum that may differ from the global one [26]. However, a proper initial solution increases the chances to find a global optimum and, if problem is convex or sufficiently smooth, it may be reached almost regardless of the initial solution. Anyhow, most of the times this is a minor issue. An important remark is that model must show C^1 continuity for DM to be applied since derivatives of the objective function and constraints must be provided. With this purpose, the developed MVEM is completely continuous.

Due to the above strengths of DMs, it is the selected method to address this OCP. Among all the existing DM methods [30], Direct Collocation (DC) is the choice for this work due to computational burden reasons. The idea behind DC is to discretize

state and control trajectories into $N+1$ time steps such that the complex original OCP translates into the selection of a large set of scalar unknowns (for a better insight into DC and DMs in general, the interested reader is redirected to [26]). To guarantee continuity between those individual unknowns and fulfilling of the model dynamics, ODEs are approximated with a numerical method–centered Euler's method in this case. Therefore, the OCP is transformed into a large–and sparse–NLP. The objective function (38) becomes:

$$\min_{\mathrm{x,u}} \left\{ \sum_{i=0}^{N-1} \dot{m}_f \left(\frac{\mathrm{x}_i + \mathrm{x}_{i+1}}{2}, \mathrm{u}_i \right) \right\} \tag{42}$$

where the subindex i refers to the states and controls at time $t_i = i \cdot \mathcal{T}/N$. The state dynamics are translated into a set of algebraic constraints:

$$f\left(\frac{\mathrm{x}_i + \mathrm{x}_{i+1}}{2}, \mathrm{u}_i \right) - \frac{\mathrm{x}_{i+1} - \mathrm{x}_i}{\Delta t} = 0 \tag{43}$$

with $i = 0, \ldots, N-1$. Similarly, $\mathrm{NO_x}$ bound (39), torque fulfilling (40) and λ limitation (41) constraints are also transcribed into algebraic constraints:

$$\sum_{i=0}^{N-1} \dot{m}_{nox} \left(\frac{\mathrm{x}_i + \mathrm{x}_{i+1}}{2}, \mathrm{u}_i \right) \Delta t - \widehat{m}_{nox} \leq 0 \tag{44}$$

$$\widetilde{T}_i - T_e \left(\frac{\mathrm{x}_i + \mathrm{x}_{i+1}}{2}, \mathrm{u}_i \right) \leq 0 \tag{45}$$

$$\widehat{\lambda}_i - \lambda \left(\frac{\mathrm{x}_i + \mathrm{x}_{i+1}}{2}, \mathrm{u}_i \right) \leq 0 \tag{46}$$

This NLP is extremely sparse due to the fact that only two consecutive time steps are related at most of these algebraic constraints (times t_i and t_{i+1}). Integral constraints are dense, but they are only a few. Therefore, the Jacobian of this NLP–the derivative of constraints with respect to all unknowns, *i.e.* discretized states and controls–is made of diagonal submatrices as it can be appreciated in figure 4. This matrix is quite big (239855×299816) but only few elements are non-zero (0.0035% in this case). This makes possible that, even after exploiting problem dimensions by discretizing states, controls and

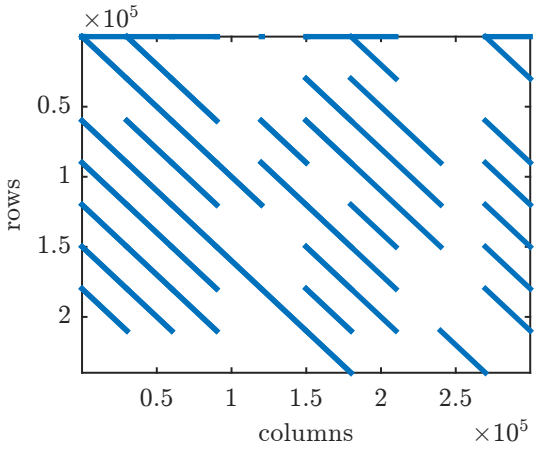

Figure 4: Jacobian matrix of the transcribed NLP. Blue dots are non-zero elements, whose number is 2488433, a 0.0035% of the total number of matrix elements.

constraints, the computational time required to compute a complete NLP iteration is 1.04 seconds with a total memory stamp of 23.1 MB.

The NLP is addressed with IPOPT, a software package for large-scale nonlinear optimization that is specially efficient for sparse problems [31]. Analytical expressions for the objective function, constraints and their exact derivatives (obtained with symbolical differentiation) are provided to IPOPT. Second derivatives may be also provided explicitly, but they have been numerically approximated with the quasi-Newton L-BFGS algorithm [32] for computational efficiency reasons. The initial solution provided to IPOPT is an arbitrary constant value trajectory for states and controls. A more sophisticated initial solution such as ECU setpoints may reduce the number of iterations to find an optimum and increase the robustness of the method. However, as long as computational time is still short with such a simple initial solution, it has been considered sufficient for this study.

4.2.1 Remarks on solution robustness

The above NLP does not intrinsically limit the rate of change of controls. In case of model discrepancies, it may happen that optimal trajectories show an oscillating behavior. This is not an issue for the fueling rate neither for the SOI. The first has a direct

impact on torque output which is constrained to the driving cycle trajectory, while the second shows a low interaction with other variables. However, EGR and VGT controls are free to show a switching policy, producing pulsating flows along the airpath and an unstable behavior. In order to improve the robustness of the method avoiding operating with a highly dynamic policy, a rate of change limitation has been imposed in the VGT control. Since this involves a constraint on the derivative of a control, it is necessary to introduce an auxiliary state:

$$x_{vgt} = u_{vgt} \tag{47}$$

which replaces the original VGT control variable. Instead, a new control must be defined:

$$u_{\delta vgt} = \frac{\partial u_{vgt}}{\partial t} \tag{48}$$

The dynamics of the additional state are driven by:

$$\dot{x}_{vgt} = u_{\delta vgt} \tag{49}$$

The rate of change limitation can be imposed by specifying upper and lower bounds to the control $u_{\delta vgt}$. In figure 5 it may be appreciated that, by applying this approach, VGT and EGR rate of change are reduced from highly oscillating situations (up) to a control policy that falls within the ±60%/s range (down). Note that EGR oscillation is indirectly reduced as a consequence of the coupling between both controls in the airpath management. For this reason, additional constraints on EGR rate of change were considered unnecessary.

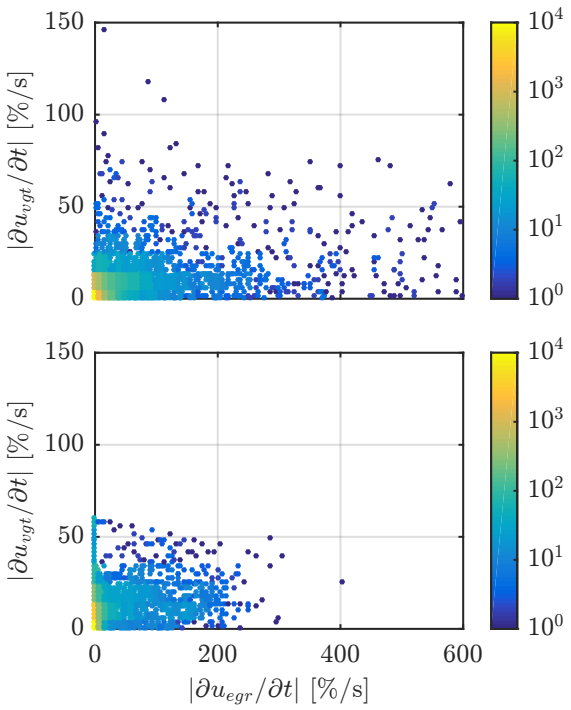

Figure 5: Derivative of the EGR and VGT optimal controls for an OCP with (down) and without (up) a constraint on the VGT rate of change. The color scale represent the number of occurrences of these points. Note the effect of the constraint not only in VGT but also in the EGR rate of change.

5 Experimental results

The presented OCP is addressed for different trade-offs between total fuel consumption and accumulated NO_x emissions. The resulting experimental fuel consumption, NO_x emission and soot generation are summarized in table 3. Optimization objective in cases #1 to #10 is fuel consumption minimization for different levels of maximum accumulated NO_x emissions $\widehat{m}_{nox}$, case #11 is fuel minimization with no specified limit for NO_x emissions, case #12 is NO_x minimization regardless of fuel consumption, and case factory is performed with the factory ECU control. Note that simulation results correspond to experimental fuel consumption (it is forced to be the same during experimental tests with a small deviation due to the feedback controller) and NO_x emission limit $\widehat{m}_{nox}$ (optimization meets all constraints within a negligible tolerance).

Despite all these cases delivered the same effective torque in simulations, model uncertainties might produce torque deviations, resulting in an unfair comparison between different strategies and factory calibration. In order to approach this issue, a feedback controller monitors the actual torque output during the test and performs a fine tune of fueling rate accordingly. Of course this simple solution somehow spoils the optimality of the control strategy, but provides a fair and clear comparison among strategies while keeping a quite similar control policy (corrections are bounded by ±0.85 mg/str). The difference between requested and delivered torque is shown in the histogram at figure 6. This histogram

Case	OCP		Experimental results		
	J	$\widehat{m}_{nox}$ [g]	m_f [kg]	m_{nox} [g]	m_{soot} [g]
#1	m_f	3	1.395	3.6	1.846
#2	m_f	4	1.344	5.2	0.892
#3	m_f	5	1.291	6.1	1.001
#4	m_f	6.4	1.268	6.8	0.847
#5	m_f	7.5	1.262	7.9	0.477
#6	m_f	9	1.254	10.5	0.382
#7	m_f	11	1.252	12.1	0.282
#8	m_f	13.3	1.252	13.9	0.304
#9	m_f	17	1.253	17.3	0.218
#10	m_f	20	1.257	22.5	0.208
#11	m_f	–	1.247	20.5	0.229
#12	m_{nox}	–	1.369	3.8	0.936
Factory	–	–	1.306	11.4	0.340

Table 3: Summary of all OCPs that have been issued and their main experimental results. The information shown is, from left to right, the case number, the minimization objective, NO_x emission limit assessed by simulation, experimental fuel consumption, actual NO_x emission and total soot generation.

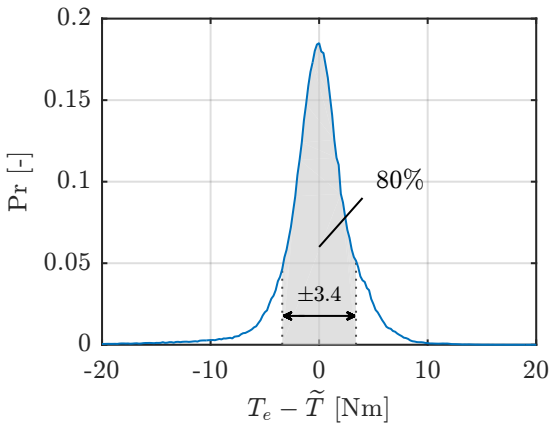

Figure 6: Histogram of the error between experimental and reference engine torque. An 80% of the measurements lie in the gray area (±3.4 Nm).

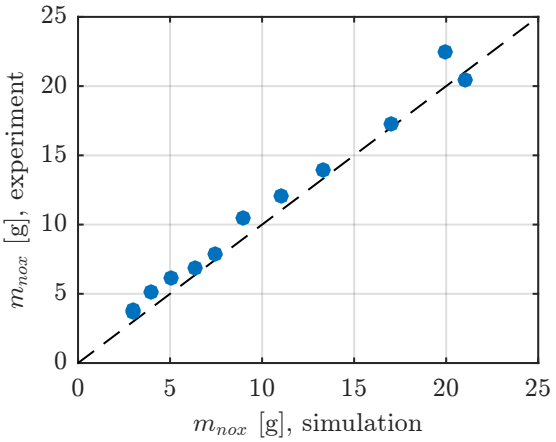

Figure 7: Correlation in accumulated NO_x emissions between OCP simulations and experiments. The average error is −0.86 g and the standard deviation is 0.75 g.

is quite close to a normal variable, with an average of −0.1 Nm and a standard deviation of 3.3 Nm. These results confirm that the different control strategies are comparable since the driving cycle is followed up to a sufficient threshold.

It can also be appreciated that accumulated NO_x emissions are pretty well correlated to the model estimations as shown in figure 7. Experimental values are slightly above of simulation results due to NO_x model inaccuracies during transients. However differences are mostly under one gram so model accuracy is acceptable in that respect.

Experimental fuel consumption and NO_x emissions at all OC strategies are depicted in the Pareto front at figure 8. Soot is also included as a reference of how different controls affect to its generation but remind that it is not included in the optimization problem explicitly but through a varying λ limit. The Pareto front shows the set of points that minimize fuel consumption for a given level of NO_x emissions, defining the frontier where the engine operates on its most efficient way: an optimal control falls within it, any suboptimal control remains over the frontier, and operation below is not possible. Factory calibration (red square) falls over the Pareto front since it is not optimal and, in fact, several experimental tests showed lower fuel consumption and NO_x emissions simultaneously.

According to the results, it is possible to decrease both NO_x emissions and fuel consumption with an OC strategy compared to the factory calibration.

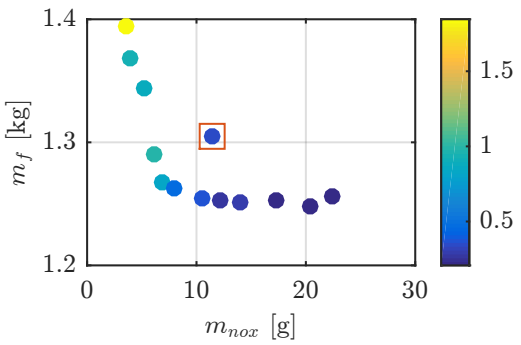

Figure 8: Pareto frontier analysis of OC experimental results for a range of NO_x limits. Factory calibration is denoted with a red square. The color scale shows the soot generation in grams.

However, those two quantities do not show the same room for improvement: NO_x emissions can be easily decreased to less than a half with the appropriate strategy, but fuel savings are not greater than a 4%. The reason for this is that torque output–which is constrained–is mostly a function of the fueling rate while NO_x generation is affected by many other factors. It may be remarked that the decrease in NO_x seems to carry an increase in soot emissions. Although this effect might limit the minimum NO_x emissions that can be achieved, these levels are raw measurements (before after-treatment) and DPF may equate the soot emission of all these control strategies, although regenerations might be necessary more often with its corresponding fuel penalty.

The Pareto front shows a tradeoff between fuel and NO_x nearby the factory calibration. However, in the extrema both variables show an asymptotic trend that can be explained as the combination of two effects: on the one hand, engine technology limits the minimum level of fuel and NO_x that can be achieved and, on the other hand, engine operation at extrema is on the limits of the model where little fitting information was available. Therefore, 1.25 kg of fuel and 6 g of NO_x are the limits for this cycle.

In order to analyze how do the OC strategies improve the factory control, histograms of four optimized controls–u_f, u_{egr}, u_{vgt} and $u_{\delta soi}$–are shown in the contours at figure 9. At first glance, it is clear that factory control is pretty different from OC. The EGR rate is much higher in OC, recirculating as much as possible (90% to 100% opening) with the exception of the case where NO_x emissions are neglected (> 15 g). Meanwhile, factory control moves along lower rates, with EGR valve position mostly between 40% and 60%. On the contrary, OC strategies tend to open the VGT delivering less pressure than factory at the intake manifold, and this effect is stronger as fuel consumption decreases. Only at very low NO_x emissions (high fuel consumption) VGT is clearly closer than at factory control. Regarding fuel consumption, it is obviously lower than factory calibration and its value rises as NO_x emissions decrease. Finally, SOI is, in general, 1° delayed compared to factory control regardless of the case, suggesting that the engine has a strong optimal pole at that SOI setting that differs from factory calibration in this cycle.

According to the above observations, OC exploits two strategies. On the one hand, it reduces pumping losses by opening the VGT. This decreases the intake manifold pressure and, therefore, λ might be penalized; however, since λ is constrained in the problem, the strategy still keeps AFR at acceptable values. The decrease of intake pressure also jeopardizes the torque reserve. ECU has no lookahead information so it has to keep some amount of intake pressure to satisfy a strong load step whenever requested with a reduced lag. On the contrary, OC knows the complete cycle in advance so that torque reserve is no longer required and avoids an unnecessary pumping effort. On the other hand, OC recirculates as much exhaust gases as possible to reduce NO_x generation at the expense of a lower combustion efficiency. In fact, to satisfy tighter NO_x constraints, VGT is also used to build backpressure and increase the EGR ratio.

Then, the OC is a question of balance between the advantage of reduced pumping losses and the penalization of high EGR rates to decrease NO_x emissions. The definition of a proper OCP can help finding the right balance to minimize fuel consumption at a given NO_x limit. In figure 10 it may be appreciated that optimal strategies are able to decide which is the most appropriate part of the cycle where low NO_x strategies benefit the most and where fuel consumption should be a priority. For example, $t = [200, 300]$ s seems specially interesting to reduce NO_x as all OC strategies do strongly recirculate exhaust gases in that phase. Later in the

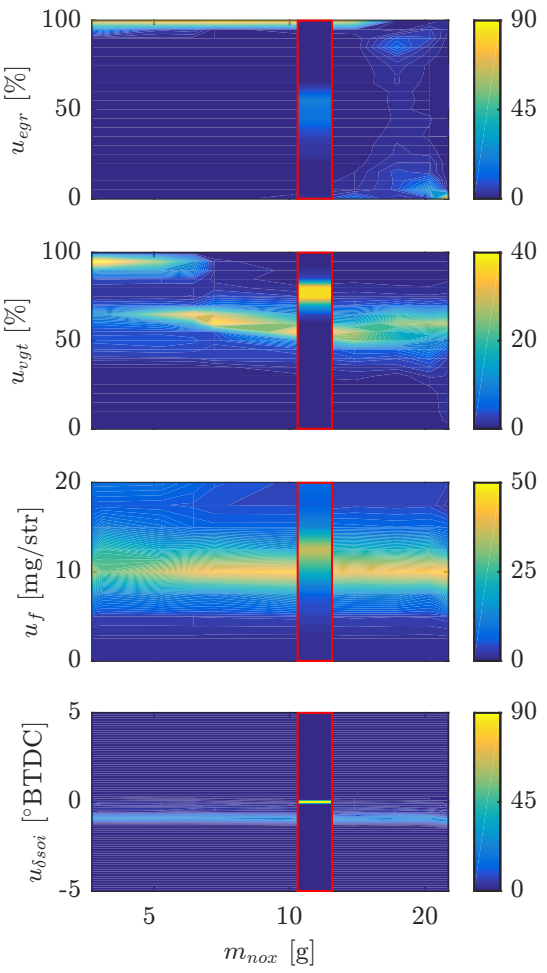

Figure 9: Histogram of controls for all 12 optimization cases that have been issued. Each plot represents a different control, the x-axis lists the total NO_x emissions of the OC strategy and the y-axis the range of operation of the control. Each column belongs to a different case. ECU control is shown at its corresponding column–according to its NO_x emission level–rounded in a red square. The color scale shows the level of occurrence of a control as a percentage.

cycle, fuel consumption becomes the target since EGR rates are low. Meanwhile, the factory control holds an intermediate EGR amount without discriminating different cycle characteristics. This discerning capability of the OC is key to find the proper balance between fuel and NO_x by exploiting the most attractive strategy at each phase of the cycle, and it is something that the factory control cannot perform as long as it lacks of lookahead information.

6 Conclusions

OC theory has been applied to the fueling rate, EGR, VGT and SOI control of a Euro 5 diesel engine in a real driving cycle. A detailed, control-oriented, experimentally validated MVEM and a methodology to calculate the optimal trajectories for these four controls making use of a DC optimization method have been presented and described above. These trajectories have been experimentally validated on a testing facility showing that OC is able to improve both fuel efficiency and NO_x emissions compared to factory control. Also, it must be remarked that it is pretty straightforward to adapt the control strategy to specific emissions requisites just by modifying the NO_x constraint limit at the formulation without any additional calibration.

According to the results, a 45% NO_x reduction is possible while keeping the factory fuel efficiency; similarly, a 4% improvement in fuel consumption may be achieved with factory NO_x and soot levels. These results demonstrate that an OC strategy may improve the performance compared to a fixed calibration, showing that there is still room for improvement in both fuel consumption and NO_x emissions with an adequate control strategy in a real driving cycle.

OC trajectories follow two strategies to achieve these results. On the one hand, pumping losses are reduced by opening the VGT at the expense of a decrease in intake pressure. It spoils the torque reserve but, as long as the cycle is known in advance, torque reserve is no longer necessary. On the other hand, EGR rates are much higher in order to reduce NO_x generation at the expense of penalizing the combustion efficiency. These two strategies are applied specifically on the parts of the cycle where they are more beneficial for the global efficiency. It may be appreciated how the control switches from strong NO_x reducing phases to high fuel efficient operation depending on the current cycle demands in the context of the complete driving cycle. The ECU has no chance to find out that cycle-specific

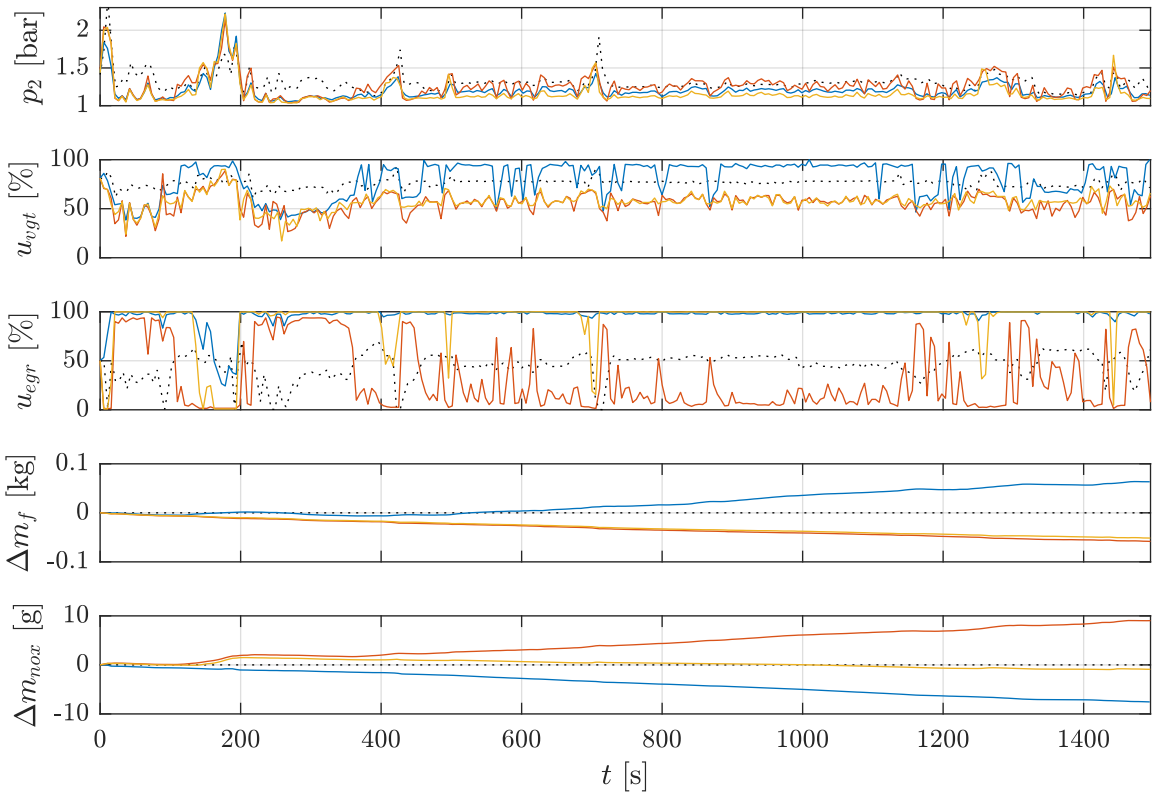

Figure 10: OC trajectories for the whole driving cycle. From top to bottom: intake manifold pressure, VGT opening, EGR position, fuel consumption difference with respect to factory and NO_x emissions compared to factory (values greater than zero mean higher mass than factory). The cases depicted in this figure and their corresponding case number at table 3 are: factory calibration (- - -), low NO_x emissions (——, case #12), low fuel consumption (——, case #11) and factory NO_x with lower fuel consumption (——, case #6).

balance and, therefore, it has to keep a reasonable tradeoff between NO_x and fuel at any condition, which is by far not the best option.

The necessity of lookahead information to perform an OC may be a significant drawback of this methodology in a practical implementation with current technology. However, it is still a valuable tool for benchmarking purposes of existing engine control strategies as a reference of how far they are from the theoretical optimum, or even to derive new heuristic rules from the OC results. OC can also be used to inspect the additional room for improvement that introduces an architecture or technology change. Scenarios where lookahead information is available are good candidates for the presented methodology, such as adaptive cruise controls with orographic information or even autonomous vehicles.

References

[1] F Payri et al. “A Challenging Future for the IC Engine: New Technologies and the Control Role”. In: *Oil & Gas Science and Technology – Revue d’IFP Energies nouvelles* 70.1 (2015), pp. 15–30.

[2] J Heywood. *Internal Combustion Engine Fundamentals*. McGraw-Hill, 1988.

[3] Arthur E Bryson and Yu-Chi Ho. *Applied optimal control: Optimization, estimation, and control*. Taylor & Francis, 1975.

[4] Patrick Hagelauer and Felix Mora-Camino. "A soft dynamic programming approach for on-line aircraft 4D-trajectory optimization". In: *European Journal of Operational Research* 107.1 (1998), pp. 87–95.

[5] Andrew Higgins, Erhan Kozan, and Luis Ferreira. "Optimal scheduling of trains on a single line track". In: *Transportation research part B: Methodological* 30.2 (1996), pp. 147–161.

[6] Christopher L Darby and Anil V Rao. "Minimum-Fuel Low-Earth Orbit Aeroassisted Orbital Transfer of Small Spacecraft". In: *Journal of Spacecraft and Rockets* 48.4 (2011), pp. 618–628.

[7] Tomas Nilsson, Anders Froberg, and Jan Aslund. "Optimal Operation of a Turbocharged Diesel Engine during Transients". In: *SAE International Journal of Engines* 5 (2012), pp. 571–578.

[8] Jonas Asprion, Oscar Chinellato, and Lino Guzzella. "Optimal control of diesel engines: Numerical methods, applications, and experimental validation". In: *Mathematical Problems in Engineering* 1 (2014), pp. 1–21.

[9] Carlos Guardiola et al. "Adaptive calibration for reduced fuel consumption and emissions". In: *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering* 230.14 (2016), pp. 2002–2014.

[10] José Manuel Luján et al. "Switching strategy between HP (high pressure)- and LPEGR (low pressure exhaust gas recirculation) systems for reduced fuel consumption and emissions". In: *Energy* 90, Part 2 (2015), pp. 1790–1798.

[11] Carlos Guardiola et al. "Optimal Control as a method for Diesel engine efficiency assessment including pressure and NOx constraints". In: *Applied Thermal Engineering* 117 (2017), pp. 452–461.

[12] C. Guardiola et al. "Modelling driving behaviour and its impact on the energy management problem in hybrid electric vehicles". In: *International Journal of Computer Mathematics* 91.1 (2014), pp. 147–156.

[13] Daniel Ambühl et al. "Explicit optimal control policy and its practical application for hybrid electric powertrains". In: *Control Engineering Practice* 18.12 (2010), pp. 1429–1439.

[14] Antonio Sciarretta, Giovanni De Nunzio, and Luis Leon Ojeda. "Optimal Ecodriving Control: Energy-Efficient Driving of Road Vehicles as an Optimal Control Problem". In: *IEEE Control Systems* 35.5 (2015), pp. 71–90.

[15] Carlos Guardiola et al. "Representation limits of mean value engine models". In: *Identification for Automotive Systems*. Springer, 2012, pp. 185–206.

[16] Johan Wahlström and Lars Eriksson. "Modelling diesel engines with a variable-geometry turbocharger and exhaust gas recirculation by optimization of model parameters for capturing non-linear system dynamics". In: *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering* 225.7 (2011), pp. 960–986.

[17] Lino Guzzella and Christopher H Onder. *Introduction to modeling and control of internal combustion engine systems*. Springer, 2004.

[18] Michael J Moran et al. *Fundamentals of engineering thermodynamics*. John Wiley & Sons, 2010.

[19] Jonas Asprion, Oscar Chinellato, and Lino Guzzella. "Optimisation-oriented modelling of the NOx emissions of a diesel engine". In: *Energy Conversion and Management* 75 (2013), pp. 61–73.

[20] Richard Bellman. "The theory of dynamic programming". In: *Bulletin of the American Mathematical Society* 60.6 (1954), pp. 503–515.

[21] Hans P Geering. *Optimal control with engineering applications*. Springer, 2007.

[22] L S Pontryagin et al. *The mathematical theory of optimal processes*. Interscience, New York, 1962.

[23] Richard Vinter. *Optimal control*. Springer, 2010.

[24] Heinz Schättler and Urszula Ledzewicz. *Geometric Optimal Control: Theory, Methods and Examples*. Springer, 2012.

[25] D Tieu, W R Cluett, and A Penlidis. “A comparison of collocation methods for solving dynamic optimization problems”. In: *Computers and Chemical Engineering* 19.4 (1995), pp. 375–381.

[26] John T Betts. *Practical methods for optimal control and estimation using nonlinear programming*. SIAM, 2010.

[27] Michael Hinze and Arnd Rösch. “Discretization of optimal control problems”. In: *Constrained Optimization and Optimal Control for Partial Differential Equations*. Springer, 2012, pp. 391–430.

[28] Richard Bellman. “Dynamic programming and Lagrange multipliers”. In: *Proceedings of the National Academy of Sciences* 42.10 (1956), pp. 767–769.

[29] F Lewis and V Syrmos. *Optimal Control*. Wiley-Interscience, 1995.

[30] Moritz Diehl et al. “Fast direct multiple shooting algorithms for optimal robot control”. In: *Fast motions in biomechanics and robotics*. Springer, 2006, pp. 65–93.

[31] Andreas Wächter and Lorenz T. Biegler. “On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming”. In: *Mathematical Programming* 106.1 (2006), pp. 25–57.

[32] Jorge Nocedal. “Updating quasi-Newton matrices with limited storage”. In: *Mathematics of Computation* 35.151 (1980), pp. 773–782.