In-Vehicle Realization of Nonlinear MPC for Gasoline Two-Stage Turbocharging Airpath Control

Thivaharan Albin, Dennis Ritter, Norman Liberda, Rien Quirynen, and Moritz Diehl

***Abstract*— Innovative charging concepts, such as two-stage turbocharging for gasoline engines, cause high demands on the process control. The open-loop process is characterized by a complex, nonlinear system behavior. In addition, the requirements on the closed-loop system are challenging: fast reference tracking has to be achieved without overshoots while respecting constraints on the turbocharger speeds in order to prevent damaging of the components. Nonlinear model predictive control (NMPC) offers a high potential for this purpose. It is capable of handling coupled multiple-input systems while achieving high control quality and respecting constraints on system states. This paper presents an NMPC scheme for a two-stage turbocharged gasoline airpath, which is based on a physically driven reduced-order model formulated as a set of differential-algebraic equations. The online optimal control algorithm uses the real-time iteration scheme and is implemented on a control prototyping platform. For validation of the algorithm, it is tested based on simulations and vehicle experiments. These experiments have been conducted on a vehicle dynamometer as well as on an automotive testing track. For this purpose, a modified production vehicle is used in which the airpath concept is implemented.**

***Index Terms*— Airpath control, engine control, nonlinear model predictive control (NMPC), turbocharging.**

I. Introduction

TO REDUCE fuel consumption and emissions for internal combustion engines, "downsizing" by the use of turbochargers is investigated. Turbochargers allow one to use energy from the hot exhaust gas in order to compress the intake air of the combustion engine to higher pressures than achievable with a naturally aspirated engine. For increasing the specific power, conventional single-stage turbocharging concepts lead to conflicting goals, concerning the dimensioning of the charging components. A high specific power on the one hand and a fast transient raise of the boost pressure on the other cannot be realized at the same time. To mitigate this tradeoff, more variability in the charging devices is used. A future promising technology in gasoline engines is the two-stage turbocharging concept.

The architecture comprises a small high-pressure (HP) stage and a large low-pressure (LP) stage. The small HP stage is capable of realizing fast transients, even though it is restricted concerning the specific power. By contrast, the large LP stage can realize a high specific power with slower transient dynamics. One of the arising challenges consists in the design of the closed-loop controller, which fulfills the high requirements on the control quality.

A. Control Requirements

The control algorithm needs to handle both turbocharger stages in a coordinated fashion, such that reference tracking for the boost pressure is made possible and disturbances are rejected. A fixed air-to-fuel ratio is maintained for the combustion of a gasoline engine. As a consequence, the fuel is always adapted to the amount of air mass trapped in the cylinder, which is called quantitative control. In this case, there is a direct correlation between the boost pressure, the amount of air, the fuel amount, and the generated torque. The boost pressure reference should thus be reached as quickly as possible, as this determines the transient acceleration capability of the vehicle. However, in case of a step reference input, the output should additionally be achieved without strong overshoots, as this negatively influences the driving behavior. The working principle of a diesel engine does not require a fixed air-to-fuel ratio. Instead, the airpath is adjusted, such that an excess of air is present all the time. This allows the torque to be only based on the amount of fuel, which is called qualitative control. Therefore, in a diesel engine, the torque and the boost pressure are decoupled, such that oscillations can be tolerated up to a certain amount.

In addition to the latter requirements on the reference tracking performance, the control algorithm should respect the upper limit constraints for the HP and LP turbocharger speeds, since exceeding these limits might damage the turbocharger. This becomes especially challenging, as the turbocharger speed is typically not even measured in a series production configuration. In summary, the following three major requirements have to be considered at the same time for gasoline airpath control:

1) fast reference tracking;
2) no oscillatory behavior for step reference;
3) while respecting limits on turbocharger speeds.

In addition, turbocharged systems typically show strongly nonlinear behavior and their dynamic behavior is very dependent on disturbance variables, such as the engine speed. For this purpose, nonlinear model predictive control (NMPC) is

Manuscript received September 16, 2016; revised March 30, 2017; accepted June 15, 2017. Manuscript received in final form July 4, 2017. The research was performed as part of the Research Unit (Forschergruppe) FOR 2401 "Optimization based Multiscale Control for Low Temperature Combustion Engines," which is funded by the German Research Association (Deutsche Forschungsgemeinschaft, DFG). The work of R. Quirynen was supported by the Research Foundation–Flanders (FWO) through the Ph.D. Fellowship. Recommended by Associate Editor K. Butts. *(Corresponding author: Thivaharan Albin.)*

T. Albin and D. Ritter are with the Institute of Automatic Control, RWTH Aachen University, 52074 Aachen, Germany (e-mail: t.albin@irt.rwth-aachen.de).

N. Liberda is with FEV Europe GmbH, 52078 Aachen, Germany.

R. Quirynen is with KU Leuven, 3000 Leuven, Belgium.

M. Diehl is with the University of Freiburg, 79110 Freiburg, Germany.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TCST.2017.2724020

investigated in this paper. NMPC is a suitable choice, as it can handle the nonlinear system dynamics and is able to respect constraints on system states with a high control quality. The overall goal is that the NMPC scheme makes use of the specific turbocharging architecture to overcome the tradeoff between fast transient raise and high power.

B. Related Work

An overview of fundamental modeling and control of airpath systems can be found in [1] and [2]. For the purpose of airpath control, a variety of control concepts have been applied, where especially model-based control shows advantages due to the arising strong nonlinearities. Examples of the latter are internal model control [3] and flatness-based control [4]. Compared with other model-based control concepts, MPC [5] is attractive, as it is able to directly consider constraints on the actuated signals and the system states. In case of airpath control, this is an important feature, since not only the actuators are constrained, but also limits on the turbocharger speeds must be respected.

Among the applied MPC concepts, online optimization was investigated as well as offline optimization-based implementations. Explicit MPC is used, e.g., in [6] based on a piecewise-affine (PWA) approximation of the system behavior. On the other hand, different MPC concepts based on online optimization have also been investigated, where different measures were used to take the nonlinearity into account. The majority of the previous publications use linear time-varying (LTV) MPC, which means that the optimal control problem (OCP) is based on a linearized model gained at the recent operating point, see [5] or [7]. More elaborate NMPC control was recently investigated in [8]–[10], where it is applied to single-stage diesel engines with a data-based model. When an (N)MPC algorithm is applied to a real-world system, all components, such as observer, problem formulation, and optimization algorithm, have to match each other. An additional important challenge is the realization of offset-free control, for which different techniques are available, such as virtual references [11] or rate-based MPC [12].

However, for the control of the more complex two-stage gasoline turbocharging architecture, only a small number of publications are available. The work in [13] presents a vehicle setup of the two-stage turbocharging architecture for gasoline engines. The authors utilize a single-input single-output closed-loop PID controller. For lower engine speeds, the HP stage is controlled (LP wastegate fully open), and for higher engine speeds, the LP stage is controlled (HP wastegate fully open). In the medium speed range, one of the two actuated values is used in a feedforward manner with lookup tables, and the other one is used in feedback. In [14], the modeling and control of the pneumatic actuation system for a two-stage gasoline airpath architecture is carried out. An MPC approach based on a PWA approximation of the turbocharging concept was presented in [5]. The PWA-based MPC works well for small load steps, e.g., for steps starting at throttled operation to a boost pressure of 1.4 bar. For steps to boost pressures larger than 1.5 bar, where the nonlinearity is stronger, the control results show considerable overshoots that are not tolerable for real-world driving.

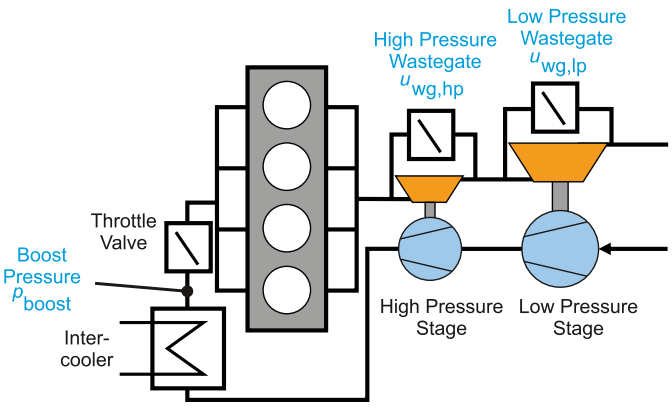

Fig. 1. System overview of the investigated two-stage turbocharging concept.

This paper is based on a previous publication of the authors [15], where initial closed-loop simulations were carried out in nominal operation. The concept has been extended, such that it allows for real-time control, e.g., including the observer, deadtime compensation, and real-time feasibility. Moreover, the control algorithm was implemented on a rapid prototyping hardware, and vehicle experiments have been carried out on a dynamometer and on the road for validation of the control algorithm.

C. Main Contributions and Outline

The main contribution of this paper is the development of an NMPC scheme that allows meeting the tough demands on the closed-loop control of the two-stage turbocharging architecture for gasoline engines. For validation of the control concept, in-vehicle experiments are conducted. The major bottleneck for the implementation of NMPC is typically the computational effort. For this purpose, a sufficiently reduced-order state-space modeling approach is proposed, which mainly relies on physical equations. The resulting model is a system of differential-algebraic equations (DAEs), covering the entire operating range. Direct optimal control using an online sequential quadratic programming (SQP)-type algorithm is used as a basis for the proposed implementation. To allow for real-time feasible computations, the combination of a code-generated real-time iteration (RTI) scheme with the presented reduced-order model is used.

This paper is organized as follows. Section II provides an overview on the setup of the system. The reduced-order modeling of the system is described in Section III. The implementation of the NMPC scheme is discussed in Section IV. Section V presents the closed-loop simulation results for the control algorithm. The validation of the controller using experiments on a vehicle dynamometer is described in Section VI, and SectionVII presents the vehicle experiments on the road while driving dynamically.

II. Two-Stage Turbocharged Gasoline Engine

A. System Setup

Fig. 1 shows the schematic setup of the investigated airpath architecture. For an experimental analysis of the system,

the depicted architecture was built up and implemented in a demonstrator vehicle (Ford Focus) with a 1.8 L four-cylinder gasoline engine. A rapid control prototyping hardware is used in the vehicle, on which the control algorithm has been implemented. The overall setup allows for real-world experiments to validate the control performance. For the validation of the system, closed-loop experiments have been conducted, using a vehicle dynamometer as well as by performing driving experiments on the road at an automotive testing center. In [16], a more detailed overview on the turbocharging system is given.

B. Available Sensors

The controlled variable of the system is the boost pressure p_{boost},[1] which is measured with a pressure sensor positioned behind the intercooler and in front of the throttle valve. Due to the high exhaust gas temperatures in gasoline engines, the application of sensors in the exhaust gas path is impractical for a series configuration due to durability and price issues. Thus, in the presented control approach, no sensor signal from the exhaust gas path is used for control purposes. Additionally used measured variables are the engine speed n_{eng} and the ambient pressure p_{amb}. Therefore, all sensors used for control are typically available in a series production configuration. For the purpose of modeling and validation of the algorithm, additional sensors are implemented in the car. For example, both turbocharger speeds and the pressure between the compressors (providing the pressure ratio over the two compressors) are measured in the demonstrator vehicle only for modeling and validation purposes.

C. Actuation System

As actuators for the control of the turbochargers, wastegates on the HP ($u_{\text{wg,hp}}$) and the LP stage ($u_{\text{wg,lp}}$) are used. Commonly, electronic wastegates are used for turbocharging, which have the advantage that they have a position feedback sensor and thus allow for accurate setting of the valve opening area. In this paper, the use of alternative, simpler, and cheaper pneumatic actuators is investigated. They do not use any additional sensor, e.g., for position feedback, which makes the control more demanding. The wastegate actuation signals correspond to a pulsewidth modulated (PWM) signal, which has an allowable operating range $u_{\text{wg,hp}} = [0 \dots 100]\%$ and $u_{\text{wg,lp}} = [0 \dots 100]\%$. It manipulates the pilot pressure, which has influence on the cross-sectional diameter of the opening. Thus, the amount of exhaust gas, which passes the turbine or the wastegate, can be adjusted. Low values, such as $u_{\text{wg}} = 0\%$, open the wastegate, and higher values, such as $u_{\text{wg}} = 100\%$, close the wastegate. In case of a fully open wastegate, the majority of the exhaust gas bypasses the turbine, whereas in case of a fully closed wastegate, all the exhaust gas flows through the turbine. In addition, a throttle valve is present in the considered architecture, which is controlled via a separate function, not part of the turbocharger control. No additional HP bypass for the compressor was used in the control concept, accepting efficiency losses at high mass flows. At very high mass flows, the HP stage does not deliver boost pressure anymore. However, with the given sizing of the components, the HP stage does not go into a safety critical mode.

[1]All relevant variables and indices are summarized further in the corresponding tables of Appendix B.

D. Engine Control Algorithm

In summary, the controlled variable is the boost pressure p_{boost} and the actuated values are the HP and LP wastegate PWM signals $u_{\text{wg,hp}}$ and $u_{\text{wg,lp}}$. Additionally, the engine speed n_{eng} and the ambient pressure p_{amb} are used as measured disturbances. The throttle valve is only used in the nonboosted region, as is typical also in a series configuration. In the boosted region, the throttle valve is set completely open for reasons of fuel efficiency. As a consequence, the boost pressure is equal to the intake manifold pressure and directly correlated with the torque of the engine. Therefore, the throttle valve is not investigated further. The setpoint for boost pressure is determined by the requested torque in a conventional manner, as given in torque oriented engine control structures. All other parameters of the engine control structure, such as ignition, injection, and camshaft position, are based on the standard calibration. The airpath NMPC control algorithm is implemented on a dSpace MicroAutoBox. Note that all other engine control functions, such as setpoint calculation and ignition timing, are also implemented on the same MicroAutoBox system.

III. Modeling of the Two-Stage Turbocharging

In the following, the nonlinear state-space model is introduced, which is used within the NMPC problem formulation. For the controller internal model, the focus is set on capturing the system dynamics while trying to keep the function outputs as well as their derivatives smooth and the state dimensions small. In a previous publication of the authors [15], a nonlinear state-space model has been presented, which is mainly driven by physical equations. In [15], the model was validated against static and dynamic measurement data. The model has shown to reproduce the measurement data quantitatively well in a big operating range concerning engine speed and engine load. These findings are summarized in the following.

A. Fundamental Equations of Two-Stage Turbocharging

The well-investigated physical equations are used as a basis for the model, and a detailed overview on turbocharger modeling can be found in [2]. The power Ψ on turbine and compressor is used to describe the power balance on the HP and the LP stage as described in (1). The expressions in (2) and (3) relate the power on the two stages of the compressor and, respectively, the turbine to the mass flow through the corresponding device and the total change of enthalpy. In the equations, the pressure ratio is given by $\Pi_t = p_{\text{ut}}/p_{\text{dt}}$ and $\Pi_c = p_{\text{dc}}/p_{\text{uc}}$. The mass flow is given by $\dot{m}$, the turbocharger speed by n_{tc}, the specific isobaric heat capacity by c_p, the polar mass moment of inertia of the turbocharger by Θ_{tc}, the isentropic exponent by κ, the isentropic efficiency by η_s, the temperature upstream of the compressor

by T_{uc}, and the upstream turbine by T_{ut}. Expressions given with $\star$ hold for the HP and the LP stage $\star = \mathrm{lp}, \mathrm{hp}$

$$\frac{d}{\mathrm{d}t}\left(\frac{1}{2}\Theta_{\mathrm{tc},\star} n_{\mathrm{tc},\star}^2\right) = \Psi_{t,\star} - \Psi_{c,\star} \tag{1}$$

$$\Psi_{c,\star} = \dot{m}_{c,\star} c_{p,\mathrm{air}} T_{\mathrm{uc},\star} \frac{1}{\eta_{s,c,\star}} \left(\Pi_{c,\star}^{\frac{\kappa_{\mathrm{air}}-1}{\kappa_{\mathrm{air}}}} - 1\right) \tag{2}$$

$$\Psi_{t,\star} = \dot{m}_{t,\star} c_{p,\mathrm{exh}} T_{\mathrm{ut},\star} \eta_{s,t,\star} \left(1 - \Pi_{t,\star}^{\frac{1-\kappa_{\mathrm{exh}}}{\kappa_{\mathrm{exh}}}}\right). \tag{3}$$

The aspirated mass flow of the engine and the fuel mass flow can be obtained, using the volumetric efficiency η_v, the volume of the cylinder V_{cyl}, gas constant R, the stochiometric air requirement λ_s, and the intake manifold temperature T_{im} as follows (see [17]):

$$\dot{m}_{\mathrm{asp}} = \frac{\eta_v p_{\mathrm{amb}} \Pi_{c,\mathrm{lp}} \Pi_{c,\mathrm{hp}} V_{\mathrm{cyl}} n_{\mathrm{eng}}}{R_{\mathrm{air}} T_{\mathrm{im}} 120} \tag{4}$$

$$\dot{m}_{\mathrm{fuel}} = \frac{\dot{m}_{\mathrm{asp}}}{\lambda_s} \tag{5}$$

where a stochiometric operation (air-to-fuel ratio $\lambda = 1$) is assumed. For calculating the mass flow through the HP and LP wastegate and through the turbines, the throttle equation can be used with A as an opening area

$$\dot{m}_{t,\star} = f_{t,\star} \frac{p_{\mathrm{dt},\star}}{\sqrt{T_{\mathrm{ut},\star}}} \tag{6}$$

$$f_{t,\star} \equiv A_{t,\star} \frac{\Pi_{t,\star}^{\frac{3}{2}}}{\sqrt{R_{\mathrm{exh}}}} \cdot \sqrt{\frac{2\kappa_{\mathrm{exh}}}{\kappa_{\mathrm{exh}}-1}\left(\Pi_{t,\star}^{\frac{-2}{\kappa_{\mathrm{exh}}}} - \Pi_{t,\star}^{\frac{-\kappa_{\mathrm{exh}}-1}{\kappa_{\mathrm{exh}}}}\right)} \tag{7}$$

$$\dot{m}_{\mathrm{wg},\star} = f_{\mathrm{wg},\star} \frac{p_{\mathrm{dt},\star}}{\sqrt{T_{\mathrm{ut},\star}}} \tag{8}$$

$$f_{\mathrm{wg},\star} \equiv A_{\mathrm{wg},\star} \frac{\Pi_{t,\star}}{\sqrt{R_{\mathrm{exh}}}} \cdot \sqrt{\frac{2\kappa_{\mathrm{exh}}}{\kappa_{\mathrm{exh}}-1}\left(\Pi_{t,\star}^{\frac{-2}{\kappa_{\mathrm{exh}}}} - \Pi_{t,\star}^{\frac{-\kappa_{\mathrm{exh}}-1}{\kappa_{\mathrm{exh}}}}\right)}. \tag{9}$$

B. Model Simplification

Various simplifications have to be done to gain a model suitable for real-time NMPC, as described in the following.

1) Rotational Kinetic Energy and Pressure Ratio: To simplify the model, the rotational kinetic energy can be correlated with the pressure ratio over its corresponding compressor. For the LP stage, an affine map with the parameters a_{lp} and b_{lp} can be fit via linear regression on measurement data

$$n_{\mathrm{tc,lp}}^2 = a_{\mathrm{lp}} \Pi_{c,\mathrm{lp}} + b_{\mathrm{lp}}. \tag{10}$$

For the HP stage, the following bilinear function has been shown to work well [15]:

$$n_{\mathrm{tc,hp}}^2 = (a_{\mathrm{hp}} n_{\mathrm{eng}} + c_{\mathrm{hp}}) \Pi_{c,\mathrm{hp}} + (b_{\mathrm{hp}} n_{\mathrm{eng}} + d_{\mathrm{hp}}). \tag{11}$$

2) Singular Perturbation: For the gasoline two-stage turbocharging, the model simplification by a singular perturbation theory is applicable as described in [18]. This results in the following equations:

$$\dot{m}_{c,\mathrm{lp}} = \dot{m}_{c,\mathrm{hp}} = \dot{m}_{\mathrm{asp}} \tag{12}$$

and

$$\dot{m}_{\mathrm{asp}} + \dot{m}_{fuel} = \dot{m}_{t,\mathrm{hp}} + \dot{m}_{\mathrm{wg,hp}} \tag{13}$$

$$= \dot{m}_{t,\mathrm{lp}} + \dot{m}_{\mathrm{wg,lp}}. \tag{14}$$

3) Modeling of the Wastegate Opening Area: Pneumatic systems are used for actuating the wastegate opening area in the given setup. The HP wastegate is actuated via underpressure and the LP stage via excesspressure, which makes the system behavior of each stage different. The PWM signal $u_{\mathrm{wg,hp}}$ adjusts the ratio of underpressure delivered by a vacuum pump, which is mixed with ambient pressure. In contrast to that, the LP stage uses the charging pressure to actuate the wastegate. The PWM signal $u_{\mathrm{wg,lp}}$ adjusts the ratio of charging pressure, which is mixed with the ambient pressure. In both cases, the opening area can be calculated by force equilibrium of the different acting forces. For gaining fast computations, the resulting opening area characteristic has been determined by simulations and was afterward approximated by a smooth function. The HP stage opening area $A_{\mathrm{wg,hp}}$ depends mainly on the actuated signal, and it correlates linearly, such that a simple model can be used

$$A_{\mathrm{wg,hp}} = \left(1 - \frac{u_{\mathrm{wg,hp}}}{100}\right) A_{\mathrm{wg,hp,max}}. \tag{15}$$

In contrast to that, the dependence of the LP stage opening area on the current boost pressure cannot be neglected. It was modeled by two sigmoid functions

$$A_{\mathrm{wg,lp}} = f_1(\Pi_{c,\mathrm{lp}} \Pi_{c,\mathrm{hp}}) f_2(u_{\mathrm{wg,lp}}) \tag{16}$$

with

$$f_i(x) = a_i + \frac{b_i}{1 + e^{\frac{-x+c_i}{d_i}}} \tag{17}$$

where a_i, b_i, c_i, and d_i are constants for $i = 1, 2$, which need to be identified. For the dynamics of the actuator, only a deadtime is considered, as detailed in Section III-B4.

4) Deadtime: When the turbocharger architecture is implemented in a vehicle, a quite considerable deadtime results. The deadtime arises between the change of the manipulated variables ($u_{\mathrm{wg,hp}}$ and $u_{\mathrm{wg,lp}}$) and the reaction of the system referring to the pressure ratios over the compressor ($\Pi_{c,\mathrm{lp}}$ and $\Pi_{c,\mathrm{hp}}$) and, consequently, also the controlled variable (p_{boost}). The main reason for the deadtime can be found in the pneumatic actuation as a change in the PWM signal must result in a pressure change inside the wastegate actuator before the wastegate position changes. An additional reason for the deadtime can be found in the volumes of the air and exhaust path as well as the components, such as the intercooler. The volumes have to be filled or depleted, and the components have to be passed by the gas. Instead of modeling all these effects

individually, they are summarized as one overall deadtime t_D, which acts directly on the input

$$u_{\mathrm{wg},\star} = u_{\mathrm{wg},\star}(t - t_D). \tag{18}$$

Based on the measurement data, the resulting deadtime is estimated to be $t_D = 0.45$ s. This value is quite large and makes the control task more challenging for NMPC.

C. Resulting State-Space Model

Within the NMPC problem formulation, the resulting state-space model is used that consists of a set of DAEs. For this purpose, let $\boldsymbol{x}(t) \in \mathbb{R}^{n_x}$ denote the differential states, $\dot{\boldsymbol{x}}(t)$ the differential state derivatives, $\boldsymbol{z}(t) \in \mathbb{R}^{n_z}$ the algebraic variables, and $\boldsymbol{u}(t) \in \mathbb{R}^{n_u}$ the control inputs and $\dot{\boldsymbol{u}}(t)$ their derivatives. The system is governed by introducing the differential states $x_1 = \Pi_{c,\mathrm{lp}}$, $x_2 = \Pi_{c,\mathrm{hp}}$, and the algebraic states $z_1 = \Pi_{t,\mathrm{lp}}$, $z_2 = \Pi_{t,\mathrm{hp}}$. The model constants can be summarized as $c_1, \ldots, c_8$, which are described further in Appendix A. Based on the latter parameters and setting $\kappa_{\mathrm{air}} = 1.4$ as well as $\kappa_{\mathrm{exh}} = 1.33$, the resulting DAE system reads as

$$\begin{aligned}\dot{\Pi}_{c,\mathrm{lp}}(t) = {} & c_1 p_{\mathrm{amb}}\left(\Pi_{t,\mathrm{lp}}^{1.5} - \Pi_{t,\mathrm{lp}}^{1.25}\right)\sqrt{\Pi_{t,\mathrm{lp}}^{-1.5} - \Pi_{t,\mathrm{lp}}^{-1.75}} \\ & - c_2 p_{\mathrm{amb}} n_{\mathrm{eng}} \Pi_{c,\mathrm{hp}}\left(\Pi_{c,\mathrm{lp}}^{1.29} - \Pi_{c,\mathrm{lp}}\right) \qquad (19)\\ 0 = {} & \Pi_{c,\mathrm{lp}}\Pi_{c,\mathrm{hp}} - c_3 \frac{1}{n_{\mathrm{eng}}}\sqrt{\Pi_{t,\mathrm{lp}}^{0.5} - \Pi_{t,\mathrm{lp}}^{0.25}} \\ & \cdot \left(\sqrt{\Pi_{t,\mathrm{lp}}} + c_4 A_{\mathrm{wg,lp}}(\Pi_{c,\mathrm{lp}}\Pi_{c,\mathrm{hp}}, u_{\mathrm{wg,lp}}(t - t_D))\right) \qquad (20)\\ \dot{\Pi}_{c,\mathrm{hp}}(t) = {} & c_5 p_{\mathrm{amb}}\Pi_{t,\mathrm{lp}}\left(\Pi_{t,\mathrm{hp}}^{1.5} - \Pi_{t,\mathrm{hp}}^{1.25}\right)\sqrt{\Pi_{t,\mathrm{hp}}^{-1.5} - \Pi_{t,\mathrm{hp}}^{-1.75}} \\ & - c_6 p_{\mathrm{amb}} n_{\mathrm{eng}} \Pi_{c,\mathrm{lp}}\left(\Pi_{c,\mathrm{hp}}^{1.29} - \Pi_{c,\mathrm{hp}}\right) \qquad (21)\\ 0 = {} & \Pi_{c,\mathrm{lp}}\Pi_{c,\mathrm{hp}} - c_7 \frac{1}{n_{\mathrm{eng}}}\Pi_{t,\mathrm{lp}}\sqrt{z_2^{0.5} - \Pi_{t,\mathrm{hp}}^{0.25}} \\ & \cdot \left(\sqrt{\Pi_{t,\mathrm{hp}}} + c_8 A_{\mathrm{wg,hp}}(u_{\mathrm{wg,hp}}(t - t_D))\right) \qquad (22)\end{aligned}$$

with the output functions

$$p_{\mathrm{boost}} = p_{\mathrm{amb}}\Pi_{c,\mathrm{lp}}\Pi_{c,\mathrm{hp}} \tag{23}$$

$$n_{\mathrm{tc,lp}} = \sqrt{(a_{\mathrm{lp}}\Pi_{c,\mathrm{lp}} + b_{\mathrm{lp}})} \tag{24}$$

$$n_{\mathrm{tc,hp}} = \sqrt{(a_{\mathrm{hp}} n_{\mathrm{eng}} + c_{\mathrm{hp}})\Pi_{c,\mathrm{hp}} + (b_{\mathrm{hp}} n_{\mathrm{eng}} + d_{\mathrm{hp}})}. \tag{25}$$

The DAE model is of index 1 and consists of two differential and two algebraic states. The output y_1 corresponds to the boost pressure p_{boost}, which forms the tracking objective. The outputs y_2 and y_3 correspond to the LP $n_{\mathrm{tc,lp}}$ and the HP $n_{\mathrm{tc,hp}}$ turbocharger speeds, which will be constrained as part of the OCP.

IV. Nonlinear Model Predictive Control and State Estimation

This section gives an overview on the entire closed-loop system, as shown in Fig. 2. The key element is the direct optimal control method, which has been applied, in this paper, to obtain a fast NMPC implementation. An extended Kalman Filter (EKF) is used to observe the state of the system, and additionally, a deadtime compensator (DTC) is implemented to account for the considerable deadtime.

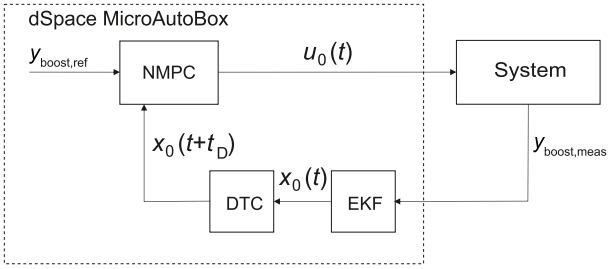

Fig. 2. Illustration of the closed-loop system on the dSpace MicroAutoBox: NMPC based on ACADO code generation with DTC and EKF.

A. Disturbance Model and State Observer

One of the goals for the closed-loop controller is to achieve offset-free tracking, even in the presence of disturbances and model-plant mismatch. More precisely, our aim is that the boost pressure y_{boost} is reference tracked without offset for the case that the reference and the disturbances are asymptotically constant. For this purpose, a pure output disturbance model can be used as discussed in [19]. The boost pressure calculation is, therefore, augmented by an additional disturbance state $d(t)$

$$y_{\mathrm{boost}}(t) = p_{\mathrm{amb}} x_1 x_2 + d(t) \tag{26}$$

where the disturbance variable is assumed to be constant over the time horizon, i.e., $\dot{d} = 0$. By augmenting the state vector

$$\bar{x}(t) = \begin{bmatrix} x(t) \\ d(t) \end{bmatrix} \tag{27}$$

the overall system equations, including the disturbance model, can be represented as

$$0 = f(\dot{\bar{\boldsymbol{x}}}(t), \bar{\boldsymbol{x}}(t), \boldsymbol{z}(t), \boldsymbol{u}(t - t_D)) \tag{28}$$

$$\boldsymbol{y} = g(\bar{\boldsymbol{x}}(t)). \tag{29}$$

All system, disturbance, and output variables are assumed to exhibit zero mean Gaussian white noise. An EKF filter is used for the estimation of the system states and disturbances, as described in [20]. The EKF updates the state estimates with a sampling time $t_{\mathrm{samp}} = 0.01$ s. For the relation between continuous time t and the discrete time index $k \in \{1, 2, 3, \ldots\}$, it follows $t = t_0 + k \cdot t_{\mathrm{samp}}$, and for the deadtime, $k_D = \lfloor t_D / t_{\mathrm{samp}} \rfloor$. Denoting $\bar{\boldsymbol{x}}_{\mathrm{es},k}^{-}$ as the *a priori* state estimate at time index k and the *a priori* error covariance as $\mathbf{P}_k^{-}$ and defining $J_{x,k} = (\partial\phi(\cdot)/\partial x_k)|_{\bar{x}_{\mathrm{es},k}, u_{k-k_D}}$ and $J_{y,k} = (\partial g(\cdot)/\partial x_k)|_{\bar{x}_{\mathrm{es},k}^{-}}$, the prediction step reads as

$$\bar{\boldsymbol{x}}_{\mathrm{es},k}^{-} = \phi(\bar{\boldsymbol{x}}_{\mathrm{es},k-1}, \boldsymbol{u}_{k-1-k_D}) \tag{30}$$

$$\boldsymbol{P}_k^{-} = \boldsymbol{J}_{x,k-1} \cdot \boldsymbol{P}_{k-1} \cdot \boldsymbol{J}_{x,k-1}^{T} + \boldsymbol{Q}_{\mathrm{KF}}. \tag{31}$$

The function $\phi(\boldsymbol{x}_i, \boldsymbol{u}_i)$ denotes the numerical simulation of the nonlinear dynamics throughout one sampling step, starting from the given state $\boldsymbol{x}_i$ and using the control inputs $\boldsymbol{u}_i$. Since the model consists of an implicit DAE system in (28), and also, an implicit integration method is needed as discussed in [21]. For calculating the Jacobian matrices, algorithmic differentiation [22] techniques are used in combination with a tailored sensitivity propagation for the implicit Runge–Kutta integration method [21]. The correction step, calculating the

a posteriori estimates for the system state $\bar{\boldsymbol{x}}_{\mathrm{es},k}$ and the error covariance $\boldsymbol{P}_k$, reads as

$$\boldsymbol{L}_k = \frac{\boldsymbol{P}_k^- \cdot \boldsymbol{J}_{y,k}^T}{\boldsymbol{J}_{y,k} \cdot \boldsymbol{P}_k^- \cdot \boldsymbol{J}_{y,k}^T + \boldsymbol{R}_{\mathrm{KF}}} \tag{32}$$

$$\bar{\boldsymbol{x}}_{\mathrm{es},k} = \bar{\boldsymbol{x}}_{\mathrm{es},k}^- + L_k \cdot \left(\boldsymbol{y}_{\mathrm{meas},k} - g\left(\bar{\boldsymbol{x}}_{\mathrm{es},k}^-\right)\right) \tag{33}$$

$$\boldsymbol{P}_k = (\mathbb{1} - \boldsymbol{L}_k \cdot \boldsymbol{J}_{y,k}) \cdot \boldsymbol{P}_k^-. \tag{34}$$

B. Deadtime Compensation

The EKF provides the current state estimate $\bar{\boldsymbol{x}}_{\mathrm{es},k}$, which can be used along with the model (28) and (29) in the NMPC algorithm to compute the next control input. However, the deadtime is typically compensated for based on a prediction, which results in a smaller problem size for optimization. A delay-free model can be considered based on the same system dynamics as (28) and (29)

$$0 = f(\dot{\hat{\boldsymbol{x}}}(t), \hat{\boldsymbol{x}}(t), \hat{\boldsymbol{z}}(t), \boldsymbol{u}(t)) \tag{35}$$

$$\boldsymbol{y} = g(\hat{\boldsymbol{x}}(t)). \tag{36}$$

The new states $\hat{\boldsymbol{x}}_{\boldsymbol{k}}$ correspond to the states $\bar{\boldsymbol{x}}_{\boldsymbol{k+k_D}}$, shifted by the deadtime. The predicted states $\hat{\boldsymbol{x}}_{\mathrm{es},k}$ need to be obtained, using the inputs which have already been applied to the plant. Note that for this purpose, the last k_D inputs need to be stored. Using $\Phi(\bar{x}_i, \boldsymbol{u}_{i-k_D}, \ldots, \boldsymbol{u}_{i-1})$ to denote the numerical integration from $i \to i + k_D$ starting from the initial state $\bar{x}_i$ with the actuated signals $\boldsymbol{u}_{i-k_D}, \ldots, \boldsymbol{u}_{i-1}$, it follows:

$$\hat{\boldsymbol{x}}_{\mathrm{es},k} = \Phi(\bar{x}_{\mathrm{es},k}, \boldsymbol{u}_{k-k_D}, \ldots, \boldsymbol{u}_{k-1}). \tag{37}$$

C. Optimal Control Problem Formulation

The NMPC scheme needs to solve one nonlinear OCP at each sampling instant. Let us introduce the following continuous time OCP formulation for the delay-free model:

$$\min_{\boldsymbol{x}(\cdot),\boldsymbol{u}(\cdot)} \frac{1}{2}\int_0^{t_{\mathrm{ch}}} F\left(\boldsymbol{x}(t), \boldsymbol{u}(t)\right) \mathrm{d}t + \frac{1}{2} F_N\left(\boldsymbol{x}(t_{\mathrm{ch}})\right) \tag{38}$$

$$\text{s.t. } 0 = \boldsymbol{x}(0) - \hat{\boldsymbol{x}}_{\mathrm{es}}, \tag{39}$$

$$0 = f(\dot{\boldsymbol{x}}(t), \boldsymbol{x}(t), \boldsymbol{z}(t), \boldsymbol{u}(t)) \quad \forall t \in [0, t_{\mathrm{ch}}] \tag{40}$$

with the following additional path constraints on the inputs and outputs of the system:

$$\underline{y}_2 - s(t) \le y_2(t) \le \bar{y}_2 + s(t) \quad \forall t \in [0, t_{\mathrm{ch}}] \tag{41}$$

$$\underline{y}_3 - s(t) \le y_3(t) \le \bar{y}_3 + s(t) \quad \forall t \in [0, t_{\mathrm{ch}}] \tag{42}$$

$$\underline{u}_{\mathrm{wg,lp}} \le u_{\mathrm{wg,lp}}(t) \le \bar{u}_{\mathrm{wg,lp}} \quad \forall t \in [0, t_{\mathrm{ch}}] \tag{43}$$

$$\underline{u}_{\mathrm{wg,hp}} \le u_{\mathrm{wg,hp}}(t) \le \bar{u}_{\mathrm{wg,hp}} \quad \forall t \in [0, t_{\mathrm{ch}}] \tag{44}$$

$$\underline{\dot{u}}_{\mathrm{wg,lp}} \le \dot{u}_{\mathrm{wg,lp}}(t) \le \bar{\dot{u}}_{\mathrm{wg,lp}} \quad \forall t \in [0, t_{\mathrm{ch}}] \tag{45}$$

$$\underline{\dot{u}}_{\mathrm{wg,hp}} \le \dot{u}_{\mathrm{wg,hp}}(t) \le \bar{\dot{u}}_{\mathrm{wg,hp}} \quad \forall t \in [0, t_{\mathrm{ch}}] \tag{46}$$

$$0 \le s(t) \quad \forall t \in [0, t_{\mathrm{ch}}], \tag{47}$$

and the stage and terminal cost functions defined as

$$F(\boldsymbol{x}(t), \boldsymbol{u}(t)) = \|y_1(t) - y_{\mathrm{boost,ref}}(t)\|_Q^2 + \|\boldsymbol{u}(t)\|_{\boldsymbol{R}}^2$$

$$F_N(\boldsymbol{x}(t_{\mathrm{ch}})) = \|y_1(t_{\mathrm{ch}}) - y_{\mathrm{boost,ref}}(t_{\mathrm{ch}})\|_{Q_N}^2. \tag{48}$$

This nonlinear OCP depends on the parameter $\hat{\boldsymbol{x}}_{\mathrm{es}} \in \mathbb{R}^{n_x}$, which denotes the current state estimate, through the initial value condition of (39). The dynamic optimization problem is defined over the control horizon t_{ch}. The change rates of the controls and the slack variable are optimized directly $\boldsymbol{u} := [\dot{u}_{\mathrm{wg,hp}}, \dot{u}_{\mathrm{wg,lp}}, s]^\top \in \mathbb{R}^3$. The slack variable $s(t)$ is used to penalize violations of the constraints in (41) and (42), which allows to always have a feasible optimization problem.

The NMPC objective consists of a least squares tracking cost, note that the reference signal $y_{\mathrm{boost,ref}}(t)$ is used as a constant value over the prediction horizon. The path constraints consist of simple bounds defined by (41)–(47). The implicit dynamics from (40) are given by the DAE system in (19)–(22). As the DAE is of index 1, the Jacobian matrix $(\partial f(\cdot)/\partial\ (z, \dot{x}))$ is invertible, which is needed by most numerical simulation methods.

D. Direct Multiple Shooting Discretization

Direct optimal control methods proceed by first forming a discrete approximation of the continuous time OCP from (38)–(47), such that one can solve the resulting tractable nonlinear program (NLP). For the sake of simplicity, an equidistant grid of N shooting intervals is considered over the control horizon consisting of time points t_i, where $t_N - t_0 = t_{\mathrm{ch}}$ and $t_{i+1} - t_i = (t_{\mathrm{ch}}/N)$ for $i = 0, \ldots, N-1$. Additionally, the notation will be simplified by restricting to a piecewise constant control parameterization $\boldsymbol{u}(\tau) = \boldsymbol{u}_i$ for $\tau \in [t_i, t_{i+1})$.

A popular approach is known as *direct multiple shooting*, which was proposed in [23]. For the OCP in (38)–(47), the following NLP results:

$$\min_{X,\,U} \frac{1}{2}\sum_{i=0}^{N-1} F\left(\boldsymbol{x}_i, \boldsymbol{u}_i\right) + \frac{1}{2} F_N\left(\boldsymbol{x}_N\right) \tag{49}$$

$$\text{s.t. } 0 = \boldsymbol{x}_0 - \hat{\boldsymbol{x}}_{\mathrm{es}} \tag{50}$$

$$0 = \boldsymbol{x}_{i+1} - \phi\left(\boldsymbol{x}_i, \boldsymbol{u}_i\right), \quad i = 0, \ldots, N-1 \tag{51}$$

including the path constraints

$$\underline{y}_2 - s_i \le y_{2,i} \le \bar{y}_2 + s_i, \quad i = 0, \ldots, N \tag{52}$$

$$\underline{y}_3 - s_i \le y_{3,i} \le \bar{y}_3 + s_i, \quad i = 0, \ldots, N \tag{53}$$

$$\underline{u}_{\mathrm{wg,lp}} \le u_{\mathrm{wg,lp},i} \le \bar{u}_{\mathrm{wg,lp}}, \quad i = 0, \ldots, N \tag{54}$$

$$\underline{u}_{\mathrm{wg,hp}} \le u_{\mathrm{wg,hp},i} \le \bar{u}_{\mathrm{wg,hp}}, \quad i = 0, \ldots, N \tag{55}$$

$$\underline{\dot{u}}_{\mathrm{wg,lp}} \le \dot{u}_{\mathrm{wg,lp},i} \le \bar{\dot{u}}_{\mathrm{wg,lp}}, \quad i = 0, \ldots, N-1 \tag{56}$$

$$\underline{\dot{u}}_{\mathrm{wg,hp}} \le \dot{u}_{\mathrm{wg,hp},i} \le \bar{\dot{u}}_{\mathrm{wg,hp}}, \quad i = 0, \ldots, N-1 \tag{57}$$

$$0 \le s_i, \quad i = 0, \ldots, N-1 \tag{58}$$

and the discrete stage and terminal cost functions read as

$$F(\cdot) = \|y_{1,i} - y_{\mathrm{boost,ref}}\|_Q^2 + \|\boldsymbol{u}_i\|_{\boldsymbol{R}}^2 \tag{59}$$

$$F_N(\cdot) = \|y_{1,N} - y_{\mathrm{boost,ref}}\|_{Q_N}^2 \tag{60}$$

with state trajectory $X = [\boldsymbol{x}_0^\top, \ldots, \boldsymbol{x}_N^\top]^\top$ and control trajectory $U = [\boldsymbol{u}_0^\top, \ldots, \boldsymbol{u}_{N-1}^\top]^\top$. For the sake of simplicity, the latter NLP does not include algebraic states in the decision variables as in [24], since they only enter the DAE system. Note that the penalization of the control change rate is implemented in the dynamic optimization software by including an extra state which denotes the original control value, while its time derivative is defined as the new control input.

E. Real-Time Iterations and Closed-Loop Stability

To solve the NLP in (49)–(58), this paper considers an SQP approach. It solves a sequence of quadratic program (QP) approximations and can converge to a locally optimal solution of the original NLP. Note that one is typically satisfied with finding a local minimizer in case of nonconvex optimization [25], especially within online algorithms for NMPC [26]. Since the OCP consists of a least squares type objective, the generalized Gauss–Newton method [27] is used. All nonlinear functions in the NLP are linearized with automatic differentiation at each SQP iteration. There are different options to efficiently solve the large structured QP [28]. In the proposed algorithm, condensing has been used in combination with qpOASES [29].

In NMPC, one needs to solve this nonlinear OCP at each sampling instant. For this purpose, specific online algorithms have been developed, such as the RTI scheme [30], which is used in our proposed implementation. The idea is to minimize the computational delay between obtaining the new state estimate $\hat{\boldsymbol{x}}_{\mathrm{es}}$ and applying the next control input. It has been shown in [30] that, under some reasonable assumptions, the stability of the closed-loop system based on the RTI scheme can also be guaranteed in presence of inaccuracies, model errors, or external disturbances.

To obtain a real-time feasible NMPC implementation on the embedded control hardware, the ACADO code generation tool has been used as presented in [21] and [31]. It is part of the open-source ACADO Toolkit, which can be downloaded from `www.acadotoolkit.org`. The code generation tool exports highly efficient, standalone C-code implementing the RTI scheme for fast optimal control. It supports the ability to exploit specific model structures as detailed in [21].

V. Simulative Assessment of NMPC

Before testing the NMPC in the vehicle, closed-loop simulations have been conducted. The main purpose of the simulations is the parameterization and validation of the scheme, concerning the aforementioned control requirements. Additionally, its performance has been compared with alternative control algorithms. The simulations are conducted on the control prototyping hardware that is also used in the vehicle. Thus, it is possible to evaluate the computation time of the NMPC algorithm, which allows to check for real-time feasibility during parameterization of the algorithm.

A. Parameterization of NMPC

In case of NMPC, quite a few parameters have to be tuned to realize closed-loop control that satisfies real-time feasibility while achieving high control quality. The control horizon is set to $t_{\mathrm{ch}} = 1.5$ s, which allows to cover the whole dynamics until steady state of the system. If the control horizon is chosen too short, the control performance decreases, and in the worst case, the turbocharger speed limits cannot be respected anymore. In order to maintain real-time feasibility, the NMPC scheme does not use the sampling time of NMPC as the discretization step size. Instead, $N = 20$ intervals are used, which results in a discretization time $t_{\mathrm{dis}} = 75$ ms for each shooting interval. In the simulations, it has shown to be more effective to use a larger discretization time along with a longer control horizon instead of using smaller discretization times and horizon lengths. The implicit Runge–Kutta method, used for simulating the DAE system, is based on a fixed integration step size of $t_{\mathrm{int}} = 37.5$ ms. The RTI scheme is used, based on one SQP iteration per time step. To solve each QP subproblem, condensing with qpOASES is used with $N_{\mathrm{QP}} = 50$, as the maximum number of active set changes.

A terminal cost in combination with a sufficiently long control horizon can be used for obtaining closed-loop stability of the resulting NMPC scheme [32]. For real-time purposes, it is advantageous to avoid the storage or online computation of a terminal region and corresponding cost for each output and disturbance signal. In the simulations and during the conducted experiments, no situations were encountered, where the terminal cost (60) was insufficiently represented. Instead, the used control horizon is sufficient for stabilization in this case, where the stage cost in (59) has been tuned mainly to result in fast reference tracking. For the EKF, the covariance matrices were chosen, such that quick disturbance estimation is achieved, as the present noise is rather weak.

With these parameterization choices, the following values result for the computation time and memory footprint of the NMPC algorithm on the prototyping hardware.

1) *Average Computation Time:* 39 ms.
2) *Maximum Computation Time:* 43 ms.
3) *Memory Usage (Whole ECU incl. NMPC):* 4105 kB.

Note that this particular NMPC scheme solves an NLP that consists of 384 optimization variables, 324 equality constraints, and 260 inequality constraints in every time step. The maximum computation time is observed, when the maximum number of QP-iterations $N_{\mathrm{QP}} = 50$ is reached, which happens, e.g., at the very first time step after a big step in the reference value. In order to always ensure that the solution can be calculated in one sampling step, the sampling time for the whole airpath control was set to $t_{\mathrm{samp}} = 50$ ms.

B. Validation of NMPC

The general functionality of the NMPC algorithm has been compared with linearized MPC schemes. A linear time invariant (LTI) and an LTV MPC controller have been implemented in simulation. The LTI scheme uses one linear model over the entire operating region, whereas the LTV MPC calculates the linearized model around the recent operating point in every time step. In the case of LTI and LTV MPC, the weighting matrices heavily affect the control action. If the weighting factors are chosen too high, overshoots in the boost pressure will result as well as violations of the turbocharger speed limits $n_{\mathrm{tc,hp}}$ and $n_{\mathrm{tc,lp}}$. For this reason, the weighting matrices have been chosen, such that reference tracking is as quick as possible, without allowing for such violations of the turbocharger speed limit and without considerable overshoots.

For one specific reference step, the LTI, LTV MPC, and NMPC can be tuned, such that they show stabilizing control in simulations, are able to consider the constraints on the turbocharger speed limits, and allow for offset-free reference

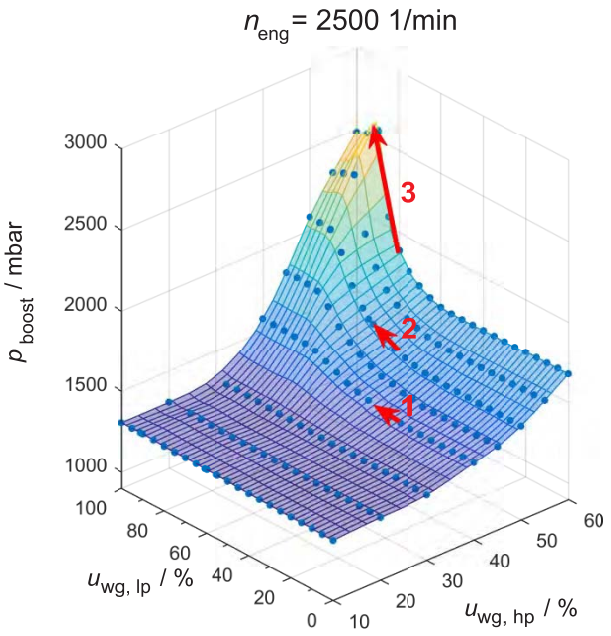

Fig. 3. Steady-state measurement data for $n_{eng} = 2500$ 1/min, along with change in boost pressure for 10% lp-wastegate actuation.

TABLE I

COMPARISON OF TIME t_{95} BETWEEN NONLINEAR MPC, LTV MPC, AND LTI MPC

	Nonlinear MPC	LTV MPC	LTI MPC
Setpoint Step to 2.2bar	1.3 s	1.8 s	3.8 s
Setpoint Step to 2.6bar	1.4 s	1.9 s	3.9 s

TABLE II

COMPARISON OF GRADIENTS AT STEADY-STATE MEASUREMENT DATA FOR $n_{eng} = 2500$ 1/min (FIG. 3)

	$u_{wg,lp}$	$u_{wg,hp}$	Δp_{boost}
1	40% $\rightarrow$ 50%	40%	23 mbar
2	60% $\rightarrow$ 70%	50%	84 mbar
3	80% $\rightarrow$ 90%	60%	680 mbar

tracking. The difference arises in the time t_{95}, which is the time needed for reaching 95% of the reference. Table I gives an overview on t_{95} for different load steps, under nominal conditions for $n_{eng} = 1500$ 1/min. All step responses have been carried out, starting from the same initial conditions.

The difference between NMPC and LTV MPC is around 0.5 s for the two control concepts, which is a considerable and perceptible amount. To explain this further, Fig. 3 along with Table II shows the steady-state behavior (with experimental data) for a constant engine speed. Depending on the operating point, the gradients change significantly. The behavior is qualitatively present in the entire operating range. This has to be considered in a suitable control scheme, especially for a large step in the boost pressure. In contrast to the other concepts, NMPC is aware of these nonlinearities in advance, such that the reference can be tracked much faster. For quick response, the NMPC activates the constraint in transient operation for the HP turbocharger speed limit (a qualitatively comparable closed-loop trace can be found in [15], where this constraint activation can be seen). For the case of LTI and LTV MPC, it is possible that the future gradient is assumed to be too flat. This will result in an aggressive control action, which causes an overshoot on the boost pressure and constraint violations. To counteract this behavior, the weighting matrices have to be adjusted accordingly. By putting less weight on the reference tracking term, this consequently leads to a slower dynamic response.

The closed-loop behavior of LTI MPC is very dependent on the operating point, the linear model that is used, and the tuning of the weights. For a small step in the boost pressure reference, a stable behavior with quite reasonable times for t_{95} can be found, comparable with the ones from LTV MPC. However, if the setpoint change is large, a tuning can still be found that results in stable behavior for this specific step in the operating point. The drawback is that the system response is rather slow, compared with LTV and NMPC. Additionally, it results in an unstable behavior when a setpoint change is applied that is considerably different from the one for which the model and the weighting factors have been tuned.

The closed-loop simulations serve as an important platform for validation of the general functionality of the different MPC schemes and parameterization with respect to characteristics, such as real-time feasibility. As can be seen in Sections VI and VII, the simulations fit qualitatively well with the experimental results that have been gained. However, the values for t_{95} do not quantitatively correspond to the values from the experiments, as different physical processes are not incorporated in the simulation model. For example, in the real operation, the steps start in the throttled operating region and, then, go to the boosted region. In the simulations, the initial condition corresponds to the boost pressure at open throttle.

VI. IN-VEHICLE VALIDATION OF THE CONTROL CONCEPT: VEHICLE DYNAMOMETER

The first step of vehicle testing was done on a vehicle dynamometer. This offers the possibility to test the airpath control for the case of a constant engine speed, which is a major disturbance variable.

In Fig. 4, the results of vehicle testing at the dynamometer with the LTI MPC at constant engine speed $n_{eng} = 2500$ 1/min are shown. As expected by simulation, one can see that small load steps can be managed reasonably well. The time t_{95} is quite high, but the closed-loop control is stable and achieves the reference without any offset. If the load step is higher than what the LTI scheme was originally tuned for, it will result in an oscillatory behavior. Going even higher, in this case, $y_{boost,ref} = 2$ bar, will result in an unstable behavior. The behavior is as expected, as the closed-loop gain will get bigger with increasing loads (see Fig. 3).

In Fig. 5, one exemplary closed-loop control result for the NMPC is depicted. In the given case, a step is applied from the nonboosted region to a high value of boost pressure of $y_{boost,ref} = 2.2$ bar at a constant engine speed of $n_{eng} = 2500$ 1/min on the dynamometer. The closed-loop controller

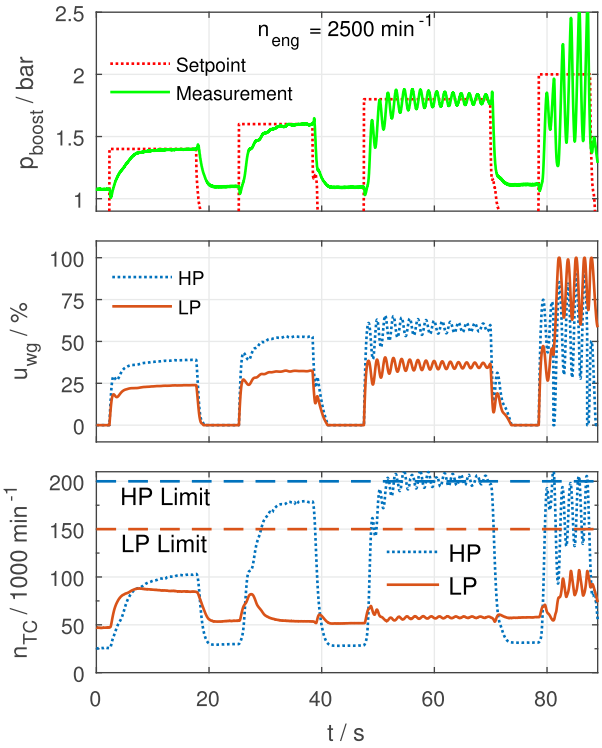

Fig. 4. Experiments on dynamometer: LTI MPC control for steps in the boost pressure reference signals at constant engine speed $n_{\text{eng}} = 2500$ 1/min.

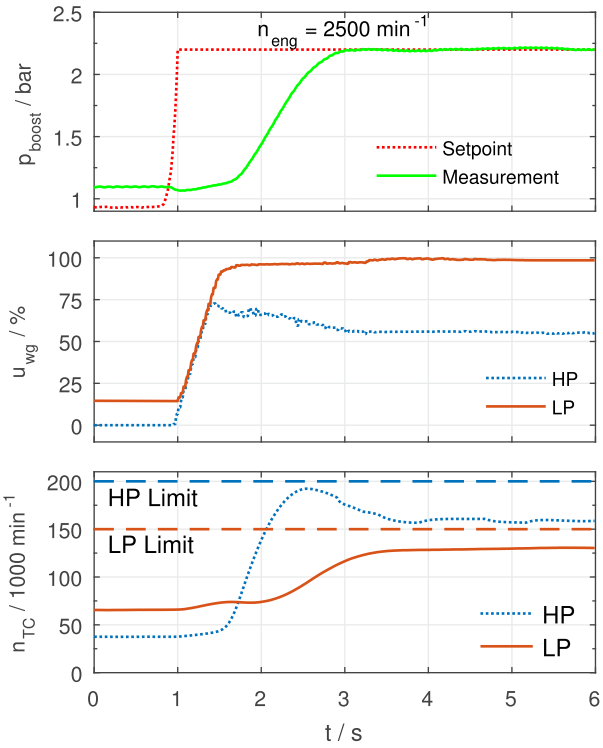

Fig. 5. Experiments on dynamometer: NMPC control for a step in the boost pressure reference signal at constant engine speed $n_{\text{eng}} = 2500$ 1/min.

is able to meet all specified criteria. It is possible to track the reference without overshoot and without offset, by rejecting all disturbances. At the same time, the controller is able to respect the HP and LP turbocharger speed limits. It can be observed that the control scheme is able to exploit the turbocharging architecture. First, the HP stage is used to quickly increase boost pressure. However, it is only used as much, such that the speed limit is not violated. Simultaneously, the speed of the LP stage is increased. Both stages are balanced, such that no overshoot is present, which is especially challenging as $y_{\text{boost}} = y_{\text{boost,ref}}$ is reached before the system goes to steady state. After reaching the setpoint, the actuation as well as the turbocharger speeds are still changing, which shows the advantage of the nonlinear predictive control.

In addition, Fig. 6 shows the control results, where the same boost pressure reference step is applied for different constant engine speeds, all measured at the dynamometer. The control requirements can be met for all engine speeds. Additionally, the pressure ratios are illustrated for each case. They show that at all three engine speeds, the HP stage is used for quick increase in boost pressure. At $n_{\text{eng}} = 2000$ 1/min, the control relies in a relatively similar fashion on the HP and the LP stage in stationary operation. Compared with that, at $n_{\text{eng}} = 3000$ 1/min, the boost pressure is realized almost completely by the LP stage. In conclusion, the control concept is able to account for the changeover between the two stages, such that the design goals of the two-stage turbocharging are realized (high dynamics with HP stage and high power with LP stage).

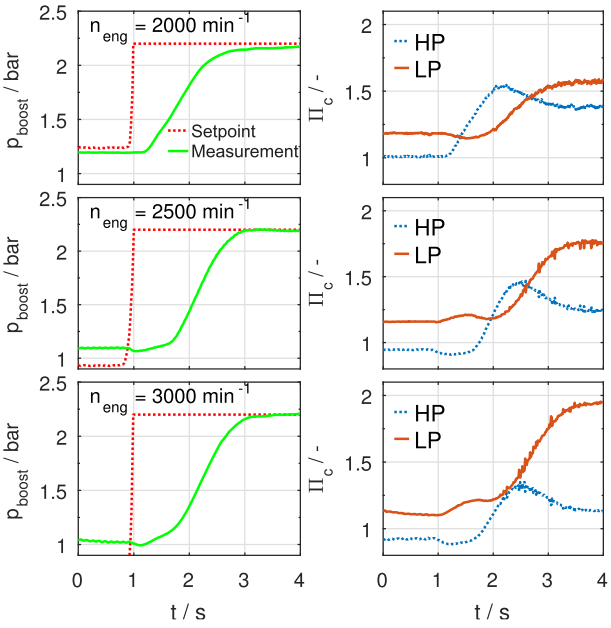

Fig. 6. Experiments on dynamometer: NMPC control for a step in the boost pressure reference signal at different constant engine speeds.

VII. In-Vehicle Validation of the Control Concept: Testing on the Road

The testing on the vehicle dynamometer provides important insights to the control system and its performance. However, as the process is highly nonlinear, random reference values and disturbance signals can have a drastic impact on the control performance. For this reason, experimental testings on the road are inevitable to demonstrate functionality. For the road testing, random drive profiles have been driven on an

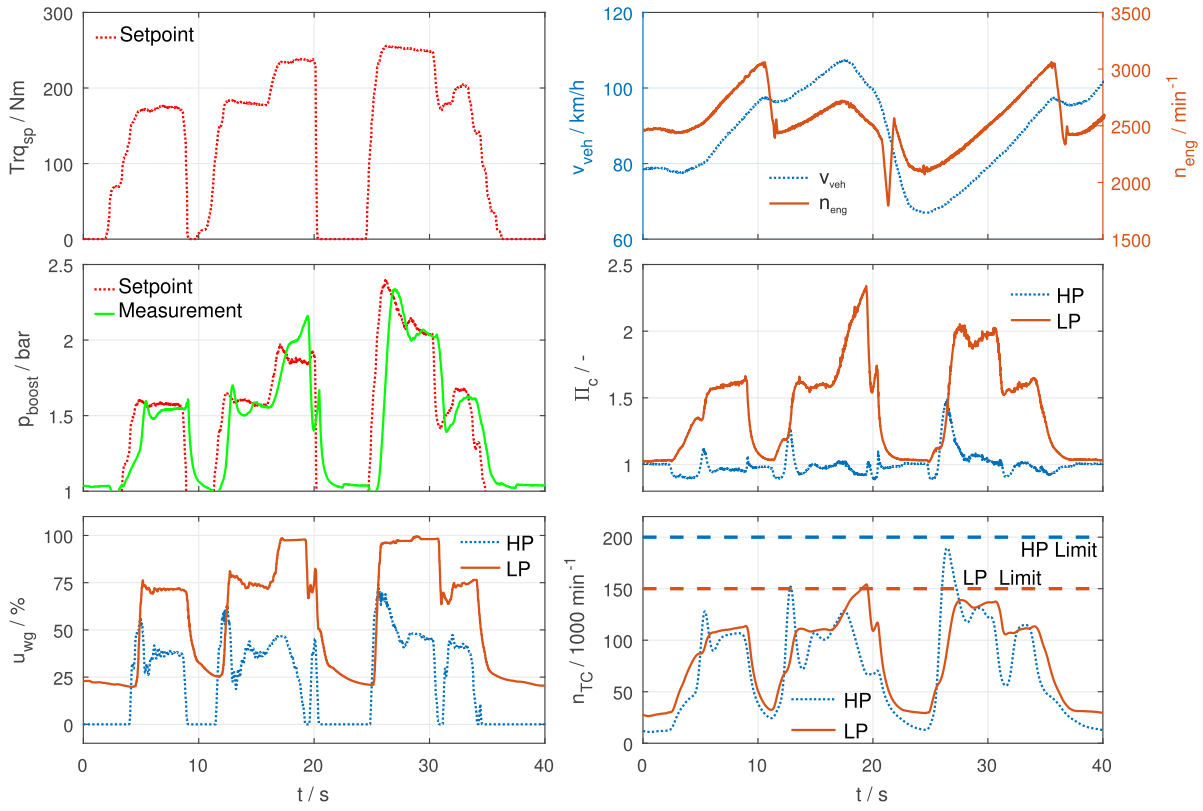

Fig. 7. Experiments on the road: NMPC control for a random driving maneuver with varying engine speed incl. gear shifts.

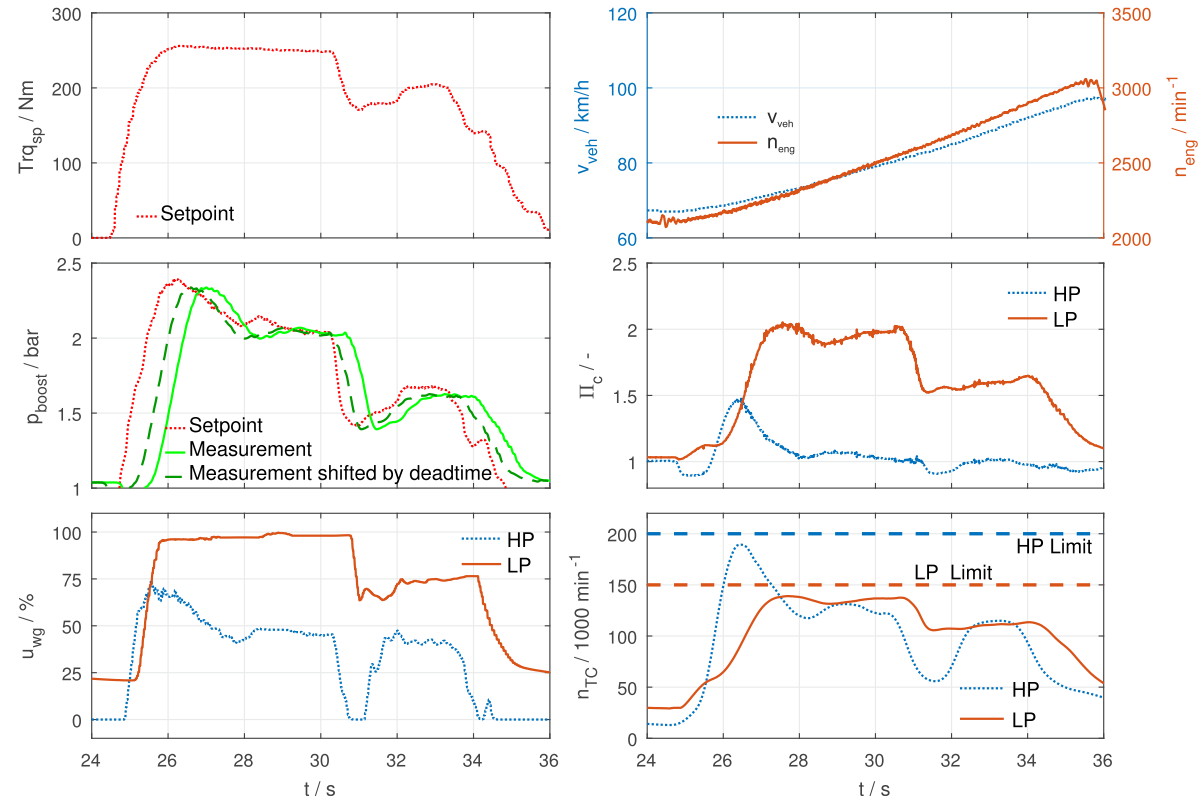

Fig. 8. Experiments on the road: detailed view on NMPC control for tip-in maneuver with varying engine speed.

automotive testing track. The experiments have shown that the developed NMPC control algorithm is able to have high performance in the entire operating range, for all drive profiles.

Fig. 7 shows an exemplary control result of the NMPC. In this case, a random driving maneuver is conducted, which includes varying engine speed with changes in gear shift.

In Fig. 7, the relevant traces for the boost pressure setpoint and measurement, wastegate actuation, turbocharger speeds, vehicle speed, engine speed, pressure ratios over compressor, and the torque setpoint are shown. The NMPC controller is able to follow the boost pressure setpoint profile accurately. Additionally, the HP stage is used to quickly increase boost pressure in transients. The controller makes use of the different stages taking into account the different values for the turbocharger speed limits. At $t = 19.5$ s, a minimal violation of the lp-speed constraint can be seen, which is due to the implementation as soft constraint and also due to the model-plant mismatch. Taking into account that the strategy is realized without using turbocharger speed sensors and only relying on the model (the plotted profile corresponds to measurement values, which were only used for validation purposes), it is a very good result. All in all, the engaging and disengaging of the different stages work even for the case of speed transient operation with random driving profiles.

For better evaluation of the control performance, Fig. 8 shows one tip-in maneuver in detail. The reference can be accurately tracked, despite the variations in the signals of boost pressure reference and engine speed. A certain lag is present between the reference and the system output, which cannot be overcome as it results from the deadtime of the system and the inertia of the turbochargers. For better evaluation, also the boost pressure output profile is plotted, but shifted by the deadtime of $t_D = 0.45$ s. Just as in the case of vehicle dynamometer testings, one can see that the control uses the HP stage to quickly increase the boost pressure. It should be stressed that the HP stage is only used in such a way that the limit values on the turbocharger speed are taken into account, which intuitively corresponds to an optimal strategy.

VIII. Conclusion

Novel airpath concepts are investigated for gasoline engines to reduce fuel consumption. One of them is the serial two-stage turbocharging. This airpath concept exhibits strong nonlinear behavior while having high demands on the control quality. The turbocharger speed limits should not be violated, and no overshoots in the boost pressure should be present. For achieving these goals, an NMPC approach is investigated in this paper. The basis of this NMPC scheme is a physically driven reduced-order state-space model formulated as a set of DAEs. The scheme uses direct optimal control, based on an online SQP type algorithm. Simulative testings have been conducted, which show the advantage over alternative approaches. Additionally, the control algorithm has been evaluated in a vehicle, where the airpath concept is implemented. The experiments have been conducted on a vehicle dynamometer and also on the road. These real-world experiments show that the control algorithm is able to fully exploit the multi-input characteristic of the two-stage turbocharging. The HP stage is used for quick pressure increases, whereas the LP stage is used for high mass flows.

Appendix A
Model Parameters

The constants $c_1, \ldots, c_8$ in the system of DAEs in (19)–(22) are defined as follows:

$$c_1 = \frac{2c_{p,\mathrm{exh}}\sqrt{T_{\mathrm{ut,lp}}}\eta_{s,t,\mathrm{lp}}A_{t,\mathrm{lp}}\sqrt{\frac{2\kappa_{\mathrm{exh}}}{\kappa_{\mathrm{exh}}-1}}}{\Theta_{\mathrm{tc,lp}}a_{\mathrm{lp}}\sqrt{R_{\mathrm{exh}}}} \tag{61}$$

$$c_2 = \frac{2\eta_v V_{\mathrm{cyl}}c_{p,\mathrm{air}}T_{\mathrm{uc,lp}}}{120R_{\mathrm{air}}T_{\mathrm{im}}\eta_{s,c,\mathrm{lp}}\Theta_{\mathrm{tc,lp}}a_{\mathrm{lp}}} \tag{62}$$

$$c_3 = \frac{120R_{\mathrm{air}}T_{\mathrm{im}}A_{t,\mathrm{lp}}\sqrt{\frac{2\kappa_{\mathrm{exh}}}{\kappa_{\mathrm{exh}}-1}}}{(1+\lambda_s)\sqrt{T_{\mathrm{ut,lp}}}V_{\mathrm{cyl}}\eta_v\sqrt{R_{\mathrm{exh}}}} \tag{63}$$

$$c_4 = \left(\frac{1}{A_{t,\mathrm{lp}}}\right) \tag{64}$$

$$c_5 = \frac{2c_{p,\mathrm{exh}}\sqrt{T_{\mathrm{ut,hp}}}\eta_{s,t,\mathrm{hp}}A_{t,\mathrm{hp}}\sqrt{\frac{2\kappa_{\mathrm{exh}}}{\kappa_{\mathrm{exh}}-1}}}{\Theta_{\mathrm{tc,hp}}(a_{\mathrm{hp}}n_{\mathrm{eng}}+c_{\mathrm{hp}})\sqrt{R_{\mathrm{exh}}}} \tag{65}$$

$$c_6 = \frac{2\eta_v V_{\mathrm{cyl}}c_{p,\mathrm{air}}T_{\mathrm{uc,hp}}}{120R_{\mathrm{air}}T_{\mathrm{im}}\eta_{s,c,\mathrm{hp}}\Theta_{\mathrm{tc,hp}}(a_{\mathrm{hp}}n_{\mathrm{eng}}+c_{\mathrm{hp}})} \tag{66}$$

$$c_7 = \frac{120R_{\mathrm{air}}T_{\mathrm{im}}A_{t,\mathrm{hp}}\sqrt{\frac{2\kappa_{\mathrm{exh}}}{\kappa_{\mathrm{exh}}-1}}}{(1+\lambda_s)\sqrt{T_{\mathrm{ut,hp}}}V_{\mathrm{cyl}}\eta_v\sqrt{R_{\mathrm{exh}}}} \tag{67}$$

$$c_8 = \left(\frac{1}{A_{t,\mathrm{hp}}}\right). \tag{68}$$

Appendix B
Abbreviations and Indices

The following abbreviations are used in this paper.

A	Cross-sectional area.
c_p	Specific isobaric heat capacity.
d	Disturbance state.
F	Force.
$\dot{m}$	Mass flow.
n	Speed.
P	Error covariance.
p	Pressure.
Q, R	Weighting matrices.
R_g	Gas constant.
s	Slack variable.
T	Temperature.
t	Time.
u	Control input.
V	Volume.
x	Differential state variable.
y	System output.
z	Algebraic state variable.
η	Efficiency.
κ	Isentropic exponent.
λ_s	Stochiometric air requirement.
Π	Pressure ratio.
Ψ	Power.
Θ	Polar mass moment of inertia.

The following indices are used in this paper.

amb	Ambient conditions.
asp	Aspirated.
c	Compressor.
ch	Control horizon.
comp	Computational.
cyl	Cylinder.
D	Deadtime.
dc	Downstream compressor.
dis	Displacement.
dt	Downstream turbine.
eng	Engine.
exh	Exhaust gas.
hp	High-pressure stage.
im	Intake manifold.
k	Discrete timestep.
$k+j\|k$	Prediction for timestep $k+j$, at time k.
kf	Kalman filter.
lp	Low-pressure stage.
ref	Reference.
s	Isentropic.
samp	Sampling.
t	Turbine.
tc	Turbocharger.
uc	Upstream compressor.
ut	Upstream turbine.
v	Volumetric.
wg	Wastegate.

APPENDIX C
NMPC CONTROLLER SETTING

The following settings have been used for the NMPC OCP formulation.

Q	100 000.
Q_N	100 000.
$R_{U,1}$	1.
$R_{U,2}$	1.
R_S	$1e9$.
$\underline{u}_{\text{wg,lp}}$	0%.
$\underline{u}_{\text{wg,hp}}$	0%.
$\bar{u}_{\text{wg,lp}}$	100%.
$\bar{u}_{\text{wg,hp}}$	100%.
$\underline{\dot{u}}_{\text{wg,lp}}$	-200 1/s.
$\underline{\dot{u}}_{\text{wg,hp}}$	-200 1/s.
$\bar{\dot{u}}_{\text{wg,lp}}$	200 1/s.
$\bar{\dot{u}}_{\text{wg,hp}}$	200 1/s.
$\bar{y}_2$	150 000 1/min.
$\bar{y}_3$	200 000 1/min.
$\underline{y}_2$	0 1/min.
$\underline{y}_3$	0 1/min.
Q_{KF}	$\text{diag}([1e-6; 1e-6; 500])$.
R_{KF}	1000.

REFERENCES

[1] L. Eriksson and L. Nielsen, *Modeling and Control of Engines and Drivelines*. Hoboken, NJ, USA: Wiley, 2014.

[2] L. Guzzella and C. Onder, *Introduction to Modelling and Control of Internal Combustion Engine Systems*. Springer, 2004.

[3] Z. Qiu, M. Santillo, M. Jankovic, and J. Sun, "Composite adaptive internal model control and its application to boost pressure control of a turbocharged gasoline engine," *IEEE Trans. Control Syst. Technol.*, vol. 23, no. 6, pp. 2306–2315, Nov. 2015.

[4] P. Kotman, M. Bitzer, and A. Kugi, "Flatness-based feedforward control of a two-stage turbocharged diesel air system with EGR," in *Proc. IEEE Int. Conf. Control Appl.*, Sep. 2010, pp. 979–984.

[5] T. Albin, D. Ritter, N. Liberda, S. Pischinger, and D. Abel, "Two-stage turbocharged gasoline engines: Experimental validation of model-based control," *IFAC-PapersOnLine*, vol. 48, no. 15, pp. 124–131, 2015.

[6] M. E. Emekli and B. A. Güvenç, "Explicit mimo model predictive boost pressure control of a two-stage turbocharged diesel engine," *IEEE Trans. Control Syst. Technol.*, vol. 25, no. 2, pp. 521–534, Mar. 2017.

[7] P. Dickinson, K. Glover, N. Collings, Y. Yamashita, Y. Yashiro, and T. Hoshi, "Real-time control of a two-stage serial VGT diesel engine using MPC," *IFAC-PapersOnLine*, vol. 48, no. 15, pp. 117–123, 2015.

[8] M. Huang, H. Nakada, K. Butts, and I. Kolmanovsky, "Nonlinear model predictive control of a diesel engine air path: A comparison of constraint handling and computational strategies," *IFAC-PapersOnLine*, vol. 48, no. 23, pp. 372–379, 2015.

[9] A. Murilo, M. Alamir, and D. Alberer, "A general nmpc framework for a diesel engine air path," *Int. J. Control*, vol. 87, no. 10, pp. 2194–2207, 2014.

[10] M. Herceg, T. Raff, R. Findeisen, and F. Allgowe, "Nonlinear model predictive control of a turbocharged diesel engine," in *Proc. IEEE Conf. Comput. Aided Control Syst. Design*, Oct. 2006, pp. 2766–2771.

[11] D. Alberer, M. Hirsch, and L. del Re, "A virtual references design approach for diesel engine control optimization," *Control Eng. Pract.*, vol. 18, no. 11, pp. 1263–1271, Nov. 2010.

[12] M. Huang, H. Nakada, S. Polavarapu, K. Butts, and I. Kolmanovsky, "Rate-based model predictive control of diesel engines," *IFAC Proc. Volumes*, vol. 46, no. 21, pp. 177–182, 2013.

[13] S. Glueck, "Charging concepts for a two-stage turbocharging gasoline engine," Ph.D. dissertation, RWTH Aachen Univ., Aachen, Germany, 2013.

[14] A. Thomasson, O. Leufvén, I. Criscuolo, and L. Eriksson, "Modeling and validation of a boost pressure actuation system, for a series sequentially turbocharged SI engine," *J. Control Eng. Pract.*, vol. 21, no. 12, pp. 1860–1870, Dec. 2013.

[15] T. Albin, D. Ritter, D. Abel, N. Liberda, R. Quirynen, and M. Diehl, "Nonlinear MPC for a two-stage turbocharged gasoline engine airpath," in *Proc. 54th IEEE Conf. Decision Control (CDC)*, Dec. 2015, pp. 849–856.

[16] F. Buchner, S. Wedowski, A. Sehr, S. Gluck, and C. Schernus, "In-vehicle optimization of 2-stage turbocharging for gasoline engines," *Int. J. Autom. Eng.*, vol. 2, no. 4, pp. 143–148, 2011.

[17] P. Moulin and J. Chauvin, "Modeling and control of the air system of a turbocharged gasoline engine," *J. Control Eng. Pract. Control*, vol. 19, no. 3, pp. 287–297, Mar. 2011.

[18] H. Khalil and J. Grizzle, *Nonlinear systems*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1996.

[19] M. Morari and U. Maeder, "Nonlinear offset-free model predictive control," *Automatica*, vol. 48, no. 9, pp. 2059–2067, 2012.

[20] M. S. Grewal and A. P. Andrews, *Kalman Filtering: Theory and Practice Using MATLAB*. Hoboken, NJ, USA: Wiley, 2008.

[21] R. Quirynen, M. Vukov, M. Zanon, and M. Diehl, "Autogenerating microsecond solvers for nonlinear MPC: A tutorial using ACADO integrators," *Optim. Control Appl. Methods*, vol. 36, no. 5, pp. 685–704, 2014.

[22] A. Griewank and A. Walther, *Evaluating Derivatives*. Philadelphia, PA, USA: SIAM, 2008.

[23] H. G. Bock and K. J. Plitt, "A multiple shooting algorithm for direct solution of optimal control problems," in *Proc. IFAC World Congr.*, 1984, pp. 242–247.

[24] D. Leineweber, *Efficient Reduced SQP Methods for the Optimization of Chemical Processes Described by Large Sparse DAE Models*, vol. 613. VDI Verlag, 1999.

[25] J. Nocedal and S. J. Wright, *Numerical Optimization* (Springer Series in Operations Research and Financial Engineering). New York, NY, USA: Springer, 2006.

[26] M. Diehl, H. J. Ferreau, and N. Haverbeke, *Nonlinear Model Predictive Control* (Lecture Notes in Control and Information Sciences). Springer, 2009, pp. 391–417.

[27] H. G. Bock, "Recent advances in parameter identification techniques for ODE," in *Numerical Treatment of Inverse Problems in Differential and Integral Equations*. Basel, Switzerland: Birkhäuser, 1983, pp. 95–121.

[28] M. Vukov, A. Domahidi, H. J. Ferreau, M. Morari, and M. Diehl, "Auto-generated algorithms for nonlinear model predictive control on long and on short horizons," in *Proc. 52nd IEEE Conf. Decision Control (CDC)*, Dec. 2013, pp. 5113–5118.
[29] H. J. Ferreau, C. Kirches, A. Potschka, H. G. Bock, and M. Diehl, "qpOASES: A parametric active-set algorithm for quadratic programming," *Math. Program. Comput.*, vol. 6, no. 4, pp. 327–363, 2014.
[30] M. Diehl, R. Findeisen, F. Allgöwer, H. G. Bock, and J. P. Schlöder, "Nominal stability of real-time iteration scheme for nonlinear model predictive control," *IEE Proc.-Control Theory Appl.*, vol. 152, no. 3, pp. 296–308, 2005.
[31] B. Houska, H. J. Ferreau, and M. Diehl, "An auto-generated real-time iteration algorithm for nonlinear MPC in the microsecond range," *Automatica*, vol. 47, no. 10, pp. 2279–2285, 2011.
[32] L. Grüne, "NMPC Without Terminal Constraints," in *Proc. IFAC Conf. Nonlinear Model Predictive Control*, 2012, pp. 1–13.

Thivaharan Albin received the Dipl.-Ing. degree in mechanical engineering from Technische Universität Braunschweig, Braunschweig, Germany, and the University of Waterloo, Waterloo, ON, Canada, in 2008, and the Ph.D. degree from the Institute of Automatic Control, RWTH Aachen University, Aachen, Germany, in 2013.

He is a Senior Engineer with the Institute of Automatic Control. He was a Visiting Researcher with ETH Zürich, Zürich, Switzerland, from 2016 to 2017, and a Visiting Professor with Peter the Great St. Petersburg Polytechnic University, Saint Petersburg, Russia, in 2017.

Dennis Ritter received the Dipl.-Ing. degree in mechanical engineering from RWTH Aachen University, Aachen, Germany, in 2014, where he is currently pursuing the Ph.D. degree with the Institute of Automatic Control.

His current research interests include the model-based predictive control of combustion engines.

Norman Liberda received the Dipl.-Ing. degree in mechanical engineering from RWTH Aachen University, Aachen, Germany, in 2011, and the Ph.D. degree from the Institute for Combustion Engines, RWTH Aachen University.

Since 2016, he has been a Gasoline Base Calibration Team Leader with FEV Europe GmbH, Aachen, Germany.

Rien Quirynen received the bachelor's degree in computer science and electrical engineering and the master's degree in mathematical engineering from KU Leuven, Leuven, Belgium, and the joint Ph.D. degree from KU Leuven and the University of Freiburg, Freiburg, Germany.

From 2014 to 2017, he was with KU Leuven and the University of Freiburg. He is currently with the Mitsubishi Electric Research Laboratories, Cambridge, MA, USA.

Dr. Quirynen received a four-year Ph.D. Scholarship from the Research Foundation–Flanders (FWO).

Moritz Diehl graduated with a Diploma in physics and mathematics at Heidelberg University, Heidelberg, Germany and Cambridge University, Cambridge, U.K. in 1999, and received his Ph.D. degree from Heidelberg University in 2001.

From 2006 to 2013, he was a Professor with the Department of Electrical Engineering, KU Leuven, Leuven, Belgium. In 2013, he moved to the University of Freiburg, Freiburg, Germany, where he is currently the Head of the Systems Control and Optimization Laboratory, Department of Microsystems Engineering, and the Department of Mathematics.