Modeling and Control of Automotive Powertrain Systems: A Tutorial

Jing Sun, Ilya Kolmanovsky, Jeffrey A. Cook, and Julia H. Buckland

Abstract—This tutorial presents an overview of key issues in electronic control of internal combustion engines for automotive passenger vehicles, and showcases the control oriented engine and aftertreatment system models that are useful in addressing these issues. Beginning with a discussion on electronic engine control systems and standard sensors and actuators, key engine control subsystems and their associated functionalities will be outlined. Models for gasoline and diesel engines, lean aftertreatment systems, and turbochargers will be described in more detail. Several representative control problems will be elaborated through examples.

I. INTRODUCTION

Since its inception in the late 1970s, electronic engine control has played a critical role in reducing emissions, improving fuel economy and enhancing performance and driveability of passenger vehicles. The introduction of new hardware innovations and advanced sensors and actuators, coupled with increasingly stringent government regulations and higher customer expectations, presents serious challenges as well as unprecedented opportunities for powertrain control. To satisfy demanding and often competing requirements, powertrain control engineers are relying to a greater extent on advanced control methodologies, effective design tools and streamlined processes. The pressure for reducing product cost and development time has also driven control system development processes towards those which favor the systematic application of model-based design and analysis tools and methodologies.

This tutorial is developed to provide an overview of fundamental problems in powertrain control, as well as associated design tools. The goal is to expose powertrain control issues to the broader control community and to stimulate interest and activity to address pressing issues and to develop effective tools. Fundamental powertrain models will be presented, representative control issues will be illustrated, and existing design tools and methodologies will be highlighted. A recent survey paper [1] provides a more comprehensive summary of powertrain control problems and results, including those for alternative powertrain systems such as the fuel cell and the hybrid electric powertrain.

This tutorial paper is organized as follows. After a brief overview of engine control features in Section II, Sections III, IV, V, and VI will cover modeling and control for port fuel injection spark ignition, gasoline direct injection stratified charge, diesel, and turbocharged automotive engines, respectively. In each section, fundamental system models will be presented, followed by a discussion of the main control issues. Examples of control problems and their associated solutions will be provided in each section to demonstrate the multi-faceted nature of automotive engine control. The paper will conclude with a short discussion on emerging powertrain trends and their implications on control system development.

II. ELECTRONIC ENGINE CONTROL OVERVIEW

Modern electronic engine control involves many functions and leverages multiple sensors and actuators. The functions of an ECU (engine control unit) range from delivering the driver demanded torque to monitoring the emission control systems on board the vehicle, and from warming up the catalyst to purging the vapor from the fuel tank. Due to the introduction of new sensors and actuators, along with more aggressive regulation, the complexity of engine control functions has dramatically increased.

Figure 1 illustrates the key elements of a typical system, including sensors and actuators. Typical sensors include crankshaft and camshaft position sensors; intake mass air flow; manifold pressure; exhaust gas oxygen sensors, both pre- and post catalyst; coolant and cylinder head temperatures. Conventional actuators include throttle control, exhaust gas recirculation, ignition and fuel injection control. Advanced technology engines may incorporate valve timing control (intake valves, exhaust valves or both), turbocharger wastegate control, cylinder deactivation control, fuel pressure control, swirl control, intake manifold runner control and fuel pressure control. Advanced sensing technologies
include in-cylinder pressure or ionization measurements to optimize combustion.

III. PORT FUEL INJECTION ENGINE CONTROL

For conventional port fuel injection (PFI) engines, three fundamental control tasks affect emissions, performance, and fuel economy: (1) air-fuel ratio (A/F) control, that is, providing the correct ratio of air and fuel for efficient combustion and for optimal aftertreatment conditioning; (2) ignition control, which refers to firing the appropriate spark plug at the precise instant required; and (3) control of exhaust gas recirculation (EGR) to reduce the formation of the oxides of nitrogen (NOx). Other functions, such as the idle speed control, knock control and on-board diagnosis (OBD), are also essential. These control functions have been greatly augmented, in terms of both performance and complexity, with the introduction of new actuators, such as electronic throttle, variable cam timing, variable displacement, etc. In this section, we will first describe a control oriented engine model which has proven very useful for control design and performance analysis, followed by a discussion of some control challenges associated with PFI engines. Due to the space limitations, we choose to elaborate on two representative control features in this section: A/F control and torque control for engines with variable cam timing.

A. Control Oriented Gasoline Engine Models

Control oriented engine models refer to the linear and nonlinear low frequency phenomenological representations that capture the essential system dynamics required for control development, along with key static behavior such as emissions and volumetric efficiency that may be obtained experimentally from steady state mapping on an engine dynamometer. Nonlinear engine models were developed from the late 1970s through the 1980s, as reported in many papers such as [2], [3]. The four-stroke engine cycle naturally divides the physical process into four events comprising intake, compression, power generation and exhaust. Models exploiting the inherently discrete nature of the system by crank angle based sampling are described, for example, in [4], [5], [6].

The mathematical representation of the conventional, naturally aspirated engine includes the following elements: (1) the throttle body, (2) the intake manifold, (3) torque generation and (4) engine rotational dynamics. The model may also include the EGR system, exhaust gas temperature and pressure dynamics, and feedgas emissions. Figure 2 shows the block diagram of a phenomenological engine model.

The intake manifold dynamics are derived from the ideal gas law:

\[ \dot{P}_i = K_i (W_a + W_{egr} - W_{cyl}) \]

where \( K_i \) depends on the intake manifold volume and temperature; \( W_a, W_{egr} \) are the mass flow rates through the throttle body and the EGR valve, respectively; and \( W_{cyl} \) is the mean value of the rate at which the charge is inducted into the cylinders. The flows through the throttle body and EGR valve are represented by a standard orifice equation:

\[ W_a = A_{th} P_t \frac{\phi \left( \frac{P_t}{P_a} \right)}{\sqrt{T_a}} \]

\[ W_{egr} = A_{egr} P_e \frac{\phi \left( \frac{P_e}{P_e} \right)}{\sqrt{T_e}} \]

where \( A_{th}, A_{egr} \) are the effective flow areas for the throttle body and EGR valve respectively; \( P_t, P_e, P_a, P_e \) are intake manifold, exhaust manifold and ambient pressures; and \( T_a, T_e \) are the ambient and exhaust temperatures. The function \( \phi \) represents the effects of the pressure ratio on the flow across the valve:

\[ \phi(x) = \begin{cases} \gamma \frac{g}{2} \left( \frac{2}{\gamma+1} \right)^{\frac{\gamma+1}{\gamma-1}} & \text{if } x \leq \left( \frac{2}{\gamma+1} \right)^{\frac{\gamma}{\gamma-1}} \\ x^{\frac{1}{2}} \left( \frac{\gamma}{\gamma-1} \right) \left[ 1 - x^{\frac{\gamma-1}{\gamma}} \right] & \text{if } x > \left( \frac{2}{\gamma+1} \right)^{\frac{\gamma}{\gamma-1}} \end{cases} \]

where \( \gamma \) is the ratio of specific heats, which takes different values for \( W_a \) and \( W_{egr} \).

The amount of charge inducted into the cylinders, \( W_{cyl} \), is a function of engine speed, intake manifold pressure and, possibly, temperature, where intake manifold temperature depends on mass air flow and EGR. \( W_{cyl} \) is generally represented as a static regression equation based on steady-state mapping data for a particular engine. One commonly used representation takes the form of

\[ W_{cyl} = (\alpha(N)P_t + \beta(N))f(T_i) \]

which is affine in \( P_t \), where \( \alpha, \beta \) are low order polynomials of \( N \), and \( f(T_i) \) is used to compensate for the air density variation due to temperature variation in the intake manifold.

Engine rotational dynamics follow the equation:

\[ \frac{\pi}{30} J_i \dot{T} = T_b - T_f \]

where \( T_b, T_f \) are the engine brake and load torques in \( Nm \), respectively, and the factor \( \pi/30 \) is due to the unit conversion of engine speed (from \( rpm \) to \( rad/sec \)). The engine brake torque, \( T_b \), is the net torque available on the crankshaft to drive the rest of the powertrain, and can be decomposed into:

\[ T_b = T_i - T_f \]

where \( T_i \) is the indicated torque, a measure of the total torque delivered to the piston by burning the fuel and \( T_f \) is the total friction which the engine has to overcome when delivering the torque to the crankshaft. The friction torque includes the pumping losses during the intake and exhaust strokes plus mechanical friction and may be regressed as a function of speed and intake manifold pressure. Brake torque is generally represented as a regressed function of \( W_{cyl}, A/F, N, \) and ignition timing, based on engine mapping data.
The engine model described by equations (1)-(6) forms the basis of control oriented engine models. Many variations and modifications have been developed to account for various hardware configurations and operating conditions and to serve specific design purposes. These models have been used successfully for design and analysis of many control subsystems, including those to be discussed in the following subsections.

B. A/F control for PFI engines

Figure 3 underscores the criticality of A/F control for PFI engines. It illustrates that high simultaneous conversion efficiencies for the three regulated species \((HC, CO, NO_x)\) occur only in a narrow band around stoichiometry for a three way catalyst (TWC), which is a standard emission control device for PFI engines. A/F control of the conventional PFI engine encompasses three main aspects: accurate estimation of air charge, compensation for fuel puddling dynamics in the intake manifold runners and on the intake valves, and closed-loop regulation of A/F for high catalyst performance. Each of these three problems is briefly discussed as follows.

1) Air charge estimation: This refers to the task of estimating the amount of fresh air that is inducted into the cylinder for each combustion cycle. This charge estimation forms the basis for fuel scheduling and A/F control. The model for \(W_{cyl}\) given by equation (4) is often used as the baseline for charge estimation, when intake manifold pressure is measured. The manifold dynamics and the influence of EGR also have to be incorporated. However, it is often insufficient to rely on this model for accurate charge estimation and precise A/F control, as volumetric efficiency and engine breathing characteristics can change with operating conditions and component aging. The robustness of the charge estimation algorithm is significantly enhanced with the incorporation of the mass air flow meter, typically a hot wire anemometer used to measure the inlet air flow rate. This air meter, while introducing direct feedback for charge estimation and significantly simplifying the estimation algorithm, imposes other issues due to its relatively slow dynamics.

A low frequency model of the induction process is described in [7]. This model has been used to develop a compensation scheme for the relatively slow dynamics of the air meter. In [8], a dynamic model incorporating intake runner acoustic and inertial effects is developed that is capable of describing the induction process in individual cylinders. Other performance enhancing algorithms developed for diesel [9] or gasoline direct injection engines [10] are also applicable to PFI engines with minor modifications.

2) Transient fuel compensation: Transient fuel characteristics for a PFI engine were first reported by Fozo and Aquino in [11]. Depending on the operating condition, part of the fuel injected at the port may not be inducted into the cylinder to participate in the immediate combustion event. Instead, a portion may stay on the wall of the intake runner and on the intake valves as fuel puddles, to be subsequently inducted into the cylinder for later combustion events. This so-called wall-wetting phenomenon acts as a disturbance to the A/F control system. An empirical model has been developed [11] to capture the fuel puddling dynamics, and this model has been exploited for transient fuel compensation by many authors.

In [12], a method of adaptive transient compensation for wall-wetting dynamics is described that accounts for varying fuel properties. The technique requires only a heated exhaust gas oxygen (HEGO) sensor\(^1\), which remains

\(^1\) A HEGO is a switching sensor that indicates if the mixture is lean or rich of stoichiometry, but not by how much.
the prevalent feedback sensor for closed-loop A/F control. The approach described in [12] uses the feedback signal to evaluate changes in A/F during driver induced transients in closed loop, and stores corrections to the compensation algorithm indexed by engine temperature for use in the next transient or during open-loop cold start operation.

3) Closed-loop A/F control: The precipitous falloff of TWC efficiency away from stoichiometry (see Figure 3) necessitates tight regulation of A/F around the stoichiometric value. Closed-loop A/F control, now a standard feature for automotive engines, was made possible by the invention of the HEGO sensor. Due to the switching nature of the sensor, most of the standard control design methodologies based on linear system theory cannot be applied directly. Consequently, many of the HEGO based A/F feedback control strategies are designed based on heuristic rules and physical insights.

Another issue that complicates the HEGO based A/F control system is the shift of the sensor switch point. In [13], it was shown that cylinder-to-cylinder A/F differences result in a closed-loop lean shift in controlled A/F due to preferential diffusion of H2 and CO across the HEGO sensor upstream of the catalyst. Typically, this effect is mitigated by biasing the A/F setpoint slightly rich, at a cost in fuel economy and conversion efficiency of the other exhaust constituents. An alternative solution is to address the root cause by eliminating or minimizing the cylinder-to-cylinder A/F variations. In [14], an approach to achieving uniform cylinder-to-cylinder A/F control for a 4-cylinder engine in the presence of injector mismatch and unbalanced air flow due to engine geometry is presented. The method recognizes that the individual cylinder representation of the fueling process describes a periodically time varying system due to the unequal distribution of A/F from cylinder to cylinder. The key features of the controller are the construction of a time-invariant representation of the process and event-based sampling and feedback. In [15], the method was extended to an 8-cylinder engine in which exhaust manifold mixing dynamics were significant.

A significant advancement in A/F feedback control capability is the introduction of the Universal Exhaust Gas Oxygen (UEGO) sensor in production vehicles. Unlike the conventional HEGO sensor, which simply switches about stoichiometry, the UEGO is a linear device that permits an actual measurement of A/F. The use of UEGO allows application of many advanced control design methodologies. Unique challenges associated with A/F control, such as the inherent transport delays and robustness over the lifetime of the vehicle, have motivated many research efforts and led to fruitful results (see [16] and the references therein).

Another issue that is closely related to A/F control is catalyst conditioning and monitoring. Control and diagnosis of catalysts using UEGO sensors are described by [17], [18]. In [19], Fiengo and co-authors use the catalyst model described in [20], along with pre- and post-catalyst UEGO sensors, to develop a controller with two objectives: to simultaneously maximize the conversion efficiencies of HC, CO and NOx, and to obtain steady-state A/F control that is robust with respect to disturbances. A series controller topology is adopted as illustrated in Figure 4. The objective of the first block, the Aft Controller, is to respond relatively quickly to A/F disturbances on the basis of measured feedgas oxygen level. The objective of the second block, the Fore Controller, is to adjust the setpoint of the fore controller, on the basis of both A/F measurements, so that the TWC achieves simultaneously high conversion efficiencies for HC and NOx. The aft controller is composed of a bias estimator and a proportional term. The bias estimator uses the upstream and downstream A/F measurements to correct the error in the upstream oxygen sensor. The proportional controller feeds back the post-catalyst UEGO sensor measurement and establishes the reference for the fore controller.

C. Torque Control for PFI engines with Variable Cam Timing

Variable cam timing provides improved performance and reduced feedgas emissions using an electro-hydraulic mechanism to rotate the camshaft relative to the crankshaft and retard cam timing with respect to the intake and exhaust strokes of the engine. In this manner, the amount of residual gas trapped in the cylinder at the end of the exhaust stroke is controlled, suppressing NOx formation [21]. In addition, VCT allows the engine designer to optimize cam timing over a wide range of engine operating conditions, providing both good idle quality (minimal overlap between the intake and exhaust events) and improved wide-open throttle performance (maximum inducted charge). Properly controlled, the variable cam can be used to operate the engine at higher intake manifold pressures, reducing pumping losses at part throttle conditions to provide a fuel economy improvement. Uncompensated, however, VCT acts as a disturbance to the breathing process, compromising drivability and substantially reducing its effectiveness in emission control.

Four versions of VCT are available: phasing only the intake cam (intake only), phasing only the exhaust cam (exhaust only), phasing the intake and exhaust cams equally (dual equal), and phasing the two camshafts independently (dual independent). A low order nonlinear model of a dual-equal VCT engine is derived in [22]. The model forms the
basis for active compensation of VCT induced cylinder air charge variation employing electronic throttle control (ETC) [23]. The balance of this section will review the coordinated ETC/VCT control for torque compensation.

The basic equations of the PFI engine model can be modified to incorporate the effects of the cam actuator on engine breathing. For the VCT engine, the mass air flow rate into the cylinders is represented as a function of cam phasing, \( \zeta_{\text{cam}} \), in addition to manifold pressure, \( P_i \), and engine speed, \( N \):

\[
W_{\text{cyl}} = (\alpha_1(N, \zeta_{\text{cam}})P_i + \alpha_2(N, \zeta_{\text{cam}}))f(T_i),
\]

where \( \alpha_1 \) and \( \alpha_2 \) are low-order polynomials in \( N \) and \( \zeta_{\text{cam}} \).

Typically, the cam timing reference, \( \zeta_{\text{ref}} \), is scheduled based on engine speed and driver demanded throttle position, \( \theta_0 \). The cam schedule reaches maximal cam retards at part throttle to provide maximal internal EGR; close to idle and at wide open throttle, the cam phasing is at zero or slightly advanced. Scheduling cam with throttle causes it to change when the pedal is depressed or released. It is this torque variation caused by the cam transient that results in undesirable engine response and driveability problems.

To minimize this torque variation, the throttle angle is comprised of the throttle position due to the driver’s request (\( \theta_0 \)) and an additive term due to the compensation (\( \theta^* \)),

\[
\theta = \theta_0 + \theta^*.
\]

The throttle flow equation is represented as functions of pressure and flow area geometry, as in the conventional engine model given by (2)-(3).

A feedforward compensator is designed to recover the drivability of the conventional engine by eliminating the effect of the cam transients on cylinder mass air flow. The algorithm employs \( \theta^* \) as a virtual actuator, according to [23]. That is, a control law is developed for \( \theta^* \) such that the rate of change of \( W_{\text{cyl}} \) coincides with that of the conventional engine. Specifically, compensation \( \theta^* \) is evaluated:

\[
\theta^* = g^{-1}\left(\frac{\partial \alpha_1}{\partial \zeta_{\text{cam}}} \dot{P_i} + \frac{\partial \alpha_2}{\partial \zeta_{\text{cam}}}, \zeta_{\text{cam}} + \frac{\phi(\dot{P_i})}{\phi(\dot{P_i})}g(\theta_0)\right) - \theta_0,
\]

where \( \dot{P_i} \) is a fictitious reference manifold pressure which should be equal to the manifold pressure of the conventional engine driven with the throttle angle, \( \theta_0 \), and engine speed, \( N \). This reference manifold pressure is generated by

\[
\dot{P_i} = K_i \left( \phi(\dot{P_i})g(\theta_0) - \alpha_1(N, 0) \dot{P_i} - \alpha_2(N, 0) \right).
\]

Figure 5 shows the reduction of the torque fluctuation during cam transients achieved by the compensation.

IV. GASOLINE DIRECT INJECTION ENGINE CONTROL

A direct injection stratified charge (DISC) engine, like a diesel, injects fuel directly into the combustion chamber. It is different from a conventional PFI engine discussed in Section III in several aspects. Most importantly, the DISC engine can, depending on speed and load, operate in one of three combustion modes: homogeneous stoichiometric (\( A/F \approx 14.64 \)), homogeneous lean (\( A/F \) between stoichiometry and about 20) or stratified (\( A/F \geq 20 \)). A homogeneous \( A/F \) mixture is achieved by injecting fuel early in the intake stroke, while stratification is achieved by injecting late, during the compression stroke [24]. The torque and emission characteristics corresponding to homogeneous and stratified operation are so distinct that different control strategies are required to optimize performance in the two regimes [25]. Note also that, in addition to the usual control variables such as throttle position, ignition timing, exhaust gas recirculation (EGR) and fueling rate, the DISC engine requires new inputs including injection timing, fuel rail pressure and swirl control at a minimum [26]. Finally, the ultra-lean \( A/F \) operation of the direct injection engine mandates the use of a lean \( NO_x \) trap (LNT) to manage emission of the oxides of nitrogen. The LNT, as a \( NO_x \) storage device, needs to be purged periodically to regenerate its storage capacity.

These special features of DISC engine operation have important control implications and lead to several unique control problems, such as mode transition between stratified and homogeneous operations, aftertreatment system management, etc. In this section, we will concentrate on the aftertreatment system for the DISC powertrain. We will first present a lean aftertreatment model, then delineate the pertinent control issues.

A. A Phenomenological LNT Model

The typical aftertreatment system for a lean-burn engine with a commonly used sensor configuration is shown in Figure 6. It consists of a conventional three-way catalytic converter (usually closely coupled to the engine for opti-
We define the instantaneous efficiency, \( \eta_s \), as:

\[
\eta_s = \frac{\dot{m}_{NOx,in} - \dot{m}_{NOx,tp}}{\dot{m}_{NOx,in}}
\]

then the storage dynamics can be described as

\[
\dot{x} = -\frac{1}{C_{LNT}} \frac{dC_{LNT}}{dt} \cdot x + \frac{1}{C_{LNT}} \cdot \eta_s \cdot \dot{m}_{NOx,in}.
\]

The storage capacity \( C_{LNT} \) for a typical trap is a function of the trap temperature, as shown in Figure 7. It can be modeled as a Gaussian function that has center \( T_m \) [\(^\circ\)C], span \( T_s \) [\(^\circ\)C], and peak capacity \( C_m \) [g]:

\[
C_{LNT} = C_m \cdot \exp \left[ -\left( \frac{T - T_m}{T_s} \right)^2 \right].
\]

For a given trap, parameters \( T_m, T_s, C_m \) can be identified from experimental data.

The instantaneous storage efficiency \( \eta_s \), on the other hand, changes as a function of the LNT state \( x \) and the trap temperature, as shown in Figure 7. In [28], the storage efficiency is described by the function

\[
\eta_s = \frac{e^{-ax} - e^{-\alpha}}{1 - e^{-\alpha}}
\]

where \( \alpha \) is a parameter that incorporates the effects of the trap temperature on storage efficiency.
physical properties, such as formulation, geometry, etc. The second term on the right hand-side \((1 - x_{o,xy})\) captures the interactions between \(NO_x\) and oxygen storage mechanisms. For LNTs with small oxygen storage capacity, \(x_{o,xy}\) can be neglected. The function \(f_\lambda(\lambda_{in}, W_a, T)\) takes into account the effect of \(A/F\) at the LNT entrance, space velocity and temperature.

The last step in trap regeneration is to convert the released \(NO_x\) into non-pollutant species. The efficiency of this process, defined as:

\[
\eta_c = \frac{\dot{m}_{NOx,r} - \dot{m}_{NOx,tp}}{\dot{m}_{NOx,r}}
\]

can be represented by an empirical function of \(x, A/F, W_a\) and exhaust temperature.

By using the indicator function \(I\{A\}\) \((I\{A\} = 1, \text{if } A\text{ is satisfied}; \ I\{A\} = 0, \text{otherwise})\) the dynamics of the LNT can be described by:

\[
\frac{dx}{dt} = -\frac{1}{C_{LNT}} \frac{dC_{LNT}}{dt}, x + I\{\lambda_{in} > 1\} \eta_s \frac{\dot{m}_{NOx,in}}{C_{LNT}} - I\{\lambda_{in} \leq 1\} \frac{\dot{m}_{NOx,r}}{C_{LNT}}
\]

for both storage and purge operation. The \(NO_x\) flow rate leaving the LNT, \(\dot{m}_{NOx,tp}\), is

\[
\dot{m}_{NOx,tp} = I\{\lambda_{in} > 1\} (1 - \eta_s) \dot{m}_{NOx,in} + I\{\lambda_{in} \leq 1\} (1 - \eta_c) \dot{m}_{NOx,r}
\]

**B. Aftertreatment Control and Adaptation**

The requirements for DISC aftertreatment control include (1) periodically running the engine rich of stoichiometry to regenerate LNT trap capacity, (2) dealing with the sulphur poisoning problem to maintain LNT efficiency, and (3) assuring that the LNT operates within its temperature window to maintain high efficiency and to avoid thermal degradation.

To achieve the best tradeoff among competing requirements such as fuel economy, emissions and driveability, the LNT control strategy must manage the purge starting time and duration, and purge conditions (such as \(A/F\)), and at the same time provide a bumpless transition between the lean and purge modes. The main challenges of LNT control stem from the lack of on-board measurements of key variables and uncertainties in the characteristics of key components. The \(NO_x\) storage capacity of the LNT, one of the most critical parameters for control design and calibration, varies dynamically. In particular, the trap is susceptible to sulfur poisoning [29] and the capacity of the trap is reduced as sulfates accumulate. In addition, ambient conditions and component-to-component variations can affect the LNT operation and lead to deteriorated performance.

In the absence of real-time measurements, the aftertreatment control has to rely on feedforward and model-based control, making the system performance vulnerable to uncertainties and model inaccuracies. In [30], it is shown that the parameters of the LNT model [27] can be identified on-line using a conventional switching exhaust gas oxygen sensor. For the model structure and uncertainty representations used in [30], a nonlinear parametric model results. An on-line recursive algorithm is developed to improve the robustness of the model-based feedforward control and to ease the computational requirement of parameter identification for the nonlinear parametric model. Persistent excitation, a condition normally required for parameter convergence, is established in [30] by changing purge thresholds.

In an effort to relax the computational intensity associated with the nonlinear parametric model used in [30], a new purge model [28] is exploited by the authors of [31] to develop an adaptive control strategy that is more feasible for real-time implementation in a computationally resource-constrained environment. By incorporating the physical properties of the system and properly choosing the structure for the LNT model and parameterization for the uncertainties, a linear parametric model is developed in [31] for on-line adaptation. Results show that, when integrated with model-based LNT control, the adaptation improves the aftertreatment control robustness by maintaining the desired tradeoffs between fuel economy and emissions.

**V. DIESEL ENGINE CONTROL**

Diesel engines offer superior fuel economy compared to their conventional gasoline counterparts. Their drawbacks are associated with higher cost and complexity of the aftertreatment system. Despite earlier skepticism, diesel engines have achieved a remarkable passenger car market penetration in Europe thanks to technology improvements. The consensus is that their penetration in North America will grow too, albeit at a slower pace due to differences in fuel cost and taxation.

Figure 8 shows a schematic of a modern diesel engine with exhaust gas recirculation.
A. Diesel Engine Model

For compression ignited, turbocharged, diesel engines operating with high EGR, the effects of gas temperature and gas composition become more dominant and additional states have to be introduced to characterize the engine dynamics. The main difference between the diesel model introduced here and the naturally aspirated gasoline engine model discussed in Section III is the introduction of new states, specifically, \( \rho_i \) (intake density), \( F_i \) (intake burned gas fraction), \( \rho_e \) (exhaust density), \( F_e \) (exhaust burned gas fraction), \( P_e \) (exhaust pressure) and \( N_{te} \) (turbocharger speed). The state equations are obtained based on the ideal gas law, mass and energy balances and the assumption of uniform distribution of pressure, composition and temperature in the intake and exhaust manifolds.

\[
\begin{align*}
\dot{\rho}_i &= \frac{W_e + W_{egr} - W_{cyl}}{V_i}, \\
\dot{P}_i &= \frac{\gamma R}{V_i} (W_i T_i + W_{egr}((1 - \epsilon) T_e + \epsilon T_{ECT}) \\
&\quad - W_{cyl} T_i - \frac{\dot{Q}_i}{c_p}), \\
\dot{F}_i &= \frac{-F_i W_e + (F_e - F_i) W_{egr}}{\rho_i V_i}, \\
\dot{\rho}_e &= \frac{W_e - W_i - W_{egr}}{V_e}, \\
\dot{F}_e &= \frac{(F_n - F_e) W_e}{\rho_e V_e}, \\
\dot{P}_e &= \frac{\gamma R}{V_e} (W_e T_{eng} - W_i T_e - W_{egr} T_e - \frac{\dot{Q}_e}{c_p}),
\end{align*}
\]

where \( W_e, T_e \), \( W_{egr}, T_{egr} \), \( W_i, T_{eng} \), \( W_{cyl}, T_i \), \( W_e, T_{eng} \) are the rates and corresponding temperatures of the flows through the compressor, EGR valve, turbine, to the cylinder and to the exhaust manifold, respectively. Here, the temperatures of \( W_{cyl}, W_i \) are assumed to be equal to the gas temperatures in the intake and exhaust manifolds respectively. \( T_{ECT} \) is the engine coolant temperature and \( \epsilon \) is the efficiency of the EGR cooler. \( \dot{Q}_i, \dot{Q}_e \) are the heat transfer rates from the intake manifold and the exhaust manifold, respectively. Turbocharger dynamics are discussed in Section VI.

The exhaust flow, \( W_e \), and burned gas fraction in the flow from the engine into the exhaust manifold, \( F_n \), appearing on the right hand side of the state equations (19) are nonlinear functions of the states and inputs:

\[
W_e(t) = W_{cyl}(t - t_1) + W_f(t - t_2), \\
F_n(t) = \min \left\{ \frac{F_i(t-t_1)W_{cyl}(t-t_1)+W_f(t-t_2)(\Phi_e+1)}{W_{cyl}(t-t_1) + W_f(t-t_2)}, 1 \right\},
\]

where \( \Phi_e \) is the stoichiometric value of \( A/F \), and \( t_1 = \frac{30}{N(t)} \) and \( t_2 = \frac{30}{N(t)} \) are introduced to account for delays, which may not be negligible at low engine speeds. The flow through the EGR valve is modeled by the orifice equations (2)-(3) given in Section III. The EGR cooler efficiency, \( \epsilon \), is a function of the flow rate, \( W_{egr} \), and is determined experimentally with engine testing.

The cylinder flow, \( W_{cyl} \), is represented as

\[
W_{cyl} = \frac{V_d N}{120RT_i} \eta_{ool}(P_i, P_e, T_i, T_e, N) P_i,
\]

where \( \eta_{ool} \) is the nonlinear volumetric efficiency function and \( V_d \) is the engine cylinder displacement volume, while the exhaust temperature model has the following form

\[
T_e = T_i + \Delta T_{rise}(N, W_f, W_{cyl}, F_i, \delta),
\]

where \( \delta \) is the timing of the main injection. To improve the transient accuracy of the exhaust temperature model, one may augment the wall temperature dynamics with a transient heat transfer model in the following form

\[
\dot{Q}_e = k \cdot (T_e - T_{wall}), \\
\dot{T}_{wall} = -\alpha T_{wall} + \alpha T_e,
\]

where \( T_{wall} \) is the wall temperature of the exhaust manifold, while \( k \) and \( \alpha \) are parameters.

The engine rotational dynamics are essentially the same as those for PFI engines. The engine torque, \( T \), can be modeled similarly as that for the PFI engines based on engine mapping data, except that a different set of variables will be used for the functional representation of \( T \):

\[
T = T_{ind}(W_f, N, \delta, F_i) + T_{pump}(P_2, P_1, N) + T_{fric}(N).
\]

The models (20), (21) and (23) represent key static characterizations of the diesel engine and they are typically embedded into the engine control strategy. These models are usually generated via regression of steady-state engine mapping data. It is important to note that having physics-based functional forms in these models enables extrapolation during transient conditions.

B. Control Problems for Diesel Engines

Diesel engines provide many challenging control problems. The number of inputs (degrees of freedom) which needs to be dynamically controlled in a diesel engine is typically between 8 and 20, depending on the engine configuration. An increase in modeling, control and calibration
complexity occurs with each degree of freedom. Diesel engine dynamics are not only highly nonlinear but they are higher order than the ones for naturally aspirated gasoline engines.

In this section, we will briefly discuss several unique diesel engine control problems. More detailed discussion on coordinated VGT and EGR control will be postponed until Section VI, after a turbocharger model is presented.

1) Static and dynamic interactions: Figure 9 illustrates the effect of static interactions for the diesel engine shown in Figure 8. Note that at operating point “b” when the EGR valve is fully open, opening the VGT results in an increase in compressor flow. However, exactly the opposite happens at operating points “a” (when the EGR valve is closed) and “c” (when the EGR valve is fully open and the VGT is more than half open). This behavior is referred to as “dc gain reversal” and it complicates control development [32].

![Figure 9](image)

Fig. 9. The steady-state dependence of compressor mass air flow, \( W_{c1} \), on the VGT position, \( \chi_{vgt} \) for different positions of the EGR valve, \( \chi_{egr} \).

The dynamic interactions are illustrated in [32], where it is shown that engine dynamics become slower when the EGR valve is more open, and that for a usual selection of outputs the system may exhibit non-minimum phase behavior. It was also shown through numerical optimal control-based analysis [33], that the optimal operation of the VGT during a tip-in may not be its immediate closing (as the purely steady-state analysis would suggest). If the VGT is closed immediately during the tip-in, the exhaust pressure may increase rapidly in advance of the intake pressure increase thereby “depressing” the volumetric efficiency, increasing pumping losses and increasing the turbo-lag. A more optimal operation of the VGT during this transient is to initially open it, then close it and reopen it again at higher rpm to prevent over-boost.

2) Selection of Sensor Configuration and Control System Architecture: In view of static and dynamic interactions in the diesel engine, the proper selection of sensor configuration and control system architecture is particularly important. Various internal variables may be used for feedback and they result in different levels of sensitivity to uncertainties and transient performance.

The simplest analysis procedure is to determine the steady-state sensitivities of key performance variables (such as fuel consumption and emissions) to the uncertainties for different sensor and controller configurations. The underlying assumption in this analysis is that a measured internal variable is maintained by the controller at the desired setpoint despite the effects of the uncertainties. In order for this analysis to lead to meaningful conclusions, the relative importance of performance variables and the expected size of uncertainties need to be established. Note also that the best sensor configuration or controller architecture may, in general, depend on the engine operating point, as was noted previously for DISC gasoline engines.

Other related procedures include the use of control-theoretic techniques such as Relative Gain Array (RGA) analysis [32] and \( \mu \)-analysis [34]. The value of \( \mu \) is computed in [34] for different sensor configurations and at different operating points wherein low \( \mu \) implies high robustness against uncertainties and small tracking errors. It is shown that although the numerical value of \( \mu \) changes with the operating point, the relative ranking of the different configurations remains the same, thus permitting the identification of the best sensor configuration across the full engine operating range.

In addition to formal procedures that consider the effect of uncertainties, the direct analysis of interactions and properties of the system may lead to an effective control architecture. In [35], the feedback architecture is designed based on consideration of available actuator authority at the optimal setpoints. It is shown that locally at these optimal setpoints, the EGR valve and the VGT become limited in their ability to independently affect the performance variables. This analysis leads to a feedback controller architecture reliant on a single integrator instead of two. In reference [36], the exhaust pressure measurement is introduced to avoid the nonminimum-phase dynamics associated with the standard sensor configuration (compressor mass air flow and intake manifold pressure) and to take advantage of the relative degree properties of the re-defined output set. This enables application of effective robust nonlinear control design techniques. References [37], [38] propose combining switching logic and PID controllers to provide fast boost pressure response with small overshoot. Reference [39] utilizes an air-fuel ratio sensor positioned after the turbine and an LQG/LTR controller for the EGR valve in an engine with a conventional turbocharger. The use of the air-fuel ratio sensor can improve the system robustness and reduce calibration effort, although transient performance may be limited due to delay and sensor dynamics.

The guidelines resulting from numerical optimal control [33] can also be useful in comparing different controller architectures in terms of their capability to generate an optimal behavior and for ease of subsequent controller calibration. For example, it is shown in [33] that the conventional decentralized architecture, wherein the VGT is controlled using a proportional plus integral feedback on
intake manifold pressure and the EGR valve is controlled using a proportional plus integral feedback on the compressor mass air flow, is limited in its ability to generate the optimal behavior.

VI. TURBOCHARGED AUTOMOTIVE POWERTRAIN SYSTEMS

Boosting intake manifold pressure increases the density of the air entering the engine and allows a smaller displacement engine to produce torque similar to that of a higher displacement naturally aspirated engine. For diesel engines which operate lean of stoichiometry, boosting is necessary to meet power density requirements.

Turbocharging is an efficient method to boost intake pressure, as it extracts energy from the exhaust gases to drive a compressor to pressurize ambient air. It has been applied to both gasoline and diesel engines for automotive applications. Given the wide range of speed and load operating conditions for automotive applications, a design challenge is to develop a system that provides adequate boost at low speed and load without creating an over-boost situation at high speed and loads. Typically, the amount of boost delivered by a turbocharger is controlled by a wastegate for gasoline engines and a VGT for diesel engines. In any event, the advantages of turbocharging are accompanied by an increase in complexity of the control design and calibration.

Complexity is also introduced by other phenomena associated with turbocharging. For example, increasing charge density increases propensity for engine knock in gasoline engines, particularly at high loads. This phenomenon is alleviated in many applications by passive or active thermal management with a charge cooling device, such as an intercooler. In direct injection engines, fuel injection control may also provide some benefit [40].

Transient response is another factor, as turbocharger inertia leads to a phenomenon known as “turbo lag”. Turbo lag describes the delay in torque response due to the time required for the turbocharger to change speed and thus affect boost pressure. In gasoline applications, control objectives for fast response to minimize this effect are tempered by limits on boost pressure overshoot, which can lead to unacceptable torque disturbances [41], [42].

This section will first describe a turbocharger model that is applicable to both gasoline and diesel applications. We will then discuss representative control problems, one for a gasoline engine with wastegate, another for a diesel engine with coordinated VGT/EGR control.

A. Turbocharger Models

The representation of the turbocharger consists of models of the compressor and turbine, and includes the dynamic coupling of the compressor and turbine. The mass flow rate through the compressor, \( W_c \), is described by

\[
W_c = f_c \left( \frac{P_b}{P_a}, N_{tc}, T_a \right). \tag{24}
\]

where \( P_b \) is the compressor exit pressure, typically referred to as boost pressure, \( P_a \) and \( T_a \) are the compressor inlet conditions, which in most cases are assumed to be ambient, and \( N_{tc} \) is the turbocharger shaft speed.

The compressor exit temperature can be calculated as

\[
T_c = T_a \left[ 1 + \frac{1}{\eta_{isen}} \left( \left( \frac{P_b}{P_a} \right)^{\frac{\gamma - 1}{\gamma}} - 1 \right) \right] \tag{25}
\]

\[
\eta_{isen} = f_{\eta} \left( \frac{P_b}{P_a}, N_{tc}, T_a \right) \tag{26}
\]

where \( \eta_{isen} \) is the isentropic efficiency of the compressor. The power consumed by the compressor, \( Power_c \), is calculated via the first law of thermodynamics,

\[
Power_c = c_{p,c} W_c (T_c - T_a)
\]

where, \( c_{p,c} \) is the specific heat at constant pressure of the air in the compressor.

The turbine is described in a similar fashion. The mass flow through the turbine, \( W_t \), is modeled as

\[
W_t = \frac{P_c}{\sqrt{T_c}} f_t \left( \frac{P_t}{P_c}, \frac{N_{tc}}{\sqrt{T_c}} \right), \tag{27}
\]

where \( P_c \) and \( T_c \) are the pressure and temperature at the inlet of the turbine, respectively, which are typically assumed equal to the exhaust manifold conditions, and \( P_t \) is the turbine exit pressure.

The turbine exit temperature is given by

\[
T_t = 1 - \left[ 1 - \left( \frac{P_t}{P_c} \right)^{\frac{\gamma - 1}{\gamma}} \right] \eta_{tisen}^{-1} T_c \tag{28}
\]

\[
\eta_{tisen} = f_{\eta} \left( \frac{P_t}{P_c}, \frac{N_{tc}}{\sqrt{T_c}} \right), \tag{29}
\]

where \( \eta_{tisen} \) is the isentropic efficiency of the turbine.

The power generated by the turbine, \( Power_t \), is calculated from the first law of thermodynamics,

\[
Power_t = c_{p,t} W_t (T_c - T_t),
\]

where \( c_{p,t} \) is the specific heat at constant pressure of the gas in the turbine.

The dynamics of the turbocharger shaft are given by

\[
\dot{N}_{tc} = \frac{Power_t - Power_c}{J_{tc} N_{tc} \frac{d\theta}{dt}}, \tag{30}
\]

where \( J_{tc} \) is the inertia of the turbocharger.

This turbocharger model can be appended to the gasoline engine model given in Section III or the diesel engine model described in Section V to form the boosted gasoline or diesel engine model, respectively.
B. Control of Boosted Gasoline Engines with Wastegate

The turbocharger model can be augmented by a wastegate model to represent a typical boosted gasoline engine. The wastegate can be modeled with the standard orifice flow equations (2)-(3), as described in Section III. Measurements needed to derive the effective orifice area may be difficult to obtain; nonetheless, an effective model can be developed with selected use of estimated variables, such as exhaust flow rate.

Model integration requires an exhaust manifold model and a model to represent the volume between the compressor and the throttle. This can be done in a fashion similar to that discussed in Section V for diesel engines.

A turbocharged system model of this type is used by the authors of [41] to analyze system characteristics and develop charge control algorithms for a wastegated turbocharged system equipped with electronic throttle. Boost pressure and intake manifold pressure are both measured and conventional decentralized PI control with feedforward on the wastegate is used to regulate these measured variables to desired setpoints, which are chosen to achieve fuel economy, emissions and driveability objectives. The control structure is shown in Figure 10.

This approach produces acceptable performance, however the wastegate is prone to saturation. Multivariable control techniques can be used to analyze the system to guide formulation of a modified controller that maintains a simple structure desirable for implementation, and yet benefits from a centralized control methodology. Such an approach is described in [43]. This technique reformulates full state feedback integral control as an output feedback control by

\[
C_{eq} = \left( -K_{sf} (sI - (A - BK_{sf}))^{-1} B + I \right) K_i \frac{K_i}{s},
\]

where \( K_{sf} \) is the gain corresponding to the plant states and \( K_i \) is the gain corresponding to the error states. Both are obtained from full state feedback design. \( A \) and \( B \) are matrices appearing in a state-variable representation of a linear model of the system.

This \( C_{eq} \) controller is examined to identify prominent behavior. Key controller characteristics are emulated with simple linear elements such as low- and high-pass filters, to achieve an approximation of the multivariable controller. Although robustness to unmeasured disturbances needs to be further explored, a simplified controller that exhibits some of the characteristics of the multivariable control is obtained.

C. Coordinated EGR-VGT control

For boosted diesel engines, coordination of EGR and VGT represents a challenge, because of their complicated close coupling (see discussion in Section V). Both linear and nonlinear controllers have been developed for coordinated control of EGR valve and VGT. Reference [44] compares several different linear and nonlinear control designs.

One control approach, [44], relies on a multivariable linear proportional-plus-integral (MIMO PI) controller for the EGR valve position, \( \chi_{egr} \), and VGT position, \( \chi_{vgt} \), which uses the measurements of the intake manifold pressure, \( P_i \), and compressor mass flow air, \( W_c \), for feedback. This controller has the following form,

\[
\begin{bmatrix}
\chi_{egr} \\
\chi_{vgt}
\end{bmatrix} = \begin{bmatrix}
\chi_{egr,ff} \\
\chi_{vgt,ff}
\end{bmatrix} + G \begin{bmatrix}
k_1 & 0 \\
0 & k_2
\end{bmatrix} \begin{bmatrix}
P_{i,d} - P_i \\
W_{c,d} - W_c
\end{bmatrix} + G \begin{bmatrix}
k_3 & 0 \\
0 & k_4
\end{bmatrix} \nu(t),
\]

(32)

where \( \chi_{egr,ff}, \chi_{vgt,ff} \) are the feedforward positions of the EGR and VGT, respectively; \( P_{i,d}, W_{c,d} \) denote the setpoints for the intake manifold pressure and compressor mass air-flow, respectively; \( G \) is a \( 2 \times 2 \) gain matrix; \( k_1, k_2, k_3 \) and \( k_4 \) are scalar gains; and \( \nu \) is the state of the integrator,

\[\nu(t + T_s) = \nu(t) + T_s \begin{bmatrix}
P_{i,d} - P_i \\
W_{c,d} - W_c
\end{bmatrix} .\]

(33)

Note that (32), (33) can represent either a MIMO PI controller or two SISO decentralized loops. In the former case, the matrix \( G \) is selected as an inverse of the (static) dc gain of the plant for different operating conditions,

\[G = G(N,W_f) = (C_y A^{-1} B_u)^{-1},\]

where \( A, B_u, C_y, D_{uy} \) is a linearization of the plant model at the operating point corresponding to the given engine speed \( N \) and fueling rate \( W_f \). For the decentralized control, \( G \) can be chosen as diagonal with each diagonal element being an inverse of the dc gain for the corresponding input/output pair. Since the matrix \( G \) decouples the plant at low frequencies, only 4 master gains, \( k_1, k_2, k_3 \) and \( k_4 \), need to be tuned on the engine. For the case of decentralized PI control, reference [33] proposes the use of a transient governor which dynamically modifies the set-points, \( P_{i,d} \) and \( W_{c,d} \) for the EGR valve and VGT positions to provide a faster tip-in response.

Another approach can be developed based on Control Lyapunov Functions (CLF) method [45]. This method is applied to a reduced order model with three states \( \bar{P}_i = \frac{P_i}{T_{amb}}, \bar{P}_c = \frac{P_c}{T_{amb}} \) and \( Power_c \), where the turbine flow
(\(W_t\)) and the EGR flow (\(W_{egr}\)) are treated as control inputs. A Lyapunov function,
\[
V = \frac{c_1}{2}(W_e - W_{e,d})^2 + \frac{c_2}{2}(\bar{P}_e - \bar{P}_{e,d})^2 + \frac{c_3}{2}(\bar{P}_{i} - \bar{P}_{i,d})^2
\]
is constructed for the closed-loop system with a feedback linearizing controller. Following the CLF design procedure\(^2\), a controller for the desired mass flow rate of EGR, \(W_{egr}\), and the desired mass flow rate through the turbine, \(W_t\), can be derived. The EGR valve and turbine flow characteristics are then inverted to backtrack the desired EGR valve and VGT positions from \(W_{egr}\) and \(W_t\). This controller enjoys robustness properties (to input uncertainties) guaranteed for the CLF controllers with infinite gain margin and 60 deg phase margin. It also suggests the advantages of using the exhaust manifold pressure measurement for feedback [45].

VII. CONTROL IMPLEMENTATION CONSIDERATIONS

There are important considerations in implementing automotive powertrain controllers that need to be taken into account during the development and technology transfer of the algorithms. They include computations, calibration and process issues, among others. Specifically, automotive micro controllers provide only limited computational resources (in terms of chronometrics and available RAM and ROM size). They are highly optimized for reliability and cost but their computational capability is nowhere near the current PC technology and does not improve with time at the rate suggested by Moore’s law. To accommodate different engine and powertrain configurations, the same control system and software are often used, but the parameters in it (numbered in thousands for a typical strategy) are tuned for the specific application. The tuning process (referred to as calibration) is usually performed by a different group of specialists (referred to as calibrators) than the control system developers. Consequently, the control algorithms must be designed concurrently with a streamlined calibration procedure for their parameters.

To be successful, control developers must have a good understanding of the above issues as well as a complete picture of the requirements and interactions for the particular control algorithm in all powertrain operating modes, inclusive of abnormal operation. A demonstration of performance and robustness improvements versus existing technologies, the considerations of scalability and modularity of control algorithms for different powertrain configurations, effects on overall customer vehicle perception and on the vehicle reliability and durability also become critical for a successful technology transfer.

As has been recently recognized in popular press including product advertisements by leading OEMs, control is becoming more and more a key product differentiator. In this light opportunities for advanced powertrain control applications should only grow not only in terms of delivering improved performance, accommodating more complex powertrain systems and satisfying more stringent requirements, but also in terms of improving control system implementation process including the reduction in calibration time and effort.

VIII. CONCLUSION

Powertrain control has been and will remain a dynamic and exciting research topic. For many years, automotive systems have been used as ideal platforms for motivating and validating theoretic control research. Hardware innovations and product demands will continue to offer opportunities for the powertrain control engineering community. We expect that the demand for education of control engineers with well balanced automotive and dynamics perspectives will grow, as will the demand for tool development that can integrate the most recent control theoretical advancements into powertrain control development processes. Many interesting research topics with direct applications to powertrain control, such as data-driven model development methodologies and tools, hybrid system analysis, integrative control strategies for complex systems, and effective embedded optimization algorithms, remain open and under-explored; and advances in these directions could have significant impact on the automotive industry.

REFERENCES


\(^2\)For a nonlinear system, \(\dot{x} = f(x) + g(x)u\), with a Control Lyapunov Function, \(V\), the CLF controller has the form \(u = -\gamma(\frac{\partial V}{\partial x})^T g(x)\).