



Optimal Control of Mass Transport Time-Delay Model in an EGR

Sandra Malik Hamze Renault

Didier Georges Grenoble INP

Emmanuel Witrant GIPSA-lab

Delphine Bresch-Pietri Mines ParisTech

Citation: Hamze, S.M., Georges, D., Witrant, E., and Bresch-Pietri, D., "Optimal Control of Mass Transport Time-Delay Model in an EGR," SAE Technical Paper 2020-01-0251, 2020, doi:10.4271/2020-01-0251.

Abstract

This paper touches on the mass transport phenomenon in the exhaust gas recirculation (EGR) of a gasoline engine air path. It presents the control-oriented model and control design of the burned gas ratio (BGR) transport phenomenon, witnessed in the intake path of an internal combustion engine (ICE), due to the redirection of burned gases to the intake path by the low-pressure EGR (LP-EGR). Based on a nonlinear AMESim® model of the engine, the BGR in the intake manifold is modeled as a state-space (SS) output time-delay model, or alternatively as an ODE-PDE coupled system, that take into account the time delay between the moment at which the combusted gases leave the exhaust manifold and that at which they are readmitted in the intake manifold. In addition to their mass transport delay, the BGRs in the intake path are also subject to state and input inequality constraints. The objective of the control problem is to track a reference output profile

of the BGR in the intake manifold, taking into account the transport delay and the state (output) and input constraints of the system. In this aim, two indirect optimal control approaches are implemented and compared, the discretize-then-optimize approach and the optimize-then-discretize approach. To account for the state inequality constraints, both methods are equipped with techniques for constrained optimization such as the augmented Lagrangian and the UZAWA methods. The necessary conditions of optimality are formulated, in each of both cases, and the resulting equations are solved numerically using the projected gradient-descent method, which ensures the non-violation of the input inequality constraints. The novelty of the work lies in considering the system's constraints and the infinite-dimensionality of the mass transport phenomenon governing it. The merits of the time-delay model and the model-based control design are illustrated on the nonlinear® AMESim model on which the mathematical model is based.

Introduction

The LP-EGR is a system which draws off some of the exhaust gases, cools them down by means of a heat exchanger, and redirects them back into the air intake system. In gasoline engines such as the H5Ft400, the technology aims at promoting fuel economy and improving knock resistance, which comes by as a natural effect of the reduction in peak in-cylinder temperatures.

Overlooking the dynamics of dilution in the volume downstream the EGR valve and assuming that fresh air and recirculated gases blend together instantaneously, the BGR is calculated as

$$BGR = \frac{Q_{EGR}}{Q_{EGR} + Q_{air}} \quad (1)$$

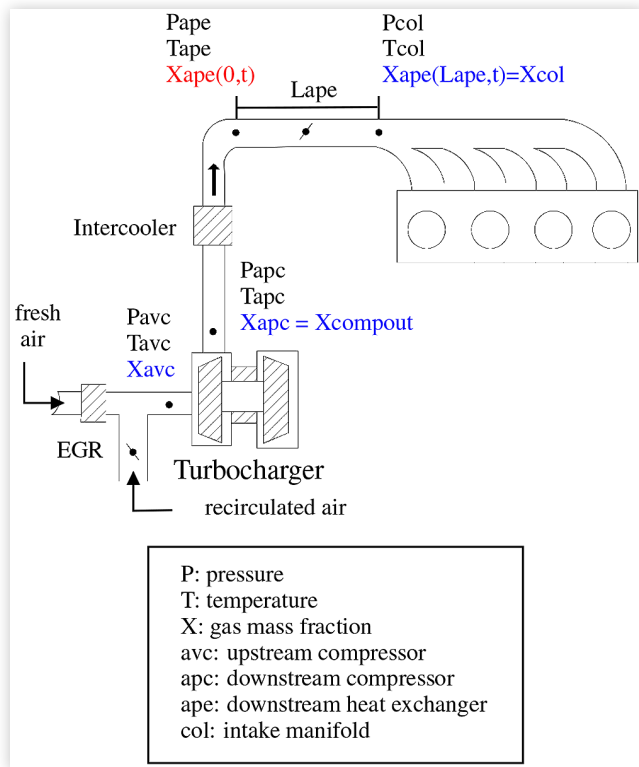
where Q_{EGR} is the mass flow rate of the recirculated exhaust air admitted through the EGR valve to the air intake line, and Q_{air} is the mass flow rate of the cold fresh air.

Understanding the dynamics of the BGR is fundamental for its control, which is why we hereby present the mathematical equations governing the physics behind the mixture of burned gases and fresh air, as it propagates in the intake line, shown in [Figure 1](#). The burned gas fraction along the line evolves progressively with time, and its dynamics vary according to the intake line components it is traversing. Two kinds of components can be distinguished: tube sections, where gas transport takes place, and control volumes, where gas mixing takes place [1].

Let $X(x, t)$ be the BGR at time t for a spatial coordinate $x \in [0, L]$. Its dynamics can be modeled using a first-order hyperbolic PDE expressed as [1]:

$$\partial_t X(x, t) + u(x, t) \partial_x X(x, t) = 0 \quad (2)$$

where $X(0, t) = X_{in}(t)$ and $X(x, 0)$ are the respective boundary and initial conditions, and $u(x, t)$ is a propagation speed.

FIGURE 1 Air intake path scheme

Knowing the boundary condition $X_{in}(t)$ at the entrance of the tube section, and considering $u(x, t) = u(t)$ (with a plug-flow assumption, considering the fluid as incompressible), the burned gas fraction at the exit of the tube section can be calculated, using equation (2) and the method of characteristics, as [2]:

$$X(L, t) = X(t - \tau_f(t)) \quad (3)$$

where $\tau_f(t)$ is the time-varying time delay due to the transport of the gas inside the tube section, calculated as [1]:

$$\tau_f(t) \approx \frac{P(t)V_{tube}}{RT(t)Q(t)} \quad (4)$$

where V_{tube} is the volume of the tube section (m^3), $R = 287.058$ ($J \cdot kg^{-1} \cdot K^{-1}$) is the gas constant, $T(t)$ is the temperature (K), $Q(t)$ is the mass flow rate ($kg \cdot s^{-1}$), and $P(t)$ is the pressure (Pa) of the mixture at time t .

Equation (3) shows that the transport of the BGR inside the tube is, in fact, not immediate and subject to time delay. Taking into account this transport time is important in the case of LP-EGR, as the distance traveled by the recirculated gas is relatively long, and neglecting it could lead to severe performance degradation, especially during transient phases.

In a control volume, such as the intake manifold, by ignoring the air fraction dynamics with respect to space, the volume-average air fraction dynamics with respect to time can be formulated as a 0-D model, implying the law of conservation of mass and expressed as [1]:

$$\dot{X}_{cv} = \frac{RT_{cv}}{P_{cv}V_{cv}} \left[\sum_{in=1}^l (X_{in}Q_{in}) - X_{cv} \sum_{out=1}^o (Q_{out}) \right] \quad (5)$$

where X_{cv} , T_{cv} (K), P_{cv} (Pa), and V_{cv} (m^3) are the control volume's respective BGR, temperature, pressure, and volume, X_{in} and Q_{in} ($kg \cdot s^{-1}$) are the respective BGR and mass flow rate of the l gases coming into the control volume, and Q_{out} ($kg \cdot s^{-1}$) is the mass flow rate of the o gases going out of the control volume.

The amount of exhaust gas being recirculated has a huge impact on whether the EGR is functioning in favor of the engine overall performance or against it, which is why controlling the BGR is essential to procure its benefits and avoid its drawbacks. For example, increasing the BGR raises the temperature of the intake mixture, leading to better vaporization of the fuel, better homogeneity of the mixture in the intake manifold, and therefore better combustion. This is highly recommended in cold weather, or directly after starting the engine to accelerate its warm-up. Nevertheless, the BGR can't be increased unlimitedly, because the higher the BGR, the more dilute the mixture in the intake manifold is. This mixture is hence poor in fuel and oxygen, and therefore low in power, which is inconvenient when high or full power is in demand [3]. So the BGR ratio can be advantageous or disadvantageous depending on the operating conditions of the engine, the driver's commands and expectations, the external environment, and several other factors. Taking into account all these factors, experts define a cartography that maps different engine operating conditions to setpoints or reference values of X_{col} , the exhaust gas mass fraction in the intake manifold. The purpose of EGR control is to manipulate the opening angle of the EGR valve such that X_{col} attains its desired reference value, at the current operating conditions, while taking into account the transport delay of the intake gas and respecting the physical constraints of the valve, all at the same time.

Different kinds of models were used in the literature to model the engine air-path and the EGR system. Most of the studies relied on mathematical and MVEM nonlinear models, inspired from the physical principles, to model the air-path and EGR system dynamics, take for example [4, 5]. Among these studies, many linearized the physical models [6], sometimes as LPV models [7] or LTV models [8], with or without model reduction. Nevertheless, these models did not explicitly take the transport delay into account, unlike [9, 1], where the delay due to the mass transport phenomena resulted in SS formulations, describing the dynamics of X_{col} with time-varying input and state delays, respectively. The model in [9] uses the input-delay model, describing X_{col} by means of a composition balance equation, for control purposes. The control input is the burned gas ratio downstream the EGR valve, delayed by the summation of the times needed to transport the gas from downstream the EGR valve to the compressor, from downstream the compressor to the intercooler, and from downstream the intercooler to the intake manifold. From the input, the EGR mass flow rate can be deduced, and the conversion of the latter to the effective valve opening is done using the Saint-Venant equation. The model in [1], on the other hand, is used for observation purposes. It divides the air intake path

into control volumes, in which the air fraction dynamics are modeled as an ODE, and into tube sections, in which the air fraction dynamics are modeled as first-order hyperbolic PDEs, which can be reformulated as a time-delay system by means of the method of characteristics. The delay in this case corresponds to the transport of the gas from downstream the compressor to the intercooler, inside the intercooler (considered here as a tube section), and from downstream the intercooler to the intake manifold.

Different control techniques have been used in the literature to control the EGR valve and the burned gas ratio in the intake manifold; take for example MPC [7], PID [10], and LQG [11], among others. Despite the efficacy of these techniques in controlling the air or burned gas fraction in the intake line, yet they suffer from a major drawback, which is their reliance on 0-D, instead of 1-D, air-path models. By doing so, they neglect the time delays induced by the transport of the gas mixture in the air-path tubes, which might have serious impacts on the efficiency of the air-path control design, and consequently on the overall engine performance. To avoid this drawback, automakers and researchers started integrating 1-D air-path models in the engine control design. In [9], a prediction-based trajectory tracking control is designed on a time-varying input delay model to determine the intake BGR for gasoline engines. In [1], where air fraction transport phenomena in a LP-EGR-equipped diesel engine are modeled by means of a cascade of first-order LPV hyperbolic systems with dynamics associated with the boundary conditions, Lyapunov-based techniques and matrix inequalities are employed to find sufficient conditions for the exponential stabilization and observation of this class of systems. Overall, 1-D model-based air-path control design is a topic which is still under-investigated.

The transport of the BGR in the air intake path is modeled, as per [1], as a set of coupled ODEs and PDEs attributed respectively to a succession of control volumes and tube sections. In this work, we consider the average BGR dynamics in the tube downstream the compressor, considered as a control volume, and model it using a 0-D model. The part of the air intake path extending from upstream the heat exchanger until the intake manifold is considered as a tube section and modeled using a 1-D model. This results in a simplified ODE-PDE coupled system, whose one-dimensional part is discretized, for control purposes, by means of the method of lines [12]. To track a reference profile of X_{col} , the EGR system is subject to two indirect optimal control approaches: discretize-then-optimize and optimize-then discretize [13]. In the former, the PDE is discretized first, leading to a delay-free traditional SS system, and the calculation of the necessary conditions of optimality follows afterwards. In the latter, on the other hand, the necessary conditions of optimality stem directly from the ODE-PDE coupled system, and then the discretization takes place to solve the resulting boundary-value problem. In both cases, however, because the BGR is limited between 0 and 100, the augmented Lagrangian method is employed to ensure that the controlled output respects this constraint. The resulting system of equations representing the necessary optimality conditions is solved numerically using the projected gradient-descent method, which makes sure that the constraints on the control input X_{avc} (which is the BGR upstream the compressor)

are respected. Although constraining the control problem is not a novel idea in the field of EGR control (take for example [14]), but explicitly accounting for the input and output constraints while considering an infinite-dimensional air-path model that accounts for its transport delays, is, to the best of our knowledge, novel. The merits of using these control techniques are evaluated on the nonlinear AMESim® model from which the mathematical model was initially extracted.

The remainder of this paper is organized as follows. Section 3 presents a delay linear system modeling of the EGR mass fraction, and section 4 formulates the EGR control problem. Section 6 recapitulates the augmented Lagrangian method, which is at the core of the considered indirect optimal control method, in its turn recapitulated in section 5. The application of the indirect discretize-then-optimize and optimize-then-discretize approaches is respectively detailed in sections 7 and 8, along with their simulation results (section 9). Finally, section 10 concludes the paper and points out future work perspectives.

EGR Linear State-Space Model

The schematic presentation of the section of the air intake path extending from the LP-EGR valve to the intake manifold is shown in [Figure 1](#). It shows the fresh air coming from the external environment mixing with the recirculated gases, admitted to the intake path through the EGR valve, in the volume upstream the compressor (AVC). The mixture of gases then traverses the compressor, the volume downstream the compressor (APC), the heat exchanger (HE), and the tube section following it (APE). Note how AVC and APC are considered as control volumes, where mere gas mixing occurs, whereas HE and APE are considered as tube sections. $X(x, t)_n$ and $X(t)_m$ are the respective burned gas fractions in the tube sections $n = \{HE, APE\}$ and control volumes $m = \{COL, APC, AVC\}$. x represents the spatial coordinate which spans an interval $[0, L_n]$ (L_n being the length of tube section n), and t represents the time coordinate which spans an interval $[t_0, \infty)$.

To calculate the burned gas fractions in the intake path, we consider the following assumptions, inspired from [1]:

- The control volumes in the EGR path are large compared to their length, which means that the mass transport phenomenon is negligible compared to the mixing phenomenon, which justifies modeling the gas mixing in these volumes as ODEs.
- the gas mixing in the heat exchanger of the intake path is considered negligible, and the heat exchanger is entirely considered as a tube section;
- the mixing dynamics in the intake manifold are negligible;

From the first two assumptions, and from [equation \(5\)](#) describing the BGR dynamics in a control volume, the BGR dynamics in the control volume downstream the compressor can be written as:

$$\dot{X}_{apc} = \alpha_{apc} [-Q_{engine} X_{apc}(t) + Q_{engine} X_{avc}(t)] \quad (6)$$

where $\alpha_{apc} = \frac{RT_{apc}}{P_{apc} V_{apc}}$, X_{apc} (%), T_{apc} (K), P_{apc} (Pa), and V_{apc} (m^3) are the respective BGR, temperature, pressure, and volume in the control volume downstream the compressor, and Q_{engine} ($kg \cdot s^{-1}$) is the mass flow rate of the gas flowing in the intake air-path, equal to $(Q_{air} + Q_{EGR})$ (assuming that the speed dynamics of the particles are fast in comparison to the dynamics of the BGR).

According to [9], the control input to the model describing the dynamics of X_{col} is X_{avc} , the burned gas ratio downstream the EGR valve, delayed by the transport time in the intake path. It is chosen as control input because, once calculated, it can be translated to EGR mass flow rate by virtue of [equation \(1\)](#), and then to effective valve opening by virtue of the Saint-Venant equation.

The third assumption implies that the dynamics of X_{col} are not expressed as an ODE, as is the case for control volumes. Instead, $X_{col}(t)$ is supposed to be equal to $X_{ape}(L_{ape}, t)$, calculated at the very end of the tube section downstream the heat exchanger. This assumption, along with [equation \(3\)](#) accounting for the transport of the BGR in the heat exchanger and in the tube section following it, can be formulated as:

$$X_{col}(t) = X_{ape}(L_{ape}, t) = X_{apc}(t - \tau_{total}(t)) \quad (7)$$

where X_{col} (%) and X_{ape} (%) are the respective BGRs in the intake manifold and in the tube downstream the heat exchanger, $\tau_{total}(t) = \tau_{ape}(t) + \tau_{HE}(t)$, $\tau_{ape}(t) = \frac{P_{ape}(t) V_{ape}}{RT_{ape}(t) Q_{engine}(t)}$ and $\tau_{HE}(t) = \frac{P_{HE}(t) V_{HE}}{RT_{HE}(t) Q_{engine}(t)}$ are the respective delays, measured in seconds, due to the transport of the gas in the tube downstream the heat exchanger and in the heat exchanger itself, L_{ape} (m), T_{ape} (K), P_{ape} (Pa), and V_{ape} (m^3) are the respective length, temperature, pressure, and volume of the tube downstream the heat exchanger, and T_{HE} (K), P_{HE} (Pa), and V_{HE} (m^3) are the respective temperature, pressure, and volume of the heat exchanger.

Therefore, the resulting SS model is an output delay model expressed as:

$$\begin{aligned} \dot{\mathcal{X}}(t) &= A(t)\mathcal{X}(t) + B(t)U(t) \\ Y &= \mathcal{X}(t - \tau_{total}(t)) \end{aligned} \quad (8)$$

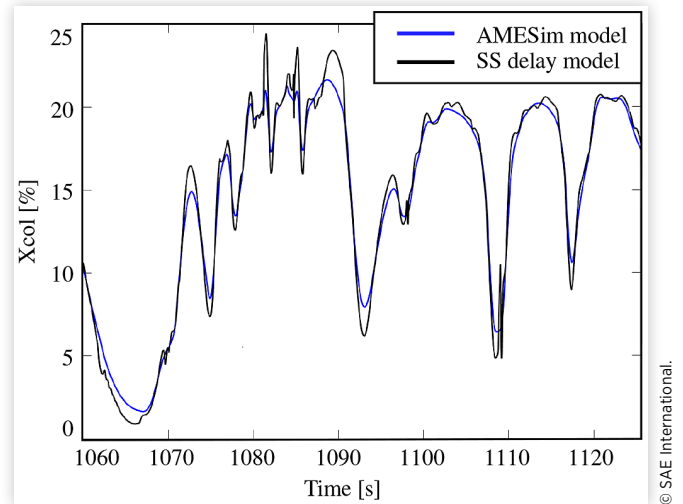
where $A(t) = -\alpha_{apc} Q_{engine}$, $B(t) = \alpha_{apc} Q_{engine}$, $\mathcal{X}(t) = X_{apc}(t)$, $U(t) = X_{avc}(t)$, and $Y(t) = X_{col}(t) = X_{apc}(t - \tau_{total}(t))$.

Note that, alternatively, in view of [\(2\)](#), these dynamics can be formulated as the following ODE-PDE system:

$$\begin{aligned} \dot{\mathcal{X}}(t) &= A(t)\mathcal{X}(t) + B(t)U(t) \\ \partial_t X(x, t) + u(t) \partial_x X(x, t) &= 0 \\ X(0, t) &= \mathcal{X}(t) \\ Y(t) &= X(1, t) \end{aligned} \quad (9)$$

where $u(t) = \frac{1}{\tau_{total}(t)}$ is the BGR propagation speed, assuming a normalized length $L_{ape} = 1$.

FIGURE 2 Xcol from AMESim® and the linear SS delay model



[Figure 2](#) shows X_{col} generated by the AMESim® model and by model [\(8\)](#). The small difference between the two plots is an indicator of the adequacy of the assumptions.

Control Problem Formulation

The objective of controlling the EGR is to drive the exhaust gas mass fraction in the intake manifold $X_{col}(t)$ to its setpoint $r(t)$, using minimum control effort U . Therefore, an objective function to minimize, with respect to U , can be expressed as:

$$\begin{aligned} J_1(t) &= \frac{1}{2} \|X_{col}(T) - r(T)\|_{L_1}^2 \\ &+ \frac{1}{2} \int_0^T \|X_{col}(t) - r(t)\|_{Q_1}^2 + \|U(t)\|_{R_1}^2 dt \end{aligned} \quad (10)$$

where $\|\cdot\|_{\beta}$ (with $\beta^T = \beta$ a positive matrix) denotes the weighted Euclidean norm, $L_1 \in \mathbb{R}_{\geq 0}$, $Q_1 \in \mathbb{R}_{\geq 0}$, and $R_1 \in \mathbb{R}_{> 0}$ are the matrix weights used to manage the trade-off between the different minimization objectives implied in the cost function J_1 , and T is the final time.

This optimization problem is subject to the infinite-dimensional dynamics described in [equation \(8\)](#) or [\(9\)](#), and to constraints on the input $U(t) = X_{avc}(t)$ and on the output/state $X_{col}(t)$, because these quantities represent gas mass fractions whose values vary between 0% and 100%. This leads to the inequality constraints

$$\begin{aligned} 0 &\leq X_{col} \leq 100 \\ 0 &\leq U \leq 100 \end{aligned}$$

Based on the above, we formulated a constrained optimization problem of a time-delay system. Before addressing the application of the indirect optimal control methods, the next sections give a general overview of their principles, with

a focus on the augmented Lagrangian method, and the UZAWA method, which are used to manage the input and state constraints.

Indirect Optimal Control Methods

The indirect methods, which are the methods adopted in the sequel for the air-path control problem, solve the optimal control problem, in continuous time, using Pontryagin Minimum Principle [15], which provides the necessary conditions for optimality, using the calculus of variations. For a detailed explanation of the indirect methods and the necessary conditions for optimality in the cases of unconstrained admissible controls and states, constrained controls, and constrained states, one can refer to [16], and for an extension to the case of constrained controls and states, one can refer to [17]. Though fundamentally developed for finite-dimensional state dynamics, these methods found their way through infinite-dimensional state dynamics, thanks to the works of [18, 19, 20, 21], and others.

As already mentioned in the introduction, the air-path control problem is approached from two different perspectives: discretize-then-optimize and optimize-then-discretize. In the discretize-then-optimize method, the discretization of the system transforms the infinite-dimensional PDE into a finite-dimensional system of ODEs, before the calculation of the first-order optimality conditions takes place. This leads to the definition of a standard optimal control problem, unlike the optimize-then-discretize method which preserves the infinite-dimensional system dynamics, as it calculates the necessary optimality conditions before discretizing the PDE. As both methods are addressed in this work, we proceed by stating the necessary conditions of optimality in each of both cases.

Necessary Optimality Conditions in Case of Finite-Dimensional Dynamics

The necessary conditions, corresponding to the finite-dimensional case, are mentioned briefly below, by referring to [16], which provides a thorough demonstration of their origins.

Let s be the state of the system, U be its control input, d be the function describing the dynamics, J be the cost function, \mathcal{F} be the terminal cost, \mathcal{G} be the cost-to-go, and t be the time falling in the interval $[t_0, t_f]$.

As a first step, the admissible state and control input are considered to be unbounded. We also assume that the initial state and time s_0 and t_0 are specified, the final time t_f is specified, whereas the final state s_f can be specified or free.

After discretization, the dynamics of the system can be described as:

$$\dot{s}(t) = d(s(t), U(t), t) \quad (11)$$

The objective of the controller is to find the optimal control input U^* that minimizes the cost function

$$J(U) = \mathcal{F}(s(t_f), t_f) + \int_{t_0}^{t_f} \mathcal{G}(s(t), U(t), t) dt \quad (12)$$

For this purpose, define the Hamiltonian

$$\mathcal{H}(s(t), U(t), p(t), t) = \mathcal{G}(s(t), U(t), t) + p^T(t) d(s(t), U(t), t) \quad (13)$$

where $p(t)$ is the adjoint state or co-state.

For all $t \in [t_0, t_f]$, the necessary conditions of optimality can be written as:

$$s^*(t) = \frac{\partial \mathcal{H}}{\partial p}(s^*(t), U^*(t), p^*(t), t) \quad (14a)$$

$$\dot{p}^*(t) = -\frac{\partial \mathcal{H}}{\partial s}(s^*(t), U^*(t), p^*(t), t) \quad (14b)$$

$$0 = \frac{\partial \mathcal{H}}{\partial U}(s^*(t), U^*(t), p^*(t), t) \quad (14c)$$

$$\left[\frac{\partial \mathcal{F}}{\partial s}(s^*(t_f), t_f) - p^*(t_f) \right]^T \delta s_f + \left[\mathcal{H}(s^*(t_f), U^*(t_f), p^*(t_f), t_f) + \frac{\partial \mathcal{F}}{\partial t}(s^*(t_f), t_f) \right] \delta t_f = 0 \quad (15)$$

where s^* , U^* , and p^* are the optimal state, control input and co-state, respectively.

Note that the necessary conditions (14a), (14b), and (15) remain unchanged in case of a constrained input. Condition (14c), however, becomes

$$\mathcal{H}(s^*(t), U^*(t), p^*(t), t) \leq \mathcal{H}(s^*(t), U(t), p^*(t), t) \quad (16)$$

for all $t \in [t_0, t_f]$ and for all admissible $U(t)$. Equation (16) is, in fact, the Pontryagin Minimum Principle.

Considering the air-path control problem, the final time is specified ($t_f = T$) implying that $\delta t_f = 0$, and the final state is free implying that $\delta s_f \neq 0$. Therefore, in this case, (15) reduces to $p^*(T) = \frac{\partial \mathcal{F}}{\partial s}(s^*(T), T)$.

Necessary Optimality Conditions in Case of Infinite-Dimensional Dynamics

For the calculation of the necessary conditions in the infinite-dimensional case, we refer to the Lagrangian-based adjoint method, from [13], which is a general optimal control approach that applies to linear and nonlinear PDEs. We consider minimizing a cost function subject to infinite-dimensional equality constraints, such as PDEs representing the system dynamics. We also consider the boundary-control case, because it is representative of the EGR control problem.

Let s be the state of the system, s_r be a bounded reference to be tracked, U be the system's control input, Z be a partial derivative operator, H be a bounded linear operator, q and r be strictly positive weights, x be the spatial coordinate falling in the interval $[0, 1]$, and t be the time coordinate falling in the interval $[t_0, t_f]$.

The dynamics of the system can be described by the following PDE:

$$\partial_t s(x, t) = Z(s(x, t)) \quad (17)$$

where $s(x, t_0) = s_0(x)$ is the initial condition, and $G_1(s(0, t)) = U(t)$, $G_1(s(1, t)) = 0$ is one of two possible boundary-control scenarios.

The objective of the controller is to find the optimal control input U^* which minimizes the cost function

$$J(U) = \int_{t_0}^{t_f} \int_0^1 |H(s(x, t) - s_r(x, t))|^2 dx dt + \int_{t_0}^{t_f} r |U(t)|^2 dt + \int_0^1 q |H(s(x, t_f) - s_r(x, t_f))|^2 dx \quad (18)$$

For this purpose, define the Lagrangian

$$\begin{aligned} \mathcal{L}(s, U, p, \lambda) = & \int_{t_0}^{t_f} \int_0^1 |H(s(x, t) - s_r(x, t))|^2 dx dt \\ & + \int_{t_0}^{t_f} r |U(t)|^2 dt + \int_0^1 q |H(s(x, t_f) - s_r(x, t_f))|^2 dx \\ & + \int_{t_0}^{t_f} \langle \lambda(t), (U(t) - G_1(s(0, t))) \rangle dt \\ & + \int_{t_0}^{t_f} \int_0^1 \langle p(x, t), \partial_t s(x, t) - Z(s(x, t)) \rangle dx dt \end{aligned} \quad (19)$$

where $p(x, t)$ is the adjoint state, $\lambda(t)$ is the Lagrange multiplier associated to the boundary control $G_1(s(0, t)) = U(t)$, and $\langle \cdot, \cdot \rangle$ denotes a scalar product.

The necessary conditions for optimality can be written as:

$$L_U(s^*, U^*, p^*, \lambda^*) = 0 \quad (20a)$$

$$L_s(s^*, U^*, p^*, \lambda^*) = 0 \quad (20b)$$

$$L_p(s^*, U^*, p^*, \lambda^*) = 0 \quad (20c)$$

$$L_\lambda(s^*, U^*, p^*, \lambda^*) = 0 \quad (20d)$$

$$p^*(x, t_f) = 2qH(s^*(x, t_f) - s_r(x, t_f)) \quad (20e)$$

$$s^*(x, t_0) = s_0(x) \quad (20f)$$

where the index of L denotes its partial derivatives with respect to the variable.

Augmented Lagrangian Method

The Augmented Lagrangian approach, proposed by [22, 23], dates back to the year 1969. It transforms a constrained optimization problem to an unconstrained optimization problem by adding a penalty term penalizing the constraint violation. However, it considers the Lagrangian, which is why it is also called the "Method of Multipliers" or the "Penalty-Multiplier Method". Since its first introduction in 1969, different variants of the Augmented Lagrangian method have been introduced to improve the quality of the solution or to solve different kinds of constrained optimization problems; take for example the alternating direction method of multipliers (ADMM) or the generalized Augmented Lagrangians (GAL) method. In the rest of this section, however, we stick to the classical Augmented Lagrangian method considering the case of equality constraints, presented in [22]. Since we are dealing with inequality constraints, we show how a simple manipulation of the constraint inequality function can transform the inequality constraint problem into an equality constraint problem, upon which the classical method can be applied.

By referring to [22], consider the equality constrained problem:

$$\begin{aligned} & \text{minimize } M(\mu) \\ & \text{s.t. } e(\mu) = 0 \end{aligned} \quad (21)$$

where $M: \mathbb{R}^n \rightarrow \mathbb{R}$ and $e: \mathbb{R}^n \rightarrow \mathbb{R}$ are assumed to be twice continuously differentiable functions. To solve the minimization problem (21), the Augmented Lagrangian method suggests using the augmented Lagrangian function

$$\mathcal{L}_a(\mu, \lambda) = M(\mu) + \lambda^T e(\mu) + \frac{c}{2} |e(\mu)|^2 \quad (22)$$

where c is a positive constant chosen sufficiently large, and λ is the Lagrange multiplier whose appropriate value is to be computed.

The reason behind this suggestion is that if μ^* is a minimum point of the function \mathcal{L}_a respecting the equality constraint $e(\mu^*) = 0$, then, according to [22], μ^* is also a minimum to M and respects the equality constraint $e = 0$. In other words, μ^* is also a solution of problem (21).

To compute the appropriate value of λ starting from an initial estimate $\lambda^{(0)}$, the following update law is used:

$$\lambda^{(k+1)} = \lambda^{(k)} + c^{(k)} e(\mu^{(k)}) \quad (23)$$

where $0 < c^{(k)} \leq c$ and $\mu^{(k)} = \underset{\mu}{\operatorname{argmin}} \mathcal{L}_a(\mu, \lambda^{(k)})$.

Not only is this method useful for minimization problems with equality constraints, but it is also able to deal with certain inequality constraints. To illustrate the incorporation of

inequality constraints in the Augmented Lagrangian method, we consider the following minimization problem:

$$\begin{aligned} & \text{minimize } M(\mu) \\ & \text{s.t. } Z_i(\mu) \leq 0, \quad i=1,2,\dots,m \end{aligned} \quad (24)$$

Under certain assumptions [24, 25], this minimization problem can be written as an equality-constrained minimization problem, similar to problem (21), where

$$e_i(\mu) = Z_i(\mu) + u_i = 0 \quad (25)$$

u_i being a slack variable equal or greater than zero. By doing so, the augmented Lagrangian in case of equality constraints (22) can be reformulated as:

$$\begin{aligned} \mathcal{L}_a(\mu, \lambda) = M(\mu) + \sum_{i=1}^m \left\{ \lambda_i^T [Z_i(\mu) + u_i] \right. \\ \left. + \frac{c}{2} [Z_i(\mu) + u_i]^2 \right\} \end{aligned} \quad (26)$$

The dual function can therefore be written as:

$$\begin{aligned} \Theta(\lambda) &= \min_{u \geq 0, \mu} \mathcal{L}_a(\mu, \lambda) \\ &= \min_{u \geq 0, \mu} \left\{ M(\mu) + \sum_{i=1}^m \left\{ \lambda_i^T [Z_i(\mu) + u_i] + \frac{c}{2} [Z_i(\mu) + u_i]^2 \right\} \right\} \end{aligned} \quad (27)$$

The first minimization of the dual function $\Theta(\lambda)$ is with respect to $u \geq 0$, and then its minimization with respect to μ follows. It is shown in [24, 25] that minimizing $\Theta(\lambda)$ with respect to $u \geq 0$ yields an optimal value of u_i equal to:

$$u_i = \max \left\{ 0, -Z_i(\mu) - \frac{\lambda_i}{c} \right\} \quad (28)$$

Substituting for this u_i in equation (27) allows its reformulation as [24, 25, 26]:

$$\Theta(\lambda) = \min_{\mu} \left\{ M(\mu) + \frac{1}{2c} \sum_{i=1}^m \left\{ \left[\max \{ 0, \lambda_i + cZ_i(\mu) \} \right]^2 - \lambda_i^2 \right\} \right\} \quad (29)$$

The minimization with respect to μ now takes place as per equation (29), yielding a value $\mu^{(k)}$ for each iteration k . Finally, in view of equations (23), (25), and (28), the appropriate value of λ_i can be computed as:

$$\lambda_i^{(k+1)} = \max \left\{ 0, \lambda_i^{(k)} + cZ_i(\mu^{(k+1)}) \right\} \quad (30)$$

This iterative solution of the constrained optimization problem is, in fact, the Uzawa algorithm. The Uzawa algorithm was first introduced in [27] and adapted to the augmented Lagrangian in [28] to become the Augmented Lagrangian

Uzawa Method (ALUM), which is nothing but the Uzawa algorithm applied to the saddle-point problem. ALUM can be seen as the projected gradient-descent method applied to the constrained dual problem, to maximize the dual function (27) for $\lambda \geq 0$. It substitutes the constrained dual problem by a sequence of unconstrained problems, and solves them iteratively in three basic steps:

- Choose an initial nonnegative Lagrange multiplier $\lambda^{(0)}$.
- Calculate $\mu^{(k+1)}$, solution of the unconstrained problem (29).
- Update the Lagrange multiplier $\lambda^{(k+1)}$, as per equation (30).

Indirect Method: Discretize-Then-Optimize Approach

In this approach, discretization of the system precedes the calculation of the necessary optimality conditions. Its objective is to transform the system into a delay-free system to which we can attribute delay-free control. System (8) is an output delay system with a time-varying delay equal to $\tau_{total}(t)$. In what follows, for simplicity, we consider a time-invariant delay, denoted as τ_{total} , which is the average of the time-varying delays $\tau_{total}(t)$. To transform system (8) into a delay-free SS system, we use equation (9) and discretize the transport PDE along the x -direction, using the method of lines [12]. Figure 3 shows the tube section's discretization points, equal to $x = idw$, and their corresponding values, $X_i(t) = X(x = idw, t)$. $X_0(t) = X(0, t) = X_{apc}(t) = \mathcal{X}(t)$ is the BGR at the entrance of the tube section and $X_N(t) = X(L, t) = X_{apc}(t - \tau_{total}) = \mathcal{X}(t - \tau_{total}) = Y(t)$ is the BGR at the exit of the tube section of length $L = 1$.

This discretization scheme results in a vector of discretization points $\mathcal{Z} = [X_1 \ X_2 \ \dots \ X_N]^T \in \mathbb{R}_{\geq 0}^N$, upon which we discretize the transport PDE in equation (9), in time using an explicit forward Euler scheme, and in space using an implicit backward discretization scheme:

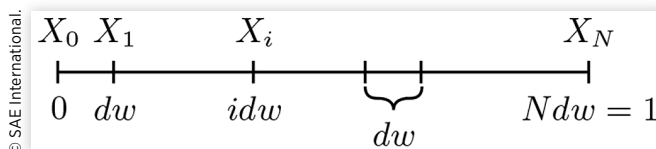
$$\dot{X}_i(t) = -\frac{1}{\tau_{total}} \frac{X_i - X_{i-1}}{dw} \quad (31)$$

$$\text{For } i=1: \dot{X}_1 = -\frac{1}{\tau_{total}dw} (X_1 - X_0) = -\frac{1}{\tau_{total}dw} (X_1 - \mathcal{X}(t))$$

$$\text{For } i=2: \dot{X}_2 = -\frac{1}{\tau_{total}dw} (X_2 - X_1)$$

Discretizing the system forward in time and backward in space results in a discretization scheme which is conditionally stable. This was shown in Von Neumann's stability analysis [29], where the stability condition is such that $0 \leq \sigma = \frac{u\Delta t}{\Delta w} \leq 1$,

FIGURE 3 Discretization of $X(x, t)$ along the x -direction



$$\begin{aligned} \frac{\partial \mathcal{H}(t)}{\partial \lambda} &= \frac{\partial \zeta_c(\lambda, I(\mathcal{XZ}(t)))}{\partial \lambda} \\ &= \frac{1}{c} \left(\max(0_{2 \times 1}, \lambda + cI(\mathcal{XZ}(t))) - \lambda \right) \end{aligned} \quad (40)$$

Uzawa method is used to solve this problem numerically, and the algorithm will execute the following steps:

1. Choose arbitrary values of $\lambda^{(0)} \in \mathbb{R}_{\geq 0}$ and $U(t)^{(0)}$, where the subscript $k \in [0, \infty)$ refers to the number of iteration.
2. Calculate $\mathcal{XZ}(t)$ using (33) and $\mathcal{XZ}(0)$
3. Calculate $p(t)$ using (37) and

$$p(T) = L_2(\mathcal{XZ}_{N+1}(T) - r(T))$$
4. Update $U(t)$

$$U^{(k+1)}(t) = U^{(k)}(t) - v_1 \cdot \frac{\partial \mathcal{H}(t)}{\partial U(t)}$$

where $\frac{\partial \mathcal{H}(t)}{\partial U(t)}$ is obtained from (39) and v_1 is a positive constant.

5. Saturate $U^{(k+1)}(t)$ by Projecting It on the Feasible Set.

$$U^{(k+1)}(t) = \begin{cases} 0 & , U^{(k+1)}(t) < 0 \\ U^{(k+1)}(t) & , 0 \leq U^{(k+1)}(t) \leq 100 \\ 100 & , U^{(k+1)}(t) > 100 \end{cases}$$

6. For each $\lambda^{(k)}$, verify that $\|U(t)^{(k+1)} - U(t)^{(k)}\| \leq \epsilon$, where ϵ is a positive constant. If condition is true, go to step 7. Otherwise, go back to step 2.
7. Update λ

$$\begin{aligned} \lambda^{(k+1)} &= \lambda^{(k)} + v_2 \cdot \frac{\partial \mathcal{H}(t)}{\partial \lambda} \\ &= \lambda^{(k)} + \frac{v_2}{c} \left(\max(0_{2 \times 1}, \lambda^{(k)} + cI(\mathcal{XZ}(t))) - \lambda^{(k)} \right) \end{aligned}$$

where v_2 is a positive constant. This update law reduces to equation (38) when $v_2 = c$.

8. Verify that $\|\lambda^{(k+1)} - \lambda^{(k)}\| \leq \epsilon$. If condition is true, terminate. Otherwise, go back to step 3.

Indirect Method: Optimize-Then-Discretize Approach

In this approach, following equation (9), the objective function (10) can be reformulated as:

$$\begin{aligned} J_3(t) &= \frac{1}{2} \|X(1, T) - r(T)\|_{L_3}^2 \\ &+ \frac{1}{2} \int_0^T (\|X(1, t) - r(t)\|_{Q_3}^2 + \|U(t)\|_{R_3}^2) dt \end{aligned} \quad (41)$$

where $L_3 \in \mathbb{R}_{\geq 0}$, $Q_3 \in \mathbb{R}_{\geq 0}$, and $R_3 \in \mathbb{R}_{> 0}$ are the weights of the cost function J_3 ,

and the inequality constraints can be reformulated as:

$$\begin{aligned} 0 &\leq X(1, t) \leq 100 \\ 0 &\leq U(t) \leq 100 \end{aligned}$$

Similar to the discretize-then-optimize approach, the constraints on $U(t)$ are taken care of using the projected gradient descent method, whereas the constraints on $X(1, t)$ are considered in the state-constraint vector $I(X(1, t)) \in \mathbb{R}_{\leq 0}^2$, expressed as:

$$I(X(1, t)) = \begin{bmatrix} X(1, t) - 100 \\ -X(1, t) \end{bmatrix}$$

To calculate the necessary optimality conditions, the augmented Lagrangian is introduced, and is expressed as:

$$\begin{aligned} \mathcal{L}(\mathcal{X}(t), \dot{\mathcal{X}}(t), U(t), X(x, t), p_1(t), p_2(t)) &= \frac{1}{2} \|X(1, T) - r(T)\|_{L_3}^2 \\ &+ \frac{1}{2} \int_0^T (\|X(1, t) - r(t)\|_{Q_3}^2 + \|U(t)\|_{R_3}^2) dt \\ &+ \int_0^T p_1^T(t) (A(t)\mathcal{X}(t) + B(t)U(t) - \dot{\mathcal{X}}(t)) dt \\ &+ \int_0^T \int_0^1 p_2^T(x, t) \left(\partial_x X(x, t) + \frac{1}{\tau_{total}} \partial_x X(x, t) \right) dx dt \\ &+ \int_0^T \zeta_c(\lambda, I(X(1, t))) dt \end{aligned} \quad (42)$$

where $p_1(t) \in \mathbb{R}$ and $p_2(x, t) \in \mathbb{R}$ are the Lagrange multipliers associated to the ODE and PDE dynamics respectively, $\zeta_c(\lambda, I(X(1, t))) = \frac{1}{2c} (\|\max(0_{2 \times 1}, \lambda + cI(X(1, t)))\|^2 - \|\lambda\|^2)$, $\lambda \in \mathbb{R}_{\geq 0}^2$ is the Lagrange multiplier, and c is a positive scalar.

To ease the calculation of the necessary optimality conditions, we first expand and clarify some terms in the Lagrangian and reformulate its double integral terms using integration by parts.

$$\begin{aligned} &\int_0^T -p_1^T(t) \dot{\mathcal{X}}(t) dt \\ &= -p_1^T(t) \mathcal{X}(t) \Big|_0^T + \int_0^T \mathcal{X}(t) \dot{p}_1^T(t) dt \\ &= -p_1^T(T) \mathcal{X}(T) + p_1^T(0) \mathcal{X}(0) + \int_0^T \frac{dp_1^T(t)}{dt} \mathcal{X}(t) dt \end{aligned}$$

$$\begin{aligned}
& \int_0^T \int_0^1 p_2^T(x,t) \partial_t X(x,t) dx dt \\
&= \int_0^1 \left(p_2^T(x,t) X(x,t) \Big|_0^T - \int_0^T X(x,t) \partial_t p_2^T(x,t) dt \right) dx \\
&= \int_0^1 \left(p_2^T(x,T) X(x,T) - p_2^T(x,0) X(x,0) \right) dx \\
&\quad - \int_0^T \int_0^1 X(x,t) \partial_t p_2^T(x,t) dt dx \\
& \int_0^T \int_0^1 p_2^T(x,t) \frac{1}{\tau_{total}} \partial_x X(x,t) dx dt \\
&= \int_0^T \left(p_2^T(x,t) \frac{1}{\tau_{total}} X(x,t) \Big|_0^1 - \int_0^1 X(x,t) \frac{1}{\tau_{total}} \partial_x p_2^T(x,t) dx \right) dt \\
&= \int_0^T \left(p_2^T(1,t) \frac{1}{\tau_{total}} X(1,t) - p_2^T(0,t) \frac{1}{\tau_{total}} X(0,t) \right) dt \\
&\quad - \int_0^T \int_0^1 \frac{1}{\tau_{total}} X(x,t) \partial_x p_2^T(x,t) dx dt
\end{aligned}$$

Consequently, the Lagrangian can be expressed as:

$$\begin{aligned}
& \mathcal{L}(\mathcal{X}(t), \dot{\mathcal{X}}(t), U(t), X(x,t), p_1(t), p_2(t)) \\
&= \frac{1}{2} \|X(1,T) - r(T)\|_{L_3}^2 + \frac{1}{2} \int_0^T \left(\|X(1,t) - r(t)\|_{Q_3}^2 + \|U(t)\|_{R_3}^2 \right) dt \\
&+ \int_0^T p_1^T(t) (A(t)\mathcal{X}(t) + B(t)U(t)) dt + p_1^T(0)\mathcal{X}(0) \\
&- p_1^T(T)\mathcal{X}(T) + \int_0^T \frac{dp_1^T(t)}{dt} \mathcal{X}(t) dt \\
&+ \int_0^1 \left(p_2^T(x,t) X(x,T) - p_2^T(x,0) X(x,0) \right) dx \\
&- \int_0^T \int_0^1 \left(\frac{1}{\tau_{total}} X(x,t) \partial_x p_2^T(x,t) + X(x,t) \partial_t p_2^T(x,t) \right) dx dt \\
&+ \int_0^T \left(p_2^T(1,t) \frac{1}{\tau_{total}} X(1,t) - p_2^T(0,t) \frac{1}{\tau_{total}} X(0,t) \right) dt \\
&+ \int_0^T \zeta_c(\lambda, I(X(1,t))) dt \tag{43}
\end{aligned}$$

From the augmented Lagrangian, the necessary conditions of optimality can be derived by studying the variation of $L(\cdot)$ with respect to $U(t)$, $\mathcal{X}(t)$, and $X(x,t)$, denoted respectively as $L_U(\cdot)$, $L_{\mathcal{X}}(\cdot)$, and $L_X(\cdot)$.

$$\begin{aligned}
& L_U(\mathcal{X}(t), \dot{\mathcal{X}}(t), U(t), X(x,t), p_1(t), p_2(t)) \\
&= \int_0^T \left(U(t)^T R_3 + B^T(t) p_1(t) \right) \delta U(t) dt
\end{aligned}$$

$$\begin{aligned}
& L_{\mathcal{X}}(\mathcal{X}(t), \dot{\mathcal{X}}(t), U(t), X(x,t), p_1(t), p_2(t)) \\
&= \int_0^T \left(A^T(t) p_1(t) - \frac{1}{\tau_{total}} p_2^T(0,t) + \frac{dp_1^T(t)}{dt} \right) \delta \mathcal{X}(t) dt \\
&\quad - p_1^T(T) \delta \mathcal{X}(T) - p_1^T(0) \delta \mathcal{X}(0)
\end{aligned}$$

$$\begin{aligned}
& L_X(\mathcal{X}(t), \dot{\mathcal{X}}(t), U(t), X(x,t), p_1(t), p_2(t)) \\
&= L_3(X(1,T) - r(T)) \delta X(1,T) \\
&\quad + \int_0^1 \left(p_2^T(x,T) \delta X(x,T) - p_2^T(x,0) \delta X(x,0) \right) dx \\
&\quad + \int_0^T \left(Q_3(X(1,t) - r(t)) \delta X(1,t) + \frac{1}{\tau_{total}} p_2^T(1,t) \delta X(1,t) \right. \\
&\quad \quad \left. + \frac{\partial \zeta_c(\lambda, I(X(1,t)))}{\partial X(1,t)} \delta X(1,t) \right) dt \\
&\quad - \int_0^T \int_0^1 \left(\frac{1}{\tau_{total}} \partial_x p_2^T(x,t) \delta X(x,t) + \partial_t p_2^T(x,t) \delta X(x,t) \right) dx dt
\end{aligned}$$

where

$$\begin{aligned}
& \frac{\partial \zeta_c(\lambda, I(X(1,t)))}{\partial X(1,t)} = \frac{\partial \zeta_c(\lambda, I(X(1,t)))^T}{\partial I(X(1,t))} \cdot \frac{\partial I(X(1,t))}{\partial X(1,t)} \\
&= \max(0_{2 \times 1}, \lambda + cI(X(1,t)))^T [1 \quad -1]^T \\
&= \max(0, \lambda_1 + cX(1,t) - 100c) \\
&\quad - \max(0, \lambda_2 - cX(1,t))
\end{aligned}$$

To obtain the necessary optimality conditions, these variations are set to zero, while admissible (i.e. $\delta \mathcal{X}(0) = \delta X(x,0) = 0$), which gives:

- $\delta U(t) = 0 \Rightarrow U(t)^T R_3 + B^T(t) p_1(t) = 0$
 $\Rightarrow U(t) = -R_3^{-1} B(t)^T p_1(t)$
- $\delta \mathcal{X}(t) = 0 \Rightarrow \frac{dp_1(t)}{dt} + A^T(t) p_1(t) = \frac{1}{\tau_{total}} p_2(0,t)$
- $\delta \mathcal{X}(T) = 0 \Rightarrow -p_1(T) = 0 \Rightarrow p_1(T) = 0$
- $\delta X(x,t) = 0 \Rightarrow \partial_t p_2(x,t) + \frac{1}{\tau_{total}} \partial_x p_2(x,t) = 0$
- $\delta X(x,T) = 0 \Rightarrow p_2(x,T) = 0$

$$\begin{aligned}
\bullet \quad \delta X(1,t) = 0 &\Rightarrow Q_3(X(1,t) - r(t)) + \frac{1}{\tau_{total}} p_2(1,t) \\
&+ \max(0, \lambda_1 + cX(1,t) - 100c) \\
&- \max(0, \lambda_2 - cX(1,t)) = 0 \\
\Rightarrow p_2(1,t) &= \tau_{total} Q_3(X(1,t) - r(t)) \\
&+ \tau_{total} (\max(0, \lambda_1 + cX(1,t) - 100c) \\
&- \max(0, \lambda_2 - cX(1,t)))
\end{aligned}$$

$$\bullet \quad \delta X(1,T) = 0 \Rightarrow X(1,T) - r(T) = 0 \Rightarrow X(1,T) = r(T)$$

For the sake of clarity, we hereby sum up the boundary-value problem resulting from the system dynamics and the necessary optimality conditions.

Equations describing the dynamics of the system:

$$\partial_t X(x,t) + \frac{1}{\tau_{total}} \partial_x X(x,t) = 0 \quad (9) \text{ revisited}$$

$$\begin{aligned}
\text{Initial condition} &: X(x, 0) \\
\text{Boundary conditions} &: X(0, t) \text{ and } X(1, t) \\
\text{Terminal condition} &: X(1, T) = r(T)
\end{aligned}$$

$$\dot{\mathcal{X}} = A(t)\mathcal{X}(t) - B(t)R_3^{-1}B^T(t)p_1(t) \quad (44)$$

$$\text{Initial condition} : \mathcal{X}(0) = \mathcal{X}_0$$

Equations describing the dynamics of the adjoint states:

$$\frac{dp_1(t)}{dt} + A^T(t)p_1(t) = \frac{1}{\tau_{total}} p_2(0,t) \quad (45)$$

$$\text{Terminal condition} : p_1(T) = 0$$

$$\partial_t p_2(x,t) + \frac{1}{\tau_{total}} \partial_x p_2(x,t) = 0 \quad (46)$$

$$\begin{aligned}
\text{Terminal condition} &: p_2(x, T) = 0 \\
\text{Boundary condition} &: p_2(1, t) = \tau_{total} Q_3(X(1, t) - r(t)) \\
&+ \tau_{total} (\max(0, \lambda_1 + cX(1, t) - 100c) - \max(0, \lambda_2 - cX(1, t)))
\end{aligned}$$

To solve this boundary-value problem, we discretize the PDEs in [equations \(9\) and \(46\)](#). Concerning the PDE in [equation \(9\)](#), as in the previous section, it is discretized forward in time and backward in space, upon [equation \(31\)](#), resulting in its discretized version [\(32\)](#). Concerning [equation \(46\)](#), it is discretized, along the x-direction, using the method of lines, in a discretization scheme similar to that shown in [Figure 3](#) (p instead of X), where $p_{2i}(t) = p_2(x = idw, t)$, $p_{20}(t) = p_2(0, t)$ is taken at the entrance of the tube section, and $p_{2N}(t) = p_2(1, t)$ is taken at the exit of the tube section of length $L = 1$. This discretization scheme results in a vector of discretization points $W = [p_{20} \quad p_{21} \quad \dots \quad p_{2N-1}]^T$, upon which we discretize

[equation \(46\)](#) in time using a backward Euler scheme, and in space using a forward discretization scheme:

$$\dot{p}_{2i}(t) = -\frac{1}{\tau_{total}} \frac{p_{2i+1} - p_{2i}}{dw} \quad (47)$$

$$\text{For } i = 0: \dot{p}_{20} = -\frac{1}{\tau_{total} dw} (p_{21} - p_{20})$$

$$\text{For } i = N - 1: \dot{p}_{2N-1} = -\frac{1}{\tau_{total} dw} (p_{2N} - p_{2N-1})$$

This discretization scheme is used in this case because p_{2N} is the known boundary and it is used to calculate all the other values p_{2i} . Consequently, the discretized version of [equation \(46\)](#) can be expressed as:

$$\dot{W}(t) = KW(t) + Dp_2^T(1,t) \quad (48)$$

$$\text{where } K = \begin{bmatrix} \frac{1}{\tau_{total} dw} & -\frac{1}{\tau_{total} dw} & & 0 \\ & \ddots & \ddots & \\ & & \ddots & -\frac{1}{\tau_{total} dw} \\ 0 & & & \frac{1}{\tau_{total} dw} \end{bmatrix} \in \mathbb{R}^{N \times N}$$

$$\text{and } D = \begin{bmatrix} 0 & \dots & 0 & \frac{-1}{\tau_{total} dw} \end{bmatrix}^T \in \mathbb{R}_{\leq 0}^N.$$

Now that the discretization of the PDEs is done, the boundary-value problem can be reformulated. By concatenating [equations \(32\) and \(44\)](#), and concatenating [equations \(45\) and \(48\)](#), the boundary-value problem can be expressed as:

$$\dot{\mathcal{X}}Z(t) = \Lambda(t) \cdot \mathcal{X}Z(t) + \Upsilon(t) \cdot PP(t) \quad (49)$$

$$\dot{PP}(t) = \Gamma(t) \cdot PP(t) + \Omega \cdot p_2(1,t) \quad (50)$$

where $\mathcal{X}Z$ and Λ are given in [equation \(33\)](#),

$$PP(t) = \begin{bmatrix} p_1(t) & W(t)^T \end{bmatrix}^T, \Omega = \begin{bmatrix} 0 & D^T \end{bmatrix}^T,$$

$$\Upsilon(t) = \begin{bmatrix} -B(t)R_3^{-1}B^T(t)^T & 0_{1 \times N} \\ 0_{N \times 1} & 0_{N \times N} \end{bmatrix}, \text{ and}$$

$$\Gamma(t) = \begin{bmatrix} -A^T(t) & \begin{bmatrix} \frac{1}{\tau_{total}} & 0_{1 \times N-1} \end{bmatrix} \\ 0_{N \times 1} & K \end{bmatrix}.$$

By referring to the boundary condition $p_2(1, t)$ corresponding to [equation \(46\)](#), [\(50\)](#) becomes

$$\begin{aligned} \dot{P}P(t) &= \Gamma(t) \cdot PP(t) - \Omega \cdot \tau_{total} \cdot Q_3 \cdot X(1, t) \\ &\quad + \Omega \cdot \tau_{total} \cdot Q_3 \cdot r(t) \\ &\quad - \Omega \cdot \tau_{total} \cdot \max(0, \lambda_1 + cX(1, t) - 100c) \\ &\quad + \Omega \cdot \tau_{total} \cdot \max(0, \lambda_2 - cX(1, t)) \\ &= \Gamma(t) \cdot PP(t) - \Omega \cdot \tau_{total} \cdot Q_3 \cdot E \cdot \mathcal{XZ}(t) \\ &\quad + \Omega \cdot \tau_{total} \cdot Q_3 \cdot r(t) \\ &\quad - \Omega \cdot \tau_{total} \cdot \max(0, \lambda_1 + c \cdot E \cdot \mathcal{XZ}(t) - 100c) \\ &\quad + \Omega \cdot \tau_{total} \cdot \max(0, \lambda_2 - c \cdot E \cdot \mathcal{XZ}(t)) \end{aligned}$$

where $E = [0_{1 \times N} \quad 1]$, and $PP(T) = [p_1(T)W(T)]^T = [0 \quad 0_{N \times 1}]^T$ is the terminal condition.

To solve this problem numerically, UZAWA method, which was used earlier, is implemented. The algorithm executes the following steps.

1. Choose arbitrary values of $\lambda^{(0)} \in \mathbb{R}_+$ and $U(t)^{(0)}$
2. Calculate $\mathcal{XZ}(t)$ using [\(49\)](#) and $\mathcal{XZ}(0)$
3. Calculate $PP(t)$ using [\(50\)](#) and $PP(T)$
4. Update $U(t)$

$$U^{(k+1)}(t) = U^{(k)}(t) - v_1 \cdot \frac{\partial \mathcal{L}(\cdot)}{\partial U(t)}$$

where $\frac{\partial \mathcal{L}(\cdot)}{\partial U(t)}$ is obtained from $L_{U(\cdot)}$ as $U(t)R_3 + B^T(t)p_1(t)$ and v_1 is a positive constant.

5. Saturate $U^{(k+1)}(t)$ by projecting it on the feasible set.

$$U^{(k+1)}(t) = \begin{cases} 0 & , U^{(k+1)}(t) < 0 \\ U^{(k+1)}(t) & , 0 \leq U^{(k+1)}(t) \leq 100 \\ 100 & , U^{(k+1)}(t) > 100 \end{cases}$$

6. For each $\lambda^{(k)}$, verify that $\|U(t)^{(k+1)} - U(t)^{(k)}\| \leq \epsilon$, where ϵ is a positive constant. If condition is true, go to step 7. Otherwise, go back to step 2.
7. Update λ

$$\begin{aligned} \lambda^{(k+1)} &= \lambda^{(k)} + v_2 \cdot \frac{\partial \mathcal{L}(\cdot)}{\partial \lambda} \\ &= \lambda^{(k)} + \frac{v_2}{c} \left(\max(0_{2 \times 1}, \lambda^{(k)} + cI(X(1, t))) - \lambda^{(k)} \right) \end{aligned}$$

where v_2 is a positive constant. This update law reduces to [equation \(30\)](#) when $v_2 = c$

8. Verify that $\|\lambda^{(k+1)} - \lambda^{(k)}\| \leq \epsilon$. If condition is true, terminate. Otherwise, go back to step 3.

Simulation Results

EGR Linear Model Control

The control designed on the EGR linear model results in a closed-loop system whose input and output are shown in [Figure 4](#), for the discretize-then-optimize and for the optimize-then-discretize approaches. Both approaches share the same solver parameters, shown in [Table 1](#). It can be noticed from [Figure 4](#) that starting from the same initial input $U^{(0)} = X_{avc} = 50\%$, both methods calculate the same optimal input $U^* = X_{avc}^*$ and result in the same output X_{col} , which perfectly tracks the reference.

While some authors prefer the discretize-then-optimize approach because the gradient stems directly from the original cost function, others prefer the optimize-then-discretize approach, because it is more precise and less sensible to the choice of the numerical solver, as the discretization comes in after the calculation of the necessary conditions of optimality. The EGR control problem, however, doesn't show a preference for any of both approaches, in terms of merits as well as demerits. From the positive point of view, the discretization in both cases seems to retain the precision conveyed by the infinite-dimensionality of the control problem. From the negative point of view, both approaches are computationally expensive.

FIGURE 4 Discretize-then-optimize and Optimize-then-discretize approaches: closed-loop system input and output

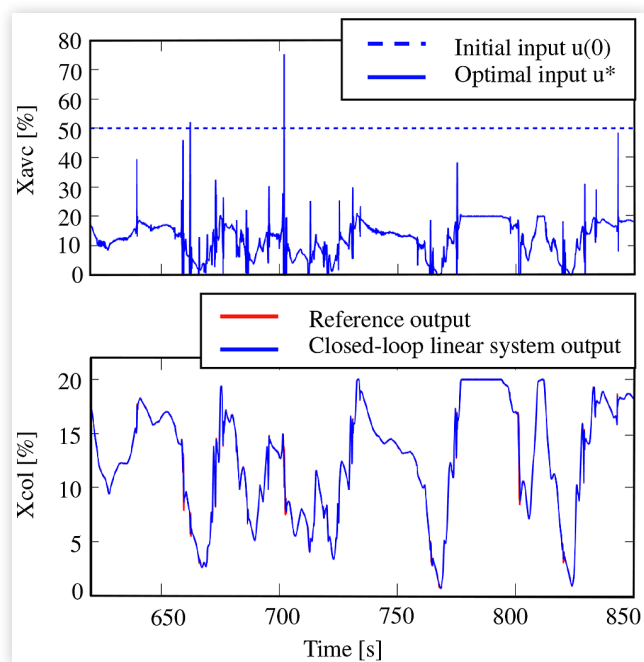


TABLE 1 Parameters of the solver

λ_0	$[0.01 \quad 0.01]^T$	c	0.1
ϵ	0.002	$L_{2,3}$	50
v_1	0.01	$Q_{2,3}$	50
v_2	0.01	$R_{2,3}$	0.01

Their iterative nature prolongs the time needed to calculate an optimal solution, which makes them not practical for an implementation on an ECU, unless subject to solutions, such as:

- increasing the value of the error tolerance ϵ to decrease the number of iterations of the UZAWA method;
- decreasing the number of discretization points in the method of lines resulting in a discrete system of smaller dimension;
- using more efficient faster numerical solvers, such as those proposed in [31] for large-scale optimization problems;
- including the constraints on the input U in the constraint vector I and using non-iterative methods, such as [32] or [33], to solve the boundary value problem.

AMESim® Nonlinear Model Control

The control input in the AMESim® nonlinear model is the EGR duty cycle (%). Therefore, verifying the optimal control law on the nonlinear AMESim® model requires converting the optimal control input X_{avc}^* to EGR duty cycle (%). This conversion is a 4-step process:

- Convert X_{avc}^* to Q_{EGR} using the ODE:

$$\dot{X}_{avc} = \alpha_{avc} \left[-(Q_{EGR} + Q_{air}) X_{avc} + Q_{EGR} X_{em} \right] \quad (51)$$

where $\alpha_{avc} = \frac{RT_{avc}}{P_{avc}V_{avc}}$, T_{avc} (K), P_{avc} (Pa), and V_{avc} (m^3) are

the respective temperature, pressure, and volume upstream the compressor, and X_{em} is the BGR of the gas arriving from the exhaust manifold. This ODE formulation stems from the fact that the volume upstream the compressor is considered as a control volume.

- Convert Q_{EGR} to S_{EGR} using the Saint-Venant equation, which calculates the mass flow rate as a function of the pressure drop across the EGR valve, as follows:

$$Q_{EGR} = \frac{S_{EGR} C_q C_m P_{up}}{\sqrt{T_{up}}} \quad (52)$$

where S_{EGR} (m^2) is the cross-sectional surface area of the EGR opening, C_q is the flow coefficient equal to 1, P_{up} (Pa) and T_{up} (K) are the respective pressure and temperature upstream the EGR valve, and C_m is the mass flow parameter expressed, by assuming a subsonic flow [34], as:

$$C_m = \sqrt{\frac{2\gamma_g}{R(\gamma_g - 1)} \left(\left(\frac{P_{down}}{P_{up}} \right)^{\frac{2}{\gamma_g}} - \left(\frac{P_{down}}{P_{up}} \right)^{\frac{\gamma_g + 1}{\gamma_g}} \right)}$$

where γ_g is the specific heat ratio of the gas, and P_{down} (Pa) is the pressure downstream the EGR valve.

- Saturate S_{EGR} between $S_{min} = 0$ and $S_{max} = 250.025 \text{ mm}^2$ to make sure that the physical limits of the actuator are respected, and smooth it using a median filter to avoid the peaks and outliers.
- Convert S_{EGR} (mm^2) to EGR duty cycle (%) using Table 2.

Processing the optimal input X_{avc}^* in this way to obtain the optimal EGR duty cycle, results in an eventually smoother input X_{avc} , shown in Figure 5 along with the system output, which well tracks the reference output. Note that the output falling between $t = 100$ s and $t = 150$ s corresponds to a stop phase in which the engine is not running. Therefore, tracking the reference output in this phase is senseless.

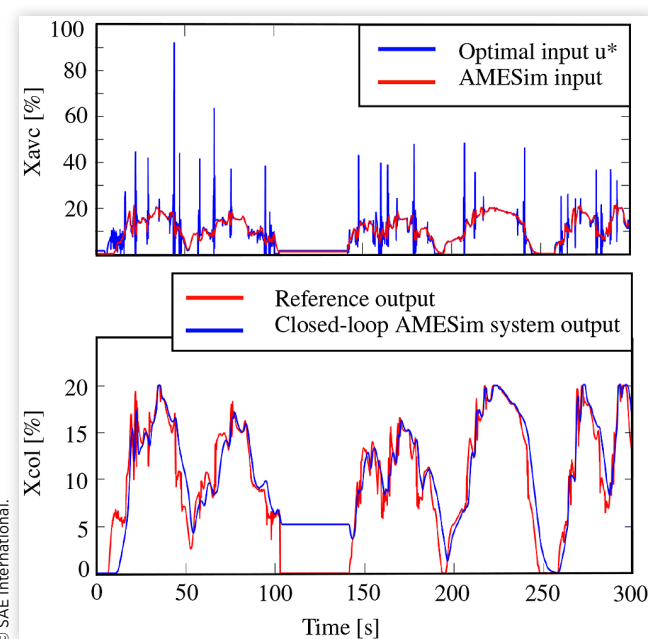
The nonlinear AMESim® model was initially controlled using a PI controller, with proportional and integral gains equal to 1. The input of the controller is the error between the current X_{col} and its reference value, and the output of the controller is the EGR duty cycle, whose value is saturated between 0 and 100 to avoid exceeding the physical limits of the valve. Figure 6 helps comparing the optimal controllers

TABLE 2 Look-up table matching S_{EGR} and EGR duty cycle

S_{EGR} (mm^2)	Duty cycle (%)
0	0
25	10
125	20
210	30
250	40
250.01	50
250.015	60
250.02	70
250.025	100

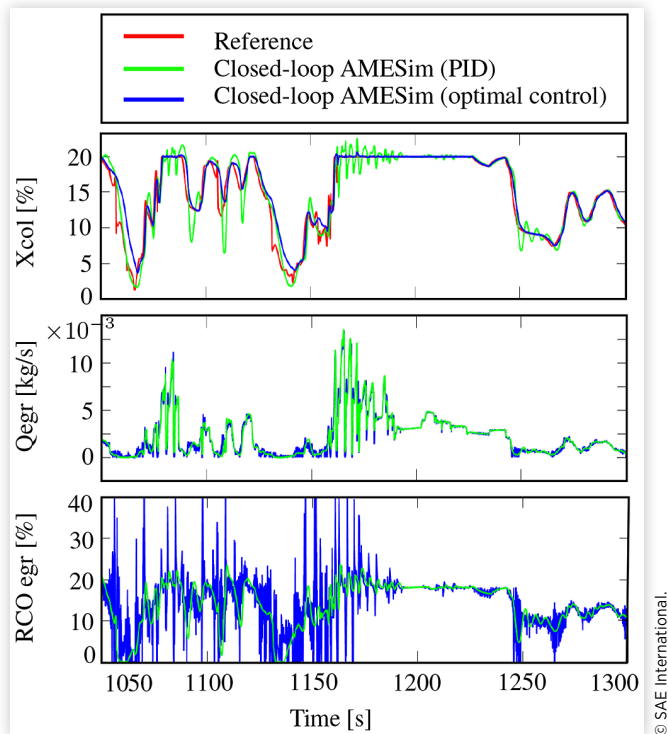
© SAE International.

FIGURE 5 AMESim® model input and output



© SAE International.

FIGURE 6 AMESim® model output and EGR mass flow rate and duty cycle



proposed in this paper with the PI controller, by showing the output X_{col} , the mass flow through the EGR Q_{EGR} , and the EGR duty cycle in both cases. The output plot shows that, while both controllers are able to track the reference output, the optimal controllers feature less overshoot and transient oscillations than the PI controller.

Conclusion and Future Perspectives

Mass transport phenomena, taking place in the pipes of the engine air-path, such as the EGR system, can be modeled as first-order hyperbolic PDEs of infinite-dimensional nature. This paper describes the BGR mass transport in the intake manifold in the form of a control-oriented time-delay model. Two indirect optimal control approaches, taking into account the infinite-dimensional nature of the model, are compared: discretize-then-optimize and optimize-then-discretize. To account for the system's input and state constraints, the controllers are equipped with constraint management techniques such as the Augmented Lagrangian Uzawa method, and their merits are demonstrated on the original AMESim® model. To deal with the computationally expensive optimization process, a future scope of this research work lies in exploring more efficient simulation and numerical tools. Also, replacing the time-invariant delay by a time-variant delay would enhance the control performance by making it more realistic.

References

1. Buenaventura, F.C., "Modélisation et Contrôle de la Boucle d'Air des Moteurs Diesel pour Euro 7," Ph.D. thesis, Université de Grenoble, 2013.
2. Witrant, E. and Niculescu, S.-I., "Modeling and Control of Large Convective Flows with Time-Delays," *Mathematics in Engineering, Science and Aerospace* 1(2):191-205, 2010.
3. Mcadams, W.H., "Method of Controlling Recycling of Exhaust Gas in Internal Combustion Engines," U.S. Patent 1,916,325, July 4 1933.
4. Van Nieuwstadt, M., Moraal, P., Kolmanovskiy, I., Stefanopoulou, A. et al., "Decentralized and Multivariable Designs for Egr-Vgt Control of a Diesel Engine," *IFAC Proceedings Volumes* 31(1):189-194, 1998.
5. Ammann, M., Fekete, N.P., Guzzella, L., and Glattfelder, A.H., "Model-Based Control of the Vgt and Egr in a Turbocharged Common-Rail Diesel Engine: Theory and Passenger Car Implementation," *SAE Transactions* 527-538, 2003.
6. Stefanopoulou, A.G., Kolmanovskiy, I., and Freudenberg, J.S., "Control of Variable Geometry Turbocharged Diesel Engines for Reduced Emissions," *IEEE Transactions on Control Systems Technology* 8(4):733-745, 2000.
7. Ortner, P. and Del Re, L., "Predictive Control of a Diesel Engine Air Path," *IEEE Transactions on Control Systems Technology* 15(3):449-456, 2007.
8. Wiese, A.P., Stefanopoulou, A.G., Karnik, A.Y., and Buckland, J.H., "Model Predictive Control for Low Pressure Exhaust Gas Recirculation with Scavenging," in *2017 American Control Conference (ACC)*, 2017, 3638-3643, IEEE.
9. Bresch-Pietri, D., Leroy, T., Chauvin, J., and Petit, N., "Prediction-Based Trajectory Tracking of External Gas Recirculation for Turbocharged Si Engines," in *2012 American Control Conference (ACC)*, 2012, 5718-5724, IEEE.
10. Xiaolong, Y., Ming, H., Jingping, L., and Biao, L., "A Robust Egr Control System for Gasoline Engine Using Pid," in *2010 International Conference on Optoelectronics and Image Processing*, 2010, Vol. 1, 116-119, IEEE.
11. Plianos, A. and Stobart, R., "Modeling and Control of Diesel Engines Equipped with a Two-Stage Turbo-System," *SAE Technical Paper 2008-01-1018*, 2008, <https://doi.org/10.4271/2008-01-1018>.
12. Ross, D.W., "Controller Design for Time Lag Systems via a Quadratic Criterion," *IEEE Transactions on Automatic Control* 16(6):664-672, 1971.
13. Hinze, M., Pinnau, R., Ulbrich, M., and Ulbrich, S., *Optimization with PDE Constraints*. Vol. 23 (Springer Science & Business Media, 2008).
14. Chauvin, J., Corde, G., and Petit, N., "Constrained Motion Planning for the Airpath of a Diesel Hcci Engine," in *Proceedings of the 45th IEEE Conference on Decision and Control*, 2006, 3589-3596, IEEE.
15. Pontryagin, L.S., Boltyanski, V.G., Gamkrelidze, R.V., and Mishchenko, E.F., *The Mathematical Theory of Optimal Processes* (New York: Interscience Publishers, 1962).
16. Kirk, D.E., *Optimal Control Theory: An Introduction* (Mineola, New York: Dover Publications, 2004).

17. Sage, A.P. and White, C.C., *Optimum Systems Control* Second Edition (Prentice Hall: Englewood Cliffs, NJ, 1977).
18. Lions, J.L., *Optimal Control of Systems Governed by Partial Differential Equations (Grundlehren der Mathematischen Wissenschaften)* First Edition (Berlin, Heidelberg: Springer-Verlag, 1971), 170.
19. Ahmed, N.U. and Teo, K.L., *Optimal Control of Distributed Parameter Systems* (Elsevier Science Inc., 1981).
20. Tröltzsch, F., "Optimal Control of Partial Differential Equations: Theory, Methods, and Applications," *American Mathematical Soc.* 112, 2010.
21. Li, X. and Yong, J., *Optimal Control Theory for Infinite Dimensional Systems* (Birkhäuser Basel, 1995).
22. Hestenes, M.R., "Multiplier and gradient methods," *Journal of Optimization Theory and Applications* 4(5):303-320, 1969.
23. Powell, M.J., "A Method for Nonlinear Constraints in Minimization Problems," *Optimization* 283-298, 1969.
24. Luenberger, D.G. and Ye, Y., *Linear and Nonlinear Programming* Fourth Edition (Springer, 2008).
25. Bertsekas, D.P., "On Penalty and Multiplier Methods for Constrained Minimization," *SIAM Journal on Control and Optimization* 14(2):216-235, 1976.
26. Iusem, A.N., "Augmented Lagrangian Methods and Proximal Point Methods for Convex Optimization," *Investigación Operativa* 8(11-49):7, 1999.
27. Uzawa, H., "Iterative Methods for Concave Programming," *Studies in Linear and Nonlinear Programming* 6:154-165, 1958.
28. Fortin, M. and Glowinski, R., *Augmented Lagrangian Methods: Applications to the Numerical Solution of Boundary-Value Problems* (Amsterdam: North-Holland, 1983).
29. Charney, J.G., Fjørtoft, R., and Neumann, J.V., "Numerical Integration of the Barotropic Vorticity Equation," *Tellus* 2(4):237-254, 1950.
30. Courant, R., Friedrichs, K., and Lewy, H., "On the Partial Difference Equations of Mathematical Physics," *Mathematische Annalen* 100(1):32-74, 1928.
31. Hager, W.W., Hearn, D.W., and Pardalos, P.M., *Large Scale Optimization: State of the Art* (Springer US, 1994).
32. Hyman, M.A., "Non-iterative Numerical Solution of Boundary-Value Problems," *Applied Scientific Research, Section B* 2(1):325-351, 1952.
33. Fazio, R., "A Noniterative Transformation Method Applied to Two-Point Boundary-Value Problems," *Applied Mathematics and Computation* 39(1):79-87, 1990.
34. Heywood, J.B. et al., *Internal Combustion Engine Fundamentals* (New York: McGraw-Hill, 1988).

Contact Information

Sandra HAMZE

sandra.hamze@renault.com

sandrahamzeh@hotmail.com

Acronyms

BGR - Burned Gas Ratio

ECU - Engine Control Unit

EGR - Exhaust Gas Return/Recirculation

ICE - Internal Combustion Engine

LP-EGR - Low-Pressure Exhaust Gas Return/Recirculation

LPV - Linear Parameter-Varying

LQG - Linear Quadratic Gaussian

LTV - Linear Time-Varying

MPC - Model Predictive Control

MVEM - Mean-Value Engine Model

ODE - Ordinary Differential Equation

PDE - Partial Differential Equation

PI - Proportional Integral

PID - Proportional Integral Derivative

SS - State-Space