ABSTRACT

This paper describes a diesel engine lean NOx trap (LNT) regeneration air to fuel ratio (AFR) control system using nonlinear model predictive control (NMPC) technique for simultaneous regeneration fuel penalty and overall tailpipe-out NOx reductions. A physics-based and experimentally validated nonlinear LNT dynamic model was employed to construct the MMPC control algorithm, which dictates the AFR value during regenerations. Different choices of NMPC cost function were examined in terms of the impact on fuel penalty and total tailpipe NOx slip amount. The cost function to achieve the best tradeoff between fuel penalty and tailpipe-out NOx was selected based on physical insights into the LNT system and oxygen storage dynamics. The NMPC regeneration AFR control system was evaluated on a vehicle simulator cX-Emissions with a 1.9L diesel engine model through the FTP75 driving cycle. Compared with the conventional LNT regeneration AFR control strategies such as a PID controller, 23.7% of regeneration fuel penalty reduction and 16% of tailpipe NOx mass reduction were observed.

INTRODUCTION

Diesel engines possess noticeable advantages in terms of efficiency, reliability, and power compared with gasoline counterparts. However, on the other hand, diesel engine emission control, especially NOx reduction, is much more challenging than that for gasoline engines. The lean burn characteristic of diesel engine is notable of higher NOx emissions compared with gasoline stoichiometric engines. It is often difficult to achieve the increasingly stringent NOx emission regulations by engine control and combustion improvement only. Catalytic aftertreatment systems have been introduced to significantly reduce the tailpipe emissions. Among a variety of selections, the most promising and commonly used NOx aftertreatment systems are lean NOx trap (LNT) for light-duty vehicles and selective catalytic reduction (SCR) for medium- and heavy-duty vehicles. In order to maintain the NOx conversion efficiency for LNTs, rich exhaust gas so far is necessary in order to purge the stored NOx [1, 2].

The main LNT operation process can be described by two steps. The first step, named NOx absorption, in which LNT traps the emitted NOx from the engine and stores them in solid state. Because the LNT storage capacity is finite, it has to be purged when the storage is full. The second step process is called NOx desorption, in which the stored NOx are released to gaseous state, and at the same time are converts to N2 by the available reductants (CO and HC). In this step, extra amount of fuel is dosed at the upstream of LNT in order to generate a rich environment that can conduce to regeneration. After the regeneration is finished, LNT regains the NOx adsorption capacity and starts trapping NOx. The LNT was demonstrated of being able to convert more than 90% of NOx emissions from engine. However, a common concern about LNTs is the associated fuel penalty during regenerations that can lead to higher overall vehicle fuel consumption. Tradeoff between LNT regeneration fuel penalty and tailpipe NOx slip has been a major challenge for LNT control.

Different LNT regeneration control strategies aiming at fuel penalty and emission reductions have been studied by researchers in both industry and academia [2-7]. For instance, in [2], the author developed an in-cylinder no-post-injection rich-lean combustion switching control for LNT regeneration purpose for fuel penalty reduction. Nakagawa et al. proposed a model-based approach to control the rich pulse timing and its duration for LNT regenerations [3]. A LNT model adaptation mechanism was used to adjust the rich pulse timing and duration to reduce the tailpipe emissions. In [5], the authors studied an adaptive LNT purge control strategy, in which on-line adaptation for the LNT capacity and LNT-in NOx flow rate models was conducted with the assistance provided by an exhaust gas oxygen sensor. The estimated LNT capacity usage percentage was then used to trigger the LNT regeneration events. Due to the highly nonlinear LNT regeneration efficiency, the AFR control during regeneration is also of importance for fuel penalty and emission reductions. However, little attention has been devoted to this aspect. In this paper, we specifically...
study the near-optimal AFR control during LNT regeneration using nonlinear model predictive control approach. Based on a simulation case study using an experimentally validated LNT model, it was observed that the NMPC AFR controller can reduce the regeneration fuel penalty up to 23.7% and the tailpipe NOx by 16% compared with a conventional LNT regeneration AFR control strategy such as PID control.

The rest of this paper is organized as the follows. In section II, a physics-based LNT model and its experimental validation are briefly described. The NMPC regeneration AFR control approach is presented in section III followed by simulation results and analyses. Conclusive remarks and future research directions are summarized in section V.

LNT MODEL AND VALIDATION

In order to exercise NMPC, a mathematical model that fairly accurately describes the dynamics of the physical process while still maintaining affordable computational burden is essential. Figure 1 shows the block diagram of the LNT model used in this paper [12]. The inputs are the exhaust gas mass flow rate and temperature, together with the composition of the feed gas, for which eight species are considered (\( \text{CO}, \text{H}_2\text{O}, \text{O}_2, \text{N}_2, \text{NO}, \text{CO}, \text{C}_n\text{H}_m, \text{SO}_2 \)). The outputs are the outlet gas temperature (assumed equal to the catalyst brick temperature) as well as the outlet mixture composition.

The input/output characterization of the LNT model results from the interaction of three subsystems, namely an oxygen storage dynamics model, a \( \text{NO}_x \) storage dynamics model, and a model for the catalyst temperature dynamics. The equations of the above subsystems, originally derived in [11], will be briefly summarized in this section. Interested readers can refer [11, 12] for details regarding the models.

OXYGEN STORAGE/RELEASE DYNAMICS

The model characterizing the oxygen storage and release dynamics is based on application of the continuity equation to the oxygen stored (in solid state) and to the oxygen present in the exhaust stream (gas phase), as in the following:

\[
\frac{dM_{O_2}}{dt} = C_{O_2} \frac{dx_1}{dt} = r_{O_2,\text{stor}} - r_{O_2,\text{rel}},
\]

\[
m_{O_2,\text{out}} = m_{O_2,\text{in}} - r_{O_2,\text{stor}} + r_{O_2,\text{rel}},
\]

where \( C_{O_2} \) is the oxygen storage capacity and \( x_1 \) represents the catalyst oxygen fill ratio.

The oxygen storage and release rate, \( r_{O_2,\text{stor}} \), and \( r_{O_2,\text{rel}} \), can be described as:

\[
r_{O_2,\text{stor}} = k_{st} \left( \frac{1-e^{\alpha x_1}}{e^{\alpha x_1} - 1} + 1 \right) m_{O_2,\text{in}},
\]

\[
r_{O_2,\text{rel}} = k_{rel} \left( \frac{1-e^{\beta x_1}}{e^{\beta x_1} - 1} \right) (m_{\text{CO}} + m_{\text{HC}})_{\text{in}},
\]

where \( k_{st} \) and \( k_{rel} \) are the two empirical constants, and the multipliers \( \alpha \) and \( \beta \) are linear functions of the catalyst temperature \( T \).

It is important to observe that the model assumes the \( \text{O}_2 \) storage and release phenomena prevailing over the \( \text{NO}_x \) ones. In particular, if storage sites are available in the catalyst and the engine exhaust gas is lean, the oxygen will be stored before the \( \text{NO}_x \). Conversely, if oxygen is stored in the LNT and the feed gas is rich, the reductants present will release and combine with the oxygen first and then \( \text{NO}_x \).

\( \text{NO}_x \) STORAGE/RELEASE DYNAMICS

The dynamics associated with the \( \text{NO}_x \) adsorption and release are described using a similar approach. Due to the fact that various parameters depend on inputs and states such as temperature, inlet mass flow rate and catalyst fill ratio, the model equations developed are highly nonlinear:

\[
\frac{dM_{\text{NO}_x}}{dt} = C_{\text{NO}_x} \frac{dx_2}{dt} = r_{\text{NO}_x,\text{stor}} - r_{\text{NO}_x,\text{rel}},
\]

\[
m_{\text{NO}_x,\text{out}} = m_{\text{NO}_x,\text{in}} - r_{\text{NO}_x,\text{stor}} + r_{\text{NO}_x,\text{rel}}.
\]

Here \( x_2 \) is the LNT NOx fill ratio and \( C_{\text{NO}_x} \) is the LNT NOx storage capacity. The \( \text{NO}_x \) storage and release rates can be expressed by:

\[
r_{\text{NO}_x,\text{stor}} = k_{st} \eta \left( \frac{1-e^{\gamma x_2}}{e^{\gamma x_2} - 1} \right) m_{\text{NO}_x,\text{in}},
\]

\[
r_{\text{NO}_x,\text{rel}} = k_{rel} \left( \frac{e^{\gamma \phi x_1} - 1}{e^{\gamma \phi x_1} - 1} \right) (m_{\text{CO}} + m_{\text{HC}})_{\text{in}},
\]

where the multiplier \( \gamma \) depends linearly on the catalyst temperature. The efficiency term \( \eta \) accounts for the oxygen available to promote the \( \text{NO}_x \) storage reactions.

The trap regeneration is modeled considering two sequential phases. First, the stored \( \text{NO}_x \) are released from the trap in presence of rich exhaust gas conditions. After that, the released \( \text{NO}_x \) is converted to \( \text{N}_2 \) by reductants. The conversion process depends on the
catalyst temperature, fill ratio, and concentration of CO and hydrocarbons available in the mixture. Therefore, the mass rate of NOx conversion to N₂ can be defined as:

\[ r_{\text{NOx, conv}} = k_{\text{conv}} \eta_{\text{conv}} r_{\text{NOx, rel}}, \]  \hspace{1cm} (5)

where the conversion efficiency is a function that includes the effects of the mentioned variables,

\[ \eta_{\text{conv}} = \left( \frac{e^{\lambda T} - e^{\lambda u}}{1 - e^{\lambda u}} \right) \left( \frac{e^{-\sigma n} - 1}{e^{-\sigma n} - 1} \right), \]  \hspace{1cm} (6)

where \( \lambda \) is a linear function of the catalyst temperature and \( u \) represents the mass of equivalent reductants (CO + \( C_n H_m \)) that is required to convert the released NOx, normalized to the corresponding stoichiometric value [11, 17, 23].

**CATALYST TEMPERATURE DYNAMICS**

The LNT temperature dynamics can be characterized by applying the energy conservation principle to the catalyst brick, assuming its temperature equal to the one of the gas contained in the system. This simplification can be justified by the small diameter of the passages inside the catalyst, which reduces the average gas velocity. With this assumption, the energy balance yields:

\[ \frac{dT}{dt} = \frac{1}{M_{\text{cat}} c_{\text{cat}}} [m_{\text{gas}} c_p (T_{\text{in}} - T) - Q_{\text{reac}} - Q_{\text{ht}}] \]  \hspace{1cm} (7)

where \( M_{\text{cat}} c_{\text{cat}} \) denotes the catalyst thermal capacity, \( c_p \) is the specific heat of the feed gas, and \( Q_{\text{ht}} \) is relative to the heat losses, mainly due to convection. The term \( Q_{\text{reac}} \) represents the total reaction enthalpy produced by the release and conversion of the stored O₂ and NOx approximated from the results of an exotherm analysis of a LNT catalyst [20].

**LNT MODEL VALIDATION**

The LNT model developed has been validated with experimental data obtained on a 4 cylinder Diesel engine equipped with an aftertreatment system consisting of a dual lean NOx traps and a bypass regeneration system [23].

The inputs to the model are the engine air mass flow rate, the air/fuel ratio, and the exhaust gas temperature. The inlet composition, in particular the NOx concentration, was determined by measurements (in particular, a NOx analyzer was available) and by elementary mass balances based on equilibrium reactions [18]. The composition of the exhaust gas (in particular, the concentrations of CO₂, CO, O₂ and hydrocarbons) is implemented in the model as a function of the air mass flow rate and air/fuel ratio, based on the steady state engine emission maps. The composition during rich operations was estimated from experimental data obtained on a Diesel fuel reformer for regeneration of LNT systems [13]. Figure 2 and Figure 3 show the profiles of the main model inputs, namely engine MAF, AFR and NOx concentration in the exhaust, as obtained in the experimental tests.
\[ \dot{x}(t) = f(x(t), u(t)), \quad (8) \]

\[ x(t): \text{ vector of system states} \]
\[ u(t): \text{ vector of system inputs}. \]

In NMPC, generally, the model equations are discretized first in order to increase the computation efficiency. The discretized system model can then be expressed as:

\[ x(k+1) = F(x(k), u(k)). \quad (9) \]

NMPC uses the discretized model to predict the system behavior in a receding horizon from current time \( t_0 \) to \( t_n \). The optimal control sequence \( \{u(1), u(2), ..., u(n)\} \) is calculated by solving the following optimization problem using numerical methods.

Cost function:

\[ j = \sum_{k=0}^{n} J(u(k), x(k)), \]

Constraints:

\[ x(k) \in X, u(k) \in U, \]

Control variable:

\[ u(1, 2, ..., n), \quad (10) \]

The closed loop NMPC is achieved by applying the first \( m \) control signals to the system.

\[ u(1:m) = \{u(1), u(2), ..., u(m)\}, \quad \text{where } (m < n). \quad (11) \]

When the \( m \)-th control output is applied to the system, NMPC starts the prediction of the next receding horizon and calculate the next optimal control sequence. At this time, the final states of the previous horizon \( x(m) \) is fed back to NMPC to be the initial states of the new horizon \( x(0) \). The architecture of the nonlinear NMPC can be expressed by the following picture.

![Figure 5 NMPC scheme](image)

In this paper, the objective of the AFR control during LNT regeneration is to minimize the fuel penalty used for \( NO_x \) desorption and the overall tailpipe \( NO_x \) emission. As indicated in [11] and [12], during LNT regeneration, the relations between AFR and \( NO_x \) desorption efficiency and tailpipe \( NO_x \) emission are highly nonlinear and time varying as well. In this case, an optimal control strategy that only considers the current state optimization will not necessarily produce the optimal performance. Because the LNT release efficiency changes with time, intuitively, to avoid wasting fuel during regeneration, more reductants should be applied when the release efficiency is high and less regeneration fuel should be used when the LNT release efficiency is low. Moreover, in order to reduce the tailpipe \( NO_x \) slip, \( NO_x \) absorption efficiency should be considered as well. To effectively address these issues and achieve the best tradeoff between fuel consumption and tailpipe \( NO_x \) reduction, knowledge or prediction of LNT release, conversion, and absorption efficiencies need to be systematically incorporated in the regeneration AFR control law. By employing a fairly accurate LNT model that captures these characteristics, MPC can therefore produce a control law (AFR) to optimally deal with these plant variations and features. The details of the NMPC regeneration AFR control design are presented in the following sub-sections.

MODEL PREDICTION

The LNT model described in section II was simplified to a three-state rich condition model in order to increase the computational efficiency. The three states are the stored oxygen fill ratio, \( x_1 \), stored \( NO_x \) fill ratio, \( x_2 \), and LNT temperature, \( x_3 \). The inputs and outputs of the simplified model can be described by the following figure.

![Figure 6 simplified model for prediction](image)

The initial values of the three states are obtained from sensors and/or models. Two independent external signals are engine speed and engine torque. These signals can be obtained from the driving cycle, assumed as constants, or predicted based on their past values. AFR is the design variable that NMPC calculates to minimize the cost function. The output of the model is the estimated \( NO_x \) desorption rate. Using this model, NMPC predicts the \( NO_x \) storage in the next \( n \) steps.

CONTROL OPTIMIZATION

The NMPC AFR control optimization is formulated as the follows.
The cost function represents the ratio between the extra fuel mass required to produce the desired AFR at the LNT upstream and the released NOₓ mass during regeneration, plus a weighted LNT NOₓ fill ratio at the end of the NMPC horizon. Both terms were selected based on physical insights into the LNT system dynamics. The first term of the cost function attempts to maximize the NOₓ desorption efficiency. The second term in the cost function is to enforce a low NOₓ fill ratio at the end of regeneration. Consider the fact that the LNT absorption efficiency is high when the NOₓ storage is low and the reductants always react with stored oxygen first before they do with the stored NOₓ during regeneration, it is desirable to have the LNT capacity regenerated as much as possible during a regeneration. The second term in the cost function therefore has the intention to maximize the LNT capacity at the end of regeneration. It is worthy to mention that the two terms in the cost function have the same order of magnitude.

Numerical optimization algorithm is implemented to minimized the cost function by specifying the optimal control sequence \{u(1), u(2), ..., u(n)\} in the horizon from \(t_0\) to \(t_n\).

**APPLY THE OPTIMAL CONTROL SEQUENCE**

The first \(m\) control inputs of the optimal control sequence \{u(1), u(2), ..., u(m)\} are applied to the exhaust gas AFR controller accordingly. After the \(m\)th control input is applied to the system, the current sensor information is fed back to the NMPC controller to calculate the initial state \(x(0)\) of the new receding horizon. And NMPC starts calculating the new optimal control sequence according to the initial state values.

The overall control scheme of NMPC for LNT regeneration AFR control is shown in the following figure.

**SIMULATION RESULTS AND ANALYSIS**

A complete Diesel engine NOₓ aftertreatment simulator named cX-Emissions¹ was built in Matlab/Simulink [13]. The aftertreatment model was coupled with a quasi steady-state vehicle model that includes simple models for engine, transmission, and powertrain components. The simulator is calibrated with data obtained from a midsize SUV powered with a 1.9l Diesel engine. This allows for the possibility of simulating the aftertreatment performance during steady-state and transient driving conditions. Figure 8 illustrates the combined vehicle-engine-aftertreatment simulation model structure.

FTP75 cycle was used in the simulation. Four different cases were compared in this study. In the first case, no regeneration was applied to the LNT. This case was set as a reference to compare with other cases using different LNT regeneration AFR controllers. It should be noted that in case 2, case 3, and case 4, same LNT regeneration trigger strategy, which was based on the LNT NOₓ fill ratio, was used.

In case 2, a PID controller from a previous work [13] is used to control the LNT regeneration AFR. The PID controller tries to regulate the NOₓ slip to zero by adjusting the AFR during regeneration. The NOₓ slip is defined by the following equation,

\[
S_{NO_x} = \frac{m_{NO_x,in} - m_{NO_x,out}}{m_{NO_x,in}}. \tag{13}
\]

\(m_{NO_x,in}\): mass flow rate of NOₓ into LNT
\(m_{NO_x,out}\): mass flow rate of NOₓ out of LNT
In case 3, the proposed NMPC control scheme was used. In this case, the weighting factor $k$ was set to zero in order to minimize the fuel cost used during $NO_x$ desorption.

In case 4, besides the fuel cost minimization, we also want to minimize the $NO_x$ fill ratio at the end of the regeneration. The weighting factor $k$ was set to 0.8 in this case.

Figure 9 and Error! Reference source not found. show the fuel consumptions at the end of the FTP cycles for each of the four cases. As expected, case 1 has the least fuel consumption because no regeneration was triggered, so no fuel penalty. The fuel consumption of 4597.4 g was defined as the reference value of the best fuel consumption for this driving cycle. Comparing case 2 and case 3, we did not find fuel consumption improvement when NMPC was applied. This result can be explained in Figure 10, which is a comparison of $NO_x$ storages during the cycle.

Table 1 compare fuel consumption and NOx emission with different AFR control strategy

<table>
<thead>
<tr>
<th></th>
<th>Fuel Penalty (g)</th>
<th>Tailpipe NOx (g)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case 1</td>
<td>4597.4</td>
<td>17.34</td>
</tr>
<tr>
<td>Case 2</td>
<td>4629.8</td>
<td>7.83</td>
</tr>
<tr>
<td>Case 3</td>
<td>4631.5</td>
<td>9.09</td>
</tr>
<tr>
<td>Case 4</td>
<td>4622.1</td>
<td>6.58</td>
</tr>
</tbody>
</table>

In Figure 10, we see the NMPC in case 3 stopped regeneration when only half or less of the stored $NO_x$ was purged. This was because the $NO_x$ release efficiency was low when the $NO_x$ storage ratio was low, and the low $NO_x$ release efficiency causes more fuel consumption to remove the $NO_x$ stored from LNT. To minimize the ratio between extra fuel and released NOx, the NMPC decided to stop the regeneration when $NO_x$ storage was low. When the $NO_x$ adsorption starts, because the $NO_x$ storage capacity was smaller, another $NO_x$ regeneration would be needed shortly. There were six regenerations in case 3, compared to five regenerations in case 2. Because some fuel has to be used to remove the stored oxygen in LNT at the beginning of the regeneration, the higher regeneration number leads to a higher overall fuel penalty.

Figure 11 is the comparison of AFR control in case 2, 3, and 4. We can see that the AFR of case 3 is higher than case 4 during regeneration. It verifies the explanations that the case 3 AFR becomes higher when the $NO_x$ storage is low. The NMPC in case 4 enforced a more complete regeneration using lower AFR value while maintaining the regeneration efficiency close to the optimum value.
The objective of NMPC on AFR control is to reduce the fuel penalty used for $N_O_x$ desorption and the tailpipe $N_O_x$ emission simultaneously. The prediction ability of NMPC enables more stored $N_O_x$ to be released from LNT when the release efficiency is high. Moreover, the introduction of the weighted $N_O_x$ fill ratio at the end of regeneration in the cost function increased the LNT absorption efficiency during lean operations and reduced the overall fuel penalty. Simulation studies using cX-Emissions show the proposed NMPC regeneration AFR controller can reduce 23.7% of regeneration fuel penalty and 16% of tailpipe $N_O_x$ emission compared with a conventional LNT regeneration AFR controller.

**FUTURE WORK**

In this paper, the NMPC was used only for the AFR ratio control during regeneration. However, similar model predictive control methodology can be employed to conduct the optimal LNT regeneration triggering control as well. It is anticipated that a dual-loop MPC with the outer loop MPC controlling the LNT regeneration timing/duration, and inner loop MPC controlling the regeneration AFR would result in a minimal overall regeneration fuel penalty while maintaining tailpipe NOx emissions under the stringent regulations. Moreover, in the NMPC algorithm used in this paper, constant LNT storage capacities were assumed. It is typical that sulfur poisoning can considerably reduce the LNT capacity during its lifetime. Adaptive MPC may provide more promising potentials in terms of ensuring long-term real-world fuel penalty and tailpipe NOx emission reductions. The authors are actively working on the above research directions and will report the results in the future publications.

**ACKNOWLEDGMENTS**

The authors would like to thank the financial support provided by the OSU CAR Industrial Consortium members to this project.

**REFERENCES**

12. P. Pisu, M. Canova, and A. Soliman, "Model-Based Fault Diagnosis of a NOx Aftertreatment System", in proceeding of the 17th IFAC, Korea, 2008.

CONTACT

Junmin Wang, Ph.D.
Assistant Professor
Department of Mechanical Engineering
The Ohio State University
Tel: 614-247-7275
Email: wang.1381@osu.edu

Ming Feng Hsieh
Department of Mechanical Engineering
Center of Automotive Research
930 Kinear Rd. Columbus OH, 43210, USA
Tel: 614-596-4161
Email: hsieh.122@osu.edu