A General NMPC Framework for a Diesel Engine Air Path

André Murilo a *, Mazen Alamir b and Daniel Alberer c

a Control Systems Department, Gipsa-Lab, Saint Martin d’Hères, France;
b Control Systems Department, Gipsa-Lab/CNRS, Saint Martin d’Hères, France;
c Institute for Design and Control of Mechatronical Systems, Johannes Kepler University Linz, Austria

(Received 00 Month 200x; final version received 00 Month 200x)

This paper presents the formulation of a parameterized Nonlinear Model Predictive Control (NMPC) scheme to be applied on a Diesel engine air path. The most important feature of the proposed controller is that it uses no structural properties of the system model. Therefore, the proposed NMPC scheme can be applied to any nonlinear system, leading to a general framework for a Diesel engine air path. Moreover, the computational burden is substantially reduced due to an optimization problem of low dimension obtained by means of the parameterized approach. Simulation results and an experimental validation are presented in order to emphasize the controller’s efficiency and the real-time implementability.

Keywords: Parameterized NMPC, Diesel Engine Air Path, General NMPC Framework, Real-Time Implementation.

1 Introduction

Model Predictive Control (MPC) is an advanced control methodology which uses explicitly the system model to predict the future evolution of the process. The basic idea of MPC consists in computing an optimal control sequence over a prediction horizon at each decision instant, by minimizing some given cost function expressing the control objective. The first control signal is scheduled to be applied to the system during the next sampling period and this optimization process is successively repeated at each sampling time.

In the last recent years, the interest of the scientific community in MPC has considerably grown as well as the number of successful industrial case studies involving MPC controllers. Originally conceived to be applied to slow dynamic processes such as chemical industries or oil refineries, MPC-like strategies are becoming more and more present in different industrial sectors (Qin and Badgwell 2003). This was made possible thanks to the advances in computer science and to some extent, the dedicated formulations that aims at using MPC for systems with fast dynamics (Findeisen 2006). An important advantage of MPC is the fact that constraints on the inputs and the states can be explicitly handled in the problem formulation (Mayne et al. 2000). These constraints such as control saturation, safety requirement and contractual specifications are unavoidable in any realistic context. The presence of these constraints in control problems renders quite often MPC the only available option for the success of the control task.

Although linear MPC theory has reached a high level of maturation, there are several situations where nonlinearities are quite strong and real processes can not be correctly represented by linear models. As a result, the linear paradigm became insufficient to describe the whole complexity of the process, and nonlinear approaches become more appropriate. This motivates the use of a more elaborated strategy, namely Nonlinear Model Predictive Control (NMPC). Thus, one of the main...
difficulties of the NMPC strategy lies in the fact that the convexity of the optimization problem is no more guaranteed as in the linear case (Camacho and Bourdons 2007). Nonlinear Programming (NLP) approaches become a natural candidate to deal with such problems. However, NLP-based solutions may be very difficult to obtain, especially when fast dynamic systems are considered, since the optimal solution must be calculated within a short sampling period.

In order to deal with fast systems, some NMPC approaches emerged in this direction such as the Real-Time Iteration scheme (Diehl et al. 2005), Interior-point method (Wächter and Biegler 2006) and Continuation/Generalized method (Ohtsuka 2004). These techniques have the interesting feature of being general methods to solve NLP to address the real-time problem of fast systems. On the other hand, such strategies have to deal with sparse and high-dimension matrices. Moreover, the whole control sequence vector is taken as decision variable, and in some cases even the state vector, leading to a huge optimization problem. Despite the fact that the available tools may be used to more efficiently solve such problems, real-time implementation may be an issue in the cases where computational resources and available memory are limited. Contrarily to the standard approaches, the parameterized NMPC scheme presented by Alamir (2006) arises as an interesting alternative for fast systems control design. This method is based on a well-structured parametrization scheme leading to a low-dimensional optimization problem, which is a quite important feature to handle real-time requirements. This is particular true for a certain class of systems, such as those encountered in mechatronics and robotics for example.

In this paper, a constrained nonlinear system corresponding to the emissions control of a Diesel engine air path is analyzed.

The interest in Diesel engines does not stop growing due to some important reasons, such as high torque and power at low engine speeds and low fuel consumption (Heywood 1988). However, Diesel engines have an important drawback that must not be ignored, which is the problem of emissions of nitric oxides (NOx) and particulate matter (PM) (Johnson 2001). Therefore, the aim is to control the level of emissions in the Diesel engine air path which is a complex and coupled system. Furthermore, the presence of constraints in the system model make predictive controllers an interesting solution for this process. Some important contributions emerged that enable real-time implementation of linear MPC to Diesel engines. In Ortner and del Re (2007), the explicit MPC method (Bemporad et al. 2002) was successfully applied in a real testbed by means of the Hybrid Toolbox software (Bemporad 2004). Another relevant work proposed by Ferreau et al. (2007) utilizes the on-line active set strategy (Ferreau et al. 2006) also enabling real-time implementation. On the other hand, both approaches are based on multi-linear representations of the engine, forcing the controller to switch between different operating points, which can be prohibitive from a computational point of view. In fact, the trade-off between the model complexity, control design and real-time implementation represents a challenging problem for the Diesel engine air path and motivates the use of more sophisticated strategies, such as those based on NMPC schemes applied to fast dynamics systems.

In this paper, a parameterized NMPC scheme is proposed for the air-path control of a real-world Diesel engine. The main point of this control design is the fact that the optimal solution is completely independent of the nonlinear model structure being used, which leads to a general framework for the Diesel engine air path. Moreover, the constraints on the control inputs are structurally taken into account in the formulation and the optimization routine is reduced to a simple low dimensional optimization problem which enables very short computation times as it is shown in the remainder of this paper. As a result, the proposed NMPC scheme can deal at the same time with complex model representations as well as the real-time implementation issues.

This paper is organized as follows. First, important aspects about the Diesel engine process and the control problem are shown in section 2. Then, section 3 introduces some definitions and notations about the parameterized NMPC approach and shows how it can be used as general framework for Diesel engines. Section 4 presents some simulation results and an experimental validation of the proposed control scheme obtained on a real-world Diesel engine. This paper ends with a conclusion and gives some guidelines for further investigations.
2 Process Description and Control Problem

The previous section showed that the emissions of NOx and PM represent an important inconvenience of Diesel engines. Unfortunately, due to the lean operation a three way catalyst cannot be applied, yet an oxidation catalyst is used to diminish carbon monoxide and unburnt hydrocarbons in the exhaust in order to meet the regulatory requirements. Although aftertreatment devices for nitric NOx exist, their use is expensive and should be avoided. In other words, NOx emissions are a main concern during engine development. The other major concern, the PM emissions, are related to the diffusive combustion process in CI engines. For PM abatement even more matured aftertreatment devices than for NOx are available – Diesel particulate filters (DPF) – and also applied in practice. However, according to its loading condition the DPF requires regeneration from time to time, leading to an increase in fuel consumption. Obviously, the lower the particulate raw emissions the less frequent this fuel consuming regeneration has to be done. Summarizing, there is a huge interest in reducing the raw emission both for NOx and PM.

In order to decrease the level of emissions, it is important to understand the air path process of the engine to formulate the control problem to be solved (Wei 2006). Figure 1 shows the complete scheme of a Diesel engine air path. The compressor pumps the fresh air into the intake manifold of the engine in order to, in combination with the intercooler, increase the density of the charge. The fuel is directly injected into the combustion chamber and is burnt with the air coming from the compressor. A part of the exhaust gas of the exhaust manifold is recirculated into the intake manifold by means of the exhaust gas recirculation (EGR), which is extremely important to reduce NOx emissions (Jacobs et al. 2003). The remaining gas that is not recirculated passes through the variable geometry turbine (VGT), which absorbs the energy from the exhaust gas to propel the compressor (which pumps the fresh air) and closes the whole cycle.

Figure 1.: Schematic view of a Diesel engine air path.

As far as the emission problem is concerned, a common control objective is to track references of the fresh Mass Air Flow (MAF) and the intake Manifold Absolute Pressure (MAP) of the engine. A reason for using MAF and MAP as control outputs is that, for instance, the engine is equipped with sensors for such signals while there are usually no on-board sensors for the emissions in commercial vehicles, or the available sensors are too slow to be used for control. The reference values for them are optimized in steady-state with respect to emissions and are mapped over the entire speed and fuel profiles. Therefore, provided that such set-points are correctly chosen, a good tracking of these variables may lead to an acceptable level of emissions. The choice and generation of the set-points of MAF and MAP to cope with emissions is a challenging problem.
and more details can be found in Nieuwstadt et al. (2000). For the following it is assumed that these values are correctly chosen and hence, the quality of the control is evaluated with respect to this auxiliary problem.

In fact, the Diesel engine is a highly nonlinear and coupled multi-variable system due to the combined effects of VGT and EGR (Nieuwstadt et al. 1998). In the literature, the Diesel engine behaviour is normally represented by designing Mean Value Models (MVM) for simulation and control design purposes. The parameters of such representation are derived by the steady state analysis of several components of the engine. Some important works concerning modeling and identification of Diesel engines are presented in Christen et al. (2001), Jung (2003), Wei (2006).

For the validation of the control scheme proposed in this paper, a 6th-order model expressing the physics of the engine was used. This model shows significant nonlinearities involving 6 dynamical equations with 5 algebraic equations and 7 polynomial maps represented by look-up tables. The combined effect of this system of equations introduces several discontinuities, dead-zones and nonlinearities which represents a challenging problem for the control design. Figure 2 illustrates an example of a look-up table used for the system model. The exhaust to intake temperature \( T_{xi} \) was modeled as a second order polynomial depending on the exhaust temperature \( T_x \) and exhaust to intake mass flow rate \( W_{xi} \). The identification process was developed using a dynamical engine test bench available at the Johannes Kepler University Linz. Since the system modeling is not the main part of this work, the description of this model is shown in appendix A. The values are omitted due to a confidential term with the University of Linz.

![Exhaust to Intake Temperature Map](image)

Figure 2.: Example of a polynomial map (look-up table) used in the system model concerning the exhaust to intake temperature \( T_{xi} \)

In this paper, the aim is to define a formulation that does not rely on any particular mathematical structure of the model. Therefore, let us define a general nonlinear system represented by the following notation:

\[
x(k + 1) = f(x(k), u(k), w(k))
\]

where \( k \) is the present instant, \( x(k) \) is the state vector, \( f : \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \times \mathbb{R}^{n_w} \rightarrow \mathbb{R}^{n_x} \) and \( n_x, n_u, n_w \) the dimensions of the state, input and measurement disturbance vectors respectively. Here it is assumed that the whole state \( x(k) \) is known, either by direct measurements if the physical model is used (Appendix A) or by estimations based on some observer design method such as proposed in Murilo et al. (2009). The control inputs \( u \in \mathbb{R}^2 \) for the Diesel engine air path are the positions of the EGR and VGT valves, in percent. The engine speed \( N_e \) (in RPM), and the fuel rate \( w_f \) (in mg/stroke), are the external inputs, which can be regarded as measured disturbances \( w \in \mathbb{R}^2 \). The output \( y \in \mathbb{R}^2 \) of the system are not composed of the emissions as mentioned before, but MAF, \( y_1 \) (in kg/h), and MAP, \( y_2 \) (in Pa). Table 1 summarizes the main parameters of the Diesel engine air path model.

The next step consists in formalizing the control problem to be solved. The control task is therefore to adjust the EGR and VGT valve positions to track the required boost pressure and the fresh air mass flow set-points. More formally, the control problem is to design a controller
Table 1.: Main Parameters of the System Model

<table>
<thead>
<tr>
<th>Notation</th>
<th>Description</th>
<th>Variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>$W_{ci}$</td>
<td>Compressor to Intake Mass Air Flow: MAF control output</td>
<td></td>
</tr>
<tr>
<td>$p_i$</td>
<td>Intake Manifold Absolute Pressure: MAP state and control output</td>
<td></td>
</tr>
<tr>
<td>$m_i$</td>
<td>Intake Mass Air Flow state</td>
<td></td>
</tr>
<tr>
<td>$p_x$</td>
<td>Exhaust Manifold Pressure state</td>
<td></td>
</tr>
<tr>
<td>$m_x$</td>
<td>Exhaust Mass Air Flow state</td>
<td></td>
</tr>
<tr>
<td>$P_c$</td>
<td>Compressor Power state</td>
<td></td>
</tr>
<tr>
<td>$T_{xif}$</td>
<td>Temperature in the Downstream of the EGR Cooler state</td>
<td></td>
</tr>
<tr>
<td>$u_{egr}$</td>
<td>Exhaust Gas Recirculation Valve Position: EGR control input</td>
<td></td>
</tr>
<tr>
<td>$u_{vgt}$</td>
<td>Variable Geometry Turbine Valve Position: VGT control input</td>
<td></td>
</tr>
</tbody>
</table>

that forces $y$ to track a desired set-point $y^d$ of MAF and MAP, namely $y^d = [y_1^d, y_2^d]$, while the control inputs $u$ must satisfy the following set of constraints:

$$ u \in [u_{min}, u_{max}] ; \ u_{min} \in \mathbb{R}^2 ; \ u_{max} \in \mathbb{R}^2 $$

$$ \delta u \in [-\delta_{max}, +\delta_{max}] ; \ \delta_{max} \in \mathbb{R}^2 $$

where the inclusions are to be considered component wise and where $\delta u$ are the increments on the inputs and can be defined at each instant $k$ as $\delta u(k) = u(k) - u(k-1)$.

It is important to highlight two relevant points related to the Diesel engine plant as shown in Figure 2. The first one concerns the open-loop behaviour. In fact, for a given stationary value of the inputs EGR and VGT the outputs MAF and MAP asymptotically converge to some steady value that depend on the input values. As a result, the system is open-loop stable. This property is an important feature for the control design, especially for the parameterized NMPC scheme which is presented in the next section. The second point is the non-minimum phase behaviour that can be observed on the output evolution. In some cases, the outputs are forced to go to the opposite direction before going to the right one, which comes from the fact that the Diesel engine is a very coupled system. The implication of this property on the NMPC design is that the prediction horizon must be sufficiently long to overcome this opposite response transient. This requirement makes the real-time issue even harder since very short prediction horizons cannot be considered and long prediction horizons are more demanding in terms of on-line computation.

3 Control Design

Once the Diesel engine particularities and the control problem are introduced, the next step consists in formalizing the parameterized NMPC approach used as a general controller to be applied to the Diesel engine air path problem. This section starts with some definitions and notations in subsection 3.1. Then the parameterized NMPC approach is formulated in subsection 3.2 while subsection 3.3 gives precisely the definition of the optimization problem. Finally, subsection 3.4 summarizes the whole control strategy.

3.1 Definitions and Notations

This section introduces some definitions and notations related to the parameterized NMPC approach as presented in Alamir (2006). First, let us consider a time-invariant dynamic model given in the following general form:

$$ x(t) = X(t, x_0, u) $$

February 27, 2014 22:17 International Journal of Control NMPC˙diesel˙ijc˙v2
This clearly results in the sampled-time state feedback law defined by:

\[ u(t) = u^{(k)}(p, x) \quad ; \quad t \in [t_{k-1}, t_k] \quad ; \quad t_k = k\tau_s \]  

\[ U_{pwc}(p, x) := (u^{(1)}(p, x) \ldots u^{(N)}(p, x)) \in \mathbb{R}^N \]  

The state trajectory under the PWC control profile \( U_{pwc}(p, x_0) \) is denoted hereafter by \( X(\cdot, x_0, p) \). More generally, for each sampling instant \( t_j = j\tau_s \ (j \in \mathbb{N}) \), the notation \( X(t, x(t_j), p) \) denotes the state trajectory of the model at instant \( t_j + t \) under the PWC control profile defined by \( U_{pwc}(p, x(t_j)) \) over the time interval \([t_j, t_j + T] \). The NMPC strategy relies on the solution at each decision instant \( t_j \) of an optimization problem which can be defined as follows:

\[ \hat{p}(x(t_j)) := \text{arg min}_{p \in \mathbb{P}} J(p, x(t_j)) \]  

\[ \text{s.t.} \quad C(p, x(t_j)) \leq 0 \]  

where \( J(p, x(t_j)) \) is some cost function defined on the system trajectory starting from the initial condition \((t_j, x(t_j))\) under the PWC control profile defined by \( U_{pwc}(p, x(t_j)) \). The condition \( C(p, x(t_j)) \leq 0 \) gathers all the problem constraints defined on the same trajectory including possible final constraints on the state. Classical NMPC formulation states that once a solution \( \hat{p}(x(t_j)) \) is obtained, the first control in the corresponding optimal sequence is applied to the system during the sampling period \([t_j, t_{j+1}] \) such as:

\[ K(x(t_j)) := u^{(1)}(\hat{p}(x(t_j)), x(t_j)) \]  

This clearly results in the sampled-time state feedback law defined by:

\[ K := u^{(1)}(\cdot, \cdot) : \mathbb{R}^{n_x} \rightarrow \mathbb{U} \]
However, when a system with fast dynamics like the Diesel engine is considered, only a finite number of iterations of some optimization process $S$ can be performed during the sampling period $[t_{j-1}, t_j]$. This lead to the following extended dynamic closed-loop system:

$$x(t_{j+1}) = X(\tau_s, x(t_j), p(t_j)) \tag{10}$$

$$p(t_{j+1}) = S^{i_t}(p^+(t_j), x(t_j)) \tag{11}$$

where $S^{i_t}$ denotes successive finite number of iterations $i_t$ of $S$ starting from the initial guess $p^+(t_j)$ which is related to $p(t_j)$ to guarantee (if possible) the translatability property (Alamir 2006). This property makes it possible to construct a new control sequence (after the prediction horizon is shifted) that is built up with the remaining part of the past control sequence being used completed with a convenient final control. The stability of the extended system (10)-(11) heavily depends on the performance of the optimizer $S$, the number of iterations $i_t$ and the quality of the model. The details about the stability issue of the proposed scheme can be seen in Alamir (2008).

The main drawback of the parameterized approach lies in the fact that there is no universal parametrization that can be applied to any system. The choice of the kind of parametrization is rather problem dependent. In a sense, this makes such approach more a way of thinking rather than a systematic design strategy. In this paper, this strategy is applied to solve the Diesel engine problem that is a quite challenging from a computational point of view. This is the aim of the forthcoming sections.

### 3.2 Parametric NMPC Formulation for Diesel Engines

The previous topic showed that NMPC is based on the prediction of the state evolution under some control profile. The latter depends on a finite number of parameters to be determined by on-line optimization. While in standard NMPC formulations (Diehl et al. 2005, Flores and Milam 2006, Ohtsuka 2004, Wächter and Biegler 2006) all the components of $u$ are taken as degrees of freedom, here a parametrization of the control sequence is defined in order to get a decision variable of low dimension which is decoupled from the choice of the prediction horizon. Since the Diesel engine model is open-loop stable, it is interesting to use the static gain map directly in the parametrization of the control profile in order to guarantee an asymptotically vanishing tracking error. By doing so, the degrees of freedom of the parametrization are used to improve the transient behaviour of the tracking error represented by some exponential terms. As a result, the parametrization of the control sequence for a Diesel engine air path may involve the stationary control $u^*$ together with a simple exponential parametrization of the future control sequence as follows:

$$u(i\tau_s + t) = Sat_{u_{\min}}^{u_{\max}} \left( u^* + \alpha_1 e^{-\lambda_1 \tau_s} + \alpha_2 e^{-q \lambda_2 \tau_s} \right) \tag{12}$$

for $t \in [(k-1)\tau_s, k\tau_s]$ where $i \in \{0, \ldots, N-1\}$, $\tau_s$ is the sampling period, $\lambda > 0$, $q \in \mathbb{N}$ are tuning parameters, $\alpha_1, \alpha_2 \in \mathbb{R}^2$ are the coefficients to be determined (see below for more details) and $Sat$ is a saturation map $Sat: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ defined as follows:

$$Sat_{u_{\min}}^{u_{\max}}(u_i) = \begin{cases} 
  u_{\min}^i & \text{if } u_i \leq u_{\min}^i \\
  u_{\max}^i & \text{if } u_i \geq u_{\max}^i \\
  u_i & \text{otherwise}
\end{cases}, \quad i \in \{1, 2\}$$

Hereafter, the $Sat$ term will be omitted to simplify the forthcoming notations. Note that by using the above expressions, the saturation constraints on the control input are structurally
respected. The exponential parametrization used above is not the only possible one. However it shows the important property of translatability as mentioned previously, since the remaining part of an open loop control trajectory after horizon shifting can be easily expressed in terms of the last parametrization. Another advantage of the exponential parametrization is that one can link the decreasing rate represented by $\lambda$ of the exponential function to the bandwidth of the control. However, despite of these advantages, it is possible to consider that many other simple parametrization could have been successfully used.

As a matter of fact, at instant $k$, the continuity of the control sequence must be guaranteed. Therefore, using $i = 0$ in (12) leads to the following constraints:

$$u^* + \alpha_1 + \alpha_2 = u(k - 1)$$

where $u(k - 1)$ is the previous value of the control input already scheduled according to the computation performed over the sampling period $[k - 1, k]$. Let us consider one step ahead in (12) with $i = 1$:

$$u^* + \alpha_1 e^{-\lambda \tau_s} + \alpha_2 e^{-q \lambda \tau_s} = u(k)$$

The difference between (14) and (13) leads to the following expression:

$$u(k) - u(k - 1) = \alpha_1 (e^{-\lambda \tau_s} - 1) + \alpha_2 (e^{-q \lambda \tau_s} - 1) = \delta_{\text{max}}$$

It is worth noting that $u(k) - u(k - 1) = \delta_{\text{max}}$ represents the constraints on the variation rates of the control inputs according to (3). Moreover, such constraints can be respected provided that one meets the following expression:

$$\alpha_1 (e^{-\lambda \tau_s} - 1) + \alpha_2 (e^{-q \lambda \tau_s} - 1) = \left( \begin{array}{c} p_1 \\ p_2 \end{array} \right) \cdot \delta_{\text{max}}$$

where $p_1, p_2 \in [-1, +1]^2$ are the first two parameters (one for each control input) used in the control parametrization. Note that the above expression simply states that the difference between two successive controls does not exceed a fraction $p \in [-1, +1]^2$ of the maximal allowable values $\delta_{\text{max}}$, in order to respect the constraints represented by relation (3). Again, starting from the structure of the control profile in (12), one can consider that $u^*$ used in the expression of the control profile can be defined as a part of the decision variable. Therefore, the new control sequence can be rewritten as follows:

$$u(i \tau_s + t, p) = \left( \begin{array}{c} p_3 \\ p_4 \end{array} \right) + \alpha_1 e^{-\lambda i \tau_s} + \alpha_2 e^{-q \lambda i \tau_s}$$

Note that the extended set of parameters in (17) is defined thus as $p \in \mathbb{R}^4$. Then, the resulting parameter vector $p$ must respect the following set of constraints.

$$\left( \begin{array}{c} p_1 \\ p_2 \end{array} \right) \in [-1, +1]^2$$

$$\left( \begin{array}{c} p_3 \\ p_4 \end{array} \right) \in \mathbb{R}^2$$

The above expression shows that the parameters $p_3$ and $p_4$ can be freely chosen in $\mathbb{R}^2$ since the control sequence $u$ in (12) is already saturated between $u_{\text{min}}$ and $u_{\text{max}}$. Now, expressions (13),
used to enforce the constraint on the final state for stability purposes. Moreover, denoted by loop control profile defined by
\[ u(k-1) = p_3 + \alpha_1^{u_1} + \alpha_2^{u_2} \]
\[ u_2(k-1) = p_4 + \alpha_1^{u_2} + \alpha_2^{u_2} \]
\[ p_1\delta_{\text{max}}^{1} = \alpha_1^{u_1}(e^{-\lambda_{\tau_s}} - 1) + \alpha_2^{u_1}(e^{-q_{\lambda_{\tau_s}}} - 1) \]
\[ p_2\delta_{\text{max}}^{2} = \alpha_1^{u_2}(e^{-\lambda_{\tau_s}} - 1) + \alpha_2^{u_2}(e^{-q_{\lambda_{\tau_s}}} - 1) \]
where the notations \( \alpha_1 = (\alpha_1^{u_1}, \alpha_1^{u_2})^T \) and \( \alpha_2 = (\alpha_2^{u_1}, \alpha_2^{u_2})^T \) are used. Therefore, the computation of \( \alpha_1 \) and \( \alpha_2 \) can be done by solving the following system of equations:
\[
\begin{pmatrix}
\alpha_1^{u_1}(p) \\
\alpha_1^{u_2}(p) \\
\alpha_2^{u_1}(p) \\
\alpha_2^{u_2}(p)
\end{pmatrix}
= 
\begin{pmatrix}
1 & 1 & 0 & 0 \\
-e^{-\lambda_{\tau_s}} - 1 & -e^{-q_{\lambda_{\tau_s}}} - 1 & 0 & 0 \\
0 & 0 & 1 & 1 \\
0 & 0 & e^{-\lambda_{\tau_s}} - 1 & e^{-q_{\lambda_{\tau_s}} - 1}
\end{pmatrix}
^{-1}
\begin{pmatrix}
\alpha_1^{u_1}(p) - p_3 \\
p_1\delta_{\text{max}}^{1} \\
p_2\delta_{\text{max}}^{2} \\
p_1\delta_{\text{max}}^{2} - p_4
\end{pmatrix}
\]
which clearly shows that parameters \( \alpha_1 \) and \( \alpha_2 \) are now dependent on \( p \). Inserting this into (17) leads to:
\[
u(i\tau_s + t, p) = \begin{pmatrix} p_3 \\ p_4 \end{pmatrix} + \alpha_1(p).e^{-\lambda_{\tau_s}i} + \alpha_2(p).e^{-q_{\lambda_{\tau_s}}i}.
\]
Expression (22) shows how the control profile depends on the parameter vector \( p \in \mathbb{R}^4 \) introduced in (16) and (17). This vector is obtained by solving an optimization problem as shown in the next section. It is important to emphasize that more exponential terms can be inserted in the control structure (22). However the use of only two of these terms are sufficient to cover the dynamics of the engine, even for the fast transients as shown in the forthcoming sections. Moreover, such unnecessary increase in the number of degrees of freedom in the control scheme affects directly the computational burden since the number of parameters to obtain by solving the optimization problem is also augmented.

3.3 Definition of the Optimization Problem

In the parameterized NMPC approach, the optimization problem must return the set \( \hat{p} \) as defined in (7). For the definition of the cost function, since the problem to be solved is the tracking of MAF and MAP set-points, it is reasonable to consider a stage cost that penalizes the error \( y - y^d \) and a terminal cost on the state \( x \) (Mayne et al. 2000). As a result, the optimization procedure consists in evaluating the cost function under the open loop control profile, applied in the system model along the prediction horizon. The aim is to find the best set of parameters that minimizes \( J \), within the available computation time. Therefore, given the desired set-point \( y^d \), the best set of parameters to be used in the definition of optimal control sequence is provided by solving, during the sampling period \([k\tau_s, (k+1)\tau_s]\), the following optimization problem:
\[
\hat{p} := \arg\min_{p \in \mathcal{P}} \left[ p_x \cdot \left\| x_f(p) - f(x_f(p), u(N)(p(-)), w) \right\|^2 + \sum_{i=0}^{N-1} \left\| Y(i, x(k), p) - Y_f(i, y^d, w) \right\|^2_{Q_y} \right]
\]
where \( Y(i, x(k), p) \) is the output at instant \( k+i \) based on the model being used under the open loop control profile defined by \( p \) over \([k, k+N-1]\) and starting from the current value of the state denoted by \( x(k) \), \( x_f(p) = X(N, x(k), p) \) (according to (4)) and \( p_x > 0 \) is a weighting coefficient used to enforce the constraint on the final state for stability purposes. Moreover, \( Y_f(i, y^d, w) \) is
the filtered version of the set-point used to avoid overshoot in the system response, namely:

$$Y_f(i, y^d, w) = y^d + e^{-3\tau_r i / t_r} \cdot [y(k) - y^d]$$

where $t_r$ is the desired response time of the closed-loop system to reach 95% of $y^d$, and $Q_y \in \mathbb{R}^{2 \times 2}$ is a positive definite weighting used to differently weight the outputs $y_1$ and $y_2$, and has the following structure.

$$Q_y = \begin{pmatrix} \rho_1 & 0 \\ \tilde{y}_1 & \rho_2 \\ 0 & \tilde{y}_2 \end{pmatrix}$$

where $\rho_1$ and $\rho_2$ are the weighting terms of $y_1$ and $y_2$ respectively and $\tilde{y}_1$ and $\tilde{y}_2$ the normalization terms of each output. Note that the terminal cost in (23) is defined as the norm of the difference of two successive values of the state, at the end of prediction horizon, under the control profile depending on $p$. In fact, this is the classical way to address stability in predictive control formulations (Mayne et al. 2000). It is worthy to underline that the computation of $p$ does not rely on any specific model structure. This renders the present solution appropriate for use as a general NMPC framework for Diesel engines or for any other nonlinear system which has a similar behavior. Figure 3.3 illustrates the parameterized NMPC scheme.

Figure 4.: Schematic view of the predictive control scheme. The cost function is evaluated according to the system requirements over the open loop profile of the control input applied in the system model along the prediction horizon. The solution of the optimization problem $\hat{p}$ is used to generate the optimal parameterized control scheme based on the exponential structure (22).

As far as the optimizer is concerned, classical solvers like Powell’s method and simplex were tested, and the obtained results were quite positive and very similar. However, a particular Sequential Quadratic Programming (SQP) routine was developed for simulations and also used for experimental validation. Since this is not the main contribution of this paper, for the sake of completeness, a brief presentation of the algorithm is provided and a detailed description can be found in Murilo (2009). Basically, the algorithm performs scalar SQP’s on each component to obtain a quadratic approximation (which corresponds to three iterations). Then, a candidate cost function is evaluated and compared with the current value of the cost function $J$, leading
to an adaptation of a trust region depending on the success or not of this operation. After each complete cycle of scalar SQP’s over the 4 parameters the results are used for the second part, the gradient procedure, to construct an approximation of the gradient along which potentially successful steps are attempted. Again, a trust region is updated along the gradient. The whole algorithm stops when the upper bound number of function evaluations $n_{fe}$ is reached by the number of iterations to satisfy the available computation time and the best value of $\hat{p}$ that minimizes the cost function is returned. As a result, the proposed routine yields an optimization algorithm that uses the model as a black-box simulator and is therefore easily re-usable if a more sophisticated and faithful process model is made available.

3.4 General NMPC Framework

The previous sections showed that the control law $u$ presented in (22) depends on the computation of the set of parameters $p$ by solving the optimization problem (23) and the solution of (21) to obtain the coefficients $\alpha_1(p)$ and $\alpha_2(p)$. To summarize, a general parameterized NMPC framework for a Diesel engine air path can be formalized as follows:

3.4.1 Set of Parameters $p$

Compute the four components $p_1, p_2, p_3$ and $p_4$ by solving an optimization problem that satisfies the tracking requirements and the stability condition.

$$\hat{p} := \arg\min_{p \in P} \left[ \rho_x \cdot \left\| x_f(p) - f(x_f(p), u^{(N)}(p(\cdot)), w) \right\|^2 + \sum_{i=0}^{N-1} \left\| Y(i, x(k), p) - Y_f(i, y^d, w) \right\|^2_{Q_x} \right]$$

3.4.2 Coefficients $\alpha_1(p)$ and $\alpha_2(p)$

Solve a simple linear system to obtain the parameters $\alpha_1(p)$ and $\alpha_2(p)$

$$\begin{pmatrix} \alpha_1^{u_1}(p) \\ \alpha_2^{u_1}(p) \\ \alpha_1^{u_2}(p) \\ \alpha_2^{u_2}(p) \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ e^{-\lambda_1 \tau_s} - 1 & e^{-q_1 \lambda_1 \tau_s} - 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & e^{-\lambda_2 \tau_s} - 1 & e^{-q_2 \lambda_2 \tau_s} - 1 \end{pmatrix}^{-1} \begin{pmatrix} u_1(k - 1) - p_3 \\ p_1 \delta_1^{max} \\ u_2(k - 1) - p_4 \\ p_2 \delta_2^{max} \end{pmatrix}$$

3.4.3 Control Law

Build the optimal PWC control profile based on the exponential structure where only the first part of the resulting control sequence is applied.

$$u(i \tau_s + t, p) = \begin{pmatrix} p_3 \\ p_4 \end{pmatrix} + \alpha_1(p) e^{-\lambda_1 i \tau_s} + \alpha_2(p) e^{-q_1 \lambda_1 i \tau_s}$$

4 Simulation Results and Experimental Validation

In this section, some simulation results and an experimental validation are presented in order to evaluate the performance of the proposed NMPC scheme and its real-time implementability. The physical model, as presented before and shown in Appendix A, was used for simulations and sampled at 10 ms. The programs were developed in C language and integrated into a SIMULINK environment via S-functions and MEX-files.

The simulation results are divided into two main sets according to the kind of set-point profiles: the sequence of step changes and a part of the New European Driving Cycle (NEDC) which is the European homologation cycle for passenger cars. The main parameters of the NMPC controller presented in the previous section are shown in Table 2.
Table 2.: Parameters used in simulations

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>100</td>
<td>$\rho_2$</td>
<td>$1 \times 10^5$</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>8</td>
<td>$q$</td>
<td>1.25</td>
</tr>
<tr>
<td>$\tau_x$</td>
<td>0.01s</td>
<td>$t_r$</td>
<td>1s</td>
</tr>
<tr>
<td>$\rho_1$</td>
<td>1</td>
<td>$\rho_2$</td>
<td>1</td>
</tr>
<tr>
<td>$\bar{y}_1$</td>
<td>1</td>
<td>$\bar{y}_2$</td>
<td>1</td>
</tr>
</tbody>
</table>

4.1 Successive Step Sequence

In this part, a successive step sequence for MAP and MAF was used to generate the setpoint profile. Figure 5 illustrates the closed-loop output behavior under the sequence of step changes in the set-points. The set-points for MAF and MAP are represented by the solid lines, EGR is saturated between 0% and 100% and VGT between 50% and 100% since the latter never works completely opened.

![Figure 5.](image)

(a) Evolution of the closed-loop output under a sequence of step changes. (b) Evolution of the 4 parameters computed by the optimization problem. Note that the transient periods are driven by $p_1$ and $p_2$ since they affect the rate of change of the control input while $p_3$ and $p_4$ are related to the steady solution.

The tracking performance is quite nice except during the interval time $t \in [165; 180]$ (see MAF) which means that the valves are saturated since the set-points are unfeasible. This phenomenon is common in Diesel engines and has already been observed in Murilo et al. (2009) which showed that, for a certain range of values of MAF and MAP, the control input can not achieve both set-points at the same time. In this case, the controller does its best to reach the closest achievable output. Figure 5 also shows the evolution of the 4 parameters that are solution of the optimization problem (23) given the available computation time of 10 ms. Note that by the definition of the parametrization, when $p_1$ and $p_2$ are saturated, it means that the constraint on the rate of change of the control input is active. A similar behaviour is observed for parameters $p_3$ for the steady solution of EGR as well. It is worth noting that the solver needs to perform only 4 function evaluations $n_{fe}$ to meet the real-time requirements, which corresponds to only 1 SQP iteration to satisfy (11), thanks to the parameterized NMPC structure and the well-posed optimization problem as presented in (23).

4.2 New European Driving Cycle - NEDC

The next reference profile to be tracked is a part of the NEDC which can be divided into two relevant periods: the urban and extra urban parts, as shown in Figure 6. The first one is
represented by the interval 0s to 250s and the second one holds on the time interval ranging from 250s to 500s. Figure 6 shows the first simulation scenario for the NEDC tracking. The weighting term $\rho_2$ has been set to 1. Note the good tracking performance in the urban and extra-urban parts and how the constraints on the inputs are correctly handled. Note also the unfeasible set-points, especially for MAP around 0s, 100s and 480s.

Figure 6.: Simulation results showing the evolution of MAF and MAP over a part of the NEDC, with $\rho_2 = 1$. Note the quality of the tracking, especially for the urban part. Note also the points where the controller is saturated and is not able to handle the output requirements at the same time since the set-points are not feasible.

Figure 7 illustrates the simulation results for two values of $\rho_2$ in (24). The first one, $\rho_2 = 100$ and the second one $\rho_2 = 0.01$. While in the first case, the tracking performance of MAP is nice, for the second one, MAF is privileged. This clearly shows the trade-off between the tracking quality of the set-points and the weighting terms that are imposed on the outputs in the definition of the cost function. Then, such penalizing terms can be modified on-line in order to achieve specific performances in different operating points of the engine according to emissions requirements. Otherwise, one can just find a good trade-off as shown in the previous scenario depicted in Figure 6.

Figure 7.: Evolution of MAF and MAP over a part of the NEDC, with $\rho_2 = 100$ (a) and $\rho_2 = 0.01$ (b).
Figure 8 shows the computation time spent to process the fixed number of function evaluations \( n_{fe} \) of the solver \( S \), which in this case, has been set to 4 as mentioned previously. Indeed, the parameterized NMPC scheme presented in this paper can perform a sufficient number of function evaluations to obtain a satisfactory solution within a very small sampling period, enabling real-time implementation in a real-world Diesel engine as will be presented in the next section.

![Computation Time Graph](image)

Figure 8.: Evolution of the computation time of the parameterized NMPC scheme. Note that the maximum allowed value 10 ms is never reached.

### 4.3 Real-Time Implementation

In order to show the efficiency of real-time implementation concerning the parameterized approach, some experiments were performed at the engine testbed of the Johannes Kepler University Linz. The platform consists of a passenger car Diesel engine fulfilling the EU4 emission standard, connected to a highly dynamical dynamometer, which simulates the load on the engine shaft. A dSPACE Autobox was used as a real-time system, running at 480 MHz and linked to MATLAB software, with a sampling period of 50 ms, keeping the same configuration for other predictive controllers that also used the test bench. This hardware architecture allows the parameterized NMPC to easily perform up to 30 function evaluations of \( S \). The programs were also developed in C language to assess the real-time performances. Again, the experiment consists by imposing a step sequence trajectory for MAF and MAP covering different operating points of the engine, as shown in Figure 9. The program was successfully embedded in the dSPACE device which performed on-line optimizations in a very efficient way due to the well-posed optimization problem. As well as for the simulation results, the parameterized NMPC approach presented a good tracking performance (maximum overshoot of 5% for MAP) and the constraints on the inputs were correctly handled.

### 4.4 Comparison with the Existing Approaches

Concerning the structure of the system model, despite the popularity of linear MPC strategies developed for the Diesel engine air path control problem (Ortner and del Re 2007, Ferreau et al. 2007), the use of multi-linear models according to operating points may be insufficient to correctly cover the whole engine dynamics. Actually, in Ferreau et al. (2007) authors mentioned that more elaborated models are needed in order to obtain better results. Moreover, for other dedicated nonlinear-like strategies (Jankovic and Kolmanovsky 1998, Plianos and Stobart 2007), authors need to use reduced-order models otherwise it would be a very hard task for the control design. In this context, the proposed NMPC scheme overcomes this drawback allowing the use of any model structure whatever its complexity.

The use of several explicit MPC controllers in parallel such as presented in Ortner and del Re (2007) increases considerably the computational cost of such approaches. Authors emphasized the difficulty to work with longer control horizons given the available computation time which was limited to 50 ms. The same issue was pointed out by Ferreau et al. (2007). In Herceg et al.
Figure 9: Experimental results showing the response of the parameterized NMPC with $N = 30$ and $\rho_x = 1000$ over several operating points of the engine. The engine platform was configured to track the outputs of MAF and MAP around its means values (119.2 and 1183.60 respectively) and therefore the normalization terms of (24) were set to 1 for both outputs.

(2006), the proposed NMPC strategy does not consider any particular algorithm to solve the resulting optimization problem in a faster way and hence is likely to be unsuitable for real-time implementation. The control scheme presented in this paper reduces substantively the computational load thanks to a well-posed optimization problem enabling a sampling period of 10 ms, computed over a complex set of nonlinear equations. Furthermore, the use of smaller sampling periods together with a more sophisticated model is quite important to deal with possible model-plant mismatches since the control inputs can be updated more frequently based on a more realistic prediction of the system model.

In Ferreau et al. (2007), the experiments were performed over a narrow operating range of the engine and important overshoots could be observed in some regions (mainly for MAF). Normalized test cycles like NEDC were also not verified. On the other hand, in the parameterized NMPC approach, the overall tracking performance presented nice results for both outputs, covering several operating points of the engine without switching between models and saving computation time. Simulation tests of the NEDC showed important results for the tracking of MAF and MAP and a weighting term can be used to find an acceptable tuning for the outputs in order to satisfy emissions requirement.

5 Conclusion and Future Works

In this paper, a general NMPC framework was proposed to address the Diesel engine air path control problem. The simulations were performed with a full nonlinear model obtaining relevant results. The controller structurally satisfied the constraints on the inputs, derivatives and available time to perform the optimization problem. It was also shown that the presented solution can be used as a sort of black-box NMPC for Diesel engines, no matter the complexity and the structure of the model being used. Finally, an experimental validation of the presented controller was also proposed to emphasize the efficiency of the parameterized control strategy to deal with the constraints on the computation time for embedded applications.

Concerning other predictive controllers, the proposed control scheme overcomes the main drawbacks related to the existing MPC strategies: exponential growth of the number of critical regions affecting directly the computation burden, several local predictive controllers running in parallel, linear, simplified or reduced models used to reproduce the Diesel engine dynamics and switching between several operating points. Then, the results presented in this paper are of great importance for the development of further implementable NMPC strategies for Diesel
engines. Future works consist in developing a generic real-time software based on the proposed NMPC approach to be applied in a wider range of nonlinear applications.

Appendix A: Physical Model

A nonlinear representation was used to describe the behavior of the considered Euro 4 passenger car Diesel engine. This nonlinear model was parameterized with measurements from an engine test bench (Alberer 2009). To summarize, the whole behavior of the Diesel engine air path can be described by a 6th-order system with a set of algebraic equations and look-up tables. The parameters of the model are listed in Table A1.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Parameter</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>Temperature</td>
<td>A</td>
<td>Effective EGR valve area</td>
</tr>
<tr>
<td>V</td>
<td>Volume</td>
<td>Pc</td>
<td>Compressor power</td>
</tr>
<tr>
<td>m</td>
<td>Mass</td>
<td>( \eta_c )</td>
<td>Volumetric efficiency</td>
</tr>
<tr>
<td>W</td>
<td>Mass Flow Rate</td>
<td>( w_f )</td>
<td>Fuel injection quantity</td>
</tr>
<tr>
<td>p</td>
<td>Pressure</td>
<td>( N_e )</td>
<td>Engine speed</td>
</tr>
<tr>
<td>( p_r )</td>
<td>Pressure ratio</td>
<td>R</td>
<td>Ideal gas constant</td>
</tr>
<tr>
<td>( \tau_{egr} )</td>
<td>Time constant of EGR</td>
<td>( \tau_{vgt} )</td>
<td>Time constant of VGT</td>
</tr>
<tr>
<td>( \kappa )</td>
<td>Specific heat ratio</td>
<td>( P_t )</td>
<td>Turbine power</td>
</tr>
<tr>
<td>( \eta_m )</td>
<td>Compressor efficiency</td>
<td>( \epsilon_p )</td>
<td>Specific heat at constant pressure</td>
</tr>
<tr>
<td>( p_a )</td>
<td>Ambient pressure</td>
<td>( V_d )</td>
<td>Total engine displacement volume</td>
</tr>
<tr>
<td>( u_{egr} )</td>
<td>EGR valve position</td>
<td>( u_{vgt} )</td>
<td>VGT valve position</td>
</tr>
</tbody>
</table>

A.1 Dynamic Equations

\[
\dot{p}_i = \frac{R \kappa}{V_i}(W_{ci}T_c - W_{ie}T_i + W_{xi}T_{xi})
\]

\[
\dot{p}_x = \frac{R \kappa}{V_x}((W_{ie} + W_f)T_e - (W_{xi} + W_{xt})T_x)
\]

\[
\dot{P}_c = \frac{1}{\tau_{vgt}}(-P_c + \eta_m P_t)
\]

\[
\dot{m}_i = W_{ci} - W_{ie} + W_{xi}
\]

\[
\dot{m}_x = W_{ie} + W_f - W_{xi} + W_{xt}
\]

\[
\dot{T}_{xif} = \frac{1}{\tau_{egr}}(-T_{xif} + T_{xi})
\]

A.2 Algebraic Equations

\[
W_{xi} = \begin{cases} 
-\frac{A(u_{egr})p_x}{\sqrt{R_T}} \sqrt{2p_r \left(1 - \frac{p_r}{p_x}\right)} & \text{if } p_x \geq p_i \\
-\frac{A(u_{egr})p_i}{\sqrt{R_T}} \sqrt{\frac{2p_r}{p_x} \left(1 - \frac{1}{p_r}\right)} & \text{if } p_i < p_x 
\end{cases}
\]

\[
T_i = \frac{V_i}{Rm_i} p_i
\]

\[
T_x = \frac{V_x}{Rm_x} p_x
\]

\[
W_{ie} = \frac{m_i N_e V_d}{V_i} \frac{2\pi}{2}
\]

\[
P_t = W_{xt} \epsilon_p T_x \left(1 - \left(\frac{p_a}{p_x}\right)^\kappa\right)
\]
A.3 Polynomial Maps / Look-up Tables

\[ T_{xi} = P_1(T_x, W_{xi}) \]
\[ A(u_{egr}) = P_2(u_{egr}) \]
\[ \eta_v = P_3(p_i, N_e) \]
\[ T_{ex} = P_4(w_f, W_{ie}) \]
\[ W_{xt} = P_5(u_{vgt}, p_x) \]
\[ W_{ci} = P_6(P_c, p_i) \]
\[ T_{ci} = P_7(W_{ci}, p_i) \]

where \( P_i(v_1, v_2, \cdots, v_{N_v}) \) represents a second order polynomial identified on the Diesel engine test bench depending on some set of variables \((v_1, v_2, \cdots, v_{N_v})\). The index \( i \) corresponds to the intake manifold, \( x \) the exhaust manifold, \( ci \) compressor to intake, \( ie \) intake to engine, \( ex \) engine to exhaust, \( xi \) exhaust to intake, \( xt \) exhaust to turbine and \( xif \) indicates the first order low pass filtered value of the temperature in the downstream of the EGR cooler.

Acknowledgment

The sponsoring of this work by the COMET K2 Center ”Austrian Center of Competence in Mechatronics (ACCM)” is gratefully acknowledged. The COMET Program is funded by the Austrian Federal Government, the Federal State Upper Austria and the Scientific Partners of ACCM.

References


of a real-world Diesel engine using an extended online active set strategy,” *Annual Reviews in Control*, 31, 293–301.


